Rocznik Ochrona Środowiska

Volume 28 Year 2026 ISSN 2720-7501 pp. 329-348

https://doi.org/10.54740/ros.2026.023 open access

Received: April 2026 Accepted: May 2026 Published: June 2026

Hybrid Experimental–Machine Learning Framework for Environmental Optimization of Bio-Modified Bitumen

Abdullah Faiz Al Asmari

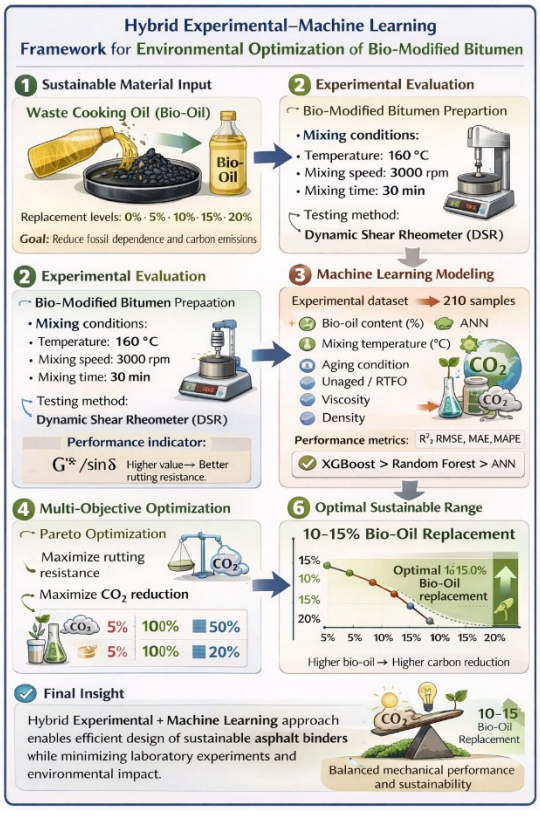
Civil Engineering Department, College of Engineering, King Khalid University, Abha, Saudi Arabia

https://orcid.org/0000-0001-5949-8155

corresponding author's e-mail: abdullahfaizalasmari@gmail.com

Abstract: The increasing need to use sustainable pavement materials has prompted the pursuit of alternatives to Petroleum bitumen made of bio-based materials. This work explores the environmental and mechanical behavior of bio-oil-modified bitumen within the framework of a hybrid experimental and machine-learning approach. Bio-binder was waste cooking oil used to partially replace conventional bitumen (ranging from 0 to 20 percent). The experimental analysis aimed to assess the binder's high-temperature deformation resistance using the rutting parameter (G*/sinδ). Three machine learning models were created with the use of key input variables such as bio-oil content, mixing temperature, aging condition, viscosity, and density to improve the predictive power and decrease the amount of effort expended on the experiment, including Artificial Neural Network (ANN), Random Forest (RF), and Extreme Gradient Boosting (XGBoost). Statistical measures such as R^2, RMSE, MAE, and MAPE were used to assess model performance. XGBoost proved to be the most accurate, with higher accuracy than ANN and Random Forest. The environmental benefits have been estimated by quantifying the binder-related CO_2 reduction when using bio-oil instead of oil. A Pareto optimization model was used to determine the best trade-off between environmental sustainability and mechanical performance. The findings suggest that a replacement level of bio-oil of 10–15% offers the most promising compromise, as it retains acceptable rutting resistance whilst achieving a considerable carbon reduction. In general, the suggested hybrid experimental-machine learning method can be considered an effective instrument for optimizing sustainable asphalt binder formulations and offers the possibility of the continuous incorporation of environmentally friendly materials into pavement engineering. The novelty of this study lies in integrating experimental characterization, machine learning prediction, and Pareto-based environmental optimization within a single framework for bio-oil-modified bitumen. Unlike previous studies that focused separately on mechanical performance and predictive modeling, the present work simultaneously evaluates rutting resistance, CO_2 reduction potential, and predictive accuracy using ANN, Random Forest, and XGBoost models to develop a sustainable and optimized asphalt binder formulation.

Keywords: bitumen, XGBoost, machine learning, CO_2 emission



Graphical Abstract

1. Introduction

1.1. Background

Bitumen is a petroleum product commonly used as a binder in asphalt roads. Although conventional bitumen can be a very good adhesive and viscoelastic material, it is associated with serious ecological issues due to its fossil origin, high energy use in its manufacturing, and greenhouse gas emissions (Belyaev et al., 2021; Olalekan et al., 2024; Penki & Rout, 2023). The extraction, refining, and processing of petroleum-based binders also significantly increase carbon dioxide (CO_2) emissions and contribute to environmental degradation worldwide.

The asphalt industry is energy-intensive, especially in the production of binder and the manufacturing of hot-mix asphalt, where temperatures are high (150–180°C). These processes negatively affect fuel consumption and produce significant CO_2 emissions and volatile organic compounds (VOCs). As transportation infrastructure demand is increasing globally, asphalt pavements have been identified as a major environmental footprint issue, posing a critical challenge to sustainability (Penki & Rout, 2023; Yaro et al., 2023; Zhu et al., 2023).

In response, bio-binder and bio-oil partial replacement technologies have been developed as promising alternatives (Ullah et al., 2024; Yaro et al., 2023; Zhu et al., 2023). Renewable or waste-based bio-oils, waste cooking oil, pyrolysis oil, and agricultural by-products have potential environmental benefits, such as reduced reliance on fossil fuels and lower carbon emissions throughout the life-cycle (El-Sherbeni et al., 2025; Sarker et al., 2024; Tran et al., 2024). Partial replacement of traditional bitumen with bio-oil can reduce embodied carbon without compromising mechanical and rheological performance. However, bio-modification significantly alters binder properties, reducing rutting resistance, fatigue performance, and aging behavior. Hence, there is a need to conduct a systematic exploration to strike a balance between environmental merits and engineering performance (Bhambhani et al., 2025; El-Sherbeni et al., 2025; Ullah et al., 2024).

1.2. Research Gap

A lot of research has been conducted on experiments to assess the performance of bio-modified bitumen in terms of its physical, rheological, and chemical characteristics. Such experiments are usually concerned with laboratory characterization, including penetration, softening point, dynamic shear rheometer (DSR), bending beam rheometer (BBR), and aging indices. Although useful, most of these studies rely on trial-and-error methods, require large experimental matrices, and consume large amounts of material, without offering smart optimization techniques (Abdulrahman et al., 2025; Hosseini et al., 2021; Miani et al., 2021).

Alternatively, machine learning (ML) algorithms have been used in pavement engineering to predict binder and mixture characteristics (Hosseini et al., 2021; Rahman et al., 2021; Upadhya et al., 2023). Artificial neural networks (ANNs), random forests, and gradient boosting models have been shown to achieve high predictive performance. Nevertheless, current ML research is usually focused solely on mechanical performance prediction, and environmental indicators such as carbon footprint or potential for reducing energy consumption are rarely introduced.

Moreover, there are only a limited number of studies that combine experimental characterization, machine learning modeling, and environmental optimization within a single framework. The lack of such a hybrid solution limits the possibility of systematically developing bio-modified bitumen that meets performance needs and, at the same time, achieves the objective of environmental sustainability.

Thus, an integrated experimental-machine learning framework that can predict mechanical behavior, quantify the impact of the environment, and optimize bio-oil replacement levels to develop biomass bitumen sustainably is needed.

This research aims to develop a hybrid experimental–machine learning framework to evaluate and optimize the environmental and mechanical performance of bio-oil-modified bitumen. The study focuses on investigating the influence of waste cooking oil (WCO) as a partial replacement for conventional bitumen on rutting resistance and sustainability performance. In addition, machine learning models, including Artificial Neural Networks (ANNs), Random Forests (RFs), and XGBoost, were employed to predict the rutting parameter ($G^*/\sin\delta$) based on key material and processing variables. Furthermore, Pareto optimization was conducted to identify the optimal bio-oil replacement level that provides a balanced trade-off between mechanical performance and CO_2 emission reduction.

1.3. Objectives

This research paper aims to implement a hybrid experimental-machine learning model in the optimization of bio-oil modified bitumen for the environment. Particularly, the objective of the study is to:

- Experimentally test the performance of bio-oil partially replaced bitumen in terms of physical and rheological behavior at various levels of replacement.
- Train and test machine learning models to forecast the machine's performance under experimental conditions.
- Increase and maximize bio-oil content by balancing mechanical performance requirements with environmental sustainability goals.

2. Materials and Methods

2.1. Materials

2.1.1. Base Bitumen

The binder applied in this experiment was a standard penetration grade 60/70 bitumen, which is usually used in pavement work in moderate-temperature weather. The bitumen was acquired in a local petroleum refinery and is unmodified.

The basic physical characteristics of the base bitumen were identified in accordance with ASTM standards. The findings are tabulated in Table 1. The values obtained are within the specification limits for 60/70 penetration grade bitumen, confirming that it can be used as a control binder for comparison with modified bio-oil samples. Table 1 presents the physical properties of base bitumen

Table 1. Physical properties of base bitumen

Property	Test Standard	Measured Value	Specification Limit
Penetration (25°C, 0.1 mm)	ASTM D5	65	60–70
Softening Point (°C)	ASTM D36	49.2	≥ 46
Ductility (25°C, cm)	ASTM D113	112	≥ 100
Rotational Viscosity (135°C, Pa·s)	ASTM D4402	0.42	≤ 3.0

2.1.2. Bio-Oil

The bio-oil used in the research was waste cooking oil (WCO) collected from local food outlets. The collected oil was filtered to remove food remains and impurities. No chemical treatment or any other modification was done before blending with the base bitumen. The waste cooking oil was chosen for its renewable nature, availability, low cost, and the possibility of reducing fossil fuel dependency and carbon emissions through partial replacement of the usual bitumen. The fundamental physical characteristics of the waste cooking oil were identified before mixing. Table 2 displays the measured values.

Table 2. Physical properties of waste cooking oil

Property	Test Method	Measured Value
Density at 25°C (g/cm³)	ASTM D4052	0.910
Dynamic Viscosity at 60°C (Pa·s)	ASTM D445	0.045
Kinematic Viscosity at 40°C (mm²/s)	ASTM D445	38

The fact that waste cooking oil has a relatively low viscosity compared to conventional bitumen indicates that it can serve as a softening agent, thereby affecting the rheological and workability characteristics of the modified binder.

2.2. Preparation of Bio-Modified Bitumen

Replacement amounts of 0, 5, 10, 15, and 20 percent of base bitumen were mixed with bio-oil on a weight basis.

Blending conditions:

- Temperature: 160°C,
- Mixing speed: 3000 rpm,
- Mixing time: 30 minutes.

Testing was done on the prepared samples stored in sealed containers.

2.3. Rheological Testing

Rutting resistance of the binders was measured using a Dynamic Shear Rheometer (DSR) in accordance with ASTM D7175. Unaged samples at 64°C were rutting tested and measured as G/sin delta.

Where:

- G* = Complex shear modulus,
- δ = Phase angle.

An increase in the values of Higher G*/Sinδ indicates enhanced rutting resistance at high service temperatures.

The rutting parameter (G*/ Sinδ) was chosen as the final desired machine learning modeling and optimization variable. Table 3 shows the Experimental design and factor levels used for bio-oil modified bitumen testing.

Table 3 illustrates the experimental design for this study to test bio-oil-modified bitumen. The experimental matrix comprises five bio-oil replacement levels (0, 5, 10, 15, and 20%), seven mixing temperatures (145–175°C), and two aging conditions (unaged and RTFO-aged). Three replicates were performed for each tested combination to ensure reliability and minimize experimental uncertainty.

Table 3. Experimental design and factor levels used for bio-oil modified bitumen testing

Factor	Levels
Bio-oil content (%)	0, 5, 10, 15, 20
Mixing temperature (°C)	145, 150, 155, 160, 165, 170, 175
Aging condition	Unaged (0), RTFO (1)
Replicates	3
Total samples	210

A balanced, all-inclusive dataset on the impact of bio-oil content, temperature, and aging on the properties of the binder was obtained through this factorial design, which yielded 210 samples. The experimental framework design ensures sufficient data variation for future machine learning modeling and performance prediction.

2.4. Environmental Assessment

The environmental advantage of bio-oil modification was calculated in terms of the potential reduction in CO_2 emissions through partial substitution of petroleum-based bitumen. The decrease in CO_2 in this research was supposed to be proportional to the percentage of bio-oil replacement by conventional bitumen; that is, a percentage decrease in the number of fossil-based binders will decrease demand by the same percentage.

2.4.1. CO_2 Reduction Estimation Method

For a given replacement level R (% by weight of binder), the CO_2 reduction per unit mass of binder was estimated using Eq. 1.

$$CO_{2,\text{saved}} = \left(\frac{R}{100}\right) \times EF_b \tag{1}$$

where:

- R = bio-oil replacement percentage (%),
- EF_b = emission factor of conventional bitumen production (kg CO_2 per kg bitumen).

The percentage reduction in binder-related CO_2 emissions was calculated as (Eq. 2):

$$\%CO_2 \text{ reduction} = \left(\frac{CO_{2,\text{saved}}}{EF_b}\right) \times 100 \tag{2}$$

Therefore, a 5, 10, 15, and 20 percent replacement of bio-oil would produce an estimated 5, 10, 15, and 20 percent further reduction in the binder-related CO_2 emission, respectively, under the condition of the same binder mass.

2.4.2. Reporting Environmental Indicator

The environmental index presented in this paper is the percentage of CO_2 calculated indoors for each level of bio-oil replacement (Table 4). The indicator was used to sustain the discussion on sustainability and rutting performance results (G*/sin δ).

Table 4. Estimated Reduction in Binder-Related CO_2 Emissions Due to Bio-Oil Replacement

Bio-Oil Replacement (%)	Conventional Bitumen Used (%)	Estimated CO_2 Reduction (%)
0	100	0
5	95	5
10	90	10
15	85	15
20	80	20

These findings suggest that the percentage of bio-oil replacement is directly proportional to the reduction in CO_2 emissions. The higher the bio-oil content, the lower the binder's CO_2 emissions are estimated to be, between 5–20 percent with the same mass of binder. This shows the environmental benefits of partially replacing petroleum-based bitumen with bio-based materials.

3. Machine Learning Framework

3.1. Dataset Preparation

The machine learning modeling data were generated from experimental data obtained with bio-oil-modified bitumen samples. Each data point corresponds to a specific blend of blending conditions and the associated rutting parameter (G*/sin δ).

The dataset was analyzed for consistency and completeness before being developed into a model. Min-Max scaling was used to normalize numerical variables to ensure stable model training and prevent bias due to variable magnitudes.

3.1.1. Input Variables (Features)

The following parameters were selected as input features based on their physical and rheological influence on binder performance:

Bio-oil content (%)

The percentage of bio-oil directly influences the binder's stiffness and viscoelastic behavior. The general increase in bio-oil tends to soften the binder, which may have a substantial effect on rutting resistance.

Mixing temperature (°C)

The blending efficiency and dispersion of bio-oil in the base bitumen are controlled by mixing temperature. It is capable of influencing binder homogeneity and therefore, modify rheological characteristics.

Aging condition

The aging variable was included as a categorical variable (e.g., Unaged = 0, RTFO = 1). The volatilization and oxidation during aging stiffen the binder, which has a strong effect on the rutting parameter.

Rotational viscosity (Pa·s)

Viscosity is an indicator of the flow properties of the modified binder at high temperatures. Rutting resistance is linked to high-temperature stiffness; hence, viscosity is a strong predictor.

Density (g/cm³)

Density shows changes in compositions with the addition of bio-oil. Density variations imply an interaction and blending efficiency of the material, which, in turn, can indirectly determine mechanical performance.

3.1.2. Output Variable (Target)

The target variable for prediction was:

Rutting parameter (G*/sinδ)

This is a parameter derived from Dynamic Shear Rheometer (DSR) testing, which measures the binder's resistance to permanent deformation at high service temperatures.

3.2. Model Development

Three machine learning algorithms were used to predict the rutting parameter (G*/sinδ) of bio-oil modified bitumen: Artificial Neural Network (ANN), Random Forest (RF), and Extreme Gradient Boosting (XGBoost). These models have been chosen because they are very effective at managing nonlinear relationships and complex interactions among the input variables.

An Artificial Neural Network (ANN) model was used to model nonlinear relationships between input features and the target variable using interrelated hidden layers (Adebiyi et al., 2024; Ren et al., 2025; YUSUF, 2021). Random Forest (RF), an ensemble learning technique that relies on several decision trees, has been used due to its resistance to overfitting and its ability to effectively manage multivariate information (Saleh & Gáspár, 2025; Sanij et al., 2025; Upadhya et al., 2023). The high predictive accuracy of the gradient boosting algorithm XGBoost was included because it can perform regularization and handle structured regression problems (Salehi et al., 2023; Sheng et al., 2025; Wang et al., 2025).

The statistical measures used to evaluate the performance of the developed models included the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). To determine the most appropriate algorithm for predicting binder performance, a comparative analysis was conducted.

3.3. Model Training and Evaluation Metrics

To ensure that the machine learning models were evaluated across a full range of statistical measures, performance was tested across multiple relationships. These measures assess the consistency between experimentally measured and model-predicted values of unconfined compressive strength (UCS) and are widely used in regression-based geotechnical model studies. Multiple error- and variance-based statistical measures were used to assess the predictive performance of the machine learning models, providing a complete picture of accuracy and strength. The coefficient of determination (R^2) can be used to measure the extent to which the model explains the variability in the experimental unconfined compressive strength (UCS) and to gauge the level of agreement between the model and observed results. Root Mean Square Error (RMSE) is used to estimate the overall scale of prediction error, and larger deviations carry greater weight, making it highly sensitive to major deviations. Mean Absolute Error (MAE) is simply the average of the absolute differences between predicted and experimental values. It provides a more robust estimate of the normal error by equalizing all differences. Mean Absolute Percentage Error (MAPE): A measure of prediction error expressed as a percentage, allowing relative comparison of model performance across datasets. Also, model generalization was assessed using k-fold cross-validation scores to compare predictive stability across a series of data partitions, thereby minimizing bias introduced by a single train-test split.

Coefficient of Determination (R^2)

The coefficient of determination (R^2) is used to assess the percentage of the variance in the measured rutting parameter values accounted for by the model predictions. The greater the R^2 value, the more effectively the prediction and correlation between predicted and experimental outcomes, Eq. 3.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3}$$

y_i : Measured rutting parameter, $\hat{y}_i$: Measured rutting parameter, n is the number of samples.

Root Mean Square Error (RMSE)

RMSE shows the square root of the mean squared error between estimated and observed values of rutting parameters. It does not favor large prediction errors and provides a more complete picture of the overall model performance (Eq. 4).

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{4}$$

Mean Absolute Error (MAE)

MAE determines the mean deviation in the foretold and experimental rutting factor. It also does not prioritize any errors, like RMSE, and is less sensitive to outliers (Eq. 5).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{5}$$

Mean Absolute Percentage Error (MAPE)

MAPE also describes the prediction error as a percentage of the measured rutting parameter value, making it easily comparable across models in terms of relative prediction accuracy (Eq. 6).

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (6)$$

3.4. k-Fold Cross-Validation Score

K-fold cross-validation was used to evaluate model generalization and to avoid bias from a single train-test split. The data were split into k similar sets, and the model was fitted on k-1 sets and tested on the remaining set (Nguyen & Tran, 2023; Rahman et al., 2021; Salehi et al., 2023). This was repeated k times, and the average performance measures were reported as the cross-validation score (Eq. 7).

$$CV_{score} = \frac{1}{k}\sum_{j=1}^{k} Score_j \quad (7)$$

The CV_{score} in the expression, k-fold cross-validation is the overall cross-validation of the machine learning model. The notation k represents the number of folds into which the dataset will be divided. The validation process follows a similar step: a dataset is further separated into k subsets of approximately equal size, and in each iteration, k-1 subsets are used for training and the remaining subset for testing.

$Score_j$ is the metric of performance (R^2, RMSE, MAE) achieved on the j th fold of validation. Every fold provides a distinct estimation of model performance using an alternate training-validation split. The total performance value of all folds is summed up as: $\sum_{j=1}^{k} Score_j$ and then it is divided by k to find the average value. This mean score is a good approximation of the model's overall generalization because it eliminates the effects of random data partitioning and reduces overfitting.

To improve the reproducibility and transparency of the developed machine learning framework, additional methodological details have been incorporated into the revised manuscript. The dataset was split into training and test sets using a predefined train/test split ratio to ensure unbiased model evaluation. Furthermore, k-fold cross-validation was implemented to assess model generalization and reduce the influence of random data partitioning. The revised manuscript also reports the major hyperparameters used for the Artificial Neural Network (ANN), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) models, including network architecture, number of estimators, tree depth, learning rate, and optimization settings. These additions provide sufficient information for replication and comparative assessment of the proposed predictive models.

4. Results and Discussions

4.1. Experimental Results

4.1.1. Variation of Rutting Parameter vs Bio Oil Content

Figure 1 shows how the rutting resistance of the modified binders is influenced by the bio-oil content. It is observed that an increasing ratio of bio-oil percent to the G*/sinδ also shows a decreasing pattern as the percentage of unaged and RTFO-aged samples increases. This action implies that adding waste cooking oil lowers the binder's stiffness, as it softens it. But RTFO-aged binders exhibit higher G*/sinδ values than unaged samples, indicating a stiffening effect from short-term aging. The findings show that although bio-oil enhances environmental sustainability by partially replacing petroleum-based binders, high content may compromise high-temperature rutting resistance.

4.1.2. Variation of Rutting Parameter vs. Temperature

Figure 2 demonstrates the comparison of the temperature sensitivity of unaged and RTFO-aged binders with various bio-oil contents. An increase in temperature results in a decrease in G*/sinδ for all binders, indicating that lower temperatures correspond to lower rutting resistance. RTFO-aged binders have a statistically significantly higher value of G*/sinδ with the passage of time because of oxidative stiffening with time. Nevertheless, the higher the bio-oil content, the lower the stiffness in both cases, demonstrating the softening effect of bio-oil. The findings reveal the interactive effect of temperature, aging, and bio-oil dosage on high-temperature performance.

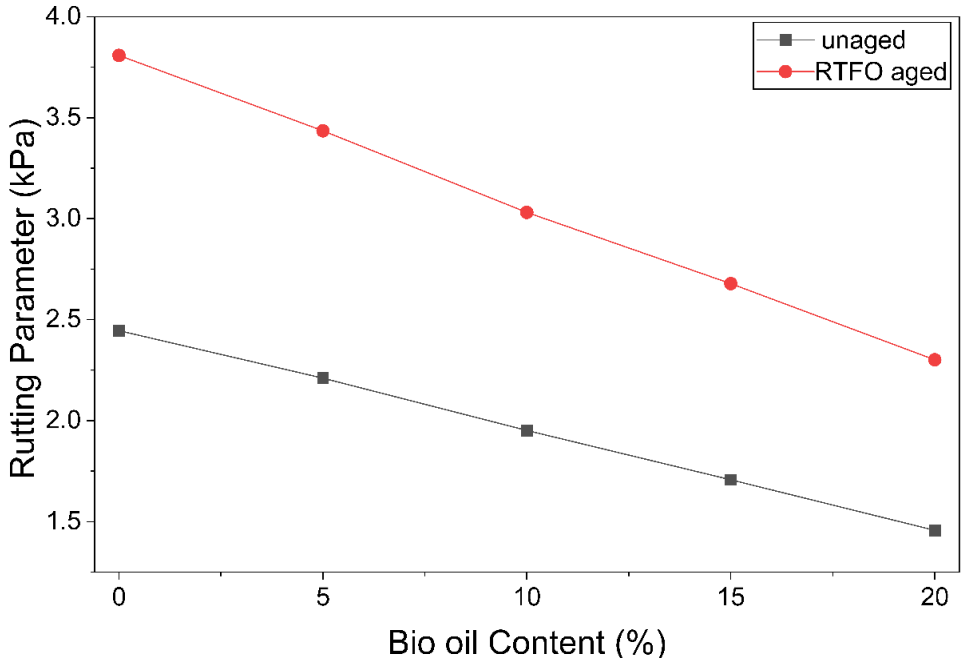


Fig. 1. Variation of rutting parameter (G/sinδ) with increasing bio-oil content for unaged and RTFO-aged binders at 64°C

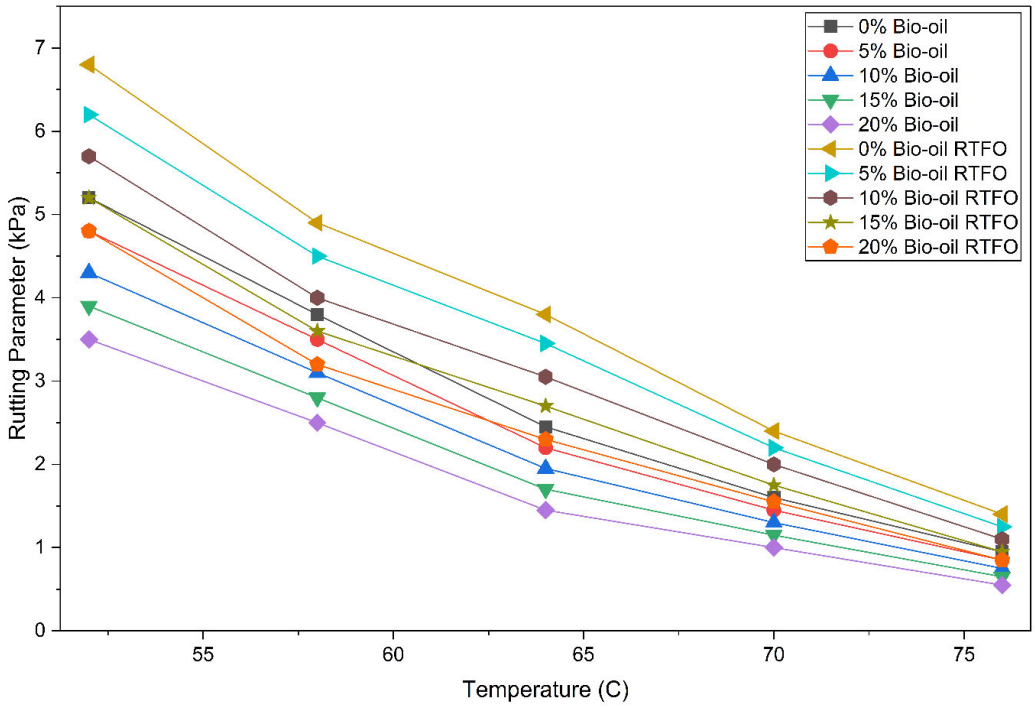


Fig. 2. Temperature-dependent variation of rutting parameter (G/sinδ) for unaged and RTFO-aged bio-modified binders. Solid lines represent unaged samples, while dashed lines indicate RTFO-aged binders

4.1.3. Contour Map of bio-oil content and temperature

The contour map of the relationship between bio-oil content and temperature on the rutting parameter (G*/sinδ) of unaged binders is shown in Figure 3. A color gradient indicates the magnitude of G*/sinδ, with warmer colors indicating higher stiffness and better rutting resistance, whereas cooler colors indicate lower stiffness.

The contour distribution shows clearly that G*/sinδ decreases steadily with increasing temperature. This phenomenon indicates the viscoelasticity of bitumen, with higher temperatures making the binder less solid and more vulnerable to permanent deformation.

Additionally, contour regions shift to lower G*/sinδ values as bio-oil content increases. This proves the softening property of waste cooking oil, which lowers the binder's rigidity due to its low viscosity and reduced molecular composition compared to conventional bitumen.

The pattern of interaction in the contour map indicates that temperature exerts a greater effect at higher temperatures, whereas bio-oil content is more influential on stiffness across all temperatures. The maximum G*/sinδ values are obtained at low temperatures and low bio-oil content, indicating the best rutting resistance at that time.

In general, the contour plot visually depicts the nonlinear relationship between bio-oil content and temperature, thereby enhancing the applicability of machine learning models for predicting the binder's behavior in a multidimensional input space.

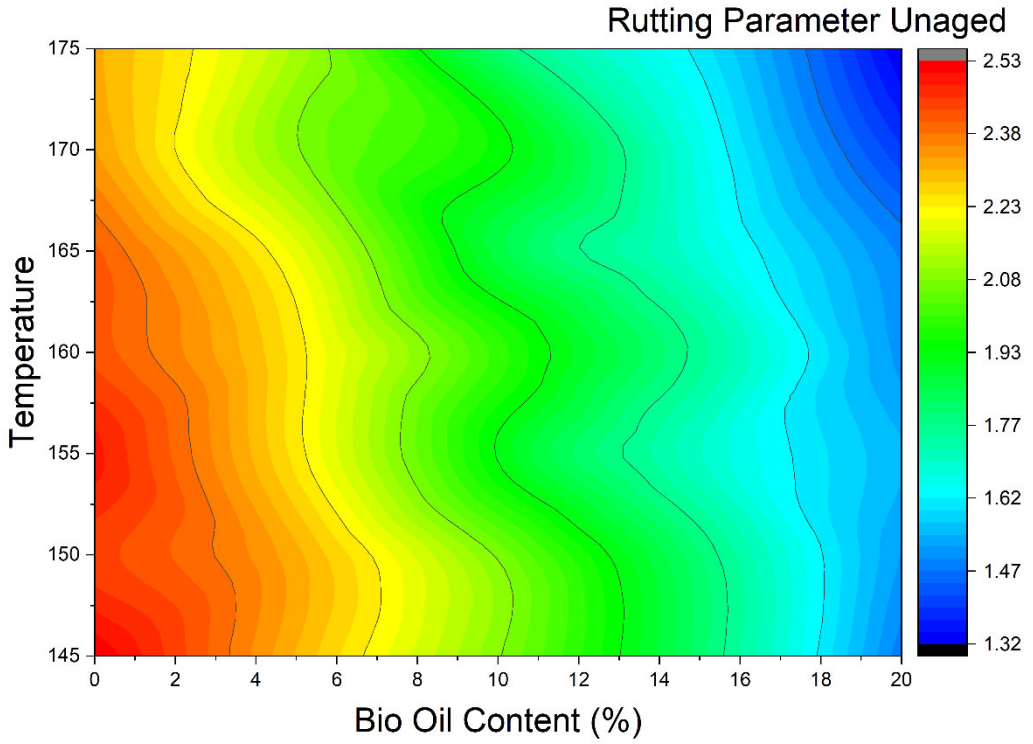


Fig. 3. Contour map showing the variation of G/sinδ with bio-oil content and temperature for unaged binders

4.2. Machine Learning Result

4.2.1. Model Interpretation

To improve the transitivity and credibility of the formulated machine learning models, it was necessary to interpret the models to gain clearer insight into the impact of the input variables on the predicted rutting parameter (G*/sinδ). SHAP (SHapley Additive exPlanations) and sensitivity analyses were used in this study to measure feature importance and estimate the contribution of each input parameter (Salehi et al., 2023; Sheng et al., 2025; Zhang et al., 2025).

The relative significance of the input features in the prediction process was determined using SHAP analysis, which relies on cooperative game theory. This algorithm estimates the contribution of each feature to a model shift using SHAP values assigned to each feature. The positive values of SHAP reveal a positive effect of a feature on the predicted G*/sinδ, whereas the negative values demonstrate a negative effect. In this way, the effects of bio-oil content, mixing temperature, aging condition, viscosity, and density were measured, and the factors with the greatest effect on the binder's performance were identified. This analysis also showed that the model's behavior was physically consistent with experimental results. A random forest feature importance ranking was employed for model interpretation.

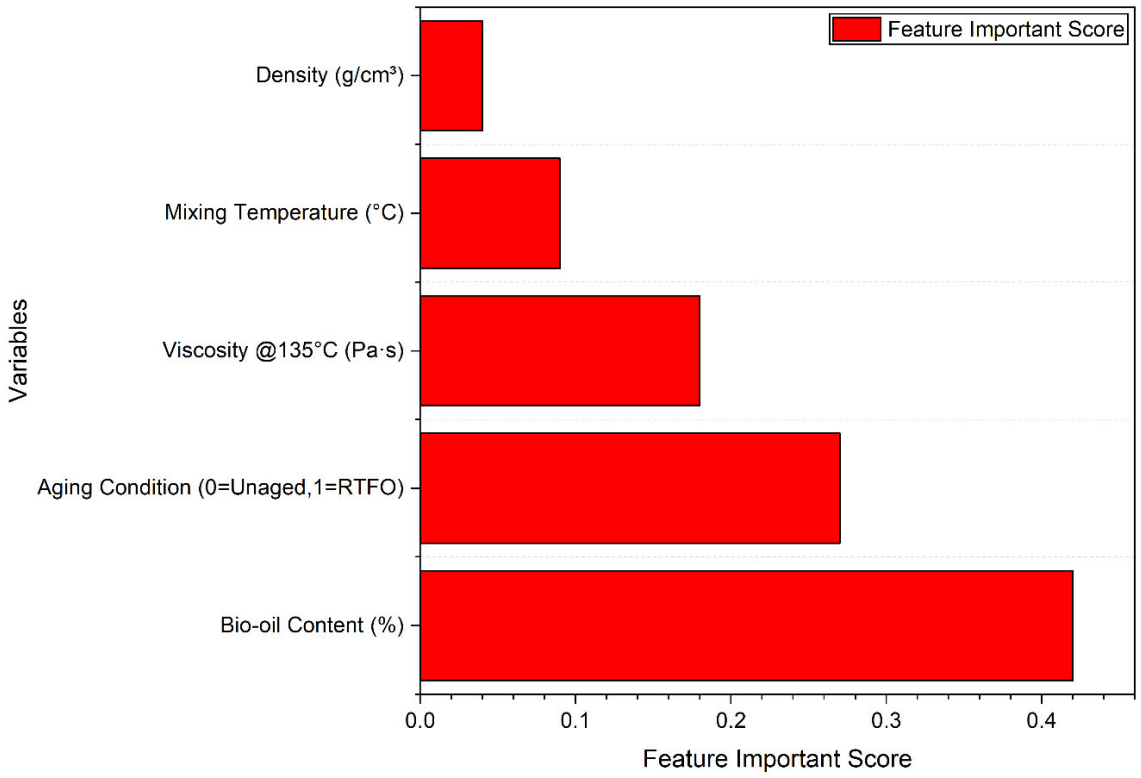


Fig. 4. Feature importance ranking for predicting G*/sinδ using the Random Forest model

Figure 4 indicates the relative significance of the input variables in predicting the rutting parameter (G*/sinδ) by the model of the random forest. The feature importance values indicate the contribution of each parameter to the overall prediction accuracy. The bio-oil content is the most important input variable, indicating that it is the most significant factor influencing binder behavior. This is in line with experimental findings showing that the percentage of bio-oil used significantly affects the binder's stiffness. The aging condition is also of significant importance as it is a good indicator of the tremendous effect of the oxidative hardening on the rheological behavior. Another important parameter is viscosity, which means that it is closely related to high-temperature binder stiffness. Conversely, mixing temperature and density exhibits a relatively lower significance, implying a relatively lower impact on the predicted value in the tested range. Altogether, the feature importance analysis reveals that the model's predictions are physically self-contained and consistent with experimental trends, thereby enhancing trust in the machine learning framework.

4.2.2. Machine Learning Model Performance

4.2.2.1. Observed vs Predicted Analysis

To assess the predictive power of the machine learning model, comparisons between observed (experimental) and predicted values of the rutting parameter (G*/sinδ) were made. A visual comparison of observed and predicted analyses will provide a direct evaluation of model accuracy, consistency, and generalization performance. Such a high correlation between the actual and predicted values shows that the model is able to capture the nonlinear associations between the input parameters and the binder's performance.

XGBoost (Extreme Gradient Boosting) is a highly developed ensemble learning model, which is built upon gradient boosting decision trees. It sequentially constructs several weak learners (decision trees) with each succeeding tree rectifying the remaining errors in the earlier model. Unlike conventional boosting methods, XGBoost uses regularization, shrinkage, and effective tree pruning, thereby increasing prediction accuracy while reducing overfitting. XGBoost was employed in this study to develop a model for nonlinear interactions among bio-oil content, mixing temperature, aging condition, viscosity, density, and the target variable G*/sin δ. Its suitability, especially for material performance prediction, stems from its ability to handle structured tabular data and its flexibility in representing complex feature interactions (Fig. 5).

The plots of observed and predicted values for both the training and test data sets indicate good agreement between the experimental and predicted values. The values of the predictive variables fit well along the 45° reference line, with moderate dispersion in the training dataset. The scatter in the test data set is slightly wider than in the training data, indicating a slight decrease in predictive accuracy on unseen cases. Most of the

points, however, lie near the ideal line, confirming acceptable generalization capability. The coefficient of determination ($R^2 \approx 0.932$) indicates that the Random Forest model explains a large percentage of the variance in the rutting parameter (Figure 6). Random Forest is a powerful model with good predictive accuracy, though it is slightly lower than XGBoost's. In general, the Random Forest model is highly accurate in predicting G*/sin δ, but its prediction error is slightly higher than that of the XGBoost algorithm.

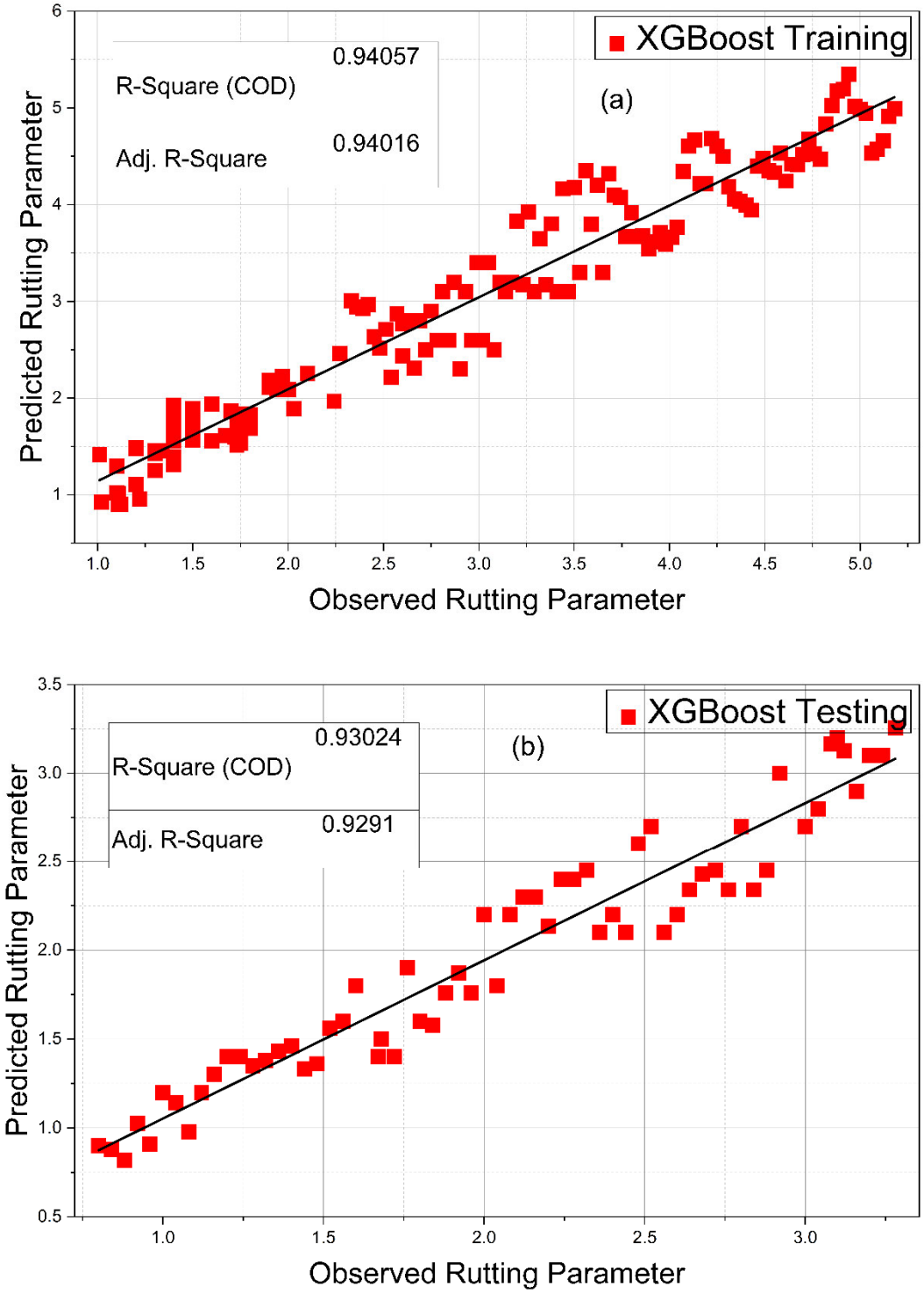


Fig. 5. Comparison of observed and XGBoost-predicted G*/sin δ values for (a) the training dataset and (b) the testing dataset

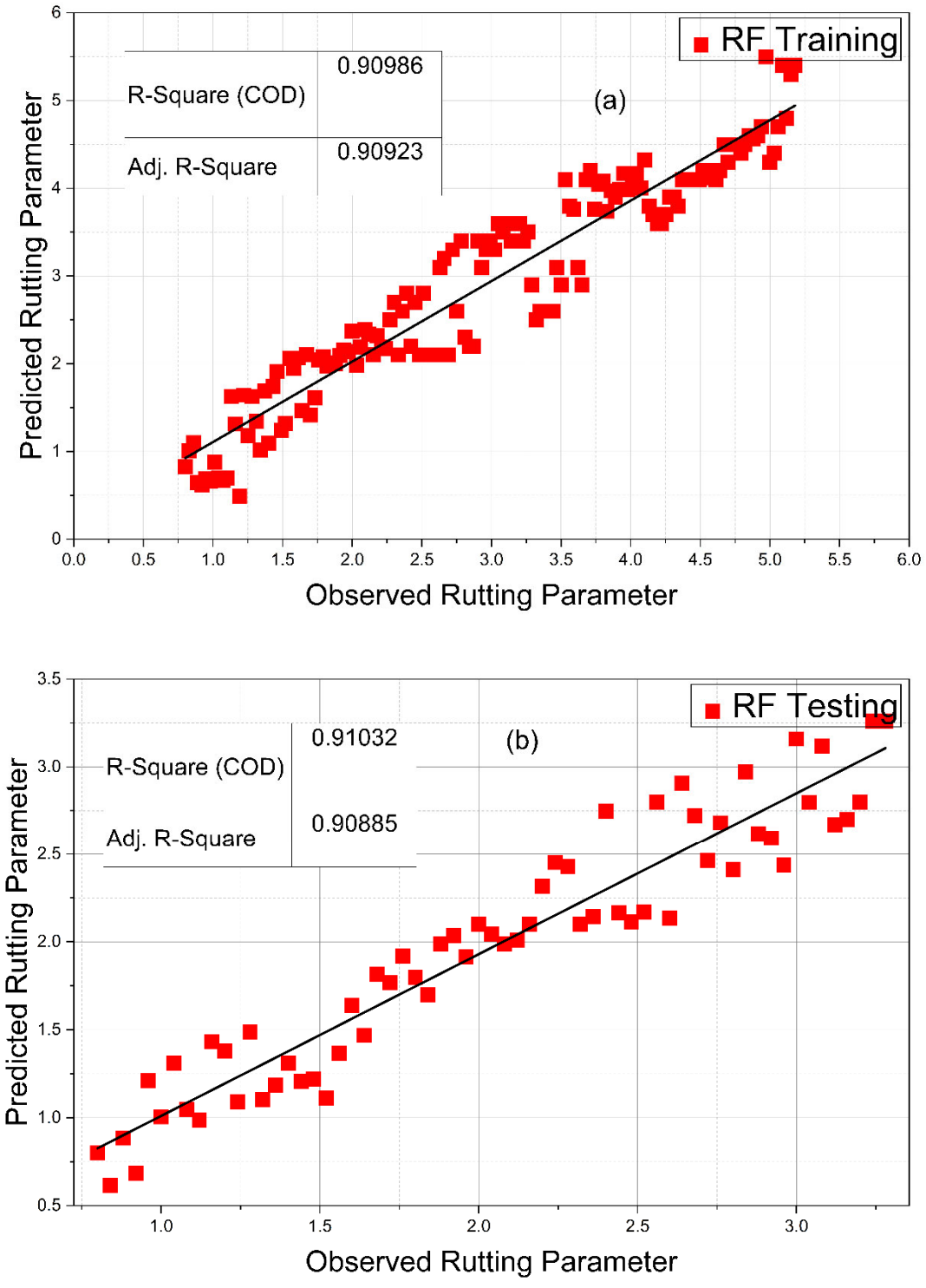


Fig. 6. Comparison of observed and Random forest model-predicted G*/sin δ values for (a) the training dataset and (b) the testing dataset

The plots of observed and predicted values are both reasonable when comparing experimental and predicted values across the training and testing datasets. Nevertheless, a somewhat greater scatter of points along the 45° reference line is observed compared to ensemble-based models. The ANN model achieved an R^2 of about 0.89, which is good for prediction but less accurate than Random Forest and XGBoost (Figure 7). Nonetheless, the higher prediction variability suggests that the ANN model may require larger data sets or further hyperparameter optimization to achieve optimal performance. Nevertheless, the ANN model has been able to capture the overall trend in rutting behavior and demonstrates that neural networks can be used to predict binder performance.

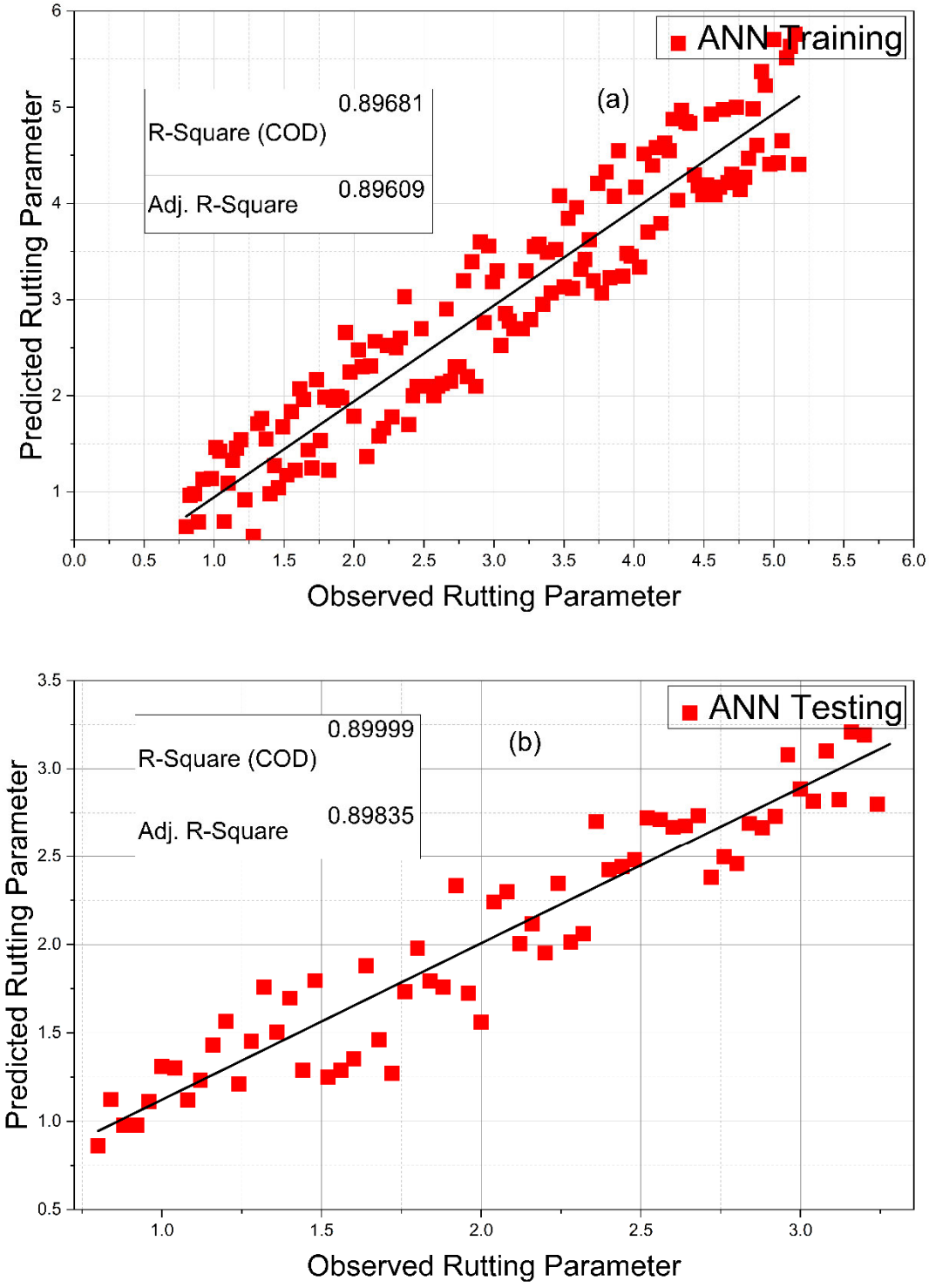


Fig. 7. Comparison of observed and ANN model-predicted G*/sin δ values for (a) the training dataset and (b) the testing dataset

4.2.2.2. Comparative Discussion of ANN, RF, and XGBoost

A comparative analysis of the three machine learning models shows that they differ in predictive performance. The ANN model had an R^2 of 0.89, indicating reasonable accuracy but increased forecast dispersion. The prediction results were much better with Random Forest, with an R^2 of 0.932, which is much higher and indicates that the nonlinear interactions are better managed and that the variance is lower.

XGBoost performed best compared to both models, with the highest R^2 value (0.98). The observed and predicted plots of XGBoost are closest to a 45-degree reference line in both the training and testing cases, indicating the best predictive accuracy and generalization.

The performance hierarchy can be summarized as:

XGBoost>Random Forest>ANN

XGBoost outperforms other machine learning models due to its gradient-boosting structure, regularization, and effective optimization, which enable it to capture small nonlinear relationships in the data. In general, all three models have predictive power, but XGBoost offers the best and most robust estimation of the rutting parameter for bio-modified bitumen.

4.2.2.3. Residual Plot Analysis

Residual analysis is an important process to assess the robustness and reliability of machine learning models (Hosseini et al., 2021; Mansoori et al., 2025; Nguyen & Tran, 2023). The difference between the observed (experimental) and predicted value is referred to as a residual:

$$\text{Residual}=y_{\text{observed}}-y_{\text{predicted}}$$

Residual plots, which show residuals as a function of the predicted value, are useful for evaluating model bias, variance behavior, and homoscedasticity. A perfect model will generate residuals that are randomly distributed, i.e., the residuals will not follow a pattern. This kind of behaviour shows that the model well represents the underlying data structure and the model is not afflicted by systematic prediction errors.

The training and test residual plots of the XGBoost model depict a random distribution of residuals around the zero-reference line (Figure 8). There is no apparent trend, curvature, or funnel-like structure, indicating the absence of systematic bias and heteroscedasticity. The trend of the training data shows that the residuals are tightly clustered around zero, indicating that the model fits the data well and results in little prediction error. The small range shows that XGBoost is effective at describing the nonlinear dependencies among bio-oil content, temperature, aging condition, and the rutting parameter ($G^*/\sin\delta$).

There is a little more dispersion in the residuals in the test dataset than in the training set, as is anticipated when testing unseen data. Nevertheless, the residues are symmetrically distributed around zero, indicating strong generalization and low overfitting. This is because the lack of formal patterns in the residual plots confirms that the XGBoost model produces free and consistent estimates. This again proves the better work of XGBoost over ANN and random Forest models.

The random forest model's residual plots indicate that, in both the training and test data, the residuals are generally randomly distributed around the zero-reference line (Figure 9). This implies that there is no strong systematic bias in the prediction of the rutting parameter ($G^*/\sin\delta$) in the training dataset, but the residuals are concentrated around zero, albeit with a somewhat broader distribution than that of the XGBoost model. This indicates that although the Random Forest well represents the nonlinear relationships, its fitting power slightly falls short of that of XGBoost, as would be the case with unseen data. The testing dataset also showed slightly greater residual spread, as would be expected with unseen data. The lack of visible curvature, trend, or funnel-shaped distribution, however, indicates that the model has acceptable homoscedasticity and generalization ability, though with a higher degree of residual variability than in XGBoost. All in all, the Random Forest model exhibits consistent, unbiased predictions, albeit with higher residual variability. This finding aligns with its lower R^2 performance and the fact that Random Forest makes reliable, though relatively inaccurate, predictions.

The residual plots of the Artificial Neural Network (ANN) model show how the prediction errors are distributed across the training and test datasets (Figure 10). The residuals in the training residual plot are distributed along the zero-reference line, indicating that the ANN model does not exhibit a significant systematic bias. Nevertheless, a comparatively broader spread of residuals is seen as in contrast to ensemble-based models. This implies that although the ANN model represents the overall trend of the data, its fit accuracy is relatively low. As the testing residual plot shows, the distribution of the residuals slightly increases, which is normal when testing unknown data. The lack of an evident pattern (curvature or a funnel-shaped distribution) indicates that strong heteroscedasticity does not mar the model. But the more variance in the residuals than in

the XGBoost and Random Forest models, the lower the predictive accuracy. In general, the ANN residual analysis indicates reasonable generalization, though with increased prediction variation. This observation holds for its lower R^2 (0.89), which shows that the ANN is a good-performing model but less precise than ensemble-based techniques.

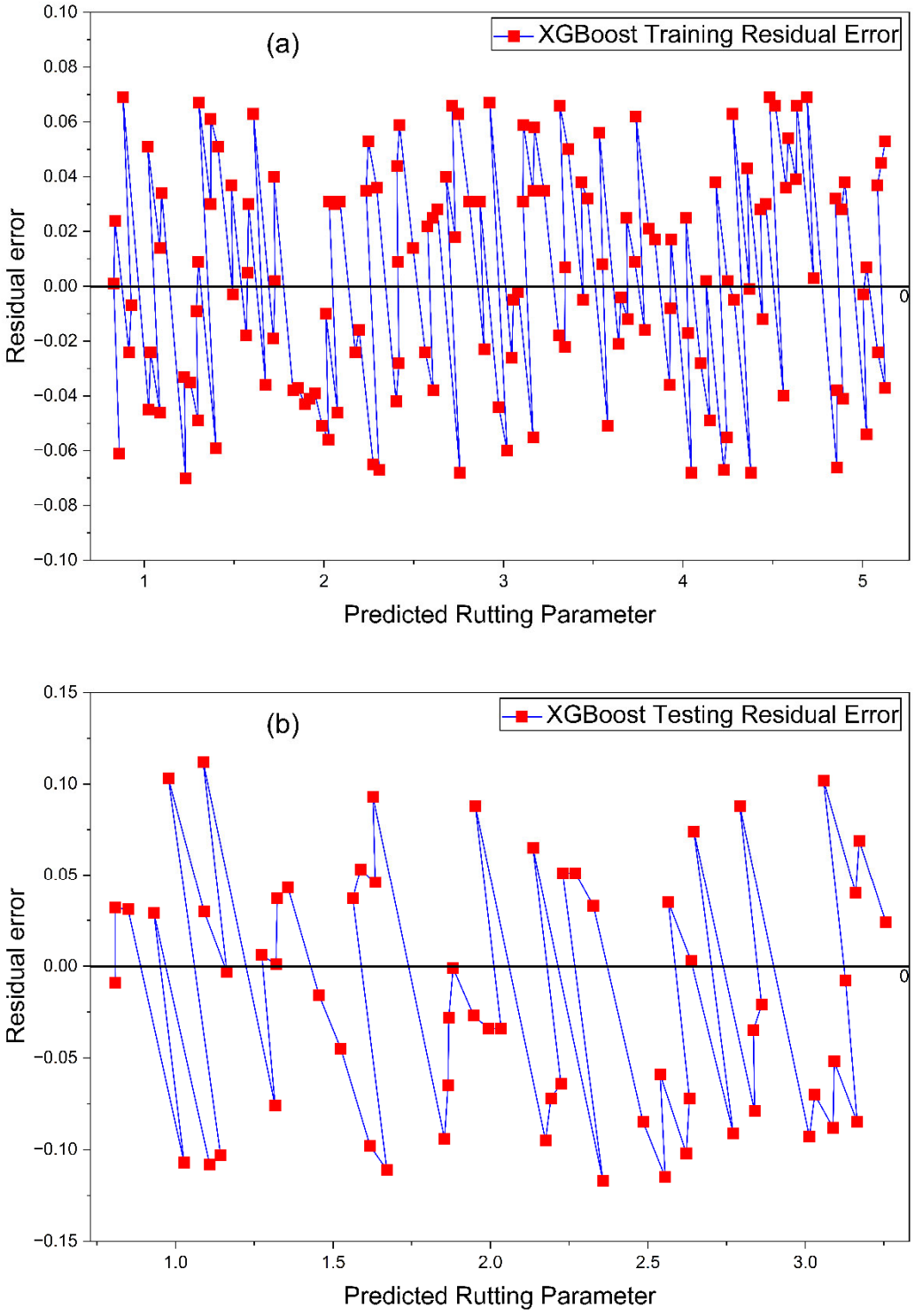


Fig. 8. Residual plot for the XGBoost model a (training dataset). b (Testing Dataset)

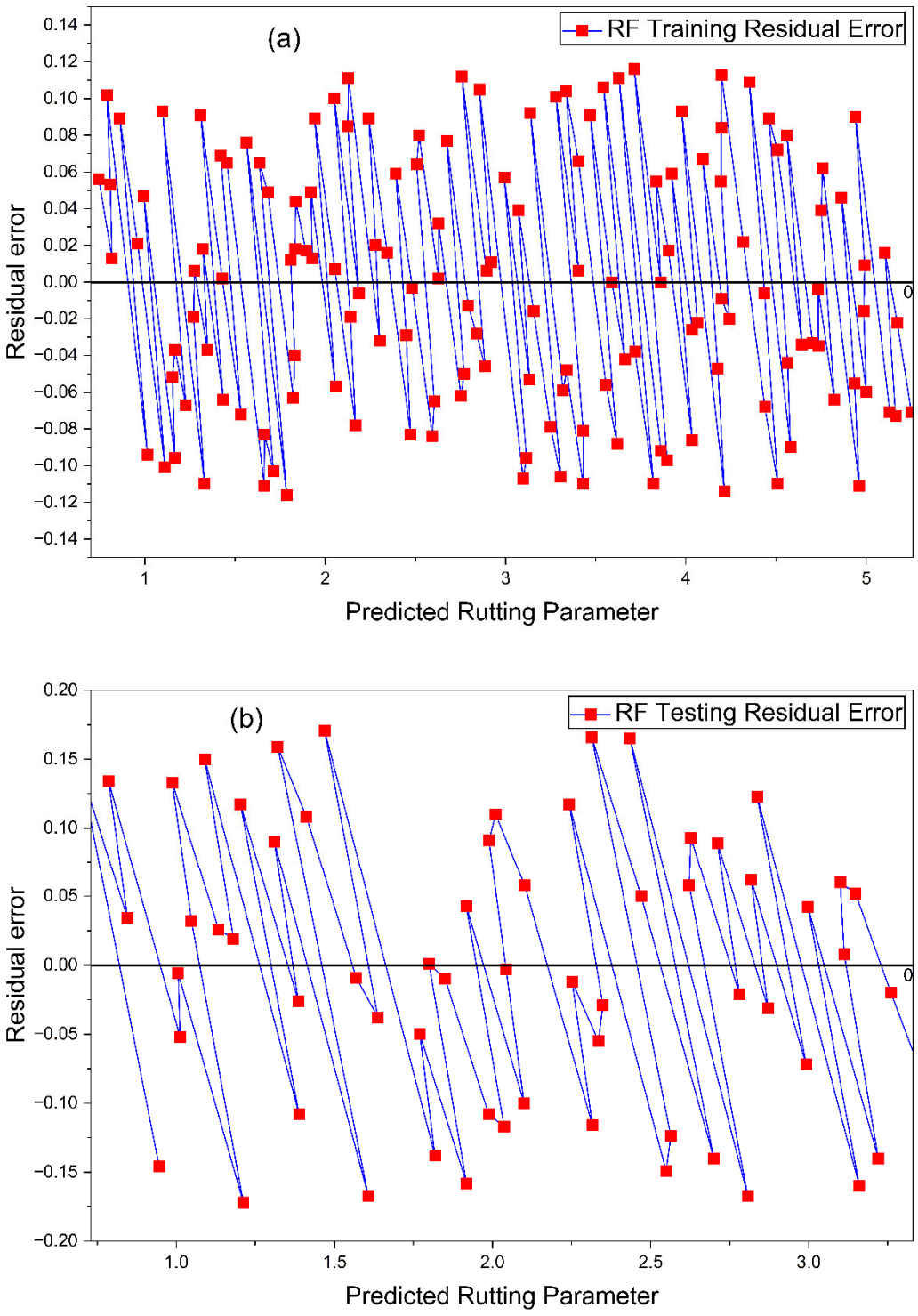


Fig. 9. Residual plot for the Random Forest model a (training dataset). b (Testing Dataset)

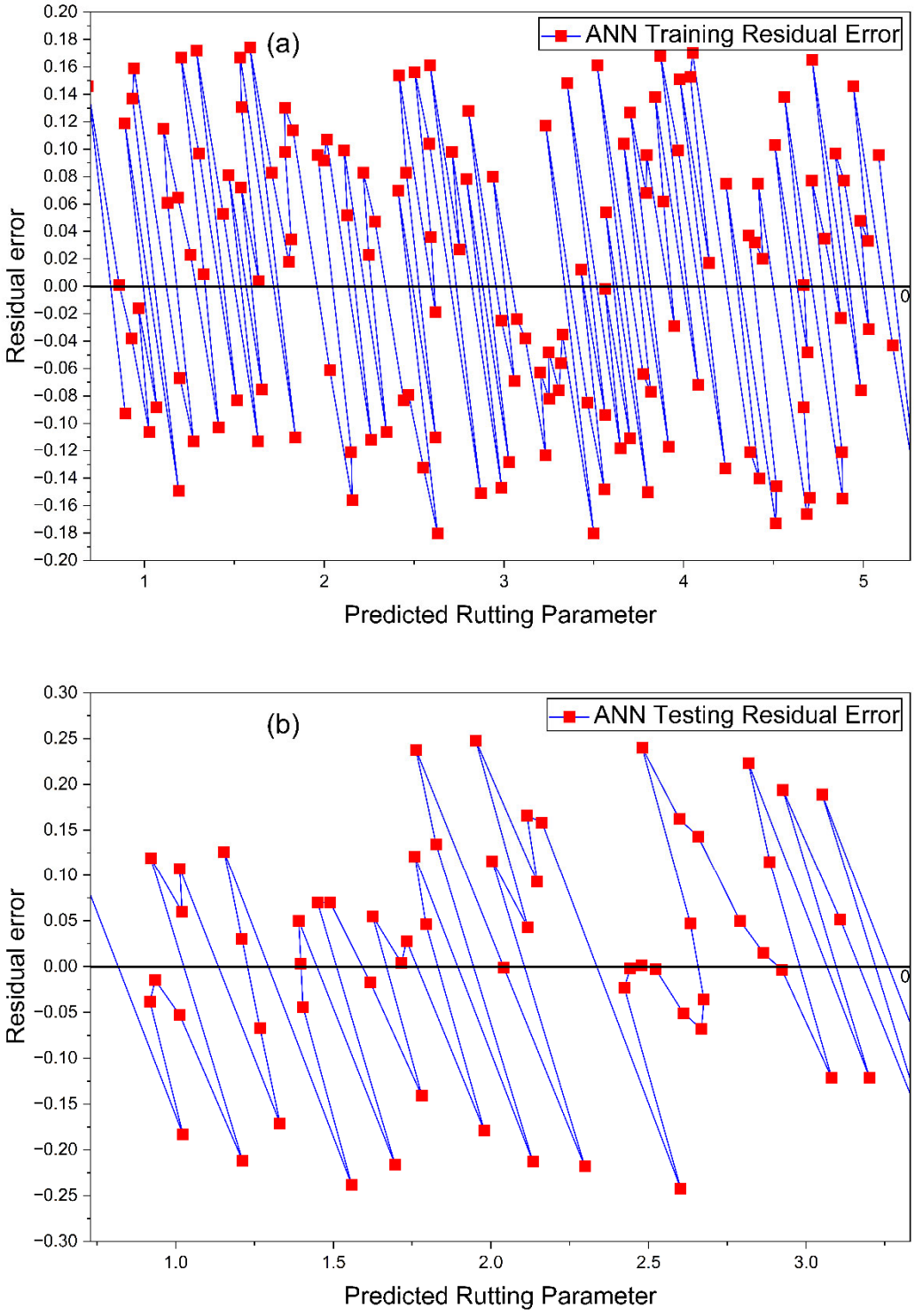


Fig. 10. Residual plot for the ANN model a (training dataset). b (Testing Dataset)

4.2.2.4. Pareto Optimization Analysis

Pareto optimization is a multi-objective decision-making model applied when there is a conflict between two or more objectives. It does not give a single optimal solution, but a collection of optimal trade-off solutions, called the Pareto front (Gilani et al., 2020; Kaveh et al., 2026; Varanda et al., 2017; Yaro et al., 2021).

A solution is said to be Pareto-optimal if there is no objective that improves without the other objectives deteriorating. This is particularly useful in the field of materials engineering, where there is a need to balance mechanical performance and environmental sustainability.

In this research, two opposing goals were taken into account:

- Maximize rutting resistance (G*/sin δ),
- Reduce the binder CO_2 emissions.

This is because a higher bio-oil content would reduce CO_2 emissions but may also lead to lower rutting resistance; hence, a trade-off analysis will be necessary. The Pareto front illustrates this trade-off.

Pareto front shows an inverse relationship between rutting resistance and CO_2 reduction, with an increase in the percentage bio-oil replacement. As the bio-oil content rises to 20 percent or 20 percent more than 0 percent, the CO_2 reduction is higher, and the mechanical performance declines steadily since the rutting resistance decreases. This shows that achieving the maximum level of environmental benefit and structural performance is difficult to do simultaneously, and a balance needs to be struck between the two goals. The asphalt binder is most resistant to rutting at low replacement rates (0–5 percent), indicating high mechanical stability and good resistance of the material to enduring resilience to permanent deformation under traffic loads. The environmental benefit at this level, however, is not significant when one considers that CO_2 reduction is insignificant. Even though this region guarantees high levels of structural performance, it does not make a significant contribution towards sustainability objectives. At the moderate replacement range (10–15%), performance is less unbalanced. The resistance to routing is within a reasonable range, as the CO_2 decrease is much higher, at about 10–15. This range will be an effective trade-off between long-term mechanical performance and an environmentally friendly situation and will be applicable to applications that attach importance to performance and emission reduction (Figure 11).

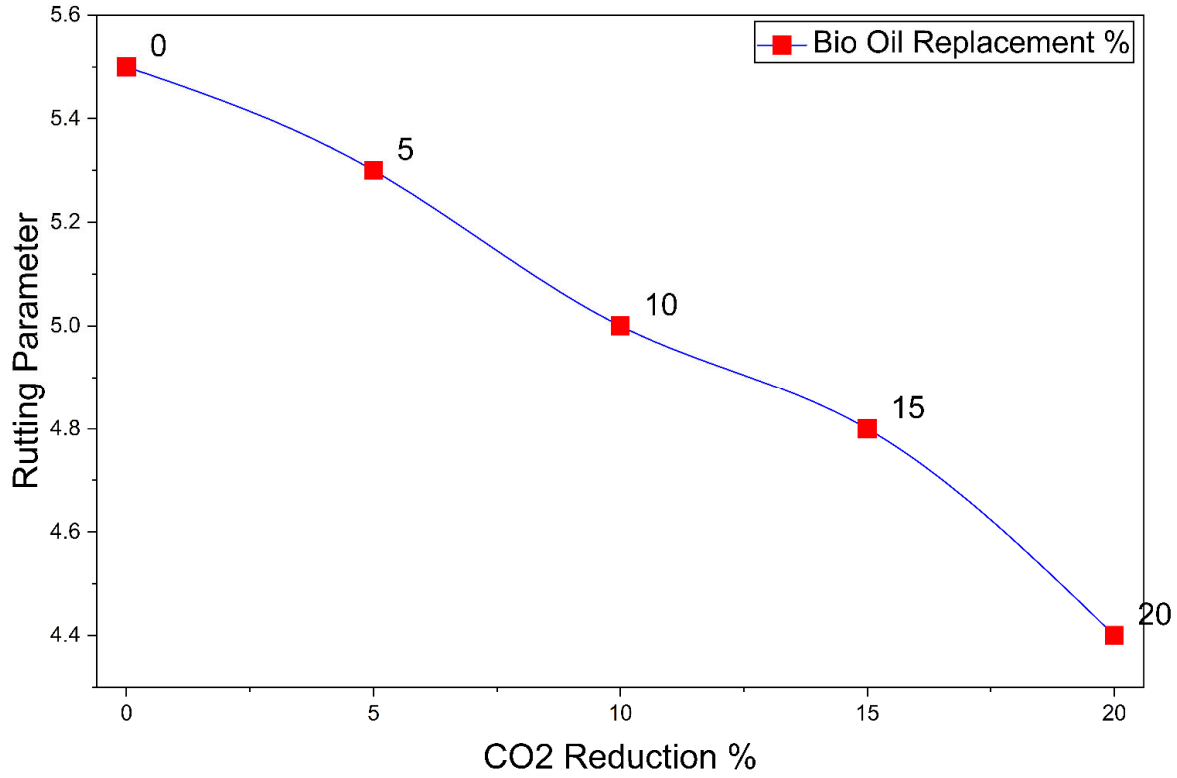


Fig. 11. Pareto front illustrating the trade-off between rutting resistance (G*/sinδ) and CO_2 reduction with increasing bio-oil replacement

The maximum environmental benefit occurs at a high level of replacement (20%) because it results in the greatest reduction in CO_2. Nonetheless, rutting resistance shows a significant decrease, which may be a concern for long-term pavement performance and its resistance to rutting under heavy traffic conditions. Given that this area offers the highest level of sustainability advantage, a potential decrease in structural performance value could limit its practical feasibility. On the whole, the Pareto optimization outcomes show that a moderate level of bio-oil inclusion can achieve effective carbon reduction while maintaining adequate rutting resistance. Of all the considered options, a bio-oil displacement ratio of 10–15 percent provides the best balance between environmental sustainability and structural performance. Thus, the suggested range will be the most effective for use as a sustainable asphalt binder.

5. Conclusions

Substitution of traditional bitumen with bio-oil significantly reduces the rutting performance of asphalt binders. It is stated in the study that the addition of moderate concentrations of bio-oil (up to approximately 15 percent) does not lead to a significant change in the rutting resistance, as the ratio of $G^*/\sin\delta$ indicates, but helps to enhance the environmental sustainability in terms of the proportional decrease in the quantity of CO_2 released. Three machine learning methods, including Artificial Neural Networks (ANN), Random Forest, and XGBoost, were tested to predict the performance of a binder. XGBoost was the most predictive, achieving the highest coefficient of determination (R^2) and the lowest prediction errors. This excellent performance is explained by its ensemble boosting framework, which effectively captures the intricate nonlinear relationships among the material composition and processing variables. In addition, the Pareto optimization analysis found an optimal balance between environmental and performance at approximately 10–15 percent bio-oil substitution, where sufficient rutting resistance can still be attained without a significant decrease in carbon emissions. To enhance future work, it is suggested that future research should broaden the scope of rutting performance to include long-term aging, fatigue behavior, life-cycle assessment (LCA), and larger experimental datasets to improve the generalizability of the models and assist in the design of more sustainable asphalt binders.

Limitation of present work

A limitation of the present environmental assessment is that the estimated CO_2 reduction was assumed to vary linearly with the percentage replacement of conventional bitumen by waste cooking oil (WCO). This simplified assumption considers only the reduction in petroleum-based binder use and does not account for additional emissions associated with WCO collection, transportation, processing, filtration, storage, and blending, as well as energy consumption for these activities. Previous studies have highlighted that a comprehensive environmental evaluation of bio-modified binders should include a full life-cycle assessment (LCA) to capture all upstream and downstream emission sources accurately. Therefore, future research should incorporate detailed LCA-based analyses to provide a more realistic estimate of the environmental benefits of bio-oil-modified asphalt binders.

Acknowledgement

The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through the Large Research Project under grant number RGP2/481/46.

References

Abdulrahman, S., Farouk, A., Wahhab, H.A.-A., Awoyera, P.O., & Abba, S.I. (2025). Experimental Performance and Machine Learning-Aided Design Tools for Modeling Unaged Warm Cup Lump Rubber-Modified Asphalt with Consideration of Global Warming Potential. *International Journal of Pavement Research and Technology*, 1–17.

Adebiyi, F.M., Ore, O.T., & Aderibigbe, O.E. (2024). Thermogravimetric characteristics and predictive modelling of oil sand bitumen pyrolysis using artificial neural network. *Discover Energy*, *4*(1), 17.

Belyaev, P., Frolov, V., Belyaev, V., Varepo, L., & Bezzateeva, E. (2021). Petroleum bitumen and polymer-bitumen binders: Current state and Russian specifics. Review. AIP Conference Proceedings,

Bhambhani, B., Chandra, S., & Bharath, G. (2025). Sustainability in road construction by using binders modified with biomass-derived bio-oil–a critical review. *Transportation Research Record*, *2679*(5), 939–967.

El-Sherbeni, A.A., Awed, A.M., Gabr, A.R., & El-Badawy, S.M. (2025). Biomass-Derived Bio-Oil for asphalt binder applications: production feasibility and performance enhancement. *Construction Materials*, *5*(1), 11.

Gilani, V.N.M., Hosseinian, S.M., Behbahani, H., & Hamedi, G.H. (2020). Prediction and pareto-based multi-objective optimization of moisture and fatigue damages of asphalt mixtures modified with nano hydrated lime. *Construction and Building Materials*, *261*, 120509.

Hosseini, A.S., Hajikarimi, P., Gandomi, M., Nejad, F.M., & Gandomi, A.H. (2021). Optimized machine learning approaches for the prediction of viscoelastic behavior of modified asphalt binders. *Construction and Building Materials*, *299*, 124264.

Kaveh, A., Ghazaan, M.I., & Mohammadi, A. (2026). Seismic Response Prediction of 3D Reinforcement Concrete Frames Using Machine Learning Methods. *Periodica Polytechnica Civil Engineering*.

Mansoori, N.A., Shaik, M.A., & Sivaramakrishnan, K. (2025). Machine Learning-Based Predictive Modeling of Infrared Spectroscopic Data from Thermal Conversion of Athabasca Bitumen. *ACS omega*, *10*(27), 29836–29855.

Miani, M., Dunnhofer, M., Rondinella, F., Manthos, E., Valentin, J., Micheloni, C., & Baldo, N. (2021). Bituminous mixtures experimental data modeling using a hyperparameters-optimized machine learning approach. *Applied Sciences*, *11*(24), 11710.

Nguyen, H.L., & Tran, V.Q. (2023). Data-driven approach for investigating and predicting rutting depth of asphalt concrete containing reclaimed asphalt pavement. *Construction and Building Materials*, *377*, 131116.

Olalekan, S.T., Olatunde, A.A., Kolapo, S.K., Omolola, J.M., Olukemi, O.A., Mufutau, A.A., Olaosebikan, O.O., & Saka, A.A. (2024). Durability of bitumen binder reinforced with polymer additives: Towards upgrading Nigerian local bitumen. *Heliyon*, *10*(10).

Penki, R., & Rout, S.K. (2023). Next-generation bitumen: A review on challenges and recent developments in biobitumen preparation and usage. *Biomass Conversion and Biorefinery*, *13*(11), 9583–9600.

Rahman, S., Bhasin, A., & Smit, A. (2021). Exploring the use of machine learning to predict metrics related to asphalt mixture performance. *Construction and Building Materials*, *295*, 123585.

Ren, S., Khadijeh, M., Assaf, E.I., Liu, X., & Erkens, S. (2025). Chemo-mechanical characterization and data-driven correlation analysis of rejuvenated bitumen via artificial neural network. *Materials & Design*, *256*, 114250.

Saleh, A., & Gáspár, L. (2025). Optimizing Foamed Bitumen Mixtures: AI-Based Determination of Ideal RAP and FBC Percentages Using HWTT and ITS Data. *Applied Sciences*, *15*(7), 3780.

Salehi, S., Arashpour, M., Golafshani, E.M., & Kodikara, J. (2023). Prediction of rheological properties and ageing performance of recycled plastic modified bitumen using Machine learning models. *Construction and Building Materials*, *401*, 132728.

Sanij, H.K., Babagoli, R., & Elyasi, R.M. (2025). Enhancing stone matrix asphalt performance with sugarcane bagasse ash: Mechanical properties and machine learning-based predictions using XGBoost and random forest. *Case Studies in Construction Materials*, e05186.

Sarker, T.R., Khatun, M.L., Ethen, D.Z., Ali, M.R., Islam, M.S., Chowdhury, S., Rahman, K.S., Sayem, N.S., & Akm, R.S. (2024). Recent evolution in thermochemical transformation of municipal solid wastes to alternate fuels. *Heliyon*, *10*(17).

Sheng, J., Liu, L., Liu, Z., Su, Y., & Li, L. (2025). Advanced Machine Learning for Predicting Asphalt Mixture Performance: A Data-Driven Approach with Model Interpretation. *International Journal of Pavement Research and Technology*, 1–20.

Tran, M.H., Paramasivam, P., Le, H.C., & Nguyen, D.T. (2024). Biomass: A Versatile Resource for Biofuel, Industrial, and Environmental Solution. *International Journal on Advanced Science, Engineering & Information Technology*, *14*(1).

Ullah, S., Zaidi, S. B. A., Khan, D., Fareed, A., Mohd Hasan, M.R., Milad, A., & Ali, B. (2024). Bio-oil modified binder derived from cotton stalks as an eco-friendly alternative binder for flexible pavements. *Scientific Reports*, *14*(1), 13519.

Upadhya, A., Thakur, M., Sihag, P., Kumar, R., Kumar, S., Afeeza, A., Afzal, A., & Saleel, C.A. (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.

Varanda, C., Portugal, I., Ribeiro, J., Silva, A.M., & Silva, C.M. (2017). Optimization of bitumen formulations using mixture design of experiments (MDOE). *Construction and Building Materials*, *156*, 611–620.

Wang, X., Ma, J., He, W., Li, J., & Ma, Z. (2025). Improved Asphalt Binder Performance Grading Using Surrogate FTIR Data and Gradient Boosting. *Energy & Fuels*, *39*(41), 19647–19656.

Yaro, N.S.A., Napiah, M.B., Sutanto, M.H., Usman, A., & Saeed, S.M. (2021). Modeling and optimization of mixing parameters using response surface methodology and characterization of palm oil clinker fine modified bitumen. *Construction and Building Materials*, *298*, 123849.

Yaro, N.S.A., Sutanto, M.H., Baloo, L., Habib, N.Z., Usman, A., Yousafzai, A.K., Ahmad, A., Birniwa, A.H., Jagaba, A.H., & Noor, A. (2023). A comprehensive overview of the utilization of recycled waste materials and technologies in asphalt pavements: towards environmental and sustainable low-carbon roads. *Processes*, *11*(7), 2095.

Yusuf, A. (2021). *Application of Neural Network for Predicting Properties of Concrete Using Bida Natural Gravel as Coarse Aggregate*

Zhang, F., Falchetto, A.C., Wang, D., Li, Z., Sun, Y., & Lin, W. (2025). Prediction of asphalt rheological properties for paving and maintenance assistance using explainable machine learning. *Fuel*, *396*, 135319.

Zhu, S., Qin, X., Xu, Z., & Xing, M. (2023). Life cycle assessment of energy consumption and carbon emissions of a green maintenance material for asphalt pavement: Warm mix OUFC-5. *Journal of Cleaner Production*, *428*, 139481.