Using Machine Learning to Forecast Pavement Condition and Roughness Based on Climate Zone

Mohammadmahdi Shahidi

The Department of Civil, Geological and Environmental Engineering

University of Saskatchewan

Saskatoon, Canada

mahdi.shahidi@usask.ca

Haithem Soliman

Associate Professor

The Department of Civil, Geological and Environmental Engineering

Saskatoon, Canada

haithem.soliman@usask.ca

Paper prepared for the session Innovative Technologies in Asset Management

2026 Transportation Association of Canada (TAC) Conference & Exhibition

Winnipeg, Manitoba

Abstract

In recent years, Machine Learning has became a useful tool for several Civil Engineering applications, including Pavement Engineering. This study presents a two-part machine learning framework for analyzing and predicting pavement performance. The Long-Term Pavement Performance (LTPP) database was used for this study. For the first part, Random Forest regression was used for multi-target prediction of pavement indicators including International Roughness Index (IRI), rutting depth, fatigue cracking, and bleeding. The model integrates environmental parameters, crack severity metrics, and climate zone classifications. In the second part, IRI progression across climate zones was modeled by training Random Forest models on age-restricted (asphalt concrete pavement < 21 years) segment data. The analysis revealled distinct deterioration trends in wet-freeze, wet-nonfreeze, dry-nonfreeze, and maritime regions. Results showed strong predictive accuracy ($R^2 > 0.9$) and significant climate-based variation in roughness progression. The findings support the integration of climate zone and surface distress analytics into data-driven Pavement Management Systems (PMS).

Keywords**:** Pavement Management Systems, Random Forest, Climate Zone, IRI, Surface Distress, LTPP, Machine Learning

1. Introduction

Pavement deterioration monitoring and forecasting are essential components of Pavement Management Systems (PMS), guiding maintenance scheduling, budgeting, and safety assessments. Traditional approaches rely on linear models or empirical thresholds that may fail to account for nonlinear interactions among environmental, structural, and traffic factors. The advancement of machine learning (ML) has enabled improved predictive modeling of pavement condition indicators, notably the International Roughness Index (IRI) and rutting.

The Mean Roughness Index (MRI), defined as the average of IRI values from both wheel paths, is a direct indicator of ride quality. MRI progression, especially under varying climatic conditions, reveals important aging behaviors relevant to long-term road performance. This study presents a two-phase investigation: (1) multi-target pavement condition prediction using Random Forest regression; and (2) MRI progression modeling across FHWA and AASHTO defined climate zones. This dual approach offers a comprehensive view of how data-driven methods can enhance condition forecasting while recognizing environmental variability.

The innovation of this study lies in its integration of climate zone classifications and distress severity measures to support targeted maintenance strategies and improve model generalizability.

The application of machine learning (ML) techniques to pavement performance modeling has expanded substantially in recent years. A growing number of studies have demonstrated the superior capability of ensemble methods such as Random Forest and gradient boosting in predicting pavement deterioration metrics like International Roughness Index (IRI) and Surface Curvature Index (SCI). These models are particularly effective at capturing nonlinear relationships, managing missing or noisy data, and scaling across diverse geographic and climatic conditions.

Marcelino et al. (2019)[4] developed a generalized ML framework using Long-Term Pavement Performance (LTPP) data, demonstrating that tree-based models can effectively generalize across varied regional

networks. Sandamal et al. (2023)[5] used explainable Random Forest models to predict IRI, incorporating cumulative traffic and pavement age as critical inputs. Ahmed (2024)[6] compared Random Forest and XGBoost algorithms for SCI prediction and showed that XGBoost achieved the highest accuracy using over 3,000 LTPP samples.

Increasingly, researchers have emphasized the role of climate variables in pavement condition modeling. Zeiada et al. (2020)[7] applied artificial neural networks to performance data from warm climate regions and identified freeze-thaw cycles, relative humidity, and extreme temperature as key contributors to surface degradation. These findings underscore the importance of integrating multiple climate variables into predictive frameworks.

However, much of the current literature remains limited by narrow spatial coverage, single-target modeling, or insufficient integration of climate data. Many existing models focus primarily on temperature or precipitation without accounting for critical seasonal phenomena such as freeze-thaw variability or snowfall intensity. This study addresses these gaps by combining multi-target prediction with climate-informed trend modeling using Random Forest regression. Five climate-related variables—mean annual temperature, annual precipitation, snowfall, freezing index, and freeze-thaw days—are used to classify pavement sections into established FHWA climate zones. By aligning prediction models with regional deterioration patterns and climate classifications, the study supports more informed, data-driven pavement management strategies.

2. Methodology

This study utilizes a two-phase modeling strategy based on cleaned and integrated Long-Term Pavement Performance (LTPP) data. The dataset included pavement sections from multiple U.S. states and Canadian provinces, focusing exclusively on asphalt concrete pavement samples with an age of less than 21 years to capture early-stage deterioration. Variables included environmental data (temperature, precipitation, snowfall, freezing index, freeze-thaw days), geometric features (pavement width), and distress metrics (rutting depth, gator cracking, bleeding).

In the first phase, a Random Forest Regressor was trained, and the dataset was divided into a training set (80%) and a test set (20%) to simultaneously predict four pavement condition indicators: MRI (Mean Roughness Index), rutting depth, gator crack area (GATOR_CRACK_A), and the probability of bleeding occurrence (expressed as a percentage). Two complementary measures were used to represent rutting depth: the maximum average lane-half depth obtained from a 1.8 m straightedge (MAX_MEAN_DEPTH_1_8_MEAN) and the maximum average lane-half depth measured relative to a wire reference line (MAX_MEAN_DEPTH_WIRE_REF_MEAN). While the wire reference measure reflects a wider and more continuous surface profile across the lane, providing a more thorough representation of rut geometry, the 1.8 m straightedge metric captures localized surface deviations over a short span, providing a quick assessment of rutting severity. Higher values indicate more severe rutting and poorer pavement performance. Both indicators measure surface depressions brought on by traffic loading. In data preprocessing, missing items were eliminated, numerical features were normalized, and categorical variables were encoded.

According to FHWA and AASHTO standards, climate zones were categorized using threshold values for temperature, precipitation, snowfall, freeze-thaw cycles, and freezing index; this framework combines AASHTO pavement design climate zone criteria with FHWA LTPP climatic classifications. Based on these

variables, observations were grouped into five categories, Dry-Freeze, Dry-Nonfreeze, Wet-Freeze, Wet-Nonfreeze, and Maritime, with the corresponding threshold ranges summarized in Table 1.

Table 1. Climatic Zone Classification Based on Environmental Variables

Climatic Zone	Mean Annual Temp (°C)	Annual Precipitation (mm)	Annual Snowfall (mm)	Freeze-Thaw Days	Freezing Index (°C-days)
Dry-Freeze	< 7	< 508	200–1000	40–90	300–1000
Dry-Nonfreeze	≥ 7	< 508	< 200	< 40	< 300
Wet-Freeze	< 7	≥ 508	≥ 1000	> 90	> 1000
Wet-Nonfreeze	≥ 7	≥ 508	< 200	< 40	< 300
Maritime	≈ 7	> 1000	Variable (300–800 est.)	50–90	300–800

Pavement condition classifications (Good, Fair, and Poor) were established using FHWA distress thresholds for bleeding and alligator cracking. Based on minor, moderate, and severe distress levels, a three-tier system was established; Table 2 provides particular threshold ranges for each group.

Table 2. three-Tier System Based on FHWA Thresholds

Class	Condition	Criteria
Good	Minor distress	GATOR_CRACK_A < 5 m^2 & BLEEDING < 5 m^2
Fair	Moderate distress	GATOR_CRACK_A 5–25 m^2 & BLEEDING 5–20 m^2
Poor	Severe distress	GATOR_CRACK_A > 25 m^2 or BLEEDING > 20 m^2

In the second phase, MRI values were modeled as a function of AGE_YEAR within specific climate zones: Dry-Nonfreeze, Wet-Nonfreeze, Wet-Freeze, and Maritime. Only states with sufficient segment-level data were retained for each zone; Arizona and California were used for dry-nonfreeze, Prince Edward Island for maritime, Minnesota, New York, Alberta, Manitoba, and Saskatchewan for wet-freeze, Florida, Georgia, Louisiana, Mississippi, Texas and Alabama for wet-nonfreeze and Ontario remained unclassified based on our thresholds. Random Forest regressors were trained for each state, and the resulting predictions were averaged to generate climate-specific roughness progression curves. Each climate zone curve is the result of combining several curves related to different states in that climate zone, and the curve of each state is the result of combining several different segment curves from that state using a Random Forest Regression model.

Finally, cross-tabulations were created to investigate the connection between pavement condition categories and climate zones. Standard accuracy metrics, such as the root mean square error (RMSE) and the coefficient of determination (R^2), were used to assess the model's performance. While RMSE evaluates

the average magnitude of prediction errors, indicating how closely estimated values match the observed outcomes, R^2 quantifies the fraction of variation in the observed data explained by the model, demonstrating overall goodness-of-fit. These criteria were combined to evaluate the model's accuracy across all target variables.

3. Results and discussion

The multi-target Random Forest regression model demonstrated high predictive accuracy across all pavement condition indicators. The model achieved an average R^2 score of 0.9019 and RMSE values of 0.2613 (m/km), 1.4732 (mm) and 1.8757 (mm) for the MRI and rutting depth metrics (MAX_MEAN_DEPTH_1_8, MAX_MEAN_DEPTH_WIRE_REF), respectively. The predicted alligator cracking area and the chance of bleeding occurrence also followed realistic patterns across climate zones, with the RMSE values of 1.5842 (area per 100 m^2 lane) and 2.5483% for gator crack area and bleeding, which are showing minor variations, suggesting regional dependencies.

Table 3. Multi-Target Model Evaluation Results

Parameters	RMSE	Average R^2 Score
MRI	0.2613 (m/km)	0.9019
MAX_MEAN_DEPTH_1_8	1.4732 (mm)	
MAX_MEAN_DEPTH_WIRE_REF	1.8757 (mm)	
GATOR_CRACK_A	1.5842 (area per 100 m^2 lane)	
BLEEDING	2.5483 %	

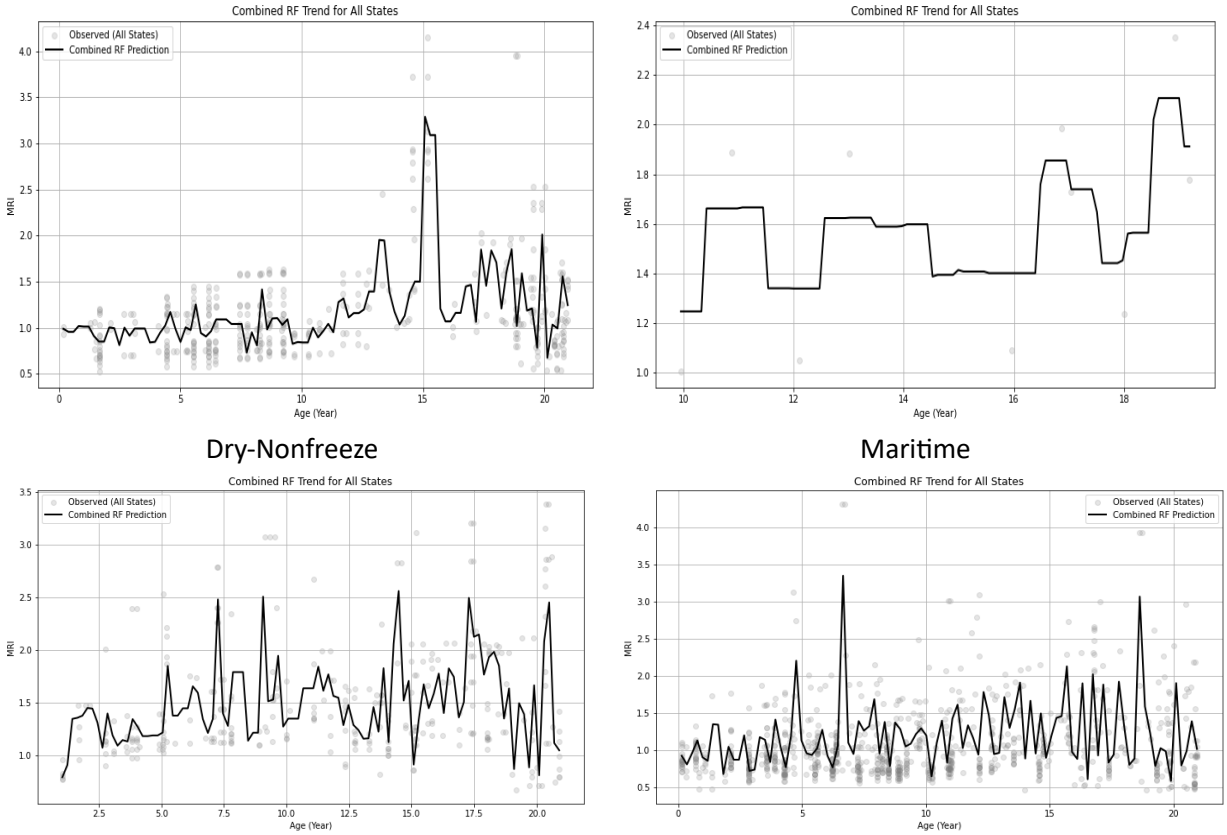
The amount of average R^2 score in the previous test models was higher (0.9772), but because of including some important parameters for the purpose of this project, like crack data in the modelling phase, the R^2 score reduced slightly to 0.9019. This reduction was expected due to overfitting risks and data volume decline following the dataset merge. Nevertheless, the model still performed well above industry benchmarks, demonstrating the feasibility of incorporating surface distress information into ML-driven PMS frameworks.

In the MRI progression analysis, climate zone-specific Random Forest regressions revealed clear differences in deterioration behaviour, Figure 1 demonstrates the combined random forest trend for all states in different climate zones. Comparison of MRI vs age trends across climate zones, which is shown in Figure 2, presents that the Wet-Nonfreeze regions exhibited gradual deterioration with consistent upward trends. Wet-Freeze zones experienced high variability, likely due to freeze-thaw cycles and moisture infiltration. Maritime regions presented smoother curves with minimal MRI escalation, likely reflecting effective drainage and stable thermal environments.

In the first ten years, the highest overall slope of the increase in MRI is related to the Wet-Freeze region, which is almost the same for this region in the second ten years. In the second ten years, the Dry-nonfreeze

and Maritime regions have maintained the same overall slope of the increase in MRI from the first ten years, but the Wet-Nonfreeze region has an increase in this slope in the second ten years.

If we want to compare the overall slope of the increase in MRI in this twenty-year period across the regions, the order of these regions from highest to lowest will be as follows: wet-freeze, wet-nonfreeze, dry-nonfreeze, maritime.



Dry-Nonfreeze Maritime

Wet-Freeze Wet-Nonfreeze

Figure 1: Random Forest Predicted MRI-Age Trends by Climate Zone

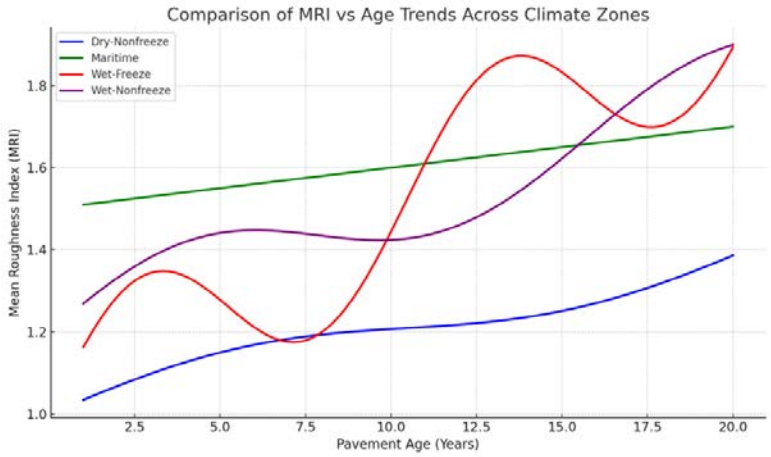


Figure 2: Comparison of Random Forest MRI Trends Across All Climate Zones

Figure 3, Pavement condition by climate zone cross-tabulation revealed that pavement segments in Wet-Nonfreeze and Dry-Nonfreeze zones exhibited the highest share of "Good" condition ratings, with over 75% of data falling into this category. In contrast, Wet-Freeze regions showed elevated rates of Poor and Fair conditions, likely due to the combined effects of freeze-thaw cycles and high precipitation. Notably, the Maritime zone, although limited in data volume, demonstrated the highest proportion of Poor conditions, emphasizing its unique deterioration pattern. These findings underscore the importance of incorporating region-specific climate segmentation into predictive maintenance models and pavement management planning.

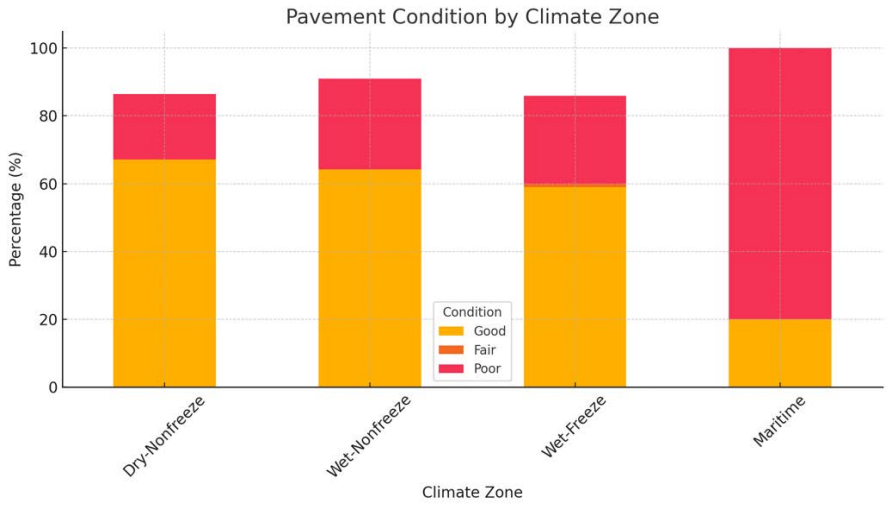


Figure3 : Pavement Condition by Climate Zone

4. Conclusions and recommendations

This study demonstrates the effectiveness of Random Forest regression in both multi-target prediction and climate-specific roughness modeling for pavement management. Predictive performance remained high across key indicators, with an average R^2 exceeding 0.9 even when incorporating surface distress features such as gator cracking and bleeding. MRI trends across climate zones highlighted deterioration patterns aligned with known environmental stressors, including freeze-thaw cycles, moisture levels, and aging mechanisms.

The findings emphasize the value of integrating climate classification and distress metrics into data-driven Pavement Management Systems. Specifically, Wet-Freeze and Wet-Nonfreeze zones require proactive rehabilitation planning due to more aggressive deterioration, while Wet-Freeze regions benefit from early preventive treatments before rapid post-10-year degradation. Maritime regions demonstrate smoother performance curves, suggesting that existing maintenance practices may be sufficient.

Recommendations

Based on the findings of this study, several practical recommendations can be made to support the implementation of data-driven pavement management strategies:

- The use of Random Forest or similar ensemble learning methods is strongly recommended for integration into national and regional Pavement Management Systems, given their ability to model complex interactions and deliver high predictive accuracy.
- Climate zone classification should be explicitly incorporated into maintenance prioritization frameworks, allowing agencies to tailor intervention strategies according to region-specific deterioration patterns.
- Future data collection efforts should be expanded to include real-time or near real-time inputs, such as those obtained from IoT-based monitoring systems, in order to improve the temporal resolution and predictive capability of pavement performance models.
- Continued research is recommended on hybrid modeling approaches that combine classification and regression techniques, particularly for multi-target pavement condition prediction, to further enhance model reliability and interpretability.

References

[1]Federal Highway Administration. *Distress Identification Manual for the Long-Term Pavement Performance Program (LTPP)*. Report No. FHWA-RD-03-031, U.S. Department of Transportation, Federal Highway Administration, McLean, Va., (2003)

[2]ASTM International. *Standard Practice for Computing International Roughness Index of Roads from Longitudinal Profile Measurements*. ASTM E1926-08, ASTM International, West Conshohocken, PA (2008)

[3]Breiman, L. “Random Forests.” *Machine Learning*, 45, 5–32 (2001)

[4]Marcelino, P., Ferreira, A., and Flintsch, G. “Machine Learning Approach for Pavement Performance Prediction.” *International Journal of Pavement Engineering*, 20(12), 1477–1489 (2019)

[5]Sandamal, K., Bandara, N., and Gunaratne, M. “Pavement Roughness Prediction Using Explainable and Supervised Machine Learning Techniques for Long-Term Performance.” *Sustainability*, 15(12), 9617 (2023)

[6]Ahmed, S. “Development of Pavement Performance Prediction Models Using Regression Methods.” *Open Journal of Civil Engineering*, 14(1), 36–52 (2024)

[7]Zeiada, W., Elseifi, M., Alavi, A. H., and Titi, H. “Machine Learning for Pavement Performance Modelling in Warm Climate Regions.” *Arabian Journal for Science and Engineering*, 45(12), 10491–10503 (2020)