Journal of Science and Transport Technology
Journal homepage: https://jstt.vn/index.php/en

Practical numerical tool for marshall stability prediction based on machine learning: an application for asphalt concrete containing basalt fiber

Ba-Nhan Phung, Thanh-Hai Le*, Minh-Khoa Nguyen, Thuy-Anh Nguyen, Hai-Bang Ly
University of Transport Technology, Hanoi 100000, Vietnam

Article info
Type of article:
Original research paper

DOI:
https://doi.org/10.58845/jstt.utt.2023.en.3.3.26-43

***Corresponding author:**
E-mail address:
hailt@utt.edu.vn

Received: 05/09/2023
Revised: 26/09/2023
Accepted: 29/09/2023

Abstract: Marshall stability (MS) is used to evaluate the resistance to settlement, deformation and displacement of asphalt concrete. However, these experiments are complex, expensive and time-consuming. Therefore, it is important to develop an alternative method to quickly determine these parameters. This paper presents a comprehensive investigation into applying machine learning techniques for predicting the MS of basalt fiber asphalt concrete. The study leverages the Gradient Boosting algorithm to establish predictive models. A database containing 128 samples is employed as the foundation for model construction. Additionally, SHAP analysis is employed to reveal the underlying variables influencing the predictive outcomes. To extend the practicality of the findings, a Graphical User Interface (GUI) is developed to facilitate easy access to the predictive tool for material engineers. The results show that the content aggregate 4.75mm is the most influential variable, followed by the content aggregate 2.36mm, the content of fiber, the content of binder, and the content aggregate 9.5mm in descending order of impact.

Keywords: Marshall stability, basalt fiber asphalt concrete, Machine Learning.

1. Introduction

Asphalt pavement, a prevalent type of roadway infrastructure, offers a range of benefits and drawbacks. It boasts advantages such as cost-effectiveness, ease of construction, and a smooth riding surface, enabling efficient transportation networks [1]. However, it contends with susceptibility to cracking, rutting, and diminished durability under heavy traffic loads and environmental factors. To address these challenges, advancements in methods and technologies have been pivotal in enhancing the quality and longevity of asphalt concrete pavement. One such innovative approach involves the incorporation of fibers to bolster the pavement's structural integrity. This method has garnered attention for its potential to mitigate cracking, improve fatigue resistance, and enhance overall performance [2]. Among the various materials available, basalt fiber emerges as a compelling choice for reinforcing asphalt pavement, owing to its exceptional advantages. Basalt fiber is derived from natural volcanic rock, offering high tensile strength and modulus, rendering it resilient against deformation under heavy loads [3]. Its resistance to corrosion and chemicals ensures prolonged pavement lifespan, even in aggressive environments. Basalt fiber's low thermal expansion coefficient and high temperature stability contribute to reduced thermal cracking susceptibility, a

common concern in asphalt pavements [4,5]. Additionally, its compatibility with asphalt binder ensures efficient integration without compromising the pavement's structural integrity.

Marshall Stability (MS) is a fundamental mechanical property that characterizes the resistance of asphalt concrete in general and fiber-reinforced asphalt concrete in particular to deformation under applied loads and temperature variations. It is a key indicator of the overall performance and durability of asphalt mixtures in various engineering applications. MS plays a pivotal role in designing and evaluating asphalt concrete mixtures. It helps assess the ability of an asphalt mixture to withstand traffic loads, temperature changes, and moisture effects without experiencing excessive deformation or failure [6]. A higher MS value indicates better resistance to rutting, shoving, and other distresses, ensuring a longer pavement service life. Several methods are employed to determine MS, with the most common being the Marshall Test, which simulates the compressive load and temperature conditions experienced by asphalt pavements. This test involves compacting cylindrical specimens, subjecting them to axial loading, and measuring the maximum load at failure. The Marshall Test is widely used due to its simplicity and practicality in replicating field conditions [7]. However, during the experimental process, the MS of both laboratory-prepared samples and field core samples is significantly influenced by the sample dimensions (diameter, height, surface roughness). Additionally, conducting tests in the laboratory and in-field conditions can be costly and time-consuming, with results being contingent upon the type of testing equipment and the expertise of the laboratory personnel. Therefore, developing an approach that combines computational modeling, machine learning, or advanced non-destructive testing could offer more efficient and cost-effective ways to determine the MS, leading to improved asphalt mixture design and pavement performance.

Hence, numerous studies have proposed applying artificial intelligence (AI) techniques to predict the mechanical behavior of asphalt concrete, with a specific emphasis on fiber-reinforced concrete. Using AI or machine learning (ML)- based models is a remarkable strategy to effectively and consistently model intricate engineering aspects and estimate output parameters [8,9]. In the study by Alireza and Salman [10], a genetic programming simulation approach was employed to predict Marshall mix design parameters of asphalt concrete. Moreover, several linear regression models have been utilized as baseline models to evaluate the presented genetic programming model. These proposed models predict Marshall mix design parameters based on particle shape, aggregate structure indices, asphalt binder viscosity, and quantity. The results substantiated the superior effectiveness of the proposed methods over resource-intensive laboratory procedures, with minimal errors quantified by root mean square error (RMSE), mean absolute error (MAE), and correlation coefficient (R) exceeding 0.9, enabling relatively accurate predictions of Marshall mix design parameters. Upadhya et al. [11] succeeded in identifying the most suitable predictive model for MS and optimal asphalt content (OAC) in soft road surface carbon fiber-reinforced asphalt concrete. Five machine learning techniques, namely Support Vector Machine (SVM), Gaussian Process Regression (GPR), Random Forest (RF), Decision Tree (DT), and M5P model -were utilized to determine the optimal predictive model. Seven statistical metrics, specifically R, MAE, RMSE, Relative Absolute Error (RAE), Relative Root Square Error (RRSE), Willmott Index (WI), and Nash-Sutcliffe Efficiency (NSE), were employed to assess the efficacy of the applied models. The results demonstrated the superior performance of the RF model over all used models, yielding R values of 0.9735, MAE of 1.1755, RMSE of 1.5046, RAE of 25.68%, RRSE of 26.93%, WI of 0.9351,

and NSE of 0.9272 in the validation phase. Additionally, sensitivity analysis highlighted the significant influence of approximately 5.0% asphalt content (AC) on MS in carbon fiber-containing asphalt concrete mixes. In another study by Upadhya et al. [12], six machine learning models based on algorithms such as Artificial Neural Network (ANN), SVM, GPR, M5P model, RF, and Random Tree (RT) were employed to predict the MS of fiber-reinforced asphalt concrete mixes. The evaluation results showcased the dominance of the ANN model in predicting MS, with R values of 0.9287 and 0.9126, MAE values of 1.7527 and 1.8702, RMSE values of 2.3305 and 2.4438, RAE values of 32.51% and 39.64%, and RRSE values of 37.30% and 43.59% for the training and validation phases, respectively. Sensitivity analysis further emphasized the pivotal role of AC and the comparable sensitivity of carbon fiber in predicting MS in carbon fiber asphalt concrete mixes. Furthermore, the prediction of MS in glass fiber asphalt concrete was pursued in a subsequent study by Upadhya et al. [13]. For this endeavor, models including ANN, RF, DT, and Adaptive Neuro-Fuzzy Inference System (ANFIS) were proposed. The results highlighted the superior performance of the ANFIS model over the remaining models, achieving R of 0.8347, RMSE of 2.7254, WI of 0.7310, NSE of 0.6951, and MAE of 1.8756. Sensitivity analysis was also performed based on the best-performing model, indicating the importance of glass fiber content and AC as significant factors in predicting MS in glass fiber-reinforced asphalt concrete mixes. These studies demonstrate the efficacy of applying artificial intelligence models to predict MS in asphalt concrete, particularly fiber-reinforced asphalt concrete mixes.

It can be seen, however, while the application of AI in civil engineering has shown remarkable success, its utilization in the domain of asphalt concrete containing basalt fiber remains relatively limited. Furthermore, extant studies concentrate on established algorithms such as ANN, RF, RT, and ANFIS, without delving into newly proposed algorithms or hybrid approaches that could potentially address these intricate challenges. Therefore, the focus of this research lies in the application of the Gradient Boosting algorithm to predict the MS of basalt fiber asphalt concrete. This algorithm has gained prominence due to its remarkable performance in handling complex and high-dimensional datasets, making it particularly suited for predicting intricate engineering parameters like MS. The core of this investigation is a database consisting of 128 samples, serving as the foundational dataset for the study's objectives. In constructing an effective model, the grid search method was harnessed in combination with a 5-fold cross-validation approach. For enhanced model interpretability, SHAP (SHapley Additive exPlanations) analysis is employed. This methodology imparts insights into the contribution of each individual feature to the model's predictions. Consequently, it elucidates the underlying relationships and mechanisms governing the MS of basalt fiber asphalt concrete. Furthermore, acknowledging the practical implications of this research, a Graphical User Interface (GUI) tailored to the requirements of material engineers is developed. This GUI serves as a user-friendly tool that allows engineers to input relevant parameters and obtain rapid predictions of MS, streamlining decision-making processes and facilitating the integration of our research findings into real-world engineering applications.

2. Database description and analysis

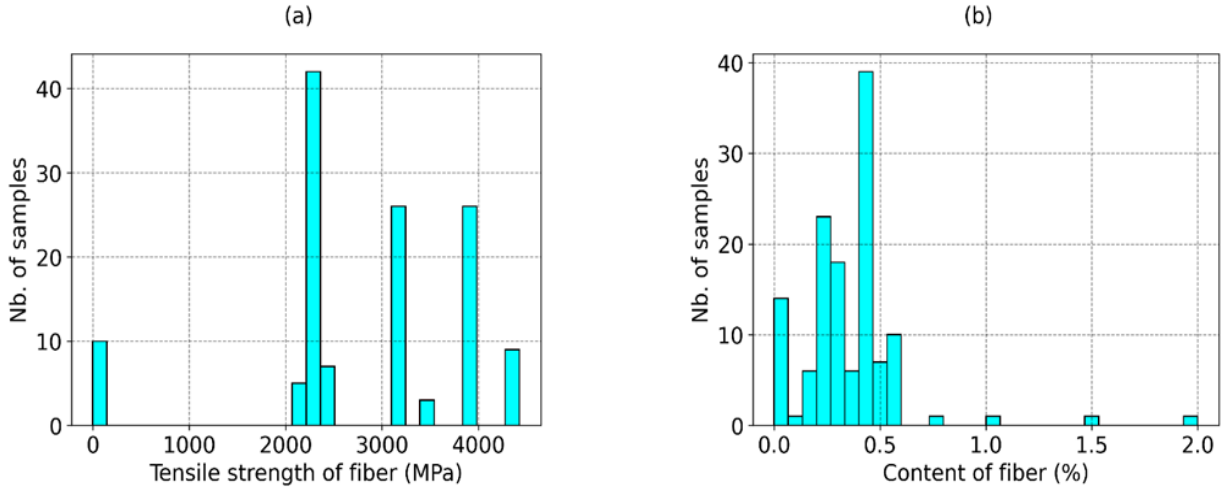
To predict the MS of basalt fiber asphalt concrete, a database comprising 128 experimental results was compiled from 18 works: Zheng et al. [14], Zhao et al. [15], Cheng et al. [16], Wang et al. [17], Cheng et al. [18], Wang et al. [19], Fan et al. [20], Tan et al. [21], Fan et al. [22], Chai et al. [23], Cheng et al. [24], Wang et al. [25], Zheng et al. [26], Zhao et al. [27], Wu et al. [28], Liu et al. [29], Huang et al. [30], and Morova [31]. For the asphalt

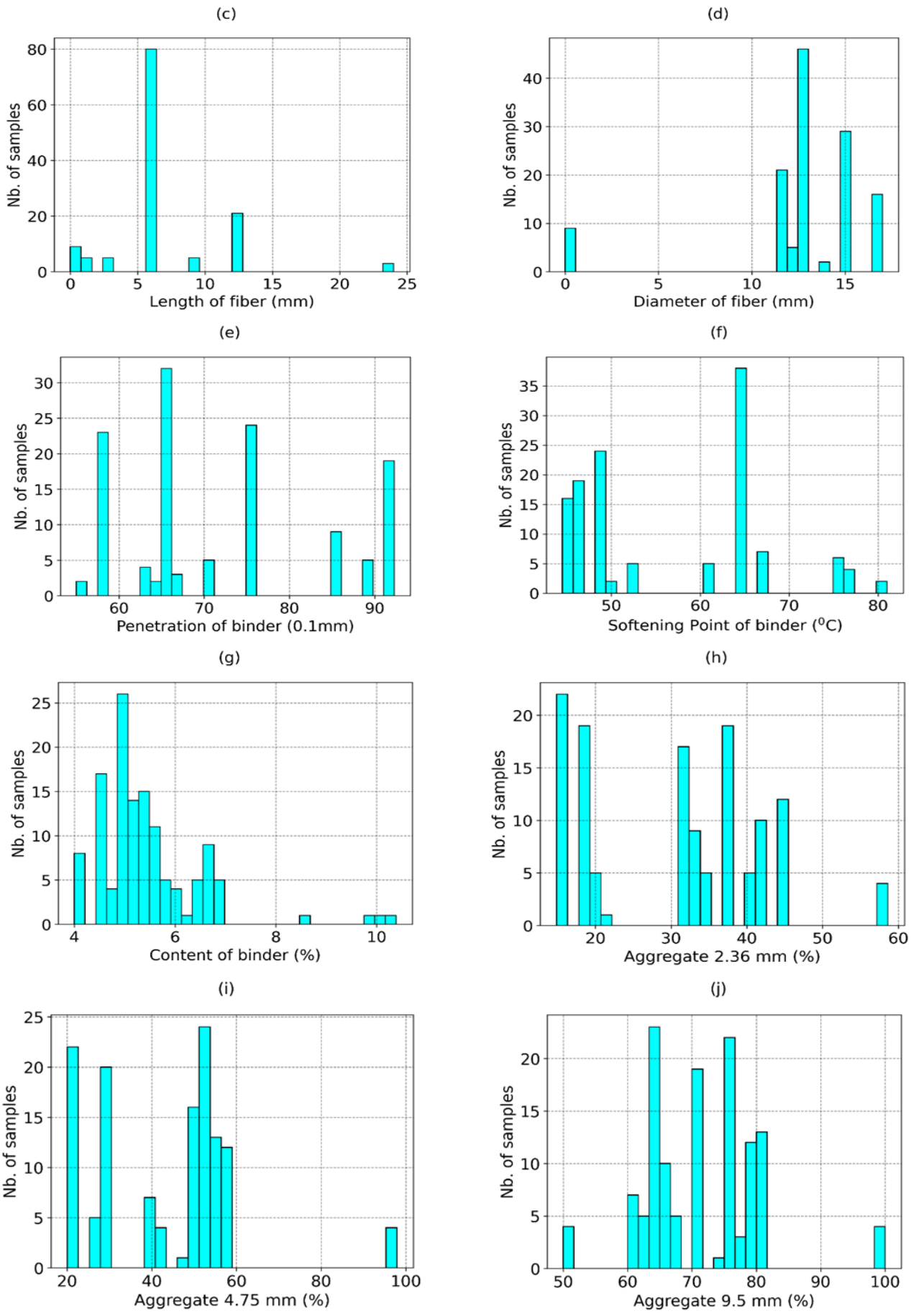
concrete samples using basalt fiber, three distinct sets of input parameters were considered: those related to basalt fiber, those related to asphalt binder, and those related to aggregate gradation.

The first group, in conjunction with basalt fiber, encompasses input parameters such as tensile strength of fiber (X_1), content of fiber (X_2), length of fiber (X_3), and diameter of fiber (X_4). These parameters are pivotal in determining the mechanical properties and overall performance of basalt fiber-reinforced asphalt concrete samples. The second group pertains to asphalt binder characteristics and includes input parameters like penetration of binder (X_5), softening point of binder (X_6), and content of binder (X_7). These properties directly influence the viscoelastic behavior, adhesion, and durability of the asphalt concrete mixture, impacting its performance under varying loading conditions and environmental contexts. The third group centers around aggregate gradation and incorporates input parameters related to aggregate 2.36 mm (X_8), aggregate 4.75 mm (X_9), and aggregate 9.5 mm (X_{10}). These sieve sizes were chosen due to their significant contributions to the overall composite framework of asphalt concrete, affecting stability, strength, and deformation resistance. A comprehensive statistical analysis of all corresponding input and output parameters with the database is presented in Table 1, including symbols, roles, units, minimum and maximum values, averages, and standard deviations.

Table 1. The input and output parameters used in the development of ML models

Parameter	Unit	Role	Average	Std	Min	Max
Tensile strength of fiber - X_1	(MPa)	Input	2800.66	1092.16	0.00	4425.00
Content of fiber - X_2	(%)	Input	0.34	0.25	0.00	2.00
Length of fiber - X_3	(mm)	Input	6.81	4.12	0.00	24.00
Diameter of fiber - X_4	(mm)	Input	12.74	3.88	0.00	17.00
Penetration of binder - X_5	(0.1mm)	Input	72.11	11.95	55.00	92.30
Softening Point of binder - X_6	(^{o}C)	Input	56.89	10.38	44.50	81.00
Content of binder - X_7	(%)	Input	5.43	1.06	4.00	10.39
Aggregate 2.36 mm - X_8	(%)	Input	30.64	11.64	14.90	58.62
Aggregate 4.75 mm - X_9	(%)	Input	43.22	16.90	20.08	97.88
Aggregate 9.5 mm - X_{10}	(%)	Input	71.34	9.00	50.00	100.00
Marshall stability - MS	kN	Output	10.65	2.58	5.69	17.70




(c)
Nb. of samples
Length of fiber (mm)
(d)
Nb. of samples
Diameter of fiber (mm)
(e)
Nb. of samples
Penetration of binder (0.1mm)
(f)
Nb. of samples
Softening Point of binder (°C)
(g)
Nb. of samples
Content of binder (%)
(h)
Nb. of samples
Aggregate 2.36 mm (%)
(i)
Nb. of samples
Aggregate 4.75 mm (%)
(j)
Nb. of samples
Aggregate 9.5 mm (%)

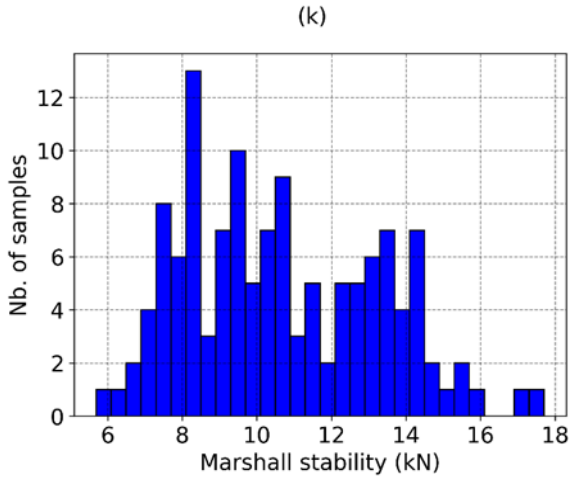


Fig. 1. Distribution of input and output parameters of the database

To gain a comprehensive understanding of the collected database's outcomes, the distribution of parameters within the database is investigated and depicted in Table 2. Furthermore, the correlations between input parameters and between input parameters and output parameters are crucial foundations for assessing the significance and importance of the input variables. Positive values denote positive correlations, while negative values signify negative correlations. Analysis from Table 2 reveals that certain pairs of input parameters exhibit relatively high correlations (such as X_1 and X_4, X_8 and X_9). However, to comprehensively assess the impact of the collected parameters on prediction values, this study preserves all ten input parameters.

The database is divided into two subsets, with 70% of the data utilized for developing machine learning models, termed the training database, and the remaining 30% employed for testing and evaluating the accuracy of the developed models, referred to as the testing database. The essence of partitioning 70% of the data for model training is to separate the testing and training portions. This implies that the testing data (30%) still needs to be discovered to the models beforehand. Hence, the predictive capability of the models can be objectively and accurately evaluated through the testing part. Furthermore, as indicated by some studies [32,33], the 70/30 ratio for data division is reasonable to ensure the reliability and representativeness of the data for machine learning models during both the training and testing processes.

Table 2. Linear statistical correlation coefficient R between parameters

R	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	MS
X_1	1	0.14	0.60	0.86	0.13	-0.24	0.19	0.14	0.12	0.06	0.21
X_2		1	0.10	0.39	-0.15	-0.04	0.03	-0.22	-0.22	-0.22	-0.26
X_3			1	0.55	-0.04	-0.06	0.40	0.38	0.37	0.23	0.01
X_4				1	0.02	-0.16	0.06	-0.09	-0.10	-0.10	0.09
X_5					1	-0.75	-0.28	0.47	0.41	0.20	0.54
X_6						1	0.32	-0.41	-0.26	0.07	-0.25
X_7							1	0.27	0.45	0.58	0.10
X_8								1	0.95	0.59	0.52
X_9									1	0.75	0.59
X_{10}										1	0.43
MS											1

3. Methods

3.1. Gradient Boosting method

Gradient Boosting [34] is a prominent machine learning technique that belongs to the family of ensemble learning methods. It has gained significant popularity due to its exceptional predictive capabilities across a wide range of applications, from classification and regression tasks to ranking and recommendation systems. The fundamental idea behind Gradient Boosting is to combine the predictions of multiple weak learners, often decision trees, into a single strong predictive model. This combination of diverse models leads to enhanced accuracy and robustness, making Gradient Boosting a valuable tool in data analysis and predictive modeling.

Unlike some other ensemble methods, Gradient Boosting focuses on minimizing the errors of the previous models in a sequential manner. It does so by assigning higher weights to the data points that were poorly predicted by the previous models. This emphasis on learning from mistakes ensures that the subsequent models are tailored to rectify the shortcomings of their predecessors. By iteratively refining the predictions and gradually adjusting the model's weights, Gradient Boosting adapts itself to the nuances of the data, resulting in improved predictive performance.

One of the distinguishing features of Gradient Boosting is its ability to handle both categorical and numerical data, making it versatile and applicable to a wide variety of databases. Advantages of Gradient Boosting include its robustness to outliers and the flexibility to handle different types of data. It's also less prone to overfitting compared to individual weak learners. However, like any machine learning technique, Gradient Boosting comes with its own set of trade-offs. It can be computationally intensive, especially when dealing with a large number of weak learners or complex databases. Additionally, parameter tuning is crucial to ensure optimal performance and avoid issues like overfitting.

3.2. Cross-validation

Cross-validation is a fundamental technique in machine learning and statistics that plays a pivotal role in assessing predictive models' performance and generalization capability. It addresses the challenge of evaluating a model's performance on new, unseen data, which is crucial to avoid overfitting and ensure reliable predictions [35].

The primary objective of cross-validation is to simulate the model's performance on independent data samples. This is achieved by partitioning the available database into multiple subsets, typically called "folds." The cross-validation process then involves training the model on a subset of the data (known as the training database) and evaluating its performance on the remaining subset (known as the validation or testing database). This procedure is repeated for each fold, ensuring that each data point is used for training and validation across different folds [36].

The most common form of cross-validation is k-fold cross-validation, where the database is divided into k subsets of approximately equal size. The training and validation process is repeated k times, with each subgroup serving as the validation set once while the remaining k-1 subsets are used for training. The results from each fold are then averaged to provide a more robust assessment of the model's performance [37].

Cross-validation offers several benefits. It provides a more accurate estimate of a model's actual performance, as it reduces the impact of variability that can arise from a single training-validation split. Moreover, cross-validation allows for better model tuning by identifying potential issues like overfitting or underfitting. It also helps to avoid the bias that may occur when using a fixed validation database [35].

3.3. Performance indices of models

The performance evaluation criteria of the ML model used in this study are RMSE, MAE, R^2, and MAPE. Details of these criteria and their formulae are given in Table 3.

Table 3. Details of the performance evaluation criteria of the ML model used in this study

Criterion	Abbreviation	Formulae
Root mean squared error	RMSE	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(M_i - \hat{M}_i\right)^2}$
Mean absolute error	MAE	$MAE = \frac{1}{N}\sum_{i=1}^{N}\left\|M_i - \hat{M}_i\right\|$
Coefficient of determination	R^2	$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(M_i - \hat{M}_i\right)^2}{\sum_{i=1}^{N}\left(M_i - \bar{M}\right)^2}$
Mean absolute percentage error	MAPE	$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left\|\frac{M_i - \hat{M}_i}{M_i}\right\| \times 100\%$

where N is the number of samples, M_i is the actual value, $\hat{M}_i$ is the predicted value, and $\bar{M}$ is the average of the actual values.

4. Results and Discussion

4.1. Hypeparameter tuning

Hyperparameter tuning is a critical phase in the development of machine learning models, and it plays a pivotal role in achieving optimal model performance. This process involves finding the ideal combination of hyperparameters that govern various aspects of the model's behavior. By adjusting these hyperparameters, aim to strike a balance between model effectiveness and complexity, ultimately leading to better generalization on unseen data [38]. The choice of hyperparameters greatly influences the model's performance and its ability to capture complex relationships within the data. For the GB model, there are several hyperparameters that significantly impact the predictive accuracy of the model. Four hyper-parameters, namely “learning rate”, “n_estimator”, “max depth”, and “min_samples_leaf” that are considered to affect the performance of the model, were selected for optimization [39]. The search domain of hyperparameters is also an important factor in solving prediction problems. The parameter “n_estimator” characterizes the number of unit trees used in the GB model, too few trees will affect the predictive power of the entire model, while too many trees can lead to overfitting means that the predictive model is too good for the learned data but not so well for new data. In addition, increasing "n_estimator" can significantly affect the computational speed of the model, making the model very complex and requiring a lot of computer memory to operate. Next, “learning_rate” represents the learning rate of the tree, with the higher the value, the faster the learning rate and vice versa. Usually with tree models, this parameter is often used and searched in the range from 0.01 to 0.3. “Max_depth” is also an important hyperparameter in tree models in general, because it reflects the depth of the tree, which means the complexity of the tree model structure when performing forecasting. The larger this hyperparameter, the greater the height, or depth, of the tree and the more complex the tree structure. Depending on the complexity of the problem, max_depth can have values from 1, 2, or more than 20. Within the GB model, "min_samples_leaf" is a good tuning parameter for decision tree depth tuning. This parameter defines the minimum number of samples required in each leaf of the decision tree used in the GB model. The use of the parameter "min_samples_leaf" is to limit the phenomenon of "overfitting" of the model, especially when the decision leaves are too small. When the leaves are too small, the model can learn the noises in the training set, leading to making false predictions on the testing set. So, using this parameter can help to reduce “overfitting” and improve the accuracy of the model. The search

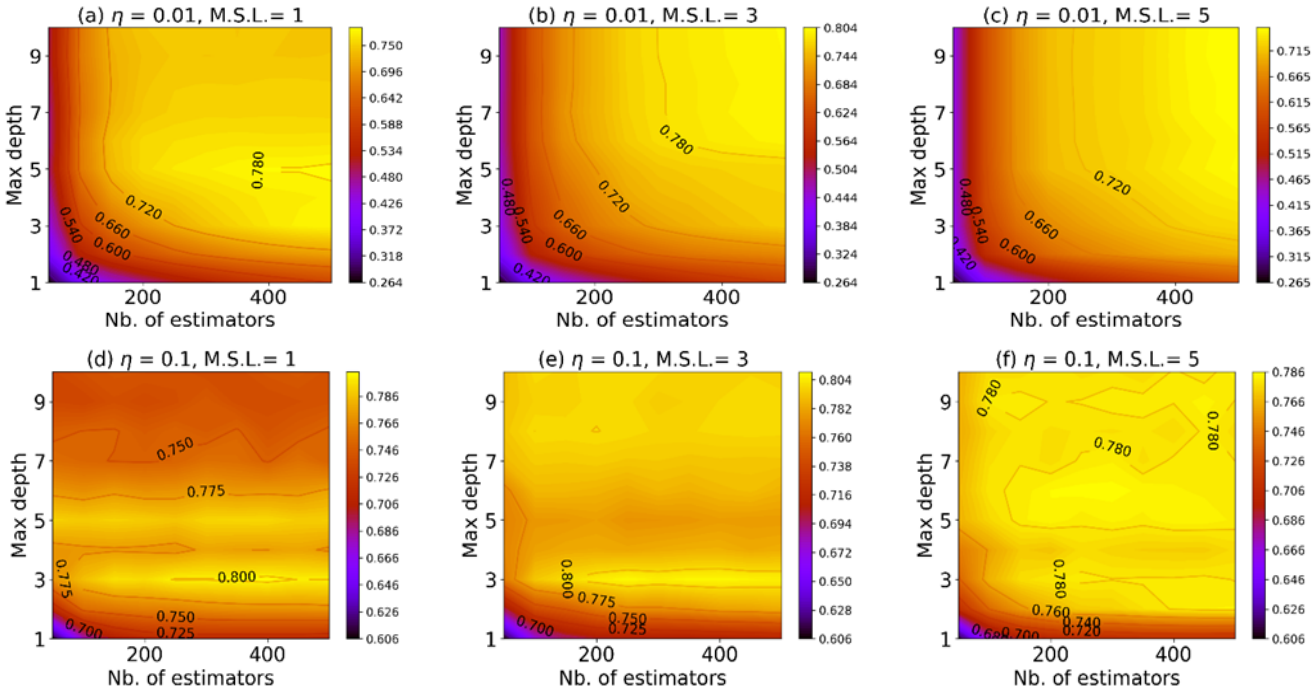
domains of these hyperparameters are shown in Table 4. The remaining hyperparameters use the default values of the Python source code.

Two critical evaluation criteria, R^2 and RMSE, are employed for hyperparameter tuning, and both computed on a validation database. These criteria are informed choices due to their ability to measure the model's predictive accuracy and error magnitudes. The selection of these criteria is refined through a 5-fold cross-validation process applied to the training database. Reserving the testing database ensured that the model's evaluation was based on previously unseen data, enhancing the reliability and robustness of the reported results. The outcomes of the hyperparameter tuning process are represented in Fig. 2 and Fig. 3. These visualizations offer an insightful depiction of how changes in hyperparameter values impact the model's performance metrics. By examining these figures, a deeper understanding of the intricate relationships between hyperparameters and model outcomes can be observed.

Throughout the process of hyperparameter tuning, notable insights were gained from the study. A significant increase in accuracy, with an R^2 value of 0.816, was achieved on the validation database when employing a learning rate of 0.3 and a min_samples_leaf value of 3. Moreover, heightened accuracy is observed as the max depth exceeds 5, underscoring the model's increased performance potential with deeper trees. Remarkably, the number of estimators appears to hold lesser significance, exerting minimal influence on accuracy when the learning rate remains fixed at 0.3. In accordance with the criteria of R^2 and RMSE, six distinct GB models were identified and selected due to their commendable performance on the validation database. However, the performance of these models on the testing database remains uncertain. To comprehensively assess their effectiveness, a comparison and evaluation of these six models have been designated based on their predictive accuracy using an independent testing database. This analytical step aims to offer deeper insights into the models' ability to generalize and their robustness across diverse subsets of data.

Table 4. Search domain of hyperparameters of the model GB

n_estimators	learning_rate	max_depth	min_samples_leaf
50 – 500	0.01 – 0.5	1 – 10	1 – 5



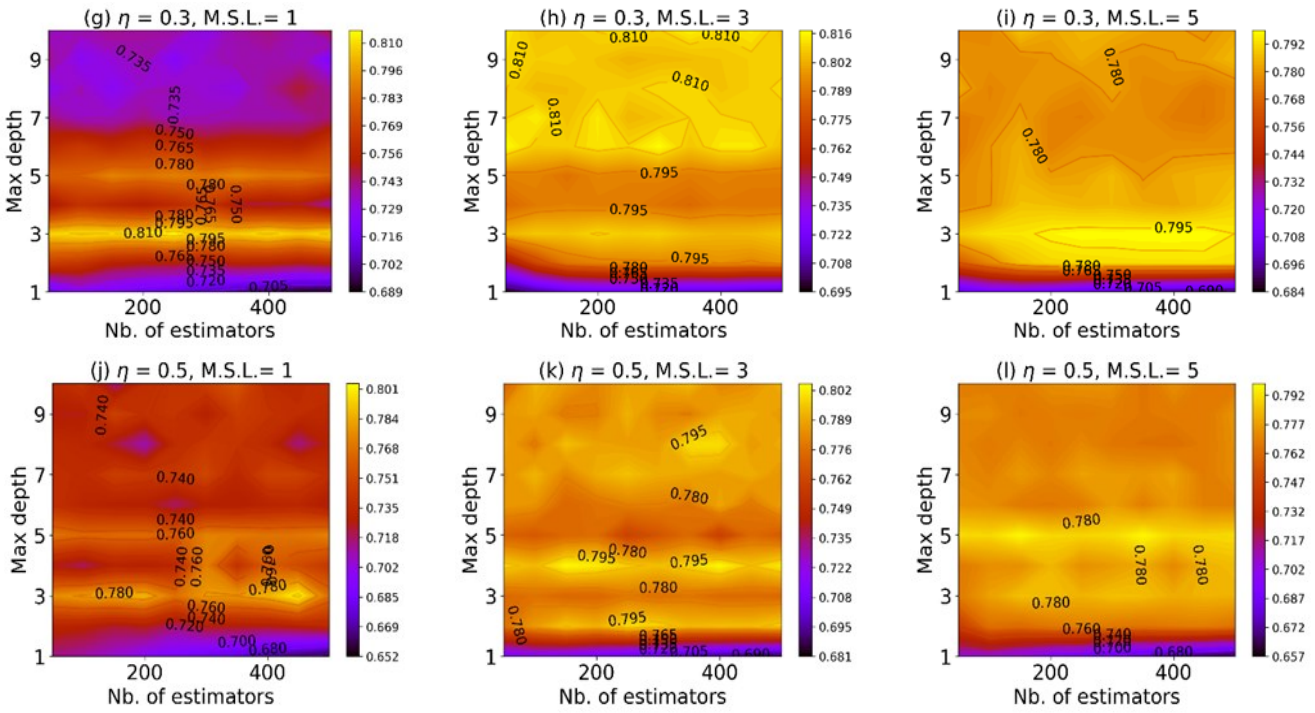
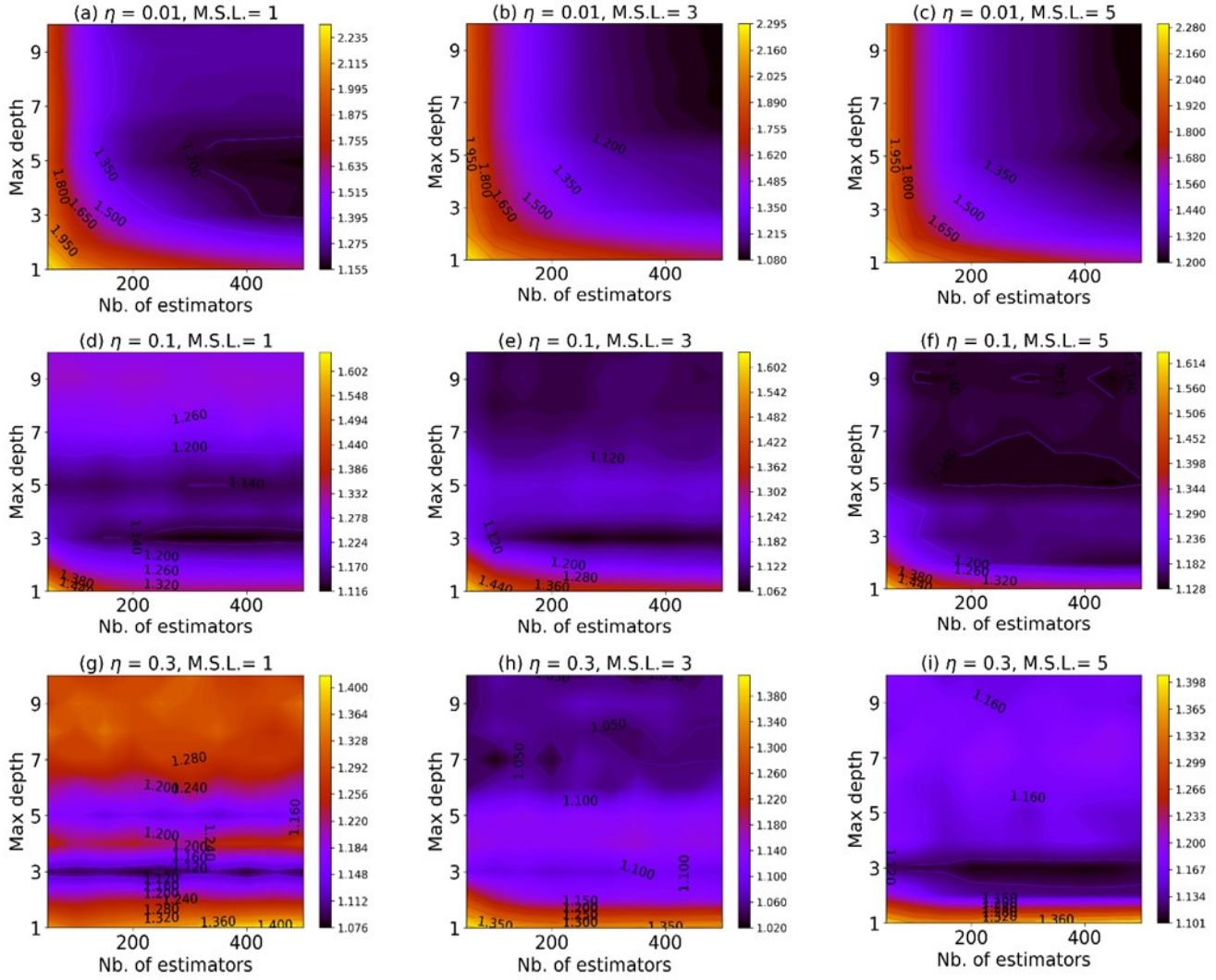


Fig. 2. Result of GB model optimization evaluated according to R^2



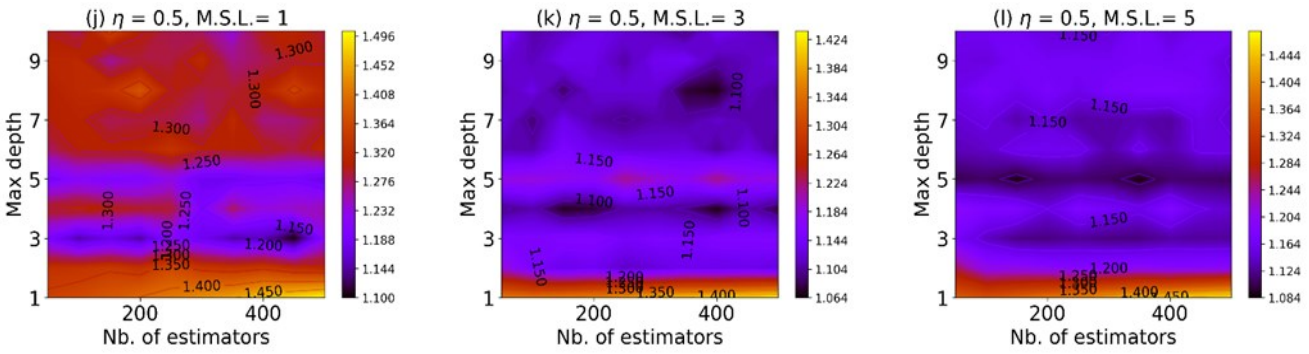


Fig. 3. Result of GB model optimization evaluated according to RMSE

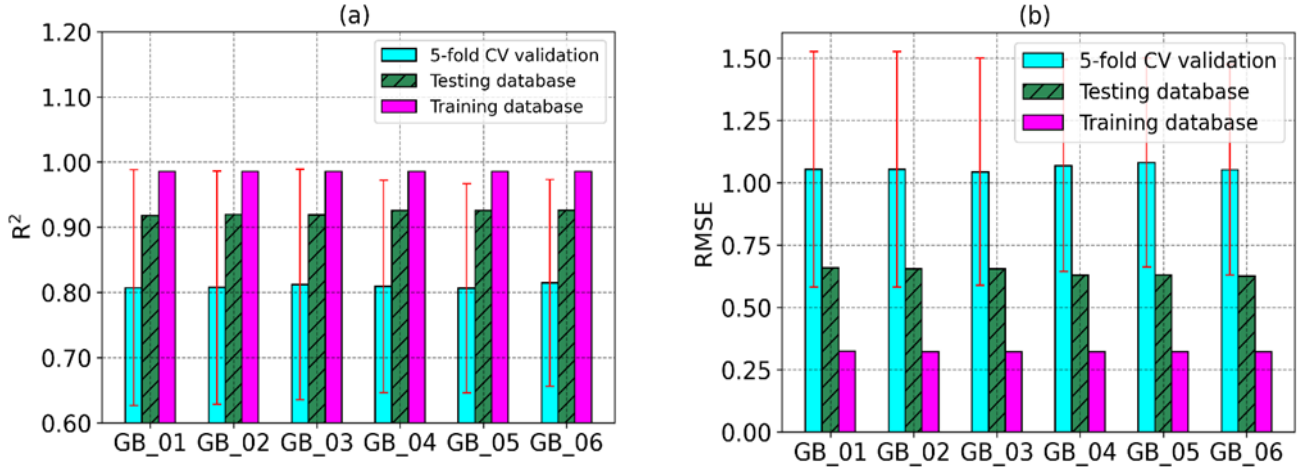


Fig. 4. Performance comparison of 6 GB models

Table 5. Summary of prediction results of 6 GB models according to RMSE and R^2

		GB_01	GB_02	GB_03	GB_04	GB_05	GB_06
Learning rate (η)		0.3	0.3	0.3	0.3	0.3	0.3
Max depth		7	7	7	6	6	6
Min samples leaf (M.S.L)		3	3	3	3	3	3
N estimators (N_e)		50	200	350	200	300	450
RMSE-5 fold CV	Mean	1.027	1.059	1.065	1.068	1.062	1.053
	Std	0.485	0.465	0.464	0.417	0.421	0.425
R^2-5 fold CV	Mean	0.813	0.806	0.805	0.810	0.812	0.813
	Std	0.183	0.179	0.177	0.159	0.160	0.163
RMSE – Train		0.323	0.322	0.322	0.322	0.322	0.322
R^2 – Train		0.985	0.986	0.986	0.986	0.986	0.986
RMSE – Test		0.661	0.658	0.655	0.628	0.631	0.615
R^2 – Test		0.918	0.918	0.919	0.926	0.925	0.928

Fig. 4 illustrates the values of two statistical criteria, R^2 and RMSE, employed to assess the accuracy of the six selected models across the validation, training, and testing databases.

Table 5 provides a comprehensive breakdown of the hyperparameter values for the chosen models, along with their corresponding predictive outcomes for the RMSE and R^2 statistics. All six models exhibit commendable performance, displaying minimal discrepancies in their efficacy across the databases. Nevertheless, the GB_06 model marginally outperforms its counterparts on the testing database, showcasing slightly enhanced performance. This observation underscores the crucial significance of evaluating models based on the testing database rather than solely relying on validation data. This practice ensures the validation of a model's generalizability

to new, unseen data points, thus reflecting its true predictive capabilities. As a result, model GB_06 was selected as the best model to present the typical results in the following section

4.2. Representative prediction results

In this section, typical results of the GB_06 model are presented to highlight the model's performance. The regression graph between the actual and predicted results for the training and testing databases is shown in Figs. 5a and 5b. A linear fit is also applied and graphed in each case, with R^2, RMSE, MAE, and MAPE values shown. The R^2 values are reported as 0.986 and 0.928 consecutively, while the RMSE values stand at 0.322 kN and 0.615 kN for the training and testing databases, respectively. Furthermore, the MAE values are measured at 0.065 kN and 0.411 kN, and the MAPE values at 0.009 and 0.041 for the training and testing databases. The results show that the proposed GB_06 model can predict the MS of basalt fiber asphalt concrete well.

Furthermore, the strong correlation between the predicted MS values and the actual MS values is confirmed through the representation of error plots. Notably, the error values for the training database are relatively modest. The majority of samples exhibit errors approximating zero, with the maximum concentration of errors situated around this range. Only two samples deviate from the range of [-1; 1] kN. Upon assessing the testing database, a substantial portion of samples also exhibit errors within the range of [-1; 1] kN. Merely three samples stray outside this range. These error distributions collectively substantiate the robust predictive capabilities of the GB_06 model (Figs. 6a and 6b).

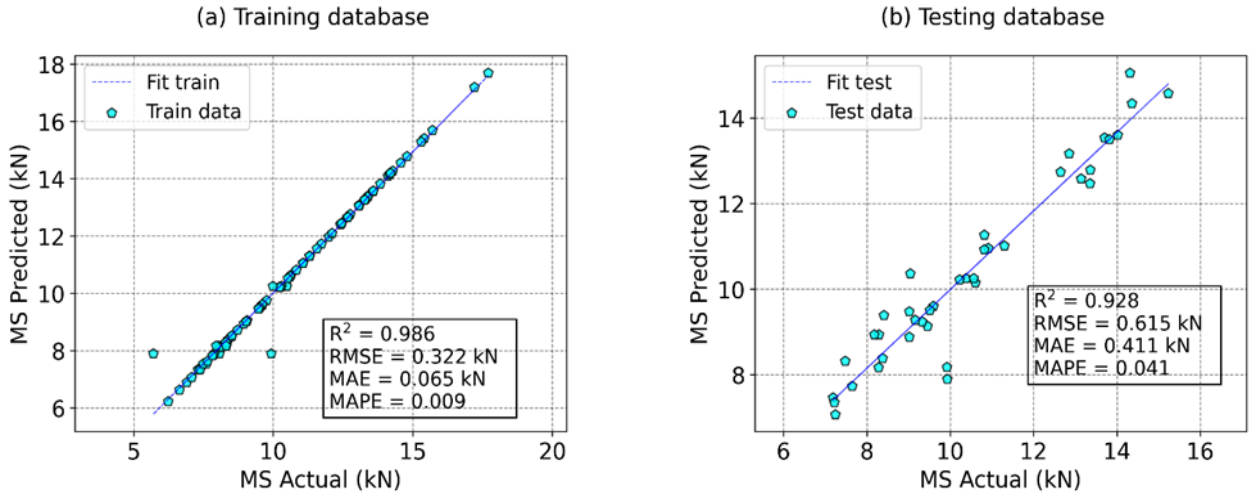


Fig. 5. Regression charts between MS prediction results and the experimental values (a) training database, (b) testing database

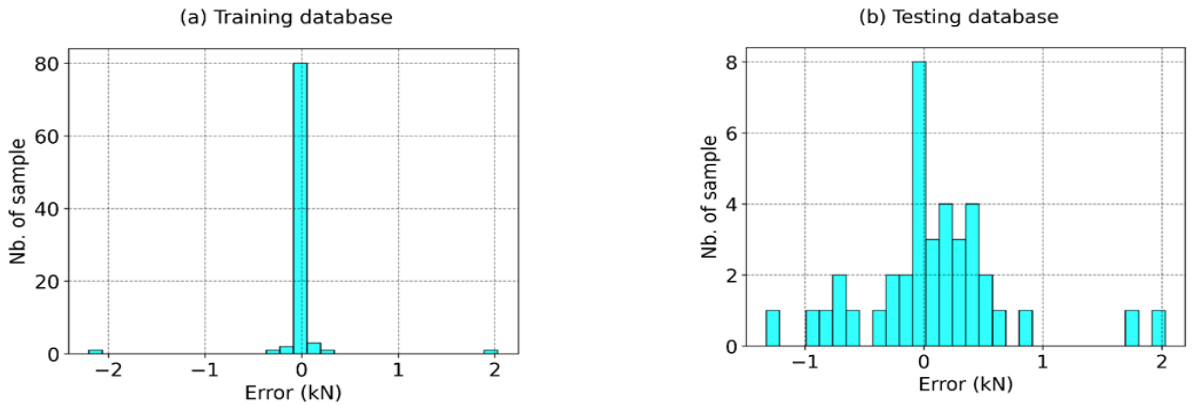


Fig. 6. Error between the MS prediction results and the experimental values: (a) training database, (b) testing database

4.3. Comparison with other machine learning models

To substantiate the performance of the selected model, this section undertakes a comparative analysis between the performance of the GB_06 model and that of several other models. Six alternative machine learning models are proposed for evaluation: specifically, the Hist Gradient Boosting (HGB), Light Gradient Boosting (LGB), K-nearest neighbors (KNN), Decision Tree (DT), Random Forest (RF), and Linear Regression (LR) models. The performance of these machine learning models is evaluated using four statistical metrics considered in this study. Column charts in Fig. 7 visually represent the values of these metrics for both the training and testing databases. Furthermore, Table 6 provides an in-depth presentation of the corresponding metric values.

Based on these metrics, it becomes evident that the GB_06 model outperforms its counterparts, exhibiting commendable performance across both the training and testing databases. Among the six models employed for comparison, the DT model displays strong performance on the training database. However, its performance substantially diminishes on the testing database, indicating a potential overfitting issue. The remaining models show lower prediction performance on the training database than on the testing database. This discrepancy suggests that these models may need to align better with the input data, or the training methodologies might be suboptimal. These findings underscore the credibility of the GB_06 model as a prudent choice for predicting the MS of basalt fiber asphalt concrete.

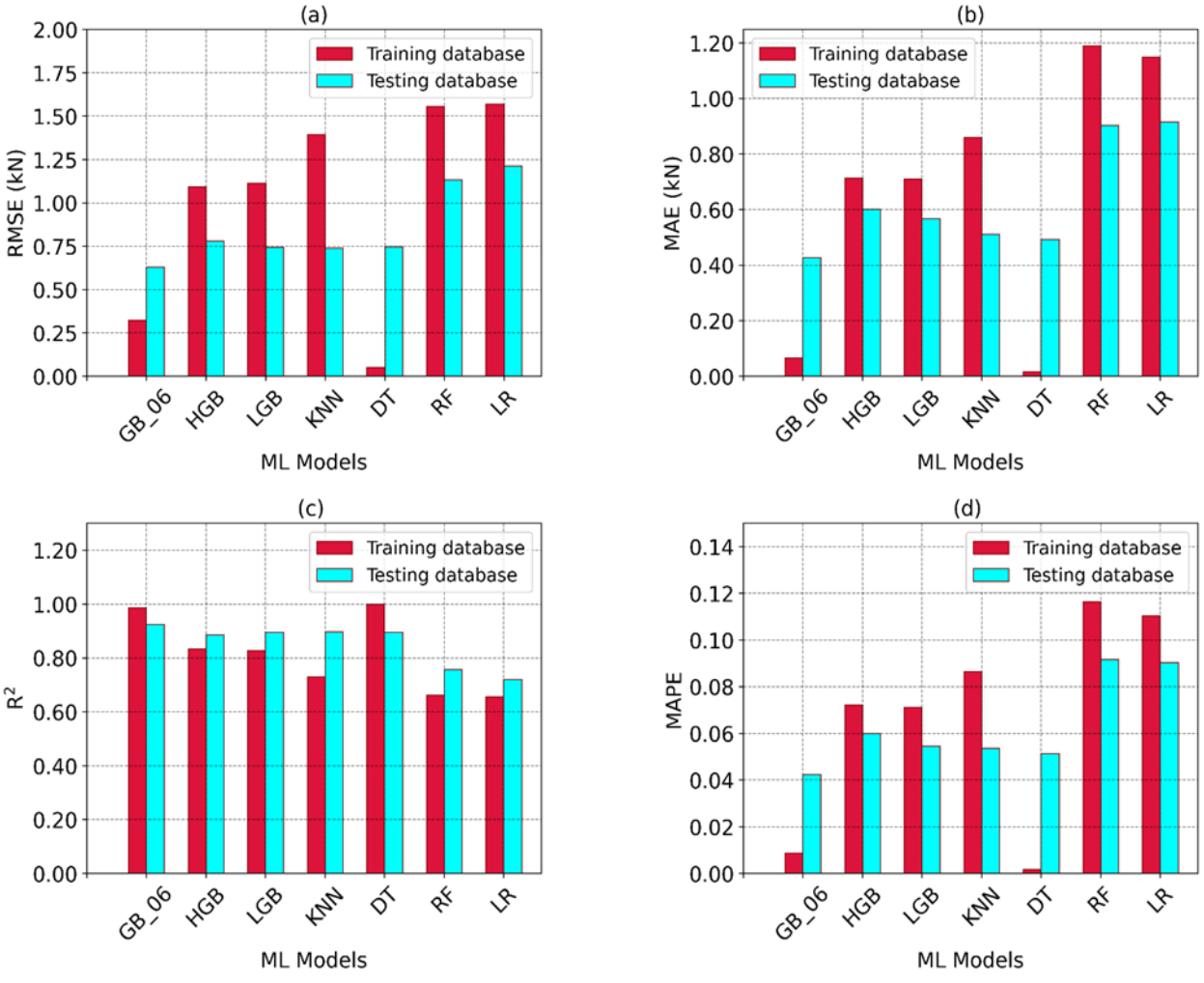


Fig. 7. Performance comparison of ML models

Table 6. Values of statistical criteria of ML models

	RMSE (kN)	MAE (kN)	R^2	MAPE	RMSE (kN)	MAE (kN)	R^2	MAPE
	Training database				Testing database			
GB_06	0.322	0.065	0.986	0.009	0.615	0.411	0.928	0.041
HGB	1.091	0.714	0.834	0.072	0.780	0.602	0.885	0.060
LGB	1.114	0.710	0.827	0.071	0.745	0.567	0.895	0.055
KNN	1.395	0.859	0.729	0.087	0.739	0.511	0.897	0.054
DT	0.052	0.015	0.999	0.002	0.746	0.493	0.895	0.052
RF	1.562	1.199	0.659	0.118	1.121	0.901	0.763	0.092
LR	1.569	1.148	0.656	0.110	1.215	0.915	0.721	0.090

4.4. Analyzing the factors influencing Marshall stability using the SHAP value technique

The intricate nature of machine learning models often results in their portrayal as black boxes, necessitating a comprehensive understanding of their inherent characteristics for seamless practical application. This comprehension facilitates the model's seamless integration into practical scenarios, enhancing its pragmatic utility. This study harnesses the SHAP (SHapley Additive exPlanations) value analysis technique as a pivotal method for conducting sensitivity analysis. This technique is instrumental in quantifying the influence of each variable on the output results, thereby providing critical insights into the model's inner workings.

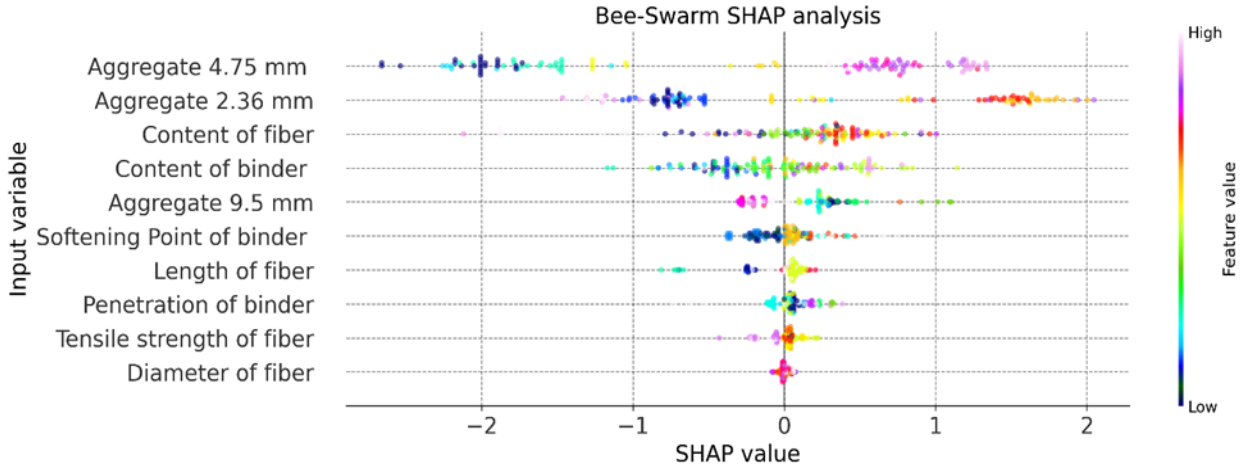


Fig. 8. The SHAP value graph analyzes the importance of input parameters

The interpreted results, depicted in Fig. 8, clearly emphasize that the content aggregate 4.75mm is the most influential variable, followed by the content aggregate 2.36mm, the content of fiber, the content of binder, and the content aggregate 9.5mm in descending order of impact. It is discerned that the grade of the aggregates a paramount role in its composition, yielding a substantial impact on the value of MS. Moreover, the presence of fiber and resin content is notably influential. It is noteworthy to underscore that additional variables such as binder quality (softening point of binder and penetration of binder) and fiber parameters (such as length of the fiber, the tensile strength of fiber, and diameter of binder) appear to exert a comparably lesser influence on MS. These variables while contributing to the overall model dynamics, demonstrate a relatively diminished role in shaping the resultant MS values. The asphalt mixture has a low content passing through the 4.75mm sieve, which means the mixture has a high proportion of coarse aggregate

content. Asphalt mixtures with coarse aggregates will have a stone-on-stone skeleton in the mixture and form a solid coarse aggregate framework. This aggregate skeleton will create a sturdy structure for the asphalt mixture, increasing the bearing capacity of the asphalt mixture [40]. In summation, the SHAP value analysis technique is potent in unraveling the intricate interplay of variables within the model's architecture. The results from this analysis furnish insights into the hierarchy of variable influence, with certain factors like mixture grade, fiber content, and resin content exerting a more profound effect on the output MS. Conversely, variables such as binder quality and fiber parameters exhibit a comparatively subdued influence. This comprehensive understanding augments the interpretability and applicability of the model, laying a robust foundation for informed decision-making in real-world applications.

4.5. Practical application

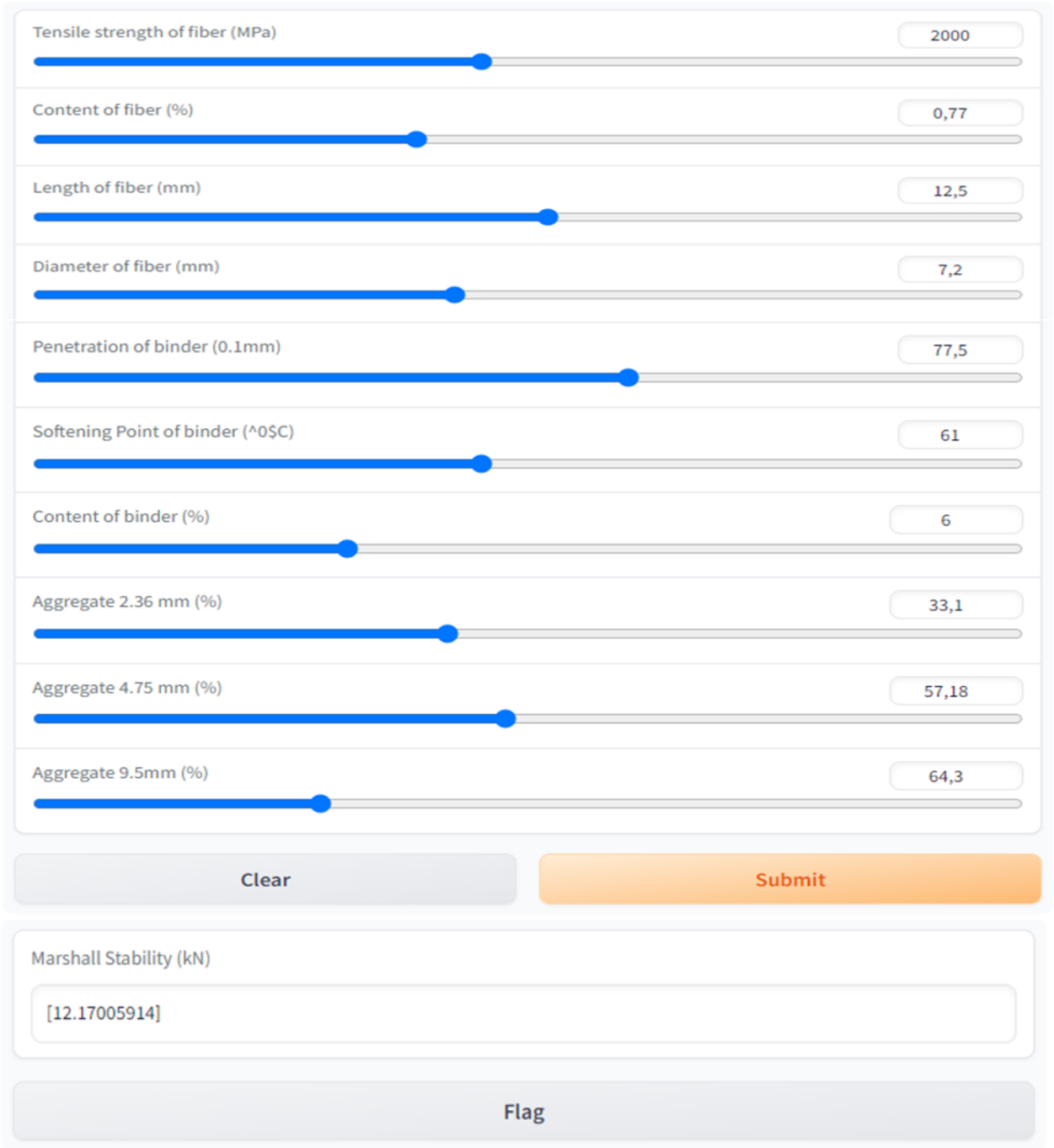


Fig. 9. GUI interface for users to quickly predict the MS of asphalt concrete

This section introduces the Graphic User Interface (GUI) for predicting the MS of basalt fiber asphalt concrete. Applying the Gradient Boosting model within the GUI framework simplifies and enhances the process of MS prediction. The GB model utilizes aggregate gradation data, fiber characteristics, and binder properties to construct a predictive model. Subsequently, this model is integrated into the GUI, allowing users to input information about the basalt fiber asphalt concrete and relevant MS parameters for prediction. Fig. 9 illustrates the main interface of the GUI, designed to be user-friendly and straightforward. Users can input values for input parameters, and the MS predicted of basalt fiber asphalt concrete is instantly displayed upon clicking the "Submit" button.

5. Conclusion

In this study, an exploration of machine learning methodologies for predicting the Marshall Stability (MS) of basalt fiber asphalt concrete has been presented. The application of the Gradient Boosting algorithm, coupled with extensive dataset analysis, underscores the effectiveness of this approach in accurate MS predictions. The integration of grid search and 5-fold cross-validation techniques further enhances the model's reliability. The use of SHAP analysis offers valuable insights into the underlying factors influencing MS, contributing to the interpretability of the predictive models.

The development of a user-friendly Graphical User Interface (GUI) not only underscores the practical applicability of the research but also demonstrates its potential for real-world implementation by material engineers. This research not only contributes to the field of asphalt concrete design but also showcases the underexplored potential of AI techniques in addressing engineering challenges.

Acknowledgment

This research is funded by the University of Transport Technology, Thanh Xuan, Hanoi, Vietnam (UTT), under grant number DTTD2022-18.

References

[1] W. Jiang, Y. Huang, A. Sha. (2018). A review of eco-friendly functional road materials. *Construction and Building Materials*, 191, 1082-1092.

[2] D. Zheng, W. Song, J. Fu, G. Xue, J. Li, S. Cao. (2020). Research on mechanical characteristics, fractal dimension and internal structure of fiber reinforced concrete under uniaxial compression. *Construction and Building Materials*, 258, 120351.

[3] H. Jamshaid, R. Mishra. (2016). A green material from rock: basalt fiber – a review. *The Journal of The Textile Institute*, 107(7), 923-937.

[4] A. Sha, Z. Liu, W. Jiang, L. Qi, L. Hu, W. Jiao, D.M. Barbieri. (2021). Advances and development trends in eco-friendly pavements. *Journal of Road Engineering*, 1, 1-42.

[5] A. Alfalah, D. Offenbacker, A. Ali, C. Decarlo, W. Lein, Y. Mehta, M. Elshaer. (2020). Assessment of the impact of fiber types on the performance of fiber-reinforced hot mix asphalt. *Transportation Research Record*, 2674(4), 337-347.

[6] E. Ozgan. (2011). Artificial neural network based modelling of the Marshall Stability of asphalt concrete. *Expert Systems with Applications*, 38(5), 6025-6030.

[7] S.A.R. Shah, M.K. Anwar, H. Arshad, M.A. Qurashi, A. Nisar, A.N. Khan, M. Waseem. (2020). Marshall stability and flow analysis of asphalt concrete under progressive temperature conditions: An application of advance decision-making approach. *Construction and Building Materials*, 262, 120756.

[8] F. Aslam, M.A. Elkotb, A. Iqtidar, M.A. Khan, M.F. Javed, K.I. Usanova, M.I. Khan, S. Alamri, M.A. Musarat. (2022). Compressive strength prediction of rice husk ash using multiphysics

genetic expression programming. *Ain Shams Engineering Journal*, 13(3), 101593.

[9] A.A. Milad, I. Adwan, S.A. Majeed, Z.A. Memon, M. Bilema, H.A. Omar, M.G. Abdolrasol, A. Usman, N.I.M. Yusoff. (2021). Development of a hybrid machine learning model for asphalt pavement temperature prediction. *IEEE Access*, 9, 158041-158056.

[10] A. Azarhoosh, S. Pouresmaeil. (2020). Prediction of Marshall mix design parameters in flexible pavements using genetic programming. *Arabian Journal for Science and Engineering*, 45, 8427-8441.

[11] A. Upadhya, M.S. Thakur, P. Sihag, R. Kumar, S. Kumar, A. Afeeza, A. Afzal, C.A. Saleel. (2023). Modelling and prediction of binder content using latest intelligent machine learning algorithms in carbon fiber reinforced asphalt concrete. *Alexandria Engineering Journal*, 65, 131-149.

[12] A. Upadhya, M.S. Thakur, P. Sihag. (2022). Predicting Marshall stability of carbon fiber-reinforced asphalt concrete using machine learning techniques. *International Journal of Pavement Research and Technology*, 17, 102-122.

[13] A. Upadhya, M.S. Thakur, N. Sharma, P. Sihag. (2022). Assessment of soft computing-based techniques for the prediction of marshall stability of asphalt concrete reinforced with glass fiber. *International Journal of Pavement Research and Technology*, 15, 1366-1385.

[14] Y. Zheng, Y. Cai, G. Zhang, H. Fang. (2014). Fatigue property of basalt fiber-modified asphalt mixture under complicated environment. *Journal of Wuhan University of Technology-Mater. Sci. Ed*, 29, 996-1004.

[15] H. Zhao, B. Guan, R. Xiong, A. Zhang. (2020). Investigation of the performance of basalt fiber reinforced asphalt mixture. *Applied Sciences*, 10(5), 1561.

[16] Y. Cheng, D. Yu, Y. Gong, C. Zhu, J. Tao, W. Wang. (2018). Laboratory evaluation on performance of eco-friendly basalt fiber and diatomite compound modified asphalt mixture. *Materials*, 11(12), 2400.

[17] W. Wang, Y. Cheng, G. Tan. (2018). Design optimization of SBS-modified asphalt mixture reinforced with eco-friendly basalt fiber based on response surface methodology. *Materials*, 11(8), 1311.

[18] Y. Cheng, L. Li, P. Zhou, Y. Zhang, H. Liu. (2019). Multi-objective optimization design and test of compound diatomite and basalt fiber asphalt mixture. *Materials*, 12(9), 1461.

[19] W. Wang, Y. Cheng, P. Zhou, G. Tan, H. Wang, H. Liu. (2019). Performance evaluation of styrene-butadiene-styrene-modified stone mastic asphalt with basalt fiber using different compaction methods. *Polymers*, 11(6), 1006.

[20] W.-X. Fan, H.-G. Kang, Y.-X. Zheng. (2010). Experimental study of pavement performance of basalt fiber-modified asphalt mixture. *Journal of Southeast University (English Edition)*, 26, 614-617.

[21] G. Tan, W. Wang, Y. Cheng, Y. Wang, Z. Zhu. (2020). Master curve establishment and complex modulus evaluation of SBS-modified asphalt mixture reinforced with basalt fiber based on generalized sigmoidal model. *Polymers*, 12(7), 1586.

[22] W.X. Fan, S.F. Zhang, L.Q. Liu. (2013). Laboratory study of marshall of basalt fiber-modified asphalt mixture. *Applied Mechanics and Materials*, 256, 1851-1857.

[23] C. Chai, Y. Cheng, Y. Zhang, B. Zhu, H. Liu. (2020). Mechanical properties of crumb rubber and basalt fiber composite modified porous asphalt concrete with steel slag as aggregate. *Polymers*, 12(11), 2552.

[24] Y. Cheng, C. Chai, Y. Zhang, Y. Chen, B. Zhu. (2019). A new eco-friendly porous asphalt mixture modified by crumb rubber and basalt fiber. *Sustainability*, 11(20), 5754.

[25] W. Wang, Y. Cheng, H. Chen, G. Tan, Z. Lv, Y. Bai. (2019). Study on the performances of

waste crumb rubber modified asphalt mixture with eco-friendly diatomite and basalt fiber. *Sustainability*, 11(19), 5282.

[26] Y.X. Zheng, Y.C. Cai, Y.M. Zhang. (2011). Laboratory study of pavement performance of basalt fiber-modified asphalt mixture. *Advanced Materials Research*, 266, 175-179.

[27] L. Zhao, J. Chen, S. Wang. (2010). Using mineral fibers to improve asphalt and asphalt mixture behavior. *Traffic and Transportation Studies 2010*, pp 1352-1360.

[28] C. Wu, L. Li, Y. Cheng, Z. Gu, Z. Lv, R. Wang, B. Guan. (2020). Effect of Diatomite and Basalt Fibers on Pavement Performance and Vibration Attenuation of Waste Tires Rubber-Modified Asphalt Mixtures. *Mathematical Problems in Engineering*, 1-13.

[29] F. Liu, A. Dong, C. Liu, W. Wu. (2018). Mix design of asphalt mixture used for the waterproof and anti-cracking layer in the rainy area of South China. *Journal of Applied Biomaterials & Functional Materials*, 16, 112-118.

[30] T. Huang, J. Chen, M. Li, Y. Tang, S. Lv, H. Yu, H. Liu, W. Yan. (2021). Investigation on three-dimensional failure criterion of asphalt mixture considering the effect of stiffness. *Construction and Building Materials*, 285, 122431.

[31] N. Morova. (2013). Investigation of usability of basalt fibers in hot mix asphalt concrete. *Construction and Building Materials*, 47, 175-180.

[32] S. Salcedo-Sanz, R.C. Deo, L. Carro-Calvo, B. Saavedra-Moreno. (2016). Monthly prediction of air temperature in Australia and New Zealand with machine learning algorithms. *Theoretical and Applied Climatology*, 125, 13-25.

[33] A. Sharma, M.K. Goyal. (2015). Bayesian network model for monthly rainfall forecast. *2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, pp 241-246.

[34] J.H. Friedman. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 29(5), 1189-1232.

[35] R. Kohavi. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Ijcai, Montreal, Canada*, pp 1137-1145.

[36] C. Schaffer. (1993). Selecting a classification method by cross-validation. *Machine Learning*, 13, 135-143.

[37] Y. Bengio, Y. Grandvalet. (2003). No unbiased estimator of the variance of k-fold cross-validation. *Advances in Neural Information Processing Systems 16 (NIPS 2003)*.

[38] R. Bardenet, M. Brendel, B. Kégl, M. Sebag. (2013). Collaborative hyperparameter tuning. Proceedings of the 30th International Conference on Machine Learning, PMLR 28(2), pp 199-207.

[39] W. Zhang, C. Wu, H. Zhong, Y. Li, L. Wang. (2021). Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization. *Geoscience Frontiers*, 12(1), 469-477.

[40] E. Brown, R.B. Mallick, J.E. Haddock, J. Bukowski. (1997). Performance of stone matrix asphalt (SMA) mixtures in the United States. *National Center for Asphalt Technology*.