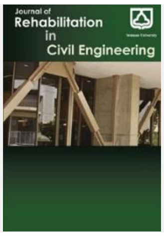


Journal of Rehabilitation in Civil Engineering

Journal homepage: https://civiljournal.semnan.ac.ir/

Machine Learning-Based Empirical Formulations for Strength Properties of Steel Fiber Reinforced Concrete

Mohammad Hossein Taghavi Parsa[1]*; Mohammad Reza Adlparvar[2]*; Morteza Esmaeili[3]

1. Ph.D. Candidate, Department of Civil Engineering, Faculty of Engineering, University of Qom, Qom, Iran
2. Associate Professor, Department of Civil Engineering, Faculty of Engineering, University of Qom, Qom, Iran
3. Professor, Department of Railway Engineering, Iran University of Science and Technology, Tehran, Iran

Corresponding author: *mh.taghavi@stu.qom.ac.ir* (M.H.T.P); *adlparvar@qom.ac.ir* (M.R.A.)

ARTICLE INFO

Article history:
Received:
Revised:
Accepted:

Keywords:
Strength properties;
Machine learning;
Empirical formulation;
Steel fiber reinforced concrete.

ABSTRACT

The accurate approximation is a benefit of the modern machine learning technique, which also disappeared the problems of traditional empirical methods, such as human and technical errors plus environmental pollution. Although there are many good samples on the state-of-the-art regarding the machine learning prediction of strength properties of steel fiber reinforced concrete, fewer articles are dedicated to proposing empirical formulations. This paper brings some novel empirical formulations to identify the strength properties of macro steel fiber-reinforced concrete. A 2650 multi-national data records are used to perform the regression, which is an exclusive dataset. This archive is the largest available dataset used in the state-of-the-art steel fiber-reinforced concrete prediction process, which is beneficial for supervised learning. Since the user must be careful regarding overtraining with such a vast resource, a successful strategy provided by the authors in previous research is utilized in which various machine learning techniques are compared to forecast the considered properties. So the Ridge, Lasso, and linear methods are used as regressors to predict the strength properties and the constants. Symbolic regression, a powerful tool for producing empirical formulations, is used for creating mathematical expressions regarding the strength properties. The performance is also evaluated based on well-known error analysis metrics. The formulations are presented for flat, waved, and hooked end fibers, the most common fibers used in construction engineering. The machine learning-driven formulations are exclusive due to the utilized strategy and the resources, and the precision of the relations are denoted, which presents the superiority to traditional methods.

E-ISSN: 2345-4423
© 2024 The Authors. Journal of Rehabilitation in Civil Engineering published by Semnan University Press.
This is an open access article under the CC-BY 4.0 license. (https://creativecommons.org/licenses/by/4.0/)

How to cite this article:
Taghavi Parsa, M. H., Adlparvar, M. R., & Esmaeili, M. (2024). Machine Learning-Based Empirical Formulations for Strength Properties of Steel Fiber Reinforced Concrete. Journal of Rehabilitation in Civil Engineering, 13(1), 27-44. https://doi.org/10.22075/jrce.2024.32743.1963

Journal Pre-proof

1. Introduction

Forecasting the strength properties of a broadly used construction material such as concrete would be possible with both direct and indirect approaches. The direct methods, as it first comes to mind, are methods in which the researchers spend huge costs and time to produce dozens of experimental specimens based on the 28-day prepared sample according to the standard and tested according to the regulations. In contrast, in the indirect methods, the scientists replace surrogate models, some numerical models developed based on the statistics to predict the properties in the non-sampled points to decrease the number of tests and, consequently, the time and cost. Many of these methods were promising in theory but were effective and efficient due to the high aggregated error values compared with the experimental outcome. One of the main reasons was the shortcomings in collecting adequate data for the primary dataset. The other was the inefficiency of the method in prediction due to some fundamental lacks. Many of the regression methods fail to face non-smooth objectives. In such cases, a perturbation in the objective functions may cause fast changes in the objective value. This phenomenon makes the objective look like a step function which is also non-differentiable. A situation of occurring such a case is in the buckling of columns. So, proposing novel methodologies to improve the predictions in such instances and testing the available alternatives is highly important. For example, developing kriging models with some local regressions in the discontinuous parts of the objective [1] is a possible alternative but needs more progress.

The background of using steel fibers in the concrete matrix backs to half a century. This research study comes with a vital context and is ongoing in state of the art. This type of reinforced concretes benefits the bridging of the fiber in the crack tips to prevent furthered propagation in the composite structure that finally leads to some ultimate possible strength properties. It must be noted that this facility will be highlighted in the performance of concrete in the post-peak cases. In some papers in the very beginning, the strength properties of the reinforced concrete are a function of two variables. These factors are the strength of unreinforced concrete and the percentage of fiber used in the mixture of the mentioned concrete. Some research works reported the strength properties to have a positive relationship with this percentage, while others recommended a second power law relation.

In one of the vanguard studies, Wafa and Ashour [2] used all four types of steel fibers consisting of crimped, duo form, hooked end, and straight with the consideration of the volume fraction in a wide range of strength. The extremums for the strengths were 41 and 115 MPa for straight and hooked end fibers. The paper became the basis for many further studies. Some development trend in this research path is illustrated in Fig. 1, in which some significant studies are noted. According to the described process, the research by Khaloo and Kim [3] investigated the compressive, tensile, and flexural strength of steel fiber-reinforced concrete (SFRC) and derived the corresponding formulations as a function of fiber percentage and the strength of unreinforced concrete.

The compressive empirical formulation predicts the experiment with a 30 percent of error, and it is relative to both the first and second power low of the volume fraction of the fibers. The structure of the identified formulations for the flexural and compressive strength is the same, with different coefficients and a similar 30 percent error compared with the experiments. Next, using experimental specimens, Nataraja et al. [4] modified the relationship by producing the global stress-strain curve. Three values of volume fraction and two for aspect ratio were considered the

Journal Pre-proof

permutations of the research study to generate the corresponding numerical formulations. The generated empirical formulae forecast the compressive case with a 25 percent of error which was an improvement at the time. The reference [5] modified the model proposed in [3] to predict the properties of high-strength SFRC (more than 60 MPa), but the error analysis revealed an error of 35 percent in the prediction, which is not surprising. Also, about 98 percent of improvement was reported in the splitting tensile strength.

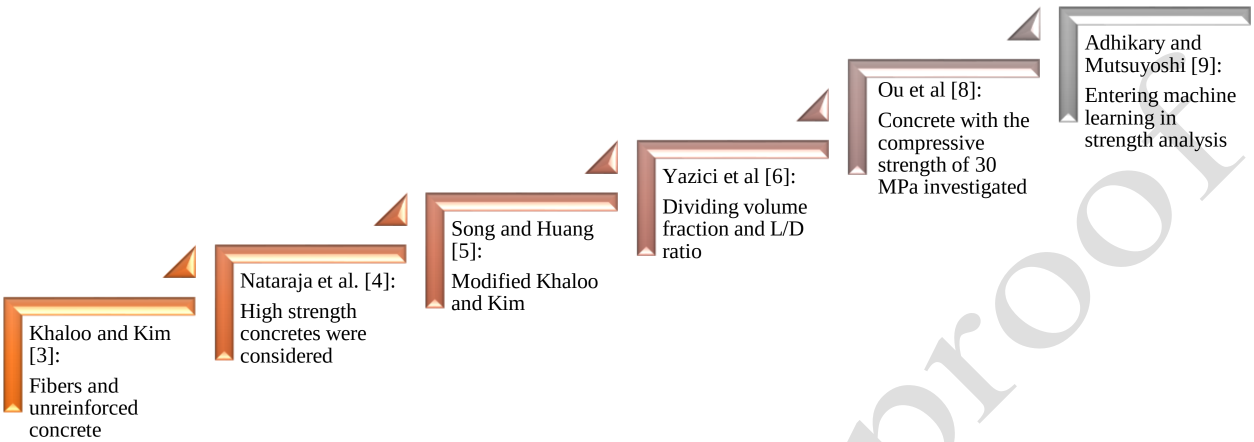


Fig. 1. Some critical research in the context of steel fiber reinforced concrete.

The effect of aspect ratio plus volume fraction was first studied in [6], where three types of this ratio are utilized in ten mixes to produce specimens. The errors for these predictions depend on the strength property, so for compressive, tensile, and flexural cases, it is correspondingly 10, 28, and 23 percent. Another study presented empirical formulations and reported error values of 10 percent, which was a considerable improvement [7]. The drop test was utilized to derive the empirical relationship for impact resistance of the SFRC in which the lightweight concrete is analyzed [10]. The formulations brought novel relations between flexural toughness and the impact energy.

The machine learning technique has been widely used in the state-of-the-art as an alternative to the traditional surrogate modeling techniques, presenting higher efficiency and accuracy. The community has broadly studied the integration of machine learning in predicting SFRC strength properties and the features of other types of concrete. Alilou and Teshnehlab [11] used a feed-forward neural network to predict the strength of concrete under compression with consideration of 3-day strength as an essential parameter. The technique is claimed to decrease the analysis duration due to using fewer samples. The artificial neural network is reported for the premiere performance in altering the prediction using the Lavenberg Marquardt gradient-based optimization [12].

The effect of the number of neurons designed in the hidden layer of the network of an artificial neural network (ANN) structure is reported to be directly influential on the outcome [13]. This study also performed an analysis of sensitivity for analyzing this parameter. More research on using ANN in predicting SRFC properties is available in [14-18]. In another study, available data from state-of-the-art and personal experiments are utilized to create empirical models predicting flexural strength [19]. The proposed model in the mentioned research had more than 80 percent confirmation from the validation technique. Some researchers suggested mathematical models for this study. For example, [20] used a numerically simplified formula derived by removing parameters such

Journal Pre-proof

as water, cement, and water-to-cement ratio. Besides, the concrete strength gain specifications are simplified by eliminating polynomial equations and replacing them with a simple mathematical formula.

The considerable efficiency of the stepwise regression is presented in [21], where newly developed formulations are proposed based on the analysis of shear capacity. The research's outcome is according to the statistical characteristics of seven single parameters. The gene expression programming -presented some of the most accurate results in the forecasting process of shear and compressive strength [22], and the keynote of the related study was that the shear strength decreases with the shear span to depth ratio. Another effective utilization of this method which led to new formulations can be found in [23]. Al-Musawi et al. [24] changed the general gradient-based solver of supported vector regression by the firefly algorithm, which is a population-based heuristic optimizer. This modification has significantly improved the efficiency of this method which was looser compared with many other regression methods [25]. Another idea for enhancing the supported vector regression in predicting the SFRC strength was the hybridization of the response surface to forecast the shear capacity [26]. This way, the supported vector machine was adjusted utilizing the corresponding response surface. The shear resistance of the SFRC has recently been investigated with multi-expression programming [27]. This method was used for analyzing the shear resistance. Also, in a recent study, Ahmadi et al. [28] proposed newly developed mathematical relations and correlations between some geometrical and material properties and the shear stress of SFRC using the gene programming method, which had high accuracy compared with many references. The well-known, fully developed machine learning techniques are compared based on standard Python implementation to find the most elite one in predicting the strength properties of SFRC [29]. Despite the available research items on the strength properties of SFRC [30], there are insufficient achievements in proposing machine learning-driven empirical formulations for the compressive, tensile, and flexural strengths of SFRC. Therefore, more symbolic machine-learning techniques must be used and developed for this purpose.

The present research is dedicated to deriving empirical formulations using data-driven supervised machine-learning techniques. Since linear regression methods are widely and effectively used in recent research on composite structures [31], three linear regression techniques are investigated to approximate an objective function. The data for these methods are prepared based on a strategy to remove outliers from the huge prepared dataset to approach the highest accuracy. The influence of each feature is presented in the paper using signed parameters. An error analysis is also available in the text. The formulations presented in this paper are prepared using symbolic regression. The rest of this research paper is structured as follows: Section 2 explains the methods used to produce the outcome, Section 3 is dedicated to the results and discussions, and finally, Section 4 highlights the concluding remarks from the results and discussions section.

2. Methodology

In this paper, three well-known regression methods from the family of linear methods are used for the prediction process. The authors used these three methods in another research in which various methods were compared for the approximation of properties of SFRC [29]. The data regarding the considerable accuracy of these methods is available in the article. The factor of effectiveness for these methods is broadly discussed for each feature. It must be noted that an exclusive strategy is used for preparing the data the authors proposed in

Journal Pre-proof

another research article, and the efficiency of this approach is proven. The symbolic regression is used to derive the empirical formulations for the strength properties. Also, the k-fold validation technique [32] ($k=5$) is used to prevent overtraining. Using more values of this parameter was not beneficial and increased the computational cost. It must be mentioned that the outcome of the training process is an average of 30 repetitions progresses to avoid the unreliability of the non-deterministic comport of the regressor algorithms.

2.1. Feature selection strategy

The strategy for data preparation and feature selection [33] is illustrated in Fig. 2. The adopted strategy prevents overtraining [34] and numerical errors to receive the best regression performance and accuracy. This strategy was proposed, and its effectiveness is proven in another research [29].

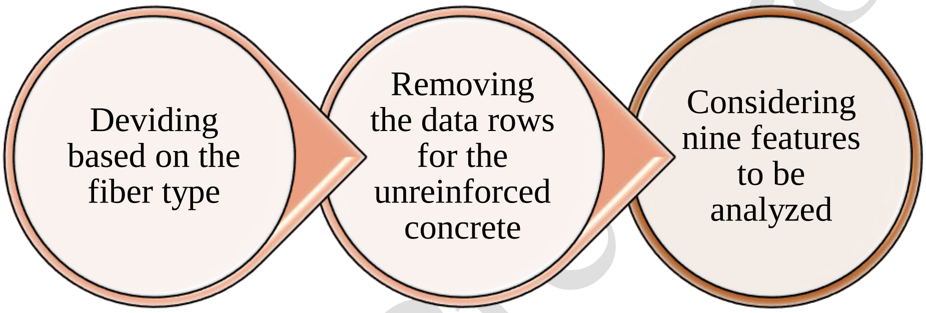


Fig. 2. The process illustration for the data preparation and feature selection.

According to the illustration, the first step is dedicated to data division based on the fiber type. In the next section, it is visible that each strength property depends on the fiber type, which is so important. After this essential step, the data for unreinforced concrete is removed from the dataset. This effort helps to remove outlier data from the input of the forecasting algorithm. Finally, the features are analyzed separately based on the fiber type and each strength property.

2.2. Symbolic regression

This method produces tree branches using problem features and mathematical operators [35]. Each of these branches is then sorted by their accuracy. Finally, the formulation with the best performance will be combined. Also, some modifications might be applied randomly to the same as a mutation operator. Then previous iterations' trees are combined and rated again, and this procedure continues until the termination criteria.

2.3. Linear regression

This method is recommended in the state-of-the-art as a general and well-established technique in a supervised branch of machine learning. The reason is the ease of implementation and availability of the code in Python script. It must be noted that the linear approximation has shown approximately good applicability in the engineering community, although the natural phenomenon is always nonlinear. In this method, a linear correlation is performed to present the regressed function as follows [36]:

$$\widehat{B} = \vartheta_0 + \vartheta_1\mu_1 + \vartheta_2\mu_2 + \ldots + \vartheta_{n-1}\mu_{n-1} + \vartheta_n\mu_n, \qquad (1)$$

Journal Pre-proof

The above formulation, $\hat{B}$ stands for the dependent variable, ϑ_n is a bias, and the independent variable denoted as μ_i .

2.4. Ridge regression

It is a kind of square-type regulation regression method and can be formulated as follows [37]:

$$\hat{F}(\vartheta) = \varsigma \sum_{e=1}^{E} g_e^2 + \arg\min \sum_{r=1}^{R} (\chi_i - \tau_i g)^2, \tag{2}$$

where, χ_i is an individual of the vector of observation, τ_i is an individual for the regression matrix, and finally g stands for the vector of regression coefficients. The term $\sum_{e=1}^{E} g_e^2$ is known as the regulation term multiplied by the hyperparameter ς is multiplied in this phrase.

2.5. Lasso regression

With a similar formulation compared with the previous regressor, this method contains a term $\arg\min$, but the L1-norm is replaced with the first term with the penalty parameter denoted as ρ. The formulation can be described as follows [38]:

$$\hat{Y}(\vartheta) = \rho \sum_{e=1}^{E} |g_c| + \arg\min \sum_{r=1}^{R} (\chi_i - \tau_i g)^2, \tag{3}$$

3. Results and discussions

The results of this prediction process are divided into three sections. The first section presents the effectively signed indicators for the considered features of the steel fiber-reinforced concrete. Secondly, the formulations derived by parametric regression are presented. Thirdly, the error analysis regarding the regressions introduced by Ridge, Lasso, and the linear method is presented as bar charts.

3.1. Feature indicators for each regressor

The effective parameters for the Linear regression, Ridge, and Lasso method are available in Table (1). In this Table, Table, MA stands for the max aggregation based on millimeters; E is Young's modulus in GPa; PCCS stands for plain concrete compression strength; PTS is the plain tensile strength. Also, fiber percentage is denoted as (%); L is the length of fiber in millimeters; The aspect ratio is represented as L/D. The feature constant indicates the constant variable in the regression formulae. The multipliers describe the value of features' positive or negative effects of a feature in the corresponding strength properties.

According to Table 1, the maximum aggregation parameter has a considerable positive effect on all strength properties for all flat fibers regardless of the regression method. However, the linear regression has proposed higher values for all constants. For the waved type fiber, the maximum aggregation harms both compressive and tensile strength of waved type fiber, while it works positively on the flexural strength. The MA parameter would have a nullifying consequence for the hooked end fiber. The constant parameter is also a perpetual positive parameter.

Regarding the flat-type fiber, Young's modulus is forecasted to positively affect the compressive strength with the minimum value of 10.8012 predicted by Lasso and the maximum of 14.5873 indicated by the Ridge method. This parameter is not entirely practical for the two other strength properties. The Ridge method reported the most pessimistic value of -13.2573 for flexural strength. In the wave-type fiber-reinforced concrete, the flexural strength benefits from Young's modulus in contrast with the flat type. This parameter is also beneficial in the tensile and flexural strength of hooked-end reinforced concrete fiber while partially unfavorable for compressive cases.

Journal Pre-proof

The plain concrete compression strength is observed to significantly and positively affect altering the compressive strength of the flat and hooked-type fibers. The most optimistic effectiveness was predicted by the Lasso method, with a value of 100.7482 for the flat fiber. The linear process predicts the hooked end with a factor of 141.1466 in compressive strength. The compressive strength would be improved by increasing this feature in the wave type fiber in the compressive strength, but not as much as the other two other types. The other values for the tensile and flexural strength are also available in Table 1.

Table 1. The parameters identified from each regressor.

Method	FT	Feat / Rel	Const.	MA (mm)	E (GPa)	PCCS (MPa)	W/C	%	L(mm)	(L/D)	PTS (MPa)
Lasso	Flat	Comp	88.6812	32.4810	10.8012	100.7482	-74.9379	29.3741	-22.2810	-14.3059	…
		Tens	2.1291	3.8588	-0.7084	-13.2339	-3.5783	12.8055	-1.7652	2.6532	26.7338
		Flex	25.7340	0.4064	-6.6852	-29.1643	-26.7981	74.2826	-0.5694	9.7344	0.1777
	Waved	Comp	110.2469	-14.3321	1.2361	15.8769	-96.6819	0.3380	-4.6043	23.2749	…
		Tens	3.2953	-0.3379	-0.0264	1.1055	0.2077	-0.5895	0.4878	0.5874	0.0628
		Flex	0.0162	1.1172	0.1636	-0.9660	-7.9564	18.0019	-3.1494	10.0503	27.1749
	Hooked	Comp	56.4388	-9.6256e-02	-2.8403e-03	1.4018e+02	-5.5487e+01	2.1091e01	-7.8441e+00	1.1552e+01	…
		Tens	4.1638	-1.1408	1.0656	-1.6700	-5.8725	5.7278	1.7960	-0.7311	6.2756
		Flex	6.6688	-0.9312	0.2434	-1.8802	-13.8507	33.4757	5.0533	-1.7062	17.9830
Linear	Flat	Comp	87.1948	37.5246	12.8602	102.1857	-77.0943	30.1850	-23.2134	-19.5005	…
		Tens	2.2060	4.1218	-0.8906	-17.1768	-3.7331	12.8664	-2.2505	3.7447	30.9267
		Flex	29.9491	3.2882	-13.2573	-37.6045	-31.2828	76.9769	-1.7381	17.0591	6.5929
	Waved	Comp	106.2378	-15.5955	6.7084	20.4515	-93.5729	0.5469	-7.5380	29.6110	…
		Tens	2.4071	-1.1185	-0.1796	2.5538	1.3043	-1.4733	0.5230	1.8571	-0.3753
		Flex	0.9745	4.4275	0.2695	-6.9930	-12.7347	20.0089	-4.6397	10.3723	32.6163
	Hooked	Comp	57.4016	-0.2196	-1.6094	141.1466	-56.0155	24.9601	-8.9083	12.9077	…
		Tens	3.8254	-1.2295	1.4281	-1.8581	-5.9369	5.7630	1.8652	-0.7794	6.5961
		Flex	6.3440	-1.8289	1.3634	-3.6366	-14.6543	36.5674	5.8133	-2.2504	17.8638
Ridge	Flat	Comp	88.7459	28.2381	14.5873	96.0972	-71.5904	30.0103	-24.3841	-16.1291	…
		Tens	2.3868	3.7944	-0.8326	-10.299	-3.8382	12.4547	-1.5332	2.2522	23.4465
		Flex	29.2814	1.5214	-11.9347	-30.4330	-28.7797	72.3811	-1.9838	12.9535	1.6131
	Waved	Comp	97.6521	-17.1898	4.2288	26.0555	-81.1298	-1.7368	-8.2408	30.0343	…
		Tens	3.0708	-0.5808	-0.0592	1.2781	0.3254	-0.8935	0.6354	0.8291	0.2263
		Flex	-0.3891	1.7940	0.2390	2.0908	-7.8479	17.9341	-4.9114	10.9892	22.8218
	Hooked	Comp	57.3501	-0.4641	-0.7945	139.5486	-55.5563	24.1628	-9.0875	12.7326	…
		Tens	4.3204	-1.1398	1.0670	-1.4407	-5.7940	5.4271	1.7873	-0.7728	5.7788
		Flex	6.9542	-1.6495	1.4568	-2.9443	-14.6110	30.7881	5.3051	-2.3622	17.6997

Journal Pre-proof

The W/C is an already negative factor in the strength of concrete and must be avoided, especially for the compressive strength.

The fiber percentage, which is known as one of the crucial factors, is also investigated. This study reports a negative influence on the tensile strength of wave fiber-reinforced concrete. The plain tensile strength, which is not available for the compressive strength, significantly affects the tensile strength of the concrete reinforced with flat fiber and the flexural strength in the waved-type reinforced structure. Overall, slight incoherence is present in these three regression methods' results regarding the sign that presents the mentioned strategy's effectiveness.

3.2. Derived parametric formulations

The predicted formulation for the compressive strength for the flat, waved, and hooked end fiber types can be expressed as follows:

$$\begin{aligned}
F_1\left(X_2,X_3,\alpha_1,\alpha_2\right) &= X_0X_2X_3\left(-\alpha_1\log\left(\log\left(X_2\right)\right)+\frac{\alpha_1}{\cos\left(\sin\left(\frac{X_4}{\log\left(X_2X_3\right)}\right)\right)}-\frac{1}{\alpha_2-X_3}\right)-\ldots\\
&\ldots-\alpha_1\log\left(-X_3\right)+\log\left(X_2X_3\right)+\log\left(X_2X_3\left(\frac{-1}{\alpha_2-X_3}+\frac{X_4}{X_0}\right)\right)+\ldots\\
&\ldots+\log\left(X_2X_3\left(X_2X_3\log\left(X_2X_3\right)-\frac{1}{\alpha_2-X_3}\right)\left(\frac{1}{\cos X_2}-\frac{1}{\alpha_2-X_3}\right)\right)+\ldots\\
&\ldots+\log\left(\alpha_2\log\left(\cos\left(X_3-\alpha_2\right)\right)\right)+\frac{\alpha_1}{\cos\left(X_2\right)}-\frac{1}{X_3}+\frac{X_4}{X_0},
\end{aligned} \tag{4}$$

$$\begin{aligned}
F_2\left(X_0,X_1,X_3,X_4,X_5,X_6,\alpha_3,\alpha_4,\alpha_5\right) &= \log\left(\cos\left(\alpha_3X_0-\frac{\cos\left(\alpha_3X_0\right)+\cos\left(X_1\right)}{X_1+X_5X_6+\sin\left(X_1\right)-\alpha_4}\right)+\alpha_5\right)-\ldots\\
&\ldots-1/\left(-\frac{X_1}{X_0X_6+0.397}-\left(X_1-\alpha_4\right)\left(X_1-X_3+2\left(X_1-\alpha_4\right)\cos\left(\log\left(X_6\right)\right)\right)\right),
\end{aligned} \tag{5}$$

$$\begin{aligned}
F_3\left(X_2,X_3,X_6,\alpha_6,\alpha_7,\alpha_8,\alpha_9\right) &= \left(\alpha_6\log\left(\frac{0.089}{X_2}\right)-\alpha_7-\frac{\alpha_6}{-\alpha_6+\frac{1}{X_3}+\frac{0.089}{X_2}}\right)/\ldots\\
&\ldots/\left(X_2-\alpha_9+i\pi-\frac{\log\left(X_2\right)^2}{X_2}\right)+2\log\left(X_2\right)-\ldots\\
&\ldots-\alpha_6\log\left(\cos\left(\frac{1}{\cos\left(X_6-0.241\right)}-\frac{1}{-X_2-0.089+\frac{1}{X_3}}\right)\right)+\frac{\alpha_6}{\cos\left(X_2\right)},
\end{aligned} \tag{6}$$

where X_0 is the maximum aggregation size in millimeters; X_1 is Young's modulus in GPa; X_2 is the plain concrete plain strength in MPa; the ratio of water to cement is denoted as X_3; fiber percentage is denoted by X_4 and the length of fiber is presented by X_5; X_6 is the

Journal Pre-proof

aspect ratio; X_7 is the Plain Tensile/Flexural strength based on the output type. The coefficients are:

$\alpha_1 = 39.497890205207, \alpha_2 = 0.166$,

$\alpha_3 = 5.43478260869565, \alpha_4 = 0.028399474521698, \alpha_5 = 0.999596762026504, \alpha_6 = 39.497890205207$,

$\alpha_7 = 144.229010561867, \alpha_8 = 1.02564102564103, \alpha_9 = 3.67624725795418$.

According to these formulations, the compressive strength of flat fiber reinforced concrete depends on PCPS and W/C. For the waved-type reinforced structure, the compressive strength also depends on L, fiber percentage, W/C in addition to the mentioned features. For flexural strength, X_2, X_3, X_5 are the influential factors. Similarly, the relations regarding the tensile strength for the flat, waved, and hooked end fibers correspond as follows:

$$F_4(X_0, X_1, X_2, X_3, X_4, X_5) = \frac{X_4}{X_5} - \log(0.044 X_3) + \frac{X_1 + X_2 + X_4}{X_0}, \tag{7}$$

$$F_5(X_6) = X_6 + 3.53356890459364, \tag{8}$$

$$F_6(X_3, X_4, X_7) = \frac{X_4}{X_7} + 7.57575757 + \frac{1}{X_3 + 0.0132}, \tag{9}$$

For the flexural strength, the corresponding formulations for the flat, waved, and hooked end fibers are:

$$\begin{aligned} F_7(X_0, X_2, X_3, X_4, X_5, X_7, \alpha_{10}) &= \frac{-X_0}{X_3 X_7 \log(X_4)} + X_4 X_5^3 X_7 + \left(\alpha_{10} X_4 + \sin\left(\frac{1}{X_5(X_0 - X_7)}\right)\right) \times \\ &\ldots \times \log\left(\frac{\dfrac{0.147}{-X_7 + \cos(X_0 - X_5)} + \dfrac{\sin(X_5)}{X_0 - X_2}}{X_4 X_5^2 X_7 (X_0 - X_2)}\right) + \frac{1}{\sin(X_5)}, \end{aligned} \tag{10}$$

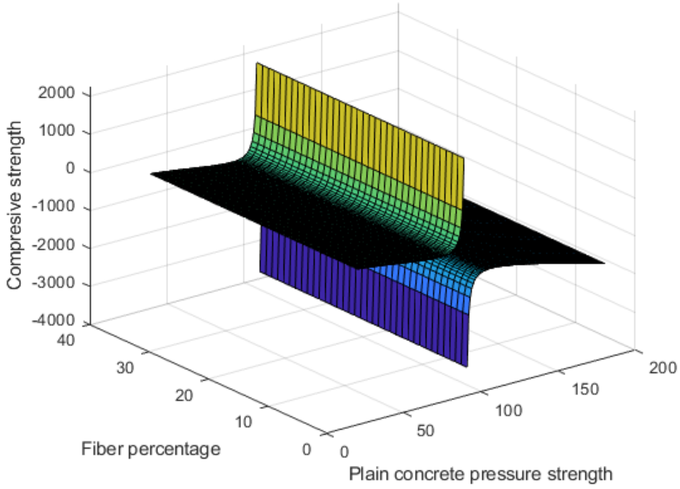
$$\begin{aligned} F_8(X_0, X_3, X_4, X_7) &= X_0 + \log(X_0) - \log(X_6 - 0.426) + \frac{X_4}{X_3} + \frac{1}{X_3} + \ldots \\ &\ldots + \frac{X_4 + X_7}{X_0^2 \left(\log\left(0.426 + \dfrac{1}{X_3}\right) + \dfrac{0.592}{X_0}\right)} \end{aligned} \tag{11}$$

$$F_8(X_3, X_4, X_6, X_7) = X_7 \left(\frac{2X_4}{X_6} + 28.5714285714286 X_7 - \frac{1}{X_3}\right) + 3.003003003003 + \frac{1}{X_3} \tag{12}$$

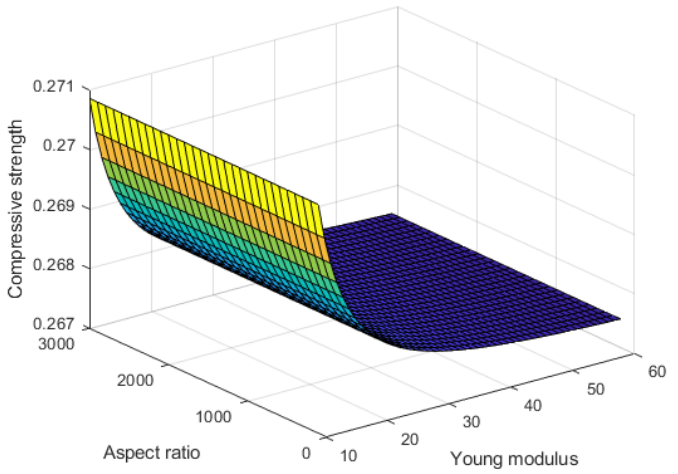
Where $\alpha_{10} = 4.97512437810945$. Fig. 3-5 shows the three-dimensional plots of the driven formulations. It must be noted that the real part of the formulations is plotted, and the imaginary parts are removed consequently. Since some of these equations are functions of more than one feature, to create three-dimensional plots, some features are considered as a constant equal to the averaged values correspondingly. These averaged values for the features $X_0, X_1, X_2, X_3, X_4, X_5, X_6, X_7$ are 13.27 mm, 58.67 MPa, 0.4

Journal Pre-proof

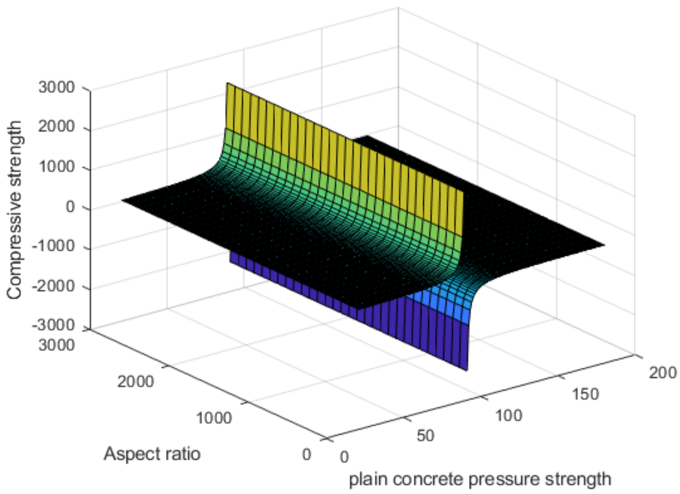
(dimensionless), 38.39 mm, 84.17 (dimensionless), 7.27 MPa for flexural strength, and 4.12 MPa for tensile strength.



(a)



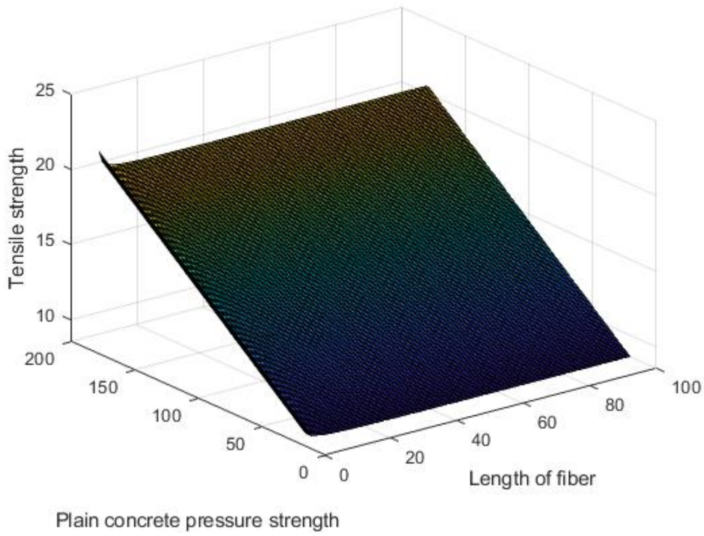
(b)



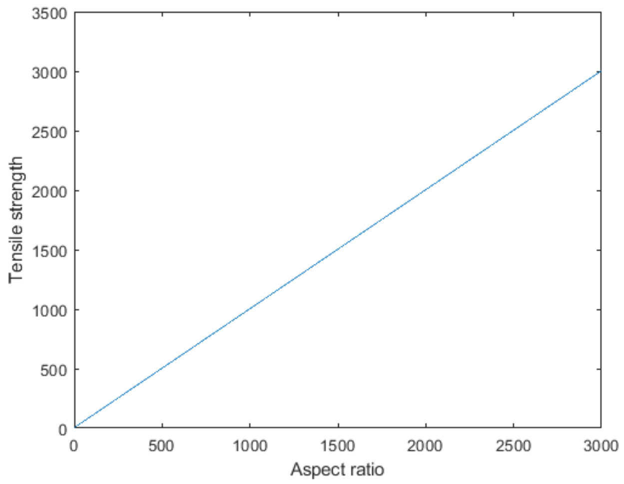
(c)

Fig. 3. Compressive strength behavior based on fiber type (a) Flat (b) Waved (c) Hooked end.

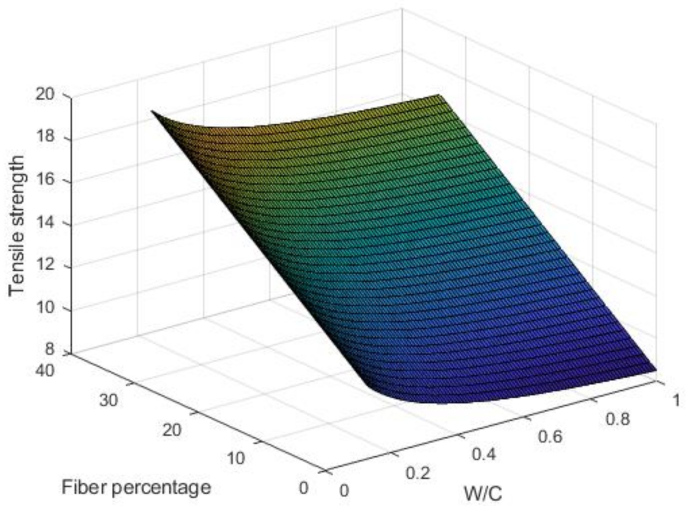
Journal Pre-proof



(a)



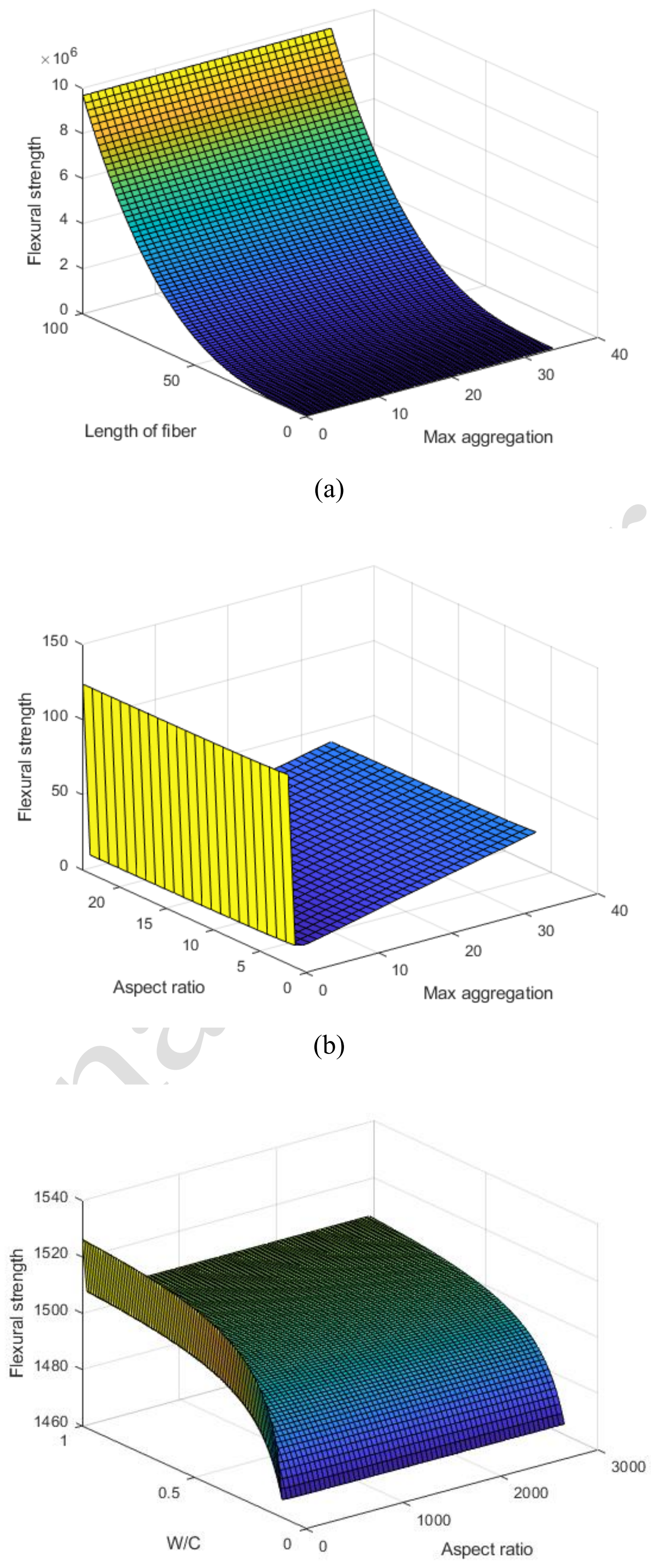
(b)



(c)

Fig. 4. Tensile strength behavior based on fiber type (a) Flat (b) Waved (c) Hooked end.

Journal Pre-proof



(c)

Fig. 5. Flexural strength behavior based on fiber type (a) Flat (b) Waved (c) Hooked end.

3.3. Error analysis

The error analysis is presented for the three types of fibers and all strength properties for the compressive, tensile, and flexural cases. Four types of errors consisting of root mean square (RMSE) [39], mean absolute percentage error (MAPE) [39], which is presented in the scale of 0.01 percent, mean absolute error (MAE) [39], and mean error (ME) [39]. It must be noted that ME is presented based on 10 MPa in the boxplot.

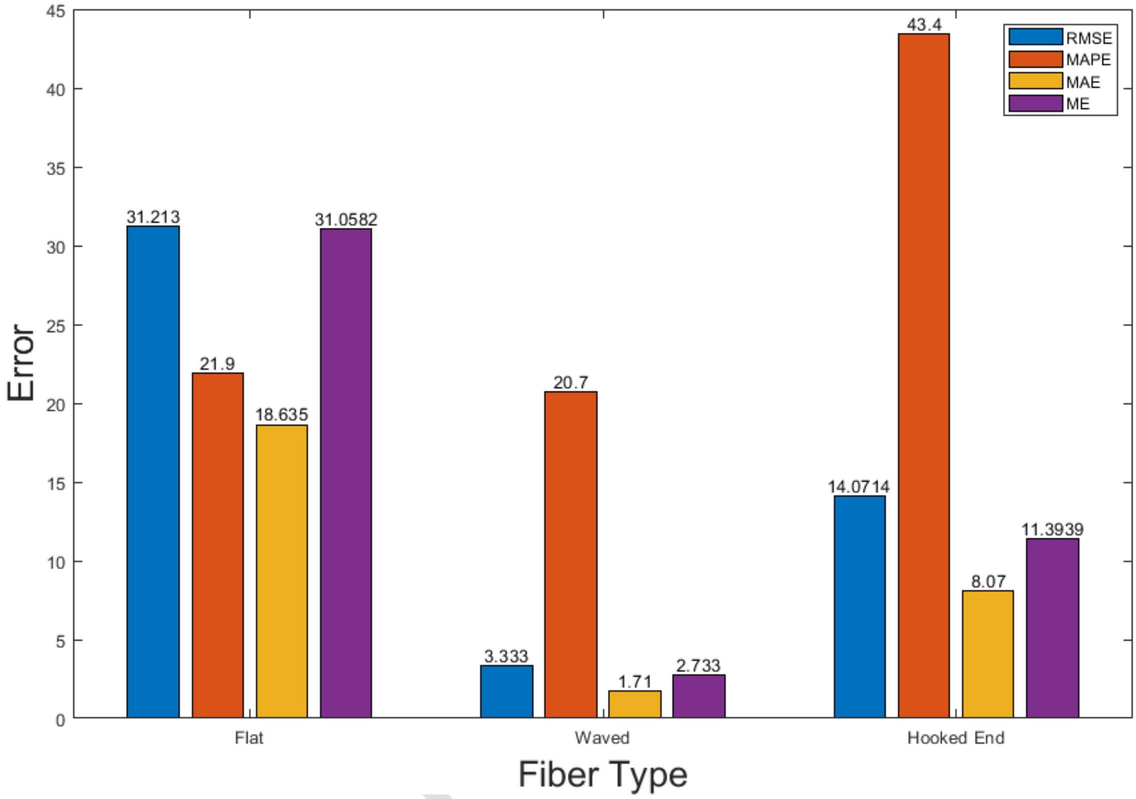


Fig. 6. The error analysis for the compressive strength.

According to Fig. 2, for the compressive case study, the value of RMSE peaked for the flat type fiber and dropped nine times smaller for the waved shape steel fiber reinforced concrete. Compared with the available formulations in the state-of-the-art, the presented formulation has an error value of 0.207 percent for MAPE for the waved fiber, which is much lower than the 25 percent, which is about 0.00828 lower [4]. Regarding the hooked end fibers, the reported error in this study is 0.0124 times lower than the 35 percent available in [5]. Another comparison for the hooked end fiber can be made by [6], where the MAPE value of 10 percent was reported, and similarly, the 10 percent of error is possible by the formulation presented in [6]. Also, the formulation shown in [7] for the hooked end fiber had a MAPE error of 30 % for the compressive strength. For the flat fiber, the MAPE error reached 0.219 %, which is acceptable compared with the available empirical formulation error analysis [3]. It must be noted that the maximum value of ME error for the compressive strength is observed for the flat type fiber, equal to 310.582 MPa. As illustrated in Fig. 7, the maximum observed error from all types is the ME for the flat fiber. The MAPE errors are already less than 1, which is approximately on the scale of 0.01 of the available formulations presented in [3-7]. Similarly, according to Fig. 8, the maximum MAPE error for the flexural strength is 0.17838 % which is approximately negligible compared with the empirical formulations presented in the state-of-the-art.

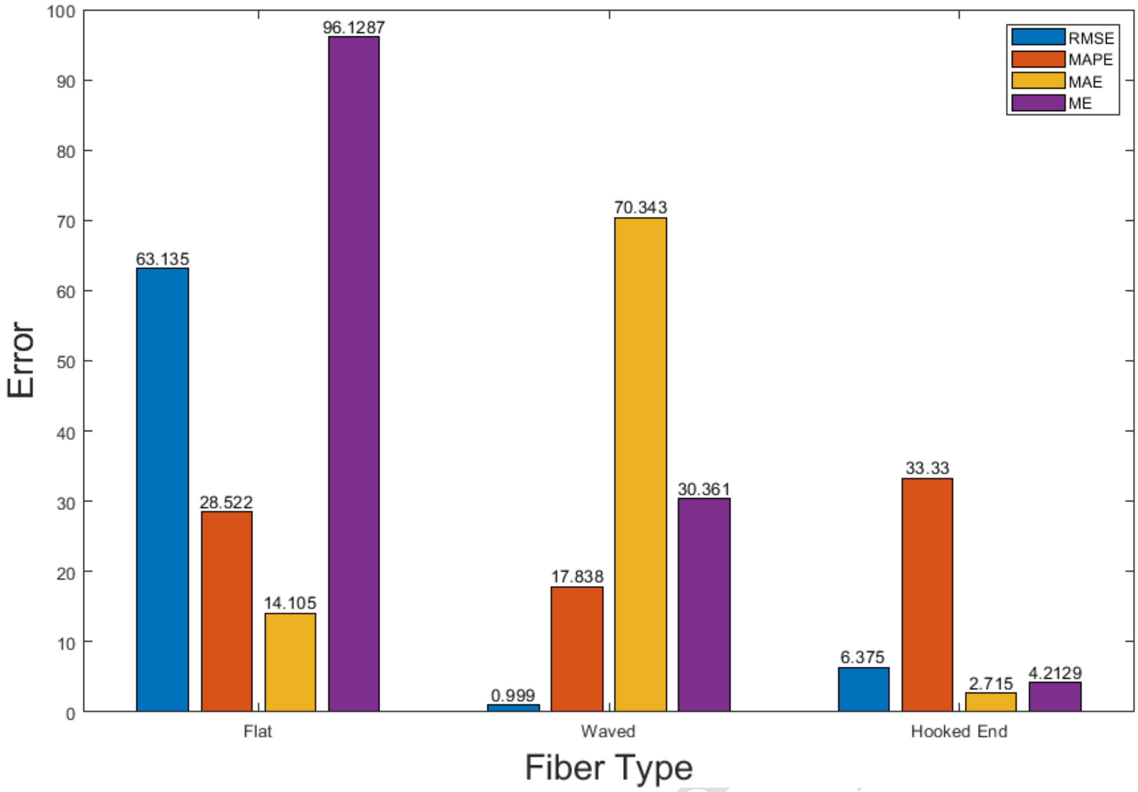


Fig. 7. Error analysis presented for the tensile strength of SFRC.

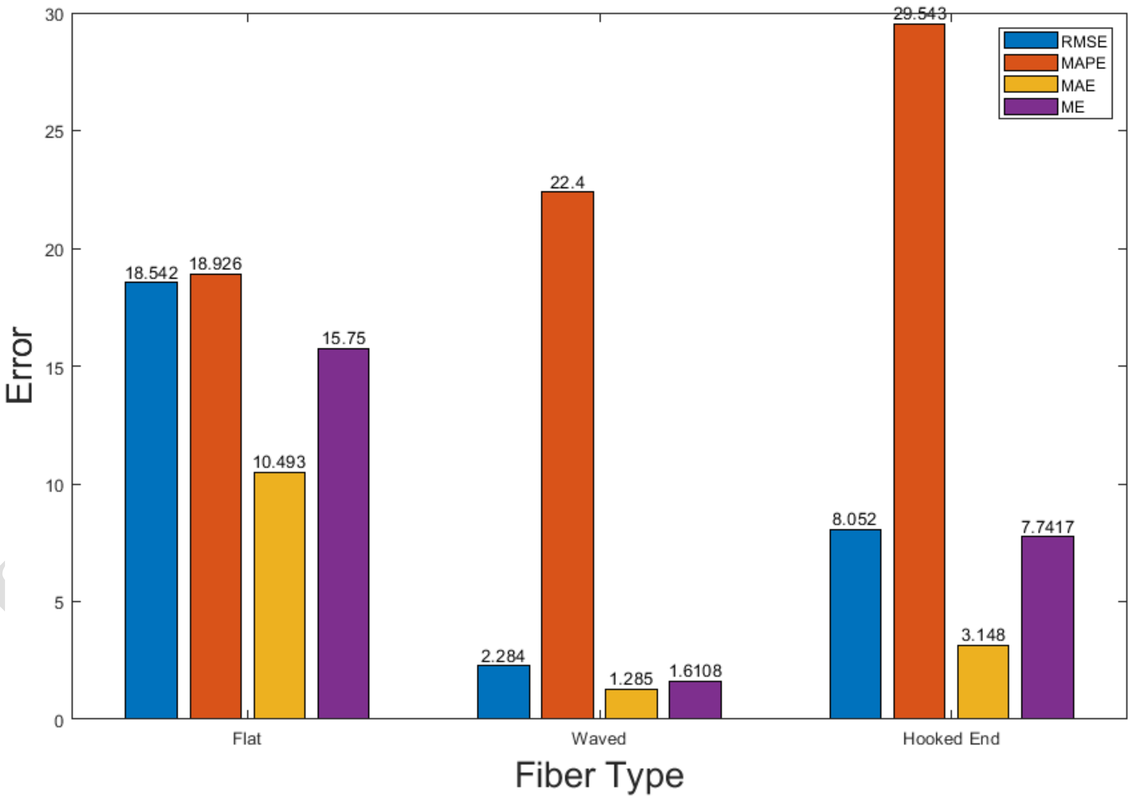
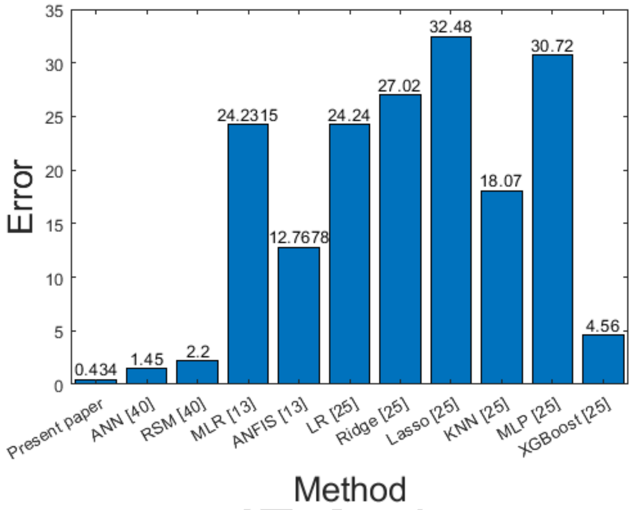


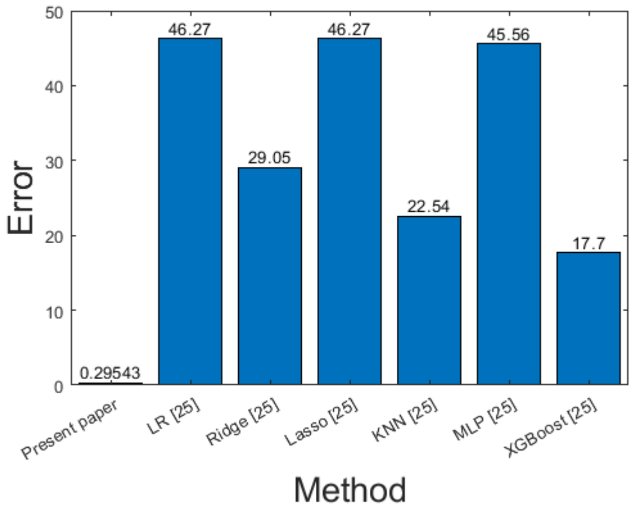
Fig. 8. The error analysis for the flexural strength.

The benefit of using the discussed strategy in section 2 compared with the recent research is in Fig. 9 for the compressive and flexural strength. It is visible that the percent of MAPE error for this research is considerably more minor than others. To compare based on MAE and RMSE errors, the reference [41] reached the corresponding values of 4.74,7.12 MPa for

Random Forest, 2.91,3.26 MPa with Gradient Boosting, and 2.76,3.10 MPa with extreme Gradient Boosting in the fifth folding stage. In the present paper, these items are equal to 8.052 and 3.148 correspondingly, which is approximately close to the mentioned recent research while considering a much bigger dataset. Please note that the values of these errors increase with the size of the considered dataset, so the method is still competitive. Another recent research [42] for predicting the compressive strength, which is comparable with the present study, reached its minimum values errors for the MAE and RMSE errors using supported vector regression (SVR) AdaBoost and SVR bagging with values equal to 4.4, 7.6 MPa, which is lower than the present research due to smaller dataset but still keeping close.



(a)



(b)

Fig. 9. Comparing errors from the present paper with recent machine learning implementations (a) Compressive strength based on MAPE error (b) Flexural strength.

4. Conclusions

This paper presented novel empirical formulations for the strength properties of SFRC using international data from state-of-the-art and well-known machine learning techniques. An error analysis is performed to evaluate the regression. Some key features of this article are as follows:

- A vast amount of data is collected from all available worldwide resources containing 2650 data sets to present the most accurate formulations compared to recent research
- The Ridge, Lasso, and Linear regression forecast the considered objective function. The influence of each feature is investigated for these methods and divided by fiber type based on the effectiveness factor
- It is observed that the influence factors for the three regression methods are approximately similar in the sign, which presents the accomplishment of data preprocessing and feature selection techniques
- The formulations are derived with the symbolic regression technique, proven more efficient than polynomials and other available techniques
- The error analysis is performed based on the four aggregative formulations containing RMSE, MAE, ME, and MAPE
- According to the presented error analysis, the MAPE error values are about 0.01 of the available empirical formulations in the state-of-the-art, presenting the effectiveness of machine learning-driven formulations and the feature selection strategy.

Funding

Any grant does not support this research.

Acknowledgment

There is no acknowledgment in this paper.

Conflict of Interest

The authors of this paper declare that they have no known competing financial interests in this article.

References

[1] Niutta, C. B., Wehrle, E. J., Duddeck, F., & Belingardi, G. (2018). Surrogate modeling in design optimization of structures with discontinuous responses. Structural and Multidisciplinary Optimization, 57(5), 1857-1869. https://doi.org/10.1007/s00158-018-1958-7

[2] Wafa FF, Ashour SA. Mechanical properties of high strength fiber reinforced concrete. ACI Mater J 89(5):449–455. 1992.

[3] Khaloo, A. R., & Kim, N. (1996). Mechanical properties of normal to high-strength steel fiber-reinforced concrete. Cement, Concrete and Aggregates, 18(2), 92-97. https://doi.org/10.1520/CCA10156J

[4] Nataraja, M. C., Dhang, N., & Gupta, A. P. (1999). Stress–strain curves for steel-fiber reinforced concrete under compression. Cement and concrete composites, 21(5-6), 383-390. https://doi.org/10.1016/S0958-9465(99)00021-9

[5] Song, P. S., & Hwang, S. (2004). Mechanical properties of high-strength steel fiber-reinforced concrete. Construction and Building Materials, 18(9), 669-673. https://doi.org/10.1016/j.conbuildmat.2004.04.027

[6] Yazıcı, Ş., İnan, G., & Tabak, V. (2007). Effect of aspect ratio and volume fraction of steel fiber on the mechanical properties of SFRC. Construction and Building Materials, 21(6), 1250-1253. https://doi.org/10.1016/j.conbuildmat.2006.05.025

[7] Thomas, J., & Ramaswamy, A. (2007). Mechanical properties of steel fiber-reinforced concrete. Journal of materials in civil engineering, 19(5), 385-392. https://doi.org/10.1061/(ASCE)0899-1561(2007)19:5(385)

[8] Ou, Y. C., Tsai, M. S., Liu, K. Y., & Chang, K. C. (2012). Compressive behavior of steel-fiber-reinforced concrete with a high reinforcing index. Journal of Materials in Civil Engineering, 24(2), 207-215. https://doi.org/10.1061/(ASCE)MT.1943-5533.0000372

[9] Adhikary, B. B., & Mutsuyoshi, H. (2006). Prediction of shear strength of steel fiber RC beams using neural networks. Construction and Building Materials, 20(9), 801-811. https://doi.org/10.1016/j.conbuildmat.2005.01.047

Journal Pre-proof

[10] Wang, H. T., & Wang, L. C. (2013). Experimental study on static and dynamic mechanical properties of steel fiber reinforced lightweight aggregate concrete. Construction and Building Materials, 38, 1146-1151. https://doi.org/10.1016/j.conbuildmat.2012.09.016

[11] Alilou, V. K., & Teshnehlab, M. (2010). Prediction of 28-day compressive strength of concrete on the third day using artificial neural networks. International journal of Engineering, 3(6), 565-576.

[12] Chopra, P., Sharma, R. K., & Kumar, M. (2016). Prediction of compressive strength of concrete using artificial neural network and genetic programming. Advances in Materials Science and Engineering, 2016. http://dx.doi.org/10.1155/2016/7648467

[13] Khademi, A., Behfarnia, K., Kalman Šipoš, T., & Miličević, I. (2021). The use of machine learning models in estimating the compressive strength of recycled brick aggregate concrete. Computational Engineering and Physical Modeling, 4(4), 1-25. https://doi.org/10.22115/CEPM.2021.297016.1181

[14] Abambres, M., & Lantsoght, E. O. (2019). ANN-based shear capacity of steel fiber-reinforced concrete beams without stirrups. Fibers, 7(10), https://doi.org/88. 10.1007/s00521-016-2417-2

[15] Ahn, N., Jang, H., & Park, D. K. (2007). Presumption of shear strength of steel fiber reinforced concrete beam using artificial neural network model. Journal of applied polymer science, 103(4), 2351-2358. https://doi.org/10.1002/app.25121

[16] Abbas, Y. M., & Iqbal Khan, M. (2016). Influence of fiber properties on shear failure of steel fiber reinforced beams without web reinforcement: ANN modeling. Latin American Journal of Solids and Structures, 13, 1483-1498. http://dx.doi.org/10.1590/1679-78252851

[17] Adhikary, B. B., & Mutsuyoshi, H. (2006). Prediction of shear strength of steel fiber RC beams using neural networks. Construction and Building Materials, 20(9), 801-811. https://doi.org/10.1016/j.conbuildmat.2005.01.04

[18] Açikgenç, M., Ulaş, M., & Alyamaç, K. E. (2015). Using an artificial neural network to predict mix compositions of steel fiber-reinforced concrete. Arabian Journal for Science and Engineering, 40(2), 407-419. https://doi.org/10.1007/s13369-014-1549-x

[19] Shende, A. M., Pande, A. M., & Manchalwar, S. (2018, September). Prediction of flexural Strength of SFRC form Concrete Strength without fibres. In IOP Conference Series: Materials Science and Engineering (Vol. 410, No. 1, p. 012023). IOP Publishing. https://doi.org/10.1088/1757-899X/410/1/012023

[20] Kabir, A., Hasan, M., & Miah, K. (2012, October). Predicting 28 days compressive strength of concrete from 7 days test result. In Proceedings of the International Conference on Advances in Design and Construction of Structures (pp. 18-22). https://doi.org/02.ADCS.2012.1

[21] Sharifi, Y., & Moghbeli, A. (2019). Stepwise Regression for shear capacity assessment of steel fiber reinforced concrete beams. Journal of Rehabilitation in Civil Engineering, 7(2), 152-163. https://doi.org/10.22075/JRCE.2018.1521.1244

[22] Kara, I. F. (2013). Empirical modeling of shear strength of steel fiber reinforced concrete beams by gene expression programming. Neural Computing and Applications, 23(3), 823-834. https://doi.org/10.1007/s00521-012-0999-x

[23] Sharifi, Y., & Moghbeli, A. (2020). New predictive models via gene expression programming and multiple nonlinear regression for SFRC beams. Journal of Materials Research and Technology, 9(6), 14294-14306. https://doi.org/10.1016/j.jmrt.2020.10.026

[24] Al-Musawi, A. A., Alwanas, A. A., Salih, S. Q., Ali, Z. H., Tran, M. T., & Yaseen, Z. M. (2020). Shear strength of SFRCB without stirrups simulation: implementation of hybrid artificial intelligence model. Engineering with Computers, 36(1), 1-11. https://doi.org/10.1007/s00366-018-0681-8

[25] Kang, M. C., Yoo, D. Y., & Gupta, R. (2021). Machine learning-based prediction for compressive and flexural strengths of steel fiber-reinforced concrete. Construction and Building Materials, 266, 121117. https://doi.org/10.1016/j.conbuildmat.2020.121117

[26] Keshtegar, B., Bagheri, M., & Yaseen, Z. M. (2019). Shear strength of steel fiber-unconfined reinforced concrete beam simulation: Application of novel intelligent model. Composite Structures, 212, 230-242. https://doi.org/10.1016/j.compstruct.2019.01.004

[27] Sarveghadi, M., Gandomi, A. H., Bolandi, H., & Alavi, A. H. (2019). Development of prediction

Journal Pre-proof

models for shear strength of SFRCB using a machine learning approach. Neural Computing and Applications, 31(7), 2085-2094. https://doi.org/10.1007/s00521-015-1997-6

[28] Ahmadi, M., Kheyroddin, A., Dalvand, A., & Kioumarsi, M. (2020). New empirical approach for determining nominal shear capacity of steel fiber reinforced concrete beams. Construction and Building Materials, 234, 117293. https://doi.org/10.1016/j.conbuildmat.2019.117293

[29] Taghavi Parsa, Adlparvar, M. R., M. H. Esmaeili, M. (2022). Machine Learning and Feature Selection Techniques for Predicting the Strength Properties of Steel Fiber Reinforced Concrete. Submitted to Engineering Structures.

[30] Taghavi Parsa, Adlparvar, M. R., M. H. Esmaeili, M. (2022). Strength properties of Fiber Reinforced Concrete Including Steel Fibers. Accepted Manuscript (WJE-12-2021-0680.R1), World Journal of Engineering.

[31] Adlparvar, M. R., & Taghavi Parsa, M. H. (2021). The Improvement of the Tensile Behavior of CFRP and GFRP Laminates at Elevated Temperatures Using Fire Protection Mortar. Journal of Rehabilitation in Civil Engineering, 9(2), 41-54.

[32] Nematzadeh, Z., Ibrahim, R., & Selamat, A. (2015, May). Comparative studies on breast cancer classifications with k-fold cross validations using machine learning techniques. In 2015 10th Asian control conference (ASCC) (pp. 1-6). IEEE. https://doi.org/10.1109/ASCC.2015.7244654

[33] Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., & Liu, H. (2017). Feature selection: A data perspective. ACM computing surveys (CSUR), 50(6), 1-45. https://doi.org/10.1145/3136625

[34] Romero, E., Sopena, J. M., Navarrete, G., & Alquézar, R. (2003, July). Feature selection forcing overtraining may help to improve performance. In Proceedings of the International Joint Conference on Neural Networks, 2003. (Vol. 3, pp. 2181-2186). IEEE. https://doi.org/10.1109/INTELCIS.2017.8260032

[35] Billard, L., & Diday, E. (2002). Symbolic regression analysis. In Classification, clustering, and data analysis (pp. 281-288). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-56181-8_31

[36] Weisberg, S. (2005). Applied linear regression (Vol. 528). John Wiley & Sons. https://doi.org/10.1002/0471704091

[37] McDonald, G. C. (2009). Ridge regression. Wiley Interdisciplinary Reviews: Computational Statistics, 1(1), 93-100. https://doi.org/10.1002/wics.14

[38] Ranstam, J., & Cook, J. A. (2018). LASSO regression. Journal of British Surgery, 105(10), 1348-1348. https://doi.org/10.1002/bjs.10895

[39] Kotulski, Z. A., & Szczepinski, W. (2009). Error analysis with applications in engineering (Vol. 169). Springer Science & Business Media. https://doi.org/10.1007/978-90-481-3570-7

[40] Shende, A. M., Yadav, K. P., Bhab,P.P., Pande, A. M. (2020). Prediction of steel fibre reinforced concrete (sfrc) strength using artificial neural network (ANN) models, response surface methodology (RSM) models and their comparative study. European Journal of Molecular and Clinical Medicine, 7(8), 306-314.

[41] Zheng, D., Wu, R., Sufian, M., Kahla, N. B., Atig, M., Deifalla, A. F., ... & Azab, M. (2022). Flexural strength prediction of steel fiber-reinforced concrete using artificial intelligence. Materials, 15(15), 5194. https://doi.org/10.3390/ma15155194

[42] Li, Y., Zhang, Q., Kamiński, P., Deifalla, A. F., Sufian, M., Dyczko, A., ... & Atig, M. (2022). Compressive Strength of Steel Fiber-Reinforced Concrete Employing Supervised Machine Learning Techniques. Materials, 15(12), 4209. https://doi.org/10.3390/ma15124209

Journal Pre-proof