Article

Estimating the Bond Strength of FRP Bars Using a Hybrid Machine Learning Model

Ran Li, Lulu Liu and Ming Cheng

1 School of Architecture and Civil Engineering, Zhengzhou University of Industrial Technology, Zhengzhou 451100, China
2 School of Civil Engineering, Central South University, Changsha 410075, China
3 China Construction Fifth Engineering Division Corp., Ltd., Changsha 410000, China
* Correspondence: ydscivil@163.com

Abstract: Although the use of fiber-reinforced plastic (FRP) rebars instead of mild steel can effectively avoid rebar corrosion, the bonding performance gets weakened. To accurately estimate the bond strength of FRP bars, this paper proposes a particle swarm optimization-based extreme learning machine model based on 222 samples. The model used six variables including the bar position (P), bar surface condition (S), bar diameter (D), concrete compressive strength (f), the ratio of the bar depth to the bar diameter (L/D), and the ratio of the concrete protective layer thickness to the bar diameter (C/D) as input features, and the relative importance of the input parameters was quantified using a sensitivity analysis. The results showed that the proposed model can effectively and accurately estimate the bond strength of the FRP bar with \( R^2 = 0.945 \) compared with the \( R^2 = 0.926 \) of the original ELM model, which shows that the model can be used as an auxiliary tool for the bond performance analysis of FRP bars. The results of the sensitivity analysis indicate that the parameter L/D is of the greatest importance to the output bond strength.

Keywords: FRP; bond strength; ELM; hybrid model; parameter importance analysis

1. Introduction

A reinforced concrete structure is an organic combination of steel and concrete materials that is widely used in buildings, roads, bridges, and marine engineering. It has the excellent compressive properties of concrete and embedding steel bars in the concrete can well compensate for the tensile properties of the concrete structure. However, corrosion is a significant issue that impairs the functionality of structures made of reinforced concrete [1]. This is because corrosion of the reinforcement may lead to serious degradation of the performance of the structure [2]. For this reason, to solve the great harm brought by reinforcement corrosion to civil engineering structures, many scholars have conducted a lot of research on the theory and prevention technologies for reinforcement corrosion. The search for non-conductive steel bars that can replace steel bars, such as fiber-reinforced polymer (FRP) bars, has been consistently identified as a good way to get at the root of the corrosion problem [3]. Due to its corrosion resistance, light weight, high strength-to-weight ratio, cost-effectiveness, ease of installation, fatigue resistance, and minimal creep deformation, the FRP bar has emerged as the most appealing alternative to the usual reinforcement inserted in reinforced concrete structures [4,5].

In FRP-reinforced concrete structures, the bond performance between the two is the key factor to determining whether they can perform mechanically and bear the external load together. Similar to ordinary reinforced concrete, the bonding action between an FRP reinforcement and concrete consists of three main components—namely, the chemical bonding force, friction force, and mechanical bite force [6,7]. The interfacial damage of
ordinary steel-mixed structures mainly occurs in concrete, specifically in the form of concrete shear damage between ribs, while the interfacial bond strength of FRP-reinforced concrete mainly depends on the interlaminar shear strength of the FRP reinforcement. Throughout numerous small-scale experiments, the interfacial connection has been shown to degrade drastically under adverse environmental conditions. There are many factors affecting the interfacial bond strength of FRP bars and concrete, such as the FRP bar type, diameter, surface form, and concrete strength [8]. Despite the increasing interest in FRP reinforcement as a reliable and viable alternative to corrosion-resistant steel, its bonding to concrete is poor. Therefore, it is crucial to accurately estimate and anticipate the bond strength of FRP reinforcement when building solid concrete structures. The primary elements influencing the bond performance of FRP bars have been the subject of several investigations over the years, mostly by direct pullout testing or beam tests [9–11]. For example, Alves et al. [12] investigated the bond strength of glass fiber-reinforced polymer (GFRP) reinforcement bars of different diameters and the results showed that the bond strength was weaker for larger diameter GFRP bars. There are also some research results showing that some key parameters such as a concrete protective layer, the bar surface condition, bar diameter, buried length, location of reinforcement, and the compressive strength of the concrete have different degrees of influence on the bond strength [13]. Researchers have created empirical models for calculating the bond strength of FRP bars based on experimental data, and certain studies have been created and included in pertinent design codes based on theoretical analyses and experimental validation [14]. However, most of these models utilize a limited set of experimental data, which makes the models accurate within these data spaces but lacking sufficient generalization to other parameter settings. Additionally, the effectiveness of the model is decreased by the need for several assumptions during the theoretical derivation of these constrained empirical models to represent the complicated nonlinear relationship between bond strength and important parameters [15]. Therefore, it is essential to investigate a reliable and effective technique for calculating the bond strength of FRP bars [16].

Data-driven methods based on machine learning algorithms have arisen as an alternative method to creating prediction models using integrated experimental data and information as a result of the advancement of computer science and the growth of relevant experimental datasets. Some intelligent algorithms such as artificial neural networks (ANN), support vector machines (SVM), multiple linear regression (MLR), random forests (RF), integrated learning (gradient augmented regression trees), etc., have been used for the prediction of FRP bond strength [17–20]. Su et al. [21] selected three machine learning algorithms — MLR, ANN, and SVM — to predict the interfacial bond strength between FRP and concrete, and the results demonstrated that the machine learning models could predict the strength accurately and effectively, with the SVM achieving the best performance among the three models. Koroglu et al. proposed a regression- and ANN-based model to estimate the bond strength of FRP bars in concrete [22]. The findings indicated that the ANN model performed better than both the Canadian Standards Association (CSA) and American Concrete Institute (ACI) models in estimating the binding strength of FRP bars in concrete. Hamed et al. [23] developed a multi-gene genetic programming (MGGP) model based on 223 records to predict the bond strength between concrete and FRP, and the model predicted the strength better than the ACI model. Sherin et al. [24] utilized ANN to predict the bond strength between self-compacting geopolymer concrete reinforced with basalt FRP bars, achieving more accurate prediction results than existing theoretical and analytical models. To compare the performance of different optimization algorithms in bond strength prediction, three optimizers were used for ANN model optimization by Mohammad et al. [25]. The RUNge Kutta optimizer (RUN)-based hybrid RUN-ANN model achieved the largest prediction correlation with $R^2 = 92\%$, and its prediction performance outperformed the mechanics-based method. To study the bond strength at high temperatures, Muhammad et al. [26] introduced the gene-expression programming model and established a traceable mathematical formula based on the trained
model for easy use. Zhou et al. [27] established an ANN model using the back-propagation neural network (BPNN) method on a large dataset with 969 data observations, and its prediction accuracy was higher than other methods in the literature. Similarly, an explicit equation was established based on the trained ANN model.

Prediction methods based on machine learning models have been successfully applied in the concrete strength estimation [28–30] and bond strength estimation of FRP reinforcement [31,32]. However, most of these methods are primitive models and the improvement of model prediction performance is limited by the selection of hyper-parameters, which can be solved using optimization algorithms. Considering the excellent performance of the extreme learning machine (ELM) model in regression analyses [33] and the fact that this method is rarely reported in bond strength prediction studies, this paper attempts to develop a hybrid model for the accurate prediction of bond strength using the ELM model as the original model and compares the prediction performance of the developed hybrid model with that of the original model.

2. Methodology

2.1. Extreme Learning Machine Network

The extreme learning machine was proposed by Huang et al. [34] based on the study of single-hidden layer feedforward neural networks. The connection weights between the input layer and the hidden layer and the threshold values of the neurons in the hidden layer are generated at random by the algorithm, and they do not need to be changed during the training process to reach a specific optimal solution [35]. Let the connection weights \( w \) be between the implicit layers of the input layer and the connection weights \( \beta \) be between the implicit layer and the output layer. Let the input matrix \( X \) and the output matrix \( Y \) of the training set have \( Q \) samples, where the output \( T \) of the network is:

\[
T = \begin{bmatrix}
\sum_{i=1}^{I} \beta_{1i}g(w_{i}x_{j} + b_{i}) \\
\sum_{i=1}^{I} \beta_{2i}g(w_{i}x_{j} + b_{i}) \\
\vdots \\
\sum_{i=1}^{I} \beta_{mi}g(w_{i}x_{j} + b_{i})
\end{bmatrix}, \quad j = 1, 2, \ldots, Q
\]  

Formula (1) can be expressed as:

\[
H \beta = T'
\]  

where \( T' \) is the transpose of the matrix \( T \), \( H \) is the output matrix of the hidden layer of the neural network, and \( \beta \) is the connection weights between the hidden layer and the output layer, which can be obtained by solving the least squares solution of Equation (3) as follows:

\[
\min_{\beta} \| H\beta - T' \|
\]  

The solution can be expressed as:

\[
\hat{\beta} = H'^{-1}T'
\]  

where \( H'^{-1} \) is the generalized inverse of the implied layer output matrix \( H \). The network structure and characteristics of the ELM model determine the short training time of the network itself. The input layer weights and thresholds however, which are produced at
random, set a cap on the model's performance. Given the fast convergence and high accuracy of the particle swarm algorithm [36,37], this paper employed the particle swarm optimization (PSO) algorithm to optimize the input layer weights and thresholds of the network structure to improve the ELM model prediction accuracy.

2.2. PSO-ELM Prediction Model

The PSO-ELM prediction model put forward in this research uses the PSO algorithm to optimize the extreme learning machine’s input layer weights and thresholds before using its preferred input layer weights and thresholds to predict the bond strength. Figure 1 depicts the flow chart for the specific PSO-ELM training procedure.

![Flow Chart](image)

**Figure 1.** The process of the Elm model optimized by the PSO algorithm.

The specific steps of the optimization algorithm are:

1. ELM randomly generates the weights \( w \) and threshold \( b \) of the input layer and determines the implied layer weights \( \beta \) from this set of weights and thresholds.

2. The particle population randomly generates \( n \) particles, and each particle in this population represents a vector of dimension \( D \). The dimension of this particle is determined by the input parameters of the ELM model, and the initial value of the acceleration factor and the maximum number of iterations is set, and the end condition of the experiment is to meet the accuracy requirement or to reach the maximum number of iterations [38].

3. Selection of inertia weights: In order to make the algorithm achieve better global search and local search capabilities, linear decreasing inertia weights are used.
where $k$ is the current iteration generation, $T_{\text{max}}$ is the maximum iteration generation, $w_{\text{start}}$ is the initial inertia weight, and $w_{\text{end}}$ is the inertia weight at the maximum number of iterations.

3. Dataset Collection and Model Building

3.1. Input and Output Variables

For machine learning models, the manually defined feature parameters and the size of the dataset are the key factors that affect the model prediction performance. Referring to and synthesizing the selection of feature parameters in Refs. [8,23], six parameters including the bar position ($P$), bar surface condition ($SC$), bar diameter ($D$), concrete compressive strength ($f_c$), the ratio of the bar depth to the bar diameter ($L/D$), and the ratio of the concrete protective layer thickness to the bar diameter ($C/D$) were selected as the input variables. When training the model, a rich and representative dataset is very important. For this purpose, a total of 222 records collected from previous studies, which are available in the literature [13], were used as the model datasets. The normal distribution curves of these variables are shown in Figure 2. For the rebar location $P$, the numbers 1 and 2 represent the top and bottom positions of the FRP rebar in the beam, respectively. The three numbers in the rebar surface condition indicate three types of indicated conditions. The statistical characteristics of these parameters and the Pearson correlation coefficient between the variable and the output are listed in Table 1. From Table 1, it can be seen that there is very little linear correlation between any of the input parameters and the output bond strength, with the highest correlation coefficient between any two variables not exceeding 0.8. This paper’s main objective, therefore, becomes examining the intricate non-linear relationship between the input and the output.
3.2. Evaluation Metrics

The correlation coefficient $R^2$ and four error metrics—root mean squared error (RMSE), mean absolute percentage error (MAPE), mean absolute error (MAE), and mean squared error (MSE)—were introduced to statistically analyze the prediction results to allow for an intuitive comparison of the performance of the models. The definitions of these evaluation metrics can be found in the literature [40–43]. For the correlation coefficient, an $R^2$ closer to one indicates a stronger correlation between the predicted results and the actual results—and the smaller the error-index is, the better the prediction effect is.

3.3. Model Building

For the single regression objective of bond strength, since the input variables were the six feature parameters, there were six input layer neurons and one output layer neuron in the ELM model. The number of the hidden layer neurons was selected using the trial-and-error method. After the trials, the model prediction reached its best performance when the hidden layer neurons numbered 35. The model structure is shown in Figure 3. Due to the limited number of data sets, this paper used 10-cross-fold validation for training. Referring to the classification method of the dataset in [44], a total of 222 samples were randomly divided into the training set and test set, where 80% of the dataset (178 samples) was used for training and the remaining 20% (44 samples) were used to test the generali-
zation ability of the model. For comparison purposes, the data set will be fixed after random classification. In other words, the training and test sets for the original and optimized models were the same.

![Figure 3. ELM model network structure.](image)

3.4. Comparison of Model Results

The sample prediction results before and after model optimization are shown in Figure 4. Overall, the optimized PSO-ELM model obtained results closer to the true bond strength than the ELM model. A linear fit was carried out for these data, and Figure 5 displays the correlation between the expected and actual values for each sample. In terms of correlation coefficients $R^2$, the correlation coefficients between the predicted and actual strengths of the PSO-ELM model all exceeded 0.94; the correlation coefficients were higher than those of the ELM model, with $R^2 = 0.931$ for the training phase and $R^2 = 0.926$ for the test phase.

![Graph showing comparison of model results](image)
Figure 4. Comparison between predicted bond strength and the actual values of two models.
Figure 5. Scatter plot of the prediction results.

Figure 6 also displays the proportion between the two models’ actual bond strengths and their projected values. The mean value of 1.014 for the PSO-ELM model is closer to one than the mean value of 1.022 for the ELM model. The PSO algorithm’s optimization improved the model’s capability for making predictions. When assessing the effectiveness of a prediction, the relative error is frequently more important than the absolute mistake. The relative prediction error of the models for the training and test phases is shown in Figure 7 and Table 2. The mean value of the relative error for the PSO-ELM model was lower for both the training and testing stages. Table 3 lists the five evaluation metrics introduced in this paper. Compared to the error statistics MSE, RMSE, and MAE of 1.549, 1.244, and 0.883 with the pre-optimization model for the test set, the optimized model achieved the smaller MSE, RMSE, and MAE of 1.170, 1.082, and 0.777, respectively. For the training set, the average relative error was reduced from 14.714% to 11.655%, and the average relative error was reduced from 15.590% to 13.953% for the test set, which further indicates that the PSO-ELM model’s predictability was indeed improved after optimization.
Figure 6. The proportion between the predicted and actual bonding strength.

Table 2. Model prediction results and relative errors of the test set.

<table>
<thead>
<tr>
<th>Test Set</th>
<th>Experimental Value</th>
<th>ELM Model</th>
<th>Relative Error(%)</th>
<th>PSO-ELM Model</th>
<th>Relative Error(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.26</td>
<td>5.80</td>
<td>10.32</td>
<td>5.63</td>
<td>7.02</td>
</tr>
<tr>
<td>2</td>
<td>11.5</td>
<td>12.60</td>
<td>9.56</td>
<td>10.02</td>
<td>12.90</td>
</tr>
<tr>
<td>3</td>
<td>4.29</td>
<td>4.74</td>
<td>10.53</td>
<td>4.52</td>
<td>5.27</td>
</tr>
<tr>
<td>4</td>
<td>4.6</td>
<td>3.72</td>
<td>19.15</td>
<td>4.60</td>
<td>0.06</td>
</tr>
<tr>
<td>5</td>
<td>5.1</td>
<td>6.40</td>
<td>25.48</td>
<td>5.96</td>
<td>16.81</td>
</tr>
<tr>
<td>6</td>
<td>3.3</td>
<td>3.12</td>
<td>5.50</td>
<td>3.02</td>
<td>8.42</td>
</tr>
<tr>
<td>7</td>
<td>2.2</td>
<td>2.51</td>
<td>13.86</td>
<td>3.35</td>
<td>52.11</td>
</tr>
<tr>
<td>8</td>
<td>16</td>
<td>14.53</td>
<td>9.21</td>
<td>15.31</td>
<td>4.29</td>
</tr>
<tr>
<td>9</td>
<td>4.61</td>
<td>4.21</td>
<td>8.61</td>
<td>4.78</td>
<td>3.68</td>
</tr>
<tr>
<td>10</td>
<td>3.3</td>
<td>2.52</td>
<td>23.52</td>
<td>3.18</td>
<td>3.71</td>
</tr>
<tr>
<td>11</td>
<td>7.58</td>
<td>7.03</td>
<td>7.22</td>
<td>6.15</td>
<td>18.85</td>
</tr>
<tr>
<td>12</td>
<td>4.84</td>
<td>5.38</td>
<td>11.14</td>
<td>4.90</td>
<td>1.34</td>
</tr>
<tr>
<td>13</td>
<td>8.9</td>
<td>8.26</td>
<td>7.17</td>
<td>11.19</td>
<td>25.68</td>
</tr>
<tr>
<td>15</td>
<td>6.02</td>
<td>5.80</td>
<td>3.71</td>
<td>5.82</td>
<td>3.39</td>
</tr>
<tr>
<td>16</td>
<td>19.14</td>
<td>16.72</td>
<td>12.64</td>
<td>17.82</td>
<td>6.88</td>
</tr>
<tr>
<td>17</td>
<td>3.74</td>
<td>4.40</td>
<td>17.67</td>
<td>5.64</td>
<td>50.90</td>
</tr>
<tr>
<td>18</td>
<td>13.39</td>
<td>12.03</td>
<td>10.13</td>
<td>11.42</td>
<td>14.70</td>
</tr>
<tr>
<td>19</td>
<td>21</td>
<td>16.72</td>
<td>20.38</td>
<td>17.82</td>
<td>15.13</td>
</tr>
<tr>
<td>20</td>
<td>4.83</td>
<td>5.38</td>
<td>11.45</td>
<td>4.78</td>
<td>1.00</td>
</tr>
<tr>
<td>21</td>
<td>5.4</td>
<td>6.07</td>
<td>12.42</td>
<td>5.06</td>
<td>6.32</td>
</tr>
<tr>
<td>22</td>
<td>3.7</td>
<td>2.24</td>
<td>39.36</td>
<td>2.69</td>
<td>27.35</td>
</tr>
<tr>
<td>23</td>
<td>4.4</td>
<td>2.62</td>
<td>40.38</td>
<td>3.11</td>
<td>29.33</td>
</tr>
<tr>
<td>24</td>
<td>3.6</td>
<td>3.89</td>
<td>8.15</td>
<td>4.64</td>
<td>28.80</td>
</tr>
<tr>
<td>25</td>
<td>3.6</td>
<td>6.88</td>
<td>91.19</td>
<td>3.59</td>
<td>0.17</td>
</tr>
<tr>
<td>26</td>
<td>6.46</td>
<td>6.48</td>
<td>0.26</td>
<td>6.28</td>
<td>2.85</td>
</tr>
<tr>
<td>27</td>
<td>3.6</td>
<td>3.51</td>
<td>2.59</td>
<td>3.77</td>
<td>4.86</td>
</tr>
<tr>
<td>28</td>
<td>5.28</td>
<td>5.38</td>
<td>1.95</td>
<td>4.78</td>
<td>9.43</td>
</tr>
<tr>
<td></td>
<td>ELM training</td>
<td>ELM test</td>
<td>PSO-ELM training</td>
<td>PSO-ELM test</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>--------------</td>
<td>----------</td>
<td>------------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>2.56</td>
<td>3.26</td>
<td>27.39</td>
<td>2.12</td>
<td>17.18</td>
</tr>
<tr>
<td>30</td>
<td>0.97</td>
<td>1.09</td>
<td>12.58</td>
<td>1.54</td>
<td>59.21</td>
</tr>
<tr>
<td>31</td>
<td>7.35</td>
<td>6.27</td>
<td>14.72</td>
<td>5.87</td>
<td>20.08</td>
</tr>
<tr>
<td>32</td>
<td>3.28</td>
<td>2.82</td>
<td>13.99</td>
<td>2.45</td>
<td>25.19</td>
</tr>
<tr>
<td>33</td>
<td>4.98</td>
<td>5.48</td>
<td>10.02</td>
<td>5.18</td>
<td>4.00</td>
</tr>
<tr>
<td>34</td>
<td>1.65</td>
<td>1.09</td>
<td>33.82</td>
<td>1.54</td>
<td>6.40</td>
</tr>
<tr>
<td>35</td>
<td>6.99</td>
<td>6.90</td>
<td>1.23</td>
<td>6.09</td>
<td>12.87</td>
</tr>
<tr>
<td>36</td>
<td>7.54</td>
<td>6.91</td>
<td>8.42</td>
<td>6.91</td>
<td>8.37</td>
</tr>
<tr>
<td>37</td>
<td>7.36</td>
<td>7.81</td>
<td>6.06</td>
<td>7.54</td>
<td>2.40</td>
</tr>
<tr>
<td>38</td>
<td>6.91</td>
<td>5.10</td>
<td>26.16</td>
<td>4.66</td>
<td>32.59</td>
</tr>
<tr>
<td>39</td>
<td>4.64</td>
<td>4.46</td>
<td>3.87</td>
<td>4.90</td>
<td>5.55</td>
</tr>
<tr>
<td>40</td>
<td>4.71</td>
<td>4.74</td>
<td>0.70</td>
<td>4.34</td>
<td>7.77</td>
</tr>
<tr>
<td>41</td>
<td>6</td>
<td>4.96</td>
<td>17.31</td>
<td>5.47</td>
<td>8.88</td>
</tr>
<tr>
<td>42</td>
<td>5</td>
<td>4.46</td>
<td>10.79</td>
<td>4.90</td>
<td>2.05</td>
</tr>
<tr>
<td>43</td>
<td>3.6</td>
<td>2.90</td>
<td>19.51</td>
<td>3.52</td>
<td>2.23</td>
</tr>
<tr>
<td>44</td>
<td>9.2</td>
<td>6.47</td>
<td>29.66</td>
<td>7.03</td>
<td>23.62</td>
</tr>
</tbody>
</table>

(a) Box plots of relative error
4. Sensitivity Analysis of Input Variables

Sensitivity analysis is a tool that helps to measure the contribution and impact of input parameters on output outcomes [45]. Sensitivity analysis was used in this study to determine the relative impact of various input factors on the target. Figure 8 illustrates the relative influence of the six parameters discussed in this research on the output results. It is clear that the parameter $L/D$ had an 88 percent relative importance to the output findings, and that this importance had a detrimental impact on bond strength while the relative importance of the other five parameters to the results was low.
5. Conclusions

To achieve bond strength prediction of an FRP bar, a hybrid-optimized PSO-ELM model was proposed in this paper. The following is a summary of the key findings.

1) With a stronger correlation coefficient and reduced prediction error, the proposed optimization model PSO-ELM can accurately capture the nonlinear relationship between the numerous input factors and the output bond strength.

2) Regarding the criteria for model evaluation, the PSO-ELM achieved a better prediction performance with a smaller MSE, RMSE, and MAE of 1.170, 1.082, and 0.777 compared to the MSE, RMSE, and MAE of 1.549, 1.244, and 0.883 for the pre-optimization ELM model. The mean value of the relative prediction error was reduced from 15.590% in the original ELM model to 13.953% in the PSO-ELM model.

3) Among the six input parameters mentioned in this paper, the parameter bar-embedded length to the bar-diameter ratio \((L/D)\) was the dominant influencing factor, with a relative importance of 88% to the output results; this variable was negative for the bond strength of the FRP bar.

Author Contributions: R.L.: Writing-original draft Software. L.L.: Methodology, Validation. M.C.: Data curation, Resources. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.
References


