

# Explainable Prediction of Asphalt Pavement Rutting Depth Using XGBoost and SHAP: Insights from the LTPP Database

Jiaxuan He

College of Urban Transportation, Beijing University of Technology, Beijing 100124, China

**Abstract:** Rutting is one of the most critical distresses affecting the performance and durability of asphalt pavements. This study developed an explainable machine learning framework for predicting asphalt pavement rutting depth by integrating XGBoost and SHAP interpretation methods based on the Long-Term Pavement Performance (LTPP) database. A total of 623 valid records from 357 control asphalt pavement sections in warm climate zones with or without freezing conditions were selected. Six key influencing factors—including Annual Average Daily Traffic (AADT), average temperature, precipitation, freeze index, humidity, and radiation—were used as input variables. The XGBoost regression model achieved excellent predictive performance on the test set, with an  $R^2$  of 0.9434, RMSE of 0.9278, and MAE of 0.4083. SHAP analysis revealed that AADT and average temperature are the dominant factors contributing 42.1% and 34.6% respectively to rut depth prediction, forming a clear “dual-core driving” mechanism. Precipitation showed moderate influence, while freeze-thaw, humidity, and radiation had relatively minor effects. The results highlight the strong nonlinear coupling between traffic load and high-temperature environment in rut development. This study not only provides a high-precision prediction model but also offers interpretable insights that can support scientific decision-making in anti-rutting pavement design, maintenance strategies, and regional management in high-traffic and high-temperature areas.

**Keywords:** XGBoost; SHAP; Asphalt Pavement Rutting.

## 1. Introduction

With the continuous growth of global traffic volume and complex changes in climate conditions, rutting on asphalt pavement, as a typical load dominant disease, has become one of the most significant challenges affecting pavement performance, durability, and driving safety. Rutting is manifested as permanent deformation of longitudinal grooves along the driving direction within the wheel track belt, which not only significantly reduces the smoothness and service level of the road surface, but also induces secondary risks such as vehicle handling instability, water accumulation slip, and even traffic accidents. In high temperature and high traffic areas, the development of ruts is particularly rapid and has become a core issue in the maintenance and management of highways and heavy-duty trunk roads in many countries [1-4].

The formation mechanism of ruts is essentially the cumulative plastic deformation process of asphalt mixture under repeated traffic loads. This process is controlled by multiple coupled factors, mainly including: (1) traffic load characteristics (such as axle load level, number of actions, traffic composition); (2) Environmental and climatic conditions (temperature, precipitation, humidity, radiation, freeze-thaw cycles); (3) Material and structural factors (asphalt performance, gradation, layer thickness, base modulus, etc.). Among them, traffic loads directly drive the shear flow and compaction deformation of the mixture by repeatedly applying shear and compressive stresses; Temperature significantly changes the viscoelastic behavior of asphalt (with a sharp decrease in modulus and increased viscosity at high temperatures), greatly reducing the material's ability to resist permanent deformation. A large number of accelerated loading tests and on-site observations

have shown that when the road surface temperature rises from about 34 °C to 70 °C, the rut depth can increase by nearly 4-5 times; At the same time, under heavy load conditions, for every doubling of axle load, the development rate of ruts often shows a non-linear acceleration trend[5, 6]. These mechanisms indicate that rutting is not the result of a single linear factor, but a complex nonlinear dynamic process under the coupling of multiple fields such as traffic temperature water environment.

The traditional rutting prediction methods mainly rely on empirical formulas (such as AASHTO empirical model, Shell model) and mechanical empirical models (such as the permanent deformation estimation module in MEPDG). These methods are typically based on limited experimental data or simplified assumptions (such as linear elastic or viscoelastic constitutive models), and exhibit significant limitations when dealing with high-dimensional, multi factor, nonlinear coupling problems: prediction accuracy decreases with prolonged service time, and adaptability to new materials, new structures, or extreme weather conditions is poor [7]. In recent years, with the advancement of data collection technology (such as LTPP long-term pavement performance database, pavement monitoring system) and the improvement of computing power, data-driven machine learning methods have gradually become an important tool for solving complex pavement engineering problems. Compared to traditional statistical regression, machine learning does not require explicit function forms between variables, and can automatically capture implicit nonlinear mapping relationships and interaction effects from massive observational data, demonstrating significant advantages in dealing with strong noise, high-dimensional, and heterogeneous data [8].

Among numerous machine learning algorithms, ensemble

learning methods based on Gradient Boosting Decision Tree (GBDT) have attracted much attention due to their excellent nonlinear fitting ability, robustness, and tolerance for noise. Among them, XGBoost (Extreme Gradient Boosting) is an efficient engineering implementation of GBDT. By introducing optimization strategies such as second-order Taylor expansion, regularization terms (L1 and L2), column sampling, row sampling, and parallel tree construction, XGBoost significantly improves training speed and generalization performance while maintaining high accuracy[9-11]. In recent years, XGBoost has been successfully applied in the field of pavement performance prediction, including modeling of multiple indicators such as International Roughness Index (IRI), fatigue cracks, and rut depth, demonstrating superior predictive capabilities compared to traditional regression, support vector machines, and random forests [12, 13]. Especially in the research of rutting prediction based on LTPP database, XGBoost and its variant models have repeatedly achieved high-precision results with  $R^2 > 0.90$ , demonstrating their potential in exploring the complex relationship between traffic environment performance [14].

However, high precision often comes with a sacrifice of model interpretability. The integrated tree models such as XGBoost are essentially "black box" models, and the weighted combination of thousands of decision trees inside makes it extremely difficult to directly understand "why a certain rut depth is predicted". This poses a significant obstacle in engineering applications: pavement engineers and management departments not only need accurate numerical predictions, but also need to know which factors are dominant drivers, how variables interact and affect each other, and the direction and magnitude of predicted changes under specific conditions, in order to support scientific maintenance decisions, material optimization, and design improvements [15].

To overcome this limitation, this study introduces the Shapley Additive Explanations (SHAP) method based on cooperative game theory for posterior interpretation of the XG Boost model. SHAP achieves a fully additive decomposition of prediction results by calculating the marginal contribution (Shapley value) of each feature under all possible feature combinations. This not only quantifies the importance of global/local features, but also reveals the positive and negative impact directions of variable values on the output and nonlinear features [16-18]. This method has demonstrated strong explanatory power in the field of civil engineering, such as predicting concrete strength and road surface smoothness, and can transform complex machine learning models into interpretable frameworks with physical significance [19, 20].

This study is based on the above background and fully utilizes the LTPP (Long Term Pavement Performance) database maintained by the Federal Highway Administration (FHWA) in the United States to screen for control asphalt pavement sections located in warm climate zones, with or without icing conditions and no maintenance or repair history. A total of 357 road sections and 623 valid records were obtained. These data cover key climate and traffic variables such as annual average precipitation, annual average temperature, freezing index, annual average relative humidity, annual average radiation, annual average daily traffic volume (AADT), as well as maximum rut depth as the target indicator[21].

The study used XGBoost regression model to construct a high-precision prediction framework for rut depth, and conducted systematic explanatory analysis through SHAP method, aiming to answer the following core questions: (1) How to rank the relative contributions of traffic load and environmental factors to rut depth in asphalt pavement dominated by warm climate? (2) Does AADT and temperature constitute a "dual core driving" mechanism, and is there a significant coupling amplification effect between the two? (3) Does the water environment (precipitation, humidity) and freeze-thaw factors only play an auxiliary role in the context of this study? (4) Can SHAP reveal the nonlinear and interactive features overlooked by traditional Pearson correlation analysis?

By combining high-precision prediction with interpretable analysis, this paper constructs a complete research paradigm of "data-driven mechanism revelation", aiming not only to achieve better predictive performance than traditional methods (target  $R^2 > 0.93$ ), but also to provide clear factors affecting road diameter and priority ranking for engineering practice, and to provide scientific basis for anti rutting design, drainage protection strategies, and regional differentiated maintenance decisions in high traffic and high temperature areas. At the same time, this framework also provides a method reference for predicting road performance by introducing more structural parameters, dynamic monitoring data, or multi-source heterogeneous data in the future.

## 2. Methods

With the widespread application of data-driven methods in the field of civil engineering, machine learning has gradually become an important tool for modeling complex engineering problems. Compared to traditional statistical methods, machine learning can automatically learn potential mapping relationships between input and output variables from data without explicitly assuming functional relationships between variables, especially when dealing with multi factor coupling, strong nonlinearity, and high-dimensional data problems. In the field of pavement engineering, the development process of ruts is influenced by multiple factors such as traffic load, temperature, moisture, and material properties. Its formation mechanism is complex and has obvious nonlinear characteristics. Traditional linear models are difficult to accurately describe this process. Therefore, introducing machine learning methods to model and analyze the depth of wheel ruts is of great significance.

Among numerous machine learning algorithms, the Gradient Boosting Decision Tree (GBDT) model based on ensemble learning is widely used due to its excellent predictive performance and strong nonlinear fitting ability. Among them, XGBoost (Extreme Gradient Boosting) is an efficient implementation of GBDT, which introduces regularization terms and second-order degree optimization strategies to effectively improve computational efficiency and generalization ability while ensuring model accuracy. Therefore, this article uses the XGBoost regression model to predict and model the depth of vehicle ruts, in order to fully explore the complex relationship between traffic loads and environmental factors.

However, traditional machine learning models are often regarded as "black box models", and their internal decision-making mechanisms are difficult to explain directly, which is not conducive to revealing the mechanisms of various influencing factors. To address this issue, this article further

introduces the Shapley Additive Explanations (SHAP) method based on cooperative game theory to explain the model. The SHAP method achieves interpretability decomposition of prediction results by calculating the marginal contributions of each feature to the model output under different combinations. It not only quantifies the importance of each variable, but also reveals the positive and negative impact directions of variable values on the results and their nonlinear characteristics. By combining the XGBoost model with the SHAP interpretation method, a research framework of "high-precision prediction interpretable analysis" was constructed, which not only improves the prediction accuracy but also deeply reveals the influence mechanism of rut depth. The core idea of the XGBoost model is to iteratively construct multiple weak learners (decision trees) and combine them weighted to minimize the objective function. The overall prediction model can be expressed as:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (1)$$

where,  $\hat{y}_i$  is the predicted value of the  $i$ -th sample,  $f_k$  represents the  $k$ th regression tree,  $F$  is the set of all regression trees.

XGBoost trains the model by minimizing the following regularization objective function:

$$L = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (2)$$

where,  $l(y_i, \hat{y}_i)$  is the loss function (square error loss is used in this study),  $\Omega(f_k)$  is the regularization term used to control the complexity of the model, and its expression is:

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (3)$$

where,  $T$  is the number of leaf nodes,  $w_j$  is the weight of the  $j$ -th leaf node,  $\gamma$  and  $\lambda$  is the regularization parameter.

In terms of parameter settings, this study combines experience and experimental adjustments to select the following model parameters: the number of weak learners (`n_estimators`) is 500, the maximum tree depth (`x_depth`) is 4, the learning rate (`learning_rate`) is 0.03, the sample sampling ratio (`subsample`) is 0.8, and the feature sampling ratio (`colsample_bytree`) is 0.8. At the same time, L1 regularization term (`reg_alpha=0.1`) and L2 regularization term (`reg_lambda=1`) are introduced to suppress overfitting and improve the model's generalization ability.

### 3. Data Description and Preprocessing

The data used for analysis is sourced from the Long Term Pavement Performance (LTPP) database. The establishment of the LTPP database can be traced back to 1987, with the main purpose of studying methods for constructing high-performance pavements under different conditions. This database covers data from over 600 road segments and includes archival data from over 1900 road segments. The data collection conducted by LTPP mainly includes two schemes: general pavement research and specific pavement research (SPS). LTPP is supervised and operated by the Federal Highway Administration (FHWA), which is also responsible for providing free data access on its website at <https://infopave.fhwa.dot.gov/>. Researchers and professionals in the field of pavement engineering can utilize this valuable

database to support their analysis and research.

This study extracted control asphalt pavement segments located in warm climate zones, with or without icing conditions, and without maintenance or repair history from the LTPP database. The data selection process generated a total of 357 control road segments, including 623 individual records that met these specific criteria. Four main data types were retrieved from the LTPP database, including climate, transportation, and performance related information. Table 1 shows the selected attributes and specific variables used in each category. The collected variables include annual average precipitation, annual average temperature, annual average icing index, annual average relative humidity, annual average radiation, maximum rut depth, and annual average daily truck traffic volume.

**Table 1.** Summary of the collected data.

Data type	Data Attribute	Description
Climate	Precipitation(mm)	Average amount of rainfall in millimeters experienced over a year.
	Average temperature(deg C)	Average temperature in degrees Celsius experienced over a year.
	Freeze Index(deg C degree days)	Index that represents the severity and duration of freezing conditions.
	Average Humidity(%)	Average relative humidity experienced over a year.
	Average radiation level(W/sq m)	Average surface incident shortwave radiation for time period
Traffic	Annual Average Daily Traff	The average number of vehicles passing through a specific location on the highway every day.
Performance	Maximum rut depth (mm)	The maximum rut depth measured is an indicator for evaluating the degree of pavement deformation caused by traffic loads.

**Table 2.** Descriptive statistical analysis

variable	mean	standard deviation	min	median	Max
Precipitation	1008.61	426.14	54.10	1069.04	3174.74
Average temperature	13.96	5.38	-3.10	13.70	27.40
Freeze Index	222.31	323.44	0.00	90.00	2797.00
Average Humidity	67.97	12.27	28.00	72.00	89.00
Average radiation level	146681.80	19263.67	71143.60	148284.00	195328.60
AADT	902.89	1024.70	0.00	500.00	7218.00
Maximum rut depth	6.39	4.14	1.00	5.00	29.00

A dataset of 623 road observation records was used for modeling, which includes 6 environmental climate and traffic load related characteristic variables (annual precipitation, annual average temperature, freezing index, annual average relative humidity, annual average radiation level, annual average daily traffic volume AADT) and 1 target variable (maximum rut depth). Among them, the maximum rut depth

is set as the dependent variable of the regression task, and the remaining six variables are input as independent variables into the machine learning model. All variables are numerical and have no categorical features.

#### 4. Results and Discussion

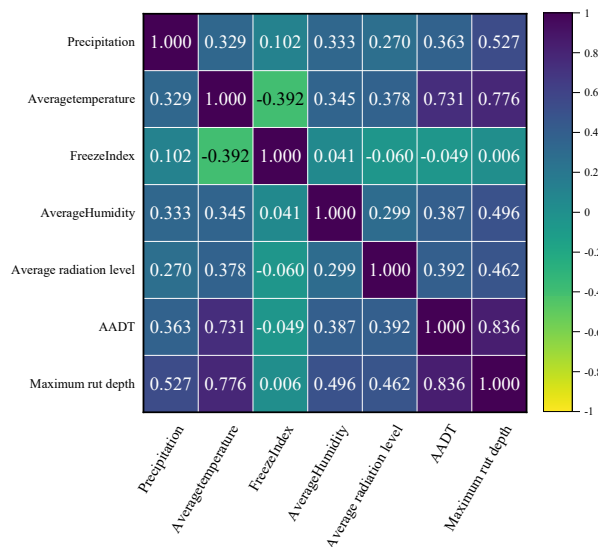


Fig 1. Correlation analysis chart

To explore the relationship between the development of road ruts and traffic loads and environmental factors, Pearson correlation coefficient was used to analyze each variable. The results in Figure 1 indicate that the overall relationship structure between variables is mainly positively correlated, and there are significant differences in the degree of influence of different factors on rut depth. The correlation coefficient between the maximum rut depth and various variables ranges from 0.006 to 0.836, indicating that different factors have a significant distribution of influence on it. Among them, traffic load shows the strongest correlation with temperature factors, while the influence of freeze-thaw factors is extremely weak. The analysis results show that the Pearson correlation coefficient between annual average daily traffic volume (AADT) and rut depth is 0.836, indicating a strong positive correlation. This indicates that traffic load is the dominant factor affecting the formation and development of ruts. From a mechanistic perspective, ruts mainly originate from the cumulative plastic deformation of asphalt mixtures under repeated vehicle loads. With the increase of traffic volume, the number of wheel loads significantly increases, and the internal shear stress of the pavement structure continues to accumulate, leading to the continuous development of irreversible deformation, ultimately manifested as deepening of ruts. Therefore, AADT not only reflects the frequency of load action, but also indirectly characterizes the total energy input borne by the road surface during service. In addition, this result is consistent with the classical pavement damage theory, which states that ruts are a typical "load dominant disease" and their development rate is closely related to the level of traffic load. The correlation coefficient between average temperature and rut depth is 0.776, showing a strong positive correlation and being an important influencing factor second only to traffic load. This result indicates that temperature conditions have a significant regulatory effect on the anti rutting performance of asphalt pavement. As the temperature increases, the viscoelastic properties of asphalt

materials change, with a decrease in modulus and an increase in viscosity, making the material more prone to permanent deformation. In high temperature environments, even under the same traffic load conditions, the road surface is more prone to generate larger rut depths. Further combining with the correlation matrix, it can be concluded that the correlation coefficient between average temperature and AADT is 0.731, indicating that there may be a certain degree of coupling between the two. This coupling relationship may originate from regional distribution characteristics or the combined influence of traffic and climate, thereby amplifying the impact of temperature on ruts to a certain extent. The correlation coefficients between precipitation, average humidity, and average radiation level and rut depth are 0.527, 0.496, and 0.462, respectively, all showing a moderate positive correlation. These environmental factors mainly affect the development of ruts through the following mechanisms: water damage, precipitation, and high humidity environments can weaken the adhesion between asphalt and aggregate, leading to delamination and reducing the overall strength and shear resistance of the material. Structural softening effect: After moisture infiltrates the pavement structure, it can cause a decrease in the strength of the base or surface layer materials, making them more prone to deformation under traffic loads. The radiation level of temperature coupling indirectly affects the development of ruts by influencing the road surface temperature, forming a synergistic effect mechanism with temperature factors. Although the above factors show a positive correlation, their correlation is significantly lower than that of traffic load and temperature, indicating that they mainly play an auxiliary or regulatory role rather than a dominant factor. The correlation coefficient between freeze-thaw index and rut depth is only 0.006, indicating almost no correlation. This indicates that within the scope of the data in this study, the impact of freeze-thaw action on the development of ruts can be ignored. The reasons for this phenomenon may include: the climate conditions in the study

area are relatively mild, and the freeze-thaw cycles are not significant; Car ruts are mainly formed by the combined effects of high temperatures and traffic loads, rather than being dominated by low temperature damage; Freeze thaw action is more likely to cause crack like diseases (such as fatigue cracks and low-temperature cracks), but has limited impact on wheel ruts. Therefore, in the subsequent modeling or variable selection process, the freeze-thaw index can be considered as a low priority variable. Based on the correlation analysis results, the formation mechanism of ruts can be classified into three dominant processes: traffic load dominant mechanism (Primary factor)

AADT determines the repeated load acting on the road surface and is the fundamental driving force for the development of ruts. The temperature control mechanism (Key environmental factor) significantly affects the deformation resistance of asphalt materials by regulating their mechanical properties. Secondary factors such as precipitation, humidity, and radiation have a secondary impact on the development of ruts by weakening structural performance or altering temperature fields. Overall, the formation of ruts is the result of the combined action of traffic loads and environmental factors, but the coupling effect of "traffic load temperature" is the most critical. The correlation analysis results have important guiding significance for practical engineering: in the rutting prediction model, AADT and temperature variables should be given priority consideration; In high-temperature and high traffic areas, emphasis should be placed on strengthening anti rutting design (such as improving the high-temperature stability of asphalt); Water damage protection (drainage design, waterproof layer) can be used as auxiliary measures; Freeze thaw factors can be appropriately weakened in the context of this study.

Subsequently, the maximum rut depth was taken as the research object, and traffic and environmental factors were selected as input variables to construct a predictive model. The SHAP method was further used for explanatory analysis of the model. The first six columns of the dataset are precipitation, average temperature, freeze-thaw index, average humidity, average radiation level, and annual average daily traffic volume (AADT). The seventh column is the target variable, rut depth. The above variables reflect the road service conditions from multiple perspectives such as water environment, thermal environment, cold climate, and traffic load. Among them, precipitation and humidity mainly affect the process of water damage, temperature and radiation affect the mechanical properties of materials, freeze-thaw index reflects the low-temperature cycling effect, and AADT directly characterizes the level of traffic load and is an important driving factor for the formation of ruts.

In the process of model construction, the data is first partitioned, and all samples are randomly divided into training and testing sets in an 8:2 ratio (random\_state=42) to ensure the reproducibility of the results. On this basis, the XGBoost regression model is selected to predict the depth of wheel ruts. The model parameters are set to: n\_estimators=500, max\_depth=4, learning\_rate= 0.03, subsample=0.8, colsample\_bytree=0.8, Simultaneously introducing L1 regularization term (reg\_alpha=0.1) and L2 regularization term (reg\_lambda=1) to improve the model's generalization ability and suppress overfitting. XGBoost, as an ensemble learning method based on gradient boosting trees, can effectively capture nonlinear relationships and

interactions between variables, making it particularly suitable for modeling complex engineering data.

To comprehensively evaluate the performance of the model, the coefficient of determination ( $R^2$ ), mean square error (MSE), root mean square error (RMSE), and mean absolute error (MAE) are used as evaluation indicators. At the same time, a 5-fold cross validation (KFold, n\_splits=5, shuffle=True, random\_state=42) was introduced to test the stability of the model. The results showed that the  $R^2$  of the model on the test set reached 0.9434, and the average  $R^2$  of cross validation was 0.9389 ( $\pm 0.0106$ ). The difference between the two was small, indicating that the model has good generalization ability and stability. In addition, the MSE is 0.8608, RMSE is 0.9278, and MAE is 0.4083, indicating that the prediction error is controlled at a low level and the model can accurately reflect the variation pattern of rut depth.

**Table 3.** Performance evaluation of test set

indicator	value
$R^2$	0.9434
MSE	0.8608
RMSE	0.9278
MAE	0.4083

**Table 4.** 50% cross validation results

indicator	mean	standard deviation
$R^2$	0.9389	0.0106
RMSE	1.0108	0.1316

In terms of model interpretation, the SHAP (Shapley Additive Explanations) method is used to analyze the importance and impact direction of each input variable. The SHAP method originates from the Shapley value theory in cooperative game theory, which calculates the marginal contribution of each feature to the model output under different feature combinations, thereby achieving interpretability analysis of complex models. Specifically, this study constructs a SHAP interpreter based on a trained XGBoost model and calculates the SHAP values of the test set samples. The importance ranking of each variable is obtained by taking the absolute value average, and a SHAP summary plot (beeswarm plot) is drawn to analyze the distribution of the impact of variable values on the prediction results.

The SHAP analysis results in Figure 2 indicate that the contribution of each variable to the depth of wheel ruts presents a distinct layered feature. Among them, the average SHAP value of AADT is 1.3868, which is the highest among all variables, indicating that traffic load is the dominant factor affecting the development of ruts. From the SHAP distribution map, it can be seen that high AADT (red) corresponds to the positive SHAP value area, and the distribution range is wide (about -2 to+4), indicating that with the increase of traffic volume, the rut depth significantly increases, and its impact has certain nonlinear characteristics, especially showing an increasing trend under high traffic levels. The SHAP value of the average temperature is 1.1764, second only to AADT. Its SHAP distribution range is larger (about -1.5 to+6), and the high-temperature samples are concentrated in the high positive contribution area, indicating that the increase in temperature will significantly reduce the deformation resistance of asphalt materials, thereby exacerbating the development of ruts and reflecting typical

temperature sensitivity behavior.

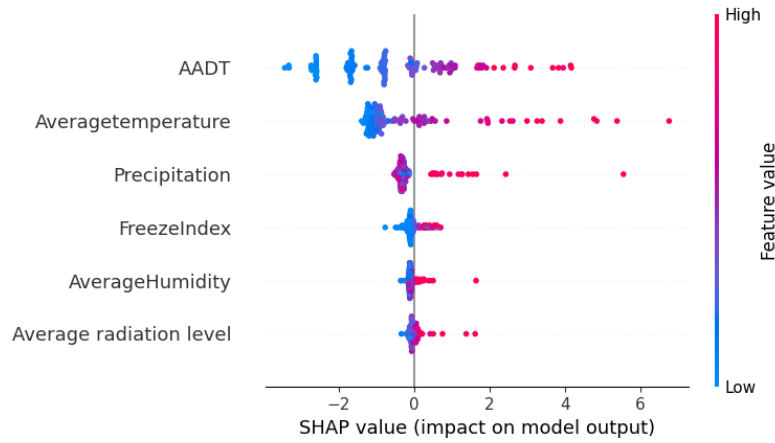


Fig 2. SHAP Figure

The SHAP value of precipitation is 0.4502, which is at a moderate level and mainly concentrated between 0 and 2. High precipitation corresponds to a positive contribution, indicating that precipitation promotes the formation of ruts through water damage mechanisms, but its impact is significantly lower than that of traffic and temperature factors. The SHAP value of the freeze-thaw index is 0.1827, although the overall contribution is small, it is not completely zero, and the high value is slightly biased towards the positive contribution area. This indicates that freeze-thaw effects may indirectly affect the development of ruts through material microstructure damage or strength degradation, reflecting a certain nonlinear or hysteresis effect. In contrast, the SHAP values of average humidity (0.1339) and radiation level (0.1169) are lower, and their distribution height is concentrated near zero, indicating that their direct impact on the depth of ruts is weaker, mainly playing an environmental regulating role, with radiation playing a more indirect role by affecting temperature.

Further combining SHAP analysis with traditional Pearson correlation analysis, it can be found that the two methods have consistency in identifying dominant factors, with AADT and average temperature being key control variables. However, the SHAP method further reveals the potential nonlinear contributions of variables such as freeze-thaw index, compensating for the shortcomings of linear analysis methods. This indicates that the development process of ruts is not a single linear mechanism, but a complex nonlinear process under the coupling effect of traffic loads and environmental factors.

After completing the SHAP value calculation, in order to further quantify the overall contribution of each feature variable to the prediction of rut depth, the absolute SHAP values of all samples were taken and averaged to obtain the global importance of each variable, and a feature importance bar chart was drawn. As shown in the figure, the importance of each variable presents a clear hierarchical structure, with AADT and average temperature significantly higher than other variables, indicating their dominant position in the development of ruts.

From a specific numerical perspective, the average SHAP value of AADT is about 1.5017, which is the highest among all variables, indicating that it contributes the most to the model's prediction results and is the primary factor affecting the depth of ruts. This result indicates that the level of traffic load directly determines the number of repeated actions that

the road surface bears, thereby controlling the cumulative development process of ruts. In practical engineering, with the increase of traffic volume, the frequency of wheel loading significantly increases, and the plastic deformation generated inside the asphalt mixture continues to accumulate, resulting in a continuous increase in the depth of ruts. Therefore, AADT not only has the highest importance in statistical significance, but also has a clear explanatory basis in mechanical mechanisms.

The importance of average temperature is second only to AADT, with an average SHAP value of 1.2340, indicating that temperature is a key environmental factor affecting the development of ruts. An increase in temperature will significantly reduce the modulus of asphalt materials, causing them to shift from elastic dominant to viscous dominant, making them more prone to permanent deformation under load. The importance of this variable in the bar chart is close to AADT, indicating that within the data range of this study, temperature and traffic load jointly constitute the "dual core driving factors" for the development of ruts, and there may be a certain coupling effect between the two.

Overall, the bar chart results clearly reveal the hierarchical relationship of the importance of each variable: AADT and average temperature form the first tier and are the dominant factors controlling the development of ruts; Precipitation belongs to the second tier and plays a secondary impact role; Freeze thaw, humidity, and radiation belong to the third tier, and their impact is relatively weak. This hierarchical structure is highly consistent with the variable influence pattern reflected in the aforementioned SHAP summary plot (beeswarm plot), further verifying the reliability of the analysis results.

Table 5. SHAP Feature importance ranking

ranking	feature name	Mean absolute SHAP value	Relative contribution rate (%)
1	AADT	1.5017	42.1
2	Average temperature	1.2340	34.6
3	Precipitation	0.4182	11.7
4	Freeze Index	0.1699	4.8
5	Average Humidity	0.1286	3.6
6	Average radiation level	0.1160	3.3

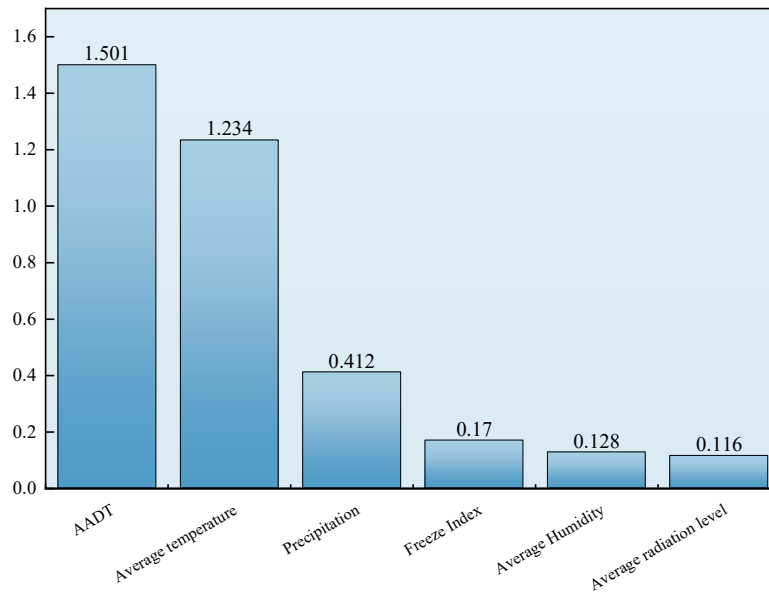


Fig 3. SHAP feature importance bar chart

In addition, this bar chart compensates for the shortcomings of relying solely on correlation analysis by conducting global statistics on SHAP values. Compared to traditional methods, SHAP can not only reflect the degree of correlation between variables and targets, but also comprehensively consider the nonlinear contributions of variables in different value ranges, thus more accurately evaluating their true impact. Therefore, the analysis results provide important basis for subsequent variable screening, model optimization, and pavement design.

Overall, the depth of ruts is mainly controlled by traffic loads and temperature factors, with water environment and freeze-thaw factors playing auxiliary roles. This conclusion not only conforms to the theory of pavement engineering, but also provides a clear direction for the prevention and control of ruts in practical engineering. For example, in high traffic and high temperature areas, the anti rutting performance of the pavement should be emphasized, and drainage design should be strengthened to reduce the impact of water damage.

## 5. Conclusion

This study is based on 623 control asphalt pavement observation records with no maintenance history in warm climate zones and partial icing conditions selected from the LTPP database. The impact mechanism of traffic load and environmental climate factors on the maximum rut depth of the pavement is systematically explored, and a high-precision and interpretable prediction model is constructed. The main conclusions are as follows:

Firstly, through Pearson correlation analysis and SHAP explanatory analysis, the results show that the annual average daily traffic volume (AADT) is the dominant factor affecting the development of ruts, with a correlation coefficient of 0.836 with rut depth, an average absolute SHAP value of about 1.50, and a relative contribution rate of 42.1%. This confirms from both statistical and model contributions that the repeated action of traffic loads is the fundamental driving force for the accumulated plastic deformation of ruts, which is in line with the basic understanding of "load dominant diseases" in classical pavement damage theory.

Secondly, the annual average temperature, as a key

environmental regulation factor, is second only to AADT in importance (with an average absolute SHAP value of 1.23 and a relative contribution rate of 34.6%), and shows a strong positive correlation with rut depth ( $r=0.776$ ). The significant decrease in modulus and increase in viscosity of asphalt materials under high temperature conditions lead to a decrease in their ability to resist permanent deformation, which is the core underlying mechanism for the accelerated development of ruts in high-temperature areas. The strong coupling relationship between AADT and average temperature ( $r=0.731$ ) further amplifies the synergistic effect of the two, jointly constituting the "dual core driving" mechanism for the development of ruts.

Thirdly, precipitation, average relative humidity, and average radiation levels show moderate to weak positive contributions (SHAP values of 0.42, 0.13, and 0.12, respectively), mainly indirectly regulating the development of ruts through water damage, structural softening, and temperature fields, which are secondary influencing factors. It is worth noting that the overall contribution of the freezing index within the data range of this study is extremely low (correlation coefficient close to 0, average SHAP value 0.17), indicating that in the warm climate dominated road sections covered by this dataset, the direct impact of freeze-thaw cycles on rut depth can be ignored, and its effect is more likely to be reflected in crack diseases rather than rut diseases.

Fourthly, using the XGBoost regression model combined with careful parameter tuning ( $n_{\text{estimators}}=500$ ,  $x_{\text{depth}}=4$ ,  $\text{learning\_rate}=0.03$ , etc.), the model achieved excellent performance on the test set with  $R^2=0.9434$ ,  $\text{RMSE}=0.9278$ ,  $\text{MAE}=0.4083$ , and a 5-fold cross validation average  $R^2$  of 0.9389, demonstrating good prediction accuracy and generalization ability. Combining XGBoost with SHAP method not only achieves high-precision prediction of wheel rut depth, but also successfully transforms the "black box" model into a physically interpretable analysis framework, clearly revealing the nonlinear contribution direction, interaction effect, and importance level of each factor.

In summary, this study confirms that the development of asphalt pavement ruts is mainly controlled by the coupling mechanism of traffic load and high temperature environment,

with water environmental factors playing an auxiliary role, while freeze-thaw factors have limited impact in the context of this study. This conclusion provides clear directional guidance for pavement design, material selection, and maintenance strategy optimization: in high traffic and high temperature areas, priority should be given to improving the high-temperature deformation resistance of asphalt mixtures (such as using high viscosity modified asphalt and optimizing gradation design); Simultaneously strengthen the design of road drainage system and waterproof layer to reduce the impact of light water damage; For cold regions with significant freeze-thaw effects, comprehensive consideration should be given to crack prevention and control, rather than simply using rut depth as a control indicator.

This study validates the potential application of ensemble learning combined with interpretable AI methods in complex pavement engineering problems, providing methodological references and theoretical foundations for subsequent multi factor coupled prediction based on larger datasets, real-time monitoring data-driven dynamic rutting estimation, and fine-grained modeling considering material micro parameters. Future work can further introduce more pavement structure parameters, vehicle types, and long-term dynamic monitoring data to enhance the universality and engineering practical value of the model.

## References

- [1] ALNAQBI A J, ZEIADA W, AL-KHATEEB G G, et al. Creating rutting prediction models through machine learning techniques utilizing the long-term pavement performance database [J]. *Sustainability*, 2023, 15(18): 13653.
- [2] AL-SABAEI A M, SOULIMAN M I, JAGADEESH A. Smartphone applications for pavement condition monitoring: A review [J]. *Construction and Building Materials*, 2024, 410.
- [3] CAO W M, LIU Q F, HE Z Q. Review of Pavement Defect Detection Methods [J]. *Ieee Access*, 2020, 8: 14531-44.
- [4] ELSEICY A, ALONSO-DíAZ A, SOLLA M, et al. Combined Use of GPR and Other NDTs for Road Pavement Assessment: An Overview [J]. *Remote Sensing*, 2022, 14(17).
- [5] GOLANBARI B, MARDANI A, FARHADI N, et al. Applications of machine learning in predicting rut depth in off-road environments [J]. *Scientific Reports*, 2025, 15(1).
- [6] ZHAO Y L, GAO Y, ZHANG K, et al. Establishment of Rutting Model of Wheel-Tracking Test for Real-Time Prediction of Rut Depth of Asphalt Layers [J]. *Advances in Civil Engineering*, 2021, 2021.
- [7] CHENG C, WANG L, ZHOU X, et al. Predicting rutting development using machine learning methods based on RIOCHTrack data [J]. *Applied Sciences*, 2024, 14(8): 3177.
- [8] ZHAO J, WANG H. Machine learning based pavement performance prediction for data-driven decision of asphalt pavement overlay [J]. *Structure and infrastructure engineering*, 2025, 21(6): 940-55.
- [9] CHENG T J, YAO C F, DUAN J N, et al. Prediction of active length of pipes and tunnels under normal faulting with XGBoost integrating a complexity-performance balanced optimization approach [J]. *Computers and Geotechnics*, 2025, 179.
- [10] MEHDARY A, CHEHRI A, JAKIMI A, et al. Hyperparameter Optimization with Genetic Algorithms and XGBoost: A Step Forward in Smart Grid Fraud Detection [J]. *Sensors*, 2024, 24(4).
- [11] TAKEFUJI Y. Beyond XGBoost and SHAP: Unveiling true feature importance [J]. *Journal of Hazardous Materials*, 2025, 488.
- [12] CHEN S, CAO J, WAN Y, et al. Enhancing rutting depth prediction in asphalt pavements: A synergistic approach of extreme gradient boosting and snake optimization [J]. *Construction and Building Materials*, 2024, 421: 135726.
- [13] GUO R, FU D, SOLLAZZO G. An ensemble learning model for asphalt pavement performance prediction based on gradient boosting decision tree [J]. *International Journal of Pavement Engineering*, 2022, 23(10): 3633-46.
- [14] LIU B, JAVED D, HU J, et al. A unified framework for asphalt pavement distress evaluations based on an extreme gradient boosting approach [J]. *Coatings*, 2025, 15(3): 349.
- [15] ERFANI A, SHAYESTEH N, ADNAN T. Data-augmented explainable AI for pavement roughness prediction [J]. *Automation in Construction*, 2025, 176: 106307.
- [16] AL-JARAZI R, RAHMAN A, AI C F, et al. Development of prediction models for interlayer shear strength in asphalt pavement using machine learning and SHAP techniques [J]. *Road Materials and Pavement Design*, 2024, 25(8): 1720-38.
- [17] ERTEN K M, GÜRFIDAN R. Regression-Based Performance Prediction in Asphalt Mixture Design and Input Analysis with SHAP [J]. *Applied Sciences-Basel*, 2025, 15(19).
- [18] LEI B, CHEN J W, YU Y, et al. An interpretable model for predicting the performances of asphalt mixtures comprising the chemistry composition of steel slag [J]. *International Journal of Pavement Engineering*, 2025, 26(1).
- [19] SANDAMAL K, SHASHIPRABHA S, MUTTIL N, et al. Pavement roughness prediction using explainable and supervised machine learning technique for long-term performance [J]. *Sustainability*, 2023, 15(12): 9617.
- [20] ADNAN T, ERFANI A. Explainable AI for Predicting Pavement Roughness under Maintenance and No-Maintenance Scenarios [J]. *Results in Engineering*, 2025: 108666.
- [21] ELKINS G E, OSTROM B. Long-term pavement performance information management system user guide [J]. 2021.