Developing Hybrid Machine Learning Models to Determine the Dynamic Modulus (E*) of Asphalt Mixtures Using Parameters in Witczak 1-40D Model: A Comparative Study

Wenjuan Xu 1,2, Xin Huang 1, Zhengjun Yang 1,3, Mengmeng Zhou 4 and Jiandong Huang 4,*

1 College of Civil Engineering, Nanjing Forestry University, Nanjing 210037, China; xuwenjuan@jsafc.edu.cn (W.X.); xinhuang@njfu.edu.cn (X.H.); yangzj0511@163.com (Z.Y.)
2 College of Landscape Architecture, Jiangsu Agricultural and Forestry Technical College, Zhenjiang 212400, China
3 Jurong Traffic and Transport Bureau, Zhenjiang 212400, China
4 School of Mines, China University of Mining and Technology, Xuzhou 221116, China; ts21020232p21@cumt.edu.cn
* Correspondence: huang@cumt.edu.cn

Abstract: To characterize the dynamic modulus (E*) of the asphalt mixtures more accurately, a comparative study was shown in this paper, combining six ML models (BP, SVM, DT, RF, KNN, and LR) with the novelly developed MBAS (modified BAS, beetle antennae search) algorithm to check the potential to replace the empirical model. The hyperparameter tuning process of the six ML models by the proposed MBAS algorithm showed satisfactory results. The calculation and evaluation process demonstrated fast convergence and significantly lower values of RMSE for the five ML models (BP, SVM, DT, RF, and KNN) to determine the E* of the asphalt mixtures. Comparing the performances of the six ML models in the prediction of the E* by the statistical coefficients and Monte Carlo simulation, the RF model showed the highest accuracy, efficiency, and robustness.

Keywords: machine learning; dynamic modulus; asphalt mixtures; BAS; hyperparameters

1. Introduction

A crucial parameter that distinguishes the flexible pavement performance conditioning on multiple temperatures and loading circumstances in the Mechanistic-Empirical Pavement Design Guide (MEPDG) is the E* (dynamic modulus) of asphalt mixtures [1-3]. To achieve an excellent pavement performance and long service life, the precise calculation of the E* is a crucial assignment in the designing process of the pavements [4–10]. In the previous studies, laboratory-based methods and predictive models were employed to determine the E* to be used in the MEPDG [4–7,11–14]. Regarding the laboratory-based methods, it is a typically time-consuming procedure because of the requirements of the notable amount of time to produce the samples, balance the temperature, and conduct the experimental processes [8,15–24]. Additionally, this complicated experimental test requires trained professional staff and costly devices (e.g., AMPT and SPT). To address the above-mentioned issues, researchers and practitioners around the world have developed several predictive models to determine the E*. Among these, one of the most frequently mentioned is the so-called Witczak models. Another example of these models is the adjusted Hirsch model, which has also been broadly used to determine the E*.

However, although these models have been applied in applications (e.g., in various software and engineering practices), their shortcomings were exposed in recent studies: many researchers challenged the accuracy of these models, particularly at the extreme high and low temperatures [4,6,25–27]. Yousefdoost et al. [28] found that the models
based on the parameter of viscosity can miscalculate the dynamic modulus, while those models based on the parameter of dynamic shear modulus of the asphalt binder can overvalue the dynamic modulus of the asphalt mixtures. Additionally, Ali et al. [7] and Georgouli et al. [12] discovered that models based on the viscosity can underestimate the dynamic modulus of the asphalt mixtures. Batioja-Alvarez et al. [29] observed that the Hirsh model can underestimate the dynamic modulus of the asphalt mixtures used in the Indiana state. It is worth noting that the inaccuracy of these predicting models originates from the limitations of the regression-based techniques. They are all predicated on linear or nonlinear regression methods. For example, the general form of a regression model to predict the $E^*$ of the asphalt mixture is based on the following equation.

$$y = f(w_i \cdot x_i)$$  \hspace{1cm} (1)

In general, it is difficult to obtain an accurate regression equation by using empirical model, so the regression equation for predicting the $E^*$ value of asphalt mixture is limited. Therefore, it is needed to employ more advanced techniques to generate more precise predictive models of $E^*$. Nowadays, machine learning (ML) techniques are widely used to solve different kinds of engineering problems because of their abilities in data processing and optimization. The principal advantage of the ML methods is that they can discover the underlying behavior of a complicated system without needing a previous understanding of the association between the input and the target variable. ML techniques have been applied for predicting the performance of the infrastructural materials and in the realm of pavement engineering. Some of these prominent studies include the discovery of the pavement distress [13], forecast of the international roughness index (IRI) [14], assessment of the segregation of asphalt pavement [15], forecast of the rutting [16], forecast of the fracture energy of asphalt mixtures [17], and calculation of the mechanical characteristics of concretes [18,19]. Regarding the weakness of the regression-based techniques for the prediction of the $E^*$, many scholars have also applied the ML techniques in recent studies, including the M5P model tree algorithm [30], deep learning-based models [31], bagged trees ensemble [32], pre-trained deep learning [33], and gradient decision tree boosting [34].

It should be noted that although the various ML techniques mentioned above were used in predicting $E^*$, there were still some problems that need to be solved: (1) only a limited number of advanced ML algorithms have been used to predict $E^*$, and the reliability and computational efficiency of other advanced ML algorithms, such as random forest (RF), have not been deeply studied in previous studies; (2) the application of ML algorithm always requires proper selection of the hyperparameters, but the optimization ability of beetle antennae search (BAS) for the hyperparameters has not been fully studied in the prediction of asphalt mixture; (3) the existing ML models may have great differences in the calculation results and efficiency when applied to predict the $E^*$ of the asphalt mixture, but there is still a lack of systematic and quantitative comparison of these algorithms.

2. Research Objective

To address the above-mentioned issues, six ML models, including the back-propagation (BP), decision tree (DT), k-nearest neighbors (KNN), logistic regression (LR), random forest (RF), and support vector machine (SVM), were investigated and compared to predict the $E^*$ of asphalt concrete in the present study. These ML algorithms were used to establish the relationships between the design parameters (including $V_{beff}$, $V_a$, $\rho_{200}$, $\rho_4$, $\rho_{3/8}$, $G^*$, and $\delta$) and $E^*$ of asphalt mixtures, and the BAS algorithm was used to adjust the hyperparameters. This study can be used as a benchmark study for the application of the ML methods to predict the $E^*$ of asphalt concrete, which is of great significance in the performance evaluation of flexible pavement in the future. The rest of this article will introduce the dataset descriptions, the application of the artificial intelligence models, the tuning of the hyperparameters, and the summarization of the conclusions.
3. Methodology

3.1. Overview of the Machines Learning (ML) Models

In the present study, six widely used ML models (including the BPNN, DT, KNN, LR, RF, and SVM models) were employed to predict the E* of the asphalt mixtures using the design factors (Vbeff, Va, ρ200, ρ4, ρ3/8, G*, and δ) from the prediction model in MEPDG. The dataset was collected from the previous studies using the same input and output parameters [35]. In the Superpave mixture design of 16 asphalt mixtures, two types of binders and eight types of aggregate gradations were used. Then, the E* of asphalt mixture was evaluated at three temperatures and three loading frequencies, and 144 laboratory data sets were constructed. It should be noted that the shear dynamic modulus and phase angle were also determined at the same temperature and loading frequency as predicted from the master curves of the binder. Table 1 shows the descriptive statistics of the training and test data sets used for the prediction.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Variables</th>
<th>Mean</th>
<th>STD</th>
<th>Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E*</td>
<td>3532.0</td>
<td>3272.3</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Vbeff</td>
<td>3.5</td>
<td>0.4</td>
<td>-0.4</td>
</tr>
<tr>
<td></td>
<td>Va</td>
<td>35.7</td>
<td>2.7</td>
<td>-0.3</td>
</tr>
<tr>
<td>Training dataset</td>
<td>ρ200</td>
<td>16.6</td>
<td>4.3</td>
<td>0.4</td>
</tr>
<tr>
<td>(the number of</td>
<td>ρ4</td>
<td>6.9</td>
<td>0.1</td>
<td>-0.3</td>
</tr>
<tr>
<td>data points is</td>
<td>ρ3/8</td>
<td>10.5</td>
<td>0.7</td>
<td>-0.3</td>
</tr>
<tr>
<td>100)</td>
<td>G*</td>
<td>2.8</td>
<td>5.2</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td>δ</td>
<td>59.1</td>
<td>12.1</td>
<td>-0.2</td>
</tr>
<tr>
<td></td>
<td>E*</td>
<td>3557.9</td>
<td>3530.2</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>Vbeff</td>
<td>3.7</td>
<td>0.6</td>
<td>-1.3</td>
</tr>
<tr>
<td></td>
<td>Va</td>
<td>36.2</td>
<td>3.0</td>
<td>0.4</td>
</tr>
<tr>
<td>Testing dataset</td>
<td>ρ200</td>
<td>17.8</td>
<td>2.7</td>
<td>1.4</td>
</tr>
<tr>
<td>(the number of</td>
<td>ρ4</td>
<td>6.9</td>
<td>0.2</td>
<td>-0.2</td>
</tr>
<tr>
<td>data points is</td>
<td>ρ3/8</td>
<td>10.5</td>
<td>0.7</td>
<td>-1.2</td>
</tr>
<tr>
<td>44)</td>
<td>G*</td>
<td>2.8</td>
<td>5.1</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>δ</td>
<td>57.5</td>
<td>12.2</td>
<td>-0.1</td>
</tr>
</tbody>
</table>

BP neural network is composed of the input layer, output layer, and hidden layer. Each layer of the BP neural network has multiple neurons, the number of which is determined by a specific model. Two or two neurons inside the layer are not interconnected, all neurons are connected with the nearby layer through a one-way connection, and the connected nodes interact with each other through weight [36–39]. The DT model is a method of approximating the values of discrete functions. This is a typical classification method [40]. The data is processed, rules and decision trees are generated by induction algorithm, and then decision analysis is made on the new data. Decision trees are essentially the process of classifying data through a set of rules. KNN algorithm calculates the distance between the unknown sample and all known samples by judging the category of the unknown sample and taking all known samples as the reference [41,42]. According to the majority-voting law, the KNN algorithm calculates the distance between the unknown sample and all known samples. The unknown samples were classified into the same category as the ones with more categories in the K nearest neighbor samples. LR algorithm can be used for classification and prediction. It can predict the probability of occurrence of future results based on the performance of historical data [43,44]. The goal of linear regression fitting is to keep the data points on a straight line as much as possible, while the goal of LR is to try to place the points of different categories on both sides of the
straight line. RF is a highly accurate prediction for the dataset but is not easy to overfit. Additionally, it can rank the importance of the resulting variables and can process both discrete and continuous data, and does not need to be normalized [45–47]. SVM is a kind of generalized linear classifier based on supervised learning. The decision boundary is the maximum boundary hyperplane of the learning sample [17]. A more detailed description of these six ML models can be found in the previous studies [48–50].

3.2. Modified Beetle Antennae Search (MBAS)

3.2.1. Beetle Antennae Search (BAS) Algorithm

BAS algorithm is a recently proposed heuristic optimization algorithm, which is inspired by the search behavior of beetles [33,40].

To determine the global optimality regarding a multidimensional space, Equation (2) should be employed as described below:

\[
\begin{align*}
\text{Minimize} & \quad f(x), \quad x = [x_1, x_2, \cdots, x_N]^T \\
\text{Maximize} & \quad f(x), \quad x = [x_1, x_2, \cdots, x_N]^T
\end{align*}
\] (2)

The following equation can be used to define the beetle’s searching behavior.

\[
b = \frac{\text{rand}(k, 1)}{\|\text{rand}(k, 1)\|}
\] (3)

\[
x_r = x^i + d^i b
\] (4)

\[
x_l = x^i - d^i b
\] (5)

The following equation can be used to define the detecting behavior of the beetle.

\[
x_{i+1} = x^i + \delta^i b \cdot \text{sign}(f(x_r) - f(x_l))
\] (6)

The following equation can be used to describe the step size of the updating equation.

\[
\delta_{i+1} = \eta \delta_i
\] (7)

where \(\eta\) is the coefficient of step size to characterize the attenuation.

3.2.2. MBAS Algorithm

For traditional BAS, the beetle’s step size is constant or decreases with each iteration. There are problems with this step-size strategy. If the given step size is not big enough, the BAS algorithm may converge slowly or fall into a local optimal state. Therefore, Levy flight and self-inertial weight were used to adjust BAS step size in this study. To improve the search efficiency, this paper named the improved the BAS to the MBAS. The following objectives can be achieved: (i) the step size can be adjusted quickly based on the value of the current fitness, and the adaptive weight can be used to reduce the oscillation; (ii) Levy flight was used to improve the step size randomly if the BAS fell into the state of local optimal.

(i). Levy flight

During the calculation process of the traditional BAS algorithm, the step size is kept constant or decreasing in one iteration, which leads to the BAS algorithm falling into the state of local optimal easily. To solve this problem, Levy flight was adopted to modify the step size of the BAS algorithm. From the early research results, Levy flight is useful to adjust step size [51–53]. In this study, if the BAS algorithm is in a locally optimal state, the step size of the beetle can be increased by using Equation (8).

\[
\delta^{(i)} = a|\text{Levy}| \otimes \delta^{(i-1)}
\] (8)
in which $\alpha$ is a parameter to describe the random value between 0 and 1. $\otimes$ represents multiplication by item; $|Levy|$ is a variance of infinite Levy distribution, and the infinite is described as $Levy \sim u = t^{-\lambda}, (1 < \lambda \leq 3)$. The Levy flight can be triggered as

$$|f^{(i)} - f^{(i-1)}| < \mu(f_w - f_b)$$

(9)

where $\mu$ is the parameter, which is determined as $10^{-5}$ in the present study.

(ii). Self-adaptive inertia weight

In the present study, the adaptive inertia weights are applied to a monotone reduction equation, as described by the following equation.

$$\delta^{i+1} = \eta^i \times \delta^i$$

(10)

where $\delta^i$ is the step of the current position; $\eta^i$ represents the adaptive inertia weight, which is determined by:

$$\eta^i = (1 - \alpha) 0.95 + \alpha \frac{f^i_w - f^i}{f^i_w - f^i_b}$$

(11)

where $f^i$ is the fitting equation of the current position; $f^i_w$ is the best fit value; $f^i_b$ is the worst fit value; $\alpha$ is the tradeoff between two parameters: the first item ($(1 - \alpha)0.95$) is the inertia weight, the second $(\alpha \frac{f^i_w - f^i}{f^i_w - f^i_b})$ is an adaptive feature. In this study, $\alpha$ was determined to be 0.2 after modification of preliminary trial calculation.

3.3. Methods for the Evaluation and Calibration

3.3.1. Determine the Predicting Performance

The parameter used to evaluate and calibrate the proposed model was Root Mean Square Error (RMSE) [44], which employs the following equation to calculate the difference between the predicted value and the actual measured value.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y^*_i - y_i)^2}$$

(12)

where $y^*_i$ and $y_i$ represent predicted value and actual measured value, respectively. $n$ indicates the number of data samples. Another parameter used to evaluate predictive performance is the correlation coefficient ($R$), which is determined by the correlation between predicted and actual values, as shown in the following equation [54].

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

(13)

where $\bar{y}^*$ and $\bar{y}$ are the mean value for the predicted ones and actual ones, respectively.

3.3.2. K-Fold Cross-Validation

For the correction of the regression model, the early studies adopted the simple substitution method, support method, retention method, and Bootstrap method [55–57]. Regarding the approaches employed to validate the training data, the most widely used k-fold cross-validation (CV) was employed in the present research [58]. Specifically, $k$ is specified as 10, taking into account the recommendations of earlier studies [59]. Therefore, the training data set is split by a factor of 10 during hyperparameter tuning. The algorithm has been trained for nine times and verified for 10 times. This process should be repeated 10 times, each time verifying with a different fold. The final result is determined to be the result with the smallest error in a single fold. Figure 1 summarizes the CV process above.
3.4. Hyperparameter Tuning

Figure 2 gives the hyperparameters of the six ML models that were tuned by the MBAS.

\[
 f(x) = \sum_{i=1}^{n} -x_i \sin(\sqrt{|x_i|})
\]  

Equation (14) shows that as \( x \) goes to infinity, \( y \) goes to minus infinity. In other words, the smaller the \( y \) value, the better the performance of the optimization algorithm. According to the above description, the hyperparameters of the 6 ML models were adjusted using the 10-fold CV and MBAS. After 10 folds, 10 hyperparametric data sets can be obtained. For these 10 subsets, the MBA algorithm needs to search the optimal hyperparameters of 9 subsets and evaluate the RMSE of 1 subset. After convergence, the optimal hyperparameter of the first iteration can be determined. Repeat this process 10 times according to the requirement of 10 CV, and verify the different sets each time. Through the above steps, 10 RMSE values can be obtained, and the smallest RMSE is selected to correspond to the optimized hyperparameter. Finally, the predictive performance evaluation test data set of the six ML models were independent of the training data set (30% of the total data set). Therefore, the test data set is only used to evaluate predictive performance. It should be noted that minimum overfitting occurs once the model has a good fit for the training and test data sets.
4. Results and Discussion

4.1. Results of the Hyperparameter Tuning

As mentioned above, the proposed MBAS algorithm was used in this study to tune the hyperparameters of the six ML models, and the optimized hyperparameters were determined for verification according to the hyperparameter set of the minimum RMSE value in the data set. Table 2 gives the optimum hyperparameters of the six ML models.

<table>
<thead>
<tr>
<th>Models</th>
<th>Hyper-Parameters</th>
<th>Empirical Scope</th>
<th>Initial Value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP</td>
<td>hidden layer num</td>
<td>[1,4]</td>
<td>{1,2,3,4}</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>hidden layer size</td>
<td>[1,20]</td>
<td>(20,10), (10,10,10), (8,6,8)</td>
<td></td>
</tr>
<tr>
<td>DT</td>
<td>min samples split</td>
<td>[1,10]</td>
<td>25</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>min samples leaf</td>
<td>[2,10]</td>
<td>50</td>
<td>8</td>
</tr>
<tr>
<td>KNN</td>
<td>neighbors num</td>
<td>[1,10]</td>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>LR</td>
<td>tol</td>
<td>$[1 \times 10^{-5},1 \times 10^{-3}]$</td>
<td>1</td>
<td>$1.18 \times 10^{-4}$</td>
</tr>
<tr>
<td></td>
<td>c inverse</td>
<td>0.1–10</td>
<td>10</td>
<td>220</td>
</tr>
<tr>
<td>RF</td>
<td>tree num</td>
<td>[1,10]</td>
<td>40</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>min samples leaf</td>
<td>[1,10]</td>
<td>40</td>
<td>1</td>
</tr>
<tr>
<td>SVM</td>
<td>C</td>
<td>[0.1,1000]</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>$[0.001,100]$</td>
<td>16</td>
<td>$4.8 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

Figure 3 gives the RMSE values of the six ML models at the various fold number.
Figure 3. RMSE results of the hyperparameter tuning: (a) BP; (b) DT; (c) KNN; (d) LR; (e) RF; and (f) SVM.

It can be observed that the minimum values of RMSE (which are 220 MPa, 300 MPa, 625 MPa, 1000 MPa, 175 MPa, and 185 MPa) can be obtained at the fold number of 8, 10, 2, 8, 8, and 8 concerning the BP, DT, KNN, LR, RF, and SVM models, respectively. Therefore, the hyperparameters at these fold numbers were employed for these six algorithms to evaluate the E* of the asphalt mixtures.

Figure 4 gives the average values of RMSE versus iteration during the hyperparameter tuning process. It should be noted that no further decrease of the RMSE values was observed after 50 iterations for all the ML algorithms.

Figure 4. Iteration vs. RMSE values.

It can be observed that different ML models showed different evolution modes regarding the RMSE values and convergence speed. The RMSE value of the LR model tended to be stable from the initial iteration, indicating that the influence of the hyperparameters on the performance of the LR algorithm has no obvious effect on the dataset of the E*. The convergence speed of the RF and SVM models was slightly slower than the others when they were used to evaluate the E* of the asphalt mixtures. The RMSE values of the RF and SVM models tended to converge at the iteration of 15 and 22, respectively. The RMSE values of the KNN and SVM models decreased more quickly compared to the other models when the BAS was employed to tune the hyperparameters, indicating that the BAS can tune the KNN and SVM algorithms more efficiently to evaluate the dynamic modulus of the asphalt mixtures. In terms of the overall effect, the BAS algorithm can efficiently find the optimized hyperparameters of these ML models.
Regarding the accuracy of the tuning hyperparameters process, it can be seen that LR and KNN models behaved the worst, indicated by the largest RMSE values (which are 1015 MPa and 520 MPa, respectively) after the convergence. The BP, SVM, DT, and RF models showed relatively small values of RMSE, indicating the BAS can tune the hyperparameters of these four algorithms reliably.

4.2. Comparison of the Predictive Performance

Figure 5 gives the results (which are described by the box-plot) of the residual between the predicted E* and actual E* of the asphalt mixtures.

![Figure 5. Residual results of various predictive models.](image)

It can be seen that the RF performed the best among all the models, indicated by the minimum residual value. The KNN and LR models showed large residuals, demonstrating that they may not be suitable to evaluate the E* of the asphalt mixtures. The other three models (BP, SVM, and DT) showed moderate performance in predicting the E* of the asphalt mixtures, with certain accuracy. The results in Figure 5 indicate enough accuracy determined from the proposed models (RF, BP, SVM, and DT) modified by the MBAS algorithm.

To more comprehensively evaluate the performance of these six ML algorithms in predicting the E* of asphalt mixture, the predicted dynamic modulus of the asphalt mixture, the actual measured dynamic modulus of the asphalt mixture, as well as the ‘1:1’ curve were plotted, as shown in Figure 6.
The predictive performance of the six models was evaluated by the two most commonly used parameters: RMSE and R, as mentioned in the previous chapter. Better predictive performance of the BP, DT, RF, and SVM models were observed compared to the others except for the few noise points. Additionally, the good predictive performance of these four models can be indicated by the lower value of RMSE and higher value of R. The results in Figure 6 can also indicate enough accuracy determined from the proposed MBAS algorithm. Regarding the RF algorithm, the best agreements can be observed between the comparison results, with the correlation coefficients of 0.9921 and 0.9977 for the testing dataset and training dataset, respectively. The worst-performing models were LR and KNN with the highest values of RMSE and lowest values of R, indicating that these baseline algorithms may not suitable to predict the E* of asphalt mixtures.

The reliability characteristics of these ML models can also be graphically represented by the Taylor diagram (which is generally used to evaluate the accuracy of the ML model), including the information of the R, standard deviation (SD), and RMSE. In general, the scatters in the Taylor diagram represent various ML models, the radial line represents the R, the horizontal and vertical axes represent the SD, and the dashed line represents the RMSE. Taylor diagram is a modification from the previous use of the scatter diagram, which can only present two indicators to represent the accuracy of the ML model. Figure 7 gives the Taylor diagram of the six ML models as they were employed to predict the E* of the asphalt mixtures.

Figure 6. Comparison of the predicted dynamic modulus and measured dynamic modulus: (a) BP; (b) LR; (c) DT; (d) KNN; (e) RF; and (f) SVM.
It can be seen that SVM, RF, DT, and BP models are the closest to the measured dynamic modulus with relatively higher R and lower RMSE, indicating the MBAS algorithm can tune the hyperparameters of these four ML models well to predict the $E^*$ of the asphalt mixtures. It should be noted that the RF model exhibited the lowest values of RMSE and SD, as well as the highest value of R, performing best in the prediction.

To clarify the randomness characteristics of different ML models, this study employed the widely used Monte Carlo simulation [60], where a random combination of input variables that fit a certain probability distribution is selected and the ML models are used to predict the output. The uncertainty of the input variable is propagated to the output by Monte Carlo simulation.

Figures 8 and 9 give the results of the Monte Carlo simulation (Number of Monte Carlo runs vs. values of R and RMSE, respectively).
Figure 9. Monte Carlo simulation (Number of Monto Carlo runs vs. value of RMSE).

All the values of R and RMSE tended to converge as the Monto Carlo runs 500 times. LR model showed the lower R and higher RMSE on both the test and training dataset after the simulation results converged, once again proving that this baseline model may not be suitable to predict the E* of the asphalt mixtures. It is worth noting that the KNN model can obtain the lowest value of R for the test dataset, but can obtain almost the highest value of R for the training set (for the values of RMSE, it also obtain the highest and lowest RMSE, respectively). The same results were presented for RMSE values and were not repeated here. It indicated that the KNN model is not very stable to predict the dynamic modulus of the asphalt mixture. Additionally, it can be observed that the RF model obtained almost the highest value of R and lowest value of RMSE regarding both the testing and training dataset, demonstrating that the RF model is more stable to predict the dynamic modulus of the asphalt mixtures.

5. Conclusions

In this study, six ML models (BP, SVM, DT, RF, KNN, and LR) were implemented and compared using the newly developed Modified BAS (MBAS) algorithm to replace the mathematical regression model in MEPDG. The conclusions can be summarized as follows.

The BAS algorithm is modified by Levy flight and inertia weight, which avoids premature convergence to local optimum and improves search efficiency compared with the traditional BAS algorithm. The calculation results show that the five ML models (BP, SVM, DT, RF, and KNN) for evaluating the E* of the asphalt mixtures have fast convergence speed and low RMSE value, which proves the effectiveness and accuracy of the MBAS algorithm. However, for dynamic modulus data sets, hyperparameters have no significant effect on the performance of LR models. The comparison results show that the MBAS algorithm can adjust the hyperparameters of SVM and KNN more effectively.

Comparing the performances of the six ML models in the prediction of asphalt mixture’s dynamic modulus by the statistical coefficients (values of R, RMSE, and SD) and Monte Carlo simulation, the RF model showed the highest accuracy, efficiency, and robustness.

In future studies, more experimental tests will be conducted and updated databases will be continuously collected to improve the applicability of the proposed model. The application of the models on a different set of asphalt concretes will be conducted to evaluate whether the model can be generally used in asphaltic materials.
Author Contributions: Conceptualization, W.X.; Z.Y. and J.H.; methodology, W.X.; validation, M.Z. and X.H.; writing—original draft preparation, W.X.; writing—review and editing, M.Z.; supervision, J.H. and X.H. 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) and the Natural Science Foundation of Jiangsu Province (Grant No. BK20210513).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

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

References