Article

Improvement of Computational Efficiency and Accuracy by Firefly Algorithm and Random Forest for Compressive Strength Modeling of Recycled Concrete

Yong Liu 1,2,3, Yang Wang 2,4, Mengmeng Zhou 4,* and Jiandong Huang 5,*

1 China Coal Research Institute, Beijing 100013, China; liuyongjill@163.com
2 CCTEG Coal Mining Research Institute, Beijing 100013, China; graupel87@163.com
3 Tiandi Mining (Yulin) Engineering & Technology Co., Ltd., Yulin 719099, China
4 School of Mines, China University of Mining and Technology, Xuzhou 221116, China
5 School of Civil Engineering, Guangzhou University, Guangzhou 510006, China
* Correspondence: ts21020232p21@cumt.edu.cn (M.Z.); jiandong.huang@hotmail.com (J.H.)

Abstract: It is an important direction for the sustainable development of pavement to mix the discarded concrete blocks with gradation according to a certain proportion after crushing, cleaning and other technological processes, partially or completely replace aggregate, and then add cement, water, and so on to make recycled concrete for pavement paving, but the traditional evaluation model for the compressive strength (CS) of recycled concrete cannot meet the requirements of efficient calculation. To address such issues, the present research proposed to apply the firefly algorithm (FA) to optimize the random forest (RF) model. The results were demonstrated by comparing the consistency of predicted and actual values, and also by analyzing the correlation coefficient (R) and root-mean-square error (RMSE). Higher R values (0.9756 and 0.9328) and lower RMSE values (3.0752 and 6.4369) for the training and test sets present the reliability of the FA and RF hybrid machine learning model. To understand the influence law of input indexes on the output index, the importance and sensitivity of variables are further analyzed. The results displayed that effective water-cement ratio (WC) and nominal maximum recycled concrete aggregate size (NMR) have the greatest impact on the output variable, with importance scores of 2.5947 and 2.4315, respectively, while the change in the recycled concrete aggregate replacement rate (RCA) has a weak influence, with an importance score of 0.4695. Introducing FA to RF for the compressive strength modeling of recycled concrete can significantly improve the computational efficiency and accuracy.

Keywords: recycled concrete; machine learning; firefly algorithm; random forest; compressive strength

1. Introduction

Sustainability, also known as the concept of sustainable development, refers to development that can meet the needs of the present generation without endangering the needs of future generations [1]. With the development of society, sustainable development has become a consensus in various industries. Hence, sustainable development is an important development direction of the construction industry. However, with the advancement of urban construction and the acceleration of urbanization, many old buildings have been demolished and rebuilt, leading to a huge amount of construction waste [2–5]. The volume of construction waste generated worldwide every year will nearly double to 2.2 billion tons by the year 2025, according to Construction & Demolition Recycling. The treatment of construction waste needs to consume a lot of manpower and material resources, and the improper treatment of construction waste will cause serious environmental pollution and waste of aggregate resources. At the same time, the exploitation of natural aggregate also leads to the destruction of forest land, environmental damage, and serious harm to the ecological balance. With the concept of sustainable development, researchers proposed to
make construction waste into recycled aggregate, recycled cement, and recycled sand to prepare recycled concrete [6]. The recycling of construction waste can solve not only the growing shortage of cement, natural aggregate, and natural sand but also the treatment cost of waste concrete, environmental pollution, and other problems [7–9], which is conducive to the sustainable improvement of the city and the construction industry [10–13]. Currently, the old cement pavement overhaul, reconstruction, and maintenance management generate a large amount of waste cement and broken concrete, causing a certain burden on the environment. The use of waste cement concrete panels in the construction of new pavement bases or surface layers can not only reduce the mining of new aggregates, save natural resources, and reduce environmental pollution, but also eliminate the environmental damage and arable land occupation caused by waste aggregate accumulation [14], which has significant economic and social benefits. The modification process is shown in Figure 1. With the demolition of old buildings and the construction of new ones, a series of issues concerning natural energy, resources, sustainable development, and environmental protection have been brought about. It is mainly manifested in two aspects: First, the demand for concrete aggregate is increasing, resulting in the large-scale exploitation of resources, consuming a lot of human and material resources. Secondly, with the establishment of new buildings and the demolition of old buildings, a large amount of construction waste will be generated, which is bound to cause great pollution to the environment and occupy a large amount of landfill land. To maximize the recycling of construction waste, the exploitation of natural resources must be minimized, which is the basis of sustainable development of the construction industry.

Figure 1. Sustainable road construction by the regeneration process [15].
In recent years, many researchers have conducted a series of research studies on the performance of green concrete, contributing an undeniable force to the application of recycled concrete. Jayakody et al. [16] researched the potential connection between the strength of concrete made of recycled aggregate over time, and the research results indicated that the aggregate replacement rate (RCA) for coarse aggregate reached 40% and the recycled concrete met the strength requirements of traditional low- and medium-strength concrete. To ensure the feasibility of recycled concrete application in the harsh environment, Koga et al. [17] studied the test method of freeze–thaw durability of recycled concrete aggregate, and the specifications of precast concrete of recycled aggregate that can be used in the freeze–thaw environment with deicing salt were described. Adesina et al. [18] analyzed the effect of utilizing recycled concrete fine particles as an engineering cement composite (ECC) material on its performance. The research results showed that replacing silica sand in ECC with recycled concrete would increase its drying shrinkage and improve its chloride ion permeability, permeable porosity, and water absorption to some extent. Hence, the application of recycled concrete as an aggregate in ECC is conducive to the sustainable improvement and progress of the construction industry. Jaawani et al. [19] studied the impact of recycled asphalt pavement (RAP) on the flexural strength, CS, durability, and other performance indexes of concrete to determine the possibility of RAP being applied in structural concrete.

The construction of highways consumes a lot of resources such as natural sand and stone which are difficult to regenerate [20–22]. The appropriate amounts of water and cement are added to the graded crushed stone, followed by mixing, compaction, and maintenance to produce the cement-stabilized macadam base, which is the main material source for the semi-rigid base and also the most widely used type of highway base at present. The specification of the compressive strength of cement-stabilized macadam base is not very high: the general requirement is higher than 5 Mpa; based on the development law of unconfined CS, the compressive modulus of resilience and splitting strength of recycled concrete and cement-stabilized macadam base are similarly recycled, so it is feasible to use recycled concrete instead of cement-stabilized macadam base. CS reflects the ultimate strength, the maximum stress when the object is damaged under the action of the external force of an object under external pressure, which reflects the road performance of recycled concrete [23–25]. Recently, more and more scholars are paying attention to the evaluation of the CS of recycled concrete. Xie et al. [24] explored the influence rule of metakaolin on the CS of recycled concrete employing the cube CS test and a scanning electron microscope, and the results showed that when the aggregate replacement rate was 30% and 50%, and the content of nano-metakaolin was 5%, the CS of recycled concrete was the maximum. Yang et al. [26] analyzed the influence of the mixing process on the CS and splitting tensile strength of regenerated concrete based on laboratory tests and t-tests, and proved that the quality of regenerated and multiply regenerated concrete with vibration mixing is better than that of first-generation concrete. Kim Gwang Hee et al. [27] analyzed the CS of recycled concrete based on data sets from existing articles and artificial neural networks (ANN) and verified the prediction accuracy of ANN by comparing the result of ANN with the maturity of the evaluation methods for the CS of recycled concrete. Naderpour et al. [28] established the CS prediction model of recycled concrete using ANN, and the research results showed that the use of ANN to analyze the CS of recycled concrete has a positive impact on improving the prediction accuracy. Zhu et al. [29] chose to establish the evaluation model based on genetic algorithm analysis, applied the model to evaluate the CS of recycled concrete, and verified the accuracy using data from the published articles. However, most of the existing studies on the CS of recycled concrete use laboratory experimental methods and single machine learning models, among which laboratory experimental methods have limitations such as long time and high cost, and the evaluation accuracy of single machine learning models has difficulty achieving satisfactory results. Thus, it is necessary to develop more accurate, easier, and cheaper hybrid machine learning models to predict the CS of recycled concrete [30,31].
The research on the CS of recycled concrete has significant implications for the green and sustainable development of concrete. Although some researchers have applied machine learning models to the study of CS in recycled concrete, the prediction accuracy of a single machine learning model cannot achieve satisfactory results. Based on MATLAB (R2021b) software, this study proposes to introduce the firefly algorithm (FA) to the random forest (RF) for the CS modeling of recycled concrete to improve prediction accuracy. Firstly, based on the influencing factors of the CS of recycled concrete and published literature, this study selected the input indexes for evaluating the CS of recycled concrete and determined the database. The feasibility of the database and indexes was verified through mathematical statistical analysis and correlation analysis. Then 10-fold cross-validation (10-fold CV) was used to analyze the effect of FA on the hyperparameters optimization of RF. Finally, the prediction accuracy of the FA and RF hybrid machine learning model for the CS of recycled concrete was studied based on consistency analysis and error analysis, and the importance and sensitivity of input variables on the output variable were analyzed to evaluate the change rule of the output index with the change in input indexes.

2. Methodology

2.1. Data Analysis

Researchers often focus on the research of prediction models, but they ignore the important role of databases in model validation. A reliable database is the basis for evaluating the effectiveness of the studied models [32,33]. This study collected relevant data from previously published literature and formed a database with input variables including effective water-cement ratio (WC), aggregate-cement ratio (AC), and RCA; among them, the selection of RCA is mainly to study the influence of recycled aggregate on the performance of recycled concrete and then study the feasibility of pavement recycling technology. Six input variables were also selected to describe the physical properties of aggregates, including aggregate size (nominal maximum RCA size (NMR), nominal maximum NA size (NMN)), density (bulk density of RCA (BDR), bulk density of NA (BDN)), and water absorption (water absorption of RCA (WAR), water absorption of NA (WAN)), and the output variable including the compressive strength only [34–71]. Table 1 performs the mathematical statistical analysis, and Figure 2 shows the distribution of ten parameters, respectively. It is obvious that the coverage of all variables is extensive and the difference between the median and mean of the input variables and the output variable is small, so the frequency distribution histogram of compressive strength is a single peak type with reasonable data distribution. The results indicate the rationality of the database collected in this study.

Table 1. Analysis of variables.

<table>
<thead>
<tr>
<th>Indexes</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Median</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>WC</td>
<td>0.29</td>
<td>0.72</td>
<td>0.45</td>
<td>0.472</td>
<td>0.088</td>
<td>0.008</td>
</tr>
<tr>
<td>AC</td>
<td>1.7</td>
<td>6.4</td>
<td>2.8</td>
<td>2.878</td>
<td>0.913</td>
<td>0.833</td>
</tr>
<tr>
<td>RCA (%)</td>
<td>0</td>
<td>90</td>
<td>45</td>
<td>39.623</td>
<td>16.936</td>
<td>286.83</td>
</tr>
<tr>
<td>NMR (mm)</td>
<td>8</td>
<td>32</td>
<td>20</td>
<td>20.591</td>
<td>5.094</td>
<td>25.953</td>
</tr>
<tr>
<td>NMN (mm)</td>
<td>10</td>
<td>32</td>
<td>20</td>
<td>22.047</td>
<td>4.991</td>
<td>24.909</td>
</tr>
<tr>
<td>BDR (Kg/m³)</td>
<td>2010</td>
<td>2880</td>
<td>2309</td>
<td>2393.084</td>
<td>124.373</td>
<td>15,468.73</td>
</tr>
<tr>
<td>BDN (Kg/m³)</td>
<td>2510</td>
<td>2970</td>
<td>2610</td>
<td>2638.591</td>
<td>81.93</td>
<td>6712.567</td>
</tr>
<tr>
<td>WAR (%)</td>
<td>1.5</td>
<td>10.9</td>
<td>5</td>
<td>5.341</td>
<td>1.955</td>
<td>3.823</td>
</tr>
<tr>
<td>WAN (%)</td>
<td>0.2</td>
<td>3</td>
<td>1.1</td>
<td>1.163</td>
<td>0.691</td>
<td>0.477</td>
</tr>
<tr>
<td>CS (mPa)</td>
<td>14.2</td>
<td>108</td>
<td>44.9</td>
<td>45.677</td>
<td>14.519</td>
<td>210.804</td>
</tr>
</tbody>
</table>
(a) WC

(b) AC

(c) RCA (%)

(d) NMR (mm)

(e) NMR (mm)

(f) BDR (Kg/m³)

(g) BDN (Kg/m³)

(h) WAR (%)
Correlation analysis is an important analytical method for studying the potential connections between variables. The correlation between variables includes three types: positive correlation, negative correlation, and uncorrelation [72,73]. Excessive positive and negative correlations between input variables can lead to multicollinearity between variables, which affects the accuracy of model evaluation. Hence, evaluating the correlation between input parameters is an indispensable step to ensure the reliability of the selection of input parameters. In this study, the Pearson correlation coefficient was selected to calculate the correlation coefficient between input variables, and the calculation formula for the Pearson correlation coefficient is as Formula (1):

$$ \rho_{XY} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} $$  \hspace{1cm} (1) 

Figure 3 displays the results of the analysis. Excessive positive and negative correlations between indexes will harm the interpretation of the output index. Obviously, the correlations between BDN and RCA, WAR and BDR, and WAN and BDN are high, and the correlation coefficients are −0.744, 0.752, and 0.688, respectively. It is worth noting that the correlations between WAN and NMR and between WAN and AC are very small, and the correlation coefficients are −0.002 and 0.003, which are close to being irrelevant. The correlation coefficients between input variables are between −0.744 and 0.688, among which the correlation coefficients between most input variables are between −0.3 and 0.4, and only a few input variables have correlation coefficients lower than −0.3 or higher than 0.4. On the whole, there is no high correlation between the input parameters. Hence, these parameters and data sets can be applied to the machine learning process to evaluate the CS of recycled concrete.
There exists a strong link between the luminous intensity of fireflies and their location and attractiveness; the higher the luminous brightness is, and the brighter the brightness is, the higher the attraction of fireflies to fireflies [74]. In practical problems, brightness represents the answer to the problem, and the quality of the answer is directly proportional to the brightness of the fireflies. During the iteration process, fireflies are drawn to brighter fireflies and constantly update their location until they find the fireflies with the highest brightness; in other words, until they find the optimal solution or meet the termination condition. In practical problems, brightness is often used to represent the objective function value, and the relationship between the two is shown in the following formula:

$$I_i = f(x), x = (x_{i1}, x_{i2}, \ldots x_{in})$$

(2)

Suppose there is a brightness difference between firefly $i$ and firefly $j$, and the brightness of firefly $i$ is higher than that of firefly $j$. The attraction between firefly $i$ and $j$ is recorded as $\beta_{ij}$, and the solution formula is

$$\beta_{ij} = \beta_0 e^{-\gamma r_{ij}^2}$$

(3)

where $\beta_0$ is the maximum attraction, usually $\beta_0 = 1$, $\gamma$ is the light absorption coefficient, and $r_{ij}$ is the Cartesian distance between firefly $i$ and firefly $j$, which is defined as

$$r_{ij} = \|x_i - x_j\| = \sqrt{\sum_{k=1}^{d} (x_{ik} - x_{jk})^2}$$

(4)
In the formula, $d$ represents the dimension of the variable, $x_i, x_j$ represent the spatial positions of firefly $i$ and firefly $j$, and firefly $i$ completes the iterative update of its position by moving towards firefly $j$. The process is as follows:

$$x_j(t + 1) = x_j(t) + \beta_{ij}(x_i(t) - x_j(t)) + \alpha \varepsilon_j$$

(5)

In the formula, $t$ represents the number of iterations, $\beta_{ij}$ is the attraction of firefly $i$ to $j$, $\alpha$ represents the step parameter, where $\alpha \in [0, 1]$, and $\varepsilon_j$ represents the random vector. Figure 4 shows the specific process of FA.

**Figure 4.** Flowchart of FA.

2.3.2. Random Forest

RF is an integrated algorithm that combines multiple classifiers to achieve the purpose of prediction. RF has the advantages of fast speed, a simple model, and high precision, so it is often applied to solve practical problems. The classification process of RF can be summarized into the following steps (Figure 5):

1. Through the bootstrap sampling method, $m$ samples are randomly selected from the training set $M$ repeatedly and with put back. Each sample is the training data of each classification tree, and the unselected samples are used to analyze the effectiveness of the model.

2. Generation of the classification tree: when the leaf nodes start to be sorted, $n$ features are randomly selected from $N$ features, and the smallest Gini index is selected as the classification point to split, forming a single classification tree with small deviation to avoid the classification error of test data.
\[ Gini(M) = 1 - \sum_{i=1}^{k} p_i^2 \]  

where \( p_i \) represents the probability of \( C_i \) appearing in sample set \( M \).

3. Repeat step 2 \( T \) times to form \( T \) CART classification trees.

4. The classification results of the samples to be tested are determined by simple voting, and the category with the largest number of votes is selected as the final category. The result is the following:

\[ H(x) = \arg \max_Y \sum_{i=1}^{k} I(h_i(x) = Y) \]  

In the formula, \( H(x) \) is the final classification, \( \arg \max \) represents the independent variable point set of the maximum function value, \( k \) represents the number of classification trees, \( I(\cdot) \) represents the indicative function, the classification results of a single tree are represented by \( h_i(x) \), and the target classification is represented by \( Y \). The diagrammatic of RF is performed as follows:

![Schematic diagram of RF.](image-url)

3. Analysis and Discussion
3.1. Result of Hyperparameter Tuning

Machine learning algorithms contain many hyperparameters; usually, some parameters can be optimized by training, and some parameters that cannot be optimized by training are called super parameters. To ensure the fidelity of the model, we need to optimize the super parameters. In the present study, 10-fold CV was selected to verify the effect of FA on the hyperparameter tuning of RF. In this method, the data sets were randomly divided into ten groups, which were taken turns as the test groups, and the remaining nine groups were taken as the training groups. After 10 cross-validations, the optimal combination of hyperparameters was obtained. Ten-fold CV makes better use of the data sets and has the advantages of improving the generalization ability of the model and avoiding overfitting and underfitting of the model.

Figure 6 shows the result of a 10-fold CV: every fold has an RMSE value of less than 6 except for the seventh fold, and the minimum RMSE value is reflected at the eighth fold, proving that FA can effectively adjust the hyperparameters of RF and the hyperparameter tuning result of the model at the eighth fold is the best [8]. It is worth noting that due to the
abnormal phenomena of overfitting and underfitting, the prediction effect of the training set and the test set was further analyzed to determine the accuracy of the FA and RF hybrid machine learning model.

3.2. Evaluating the Prediction Performance

In this study, the consistency of the training set and that of the test set were analyzed to verify the accuracy of the FA and RF hybrid machine learning model in predicting the CS of recycled concrete. Figure 7 shows the prediction performance of the hybrid machine learning model on the CS, and the histogram in the figure represents the error between the predicted values and actual values. From Figure 7a, we can clearly see that the predicted values and the actual values almost coincide. There exists a high consistency between the predicted values and the actual values of the vast majority of samples in Figure 7b, and only a slight deviation between the predicted values and the actual values of individual samples. It is not difficult to distinguish that the consistency between the predicted and measured values is satisfactory. Although the training set has more errors compared to the test set, and there are several samples with significant deviations between the predicted and actual values, the overall consistency between the predicted values and measured values is acceptable, proving that the prediction accuracy of FA and RF on the CS of recycled concrete is also acceptable.
To further evaluate the predictive accuracy of the hybrid machine learning model from a quantitative analysis perspective, this study decided to analyze the RMSE value, representing the error, and the R value, representing consistency, separately, and the calculation of these two values was also carried out on the datasets for model training and validation. The RMSE and R values can be calculated using the following formulas:

\[ RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i^* - y_i)^2} \]  

\[ R = \frac{\sum_{i=1}^{n} (y_i^* - \bar{y}^*)(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (y_i^* - \bar{y}^*)^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} \]  

In the formula, \( y_i^* \) represents the predicted value, \( y_i \) represents the actual value, \( \bar{y}^* \) represents the average of the predicted value, \( \bar{y} \) represents that of the actual value, and \( n \) represents the sample capacity.

From Figure 8, we can see that there is an acceptable agreement between the predicted and measured values, and the CS values (predicted and measured) are concentrated in 20–70 MPa, with only a few points reaching around 100 MPa. The high consistency between the predicted and measured values can be verified through quantitative analysis of the parameters in the training and testing sets, and the parameters selected in this study are the R values and the RMSE values. It is worth noting that the R values of the data sets used for model training and model effectiveness verification are 0.9756 and 0.9328, respectively, and the RMSE values are 3.0752 and 6.4369, respectively, which shows that both the training set and the test set have a higher R value and a lower RMSE value. The above results show that the proposed FA and RF hybrid machine learning model for predicting the CS of recycled concrete does not have the phenomenon of overfitting. The satisfactory R values and RMSE values once again validate the reliability of the FA and RF hybrid machine learning model for the evaluation of the CS of recycled concrete.
3.3. Significance and Sensitivity Analysis of Input Variables

To put forward valuable suggestions for civil engineers to prepare recycled concrete with high CS, it is necessary to further analyze the influence law of input variables on output variables. Hence, this study analyzed the importance and sensitivity of variables based on the FA and RF hybrid machine learning model. The importance and sensitivity analysis results of the input indexes on the CS strength output of recycled concrete by the FA and RF hybrid model are performed in Figures 9 and 10.

![Figure 9. Importance analysis of input indexes.](image)

- **WC**: 2.5947
- **NMR**: 2.4315
- **AC**: 1.5348
- **BDR**: 1.3212
- **WAR**: 1.2847
- **BDN**: 1.0615
- **WAN**: 1.0567
- **NMN**: 0.8901
- **RCA**: 0.4695

**Figure 8.** Comparison between the predicted and actual values: (a) training set; (b) test set.

**Figure 9.** Importance analysis of input indexes.
negative correlation between them and the compressive strength. As a whole, AC, NMR, and NMN are negatively correlated with the CS of recycled concrete (see Figure 10b,d,e), while the influence of WAR and WAN on the CS of recycled concrete is random.

Figure 9. Importance analysis of input indexes.

Figure 10. Cont.
Figure 10. Sensitivity analysis of input variables: (a) WC, (b) AC, (c) RCA, (d) NMR, (e) NMN, (f) BDR, (g) BDN, (h) WAR, and (i) WAN.

The importance of the nine input indexes for the CS of recycled concrete is shown in Figure 9. As can be seen from Figure 9, the importance of the nine input variables to the CS of recycled concrete decreases in the order of WC, NMR, AC, BDR, WAR, BDN, WAN, NMR, and RCA. Among the input variables, WC and AC have the highest important score for the CS of recycled concrete; it is not difficult to understand that WC and AC affect
the output indexes by affecting its compactness, which has been confirmed in previous studies [75,76]. This indicates that engineers need to pay more attention to the indexes of WC and AC in the subsequent preparation of recycled concrete to achieve the purpose of enhancing the CS of recycled concrete by controlling compactness. In the physical properties of aggregates, the influence of size and density on the output indexes is greater than that of water absorption. Interestingly, RCA has the lowest importance score of 0.4695 for the CS of recycled concrete. The above result shows that RCA has a slight effect on the CS of recycled concrete, which proves the feasibility of recycled aggregate replacing natural aggregate [77,78]. However, the low importance score of RCA will also mistakenly lead readers to believe that it is feasible to use recycled aggregate to completely replace natural aggregate without too much impact on the CS of recycled concrete. The use of recycled aggregate instead of natural aggregate is not only beneficial to reduce resource exploitation but also helpful to the disposal of construction waste. To further explore the influence of RCA and other input variables on the CS of recycled concrete, the sensitivity of input variables on output variables was analyzed.

As shown in Figure 10a, WC within the range of 0–0.3 has little influence on the CS of recycled concrete, and when WC exceeds 0.5, there exists a negative correlation between WC and CS. It is clear that (Figure 10c), with the increase of RCA, the CS variation trend of recycled concrete increases first and then rapidly decreases. When RCA is less than 20%, the increase of RCA has a positive effect on the CS of recycled concrete. Therefore, engineers can focus on RCA to improve the CS of recycled concrete while reducing the cost. In the beginning, BDR and BDN have no significant influence on the compressive strength (see Figure 10f,g), but when their values rise to a certain extent, there exists a negative correlation between them and the compressive strength. As a whole, AC, NMR, and NMN are negatively correlated with the CS of recycled concrete (see Figure 10b,d,e), while the influence of WAR and WAN on the CS of recycled concrete is random.

4. Conclusions

Recycled concrete is a kind of green concrete processed from construction waste which plays an important role in the green and sustainable development of concrete. To further expand the utilization scope of recycled concrete, this research proposed to utilize recycled concrete to replace cement-stabilized macadam base. To verify the feasibility of recycled concrete in practical engineering, this study analyzes the impact of various factors on its compressive strength based on an artificial intelligence model made up of FA and RF. Through the study of intelligent optimization algorithms in the CS evaluation of recycled concrete, the following conclusions can be obtained:

1. There is no high positive and negative correlation between the input variables, which proves that the multicollinearity phenomenon will not occur and affect model prediction, indicating that the selection of input indicators is reasonable.
2. FA shows a strong acceptability for the optimization effect of the hyperparameters of RF, and the convergence of RMSE and the results of 10-fold CV can prove the above conclusion.
3. The predicted and measured values of the CS are highly consistent, both with satisfactory R values and RMSE values, proving that introducing the firefly algorithm to the random forest for the CS modeling of recycled concrete can improve the computational accuracy.
4. Importance and sensitivity analysis of variables shows that WC and NMR show higher importance for the CS of recycled concrete, while RCA is the index with the lowest importance score.

In the future, researchers need to invest their research efforts in optimizing the mechanical and economic properties and achieving other interrelated goals of recycled concrete to promote the application of recycled concrete in pavement materials and promote the sustainable development of the construction industry. Furthermore, in addition to the hybrid
algorithm of RF and FA, more evolutionary learning algorithms should be considered to improve the reliability of the simulation process.

**Author Contributions:** Conceptualization, Y.L. and Y.W.; methodology, M.Z.; software, J.H.; validation, Y.L., Y.W. and M.Z.; investigation, Y.L.; writing—original draft preparation, M.Z.; writing—review and editing, J.H.; supervision, J.H.; project administration, Y.L.; funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (Grant No. 52108426), the Faculty Start-up Grant of China University of Mining and Technology (Grant No. 102520282), Natural Science Foundation of Jiangsu Province (Grant No. BK20210513), the Cultivation Base of Shanxi Key Laboratory of Mining Area Ecological Restoration and Solid Waste Utilization, Shanxi Institute of Technology (Grant No. SXIT202302), and the Key Laboratory of Geological Safety of Coastal Urban Underground Space, Ministry of Natural Resources (BHKF2022Y02).

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare that they have no conflict of interest in this work.

**References**


33. Huang, J.; Xue, J. Optimization of svr functions for flyrock evaluation in mine blasting operations. Environ. Earth Sci. 2022, 81, 434. [CrossRef]


73. Huang, J.; Zhou, M.; Zhang, J.; Ren, J.; Vatin, N.I.; Sabri, M.M.S. The Use of GA and PSO in Evaluating the Shear Strength of Steel Fiber Reinforced Concrete Beams. KSCE J. Civ. Eng. 2022, 26, 3918–3931. [CrossRef]
74. Huang, J.; Zhang, J.; Gao, Y. Intelligently predict the rock joint shear strength using the support vector regression and Firefly Algorithm. Lithosphere 2021, 2021, 2467126. [CrossRef]


Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.