IMPROVING DESIGNING MODELS AND DEVELOPING NEW M&R DECISION PROCESS FOR FLEXIBLE PAVEMENTS

A Dissertation

by

MAHMOOD TABESH

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Maryam S Sakhaei Far
Committee Members,	Robert Lytton
	Dan Zollinger
	Phil Lewis
Head of Department,	Robin Autenrieth

December 2021

Major Subject: Civil Engineering

ABSTRACT

Pavements play a vital role in the transportation infrastructure in the United States. Pavement performance modeling is an essential step in pavement design and management. Recently, several software and tools have been developed to help to design a pavement at the project level. Pavement mechanistic-empirical (ME) design is one of the AAHSHTOWare Design software built to design new and rehabilitated pavements with flexible, rigid, and composite structures. The nationally calibrated performance models in Pavement ME do not well represent the construction and materials specifications, traffic, and climate conditions specific to each state and cannot precisely reflect the pavement performance. On the other hand, at the network level, the pavement performance should be monitored regularly, and maintenance and rehabilitation (M&R) treatments should be planned to keep the pavement in good condition. An acceptable treatment policy maximizes the service life and returns the benefits of the constructed pavement. The goal of this research is to enhance designing models for the flexible pavement of Oklahoma and develop a new M&R decision process using the surface roughness and structural capacity of the pavement section.

The nationally calibrated models show an improper prediction performance and a significant bias, which asserts the necessity of local calibration. Local calibration of Pavement Mechanistic-Empirical (ME) software improved the pavement performance prediction models and optimized the performance models for the pavement network of Oklahoma. The locally calibrated coefficients for distress and IRI models were determined for the Oklahoma pavement system. As a result, the error in performance prediction models was reduced through the calibration process.

The distress and IRI models show that the calibrated coefficients improve Pavement ME predictions and the design of flexible pavements in Oklahoma.

The second objective of this research was developing a new maintenance and rehabilitation decision process which considers the stochasticity of the pavement performance prediction and suggests the optimized maintenance activities for the given section by implementing a newly developed predictive model. The developed M&R decision method is a Markov Decision Process that employs IRI from a newly developed IRI prediction model and structural number from historical data. The IRI prediction model predicts the IRI with high accuracy by having the structural number, road class, climate condition, traffic load, and subgrade and structural information. Several advanced machine learning techniques were investigated, and the best model was implemented in the MDP M&R method. This model considers the M&R activities from pavement history, which affects the pavement deterioration rate, and suggests an M&R Policy for the given pavement system. By improving predictions and developing effective maintenance decision policies, machine learning algorithms can optimize maintenance and rehabilitation interventions and reduce maintenance costs.

ACKNOWLEDGMENTS

I would like to thank my committee chair, Dr. Sakhaei FAr, and my committee members, Dr. Lytton, Dr. Zollinger,and Dr. Lewis for their guidance and support throughout the course of this research.

Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience.

Finally, thanks to wife for her patience and love and to my mother for her encouragmenmt.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting Dr. Sakhaei Far, Dr. Lytton and Dr. Zollinger from the Department of Civil Engineering and Dr. Lewis from the College of Architecture.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

This work was made possible in part by Oklahoma Department of Transportation under Grant Number ODOT SPR Item Number 2277. The contents do not necessarily reflect the official views and policies of the Oklahoma DOT. This publication does not constitute a standard, specification, or regulation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Pavements play a vital role in the transportation infrastructure in the United States (US). About 80% of transportation and 68% of the cargo transportation were through roads and highways (BTS 2019). The state system of highways in Oklahoma encompasses 2.5 percent of the US roadways. On average, passenger vehicles, buses, and trucks traveled more than 73.7 million vehicle miles each day in 2018 on the state-owned highway systems. Over 90% of roadways are surfaced with asphalt concrete pavement (ODOT 2019). The FHA data lists 33% of Oklahoma's roadways in poor condition, which ranks Oklahoma as the second-worst state in overall road condition. Oklahoma’s roadways' current condition asserts the need for research projects to improve the pavement design practice and optimize the life cycle of pavement network in the state.

Pavement performance modeling is an essential step in pavement design and management from project to network levels. Several software and tools have been developed to help design a pavement at the project level. Pavement mechanistic-empirical (ME) design is one of the AAHSHTOWare Design software built to design new and rehabilitated pavements with flexible, rigid, and composite structures. The nationally calibrated performance models in Pavement ME do not well represent the construction and materials specifications, traffic, and climate conditions specific to each state and cannot precisely reflect the pavement performance (AASHTO 2010). Applying Pavement ME for designing flexible pavements with nationally calibrated performance prediction models may result in overestimation/underestimation of the asphalt thickness and properties that impose extra costs on the pavement construction overdesign early deterioration of the pavements. Therefore, many state agencies are trying to improve software outcomes by implementing different local materials, construction, climate, and traffic characteristics to evaluate and calibrate performance models.

On the other hand, at the network level, the pavement performance should be monitored regularly, and maintenance and rehabilitation (M&R) treatments should be planned to keep the pavement in good condition. An acceptable treatment policy maximizes the service life and returns the benefits of the constructed pavement. A proper pavement management system should provide a comprehensive, accurate perspective of the entire network to the manager to analyze the required amount of funds to extend the pavement network and keep the current network in the best possible shape. Determining the most effective treatment type and the most appropriate treatment schedule needs detailed information about the pavement structure and material properties, traffic loads, subgrade conditions, and climate information.

1.1. Objectives and Scope

The goal of this research is to improve designing models for the flexible pavement using case studies in Oklahoma and develop a new M&R decision process using the surface roughness and structural capacity of the pavement sections. This study has two main objectives.

The first one is local calibration and implementation of AASHTOWARE pavement ME performance models for Oklahoma pavement systems. The calibration effort for flexible pavements in the state of Oklahoma includes: a) compiling local information collected from previous studies for ODOT and national databases and developing a material database that is compatible with the Pavement ME Design software; b) calibrating Pavement ME Design by adjusting the distress model coefficients to eliminate the bias between predicted and measured pavement performance; and c) developing an analysis tool, named "INput-ME" , for converting and processing the local traffic data, climatic data, and material properties into Pavement ME Design input formats.

The second one is developing an intelligent M&R decision-making method for flexible pavements. The decision-making model has been developed based on the pavement structural capacity and pavement roughness and recommends the optimal M&R, which increases the benefit over the cost of the pavement. The developed M&R decision method is a Markov Decision Process that employs IRI from a newly developed IRI prediction model and structural number from historical data. The IRI prediction model predicts the IRI with high accuracy by having the structural number, road class, climate condition, traffic load, and subgrade and structural information. Several advanced machine learning techniques were investigated, and the best model was implemented in the MDP M&R method. This model considers the M&R activities from pavement history, which affects the pavement deterioration rate, and suggests an M&R Policy for the given pavement system. By improving predictions and developing effective maintenance decision policies, machine learning algorithms can optimize maintenance and rehabilitation interventions and reduce maintenance costs.

This research will be of interest to state road agencies with road maintenance responsibilities, particularly ODOT. By improving predictions and developing effective M&R decision policy, machine learning algorithms can optimize maintenance and rehabilitation interventions and reduce maintenance costs. Moreover, the approach presented for M&R decision planning is generic. It can be applied using different pavement performance indicators such as the Pavement Condition Index (PCI) and Present Serviceability Index (PSI) according to road agencies' needs and goals, leveraging machine learning applications' potential.

This study is outlined in 8 chapters. Chapter 2 presents the literature reviews of Pavement ME calibration efforts, pavement performance prediction modeling, and pavement maintenance and rehabilitation planning. Chapter 3 presents the local calibration methodology and process, discussing the calibration effect on the prediction models and development of Input ME. An overview of the Machine Learning models was presented in chapter 4. Chapter 5 explains the M&R decision-making method, including the dataset and explanatory data analysis, the IRI fitting model, IRI prediction models, and the Markov Decision model. Chapter 6 presents the machine learning model developments. Chapter 7 presents the plan for developing the Markov decision model. Finally, Chapter 8 presents the conclusions and recommendations based on the findings of this study.

2. LITERATURE REVIEWS

This chapter provides a comprehensive literature review on the following topics:

- Pavement ME local calibration
- Maintenance and rehabilitation (M&R) planning
- International Roughness Index (IRI)
- Structural Number (SN)

2.1. Pavement ME local calibration

Pavement mechanistic-empirical (ME) design is one of the AAHSHTOWare Design software built to design new and rehabilitated pavements with flexible, rigid, and composite structures. The mechanical-empirical design supports AASHTO's Mechanistic-Empirical Pavement Design Guide (MEPDG), which was generated under the National Cooperative Highway Research Program (NCHRP) 1-37A project (ARA 2004). This software predicts different pavement distress types, including top-down and bottom-up fatigue cracking, rutting and thermal cracking, and international roughness index (IRI) in the flexible pavements. This approach's mechanistic part calculates the cumulative damages over time based on the pavement responses to the traffic and environmental loads. Then the calculated damage will be transferred to pavement distresses through the existing empirical functions. These functions were calibrated and validated using the data gathered from the Long-Term Pavement Program (LTPP) pavement sections across the nation (Li et al. 2011).

The hierarchical input levels are features of Pavement ME design, which categorizes the input parameters in three distinct levels (ARA 2004). Designers can provide their input values from different levels based on the required data, the importance of the project, and allocated design time and budget. Level 1 data is calculated directly by in-situ or lab tests. This level provides the most precise data for input parameters with the highest data collection cost and the lowest uncertainty. Level 2 data has an intermediate level of accuracy and certainty. These data are not project-specific and can be estimated from correlation equations or the agency's database. Level 3 data have the lowest accuracy level and are based on the global or regional surveyed values. This data-level provides the least information about the input values in a specific project but is available at the lowest cost. In some input parameter values, which have low variance concerning the change in location and time, the level 3 data are excellent and reliable sources for designers. Designers, in any given project, may choose a combination of different levels of input parameters. Schwartz et al. conducted a global sensitivity analysis to capture the effect of variability of the design inputs on the performance prediction models (Schwartz et al. 2011). The calibration effort affirms the necessity and priority of the local calibration efforts' level of input variables. In designing flexible pavements, the variation of hot mix asphalt (HMA) surface layer properties have the highest effect on the predicted pavement performance. The traffic and climate data are listed as the next sensitive variables that impact the predicted performance variation.

The nationally calibrated performance models in Pavement ME do not well represent the construction and materials specifications, traffic, and climate conditions specific to each state and cannot precisely reflect the pavement performance (AASHTO 2010). Using Pavement ME for designing flexible pavements with nationally calibrated performance prediction models may result in overestimation/underestimation of the asphalt thickness and properties that impose extra costs on the pavement construction due to overdesign or early deterioration of the pavements.

Therefore, many state agencies are trying to improve the software outcomes by implementing different local materials, construction, climate, and traffic characteristics to evaluate and calibrate the performance models (Robbins et al. 2017).

The local calibration efforts of the MEPDG and Pavement ME reported by different state DOTs show an improvement in most of the performance prediction models (Darter et al. 2014; Kim et al. 2011; Williams et al. 2013; Hall et al. 2011; Kim et al. 2014; Li et al. 2009). The rutting and bottom-up fatigue cracking and IRI models were calibrated for the state of Arizona, and the calibration process substantially reduced the prediction error of the models. However, the transverse cracking model does not function properly before and after the calibration (Darter et al. 2014). The calibration effort for Colorado's state shows a noticeable improvement in the prediction of bottom-up fatigue cracking and IRI models, but the prediction errors for the rutting model increased (Mallela et al. 2013). The local calibration of rutting and bottom-up fatigue cracking models significantly reduced the prediction error and increased the models' accuracy for the flexible pavements of North Carolina (Kim et al. 2011). Top-down fatigue (longitudinal) cracking model was studied in a few calibration efforts, and a poor performance before and after the calibration effort was reported (Williams et al. 2013; Kim et al. 2014; Li et al. 2009). Generally, the rutting and bottom-up fatigue cracking and IRI models compared to the transverse and top-down fatigue cracking exhibit a better response to the local calibration process (Robbins et al. 2017).

The previous studies on the validation of Pavement ME distress models for flexible pavements of the state of Oklahoma assert the necessity of the local calibration and improving the accuracy of these models (Wang et al. 2014). The Oklahoma Department of Transportation has invested in multiple studies to improve the design of new flexible pavements and the rehabilitation strategies and preparing the database for material properties required by Pavement ME design (Nobakht et al. 2018; Nobakht et al. 2017; Hossain et al. 2011; Cross et al. 2011; Cross et al. 2007).

2.2. Maintenance and rehabilitation (M&R) planning

Pavement management system (PMS) can be defined as tools that can help the managers to beneficially decide for planning, designing, constructing, maintaining, and or reconstructing the pavements considering the agency and user benefits and costs (AASHTO 1990). There are different types of pavements with various service load levels and other environmental conditions in a network of pavements at the project, district, state, and national levels.

Determining the most effective treatment type and the most appropriate treatment schedule needs detailed information about the pavement structure and material properties, traffic loads, subgrade conditions, and climate information. This type of analysis can be funded for a limited number of sections at a project level. The project level needs information obtained through destructive tests by getting samples from the pavement layers and non-destructive evaluations (Santos et al. 2013). At the network level, the entire network would be considered and analyzed, and the proper treatments with respect to the available budget should be determined (Ferreira et al. 2002). The network-level analysis determines the treatment plan for the network by the cheapest and most informative pavement information. It estimates the impact of treatment scenarios on the funding and overall condition of the pavements. A proper pavement management system should provide a comprehensive, accurate perspective of the entire network to the manager so they can analyze the required amount of funds for extending the pavement network and keeping the current network in the best possible shape.

The transportation departments collect cost-effective and informative information from the annual condition surveying by developing new technologies and equipment, including pavement roughness, pavement distresses and cracking, falling weight deflectometer (FWD), and rolling weight deflectometer (RWD). At the network level, the managers should know the current condition and a reasonable estimation of the pavement's future status. International Roughness Index (IRI) and Pavement Condition Index (PCI) are two commonly used performance metrics to represent the pavement performance at both network and project levels (Fwa et al. 1991; Abaza 2005; Jorge et al. 2012). Researchers have suggested methods for using IRI for evaluating the pavement condition and impact and cost analysis of M&R strategies (Kelly et al. 2016; Saliminejad et al. 2013; Chopra et al. 2017; Hafez et al. 2019; Kırbaş et al. 2016). Butt et al. (1994) used a stochastic Markovian process for predicting pavement performance. Other researchers have used duration models (Prozzi et al. 2000) or other stochastic modeling methods such as power spectral density (Sun 2003), Bayesian models (Liu et al. 2014), deterministic regression, or incremental nonlinear approaches(Abaza 2004; Jannat et al. 2014; Prozzi et al. 2004), and artificial neural networks(Roberts et al. 1998; Lou et al. 2001; Yang et al. 2003; Kargah-Ostadi et al. 2010).

2.3. International Roughness Index (IRI)

The International Roughness Index (IRI) indicates pavement roughness, which is as low as 0.5 m/km for very smooth pavements and may reach 3 m/km for very rough and deteriorated pavements. Many transportation agencies use this index as a quality assurance criterion after the construction and as an index indicating the need for maintenance and rehabilitation at the pavement's terminal life (Perera et al. 2002). Pavement ME Design uses the IRI as the criterion for pavement design (ARA 2004). Also, studies show a strong correlation between the Present

Serviceability Rating or the Present Serviceability Index and IRI, which asserts IRI's application form indicating the pavement's serviceability over time (ARA 2004). Due to the IRI's importance as a pavement performance criterion, and the data's availability, the IRI modeling and prediction have been a topic of interest in many studies (Abdelaziz et al. 2020). Some of the IRI prediction models use the traffic characteristics and structural parameters as independent variables (George 2000; Albuquerque et al. 2011). Others use pavement distresses and site conditions as prominent features in the prediction model (ARA 2004; Khattak et al. 2014). The independent variables for IRI prediction models mainly include the pavement age, initial IRI, pavement distresses condition, climate features, soil properties, traffic condition, and structural parameters.

Researchers have applied different mathematical forms to fit a curve and predict the IRI change as a function of time. Linear and nonlinear models were the main regression tools for the prediction of IRI. The MEPDG models (ARA 2004; AASHTO 2015) are the well-known models that linearly fit the IRI to the independent variables, including the pavement distresses and structural factors. The performance of these models was thoroughly investigated in the previous chapters. (Khattak et al. 2014) studied the effect of the overlay on Louisiana's flexible pavements and developed an IRI prediction model using nonlinear regression analysis. The model depends on the road's functional class, cumulative equivalent single axle load, overlay thickness, temperature, precipitation, and deviation of the IRI after the treatment. This model gives the goodness of fit of %47 using 623 observations. The structural number of the pavement and age and traffic load levels was manipulated by (George 2000) to predict the IRI for Mississippi's flexible pavements. The model yields R^2 of %35 using 690 observations. (Al-Suleiman et al. 2003) used the age as the only parameter for predicting IRI and calibrated the regression parameters for slow and fast traffic lanes of flexible pavements. The first model gives an R^2 of %80 and the second model reach an R^2 of %61 using 440 observations. Other studies applied linear and nonlinear regression models with different input factors to predict the IRI of the flexible pavements (Zhou et al. 2008; Gulen et al. 2001).

Lin et al. (2003) used neural networks for the prediction of flexible pavement IRI from pavement distresses using 125 flexible pavement section of Taiwan pavement management system. This model shows the goodness of fit of %94. Choi et al. (2004) applied neural network for predicting the flexible pavement IRI using pavement properties, including structural number, AC thickness and material properties, and traffic load. The model used 117 sections of LTPP database and showed R^2 of %71. (Chandra et al. 2013) compared the linear regression performance, nonlinear regression, and ANN models in predicting the IRI from pavement distresses. The ANN model shows better performance compared to the linear and nonlinear regression models. The goodness of fit values were reported as R^2 and Mean Square Error (MSE), which were %86 and 0.22 for the training trial and %76 and 0.43 for the testing trial. Ziari et al. (2016b) used a complex polynomial model named Group Method of Data Handling (GDMH) to predict the flexible pavement roughness in short and long terms and compared the polynomial model results with a deep neural network model. The independent variables include age, traffic load, Annual average precipitation and temperature, AC layer thickness, and the pavement's total thickness. The ANN model shows the accuracy of higher than %90 in terms of R^2, and the GDMH model shows a performance between 80% to 90%. The ANN model was used in several other studies as a strong tool for IRI prediction (Kargah-Ostadi et al. 2010; Hossain et al. 2019; Mazari et al. 2016; La Torre et al. 1998). Other machine learning methods that were used for the prediction of IRI include random forest regression (Gong et al. 2018; Marcelino et al. 2019), fuzzy and gray model (Wang et al. 2011), Radial Basis Function (RBF) networks, and Support Vector Machine

(SVM) (Kargah-Ostadi 2014). Table 2-1 summarizes selected studies that implemented machine learning techniques, their independent variables, and the models' statistical test results.

2.4. Structural Number (SN)

The pavement structural integrity can be characterized by the structural number (SN). This number is an index that represents the overall strength of the pavement structure and the load-carrying capacity of the pavement (AASHTO 1993). The Equation (1) defines this index. It is a function of pavement layer thickness multiplied by coefficients relative to each layer's contribution to the pavement's structural strength. These coefficients were developed during the AASHTO road test, and SN has been used as a pavement design parameter in the AASHTO Design guideline.

$$SN = \sum a_i h_i m_i \quad (1)$$

Where

a_i = Structural coefficient of layer i,

h_i = layer thickness of layer i (in.), and

m_i = drainage coefficient of granular materials in layer i.

Table 2-1- Summary of studies applying machine learning techniques for IRI prediction

Study	Model Type	Age	IRI_0	Distresses	Climate Factors	SN	AC	Base	Sub	Traffic	Statistics	Source of Data
(La Torre et al. 1998)	ANN	✓			✓		✓	✓			N=144, R^2=0.63	LTPP
(Lin et al. 2003)	ANN			✓							N=125, R^2=0.94	Taiwan PMS
(Choi et al. 2004)	ANN					✓	✓			✓	N=117, R^2=0.71	LTPP
(Kargah-Ostadi et al. 2010)	ANN	✓	✓		✓		✓		✓		N= 8, R^2=0.95	LTPP
(Chandra et al. 2013)	L Regression			✓							N=510	India PMS
	NL Regression										N=510, R^2=0.79	
	ANN Modeling										N=510, R^2=0.76	
(Liu et al. 2014)	ANN	✓			✓		✓	✓	✓	✓	N=88, R^2=0.99	LTPP
(Kargah-Ostadi 2014)	ANN	✓	✓		✓		✓	✓	✓	✓	N=3361, R^2=0.94	LTPP
	RBF Network										N=3361, R^2=0.91	LTPP
	SVM										N=3361, R^2=0.9	LTPP
(Ziari et al. 2016a)	SVM	✓			✓		✓	✓		✓	N=205, R^2=0.91	LTPP
(Ziari et al. 2016b)	ANN	✓			✓		✓	✓		✓	N=205, R^2=0.9	LTPP
(Ziari et al. 2016c)	GDMH & ANFIS	✓			✓		✓	✓		✓	N=205, R^2=0.9	LTPP
(Mazari et al. 2016)	ANN	✓				✓				✓	R^2=0.99	LTPP
(Gong et al. 2018)	Random Forest	✓	✓	✓	✓		✓			✓	N=1990, R^2=0.97	LTPP
(Hossain et al. 2019)	ANN				✓					✓	N= 10 Section MSE= .028-0.058	LTPP
(Marcelino et al. 2019)	Random Forest				✓	✓	✓	✓		✓	N=7, R^2=0.93	LTPP
(Abdelaziz et al. 2020)	ANN	✓	✓	✓							N=7, R^2=0.75	LTPP
	L-NL Regression										N=7, R^2=0.57	LTPP

Notes: ANN: Artificial Neural network, L: Linear, NL: Nonlinear, RBF: Radial Basis Function, SVM: Support Vector Machine, ANFIS: Adaptive Network-Based Fuzzy, IRI0: Initial IRI, SN: Structural Number, AC: Asphalt Concrete Properties, Base: Baselayer Properties, Sub: Subgrade properties, N: Number of Observation, PMS: Pavement management System, LTPP: Long Term Pavement Performance

The pavement structural capacity can be estimated from the Falling Weight Deflectometer (FWD) data (AASHTO 1993). The current structural capacity is a function of the pavement modulus of all layers.

$SN_{eff} = 0.0045 * D * E_p^{0.333}$ (2)

Where

SN_{eff} = Current structural number of the pavement

D = Total thickness of pavement

E_p = Existing total pavement modulus

Wimsatt (1999) proposed a method for estimating the current pavement modulus. The subgrade modulus is a function of the deflection under the 7th sensor of the deflection basin (AASHTO 1993). However, the Wimsatt model gives the ratio of E_p to $E_{subgrade}$ through a simple regression model presented in Equation (3) . This method was of interest to researchers as it does not apply the complicated numerical-iterative method of the AASHTO Guide.

$$E_p / E_{subgrade} = 516.94 * (W7/W1)^{5/2} - 214.46 * (W7/W1)^{3/2} - 6.143 * (W7/W1) + 1.0826 * (W7/W1)^{1/2} \quad (3)$$

Where:

$E_p / E_{subgrade}$ = ratio of total pavement modulus to subgrade modulus, and

$W1,7$ = deflection under 1st and 7th sensors in mils.

The peak deflection under the loading plate is an essential feature in the FWD data. This deflection includes the compression of the pavement structure and deflection of the subgrade. Rohde (1994) concludes that the surface deflection at approximately 1.5 times the total pavement thickness is related to the subgrade layer and suggested the Structural Index of pavement (SIP) as the difference between the peak deflection and deflection at 1.5 times the total thickness of the pavement.

$SIP = D_0 - D_{1.5HP}$ (4)

where:

D_0 = Peak deflection under 9000 lb FWD load

$D_{1.5Hp}$ = Surface deflection at an offset of 1.5 times of Hp

Hp = Total pavement thickness

Rohde (1994) asserted that the SIP index is highly related to the stiffness of the pavement and, subsequently, with a structural number. A total number of 7776 pavement sections with different layer types, thickness, and stiffness were used, and the relation between the SIP and SN derived from the AASHTO model was investigated. It was found that the SN of the pavement can be estimated by having the total thickness of the pavement and SIP index using the Equation (5).

$SN = k_1 SIP^{k2} Hp^{k3}$ (5)

where:

SN = Structural number of pavements;

SIP = Structural Index of Pavement (microns);

Hp = Total pavement thickness(mm); and

k_1, k_2, k_3 = Regression coefficients = 0.4728, -0.4810 and 0.7581.

The results of the model proposed by Rohde (1994) shows very high accuracy in the estimation of SN from FWD data. This method is simple, and it can be easily implemented in the IRI prediction moles.

Many researchers believe that the structural capacity provides valuable information for project-level decisions and to some extent, in network-level M&R planning and prioritization (Haas et al. 1994; Gedafa et al. 2014). However, due to the expenses of deriving and analyzing data, the structural capacity is mostly evaluated at project levels. The FWD test is one of the most popular non-destructive tests for the assessment of pavement conditions. These test results can be used to back-calculate the modulus of different layers and estimate SN. Since performing the FWD test is expensive and time-consuming, FWD data is mostly available at a project level. Rolling weight deflectometer (RWD) is new testing equipment that can collect the pavement surface deflection data at highway speed. (Gedafa et al. 2010) discussed that the results of FWD and RWD tests are perfectly correlated, and RWD data can be collected from the pavement network at a reasonable time and cost. Thus, the RWD data will be available at the network level, and pavement structural capacity can be implemented in the M&R decision making process at the network level.

3. PAVEMENT ME LOCAL CALIBRATION PROCESS*

3.1. Methodology

The AASHTO Guide for the local calibration of the Mechanical-Empirical pavement design guide defines a guideline for the calibration of Pavement ME design according to the local material, climate, and pavement structural design (AASHTO 2010). The calibration effort for the state of Oklahoma follows the defined guideline, and it includes the following steps:

- Estimating the sample size for each distress type;
- Selecting the desired roadway segments;
- Selecting the hierarchical input levels for use in local calibration;
- Extracting and evaluating test section data;
- Analysing the sections using Pavement ME;
- Assessing bias for experimental sections; and
- Determining local calibration coefficients to eliminate the bias and improve the error

---

* Part of the data reported in this chapter is reprinted with permission from the following studies:

1- Cross, Stephen A, Robel Gibbe, and Nirajan Aryal. 2011. "Development of a flexible pavement database for local calibration of the MEPDG, Part 2 evaluation of ODOT SMA mixtures." In. Oklahoma Department of Transportation, Oklahoma City, OK.

2- Cross, Stephen A, Yatish Jakatimath, and Sumesh KC. 2007. "Determination of dynamic modulus master curves for Oklahoma HMA mixtures." In. Oklahoma Department of Transportation, Oklahoma City, OK.

3- Sakhaeifar, Maryam S, Y Richard Kim, and Pooyan Kabir. 2015. "New predictive models for the dynamic modulus of hot mix asphalt." *Journal of Construction Building Materials* 76:221-31.

4- Sakhaeifar, Maryam S., David Newcomb, Mona Nobakht, B.Shane Underwood, Padmini P. Gudipudi, and Jeff Stempihar. 2015. "Selection of long lasting rehabilitation treatment using life cycle cost analysis and present serviceability rating." In ODOT SP&R Item Number 2261. Oklahoma Department of Transportation, Materials and Research Division Oklahoma City, OK.

5- Sakhaeifar, Maryam S., Y. Richard Kim, and Pooyan Kabir. 2015. "New predictive models for the dynamic modulus of hot mix asphalt." *Construction and Building Materials* 76:221-31. doi: https://doi.org/10.1016/j.conbuildmat.2014.11.011.

6- Hossain, Zahid, Musharraf Zaman, Curtis Doiron, and Steven Cross. 2011. "Development of flexible pavement database for local calibration of MEPDG." In. Oklahoma Department of Transportation, Oklahoma City, OK.

3.1.1. Minimum Sample Size

The minimum number of required projects for Pavement ME distress and IRI model validation and calibration efforts depends on design reliability level and tolerable bias for each type of distress and can be calculated using Equation (6).

$$n = \left(\frac{t_{\alpha/2} s}{E} \right)^2 \quad (6)$$

Where n is the required number of the samples, s is the sample standard deviation of the mean values, E is the tolerable bias, and $t_{\alpha/2}$ is determined at a 90 percent confidence interval. Table 3-1 shows the distress/IRI threshold values and tolerable bias and standard deviation of the sample mean for the global Pavement ME models and the minimum number of projects required for calibration effort.

Table 3-1- Estimated number of pavement projects required for pavement ME validation and local calibration

Performance Prediction Models	Threshold Values	Standard Deviation	Tolerable Bias	Minimum Number of Project
Alligator cracking	20% lane area	5.3	2.5	13
Transverse cracking	630 ft/mi	235	100	16
Rutting	0.4 inch	0.11	0.05	14
IRI	169 inch/mi	20.3	10	12

Based on the statistical model, the minimum number of sample sizes for flexible pavements' calibration is 16. However, the local calibration guideline recommends using at least 30 pavement segments for flexible pavements (AASHTO 2010).

3.1.2. Selection of Roadway Segment

The experimental pavement sections for the local calibration effort represent the typical Oklahoma pavement design and construction practices and cover different pavements in poor, moderate, and good conditions. The LTPP database includes 53 sections of Oklahoma's flexible pavements. These sections spanned the construction age from newly constructed to older constructed pavements and rehabilitated sections. In addition to existing LTPP sections, nine sections from the previous Oklahoma pavement research study (Sakhaeifar, Newcomb, et al. 2015) were used for the calibration of Pavement ME Design. Pavement sections are located throughout the state of Oklahoma. However, to cover different climate conditions, three LTPP sections from Sherman and Ochiltree counties located in an uppermost part of Texas and one section from Norton County, Kansas, were also considered for the calibration effort performed in this study.

Oklahoma's state has been split by the interstate highway of I-35 into two regions of east and west. The east region has higher precipitation and average annual temperature, and the west region has lower precipitation and average annual temperature. The average annual temperature is between 56 to 63 ˚F, putting the two regions in the Non-Freeze climate category. Thus, to have better calibration results, the flexible pavement sections were divided into two groups of west and east regions. Figure 3-1 shows the roadway sections' location for local calibration effort of flexible pavements in Oklahoma, and Figure 3-2 shows typical flexible pavement groups used in the Pavement ME calibration effort in this study.

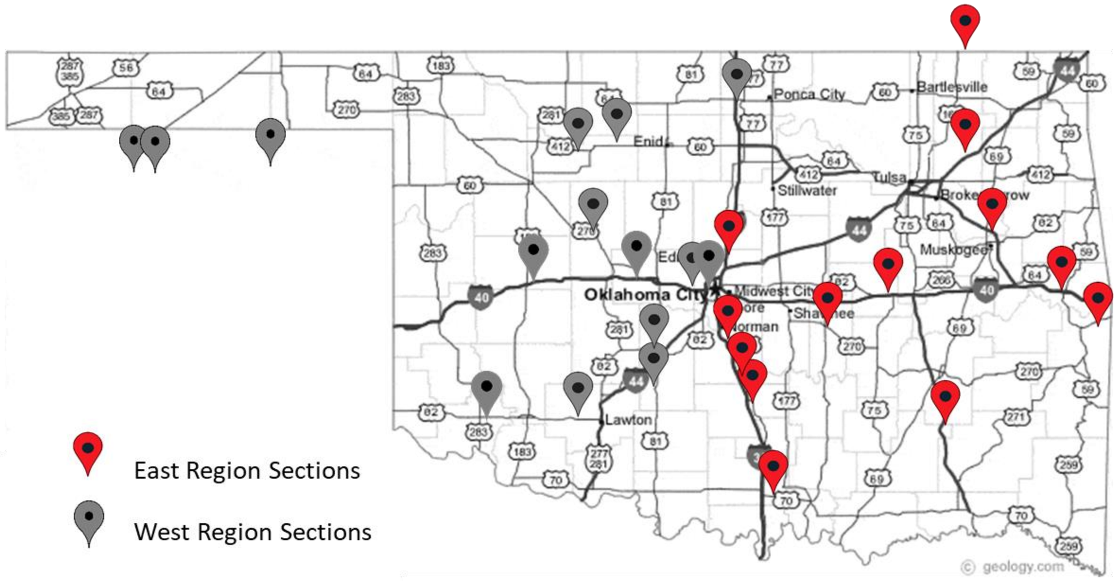


Figure 3-1- Location of roadway sections identified for local calibration effort of flexible pavements in Oklahoma

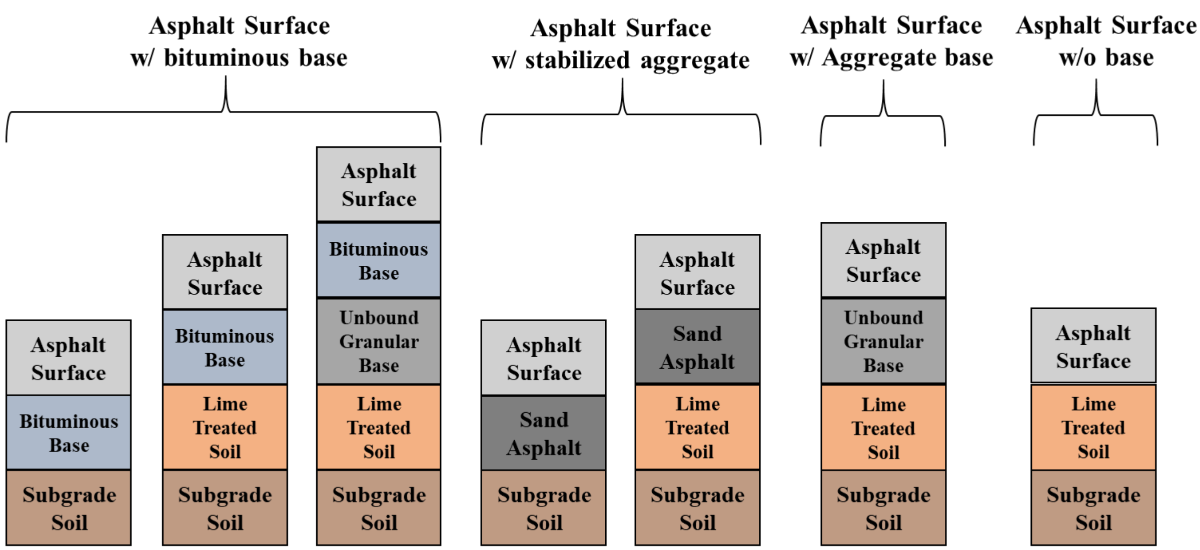


Figure 3-2- The representative flexible pavement sections used in local calibration effort

The selected pavement sections cover different types of roads and construction age. Many segments were constructed during the 1980s, and they experienced various kinds of rehabilitation and treatments. Besides, the selected sections mostly are the parts of principal arterial roads, which carry most of the traffic in Oklahoma. Figure 3-3 shows the road functional class of selected sections for the Pavement ME calibration process. The AC layer thickness has a significant effect on the performance of flexible pavements.

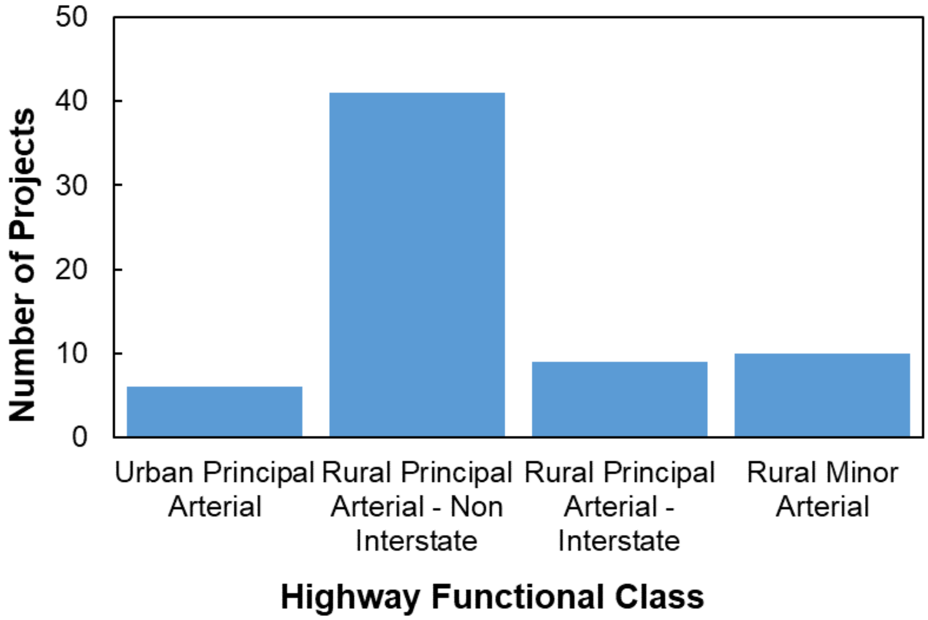


Figure 3-3- Highway functional class of projects selected for ODOT Pavement ME calibration process

Table 3-2 summarizes the distribution of layer thickness and the subgrade soil type of the selected sections. About 40 of the experimental sections have an AC layer less than 8 in. and the base thickness and subgrade soil types are almost equally distributed.

Table 3-2- Summary of the experimental section based on the layer thicknesses

HMA Thickness (in.)	Base Thickness (in.)	Number of Sections	
		w/ Coarse-Grained Subgrade Soil	w/ Fine-Grained Subgrade Soil
<8	< 6	9	13
	≥ 6	13	5

≥8	< 6	9	4
	≥ 6	5	7

3.1.3. Data Extraction and Evaluation

Extracting and evaluating the pavement construction data and preparing the master calibration database of material properties is essential and vital in the calibration process. The required data for the calibration of Pavement ME include asphalt concrete material properties, base and Subgrade/foundation field conditions and design properties, traffic information characteristics, climate, and environmental data. For this project, most of the required data were obtained from the LTPP database and previous research studies (Nobakht et al. 2017; Nobakht et al. 2018; Sakhaeifar, Kim, et al. 2015; Sakhaeifar, Newcomb, et al. 2015; Sakhaeifar, Richard Kim, et al. 2015; Cross et al. 2011; Cross et al. 2007; Hossain et al. 2011). The researchers obtained additional data, as needed, from climate and soil databases, including the Oklahoma water resource board and the U.S. Department of Agriculture (USDA) Natural Resources Conservation Service (NRCS) database. The missing and unreliable data were replaced by typical values recommended in the level 3 database.

3.2. Pavement Performance Model Calibrations

The LTPP and non-LTPP sections were analyzed using Pavement ME and collected input data sets. The software was run to analyze different pavement sections in various construction stages. Each section was considered as a new construction project until before the rehabilitation or reconstruction activity. After this stage, the pavement was analyzed as an overlay project, and the related models were calibrated. The predicted distress and IRI models were evaluated and compared with the measured distress values, and the accuracy and bias terms of each model were determined. The rutting and IRI models show better performance compared to fatigue bottom-up and top-down and thermal cracking models. The reason for bias and error terms in measured versus predicted performance prediction values mainly comes from inaccurate input data, error in the distress survey, and accuracy of performance prediction models.

The local calibration process involves investigating the causes for poor fit and bias in Pavement ME nationally calibrated models and modifying the calibration coefficients of distress and IRI models. The calibration process aims to reduce the bias and standard error of estimate (SEE) for the measured values versus predicted ones by changing the calibration coefficients for each model and finding the best set of coefficients suitable for the condition selected pavements in Oklahoma.

Calibration coefficients can be divided into two groups; 1) a portion of coefficients causes the reduction in the bias term of local calibration effort, and 2) another portion causes the decrease of the standard error term and an increase in the precision of estimates. The coefficients that correspond to eliminate bias were calibrated outside of Pavement ME through Microsoft Excel Solver, and a narrow down iterative approach was used to reduce the standard error. Figure 3-4 shows a schematic procedure for the calibration approach for the rutting model. This approach consists of finding the minimum standard error of estimate (SEE) by using a wide range of combinations for calibration coefficients and increasing the precision of the coefficients by narrowing the range of coefficients and finding the set of calibration coefficients that correspond to the minimum SEE.

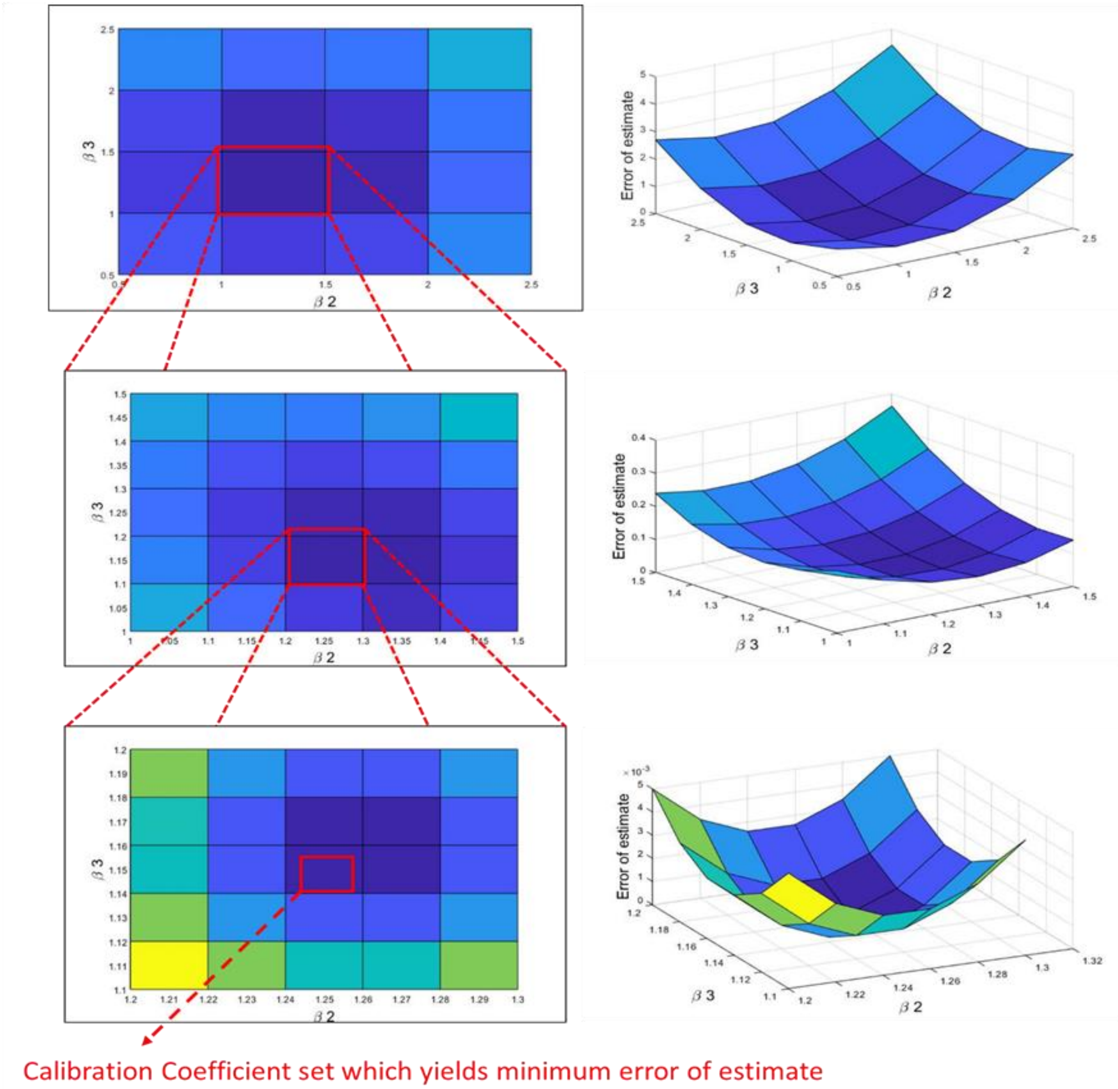


Figure 3-4- A schematic narrow down procedure for finding the rutting model calibration coefficients

In the following, the performance prediction models adopted in the Pavement ME design will be explained, and the method of calibration for each model will be elaborated.

3.2.1. Rutting Model

Surface distress in rutting is caused by plastic or permanent deformation in HMA, unbound layers, and foundation soil. The approach utilized in the Pavement ME is based upon calculating the incremental rutting within each sublayer. Rutting is estimated for each season in the middle of each layer in the pavement structure. The plastic deformation for each season is the summation of plastic vertical deformations in each layer. The model for calculating the total permanent deformation uses the plastic vertical strain under specific pavement conditions for the total number of trucks (AASHTO 2015). Conditions vary from month to month, so the “strain hardening” approach can be used to incorporate those plastic vertical strains within each month to estimate the cumulative deformation. The rate of accumulation of plastic deformation is measured in the laboratory using repeated load permanent deformation triaxial tests for both HMA mixtures and unbound materials. Among the five calibration coefficients affecting the rutting models, "$\beta r1$", "βgb" and "βsq" correspond to the amount of rutting in AC, granular base, and subgrade layers and were calibrated outside of the Pavement ME software by using the Microsoft Excel Solver. The calibration coefficients of "$\beta r2$" and "$\beta r3$" correspond to the effect of temperature and traffic on the AC layer, which was calibrated by numerous Pavement ME runs using the narrow down approach. The recommended calibration coefficients that correspond to the rutting of flexible pavements in east and west regions are presented in Table 3-3.

3.2.2. Fatigue Cracking Model

Two types of load-related cracks that are predicted by Pavement ME Design include alligator cracking and longitudinal cracking. The Pavement ME Design assumes that alligator cracks are triggered at the bottom of HMA layers and propagate to the surface under cyclic traffic load. In contrast, longitudinal cracks are assumed to initiate at the surface. The amount of fatigue damage to the asphalt concrete layer is a function of an allowable number of axle-load applications and corresponds to both types of load-related cracks (i.e., alligator and longitudinal)

(AASHTO 2015). The calibration coefficients which adjust the fatigue damage, i.e., "*βf1*", "*βf2*" and "*βf3*" were calibrated by iterative approach. The coefficients of *C1*, *C2*, and *C3* correspond to the empirical transfer functions, which estimate the amount of top-down and bottom-up cracking of flexible pavements as a function of cumulative fatigue damage. These coefficients can individually be calibrated for each type of fatigue cracking by optimizing through Microsoft excel solver. The fatigue damage and top-down and bottom-up fatigue cracking calibration coefficients for east and west regions are presented in Table 3-3.

3.2.3. Transverse Cracking Model

Transverse cracking happens due to the HMA surface's shrinkage at low temperatures, thermal fatigue at medium temperatures, or reflective cracks caused by cracks beneath the HMA layer's surface or load repetition. The primary reason for transverse cracking is contraction strains induced by temperature drop, which leads to thermal stress in the surface layer. The depth of the crack induced by given thermal/cooling cycles is calculated through the Paris law. The amount of transverse cracking is predicted through an empirical equation as a crack depth function in the Pavement ME Design. Previously, the transverse cracking calibration coefficients were only dependent on the hierarchical level of input data. In the new version released in Jun 2018, the transverse cracking calibration coefficients are functions of the Mean Annual Average Temperature (MAAT) and calibrated through the iterative approach. The resulting coefficients for the MAAT less and more than 57°F are presented in Table 3-3.

3.2.4. IRI Model

In Pavement ME design, the IRI model is a function of all distresses and is calculated by an empirical equation calibrated by LTPP sections. To calibrate the IRI model, the causes of poor goodness of fit and bias of Pavement ME nationally calibrated models were investigated. The IRI model's local calibration coefficients were modified as needed based on the information derived from distress models after the calibration. For this model, the site factors for each section were calculated. Using Microsoft Excel Solver, the coefficients were calibrated, and the best combination was obtained through this process.

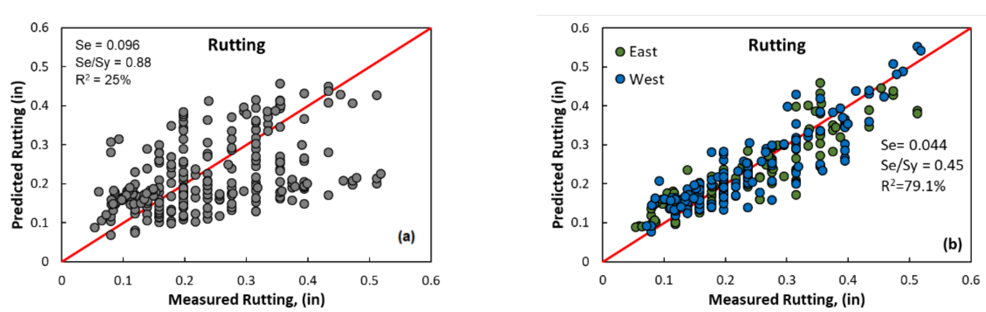
Table 3-3- Summary of calibration coefficients for Oklahoma flexible pavement design

Model		Parameter		East Region	West Region
Rutting	AC layer	β_1		0.79	0.21
		β_2		0.53	0.74
		β_3		1.48	1.03
	Granular base	βgb		0.15	0.23
	Subgrade	βsq		1.29	1.03
Fatigue Damage		$k1$		4.2	3.56
		$k2$		3.62	4.18
		$k3$		1.4	2.2
Top-down Cracking		C_1		6.6	6.1
		C_2		4.5	4.23
		C_4		723	723
Bottom-up Cracking		C_1		3.26	4.12
		C_2	h_{AC} <5 in	2.16	2.16
			5 in< h_{AC} <14 in	0.867+0.2583* h_{AC}	0.867+0.2583* h_{AC}
			h_{AC} >14 in	4.5	4.5
		C_3		6000	6000
Thermal Cracking		K	MAAT<=57 °F	$3 * 10^{-7} * MAAT^{4.0319} - 54$	$3 * 10^{-7} * MAAT^{4.0319} - 23$

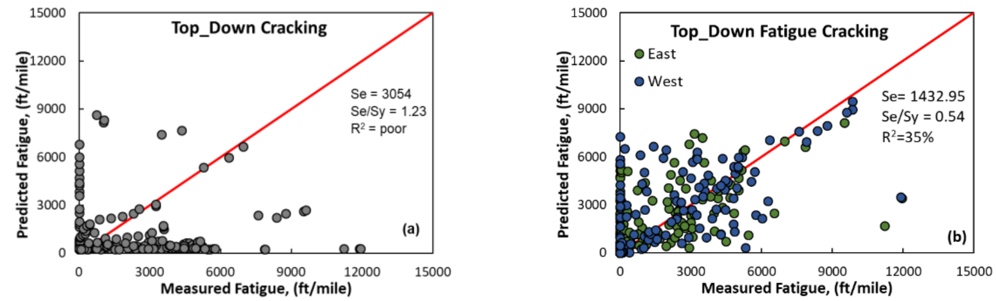
		MAAT>57 °F	$0.13 * MAAT^2 - 11.68 * MAAT + 100$	$0.13 * MAAT^2 - 11.68 * MAAT + 78$
IRI	C_1		5.23	6.46
	C_2		0.127	0.187
	C_3		0.013	0.0098
	C_4		0.0128	0.023

3.3. Discussion of Calibration Analysis Results

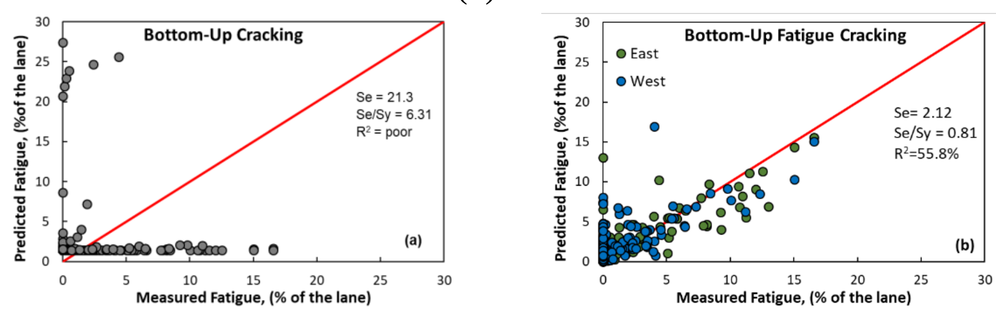
The nationally calibrated models show an improper performance and a significant bias before the local calibration process. Figure 3-5 summarizes the comparison between the predicted and measured values of distress and IRI performance for Oklahoma's flexible pavements before and after the local calibration process. Table 3-4 presents the Pavement ME performance model evaluation before and after the local calibration. The parameters that have been used for the statistical assessment of performance models include the coefficient of determination, R^2, the standard error of the estimate, *Se*, and the ratio of the standard error of estimate to the standard deviation of measured performance, *Se*/*Sy*. The analysis results show improvement in all distress and IRI models' statistical measurements after the calibration process. Generally, rutting and IRI models have a lower error and higher accuracy than fatigue and transverse cracking models.



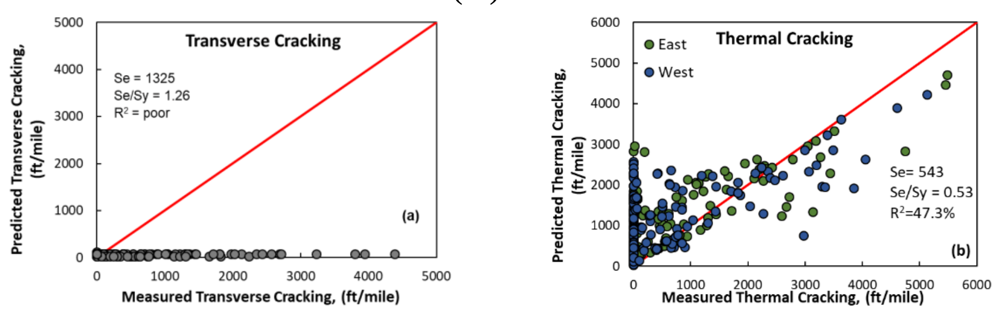
(i)



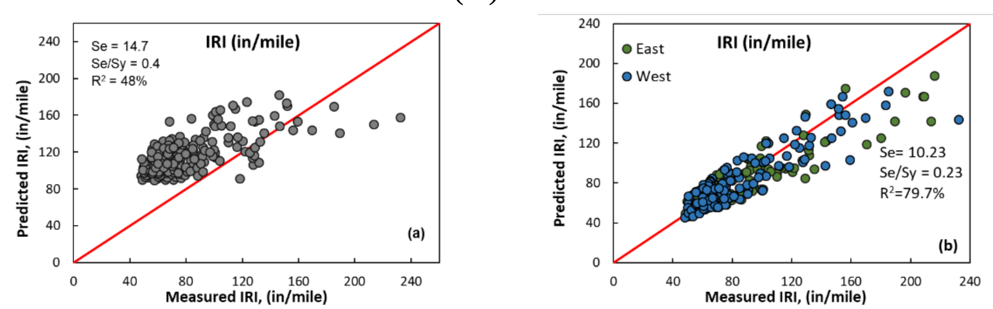
(ii)



(iii)



(iv)



(v)

(a) Before Calibration (b) After Calibration

Figure 3-5 -Comparisons between measured and predicted (i) rutting, (ii) top-down cracking, (iii) bottom-up cracking, (iv) thermal cracking; and (v) IRI models (a) before and (b) after the local calibration process

The R^2 value for the rutting model increased by 31% after the calibration, and *Se* decreased by 0.02 in. The *Se*/*Sy* term indicates the variation of error compared to the variation of measured data decreased by 0.2. Fatigue Cracking models generally underestimate the observed fatigue distresses and shows poor performance before the calibration process. After the calibration, the R^2 of bottom-up fatigue cracking increased to 36%, and the standard error decreased to 3.04% of the total lane area. Also, the term *Se*/*Sy* decreased from 6.3 to 0.9. The top-down fatigue cracking standard error decreased after the calibration, but *Se*/*Sy* increased that affirms the local calibration's infectivity for this performance model.

Pavement ME thermal cracking results are not consistent with Oklahoma's measured values. Even after the calibration, the thermal-cracking model cannot accurately predict the cracking for flexible pavements and shows a significant error from the observed cracking in the field. The same performance was reported for the MEPDG transverse cracking model conducted by other local calibration efforts (Ceylan et al. 2015; Hall et al. 2011; Darter et al. 2014).

Table 3-4- Pavement ME performance model evaluations before and after the local clibration effort

Model	**Model Calibration Status**	**Number of Observation**	**R^2 (%)**	***Se***	***Se*/*Sy***
Rutting (in.)	National	278	25%	0.096	0.88
	local	278	79.1%	0.044	0.45
Top-down Cracking (ft./mile)	National	342	poor	3054.13	0.123
	local	342	35%	1432.95	0.54
Bottom-up Cracking (%)	National	322	poor	21.3	6.31
	local	322	55.8%	2.12	0.81
Transverse Cracking (ft./mile)	National	278	Poor	1325	1.26
	local	278	47.3%	453	0.53
IRI (in./mile)	National	366	48%	14.7	0.4
	local	366	79.7%	10.23	0.23

The IRI model in Pavement ME shows an acceptable performance in the prediction of roughness in flexible pavements. The local calibration increased the accuracy of prediction to 64% and decreased the standard error to 12.27 in/mile. However, there was a slight drop in Se/Sy from 0.4 to 0.31.

The selected pavement sections identified for calibration study cover various pavement structures, construction age, traffic, and environmental conditions. To better understand the effect of local calibration effort on different pavement designs and assess the performance models' accuracy, a sensitivity analysis was required. Among the numerous pavement design inputs, the critical parameters of AC thickness, pavement age, and traffic were selected, and the performance prediction models error before and after the calibration were investigated. Table 3-5 summarizes sensitivity analysis on the rutting, bottom-up fatigue cracking, and IRI models. This table shows the percentage of standard error (SE) reduction after calibration effort for various traffic load categories, AC thickness, and pavement age.

The rutting model's error before the calibration for flexible pavements with AC thickness less than 4 in. was significantly high, and the calibration effort reduced the error by almost 70%. However, the rutting model showed a good performance for thick pavements, and the calibration effort reduced the standard error by approximately 20%.

Pavements with the age of 5 to 15 years showed a very high variation in error and the calibration effort reduced both the average of error for all pavement ages. The average error in the rutting model increases by increasing the traffic volume before and after the calibration.

However, the calibration effort essentially decreased the error, particularly for the roads with a medium level of traffic volume.

Table 3-5 - Sensitivity analysis of the prediction error reduction after calibration effort for selected pavement performance models

Variable	Variable Range	Prediction Error Reduction		
		IRI	Rutting	Bottom-Up Fatigue Cracking
Traffic Volume	AADTT<300	75.8%	25.0%	40.7%
	300<AADTT<1200	56.1%	20.3%	7.6%
	AADTT>1200	36.8%	33.3%	11.2%
Pavement Age	Age<5	69.3%	8.3%	43.5%
	5<Age<10	74.7%	34.2%	-1.4%
	10 <Age <15	52.7%	49.1%	3.0%
	15<Age<20	75.6%	6.3%	-7.7%
AC Layer Thickness (D), (inch)	H<2	37.8%	17.8%	14.8%
	2<H<4	69.4%	57.6%	20.9%
	4<H<6	77.0%	19.6%	26.8%
	6<H<8	73.6%	19.0%	43.1%
	H>8	74.5%	23.7%	31.3%

*(Positive values indicate a reduction in error and negative indicate an increase in error)

The bottom-up fatigue cracking model variation decreased for the whole range of AC thickness after the calibration process. However, the error for thicknesses ranging from 4 to 6 inches was still high after the calibration. The fatigue cracking model's average error at the early age of flexible pavements was reduced up to 40% after the calibration. For pavements older than 15 years, the fatigue cracking model cannot properly predict the observed values of cracking, and the error average increased. The high error is mostly due to the overprediction of fatigue cracking at older ages compared to observed values on the field. The calibration effort generally decreased the average error for the fatigue cracking model in all traffic volume levels.

Before the calibration, the error of IRI predictions increased by increasing the AC thickness. The calibration effort resulted in the error reduction, particularly for the pavements with a thick AC layer. This can reduce the error in distress models, especially the fatigue cracking model, and adjust the IRI model's corresponding calibration coefficients. Besides, for all the pavement ages, the average of the IRI model prediction decreased. For the low traffic roads, the IRI model underestimates the road surface roughness using national calibration coefficients while the calibration effort reduces the average of IRI prediction error by 75%.

3.4. Development of Input-ME Analysis Tool

Pavement ME requires an extensive amount of input data. Availability and accuracy of input data is a serious concern of the designers. To facilitate using the Pavement ME, interface software was developed. This software provides Pavement ME input data based on the Oklahoma material, traffic, and climate properties. A database was provided from the Oklahoma pavement materials, traffic, and climate data. INput-ME reads the required data from the database and generates input data compatible with Pavement ME import files. This software is developed by using QT creator. Qt Creator provides a cross-platform, complete integrated development environment (IDE) based on the C++ programming language. It is a tool for application developers to create applications for multiple desktops, embedded, and mobile device platforms.

Figure 3-6 shows the main window of INput-ME. Input ME includes three modules; (a) material, (b) traffic, and (c) calibration coefficients. Each module can process and compile specified data for the corresponding Pavement ME inputs to design a pavement in Oklahoma. Each dialog provides the information for three different levels of hierarchy input. At level 1, the user inserts the required measured values of properties for each parameter. At level 2, the Oklahoma-specified data collected from different sources will be offered to the user. The user can edit the provided data and insert its values for each parameter. At level 3, the software provides the data and locks the related cell while the user can review and export it as an input file for Pavement ME.

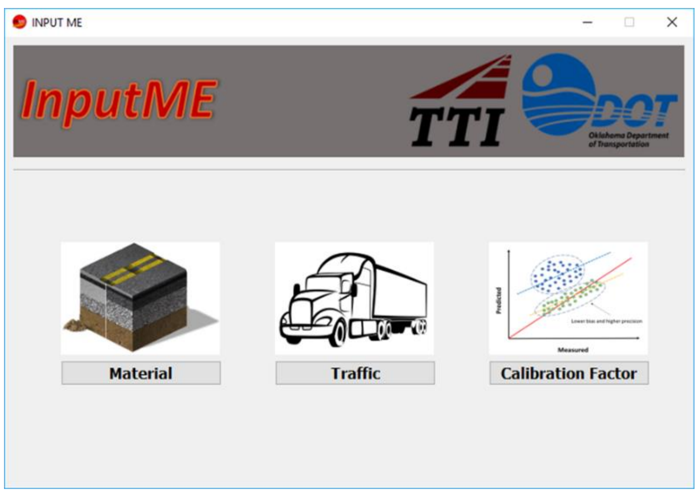


Figure 3-6 - The main window of INput-ME

The material module classifies the materials according to Pavement ME Design. This module enables the user to select materials for the subgrade soil, non-stabilized granular material, bound treated base material, asphalt concrete, and cement concrete materials.

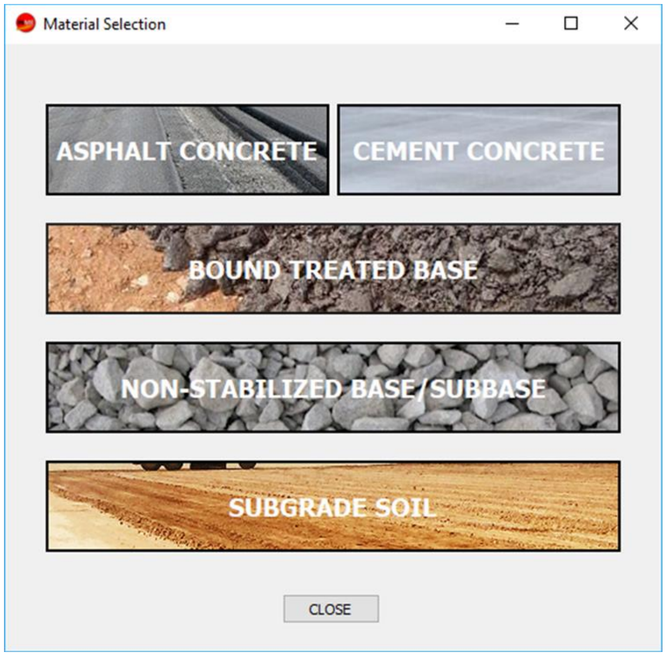


Figure 3-7- The material dialog

The subgrade soil dialog consists of two tabs, which indicate the sources of the provided information. Subgrade soil information was obtained through the LTPP database and a national database of subgrade soil developed under the NCHRP 9-23 A project. This database contains the soil properties for subgrade materials needed as input in Pavement ME. The database contains the parameters describing the soil-water characteristic curves (SWCC), which are the critical parameters in the implementation of MEPDG Level 1 environmental analysis. It also includes the measured soil index properties needed by EICM in all three hierarchical pavement design levels.

Figure 3-8 shows the binder properties and mix gradation in asphalt concrete dialog. This dialog provides information for the asphalt concrete material in the three levels of hierarchy. At level 1, the user inserts the essential measured values of properties for each parameter. The mix designs used in the database include the common mixes used in the state of Oklahoma. The provided Level 2 data are gathered from the dynamic modulus and the creep compliance test results reported in the previous ODOT research studies (Sakhaeifar, Newcomb, et al. 2015; Cross et al. 2011; Hossain et al. 2011). There is a wide range for the sieve analysis and gradation for each asphalt concrete mix in the ODOT research. However, the most used gradation was provided as Level 2 in the Input-ME database. The binder properties include PG 64-22, PG 70-22, and PG 76-22, which are the common binder types used in the asphalt concrete of Oklahoma's state.

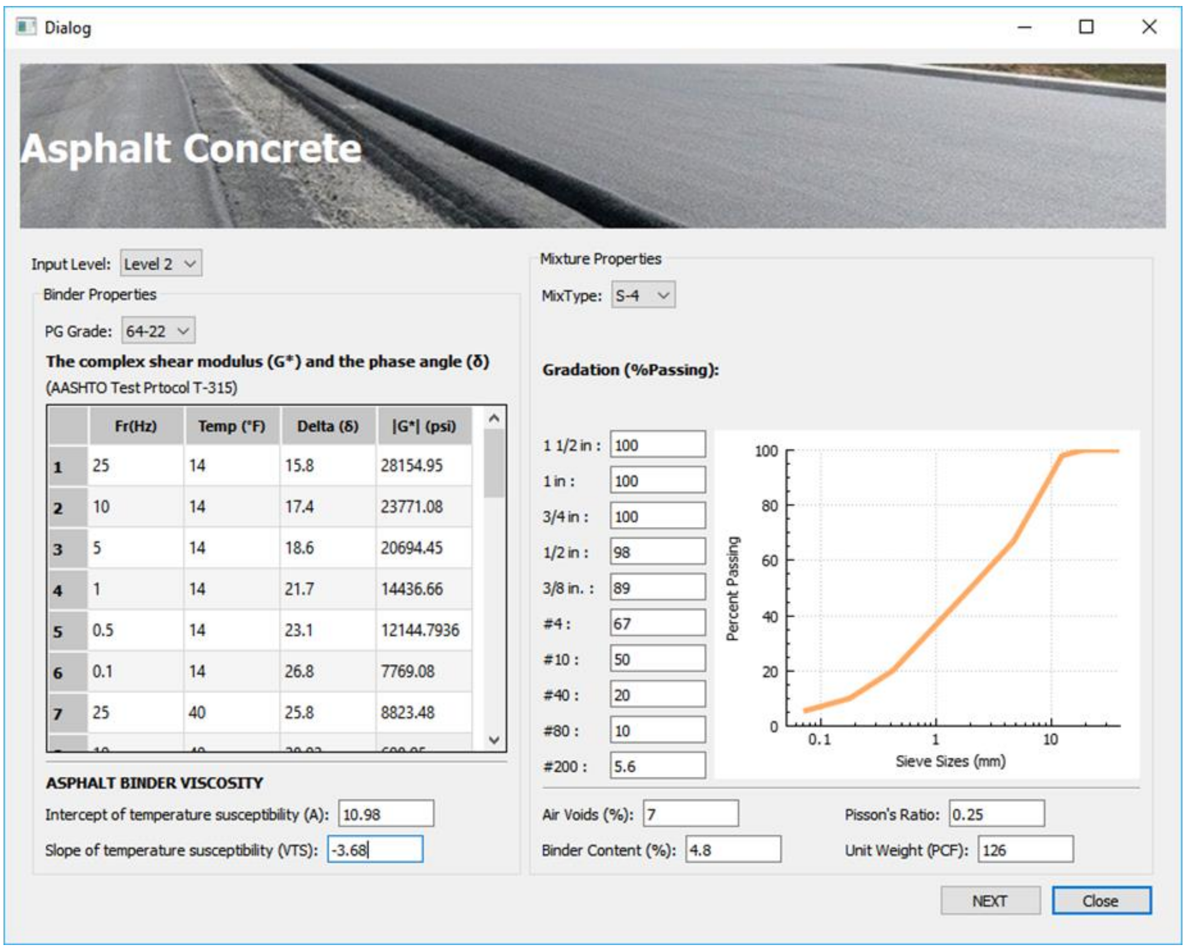


Figure 3-8- Binder properties and mix design gradation in asphalt concrete

In the traffic module, the information including AADTT, vehicle class distribution, axle load distribution, growth rate distribution, monthly adjustments, and axle per truck distribution based on the road class in the state of Oklahoma are processed and compiled. Figure 3-9 shows the general traffic data and vehicle distribution for the selected road type. These data are common, typical values, and it can be edited by the user if needed. Figure 3-10 shows the axle per truck data tab in the traffic dialog.

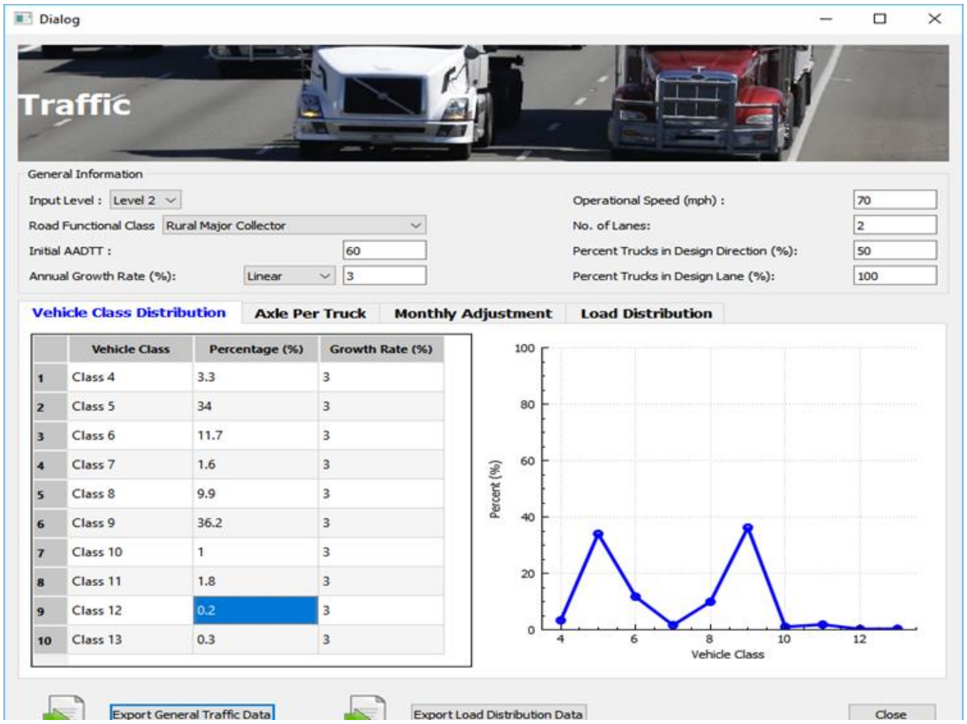


Figure 3-9- The general traffic and vehicle class distribution tab in the traffic module

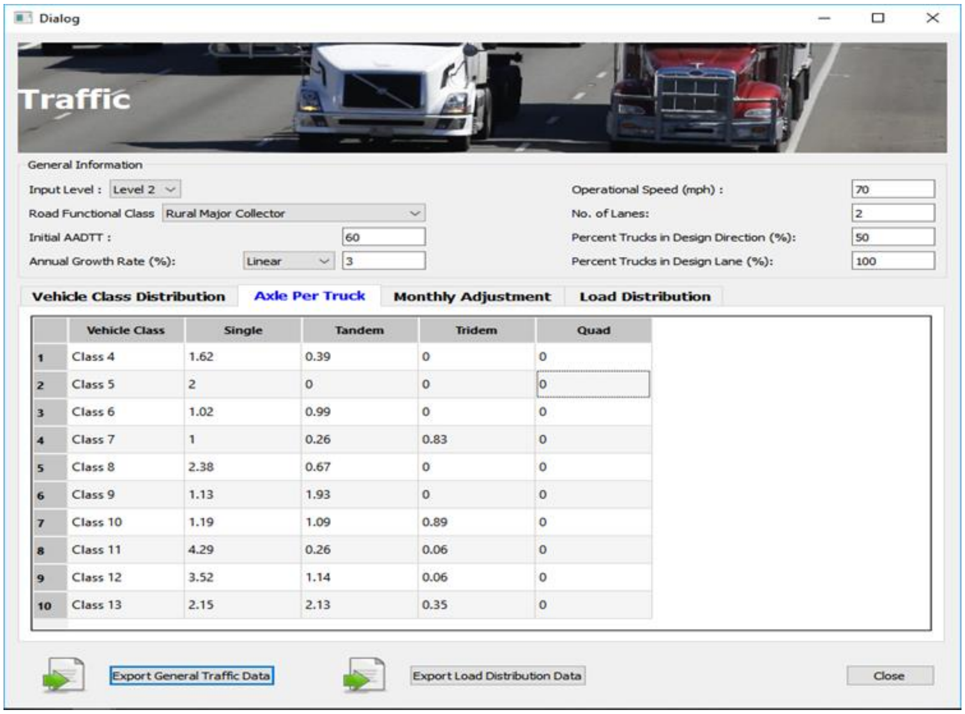


Figure 3-10- The axle per truck data tab in the traffic module

Figure 3-11 shows the calibration coefficients of the rutting model for East Oklahoma flexible pavements. The calibration coefficients for the flexible and rigid pavements will be generated in an XML file format readable by the Pavement ME by clicking on the Export Data button. This module provides a summary of the results found in this project. The calibration coefficients were provided in this module for flexible pavement designs. This calibration effort was determined for two sets of calibration coefficients for the east and west regions of Oklahoma's state.

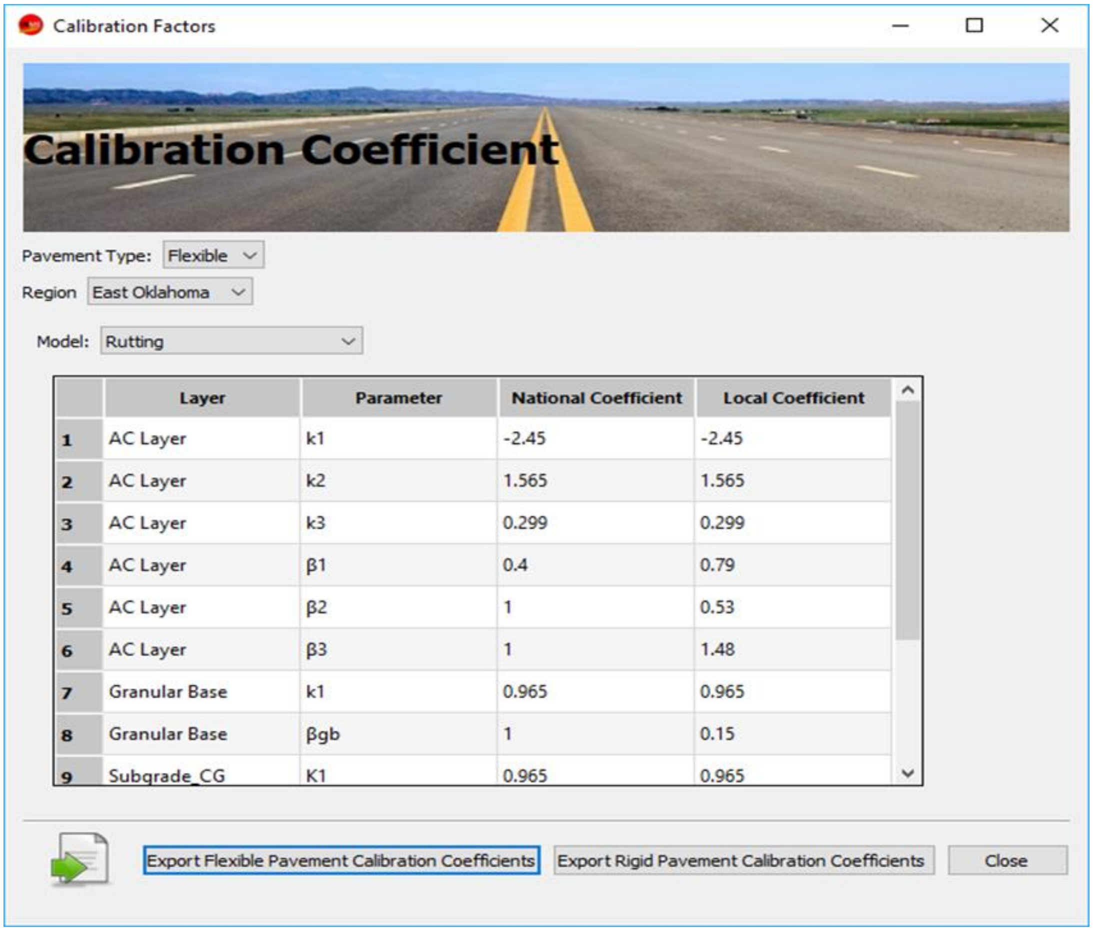


Figure 3-11- The calibration coefficient module

4. OVERVIEW OF THE MACHINE LEARNING METHODS

The intrinsic complexity of the pavement properties and parameters that affect the pavement behavior during the pavement lifetime asserts the need to try machine learning prediction models and compare each model's performance in predicting the IRI. Each machine learning algorithm has its algorithm for learning and predicting the data. Studying the mathematics behind the machine learning algorithms helps to get the best performance out of each model. Each model's performance and accuracy vary based on several factors related to the dataset, including the type of the features, distribution, and correlation between variables, amount and severity of the outliers, number of available observations, etc. The tuning parameters, the type of loss and kernel functions, and validation techniques need to be studied and investigated before applying the ML technique. In the following, selected ML techniques used in predicting the IRI will be explained.

4.1. Generalized Linear Model (GLM)

A generalized linear model is a general form of linear regression. The expected value of Y is not directly estimated from the linear combination of explanatory features X. In linear regression, Y has a normal error distribution, which is not a good fit in many regression analyses where the response does not show a constant change by the changes in the predictor variables (Nelder et al. 1972). For example, in the pavement performance prediction, the difference in performance is lower at the early age of the pavement life; however, increasing the age increases the rate of pavement performance deterioration. Thus, the linear and polynomial models are not a good fit for pavement performance prediction. The GLM relates the linear model to the response variable via a link such as a log, logit, inverse functions, while the variance of the expected value is not constant.

One technique that improves the regression models is adding more useful information by improving overfitting problem in a model (Zou et al. 2005). This process is named regularization. In regularization, the regression coefficients' weights will be reduced by penalizing the sum of the coefficients' magnitude. In IRI prediction models, the input data consist of different variable types related to climate, pavement performance, pavement structure, and traffic. These variables may be correlated, and the regularization technique can help reduce the effect of variable correlation and automatically select the most informative ones.

LASSO and RIDGE regression are two forms of regularization techniques widely used in regression analysis (Fu 1998). LASSO regression adds a penalty on the number of non-zero regression coefficients using the L1 norm of the weight's vector. The RIDGE regression has a penalty on significant regression coefficients using the L2 norm of the weight's vector in the cost function. Usually, a combination of LASSO and RIDGE regression works effectively and gives the benefits of both L1 and L2 norm of the weights (Fu 1998). The GLM elastic net (Glmnet) is a form of GLM that incorporates the combination of LASSO and RIDGE regressions. The form of the loss function at the Glmnet model is shown in Equation (7).

$$Loss = \frac{1}{N}\sum_{i=1}^{n} w_i l(y_i, g^{-1}(\beta_0 + \beta^T x_i)) + \lambda\left[(1-\alpha)\left\|\beta^T\right\|_2^2 / 2 + \alpha\left\|\beta\right\|_1\right] (7)$$

Where β is the estimation parameters for explanatory variables, g is a link function chosen during the model tuning. The elastic-net penalty is controlled by α and λ. The parameter α balances the weights between two penalty forms, including LASSO with α=1 and RIFGE with α=0. The parameter λ determines the overall strength of the penalty. These two parameters are considered as tuning parameters and should be adjusted during training and evaluation steps. The GLMnet helps linear regression models better handle collinearity and correlation among the predictors and reduces the overfitting chance. It also helps the linear model be used more accurately in the results derived from the small sample sizes.

4.2. Support Vector Machine (SVM) Regression

Support vector machine regressor is a part of the support vector machine that is mainly used for classification, but the algorithm and the approach applied for support vector machines are also helpful in using it as a regressor (Alpaydın 2020). Researchers used this model to predict the pavement performances (Yan et al. 2011; Ke-zhen et al. 2011) and showed a robust performance, especially as the IRI prediction model (Ziari et al. 2016a; Kargah-Ostadi 2014).

The linear regression models work by minimizing the RMSE between the predicted and measured values and estimates the population means of dependent variable Y, which is sensitive to extreme observations and outliers. In support vector regression, the model can be estimated by minimizing the error using the new loss function known as ε-insensitive loss. The loss function used in the SVMR is robust to extreme observations and ignores any observation with small residuals. The observations that are not ignored by the SVMR is known as the support vectors. The Equation (8) shows the Loss function at the SVMR model that considers a penalty on large regression coefficients.

$$Loss = \frac{C}{N}\sum_{i=1}^{n} l(y_i, \beta_0 + \beta^T x_i) + \lambda \left\| \beta^T \right\|_2^2 \quad (8)$$

In this model, *C* is known as the cost parameter. As *C* increases, the models give more weight to the ε-insensitive loss and works more like an unpenalized method. On the other hand, as *C* decreases, the model removes the features with broad estimation parameters. Thus, the tuning parameters that must be tuned in the SVMR model are *ε*, which controls the number of close observations ignored by the model, and *C*, which balances the weight between the training error and the penalty term. Usually, there is a strong relationship between these two tuning parameters. It generally suffices to set one at some default value and adjust the other one (Cherkassky et al. 2004). Another term in SVMR is the kernel that transforms the features to make better predictions and increase the SVMR model's flexibility. Different SVM algorithms implement different kernel functions such as linear, nonlinear, polynomial, radial basis function (RBF), and sigmoid. The best kernel function can be selected in the model evaluation using a cross-validation technique (Alpaydın 2020).

4.3. Multivariate Adaptive Regression Splines (MARS)

Multivariate adaptive regression splines (MARS) form a linear regression model that automatically incorporates the nonlinearity and interactions between the features (Friedman et al. 2001). MARS acts as a polynomial model that can catch both the degree of nonlinearity and the features interaction simultaneously. Attoh-Okine et al. (2003) used MARS as a new model for the prediction of flexible pavement roughness. Since the MARS model is a useful tool for fitting nonlinear multivariate functions, the generated model was fitted perfectly to the IRI data. The model gave a roughness equation based on the available input information and evaluated each variable's contribution to the roughness equation. Attoh-Okine et al. (2009) applied the MARS regression technique for the doweled pavement performance modeling. It is reported that the MARS model has a better performance compared to traditional regression analysis and is more robust to the anomaly and outlier in the input variables.

This model breaks the range of independent variables, X into n number of bins, and fits the best line for each bin. Figure 4-1 shows the schematic of the MARS model fitted to nonlinear data. In this example, the MARS model consists of two knots that break the independent variable into three bins. A linear model is fitted to each bin, which reduces the amount of total error compared to a linear regression model. The fitted model to each bin could be linear or multivariate nonlinear. By this technique, the MARS change the continuous variables into clusters that optimally can be estimated by a nonlinear function.

The Equation (9) can represent the MARS model.

$$y = c_0 + \sum_{i=1}^{N} c_i \prod_{j=1}^{Ki} b_{ji}(x_{v(j,i)}) \quad (9)$$

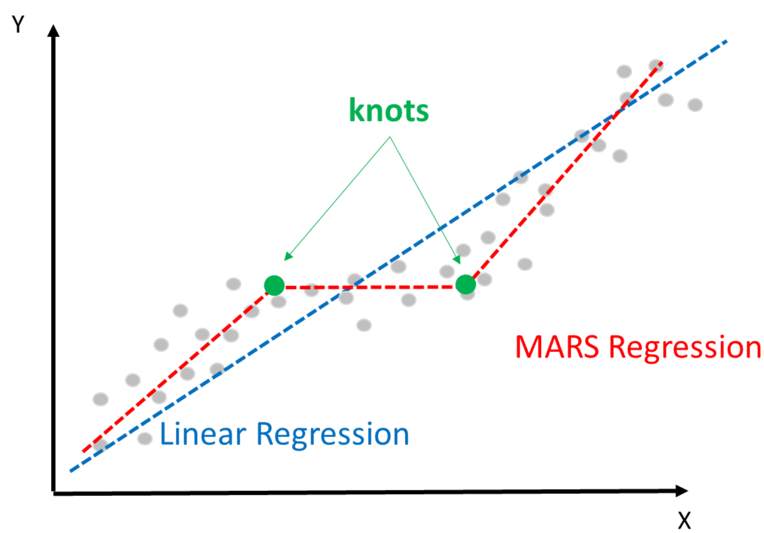


Figure 4-1- Schematic of MARS vs. linear regression fitted models

Where y is the dependent variable, c_i is a vector of regression coefficients, $b_{ij(x_{v(j,i)})}$ is the truncated basis functions, $v(i,j)$ is the index of the independent variable used in the i^{th} term of the j^{th} product, and Ki is the parameter that controls the degree of the freedom of the model. The number of bins can be tuned during model training by using the knots parameters, which indicate the number of breaks in the model. Besides, the degree of freedom is another tuning parameter that controls the nonlinearity and interaction between the features.

The MARS regression model works well with many predictor variables and automatically detects the interactions between variables. Despite its complexity, it has a fast and efficient algorithm and is robust to outliers. The limitations of the MARS model include the disability of handling the missing data and susceptibility to overfitting. Also, it is more difficult to understand and interpret than other analytical models.

4.4. Neural Network Model

Artificial Neural networks (ANN) are one of the well-known machine learning techniques used by researchers as a prediction model(Kargah-Ostadi 2014; Liu et al. 2014; Ziari et al. 2016b; Hossain et al. 2019; Abdelaziz et al. 2020). ANN includes simple elements named neurons that contribute to a mathematical process that contains the interaction of the features and transfers the results through transfer functions to increase the prediction accuracy. A linear combination of features is passed through a nonlinear transformation in successive layers. Fitting a curve under an external neural network containing a single neural layer can be represented in Equation (10).

$$loss = \frac{1}{N}\sum_{i=1}^{n} l(y_i, \beta_0 + \sum_{k=1}^{H} \beta_k h_k(X)) \quad (10)$$

The $h(X)$ is a nonlinear transformation of the linear combination of the features. H is the number of hidden neurons in the first layer, and βi is the estimation coefficients or the network's weights vector. By adding a new layer to the model, the same form will be followed, in which the output of the neurons of the previous layer is the input features for the next layer. By adding more layers, the features' interaction level will increase, which usually increases the accuracy of the prediction. On the other hand, increasing the neural network model's complexity results in overfitting and increasing the variance term of the error. The parameters which should be optimized for the NN model include the number of layers, the number of nodes in each layer, the learning rate, and the activation function type. The learning rate controls the gradient descent algorithm's step size and should be selected through model evaluation. Lower learning rates require more training iterations, and higher learning rates lead to divergence of the model. The NN common activation functions are linear, sigmoid, and rectified linear unit (ReLU) functions.

The ANN models are black boxes, meaning that it does not give information on each parameter's contribution in predicting the dependent variable. Also, due to the complexity of the model, it is computationally expensive and time-consuming. The ANN model highly depends on the training data, and the chance of the overfitting is higher than the other ML models; However, by increasing the observations, the ANN model's accuracy significantly increases (Alpaydın 2020).

4.5. XGboost Tree

Extreme gradient boosting (XGBoost) is a form of tree-based machine learning algorithms recently developed by (Chen et al. 2016). XGboost has extensively been used as a supervised learning method for structured and tabular data (Nwanganga et al. 2020). Researchers have used the XGboost model for enhancing the distress models of the pavement ME (Gong et al. 2019), analysis of asphalt overlay performance (Zhang et al. 2020), and aggregate shape classification (Pei et al. 2020).

The inherent structure of the pavement prediction features is a good fit for XGBoost tree regression. The sub algorithms implemented in the XGBoost tree are decision tree regression, bagging, and gradient boosting. The decision tree algorithm model learns how to assign the values through a tree decision process of the selected features. For example, based on the given observations, the model learns the closest clusters of the observations with almost similar features and assigns an average value to the pavement performance of any observation fitted in that cluster. The bagging and bootstrapping algorithm help increase the tree model's accuracy by repeating the tree decision process and averaging the outcomes. Random forest is a particular case of a decision tree with a bagging algorithm in which a random subset of features will always be used in the decision process. It decreases the computation time substantially and increases accuracy. The boosting algorithm helps the model by giving feedback to the model after each iteration of the decision. The feedback can be defined as the magnitude of the error for the assigned value to the observation, and by decent gradient algorithm, the features 'contribution to the error will be identified. Figure 4-2 shows a summary of the XGBoost tree sub algorithms and the XGBoost model's evolution from decision tree models.

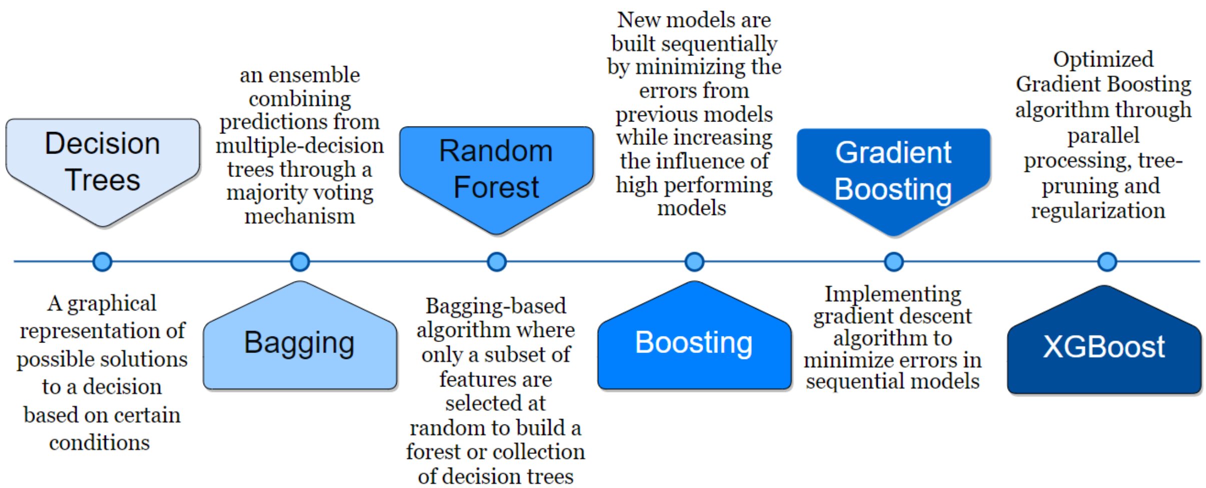


Figure 4-2- Evolution of XGBoost Algorithm from Decision Trees. Reprinted from (Kumar 2019)

The XGBoost tree model ensemble tree-based method works well with the tabular data and is one of the best models for pavement performance prediction. Compiling a comprehensive pavement database is always challenging due to the complexity of datasets and missing information. A subset of the data in some of the sections can be missed due to the differences in the strategies adopted by highway agencies in the data collection process and extension of the pavement network. For example, FWD data can be missed in some pavement segments for a few sections. The XGBoost model can handle the missing values in model since it does not necessarily need all the independent variables in the decision process. Thus, the number of observations increase and none of the sections will be discarded due to the lack of information on some features. Other benefits of the XGBoost tree model are parallelization, regularization, and cross-validation, leading to increase the computational speed and avoid overfitting in the model.

The XGBoost model parameters that should be optimized are the maximum depth of the tree models, the total number of trees to grow in each cycle, the minimum number of variables used in the tree model, and the minimum number of samples leaf of the tree and few others.

4.6. Markov Decision Process (MDP)

Markov Decision Process is a framework that were used for solving reinforced learning problems. The Markov decision process (MPD) models have been introduced as a vital tool for optimizing decisions in decision-making problems (Bellman 1957). Butt et al. (1994) applied a stochastic MDP to modeling the infrastructure deterioration. The developed model can dynamically determine the optimized M&R plan at the network level. However, the pavement performance prediction and the model has limited pavement type and section. This causes the lack of capturing the uncertainty in the M&R plan. Recent research on M&R plan models using the MDP shows promising results and asserts the feasibility and flexibility of MDP models. Narh-Dometey (2019) developed an M&R policy method by applying the MDP model using the pavement roughness and road class. This policy can balance the user and agency costs and optimize the maintenance plan of the pavement. Markov models were used to assess the rehabilitated pavement's structural potential and the cost-effectiveness of the rehabilitation policy (Osorio-Lird et al. 2018).

Reinforcement learning is one of three basic machine learning paradigms, alongside supervised and unsupervised learning processes. Reinforcement learning (RL) is an area concerned with how a decision-maker, referred to as an agent, must take actions in a different state of an environment to maximize some notion of cumulative reward (Sutton et al. 2018). Figure 4-3 shows the learning cycle in reinforced learning.

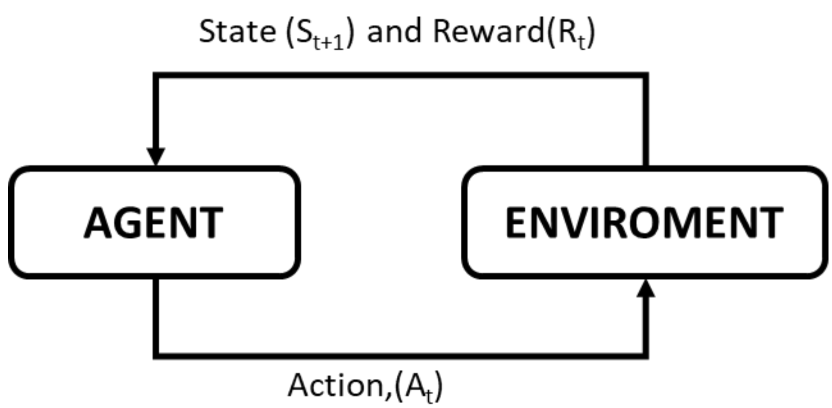


Figure 4-3- Typical Reinforcement Learning Cycle. Adopted from (Sutton et al. 2018)

The environment is the setting that compromises any possible state that the agent can get and interacts with other states by taking possible actions. The state is a specific snapshot of the environment at any time-step. Taking action in any state causes the agent to move to a new state or stay in the same state. Any action at any state can give a reward (negative or positive) to the agent, and the reward accumulates after taking a series of actions. Moving from one state to another state is named state transition, which can work in a deterministic or stochastic environment. The transition probability is the probability of reaching state St+1 by taking action At in the state St.

A Markova Decision Process (MDP) model is a reinforce learning model that contains a set of possible states (S), a set of possible actions (A), a reward function R(s, a), and a transition function T(s, a). An MDP has Markovian property in which the transition from state S(t) to S(t+1) is entirely independent of the past, and the S(t) itself can have all the necessary information for the decision.

By having the MDP model, the agent can explore the environment by taking actions and visiting different states, and accumulating rewards. After exploring the environment, each state can get a value that determines the goodness of being in that state and is a function of expected total reward gained from that state. MDP model aims to find the value function and the optimum policy. An optimum policy is an instruction for the agent that determines the best action at each state, which leads to the maximum reward.

Several algorithms have been developed for solving the MDP and finding the optimum policy. Q-learning is a model-free reinforcement learning algorithm that was extensively used to solve reinforcement learning problems that do not require an explicit definition of a Markov decision process. It trains an agent to learn an optimal policy from a dynamic environment, and the optimal learned policy tells the agent what action to take at each state. Given an environment, let S denote the state space, and A means the action space. In Q-learning, the expected total long-term reward in a given state s, and an action a is predicted by the Q-function *Q(s, a).* The agent should take the optimal action (s) for a state s such that the expected long-term reward is maximized. The optimum policy can be derived from an Equation(11).

$$\pi(s) = \arg\max_a Q(s, a) \quad (11)$$

where π denotes the optimal policy.

Q-learning applies a modified form of a Bellman equation to learn the optimal policy. If we only consider 1-step transition ⟨st, at, rt, st+1⟩ where r is the immediate reward of the current step, and st+1 is the next state, then we have:

$$Q(s_t, a_t) = r_t + \gamma \max Q(s_{t+1}, a) \quad (12)$$

where $\gamma \in [0,1]$ is the discount factor specifying how far ahead in time the algorithm should look. This means the optimal long-term reward (st, at) for the current state st and action at , is the immediate reward plus a discounted optimal long-term reward for the next state.

Q-Learning algorithm is difficult to deal with large and continuous state space. There are methods to deal with this problem, such as discretization techniques, but the implementation is hard, and the performance could be decreased. Another method is to use a neural network to solve the Q-function in the Q-learning.

4.7. Deep Q-Learning

The Deep Q-learning is an extension of the Q-Learning algorithm by modeling the Q-function (s, a) as a (deep) neural network. In this approach, the Q-function is a complicated composition of various parameterized functions, taking the input s, a, and making predictions of long-term utility. A loss function L is designed to measure how well Q makes the prediction. Finally, the parameters of Q are trained by calculating gradients of L and applying optimization. The objective of Q-learning is to make the iterative process converge, and one choice of the loss function is the "squared difference."

$$L = \left[Q(s_t, a_t) - \left(r_t + \gamma \max_a Q(s_{t+1}, a) \right) \right]^2 \quad (13)$$

Where max (s_{t+1}, a) is evaluated using the current NN prediction model, and $r_t + \gamma$ max $Q(s_{t+1}, a)$ is the target value in the regression model. Q-learning needs to consider the exploration-exploitation tradeoff like other reinforcement learning algorithms. First, It is needed to explore the environment to get a complete picture of transition and outcomes, which means some random action will be chosen rather than the optimal action; on the other hand, the current optimal action should always be executed to train the agent effectively. In this condition, a mixed strategy can be the solution by defining a trade-off parameter $\epsilon \in [0,1]$. The agent has probability ϵ to randomly choose an action for the loss, rather than optimal action. Also, ϵ dynamically will be changed over time.

•When training beings, the Q-function is not trained enough to make a good prediction. In this case, going for the optimal action is not useful. Therefore, larger ϵ at the early stages are needed.

•As the training goes on, the Q-function gains more predictive power, and trust in the predicted utilities gradually increases. Therefore, ϵ should decrease over time.

To achieve the above heuristics, three parameters ϵstart, ϵmin, and decay rate were defined. ϵ_t at t[he] state transition decreased by a factor of ϵ decay after every state transition be calculated by the following:

$$\varepsilon_t = \max(\varepsilon_{\min}, \varepsilon_{decay}^{t-1}\varepsilon_{start}) \quad (14)$$

In deep, Q-Learning is implemented by the optimizer, and the learning rate α is the learning rate of the optimizer. This value is a hyperparameter and will be tuned for each problem.

5. AN INTELLIGENT M&R DECISION-MAKING METHOD FOR FLEXIBLE PAVEMENTS

Pavement performance modelling is an essential step in pavement design and management from project to network levels. The pavement performance should be monitored regularly, and maintenance and rehabilitation (M&R) treatments should be planned to keep the pavement in good condition and maximize the service life and returned benefits of the constructed pavement (Smith et al. 1993). A proper pavement management system should provide a comprehensive, accurate perspective of the entire network to the manager to analyze the required amount of funds for extending the pavement network and keeping the current network in the best possible condition. Determining the most effective treatment type and the most appropriate treatment schedule needs detailed information about the pavement structure and material properties, traffic loads, subgrade conditions, and climate information. A comprehensive data collection for a network of pavements is costly and unnecessary since the managers should balance the study and experimental phase cost and the benefit that will be obtained (Mbwana et al. 1996). Thus, the network level's objective is to outline the treatment plan for the network by the cheapest and most informative pavement information and determine the impact of treatment scenarios on the funding and overall condition of pavements.

In this chapter, a new M&R decision process for flexible pavements was developed. The decision-making process considers the stochasticity of the pavement performance prediction and suggests the optimized M&R activities for the given section by implementing a newly developed predictive model. The experimental sections were selected from Oklahoma and few other states, and the required data were extracted from the long-term pavement performance (LTPP) database.

The M&R decision method is a Markov Decision Process (MDP) which employs IRI from a newly developed IRI prediction model and structural number from historical data. The IRI prediction model predicts the IRI with high accuracy by having the structural number (SN), road class, climate condition, traffic load, and subgrade and structural information. To get a high accuracy for the IRI prediction, several advanced machine learning models were developed, and the best model with the highest accuracy was implemented in the MDP model. This M&R decision process incorporates the most informative available information from the surface and underneath layers of the pavement. The model is compositionally cheap, and the application of the model can be extended to North America's flexible pavement networks.

The current methods of M&R planning are mainly using the performance metrics driven from a selected number of sections, including PCI or International IRI. The mentioned metrics are a good indicator of the current pavement condition and can be used to categorize a network into different subgroups based on the severity of the surface distresses. Using a structural number together with IRI in M&R decisions increases the reliability of the decisions and lead to better budget allocation management. In this research chapter, an MDP model was generated for optimizing the M&R policy selection using IRI and effective SN of the flexible pavements. The MDP model gives a pavement management planning based on the pavement structural integrity and pavement roughness. The results of this planning can be used to recommend the optimal M&R treatment, which increases the benefit over the cost of the pavement. This model dynamically predicts the pavement's roughness during the desired time frame using the developed IRI prediction model.

Figure 4-1 presents a summary of the proposed methodology for developing the MDP M&R planning model for the flexible pavement using the data collected from the LTPP database. The experimental sections for this research chaper were extracted from the long-term pavement performance (LTPP) database (InfoPave 1988), NCHRP 9-23B soil database (Zapata et al. 2012), and MERRA climate database. The input data for the IRI prediction and MDP M&R decision-making models include pavement structural properties, pavement performance characteristics, subgrade soil properties, and climate and traffic properties. A comprehensive exploratory data analysis of the features is presented in section 5.1.

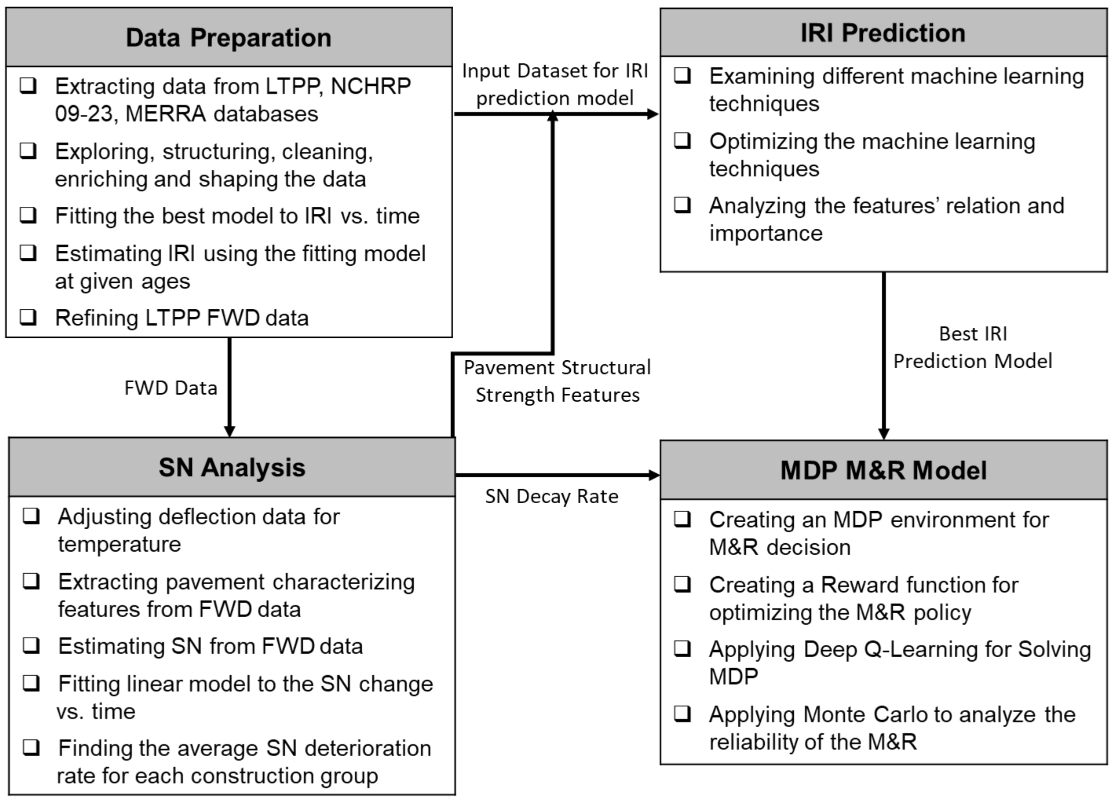


Figure 5-1- The proposed methodology for the MDP M&R decision model

In order to have a reliable M&R strategy, predicting the pavement IRI deterioration before and after applying any kind of treatment is essential. The IRI prediction model aims to predict the IRI change versus time given the structural properties of the pavement section, including structural features, FWD data, subgrade soil properties, climate, traffic information, road type, etc. In the following, the dataset used in this chapter will be presented and IRI prediction model, SN analysis, and MDP model will be explained.

5.1. Dataset and Exploratory Data Analysis (EDA)

In this section, the dataset used in for the M&R planning method was explained, and an exploratory data analysis was conducted to obtain detailed information about datasets, variables, and the relationship between different features.

The LTPP program was established to collect the pavement performance data as one of the major research areas of the Strategic Highway Research Program (SHRP). The LTPP program monitored and collected pavement performance data. The collected data include information on seven modules: Inventory, Maintenance, Monitoring (Deflection, Distress, and Profile), rehabilitation, materials testing, traffic, and climate. These in-service pavement sections are classified in the LTPP program as General Pavement Studies (GPS) and Specific Pavement Studies (SPS). The GPS program is a series of studies on selected in-service pavements structured to develop a comprehensive national pavement performance database. These studies are restricted to pavements that incorporate materials and designs representing good engineering practices and strategic importance. SPS's objective is to provide a more detailed and complete base of data to extend and refine the results obtained from the GPS studies.

5.1.1. Pavement Sections

The state of Oklahoma has 38 flexible pavement sections in the LTPP study. These sections are spanned across the state and are a good candidate for the state's road functional class, traffic load, and environmental characteristics. However, increasing the number of sections can help reduce errors in the prediction models. Thus, the LTPP sections from the neighbor states were added to the existing Oklahoma LTPP sections. A total number of 419 flexible pavement sections from Oklahoma, Texas, Arkansas, Missouri, Kansas, Colorado, and New Mexico were selected. Figure 5-2 shows the selected LTPP sections used for the IRI prediction and M&R decision planning. These sections have a good variety in layer thickness, road functionality, the load of traffic, climate, and experienced M&R.

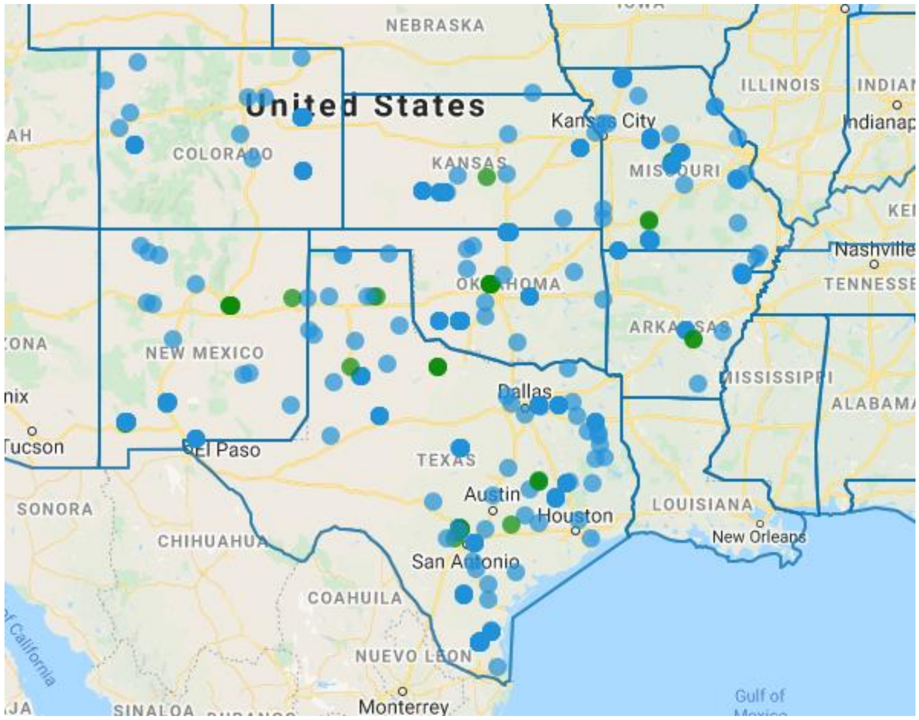


Figure 5-2- Location of the LTPP sections used in this research chapter. Blue Section: out of study sections. Green Sections: active sections. Reprinted from (InfoPave 1988; FHWA 1988)

5.1.2. Feature Statistics

The main factors affecting the roughness of flexible pavement can be categorized as material properties, structural properties, climate, environmental conditions, and loading volume. The pavement properties were collected and prepared for each section from various sources. Table 5-1 and Table 5-2 present the list of derived properties and the source of data used for developing the IRI prediction and MDP M&R decision making models. In Table 5-1, the numerical type features with their minimum, maximum, mean, and standard deviation values were presented. IRI is the dependent variable, and the rest of the features are the independent variables in IRI prediction models. The structural properties include the asphalt concrete and total thickness of the pavement, IRI, and FWD-related features collected from the LTPP database. The climate features include the average annual temperature, average annual precipitation, average annual freezing index, and average annual evaporation. The freezing index is the cumulative number of degree-days when air temperatures are below 32 degrees Fahrenheit. This index is a standard metric for determining the freezing severity during winter season, which triggers thermal distresses in the pavement. The average difference between precipitation and evaporation rate corresponds to the wetting/drying cycles of the subgrade. Subgrade of the pavements with a higher average difference between evaporation and precipitation has a higher settlement/swelling rate, which leads to weakening pavement structure and increasing pavement roughness (Byrne et al. 2009).

Table 5-1 - List of numerical pavement properties used in prediction and M&R models

Parameter	Description	Source	Min	Mean	Max	SD
IRI (in/mi)	International Roughness Index	LTPP	24.15	255.7	82.16	34.5
Age (Year)	Pavement age	LTPP	0.08	4.64	21.21	4.01
SN_{eff} (in)	Effective structural number	Equation (2)	1.2	6.1	19.6	2.8
SIP (micron)	Structural index of pavement	Equation (4)	8.4	156.64	1255.6	155.1
PD1.5 (microns)	Peak Deflection at 1.5 times D_{total}	LTPP FWD	0	81.78	420.29	52.23

Parameter	Description	Source	Min	Mean	Max	SD
D_{AC} (in)	Asphalt concrete thickness	LTPP	1.2	6.4	20.5	3.4
D_{total} (in)	Overall pavement's thicknesses	LTPP	6.4	20.9	54	8.2
AADTT	Annual Average Daily Truck Traffic	LTPP	56	629	2996	531
Temp (°F)	Annual average temperature	MERRA	44.8	62.36	77.5	7.1
FI (°F-Days)	Annual average Freezing Index	MERRA	0	72.88	657.0	103.84
Precip (in)	Water equivalent of precipitation over the year	MERRA	5.08	32.8	80.2	14.2
Evap (in)	Surface evaporation over year	MERRA	6.91	29.1	50.3	10.07

Figure 5-3 shows the correlation between numerical variables in the dataset. Data correlation is a tool to investigate the relationship between multiple variables and attributes in the dataset. If the dataset shows perfectly positive or negative attributes, the impact of the multicollinearity problem on the model's performance will increase. Multicollinearity is defined when one predictor variable in a regression model can be linearly predicted from the other predictors and can lead to skewed or misleading results (Farrar et al. 1967).

To help with the multicollinearity, regularization techniques can be used. These techniques were applied in machine learning models in this study. Also, decision trees and boosted tree algorithms, such as the XGBoost model, can perfectly handle the multicollinearity by choosing the best predictors in their decisions(Chen et al. 2016).

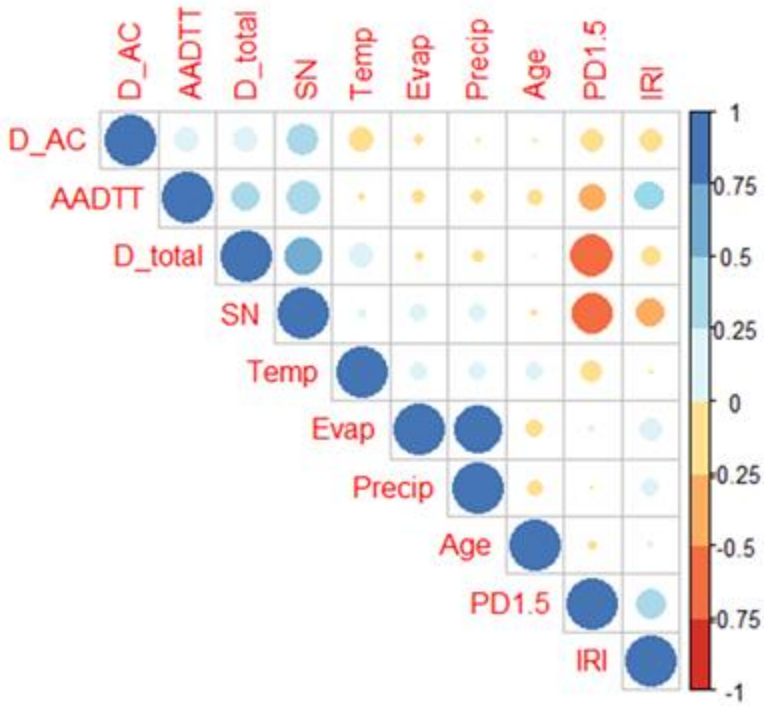


Figure 5-3 - Correlation between numerical variables

Pavement roughness shows a positive correlation with deflection magnitude, precipitation, evaporation, age, and AADTT. The pavement's average roughness decreases by an increase in asphalt and total thicknesses, structural number, and temperature. Figure 5-4 shows the relation between the structural number estimated from FWD data and pavement roughness values. Pavement roughness is higher for structurally weaker sections, and the presence of SN in predictors features can help boost the accuracy of the IRI prediction models.

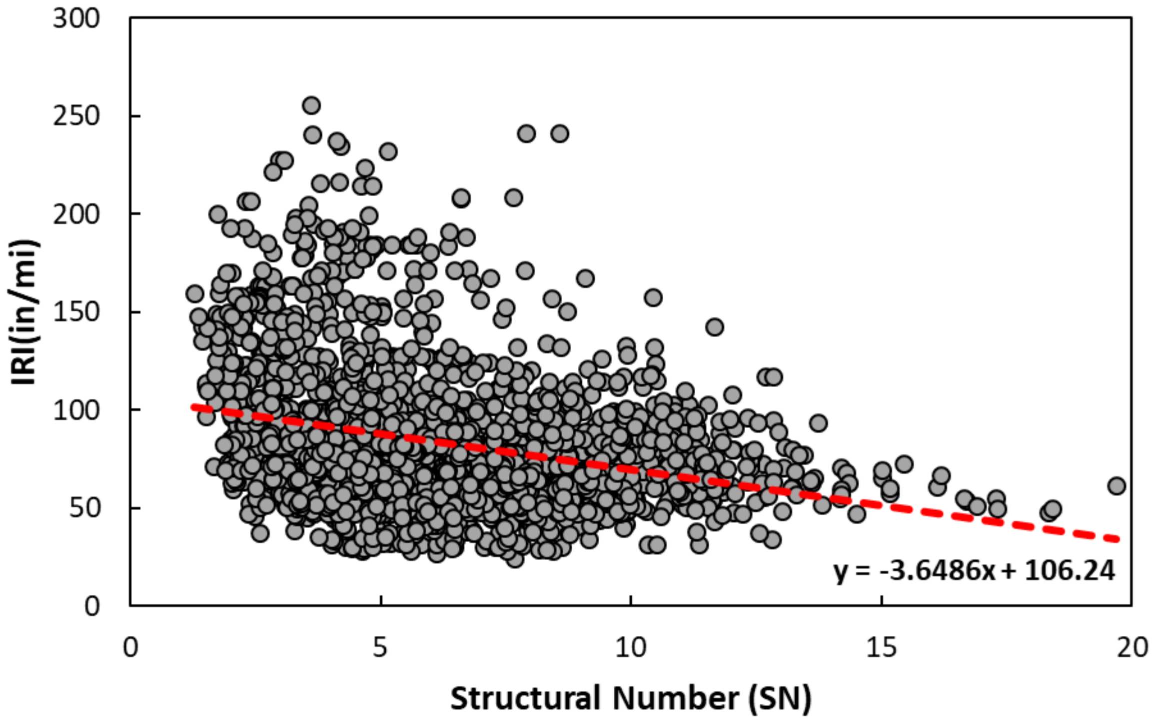


Figure 5-4- Sensitivity of the SN estimated from FWD data to pavement roughness

Table 5-2 presents categorical type properties, each category's levels, and the mean value of the IRI of the level. The impact of categorizing the observation into several levels in each variable has been tested through statistical mean difference tests. For two-level variables including Base_type, PL, and Road_Class, a standard two-sample t-test, and the M&R variable with three levels, the one-way ANAVOA test was used. The p-value of statistical tests for all categorical variables was less than %5, indicating a significant difference in mean IRI value between levels.

Table 5-2 - List of categorical pavement properties used in prediction and M&R models

Parameter	Description	Source	P-value	Level	N	Mean IRI
M&R	Past construction and M&R experiences	LTPP	2.2e-6	No M&R	1736	83.3
				Maintenance	393	99.22
				Rehabilitation	339	68.77

Base_type	Treated base/ Granular base	LTPP	4.17e-3	Granular Base	1467	87.9
				Treated base	1001	77.92
PL	Plasticity of the Subgrade soil	NCHRP 9-23B	5.82e-3	Non-Plastic	899	77.5
				Plastic	1569	87.5
Road_Class	Road classification	LTPP	5.881e-4	Rural	2211	82.24
				Urban	257	97.69

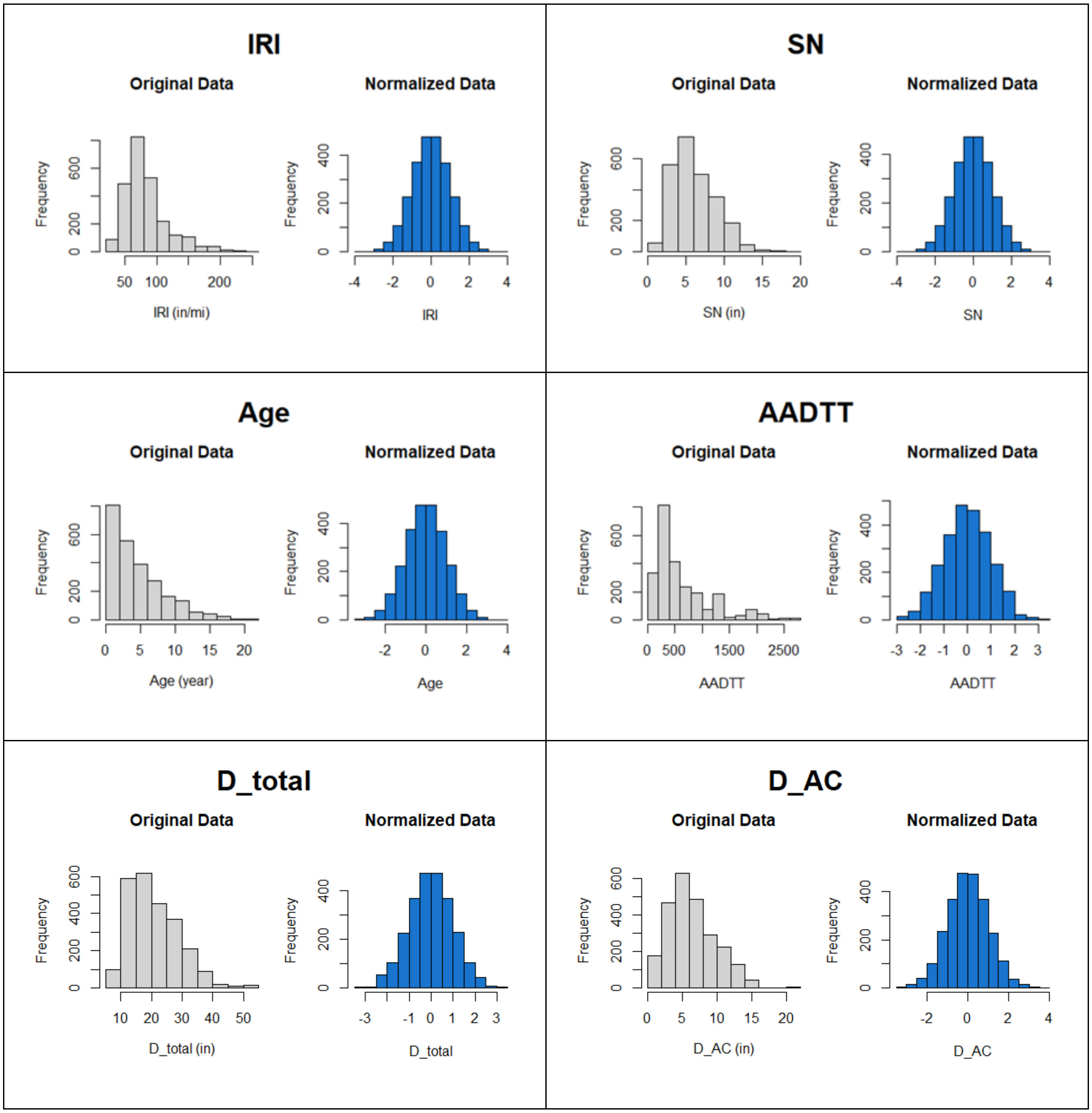
The average IRI for the pavement with no M&R is higher than pavements under maintenance. It shows that maintenances can reduce the increase in roughness but generally cannot reduce the roughness in the long term. Applying rehabilitation reduces the average roughness and helps to keep the pavement structure in better condition.

The average roughness for pavements with treated base is lower compared to pavements with granular bases. Due to the granular layer's low strength, the chance of failure of the pavement structure under cracking and rutting is higher. Treated bases are a better alternative to the conventional granular base and provide higher pavement structure strength. The effect of the subgrade's plasticity is in the swelling and settlement severity cycles during the wet/dry seasons. Plastic subgrade causes higher swelling severity and deteriorates the pavement structure.

5.1.3. Data Cleaning and Processing

One of the LTPP database problems is the inconsistency between the survey dates and FWD test dates. To handle this problem, the missed IRI value for each date were estimated from the fitted model explained in section 5.1.

The numerical variables mostly have non-normal distribution. Also, there is a big difference between the scale of the numerical variables. A transformation of the independent variables and having the same scale between variables will help achieve the model residuals' normality. Features with non-normal and skewed distributions were transformed and normalized to boost the performance of machine learning algorithms. Figure 5-5 shows some examples of the row data distribution and normalized distributions after applying Yeo Johnson's power transformation.



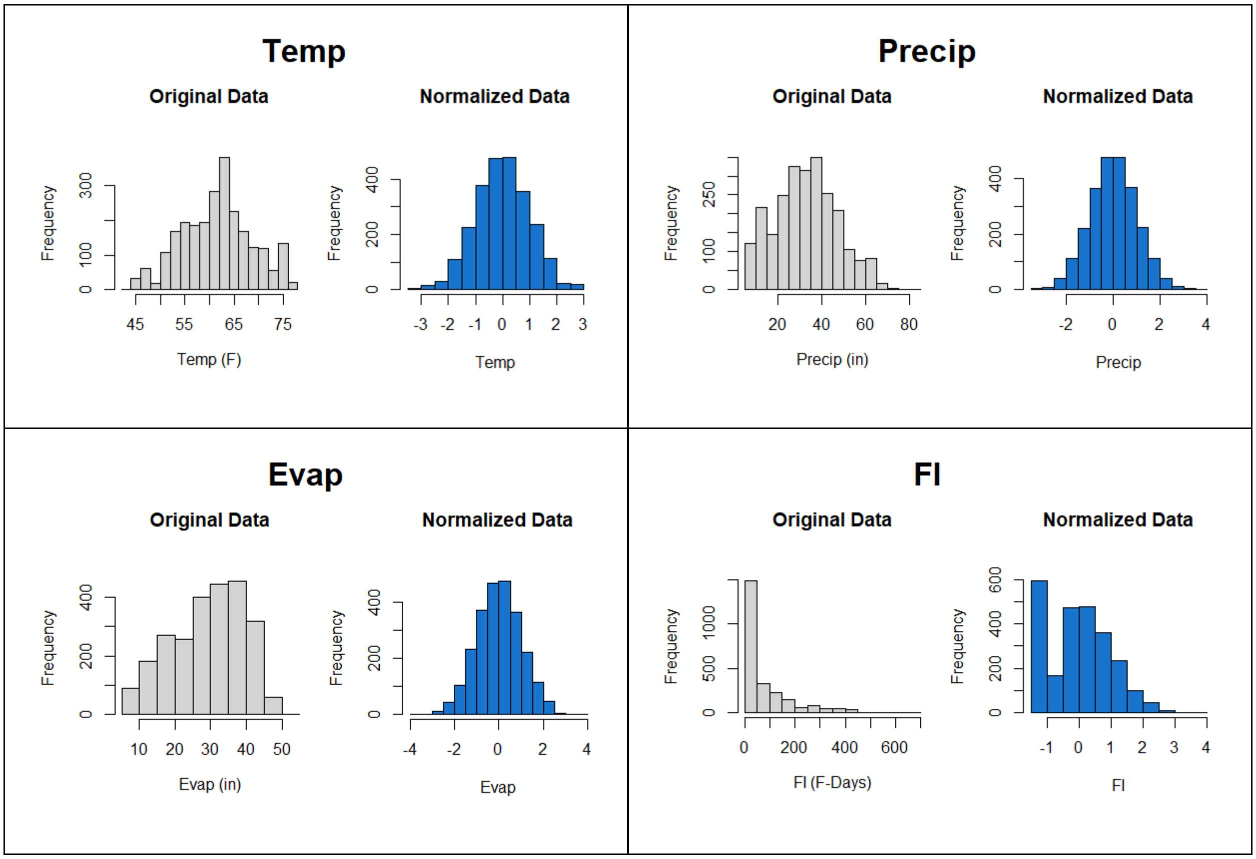


Figure 5-5- The distribution of the original and normalized numerical variables

A few of the features used in IRI prediction models were categorical. Most machine-learning algorithms (e.g., XGBoost, logistic regression, support vector machine, neural network) require all input variables and output variables to be numeric; thus, the categorical variables need to be encoded. Therefore, the one-hot encoding process was applied to categorical features such as Base_type, M&R, Road_Class, and PL to make them act as numerical variables in machine learning models.

After data processing, the data was prepared to be used in machine learning algorithms. After removing the outlier and missing data from the dataset, data were reduced to 2468 IRI data points. 75% of the dataset was randomly selected for model training and evaluation, and 25% of the data were kept as test data for evaluating and comparing the prediction models' performance. The results of the prediction models were presented in the next section.

5.2. IRI Fitting Model

The pavement roughness in many of the LTPP sections was not reported in a standard timeframe. On the other hand, there is variation in the DateTime of the IRI survey and the FWD test. To overcome the variation between measured pavement roughness survey date and visit date for FWD test, this research followed a similar method suggested by the literature (Elhadidy et al. 2019). For each section, a regression model was fitted to the measured IRI and pavement age, considering the traffic opening date as the first date. For the pavements experiencing rehabilitation and reconstruction treatments, the IRI model was refreshed, and a new model was fitted to the reconstructed pavement. Exponential and sigmoidal functions were used as proper fitting curves for predicting distresses and material behaviors in pavement studies (Ling et al. 2019). The mathematical form of the IRI fitting model used in this study is presented in Equation (15).

$$IRI = IRI_0 + \alpha e^{-(\frac{\rho}{t})^{\beta}} \quad (15)$$

Where *IRI* is the international roughness index (m/km), IRI_0 is the initial roughness of the pavement, and α, β, ρ are the maximum threshold, scale parameter, and shape parameter of the fitting curve. Figure 3 shows IRI change versus time for section 0502, construction number 3 fitted by the proposed model.

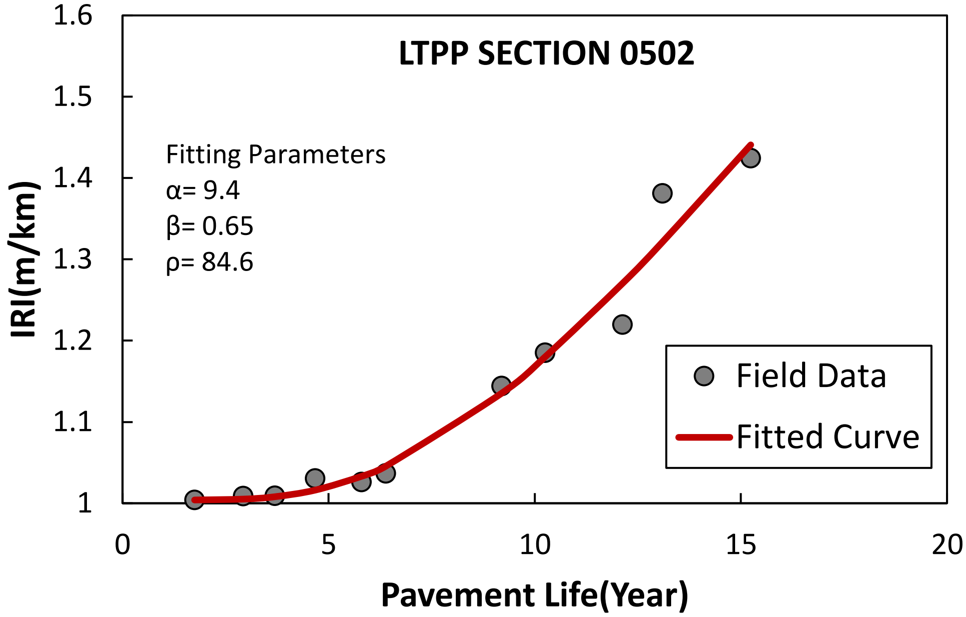


Figure 5-6- Fitted model to the IRI change versus time for LTPP section 0502-Construction Number 3

This model was fitted to the experimental sections, and the best fitting parameters were determined for each section. The root mean square error (RMSE) and R^2 for fitting the model to all experimental sections is 0.078 m/km and 93%. The statistics show that this model can accurately be fitted to the IRI change as a function of time.

The proposed model can perfectly be fitted to the IRI field data at separate construction groups. Figure 5-2 shows the IRI change versus time and the M&R history for LTPP section 0607, located in the state of Oklahoma.

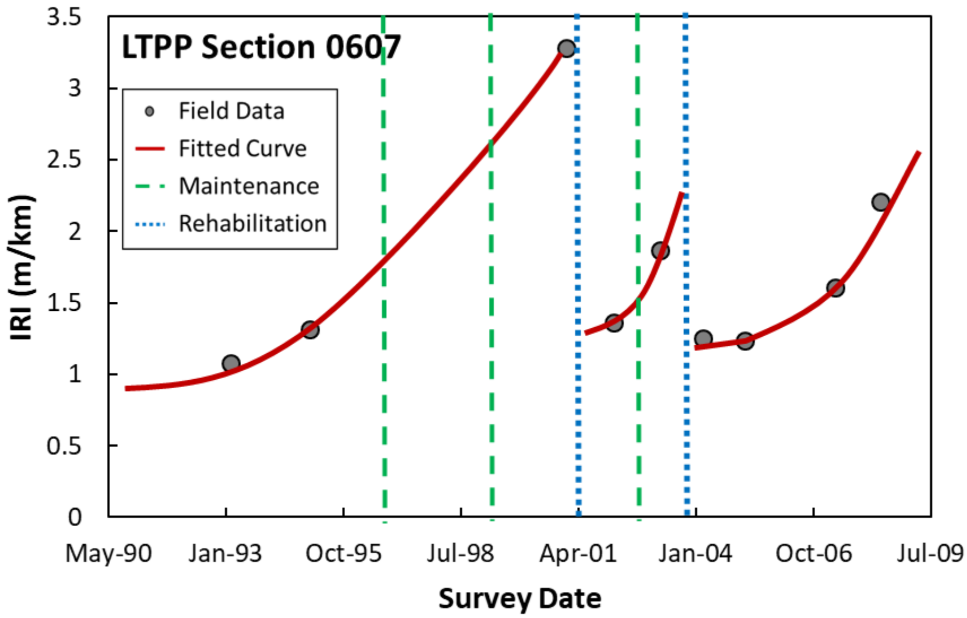


Figure 5-7- Model fit to the measured IRI at LTPP section 0607

By knowing the fitting parameters, the IRI value versus time can be plotted, and the pavement roughness at any given time after the construction can be estimated. The developed fitted model was then used for handling the inconsistency in reported IRI survey date and FWD test date by proving the IRI value at any given date of the pavement age.

5.3. Extracting Pavement Structural Properties

In the next step, the FWD was processed, and the required parameters were extracted. The FWD data were adjusted to the reference temperature of 68°F. Figure 5-8 presents the effect of temperature adjustment on the average calculated SN from FWD data. For the FWD tests at temperatures higher than the adjustment, the factor is greater than 1, which means the adjusted SNs have higher values after eliminating the effect of temperature. On the other hand, at lower temperatures, the calculated SN values should be decreased to be adjusted to the reference temperature. After temperature adjustment, the effective structural number (SN) was calculated using Equation (5).

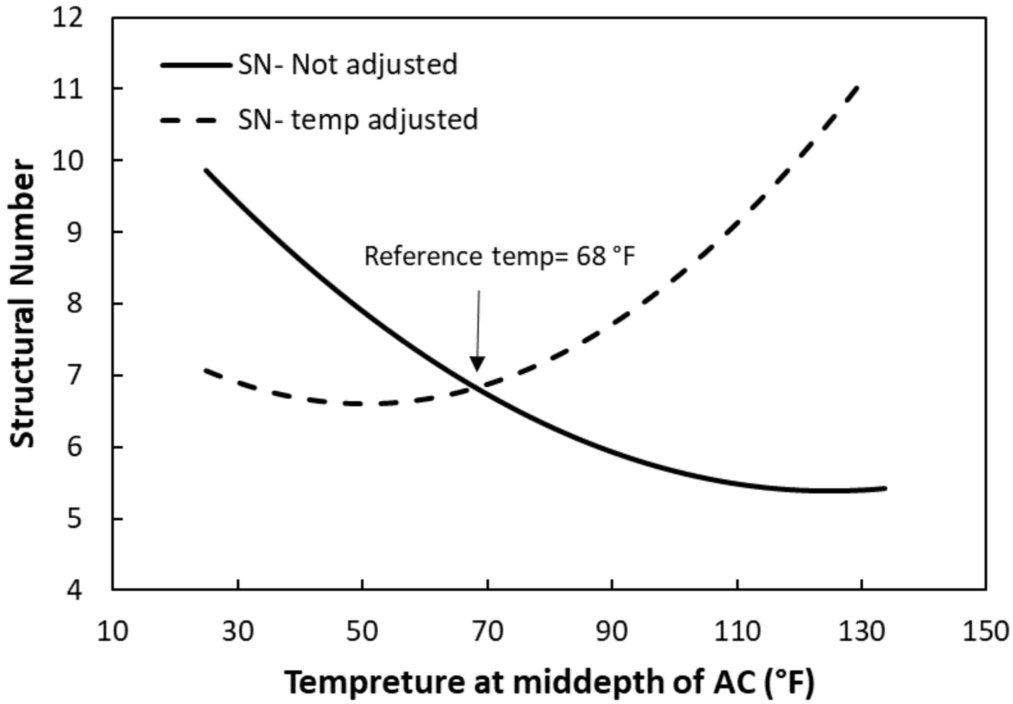


Figure 5-8 - Effect of temperature adjustment on calculated SN from FWD data

For each section, the structural number change versus time was determined and fitted with a linear model. The extracted features from FWD data were used as input into the IRI prediction model, and the rate of SN deterioration was used in the MDP decision-making model. Figure 5-9 shows the structural number for LTPP section 0114 calculated from FWD data and the deterioration rate of the SN per year.

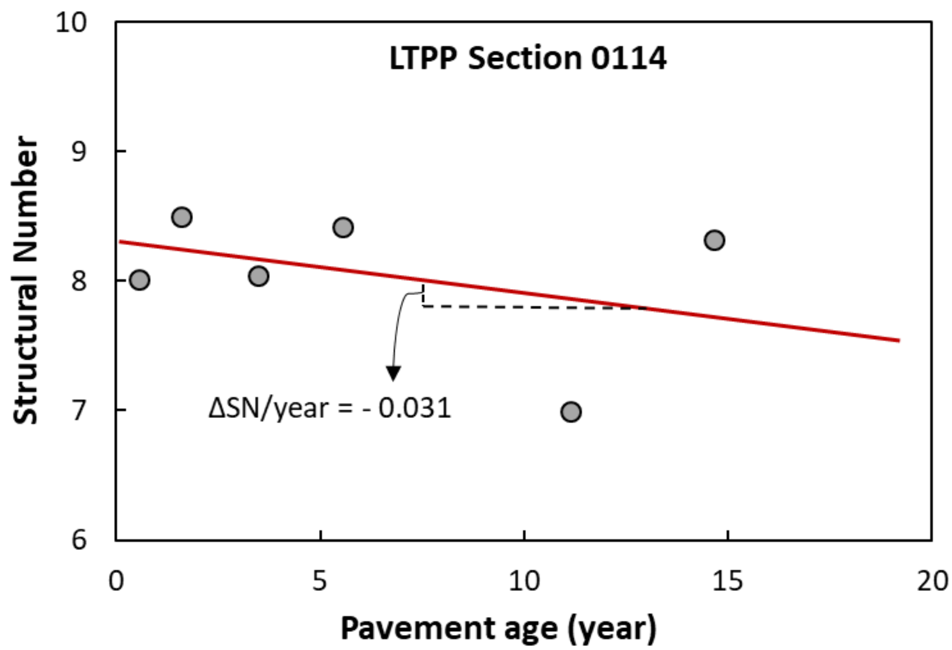


Figure 5-9- SN history for LTPP section 0114

The developed MDP model requires the average deterioration rate for estimating the structural number of next years. The rate of deterioration helps the model to predict the pavement condition and roughness better. This parameter contributes to understand the structural strength of the pavement depreciation during its lifetime and how the M&R plans can help to extend its life. Table 5-3 shows the statistics of Structural Number deterioration rate for three different construction groups, including 1) pavements with no M&R; 2) pavements with maintenance (such as crack sealing, tack coat, fog seal Coat, etc.), and 3) pavements with rehabilitation (such as overlay with AC, mill existing pavement and overlay with HMAC, RAP, WMAC, etc.). The p-value obtained from the one-way ANOVA test is 0.016, which indicates a significant difference in the distribution of SN deterioration rate in three different groups assuming 5% confidence.

Table 5-3- Statistics of structural number change per year for three groups of pavement M&R experience

Group	M&R	Min (ΔSN/year)	Max (ΔSN/year)	Mean (ΔSN/year)	SD (ΔSN/year)
1	No M&R	-2.10	0.718	-0.29	0.61
2	Maintenance	-0.511	0.418	-0.052	0.21
3	Rehabilitation	-0.529	0.868	-0.16	0.35

The M&R experience has considerable influence on the mean of the structural number deterioration rate. Pavements with no M&R show the highest rate of change, which is because of the deterioration of the pavement structure during its lifetime and lack of proper maintenance for the pavement. Applying routine maintenance can decrease the deterioration rate of the pavement structure. Rehabilitation of the pavement also helps the sustainability of the pavement structure.

5.4. IRI Prediction Models

This model aims to predict the IRI change versus time given the pavement section's structural properties, including structural features, FWD data extracted features, subgrade soil properties, climate and traffic information, road type, etc. The intrinsic complexity of the pavement properties and parameters which affect the pavement behavior during the pavement lifetime asserts the need to try different machine learning predictive models. Each model's performance and accuracy vary based on several factors related to the dataset, including the type of the features, distribution, and correlation between variables, amount and severity of the outliers, number of available observations, etc. In this study, generalized linear (GLMnet), Support Vector Regression (SVM), Multivariate Adaptive Regression Splines (MARS), Artificial Neural Network (ANN), and XGBoost models were investigated and used for predicting the IRI of the flexible pavements. Each method incorporates a different mathematical technique in the process, which was explained in the previous sections. All the proposed models were optimized to get the best performance out of them, and the best prediction model was implemented in the MDP M&R decision model. The results of the prediction models were discussed in the following section.

5.5. M&R Decision-Making Model

Pavement deterioration is a process of reducing the carrying capacity of the pavement. The surface condition is the primary and essential indicator of the structural integrity of the pavement. Pavement roughness is susceptible to the deterioration of the pavement structure and can be used as a parameter in M&R policy decisions. However, the surface roughness may not fully capture the structural integrity since, in many cases, the distresses are internally growing, but the surface does not show any signs of deterioration. These kinds of instabilities can be captured by FWD non-destructive test. The structural number is an index that can be estimated from the FWD data with a reliable level of confidence. Thus, IRI and SN act as two powerful indexes in predicting pavement performance and the effect of M&R strategies. The MDP model in this study aims to develop a stochastic pavement management planning based on the pavement structural integrity and pavement roughness.

The developed MDP model recommends the maintenance and rehabilitation policies helping from a pavement roughness index (IRI) and a structural integrity index (SN) deterioration, and the optimized policy which will be determined based on cost analysis of the plan. In the following, the elements of the developed model will be explained. The MDP model will use the developed IRI prediction machine learning model, rate of SN deterioration, and predicted traffic load as inputs. The optimized M&R plan will be introduced by implementing a proper reward as a function of the treatment benefit/cost. In the following, the detailed methodology and initial results will be explained.

To consider the effect of M&R on the IRI change and pavement deterioration, the period of the IRI survey were divided into three major groups based on the applied treatment:

1- Pavement with no applied M&R treatment;

2- Pavement with regular maintenances such as crack sealing, cheap seal, slurry seal coating, etc., and

3- Pavement with rehabilitation treatments such as milling and AC overlaying and resurfacing.

Several algorithms have been developed for solving the MDP and finding the optimum policy. Q-learning is a model-free reinforcement learning algorithm that was extensively used to solve the reinforcement learning problems that do not require an explicit definition of a Markov decision process. It trains an agent to learn an optimal policy from a dynamic environment, and the optimal learned policy tells the agent what action to take at each state. The MDP experimental model for this study was explained in the following sections.

6. MODEL DEVELOPMENT AND DISCUSSION

After data perpetration, several machine learning models were used for the prediction of the IRI. Each model has advantageous that can help to increase prediction accuracy. The algorithm behind each machine learning model and its advantages and disadvantages were discussed in section 4. Each machine learning model has parameters that govern and tune the training process. These parameters are named hyperparameters and need to be optimized during the training process. A 5 fold cross-validation was used for hyperparameters tuning and model validation. The total data were randomly divided into training and testing datasets with a ratio of 0.75 and 0.25, respectively. Machine learning models were trained on the training datasets, and their performances were evaluated using the testing dataset. In the following, the tuning results and the performance of the models on training and test dataset were presented. In the end, the best model that can be used for IRI prediction was introduced.

6.1. Generalized Linear Model (GLM)

Generalized Linear Models (GLM) extend the linear modeling framework to variables that are not normally distributed. If the relationship between independent variables (X) and dependent variables does not look linear or the dependent variable (Y) variance does not look constant regarding X, the GLM model can be helpful. For example, in the pavement performance prediction, the change in performance is lower at the early age of the pavement life; however, increasing the age increases the rate of pavement performance deterioration. The relationship between the IRI and structural number is shown in Figure 5-6 cannot be adequately explained by linear fitted models. The GLM relates the linear model to the response variable via a link such as a log, logit, inverse functions, while the variance of the expected value is not constant. For this model, the logarithmic link function was used.

The regularization technique, explained in section 4, was used for reducing the complexity of the model. This technique reduces the model parameters weights and increases the model flexibility. Thus, the model shows a lower error variance on the testing dataset. The regularization parameters, α, and λ, are part of the model hyperparameters that need to be optimized during the model training. The parameter α balances the weights between two penalty forms, including LASSO with $\alpha=1$ and RIDGE with $\alpha=0$. The parameter λ determines the overall strength of the penalty. By Increasing the λ, the model parameters' weights will be reduced, and less important features will be zeroed out. In other words, the λ keeps the important features and removes less correlated features from the model. A set of α = [0, 0.01, 0.1,0.5,1] and λ = [1e-4, 5e-4, 1e-3, 1e-2] was optimized based on the best performance by the lowest root mean square error. For α= 0.5 and λ = 1e-4, the GLM model shows the best performance with a root mean square error (RMSE) of 0.28. Figure 6-1 shows the hyperparameter optimization for the GLM model developed for the pavement performance prediction. The optimized model at λ = 1e-4 removed the Temp, Road_Class, Base_Type features.

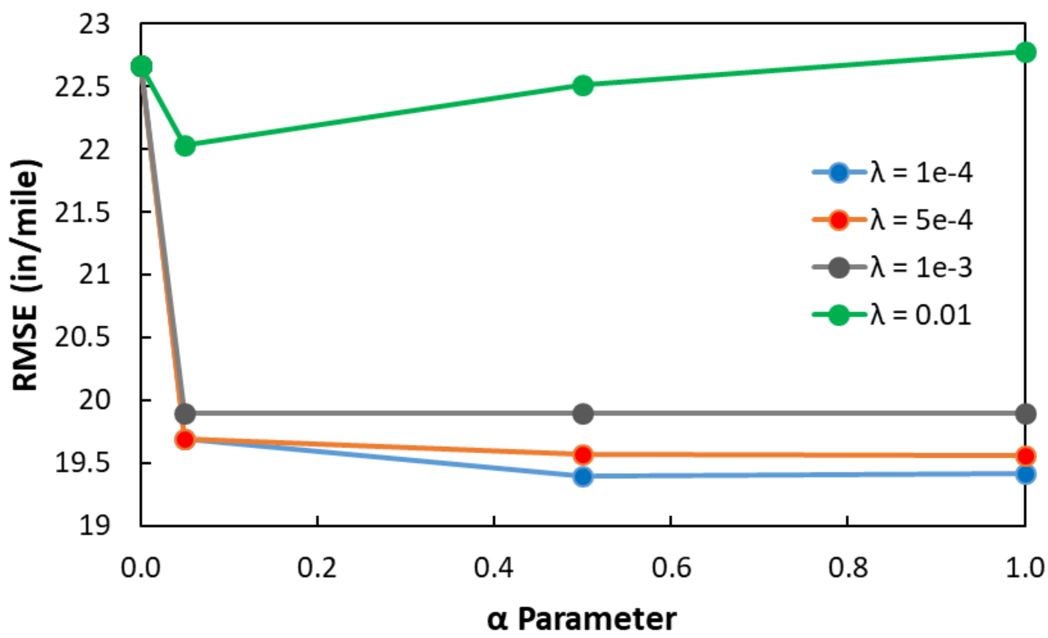
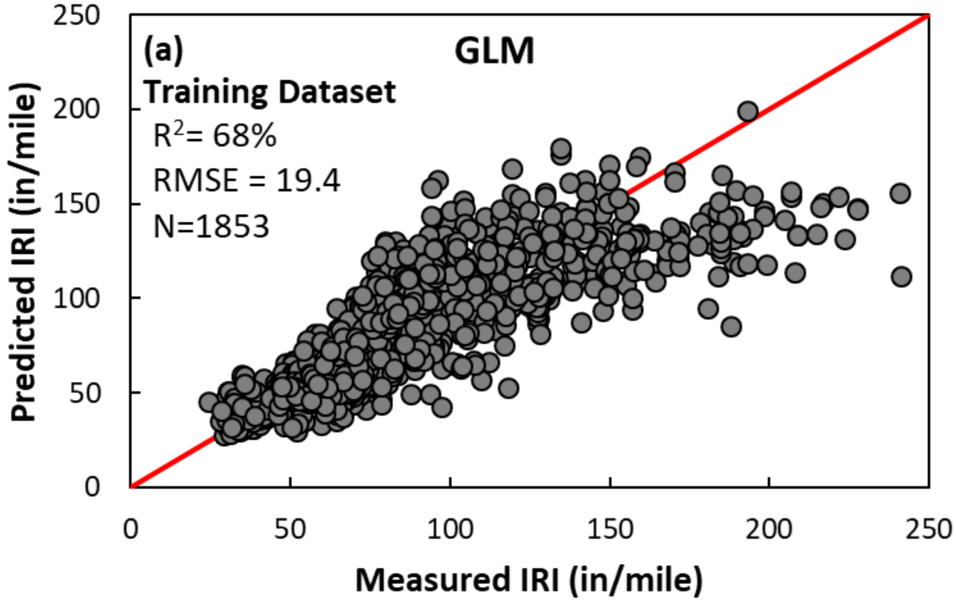


Figure 6-1- Hyperparameter optimization for GLM model

The model was trained and evaluated using 5-fold cross-validation, and root mean square error as an accuracy metric by finding the best tuning parameter. Figure 6-2 shows the prediction results for the training and testing datasets. The GLM model shows RMSE of 19.4 in/mile and coefficient of determination of %68. The model prediction using the test dataset shows RMSE of 18.8 in/mile and R^2 of %74. This model performs better than the linear regression models; however, for higher IRI values, the GLM model underpredicts the IRI. This can be because of a lack of enough data on the high values of IRI. The number of observations with high values of IRI is much less than lower values, and this is the reason for data imbalance error in the GLM model. Other ML models can be used to reduce the error of the prediction in higher IRI values.



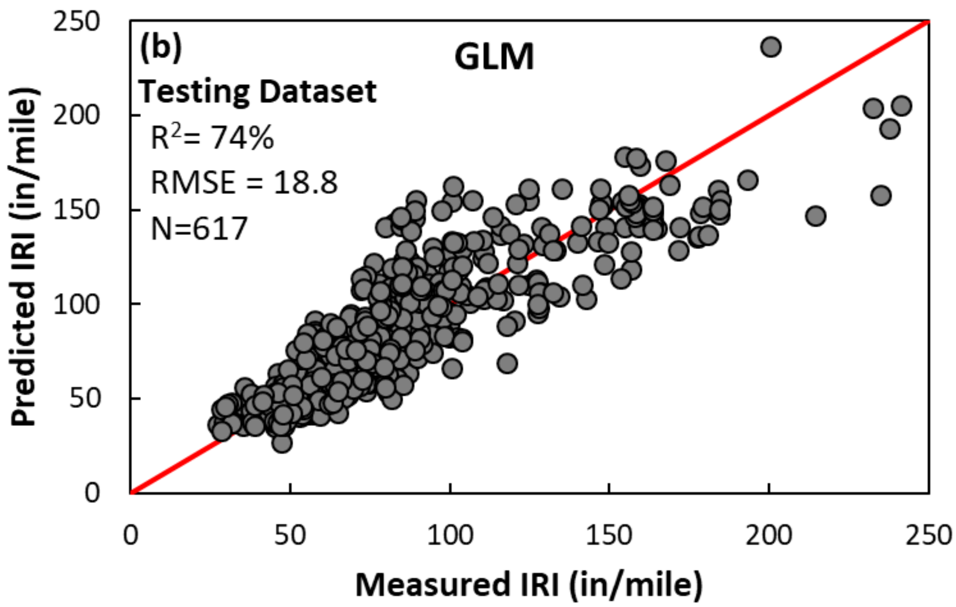


Figure 6-2- Pavement performance prediction by GLM model for (a) training(%75) and (b) testing(%25) datasets

6.2. Support Vector Machine (SVM) Regression

Support Vector Machine is a machine learning method that can be used for classification and regression problems. In this model, a non-linear function maps the dependent variables into high dimensional kernel named feature space. This mapping helps to reduce the effect of outliers on the prediction values. The SVM regression is a nonparametric method looking for the best fit to the training data to construct a mapping function while the ability to generalize to unseen data is maintained. Thus, they can fit a large number of functional forms and does not assume anything about the form of the mapping function other than patterns that best represent the dependent variable. In the IRI prediction model, an SVM regression model with nonlinear kernel can get the nonlinear relation between the roughness and pavement features at high IRI values. The SVMR model developed for the prediction of IRI has a radial kernel that works better than other choices in this example. Figure 6-3 shows the hyperparameter optimization for the SVMR model developed for the pavement performance prediction. A set of C = [0.01, 0.1, 1, 10, 50] and ε=[0.1, 0.5, 0.1] were optimized based on the best performance considering the lowest root mean square error. In the SVM model, ε controls the number of close observations ignored by the model, and C balances the weight between the training error and the penalty term. For C= 10 and ε = 1, the SVMR model shows the best performance with a root mean square error (RMSE) of 10 in/mile.

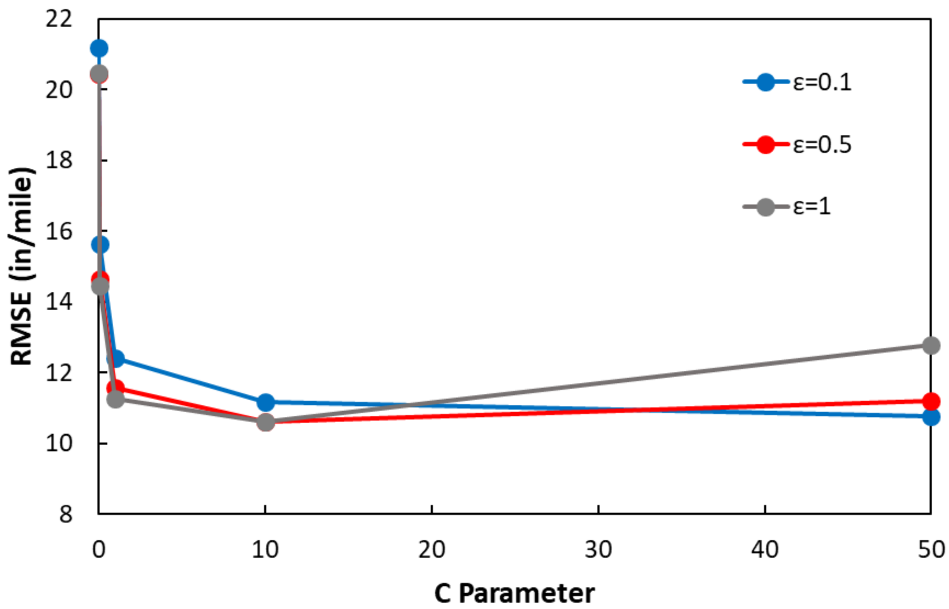
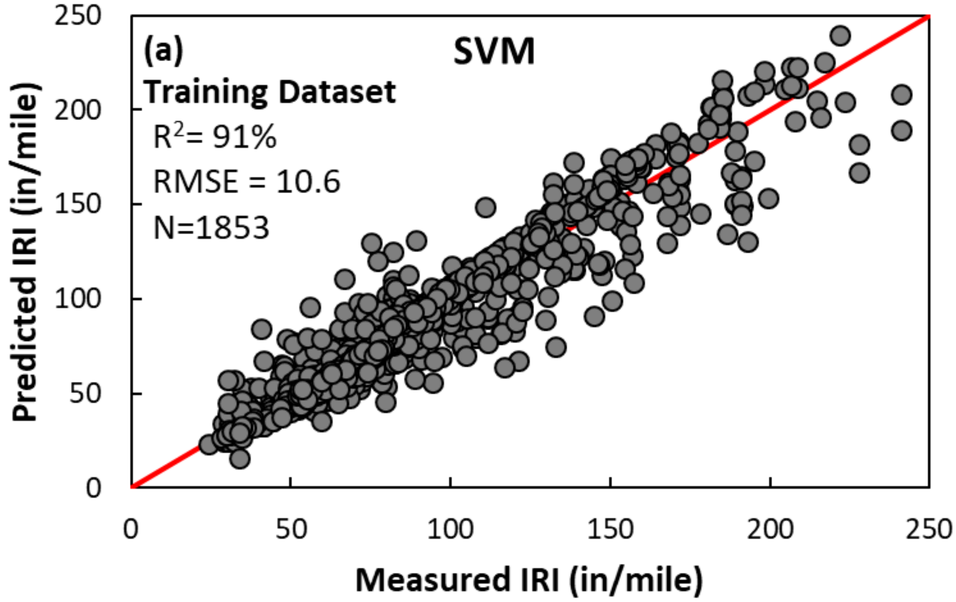


Figure 6-3 - Hyperparameter optimization for SVMR model

The mole was trained and evaluated using 5-fold cross-validation, and root mean square error as an accuracy metric by finding the best tuning parameter. Figure 6-4 shows the prediction results for the training and testing datasets. The SVM model shows RMSE of 10.6 in/mile and a coefficient of determination of 91%. The model's prediction using the test dataset shows an RMSE value of 17.19 in/mile and R^2 of 77%. This model works better than the GLM model, especially for higher values of IRI. This is because of the radial kernel, which better represents the high IRI values in the prediction.



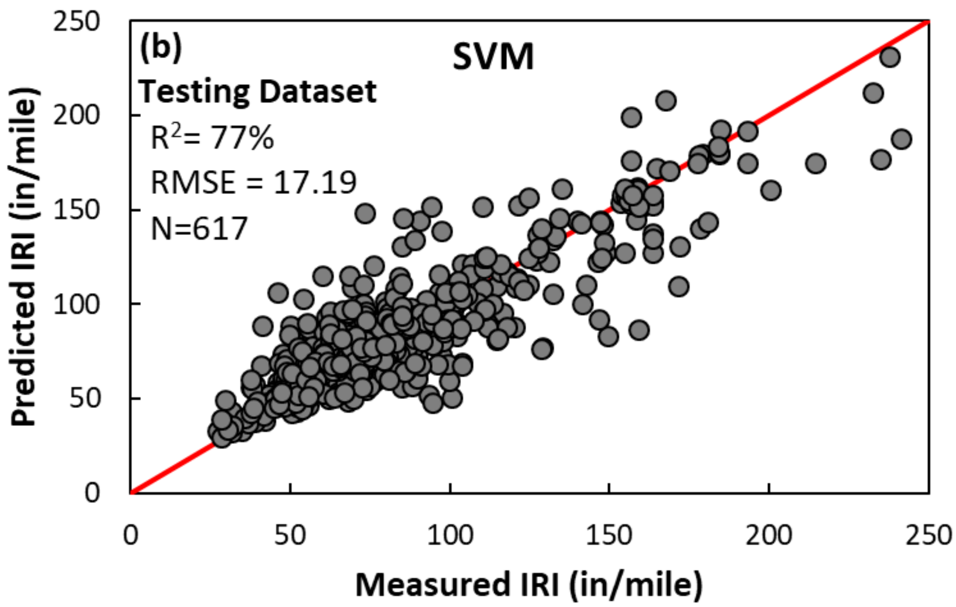


Figure 6-4- Pavement performance prediction by SVM model for (a) training(75%) and (b) testing(25%) datasets

6.3. Multivariate Adaptive Regression Splines (MARS)

Multivariate Adaptive Regression Splines (MARS) is a non-parametric function estimation technique that shows excellent performance for fitting non-linear multivariate functions. This model breaks the range of independent variables, X into n number of bins, and fits the best line for each bin. A linear model is fitted to each bin, which reduces the amount of total error compare to a linear regression model. The fitted model to each bin could be linear or multivariate nonlinear. By this technique, the MARS change the continuous variables into clusters that optimally can be estimated by a nonlinear function. The number of bins can be tuned during model training by using the knots parameters, which indicate the number of breaks in the model. In addition, the degree of freedom is another hyper tune parameter that controls the level of nonlinearity and interaction between the features. Figure 6-5 shows the parameter hyperparameter optimization for the MARS model developed for the pavement performance prediction. The number of breakpoints = 5, 6, 7, …, 20, and polynomial degrees of 1,2,3 and 4 were chosen for optimization using the RMSE metric. For knots= 10 and n = 3, the MARS model shows the best performance with root mean square error (RMSE) of 17.14 in/mile.

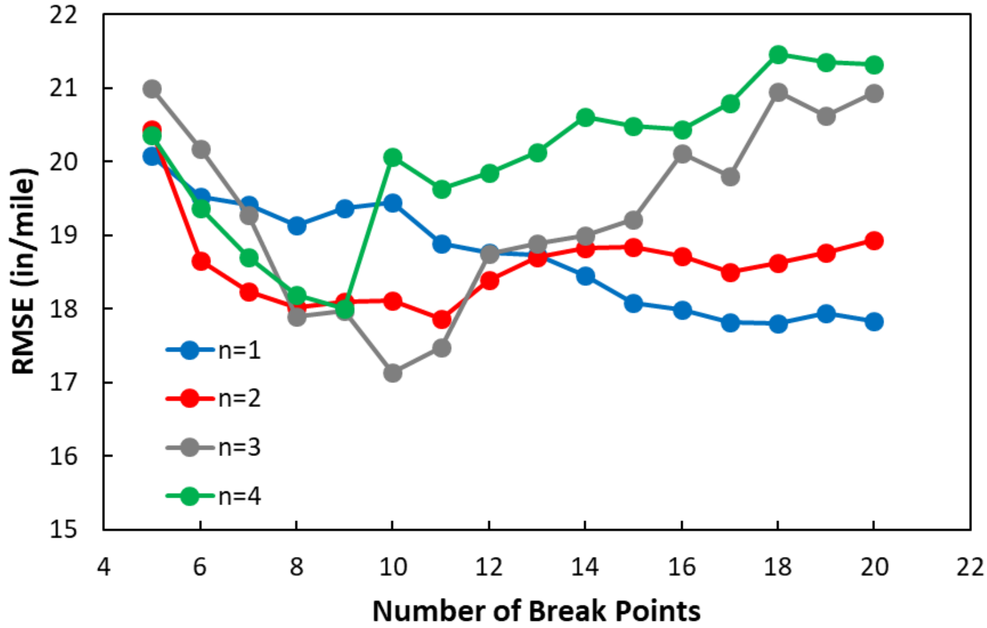
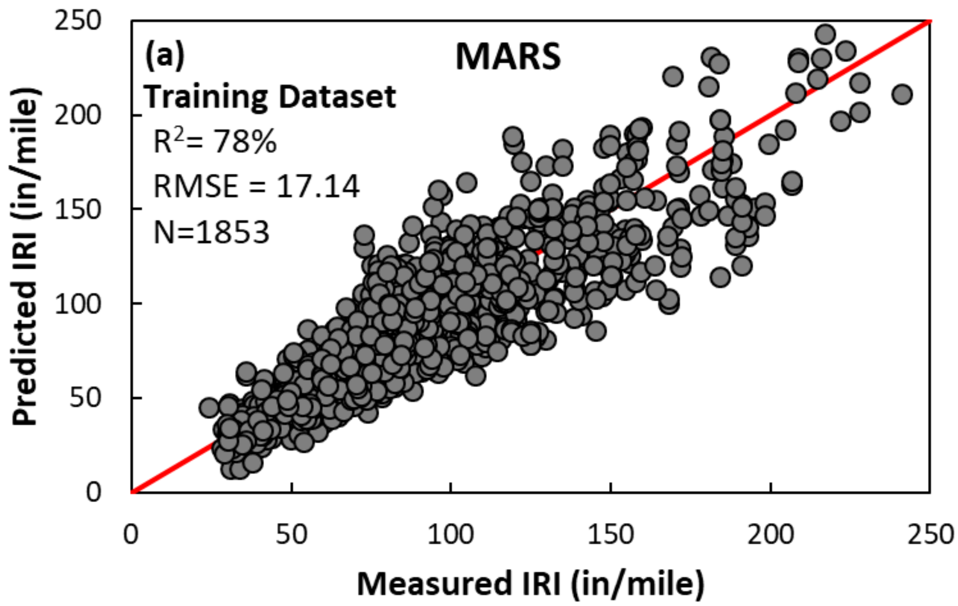


Figure 6-5- Hyperparameter optimization for MARS model

The model was trained and evaluated using 5-fold cross-validation and root mean square error as an accuracy metric by finding the best tuning parameter. Figure 6-6 shows the prediction results for the training and testing datasets. The MARS model shows RMSE of 17.14 in/mile and the coefficient of determination of 78%. The model's prediction using the test dataset shows RMSE of 17.6 in/mile and R^2 of 80%. This model does not show good performance, such as other machine learning techniques. One of the reasons could be the presence of categorical variables. The MARS model gives low bias but high variance. This is because of the overfitting problem in this model. The overfitting problem was compensated by hyperparameter optimization, but the model still shows the high variance in prediction results using the test dataset.



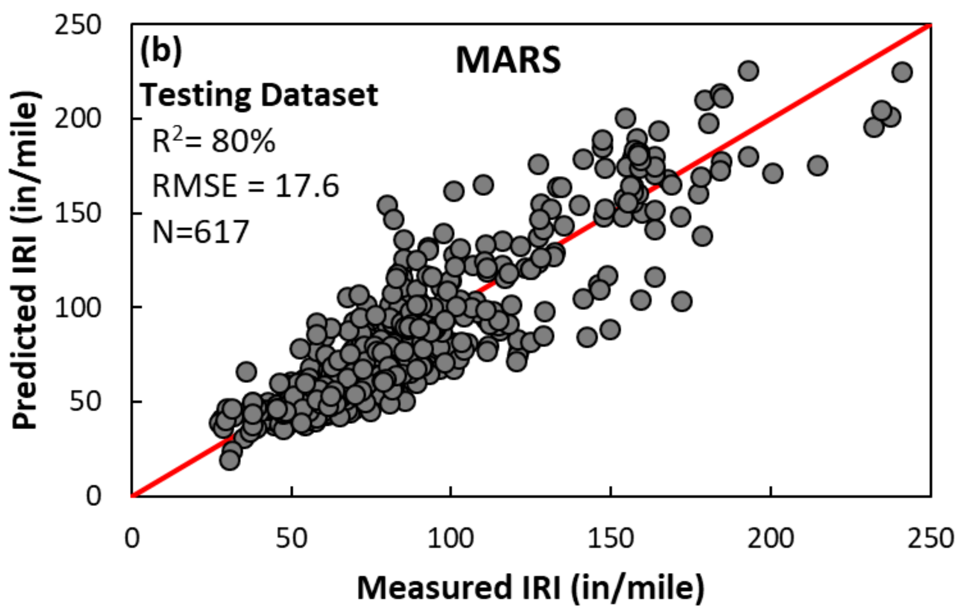


Figure 6-6- Pavement performance prediction by MARS model for (a) training(%75) and (b) test(%25) datasets

6.4. Artificial Neural Network(ANN)

An ANN model was developed for the prediction of IRI in this study. Researchers often use neural networks for prediction IRI because of their predictive accuracy(Roberts et al. 1998; Kargah-Ostadi 2014). One of the strengths of neural networks is their ability to model highly nonlinear data. Neural networks were shown to produce results similar to MARS in modeling nonlinear functions(De Veaux et al. 1993). The neural network results are susceptible to model complexity. The number of layers, the number of hidden units in each layer, and the type of transfer functions are essential parameters that should be tuned during the training and evaluation. After the hyper tuning process, the optimized model with the lowest RMSE error was developed. The optimum model compromise of 3 layers with 9,12 and 9 neurons. The ReLU function was chosen as the activation function of all layers in the ANN model.

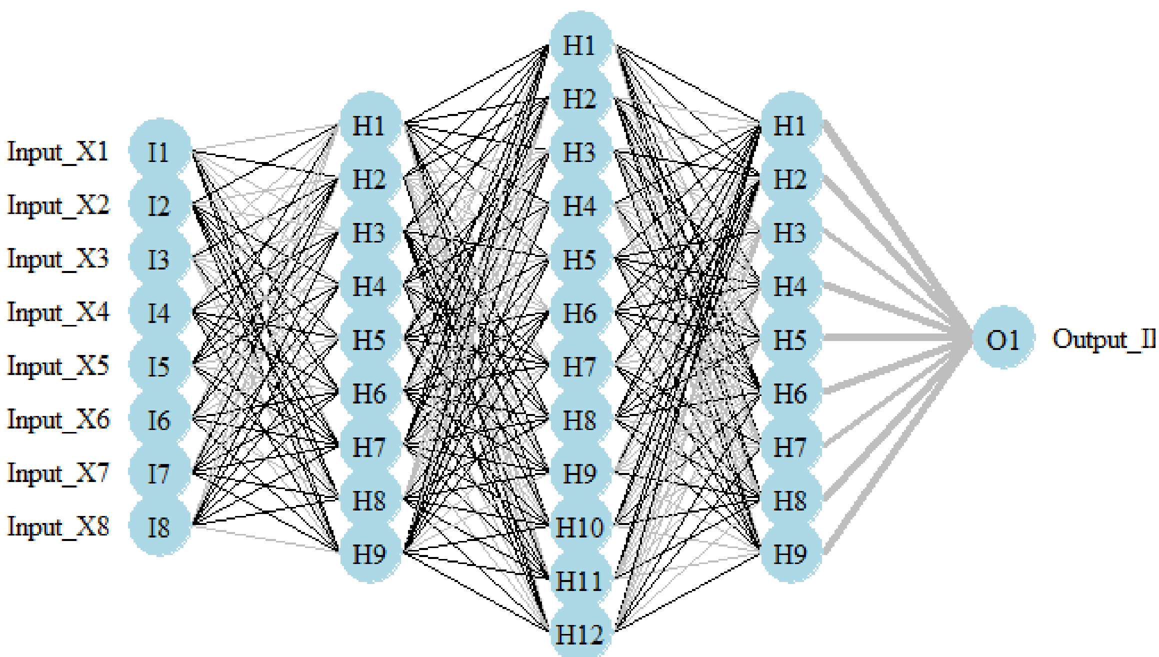
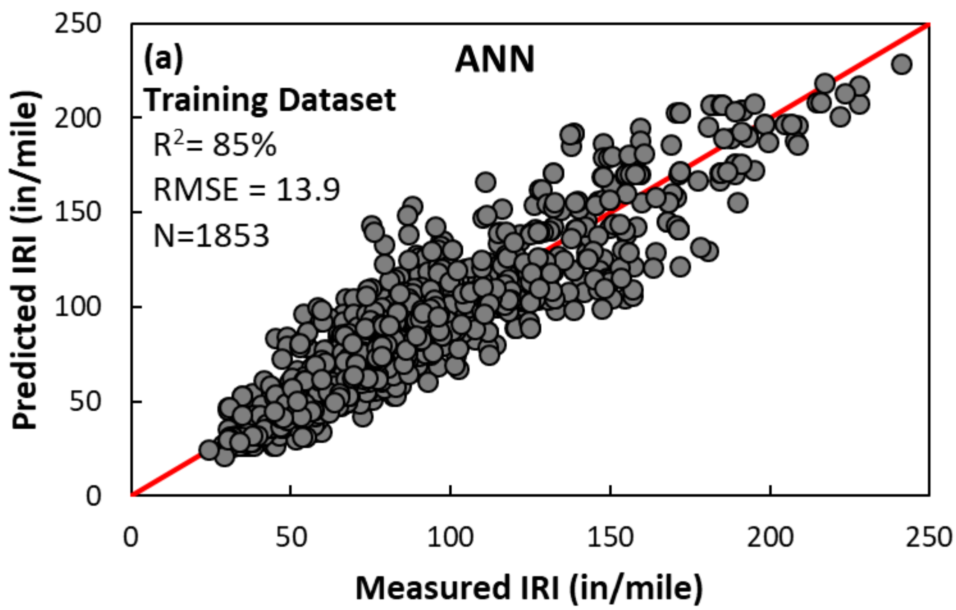


Figure 6-7- ANN model plot developed for IRI prediction

The model was trained by finding the best ANN setup, and the root mean square error was used as an accuracy metric. Figure 6-8 shows the prediction results for the training and testing datasets. The ANN model shows RMSE of 13.9 in/mile and coefficient of determination, R^2, of 85%. The model performance using the test dataset shows RMSE of 16.2 in/mile and R^2 of 81%. This model shows good performance compared to previous models and can be used as a prediction model to predict pavement performance during its lifetime. However, we need better accuracy for the IRI prediction model to increase the MDP decision model performance.



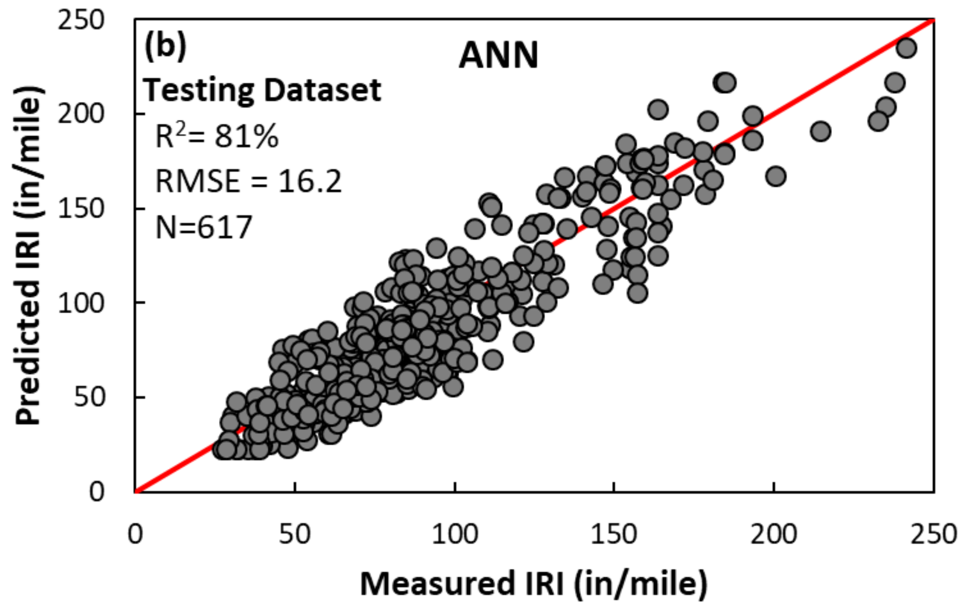


Figure 6-8- Pavement performance prediction by ANN model for (a) training(%75) and (b) testing(%25) datasets

6.5. XGBoost

The XGBoost tree model ensemble tree-based method works well with the tabular data and is one of the best models for pavement performance prediction. The XGBoost model can handle the model's missing values since it can skip the missed variable and work with other existing variables. Thus, the number of observations increases, and none of the sections will be discarded due to lack of information on some features. Other benefits of the XGBoost tree model are parallelization, regularization, and cross-validation, leading to increasing the computational speed and avoiding overfitting in the model.

The parameters of the XGBoost model that should be optimized are the maximum depth of the tree models, total number of trees to grow in each cycle, the minimum number of variables used in the tree model, minimum number of samples in the batch of observations, or measured data and few others. Figure 6-9 shows the parameter hyperparameter optimization for the XGBoost model developed for the pavement performance prediction. The maximum depth of the tree, including 3, 5, 7, 10, and a maximum number of iterations of 10 to 3000, were chosen for optimization using the RMSE metric. For the maximum depth of 5 and n value considered 3000, the XGBoost model shows the best performance with root mean square error (RMSE) of 5.92 in/mile.

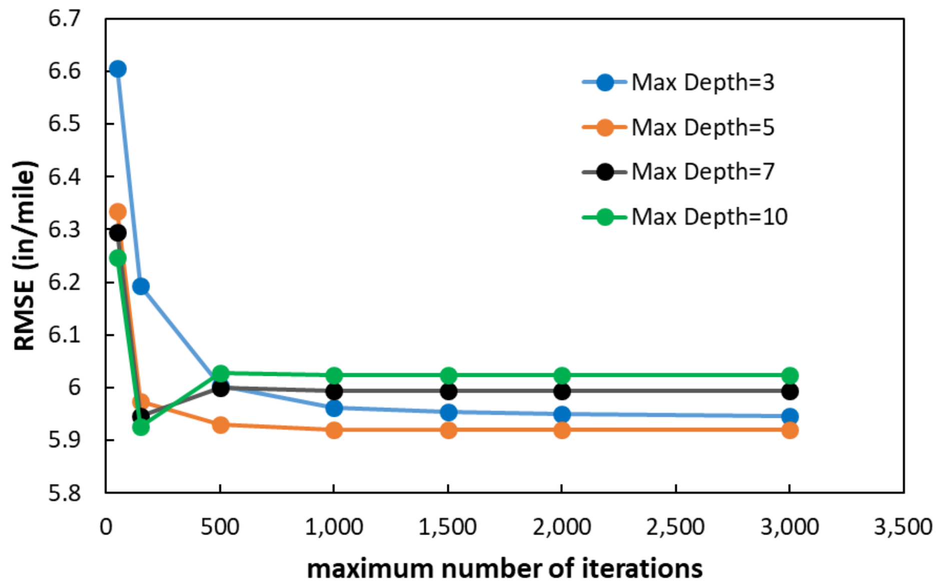


Figure 6-9- Hyperparameter optimization for XGBoost model

The model was trained by finding the best XGBoost parameters, and root mean square error was used as an accuracy metric. Figure 6-10 shows the prediction results for the training and testing datasets. The XGBoost model shows RMSE of 5.92 in/mile and the coefficient of determination of 95%. The model's prediction using the test dataset shows RMSE of 13.1 in/mile and R2 of 91%. This model shows outstanding performance and can be used as a prediction model for predicting pavement performance during its lifetime.

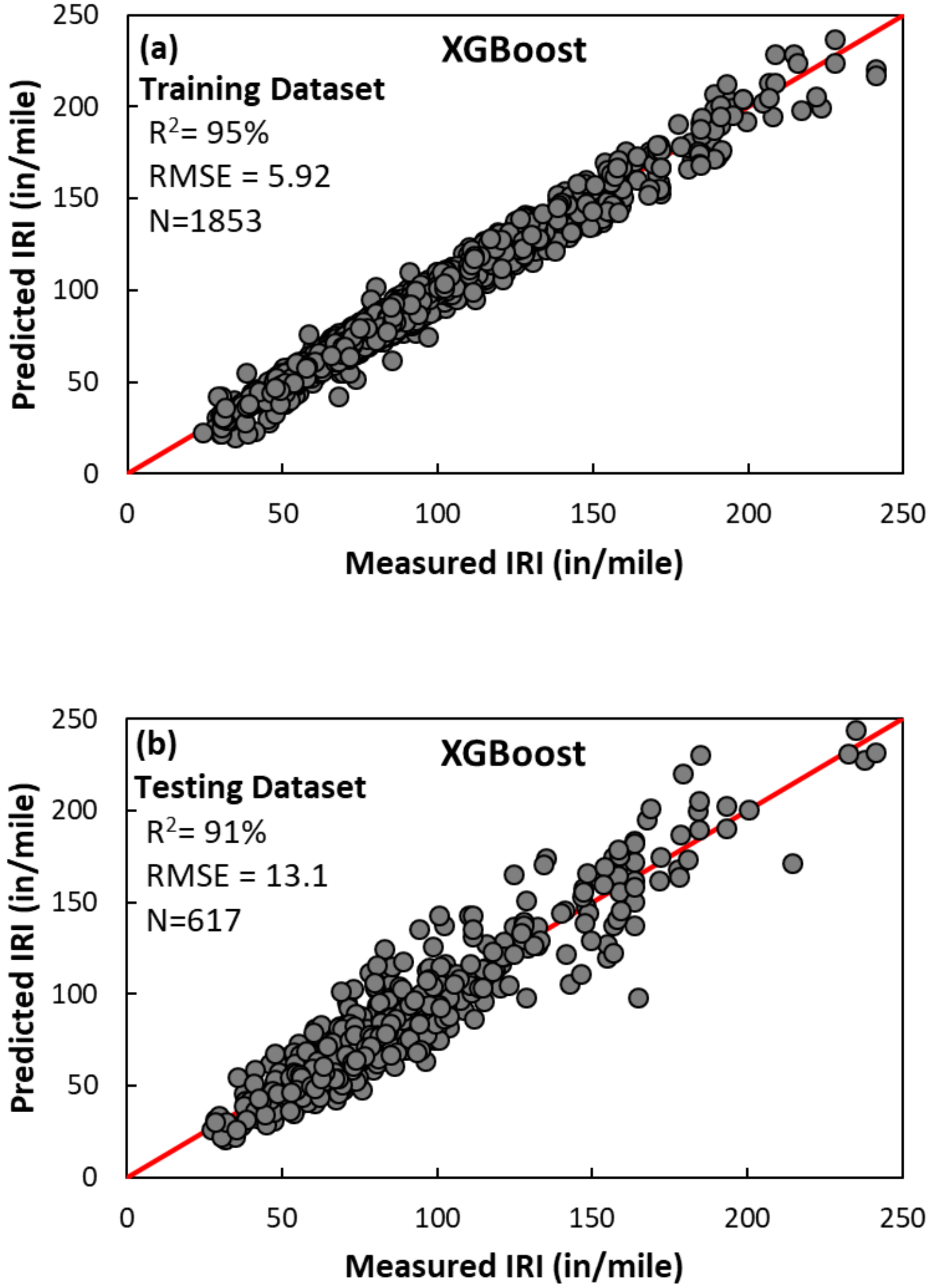


Figure 6-10- Pavement performance prediction by XGBoost model for (a) training(%75) and (b) testing(%25) datasets

6.6. Performance of New Developed Predictive Models

After developing the prediction models using different machine learning techniques, the models' performance needs to be compared, and the contribution of the features in each model being investigated. Table 6-1 shows the input variables used in the final developed model after regularization and removing less important and correlated features. Removing these features from input variables decrease model complexity and increase the accuracy of the prediction.

Table 6-1- Selected features for the final developed models

Performance		**GLM**	**SVM**	**MARS**	**ANN**	**XGBoost**
Training RMSE (in/mile)	N=1853	19.4	10.6	17.14	19.9	5.92
Training R^2	(75%)	68%	91%	78%	78%	95%
Testing RMSE (in/mile)	N=617	18.8	17.19	17.6	16.2	13.1
Testing R^2	(75%)	78%	77%	80%	81%	91%
Input Variables		**GLM**	**SVM**	**MARS**	**ANN**	**XGBoost**
Age		✓	✓	✓	✓	✓
SNeff		✓	✓	✓	✓	✓
SIP		--	--	--	--	--
PD1.5		✓	✓	✓	✓	✓
D_{AC}		--	--	✓	✓	✓
D_{total}		--	--	✓	✓	--
AADTT		✓	✓	✓	✓	✓
Temp		--	✓	--	--	--
FI		✓	--	--	--	--
Precip		✓	✓	✓	✓	✓
Evap		--	--	✓	--	--
M&R		✓	✓	✓	✓	✓
Base_type		--	✓	✓	--	✓
Plasticity		--	✓	✓	--	✓
Road_Class		--	--	✓	--	--

Age, effective structural number (SNeff), Deflection at 1.5 times of pavement thickness (PD1.5), traffic load (AADTT), precipitation, and M&R history were used in all of the models. SIP had a very high correlation with the structural number and was removed from the input variables. Asphalt and total Pavement thickness were used in MARS, ANN, and XGBoost model. The SVM and GLM models allocated very low weights to these features and removed them from the prediction model. This could be because of the correlation of these variables with SN of the pavement. The correlation between the IRI and temperature was not as good as other environmental characteristics, and thus, the average temperature was not selected as a good predictor in IRI prediction models.

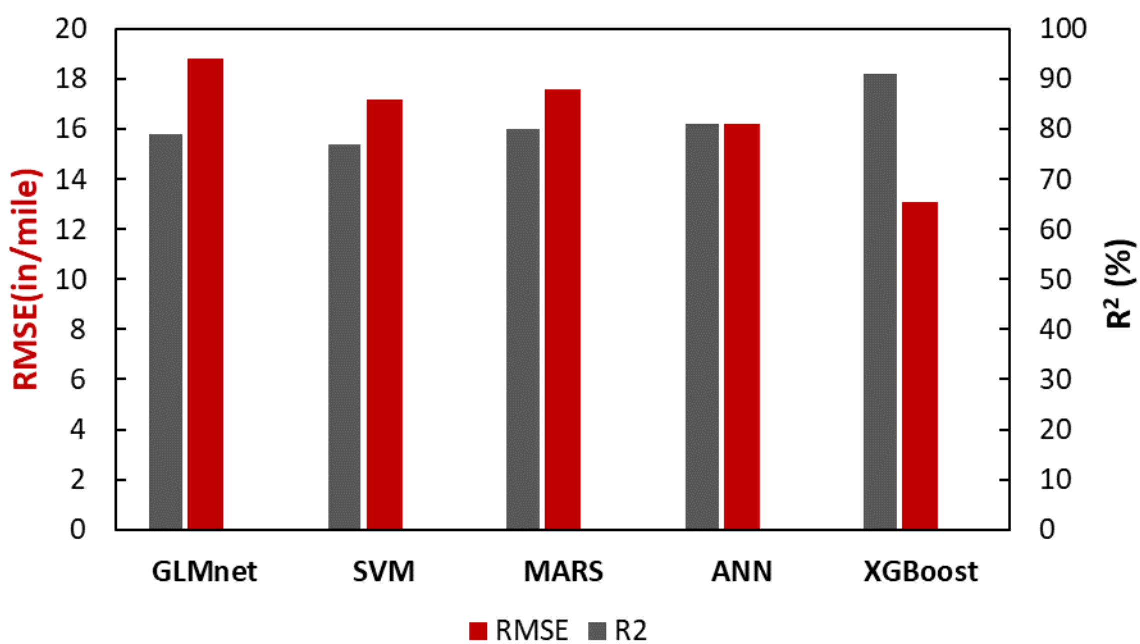


Figure 6-11 - Comparison of the performance of new developed predictive models

Figure 6-11 shows the comparison between the performance of IRI prediction models. Most of the newly developed models show acceptable performance. However, the XGBoost model shows the best RMSE and the highest R^2 between the developed models. This is because of the novel cutting edge algorithms that were implemented by this model. The tree-based regression, boosting algorithm, cross-validation, regularization of the features helps to have a robust, accurate model that can precisely predict the IRI based on the available information collected for the pavement section.

The next step is an assessment of how the input variables are working with the model. The developed models and the contribution of each input variable should be investigated. The XGBoost model does not provide an estimation parameter or weight for each feature, which can be used to interpret the relationship between a specific feature and the supervisor. The feature importance in XGBoost is a parameter that indicates the contribution of that feature in the model's decision tree process and performance. One of the essential metrics is gain, which implies the corresponding input variable's relative contribution to the model calculated by taking each feature's contribution to each tree model. The sum of the gain of all input variables is equal to one. When compared to another feature, a higher value of this metric implies it is more important for generating a prediction. Figure 6-12 shows the feature importance of the developed XGBoost model for the prediction of IRI. Features that contribute most to predicting IRI by the XGBoost model are effective structural number and peak deflection at 1.5 times of pavement thickness.

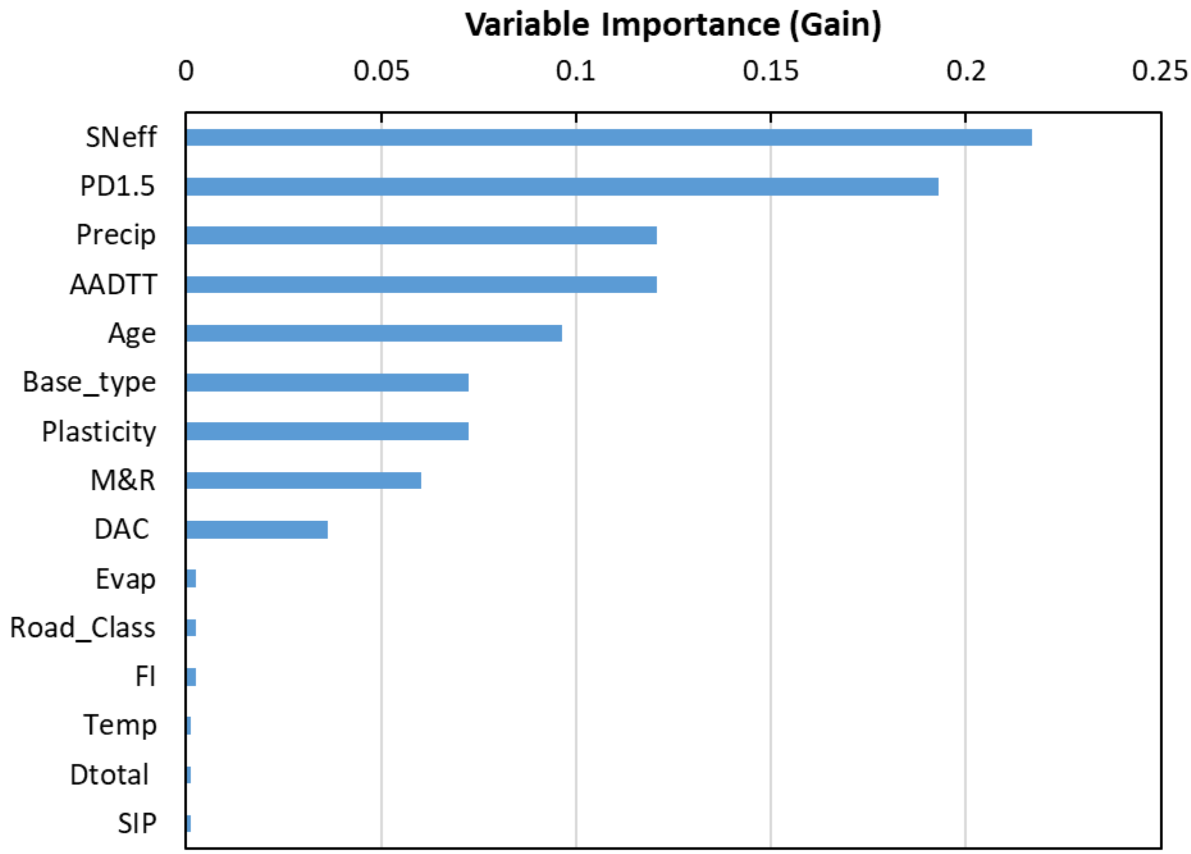


Figure 6-12 - Variable Importance derived from the XGBoost model

7. M&R STRATEGY

Pavement deterioration is a process of reducing the carrying capacity of the pavement. The surface condition is the primary and essential indicator of the structural integrity of the pavement. Pavement roughness is susceptible to the deterioration of the pavement structure and can be used as a parameter in M&R policy decisions. However, the surface roughness may not fully capture the structural integrity since, in many cases, the distresses are internally growing, but the surface does not show any signs of deterioration. These kinds of instabilities can be captured by FWD non-destructive test. The structural number is an index estimated from the FWD data with a reliable level of confidence. Thus, IRI and SN act as two powerful indexes in predicting pavement performance and the effect of M&R strategies.

A Markova Decision Process (MDP) model is a reinforced learning model that contains a set of possible states (S), a set of possible actions (A), a reward function R (s, a), and a transition function T (s, a). By having the MDP model, the agent can explore the environment by taking action, visiting different states, and accumulating rewards. After exploring the environment, each state can get a value that determines the goodness of being in that state and is a function of the expected total reward gained from that state. MDP model aims to find the value function and the optimum policy. An optimum policy is an instruction for the agent that determines the best action at each stage, which leads to the maximum reward.

The developed MDP model recommends the maintenance and rehabilitation policies helping from a pavement roughness index (IRI) and a structural integrity index (SN) deterioration, and the optimized policy with be determined based on cost and future benefit of the plan. In the following, the elements of the developed model will be explained.

7.1. MDP Environment

The pavement network has an infinite number of pavement sections with different pavement types, traffic load, functional class of the road, current roughness and SN, and many other features that may vary or not during its lifetime. To formalize this environment and use that in the MDP model, the features were divided into variable and non-variable features during the pavement lifetime. The non-variable features are the features that show low or no changes during the design period, including the annual average climate parameters, road functional class, pavement type, pavement base type, and subgrade soil plasticity. The pavement's variable features are the pavement roughness, pavement structural number, pavement surface thickness, and traffic. The environment compromises infinite pavement states with variable and non-variable features and a series of M&R decisions as actions. This kind of environment cannot be implemented in simple MDP models, and exceptional cases of the MDP algorithm, such as deep Q-learning, should be applied. The MDP has a sequential frame, and the goal is to maximize the benefit/cost of M&R plans.

7.2. MDP States

The structural number and traffic volume are the features that are changing during the pavement lifetime and have a significant effect on the ride quality of the pavement, and together with the pavement roughness index, can be used in M&R policy optimization. The states in the developed MDP model have a vector of variable pavement features with respect to time. The states' vector can be defined as St =<IRIt, SNt, KESALt, CN> in which t denoted time up to the design period and IRIt, SNt, and KESALt are the predicted pavement roughness, structural number and traffic in the given time. CN is the construction number that will be set based on the applied M&R plan.

7.3. Transition Matrix

The developed MDP has a stochastic transition, and the developed XGBoost model will determine the resulting states from the given pair of the state and action. The developed XGBoost model gives the prediction of the IRI vs. time considering the pavement features. The model will implement the pavement features into the developed model and respond by distributing estimated IRI for the next year. The annual increase in the KESAL determines KESALt+1, and the SN deterioration rate determines the SNt+1.

7.4. MDP Actions

The considered actions in the MDP model is including M&R decisions. The actions are defined as 1-Do Nothing, 2- Maintenance, 3-Rehabilitation, and 4- reconstruction. By acting "Do nothing", the pavement will deteriorate, and next year state's vector will be determined by the XGBoost model, SN deterioration rate, and traffic increase rate. The maintenance action is the regular pavement maintenances, including the crack sealing, fog seal, slurry seal, patching that affects the ride quality and help reduce the rate of deterioration but does not affect the structural integrity of the pavement. The IRI will be updated the minimum thresholds by taking this action, but the SN deterioration will get just a discount factor. The resurfacing action is including full depth and partial milling and resurfacing, cold-in place resurfacing, and overlay. The roughness index will be updated to the minimum value by taking this action, and the SN number will get the maximum threshold. The maximum threshold of the SN will be estimated from the initial pavement condition and SN. The pavement will be set to a new pavement with the new feature by taking the reconstruction action. It will be assumed that the pavement will be reconstructed with the same layer thicknesses. In this case, the IRI and SN number will be updated to the min and max values, and the construction number will be set to zero. Table 7-1 shows the impact of taking actions on the MDP state features and updating the method at each transition.

Table 7-1- Impact of taking MDP actions on the MDP state features

Action	Impact on Features
Do Nothing	• IRI_{t+1} from XGBoost model • SN_{t+1} from SN deterioration rate • CN remains the same
Maintenance	• IRI_{t+1} will be set to the min value • SN_{t+1} from SN deterioration rate considering a discount factor • CN number will be updated
Rehabilitation	• IRI_{t+1} will be set to min • SNt+1 from SN deterioration rate considering a discount factor • CN number will be updated
Reconstruction	• IRI_{t+1} will be set to min • SN_{t+1} will be set to Max value • CN number will be set to 0

7.5. MDP Rewards

The benefit/cost ratio has been assigned as the reward function for the suggested MDP model. The optimal policy derived from the MDP model has the highest benefit/cost ratio between all the suggested M&R plans. IRI value of 3.5 m/km has been selected as the threshold for a poor pavement condition that needs immediate action (Smith et al. 2004). The area between the predicted IRI curve and the threshold value has been selected as a benefit of the M&R and has been assigned as AUPC in cost over benefit calculation. Agency cost of the M&R strategies depends on various parameters such as extension and material of treatment. Since the model is comparing the benefits/cost of strategies, the actual costs of the treatments are not required. An estimation of the relative costs of the treatment types has been presented in the literature(Wilde et al. 2014; Gu et al. 2019). Table 7-2 shows the estimated relative action cost of the treatment actions used in the MDP model.

Table 7-2- Relative cost of M&R actions

	Relative action cost
Do Nothing	0
Maintenance	1
Rehabilitation	3.2
Reconstruction	12

The user's cost was neglected in calculating the strategy costs due to uncertainties in the work zone type, the number of lanes, and the duration of work. The Equation (16) was used for calculating the benefit over the cost of the treatment strategies. These strategies

$$\text{Reward} = \frac{AUPC}{NPV} \tag{16}$$

Where:

$$NPV = \text{Initial Cost} + \sum_{k=1}^{n} TC_k * \left[\frac{1}{(1+i)^n} \right] \tag{17}$$

$AUPC =$ Area under the performance curve

TC = Treatment Cost

i = discount rate = %2

n = Analysis Period

The calculated reward for each strategy was used in the MDP model, and the best action at each state was determined.

7.6. MDP Q-function

In Q-learning, the expected total long-term reward in a given state s, and an action a is predicted by the Q-function Q(s, a). The agent should take the optimal action (s) for a state s such that the expected long-term reward is maximized. The Equation (12) gives the Q-function that needs to be solved for the given environment. Q-Learning algorithm is difficult to deal with large and continuous state space like our pavement environment. There are methods to deal with this problem, such as discretization techniques, but the implementation is complex, and the performance could be decreased. Another method is to use a neural network to solve the Q-function in the Q-learning. Deep Q-learning is an extension of the Q-Learning algorithm by modeling the Q-function (s, a) as a (deep) neural network. In this approach, the Q-function is a complicated composition of various parameterized functions, taking the input s, a, and making predictions of long-term utility. The deep Q-learning process were well explained in section 4.7.

The deep Q-learning Algorithm

Initialize input parameters of Q, replay memory, epsilon, episode number.

Loop for each episode in range *episodes*:

- Reset environment, step_size;
- Reset Rewards
- While the step is less than steps_size:
    - Choose action a:
        - With prob ϵ, random choose action
        - With prob(1-ϵ), chose a that maximize Q(s,a)
    - Apply action <a> and get the next state from XGBoost model, and reward from reward function
    - Add s_t, a, r_t, s_{t+1} to replay memory.
    - If enough sample is in replay memory:
        - Sample a batch of state transition.
        - Update Q through ANN and loss function of MSE
    - Decay ϵ
    - Update rewards
- End

End

Figure 7-1 - The deep Q-Learning algorithm used for solving the Q-function.

Figure 7-1 shows the algorithm used for solving the Q-function by the deep Q-learning method. First, the Q (s, a) will be initialized to 0 for all the states. The replay memory is an array of the temporary state and rewards that will be fed to the ANN as a batch of input observations. Epsilon is a trade-off parameter for exploration vs. exploitation rate explained in section 4.7. The number of episodes determined the simulation iteration and were set to 1000000. Step size in this problem is the duration of the M&R planning and was set to 20 years.

The objective of deep Q-learning is to make the iterative process and update the Q-function. The updating process will be stopped when the Q-function converges to a fixed value. The difference between the Q-function values at each iteration is an indicator of Q-function convergence. Figure 7-2 shows the normalized gradient of the Q-function, indicating convergence of the q-values to fixed values. By having the Q-function and using Equation (11), the optimum policy at each state can be determined.

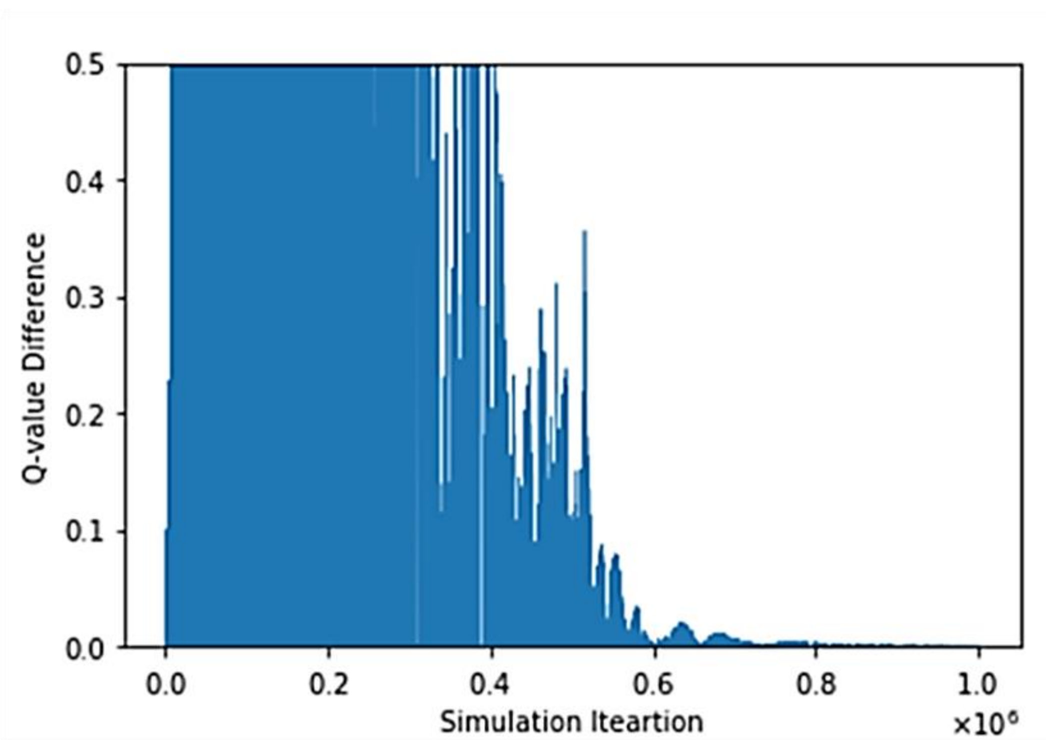


Figure 7-2- Convergence of Q-function gradient

7.7. Results and Discussion

After convergence of the Q-function, the optimum policy and best action at each state can be determined. Figure 7-3 shows the predicted IRI and optimum treatment type and time for LTPP section 0124. The blue region shows the area between the poor IRI condition threshold and predicted IRI. This area representing the benefit of the treatment plan and is equal to 41.185 (m/km) *(year). This area was used as AUPC in Equation (16). The present value of the treatment was calculated using Equation (17) and is equal to 24.3, and the total reward of the optimum policy is equal to 0.59. The MDP model is suggesting only one optimum policy, and alternative policies cannot be generated.

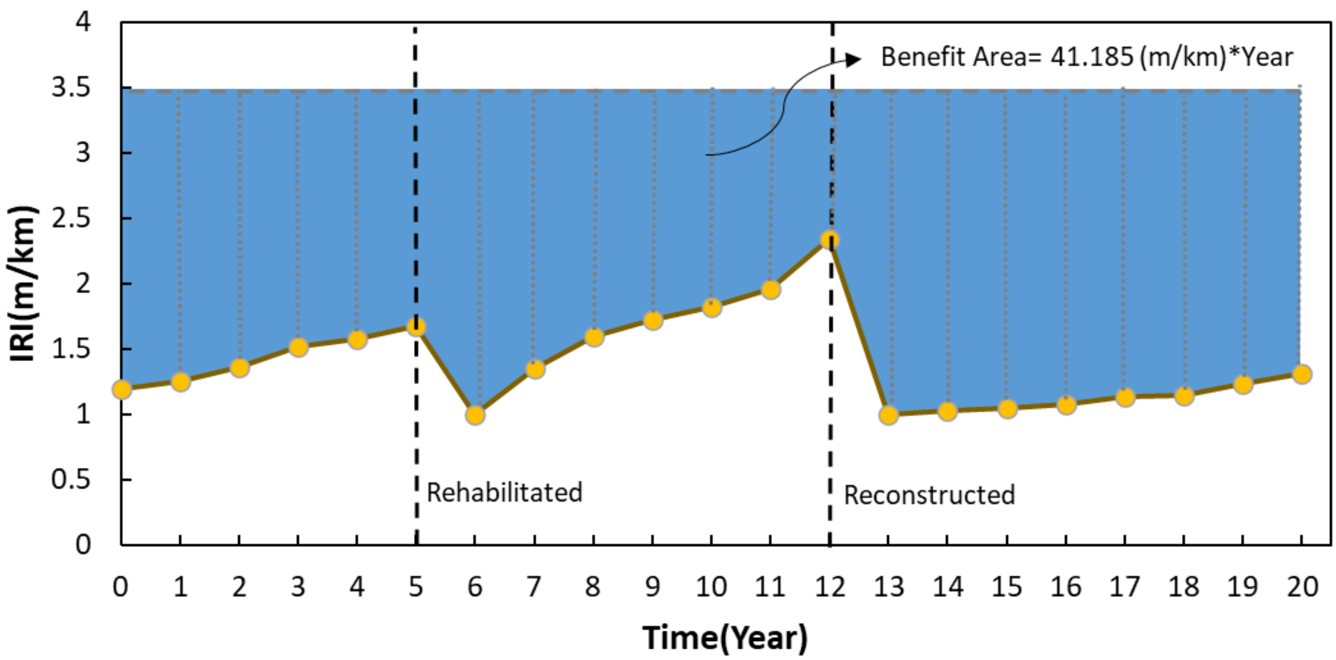


Figure 7-3 - Predicted IRI and treatment time for LTPP section 0124

The MDP model should be run for each section separately. This model needs to be trained based on the characteristics of each section and gives the optimum treatment for the given section. Figure 7-4 shows the MDP suggested treatment plans and actually placed treatments for selected

LTPP sections. The MDP model has mostly suggested rehabilitation as a treatment for the given sections. Information on the type of the sections and benefit over cost (B/C) ratio for treatments suggested by MDP and actually performed treatments were presented in Table 7-3. The M&R plan suggested for section 0124 includes a rehabilitation at year five and a reconstruction at year 12. The reason for reconstruction was poor SN that showed a defect in the structure of the pavement. The B/C for the treatment suggested by the MDP plans has a higher benefit than the performed plans. This also can be found for sections 0607 and 0118.

Table 7-3- MDP and LTPP M&R plan benefit/cost

LTPP section		Section 0607	Section 0118	Section 0124
Layer Information		16“ Unbound (granular)base 9" Bound treated base 3.8 “ HMAC	8” Lime Treated Subbase 3.6 “ Unbound (granular)base 8" Bound treated base 4.8 “ HMAC	6” Lime Treated Subbase 4 “ Bound (treated) Base 6.6 “ HMAC
Benefit/Cost of M&R Plan	MDP M&R	0.95	1.23	0.59
	LTPP M&R	0.91	1.12	0.23

The MDP model did not suggest maintenance because of the inaccuracy of the IRI reports by LTPP. In most of the LTPP sections, the observed pavement roughness after applying the maintenance increased. This could be due to the inaccuracy of the surveying or the inefficiency of the applied maintenances. The prediction model uses the provided data from LTPP and follows the given IRI patterns. The MDP model predicts an increase in IRI after taking maintenance as the action. This will reduce the area under the curve and benefit of the plan. Thus, the model will not select maintenance as the optimum treatment.

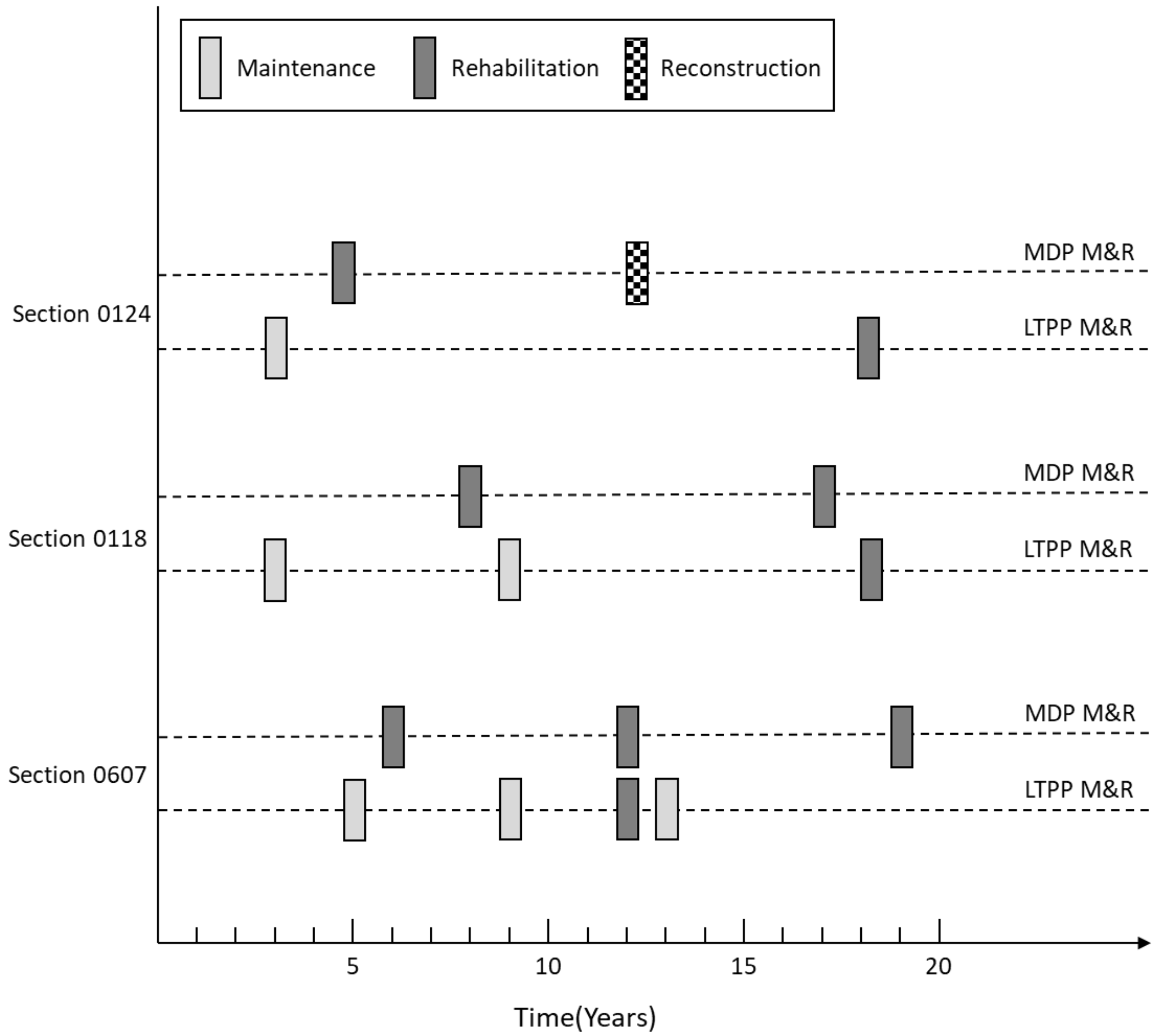


Figure 7-4- MDP suggested treatment plans and actually placed treatments for selected LTPP sections

The MDP model can suggest a plan for the treatment of the flexible pavements. The suggested plans are comparable with the actually performed treatments. The model can be improved by using more data, considering both users and agency costs, differentiating between the type of the rehabilitations, and customizing the costs based on the given treatments.

8. CONCLUSION AND RECOMMENDATION

This research study was aimed to enhance designing models for the flexible pavement of Oklahoma and develop a new M&R decision process using the surface roughness and structural capacity of the pavement section.

Pavement mechanistic-empirical (ME) design is one of the AAHSHTOWare Design software built to design new and rehabilitated pavements with flexible, rigid, and composite structures. The nationally calibrated performance models in Pavement ME do not well represent the construction and materials specifications, traffic, and climate conditions specific to each state and cannot precisely reflect the pavement performance. Local calibration of Pavement Mechanistic-Empirical (ME) software improved the pavement performance prediction models and optimized the performance models for the pavement network of Oklahoma. In chapter 3, the calibration effort of Pavement ME design prediction models using local inputs and performance data for the state of Oklahoma was presented. A total number of 66 sections from LTPP and few more asphalt pavement sections in Oklahoma were identified for this project's purpose. The selected projects represent flexible pavement construction practices in the state and cover various pavement conditions, construction age, and environmental conditions. The material, structural, and traffic data were gathered from LTPP, Oklahoma, and NCHRP datasets. The material input data were evaluated, and the most accurate available data was selected. For each pavement section, the Pavement ME design analysis was conducted. The predicted values from distress and IRI models were evaluated and compared with the measured ones, and the accuracy and bias of each model were determined. The nationally calibrated models show an improper prediction performance and a significant bias, which asserts the necessity of local calibration. The rutting and IRI models show better performance compared to fatigue bottom-up and top-down, and thermal cracking models. The reason for bias and error in the measured versus predicted distress values mainly comes from inaccurate input data, error in the distress survey, and accuracy of prediction models. By the calibration effort, the error in performance prediction models was reduced. The locally calibrated coefficients for distress and IRI models were determined for the Oklahoma pavement system. The distress and IRI models show that the calibrated coefficients improve Pavement ME predictions and the design of flexible pavements in Oklahoma. Rutting and IRI models show lower error and higher accuracy than the fatigue and transverse cracking models after the calibration. Fatigue cracking models underestimate fatigue distress before the calibration, but these models show better performance after the calibration. The results of Pavement ME transverse cracking are not consistent with the measured values in Oklahoma. Even after the calibration, the transverse cracking model cannot correctly predict the amount of cracking.

The second objective of this research was to develop a new maintenance and rehabilitation decision process which considers the stochasticity of the pavement performance prediction and suggests the optimized maintenance activities for the given section by implementing a newly developed predictive model. The developed M&R decision method is a Markov Decision Process that employs IRI from a newly developed IRI prediction model and structural number from historical data. The IRI prediction model predicts the IRI with high accuracy by having the structural number, road class, climate condition, traffic load, and subgrade and structural information. A total number of 419 flexible pavement sections from the states of Oklahoma, Texas, Arkansas, Missouri, Kansas, Colorado, and New Mexico were selected, and the input variables for the machine learning model were extracted from the selected sections. Several machine learning models, including GLM, SVM, ANN, MARS, and XGBoost, were developed and compared to predict the IRI. The XGBoost model shows an RMSE of 5.92 in/mile and the coefficient of determination of 95%. The model's prediction using the test dataset shows RMSE of 13.1 in/mile and R2 of 91%. This model shows excellent performance and can be used as a prediction model for predicting pavement performance during its lifetime.

By improving predictions and developing effective maintenance decision policies, machine learning algorithms can optimize maintenance and rehabilitation interventions and reduce maintenance costs. A Markov Decision Processing (MDP) model for suggesting the optimum treatment for the given flexible pavement section has been developed. This model considers the M&R activities from pavement history, which affects the pavement deterioration rate, and suggests an M&R Policy for the given pavement system. The MDP model can suggest a plan for the treatment of the flexible pavements. The suggested plans are comparable with the actually performed treatments. The model can be improved by using more data, considering both users and agency costs, differentiating between the rehabilitations, and customizing the costs based on the given treatments. The results of this planning can be used to recommend the optimal M&R treatment, which increases the benefit over the cost of the pavement. This model dynamically predicts the pavement's roughness during the desired time frame using the developed IRI prediction model.

REFERENCES

AASHTO. 1990. *AASHTO Guidelines for Pavement Management Systems*. Washington D.C: American Association of State Highway and Transportation Officials.

AASHTO. 1993. *AASHTO Guide for Design of Pavement Structures*. Washington, D.C: American Association of State Highway and Transportation Officials.

AASHTO. 2010. *Guide for the local calibration of the mechanistic-empirical pavement design guide*. Washington, DC: American Association of State Highway and Transportation Officials.

AASHTO. 2015. *Mechanistic-empirical pavement design guide: A manual of practice*. Second ed. Washington, DC: American Association of State Highway and Transportation Officials.

Abaza, Khaled. 2004. "Deterministic performance prediction model for rehabilitation and management of flexible pavement." *International Journal of Pavement Engineering - INT J PAVEMENT ENG* 5:111-21. doi: 10.1080/10298430412331286977.

Abaza, Khaled A. 2005. "Performance-Based Models for Flexible Pavement Structural Overlay Design." *Journal of Transportation Engineering* 131 (2):149-59. doi: doi:10.1061/(ASCE)0733-947X(2005)131:2(149).

Abdelaziz, Nader, Ragaa T. Abd El-Hakim, Sherif M. El-Badawy, and Hafez A. Afify. 2020. "International Roughness Index prediction model for flexible pavements." *International Journal of Pavement Engineering* 21 (1):88-99. doi: 10.1080/10298436.2018.1441414.

Al-Suleiman, Turki I., and Adnan M. S. Shiyab. 2003. "Prediction of Pavement Remaining Service Life Using Roughness Data—Case Study in Dubai." *International Journal of Pavement Engineering* 4 (2):121-9. doi: 10.1080/10298430310001634834.

Albuquerque, Fernando S, and Washington Peres Núñez. 2011. "Development of roughness prediction models for low-volume road networks in northeast Brazil." *Transportation Research Record* 2205 (1):198-205.

Alpaydın, Ethem. 2020. *Introduction to machine learning*: MIT press.

ARA, Inc., ERES Consultants Division. 2004. "Guide for Mechanistic–Empirical Design of new and rehabilitated pavement structures." In *NCHRP Project*. Washington, DC.

Attoh-Okine, Nii O, Stephen Mensah, and M Nawaiseh. 2003. A new technique for using multivariate adaptive regression splines (MARS) in pavement roughness prediction. Paper presented at the Proceedings of the Institution of Civil Engineers-Transport.

Attoh-Okine, Nii O., Ken Cooger, and Stephen Mensah. 2009. "Multivariate adaptive regression (MARS) and hinged hyperplanes (HHP) for doweled pavement performance modeling." *Construction and Building Materials* 23 (9):3020-3. doi: https://doi.org/10.1016/j.conbuildmat.2009.04.010.

Bellman, Richard. 1957. "A Markovian decision process." *Journal of mathematics and mechanics*:679-84.

BTS. 2019. "Pocket Guide to Transportation. ." In, edited by Bureau of Transportation Statistics. U.S Department of Transportation.

Butt, Abbas A., M. Y. Shahin, Samuel H. Carpenter, and James V. Carnahan. 1994. Application of Markov Process to Pavement Management Systems at Network Level.

Byrne, M., D. Albrecht, J.G. Sanjayan, and J. Kodikara. 2009. "Recognizing Patterns in Seasonal Variation of Pavement Roughness Using Minimum Message Length Inference." *Computer-Aided Civil and Infrastructure Engineering* 24 (2):120-9. doi: 10.1111/j.1467-8667.2008.00575.x.

Ceylan, Halil, Sunghwan Kim, Orhan Kaya, and Kasthurirangan Gopalakrishnan. 2015. "Investigation of AASHTOWare Pavement ME Design/Darwin-ME™ performance prediction models for Iowa pavement analysis and design." In. Iowa Department of Transportation, Ames, IA.

Chandra, Satish, Chalumuri Ravi Sekhar, Anish Kumar Bharti, and B. Kangadurai. 2013. "Relationship between Pavement Roughness and Distress Parameters for Indian Highways." *Journal of Transportation Engineering* 139 (5):467-75. doi: doi:10.1061/(ASCE)TE.1943-5436.0000512.

Chen, Tianqi, and Carlos Guestrin. 2016. "XGBoost: A Scalable Tree Boosting System." In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–94. San Francisco, California, USA: Association for Computing Machinery.

Cherkassky, Vladimir, and Yunqian Ma. 2004. "Practical selection of SVM parameters and noise estimation for SVM regression." *Neural networks* 17 (1):113-26.

Choi, Jae-ho, Teresa M. Adams, and Hussain U. Bahia. 2004. "Pavement Roughness Modeling Using Back-Propagation Neural Networks." *Computer-Aided Civil and Infrastructure Engineering* 19 (4):295-303. doi: 10.1111/j.1467-8667.2004.00356.x.

Chopra, Tanuj, Manoranjan Parida, Naveen Kwatra, and Jyoti Mandhani. 2017. "Development of pavement maintenance management system (PMMS) of urban road network using HDM-4 model." *International Journal Of Engineering & Applied Sciences* 9:14-. doi: 10.24107/ijeas.286473.

Cross, Stephen A, Robel Gibbe, and Nirajan Aryal. 2011. "Development of a flexible pavement database for local calibration of the MEPDG, Part 2 evaluation of ODOT SMA mixtures." In. Oklahoma Department of Transportation, Oklahoma City, OK.

Cross, Stephen A, Yatish Jakatimath, and Sumesh KC. 2007. "Determination of dynamic modulus master curves for Oklahoma HMA mixtures." In. Oklahoma Department of Transportation, Oklahoma City, OK.

Darter, Michael I, Leslie Titus-Glover, Harold Von Quintus, Biplab B Bhattacharya, and Jagannath Mallela. 2014. "Calibration and implementation of the AASHTO Mechanistic-Empirical Pavement Design Guide in Arizona." In. Arizona Department of Transportation.

De Veaux, Richard D, DC Psichogios, and LH Ungar. 1993. "A comparison of two nonparametric estimation schemes: MARS and neural networks." *Computers & chemical engineering* 17 (8):819-37.

Elhadidy, Amr A., Sherif M. El-Badawy, and Emad E. Elbeltagi. 2019. "A simplified pavement condition index regression model for pavement evaluation." *International Journal of Pavement Engineering*:1-10. doi: 10.1080/10298436.2019.1633579.

Farrar, Donald E., and Robert R. Glauber. 1967. "Multicollinearity in Regression Analysis: The Problem Revisited." *The Review of Economics and Statistics* 49 (1):92-107. doi: 10.2307/1937887.

Ferreira, Adelino, António Antunes, and Luı´s Picado-Santos. 2002. "Probabilistic Segment-linked Pavement Management Optimization Model." *Journal of Transportation Engineering* 128 (6):568-77. doi: doi:10.1061/(ASCE)0733-947X(2002)128:6(568).

FHWA. "LTPP InfoPave." https://infopave.fhwa.dot.gov/.

Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. 2001. *The elements of statistical learning*. Vol. 1: Springer series in statistics New York.

Fu, Wenjiang J. 1998. "Penalized Regressions: The Bridge versus the Lasso." *Journal of Computational and Graphical Statistics* 7 (3):397-416. doi: 10.1080/10618600.1998.10474784.

Fwa, Tien F., and Kumares C. Sinha. 1991. "Pavement Performance and Life-Cycle Cost Analysis." *Journal of Transportation Engineering* 117 (1):33-46. doi: doi:10.1061/(ASCE)0733-947X(1991)117:1(33).

Gedafa, D., M. Hossain, R. Miller, and D. Steele. 2010. "Network Level Testing for Pavement Structural Evaluation Using a Rolling Wheel Deflectometer." *Journal of Testing and Evaluation* 38 (4):439-48. doi: https://doi.org/10.1520/JTE102320.

Gedafa, Daba S., Mustaque Hossain, Richard Miller, and Thomas Van. 2014. "Network-level flexible pavement structural evaluation." *International Journal of Pavement Engineering* 15 (4):309-22. doi: 10.1080/10298436.2013.784766.

George, KP. 2000. "MDOT pavement management system: prediction models and feedback system." In.: Mississippi. Dept. of Transportation.

Gong, Hongren, Yiren Sun, and Baoshan Huang. 2019. "Gradient Boosted Models for Enhancing Fatigue Cracking Prediction in Mechanistic-Empirical Pavement Design Guide." *Journal of Transportation Engineering, Part B: Pavements* 145 (2):04019014. doi: doi:10.1061/JPEODX.0000121.

Gong, Hongren, Yiren Sun, Xiang Shu, and Baoshan Huang. 2018. "Use of random forests regression for predicting IRI of asphalt pavements." *Construction and Building Materials* 189:890-7. doi: https://doi.org/10.1016/j.conbuildmat.2018.09.017.

Gu, Fan, and Nam Tran. 2019. "Best Practices for Determining Life Cycle Costs of Asphalt Pavements." In.

Gulen, Sedat, Karen Zhu, John Weaver, Jie Shan, and William Flora. 2001. "Development of improved pavement performance prediction models for the indiana pavement management system." *Joint Transportation Research Program*:69.

Haas, R. C. G., W. Ronald Hudson, and John P. Zaniewski. 1994. "Modern pavement management."

Hafez, Marwan, Khaled Ksaibati, and Rebecca Atadero. 2019. "Best practices to support and improve pavement management systems for low-volume paved roads." *International Journal of Pavement Engineering* 20 (5):592-9. doi: 10.1080/10298436.2017.1316648.

Hall, Kevin D, Danny X Xiao, and Kelvin CP Wang. 2011. "Calibration of the Mechanistic–Empirical Pavement Design Guide for flexible pavement design in Arkansas." *J Transportation Research Record* 2226 (1):135-41.

Hossain, MI, LSP Gopisetti, and MS Miah. 2019. "International roughness index prediction of flexible pavements using neural networks." *Journal of Transportation Engineering, Part B: Pavements* 145 (1):04018058.

Hossain, Zahid, Musharraf Zaman, Curtis Doiron, and Steven Cross. 2011. "Development of flexible pavement database for local calibration of MEPDG." In. Oklahoma Department of Transportation, Oklahoma City, OK.

InfoPave, LTPP. https://infopave.fhwa.dot.gov/.
Jannat, Gulfam, X. X. Yuan, and Medhat Shehata. 2014. "Development of Regression Equation for Local Calibration of MEPDG Rutting and IRI using Ontario's Long Term PMS Data." *International Journal of Pavement Engineering*. doi: 10.1080/10298436.2014.973024.
Jorge, Diana, and Adelino Ferreira. 2012. "Road network pavement maintenance optimisation using the HDM-4 pavement performance prediction models." *International Journal of Pavement Engineering* 13 (1):39-51. doi: 10.1080/10298436.2011.563851.
Kargah-Ostadi, Nima. 2014. "Comparison of Machine Learning Techniques for Developing Performance Prediction Models." In *Computing in Civil and Building Engineering (2014)*, 1222-9.
Kargah-Ostadi, Nima, Shelley M. Stoffels, and Nader Tabatabaee. 2010. "Network-Level pavement roughness prediction model for rehabilitation recommendations." *Transportation Research Record* 2155 (1):124-33. doi: 10.3141/2155-14.
Ke-zhen, Yan, Huarong Liao, Honghui Yin, and Likui Huang. 2011. "Predicting the Pavement Serviceability Ratio of Flexible Pavement with Support Vector Machines." In *Road Pavement and Material Characterization, Modeling, and Maintenance*, 24-32.
Kelly, Gregory, Deborah Delaney, Gary Chai, and Sherif Mohamed. 2016. "Optimising local council's return on investment from annual pavement rehabilitation budgets through targeting of the average pavement condition index." *Journal of Traffic and Transportation Engineering (English Edition)* 3 (5):465-74. doi: https://doi.org/10.1016/j.jtte.2016.09.008.
Khattak, Mohammad Jamal, Mohammad Abdullah Nur, Mohammad Reza-Ul-Karim Bhuyan, and Kevin Gaspard. 2014. "International roughness index models for HMA overlay treatment of flexible and composite pavements." *International Journal of Pavement Engineering* 15 (4):334-44. doi: 10.1080/10298436.2013.842237.
Kim, Sunghwan, Halil Ceylan, Di Ma, and Kasthurirangan Gopalakrishnan. 2014. "Calibration of pavement ME Design and Mechanistic-Empirical Pavement Design Guide performance prediction models for Iowa pavement systems." *Journal of Transportation Engineering* 140 (10):04014052.
Kim, Y Richard, Fadi M Jadoun, Tian Hou, and Naresh Muthadi. 2011. "Local calibration of the MEPDG for flexible pavement design." In. North Carolina State University. Dept. of Civil, Construction, and Environmental Engineering.
Kırbaş, Ufuk, and Mustafa Karaşahin. 2016. "Performance models for hot mix asphalt pavements in urban roads." *Construction and Building Materials* 116:281-8. doi: https://doi.org/10.1016/j.conbuildmat.2016.04.118.
Kumar, Aditya. "XGBoost Algorithm: The New King." https://medium.com/@adityakumar24jun/xgboost-algorithm-the-new-king-c4a64ea677bf.
La Torre, Francesca, Lorenzo Domenichini, and Michael I Darter. 1998. Roughness prediction model based on the artificial neural network approach. Paper presented at the Fourth International Conference on Managing Pavements.
Li, Jianhua, Linda M Pierce, and Jeff Uhlmeyer. 2009. "Calibration of flexible pavement in Mechanistic–Empirical Pavement Design Guide for Washington state." *Journal of the Transportation Research Record* 2095 (1):73-83.

Li, Qiang, Danny X Xiao, Kelvin CP Wang, Kevin D Hall, and Yanjun Qiu. 2011. "Mechanistic-Empirical Pavement Design Guide (MEPDG): a bird's-eye view." *Journal of Modern Transportation* 19 (2):114-33.

Lin, Jyh-Dong, Liang-Hao Hsiao, and Graduate Student. 2003. Correlation Analysis Between International Roughness Index (IRI) and Pavement Distress by Neural Network.

Ling, Meng, Yu Chen, Sheng Hu, Xue Luo, and Robert L. Lytton. 2019. "Enhanced model for thermally induced transverse cracking of asphalt pavements." *Construction and Building Materials* 206:130-9. doi: https://doi.org/10.1016/j.conbuildmat.2019.01.179.

Liu, Litao, and Nasir G. Gharaibeh. 2014. "Bayesian Model for Predicting the Performance of Pavements Treated with Thin Hot-Mix Asphalt Overlays." *Transportation Research Record* 2431 (1):33-41. doi: 10.3141/2431-05.

Lou, Z., M. Gunaratne, J. J. Lu, and B. Dietrich. 2001. "Application of Neural Network Model to Forecast Short-Term Pavement Crack Condition: Florida Case Study." *Journal of Infrastructure Systems* 7 (4):166-71. doi: doi:10.1061/(ASCE)1076-0342(2001)7:4(166).

Mallela, Jagannath, Leslie Titus-Glover, Suri Sadasivam, Biplab Bhattacharya, Michael Darter, and Harold Von Quintus. 2013. "Implementation of the AASHTO mechanistic-empirical pavement design guide for Colorado." In. Colorado. Dept. of Transportation. Research Branch, Denver, CO: Colorado. Dept. of Transportation. Research Branch.

Marcelino, Pedro, Maria de Lurdes Antunes, Eduardo Fortunato, and Marta Castilho Gomes. 2019. "Machine learning approach for pavement performance prediction." *International Journal of Pavement Engineering*:1-14. doi: 10.1080/10298436.2019.1609673.

Mazari, Mehran, and Daniel D Rodriguez. 2016. "Prediction of pavement roughness using a hybrid gene expression programming-neural network technique." *Journal of Traffic and Transportation Engineering (English Edition)* 3 (5):448-55.

Mbwana, John R, and Mark A Turnquist. 1996. "Optimization modeling for enhanced network-level pavement management system." *Transportation Research Record* 1524 (1):76-85.

Narh-Dometey, Anita. 2019. "A Markovian Optimization Model for Pavement Maintenance Using Policy Iteration Algorithm with Discounted Road-User and Agency Costs." Morgan State University.

Nelder, J. A., and R. W. M. Wedderburn. 1972. "Generalized Linear Models." *Journal of the Royal Statistical Society: Series A (General)* 135 (3):370-84. doi: 10.2307/2344614.

Nobakht, Mona, Maryam S. Sakhaeifar, and David E. Newcomb. 2017. "Selection of structural overlays using asphalt mixture performance." *Journal of Materials in Civil Engineering* 29 (11):04017209. doi: 10.1061/(ASCE)MT.1943-5533.0002070.

Nobakht, Mona, Maryam S. Sakhaeifar, David E. Newcomb, and Shane Underwood. 2018. "Mechanistic-Empirical methodology for the selection of cost-effective rehabilitation strategy for flexible pavements." *International Journal of Pavement Engineering* 19 (8):675-84. doi: 10.1080/10298436.2016.1199878.

Nwanganga, Fred, and Mike Chapple. 2020. *Practical Machine Learning in R*: Indianapolis, Indiana : John Wiley & Sons, Incorporated.

ODOT. 2019. "Update on Oklahoma Bridges & Highways." In. Oklahoma, US: Oklahoma Department of Transportation.

Osorio-Lird, Aleli, Alondra Chamorro, Carlos Videla, Susan Tighe, and Cristina Torres-Machi. 2018. "Application of Markov chains and Monte Carlo simulations for developing pavement performance models for urban network management." *Structure and Infrastructure Engineering* 14 (9):1169-81. doi: 10.1080/15732479.2017.1402064.

Pei, Lili, Zhaoyun Sun, Ting Yu, Wei Li, Xueli Hao, Yuanjiao Hu, and Chunmei Yang. 2020. "Pavement aggregate shape classification based on extreme gradient boosting." *Construction and Building Materials* 256:119356. doi: https://doi.org/10.1016/j.conbuildmat.2020.119356.

Perera, RW, and SD Kohn. 2002. "Issues in pavement smoothness." *Transportation Research Board, Washington, DC*.

Prozzi, J. A., and S. M. Madanat. 2004. "Development of Pavement Performance Models by Combining Experimental and Field Data." *Journal of Infrastructure Systems* 10 (1):9-22. doi: doi:10.1061/(ASCE)1076-0342(2004)10:1(9).

Prozzi, Jorge A., and Samer M. Madanat. 2000. "Using Duration Models to Analyze Experimental Pavement Failure Data." *Transportation Research Record* 1699 (1):87-94. doi: 10.3141/1699-12.

Robbins, Mary M, Carolina Rodezno, Nam Tran, and D Timm. 2017. "Pavement ME Design–A summary of local calibration efforts for flexible pavements." In. National Center for Asphalt Technology.

Roberts, Craig A., and Nii O. Attoh-Okine. 1998. "A Comparative Analysis of Two Artificial Neural Networks Using Pavement Performance Prediction." *Computer-Aided Civil and Infrastructure Engineering* 13 (5):339-48. doi: 10.1111/0885-9507.00112.

Rohde, Gustav T. 1994. "Determining pavement structural number from FWD testing." *Transportation Research Record* (1448).

Sakhaeifar, Maryam S, Y Richard Kim, and Pooyan Kabir. 2015. "New predictive models for the dynamic modulus of hot mix asphalt." *Journal of Construction Building Materials* 76:221-31.

Sakhaeifar, Maryam S., David Newcomb, Mona Nobakht, B.Shane Underwood, Padmini P. Gudipudi, and Jeff Stempihar. 2015. "Selection of long lasting rehabilitation treatment using life cycle cost analysis and present serviceability rating." In *ODOT SP&R Item Number 2261*. Oklahoma Department of Transportation, Materials and Research Division Oklahoma City, OK.

Sakhaeifar, Maryam S., Y. Richard Kim, and Pooyan Kabir. 2015. "New predictive models for the dynamic modulus of hot mix asphalt." *Construction and Building Materials* 76:221-31. doi: https://doi.org/10.1016/j.conbuildmat.2014.11.011.

Saliminejad, Siamak, and Nasir G. Gharaibeh. 2013. "Impact of Error in Pavement Condition Data on the Output of Network-Level Pavement Management Systems." *Transportation Research Record* 2366 (1):110-9. doi: 10.3141/2366-13.

Santos, João, and Adelino Ferreira. 2013. "Life-cycle cost analysis system for pavement management at project level." *International Journal of Pavement Engineering* 14 (1):71-84. doi: 10.1080/10298436.2011.618535.

Schwartz, Charles W, Rui Li, Sunghwan Kim, Halil Ceylan, and Kasthurirangan Gopalakrishnan. 2011. "Sensitivity evaluation of MEPDG performance prediction." In. Transportation Research Board of the National Academies.

Smith, James T., and Susan L. Tighe. 2004. "Assessment of Overlay Roughness in Long-Term Pavement Performance Test Sites: Canadian Case Study." *Transportation Research Record* 1869 (1):126-35. doi: 10.3141/1869-15.

Smith, Roger, Thomas Freeman, and Olga Pendleton. 1993. "Pavement maintenance effectiveness." In.: Strategic Highway Research Program, National Research Council Washington, DC.

Sun, Lu. 2003. "Simulation of pavement roughness and IRI based on power spectral density." *Math. Comput. Simul.* 61 (2):77–88. doi: 10.1016/s0378-4754(01)00386-x.
Sutton, Richard S, and Andrew G Barto. 2018. *Reinforcement learning: An introduction*: MIT press.
Wang, Kelvin C.P., and Qiang Li. 2011. "Pavement Smoothness Prediction Based on Fuzzy and Gray Theories." *Computer-Aided Civil and Infrastructure Engineering* 26 (1):69-76. doi: 10.1111/j.1467-8667.2009.00639.x.
Wang, Kelvin CP, Qiang Joshua Li, Stephen A Cross, and Jeff Dean. 2014. Pavement ME Design distress modeling for Oklahoma. Paper presented at the T&DI Congress: Planes, Trains, and Automobiles, Orlando, Florida.
Wilde, W., L. Thompson, and T. J. Wood. 2014. Cost-Effective Pavement Preservation Solutions for the Real World.
Williams, R Chris, and R Shaidur. 2013. "Mechanistic-Empirical Pavement Design Guide calibration for pavement rehabilitation." In.: Oregon. Dept. of Transportation. Research Section, Salem, Ore. and Federal Highway Administration, Washington D.C.
Wimsatt, Andrew J. 1999. "Direct Analysis Methods for Falling Weight Deflectometer Deflection Data." *Transportation Research Record* 1655 (1):135-44. doi: 10.3141/1655-18.
Yan, Ke Zhen, and Zou Zhang. 2011. "Research in Analysis of Asphalt Pavement Performance Evaluation Based on PSO-SVM." *Applied Mechanics and Materials* 97-98:203-7. doi: 10.4028/www.scientific.net/AMM.97-98.203.
Yang, Jidong, Jian John Lu, Manjriker Gunaratne, and Qiaojun Xiang. 2003. "Forecasting overall pavement condition with neural networks: Application on Florida highway network." *Transportation Research Record* 1853 (1):3-12. doi: 10.3141/1853-01.
Zapata, Claudia, and Carlos E Cary. 2012. "Integrating the national database of subgrade soil-water characteristic curves and soil index properties with the MEPDG." *National Cooperative Highway Research Program Project*.
Zhang, Miaomiao, Hongren Gong, Xiaoyang Jia, Rui Xiao, Xi Jiang, Yuetan Ma, and Baoshan Huang. 2020. "Analysis of critical factors to asphalt overlay performance using gradient boosted models." *Construction and Building Materials* 262:120083.
Zhou, Guoqing, and Linbing Wang. 2008. "Analysis of Flexible Pavement Distresses on IRI Model." In *Pavements and Materials*, 150-60.
Ziari, Hasan, Mojtaba Maghrebi, Jalal Ayoubinejad, and S. Travis Waller. 2016a. "Prediction of Pavement Performance: Application of Support Vector Regression with Different Kernels." *Transportation Research Record* 2589 (1):135-45. doi: 10.3141/2589-15.
Ziari, Hasan, Jafar Sobhani, Jalal Ayoubinejad, and Timo Hartmann. 2016b. "Prediction of IRI in short and long terms for flexible pavements: ANN and GMDH methods." *International Journal of Pavement Engineering* 17 (9):776-88. doi: 10.1080/10298436.2015.1019498.
Ziari, Hasan, Jafar Sobhani, Jalal Ayoubinejad, and Timo Hartmann. 2016c. "Analysing the accuracy of pavement performance models in the short and long terms: GMDH and ANFIS methods." *Road Materials and Pavement Design* 17 (3):619-37. doi: 10.1080/14680629.2015.1108218.
Zou, Hui, and Trevor Hastie. 2005. "Regularization and variable selection via the elastic net." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67 (2):301-20. doi: 10.1111/j.1467-9868.2005.00503.x.