Journal of Engineering and Computer Sciences
Qassim University, **Vol. 14, No. 1, pp. 1-17 (January 2021 /Rabi II 1441H)**

Comparison of random forest regression and multiple linear regression for predicting pavement roughness in dry climate countries

Fawaz Alharbi

Department of Civil Engineering,
College of Engineering, Qassim University,
51452 Buraydah, Qassim, Saudi Arabia,
Email f.a@qec.edu.sa

(Received 6/8/2021, accepted for publication 10/11/2021)

Abstract. An accurate prediction of pavement performance results in an effective plan for managing the highway network in cost-effective management and future maintenance strategies. However, in some dry climate countries, the available resources are not enough to conduct a periodic evaluation for their highways and apply a suitable maintenance action at a proper time. Therefore, the objective of this study is to utilize the available data in the Long-Term Pavement Performance (LTPP) program for pavement sections that are in dry-non-freeze zones. The selection of LTPP pavement sections in dry-non-freeze zones attempts to represent the pavement performance in dry climate countries. The International Roughness Index (IRI) was used as a performance indicator because it reflects the level of riding quality, the comfort of road users, and the level of pavement condition. The random forests (RF) and multiple linear regression (MLR) models predict the IRI for flexible pavements from pavement age, traffic and climate data, pavement distress data, and structural properties. The results show that the coefficient of determination (R^2) in the MLR model is 0.70, whereas the RF model yields a relatively higher R^2 value of 0.85. Also, the results of the RF model show that the initial pavement roughness was the most significant variable that impacted the pavement roughness, as well as, pavement thickness, pavement age, and truck volume have a high impact on the IRI value.

***Keywords**:* Pavement Management Systems, Random Forest Regression, Multiple Linear Regression, Pavement Roughness, LTPP data.

1. Introduction

Highway agencies invest millions of dollars per year in assessing the pavement conditions of the network and applying the required improvements such as preventive maintenance, rehabilitation, or reconstruction to make the roadway network meets the legislative, agency, and general requirements. Using the pavement management system (PMS) can assist the highway agencies in managing the roadway network, so the pavement conditions across the road network can maintain at an acceptable condition level [1]. The component of PMS includes inventorying pavement assets, performance prediction models to predict future conditions, which improve the efficiency of decision-making and visualizing tools to present the pavement conditions overall highway network. The most critical component in the PMS is performance prediction models, which support the decision-making process for different maintenance scenarios, as shown in Fig. 1. The weakness in this component can affect the accuracy of selecting maintenance actions that should be taken [2]. These prediction models relate a pavement performance indicator to many explanatory factors to create a simple function and determine influencer factors on the performance indicator. However, the ability and strength of the prediction models vary either because of insufficient collected data or lack of statistical models [3]. Recently, many studies have utilized sophisticated prediction models in the PMS Field. The remarkable finding is capturing and evaluating the uncertainty in the model either from material properties, climate conditions, and traffic loading data which can affect the model performance [4]. In addition, these factors are difficult to predict, which can affect pavement deterioration over the years. Therefore, a reliable pavement performance prediction model has required for the development of any pavement management system. This study utilized multiple linear regression (MLR) and random forests regression (RFR) models to predict the pavement roughness and compared their performance.

The selected pavement performance indicator is pavement roughness because highway agencies widely utilize it as an indicator of riding quality. The international roughness index (IRI) reflects the pavement roughness, which is measured based on road profiler in units (m/km or in/mile) [5]. Therefore, the pavement roughness should be acceptable to ensure there is a suitable riding quality for both road users and goods. Moreover, the pavement roughness implies vehicle-operating costs in fuel consumption, maintenance cost, and tire wear cost. As a result, the fuel consumption will increase by 3 percent for each 1 m/km [6].

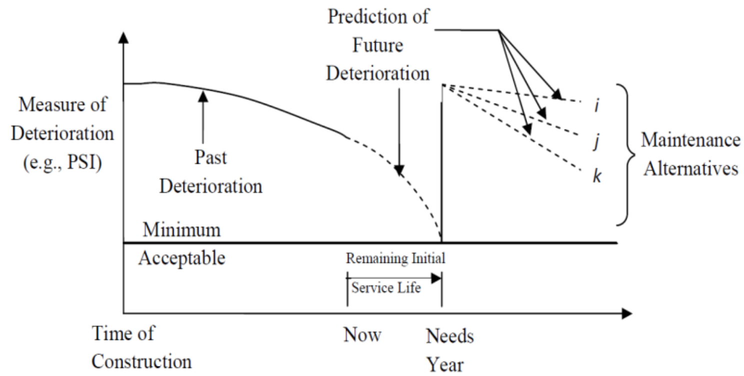

Fig. (1). Pavement performance over years [1]

This study utilized the LTPP database to predict the pavement roughness from pavement age, traffic volume, climate data, pavement thickness, and considered pavement distress. The Federal Highway Administration (FHWA) established the LTPP to collect all related data about pavements structure in North America at around 2500 pavement sections located in different climate zones [7]. The availability of this data has helped the transportation agencies save budgets by collecting distress data that needs a lot of workers and equipment. Many prediction models utilized the LTPP data for specific counties or States under particular conditions in North America. However, this study focuses on the particular climatic zone, which is dry-non-freeze climate zones, because this climatic zone represents many of the low-income countries. Therefore, the study considers a vital effort to find an alternative way to develop prediction models for the highway in low-income countries with insufficient resources to collect and analyze data.

2. Research Objectives

The main objective of this research is to predict the IRI of asphalt pavements located in dry-non-freeze climatic zones. There are two main steps for achieving the research objective:

1. Extracting all available data in dry-non-freeze climate zone from the LTTP database in the United States.
2. Develop the conventional multiple linear regression model to predict the pavement roughness.

3. Develop the random forest regression model to predict the pavement roughness.
4. Determine all the critical variables that could influence IRI values and compare the results of the two developed models.

All the selected pavement sections used in the analysis have never been exposed to any maintenance activities, so the obtained data were just years before rehabilitation or significant maintenance action.

3. Methods

Predicting an accurate pavement performance model can help decision-makers and agency engineers to make proper decisions for maintenance and rehabilitation activities to improve the pavement performance in the highway network. However, predicting a robust pavement performance is not an easy task because the prediction model must include influence factors. In this study, multiple linear regression and random forest regression models were used to predict the pavement roughness of asphalt pavement in the dry-non-freeze climate zone in the United States.

3.1. Data Collection and Preparation

The data used in this research was obtained from the Long-Term Pavement Performance (LTPP) program. The FHWA established the LTPP program to collect pavement performance data as one major part of the Strategic Highway Research Program (SHRP). According to the FHWA, 2,500 pavement test sections in the North American highways are managed and monitored by the LTPP program. The LTTP program includes all related information about each pavement section, such as inventory, maintenance/rehabilitation activities, pavement condition, material, traffic, and climate [7]. In addition, the LTPP program divided the climate conditions in the United States into four zones: Wet-Freeze, Dry-Freeze, Wet No-Freeze, and Dry No-Freeze as shown in Fig. 2.

In this study, the collected data from the LTTP includes data until the year 2018 for asphalt pavement sections located in the dry-non-freeze zone. Therefore, the study covers 81 pavement sections in California, Arizona, New Mexico, Texas, Nevada, and Utah, where the climate zone is dray-non-freeze. The number of LTPP asphalt pavement sections at each state is listed in Table 1.

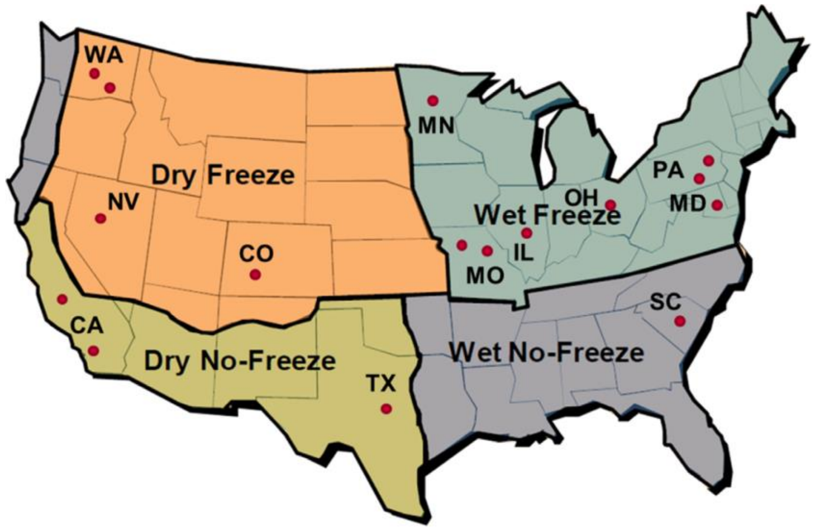

Fig. (1) Distribution of climatic zones in the United States [8]

Table (1). Number of selected LTPP sections at each State

State	Number of LTPP asphalt pavement Sections
California	31
Arizona	20
New Mexico	7
Texas	13
Nevada	9
Utah	1

The selected explanatory variables for predicting pavement roughness are based on previous studies in the literature. For instance, Gong et al. (2018) predicted pavement roughness of flexible pavements based on traffic, climate, structure, maintenance, and distress data from 11,000 data samples from the LTPP database. Likewise, Lucey et al. (2019) employed the LTPP data to predict the pavement roughness by using traffic data, pavement age, and structural properties as input variables for the prediction models [9].

Hossain et al. (2020) used the traffic and climate data from the LTPP database to predict the pavement roughness in ten sites in the United States [5]. Therefore, the collected data in this study include asphalt pavement thickness, pavement age, number of lanes, annual average precipitation, annual average temperature, annual average daily traffic, annual average daily truck traffic, fatigue (alligator cracking), wheel path longitudinal cracking, non-wheel path longitudinal cracking, transverse cracking, rutting, IRI, and initial IRI values. These data are summarized in Table 2. Each row in the data set has several attributes such as section ID, highway classification, number of lanes, structural properties, pavement distress, and climate data. The sections considered in the analysis were all the sections that had not been subjected to any maintenance or rehabilitation actions to get accurate forecasting models.

Table (1). The mean values of the considered variables

Variable	Mean
asphalt pavement thickness (mm)	195.5 mm
pavement age (year)	10.76 years
number of lanes (count)	2 lanes
annual average precipitation (mm)	285 mm
annual average temperature (^{0}C)	17 ^{0}C
annual average daily traffic (AADT)	5870 vehicles
annual average daily truck traffic (AADT)	1609 trucks
fatigue (m^2)	24.4
wheel path longitudinal cracking (m)	15.36 m
non-wheel path longitudinal cracking (m)	57.2 m
transvers cracking (count)	17.37
rutting (mm)	6.61 mm
international pavement roughness (m/km)	1.91 m/km
international pavement roughness (m/km)	0.96 m/km

3.2. Multiple Linear Regression (MLR)

Highway engineers have commonly used a multiple linear regression model to forecast pavement performance and generate a straightforward equation that engineers can use. Several studies adapted MLR models for predicting pavement performance. For instance, Chen et al. (2016) used a multiple linear regression model to predict the pavement condition of the highway network, which is important to plan future maintenance actions [10]. Other researchers utilized MLR for developing predictive models for pavement roughness based on pavement age and distress. For example, Abdelaziz et al. (2020) applied MLR to evaluate the impact of pavement thickness and traffic volume on pavement roughness as an indicator of pavement performance [11].

In this study, the R software developed the multiple linear regression to predict the pavement roughness. First, the function lm in the R package was utilized to fit the MLR model. Then, to select the best independent variables for the model, the stepwise regression method gave the candidate variables for predicting pavement roughness. Then, the multiple linear regression model is developed to evaluate the relationship between the pavement roughness and the independent variables. The significant independent variables influencing pavement roughness in asphalt pavement were initial IRI, rutting, transverse cracking, fatigue (alligator cracking), pavement age, asphalt layer thickness, and truck loading.

3.3. Random Forest Regression (RFR)

Random forest trees are decision trees constructed randomly based on bootstrapping or bagging with random feature selections and can be used for classification and regression. Unlike the MLR, the RFR model does not require any previous assumptions to be met, such as normality and variance homogeneity [12]. The performance of random forest trees works better than the single decision trees because many decision trees give less noise and then more accurate results [13]. In addition, the RFR considers the standard way to aggregate the outcomes of many trees with good predictive performance to produce a model that can deal with extensive data set, no over-fitting, and high accuracy results [3].

Cheng et al. (2019) illustrated the process of the RFR working in three steps [14]:

1. Different samples are randomly selected from the original database by using the bootstrapping technique.
2. From every single tress, multiple decision trees are built until they reach the maximum depth.
3. Combine the decision trees based on the majority voting strategy (overall, the tress in the forest). The outcome is the most predicted value across the trees; the process is illustrated in Fig. 3.

The random forest model has been applied recently in the transportation field to solve prediction and classification applications. Cheng et al. (2019) summarized 19 studies done since 2012 in four categories: travel choice behavior, traffic accident prediction, traffic flow/time prediction, and pattern recognition [14]. Their results show that the RF method can deal with different data types, nonlinear fitting relationships, and effective prediction problems. Despite the popularity of the RF method in the transportation field, its applications in the pavement management system field are limited. For example, Gong et al. (2018) and Marcelino et al. (2019) utilized the RF method to predict the IRI of flexible pavement based on structural properties, traffic data, and climate conditions [3] and [15]. In 2020, a research used the RF method to predict the structural capacity in the flexible pavement from surface deflection and air temperature [16].

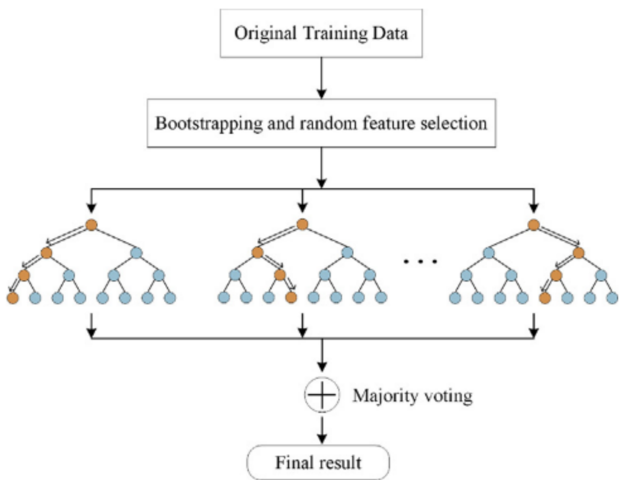

Fig.(2). Explanation of the random forest regression algorithm (Cheng et al., 2019)

In this study, the function random Forest of the R software was used to develop the RF model. The random Forest function requires selecting the number of variables (m) at each node in the trees to decide. The number of candidate variables at each split was four, and the number of trees for fitting the model was 500. The type of random forest was a regression, and it provides tools for assessing the model performance, which are residual errors and percentage of explained variance. In addition, the RF provides the variable importance for each predictor variable, which can be used to compare with other models.

For evaluating the performance of the two models, multiple linear regression and the random forest regression, the coefficient of determination (R2) and means square error (MSE) were utilized to assess and compare the performance of these models. The coefficient of determination (R2) represents the strength of correlation between the predicted and observed values, and it ranges between 0 and 1 (0 means no correlation, one means observed and predicted values are in agreement) [17]. The R^2 and MSE are defined in equations 1 and 2, respectively.

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

$$MSE = \frac{1}{n} \sum_i^n (y_i - \hat{y}_i)^2 \tag{2}$$

Where y_i is the actual value, $\hat{y}_i$ is predicted value, $\bar{y}$ is the average value, and n is the number of observations.

3.4. Essential Variables in Random Forest Regression

The RF algorithm has two valuable advantages over the other machine learning algorithms; high accuracy model among the most communally used models, and the importance of variables can be determined and evaluated [9]. Understanding which predictor variables are the most crucial task to improve and enhance the prediction models. The benefits of determining the importance of each variable are embodied in; a) removing the unreliable variables that impact the final prediction model, b) decreasing the cost of collecting and saving the data, c) improving the capability of the machine learning process [18]. Saha et al. (2016), Gong et al. (2018), and Cheng et al. (2019) presented the critical variables that show the contribution of each variable in the final prediction model [19], [3] and [14]. An RF typically uses the Gini impurity index to calculate the best split choice, which calculates the impurity of a particular variable concerning other impurities [20].

In this study, the Gini impurity index was employed to identify significant variables that affect pavement roughness. For the factor used to create the split, the Gini impurity index decrease at an internal node is calculated. A particular variable's significant value is considered the average decrease in Gini impurity index overall tress in the forest [14]. The calculation of the Gini impurity index for splitting in candidate variable Xi with different categories LJ is represented in Equation 3. When the Gini impurity indices are determined for each splitting variable choice, the separation is performed on a variable that has a higher value of the Gini impurity index.

$$G(X_i) = \sum_{j=1}^{J} P(X_i = L_j)\left(1 - P(X_i = L_j)\right) = 1 - \sum_j^J P(X_i = L_j)^2 \tag{3}$$

4. Results

4.1. Multiple Linear Regression

The multiple linear regression model was performed to predict the pavement roughness of flexible pavement in the dry-non-freeze area. A stepwise regression test was conducted to determine the most influential factors that could influence the pavement roughness. The MLR model selected significant factors based on a 95% confidence level with a p-value ≤ 0.05. The final MLR model includes seven factors: initial IRI value, rutting, pavement age, transverse cracking, fatigue, trucks load, and AC thickness, as shown in Table 3. At the same time, the annual average temperature and precipitations are not statistically significant in this model. However, as shown in Table 3, the coefficients of initial IRI, rutting, pavement age, fatigue, transverse cracking, and the truck volume are positively correlated, indicating that the roughness of asphalt pavement will increase with increasing any one of these variables. On the contrary, the pavement roughness will get better (less roughness) with increasing asphalt layer thickness as it appears as a negative relationship.

Table (2). Multiple linear regression model results

Variable	Estimate	Standard Error	t Stat	p-value
Intercept	0.314	0.079	3.967	8.626E-05
Initial roughness (m/km)	0.790	0.043	18.193	2.008E-54
Rutting (mm)	0.009	0.003	2.986	0.003
Pavement age (year)	0.008	0.002	4.431	1.212E-05
Transvers cracking (count)	0.003	0.001	3.832	1.47E-04
Fatigue (m^2)	0.001	0.000	5.041	6.99E-07
AC thickness (mm)	-0.001	0.000	-2.420	0.016
Truck (AADTT)	2.249E-05	1.084E-05	2.074	0.0387

The goodness of fit of the multiple linear regression model is represented by a coefficient of determination $R^2 = 0.70$, which means that the model can explain 70% from the analyzed data. Furthermore, the model residual plot is normally distributed, as shown in Fig. 4.

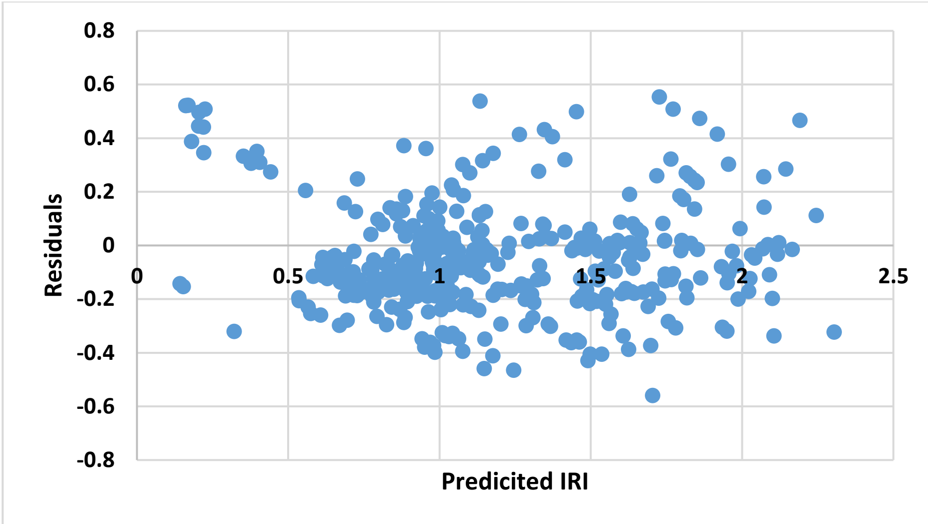

Fig. (3) Residuals versus predicted values

The relationship between the measured pavement roughness and predicted pavement roughness from the model is presented in Fig. 5.

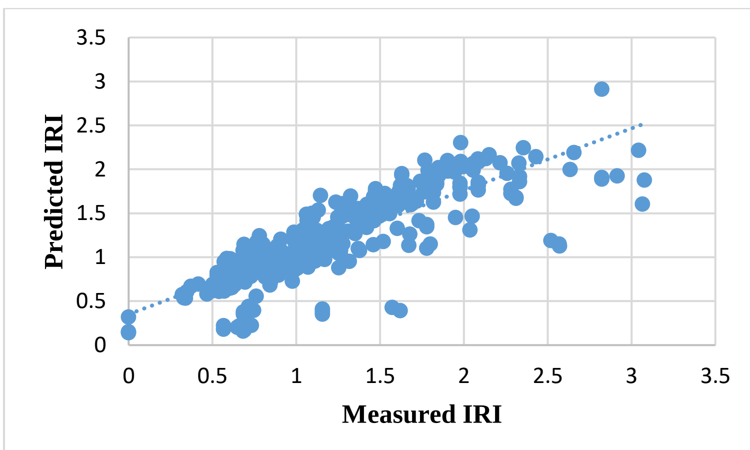

Fig. (4). Relationship between the measured and predicated IRI from the MLR Model

4.2. Random Forest Regression

The random forest model was developed to predict the pavement roughness based on 13 variables: initial IRI, AC layer thickness, pavement age, traffic loading, pavement distress (cracking, fatigue, and rutting), and climate data (temperature and precipitation). The percentage of the variance of the final model is 85%, and the fitted line between predicted and observed IRI is shown in Fig. 6.

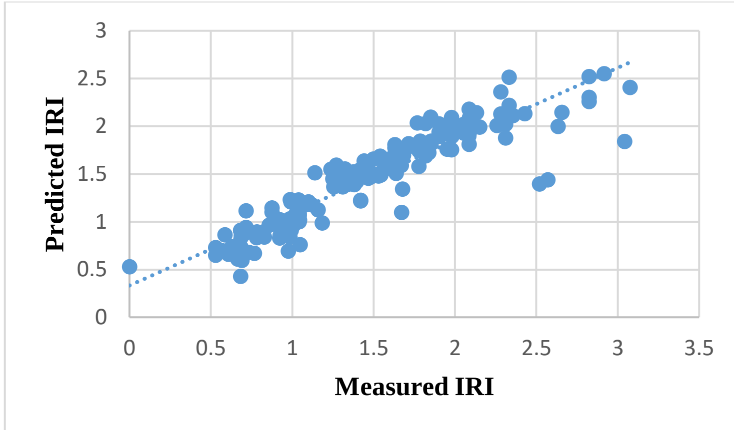

Fig. (6). Relationship between the measured and predicated IRI from the RFR model

The importance of variables is the critical output from the random forest model. A variable with a higher percentage of increase in mean square error means this variable has a higher impact on the model output. Fig. 7 shows the ranking of variables based on their importance to the pavement roughness in asphalt pavement type. The initial pavement roughness value is the most critical variable in predicting pavement roughness. That indicates pavement with low roughness value at an early age will be in good condition for a long time. Also, the results show the thickness of the asphalt layer, pavement age, transverse cracking, and the number of trucks are significant in predicting pavement roughness compared to climate factors as compared to the other pavement distress (Fatigue, rutting, and longitudinal cracking), traffic loads and the number of lanes.

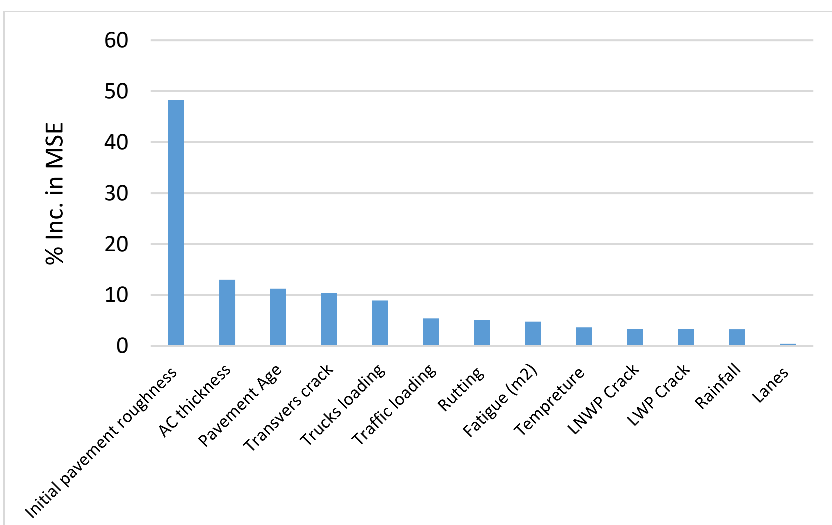

Fig. (7). Importance variables to the pavement roughness

The RFR model represents the model output in a tree to illustrate the procedure of predicting the response. For example, Fig. 8 shows a diction tree for predicting pavement roughness based on seven variables: initial pavement roughness, pavement age, rutting, AC thickness, longitudinal wheel path cracking, fatigue (alligator cracking), and transverse cracking. A diction is taken at each node from top to down until getting the predicted value.

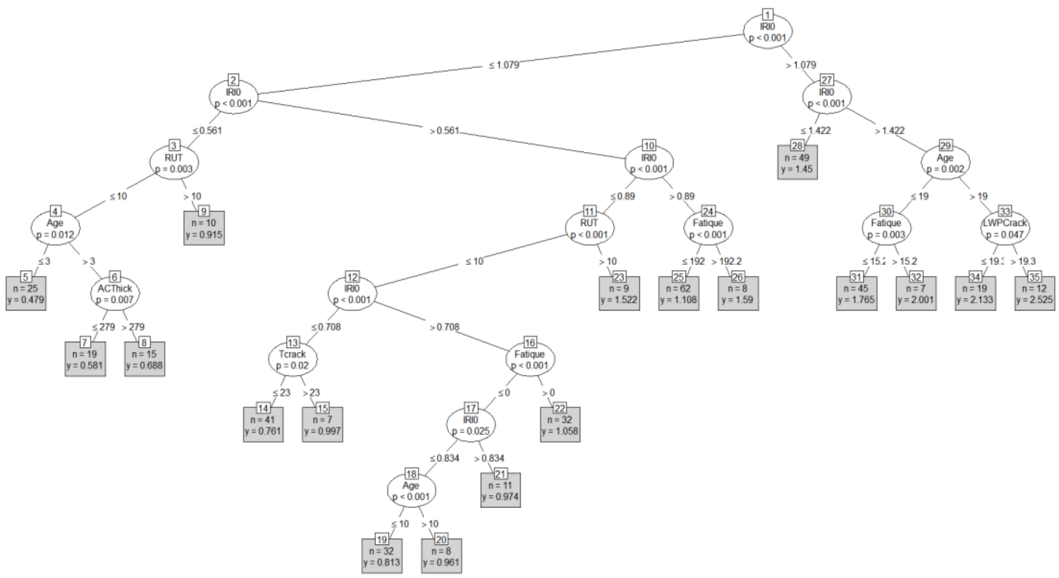

Fig. (8). The random forest trees for predicting IRI values

4.3. Comparison of random forest regression and multiple linear regression

The results of the comparison between RFR and MLR show that their results are acceptable; however, the RFR model provided better performance in predicting pavement roughness than the MLR model. The R^2 values of RFR and MLR are 85% and 70%, respectively, in predicting the pavement roughness in asphalt pavement in the dry-non-freeze climate area. Moreover, the mean square error of RFR and MLR models are 0.055 and 0.091, respectively.

The goodness of fit in the RFR model is better because of its ability to deal with nonlinear relationships between pavement roughness and the explanatory variables. The results of the two models indicated different factors that cause pavement roughness in asphalt pavement. For instance, the initial pavement roughness has been found as an essential factor affecting pavement roughness, which is consistent with a study done by [3]. In addition, the thickness of the asphalt layer and truck loading were identified as important factors in the RFR model. In contrast, these factors were not crucial in the MLR model, and that is because of the nonlinear relationships between these factors and pavement roughness. In addition, in the MLR analysis, the climate variables (temperature and rainfall) were not significant, indicating that the relationships between these data and pavement roughness were nonlinear associations.

5. Conclusion

Predicting pavement performance is the main component of pavement management systems and plays a major role in distributing the obtained money for maintenance and rehabilitation projects. This study aims to predict the IRI for flexible pavement that is located in dry-non-freeze climatic zones, which can be used in countries with the same climate conditions. The analyzed data were obtained from the LTPP database for all pavement sections that have not been subjected to any minor or major maintenance activities. The study used random forests regression (RFR) and multiple linear regression (MLR) to predict the pavement roughness from initial IRI value, traffic, climate, pavement age, layer thickness, and distress. The results show the RFR got $R^2 = 0.85$, whereas the MLR model ($R^2 = 0.70$).

Therefore, we can conclude that RFR is a more accurate model and can be deal with nonlinearity and extensive data. Also, rather than other machine learning algorithms, which are considered black-box models, the output of the RFR model can determine the relative importance variables.

This study determined the initial IRI value, pavement age, pavement thickness, and truck loading as essential variables affecting pavement roughness. Finally, the developed prediction models in this study can be considered performance models for the needed countries located in the same climatic zones until they build their models based on the locally collected data.

6. References

[1] R. Haas, W. R. Hudson and Z. John, Modern Pavement Management, Malabar, FL: Krieger Publication Company, 1994.

[2] Q. Dong, B. Huang and S. Richards, "Calibration and Application of Treatment Performance Models in a Pavement Management System in Tennessee," *Journal of Transportation Engineering* , p. 141(2), February 2015.

[3] H. Gong , Y. Sun , X. Xiang and B. Huang, "Use of random forests regression for predicting IRI of asphalt pavements," *Construction and Building Materials,* pp. 890-897, 2018.

[4] J. Prozzi, "Modeling pavement performance by combining field and experimental data," 2001.

[5] M. Hossain, L. Gopisetti and M. Miah, "Artificial neural network modelling to predict international roughness index of rigid pavements," *International Journal of Pavement Research and Technology,* 2020.

[6] I. Zaabar and K. Chatti, "Estimating vehicle operating costs caused by pavement surface conditions," *Transportation Research Record,* pp. 63-76, 2014.

[7] FHWA, "https://cms7.fhwa.dot.gov/research/long-term-infrastructure-performance/ltpp/long-term-pavement-performance," 2020. [Online].

[8] W. Zhang, S. Shen, A. Faheem, P. Basak, S. Wu and L. Mohammad, "Predictive quality of the pavement ME design program for field performance of warm mix asphalt pavements," *Construction and Building Materials,* vol. 131, pp. 400-410, 2017.

[9] J. Lucey, A. Fathi and M. Mazari, "Predicting Pavement Roughness as a Performance Indicator Using Historical Data and Artificial Intelligence," in *International Airfield and Highway Pavements Conference 2019*, Chicago, 2019.

[10] X. Chen, Q. Dong, H. Zhu and B. Huang, "Development of distress condition index of asphalt pavements using LTPP data through structural equation modeling," *Transportation Research Part C,* pp. 58-69, 2016.

[11] N. Abdelaziz, R. AbdElhakim, S. El-Badawy and H. Afify, "International Roughness Index prediction model for flexible pavements," *International Jouranl of Pavement Engineering ,* pp. 88-99, 2020.

[12] P. Smith, S. Ganesh and P. Liu, "A comparison of random forest regression and multiple linear regression for prediction in neuroscience," *Journal of Neuroscience Methods,* pp. 85-91, 2013.

[13] J. Wu and H. Xu, "Annual Average Daily Traffic Prediction Model for Minor Roads at Intersections," *Journal of Transportation Engineering, Part A: Systems,* p. 145(10), 2019.

[14] L. Cheng, X. Chen, J. De Vos, X. Lai and F. Witlox, "Applying a random forest method approach to model travel mode choice behavior," *Travel behaviour and society,* pp. 1-10, 2019.

[15] P. Marcelino, M. Antunes, E. Fortunato and M. Gomes, "Machine learning approach for pavement performance prediction," *International Journal of Pavement Engineering,* pp. 1-14, 2019.

[16] N. Karballaeezadeh, H. Therani, D. Mohammadzadah and S. Shamshirband, "Estimation of Flexible Pavement Structural Capacity Using Machine Learning Techniques," *preprint,* 2020.

[17] M. Rahman and R. Tarefder, "Dynamic modulus predictive model based on artificial neural network for the superpave asphalt mixtures of New Mexico," 2017.

[18] H. Han, X. Guo and H. Yu, "Variable Selection Using Mean Decrease Accuracy And Mean Decrease Gini Based on Random Forest," in *7th ieee international conference on software engineering and service science (icsess)*, 2016.

[19] D. Saha, P. Alluri and A. Gan, "A random forests approach to prioritize Highway Safety Manual (HSM) variables for data collection," *Journal of Advanced Transportation*, pp. 522-540, 2016.

[20] V. Rodriguez-Galiano, B. G. Rogan, M. Chica-Olmo and J. Rigol-Sanchez, "An assessment of the effectiveness of a random forest classifier for land-cover classification," *Journal of Photogrammetry and Remote Sensing*, pp. 93-104, 2012.

[21] Q. Dong and B. Huang, "Evaluation of effectiveness and cost-effectiveness of asphalt pavement rehabilitations utilizing LTPP data," *Journal of Transportation Engineering*, pp. 681-689, 2012.

Journal of Engineering and Computer Sciences
***Qassim University*, Vol. 14, No. 1, pp. 1-17 (January 2021 /Rabi II 1441H)**

مقارنة الانحدار للغابات العشوائية والانحدار الخطي المتعدد للتنبؤ بخشونة الرصف في البلدان ذات المناخ الجاف

فواز بن علي الحربي
قسم الهندسة المدنية، كلية الهندسة، جامعة القصيم، المملكة العربية السعودية
f.a@qec.edu.sa

ملخص البحث. ينتج عن التنبؤ الدقيق لأداء الرصف خطة فعالة لإدارة شبكة الطرق السريعة في إدارة فعالة من حيث التكلفة واستراتيجيات الصيانة المستقبلية. ومع ذلك ، في بعض البلدان ذات المناخ الجاف، لا تكفي الموارد المتاحة لإجراء تقييم دوري لطرقها السريعة وتطبيق إجراءات الصيانة المناسبة في الوقت المناسب. لذلك، فإن الهدف من هذه الدراسة هو الاستفادة من البيانات المتاحة في برنامج أداء الرصف طويل المدى (LTPP)لأقسام الرصف الموجودة في مناطق جافة غير متجمدة. يهدف اختيار قطاعات الرصف (LTPP) في المناطق الجافة غير المجمدة تمثيل أداء الرصف في البلدان ذات المناخ الجاف. تم استخدام مؤشر الخشونة الدولي (IRI) كمؤشر للأداء لأنه يعكس مستوى جودة الركوب وراحة مستخدمي الطريق ومستوى حالة الرصف. تتنبأ نماذج الغابات العشوائية (RF) والانحدار الخطي المتعدد (MLR) بـ IRI للرصفيات المرنة من بيانات عمر الرصف وحركة المرور والمناخ وبيانات الرصف الأولية والخصائص الهيكلية. أظهرت النتائج أن معامل التحديد (R^2) في نموذج MLR هو 0.70 ، في حين أن نموذج الغابات العشوائية RFR ينتج عنه قيمة R^2 أعلى نسبيًا تبلغ 0.85. كما أظهرت نتائج أن نموذج الغابات العشوائية RFR أن خشونة الرصف الأولية كانت من أهم المتغيرات التي تؤثر على خشونة الرصف، كما أن سمك الرصف وعمر الرصف واعداد الشاحنات لها تأثير كبير على قيمة IRI.