# STOCHASTIC PERFORMANCE AND MAINTENANCE OPTIMIZATION MODELS FOR PAVEMENT INFRASTRUCTURE MANAGEMENT

by

**Mohamed S. Yamany**

**A Dissertation**

*Submitted to the Faculty of Purdue University*

*In Partial Fulfillment of the Requirements for the degree of*

**Doctor of Philosophy**

Lyles School of Civil Engineering

West Lafayette, Indiana

May 2020

# THE PURDUE UNIVERSITY GRADUATE SCHOOL
# STATEMENT OF COMMITTEE APPROVAL

**Dr. Dulcy M. Abraham, Chair**

Lyles School of Civil Engineering

**Dr. Samuel Labi**

Lyles School of Civil Engineering

**Dr. Mario Ventresca**

School of Industrial Engineering

**Dr. Tommy E. Nantung**

Indiana Department of Transportation

**Approved by:**

Dr. Dulcy M. Abraham

***To Allah,***

***To My Parents,***

***My Wife Nashwa, and***

***My Children Abdelrahman,***

***Habibah and Mazen.***

# ACKNOWLEDGMENTS

My highest gratitude goes to the Lord, Allah, for bestowing on me his grace during my Ph.D. studies and for making that wonderful journey one of spiritual growth as well.

I would like to thank Professor Dulcy M. Abraham, my advisor for her endless patience, guidance, and encouragement throughout my graduate studies at Purdue University.

I would also like to thank my committee members: Professor Samuel Labi, Professor Mareo Ventresca, and Dr. Tommy E. Nantug for their constant support and guidance throughout my Ph.D. program. I am grateful for the time taken to help me accomplish this work, as well as the support in navigating graduate school and starting my career. I would also like to thank Professor Amr A. Kandil.

I would like to acknowledge the Egyptian Government Graduate Research Fellowship that funded me for my Ph.D. The contents of this dissertation reflect the views of the author, who is responsible for the facts and the accuracy of the data presented herein. The contents do not reflect the official views or policies of the Egyptian Government.

I would like to thank and acknowledge Dr. Rebecca McDaniel - Technical Director, North Central Superpave Center at Purdue University, Todd Shields - Statewide Asset Management Engineer, Indiana Department of Transportation, and Nick Cosenza – Pavement Design Engineer, INDOT for their assistance in deploying the survey questionnaires of this research. In addition, I would like to thank the individuals from State Transportation Agencies across the U.S. for participating in the surveys.

I would like to thank my wonderful friends here at Purdue University. I am also grateful for the support of my parents, wife Nashwa, and children Abdelrahman, Habibah and Mazen, who continue to push me to reach my full potential.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| AADT | Annual Average Daily Traffic |
| AADTT | Annual Average Daily Truck Traffic |
| AAFI | Annual Average Freezing Index |
| AAP | Annual Average Precipitation |
| AASHTO | American Association of State Highway and Transportation Officials |
| AAT | Annual Average Temperature |
| AI | Artificial Intelligence |
| ANNs | Artificial Neural Networks |
| ASCE | American Society of Civil Engineers |
| CCI | Composite Condition Index |
| CI | Crack Index |
| DOT | Department of Transportation |
| DR | Distress Rating |
| ESALs | Equivalent Single Axle Loads |
| FHWA | Federal Highway Administration |
| GA | Genetic Algorithm |
| HMA | Hot Mix Asphalt |
| HMMs | Hidden Markov Models |
| INDOT | Indiana Department of Transportation |
| IRI | International Roughness Index |
| LDC | Length of Duty Cycle |
| LTPP | Long-Term Pavement Performance |
| M&R | Maintenance and Rehabilitation |
| MAPE | Mean Absolute Percent Error |
| MCR | Micro-Surfacing |
| MLE | Maximum Likelihood Estimate |
| MOE | Measure of Effectiveness |
| MOGA | Multi-Objective Genetic Algorithm |

| | |
|---|---|
| NCS | Number of Condition States |
| NHCCI | National Highway Construction Cost Index |
| OLS | Ordinary Least Squares |
| OVR | Thin HMA Overlay |
| PCC | Portland Cement Concrete |
| PCI | Pavement Condition Index |
| PCR | Pavement Condition Rating |
| PI | Performance Indicator |
| PJ | Performance Jump |
| PM | Preventive Maintenance |
| PSI | Pavement Serviceability Index |
| PSR | Present Serviceability Rating |
| $R^2$ | Coefficient of Determination |
| RMSE | Root Mean Square Error |
| RUT | Rutting |
| SHA | State Highway Agency |
| SL | Service Life |
| SMEs | Subject Matter Experts |
| STAs | State Transportation Agencies |
| TPM | Transition Probability Matrix |
| UTBWC | Ultra-Thin Bonded Wearing Course |

# ABSTRACT

Highway infrastructure, including roads/pavements, contributes significantly to a country's economic growth, quality of life improvement, and negative environmental impacts. Hence, highway agencies strive to make efficient and effective use of their limited funding to maintain their pavement infrastructure in good structural and functional conditions. This necessitates predicting pavement performance and scheduling maintenance interventions accurately and reliably by using appropriate performance modeling and maintenance optimization methodologies, while considering the impact of influential variables and the uncertainty inherent in pavement condition data.

Despite the enormous research efforts toward stochastic pavement performance modeling and maintenance optimization, several research gaps still exist. Prior research has not provided a synthesis of Markovian models and their associated methodologies that could assist researchers and highway agencies in selecting the Markov methodology that is appropriate for use with the data available to the agency. In addition, past Markovian pavement performance models did not adequately account for the marginal effects of the preventive maintenance (PM) treatments due to the lack of historical PM data, resulting in potentially unreliable models. The primary components of a Markov model are the transition probability matrix, number of condition states (NCS), and length of duty cycle (LDC). Previous Markovian pavement performance models were developed using NCS and LDC based on data availability, pavement condition indicator and data collection frequency. However, the selection of NCS and LDC should also be based on producing pavement performance models with high levels of prediction accuracy. Prior stochastic pavement maintenance optimization models account for the uncertainty of the budget allocated to pavement preservation at the network level. Nevertheless, variables such as pavement condition deterioration and improvement that are also associated with uncertainty, were not included in stochastic optimization models due to the expected large size of the optimization problem.

The overarching goal of this dissertation is to contribute to filling these research gaps with a view to improving pavement management systems, helping to predict probabilistic pavement performance and schedule pavement preventive maintenance accurately and reliably. This study

reviews Markovian pavement performance models using various Markov methodologies and transition probabilities estimation methods, presents a critical analysis of the different aspects of Markovian models as applied in the literature, reveals gaps in knowledge, and offers suggestions for bridging those gaps. This dissertation develops a decision tree which could be used by researchers and highway agencies to select appropriate Markov methodologies to model pavement performance under different conditions of data availability. The lack of consideration of pavement PM impacts into probabilistic pavement performance models due to absence of historical PM data may result in erroneous and often biased pavement condition predictions, leading to non-optimal pavement maintenance decisions. Hence, this research introduces and validates a hybrid approach to incorporate the impact of PM into probabilistic pavement performance models when historical PM data are limited or absent. The types of PM treatments and their times of application are estimated using two approaches: (1) Analysis of the state of practice of pavement maintenance through literature and expert surveys, and (2) Detection of PM times from probabilistic pavement performance curves. Using a newly developed optimization algorithm, the estimated times and types of PM treatments are integrated into pavement condition data. A non-homogeneous Markovian pavement performance model is developed by estimating the transition probabilities of pavement condition using the ordered-probit method. The developed hybrid approach and performance models are validated using cross-validation with out-of-sample data and through surveys of subject matter experts in pavement engineering and management. The results show that the hybrid approach and models developed can predict probabilistic pavement condition incorporating PM effects with an accuracy of 87%.

The key Markov chain methodologies, namely, homogeneous, staged-homogeneous, non-homogeneous, semi- and hidden Markov, have been used to develop stochastic pavement performance models. This dissertation hypothesizes that the NCS and LDC significantly influence the prediction accuracy of Markov models and that the nature of such influence varies across the different Markov methodologies. As such, this study develops and compares the Markovian pavement performance models using empirical data and investigates the sensitivity of Markovian model prediction accuracy to the NCS and LDC. The results indicate that the semi-Markov is generally statistically superior to the homogeneous and staged-homogeneous Markov (except in a few cases of NCS and LDC combinations) and that Markovian model prediction accuracy is significantly sensitive to the NCS and LDC: an increase in NCS improves the prediction accuracy until a certain NCS threshold after which the accuracy decreases, plausibly due to data overfitting. In addition, an increase in LDC improves the prediction accuracy when the NCS is small.

Scheduling pavement maintenance at road network level without considering the uncertainty of pavement condition deterioration and improvement over the long-term (typically, pavement design life) likely results in mistiming maintenance applications and less optimal decisions. Hence, this dissertation develops stochastic pavement maintenance optimization models that account for the uncertainty of pavement condition deterioration and improvement as well as the budget constraint. The objectives of the stochastic optimization models are to minimize the overall deterioration of road network condition while minimizing the total maintenance cost of the road network over a 20-year planning horizon (typical pavement design life). Multi-objective Genetic Algorithm (MOGA) is used because of its robust search capabilities, which lead to global optimal solutions. In order to reduce the number of combinations of solutions of stochastic MOGA models, three approaches are proposed and applied: (1) using PM treatments that are most commonly used by highway agencies, (2) clustering pavement sections based on their ages, and (3) creating a filtering constraint that applies a rest period after treatment applications. The results of the stochastic MOGA models show that the Pareto optimal solutions change significantly when the uncertainty of pavement condition deterioration and improvement is included.

# CHAPTER 1. INTRODUCTION

Highway infrastructure, including roads/pavements, contributes significantly to the economic growth of countries, the improvement of life quality, and the impacts on environment. Hence, highway agencies are striving to manage their infrastructure assets efficiently in order to maintain them in good functional and structural conditions while utilizing their limited resources optimally.

## 1.1. Motivation

Roads provide mobility, connectivity and accessibility throughout countries (Labi et al. 2019). In the United States, for example, in 2017 roads were expanded to over 4 million miles, comprising various functional classes, e.g., interstates, non-interstate, urban and rural and ranging from multi-lane interstates to residential streets. In 2019, the share of the transport sector to the U.S. Gross Domestic Product (GDP) was estimated as $2,249.4 billion (10.71% of GDP) ("Trading Economics, United States" 2019). Roads contribute significantly to the U.S. transport and is expected to contribute more in the future. The number of travelled lane miles in 2017 was found to be 3.2 trillion, which is more than 300 round trips between Earth and Pluto (ASCE Infrastructure Report Card 2017). Freight demand is expected to increase from 17 billion tons in 2012 to 25.3 billion tons in 2045 (FHWA 2018), which requires more expanded roads in good condition.

The ASCE Infrastructure Report Card (2017) indicated that the roads in the U.S. are generally in "poor" condition, which translates on the Pavement Condition Index (PCI) scale into an average score of 3 out of 10. A total of $121 billion (ASCE Infrastructure Report Card 2017) or $300 per American driver (Wright 2016) were estimated as extra annual expenses on vehicle operation and repair due to poor condition of pavements. Additionally, more than 35,000 people were killed in vehicle crashes in 2015 (ASCE Infrastructure Report Card 2017) for several reasons, among them is the poor condition of pavements (Chen et al. 2019). However, a spending of $1 on highways (roads and bridges) maintenance and improvement could return $5.20 in the form of lower vehicle maintenance costs, decreased delays, reduced fuel consumption, improved safety, lower road and bridge maintenance costs, and reduced emissions as a result of improved traffic flow (ASCE Infrastructure Report Card 2017).

The U.S. public spending on highways, provided by federal, state and local governments, covers capital and operation and maintenance expenditures. The overall public spending on highways dropped from almost 1.8% of GDP in the 1950s to about 0.9% of GDP in the 1980s, when it has plateaued until 2017 (Office of Management and Budget 2018). Capital spending on highways has been decreasing since 2003, while operation and maintenance spending on highways has flattened. Although the federal spending has declined recently in 2017, the state and local spending has increased marginally (Office of Management and Budget 2018). Hence, the U.S. highways are suffering from chronic underfunding (ASCE Infrastructure Report Card 2017).

The significant role that roads play (connectivity, accessibility and mobility) necessitates maintaining roads in good functional and structural conditions; however, the ASCE Infrastructure Report Card (2017) announced that the roads in the U.S. are in “poor” condition resulting in several implications such as an annual additional cost of $121 billion for vehicle operations and repair. Consequently, U.S. roads are in a profound need to be repaired and upgraded, with chronic underfunding in addition to a mammoth backlog of $420 billion in highway maintenance (ASCE Infrastructure Report Card 2017).

## 1.2. Problem Statement

Highway agencies need an efficient pavement management system to effectively allocate their limited resources to optimally selected projects at optimal times and to the most cost-effective maintenance treatments. An efficient pavement management system should assist in predicting future pavement condition incorporating the positive and negative effects of relevant influential factors, while accounting for the uncertainty inherently attributed to pavement condition data. Additionally, it should support the optimization of pavement maintenance and rehabilitation treatments taking into account the stochasticity and randomness associated with the data and the decision variables considered in optimization models.

Despite the enormous efforts made in previous research on probabilistic modeling of pavement performance and optimization of pavement maintenance and rehabilitation (M&R), the relevant body of knowledge and body of practice still have gaps. Markov chains process has been extensively used for probabilistic modeling of pavement performance; however, the literature lacks a synthesis of Markovian pavement performance methodologies and models that would help researchers and highway agencies select the appropriate Markov methodology for their data. Prior Markovian pavement performance models did not adequately account for the marginal effects of the preventive maintenance (PM) treatments owing to data absence and limitation. The consideration of PM impacts into Markovian pavement performance models is of paramount importance in the decision-making of M&R. Moreover, previous Markov models were developed for pavement performance using a number of condition states (NCS) and a length of duty cycle (LDC) (two main components of Markov chain models) on the basis of data availability, pavement condition indicator and data collection frequency. Nevertheless, the selection of the NCS and LDC should be also based on the resulting prediction accuracy of pavement performance.

Stochastic optimization of pavement maintenance considers the uncertainty inherently attributed to pavement condition data and decision-making variables when scheduling maintenance interventions for a road network. In the literature, stochastic optimization models developed for pavement maintenance account for the uncertainty of budget constraint only. Other variables, such as pavement condition deterioration and improvement, are also associated with uncertainty, but were not adequately considered in stochastic optimization models due to the expected large number of combinations of solutions.

## 1.3. Research Questions

The dissertation contributes to solving the aforementioned problem and bridging the research gaps mentioned above. The research questions answered by this dissertation fall into two categories: modeling of pavement performance and optimization of pavement preventive maintenance.

**Modeling of Pavement Performance:** What are the appropriate types and methodologies of Markov chains for modeling pavement performance? How can the impacts of preventive maintenance be incorporated into probabilistic pavement performance models? What is the statistical significance of the number of condition states (NCS) and the length of duty cycle (LDC) for the prediction accuracy of Markovian pavement performance models, and for the selection of Markovian methodology and model types?

**Optimization of Pavement Preventive Maintenance:** What is the effect of considering the uncertainty of pavement condition deterioration and improvement on the decision-making of pavement maintenance? How can the computational complexity in terms of the number of combinations of solutions, be reduced for stochastic Multi-objective Genetic Algorithm (MOGA) models used for road network preventive maintenance optimization?

## 1.4. Research Objectives

The overarching goal of the current research is to develop methodologies and pavement performance and optimization models that will enhance pavement management systems and assist in accurately predicting future pavement condition and effectively allocating the limited resources of highway agencies to maintain pavements in a desired condition. The main objective of this research is to provide highway agencies with a stochastic pavement management system capable of predicting stochastic pavement condition over the long term and constructing pavement maintenance schedules that consider the uncertainty associated with decision-making variables. In the course of achieving this main objective, four research objectives have been developed as follows:

1- Synthesize the literature on Markovian pavement performance models. Develop a decision tree to be used for selecting the appropriate Markov methodology and its design parameters for pavement performance modeling.
2- Develop and validate a hybrid approach to incorporate the impacts of preventive maintenance (PM) into probabilistic pavement performance models. Create and validate a non-homogenous Markov model to predict probabilistic pavement condition, considering the influence of preventive maintenance.
3- Compare different Markovian methodologies and models along with various combinations of number of condition states (NCS) and length of duty cycle (LDC) for pavement performance. Identify the statistical significance of the NCS and LDC for the prediction accuracy of probabilistic pavement condition using Markov chains.
4- Develop stochastic pavement preventive maintenance optimization models at the road network level. Develop and implement approaches to reduce the number of combinations of solutions of the proposed stochastic optimization models.

## 1.5. Research Overview

Figure 1.1 outlines the general framework for the four research objectives of the current research. First, as in previous research, Markovian pavement performance methodologies and models have been critically reviewed and synthesized to summarize the state-of-the-art Markov chains used for modeling probabilistic pavement performance. Besides, a decision-tree instrument was constructed to help future researchers and highway agencies select the appropriate Markov chain methodology and its design parameters for managing their pavements.

Second, a hybrid approach was developed to incorporate the marginal effects of PM into probabilistic pavement performance models. To apply and validate the developed hybrid approach, survey questionnaires were designed and deployed to State Transportation Agencies (STAs) across the United States. In addition, condition data of interstate flexible pavements (black-topped roads that include asphalt and composite) across the Midwestern states were retrieved from the Long-Term Pavement Performance (LTPP) database to implement the hybrid approach, and to develop and validate a non-homogenous Markov model for pavement performance.

Third, the condition data collected for interstate flexible pavements across the Midwestern states were used to build Markov chain models using various Markovian methodologies along with different combinations of number of condition states (NCS) and length of duty cycle (LDC). Then the developed Markov models were compared in terms of their predictive power to assess the statistical significance of NCS and LDC for the prediction accuracy of Markovian pavement performance models.

Finally, stochastic optimization models were developed to optimally schedule the preventive maintenance treatments at road network level. The Multi-Objective Genetic Algorithm (MOGA) was used to satisfy the objective functions of minimum maintenance costs and minimum deterioration of pavement network and to obtain global optimal solutions. The stochastic MOGA models were demonstrated in the context of interstate flexible pavements. Three approaches were proposed and implemented to overcome the expected large size of the optimization problem when using stochastic MOGA models to optimize pavement preventive maintenance at road network level.

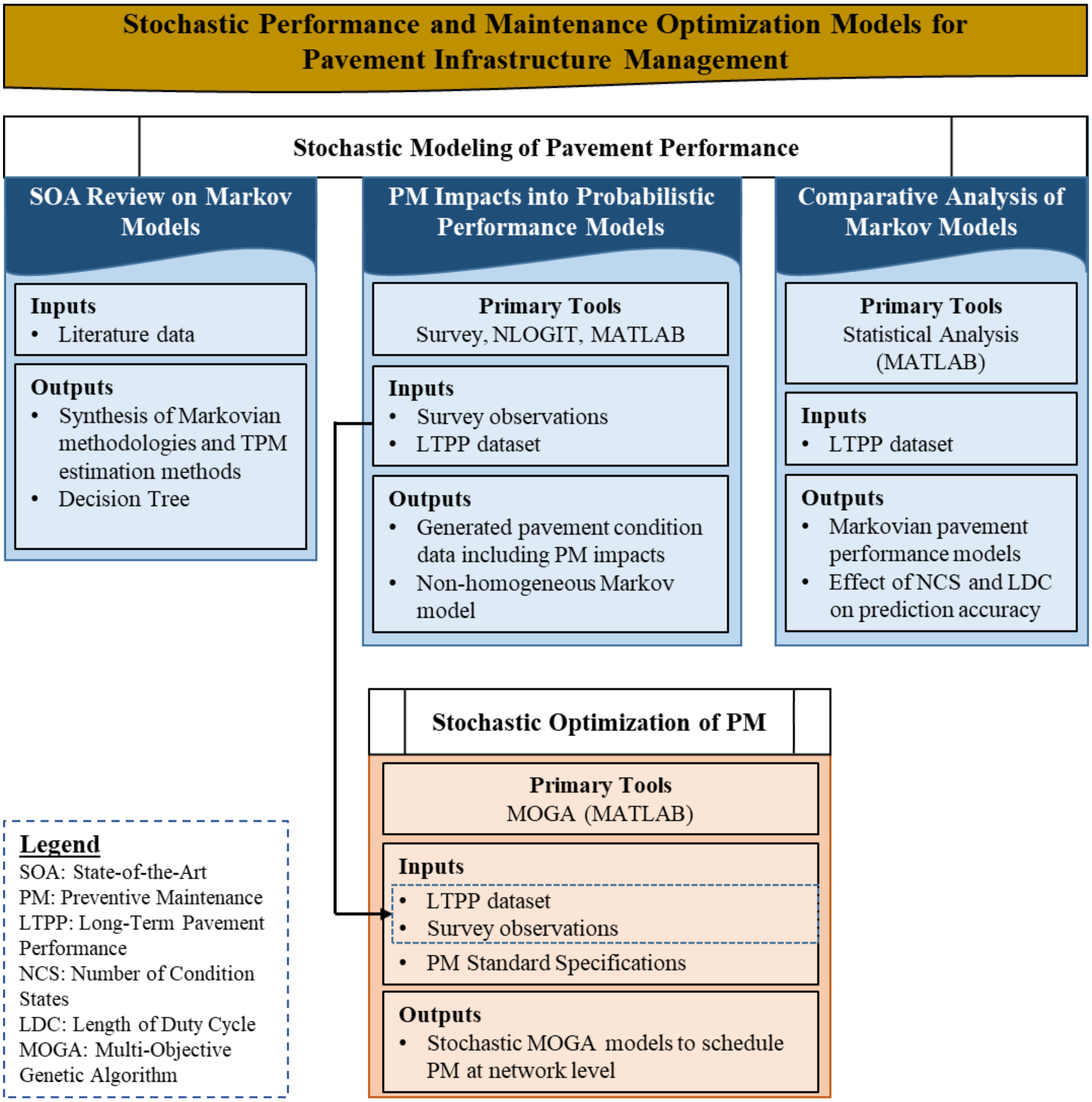


Figure 1.1. Research framework

### 1.6. Organization

This dissertation consists of six chapters and follows the “multiple publications” format. Each of the Chapters 2, 3, 4 and 5 has its own introduction, literature review, methodology, results and discussion, and conclusion sections. Significant portions of these chapters have been submitted or are in preparation for submission for review and publication in peer reviewed journals and/or refereed conferences. Chapter 1 introduces the motivation for the current research, problem statement and research objectives. Chapter 2 reviews the previous research efforts with regard to modeling of pavement performance and discusses and synthesizes the use of Markov chains in the probabilistic modeling of pavement performance. Chapter 3 introduces a hybrid approach to incorporate preventive maintenance impacts into probabilistic pavement performance models. In addition, chapter 3 discusses the results of the development and validation of non-homogeneous Markov model for pavement performance. ***This chapter is under review in the ASCE Journal of Transportation Engineering, Part B: Pavements, 2020, Mohamed S. Yamany, Dulcy M. Abraham, Hybrid Approach to Incorporate Preventive Maintenance Effectiveness into Probabilistic pavement Performance Models. Tables and figures captions were modified to maintain the form of the dissertation.***

Chapter 4 discusses the evaluation and comparison of different Markovian techniques used to model the probabilistic performance of pavement infrastructure. This chapter analyzes the prediction accuracy of Markovian methodologies and models when accounting for the number of condition states and length of duty cycle (two main components of Markov models). ***This chapter is under review in the ASCE Journal of Infrastructure Systems, 2020, Mohamed S. Yamany, Dulcy M. Abraham, and Samuel Labi, Comparative Analysis of Markovian Methodologies for Modeling Infrastructure System Performance. Tables and figures captions were modified to maintain the form of the dissertation.***

Chapter 5 discusses the stochastic optimization models that were developed using the multi-objective genetic algorithm to schedule the preventive maintenance treatments for road network. This chapter compares the results of the developed stochastic models with the typical deterministic models and highlights the contributions that stochastic optimization models can provide to the body of knowledge and to highway decision-makers. Chapter 6 presents the conclusions of the dissertation, the contributions to the body of knowledge and practice, and the limitations of the current research and recommendations for future research.

# CHAPTER 2. PRIOR RESEARCH ON PAVEMENT PERFORMANCE MODELS

This chapter reviews prior research on pavement performance models and presents the state-of-the-art probabilistic modeling of pavement performance using Markov chain techniques. In addition, it discusses the limitations and drawbacks of prior Markovian pavement performance models and identifies the knowledge and practice gaps.

## 2.1. Pavement Condition and Distresses

The purpose of a pavement system is to provide smooth surface over which vehicles may safely pass under all climatic conditions for the specific performance period of pavement. Figure 2.1 shows the basic components of the typical pavement system.

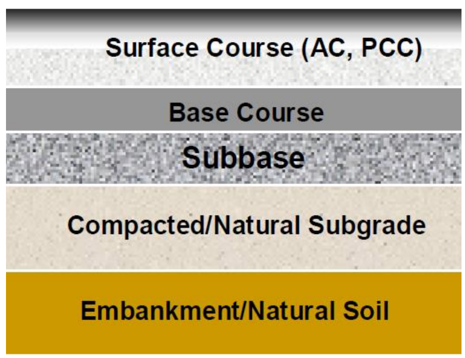



Figure 2.1. Basic components of pavement system (adapted from Christopher et al. 2006)

The types of pavements according to their construction materials or surface types are flexible or Asphalt, rigid (Portland Cement Concrete, PCC), composite (Asphalt and PCC), and unpaved. Flexible pavements have an asphaltic surface layer, with no underlying Portland cement slabs. The asphaltic surface layer may consist of high quality, hot mix asphalt (HMA) concrete, or it may be some type of lower strength and stiffness asphaltic surface treatment. In either case, flexible pavements rely heavily on the strength and stiffness of the underlying unbound layers to supplement the load carrying capacity of the asphaltic surface layer. Figure 2.2 displays the various common flexible pavement sections. Rigid pavements have a surface course of Portland Cement Concrete (PCC). The PCC slabs constitute the dominant load-carrying component in a rigid pavement system. Figure 2.3 shows the typical structural layers of rigid pavements. Composite pavements combine elements of both flexible and rigid pavement systems, usually consisting of an asphaltic concrete surface placed over PCC or bound base. Unpaved roads or naturally surfaced roads simply are not paved, relying on granular layers and the subgrade to carry the load. Seal coats are sometimes applied to improve their resistance to environmental factors.

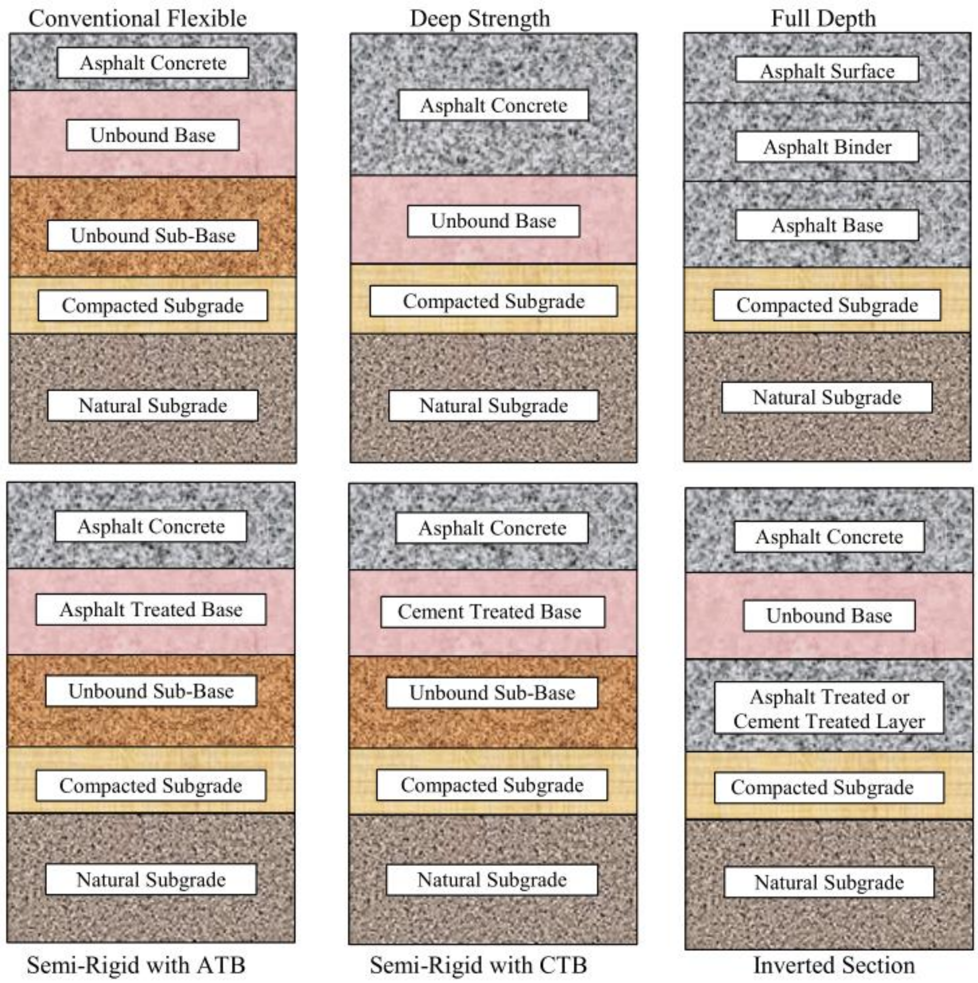



Figure 2.2. Common flexible pavement profiles (adapted from NCHRP 1-37A Design Guide, 2002).

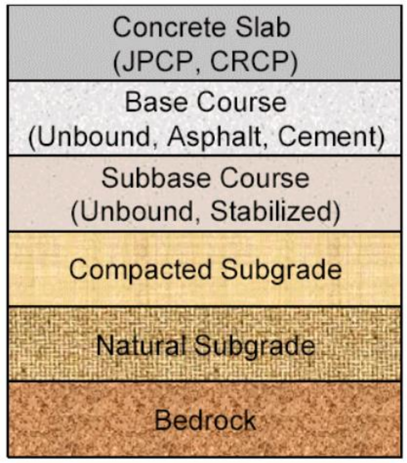



Figure 2.3. Rigid pavement profile (adapted from NCHRP 1-37A Design Guide, 2002)

Since the research of this dissertation was demonstrated using interstate flexible pavement data, this section briefly discusses the distresses associated with the flexible/asphalt pavements. The discussion is based on a review of the technical report "Maintenance Technical Advisory Guide." developed by the State of California Department of Transportation (2008). The common distresses in flexible pavements are placed into five categories: (1) cracking, (2) deformation, (3) deterioration, (4) mat problems, and (5) problems associated with seal coats. Figures 2.4 to 2.8 show examples of these distresses.

1- Cracking:

- Fatigue: cracks in asphalt layers that are caused by repeated traffic loadings. The cracks indicate fatigue failure of the asphalt layer. When cracking is characterized by interconnected cracks, the cracking pattern resembles that of an alligator's skin or chicken wire. Therefore, it is also referred to as alligator cracking.
- Longitudinal: cracks that are approximately parallel to pavement centerline and are not in the wheel path. Longitudinal cracks are non-load associated cracks. Location within the lane (wheel path versus non-wheel path) is significant. Longitudinal cracks in the wheel path are normally rated as Alligator 'A' cracking.
- Transverse: cracks that are predominately perpendicular to pavement centerline and are not located over Portland cement concrete joints. Thermal cracking is typically in this category.

- Reflective: cracks in HMA overlay surfaces that occur over joints in concrete or over cracks in HMA pavements.
- Block: pattern of cracks that divides the pavement into approximately rectangular pieces. Rectangular blocks range in size from approximately 0.1 square yard to 12 square yards.
- Edge: crescent-shaped cracks or fairly continuous cracks that intersect the pavement edge and are located within 2 feet of the pavement edge, adjacent to the unpaved shoulder. Includes longitudinal cracks outside of the wheel path and within 2 feet of the pavement edge.

Fatigue cracking

Longitudinal

Transverse

Reflective


Block

Edge

Figure 2.4. Types of cracks in flexible/asphalt pavements (Maintenance Technical Advisory Guide 2008)

2- Deformation

- Rutting: longitudinal surface depression that develops in the wheel paths of flexible pavement under traffic. It may have associated transverse displacement.
- Corrugation: transverse undulations appear at regular intervals due to the unstable surface course caused by stop-and-go traffic.
- Shoving (Wash Boarding): a longitudinal displacement of a localized area of the pavement surface. It is generally caused by braking or accelerating vehicles, and is usually located on hills or curves, or at intersections. It also may have vertical displacement.
- Depressions: small, localized surface settlement that can cause a rough, even hazardous ride to motorists.
- Overlay Bumps: in newly overlaid pavements, bumps occur where cracks in old pavements were recently filled. This problem is most prevalent on thin overlays.

3- Deterioration

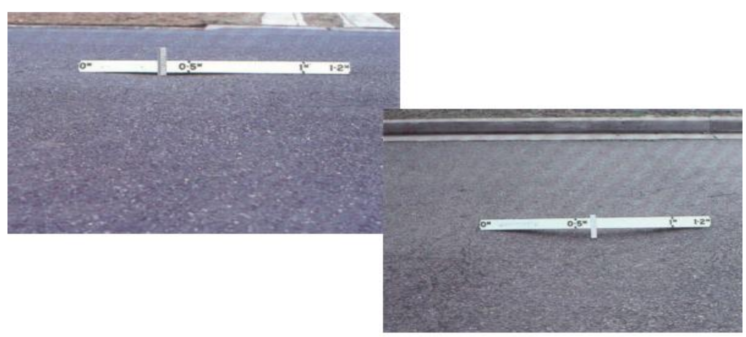
- Delamination: loss of a large area of pavement surface. Typically, there is a clear separation of the pavement surface from the layer below. Slippage cracking may often occur as a result of poor bonding or adhesion between layers.
- Potholes: bowl-shaped holes of various sizes in the pavement surface. Minimum plane dimension is 6 inches.
- Patching: portion of pavement surface, greater than 0.1 square yard, that has been removed and replaced or additional material applied to the pavement after original construction.
- Raveling**:** wearing away of the pavement surface in high-quality hot mix asphalt concrete that may be caused by the dislodging of aggregate particles and loss of asphalt binder.
- Stripping: the loss of the adhesive bond between asphalt cement and aggregate, most often caused by the presence of water in asphalt concrete, which may result in raveling, loss of stability, and load carrying capacity of the HMA pavement or treated base.
- Polished Aggregate: surface binder worn away to expose coarse aggregate.
- Pumping: seeping or ejection of water and fines from beneath the pavement through cracks.

Rutting

Corrugation

Shoving (Wash Boarding)

Overlay Bumps



Depressions

Figure 2.5. Types of deformation in flexible/asphalt pavements (Maintenance Technical Advisory Guide 2008)

Delamination

Potholes

Patching

Raveling

Stripping

Polished Aggregate

Pumping

Figure 2.6. Types of deterioration in flexible/asphalt pavements (Maintenance Technical Advisory Guide 2008)

4- Mat Problems

- Segregation: separation of coarse aggregate from fine aggregate as a result of mishandling of the mix at several points during mix production, hauling, and placing operations. Segregation leads to non-uniform surface texture and non-uniform density.
- Checking: short transverse cracks, usually 1 inch to 3 inches in length and 1 inch to 3 inches apart, which occur in the surface of the HMA mat at some time during the compaction process. The cracks do not extend completely through the depth of the course but are only 3/8 to ½ inch deep.
- Bleeding: excess bituminous binder occurring on the pavement surface. May create a shiny, glass- like, reflective surface that may be tacky to the touch. Usually found in the wheel paths.

Segregation

Checking

Bleeding

Figure 2.7. Types of mat problems in flexible/asphalt pavements (Maintenance Technical Advisory Guide 2008)

5- Problems associated with seal coats

- Raveling (Rock Loss): wearing away of the pavement surface in seal coats.
- Segregation: separation of coarse aggregate from fine aggregate as a result of mishandling of the mix at several points during mix production and placing operations. Segregation leads to non-uniform surface texture.

- Bleeding/Fat Spots: excess binder occurring on the surface treated pavements. May create a shiny, glass-like, reflective appearance. Fat spots are localized bleeding.
- Delamination: loss of portion of pavement surface treatment. Usually there is a clear separation of the surface treatment from the layer below.

Raveling (Rock Loss)

Segregation

Bleeding/Fat Spots

Delamination

Figure 2.8. Problems associated with seal coats (Maintenance Technical Advisory Guide 2008)

## 2.2. Pavement Performance Models

Pavement performance is the pavement serviceability pattern over a period of time, where the serviceability implies the ability of pavements to meet the demand for traffic under their current conditions. Pavement performance models which are critical for effective pavement maintenance and rehabilitation (M&R) decision-making require reliable and accurate pavement condition predictions.

### 2.2.1. Pavement Condition Indicators/Indices

The performance of pavements is identified by measuring and observing their condition over their lifetime. Several types of indices or indicators are used for characterizing pavement condition.

Pavement condition indicators can represent pavement structural condition (e.g., pavement structural number or distress score), pavement functional condition (e.g., International Roughness Index (IRI) or riding quality), or both pavement structural and functional conditions (e.g., Pavement Condition Index (PCI) or Pavement Condition Rating (PCR)). Each State Highway Agency (SHA) in the U.S. uses different pavement condition indices based on its policy and pavement management system. For instance, the state of Ohio uses the PCR (Rajagopal 2006), whereas, the states of Indiana, Illinois, Connecticut, Colorado, and California use the IRI (Bektas et al. 2014). The IRI, PCR and Rutting (RUT) were used to measure the effectiveness of rehabilitation treatments for the state of Indiana in the study by Labi et al. (2006), and to predict pavement condition in the studies by Sarwar and Anastasopoulos (2016 and 2017). The majority of SHAs and past studies (see Table 2.1) use the IRI to measure the surface condition of Asphalt or flexible pavements.

### 2.2.2. Factors Affecting Pavement Condition

Past studies have found several factors that affect the deterioration of pavement performance. These factors are placed into five categories (Figure 2.9): traffic, climate, material, M&R, and other. As shown in Table 2.1, the most common variables affecting pavement condition are pavement age, climatic conditions, and traffic loading. Pavement age has been recognized as the most statistically significant variable in predicting pavement performance (Abaza 2004; Kim and Kim 2006; Rajagopal 2006). Ahmed et al. (2016) used traffic loading and climate conditions to predict pavement deterioration. Other variables that were employed to predict pavement performance include subgrade resilient modulus (Hong and Somo 2001; Abaza 2004), construction quality (Rose et al. 2018), and pavement treatment expenditure (Montgomery et al. 2018).

### 2.2.3. Pavement Performance Modeling Approaches

Pavement performance has been modeled using deterministic, Artificial Intelligence (AI) and probabilistic approaches (Table 2.1). Deterministic models (e.g., Abaza 2004; Chu and Durango-Cohen 2008) assume that pavement condition can be predicted exactly. Therefore, they do not consider the inherent uncertainty and randomness of pavement condition. Although the Artificial

Neural Networks (ANNs) models, the most common AI approach, have been developed in previous research (Plati et al. 2016; Amin and Amador-Jiménez 2017; Yamany et al. 2019b; Yamany et al. 2020b) to predict pavement performance, the interpretation of their findings is not easy and they are considered as black boxes (García de Soto et al. 2018).

Unlike deterministic and AI models, probabilistic models duly acknowledge the uncertainties inherently attributed to pavement condition data, and their results are easy to comprehend and interpret. Probabilistic models result in more reliable and robust pavement condition predictions than deterministic models (Rose et al. 2018; Qiao et al. 2019). Markov chains is the most extensively used probabilistic approach to model pavement condition. One category of Markov chains models is the non-homogeneous Markov models, which are accurate and realistic in estimating and predicting pavement performance because they account for the non-stationary nature of pavement deterioration. However, these models require a large amount of historical data, which is hindered by limited resources for data collection, storage and management.

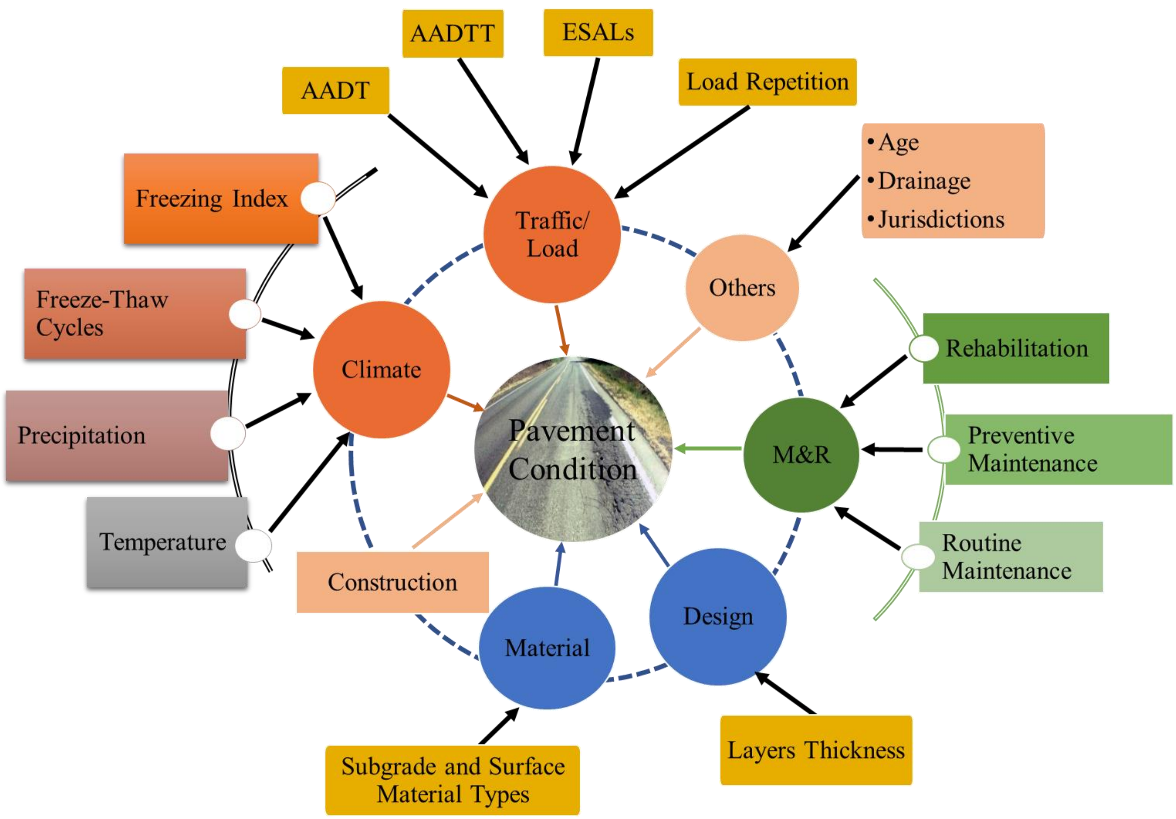



Figure 2.9. Factors affecting pavement condition

Table 2.1. Summary of insights from past research [adapted from Yamany et al. (2020b)]

| AUTHOR(S) | YEAR | STUDY REGION | USA/INT. | PAVEMENT SURFACE TYPE | CONDITION INDICATOR | ROAD CLASS | DATA SOURCE | MODELING METHODOLOGY | EXPLANATORY VARIABLES |
|---|---|---|---|---|---|---|---|---|---|
| **ABAZA** | 2004 | – | USA | AC | PSI | – | AASHTO | FP | PA, MR, ESALs, SN |
| **ABDELAZIZ ET AL.** | 2018 | – | USA, Canada | AC | IRI | – | LTPP | FP, ANN | PA, Initial IRI, Transverse and Alligator Cracks, RUT |
| **AHMED ET AL.** | 2016 | Indiana | USA | AC, PCC | IRI | IS, NHS-NIS | INDIPAVE2000, InDoT 2011a | FP | AADT, AAFI, Time |
| **ALEADELAT ET AL.** | 2018 | Wyoming | USA | AC | PSI | Locals | Pathway services Inc. 2014 | FP | PCI, IRI, Rut |
| **AL-SULEIMAN AND SHIYAB** | 2003 | Dubai | UAE | AC | IRI | Majors | In situ gathered data | FP | PA, PSI |
| **AMIN AND AMADOR-JIMÉNEZ** | 2016 | Montreal City | Canada | AC, PCC | PCI | Collectors, Arterials | Montreal City database | ANN (BPN, GDR algorithms) | AADT, ESALs, SN, PA, Layer Thickness |
| **BEKTAS ET AL.** | 2014 | Iowa | USA | AC, PCC, others | PCI-2 | IS, Non-NHS | Iowa database | FP | Cracking, Riding, Faulting Indices |
| **BIANCHINI AND BANDINI** | 2010 | Minnesota | US (Midwest) | AC | PSI | Rural | MnRoad | ANN/feedforward, Fuzzy Logic | SCI, Deflection RUT, Layer Thickness, ESALs |
| **DALLA ROSA ET AL.** | 2017 | Texas | USA | AC | IRI | Rural, Urban | TxDOT database | FP | PA, Initial IRI, Climate Zone, ESALs, Layer Thickness |
| **HONG AND PROZZI** | 2010 | Minnesota | USA (Midwest) | AC | IRI | Rural | MnRoad | FP, RP, RE | ESALs, Layer Thickness, PA, Frost Heave, Maintenance, Asphalt Mixture |
| **KHAN ET AL.** | 2014 | Queensland | Australia | AC | DR | – | Queensland (TMR-QLD database) | Markov chain | PA, Layer Thickness, AADT, ESALs |
| **KHATTAK ET AL.** | 2014 | Louisiana | USA | Composite | IRI | IS, Arterials, Collectors | Louisiana DOT database | FP | ESALs, Pre- and Post-treatment IRI, PI, TI |
| **KIM AND KIM** | 2006 | Georgia | USA | AC | PACES | State, IS | PACES database of Georgia | FP | PA, AADT |
| **LA TORRE ET AL.** | 1998 | – | USA | AC | IRI | – | LTPP | ANN | ESALs, Freezing Index, Air Voids, Base Thickness, Precipitation, PA |
| **LABI ET AL.** | 2006 | Indiana | USA (Midwest) | AC | IRI, PCR, RUT | IS, NIS | INDOT, Indiana County Flow Map, NOAA | FP | AADT, %truck, ESALs, Freeze Thaw Cycles, Freezing Index, Micro-surfacing Treatment |
| **LOU ET AL.** | 2001 | Florida | USA | AC, PCC | CI | – | FDOT database | ANN (BPN) | Maintenance Cycle, PA |
| **LUO** | 2013 | Ohio | USA (Midwest) | AC | PCR | – | Ohio DOT database | Auto-regression | PA, Past Pavement Condition |
| **MAZARI AND RODRIGUEZ** | 2016 | *Multiple Regions | USA, Canada | AC | IRI | – | LTPP | GEP, ANN, Hybrid GEP and ANN | PA, ESALs, Initial IRI |
| **ELDIN AND SENOUCI** | 2011 | Oregon | USA | AC | CR | – | Oregon DOT database | ANN (BPN) | RUT, Bleeding, Alligator &Transverse, and Block Cracking, patching/pothole, Raveling |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **PLATI ET AL.** | 2016 | – | Greece | AC | ε | – | In situ gathered data | FP, ANN | Deflection, Layer Thickness |
| **PROZZI AND MADANAT** | 2003 | – | USA | AC | PSI | – | AASHO Road Test 1962 | FP, RE | ESALs, Frost Gradient, Layer Thickness |
| **RAJAGOPAL** | 2006 | Ohio | USA (Midwest) | AC, PCC | PCN | Urban Majors, Minors | Cincinnati database | FP | PA, Surface Condition, Environmental Factors |
| **ROSE ET AL.** | 2018 | | India | AC | Raveling, Pothole, Edge Failure | Rural | Binu's (2012) | Probabilistic | PA, Construction Quality |
| **SAEED ET AL.** | 2017 | – | – | Multiple Types | – | Multiple Classes | – | FP | Traffic Loading, Freezing Index |
| **SANDRA AND SARKAR** | 2013 | Rajasthan state | India | AC | IRI | NHS, State, Major District | In situ gathered data | FP | Cracking, Potholes, Patching, RUT, Raveling |
| **SILVA ET AL.** | 2000 | Michigan | USA (Midwest) | AC | PASER | – | Michigan Counties | Logistic growth, Markov chain | PA |
| **TERZI** | 2007 | – | USA | AC | PSI | – | AASHTO | ANN (LMBPN) | IRI, Cracking, Patching, RUT |
| **XU ET AL.** | 2015 | Kentucky | USA | AC | IRI, Cracking, Raveling | IS Parkways | KTC | FP | AADT, PA, Cracking, Raveling, IRI |
| **YANG ET AL.** | 2003 | Florida | USA | AC, PCC | Cracking, Riding, RUT | – | FDOT database | ANN (BPN) | PA, Maintenance Cycle |
| **ZIARI ET AL.** | 2016 | – | USA | AC | IRI | – | LTPP | ANN, GMDH | PA, AAP, AADT, AAFI, AAT, AADTT, ESALs, ST, PT |

Note: CI: Crack Index, BPN: Backpropagation Neural, FDOT: Florida Department of Transportation, CR: Condition Rating, PSI: Present Serviceability Index, LMBPN: Levenberg-Marquardt Backpropagation Neural, GMDH: Group Method of Data Handling, AAP: Annual Average Precipitation, AAT: Annual Average Temperature, AAFI: Annual Average Freezing Index, ESALs: Equivalent Single Axle Loads, AADT: Annual Average Daily Traffic, AADTT: Annual Average Daily Truck Traffic, TxDOT: Texas Department of Transpiration, GEP: Gene Expression Programming, SN: Structural Number, PASER: Pavement Surface Evaluation and Rating, MnRoad: Databased of Minnesota Road Test Project, RE: Random Effects Model, RP: Random-parameter Regression, FP: Fixed-parameter Regression, NHS: National Highway System, Non-NHS: Non- National Highway System, PCI-2: Pavement Condition Index Developed for Iowa state, FWD: Falling Weight Deflectometer, SCI: Surface Curvature Index, NOAA: National Oceanic and Atmospheric Administration, INDOT: Indiana Department of Transportation, NHS-NIS: National Highway System-Non Interstates, IS: Interstate Highways, INDIPave: Database for Indiana state, PCR: Pavement Condition Rating, PI: Precipitation Index, TI: Temperature Index, PCI: Pavement Condition Index, AC: Asphalt Concrete pavements, PCC: Portland Cement Concrete pavements, PA: Pavement Age, and *Multiple Regions: Indiana, Iowa, Maryland, New Jersey, New York, Tennessee, Arkansas, and Oklahoma in United States, New Brunswick and Prince Edward Island in Canada.

## 2.3. Probabilistic Modeling of Pavement Performance

Reliable models of pavement deterioration play a crucial role in effective decision-making for maintaining and rehabilitating this class of infrastructure. Probabilistic modeling approaches have gained popularity in pavement deterioration modeling because they account not only for the stochastic nature of pavement behavior and deterioration factor variations but also for the imperfections and inadequacy of pavement condition data in certain situations. One of these approaches, Markov chains, has been used extensively to model the probabilistic performance of pavements through an interesting variety of methodological tweaks in the Markov model structure. Unfortunately, the current literature lacks a synthesis of Markovian models and their associated methodologies, as used in this manner. It is anticipated that a comprehensive synthesis of these models and their various forms can provide some insight into the variations of Markov model forms and methodologies, and the appropriate Markov model type to use for pavement deterioration modeling under given conditions of data types and availability. To address this issue, this section reviews Markovian models that were used in the literature to model pavement deterioration and the methodologies used to estimate the transition probabilities which are a key feature of Markov models. This section presents a critical analysis of various aspects of Markovian models as they were applied in the literature, reveals gaps in knowledge, and offers suggestions to address these gaps. This section also presents a proposed decision tree that an infrastructure agency could use to select appropriate Markov model type and methodology, to model the deterioration of a given pavement under given conditions of data availability.

### 2.3.1. Introduction

Pavement condition is evaluated with respect to its structural and functional capacities. Pavement structural capacity refers to its load-carrying strength, while pavement functional capacity refers to its level of service provided to roadway users. These structural and functional capacities are represented by condition indices/indicators such as international roughness index (IRI) and present serviceability rating (PSR). A closely related concept is pavement performance which refers to, according to AASHTO (1993), the trend of pavement serviceability over a period of time, where the serviceability indicates the ability of pavements to serve the traffic demand in the existing condition. Pavement performance models that are essential for effective decision-making of pavement maintenance and rehabilitation (M&R) need reliable and accurate predictions of pavement condition. The reliability and accuracy of condition predictions hinge on the quality and availability of pavement condition data and the modeling methodology.

Pavement performance models can be deterministic or probabilistic. Unlike deterministic models (Yamany et al. 2019b; Yamany et al. 2020b) the probabilistic models account for the variability and uncertainty in pavement condition data. These variability and uncertainty stem from: (a) measurement errors; (b) randomness of pavement deterioration; (c) inability to model the true deterioration process; (d) difficulties in quantifying the effect of all significant relevant variables; and (e) potential bias associated with the models built by using subjective expert judgment (Li 2005; Rose et al. 2018). Probabilistic models are categorized as follows: econometric, Markov chain, and reliability analysis models (Li 2005; Porras-Alvarado et al. 2014). Another way to classify probabilistic models is the criterion for change: state-based vs. time-based. State-based models, e.g., Markov processes, estimate the probability that pavement condition changes from one state to another in a given time period. Time-based models, e.g., duration models, estimate the probability of the time taken by pavement to change its condition state (Ford et al. 2011).

Although Markov models are the most commonly used probabilistic method for pavement performance modeling, the current relevant literature lacks a synthesis of Markovian models and their associated methodologies. Such a comprehensive synthesis of these models and their various forms can provide insights into the variations of Markov model forms and methodologies, and the appropriate Markov model type to use for pavement deterioration modeling under given conditions of data types and availability. As such, section 2.3 presents a state-of-the-art review for the probabilistic modeling of pavement performance using Markov chains. It discusses the properties and assumptions of Markov chain models, the categories of Markov chain models and the methods of estimating pavement transition probabilities. In addition, section 2.3 introduces a critical assessment for prior Markovian pavement performance models. Based on the insights from the literature, a decision tree is proposed to help future researchers and highway agencies select their appropriate Markov models for their pavements. Finally, this section highlights the existing gaps in the pertinent knowledge and suggests future research solutions and methodologies to bridge these gaps.

### 2.3.2. Markov Chain Models: Properties and Assumptions

Markov chain models consist of three main components: condition state vector ($S$), duty cycle or transition period, and transition probability matrix (TPM). Figure 2.10 depicts a graphical representation example of general Markov transition probabilities with condition states in nodes and transition probabilities on arrows.

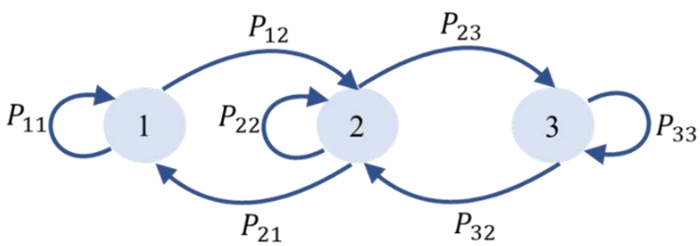



Figure 2.10. Transition probabilities diagram

The state space in this example is $X = \{1, 2, 3\}$, and the transition probability matrix is as follows:

$$\text{TPM} = \mathbb{P} = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix} \qquad (2.1)$$

In pavement performance models, the state space represents different pavement conditions measured by composite condition indices (e.g., pavement condition index (PCI)), individual distresses (e.g., cracking), or remaining service lives (Porras-Alvarado et al. 2014). The condition state vector is a list of probability distributions corresponding to the pertinent state space; $S^i = \{S^{i1}, S^{i2}, S^{i3}, \ldots\}$; where $S^i$ is the condition state vector at time $i$, any $S^{iX} \leq 1$, and $\sum_X S^{iX} = 1$. Markov models estimate the future pavement condition ($S^{i+1}$) based on the current pavement condition ($S^i$) according to the memoryless property of Markov process and the transition probabilities of pavement deterioration and improvement (TPM ); $S^{i+1} = S^i \times \text{TPM}$.

Pavement condition bounces across three phases: (1) stays at its current state $i$, (2) transits to lower states $i+1, i+2, \ldots$ etc., or (3) transits to upper states $i-1, i-2, \ldots$ etc., when maintenance or rehabilitation treatment is implemented. The condition state vector is comprised of a number of condition states defined by their probability distributions. The number of condition states depends on data availability (Martin and Kadar 2012), and it needs to be chosen prudently to capture the entire pavement condition over its lifespan (Porras-Alvarado et al. 2014). In pavement performance models, typically 10 condition states (from 1 to 10) are assumed; where state 1 represents the best condition, and state 10 represents the worst condition. However, past research assumed different numbers of pavement condition states such as 20 states (Macleod and Walsh 1998) and four states (Porras-Alvarado et al. 2014). The probability distribution of each condition state is calculated as the percentage of the number of pavement sections or the number of pavement lane-miles that lies within each state to the total size of pavement network.

The duty cycle is the duration during which pavement section transits from a condition state $(i)$ to another state $(j)$ with a corresponding probability $(P_{ij})$. The duty cycle can be a continuous time as in continuous-time Markov chain or a discrete time as in discrete-time Markov chain. Most prior studies assume discrete transition times for pavement performance models (Abaza 2016b; Abaza et al. 2004; Kobayashi et al. 2010; Pérez-Acebo et al. 2018). The selection of the duty cycle length depends on the analysis level, pavement deterioration rate and pavement inspection intervals. Prior research (Abaza 2016a, b; Abaza and Murad 2010; Butt et al. 1987; Li et al. 1996; Pulugurta et al. 2009) reported that a duty cycle of one-year length for the entire pavement lifespan is reasonable since most agencies monitor their infrastructure annually. The length of the duty cycle can be of fixed value other than one-year (Pérez-Acebo et al. 2018) or of varying values corresponding to different pavement deterioration rates.

The common assumptions of Markov chain models for pavement condition prediction include pavement deterioration is a discrete process, whereas it is continuous in nature. The duty cycle is one year because most highway agencies inspect their pavements annually. Pavement condition states can only move to one state lower every duty cycle. In other words, in the square TPM matrix, $P_{i,i}$ and $P_{i,i+1}$ are the only existent probabilities in each row of the matrix; where $i$ is the state number, and $P_{i,i} + P_{i,i+1} = 1$. The effect of the maintenance and rehabilitation treatments is not considered in estimating pavement transition probabilities, i.e. $P_{i,i-1} = 0$. The last state $(n)$ is an absorbing state, i.e., $P_{n,n} = 1$, because it is the worst condition state pavements can occupy.

### 2.3.3. Types of Markov Chain Models

Based on the assumptions of the transition probability matrix and the dependent variable (i.e. pavement condition), Markov chain models can be categorized as follows: homogeneous Markov, staged-homogeneous Markov, non-homogeneous Markov, semi-Markov, and hidden Markov models. Figure 2.11 shows the types of Markov chain models and the corresponding TPM estimation methods for pavement performance.

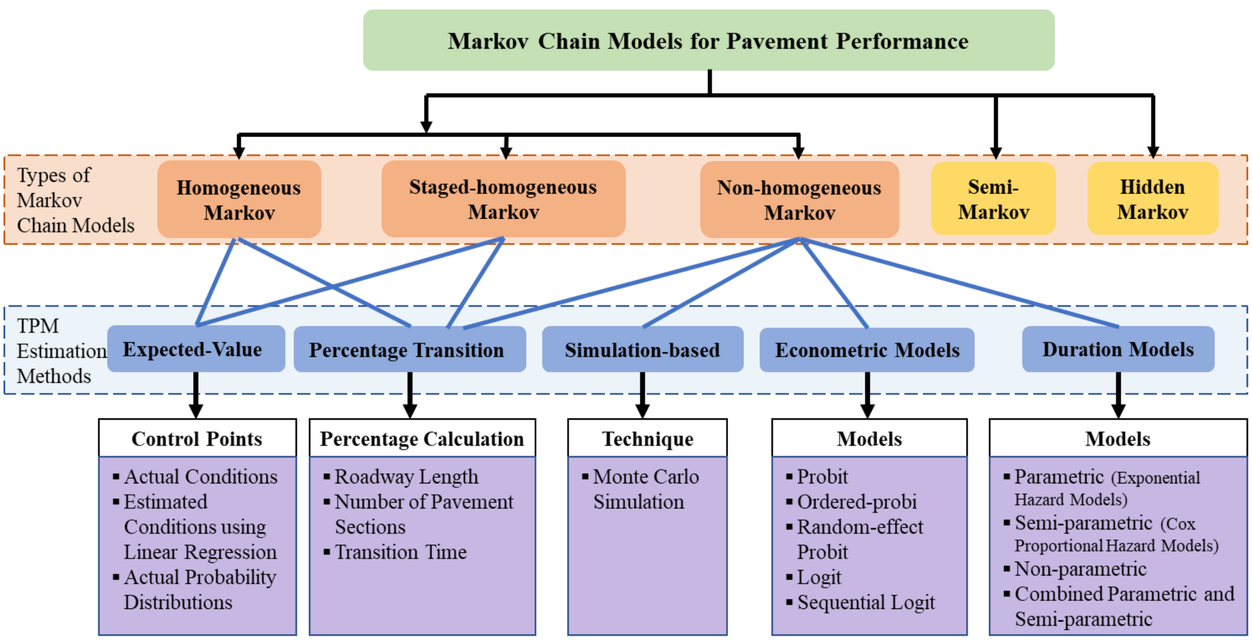



Figure 2.11. Markov Chain models and TPM estimation methods for pavement performance

#### 2.3.3.1. Homogeneous Markov Models

These models are time-independent, do not require large amounts of historical data, and are computationally simpler for pavement condition prediction. The data needed for these models is the pavement condition observations of two successive transitions. The future pavement condition after a period of time $t$ is calculated by multiplying the current probability distributions ($S^0$) by the TPM ($\mathbb{P}$) raised to the power $t$; $S^t = S^0 \times \mathbb{P}^t$.

The TPM of homogeneous Markov models is estimated using the expected-value or percentage transition methods. The expected-value method estimates $P_{ij}$ by minimizing the difference between the predicted pavement condition using Markov models and predetermined control points. These control points could be: [1] the actual pavement condition, [2] the predicted pavement condition using simple linear regression analysis, or [3] the actual probability distributions of pavement condition. The percentage transition method estimates $P_{ij}$ as the percentage of pavement sections (number of pavement sections, total length of pavement sections, or total remaining service lives of pavement sections) that have moved from state $i$ to state $j$ during the time $t$ to the total pavement sections that were originally in state $i$.

Ortiz-García et al. (2006) used the expected-value method to derive TPMs for homogeneous Markov models to predict pavement cracking, raveling, roughness and rutting. To estimate the TPM, they minimized the difference between the models' predictions and each control point. Ten condition states and 1-year duty cycle were assumed for the Markovian model structure. Different pavement performance patterns were assumed, and then pavement condition data was generated for 20 years. Pavement condition was predicted using the minimization with respect to each type of control points and then compared with the actual observations. For the control points 1 and 2, the predicted pavement conditions were found to be different from the actual observations, but for the third type of control points, the predicted and actual probability distributions were found to be similar. Hence, the excepted-value method is considered of high reliability in estimating pavement TPM when the third type of control points is used. Wang et al. (1994) derived TPMs for homogeneous Markovian models using the percentage transition method to predict pavement cracking and roughness in Arizona. Pavement sections were categorized into various groups based on traffic volume and weather condition to account for the variation in the data of pavement sections. Wang et al. (1994) found that the developed pavement performance curves using the percentage transition method match the actual performance curves. Pulugurta et al. (2009) developed homogeneous Markov models to predict pavement distress ratings and PCR in the state of Ohio. Ten condition states and 1-year duty cycle were assumed for the Markov model structure. The transition probabilities were estimated using the percentage transition method and the methodology introduced by Wang et al. (1994). Pavements were grouped based on their respective traffic volume, weather condition, and treatment type. The estimated TPM was found to be overestimated during the latter ages of pavement, and so the researchers used the statistical imputation to overcome this overestimation.

Abaza and Murad (2010) and Abaza (2014) estimated the TPM for homogeneous Markovian models to optimize pavement rehabilitation treatments and predict pavement distress rating (DR) in Palestine. Ten condition states and 1-year duty cycle were assumed for the Markovian models' structure. The TPM was estimated as the percentage of the number of remaining pavement sections in each state at the end of the duty cycle to the total number of sections at the beginning of the duty cycle. Butt et al. (1987) had found that pavement deterioration rates do not change significantly during a period of 5 or 6 years, and thus Abaza and Murad (2010) assumed an analysis period of 5 years for their model. Abaza (2014) investigated the sensitivity of the TPM to pavement section lengths 10, 30, 50 and 100m, and he concluded that the transition probabilities become more unstable when pavement section length increases. As a result, he recommended using shorter pavement sections to avoid instability in TPM values.

Abaza (2004) developed homogeneous Markov models to predict pavement present serviceability index (PSI). He estimated the TPM as the ratio between the actual transition time that each state takes to move to the next state and the duty cycle. He assumed five condition states and one and two years for the duty cycle. The transition time was interpolated from a deterministic pavement performance curve that represents the relation between pavement PSI and respective equivalent single axle loads (ESALs) using the AASHTO's design methodology (AASHTO 1993). The estimated transition probabilities were found to be consistent with the engineering intuition. Table 2.2 shows the key studies that used the expected-value and the percentage transition methods to estimate TPMs for homogenous Markov models for pavements.

Table 2.2. Prior homogenous Markov models and associated TPM estimation methods

| Expected-value | | Percentage Transition | |
|---|---|---|---|
| Control Points | Key Studies | Percentage Calculation | Key Studies |
| Actual pavement condition | Shafahi and Hakhamaneshi (2009) | Percentage of pavement length that transits to different states | Wang et al. (1994); Pulugurta (2007); Chou et al. (2008); Pulugurta et al. (2009); Hassan et al. (2017a; b); Osorio-Lird et al. (2018), |
| Predicted pavement condition using linear regression | Ranjith et al. (2011) | Percentage of the number of pavement sections that transit to different states | Macleod and Walsh (1998); Panthi (2009); Mandiartha et al. (2010); Abaza and Murad (2010); Abaza (2014); Pérez-Acebo et al. (2018) |
| Actual probability distributions of pavement condition | Ortiz-García et al. (2006); Chun et al. (2012); Porras-Alvarado et al. (2014) | Percentage of the duty cycle to the transition time | Abaza (2004) |

Although the homogenous Markov models are computationally easy, they suffer from several drawbacks. The results of homogeneous Markov models can be questionable because of their stationary assumption (Li 2005). This assumption ignores the change in pavement deterioration rate due to the increase in traffic loading and the degradation of pavement structural capacity (Abaza 2016a). Additionally, the homogeneous Markovian models do not account for the impact of the exogenous variables such as traffic loads and environmental conditions. To overcome this limitation, pavement sections can be segmented based on pavement attributes such as pavement age, traffic loading, and climate severity. However, the pavement section segmentation decreases the sample size which in turn lowers the accuracy of Markovian models. Homogeneous Markov models could yield an overestimation of pavement condition over the entire lifetime of pavement (Durango 2002) or its latter ages (Pulugurta et al. 2009). This overestimation could lead to insufficient M&R actions during pavement life. Statistical imputation techniques were recommended by Pulugurta et al. (2009) to avoid this expected overestimation.

The expected-value and transition percentage methods are typically used to derive constant TPMs for the homogeneous Markov models of pavement condition prediction. These methods require two consecutive transitions of pavement conditions, which is insufficient to capture the historical behavior of pavements. To overcome this limitation, Ortiz-García et al. (2006) suggested calculating the average transition probabilities for more than one duty cycle. Ranjith et al. (2011) concluded that the expected-value method is more accurate than the percentage transition in estimating the TPM for modeling timber bridges' elements. In the expected-value method, the methodology of minimizing the difference between the estimated pavement condition using simple linear regression and using Markov models is unreliable since the relationship between pavement condition and pavement age is non-linear, and pavement age is not the only variable influences pavement condition.

#### 2.3.3.2. Staged-homogeneous Markov Models

Butt et al. (1987) introduced this type of models to overcome the limitation of data unavailability when developing a non-homogeneous Markov model. Staged-homogeneous Markov models involve dividing the analysis period into zones, each of 5 or 6 years at maximum. Pavement sections are sorted and grouped based on their ages. Homogeneous TPMs are established for every zone. The future condition of pavement section at any time $t$ is calculated by multiplying the current probability distributions of this section by the TPM of every zone until the time $t$, i.e., $S^t = S^0 \times \mathbb{P}_1{}^z \times \mathbb{P}_2{}^z \times \ldots \times \mathbb{P}_t{}^{t-nz}$; where $\mathbb{P}_1{}^z$ is the TPM of the first zone raised to power $z$, and $z$ is the zone size in years; $\mathbb{P}_t{}^{t-nz}$ is the TPM of the zone that includes the time $t$, and $n$ is the number of zones until the zone that includes $t$. In the staged-homogeneous Markov models the TPM is estimated using the expected-value or percentage transition methods.

Butt et al. (1987) developed a staged-homogeneous Markov model to predict pavement PCI using data from the PAVER database. They assumed 10 condition states with 1-year duty cycle for the Markov model structure. The zone size was assumed to be 6 years. The TPM was estimated using the expected-value method by minimizing the difference between the actual and predicted pavement condition. The developed model was validated by comparing its predictions with the actual observations and with the predictions from a previous homogeneous Markov model that was developed by Keane and Wu (1985) in collaboration with the U.S. Army Construction Engineering Research Laboratory (USA-CERL). The results showed that the staged-homogeneous Markov model of Butt et al. (1987) outperforms its homogeneous counterpart of Keane and Wu (1985). Abaza (2016a) presented staged-homogeneous Markov models to predict pavement deterioration rate (DR) using data spanning from 1998 to 2015 for a major urban arterial road in Palestine. Two models were built: three-year and five-year staged-homogeneous Markov models. The TPMs were estimated using the percentage transition method with the transition probability equals the proportion of the number of pavement sections that transits from one state to another. The TPM was assumed to change by a constant *C* every stage/zone (3 or 5 years) in both models. The constant C was assumed to take on values greater than 1, and its value was exactly determined by minimizing the difference between the actual and predicted transition probabilities/DRs. Both models of Abaza (2016a) were found to be statistically reliable in predicting pavement condition. However, the three-year staged-homogeneous Markov model was found to be superior to the other model with respect to the sum square errors (SSE). Table 2.3 presents the key studies that used staged-homogeneous Markov models and their corresponding TPM estimation methods for pavement performance modeling.

Table 2.3. Prior staged-homogenous Markov models and associated TPM estimation methods

| **Expected-value** | | **Percentage Transition** | |
|---|---|---|---|
| Control Points | Key Studies | Percentage Calculation | Key Studies |
| Actual pavement condition | - | Percentage of pavement length that transits to different states | - |
| Predicted pavement condition using linear regression | Butt et al. (1987) | Percentage of the number of pavement sections that transit to different states | Abaza (2016a) |
| Actual probability distributions of pavement condition | - | Percentage of the duty cycle to the transition time | - |

The staged-homogeneous Markov models have two advantages. First, they are more reliable than the homogeneous Markov models in pavement condition prediction because they account for the non-stationary process of pavement deterioration. Second, they require relatively limited amounts of data. However, since the staged-homogeneous Markov models use the expected-value and percentage transition methods to estimate TPMs, they suffer from the limitations of these TPMs' estimation methods (discussed earlier). Unlike the staged-homogeneous Markov models found in the literature and because pavement deterioration rate varies over time, the analysis period should be divided into unequal zones based on pavement performance curve and its respective rate of deterioration, not into constant zones.

2.3.3.3. Non-homogeneous Markov Models

These models are time-dependent and consider the non-stationary property of pavement deterioration process. These models account for the uncertainty inherently attributed to explanatory variables such as traffic loads and weather conditions (Abaza 2016a; Kobayashi et al. 2010; Li 1997, 2005). Although non-homogeneous Markov models fit realistically the random behavior of pavement condition over time, they have not been adopted widely in pavement performance modeling because they require extensive computation and large amounts of data. The future condition of pavement at any time $t$ is calculated by multiplying the current probability distributions of this pavement by the TPM of every duty cycle until the time $t$, i.e., $S^t = S^0 \times \mathbb{P}_1 \times \mathbb{P}_2 \times \ldots \times \mathbb{P}_t$. In non-homogeneous Markov models, the TPM is estimated using one of the following methods: percentage transition, simulation-based, econometric models or duration models.

Abaza (2017a) developed a non-homogeneous Markov model to predict pavement DR in Palestine. Ten condition states and 1-year duty cycle were assumed for the Markovian model structure. He used the percentage transition method to estimate the TPM of the first duty cycle. The remaining TPMs were calculated by multiplying the TPM of the first duty cycle by the two factors: traffic loads and pavement structure number. The developed models yielded transition probabilities comparable with the actual data, which ascertains that the change in pavement deterioration rate due to traffic loading should be considered when modeling pavement performance. Furthermore, the TPM can be estimated using the simulation-based method in which transition probabilities are expressed in terms of the percentiles of pavement condition states. Li (1997) and Li et al. (1996) used this method to develop non-homogeneous Markov models to predict pavement condition. The pavement deterioration formula developed in the model of Ontario Pavement Analysis of Cost (2000) was employed, and the impact of the ensuing explanatory variables was included: material modulus and thickness of each pavement layer, subgrade modulus, annual average daily traffic (AADT), traffic growth rate, truck percentage, number of traffic lanes in each direction, and ESALs. Using Monte Carlo simulation, the transition probabilities of pavement condition were estimated assuming that the studied variables follow the standard normal distribution. The researchers further checked the sensitivity of the transition probabilities to the independent variables considered in their study, and they found that pavement transition probabilities are significantly sensitive to traffic growth rate, subgrade strength and pavement layer thickness.

The econometric models are recommended for TPM estimation to reflect the historical behavior of pavement condition based on large amounts of historical data. These models associate pavement deterioration with the influential pertinent explanatory variables. Also, they yield pavement condition predictions that are more accurate than that obtained from the abovementioned methods, percentage transition and simulation-based (Madanat et al. 1995, 1997; Yang et al. 2005). The key econometric models that are used in Markovian pavement performance models include probit, logit and ordered-probit. Probit and logit models are employed to statistically model pavement condition states as discrete variables. They are grouped into binary and multinomial models based on the number of outcomes of the model. These models assume a latent continuous dependent variable ($U$) that takes values from $-\infty$ to $\infty$, and correlates with an explanatory variables vector ($X$). The probability of selecting a specific choice or for the outcome to be equal to a specific value depends on the estimated $U$ for all choices or for all expected values. Equation 2.2 shows the estimation of the probit or logit models; where $P(i)$ is the probability of choice $i$, $I$ is the total number of choices, $n$ is the number of observations, $\beta$ is the model parameter, and $\varepsilon$ is the error term.

$$P_n(i) = P(\beta_i X_{in} - \beta_I X_{In} \geq \varepsilon_{In} - \varepsilon_{in}) \quad \forall\, I \neq i \qquad (2.2)$$

In probit models, the error term follows the standard normal distribution ($\Phi$), whereas, in logit models, it follows the logistic distribution. The maximum likelihood estimation (MLE) method is used to estimate models' parameters ($\beta$) by maximizing the log-likelihood function that is illustrated in Equations 2.3 and 2.4 for probit and logit models, respectively.

$$\begin{aligned} Log\_Likelihood &= LL = \ln P(i) \\ &= \sum_{n=1}^{N} \delta_i \, ln\, \Phi(\beta_1 X_{1n} - \beta_2 X_{2n}) + (\delta_i - 1)\, ln\, \Phi(\beta_1 X_{1n} - \beta_2 X_{2n}) \end{aligned} \qquad (2.3)$$

$$LL = \sum_{n=1}^{N}\sum_{i=1}^{I}\delta_{in}\left[(\beta_i X_{in}) - ln\sum\nolimits_{\forall I} EXP(\beta_I X_{In})\right] \quad (2.4)$$

where $\delta_i$ is the value of the choice $i$. Yang et al. (2005) developed a non-homogeneous Markov model to predict pavement cracking rate using logit models to estimate the TPM. Pavement condition data were retrieved from the state of Florida during the period from 1986 to 2003. The pavement age, ESALs, crack index (CI) and a number of rehabilitation cycles were found to be statistically significant in estimating pavement TPMs. The researchers set the pavement data of 2003 aside to build a homogeneous Markov model, and then compare its results with that of the non-homogeneous Markov model. The values of the validation measures: average absolute error (AAE), root mean square error ($RMSE$) and coefficient of determination ($R^2$), demonstrated the superiority of the non-homogeneous Markov model to the homogeneous Markov model. The study conducted by Yang et al. (2005) assures that pavement condition propagates due to exogenous variables that should be taken into account when developing Markovian prediction models for pavements. Yang et al. (2006) developed an artificial neural networks (ANNs) model to predict pavement CI. The research team compared the results of the ANNs model with that of their non-homogeneous Markov model of 2005. Based on the values of the same validation measures they used in 2005 with respect to both models, they found that both models have a similar performance for a single-year prediction, but the non-homogeneous Markov model was found to be more accurate than the ANNs for multiple-year predictions.

Ordered-probit models estimate discrete and ordered dependent variables when the order matters. Equations 2.5 shows the estimation of ordered-probit models.

$$P(C_n = k) = \Phi(\Psi_k - \beta X_n) - \Phi(\Psi_{k-1} - \beta X_n) \quad (2.5)$$

where $C_n$ is the choice of $n$ observations, $k$ is the choice value $(0, 1, \dots, K)$, $\Phi$ is the cumulative distribution function, and $\Psi_k$ is the order of the choice $k$. The MLE method is used to estimate the model's parameters $(\beta, \Psi)$. Li (2005) developed ordered-probit and sequential logit models to estimate the TPM for pavement PSI prediction. The sequential logit model is a series of independent binary logit models. Unlike ordered-probit models, sequential logit models account for the dependency between condition states. Pavement structure and environment relevant variables and traffic loading were considered in these models. Data from the AASHO Road Test was employed for models' validation. These models were compared with prior three models namely, the non-homogeneous and homogenous Markov models of Butt et al. (1987) and Wang et al. (1994), respectively, and the duration model of Prozzi and Madanat (2000). The ordered-probit and sequential logit models were found to be reliable in the prediction of pavement PSI, and more accurate than the prior three models.

Duration models assume that the transition probability of pavement condition is the probability distribution of the time elapsed until pavement changes its condition state. Duration models are effective in estimating the TPMs if relevant data are available for more than 10 years. Also, they account for censored data that is inherently associated with infrastructure data collection (Mauch and Madanat 2001). In the duration models, the data is considered either left-censored, right-censored or interval censored if the duration of leaving a given state is less than a certain value, greater than a certain value or on an interval between two values, respectively. The estimation of pavement transition probabilities using duration models is presented in Equations 2.6 and 2.7 based on the study of Mishalani and Madanat (2002).

$$R(t, \Delta t) = P(t < T < t + \Delta t \mid T > t) \tag{2.6}$$

$$R(t, \Delta t) = \frac{P(t < T < t + \Delta t)}{P(T > t)} = \frac{F(t + \Delta t) - F(t)}{1 - F(t)} = \frac{F(t + \Delta t) - F(t)}{S(t)} \tag{2.7}$$

where $R(t, \Delta t)$ is the transition probability from state 1 to state 0 during the time $\Delta t$ conditional on the observed state 1 at time $t$, $F(t)$ is the cumulative distribution function of the duration random variable $T$, $S(t)$ is the survivor probability. When $\Delta t$ approaches Zero the transition probability is called hazard rate. The hazard rate is estimated using parametric, semi-parametric or nonparametric models. In parametric models, the hazard rate follows a pre-specified distribution such as the normal or exponential distribution, which is a limitation of these models (Mishalani and Madanat 2002). Semiparametric models relax the limitation of the parametric models and determine the distribution of the hazard rate based on the actual data. Unlike parametric models, Semiparametric models relate the hazard rate to its pertinent exogenous variables. Nonparametric models neither assume distribution function nor derive a specific relation between the hazard rate and its exogenous variables (Mauch and Madanat 2001), but it mainly depends on the training dataset. Kobayashi et al. (2010) developed condition prediction models for pavement IRI, rutting and cracking. For the estimation of pavement TPMs, they implemented the duration models to account for the irregularities in pavement inspection periods. Pavement condition was discretized to five condition states. Four hazard models following the exponential distribution were developed to calculate the transition probabilities of states 1, 2, 3 and 4. The ESALs and the structural number (SN) variables were included in the hazard models. Data from a national highway in Korea was used for models' validation. The results showed that the predicted transition probabilities fit the actual observations. Table 2.4 shows a summary of the TPM estimation methods for non-homogeneous Markovian models along with the key studies that used these methods for pavement condition predictions.

Li (2005) stated that the simulation-based method is less expensive with respect to the computation process and data collection than the transition percentage method when they are used for non-homogeneous Markovian models. Prior researchers such as Li et al. (1996) and Li (1997) used the simulation-based method to estimate pavement TPMs; however, they were limited to the assumption that the explanatory variables follow the standard normal distribution. The econometric models link relevant explanatory variables to a latent continuous variable that is further used to estimate the discrete dependent variable (condition states). A methodology which simulates the latent nature of pavement deterioration process. Since these models consider the effect of pertinent independent variables in the estimation of pavement TPMs, the segmentation of pavement sections that is recommended to capture the impact of exogenous variables in other models, is not necessary. These econometric models utilize the MLE method to estimate models' parameters, thus they need an extensive amount of data. The MLE method assumes the standard normal distribution or logistic distribution and the homoscedasticity. If these assumptions are violated, the computation process becomes complex, and the accuracy of the models becomes questionable. Also, the interpretation of parameters estimated by the MLE is difficult compared with that by the Ordinary Least Squares (OLS). The econometric models assume that condition states are independent and identically distributed; however, future condition states of pavements hinge on the current and previous historical condition states. Additionally, these models do not account for data censoring that results from the infrequent or lack of pavement condition inspections (Mishalani and Madanat 2002). Madanat et al. (1997) developed a random-effect probit model to predict the condition of bridge decks. They found that when probit models were associated with random-effect models, they were able to capture the heterogeneity attributed to infrastructure data and yield more accurate predictions than when using probit models only. Future research is encouraged to explore the association of random effect with the probit models developed for pavement condition prediction to account for the heterogeneity that is attributed to pavement condition data.

Table 2.4. Prior non-homogenous Markov models and associated TPM estimation methods

| **Percentage Transition** | | **Simulation-based** | **Econometric Models** | | **Duration Models** | |
|---|---|---|---|---|---|---|
| Percentage Calculation | Key Studies | Key Studies | Models | Key Studies | Hazard Rate Technique | Key Studies |
| Percentage of pavement length that transits to different states | - | Li et al. (1996); Li (1997) | Probit | - | Parametric | Mishalani and Madanat, (2002) |
| Percentage of the number of pavement sections that transits to different states | Abaza (2017a) | | Ordered-Probit | Madanat et al. (1995); Li (2005); Yamany and Abraham (2020a, b) | Semi-parametric | Mauch and Madanat (2001) |
| Percentage of duty cycle to the transition time | - | | Random-effect Probit | Madanat et al. (1997) | Non-parametric | Kobayashi et al. (2010); Madanat et al. (2005) |
| | | | Logit | Yang et al. (2005, 2006) | | |
| | | | Sequential Logit | Li (2005) | Combined parametric and semi-parametric | Yang et al. (2013) |

Duration models are recommended for pavement performance estimation because the initiation time of pavement distresses is highly variable (Madanat et al. 2005), and they account for the irregularity inherently attributed to pavement condition inspections (Kobayashi et al. 2010). Duration models are appropriate in estimating pavement transition probabilities if frequent and continuous observations over a long time (i.e. 20 years) are available (Mauch and Madanat 2001). The common assumption of hazard and survivor models is that each condition state lowers down by only one state during a duty cycle (Khan et al. 2014), which disregards the condition states that deteriorate by more than one state. Data collected for a short window (i.e. less than 10 years) is usually left-censored (Mauch and Madanat 2001). As such, if the duration models are to be used to estimate pavement transition probabilities, data should be collected for a long time (i.e. 20 years) to reduce the potential data censoring. Based on the guidance from prior research, survival models are preferable for pavement condition models because they relax the assumption of the econometric models (condition states are independent and identically distributed). Also, to avoid data left-censoring, condition states can be assumed to transit midway between two consecutive inspection times (Mishalani and Madanat 2002). Lethanh and Adey (2012) used the Bayesian estimation approach to estimate the parameters of the econometric and duration models and found it to be more accurate than the MLE method.

#### 2.3.3.4. Semi-Markov Models

Unlike staged-homogeneous Markov models, semi-Markov models estimate the TPMs of pavement condition by dividing pavement lifetime into uneven intervals (holding times: the times that pavements take to completely leave their current states) corresponding to pavement performance curve. Figure 2.12 shows a graphical representation of pavement deterioration over time, in which condition states may take different holding time lengths until migrating to other condition states. To estimate a TPM for each interval, the holding time is assumed to follow a specific probability distribution. Semi-Markov models assume that holding times could follow any continuous-time distribution, so they are more flexible than the traditional Markov models that assume that holding times follow exponential distribution (Thomas and Sobanjo 2012).

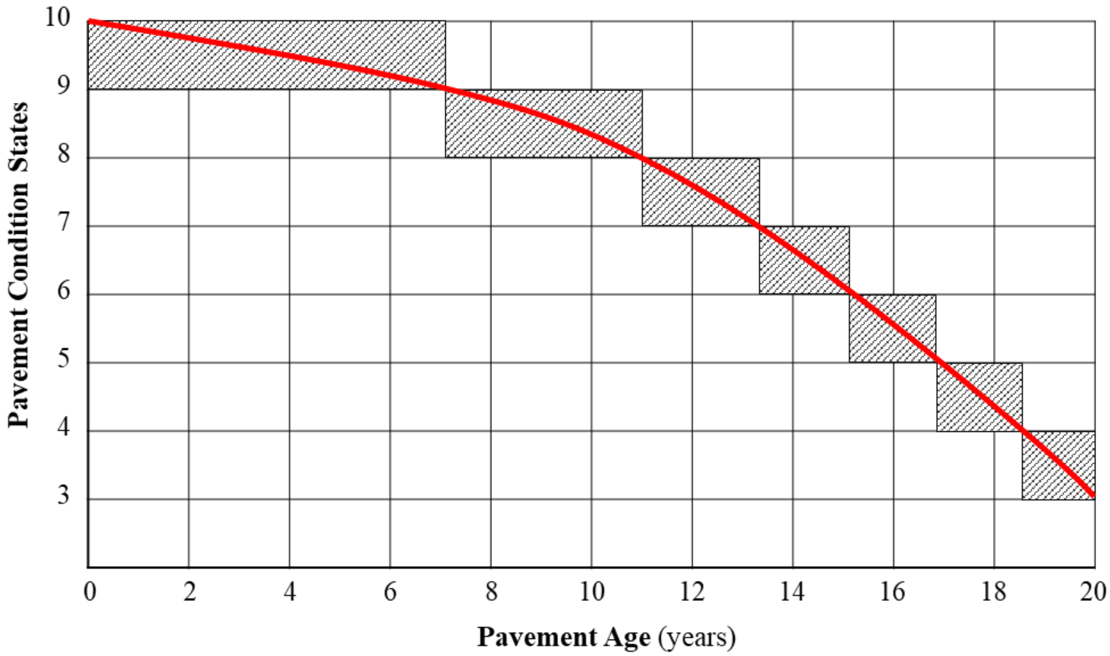


Figure 2.12. Graphical representation of the holding times of pavement condition states

Thomas and Sobanjo (2012) developed semi-Markov and homogeneous Markov models to predict pavement CI in the state of Florida. Data were obtained from the Florida Department of Transportation (FDOT) for more than 20 years. Fifty percent of the data was retained for validation and assessment of models' performance. Due to data limitation, seven condition states were created (from state 10 to state 4). The holding times were assumed to follow the Weibull distribution. The parameters of the Weibull distribution were estimated by minimizing the difference between the estimated and actual probability distributions. Wang et al.'s methodology of 1994 was used to estimate the homogeneous Markov model. Monte Carlo simulation was used to generate the TPMs and the probability distributions for both models. Both models were found to be statistically significant in terms of the Chi-square test statistic; however, the semi-Markov model was found to outperform its counterpart. Additionally, the semi-Markov model was found to over-predict pavement condition during the 7–11 years period due to data limitation during this period.

Semi-Markov models outperform homogeneous and staged-homogeneous Markov models because they relax the assumption of stationary transition probabilities; however, they require more extensive data to estimate the distribution of the holding times. With the continuous increase in the collected pavement condition data, semi-Markov models could be computationally less expensive than non-homogeneous Markov models (Nesbit et al. 1993). It is difficult to apply semi-Markov models at the pavement network level because the holding times may follow different distributions for different pavement sections (Ferreira and Santos 1999).

#### 2.3.3.5. Hidden Markov Models

Hidden Markov Models (HMMs) assume pavements have two types of condition states: observed states and hidden states. All pavement distresses such as cracking and potholes that can be inspected and measured are observable, whereas pavement condition indices such as PCI and PSR are unobservable or hidden states. Figure 2.13 shows the structural and temporal representation of the HMMs. The transition probabilities of the hidden states (i.e., $H_1, H_2, H_3$) are estimated using the data of the observed states (i.e., $O_1, O_2, O_3$) and the emission probabilities.

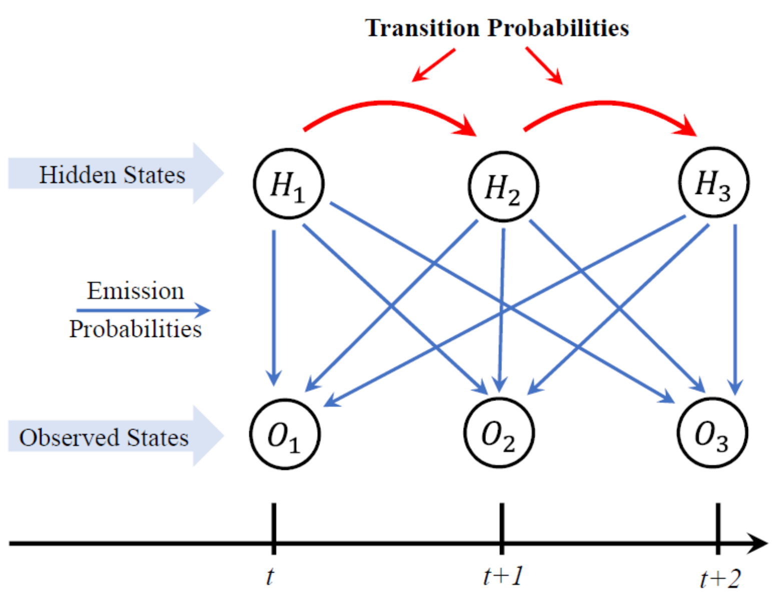



Figure 2.13. Hidden Markov Model diagram

Lethanh and Adey (2012) developed an HMM to predict pavement composite condition index (CCI) when data is incomplete. They assumed that pavement CCI represents roughness and cracking indicators. Pavement roughness data was assumed to be complete while cracking data was assumed to be incomplete. The hidden states were expressed by the CCI, whereas the observed states were expressed by the roughness and cracking indicators. The transition probabilities of the observed states were estimated using the multi-stage hazard methodology of Tsuda et al. (2006). The probability distribution of the hidden states was assumed to follow the exponential distribution. The probability distribution, transition probabilities, and exponential rate were estimated using the MLE method and the expectation-maximum algorithm. Data were acquired from Vietnam for Years 2001 and 2004. Pavement sections were grouped into five condition states based on the CCI. The variables traffic volume and pavement thickness were found to be statistically significant in estimating the transition probabilities of pavements in state 2, whereas the pavement thickness was the only statistically significant variable in estimating the transition probabilities of pavements in states 3 or 4. The estimated deterioration rates for pavements in states from 1 to 5 were found to be higher than the typical deterioration rates for typical pavements in Vietnam. To test the capability of the developed model for an incomplete data scenario, only pavement roughness data was used to estimate pavement deterioration rates. The estimated deterioration rates were found to be approximately similar to that when the entire data was used.

Lethanh and Adey (2013) extended their prior work to study to examine the accuracy of predicting pavement condition against the amount of available data. They assumed that the pavement CCI involves the roughness and texture depth of pavement. The transition probabilities of pavement roughness and texture depth were assumed to follow the exponential distribution. Lethanh and Adey (2013) used the data of their 2012 study. Four scenarios of incomplete data were created. The entire roughness data was used in all scenarios, but 100%, 50%, 25% and 10% of the texture depth data was used in scenarios 1, 2, 3 and 4, respectively. The results showed that the total duration taken by newly constructed pavements to move to state 5 is 14 years, which was consistent with pavement deterioration trends in Vietnam. Also, the predicted deterioration rates for pavements in states 1 and 2 were found to be similar across all scenarios. The predicted deterioration rates for pavements in states 3 and 4 for scenarios 2, 3 and 4 were found to be higher than that for scenario 1. These results indicate that the accuracy of pavement condition prediction improves if greater amounts of data are used for the HMM. Additionally, with 50% of the texture depth data the model was capable to predict pavement condition with 3% deviation from the predictions when the entire texture depth data was used. These results indicate that the required amount of data for modeling pavement performance can be reduced if the HMM methodology is used.

Lethanh et al. (2014) presented an HMM model to estimate pavement cracking rates and potholes for heavy traffic urban roadways in Japan from 2007 to 2011. The cracking rates were modeled using the Markov model developed by Kobayashi et al. (2012), while the potholes number was modeled using the Poisson process. Markov Chain Monte Carlo simulation and Gipp's sampler algorithm (i.e., Bayesian estimation approach) were used to estimate the models' parameters. Pavement sections were categorized into five condition states based on the cracking rates; where state 1 represents the lowest cracking rates and state 5 represents the highest cracking rates. The estimated deterioration rates associated with pavements in state 1 were found to be high with a holding time of 7 years. The probability of potholes occurrence is negligible during the first condition state but it goes up during the latter condition states.

Prior research (Lethanh et al. 2014; Lethanh and Adey 2012, 2013) developed HMMs to predict pavement condition when data is incomplete. Lethanh et al. (2014) did not test the validity of their model with actual data, while Lethanh and Adey (2012 and 2013) used data for only 2 years which may not be sufficient to capture the historical behavior of pavements. Additionally, the estimated transition probabilities were assumed to follow the exponential distribution, which means that pavement deterioration rates were assumed constants. Hence, further research is required to investigate the results of the HMMs when more extensive pavement condition data is employed, and other distribution functions are used.

### 2.3.4. Decision Tree

Based on the guidance and insights gained form the literature, a decision tree was developed to assist pavement asset managers in the selection of appropriate Markovian methodologies and TPM estimation methods. The criteria for selecting Markovian methodologies and TPM estimation methods are data availability and model assumption. Figure 2.14 shows the developed decision tree that will help highway agencies and future researchers choose the Markov methodologies that are appropriate for their data availability and desired level of accuracy and reliability. It can be noticed that if only two consecutive transitions of pavement condition are available, then the appropriate Markov methodology is the homogeneous one. On the other hand, if an extensive historical pavement condition data is available, including observations of the potentially influential variables, then the non-homogeneous Markov models are recommended in order to obtain more accurate and reliable pavement condition prediction models. In addition, the developed decision tree recommends TPM estimation methods for use in Markov models. If the historical pavement condition data are available but there is no information on explanatory variables exists, then the percentage transition method can be used with non-homogeneous Markov models, which may lead to less reliable models because of the lake of consideration of the explanatory variables.

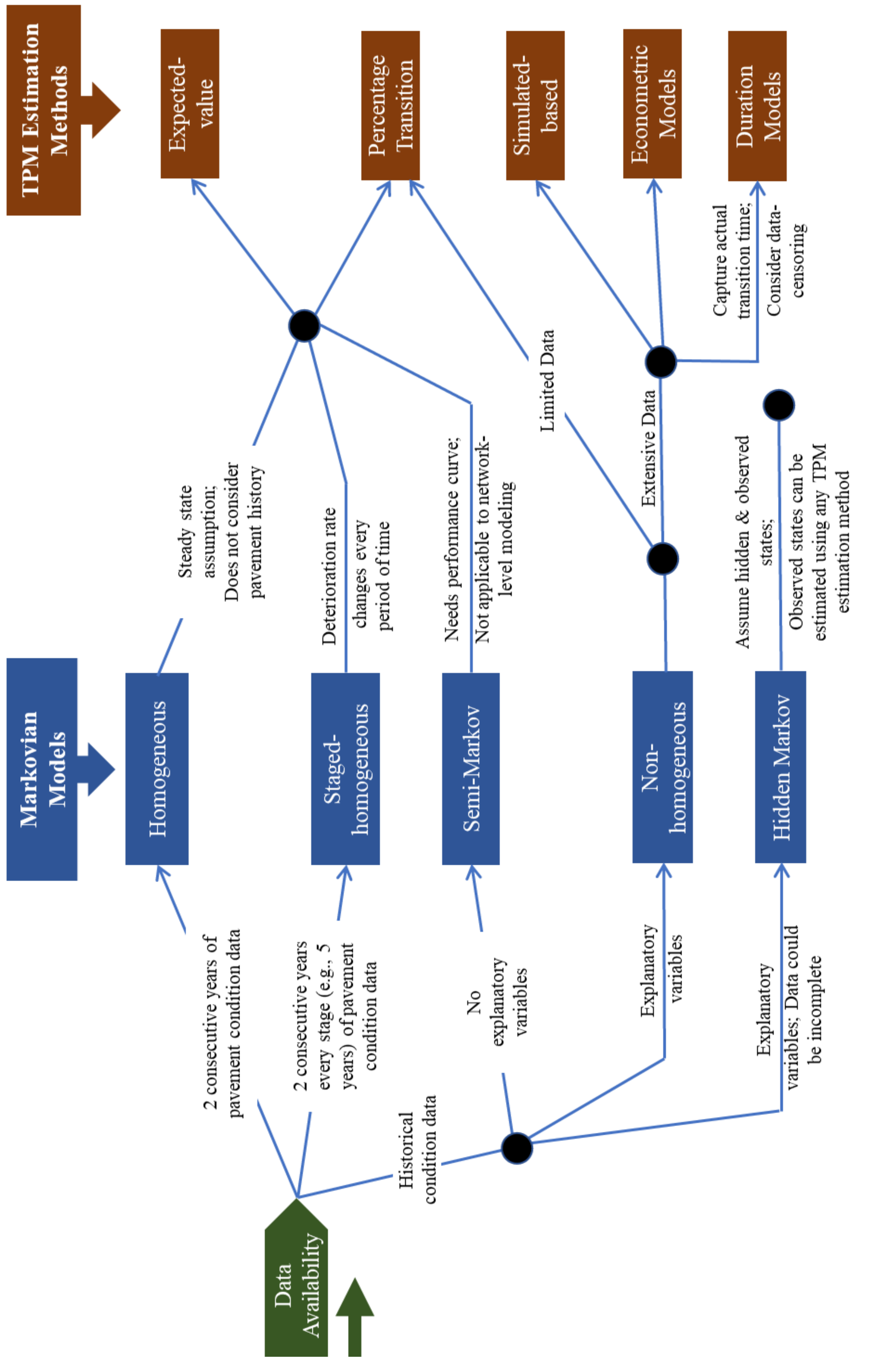



Figure 2.14. Decision tree for selection of Markov methodologies and TPM methods

## 2.4. Summary

The accuracy and reliability of pavement condition prediction depends on the employed Markov model type, the TPM estimation method, the correlated explanatory variables, and the quality and available amount of data. Markovian models and TPM estimation techniques need varying amounts of data. Some models such as non-homogenous Markov models need large amounts of data to yield accurate predictions, but that comes at the expense of data collection, storage, and management. On the other hand, other models such as homogeneous Markov models need smaller amounts of data and they are computationally more economical, but that is at the expense of prediction accuracy. Although the literature is rich in the discussion of Markov pavement performance models, several limitations were found. Previous studies assumed that the impact of pavement maintenance and rehabilitation can be captured in Markovian models by updating the condition state vector (S) every period of time when pavement condition observations are available. This assumption is valid only for short-term predictions and necessitates frequent monitoring of pavement condition. For long-term (during rehabilitation lifecycle) predictions, the effect of pavement preventive maintenance should be considered when estimating pavement transition probabilities. Additionally, prior research focused on estimating the TPM for Markov models, but exhibited gaps in estimating the number of condition states, the length of duty cycle, and the probability distributions. Further research is needed to estimate the impact of the number of condition states and the length of duty cycle on Markov model prediction accuracy and on the decision-making regarding the programming of pavement maintenance and rehabilitation treatments. The Bayesian estimation approach is more accurate than the MLE method in determining the globally optimal solution for the parameters of econometric and duration models (Lethanh and Adey 2012). Future research is needed to further investigate the accuracy of the estimated parameters in econometric and duration models when using the MLE and Bayesian estimation approaches. To contribute in bridging some of the abovementioned research gaps, Chapter 3 in this dissertation, provides a hybrid approach to account for the effect of maintenance into Markovian pavement performance models, while Chapter 4 presents a comparative analysis of Markovian methodologies regarding the prediction accuracy of pavement condition for different combinations of number of condition states and lengths of duty cycle.

# CHAPTER 3. HYBRID APPROACH TO INCORPORATE PREVENTIVE MAINTENANCE EFFECTIVENESS INTO PROBABILISTIC PAVEMENT PERFORMANCE MODELS

[A version of this chapter is under review at the Journal of Transportation Engineering, Part B: Pavements].[1]

Various methodologies are being developed to build and improve probabilistic pavement performance models that have high prediction capabilities. However, the effectiveness of preventive maintenance (PM) has not been adequately considered in such models due to the lack of historical PM data. Consequently, the predicted pavement condition is erroneous and often biased, which leads to non-optimal M&R decisions. This chapter introduces and validates a hybrid approach to incorporate the impact of PM into probabilistic pavement performance models when historical PM data is absent. The types of PM treatments and their times of application are estimated using two approaches: (1) Analysis of the state of practice of pavement maintenance through literature and expert surveys, and (2) Detection of PM times from probabilistic pavement performance curves. Using a newly developed optimization algorithm, the estimated times and types of PM treatments are integrated into pavement condition data. A non-homogeneous Markovian pavement performance model is developed by estimating the transition probabilities of pavement condition using the ordered-probit method. The developed hybrid approach and performance models are validated using cross-validation with out-of-sample data and through surveys of subject matter experts in pavement engineering and management. The results show that the hybrid approach and models developed predict probabilistic pavement condition incorporating PM effects with an accuracy of 87%.

## 3.1. Introduction

Highway agencies need effective pavement management systems to efficiently allocate their limited resources. Effective pavement management system requires pavement performance

[1] Yamany. M.S., and Abraham, D.M. Hybrid Approach to Incorporate Preventive Maintenance Effectiveness into Probabilistic Pavement Performance Models. Submitted to Journal of Transportation Engineering, Part B: Pavements, ASCE.

models of accurate and reliable predictability for pavement condition and performance, incorporating the positive and negative stochastic and deterministic effects of relevant influential factors. Pavement performance has been modeled using deterministic, Artificial Intelligence (AI) and probabilistic approaches. Deterministic models (e.g., Abaza 2004; Chu and Durango-Cohen 2008) assume that pavement condition can be predicted exactly. Therefore, they do not consider the inherent uncertainty and randomness of pavement condition. Although the Artificial Neural Networks (ANNs) models, the most common AI approach, have been developed in previous research (Plati et al. 2016; Amin and Amador-Jiménez 2017; Yamany et al. 2019b; Yamany et al. 2020b) to predict pavement performance, the interpretation of their findings is not easy and they are considered as black boxes (García de Soto et al. 2018).

Unlike deterministic and AI models, probabilistic models duly acknowledge the uncertainties inherently attributed to pavement condition data, and their results are easy to comprehend and interpret. Probabilistic models result in more reliable and robust pavement condition predictions than deterministic models (Rose et al. 2018; Qiao et al. 2019). Markov chains is the most extensively used probabilistic approach to model pavement condition. One category of Markov chains models is the non-homogeneous Markov models, which are accurate and realistic in estimating and predicting pavement performance because they account for the non-stationary nature of pavement deterioration. However, these models require a large amount of historical data, which is hindered by limited resources for data collection, storage and management.

Despite their superiority, past Markovian pavement performance models did not account for the influence of preventive maintenance (PM) on pavement performance because of the lack of PM data. PM is implemented on pavement surface to retard pavement deterioration but does not enhance its structural strength. Examples of PM treatments include thin overlay and micro-surfacing. Lack of consideration of the effectiveness of PM in Markovian pavement performance models could result in incorrect condition predictions, compromising on the reliability of model predictability, and/or failure to recognize the correlation between PM and pavement condition evolution. Additionally, failure to take PM effects into consideration could ultimately lead to less cost-effective and non-optimal maintenance and rehabilitation (M&R) strategies for pavements.

This chapter introduces and validates a hybrid approach to incorporate PM impact into probabilistic pavement performance models when historical PM data is absent or insufficient. The methodology is demonstrated using pavement condition data retrieved from the Long-Term Pavement Performance (LTPP) database for interstate flexible pavements. Table 3.1 shows the descriptive statistics of the key variables on which this data was collected. A non-homogeneous Markovian pavement performance model was developed and validated to estimate and predict pavement condition probabilistically. The methodology and performance model developed with the inclusion of PM impacts are useful for enhancing pavement management systems.

## 3.2. Prior Studies on Probabilistic Pavement Performance Modeling using Markov Chains

Markov chains have been extensively utilized to model pavement performance probabilistically. The accuracy of Markovian pavement performance models depends on the employed methodology, explanatory variables, and data quality and availability. Markovian models developed for pavement performance can be categorized into homogeneous Markov, staged-homogeneous Markov, semi-Markov, hidden Markov, and non-homogeneous Markov models.

Homogeneous Markov models are time-independent, i.e., the transition probabilities of pavement condition are constant over time. This type of models requires observations of pavement condition for only two successive transitions. Hence, several previous researchers including Wang et al. (1994), Macleod and Walsh (1998), Chou et al. (2008), Pulugurta et al. (2009), Mandiartha et al. (2010), Abaza and Murad (2010), Abaza (2014), Hassan et al. (2017a, b), Osorio-Lird et al. (2018) and Pérez-Acebo et al. (2018) chose to develop homogeneous Markov pavement performance models because of data limitation. Nevertheless, homogeneous Markov models suffer from a serious limitation because they assume that pavements deteriorate in a stationary or steady state process, which is contrary to the continuous natural change in pavement deterioration rates over time (Butt et al. 1987; Abaza 2016a).

Staged-homogeneous Markov models were introduced by Butt et al. (1987). These Markov models assume that the transition probabilities of pavement condition do not change significantly during a time interval or stage of five or six years, which is close to reality. Therefore, they consist of multiple homogeneous Markov models, each intended for each stage. Although staged-homogeneous Markov models are more reliable in predicting pavement condition than homogeneous Markov models, they fall short of capturing changes in pavement condition developed during the presumed stages. They also require more data than homogeneous Markov models, leading few studies to use this kind of Markov models such as Butt et al. (1987) and Abaza (2016a).

Compared to staged-homogeneous Markov models, semi-Markov models assume that the transition probabilities of pavement condition should vary according to the change in pavement performance over uneven intervals (holding times). Since semi-Markov models further relax the assumption of stationary transition probabilities by considering changes in pavement deterioration rates at unequal stages, they outperform the homogeneous and staged-homogeneous Markov models (Thomas and Sobanjo 2012). However, semi-Markov models require additional pavement condition data to estimate the holding times or the length of the uneven intervals. These models may overestimate pavement condition when insufficient pavement condition data is used (Thomas and Sobanjo 2012). Besides, these models cannot capture the incremental continual changes in pavement condition.

Hidden Markov models assume two condition state types: observed and hidden. Pavement distresses that can be inspected and measured such as cracking, potholes and roughness represent observed condition states, whereas pavement condition indices such as International Roughness Index (IRI) and Present Serviceability Rating (PSR) are unobservable or hidden condition states. The transition probabilities of hidden states are estimated using information on the observed states. Since hidden Markov models assist in mapping the relationship between the observed and hidden states, they can be employed to estimate pavement condition when pavement condition data are incomplete (Lethanh and Adey 2012, 2013; Lethanh et al. 2015).

Non-homogeneous Markov models are time-dependent, i.e., the transition probabilities of pavement condition change over time. In addition to capturing the unsteady effect of relevant explanatory variables, non-homogeneous Markov models consider the nonstationary property or the changes in pavement deterioration rate over time. Although these models accurately fit the random behavior of pavements compared to the other counterpart models, they requisite extensive computation and large amounts of data. Many highway agencies are hindered from taking advantage of these models until they have sufficient historical data to estimate such models. Although prior non-homogeneous Markov models (Madanat et al. 1995b; Li et al. 1996; Yang et al. 2005, 2006; Kobayashi et al. 2010; Tabatabaee et al. 2013; Abaza 2017a) outperform their counterparts by capturing the non-stationary nature of pavement deterioration, they assume that pavement condition deteriorates over time and does not improve, i.e., they do not account for PM effects over pavement life, which could lead to less cost-effective and non-optimal M&R decisions.

### 3.3. Research Methodology

This study uses a novel hybrid approach that incorporates PM impact into probabilistic pavement performance models. Six major research tasks form the core of this approach (Figure 3.1): data collection and analysis; data simulation; estimation of initial times for PM treatments; data generation; estimation of transition probability matrix (TPM); and validation of the approach and models developed, mathematically and by a survey of Subject Matter Experts (SMEs) with the experience in pavement engineering and management. The ensuing subsections discuss the proposed methodology and its components in detail.


Figure 3.1. Research methodology

### 3.3.1. Data Collection and Analysis & Survey 1

In the United States, interstate highways have the highest percentage of vehicle miles traveled compared to other functional classes. Of these interstate roads, 76% are flexible asphalt and composite roadways (FHWA 2017). Hence, more historical data is expected to be collected, stored and managed for interstate roadways across the U.S. Pavement condition data for interstate flexible pavements (black-topped roads that include asphalt and composite) from 1989 to 2016 were retrieved from the LTPP database. Details about the LTPP database can be accessed through the webpage of the U.S. Department of Transportation, Federal Highway Administration. To ensure the applicability of the proposed methodology to more than one state/location, data can be collected from multiple states. One criterion for the selection of appropriate states is that the selected states should be located in the same climate region so that these states most likely use similar standard specifications for design, construction, and/or maintenance and rehabilitation of pavements. According to the USGCRP 2018, the U.S. is divided into 10 climate regions (Reidmiller et al. 2018). In this study, transportation agencies from the Midwest region provided the highest percentage (about 63%) of responses to Survey 1 (conducted as part of this study to collect information on pavement condition and maintenance treatments). Therefore, pavement condition data were collected from the eight Midwest states: Indiana, Illinois, Wisconsin, Michigan, Ohio, Minnesota, Iowa and Missouri.

According to Survey 1, which was created and sent to the 50 State Transportation Agencies (STAs) to collect data regarding pavement condition and PM, around 58% of the respondents use one condition indicator to represent pavement condition. Moreover, about 60% of the respondents employ the IRI as the pavement condition indicator. Further details about Survey 1 will be discussed later. Hence, the collected data includes information about one pavement condition indicator, namely IRI as the response variable. The collected data includes six explanatory variables as listed in Table 3.1. Although other variables, such as the pavement structure (number, type and thickness of pavement layers), affect pavement condition, they were not considered in this dissertation because STAs have different standards and specifications for the design, construction and maintenance of pavements.

Data were pre-processed by detecting and cleaning the extreme outliers and high-leverage points using the methodology employed by Belsley et al. (1980), Luo (2013) and Ahmed et al. (2016). The outlier point has a studentized residual out of $\pm 3$, whereas the high-leverage point has a statistical leverage greater than $2p/n$, where $p$ is the number of independent variables including the constant term, and $n$ is the number of observations (1159). The number of observations was 966 after data had been cleaned. Table 3.1 shows the descriptive statistics of the data collected.

Table 3.1. Descriptive statistics

| Variable | Description | Mean | STD | Min | Max |
|---|---|---|---|---|---|
| IRI | International Roughness Index (in/mi) | 80.9 | 27.8 | 31.3 | 194.9 |
| Age | Years since construction or rehabilitation | 9.6 | 6.6 | 0.0 | 27.0 |
| AAP | Annual Average Precipitation (inches) | 38.1 | 7.7 | 24.4 | 61.4 |
| AAT | Annual Average Temperature (°F) | 50.5 | 3.3 | 42.8 | 59.7 |
| AAFI | Annual Average Freezing Index (°F days) | 788 | 393 | 70 | 1924 |
| AADTT | Annual Average Daily Truck Traffic | 2169 | 802 | 249 | 5115 |
| ESALs | Equivalent Single Axle Loads (18-Kip) | 1115 | 534 | 122 | 3195 |

### 3.3.2. Data Simulation

Since the data acquired for interstate flexible roads are limited, an extensive amount of data was generated through simulation to implement the proposed methodology and estimate the non-homogenous Markovian models. Data was simulated to facilitate the incorporation of the PM treatments into pavement condition data. To simulate pavement condition in terms of IRI, a regression model was developed to estimate the IRI in relation to the statistically significant explanatory variables. To ensure the validity of the regression model, ten percent of the gathered data was reserved for subsequent use in a cross-validation process. The Root Mean Square Error (RMSE) was calculated to check the performance of the developed model. The smaller the RMSE value, the greater the accuracy and better fit of the model.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i}^{N}(IRI_{Actual} - IRI_{Predicted})^2} \tag{3.1}$$

where $N$ is the number of observations set aside for validation purposes. The probability distribution of the explanatory variables was identified by fitting different probability distributions for each variable. The probability distribution that results in the least negative log-likelihood was selected. Finally, pavement sections were generated using the developed regression model and selected probability distributions for the explanatory variables.

### 3.3.3. Initial Times for Preventive Maintenance Treatments

Three approaches were used to estimate the initial times for PM treatments over pavement service life: literature search, questionnaire Survey 1, and detection of PM times from probabilistic pavement performance curves. Relevant literature was investigated to find information on the actual or recommended PM treatments and their timings.

Survey 1 was designed and sent out to 50 STAs in the U.S. to seek information about practices in the pavement condition and maintenance. The questions of this survey covered the following themes: (1) PM treatments applied to interstate flexible pavements, (2) Timings of PM treatments, (3) Criteria for selecting treatments and their timings, (4) Pavement condition indicators, and (5) Short-term and long-term effectiveness of PM treatments. Eighteen STAs responded to the survey including five from the Midwest states (Figure 3.2). The respondents have expertise in pavement engineering and management ranging from assistant pavement design engineer to director of pavement asset management.

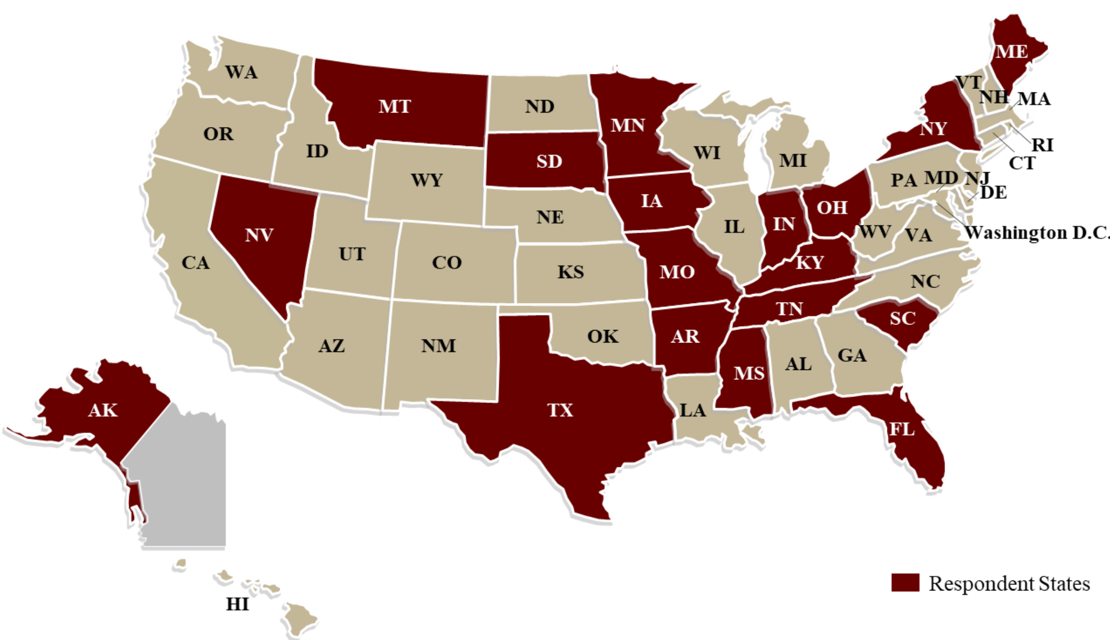



Figure 3.2. Geographical representation of STAs that responded to Survey 1

To detect the times of PM treatments, pavement performance curves were developed using the actual pavement condition data. A probabilistic model (Model A) was proposed to account for the uncertainty attributed to the data, and to capture more detail in pavement performance. Pavement IRI was discretized into $n$ number of condition states based on the IRI range, available observations and representation percentage of the condition states. The first state represents best condition, whereas the last state $n$ denotes worst condition.

The ordered-probit approach was used to develop Model A because the condition states are ordinal discrete variables. This approach assumes that the probability of occurrence of an outcome can be estimated in relation with a latent variable $U$ and a threshold parameter $\mu$. $U$ is a latent continuous variable takes on values $\infty$ to $-\infty$. Whilst $\mu$ is a cutoff value that separates the probabilities of possible outcomes of $Y$ (e.g. $y_1, y_2, \ldots$), and is estimated by mapping the relationship between the outcomes of $Y$ and $U$. If the first threshold begins at zero, $\mu_0 = 0$, the required number of thresholds is the number of outcomes minus 2. For further details on the ordered-probit approach the reader is referred to Washington et al. (2011).

The statistical significance of the ordered-probit model was assessed using the Likelihood Ratio test by calculating the Chi-square ( $\chi^2$) as

$$\chi^2 = -2\left[LL(\beta_R) - LL(\beta_U)\right] \qquad (3.2)$$

where $LL(\beta_R)$ is the log-likelihood of restricted model in which predictors' parameters are set to zero. Whereas $LL(\beta_U)$ is the log-likelihood of unrestricted model in which predictors' parameters are counted. Models are deemed statistically significant if the estimated $\chi^2$ is greater than a calculated critical $\chi^2$ at a 95% confidence level (Washington et al. 2011).

Ten percent of the actual data was retained to cross-validate Model A. The RMSE was calculated to assess the model's validity as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{1}^{N}(CS_{Actual} - CS_{Predicted})^2} \tag{3.3}$$

where $CS$ is pavement condition state $(1 \dots n)$, and $N$ is the number of observations retained for validation purposes.

### 3.3.4. Data Generation Including Types and Times of Preventive Maintenance

The estimated initial times of PM were integrated into the simulated pavement condition data. An optimization model was created to guarantee that the generated data is comparable to the actual LTPP data. The objectives of the optimization model are: (1) to ensure that PM treatments are assigned to pavement sections at times aligned with the initial times estimated for the PM, and (2) to meet the limited funds allocated to pavement PM. The design variables of the optimization model are PM treatment types, application times, costs, effectiveness, pavement condition threshold, and available annual funds. The data on the design variables were obtained from the collected LTPP data, literature search, questionnaire Survey 1, and detection of PM times from probabilistic pavement performance curves. The future value of PM costs was calculated as:

$$FV = PV \times (1 + r)^n \tag{3.4}$$

where FV and PV denote the future and present values of PM costs in USD, respectively; $r$ represents the interest rate, and $n$ is the number of periods/years. The annual interest rates were obtained from the National Highway Construction Cost Index (NHCCI 2019).

Half of the STAs that responded to Survey 1 reported using the worst-first approach among 15 other criteria to make decisions on pavement maintenance. Hence, a greedy algorithm (Figure 3.3) was designed and used to optimize the types and times of PM treatments for pavement sections eligible for PM. To obtain pavement performance comparable to the pavement performance developed using the LTPP data, different percentages of the required annual funds were specified to be the available $Budget(i)$ in the greedy algorithm (Figure 3.3). For each percentage of annual funds, a complete cycle of the greedy algorithm was run. Then, ordered-probit Models B were developed using the data resulted from the optimization runs. The reduction in the percentage of funds decreases the amount of PM assigned to pavement sections, and hence affects the overall pavement performance. The percentage of funds that resulted in performance curves best fit the performance curves developed using the LTPP data was selected to build the final Model C.

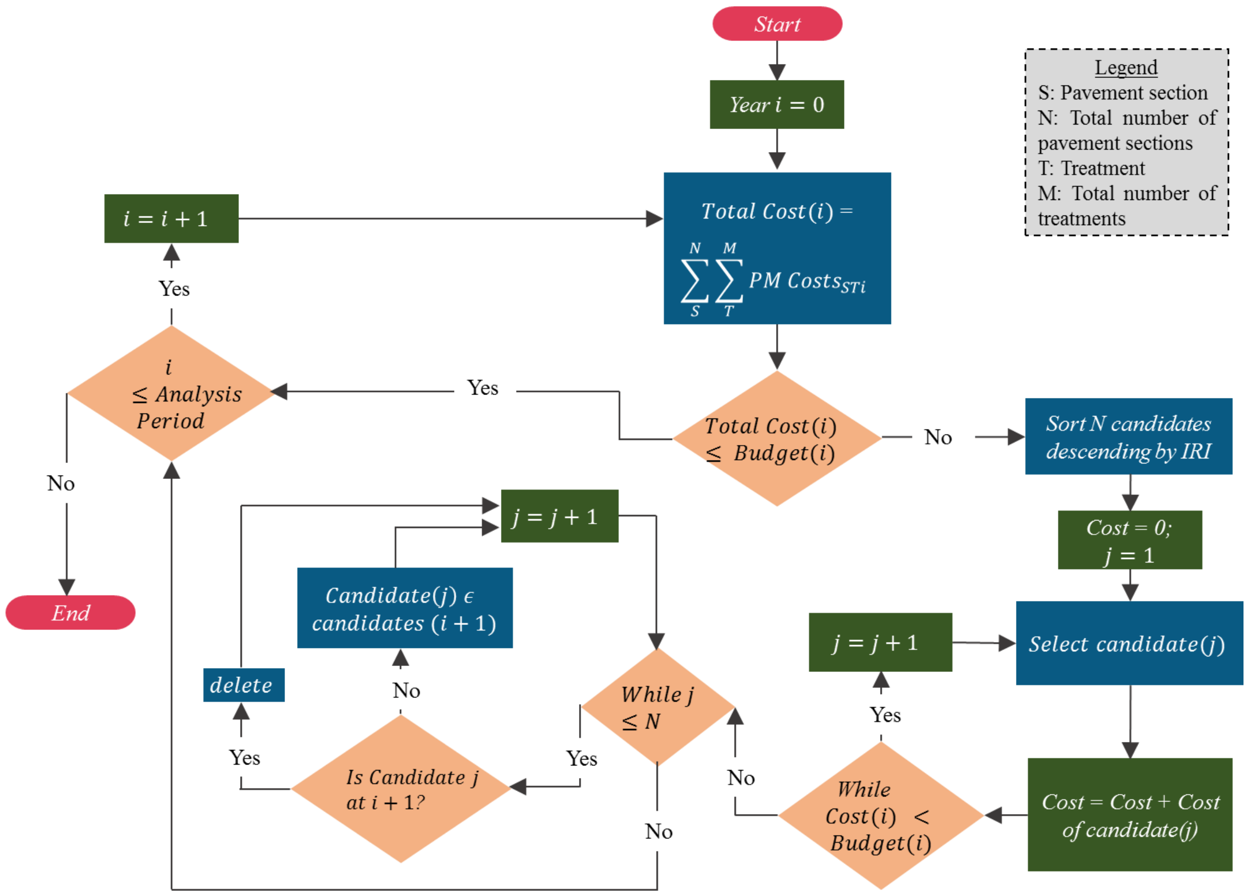


Figure 3.3. Greedy algorithm flow chart

### 3.3.5. Transition Probability Matrix Estimation

Using the generated pavement condition data that includes PM data, an ordered-probit Model C was developed using the same methodology employed to develop Model A. Model C was then validated using cross-validation with a 10% out-of-sample data to verify its predictability. Then, it was used to determine the condition state vector ($S^i = \{S^{i1}, S^{i2}, S^{i3}, \dots S^{in}\}$) at each year $i$; where $S^{in}$ is the probability of pavement condition being in state $n$ at year $i$. The transition probability matrix (Equation 3.5) of pavement condition states at each time $i$ was then calculated by dividing the condition state vector $S^i$ at each year $i$ by the condition state vector $S^{i-1}$ at each year $i-1$.

$$\mathbb{P} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & \cdots & P_{1n} \\ P_{21} & P_{22} & P_{23} & \cdots & P_{2n} \\ P_{31} & P_{32} & P_{33} & \cdots & P_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ P_{n1} & P_{n2} & P_{n3} & \cdots & P_{nn} \end{bmatrix} \tag{3.5}$$

where $\mathbb{P}$ is a square matrix of pavement transition probabilities. The elements on the diagonal of this matrix, i.e., $P_{11}, P_{22}$ … $P_{nn}$ represent the probabilities that pavements will stay at the same condition state after one year. The elements above the diagonal such as $P_{12}$ and $P_{23}$ denote the probabilities that a pavement will transition from its present condition state to a worse condition state after one year (e.g., from state 2 to state 3 as denoted by $P_{23}$). Elements below the diagonal like $P_{21}$ and $P_{32}$ represent the probabilities that pavements will migrate from their current condition state to better condition states after one year (e.g., from state 2 to state 1 as denoted by $P_{21}$). These elements (below diagonal) are typically specified to be equal to zero, indicating that the effect of pavement maintenance is not considered in Markov models developed for pavement condition prediction (Ortiz-García et al. 2006). The hybrid approach proposed in this research will therefore help to estimate the *P* values below the diagonal, which represent the probability of pavement condition improvement due to PM applications.

## 3.4. Results and Discussion

The results of the research methodology and models developed are presented and discussed in the following subsections. The validation of the methodology and models is presented as well.

### 3.4.1. Data Simulation

To simulate the actual pavement condition, regression models were developed to estimate the IRI in relation to influential independent variables. These regression models were created with different mathematical formulations (e.g., Linear, Power, and Exponential). Two explanatory variables, *Cum. AAFI* (i.e., AAFI×Age) and *Cum. AADTT* (i.e., AADTT×Age) were created to capture the interaction between the variables listed in Table 3.1. The *Cum. AADTT* was also created as a proxy for the cumulative AADTT data that was not available. The best statistical model was selected based on the coefficient of determination ($R^2$) and the t-statics and p-values of the independent variables at a 95% confidence level. The exponential multiple regression model represented in Equation 3.6 was selected to simulate pavement condition.

$$\text{IRI} = Exp(4.0875 + 0.12224\, Cum.\, AAFI + 0.07876\, Cum.\, AADTT) \tag{3.6}$$

where *Cum. AAFI* is equal to AAFI × pavement age in $10^4$ °F days; and *Cum. AADTT* is equal to AADTT × pavement age in $10^4$ trucks.

The t-statistics and p-values for the *Cum. AAFI* and *Cum. AADTT* were found to be 5.66 and <0.05, and 10.96 and <0.05, respectively, indicating their statistical significance at a 95% confidence level ($t_{critical}$ = 1.96). However, at the same confidence level, the other variables listed in Table 3.1, (AAP, AAT and ESALs) were found to be statistically insignificant with t-statistics smaller than $t_{critical}$.

The exponential multiple regression model (Equation 3.6) was validated using the reserved 10% of the actual pavement condition data. The RMSE was found to be 18.72 in/mi. Thus, the model has a satisfactory performance in predicting pavement IRI with respect to the *Cum. AAFI* and *Cum. AADTT*. To further check the model predictability, the estimated RMSE was compared with that of prior similar models. Past pavement condition prediction models of Ziari et al. (2016) and Dalla Rosa et al. (2017) have RMSE values of 20 in/mi. and 22.17 in/mi., respectively. Although these models differ from the current model regarding the data used in models' development and explanatory variables, the estimated RMSE of the current model indicates that its predictability is consistent with the literature.

Different probability distribution functions were tested to fit pavement age, *AAFI* and *AADTT*. The optimal function was chosen based on the value of the negative log-likelihood. The optimal distribution functions were chosen based on the minimum value of the negative log-likelihood. The minimum values of the negative log-likelihood for pavement age, *AAFI* and *AADTT* were found to be equal to $3.04\times10^3$, $7.05\times10^3$ and $7.61\times10^3$, respectively, with the Gaussian Mixture functions of 3, 3 and 2 components, respectively for each variable. Table 3.2 shows the properties of the Gaussian Mixture functions selected for those three variables.

Table 3.2. Properties of Gaussian Mixture functions selected for pavement age, AAFI and AADTT

| Variable | Gaussian Mixture Properties | | | | |
|---|---|---|---|---|---|
| | Number of Components | Component | Proportion | Mean | Standard Deviation |
| Pavement age | 3 | 1st | 0.145 | 01.55 | 00.39 |
| | | 2nd | 0.484 | 07.06 | 07.95 |
| | | 3rd | 0.369 | 16.22 | 29.52 |
| AAFI | 3 | 1st | 0.271 | $0.36\times10^3$ | $1.82\times10^4$ |
| | | 2nd | 0.416 | $0.76\times10^3$ | $1.55\times10^4$ |
| | | 3rd | 0.311 | $1.19\times10^3$ | $1.08\times10^5$ |
| AADTT | 2 | 1st | 0.919 | $2.00\times10^3$ | $2.41\times10^5$ |
| | | 2nd | 0.081 | $4.05\times10^3$ | $4.78\times10^5$ |

The Indiana Department of Transportation (INDOT) owns and operates approximately 5,500 miles of interstate roadways. Hence, to generate a reasonable number of pavement sections for the current research, 5,500 interstate flexible pavement sections (each one mile long) were generated according to the identified probability distributions and the developed exponential multiple regression model.

### 3.4.2. Initial Times for Preventive Maintenance Treatments

Relevant literature was reviewed for information on the types and times of PM treatments. Table 3.3 summarizes the types and times of PM treatments applied to flexible pavements.

Table 3.3. Types and times of PM treatments

| Treatment | Geoffroy (1996) | Mamlouk and Zaniewski (1999) | Labi and Sinha (2003) | Peshkin et al. (2004) | Ong et al. (2010) | INDOT Design Manual (2013) |
|---|---|---|---|---|---|---|
| **Crack Sealing (Rout and Seal)** | N/A | 8 and 16 yrs | 3 yrs for 1st application and every 4 yrs afterwards | 1 to 3yrs | IRI < 60.9in./mi | Years of 3, 6, 9, 12, 15, 18 |
| **Crack Filling** | Every 4 yrs | N/A | N/A | N/A | IRI < 60.9in/mi. | N/A |
| **Fog Seal** | N/A | 3 and 6 yrs | N/A | 0 to 3 yrs | N/A | N/A |
| **Chip Seal** | N/A | 9 and 18 yrs | 7 yrs | 2 to 5 yrs | IRI < 73.5in/mi. | N/A |
| **Slurry Seal** | N/A | 9 and 18 yrs | N/A | 2 to 6 yrs | IRI < 73.5in/mi. | N/A |
| **Scrub Seal** | N/A | N/A | N/A | 2 to 6 yrs | N/A | N/A |
| **Sand Seal** | N/A | N/A | 12 yrs | N/A | N/A | N/A |
| **Micro-surfacing** | N/A | N/A | 15 yrs | 3 to 7 yrs | N/A | IRI< 130 in/mi. |
| **Ultrathin Bonded Wearing Course (UBWC)** | N/A | N/A | N/A | 2 to 6 yrs | N/A | IRI < 140 in/mi. |
| **Thin HMA Inlay or Overlay less than or equal to 1.5 in** | 12 yrs | N/A | 17 to 20 yrs | 5 to 8 yrs | IRI<79.9in/mi. | IRI < 150 in/mi. |

*Note: N/A indicates that the corresponding treatment is not considered as preventive maintenance, or there is no information presented about this treatment in that study.*

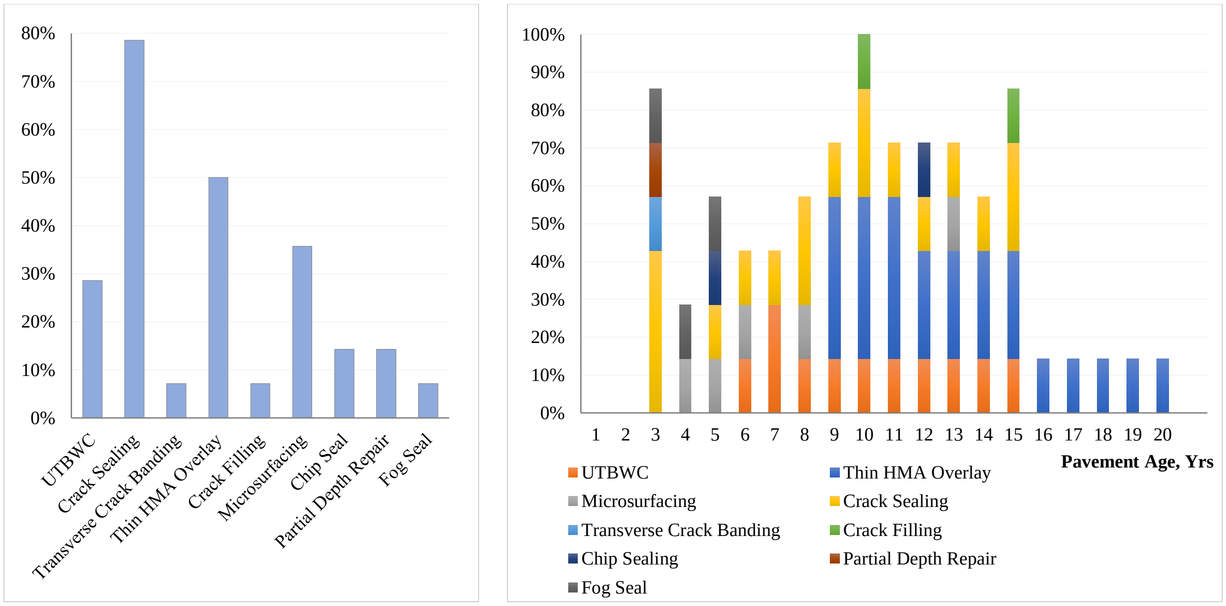
The effectiveness of PM treatments or the improvement in pavement condition due to PM was investigated from the relevant literature. The improvement or performance jump in pavement IRI after the implementation of micro-surfacing, UTBWC and thin HMA overlay were calculated as provided in Labi and Sinha (2003), Ong et al. (2010), and Ji et al. (2015). According to the INDOT design manual (2013), crack sealing, crack filling, fog seal, and chip seal do not improve the condition of pavement surface, but they help in retarding pavement deterioration.

The results of Survey 1 showed that about 58% of the responding STAs use one pavement condition indicator to assess pavement condition and make respective maintenance decisions. However, some STAs use up to four pavement condition indicators. For instance, Texas DOT uses the IRI, ride quality index, condition score and distress index. In addition, the survey results indicated that among various pavement condition indicators such as Pavement Condition Index

(PCI) and Pavement Serviceability Index (PSI), the IRI is the most widely used (60%) across the responding STAs.

Survey 1 also helped to generate information on PM treatments used in maintaining interstate flexible pavements. Figure 3.4(a) shows the distribution of PM treatments used across the STAs respondents. Crack sealing was reported as the most frequently used PM treatment, most likely due to the low cost of its application and recognized effectiveness in retarding pavement deterioration (Peshkin et al. 2004). The second most frequently used treatments are thin overlay, micro-surfacing and UTBWC. Figure 3.4(b) shows that different PM treatments implemented over time until pavement age of 20 years [typical pavement design life (Morian et al. 2005; Ceylan et al. 2009; Santos and Ferreira 2013)]. The first PM application is due at the age of 3 years when most STAs respondents reported using crack sealing. The most frequent times for PM treatments are 3 years – usually for first application of crack sealing, 9 to 13 years – nearly mid-age of pavement, and 15 years.

The STAs were asked to identify their criteria for choosing the type and time of application of PM treatments. They were also asked to rate their criteria on a scale from 1 to 5; where 1 is the least important, and 5 is the highest important. Sixteen pavement maintenance decision criteria were identified by the participants in Survey 1, and the most important criteria were found to be "pavement condition" and "maintenance cost" with mean importance ratings equal to nearly 4.1 and 3.1, respectively.




a) PM Treatments (17 respondents) b) Times for PM (15 respondents)

Figure 3.4. Distribution of types and times of PM treatments across STAs

The main goal of PM is to improve pavement surface condition and keep pavements in a state of good repair. One way to identify the actual times of PM is to investigate pavement performance curves and detect the times when pavement condition starts to improve. Thus, pavement IRI was first discretized into five condition states as follows: state 1 $(IRI \leq 55)$, state 2 $(55 < IRI \leq 70)$, state 3 $(70 < IRI \leq 95)$, state 4 $(95 < IRI \leq 120)$, and state 5 $(IRI > 120)$. The number of condition states and their ranges were determined based on the available observations, IRI variation, and the percentage of each state from the data. An ordered-probit Model A was then developed to estimate the probability of pavement condition being in each condition state. Finally, pavement performance curves were developed from which the actual times of PM treatments were detected.

Table 3.4 shows the estimation results of the developed Model A. Based on their t-statistics and p-values, the *Cum. AADTT* and *Cum. AAFI* along with the threshold parameters were found to be statistically significant at a 95% confidence level. The estimated $\chi^2$ was found to be equal to 313.59, which is greater than the critical $\chi^2$ (11.07) indicating that Model A is statistically significant at a 95% confidence level. The positive signs of the *Cum. AADTT* and *Cum. AAFI* indicate that an increase in *AADTT* or *AAFI* and/or pavement age increases the probability of pavement deterioration over time, which is in line with the engineering intuition and prior Markovian models discussed earlier.

Table 3.4. Estimation results of Model A

| **Variable** | **Description** | **Parameter Estimate** | **t-statistic** | **p-value** |
|---|---|---|---|---|
| Constant | | 0.32143 | 5.11 | <0.001 |
| Cum. AADTT | Cumulative Annual Average Daily Truck Traffic in 10,000 Trucks | 0.29414 | 10.25 | <0.001 |
| Cum. AAFI | Cumulative Annual Average Freezing Index in 10,000 °F days | 0.43719 | 6.22 | <0.001 |
| $\mu_1$ | Threshold Parameters | 0.83518 | 20.12 | <0.001 |
| $\mu_2$ | | 2.15401 | 44.12 | <0.001 |
| $\mu_3$ | | 2.83864 | 45.13 | <0.001 |
| Log-likelihood of restricted model = -1422.78 | | | | |
| Log-likelihood of unrestricted model = -1265.98 | | | | |
| Estimated $\chi^2 = -2\,[LL(\beta_R) - LL(\beta_U)] = 313.59$ | | | | |
| Degree of freedom (unrestricted number of parameters – restricted number of parameters) = 5 | | | | |
| Critical $\chi^2$ at a 95% confidence level = 11.07 | | | | |

Model A was used to predict pavement condition using the 10% retained data; then the predictions were compared with the actual pavement conditions. The RMSE value was found to be 0.75, indicating that the model can predict the mean condition state of pavement with a mean error of 0.75 or approximately 15 in/mi. Compared with the models of Ziari et al. (2016) and Dalla Rosa et al. (2017) of RMSE of 20 in/mi. and 22.17 in/mi., respectively, Model A can predict pavement condition probabilistically with satisfactory accuracy.

Probabilistic pavement performance curves were developed for the five condition states over pavement age (Figure 3.5). The probability of pavement condition being in states 1 and 2 (best condition) decreases over time, whereas the probability of being in state 5 (worst condition) increases over time. The probability of pavement condition being in states 3 and 4 increases over time during the early age (less than 10 years) of pavement life, but gradually drops after different points in time.

A closer investigation of the pavement performance curves in Figure 3.5 helps detect the actual times for PM applications. The annotated boundaries a, b and c show the time windows for possible application of PM. Figure 3.5 indicates that the probability of pavement condition being in states 3 or 4, i.e., good or fair conditions increases over time from the beginning of pavement life until the times when these probabilities start to decrease. These are the times when PM treatments might have been implemented to keep pavement condition at a desirable state. The last applicable time to implement PM, when pavement is most likely to be in good condition (70 < IRI <= 95 in/mi.), is around the 10th year of pavement age. The last chance to apply PM could be at about 15 years of age when the probability of being in state 5 begins to exceed the probability of being in state 4.

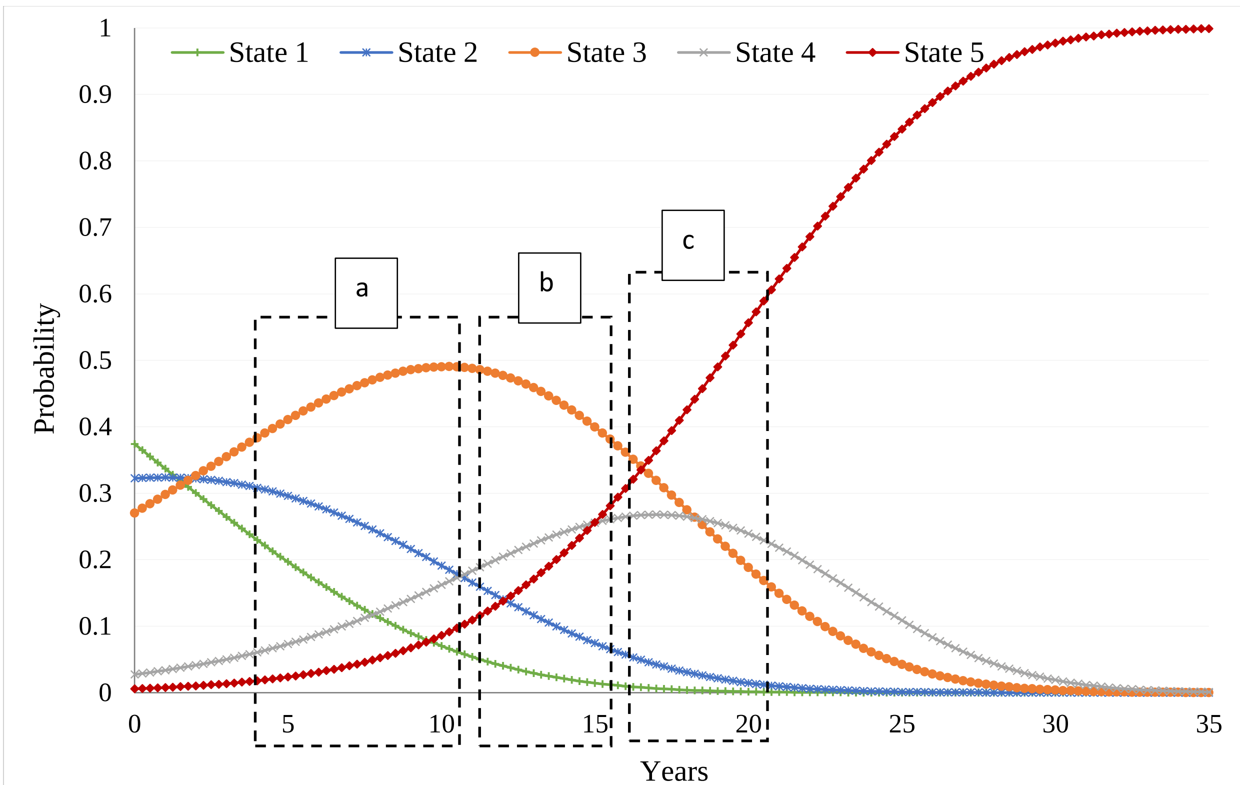



Figure 3.5. Probabilistic pavement performance curves using Model A

The probabilities of condition states at each time $i$ shown in Figure 3.5 represent the state vector $S^i$ of pavement condition, where $i$ denotes the time from 0 to pavement age. The transition probabilities of pavement condition states were calculated by dividing each consequent state vectors $(S^{i+1}/S^i)$. Also, the transition probability curves for condition states 3, 4 and 5 were developed. Then approximate ranges of pavement age during which PM treatments might have been implemented have been identified. The range of 4 to 10 years is the time window to apply PM that results in the highest probability of pavements in state 3 to stay in the same state or move to states 2 or 1. The range of 11 to 15 years is the time window to apply PM that yields the highest probability of pavements in state 4 to stay in the same state or transition to states 3, 2 or 1. The range of 16 to 20 years is the time window to apply PM that results in the highest probability of pavements in state 5 to move to states 4, 3, 2, or 1.

The effectiveness of the PM, i.e., the improvement in pavement condition after PM implementation was calculated from the estimated transition probabilities of condition states 3, 4 and 5 as they are the candidate states for PM. Equation 3.7 shows the estimation of pavement condition improvement or the drop in the IRI after PM application.

$$D_{IRI,i,t} = \sum_{j=1}^{j=i-1} P_{i,i-j} \times \left(IRI_i - IRI_{i-j}\right), \qquad \forall\, i = 3, 4 \text{ and } 5 \qquad (3.7)$$

where $\mathrm{D}_{\mathrm{IRI},i,t}$ is the drop in the IRI value (in/mi.) for pavement in condition state $i$ at time $t$; $j$ takes on the values from 1 to 4; $P_{i,i-j}$ is the transition probability of pavement condition from state $i$ to state $i-j$. $\mathrm{IRI}_i$ and $\mathrm{IRI}_{i-j}$ are the values of IRI (in/mi.) for pavements in condition states $i$ and $i-j$, respectively. The results showed that the highest improvement in pavement condition occurs when pavements in state 3 receive PM as early as 4 years of age, and when pavements in states 4 and 5 receive PM at approximate ages 13 and 17 years, respectively.

Based on the literature search (Table 3.3) and results of Survey 1, the most widely used PM treatments for interstate flexible pavements are crack sealing, micro-surfacing, thin overlay, and UTBWC. Since the crack sealing does not cause improvement in pavement condition, the micro-surfacing, thin overlay and UTBWC were used for pavement condition data generation. Based on the literature search, the micro-surfacing, UTBWC and thin overlay treatments can be implemented if pavement IRI < 130 in/mi., < 140 in/mi. and < 150 in/mi., respectively. The initial times for these treatments were assumed to follow the normal distribution, which is commonly used to represent uncertainties in engineering problems. The standard deviation of the time for each treatment was estimated as one-fourth of the range (Taylor 2013) of treatment application times. The ranges and standard deviations of application times in years for the micro-surfacing, UBWC and thin overlay are 4 – 10 and 1.5, 11 - 15 and 1, and 16 - 20 and 1, respectively for each treatment type. Finally, the effectiveness of these PM treatments was calculated as the average of the estimated drop in IRI ($D_{IRI,i,t}$) and the effectiveness calculated from Labi and Sinha (2003), Ong et al. (2010) and Ji et al. (2015).

### 3.4.3. Data Generation Incorporating Types and Times of Preventive Maintenance

The developed greedy algorithm was employed to optimize PM treatments subject to different percentages of funds (100% to 40%) allocated for PM. The optimization of PM at each percentage of funds results in pavement condition data includes *IRI, Cum. AADTT, Cum. AAFI*, and optimal times and types of PM treatments. These data were then used to develop ordered-probit models (Models B) to estimate pavement performance. To be comparable with Model A, the explanatory variables used to create Models B are *Cum. AADTT* and *Cum. AAFI*, without considering the effect of PM.

Using Models B, pavement performance curves were developed, closely investigated, and compared to those of Model A. In addition, the RMSE was calculated to measure the difference between the estimated pavement performances using the models developed at different percentages of funds (Models B) and Model A. The findings indicated that the decrease in the percentage of PM funds from 100% to 60% associates with pavement performance more comparable to that of Model A and results in a reduction in the RMSE. Whereas, the percentages below 60% correspond to pavement performance less comparable to that of Model A and yields an increase in the RMSE. Hence, the 60% funding was selected to generate the final dataset used in the subsequent analysis.

### 3.4.4. Transition Probability Matrix Estimation

The finally generated data was used to develop an ordered-probit model (Model C) to estimate and predict the probability of pavement condition states. IRI was discretized into five condition states as follows: state 1 ($IRI \leq 60$), state 2 ($60 < IRI \leq 70$), state 3 ($70 < IRI \leq 80$), state 4 ($80 < IRI \leq 100$), and state 5 ($IRI > 100$).

The estimation results shown in Table 3.5 highlight the statistical significance of the *Cum. AADTT*, *Cum. AAFI*, micro-surfacing *(MCR)*, *UTBWC*, and thin HMA overlay *(OVR)*, as well as the significance of the threshold parameters ($\mu_1$, $\mu_2$, $\mu_3$). The estimated $\chi^2$ value was found to be 60417.08, which is greater than the critical $\chi^2$ (15.05) indicating the statistical significance of Model C at a 95% confidence level. The positive signs of the *Cum. AADTT* and *Cum. AAFI* indicate their negative effect on pavement condition, while the negative signs of the PM treatments (*MCR*, *UTBWC* and *OVR*) imply their positive impact on pavement condition.

Table 3.5. Estimation results of Model C

| **Variable** | **Description** | **Parameter Estimate** | **t-statistic** | **p-value** |
|---|---|---|---|---|
| Constant | | 0.35140 | 41.17 | <0.001 |
| Cum. AADTT | Cumulative Annual Average Daily Truck Traffic in 10,000 Trucks | 0.78778 | 146.01 | <0.001 |
| Cum. AAFI | Cumulative Annual Average Freezing Index in 10,000 °F days | 1.51216 | 97.46 | <0.001 |
| MCR | 1 if micro-surfacing is implemented, 0 otherwise | -0.94987 | -43.70 | <0.001 |
| UTBWC | 1 if UTBWC is implemented, 0 otherwise | -0.99930 | -31.95 | <0.001 |
| OVR | 1 if thin HMA overlay is implemented, 0 otherwise | -4.19084 | -98.07 | <0.001 |
| $\mu_1$ | Threshold Parameters | 1.77384 | 217.42 | <0.001 |
| $\mu_2$ | | 3.56871 | 360.12 | <0.001 |
| $\mu_3$ | | 6.03055 | 300.36 | <0.001 |
| Log-likelihood of restricted model = -84142.73 | | | | |
| Log-likelihood of unrestricted model = -53934.19 | | | | |
| Estimated $\chi^2 = -2\,[LL(\beta_R) - LL(\beta_U)]$ = 60417.08 | | | | |
| Degree of freedom (unrestricted number of parameters – restricted number of parameters) = 8 | | | | |
| Critical $\chi^2$ at a 95% confidence level = 15.05 | | | | |

Using Model C, the non-homogeneous transition probability of pavement condition states can be calculated as follows:

$$U_i = 0.35140 + 0.78778 Cum.AADTT + 1.5121 Cum.AAFI - 0.94987\, MCR - 0.99930\, UTBWC - 4.19084 OVR$$

$$S_i = [\mathcal{N}(-U_i), \mathcal{N}(1.77384 - U_i) - \mathcal{N}(-U_i), \mathcal{N}(3.56871 - U_i) - \mathcal{N}(1.77384 - U_i), \mathcal{N}(6.03055 - U_i) - \mathcal{N}(3.56871 - U_i), 1 - \mathcal{N}(6.03055 - U_i)]$$

$$\mathbb{P} = \text{inverse}\ (S_i) \times S_{i+1}$$

Where $U$ is the utility function or the latent variable, $\mathcal{N}$ is the normal probability distribution, $i$ is the time in years, $S_i$ is the $1 \times 5$ condition state vector at time $i$, and $\mathbb{P}$ is the transition probability matrix.

Figure 3.6 presents the probabilistic pavement performance curves for condition states 1, 2, 3, 4 and 5 represented by curves S1c, S2c, S3c, S4c and S5c, and S1p, S2p, S3p, S4p and S5p, respectively for each condition state and for using current and past studies. These probability curves were developed at the mean value of *AAFI* (772 °F days) while the *AADTT* increases by a truck growth rate of 2.8% over pavement age. Using the current study, the expected pavement age is 32 years when pavement condition transitions to state 5 (IRI > 100 in/mi.). The impact of PM on pavement condition is more significant if implemented on pavements in states 3 or 4 rather than in states 1 or 2. STAs should therefore implement PM treatments on pavement surfaces when pavement roughness exceeds IRI of 70 in/mi. for superior maintenance effectiveness.

Figure 3.6 shows that with a probability of 38% the newly constructed pavements have an IRI less than or equal to 60 in/mi. The probability of a pavement being in condition state 2 (60 < IRI ≤ 70 in/mi.) decreases at different rates from the age of 6 to 12 years due to the effect of the expected application of PM treatments such as *MCR* or *UTBWC*. However, this probability begins to increase at the age of 17 years because of the effect of the expected application of PM treatments such as *OVR*. The probability of pavement condition being in state 3 (70 < IRI ≤ 80 in/mi.) decreases over pavement age from 9 to 16 years, during which the probability of being in state 4 (80 < IRI ≤ 100 in/mi.) increases with fluctuating deterioration rates. At the age of 32 years, pavements are expected to be in condition state 5 with IRI greater than 100 in/mi.

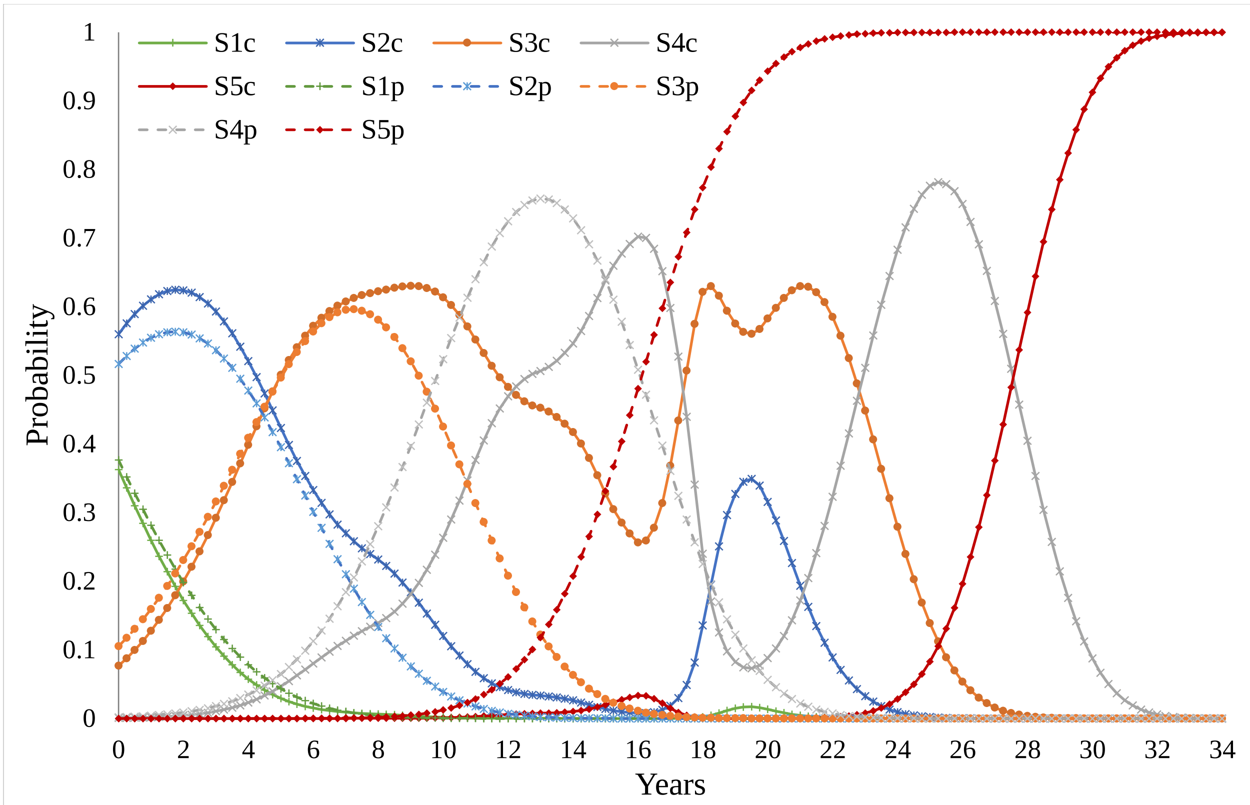



Figure 3.6. Probabilistic pavement performance curves using current (c) and prior (p) studies

***Marginal Effects***

To analyze the impact of the significant explanatory variables on pavement condition, their marginal effects were estimated following the methodology presented by Washington et al. (2011). Figure 3.7 demonstrates the marginal effects of *Cum. AADTT*, *Cum. AAFI*, *MCR*, *UTBWC* and *OVR* on the five pavement condition states: S1c, S2c, S3c, S4c and S5c. Besides, Figure 3.7 shows the marginal effects of *Cum. AADTT* and *Cum. AAFI* on the five condition states: S1p, S2p, S3p, S4p and S5p when the impact of PM is not considered. The estimated marginal effects indicate the expected change in the probability of pavement condition state when the *Cum. AADTT* and/or *Cum. AAFI* change (decease/increase) by one unit, or when the *MCR*, *UTBWC* and/or *OVR* are implemented (1) or not (0). Since the *Cum. AADTT* and *Cum. AAFI* were assumed in 10,000 in the models' estimation, one unit of the *Cum. AADTT* and *Cum. AAFI* is equal to 10,000 Trucks and 10,000 °F days, respectively.

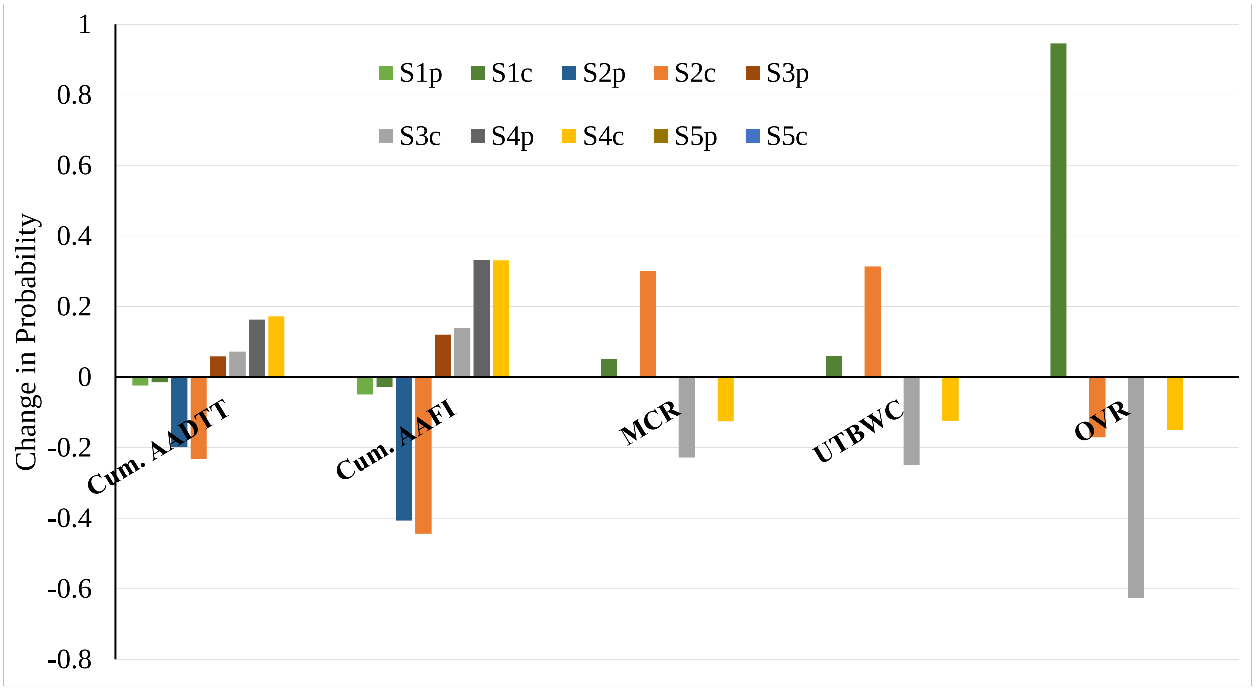



Figure 3.7. Marginal effects of explanatory variables

The *Cum. AADTT* and *Cum. AAFI* have negative effects on the probabilities of condition states 1 and 2, but positive effects on the probabilities of the rest of the condition states. Conversely, the *MCR*, *UTBWC* and *OVR* have positive effects on the probability of the condition states 1 and 2, but negative effects on the probabilities of the rest of the condition states. Figure 3.7 implies that an increase of 10,000 trucks could result in decreasing the probability of pavements being in states 1 and 2 (S1c and S2c) by about 1% and 22%, respectively, whereas increasing the probability of pavements being in states 3 and 4 (S3c and S4c) by 7% and 17%, respectively. Figure 3.7 also indicates that an increase of 10,000 °F days in the *Cum. AAFI* could result in a 44% reduction in the probability of pavements being in state 2 (S2c). Figure 3.7 shows that the marginal effects of the *Cum. AADTT* and *Cum. AAFI* on condition state 1 (S1p) are overestimated when the impact of PM is not accounted for. On the other hand, their marginal effects on the rest of the condition states are underestimated when the impact of PM is not incorporated. Such incorrect estimated marginal effects yield inaccurate prediction of pavement condition and, therefore, non-optimal M&R decision-making.

The implementation of PM treatments increases the probability of pavements being in states 1 and 2, whilst decreasing the probability of being in states 3, 4 and 5. The effects of *MCR* and *UTBWC* are quite similar; however, they are recommended at different pavement ages and conditions. The *MCR* should be used in earlier pavement ages than the *UTBWC*. The *OVR* significantly improves the condition of pavements in states 3 or 4 (70 < IRI ≤ 100 in/mi.). In other words, when *OVR* is implemented on pavements of condition states 3 or 4, they are expected to move to condition state 1 with a probability of more than 90%.

### 3.5. Methodology and Model Validation

The methodology and Model C were validated mathematically using the cross-validation with the actual or out-of-sample data and using a survey of SMEs with strong background in pavement engineering and management. Model C was cross-validated by comparing its predictions with the actual pavement conditions in terms of RMSE and MAPE. A new set of data (10% of the data used to develop Model C) was generated randomly to be used in the validation. The estimated RMSE of Model C was found to be equal to 0.65, which is equivalent to a MAPE of 13%. This indicates that Model C is robust and capable of predicting the mean pavement condition state with a validity percentage of 87%. This, in turn, emphasizes the reliability of the proposed methodology.

Two questionnaire surveys (Surveys 2 and 3) were designed and deployed to SMEs from eight Midwestern states to evaluate the results of the current study compared to the previous studies. The eight Midwestern states include Indiana, Illinois, Wisconsin, Michigan, Ohio, Minnesota, Iowa and Missouri. The SMEs were chosen based on their experience in pavement engineering and management to ensure that the results of the proposed methodology and Model C are consistent with the engineering practice and judgement. The results of using the proposed methodology and Model C were compared to the results of using prior non-homogeneous Markov models such as Yang et al. (2005), Kobayashi et al. (2010), Tabatabaee et al. (2013) and Abaza (2017a). The common assumptions of prior non-homogeneous Markov models are: (1) pavement condition deteriorates continually over time and does not improve, and (2) the effect of PM is not considered in pavement performance prediction. Under these assumptions, a non-homogeneous Markov model was developed using the generated data from the current study. Pavement performance curves were then simulated using the models of the current and previous research.

The purpose of Survey 2 is to assess the trend of the pavement performance curves of the five condition states estimated using the models of current and past studies. Each condition state curve was divided into segments according to the change of deterioration rate. The SMEs were requested to assign a score from 0 to 3 for each curve segment; where the score of 0, 1, 2, or 3 indicates that whether the curve segment strongly disagrees, disagrees, agrees, or strongly agrees with the actual pavement behavior, respectively. Figure 3.6 displays the probabilistic pavement performance curves developed using the current (S1c, S2c, S3c, S4c and S5c) and prior (S1p, S2p, S3p, S4p and S5p) studies. The aim of Survey 3 is to evaluate the marginal effects of the influential explanatory variables on pavement condition, which were estimated based on the proposed methodology and models. The SMEs were asked to give scores similar to those of Survey 2 to the estimated marginal impacts of the influential independent variables based on their engineering and practical perspectives. Figure 3.7 shows the marginal effects of *Cum. AADTT, Cum. AAFI, MCR, UTBWC* and *MCR* estimated for each condition state using the current study (S1c, S2c, S3c, S4c and S5c) and prior research (S1p, S2p, S3p, S4p and S5p).

As shown in Figure 3.6, the expected pavement age using the model of prior research is 20 years, which is consistent with the estimate of most prior research. However, the expected pavement age using Model C of the current study is 32 years. This difference in pavement age is the expected extension in pavement service life due to different PM treatments implemented over pavement life span, which was not adequately considered in previous research. It is worth noting that at the times when PM is not expected (e.g., up to the age of 6 years), the trends of pavement performance based on the current and past research are comparable. In addition, pavements in condition state 1 deteriorate similarly using both models. However, the deterioration rates of pavements in the remaining condition states are different, particularly at the times of expected PM. Following the previous studies, pavements are shown most probably to be in poor condition (IRI > 100 in/mi.) at ages greater than 16 years. On the contrary, following the current model, pavements are more likely to be in poor condition around the age of 27 years.

Six SMEs from the states of Indiana, Iowa and Minnesota and from the National Center for Pavement Preservation (NCPP) responded to Surveys 2 and 3. They confirmed that the pavement condition and marginal effects of the influential variables estimated using Model C are consistent with the state of practice. The SMEs' overall assessment of the estimated pavement performance curves for the five condition states using the current and prior research is presented in Table 3.6. The pavement performance curves developed using the newly developed hybrid approach are more compatible with the actual pavement behavior than those based on past studies. The average scores provided by the SMEs to the estimated marginal effect of *Cum. AADTT, Cum. AAFI, MCR, UTBWC* and *MCR* are 1.90, 1.75, 2.00, 1.92, and 1.80, respectively, demonstrating the consistency of the estimated marginal effects with the actual practice.

The SMEs from the state of Indiana agree with the estimated pavement performance of condition state 1 (less than 38% of new pavements are expected to have IRI less than 60 in/mi.) since the target value of IRI for newly constructed pavements in Indiana is 70 in/mi. On the contrary, the SMEs from the states of Iowa and Minnesota disagree with the SMEs from Indiana as the acceptable IRI of newly built or resurfaced pavements in Minnesota is about 35 in/mi. The SMEs from Iowa mentioned that the percentage of new pavements with an IRI of less than 60 in/mi. is greater than 38%, and pavements more than 8 years old may have IRI less than 60 in/mi. The SMEs from Indiana agree with the expected improvement in the performance curves of states 2 and 3 owing to the impact of PM. In contrast, Iowa's SMEs anticipated higher probabilities of condition states 2 and 3 for pavements below 12 years old than those shown in Figure 3.6.

Table 3.6. Assessment of SMEs to pavement performance curves of current and past research

| **Condition State** | **Average Scores of SMEs Assessment** | |
|---|---|---|
| | **Current Study** | **Past Studies** (e.g., Yang et al. 2005; Kobayashi et al. 2010; Tabatabaee et al. 2013; Abaza 2017a) |
| 1 | 2.50 | 2.25 |
| 2 | 2.20 | 1.80 |
| 3 | 2.00 | 1.45 |
| 4 | 1.90 | 1.33 |
| 5 | 2.10 | 1.36 |

*Note: Score of 0, 1, 2, or 3 indicates that the performance curve of a condition state strongly disagrees, disagrees, agrees, or strongly agrees with the state of practice, respectively.*

The SMEs from Indiana indicated that the marginal impact of the *Cum. AADTT* is reasonable, whilst the SMEs from Iowa stated that the *Cum. AADTT*'s effect is higher than expected. The

Indiana's SMEs stated that the impact of the *Cum. AAFI* should be lower than the estimated value (22% decrease in probability of condition state 2) because of the application of binder of performance grade (PG) of low temperature and/or freeze-thaw resistant material that limits the deterioration of pavement condition. Similarly, they indicated that if *MCR* is applied to pavement surface, the pavement condition is less likely to be in state 1 (IRI < 60 in/mi.), which agrees with the estimated marginal effects. Furthermore, they stated that pavements are most likely to be in condition states 1 or 2 if *OVR* is implemented.

### 3.6. Summary

The hybrid approach proposed in this study helps incorporate the effectiveness of PM into probabilistic pavement performance models when the historical PM data is absent or insufficient. This approach is comprised of six major tasks: data collection and analysis; data simulation; estimation of initial times for preventive maintenance treatments; data generation; estimation of transition probability matrix (TPM); and validation of the approach and models developed, both mathematically and through a survey of SMEs in pavement engineering and management.

Pavement condition data were collected form the LTPP database for interstate flexible pavements from eight Midwestern United States. Data were simulated through developing an exponential multiple regression model and determining the probability distribution of the dependent and independent variables. The traffic and climate loadings were found to be the statistically significant variables affecting pavement condition in the absence of historical PM data. The initial times for PM treatments were determined through literature search, Survey 1 of STAs to collect data of pavement condition and PM, and detection of PM times from probabilistic pavement performance curves. Based on the literature search and results of Survey 1, the most common effective PM treatments were found to be micro-surfacing, ultra-thin bonded wearing course (UTBWC), and thin HMA overlay.

An ordered-probit Model A was built to develop probabilistic pavement performance curves used to identify the approximate probable times for PM application, and to estimate the effectiveness of PM treatments. Eventually, the initial times for the PM treatments were estimated as 4 – 10 years with 1.5-year standard division (STD) for micro-surfacing, 11 - 15 years with 1-year STD for UTBWC, and 16 – 20 years with 1-year STD for thin overlay. Since the implementation of PM is constrained by limited funding, a greedy algorithm was developed to prioritize PM schedule based on pavement condition and treatment costs, and under the constraint of the estimated initial times for PM. The amount of funding was specified as a percentage (100% to 40%) of the total funding required to implement PM to pavement sections at the estimated initial times. For each percentage of funding, the greedy algorithm results in pavement condition data used to develop ordered-probit Models B. Probabilistic pavement performance models were created using Models B and then compared with that of Model A. The results showed that at a 60% funding the simulated data including types and times of PM treatments performs comparably to the actual data.

An ordered-probit Model C was developed to estimate the non-homogeneous transition probabilities of pavement condition incorporating PM impact. The traffic and climate loadings, as well as the PM treatments micro-surfacing, UTBWC and thin HMA overlay, were found to be statistically significant. The statistical significance of PM in pavement performance prediction emphasizes the necessity of collecting and managing PM data. Probabilistic pavement performance curves were developed, and the marginal effects of the explanatory variables were estimated. It was noticed that the lack of inclusion of the effect of PM on pavement condition causes an underestimation of the condition and remaining service life of pavement, which could lead to erroneous and non-optimal pavement M&R decisions.

The hybrid approach and the developed non-homogeneous Markov model were validated mathematically through the cross-validation with the actual/out-of-sample data, and practically using two surveys (Surveys 2 and 3) sent to the SMEs in pavement engineering and management. The cross-validation was performed to ensure that the predicted pavement condition is comparable to the actual pavement condition. The RMSE was found to be equal to 13%, indicating that the methodology and models developed are reliable and accurate in probabilistic pavement condition prediction. Survey 2 was deployed to the SMEs to assess the trends of pavement performance curves developed using the current and prior research. Survey 3 was used to assess the estimated marginal effects of the explanatory variables. The SMEs' validation revealed that pavement performance curves developed using the current study are more accurate and practical than those developed using prior studies. Furthermore, although the SMEs do not strongly agree with some trends of pavement performance curves such as that of state 4, their overall evaluations range from agreement to strong agreement.

# CHAPTER 4. COMPARATIVE ANALYSIS OF MARKOVIAN METHODOLOGIES FOR MODELING PAVEMENT PERFORMANCE

[A version of this chapter is under review at the Journal of Infrastructure Systems, ASCE].[2]

A variety of Markov chain methodologies have been used to develop stochastic performance models for infrastructure systems. These are the homogeneous, staged-homogeneous, non-homogeneous, semi-, and hidden Markov. The primary components of a Markov model are the transition probability matrix, condition states, and duty cycle. This chapter hypothesizes that the number of condition states (NCS) and length of duty cycle (LDC) significantly influence the prediction accuracy of a Markov model, and that the nature of such influence varies across the different Markov methodologies. In previous studies on Markovian performance modeling, not only is this hypothesis unanswered but also there is lack of comparison across the different Markov methodologies. Addressing these questions can be beneficial to practitioners who seek guidance on selecting appropriate Markov models for modeling the performance of their road infrastructure networks. In a bid to throw light on this issue, this chapter develops and compares the Markovian performance models using empirical data. This chapter also investigates the sensitivity of the Markovian model prediction accuracy to NCS and LDC. The results indicate that the semi-Markov is generally statistically superior to the homogeneous and staged-homogeneous Markov (except in a few cases of NCS and LDC combinations) and that the prediction accuracy of Markovian models is significantly sensitive to NCS and LDC: an increase in NCS improves the prediction accuracy until a certain NCS threshold after which the accuracy diminishes, plausibly due to data overfitting. In addition, an increase in LDC improves the prediction accuracy when the NCS is small. The results can help guide highway agencies and future researchers in selecting appropriate Markovian methodologies, and for a selected methodology, in deciding the appropriate number of condition states and duty cycle length.

[2] Yamany. M.S., Abraham, D.M., and Labi, S. Comparative Analysis of Markovian Methodologies for Modeling Infrastructure System Performance. Submitted to Journal of Infrastructure Systems, ASCE.

## 4.1. Introduction

Highway agencies need reliable infrastructure performance models so they can predict the future condition of the infrastructure more confidently. These future predictions serve as a basis for developing realistic and efficient maintenance and rehabilitation (M&R) schedules and strategies to manage their limited resources and keep their infrastructure in a state of good repair (Walls III and Smith 1998; Peshkin et al. 2004). Infrastructure performance models can be categorized as: deterministic, Artificial Intelligence (AI) and probabilistic (Yamany and Elwakil 2019, 2020; Yamany et al. 2019b; Yamany et al. 2020b). Deterministic models have been widely used for infrastructure performance but do not account for the uncertainty associated with infrastructure condition data. AI models, specifically Artificial Neural Networks (ANNs) have been successful in infrastructure condition prediction but are considered as black boxes and therefore tend to yield results that are difficult to interpret (García de Soto et al. 2018; Yamany et al. 2020a, b).

Unlike deterministic and AI models, probabilistic models explicitly consider the inherent uncertainty associated with infrastructure condition data, and therefore they yield results that are relatively more robust and intuitive (Saeed et al. 2017; Qiao et al. 2019). Greene and Henscher (2010) argued that although probabilistic models may be less precise, they are more robust than deterministic models. Of the probabilistic techniques for performance modeling, the Markov chain is the most widely used. Based on the reviewed literature, Markov chain methodologies can be categorized as follows: homogeneous, staged-homogeneous, non-homogeneous, semi-, and hidden Markov. The primary components of a Markov model are the transition probability matrix (TPM), condition states, and the duty cycle or step time. The transition probability matrix is a $n \times n$ square matrix that describes the probability of infrastructure assets migrating from one condition state to another; where $n$ is the number of condition states. The condition states are discrete ratings of infrastructure condition, such as excellent, fair and poor. The duty cycle is the time interval or the frequency of data collection, which is typically one year. Figure 4.1 depicts the transition probabilities in Markov models, where $P_{ii}$ is the probability that an infrastructure component of condition state *i* stays in the same state after one unit of duty cycle, and $P_{in}$ is the probability that an infrastructure component of condition state *i* transitions to state *n* after one unit of duty cycle.

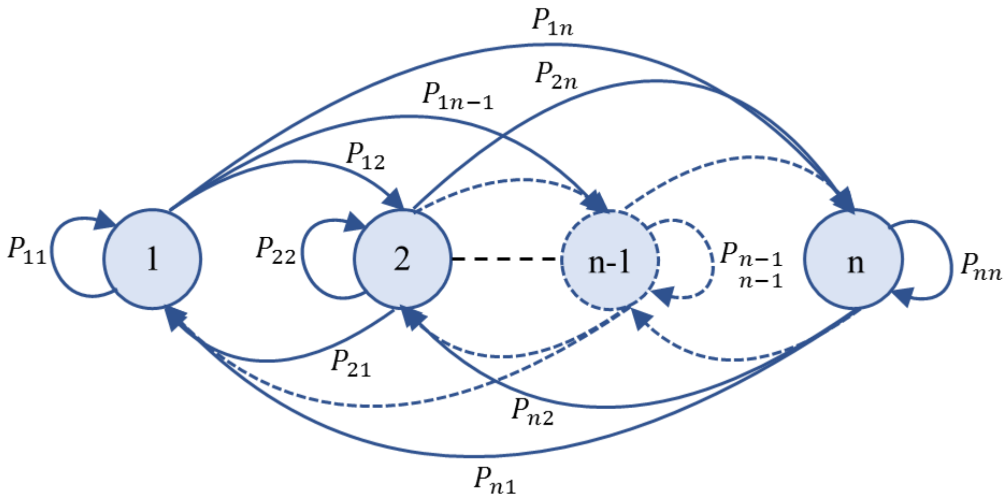


Figure 4.1. Transition probability diagram

Markov models estimate the future condition ($S^{i+1}$) of an infrastructure component based on its current condition ($S^{i}$) (the memoryless property of Markov process) and its deterioration and improvement TPM; $S^{i+1} = S^{i} \times \text{TPM}$, where the TPM can be written as follows:

$$\text{TPM} = \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{bmatrix} \qquad (4.1)$$

The current chapter begins with a review of the literature on the use of Markovian techniques for infrastructure deterioration modeling with the emphasis on pavements. This is followed by the research methodology where the data collection and study hypotheses are described. To circumvent potential statistical bias often associated with the use of different sets of condition data as evidenced in past studies, the comparative analysis in this chapter used only one set of data. In the methodology, the following hypotheses are presented: (i) the different Markov methodologies yield models with different prediction accuracies, (ii) the Markov model performance is significantly influenced by the integrity of the underlying TPM, the number of condition states (NCS), and the length of duty cycle (LDC), and (iii) the nature of these influences varies across the different combinations of NCS and LDC levels. This chapter presents the methods used for the comparative and sensitivity analysis, discusses the results, and explains the practical benefits of the results.

### 4.2. Prior Research on Using Markovian Techniques for Pavement Deterioration Modeling

A number of previous research studies have reported the superiority and validity of Markov models in terms of their prediction accuracy of infrastructure condition based on two criteria: (1) the consistency of the model assumption with the natural behavior of infrastructure deterioration; and (2) the validation of the model results.

Abaza (2016a, b), Abaza (2017a) and Pérez-Acebo et al. (2019) stated that the homogeneous Markov technique generally has the lowest prediction accuracy (compared to other Markov techniques) because it makes the rather restrictive assumption that the transition probability of pavement condition is fixed over the pavement lifetime. The homogeneous Markov methodology, however, presents the only choice when data are limited, for example, when data are available for only two consecutive transitions of pavement condition. The staged-homogeneous Markov technique is generally used to simulate the actual deterioration more realistically compared to the homogeneous Markov technique. This technique assumes that the pavement condition does not change significantly over a period of 5 - 6 years (Butt et al. 1987; Abaza 2016a), and therefore inherently assumes that the transition probabilities of pavement condition change every fixed period of time or stages. For this reason, the staged-homogeneous Markov methodology is likely to yield more accurate predictions compared to the homogeneous Markov methodology. Nevertheless, the period of 5 or 6 years may be too short or too long for pavements with slow or fast deterioration rates, respectively, and may result in overestimation or underestimation of pavement condition.

The semi-Markov technique (Nesbit et al. 1993; Thomas and Sobanjo 2012) assumes that over the pavement age, the transition probabilities change from one holding time to another, and that the holding times may differ in their durations based on the different deterioration rates. For this reason, the semi-Markov methodology is likely to have higher prediction accuracy compared to the staged-homogeneous and homogeneous Markov methodologies. The non-homogeneous Markov methodology (Yang et al. 2005, 2006; Kobayashi et al. 2010; Tabatabaee et al. 2013; Abaza 2017a) assumes that the pavement condition transition probabilities vary continually over pavement lifespan. Also, the non-homogeneous Markov methodology considers the impact of explanatory variables on the transition probabilities, and therefore is more powerful in capturing data variability. Therefore, this methodology is expected to yield the most reliable predictions of pavement deterioration. The hidden Markov methodology assumes two condition states: hidden and observed, and is used to estimate the probability of hidden condition states when data on the observed condition states are provided. Therefore, the hidden Markov technique is typically used when data are incomplete (Lethanh and Adey 2012, 2013; Lethanh et al. 2015).

Previous research studies such as Yang et al. (2005 and 2006) and Thomas and Sobanjo (2012) reported that their Markov models are robust and can provide accurate predictions by comparing their predictions with actual data. Other research studies, for example, Butt et al. (1987) and Li (2005), validated their Markov models by contrasting their prediction accuracy with that of the Markov models developed by Keane and Keane (1985), Wang et al. (1994) and Madanat et al. (1995a, b). Notwithstanding, Butt et al. (1987) and Li (2005) used condition data for pavements that was not the same as those used by Keane and Keane (1985), Wang et al. (1994) and Madanat et al. (1995a, b). It is worth noting that different sets of pavement condition data have different variations that influence the prediction accuracy of the respective Markov models. Moreover, the use of one Markov methodology across different datasets does not necessarily yield the same prediction accuracy across these datasets. Hence, the prediction power of Markovian models should not be judged across distinct sets of data.

The number of condition states (NCS) in Markovian pavement performance models has been assumed based on the availability of data (Martin and Kadar 2012) and a specific indicator of pavement condition. Typically, 10 condition states (NCS = 10) have been used in Markovian pavement performance models (Li et al. 1996; Yang et al. 2005, 2006; Ortiz-García et al. 2006; Abaza 2014, 2017a). Some researchers used NCS values as large as 20 (Macleod and Walsh 1998) and as small as 3 (Pérez-Acebo et al. 2019). Pérez-Acebo et al. (2018 and 2019) developed homogeneous Markovian models for flexible and rigid pavements in the Republic of Moldova: 6 condition states were used for flexible pavements, while 3 - 4 condition states were used for each of the three sub-classes of rigid pavements. Pérez-Acebo and his team selected the NCS based on the guidance from the previous research by Odoki and Kerali (2000) and Adedimila et al. (2009). Nevertheless, these two studies by Pérez-Acebo et al. (2018 and 2019) did not address the significance of the NCS on the prediction accuracy of their Markovian pavement performance models, which indicates that their selected NCS might not have yielded models with the highest possible prediction accuracy. It is worth mentioning that the NCS should be selected not only on the basis of data availability but also to adequately capture the changes in pavement condition over its lifetime (Porras-Alvarado et al. 2014).

The length of duty cycle (LDC) in Markovian pavement performance models is the interval between data collection. The typical LDC value reported in the literature is one year because most highway agencies collect pavement condition data annually (Butt et al. 1987; Abaza and Murad 2010; Abaza 2016a, b). However, some researchers assumed 2-year duty cycle for their Markovian pavement performance models (Abaza 2004; Hassan et al. 2017a, b). There are yet others, such as Pérez-Acebo et al. (2018 and 2019) that used a half-year duty cycle (because their data were collected biannually) but still added the caveat that a half-year duty cycle might be rather short and could lead to data overfitting. The current research recognizes that the LDC should be selected on the basis of not only the data collection frequency, but also the influence on the model’s prediction accuracy.

## 4.3. Research Methodology

The chapter’s conceptual framework (Figure 4.2) has five main parts: (1) development of Markovian models, (2) establishing the NCS and LDC combinations, (3) creating 24 homogeneous, 40 staged-homogeneous and 8 semi-Markov Models, each with different specifications and assumptions, (4) comparative analysis of the Markovian models within each Markov methodology, and (5) comparative analysis of Markovian models across all Markov methodologies.

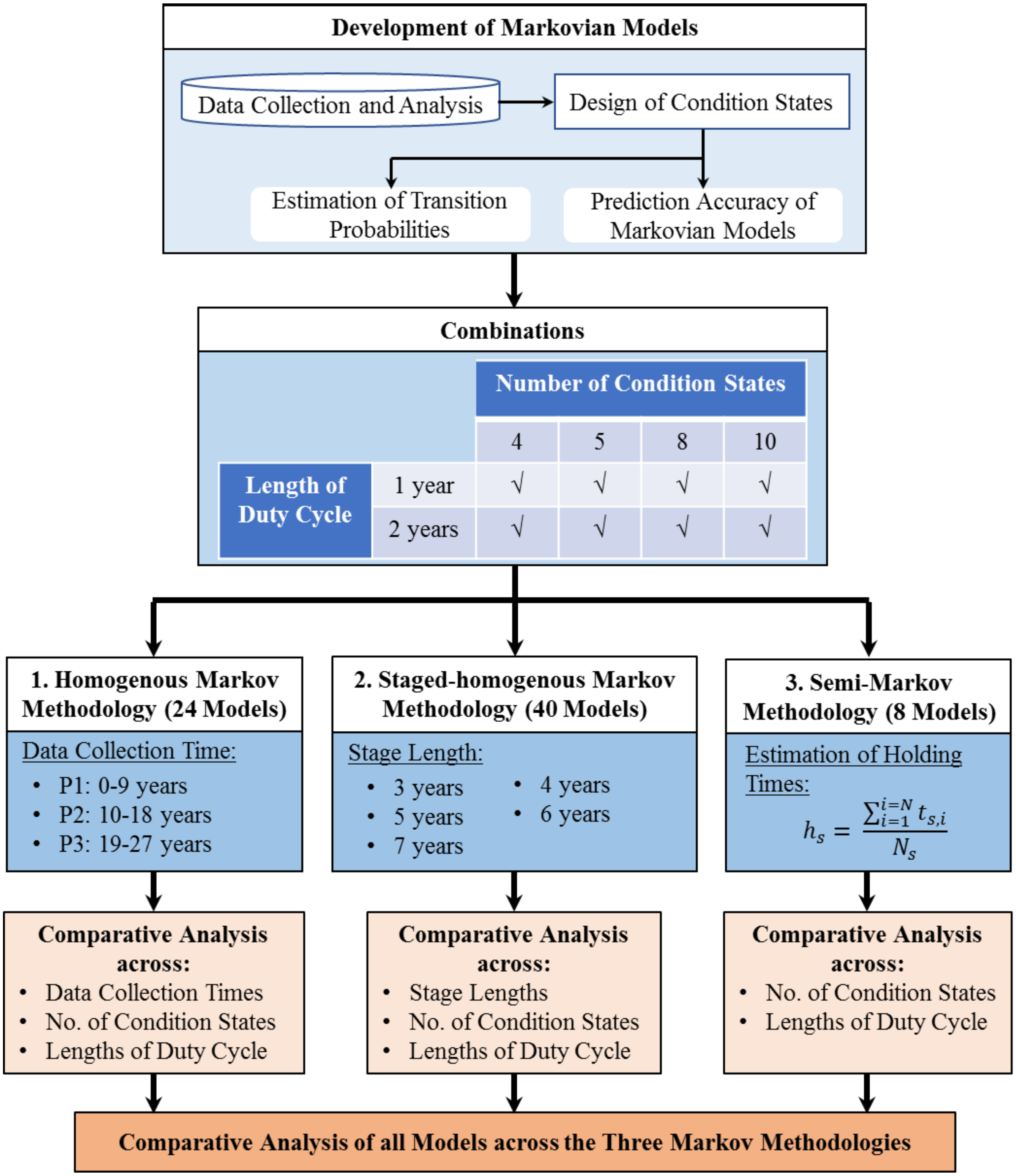



Figure 4.2. Conceptual framework

### 4.4. Development of Markovian Models

The three primary components of a Markov model are the transition probability matrix, the duty cycle, and the condition states. The estimation and settings of these three Markovian elements can influence profoundly the estimation and prediction accuracy of Markov models. In this chapter, different Markov models with different NCS and LDC configurations were developed, and their prediction accuracy is calculated and compared to each other.

### 4.4.1. Data Collection and Analysis

Interstate flexible pavement condition data were acquired from the LTPP database for eight Midwestern states (Indiana, Illinois, Wisconsin, Michigan, Ohio, Minnesota, Iowa and Missouri). The data were cleaned by deleting the extreme outliers and high-leverage points (Belsley et al. 1980; Ahmed et al. 2016), resulting in 966 observations. Table 4.1 shows the descriptive statistics of the variables considered in the analysis.

Table 4.1. Descriptive statistics

| **Variable** | **Description** | **Mean** | **STD** | **Min** | **Max** |
|---|---|---|---|---|---|
| IRI | International Roughness Index (in/mi) | 80.9 | 27.8 | 31.3 | 194.9 |
| Age | Years since construction or rehabilitation | 9.6 | 6.6 | 0.0 | 27.0 |
| AAP | Annual Average Precipitation (inches) | 38.1 | 7.7 | 24.4 | 61.4 |
| AAT | Annual Average Temperature (°F) | 50.5 | 3.3 | 42.8 | 59.7 |
| AAFI | Annual Average Freezing Index (°F days) | 788 | 393 | 70 | 1924 |
| AADTT | Annual Average Daily Truck Traffic | 2169 | 802 | 249 | 5115 |
| ESALs | Equivalent Single Axle Loads (18-Kip) | 1115 | 534 | 122 | 3195 |

To fill some gaps in the database, data input was carried out using simulation. Statistically significant explanatory variables, with their respective probability distributions, were used in the simulation.

### 4.4.2. Design of Condition States

This section establishes the NCS, the range of each condition state (in terms of lower and upper values of pavement condition indicator), and the LDC. In the past literature on Markovian pavement performance models, different NCS ranging from 3 to 20 were used, and these were based on the adopted pavement condition indicator, data availability and the level of detail required in condition prediction. The LDC values spanned from 6 months to 2 years based on the frequency of data collection and the highway agency policy.

Due to data limitation, the NCS used in this study is 4, 5, 8 and 10, and the LDC ranges from 1 to 2 years (these are the most widely used duty cycles in the literature). To build Markov chain models, the pavement IRI values were discretized into *n* condition states (*n* = 4, 5, 8 and 10). The range of the intervals or the cut-offs of the condition states in terms of IRI values were designed to satisfy two requirements: (1) equal range widths for all the condition states as indicated by Butt et al. (1987), Odoki and Kerali (2000) and Pérez-Acebo et al. (2018), and (2) a sufficient percentage of number of observations in each condition state to achieve significant results (Pérez-Acebo et al. 2019). The minimum percentage of observations in each condition state was specified as 10% for the 4 or 5 condition states, and 5% for the 8 or 10 condition states. In cases where the minimum percentage of observations for a condition state was not achieved, either the range of that state was extended or that state was combined with the lower or higher neighboring state as recommended by Pérez-Acebo et al. (2019). Figure 4.3 presents the distribution of the discretized pavement condition values in terms of IRI. Table 4.2 presents the range and percentage of observations of the final design of each condition state, for each NCS value used in the analysis.

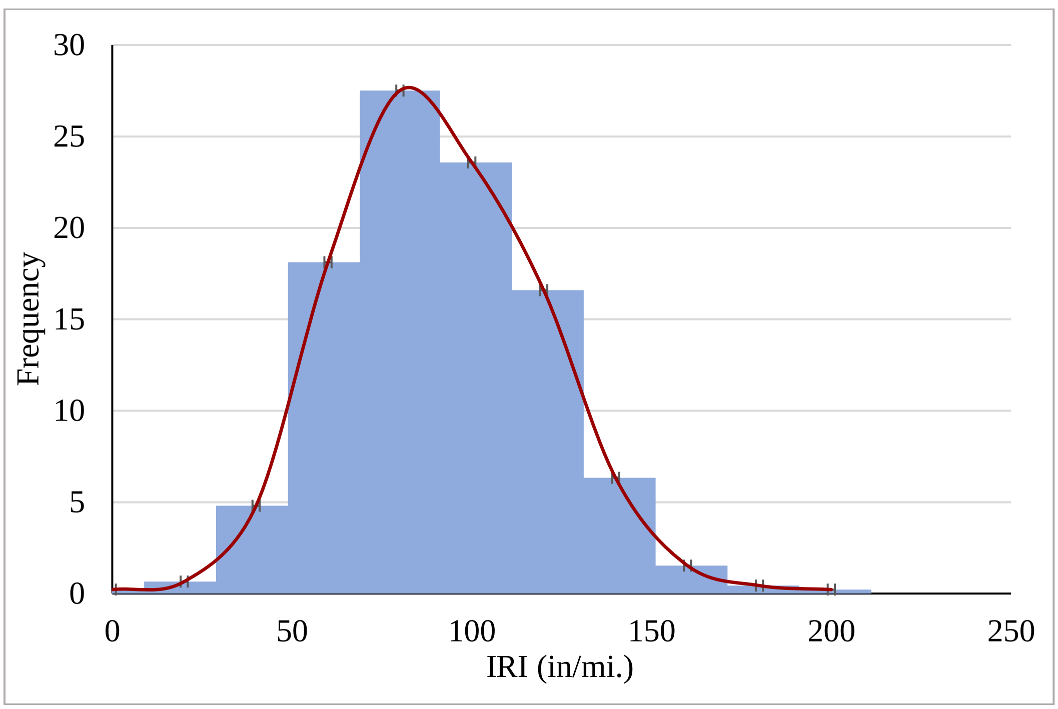



Figure 4.3. Distribution of pavement condition in terms of IRI

Table 4.2. Discretization of IRI into NCS

| NCS | | Condition State | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 4 | Range | $\text{IRI} \leq 60$ | $60 < \text{IRI} \leq 75$ | $75 < \text{IRI} \leq 100$ | $\text{IRI} > 100$ | - | - | - | - | - | - |
| | P | 9.52 | 38.19 | 31.84 | 20.33 | - | - | - | - | - | - |
| 5 | Range | $\text{IRI} \leq 60$ | $60 < \text{IRI} \leq 70$ | $70 < \text{IRI} \leq 80$ | $80 < \text{IRI} \leq 100$ | $\text{IRI} > 100$ | - | - | - | - | - |
| | P | 9.52 | 26.49 | 20.72 | 22.63 | 20.33 | - | - | - | - | - |
| 8 | Range | $\text{IRI} \leq 60$ | $60 < \text{IRI} \leq 65$ | $65 < \text{IRI} \leq 70$ | $70 < \text{IRI} \leq 75$ | $75 < \text{IRI} \leq 80$ | $80 < \text{IRI} \leq 90$ | $90 < \text{IRI} \leq 100$ | $\text{IRI} > 100$ | - | - |
| | P | 9.52 | 13.18 | 13.26 | 11.19 | 9.33 | 13.55 | 8.97 | 20.33 | - | - |
| 10 | Range | $\text{IRI} \leq 60$ | $60 < \text{IRI} \leq 64$ | $64 < \text{IRI} \leq 68$ | $68 < \text{IRI} \leq 72$ | $72 < \text{IRI} \leq 76$ | $76 < \text{IRI} \leq 80$ | $80 < \text{IRI} \leq 86$ | $86 < \text{IRI} \leq 93$ | $93 < \text{IRI} \leq 100$ | $\text{IRI} > 100$ |
| | P | 9.52 | 9.95 | 11.49 | 9.76 | 8.47 | 7.29 | 8.86 | 7.75 | 5.92 | 20.33 |

*Note: P is the percentage of observations in each condition state; IRI value is in in/mi.*

### 4.4.3. Estimation of Transition Probabilities

Five methods have been used in the literature to estimate the transition probabilities of pavement condition, namely the expected-value, the percentage transition, the simulation-based, the econometric models, and the duration models (Ortiz-García et al. 2006; Abaza 2016a; Abaza 2017a; Yamany et al. 2019a). Of these five methods, the percentage transition is the most common, and was used in this study. The percentage transition method estimates the transition probabilities as the pavement proportions that transition from one state to another in one unit of duty cycle. The pavement proportions are calculated in terms of the number of pavement sections of a fixed length (Abaza 2014; Pérez-Acebo et al. 2018) or the cumulative length of pavement sections of different lengths (Hassan et al. 2017a, b; Osorio-Lird et al. 2018). In this chapter, the transition probability of a condition state $i$ is calculated as the proportion of the number of 1-mile-long pavement sections that transitions from one state to another:

$$P_{ij} = \frac{N_{i,t-1} - N_{i,t}}{N_{i,t-1}} \qquad (4.2)$$

where $P_{ij}$ is the probability of pavement condition transitioning from state $i$ to any other state $j$ after one unit of duty cycle, $N_{i,t}$ is the number of 1-mile-long pavement sections in state $i$ at time $t$, and $N_{i,t-1}$ is the number of 1-mile-long pavement sections in state $i$ at time $t$ minus one unit of duty cycle.

### 4.4.4. Prediction Accuracy of Markovian Models

In the current research, the accuracy of pavement condition prediction is the criterion for comparing the different Markov models and methodologies. The prediction accuracy is expressed in the values of the mean absolute percent error (MAPE, %) and the root mean square error (RMSE, in/mi.) which compare the actual and predicted pavement condition. Here, the RMSE is estimated by a unit of condition state in each Markov model. To compare the RMSE values of the Markov models with different NCS and cutoffs, the computed RMSE (Equation 4.4) was adjusted by the term $\frac{(PI_{max}-PI_{min})}{n}$; where $n$ is the NCS of each Markov model, and $(PI_{max} - PI_{min})$ is the range of pavement condition (PI is the pavement condition indicator).

$$\text{MAPE} = \left(\frac{1}{N}\sum_{i}^{N}\left(\frac{\left|E_{i,Actual} - E_{i,Predicted}\right|}{E_{i,Actual}}\right)\right) \times 100 \qquad (4.3)$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i}^{N}\left(E_{i,Actual} - E_{i,Predicted}\right)^2} \qquad (4.4)$$

where $E$ is the expected value of pavement condition $(P_1 \times 1 + P_2 \times 2 + \cdots + P_n \times n)$, $P$ is the probability of being in states 1, 2, …or $n$; $n$ is the NCS of each Markov model; and $N$ is the number of observations.

## 4.5. Homogeneous Markov Models

Homogeneous Markov models assume that pavements deteriorate at a constant rate over their design life. This assumption can be considered unrealistic because in reality, the deterioration of pavements has been determined to be non-linear with respect to time. Due to the relative simplicity of their computational analysis, these models have been used in pavement management systems at highway agencies such as the Arizona DOT. These models do not require large amounts of historical data; rather they require data on only two transitions of pavement condition.

Homogeneous Markov models have been developed in past research for prediction using pavement condition data that are readily available for model estimation. These data were collected for two consecutive pavement condition transitions during any time period of the pavement life. A road network consists of pavement/road sections of different ages, and the pavement section age is defined as the number of years since construction, reconstruction or last rehabilitation. Due to differences in the deterioration rates over pavement age, using data from the early ages to predict pavement condition may yield a different prediction accuracy compared to using data from the late ages. Therefore, the time of data collection should be considered when assessing the prediction accuracy of the homogeneous Markov models. The current research hypothesizes that using pavement condition data collected at the early ages could yield low accuracy of pavement condition predictions compared to using data collected at the late ages. In this study, pavement sections were categorized into three age-based cohorts: (P1) 0 – 9th year cohort, (P2) 10 – 18th year cohort, and (P3) 19 – 27th year cohort (Figure 4.4). A cohort is a collection of pavements that exist in the network at a given time. Pavements in the P1 cohort are those constructed or in-service during the early ages (0-9 years) of the pavement network, while pavements in the P2 or P3 cohorts are those constructed or in-service during the middle (10-18 years) or the late ages (19-27 years), respectively of the pavement network. Homogeneous Markov models were developed using data of each pavement cohort (P1, P2 and P3), and for each of two duty cycles (one-year and two-year) and number of condition states ranging from 4 to 10. Therefore, twenty-four homogeneous Markov models were established and analyzed.

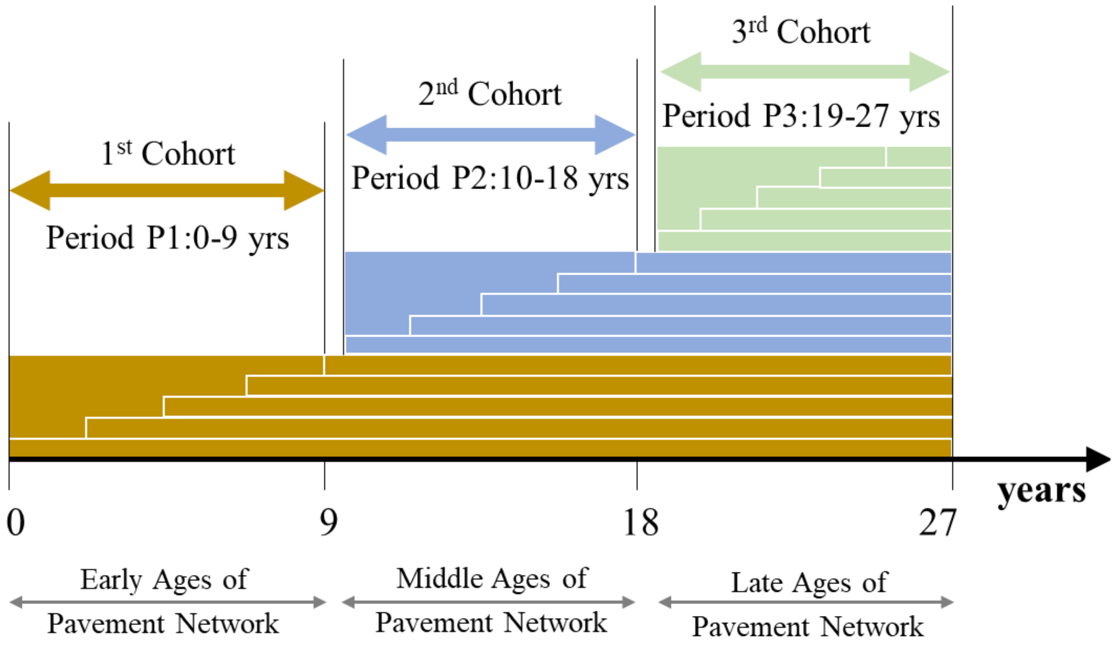
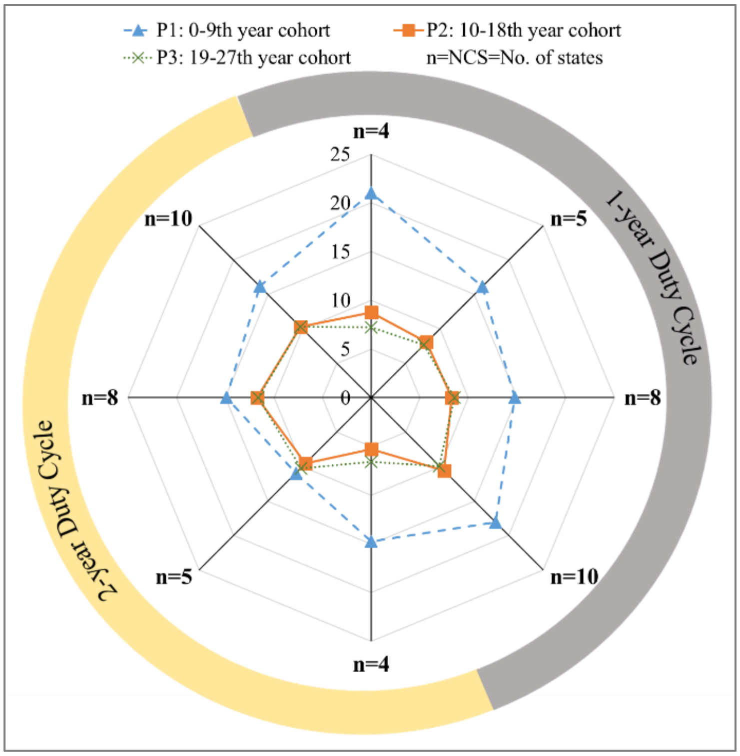


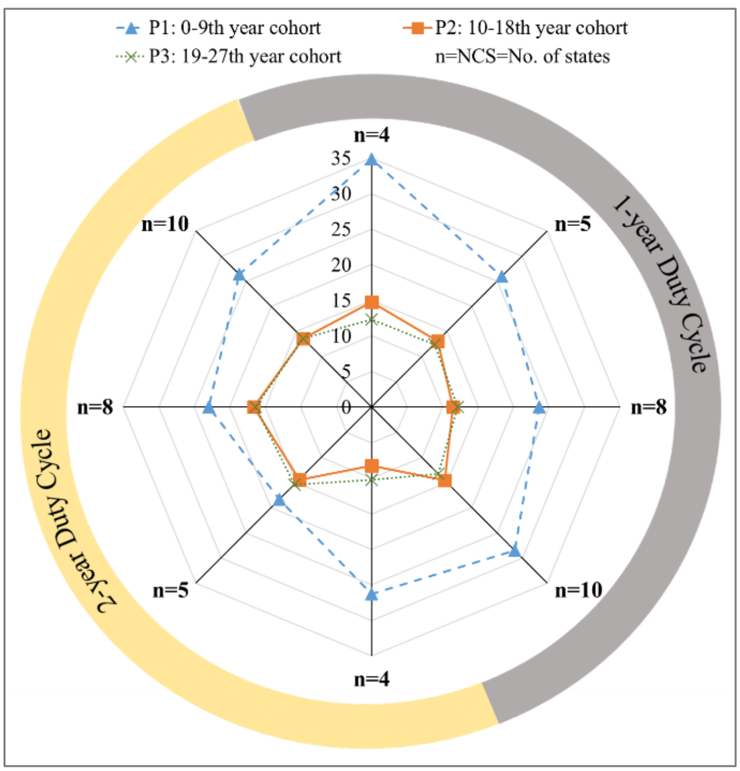

Figure 4.4. Cohorts in the pavement network

#### 4.5.1. Significance of Time of Data Collection

Figures 4.5(a and b) present the estimated MAPE and RMSE for the homogeneous Markov models developed during the three periods or pavement cohorts (P1, P2 and P3). For the homogeneous Markov technique, models developed using pavement condition data collected at the early ages of pavement network (P1, 0-9 years) yielded the least accurate predictions across all the NCS and LDC combinations. This could be due to the fact that the deterioration of pavement condition begins with minimal rates over the early ages of pavement life. In addition, according to the distribution of the pavement age variable, most of the pavement sections in the 0-9th year cohort were newly constructed (1 or 2 years old). Hence, using data gathered for the early ages to forecast pavement condition for the remaining ages (up to the age of 27 years) raises the prediction error. On the contrary, the homogenous Markov models developed using data collected during P2 (10-18 years) or P3 (19-27 years) have similar lower prediction errors (MAPE = 5% to 12%).




(a)




(b)

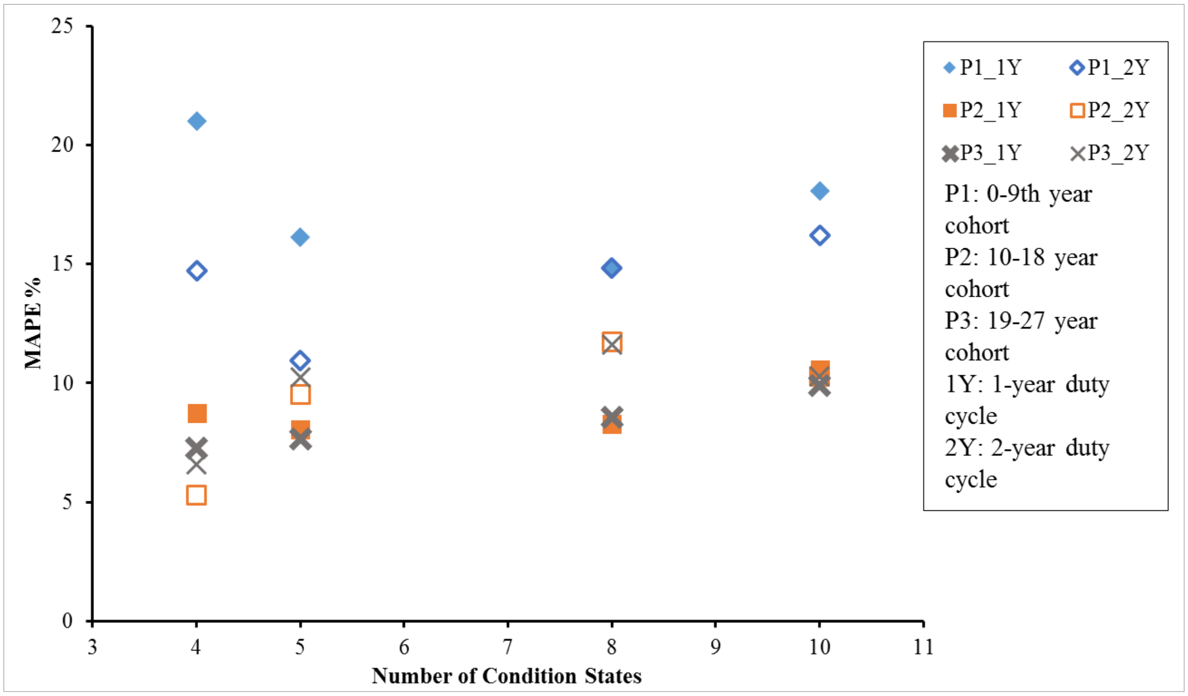
Figure 4.5. Variation of prediction accuracy of homogeneous Markov models across periods of data collection and different NCS and LDC combinations: (a) MAPE; and (b) RMSE

### 4.5.2. Significance of Number of Condition States

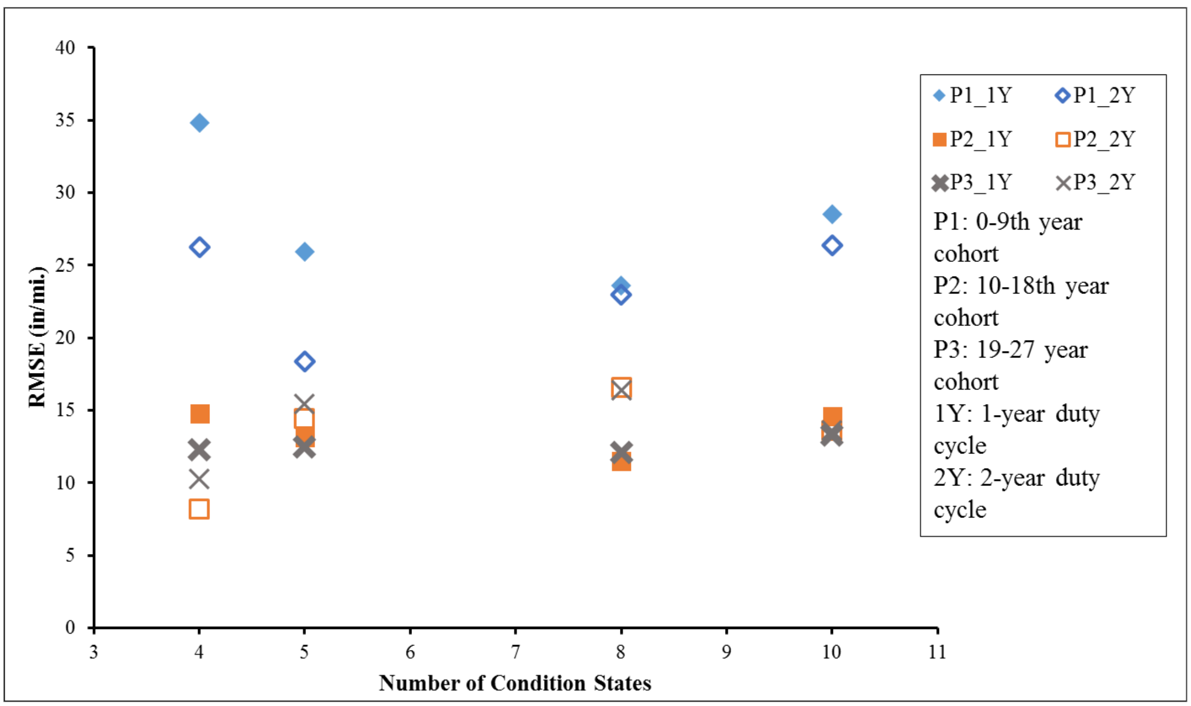
Figures 4.6(a and b) show that the prediction accuracy of homogeneous Markov models developed during P1 increases (i.e., MAPE and RMSE decrease) when the NCS increases from 4 to 8 at 1-year duty cycle. However, the prediction accuracy decreases when the NCS becomes greater than 8, which could be an indication of overfitting for the actual pavement condition data. Furthermore, this change in prediction accuracy when the NCS is greater than 8 suggests a limit to increasing the NCS. Such a limit should be determined for each unique dataset. These figures indicate that the prediction accuracy improves when the NCS rises from 4 to 5 at a 2-year duty cycle but decreases when the NCS is greater than 5. This implies that the upper limit of the NCS depends on the LDC. These findings support the hypothesis that for the homogenous Markov models, the number of condition states (NCS) affects the prediction accuracy of pavement condition.

### 4.5.3. Significance of Length of Duty Cycle

Figures 4.6(a and b) indicate that the homogeneous Markov models developed during the early years (P1), assuming a 2-year duty cycle, are more accurate compared to their 1-year duty cycle counterpart. It is worth noting that homogeneous Markov models are built using data for two consecutive transitions of pavement condition. Besides, there is relatively little variation in the pavement condition during its early age. Therefore, increasing the LDC helps in capturing a larger change in the pavement condition, which in turn may lead to more accurate predictions of pavement condition. During P2 and P3, increasing the LDC from 1 to 2 years increases the prediction errors, except when the NCS is 4. This indicates that, when a large NCS is used, a short transition/duty cycle should be adopted. On the other hand, when the NCS is small (e.g., 4 states), a long transition/duty cycle should be used to capture the variations in the deterioration of pavement condition from one state to another.




(a)




(b)

Figure 4.6. Variation of prediction accuracy of homogeneous Markov models across NCS and LDC and different periods of data collection: (a) MAPE; and (b) RMSE

## 4.6. Staged-homogeneous Markov Models

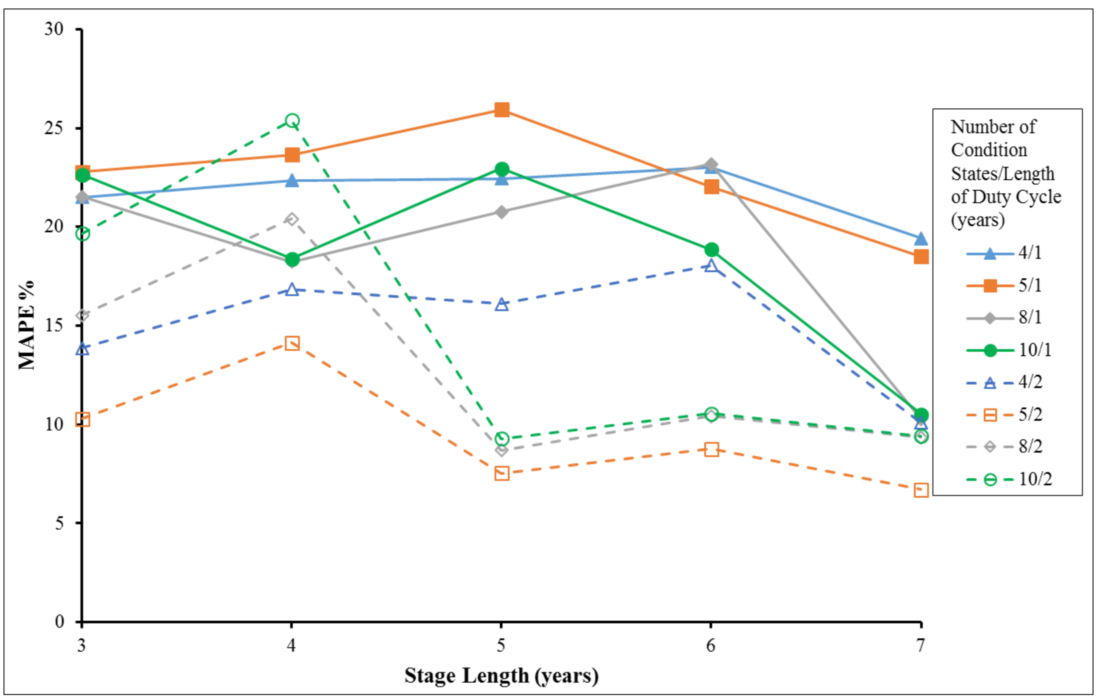
Staged-homogeneous Markov models were first developed by Butt et al. (1987) to partially relax the assumption of the homogeneous Markov models related to the constant deterioration rate of pavement condition. Staged-homogeneous Markov models assume that the pavement condition deteriorates at varying rates at every stage or period of time. In other words, staged-homogeneous Markov models can be defined as a series of homogeneous Markov models, each with different estimates of the transition probability matrix over the pavement life span. For pavements, the stage length in the staged-homogeneous Markov models has been assumed to be 5 or 6 years because pavement condition rarely changes significantly over periods less than 5 - 6 years (Butt et al. 1987). Nevertheless, this is not always the case, particularly when the NCS and LDC are taken into consideration. Staged-homogeneous Markov models are more likely to have greater accuracy in predicting pavement condition compared to homogenous Markov models, because they partly consider the continuous change in pavement condition from one stage to another. However, they still do not capture the continual non-stationary nature of pavement condition during the presumed stage length.

The current research hypothesizes that the length of the stage has a significant effect on the prediction accuracy of the staged-homogeneous Markov models for different NCS and LDC combinations. Hence, staged-homogeneous Markov models were developed for each duty cycle (one-year and two-year), each number of condition states ranging from 4 to 10, and each stage size from 3 to 7 years. Therefore, a total of 40 staged-homogeneous Markov models were established and analyzed.

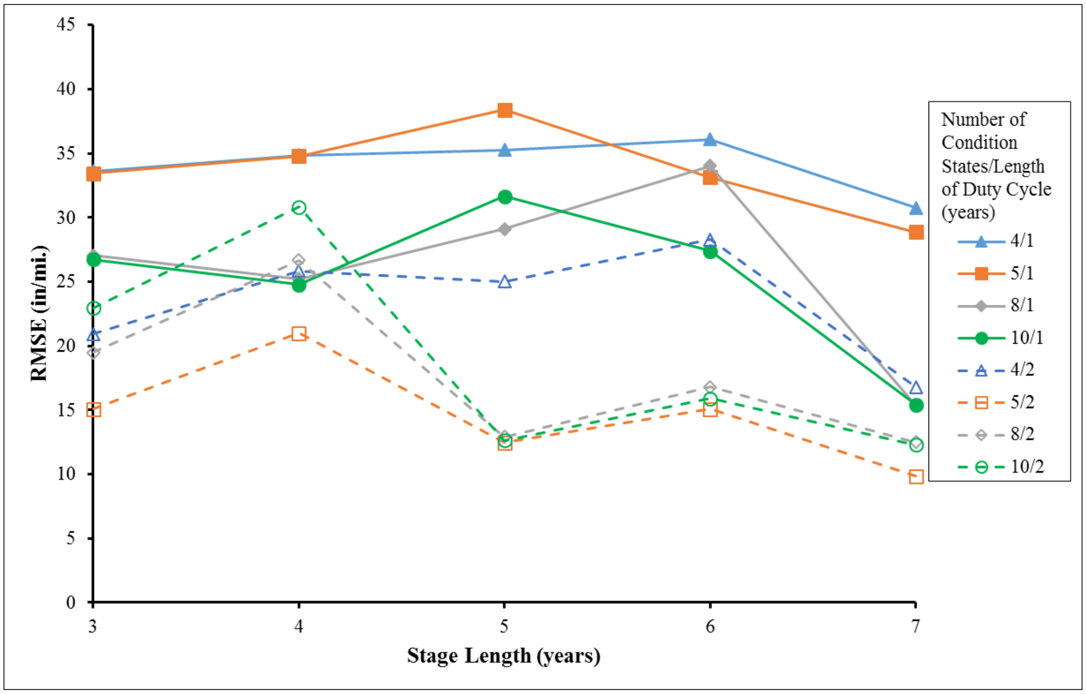
### 4.6.1. Significance of Length of Stage

Figures 4.7(a and b) present the estimated MAPE and RMSE of the predicted pavement condition using different staged-homogeneous Markov models at different lengths of stage. It can be observed that the most accurate predictions for pavement condition can be achieved when the stage length is 7 years for almost all NCS and LDC combinations. On the contrary, the prediction accuracy was found to be the worst when the stage length is 4 years for all NCS at 2-year duty cycle, and when the stage length is 5 or 6 years for all NCS at 1-year duty cycle. In past studies

(Butt et al. 1987; Abaza 2016a, b), the typical design of the staged-homogeneous Markov models is: 10 condition states, 1-year duty cycle and 5 - 6 years of stage. Abaza (2016a) developed two 3- and 5-year stage Markov models at NCS and LDC of 10 states and 1 year, respectively, and found that the 3-year stage model outperforms the 5-year stage model. The findings of this research indicate that the staged-homogeneous Markov models of 3-year stage are more accurate compared to that of 5-year stage, when the NCS and LDC are equal to 10 states and 1 year, respectively. This is similar to Abaza (2016a)'s findings but goes further to test the significance of the stage length with several NCS and LDC combinations. Based on the results of the current study, the typical design of staged-homogeneous Markov models can be improved if different stages, such as the 7-year stage, are adopted. These findings support the notion that the length of the stage in staged-homogeneous Markov models has a significant impact on the accuracy of pavement condition predictions.



(a)



(b)

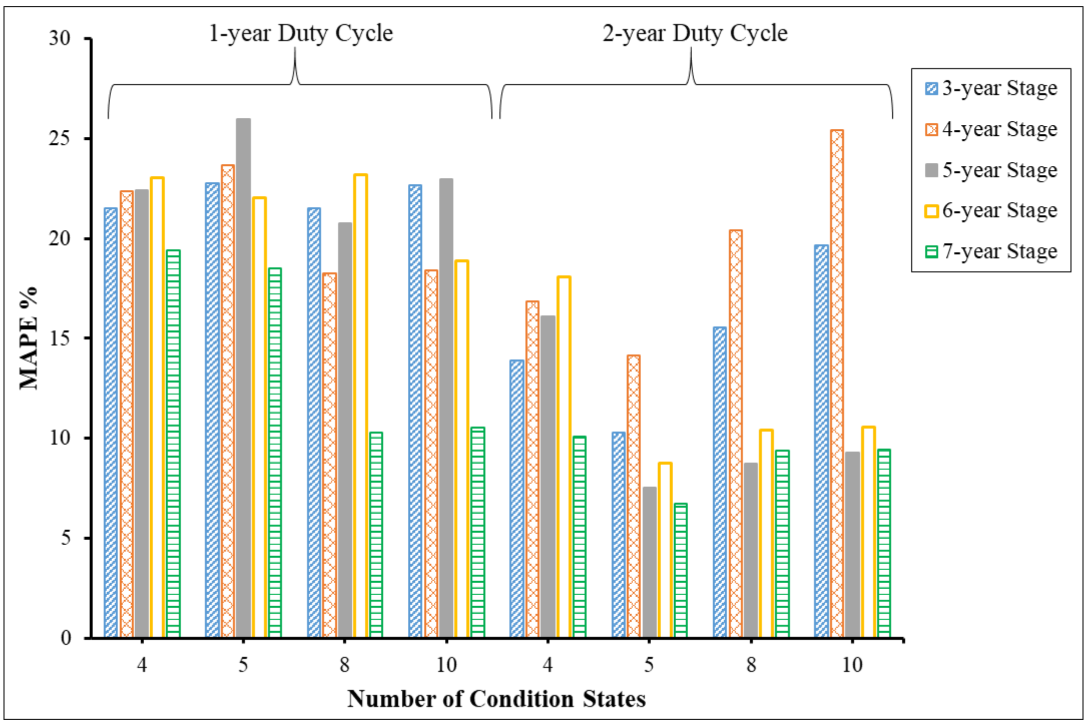
Figure 4.7. Variation of prediction accuracy of staged-homogeneous Markov models of different lengths of stage with NCS and LDC combinations: (a) MAPE; and (b) RMSE

### 4.6.2. Significance of Number of Condition States

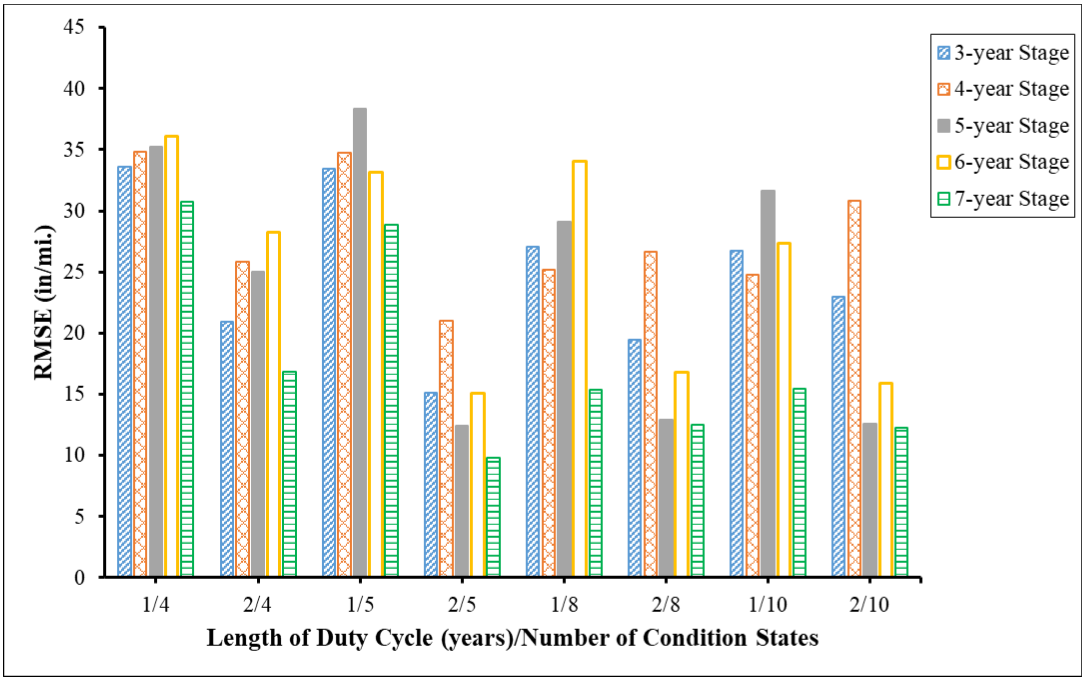
Figure 4.8(a) presents the prediction errors in terms of MAPE of the predicted pavement condition using different staged-homogeneous Markov models and with different NCS and LDC combinations. This figure clearly indicates that the most accurate predictions are obtained when the staged-homogeneous Markov models are designed using 5 condition states, 2-year duty cycle and 7-year stage. At the stage of 7 years and 1-year duty cycle, the rise in the NCS increases the prediction accuracy, but when the NCS is greater than 8 the prediction accuracy begins to decline. On the basis of these results, the typical design of staged-homogeneous Markov models (10 condition states, 1-year duty cycle and 5- or 6-year stage) can be improved if 5 condition states and 2-year duty cycle are used instead, at the same stage length of the typical design. Such results reinforce the argument that the NCS does influence the prediction accuracy of staged-homogeneous Markov models for pavement condition.

### 4.6.3. Significance of Length of Duty Cycle

As observed in Figure 4.8(b), for a 1-year duty cycle, the best configuration of the staged-homogeneous Markov models is 10 condition states with a 7-year stage. On the other hand, for a 2-year duty cycle, the best design is 5 condition states with a stage length of 7 years. It can be indicated that the staged-homogeneous Markov models of 2-year duty cycle are more likely to yield higher accurate predictions than that of 1-year for all combinations of NCS and stage lengths except for the 4-year stage with 8 or 10 condition states. These findings suggest that the LDC does have a significant effect on the prediction accuracy of staged-homogeneous Markov models for pavement condition.




(a)




(b)

Figure 4.8. Variation of prediction accuracy of staged-homogeneous Markov models across NCS and LDC combinations and with different stage lengths: (a) MAPE; and (b) RMSE

### 4.7. Semi-Markov Models

The holding time, which is the period of time that infrastructure asset remains in a specific condition state, is a specific component of the semi-Markov models. In this research, the holding time of a condition state $s$ is estimated as the average value of the holding times of that condition state for all pavement sections:

$$h_s = \frac{\sum_{i=1}^{i=N} t_{s,i}}{N_s} \tag{4.5}$$

where $h_s$ is the mean holding time for condition state $s$, $t_{s,i}$ is the holding time of condition state $s$ for pavement section $i$, $N$ is the total number of pavement sections in the road network, and $N_s$ is the number of pavement sections that were in condition state $s$. Table 4.3 shows the computed holding times for each condition state at each NCS (4, 5, 8 and 10). The holding times of the last state in each group of condition states (e.g., state 4 in the group of 4 condition states) are not calculated because whenever pavement section enters the last state, it stays there until major rehabilitation or reconstruction is carried out.

Table 4.3. Holding times of condition states

| **Holding Times** (years) | **Number of Condition States** | | | |
|---|---|---|---|---|
| | **4** | **5** | **8** | **10** |
| $h_1$ | 1.00 | 1.00 | 1.00 | 1.00 |
| $h_2$ | 4.00 | 2.80 | 1.39 | 1.10 |
| $h_3$ | 4.64 | 2.73 | 1.59 | 1.36 |
| $h_4$ | - | 3.77 | 1.51 | 1.28 |
| $h_5$ | - | - | 1.38 | 1.21 |
| $h_6$ | - | - | 2.28 | 1.15 |
| $h_7$ | - | - | 2.03 | 1.50 |
| $h_8$ | - | - | - | 1.58 |
| $h_9$ | - | - | - | 1.48 |
| $h_{10}$ | - | - | - | - |

Eight semi-Markov models were developed for each duty cycle (one-year and two-year) and each number of condition states: 4, 5, 8 and 10.

### 4.7.1. Significance of Number of Condition States

Figure 4.9 presents the estimated prediction errors (MAPE and RMSE) for the eight semi-Markov models that were developed. The best configuration of the semi-Markov models (that is, the one with the lowest MAPE and RMSE) has 5 condition states and a 2-year duty cycle. The worst design for the semi-Markov models is 4 condition states and 1-year duty cycle. Figure 4.9 also implies that the increase in the NCS when the duty cycle is 1 year boosts the semi-Markov model prediction accuracy. Nevertheless, when the NCS is greater than 8 the prediction accuracy reduces. Likewise, the prediction accuracy of the semi-Markov models at 2-year duty cycle improves if the NCS increases from 4 to 5, but begins to decrease when the NCS exceeds 5. This observation is suggestive of the existence of interactions between the NCS and LDC. Overall, the results support the hypothesis that the NCS has an effect on the semi-Markov model prediction accuracy.

### 4.7.2. Significance of Length of Duty Cycle

As Figure 4.9 indicates, for semi-Markov models, the prediction accuracy of a 2-year duty cycle is significantly higher than that of a 1-year duty cycle, at 4 or 5 condition states. Conversely, at 8 or 10 condition states the semi-Markov models of a 1-year duty cycle outperform those of a 2-year duty cycle. Previous semi-Markov models by Nesbit et al. (1993) and Thomas and Sobanjo (2012) used 10 condition states and a 1-year duty cycle. The results of the current research imply that the semi-Markov models with the NCS and LDC used in the previous research yield more accurate pavement condition predictions than others with different NCS and LDC combinations. Yet still, the analysis of the current study shows that more accurate predictions could be obtained if the semi-Markov models use the following designs: 2-year duty cycle with 4 or 5 states or 1-year duty cycle with 8 states. These findings support the argument that for semi-Markov models, the LDC has an impact on the prediction accuracy of pavement condition.

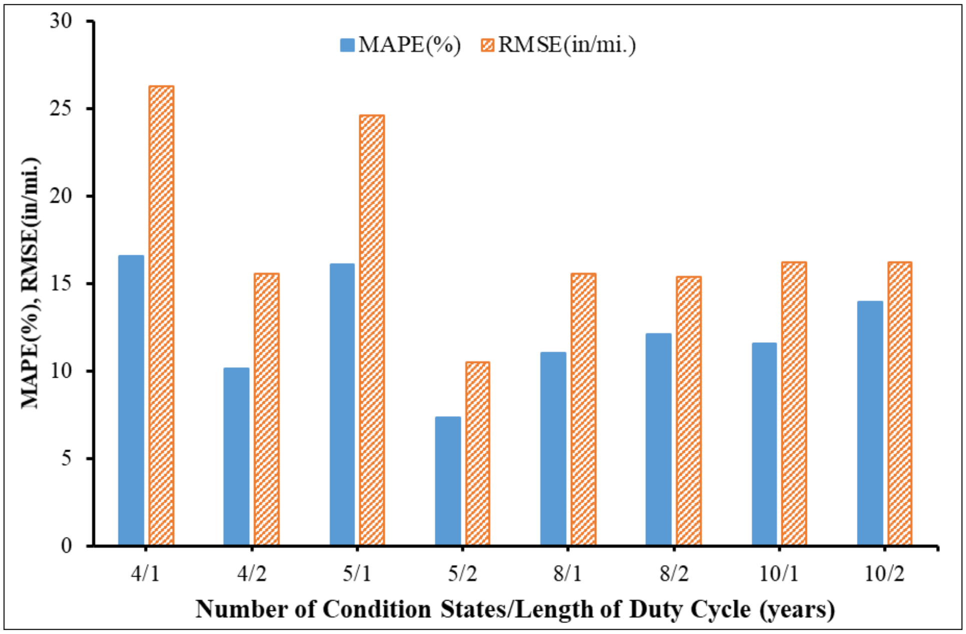



Figure 4.9. Variation of prediction accuracy of semi-Markov models at different NCS and LDC combinations

## 4.8. Comparative Analysis of Markovian Models across the Markov Methodologies

This section discusses the prediction accuracy of Markov models of different methodologies (homogeneous, staged-homogeneous and semi-Markov) with different NCS and LDC combinations. Figure 4.10 shows that for all three Markov methodologies, there is a decrease in the prediction error when the NCS increases from 4 to 8. However, when 10 condition states are used, their prediction errors increase. This indicates that there seems to exist an upper limit beyond which an increase in the NCS is not desirable. In future research, this threshold must be determined for each Markov methodology.

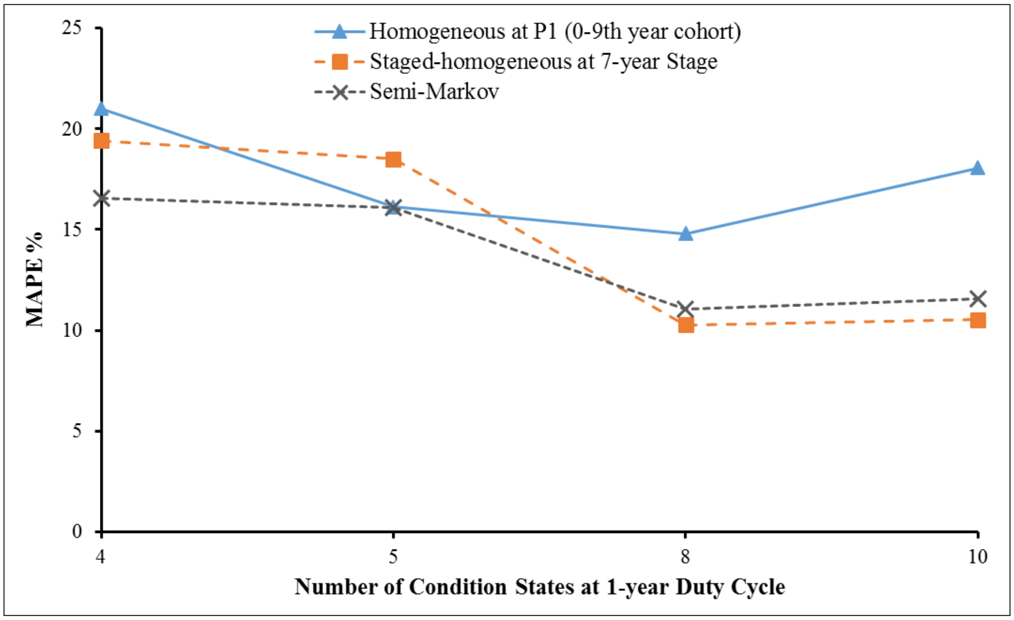



Figure 4.10. Prediction accuracy of the three Markov methodologies at different NCS and 1-year duty cycle

Table 4.4 and Figure 4.11 summarize the best designs of Markov models of different methodologies and their corresponding prediction errors (MAPE/RSME). The Markov models with the designs shown in Table 4.4 have comparable prediction accuracy for pavement condition. It can be seen that most of the best models consist of 5 condition states and 2-year duty cycle. However, the homogeneous Markov model with the design of 4 condition states and 2-year duty cycle using data collected during the period P2 has the highest prediction accuracy of approximately 95% (i.e., MAPE = 5.29%).

Table 4.4. Summary of the highest performing Markovian models

| **Markov Models** | | **Design** | | **Prediction Error** |
|---|---|---|---|---|
| **Methodology** | **Periods of Data Collection** | **NCS** | **LDC** (years) | **MAPE/RMSE** |
| **Homogeneous Markov** | P1 (0-9 yrs) | 5 | 2 | 10.94/18.37 |
| | P2 (10-18 yrs) | 4 | 2 | 5.29/8.20 |
| | P3 (19-27 yrs) | 4 | 2 | 6.57/10.25 |
| **Staged-homogeneous Markov** | every 3 yrs | 5 | 2 | 10.28/15.08 |
| | every 4 yrs | 5 | 2 | 14.14/20.99 |
| | every 5 yrs | 5 | 2 | 7.53/12.46 |
| | every 6 yrs | 5 | 2 | 8.77/15.08 |
| | every 7 yrs | 5 | 2 | 6.72/9.84 |
| **Semi-Markov** | 0-27 yrs | 5 | 2 | 7.33/10.49 |

Figure 4.11 indicates that in most Markov models, the semi-Markov models outperform the staged-homogeneous Markov models which, in turn, outperform the homogeneous Markov counterparts. These results are consistent with the literature (Butt et al. 1987; Thomas and Sobanjo 2012; Abaza 2016a, b). However, some instances shown in Table 4.4 and Figure 4.11 (marked a and b in Figure 4.11) imply that the homogeneous and staged-homogeneous Markov models outperform the semi-Markov models when different NCS, LDC, data collection time and stage length configurations are used. These findings confirm the hypothesis of the current study that the NCS and LDC have a significant impact on the prediction accuracy of Markov models for pavement condition. Furthermore, the time of data collection in the homogeneous Markov models and the stage size in the staged-homogeneous Markov models were found to have significant effects on the accuracy of pavement condition predictions. The homogenous and staged-homogeneous Markov models can produce more accurate predictions compared to the semi-Markov models when NCS, LDC, data collection time and stage size are considered in the design parameters of Markov models. Also, it is worth mentioning that they are less computationally expensive compared to semi-Markov models.

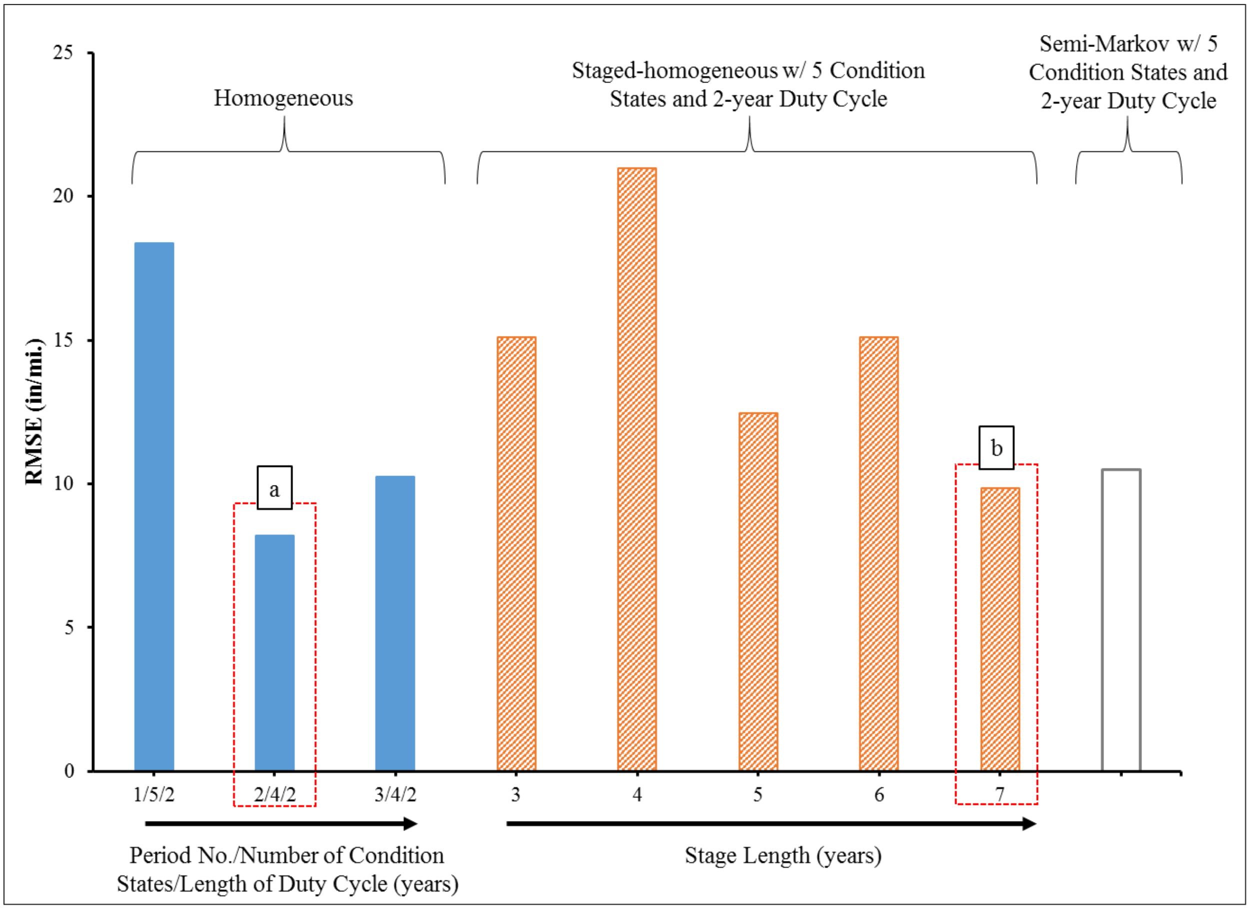



Figure 4.11. Best designs for homogeneous, staged-homogeneous and semi-Markov models

The nomograph shown in Figure 4.12 was constructed as a decision-making tool to select the appropriate Markov technique given the NCS and LDC. This graph was designed using the results of the current research for the semi-Markov models, the homogenous Markov models developed at the period P3 (i.e., when data are collected during late pavement ages), and the staged-homogeneous Markov models built at the typical stage size of 5 and 6 years. This graph shows, for example, that if a 1-year duty cycle and 5 condition states are used to build a Markov model, say M, the most accurate methodology for predicting pavement condition is the homogeneous Markov.

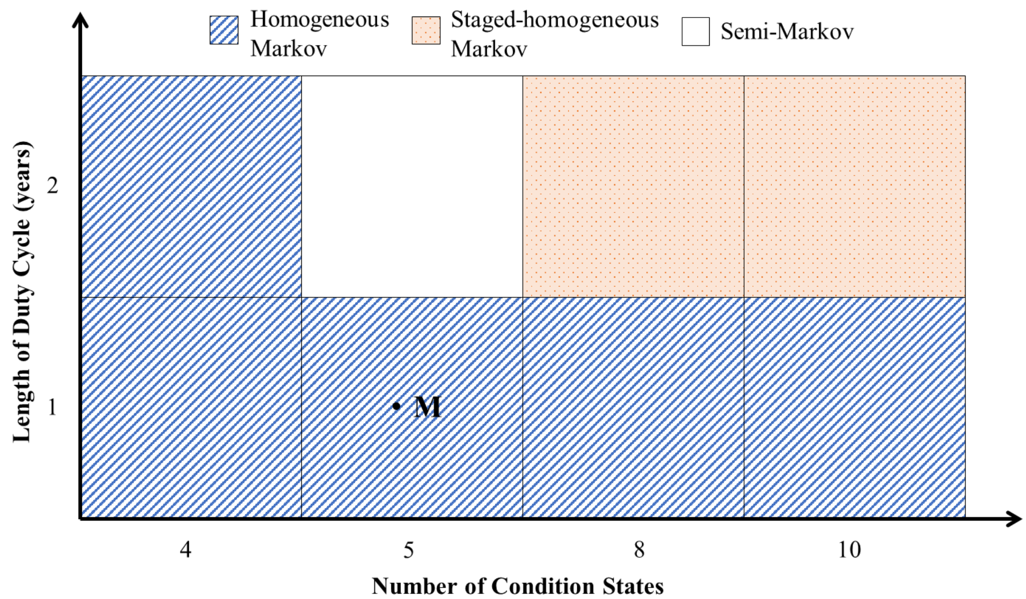



Figure 4.12. Nomograph for Markov methodology selection

## 4.9. Summary

This chapter analyzed the statistical significance of the design parameters of Markovian pavement performance models for the prediction accuracy of pavement condition, namely the number of condition states (NCS), the length of duty cycle (LDC), the periods of data collection (P1, P2 and P3) in the homogeneous Markov models and the stage size in the staged-homogeneous Markov models. A comparative analysis was carried out for the different Markovian pavement performance modeling techniques, and for each technique, the different configurations of their design parameters.

The results suggest that in some instances the semi-Markov models outperform the staged-homogeneous and the homogeneous Markov counterparts, which is consistent with the literature. Nevertheless, the staged-homogeneous and homogeneous Markov models were found to be superior to the semi-Markov models when using specific NCS and LDC. The results also showed that the NCS and LDC significantly affect the prediction accuracy of each of the three methodologies: homogeneous, staged-homogeneous and semi-Markov. It was found that increasing the NCS increases the prediction accuracy of the three Markov methodologies until the NCS reaches 8 (for a 1-year duty cycle) and 5 (for a 2-year duty cycle). Beyond these thresholds, the prediction accuracy begins to decrease. Such reductions in the prediction accuracy could be attributed to data overfitting as the NCS increases. Future researchers and highway agencies should therefore investigate whether the critical NCS and LDC identified in this research are really the best choice in terms of model prediction accuracy.

In the homogeneous and staged-homogeneous Markov models, the pavement cohorts (P1, P2 and P3) at data collection time and stage size were found to have significant impacts on the predictive accuracy of Markov models. The homogeneous Markov models were found to have high prediction accuracy when using data collected during the middle or late pavement ages. Unlike the early ages, the middle and late ages include a variety of pavement sections with different pavement ages and varying conditions, which are more useful for the prediction accuracy of statistical models. The results also confirm that the use of a stage length other than the typical length (5 or 6 years) for the staged-homogeneous Markov models yields more accurate predictions. For example, the staged-homogeneous Markov model with 7-year stage length, 5 condition states and 2-year duty cycle is more accurate than that with 5- or 6-year stage length. Therefore, when designing homogeneous and staged-homogeneous Markov models, more attention should be paid to the data collection time and the stage length, respectively.

# CHAPTER 5. STOCHASTIC OPTIMIZATION OF PAVEMENT PREVENTIVE MAINTENANCE

While seeking to improve the condition of their pavement network by implementing maintenance and rehabilitation (M&R) interventions, highway agencies are challenged by limited funding resources. As a means to manage their limited resources, they develop M&R strategies that contribute to the 3R concept: the right treatment at the right time for the right pavement section. Deterministic optimization models that consider decision variables to be fixed values have been widely developed and used to schedule M&R for road networks. On the other hand, current stochastic optimization models, which account for the uncertainty of the M&R decision variables, consider the uncertainty of budget constraint only and do not account for decision variables, such as the deterioration of pavement condition and the improvement of pavement condition following maintenance interventions, which may also have uncertain outcomes. The selection of optimal timings and types of maintenance treatments for a pavement network over the long-term (pavement design life) without considering the uncertainty of expected pavement condition and maintenance effectiveness can lead to mistiming of maintenance applications and can therefore result in less optimal alternatives.

This chapter develops stochastic pavement maintenance optimization models accounting for the uncertainty of the deterioration and improvement of pavement condition as well as the budget constraint. The objectives of the models are to minimize the overall road network deterioration, while at the same time minimizing the total maintenance costs of road network during a planning horizon of 20 years [typical pavement design life (Morian et al. 2005; Ceylan et al. 2009; Santos and Ferreira 2013)]. Multi-objective Genetic Algorithm (MOGA) is used due to its robust search capability resulting in optimal or near-optimal global solutions. To reduce the large size of the stochastic MOGA optimization problem at the network level, three approaches were proposed and applied to interstate flexible pavements across the Midwestern States. These approaches are: (1) identifying and adopting the most commonly used maintenance treatments (2) clustering pavement sections based on pavement age, and (2) creating a filtering constraint that applies a rest period after treatment applications. The results of the current study show that the Pareto optimal solutions significantly change when considering the uncertainty of pavement condition deterioration and improvement. The developed stochastic MOGA models can provide highway agencies with probabilistic Pareto optimal solutions that account for the expected uncertainty in pavement condition data.

## 5.1. Introduction

This section presents the different pavement maintenance and rehabilitation (M&R) treatments that have been used by highway agencies to keep their infrastructure in a state of good condition. Besides, it discusses pavement M&R strategies, measures of effectiveness of maintenance treatments and levels of M&R decision-making.

### 5.1.1. Pavement Maintenance and Rehabilitation

Pavement rehabilitation is defined as the activities that restore the original pavement serviceability (Humphries and Ma 2004) and implemented to structurally deficient pavements. On the other hand, pavement maintenance or preservation is defined, according to Federal Highway Administration (FHWA) and the U.S. Department of Transportation, as "a program employing a network level, long-term strategy that enhances pavement performance by using an integrated, cost-effective set of practices that extend pavement life, improve safety and meet motorist expectations." A pavement maintenance program consists primarily of three components:

1. Minor rehabilitation (non-structural)
2. Preventive maintenance
3. Routine maintenance

An effective maintenance program addresses the pavement while it is still in good condition. A cost-effective treatment in a timely manner restores pavements to their original conditions. By doing so, the cumulative costs of such treatment are substantially lower than reconstruction or major rehabilitation over pavement life (Wilde et al. 2014). In addition, the disruption of traffic is less for more frequent and minimal treatments in comparison to the reconstruction or major rehabilitation interventions. The main objectives of pavement maintenance are preventing moisture and/or debris from infiltrating into the pavement through cracks and/or joints and reducing or preventing deterioration due to environmental effects. Table 5.1 shows the general guidelines of applying the different types of maintenance and rehabilitation activities to pavements, and the effects of each activity on: pavement strength, aging and serviceability.

Table 5.1. Pavement preservation guidelines (adapted from Wilde et al. 2014)

| | **Activity** | **Increase Strength** | **Reduce Aging** | **Restore Serviceability** |
|---|---|---|---|---|
| | Major Rehabilitation | X | X | X |
| Pavement Preservation | Minor Rehabilitation | | X | X |
| Pavement Preservation | Preventive Maintenance | | X | X |
| Pavement Preservation | Routine Maintenance | | | X |
| | Corrective Maintenance | | | X |

Preventive maintenance should be applied to pavements in good condition that have significant remaining service life (RSL). It applies cost-effective treatments to the surface or near surface of structurally sound pavements to preserve, retard future deterioration, and maintain or improve the functional condition of the highway surface. The pavement engineer must rely on knowledge of pavement deterioration processes, engineering judgment, time, and traffic levels to determine the timing and type of the required preventive maintenance treatments. Since the stochastic optimization models developed in this dissertation are demonstrated within the context of interstate flexible/asphalt pavements, examples of preventive maintenance treatments applied to flexible pavements might include the following:

- Crack sealing
- Chip seals
- Micro-surfacing
- Ultra-Thin Bonded Wearing Course (UTBWC)
- Thin hot-mix asphalt (HMA) overlay

Crack sealing is a treatment used to prevent water and debris from entering cracks in the pavement. It may require routing to clean the crack and create a reservoir to hold the sealant. Crack sealing should be applied as needed whenever cracks are observed (Figure 5.1). Chip seals are typically constructed by spraying a layer of asphalt emulsion binder on a roadway and then embedding finely graded aggregate (3/8-inch crushed rock) into it (Figure 5.2). The aggregate is rolled after being evenly dispersed on pavement surface. Chip seal is primarily used on low-volume roads (average daily traffic < 5,000 and 10,000 vehicle per day for rural and urban roads, respectively) (Peshkin et al. 2011) because it creates excessive noise and loose chips can break windshields (Lee and Shields 2010).

Figure 5.1. Crack Sealing

Figure 5.2. Chip Seal

Micro-surfacing treatments involve the laying of a mixture of crushed mineral aggregate, polymer-modified asphalt emulsion, mineral filler, water, and an additive to control hardening of the mixture. A self-propelled pug mill mixes the components and lays the mix immediately after mixing (Figure 5.3). No compaction is required, and the finished surface can generally be opened to traffic soon after placement [1hour (Peshkin and Hoerner 2005)]. A micro-surfacing layer may be as thin as 3/8 inch and is capable of filling wheel ruts up to 1.5 or 2 inches deep (Labi et al. 2006).

Figure 5.3. Micro-surfacing

Ultra-Thin Bonded Wearing Course (UTBWC) is the application of a warm polymer modified emulsion membrane followed immediately with an ultra-thin wearing course (3/8-inch crushed rocks) (Figure 5.4). The surface is then compacted, and no traffic is allowed until pavement has cooled (Lee and Shields 2010). The UTBWC can be as thin as 0.5 to 1.5-inch-thick (Midland Asphalt Materials Inc. 2018).

Figure 5.4. Ultra-Thin bonded Wearing Course (UTBWC)

Thin HMA overlay consists of one thin layer of HMA pavement (generally 1.5 inches). The existing pavement is first profile-milled, then a light tack coat is placed prior to placing a uniform lift of the HMA mixture (Figure 5.5). Finally, the surface is compacted by rolling to achieve the desired density. The finished surface can be opened to traffic immediately after the rolling operation is completed (Lee and Shields 2010).

Milled Surface Finished Surface

Figure 5.5. Thin HMA Overlay

### 5.1.2. Pavement Maintenance Strategies

Highway agencies develop pavement maintenance strategies to maintain their assets at desired condition levels and to optimize their limited financial resources. These maintenance strategies determine which pavement sections need maintenance, what type of maintenance treatments are needed, and at which times over the life of each pavement section. The main goal of maintenance strategy is to attain as much improvement as possible in the condition of the entire pavement network at the lowest cost possible. To this end, life-cycle cost analysis is usually adopted to compare different strategies from the economic perspective and determine the most cost-effective one over a planning horizon. If the available funding is less than those identified as needed for any year during the analysis period, prioritization/ranking models (Wong et al. 2003; Kulkarni et al. 2004) or optimization models (Lampety et al. 2010; Irfan at al. 2012; Abaza et al. 2004; Madanat et al. 2006; Elhadidy et al. 2015; Aleadelat et al. 2018; Augeri et al. 2019; Guo et al. 2020; Sindi and Agbelie 2020) can be used to optimally allocate the available funding to achieve the greatest overall return on investment at the system-level performance.

### 5.1.3. Levels of Decision-Making for Pavement Maintenance

Decisions regarding pavement maintenance are made either at the network level or at the project level. At the network level, the decision is made to choose different pavement sections to be maintained in order to improve the entire roadway network efficiency/performance. The objective of this level of optimization is to maximize pavement network condition under the constraint of limited budget. Whereas at the project level, the decision is made to determine the optimal condition and time for maintenance application and the cost-effective maintenance treatments. The optimization objective at this level is to maximize the cost-effectiveness of pavement maintenance under the constraints of available budget and allowable level of pavement condition (M&R triggers/thresholds). Figure 5.6 shows the levels of decision-making for pavement maintenance. From the left to the right-hand side (Figures 5.6(a, b, c)), at the network-level a group of pavement sections is first selected for maintenance. Then, for each pavement section a project-level optimization is carried out to select the cost-effective treatments. Finally, another project-level optimization is conducted to optimize the maintenance timing to ensure implementing the right treatment at the right time.

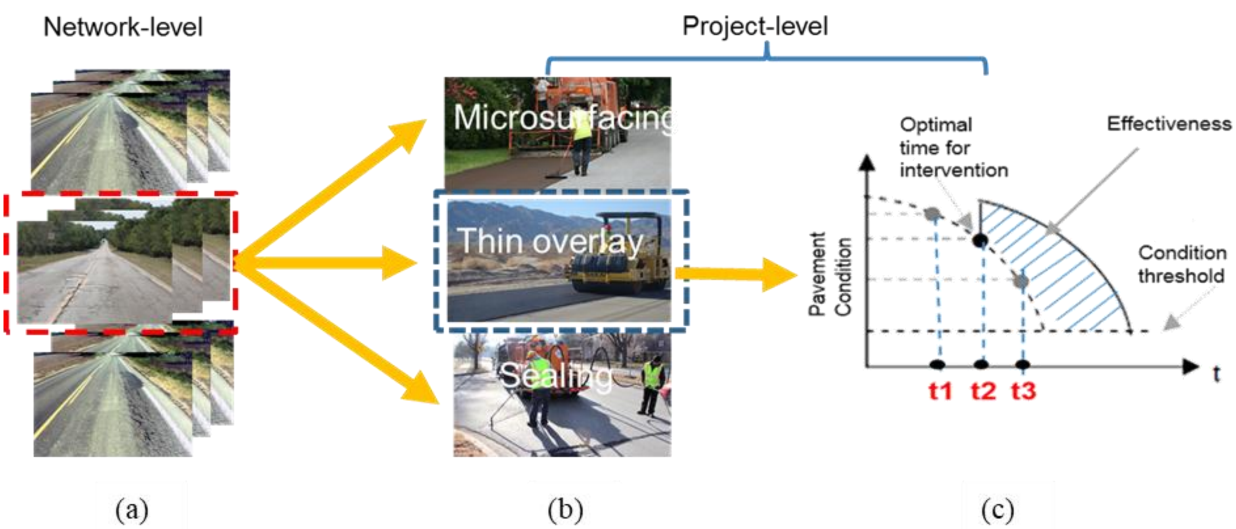
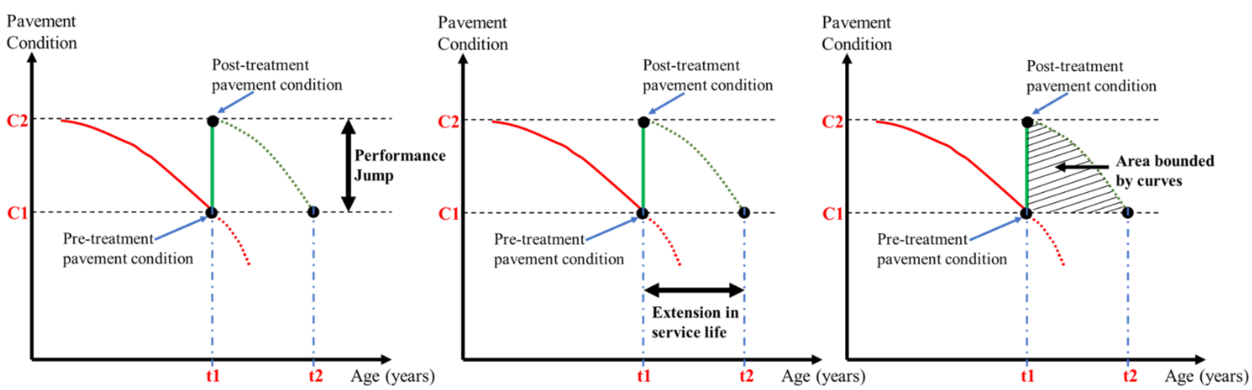



Figure 5.6. Levels of decision-making for pavement maintenance

### 5.1.4. Measures of Pavement Maintenance Effectiveness

At the project-level optimization of pavement maintenance, the objective function is usually to maximize the benefits/effectiveness of pavement maintenance by selecting the optimal treatment and timing. In order to measure the maintenance effectiveness, three measures of effectiveness (MOE) have been found in the literature: performance jump (PJ), extension in service life (SL), and area bounded by performance curves. As shown in Figure 5.7(a), the performance jump is the instant improvement in pavement condition after the treatment implementation. This improvement may be an increase in pavement condition if a decreasing indicator (a condition indicator such as the pavement condition rating (PCR); the decrease in its value means more pavement deterioration) is used, or a drop in pavement condition if an increasing indicator (a condition indicator such as the international roughness index (IRI); the increase in its value means more pavement deterioration) is used. The extension in service life is the elapsed time, $(t_2 - t_1)$ until a treated pavement section returns to the pre-treatment condition $C_1$ (Figure 5.7(b)). The third MOE is the area bounded by performance curves. In decreasing indicators, it is the area under the post-treatment performance curve (Figure 5.7(c)). In increasing indicators, it is the area above the post-treatment curve. The third MOE captures the instant improvement of pavement condition, the extended service life, and the pavement deterioration rate after treatment. The PJ is used for the short-term assessment of pavement maintenance effectiveness, whereas extension in SL and area bounded by performance curves are used for the long-term measurement of pavement maintenance effectiveness. The measure of area bounded by performance curves has been widely utilized in estimating the effectiveness of pavement maintenance and rehabilitation (Lamptey et al. 2008; Khurshid 2010; Khurshid et al. 2011;  Khurshid et al. 2014).




(a) (b) (c)

Figure 5.7. Measures of effectiveness: (a) Performance jump; (b) Extension in service life; and (c) Area bounded by performnce curves

### 5.1.5. Approaches of Pavement Maintenance Scheduling

Highway agencies develop their pavement maintenance strategies according to either an age-based approach or a condition-based approach. The age-based approach identifies the timing of pavement maintenance without requiring regular monitoring of pavement condition. This approach can result in implementing maintenance either earlier or later than pavement reaches the optimal condition for maintenance. Most highway agencies in the U.S. use this approach to avoid the short-term cost of monitoring; however, the mistiming of pavement maintenance typically associated with this approach is expensive and can result in waste of resources and decrease of treatment effectiveness. On the contrary, the condition-based approach depends largely on regular monitoring of pavement condition whereby the true time for treatment is determined. This approach is more expensive than the age-based approach due to the additional cost of monitoring of pavement condition.

### 5.1.6. Optimization of Pavement Maintenance

Optimization models are mainly composed of three components: (1) the decision variable, (2) the objective function or the desired goal, and (3) the constraints that are the upper and lower limits for each variable in the optimization models. The most common objective function is the minimization of M&R costs (Abaza et al. 2004; Madanat et al. 2006). Another objective function is the maximization of road network quality or performance (Abaza et al. 2001; Abaza 2006). Yang et al. (2013) developed an optimization model to establish pavement maintenance strategy considering two objective functions: minimization of maintenance cost and maximization of pavement network performance. The most common constraints are the limited budget allotted for M&R interventions and the allowable level of pavement condition or M&R thresholds (Abaza 2006; Ferreira et al. 2009; Lampety et al. 2010; Irfan at al. 2012; Jorge and Ferreira 2012).

Pavement maintenance optimization models can be placed into two main categories: deterministic and stochastic. Deterministic optimization models assume that the optimization design parameters are of fixed values and does not consider the randomness and uncertainty that are inherently attributed to pavement data. Unlike deterministic optimization models, stochastic optimization models have been developed for pavement maintenance to account for the uncertainty inherently associated with pavement data, and to provide the decision makers with more accurate tools that lead to more effective decisions and suits their level of risk/uncertainty. This chapter develops stochastic multi-objective optimization models to schedule pavement preventive maintenance for roadway network. Because the Multi-Objective Genetic Algorithm (MOGA) approach has a robust search for global optimal solutions (compared to other optimization techniques that will be discussed in the following section 5.2), it was used to schedule preventive maintenance interventions for road network while maximizing the functional performance of pavement network and minimizing the total life cycle cost of preventive maintenance.

## 5.2. Prior Research on Pavement Maintenance Optimization

Optimization models can be categorized based on the number of objective functions into two categories: (1) single-objective optimization models that solve for one objective function, and (2) multi-objective optimization models that solve for multiple objective functions. The single-objective optimization models have been employed by pavement researchers and practitioners to avoid the complexity and computational expenses of multi-objective optimization models. However, the solutions obtained from these single-objective models are most likely suboptimal compared to those derived from multi-objective optimization models (Fwa et al. 2000; Meneses et al. 2013). The main challenge of the single-objective optimization models is the selection of one objective function that rationally represents the other objective functions and ensuring that the other objective functions are satisfied (Fwa et al. 2000; Wu and Flintsch 2009).

Multi-objective optimization modeling of pavement M&R have been developed by researchers to overcome the limitations of the single objective models. Fwa et al. (2000) developed an optimization model with three objective functions: (a) the maximization of the work production, (b) the minimization of the total maintenance cost, and (c) the maximization of overall network pavement condition. The model was applied to four highway classes, each one with three need-urgency levels (high, medium and low), considering four M&R interventions and a planning timespan of 45 working days. Wang et al. (2003) developed an optimization model with two objectives: (a) the maximization of the total M&R effectiveness, and (b) the minimization of the total M&R disturbance cost. The model was applied to a small network of 10 road sections considering a planning time-span of 5 years. Meneses et al. (2013) developed a Multi-Objective Decision-Aid Tool (MODAT) and tested it with data from the Estradas de Portugal's Pavement Management System. Two objectives were considered in this study: (a) minimization of agency costs (M&R costs), and (b) minimization of user costs. Meneses and his team used the deterministic pavement performance model from the AASHTO flexible pavement design. Wu and Flintsch (2009), Yang et al. (2013) and Denysiuk et al. (2017) developed optimization models with two objectives: (a) the maximization of pavement network condition, and (b) the minimization of the total M&R cost.

Several parameters are used in the optimization models of pavement M&R such as time, cost, pavement condition, maintenance effectiveness, and maintenance treatments. These parameters are continuous or discrete, and linear or non-linear variables. Based on the type of these parameters, several optimization techniques have been employed in the literature. Examples of these techniques are linear programming, non-linear programming, integer programming, mixed-integer linear programing and Genetic Algorithm. Golabi et al. (1982) used linear programming for solving a homogeneous system problem with the decision variables being the proportions of facilities that need a specific M&R activity at a certain state. Integer programming is another approach to model the infrastructure maintenance problem over a certain planning horizon (Fwa et al. 1988; Jacobs 1992; Wang et al. 2003; Irfan et al. 2012). Medury and Madanat (2014) used a mixed-integer linear programming to optimize pavement M&R at the project and network levels. The Genetic Algorithm (GA) was used in the studies of Chan et al. (1994), Yang et al. (2013), Elhadidy et al. (2015) and Santos et al. (2019) to establish optimal pavement M&R strategies for highway network. Sindi and Agbelie (2020) used the GA method to develop M&R strategies for road network, and they reported that the use of GA results in reliable and accurate global optimal solutions when used at the road network level.

Pavement maintenance optimization models can be placed into two main categories: deterministic and stochastic. Deterministic models are those that assume that the optimization design parameters (input variables) are of fixed/certain values and does not consider the randomness and uncertainty that are inherently attributed to pavement condition data. Such an assumption may be reasonable in some situations in which the deployed data is of high certainty. Prior research has presented a number of deterministic optimization models for pavement M&R such as Fwa et al. (2000), Abaza et al. (2001), Ouyang and Madanat (2004) and Santos et al. (2018; 2019). Deterministic optimization models cannot produce accurate solutions before pavement condition reaches a steady state. Li and Madanat (2002) realized that pavement condition reaches a steady state at the time of the first resurfacing. Thus, they developed a deterministic optimization model to optimize pavement resurfacing. Deterministic optimization models for pavement maintenance can be effective for instantaneous or short-term (e.g., 3 years) decision-making when data are accurately determined, and at the time of steady state of pavement condition (Li and Madanat 2002). For any decision-making between present and time of the steady state, stochastic optimization models need to be considered.

Unlike deterministic optimization models, stochastic optimization models have been developed for pavement maintenance to account for the uncertainty inherently associated with pavement condition data, and to provide decision makers with more accurate tools and additional insights that lead to more cost-effective decisions. Stochastic optimization approaches attempt to achieve the best expected objective value over all possible realizations of randomness/uncertainty. Thus, stochastic optimization models can capture the uncertainty in data and the variability associated with multiple optimal solutions. Failure to include these uncertainties may lead to expensive, even disastrous consequences if the anticipated situation is not realized (Gao and Zhang 2008). Li and Puyan (2006) built a stochastic optimization model to select optimal pavement projects for rehabilitation and they considered the uncertainty of budget constraints. Gao and Zhang (2008 and 2011) accounted for the uncertainty of the budget constraints in their pavement maintenance optimization models. They assumed that pavements deteriorate linearly, and the maintenance improvements can be measured in terms of the performance jump. Wu and Flintsch (2009) developed a chance-constrained programming model to control the probability of going over budget for network-level M&R scheduling. Tables 5.2 encapsulates the key studies found in the literature with respect to stochastic optimization of pavement M&R. This table shows the prior stochastic optimization models, objective functions, constraints, methodology, and uncertainty considerations.

Table 5.2. Prior stochastic optimization models for pavement maintenance

| **Research Team** | **Objective Functions** | **Constraints** | **Methodology** | **Uncertainty Considerations** |
|---|---|---|---|---|
| Gao and Zhang (2008) | Minimize maintenance cost | Pavement condition threshold; budget | Scenario method | Budget |
| Wu and Flintsch (2009) | Maximize service level of pavement network and minimize maintenance cost | Budget | Chance-constrained programming | Budget |
| Li et al. (2010) | Minimize maintenance cost and work zone time | Budget | Heuristic model using lagrange relaxation technique | Budget |
| Gao and Zhang (2011, 2013) | Minimize maintenance cost | Pavement condition threshold; budget | Multi-stage of linear stochastic programming; Augmented Lagrangian Decomposition (ALD) method | Budget |
| Ameri et al. (2019) | Maximize service level of pavement network and minimize maintenance cost | Pavement condition threshold; budget | Two-stage stochastic model with integer programming using General Algebraic Modeling System (GAMS) software | Budget |

Pavement maintenance strategies are developed using optimization models that vary according to several considerations, such as objectives functions and constraints, decision variables, and data certainty or uncertainty. Based on the discussed literature, the main decision variables in pavement maintenance optimization models are pavement condition, maintenance effectiveness, maintenance time, maintenance costs, and available budget. Prior stochastic pavement maintenance optimization models (such as those presented in Table 5.2) considered the uncertainty of budget constraint but did not consider the uncertainty of pavement degradation and improvement, and the time and cost of maintenance interventions. The selection of optimal timings and types of maintenance treatments for a pavement network over the long-term (pavement design life) without considering the uncertainty of expected pavement condition and maintenance effectiveness has the potential to lead to mistiming of maintenance applications and can therefore result in less optimal alternatives. Hence, this chapter develops stochastic pavement maintenance optimization models with the consideration of budget uncertainty in addition to the uncertainty of predicted pavement condition and maintenance effectiveness.

## 5.3. Research Methodology

Figure 5.8 shows the proposed research framework to achieve the research objective of this chapter, in particular the stochastic optimization of pavement preventive maintenance, considering the uncertainties of budget constraint and pavement condition deterioration and improvement.

### 5.3.1. Data Collection and Analysis

A questionnaire survey was designed (see Appendix A.1) and deployed to the 50 State Transportation Agencies (STAs) to collect data regarding pavement condition, preventive maintenance (PM) treatments and criteria for scheduling pavement maintenance. Eighteen STAs responded to the survey including five STAs from the Midwest states. Around 58% of the respondents use one condition indicator to represent pavement condition. Moreover, about 60% of the respondents use IRI as the pavement condition indicator. Further information on the results of the survey can be found in Chapter 3. The data were collected for a single pavement condition indicator, namely IRI as the response variable. The collected data includes six explanatory variables (pavement age, annual average freezing index, annual average daily truck traffic, etc.) as shown in Table 5.3. Although other variables, such as the pavement structure (number, type and thickness of pavement layers), affect pavement condition, they were not considered in the current research because STAs have different standards and specifications for the design, construction and maintenance of pavements.

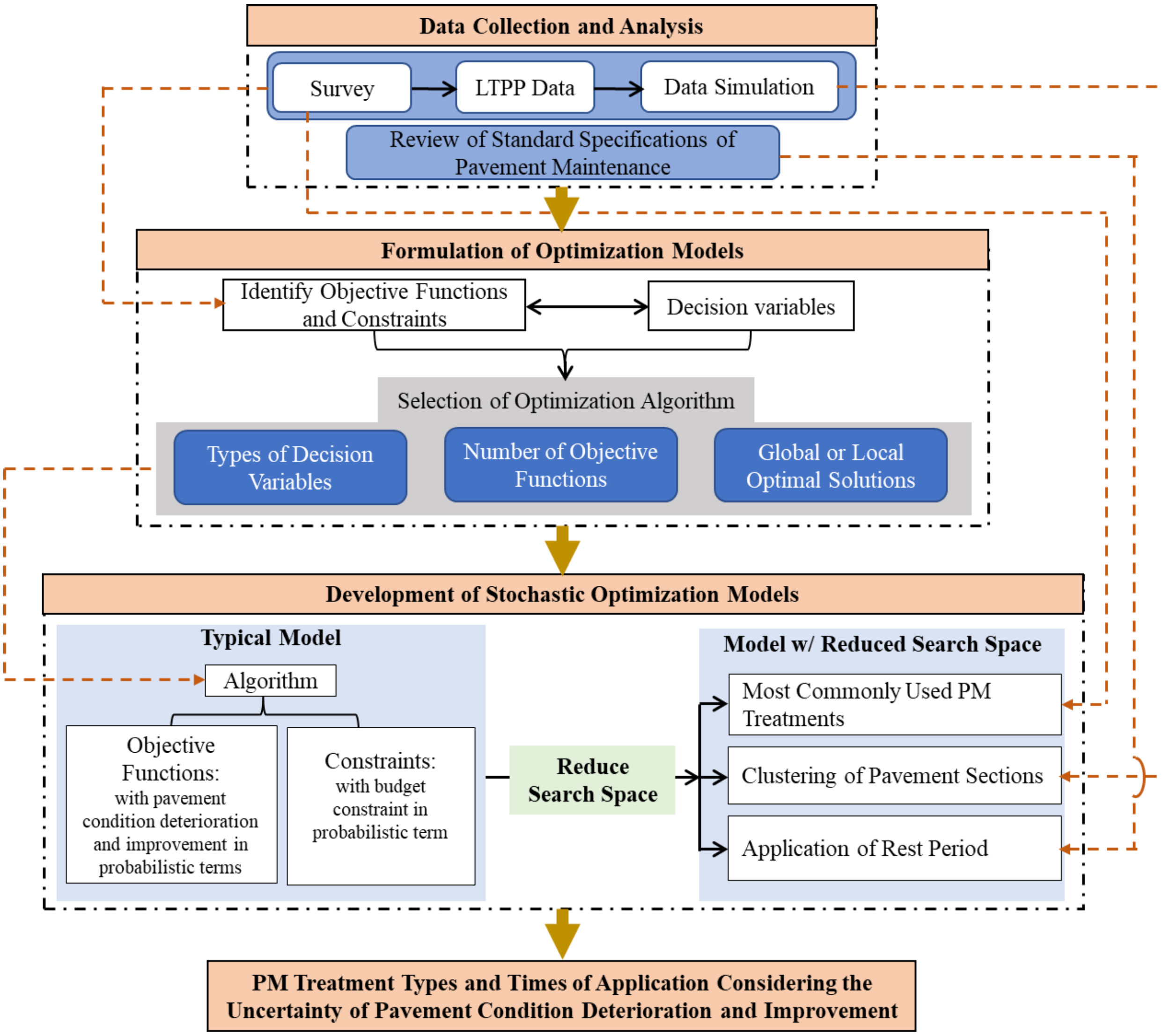



Figure 5.8. Conceptual research framework

Pavement condition data were acquired for the interstate flexible pavements from the LTPP database across eight Midwestern states (Indiana, Illinois, Wisconsin, Michigan, Ohio, Minnesota, Iowa and Missouri). The data were cleaned by deleting the extreme outliers and high-leverage points (Belsley et al. 1980; Ahmed et al. 2016), resulting in 966 observations. Table 5.3 shows the descriptive statistics of variables considered in the analysis.

Table 5.3. Descriptive statistics

| Variable | Description | Mean | STD | Min | Max |
|---|---|---|---|---|---|
| IRI | International Roughness Index (in/mi) | 80.9 | 27.8 | 31.3 | 194.9 |
| Age | Years since construction or rehabilitation | 9.6 | 6.6 | 0.0 | 27.0 |
| AAP | Annual Average Precipitation (inches) | 38.1 | 7.7 | 24.4 | 61.4 |
| AAT | Annual Average Temperature (°F) | 50.5 | 3.3 | 42.8 | 59.7 |
| AAFI | Annual Average Freezing Index (°F days) | 788 | 393 | 70 | 1924 |
| AADTT | Annual Average Daily Truck Traffic | 2169 | 802 | 249 | 5115 |
| ESALs | Equivalent Single Axle Loads (18-Kip) | 1115 | 534 | 122 | 3195 |

As the data collected for the interstate flexible roads is not sufficient to carry out the proposed study, a large amount of data has been produced by simulating the collected pavement condition data. An exponential multiple regression model has been developed to estimate pavement condition (IRI, in/mi.) in relation to the influential independent variables that have been found to statistically significant. To simulate the statistically significant explanatory variables different probability distributions were fitted for each variable, and the distributions that resulted in the minimum negative log-likelihood value were used. The Indiana Department of Transportation (INDOT), as one of the states from which pavement condition data were collected, owns and operates approximately 5,500 miles of interstate roads. Hence, to generate a reasonable number of pavement sections for the current research, data on 5,500 interstate flexible pavement sections (each one mile long) were generated according to the identified probability distributions and the developed exponential multiple regression model.

The questionnaire survey also collected data on PM treatments and criteria for scheduling pavement maintenance (selecting the right treatment at the right time for the right pavement section). Figures (5.9 and 5.10) show the results of the survey with respect to the different PM used and the criteria for maintenance scheduling identified by the responded STAs, respectively. Figure 5.9 indicates that the most commonly used treatments are crack sealing, Thin HMA overlay, Micro-surfacing and UTBWC. Figure 5.10 shows that the most important criteria for scheduling pavement maintenance are pavement condition and maintenance costs.

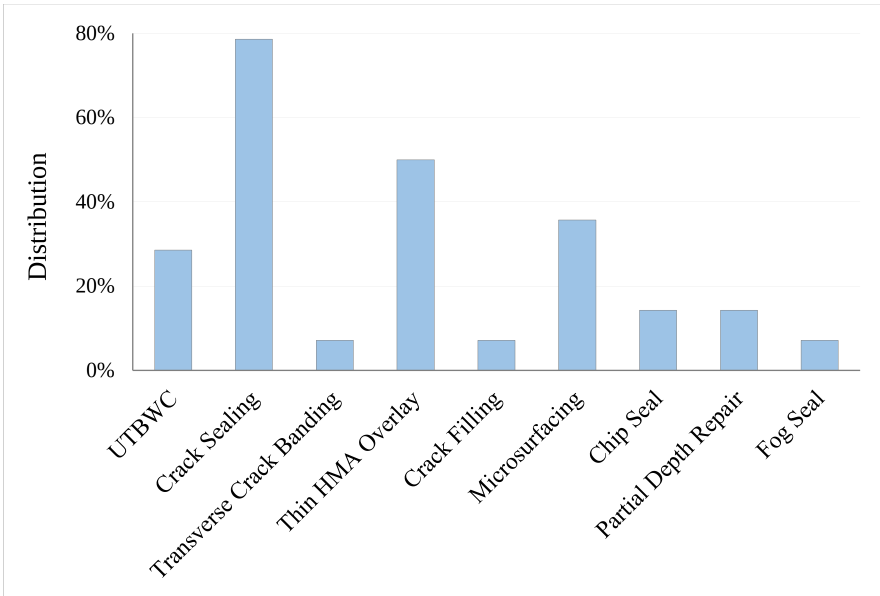



Figure 5.9. Pavement preventive maintenance treatments used by the responded STAs

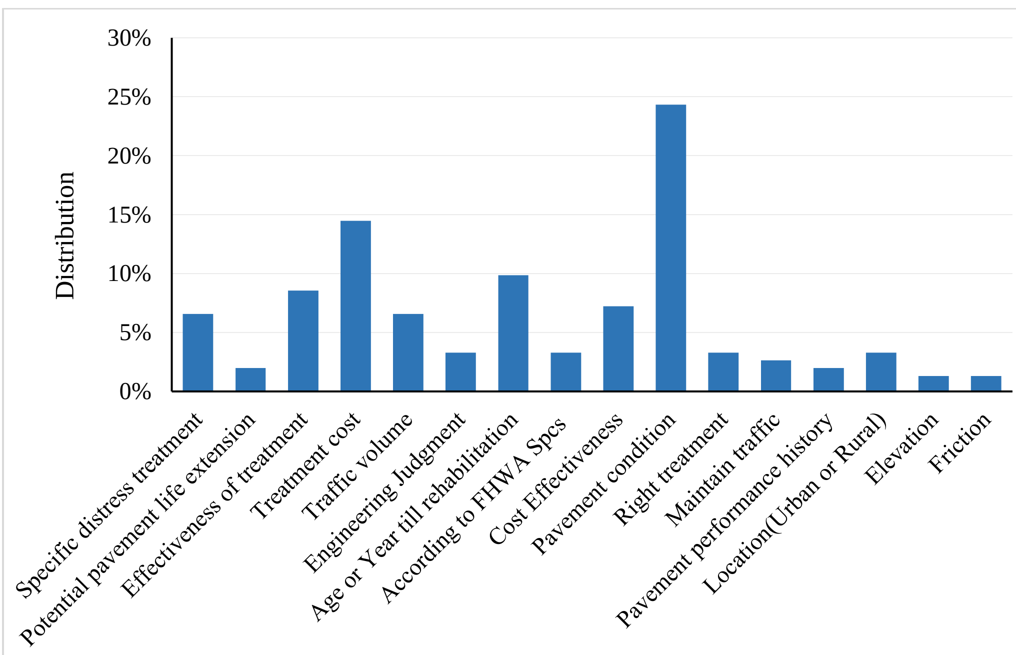



Figure 5.10. Criteria for scheduling PM treatments considered by the responded STAs

### 5.3.2. Formulation of Optimization Models

*Identification of model's objective functions and constraints:* the results of the survey indicate that the main objectives of the STAs are to improve pavement condition at the lowest cost of maintenance. This study identifies the objective functions of the optimization models as: (1)

minimizing the overall deterioration of road network in terms of IRI over an analysis period of 20 years [typical pavement design life (Morian et al. 2005; Ceylan et al. 2009; Santos and Ferreira 2013)], and (2) minimizing the overall maintenance costs of road network over the 20-year analysis period. As STAs suffer from the limited and shrinking funding for pavement preservation, this research considers the limited annual budget to be the first and main constraint for the optimization models. In addition, the current study focuses on the optimization of pavement PM only during a rehabilitation lifecycle (between rehabilitations or rehabilitation and re/construction). Therefore, the second constraint identified in this study is the threshold for the rehabilitation or reconstruction of pavement.

5.3.2.1. Model Variables and Parameters

*Decision Variables:* Treatment actions $x_{i,j}$; where $x_{i,j}$ is the PM treatment action for pavement section *i* at the time *j*. $x_{i,j}$ = {0, 1, 2, 3, …, *U*}; where 0 is do nothing action, and 1 is the first PM treatment candidate, whereas *U* is the PM treatment candidate number ***U*.**

*Other Variables and Parameters:*

- Pavement section of 1-mile long in road network *i*; where $i = \{1, 2, \ldots, I\}$; where *I* is the total number of 1-mile pavement sections, which is equal to 5,500 sections (the total number of the 1-mile pavement sections in the road network demonstrated by this study).
- Time *j*: is the time in years for PM treatment applications: $j = \{0, 1, 2, \ldots T\}$; where *T* is the analysis period, which is equal to 20 years [typical pavement design life (Morian et al. 2005; Ceylan et al. 2009; Santos and Ferreira 2013)].
- Condition of pavement section *i* at the scheduling time ($j = 0$): $\mathrm{Cond}_{i,0}$. The values of the $\mathrm{Cond}_{i,0}$ were extracted from the data simulated in the section 5.3.1.
- Cost of different PM treatments: $\mathrm{Cost}(x_{i,j})$. The cost of each treatment action was determined from the literature (Wang 2013; Roadresource 2019) in the 2019 dollars.
- Treatment effectiveness: $\mathrm{Eff}(x_{i,j})$. The performance jump (PJ) is the MOE that is most commonly used to measure the effectiveness of pavement treatments (Gao and Zhang 2008; Gao and Zhang 2011; Elhadidy et al. 2015), and thus was used in this study as the

$\text{Eff}(x_{i,j})$ of the different PM treatment actions. The $\text{Eff}(x_{i,j})$ for each value of $x$ was assumed to be constant for any pavement section *i* at any time *j*.

- Deterioration rate of pavement section *i* at time *j*: $\text{Deter}_{i,j}$. The deterioration rates are calculated in terms of IRI according to Chen and Zhang (2011) as $\text{IRI}_{i,j} = 50.43 \times EXP(0.0539 \times t_{i,j})$, in/mi.; where $t_{i,j}$ is equal to $\text{Age}_{i,0} + j$; where $\text{Age}_{i,0}$ is the age of pavement section *i* at the time of scheduling zero in years.
- Discount rate (d): an interest rate used to account for the value of money over time and is specified in this study as d = 2.25% (NHCCI 2019).
- Available Annual budget ($\text{B}_j$): the amount of funding available to STA at each year *j* to preserve its pavement network.
- Threshold for pavement rehabilitation ($\text{IRI}_{\text{max}}$): the maximum allowable IRI in in/mi. for pavement sections before the rehabilitation is due. According to the FHWA (2017) $\text{IRI}_{\text{max}}$ is set at 170 in/mi. in this study.

5.3.2.2. Mathematical Formulation of the Optimization Problem

Minimize: $\text{F}\ (x) = [\text{f}_1(x), \text{f}_2(x)]^{\text{T}}$ (vector of objective functions)

where $\text{f}_1(x)$ is the first objective function: minimizing the overall deterioration of pavement network in terms of IRI, in/mi.; whereas $\text{f}_2(x)$ is the second objective function: minimizing the total maintenance cost of pavement network.

$$f_1(x_{i,j}) = \sum_{i=1}^{I}\sum_{j=1}^{T} \text{Deter}_{i,j} - \sum_{i=1}^{I}\sum_{j=1}^{T} \text{Eff}(x_{i,j})$$

$$f_2(x_{i,j}) = \sum_{i=1}^{I}\sum_{j=1}^{T} \frac{1}{(1+d)^j} \times \text{Cost}(x_{i,j})$$

Subject to: $\text{C}\ (x) = [\text{C}_1(x), \text{C}_2(x)]^{\text{T}}$ (vector of constraints)

where $\text{C}_1(x)$ is the budget constraint function; and $\text{C}_2(x)$ is the constraint function of pavement condition threshold.

$$C_1(x_{i,j}) = \sum_{i=1}^{I} \frac{1}{(1+d)^j} \times \mathrm{Cost}(x_{i,j}) \quad \leq \mathrm{B}_j$$

$$C_2(x_{i,j}) = \mathrm{Cond}_{i,o} + \sum_{1}^{j} \mathrm{Deter}_{i,j} - \mathrm{Eff}(x_{i,j}) \leq \ \mathrm{IRI}_{max}$$

$\forall\ i$ from 1 to $I$ pavement sections, and

$j$ from 1 to 20-year analysis horizon

$g\ \ (x) = [\,],\ h\ (x)\ \ =\ [\,]$ (equality and inequality constraints)

$lb\ (x) = 0,\ ub\ (x)\ =\ U$ (lower and upper bounds)

5.3.2.3. Selection of Optimization Algorithm

According to the formulation of the optimization model, the objective functions are non-linear. In addition, the optimization problem includes combination of continuous and discrete variables. This optimization problem is complex, constrained (2 constraints), has non-linear objective functions, and has continuous variables and variables restricted to be integer-valued (treatment actions), which could lead to multiple local optimal solutions. Therefore, global optimal solution methods should be used.

Genetic Algorithm (GA) is a global optimal solution method based on natural selection - the process that drives biological evolution. GA can deal with problems with discrete and continuous variables. GA can handle complex, constrained and unconstrained, and non-linear problems (Fwa et al. 2000; Deshpande et al. 2010; Santos et al. 2019). The GA can be applied to solve a variety of optimization problems that are not well suited for standard optimization algorithms, including problems in which the objective function is discontinuous, non-differentiable, stochastic, or highly nonlinear. The GA can address problems of mixed integer programming, where some components are restricted to be integer valued. The GA differs from the classical derivative-based optimization algorithms (e.g., linear and mixed-integer linear programming) in two main ways. Classical algorithms generate single point at each iteration and choose the next point in the sequence by a deterministic computation, while GA repeatedly modifies a population of individual solutions generated in a number of steps (Plati et al. 2017; Sindi and Agbelie 2020). At each step, the GA randomly selects individuals from the current population to be parents and uses them to produce children for the next generation. Over successive generations, the population evolves toward an optimal solution.

Previous studies by Chan et al. (1994), Yang et al. (2013), Elhadidy et al. (2015), Santos et al. (2019) and Sindi and Agbelie (2020) used the GA to develop optimal pavement M&R strategies for highway network because the GA provides more robust and precise global optimal solutions compared to other classical derivative-based optimization algorithms. Sindi and Agbelie (2020) concluded that the GA is associated with high accurate global solutions for selecting road network M&R interventions and solving large optimization problems. Since the GA has been shown to effectively solve pavement maintenance optimization problems at network level, it was used to solve the current optimization problem.

The GA uses two main groups of processes (Figure 5.11) at each running step to reproduce next generation:

- Group 1: Selection rules select the individuals or parents that contribute to the population at the next generation.
- Group 2: Modification rules comprise two processes: (i) Crossover rules combine two parents to form children for the next generation, and (ii) Mutation rules apply random changes to individual parents to form children.

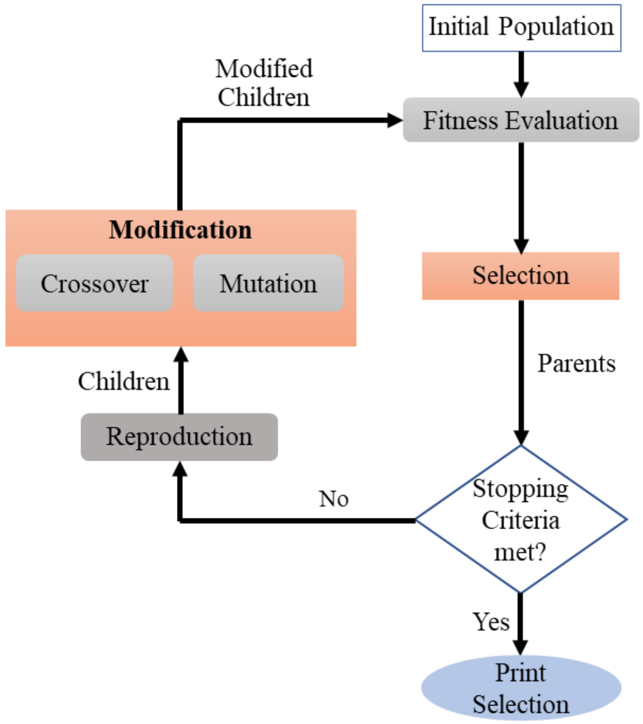



Figure 5.11. Typical GA sequence of procedures

GA algorithms have several estimation parameters: crossover fraction ($P_c$), mutation rate ($P_m$), population size, number of generations and stopping criteria. The values of the GA parameters are specific to each optimization problem and can be determined by trials (Reddy et al. 2004). Consequently, a sensitivity analysis was carried out for a range of values (from literature and typical practice) for each parameter, and the values that contribute to the best fitness function (the lowest objective function values in this study) were chosen.

The optimization problem of this study has two objective functions; therefore, the multi-objective optimization process is used to help search for a diverse set of solutions with these two objectives that can be optimized at one time. The Multi-Objective Genetic Algorithm (MOGA) "*gamultiobj*" provided by the MathWorks developer in the MATLAB software was used, which uses a controlled elitist GA that favors individuals with better fitness value and increases population diversity. Diversity required for model convergence is maintained by controlling the elite members of the population as the algorithm progresses. Two options, ParetoFraction and DistanceMeasureFcn, control the elitism. ParetoFraction limits the number of individuals on the Pareto front (elite members). The distance function, selected by DistanceMeasureFcn, helps to maintain diversity on a front by favoring individuals that are relatively far away on the front. The algorithm stops if the spread, a measure of the movement of the Pareto front, is small.

#### 5.3.2.4. Encoding of Solutions

The most frequently used GA encodings are binary and integer representations. Binary encoding uses the ‘0’ or ‘1’ binary digit. The binary strings of 0s and 1s represent the genes that are concatenated to form the full chromosome (solution or decision variable values). Binary strings are useful because the GA operations with them are easier to explain. Whereas, integer encoding represents the decision variable in its actual values [treatment actions (0-$U$)]. Figure 5.12 displays the binary and integer representations for the current optimization solutions.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0000 | 0001 | 0000 | 0101 | 0001 | 0000 | 0001 | 1001 | 0000 | 0011 | ------------- | 0001 |

(a)

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 0 | 5 | 1 | 0 | 1 | 9 | 0 | 3 | ------------- | 1 |

(b)

Figure 5.12. Encoding of an arbitrary solution for one pavement section over 20 years: (a) Binary coding; and (b) Integer coding

The effectiveness of the GA when used in real-world problems depends heavily on choosing the appropriate encoding representation (Santos et al. 2019). Although the binary representation is simpler than the integer, significant computational efforts are needed to convert integer values (treatment actions) to binary values. The integer encoding was therefore used to represent the solutions of the current stochastic MOGA model (treatment actions of every pavement section in the network over 20 years). Each possible MOGA solution is represented by an $I \times T$ matrix (gene); where $I$ is the number of pavement sections in a road network; $T$ is the analysis horizon (20 years in the current optimization problem); and the allele values of each gene are a combination of integer values (treatment actions range from 0 to $U$).

### 5.3.3. Development of Stochastic Optimization Models

The stochastic optimization models are developed using the selected optimization algorithm, i.e., MOGA. One major challenge commonly encountered in MOGA optimization models is the corresponding high computational expenses or slow convergence (Augeri et al. 2019). This is mainly because of the random selection of generations and the number of iterations that GA runs until optimal solutions are reached. When "optimal" is stated hereinafter it indicates the best solution found by the MOGA, which may not be globally optimal. In the current MOGA model, the size of the problem is expressed by the number of combinations in terms of number of pavement sections (5,500), analysis period (20 years) and the number of treatment actions [1 action (do nothing) + 9 actions (PM treatments identified by the STAs in response to the survey)]. The number of combinations in this case would be (10 actions)$^{(5{,}500 \text{ sections} \times 20 \text{ years})}$, which is equal to $10^{110{,}000}$. Moreover, when the budget constraint and pavement condition deterioration and improvement are defined in probabilistic terms, the MOGA algorithm requires large number of iterations to reach an optimal solution. A large number of iterations (e.g., 1,000 runs) is needed to sufficiently capture the probability distribution of the input variables (budget constraint and pavement condition deterioration and improvement) and build the probability distribution of the output (optimal solutions).

One solution to reach a feasible number of combinations for the proposed stochastic MOGA model is to reduce the search space for the MOGA algorithm. This study proposes three approaches to reduce the MOGA search space: (1) identifying and considering the most commonly used PM treatments; (2) clustering pavement sections; and (3) application of rest period.

(1) Most commonly used PM treatments

In order to reduce the search space for the stochastic MOGA model, the results of the survey (discussed earlier in section 5.3.1) were investigated in order to select the most commonly used PM treatments identified by the STAs respondents. From the survey results, the most commonly used PM treatments are crack sealing, Thin HMA overlay, Micro-surfacing and UTBWC. Crack sealing does not improve pavement condition (roughness, IRI) (Lee and Shields (2010), thus only Micro-Surfacing, UTBWC and Thin HMA overlay are considered in this study. Hence, four treatment actions were considered in this study as follows: 0 (do nothing), 1 (Micro-surfacing), 2 (UTBWC) and 3 (Thin HMA overlay). This approach has reduced the number of treatment actions from 10 to 4 by considering only the most widely used PM treatments, which in turn reduces the stochastic MOGA search space.

(2) Clustering of pavement sections

Pavement family/group concept has typically been used for clustering pavement sections for pavement performance and maintenance optimization models. Pavement families may be defined by the pavement type (e.g., asphalt or concrete), pavement use (e.g., roadways or parking), functional classification (e.g., interstates or locals), and/or any other similar characteristics. This study implements the proposed stochastic MOGA for the interstate flexible pavements; however, the number of interstate flexible pavement sections is 5,500 1-mile long (as identified earlier in section 5.3.1). Thus, the pavement sections are further grouped in order to reduce the search space for the proposed stochastic MOGA model. One way of clustering the 5,500 pavement sections is by a common characteristic among these sections.

This study clusters pavement sections into groups with a similar age range in years. Previous research grouped pavement sections into classes of age ranges of 5 years (Ahmed et al. 2015) and 3 years (Nunez and Shahin 1986; Khattak et al. 2013). Since the data collected for this study include pavement sections ranging from 0 to 27 years of age, an interval of 4 years was identified as the age range for pavement groups. Thus, the pavement sections (5,500) were grouped according to their ages into 7 groups: (1) from 0 to 3; (2) from 4 to 7; (3) from 8 to 11; (4) from 12 to 15; (5) from 16 to 19; (6) from 20 to 23; and (7) from 24 to 27 years. The number of pavement sections in each group at the scheduling time ($j = 0$) is 1,107, 1,345, 1,286, 686, 555, 384 and 137 for groups from 1to 7, respectively. This approach has reduced the number of pavement sections from 5,500 individuals to 7 groups (each with a 4-year range), which in turn reduces the stochastic MOGA search space.

It is worth noting that the pavement sections of the same age range may not have the same condition. In addition, not all pavement sections that are in the same age range (even if they are of the same condition) technically receive PM treatments at the same time. This study therefore adds the decision variable "group percentage" (percentage of each pavement group that receives treatment actions 0, 1, 2, or 3) to the optimization problem. This percentage can take on any value from 0 to 100 percent as determined by the MOGA algorithm.

(3) Application of Rest Period

To further reduce the search space for the proposed stochastic MOGA model, this study suggested applying a rest period to the search space. The rest period is defined as the minimum time in years between two PM treatments. The purpose of applying the rest period concept is to refrain the MOGA algorithm from searching for a solution over a specific period starting from the time of a previously selected solution. In other words, if the MOGA optimization algorithm chooses to apply Micro-surfacing for a particular pavement section at year *j*, no treatments should be applied to this pavement section at any time during a specified rest period after year *j*. Figure 5.13 shows a schematic diagram for the concept of rest period.

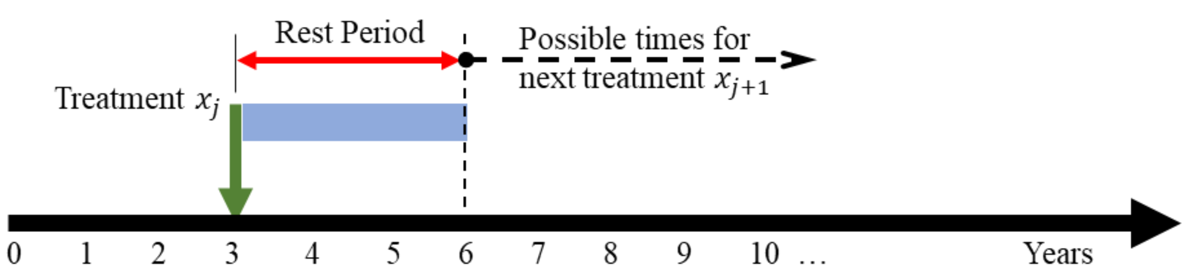



Figure 5.13. Rest period and possible times for treatments

The length of the rest period can be either the service life or the warranty period of each PM treatment. This study considers the rest period to be the warranty period of each PM treatment. According to the specifications of pavement maintenance provided by the FHWA or STAs such as INDOT, contractors are responsible for the warranted preventive maintenance treatments. In this study, Micro-surfacing, UTBWC and Thin HMA overlay are warranted PM treatments (FHWA 2017; INDOT 2020). The warranty period for each of the three PM treatments is 3 years from the date of the substantial completion of the treatment application (FHWA 2017; INDOT 2020). The rest period was therefore specified as 3 years for each of the three PM treatments. To apply the rest period to the MOGA algorithm, a filtering constraint was created to force the algorithm to assign a value of zero (Do nothing action) during the rest period. The significance of applying the rest period constraint is in twofold: (1) reducing the search space for the stochastic MOGA algorithm, and (2) avoiding the erroneous and impracticable results of the MOGA model that might suggest an optimal solution during the warranty period.

## 5.4. Results and Discussion

This section presents the results and discussions of the stochastic MOGA optimization models developed following the implementation of the proposed research framework. It presents the estimation of the MOGA parameters, the results of the stochastic MOGA optimization models, and the performance curves of pavement groups over the analysis period.

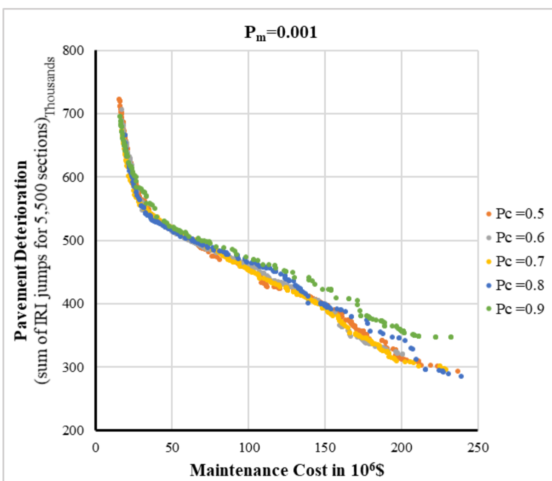
### 5.4.1. MOGA Parameters

The values of the MOGA parameters are specific to optimization problem. Therefore, a sensitivity analysis was carried out for a range of values (from literature and/or as a typical practice) for each parameter (Table 5.4), and the values that contribute to the best fitness function (in this study, the lowest values for objective functions) were chosen.

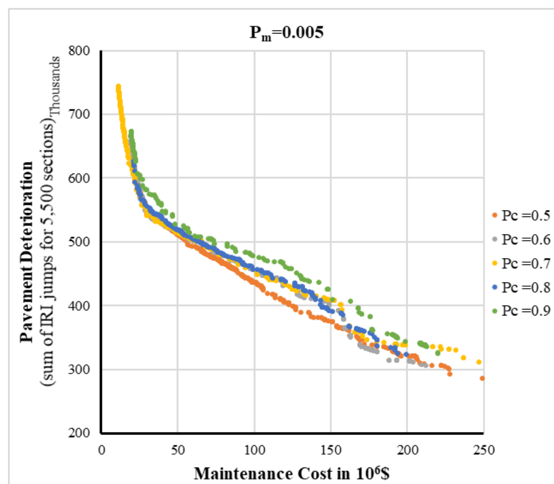
Table 5.4. MOGA parameters values from literature and common practice, and specified ranges for the sensitivity analysis

| **MOGA Parameter** | **Previous Studies** | **Common Practice** | **Ranges for Sensitivity Analysis** |
|---|---|---|---|
| Crossover Fraction | 0.6 (Elhadidy et al. 2015) | 0.8 (MATLAB) | 0.6 – 0.9 |
| Mutation Rate | 0.01 (Elhadidy et al. 2015) | 0.05 (MATLAB) | 0.001 – 0.1 |
| Population Size | 10 times the number of decision variables: 2800 (Storn 1996; Chen et al. 2015) | 50 for a number of decision variables ≤ 5, or 200 otherwise (MATLAB) | 1,000 – 5,000 |
| Number of Generations | - | 100 (MATLAB) | 0 - 250 |

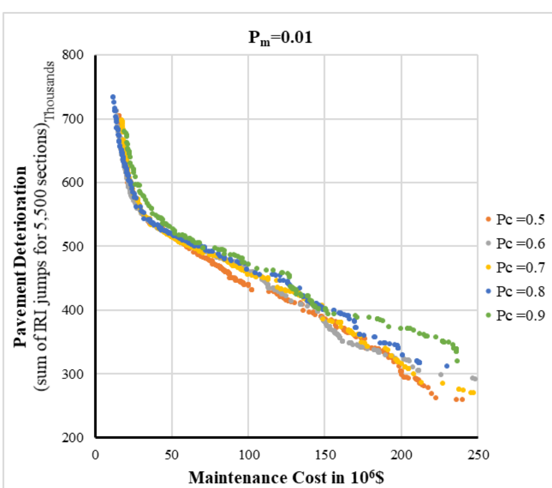
***Selection of Crossover Fraction:*** The default value for the crossover fraction ($P_c$) in the MOGA algorithm provided in the MATLAB is 0.8. However, to ensure that the selected crossover fraction yields the best values for the objective functions, a sensitivity analysis was performed for the $P_c$ values ranging from 0.5 to 0.9 for each mutation rate ($P_m$) (ranging from 0.001 to 0.1), population size equals to 3000 and number of generations equals to 50. Figures 5.14(a-e) present the Pareto frontiers for the different $P_c$ values at the $P_m$ values 0.001, 0.005, 0.01, 0.05 and 0.1, respectively. The Pareto frontiers show that the crossover fractions of 0.5 and 0.6 values yield comparable best results (lowest values for both objective functions) for all $P_m$ values. However, the value of 0.6 was selected in order to ensure an enough crossover of the population; the same value ($P_c = 0.6$) was adopted by Elhadidy et al. (2015).



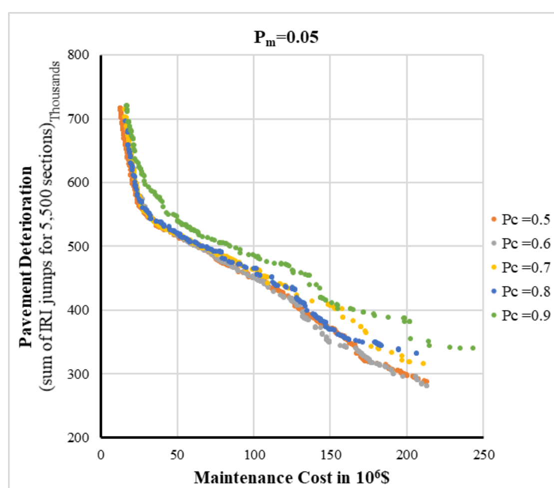

(a)



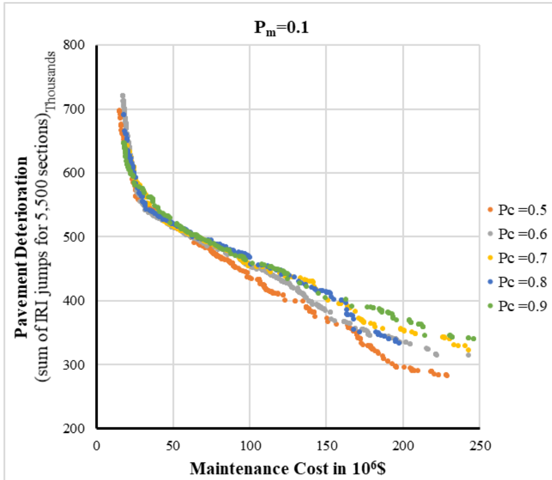

(b)




(c)




(d)




(e)

Figure 5.14. Pareto frontiers for the objective functions for the $P_c$ values from 0.5 to 0.9 and $P_m$ values: (a) $P_m$ =0.001; (b) $P_m$ =0.005; (c) $P_m$ =0.01; (d) $P_m$ =0.05; and (e) $P_m$ =0.1

***Selection of Mutation Function and Rate:*** The Adaptive Feasible function "mutationadaptfeasible" was used as the mutation function as it is typically used when a MOGA problem includes constraints and to randomly generate adaptive directions for the last successful or unsuccessful generation. The default value for the mutation rate ($P_m$) in the MOGA algorithm provided in the MATLAB is 0.05; however, to ensure that the used mutation rate yields the best values for the objective functions, a sensitivity analysis was carried out for the $P_m$ values ranging from 0.001 to 0.1, and the best value was selected. Figure 5.15 shows the results of using the abovementioned values of the $P_m$ range at the selected $P_c$ value (0.6). It can be noticed that similar optimal Pareto frontiers were obtained for all $P_m$ values and the $P_c$ value of 0.6. It is worth mentioning that low mutation rates reduce the possibility of exploring new solutions, whereas high mutation rates increase the MOGA computational time (i.e., slow MOGA convergence) (Morcous and Lounis 2005). The mutation rate value was therefore chosen to be 0.01 as used in the MOGA model by Elhadidy et al. (2015).

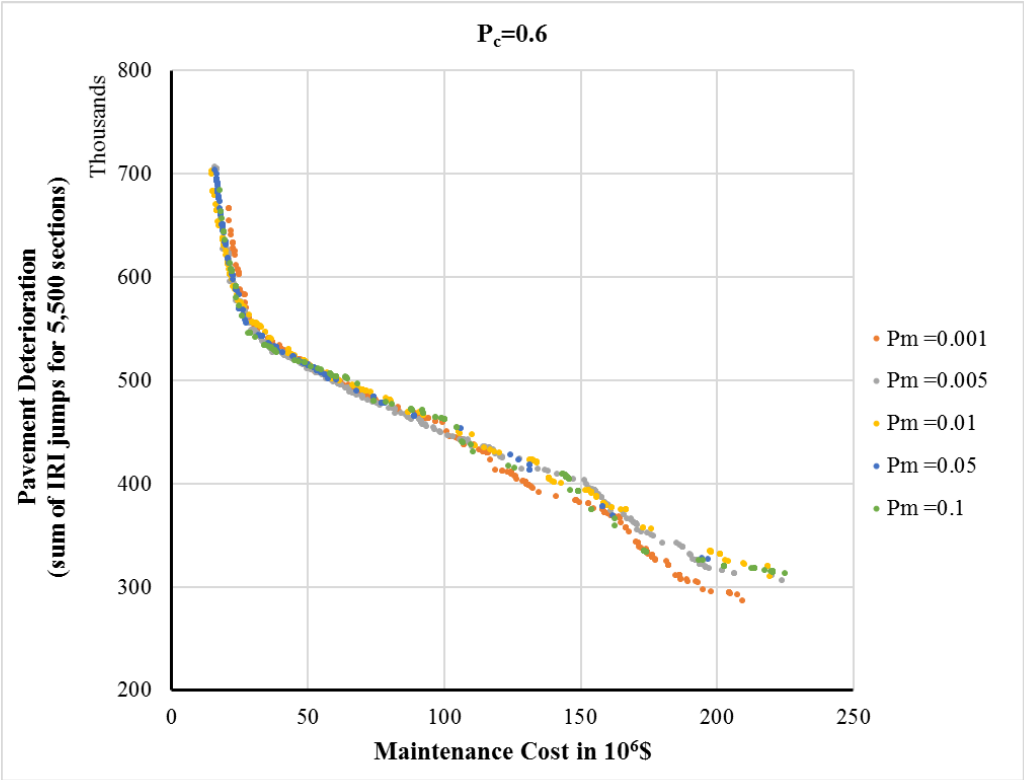



Figure 5.15. Pareto frontiers for the objective functions for the $P_m$ values from 0.001 to 0.1 and $P_c$ value of 0.6

***Selection of Population Size and Number of Generations:*** The MOGA algorithm provided in the MATLAB sets the default value for the population size as 50 for a number of decision variables less than or equal to 5, or 200 otherwise. Based on the guidelines from prior research (Storn 1996; Chen et al. 2015), the population size of the evolutionary algorithms, including MOGA, can be specified as ten times the search space dimensionality (i.e., number of decision variables). In addition, the population size of MOGA model can be determined based on trials (Reddy et al. 2004). Thus, a sensitivity analysis was conducted to select the population size that yields the shortest Euclidean distance (the lowest values for the two objective functions in the current study). The population size was specified to take on the range of values around 10 times the number of decision variables ($10 \times 280 = 2{,}800$). Therefore, the population size was specified to range from 1,000 to 5,000. Figure 5.16 shows the calculated Euclidean distance (Equation 5.1) of the two objective functions when the population size ranges from 1,000 to 5,000 at the $P_c$ of 0.6, $P_m$ of 0.01 and number of generations equal to 50. Figure 5.16 indicates that at the population size of

4,000, the Euclidean distance is the shortest (the lowest values for the two objective functions), and therefore the population size was specified as 4,000 for the current MOGA problem.

$$\text{Euclidean distance} = \sqrt{(X - X_{min})^2 + (Y - Y_{min})^2} \qquad (5.1)$$

where X is the value of the first objective; while Y is the value of the second objective.

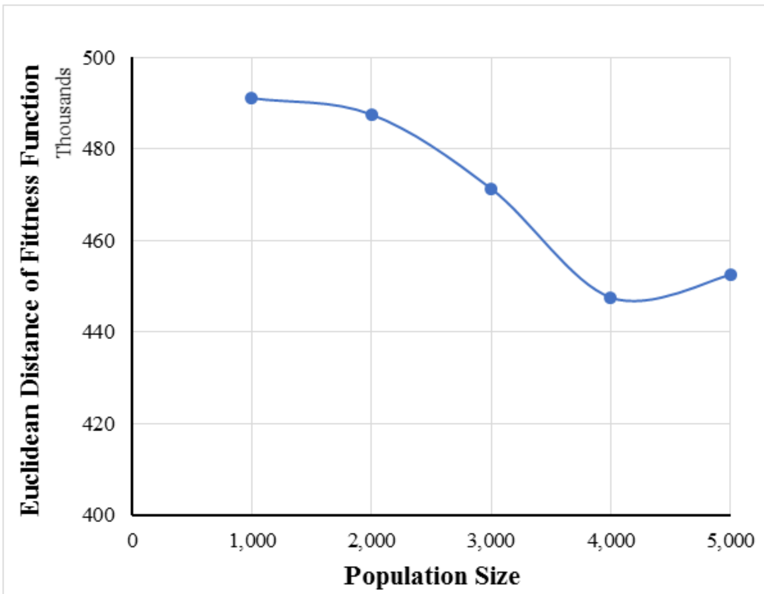



Figure 5.16. Population size versus Euclidean distance of the two objective functions

The default value for the number of generations in the MOGA provided in the MATLAB is 100; however, a sensitivity analysis was performed to ascertain that the number of generations used in the current MOGA problem yields the shortest Euclidean distance of the two objective functions. The number of generations was identified to take on the values from 0 to 250 at the selected values of the other MOGA parameters: $P_c$ = 0.6, $P_m$ = 0.01 and population size =4,000. Figure 5.17 displays the number of generations versus the calculated Euclidean distance of both objective functions (Equation 5.1). The number of generations was therefore selected to be equal to 100 because this value corresponds to the shortest Euclidean distance.

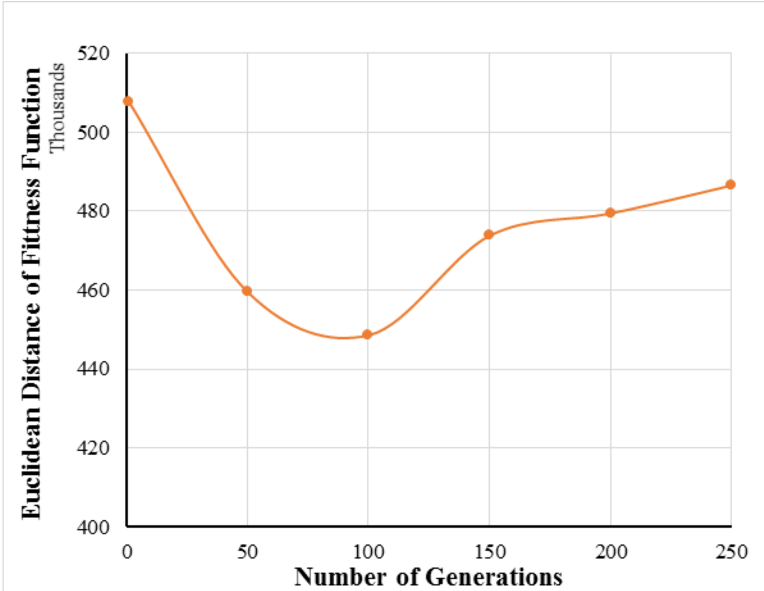



Figure 5.17. Number of generations versus Euclidean distance of the two objective functions

***MOGA Stopping Criteria and Convergence:*** The stopping criterion of MOGA can be determined on the basis of the convergence of any of the following measures: (i) the objective-function value of the best genotype in each generation; (ii) the average objective-function value of parent genotypes in each generation; and (iii) the average objective-function value of offspring genotypes in each generation (Fwa et al. 1996 and 1997). The MOGA stopping criterion can be established based on a certain number of generations (Fwa et al. 1996; 1997). The stopping criterion of number of generations was set to be 50 by Li and Wang (2019), 100 by Fwa et al. (1996) or 150 by Fwa et al. (1997). The terminating criterion of number of generations for the current MOGA models was specified to be equal to 100, at which the shortest Euclidean distance was obtained (Figure 5.17).

### 5.4.2. Stochastic MOGA Optimization Models

In this study, stochastic MOGA optimization models were developed taking into account the uncertainty of the variables: budget constraint, pavement deterioration and improvement in pavement condition after the application of PM treatments. These variables were assumed to follow the normal distribution, which commonly used to represent uncertainties in engineering problems. The mean and standard deviation values of the budget constraint are equal to \$9M and \$225,000, respectively, while the mean and standard deviation values for pavement condition deterioration and improvement (drop in IRI) are shown in Table 5.5.

Table 5.5. Means and standard deviations of pavement condition deterioration and improvement

| **Deterioration (in/mi.)** | | | **Improvement (in/mi.)** | | |
|---|---|---|---|---|---|
| **Pavement Group** | **Mean** | **STD** | **Treatment Action** | **Mean** | **STD** |
| 1 | Varies over time according to Chen and Zhang (2011): $\text{IRI} = 50.43 \times EXP(0.0539 \times t)$ | 1.0 | Do Nothing | 0 | 0.0 |
| 2 | | 1.5 | Micro-surfacing | 10 | 1.25 |
| 3 | | 2.0 | UTBWC | 15 | 2.0 |
| 4 | | 2.5 | Thin HMA Overlay | 20 | 2.5 |
| 5 | | 3.0 | | | |
| 6 | | 3.5 | | | |
| 7 | | 4.0 | | | |

1. Stochastic MOGA Considering the Uncertainty of Budget Constraint

A stochastic MOGA model was developed to account for the uncertainty of budget constraint only. Figure 5.18 shows the Pareto frontiers with probabilities of 5%, 50% and 95% of the budget constraint less than or equal to the values: *mean - 2STD*, *mean* and *mean + 2STD*, respectively. Figure 5.18 indicates that there is an insignificant difference between the Pareto frontiers with the probabilities of 5% or 95% and the probability of 50% (equivalent to the deterministic MOGA), which implies that the influence of the uncertainty of budget constraint on the optimal solutions is not significant. This is because the specified budget variation of $225,000 may be too small to affect the variation of the optimal solutions. However, it can be concluded that if the variation in the budget constraint is expected to be within the value specified herein (5% of the available fund or $225,000), there will be an insignificant difference whether a stochastic or deterministic (with a 50% probability) MOGA model is being developed.

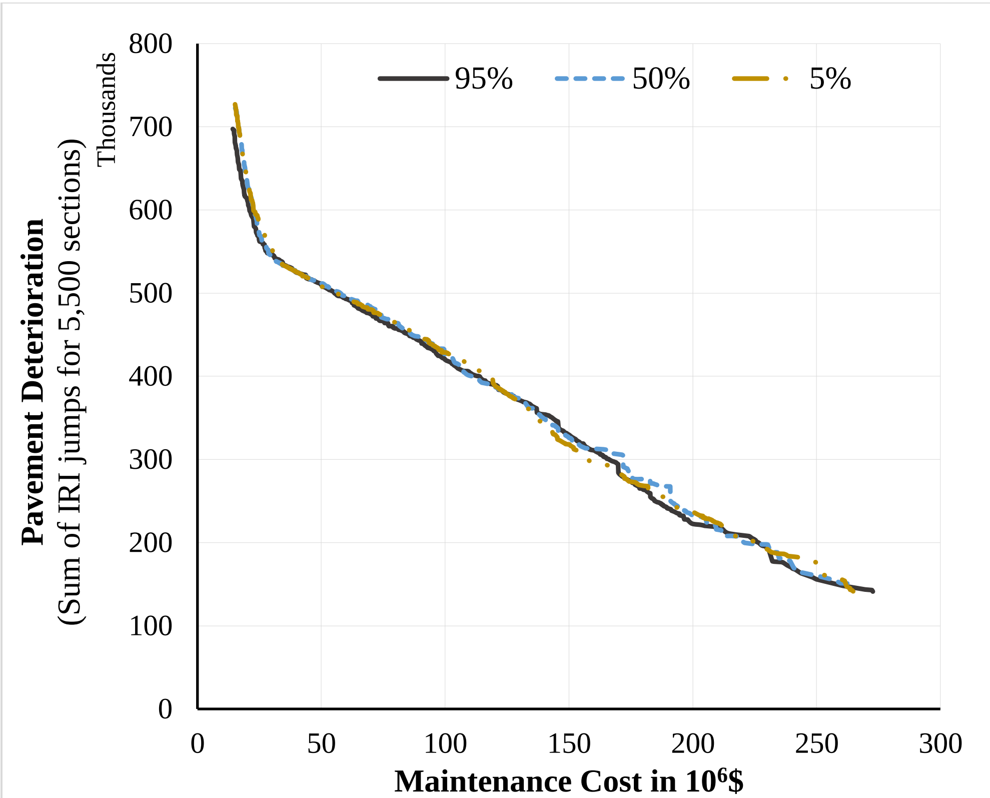



Figure 5.18. Pareto frontiers for the stochastic optimal solutions considering the uncrtainty of budget constraint

2. Stochastic MOGA Considering the Uncertainty of Pavement Condition Deterioration and Improvement and Budget Constraint

In order to estimate the impact of the uncertainty of pavement condition deterioration and improvement on the PM scheduling for pavement network, the stochastic MOGA model is run using the means and standard deviations set out in Table 5.5 for 1,000 runs. Due to the large number of the decision variables (280) and the consideration of probability distributions, the MOGA model is expected to be expensive in computation. As such, a number of combinations/scenarios for the change in pavement condition deterioration and improvement were created as shown in Table 5.6. The change in each variable was assumed to be equal to $\pm 2STD$ corresponding to the 95% and 5% probabilities. The deterioration of pavement condition is expected to change by $+2STD$ and $-2STD$ from the expected mean value with probabilities of 95% and 5%, respectively. Whereas, the improvement in pavement condition after treatment implementation is expected to change by $-2STD$ and $+2STD$ from the expected mean value with probabilities of 95% and 5%, respectively.

Table 5.6. Combinations/scenarios of pavement condition deterioration and improvement

| | | Expected Condition Improvement | |
|---|---|---|---|
| | | Increase (5%) | Decrease (95%) |
| **Expected Condition Deterioration** | Increase (95%) | II | ID |
| | Decrease (5%) | DI | DD |

Figure 5.19 shows Pareto frontiers for the optimal solutions of the stochastic MOGA model when the uncertainty of budget constraint and pavement condition deterioration and improvement are considered. It shows the Pareto frontiers for each scenario presented in Table 5.6 and the 50% probability (mean values). All of the Pareto frontier curves of the four scenarios (representing the stochastic MOGA) have trends similar to that of the 50% probability (deterministic MOGA). Nevertheless, Figure 5.19 indicates a significant difference in the values of the objective functions of the stochastic MOGA from that of the deterministic MOGA, which, in turn, underscores the need to account for the uncertainty associated with pavement condition deterioration and improvement in pavement maintenance optimization models. The stochastic MOGA models provide optimal solutions that indicate at each optimal maintenance cost there are multiple expected pavement deterioration values. This highlights the importance of considering the uncertainty of pavement condition in the MOGA pavement maintenance optimization models.

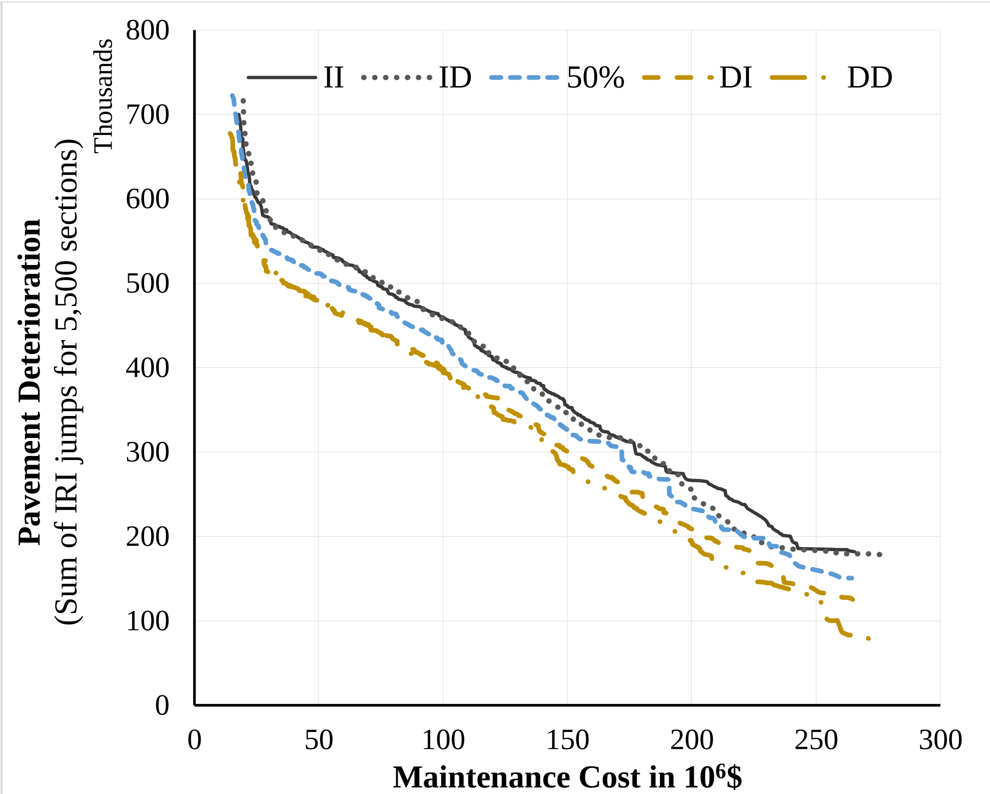



Figure 5.19. Pareto frontiers for the stochastic optimal solutions considering the uncrtainty of condition deterioraiton and improvement and the uncertainty of budget constraint

If the decision-maker chooses to adopt a solution at a minimum maintenance cost of, for example, $120M, this solution corresponds to an expected pavement network deterioration of 387,000 in/mi. with a 50% probability. The $120 million is approximately equal to $120/20 = $6 million per year for 20 years. While, a 387,000 in/mi. deterioration is approximately equal to 387,000/5,500 = 70 in/mi., an average deterioration for each pavement section (assuming that all pavement sections deteriorate at similar levels). Table 5.7. shows the PM schedule (treatment actions and the corresponding percentage of each pavement group receiving treatment actions) for the 7 pavement groups over an analysis period of 20 years. For instance, for pavement group 4, five treatment actions are scheduled over the 20-year analysis period: three thin HMA overlay applications – one each at Year 1 (20% of pavements in group 4), 5 (50% of pavements in group 4) and 17 (40% of pavements in group 4); one Micro-surfacing application at Year 9 for 50% of pavements in group 4; and one UTBWC application at Year 13 for 20% of pavements in group 4.

Table 5.7 (a). Preventive maintenance for pavement groups over 20-year analysis period and the respective percentage of pavement groups that receive PM treatments

| Year / **Pavement Group** (age range) | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** (00-03yrs) | 3 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| **2** (04-07yrs) | 2 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 3 | 0 |
| **3** (08-11yrs) | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 |
| 4 (12-15yrs) | 3 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 3 | 0 | 0 | 0 |
| **5** (16-19yrs) | 1 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 3 | 0 | 0 | 0 |
| **6** (20-23yrs) | 2 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 3 | 0 | 0 | 0 |
| **7** (24-27yrs) | 1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |

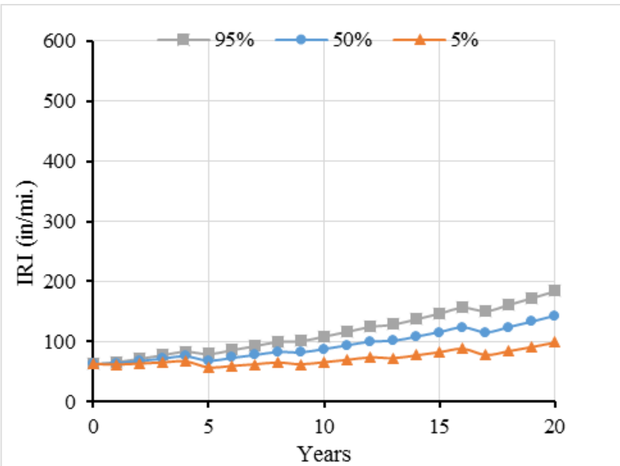
*Note: 0: Do nothing, 1= Micro-surfacing, 2= UTWBC, and 3= thin HMA overlay*

Table 5.7 (b). Preventive maintenance for pavement groups over 20-year analysis period and the respective percentage of pavement groups that receive PM treatments

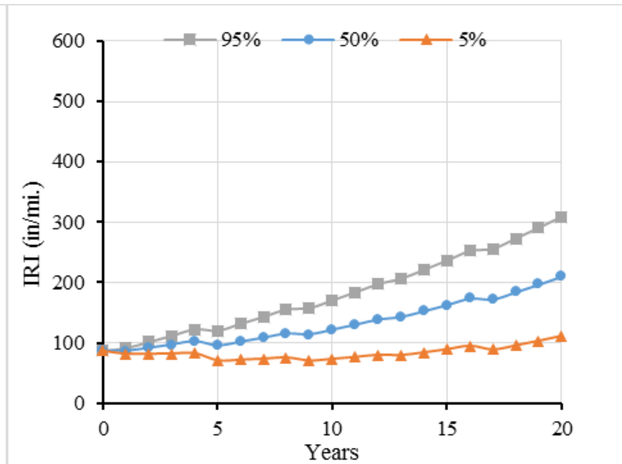
| Year / **Group** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 0.1 | 0 | 0 | 0 | 0.6 | 0 | 0 | 0 | 0.4 | 0 | 0 | 0 | 0.2 | 0 | 0 | 0 | 0.9 | 0 | 0 | 0 |
| **2** | 0.3 | 0 | 0 | 0 | 0.3 | 0 | 0 | 0 | 0.2 | 0 | 0 | 0 | 0 | 0 | 0.2 | 0 | 0 | 0 | 0.4 | 0 |
| **3** | 0.2 | 0 | 0 | 0.4 | 0 | 0 | 0 | 0.8 | 0 | 0 | 0 | 0.3 | 0 | 0 | 0 | 0.2 | 0 | 0 | 0 | 0.2 |
| **4** | 0.2 | 0 | 0 | 0 | 0.5 | 0 | 0 | 0 | 0.5 | 0 | 0 | 0 | 0.2 | 0 | 0 | 0 | 0.4 | 0 | 0 | 0 |
| **5** | 0.5 | 0 | 0 | 0 | 0.3 | 0 | 0 | 0 | 0.3 | 0 | 0 | 0 | 0.6 | 0 | 0 | 0 | 0.2 | 0 | 0 | 0 |
| **6** | 0.6 | 0 | 0 | 0 | 0.5 | 0 | 0 | 0 | 0.5 | 0 | 0 | 0 | 0.7 | 0 | 0 | 0 | 0.3 | 0 | 0 | 0 |
| **7** | 0.4 | 0 | 0 | 0 | 0.3 | 0 | 0 | 0 | 0.2 | 0 | 0 | 0 | 0.3 | 0 | 0 | 0 | 0.2 | 0 | 0 | 0 |

### 5.4.3. Pavement Performance Curves Using Stochastic MOGA Models

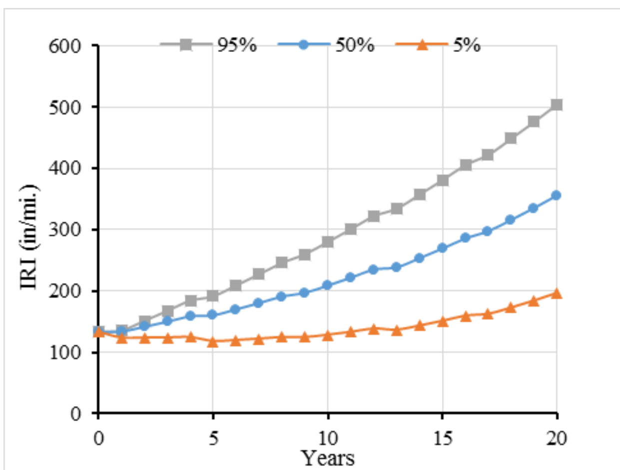
The optimal solution chosen and provided in Table 5.7 is further used to investigate the impact of the uncertainty of pavement condition deterioration and improvement on pavement performance. Figures 5.20(a, b and c) show the performance curves of the pavement groups: 1 (1107 pavement sections), 4 (686 pavement sections) and 7 (137 pavement sections), respectively, as examples of the early (0 - 3 years), middle (12 -15 years) and late (24 – 27 years) age sets of pavement sections. For each set of ages, the performance curves were developed at 95%, 50% and 5% probabilities. Figures 5.20(a, b and c) imply that there is a significant difference between pavement performance with a 50% probability (deterministic MOGA) and 95% and 5% probabilities (stochastic MOGA). Moreover, the variation in the expected pavement performance in terms of IRI values for early age pavements (group 1) is lower than that for middle age pavements (group 4), which, in turn, have lower pavement performance variations than for late age pavements (group 7).




(a) Group #1: 0-3 years




(b) Group #4: 12-15 years




(c) Group #7: 24-27 years

Figure 5.20. Performance curves with 95%, 50% and 5% probabilities for pavement groups: (a) Group #1; (b) Group #4; and (c) Group #7

## 5.5. Summary

Prior stochastic pavement maintenance optimization models considered the uncertainty of budget constraint but due to the associated additional computational complexity in terms of the number of combinations did not adequately account for the uncertainty of pavement condition deterioration and improvement. The selection of optimal timings and types of maintenance treatments for a pavement network over the long-term (pavement design life) without considering the uncertainty of expected pavement condition and maintenance effectiveness has the potential to lead to mistiming of maintenance applications and can therefore result in less optimal alternatives. This chapter discussed the development of stochastic pavement maintenance optimization models with the consideration of budgetary uncertainty in addition to the uncertainty of expected pavement condition deterioration and improvement. The Multi-Objective Genetic Algorithm (MOGA) method was used to obtain the optimal or near-optimal global solutions for the two objective functions: minimum total preventive maintenance costs and minimum overall pavement network deterioration. The stochastic MOGA models were designed with two constraints: a limited budget and a pavement condition threshold. The results showed that deterministic MOGA models provide one PM schedule for each expected total maintenance cost. Whereas, stochastic MOGA models offer multiple PM schedules for each expected total maintenance cost, each PM schedule is associated with the probability of a corresponding pavement network deterioration. At a specific maintenance cost for pavement network using deterministic MOGA model, pavement will have one expected value of condition at each year. The developed stochastic MOGA models provide decision-makers with multiple optimal maintenance schedules that may be appropriate to their level of risk/uncertainty.

# CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS

Highway infrastructure, including roads/pavements contributes significantly to a country's economic growth, life quality improvement and negative environmental impacts. Highway agencies therefore aim to manage their infrastructure assets efficiently to keep them in good condition while making optimal use of their limited and dwindling resources. This dissertation aimed to bridge the gaps in the body of knowledge and practice regarding pavement management systems, particularly pavement performance modeling and maintenance optimization. The first section of this chapter provides an overview and summary of the research, while the second section provides a summary of the findings of the research. The third section of this chapter discusses the significance of this research to the body of knowledge and practice. The fourth section underlines the limitations of this study, whilst the fifth section proposes recommendations for future research.

## 6.1. Summary of the Research

The overarching goal of this dissertation is to enhance pavement management systems by contributing to pavement performance modeling and pavement maintenance optimization. The research conducted in this dissertation was demonstrated within the context of interstate flexible/Asphalt pavements. Pavement condition data were collected from the LTPP database for interstate flexible pavements from eight Midwestern states in the U.S. Also, three questionnaire surveys were developed and deployed to subject matter experts (SMEs) from the U.S. State Transportation Agencies (STAs), with background and experience in the field of pavement engineering and management.

The first component of this dissertation presented a state-of-the-art review for the probabilistic modeling of pavement performance using Markov chains. Based on the reviewed literature, Markov chain techniques were categorized as follows: homogeneous, non-homogeneous, staged-homogeneous, hidden, and semi-Markov. Various methods used in the literature to estimate the transition probability matrix (TPM) of pavement condition were synthesized and discussed. The TPM estimation methods include the expected-value, percentage transition, simulation-based methods, and econometric and duration models. Furthermore, based on guidance and insights acquired from prior research, this doctoral research developed a decision tree for selecting the appropriate Markov technique and TPM estimation method for probabilistic modeling of pavement performance.

The second component of this dissertation introduced a hybrid approach to incorporate the effectiveness of preventive maintenance (PM) into probabilistic pavement performance models in the absence or insufficiency of historical PM data. The approach consists of six major tasks: data collection and analysis; data simulation; estimation of initial times for preventive maintenance treatments; data generation; estimation of transition probability matrix; and validation of the approach and developed models. Pavement condition data were collected from the LTPP database for interstate flexible pavements from eight Midwestern states. Data were simulated by developing an exponential multiple regression model and determining the probability distribution of dependent and independent variables. The initial times for PM treatments were determined through literature search, a survey (Survey 1) of SMEs to gather information on pavement condition and PM, and detection of PM times from probabilistic pavement performance curves.

An ordered-probit Model A was built to develop probabilistic pavement performance curves that were utilized to identify the approximate probable times for PM applications, and to estimate the effectiveness of PM treatments. Since the implementation of PM is restricted by limited funding, a greedy algorithm was developed to prioritize PM schedule based on pavement condition and treatment costs, and under the constraint of the estimated initial PM times. The amount of funding was specified as a percentage (100% to 40%) of the total funding needed to implement PM for all roadway sections of the road network at the estimated initial times. For each percentage of funding, the greedy algorithm resulted in pavement condition data that were used to develop ordered-probit Models B. Probabilistic pavement performance curves were created using the developed Models B and compared with that of Model A. An ordered-probit Model C was developed to estimate the non-homogeneous transition probabilities of pavement condition incorporating PM impact. The statistical significance of PM in pavement performance prediction found in Model C demonstrates the necessity of collecting and managing PM data. The probabilistic pavement performance curves with PM effects were developed, and the marginal effects of the explanatory variables were estimated.

The hybrid approach and the developed non-homogeneous Markov model were validated mathematically through cross-validation with the actual/out-of-sample data and practically using two surveys (Surveys 2 and 3) sent to the SMEs in pavement engineering and management. The cross-validation was performed to ensure the predicted pavement condition is comparable to the actual pavement condition. Survey 2 was deployed to the SMEs to assess the trends in pavement performance curves developed using current and prior research, while Survey 3 was used to assess the estimated marginal effects of the explanatory variables.

The third component of this dissertation addressed the statistical significance of the design parameters of Markovian models for the pavement condition prediction accuracy. These design parameters are the number of condition states (NCS), the length of duty cycle (LDC), the data collection time at pavement cohorts (P1, P2 and P3) in the homogeneous Markov models, and the stage length in the staged-homogeneous Markov models. A comparative analysis was carried out for the various Markovian techniques, and for each technique, the different combinations of Markov design parameters. The results of the comparative analysis provide guidance to future researchers and highway agencies to determine whether the NCS and LDC used for their Markov models are associated with the most accurate Markovian pavement performance models. Moreover, when designing homogeneous and staged-homogeneous Markov models, the data collection time and the stage length, respectively, should be given greater attention.

The fourth component of this dissertation discussed the development and significance of stochastic optimization models for the network-level scheduling of pavement preventive maintenance. Multi-objective Genetic Algorithm (MOGA) optimization approach was used to obtain global optimal solutions while accounting for the two main objectives of effective maintenance strategies: minimum maintenance cost and maximum pavement performance over pavement lifecycle [typically 20 years (Morian et al. 2005; Ceylan et al. 2009; Santos and Ferreira 2013)]. Pavement condition deterioration and improvement and budget constraint were specified in probability distribution terms to account for their inherent uncertainty. In the past, stochastic MOGA models were deemed to be highly computationally complex for pavement maintenance optimization, hindering prior research from considering decision variables, such as the pavement condition deterioration and improvement, in probabilistic terms. The current study proposed three approaches to incorporate the uncertainty of pavement condition deterioration and improvement while avoiding high computational complexity to achieve feasible number of combinations. These three approaches are (1) using PM treatments that are most commonly used by STAs, (2) clustering pavement sections into seven groups based on their ages, and (3) creating a filtering constraint by applying the notion of rest period after pavement treatment applications to reduce the search space for the stochastic MOGA algorithm. Whenever a pavement section receives a treatment, no further treatment is provided to this section during a rest period specified as three years in this study. The parameters of the MOGA model (crossover fraction, mutation rate, population size and number of generations) were first determined using sensitivity analysis and based on the lowest values for the fitness functions. The stochastic MOGA models were developed using these estimated MOGA parameters.

## 6.2. Summary of the Results

The research conducted in this dissertation answers the research questions and accomplishes the research objectives outlined in Chapter 1. Table 6.1 summarizes the research objectives, analyses performed and findings from the analyses.

Table 6.1. Research objectives, analyses performed and summary of the findings

| Research objectives | Analyses performed | Summary of the findings |
|---|---|---|
| State-of-the-art review for Markov models | • Synthesis of Markov models and TPM estimation methods<br>• Decision Tree | • Homogeneous Markov models require observations of pavement condition for only two successive transitions<br>• Staged-homogeneous Markov models need two consecutive transitions of pavement condition every stage (5-6 years)<br>• Semi-Markov models need pavement performance curves<br>• Hidden Markov models are used when pavement condition data are incomplete |

| | | |
|---|---|---|
| | | • Non-homogeneous Markov models require extensive historical pavement condition data and consider explanatory variables effects<br>• If percentage transition method is used to estimate the TPM in non-homogeneous Markov models, the effect of explanatory variables is not incorporated, leading to less reliable models |
| Incorporating PM impacts into probabilistic pavement performance models | Data Simulation | The traffic and climate loadings were found to be the statistically significant variables affecting pavement condition in the absence of historical PM data |
| | Initial times and types of PM (literature, Survey 1 and ordered-probit Model A) | The initial times for the PM treatments were estimated as 4 – 10 years with 1.5-year standard deviation (STD) for micro-surfacing, 11 - 15 years with 1-year STD for UTBWC, and 16 – 20 years with 1-year STD for thin overlay |
| | Data generation (Greedy Algorithm optimization and ordered-probit Models B) | At a 60% funding the simulated data including types and times of PM treatments performs comparably to the actual data |
| | Ordered-probit model C | • Traffic and climate loadings, as well as the PM treatments micro-surfacing, UTBWC and thin HMA overlay, were found to be statistically significant<br>• The lack of inclusion of the effect of PM on pavement condition causes an underestimation of the condition and remaining service life of pavement, which could lead to erroneous and non-optimal pavement M&R decisions |
| | Validation | • The MAPE of the non-homogenous Markov model developed using the introduced hybrid approach was found to be equal to 13%, indicating that the methodology and models developed are reliable and accurate in probabilistic pavement condition prediction<br>• The SMEs' validation revealed that pavement performance curves developed using the current study are more accurate and practical than those developed using prior studies<br>• Although the SMEs do not strongly agree with some trends in pavement performance curves such as that in condition state 4, their overall evaluations range from agreement to strong agreement |
| Comparative analysis of Markovian methodologies | Statistical comparative analysis between homogeneous, staged-homogeneous and semi-Markov models with respect to NCS, LDC, data collection time and stage length | • For some NCS and LDC combinations, the semi-Markov models outperformed the staged-homogeneous and the homogeneous Markov counterparts, which is consistent with the literature<br>• The staged-homogeneous and homogeneous Markov models were found to be superior to the semi-Markov models when using specific NCS and LDC |

| | | |
|---|---|---|
| | | • The NCS and LDC significantly affect the prediction accuracy of homogeneous, staged-homogeneous and semi-Markov models<br>• Increasing the NCS increases the prediction accuracy of the three Markov methodologies until the NSC reaches 8 (for a 1-year duty cycle) and 5 (for a 2-year duty cycle). Beyond these thresholds, the prediction accuracy begins to decrease<br>• In the homogeneous and staged-homogeneous Markov models, the data collection time at different pavement cohorts (P1, P2 and P3) and stage size were found to have significant impacts on the predictive accuracy of Markov models<br>• Homogeneous Markov models were found to have high prediction accuracy when using data collected during the middle (10-18 years) or late (19-27 years) ages<br>• The use of a stage length other than the typical length (5 or 6 years) for the staged-homogeneous Markov models yields more accurate predictions of pavement conditions |
| Stochastic pavement PM optimization | MOGA sensitivity analysis | • MOGA parameters were selected by performing sensitivity analyses and their values are as follows:<br>  ○ Crossover fraction: 0.6<br>  ○ Mutation rate: 0.01<br>  ○ Population size: 4000<br>  ○ Number of generations: 100<br>• These MOGA parameters can be used in future similar MOGA models for optimizing pavement maintenance at network level |
| | Stochastic MOGA optimization model considering the uncertainty of pavement condition deterioration and improvement | • The uncertainty or variation of pavement condition deterioration and improvement significantly influences the optimal Pareto frontiers for PM scheduling<br>• This uncertainty is more significant in late age (>20 years) than early age (<7 years) pavements<br>• Deterministic MOGA models provide one PM schedule for each expected total maintenance cost. Whereas, stochastic MOGA models offer multiple PM schedules for each expected total maintenance cost, each PM schedule is associated with the probability of a corresponding pavement network deterioration. This provides decision makers with multiple choices that suit their level of risk/uncertainty |

## 6.3. Limitations of the Study

The current research was carried out for interstate flexible pavements using pavement condition data collected from the LTPP database from Midwestern states. Therefore, the results and their implications are specific to the collected data, its quantity, quality and variation. Since pavement condition data were collected from various Midwest states, data variation and unobserved heterogeneity may be higher because STAs have different standards and specifications for pavement design, construction, and maintenance. Using data from different states may also result in spatial heterogeneity. In addition, the results are limited to the condition data and preventive maintenance treatments of interstate flexible pavements.

This dissertation used the international roughness index (IRI) as the sole pavement condition indicator for modeling pavement performance and optimizing PM. However, the results of Survey 1 show that some STAs use pavement condition indicators and distresses other than the IRI, such as PCI, PSI, rutting and crack index, either individually or in combination, to assess pavement condition and to make M&R decisions. Moreover, the findings of Survey 1 reveal that half of the responding STAs apply the worst-first approach to pavement maintenance decision-making. Accordingly, a greedy algorithm was designed for the optimization of pavement PM in the hybrid approach developed to account for the PM impacts in probabilistic pavement performance models. Furthermore, only three PM treatments (micro-surfacing, UTBWC and thin HMA overlay) were considered in the hybrid approach and the stochastic MOGA optimization models based on the reviewed literature and results of Survey 1, as these PM treatments were recognized as the most commonly used among highway agencies.

Two independent variables (*cum. AADTT* and *cum. AAFI*) were found to be the only statistically significant variables contributing to pavement deterioration possibly because of the quality and limited amount of data. However, other explanatory variables, such as pavement layer thickness, modulus of subgrade, construction quality and drainage, also have varying impacts on pavement deterioration. The lack of consideration of such influential variables may lead to omitted variable bias, and unreliable and erroneous results. The ordered-probit method was used for the development of pavement performance models in the hybrid approach; however, this method assumes that pavement condition data are normally distributed, which is often untrue.

Due to data limitations, this study adopts only eight different combinations of NCS and LDC to conduct the comparative analysis for Markovian methodologies. Nonetheless, more combinations should be created to further investigate the effect of the NCS and LDS on the prediction accuracy of pavement condition using various Markovian techniques. The percentage transition method was used to estimate the transition probabilities of pavement condition. Other methods which may be more computationally intensive, like the expected-value and econometric models, may yield more reliable estimates for the transition probabilities of pavement condition. In the developed semi-Markov models, the holding times were computed as the average time for pavement sections to stay in a particular condition state prior to migrating to the following states, which is often untrue. To calculate the average holding times, only the historical pavement condition data was used, without taking into account the effect of the potential influential variables on pavement condition.

To overcome the computational complexity of the stochastic MOGA optimization models for pavement PM at the network level, pavement sections were clustered into 7 groups based on their ages, which may not fully capture all incremental changes in pavement condition over time. Another approach proposed to reduce the expected large number of combinations of the stochastic MOGA solutions was to apply a rest period of 3 years after any application of PM treatment; however, the rest period should be correlated with the different PM treatments. The uncertainty of the budget constraint and pavement condition deterioration and improvement was represented in the stochastic MOGA by the normal distribution, commonly used to represent uncertainties in engineering problems. Nevertheless, the uncertainty associated with these decision variables should be estimated on the basis of the actual data.

## 6.4. Contributions of the Research

This research makes various contributions to the body of knowledge and the body of practice in the area of pavement infrastructure asset management. The overall contribution of this research is to enhance probabilistic pavement performance modeling and decision-making for pavement preventive maintenance.

### 6.4.1. Contributions to the Body of Knowledge

The comprehensive synthesis of Markov methodologies and models provides insights into the selection of Markov methodologies and TPM estimation methods for pavement deterioration modeling under given conditions of data types and availability. This dissertation presented critical analyses of various aspects of Markovian models as they were applied in the literature, revealed research gaps in the relevant body of knowledge, and offered suggestions to address these gaps for future research. Also, it introduced a decision tree for selecting the appropriate Markov technique and TPM estimation method, which provides researchers with guidance and decision support in selecting appropriate probabilistic techniques for modeling pavement deterioration in a robust manner.

Probabilistic pavement performance models should be developed using quality pavement condition data. The first process of building pavement performance models is data cleaning by removing the observations of extreme outliers. Past studies (e.g., Wang et al. 1994; Thomas and Sobanjo 2012; Pérez-Acebo et al. 2018, 2019) considered observations of improved pavement condition as outliers and, consequently, such observations were removed from the dataset primarily because historical PM data were absent or insufficient. Some PM treatments, however, are used to maintain pavement surfaces at their current condition, making it difficult to identify and/or remove such observations. Hence, data used in prior studies are skewed and less reflective of reality, resulting in underspecified, erroneous and unreliable models suffering from the omitted variable bias. Moreover, although the validation process of most previous models shows their highly accurate predictability for pavement condition, these models have used out-of-sample data that are part of the data cleaned from observations of improved pavement condition. Therefore, prior models are expected to be less accurate when the validation is performed using data containing observations of improved pavement condition. The current study contributes to the body of knowledge by introducing a hybrid approach to incorporate the effect of PM into probabilistic pavement performance models when historical PM data are absent or insufficient. In such cases, the observations of improved pavement condition were not deleted, but rather the types and times of PM treatments that probably associate with these observations were investigated using three complementary methods: literature search, survey of STAs to collect data on pavement condition and PM, and detection of PM times from probabilistic pavement performance curves.

This study indicated that three PM treatments: micro-surfacing, UTBWC and thin HMA overlay are statistically significant for pavement condition prediction, along with variables such as traffic and climate loadings. This, in turn, eliminates the omitted variable bias, as well as improves models' predictability when cross-validated with data containing observations of improved pavement condition.

Previous research (Thomas and Sobanjo 2012; Abaza 2016a, 2017a) highlighted the need to assess the predictive accuracy of different Markov models based on the estimation of the transition probability matrix (a key component of Markov models). Nevertheless, the influence of the number of condition states (NCS) and length of duty cycle (LDC) (key components of Markov models as well) were not investigated. Hence, this research explored the statistical significance of the NCS and LDC on the performance of Markov modeling techniques based on their predictive accuracy. This dissertation introduced two new design parameters (NCS and LDC) for Markov chain models for pavement infrastructure. In addition, this study used a consistent set of pavement condition data (for interstate flexible pavements collected from the LTPP database) to avoid the potential bias associated with comparing different statistical models across different data sets. The findings of this research are of paramount importance for future relevant research as it adds new design parameters to the design and development of Markov models for pavement condition prediction.

In addition, this research contributes to the body of knowledge by proposing three approaches to overcome the computational complexity of stochastic MOGA models for network-level pavement PM schedule. These three approaches are (1) using PM treatments that are most widely used by STAs, (2) clustering pavement sections into 7 groups based on their ages, and (3) creating a filtering constraint that forces the stochastic MOGA searches to assign zero value (Do Nothing decision) during a rest period (specified as 3 years in this study) following PM treatment applications. These three approaches assisted in investigating the effect of the uncertainty of pavement condition deterioration and improvement on the global optimal PM schedule, the expected overall lifecycle maintenance costs and the corresponding total deterioration of pavement network.

### 6.4.2. Contributions to the Body of Practice

This dissertation has also made significant contributions to the body of practice. It developed a decision tree, which could be used by highway agencies to choose the appropriate Markov methodology and TPM estimation method for modeling the deterioration of their pavements under the data availability condition.

Highway agencies are working to maintain their pavements in a state of good condition by developing and implementing effective PM strategies. Still, they are challenged by the collection and management of PM data mainly due to limited resources. As a result, Departments of Transportation such as Indiana (Gulen et al. 2001) and Arizona (Zaghloul et al. 2006) have built and used pavement performance models that do not account for the impact of PM treatments, despite their prominent influence on pavement condition, due to the absence or insufficiency of historical PM data. Such models, which do not reflect the correct evolution of pavement condition, would lead to erroneous prediction of pavement condition and remaining service life, and could ultimately result in non-optimal and less cost-effective M&R decisions. The current research contributes to the body of practice by providing highway agencies with a hybrid approach to improve their pavement performance models by considering the significance of PM when historical PM data are not available or insufficient. The proposed hybrid approach helps highway agencies develop probabilistic pavement performance models that account for the impact of PM, given their limited funding for data collection and management. Although this study makes a significant contribution to highway agencies that either have limited PM data or lack historical PM data, it can also contribute to the agencies that have sufficient historical PM data. Such agencies can use the proposed hybrid approach along with their current PM data, and modify the initial times for PM treatments in the model based on the actual application times of PM treatments. Also, pavement performance models A, B and C can be assessed and adjusted on the basis of their validation with the actual historical PM data. As a result, these agencies may collect future PM data only to validate and improve the approach and pavement performance models, thereby decreasing the frequency of future PM data collection, which, in turn, would reduce the costs of data collection, storage and management.

Highway agencies use their specific or the typical pavement condition indicators that may be discrete (e.g., present serviceability rating (PSR): 0-5) or continuous (e.g., IRI: greater than 0) such as those used in the states of Ohio and Minnesota, respectively. To develop Markov models, highway agencies use the same discrete indicators or the discretized form of the continuous indicators as the condition states without considering the effect of the NCS on the prediction accuracy of Markovian pavement performance models. Moreover, highway agencies tend to assume that the LDC is the frequency of data collection, typically 1 year, or the frequency of data available for model development. The current research indicates that the NCS in Markovian pavement performance models has significant influence on the accuracy of pavement condition prediction. Additionally, this study reveals that for a specific NCS, for instance 4 states, even though pavement condition data may be available every year, using pavement condition data every two years could yield more accurate predictions. These findings can help highway agencies relate the frequency of data collection, the pavement condition indicators that represent the NCS and the prediction accuracy of pavement performance models. This new relationship could contribute to potential cost savings in data collection and management, as well as to improving the prediction accuracy of pavement condition, which in turn will improve the decision-making of M&R interventions.

Furthermore, this research contributes to the body of practice by developing stochastic MOGA models to schedule pavement PM at the network level. These models consider the uncertainties of budget constraint and pavement condition deterioration and improvement, so that decision-makers (risk-averse or risk-taking) can have more tools and information to plan their maintenance interventions according to their level of risk/uncertainty. By considering the uncertainties of these decision variables, decision-makers will be able to choose their optimal solutions to ensure that future condition of pavement network meets the required standards with a probability of 95%.

## 6.5. Recommendations for Future Research

To overcome the limitations of this study, future research may investigate the applicability of the current research to other types of pavement surfaces and different functional classes, and/or using data from other states across the U.S. Different functional classes and pavement surface types can be considered to investigate whether the uncertainty of pavement condition deterioration and improvement could affect PM schedules and/or probabilistic pavement performance. The proposed hybrid approach and developed pavement performance models should use pavement condition data collected from a single state to reduce data variation. If highway agencies do not have sufficient amount of data to use with the proposed hybrid approach, they may incorporate data from other highway agencies that adopt similar standards and specifications and are located in similar climatic zones. Future research should also account for the potential unobserved heterogeneity arising from the inconsideration of all influential variables, by using random parameters regression models to test for the randomness of the parameters of the developed ordered-probit models. Besides, future research may carry out a spatial transferability test to account for the possible heterogeneity resulting from the use of data from different locations/states.

Furthermore, future research may investigate the use of pavement condition indicators, such as PCI, PSI, rutting and crack index, either individually or in combination, for the application of the current research to account for the STAs that use condition indicators other than the IRI. Also, some of the respondents to Survey 1 stated that they use criteria other than the worst-first to optimize pavement maintenance interventions. Thus, future research could use maintenance decision criteria such as cost-effectiveness and traffic volume to perform the data simulation required in the proposed hybrid approach developed to account for the PM impacts in probabilistic pavement performance models. Survey 1 may be deployed again to the STAs that have not responded to collect more data about the types and timings of pavement PM so that every possible maintenance intervention is considered in the application of the hybrid approach and the stochastic MOGA models. Due to the lack of data on pavement preventive maintenance, the proposed hybrid approach and the non-homogeneous Markov model were validated through surveys of SMEs (Surveys 2 and 3) to gauge the practicality of the results in terms of the probabilistic pavement performance curves and marginal effects of the explanatory variables. Nevertheless, future research may further validate the approach and non-homogenous Markov model using actual historical PM data from one or more state DOTs when PM data are accessible.

The significance of more explanatory variables on probabilistic pavement performance should be explored when more pavement condition data are accessible. The consideration of all potentially influential variables (e.g., pavement layer thickness, modulus of subgrade and construction quality) contributes to eliminating the omitted-variable bias, otherwise ordered-probit models with random parameters should be developed to reduce this bias. To avoid the assumption of normal distribution in the ordered-probit method, other modeling techniques such as artificial neural networks (ANNs) may be used to apply the proposed hybrid approach and develop probabilistic pavement performance models that incorporate PM impacts.

In addition, researchers can explore the significance of additional combinations of NCS and LDC that may shed more light on the conclusions reached in this dissertation regarding the selection of appropriate Markovian methodologies and pavement condition prediction accuracy. Future research may also investigate the use of the expected-value method (more reliable than the percentage transition) to estimate the transition probabilities of pavement condition used in the comparative analysis of Markovian techniques. Moreover, the duration modeling method could be explored to estimate the holding times of semi-Markov models, and to account for the influence of explanatory variables on the estimated holding times, as this might help produce more reliable predictions for pavement condition. The current research investigated the prediction accuracy of pavement condition using different Markovian methodologies (homogeneous, staged-homogeneous and semi-Markov) based on the values of the performance measures: MAPE and RMSE. Nevertheless, future work can further compare the prediction accuracy of pavement condition using these Markovian methodologies by evaluating pavement performance curves developed using these methodologies. Regarding the stochastic MOGA optimization models, future research is recommended to create more pavement groups (more than the 7 groups used in this study) to capture more detail in pavement deterioration and improvement. It is also recommended that the rest period be estimated as the service life of each PM treatment (with a minimum time equal to the warranty period of that treatment) and then used in the development of stochastic MOGA models. The values of the MOGA parameters were determined by performing a sensitivity analysis that takes a large number of iterations and may result in a set of parameters that are not associated with the best fitness functions (minimum network deterioration and maintenance costs). Other approaches, such as hyper-heuristic or adaptive parameters, may therefore be used in future research to automatically tune MOGA parameters.

# LIST OF REFERENCES


Abaza, K. A. (2004). “Deterministic performance prediction model for rehabilitation and management of flexible pavement.” *International Journal of Pavement Engineering*, 5(2), 111-121.

Abaza, K. A. (2006). “Iterative linear approach for nonlinear nonhomogeneous stochastic pavement management models.” *Journal of Transportation Engineering*, 132(3), 244–256.

Abaza, K. A. (2014). *“Derivation of pavement transition probabilities using discrete-time Markov chain.”* Transportation Research Board 93rd Annual Meeting, Washington, DC.

Abaza, K. A. (2016a). “Simplified staged-homogenous Markov model for flexible pavement performance prediction.” *Road Materials and Pavement Design*, 17(2), 365-381.

Abaza, K. A. (2016b). “Back-calculation of transition probabilities for Markovian-based pavement performance prediction models.” *International Journal of Pavement Engineering*, 17(3), 253-264.

Abaza, K. A. (2017a). “Empirical approach for estimating the pavement transition probabilities used in non-homogenous Markov chains.” *International Journal of Pavement Engineering*, 18(2), 128-137.

Abaza, K. A. (2017b). “Empirical Markovian-based models for rehabilitated pavement performance used in a life cycle analysis approach.” *Structure and Infrastructure Engineering*, 13(5), 625-636.

Abaza, K. A. (2019). “Simplified empirical approach for predicting the remaining strength factor used in pavement rehabilitation applications.” *Cogent Engineering*, 6(1), 1571003.

Abaza, K. A., Ashur, S. A., Abu-Eisheh, S. A., and Rabay’a, A. (2001). “Macroscopic optimum system for management of pavement rehabilitation.” *Journal of Transportation Engineering*, 127(6), 493–500.

Abaza, K. A., Ashur, S. A., and Al-Khatib, A. K. (2004). “Integrated pavement management system with a Markovian prediction model.” *Journal of Transportation Engineering*, 130 (1), 24-33.

Abaza, K.A. and Murad, M. (2010). “Pavement rehabilitation project ranking approach using probabilistic long-term performance indicators.” *Transportation Research Record*, 2153(1), 3-12.

Abed, A., Thom, N., and Neves, L. (2019). "Probabilistic prediction of asphalt pavement performance." *Road Materials and Pavement Design*, 20(sup1), S247-S264.

Adedimila, A. S., Olutaiwo, A. O., and Kehinde, O. 2009. "Markovian probabilistic pavement performance prediction models for a developing country." *Journal of Engineering and Applied Science,* 4(1), 13–26.

Abdelaziz, N., Abd El-Hakim, R.T., El-Badawy, S.M., and Afify, H.A. (2018). "International Roughness Index prediction model for flexible pavements." *International Journal of Pavement Engineering*, 1-12.

Ahmed, A. A. (2010). "Ultra-Thin Bonded Wearing Course performance update." Minnesota (No. MN/RC 2009-30).

Ahmed, A., Bai, Q., Lavrenz, S., and Labi, S. (2015). "Estimating the marginal cost of pavement damage by highway users on the basis of practical schedules for pavement maintenance, rehabilitation and reconstruction." *Structure and Infrastructure Engineering*, 11(8), 1069-1082.

Ahmed, A., Saeed, T.U., and Labi, S. (2016). "Estimation of rest periods for newly constructed/reconstructed pavements." *Transport*, 31(2), 183-191.

Alaswadko, N., Hassan, R., Meyer, D., and Mohammed, B. (2019). "Probabilistic prediction models for crack initiation and progression of spray sealed pavements." *International Journal of Pavement Engineering*, 20(1), 1-11.

Aleadelat, W., Saha, P., and Ksaibati, K. (2018). "Development of serviceability prediction model for county paved roads." *International Journal of Pavement Engineering,* 19(6), 526-533.

Aleadelat, W., Wulff, S., and Ksaibati, K. (2019). "Development of performance prediction models for gravel roads using Markov chains". *American Journal of Civil Engineering*, 7(3), 73-81.

Al-Suleiman, T.I., and Shiyab, A.M. (2003). "Prediction of pavement remaining service life using roughness data—case study in Dubai." *International Journal of Pavement Engineering*, 4(2), 121-129.

Ameri, M., Jarrahi, A., Haddadi, F., and Mirabimoghaddam, M. H. (2019). "A two-stage stochastic model for maintenance and rehabilitation planning of pavements." *Mathematical Problems in Engineering*, 2019, 1-15.

American Association of State Highway Officials (AASHO). The AASHO Road Test. Highway Research Board, Washington, D.C., 1962.

American Association of State Highway and transportation Officials (AASHTO) (1993). Guide for Design of Pavement Structures. Vol. 1, Washington, DC.

Amin, S.R., and Amador-Jiménez, L.E. (2017). “Backpropagation neural network to estimate pavement performance: dealing with measurement errors.” *Road Materials and Pavement Design*, 18(5), 1218-1238.

ASCE Infrastructure Report Card. (2017). Infrastructure Report Card, Roads.

Augeri, M. G., Greco, S., and Nicolosi, V. (2019). “Planning urban pavement maintenance by a new interactive multiobjective optimization approach.” *European Transport Research Review*, 11(1), 17.

Bektas, F., Smadi, O. G., and Al-Zoubi, M. (2014). “*Pavement management performance modeling: Evaluating the existing PCI equations*.” InTrans Project Reports, Institute for Transportation, Iowa State Univ., Ames, IA.

Belsley, D.A., Kuh, E., and Welsch, R.E. (1980). “*Regression diagnostics: identifying influential data and sources of collinearity*.” Wiley Series in Probability and Statistics.

Bianchini, A., and Bandini, P. (2010). “Prediction of pavement performance through Neuro-Fuzzy Reasoning.” *Computer-Aided Civil and Infrastructure Engineering*, 25(1), 39–54.

Butt, A.A., Shahin, M.Y., Feighan, K.J., and Carpenter, S.H. (1987). “Pavement performance prediction model using the Markov process.” *Transportation Research Record*, 1123, 12-19.

Ceylan, H., Coree, B., and Gopalakrishnan, K. (2009). “Evaluation of the mechanistic-empirical pavement design guide for implementation in Iowa.” *Baltic Journal of Road and Bridge Engineering*, 4(1), 5.

Chan, W. T., Fwa, T. F., and Tan, C. Y. (1994). “Road-Maintenance planning using genetic algorithms. I: formulation.” *Journal of Transportation Engineering*, 120(5), 693–709.

Chen, C., and Zhang, J. (2011). “*Comparisons of IRI-based pavement deterioration prediction models using New Mexico pavement data.*” Conference: Geo-Frontiers Congress 2011

Chen, S., Montgomery, J., and Bolufé-Röhler, A. (2015). “Measuring the curse of dimensionality and its effects on particle swarm optimization and differential evolution.” *Applied Intelligence*, 42(3), 514-526.

Chen, S., Saeed, T. U., Alqadhi, S. D., and Labi, S. (2019). “Safety impacts of pavement surface roughness at two-lane and multi-lane highways: accounting for heterogeneity and seemingly unrelated correlation across crash severities.” *Transportmetrica A: transport science*, 15(1), 18-33.

Chou, E., Pulugurta, H., and Datta, D. (2008). “*Pavement forecasting models.*” No. FHWA/OH-2008/3. Ohio. Dept. of Transportation.

Chu, C.Y. and Durango-Cohen, P.L. (2008). “Empirical comparison of statistical pavement performance models.” *Journal of Infrastructure Systems*, 14(2), 138-149.

Chun, P., Inoue, T., Seto, D., and Ohga, M. (2012). “Prediction of bridge deterioration using GIS-based Markov transition matrix.” *Society for Social Management Systems Internet Journal*, Vol. 8.

Christopher, B. R., Schwartz, C. W., Boudreaux, R., and Berg, R. R. (2006). “*Geotechnical aspects of pavements.*” Technical Report, No. FHWA-NHI-05-037, United States. Federal Highway Administration.

Dalla Rosa, F., Liu, L., and Gharaibeh, N.G. (2017). “IRI prediction model for use in network-level pavement management systems.” *Journal of Transportation Engineering*, Part B: Pavements, 143(1), 04017001.

Denysiuk, R., Moreira, A. V., Matos, J. C., Oliveira, J. R., and Santos, A. (2017). “Two-stage multiobjective optimization of maintenance scheduling for pavements.” *Journal of Infrastructure Systems*, 23(3), 04017001.

Deshpande, V. P., Damnjanovic, I. D., and Gardoni, P. (2010). “Reliabilitybased optimization models for scheduling pavement rehabilitation.” *Computer-Aided Civil Infrastructure Engineering*, 25(4), 227–237.

Donev, V. and Hoffmann, M. (2019). “Condition prediction and estimation of service life in the presence of data censoring and dependent competing risks.” *International Journal of Pavement Engineering*, 20(3), 313-331.

Durango, P. L. (2002). “*Adaptive optimization models for infrastructure management.*” University of California, Berkeley, Issue Number: 3082172.

Eldin, N., and Senouci, A. (2011). “Condition rating of rigid pavements by neural networks.” *Canadian Journal of Civil Engineering*, 22, 861-870.

Elhadidy, A. A., Elbeltagi, E. E., and Ammar, M. A. (2015). “Optimum analysis of pavement maintenance using multi-objective genetic algorithms.” *HBRC Journal*, 11(1), 107-113.

Elnashar, G., Bhat, R. B., and Sedaghati, R. (2019). “Modeling pavement damage and predicting fatigue cracking of flexible pavements based on a combination of deterministic method with stochastic approach using Miner's hypothesis.” *SN Applied Sciences*, 1(3), 229.

Ferreira, A. J. L., Meneses, S. C. N., and Vicente, F. A. A. (2009). “*Alternative decision-aid tool for pavement management.”* In Proceedings of the Institution of Civil Engineers-Transport (Vol. 162, No. 1, pp. 3-17). Thomas Telford Ltd.

Ferreira, A., and Santos, L. P. (1999). “*Pavement performance modelling: State of the art.*” In Proceedings of Seventh International Conference on Civil and Structural Engineering Computing (pp. 157-264).

FHWA (2017). “Construction, Highway Quality Compendium, Pavement Warranties Yield Innovation, Quality.” Data was accessed on March 5, 2019. https://www.fhwa.dot.gov/construction/pubs/hif07012/11.cfm

FHWA (Federal Highway Administration) (2017). “Highway Statistics 2017.” Retrieved from: https://www.fhwa.dot.gov/policyinformation/statistics/2017/

FHWA. (2018). “Freight Economy | Federal Highway Administration.” <https://www.fhwa.dot.gov/freighteconomy/> (Nov. 16, 2018).

Ford, K. M., Arman, M., Labi, S., Sinha, K. C., Shirole, A., Thompson, P., and Li, Z. (2011). “Methodology for estimating life expectancies of highway assets.” Washington, DC: National Cooperative Highway Research Program.

Fuentes, L., Camargo, R., Arellana, J., Velosa, C., and Martinez, G. (2019). “Modelling pavement serviceability of urban roads using deterministic and probabilistic approaches.” *International Journal of Pavement Engineering*, 1-10.

Fwa, T. F., Chan, W. T., and Tan, C. Y. (1996). “Genetic-algorithm programming of road maintenance and rehabilitation.” *Journal of Transportation Engineering*, 122(3), 246-253.

Fwa T. F., Chan W. T., and Hoque K. Z. (2000). “Multiobjective optimization for pavement maintenance programming.” *Journal of Transportation Engineering*, 126(5), 367–374.

Fwa, T. F., Sinha, K. C., and Riverson, J. D. (1988). “Highway routine maintenance programming at network level.” *Journal of Transportation Engineering*, 114(5), 539-554.

Fwa, T. F., Tan, C. Y., and Chan, W. T. (1997). “Backcalculation analysis of pavement-layer moduli using genetic algorithms.” *Transportation Research Record*, 1570(1), 134-142.

Galvis Arce, O. D., and Zhang, Z. (2019). “Skid resistance deterioration model at the network level using Markov chains.” *International Journal of Pavement Engineering*, 1-9.

Gao, L., and Zhang, Z. (2008). “Robust optimization for managing pavement maintenance and rehabilitation.” *Transportation Research Record: Journal of the Transportation Research Board*, 2084, 55–61.

Gao, L., and Zhang, Z. (2011). “*Optimal infrastructure maintenance scheduling problem under budget uncertainty*.” Southwest Region University Transportation Center, Center for Transportation Research, the University of Texas at Austin.

Gao, L., and Zhang, Z. (2013). “Management of pavement maintenance, rehabilitation, and reconstruction through network partition.” *Transportation Research Record*, 2366(1), 59-63.

García de Soto, B., Bumbacher, A., Deublein, M., and Adey, B.T. (2018). “Predicting road traffic accidents using Artificial Neural Network models.” *Infrastructure Asset Management*, 1-13.

Geoffroy, D. N. (1996). “*NCHRP synthesis of highway practice 223: cost-effective preventive pavement maintenance.*” Transportation Research Board, National Research Council, Washington, DC.

Golabi, K., Kulkarni, R., and Way, G. (1982). “A statewide pavement management system.” *Interfaces*, 12(6), 5–21.

Gulen, S., Zhu, K., Weaver, J., Shan, J., and Flora, W. (2001). “*Development of improved pavement performance prediction models for the Indiana pavement management system.*” Joint Transportation Research Program, 69.

Guo, F., Gregory, J., and Kirchain, R. (2020). “Incorporating cost uncertainty and path dependence into treatment selection for pavement networks.” *Transportation Research Part C: Emerging Technologies*, 110, 40-55.

Greene, W. H., and Hensher, D. A. (2010). “Ordered choices and heterogeneity in attribute processing.” *Journal of Transport Economics and Policy (JTEP)*, 44(3), 331-364.

Hassan, R., Lin, O., and Thananjeyan, A. (2017a). "A comparison between three approaches for modelling deterioration of five pavement surfaces." *International Journal of Pavement Engineering*, 18(1), 26-35.

Hassan, R., Lin, O., and Thananjeyan, A. (2017b). "Probabilistic modelling of flexible pavement distresses for network management." *International Journal of Pavement Engineering*, 18(3), 216-227.

Hong, F. and Prozzi, J.A. (2010). "Roughness model accounting for heterogeneity based on in-service pavement performance data." *Journal of Transportation Engineering*, 136(3), 205-213.

Hong, H.P. and Somo, S. (2001). "Probabilistic assessment of pavement conditions." *Canadian Journal of Civil Engineering*, 28(5), 813–822.

Humphries, J. A., and Ma, E. (2004). "*Optimization of pavement preservation activities*." The 2004 Annual Conference of the Transportation Association of Canada, Quebec City, Quebec.

Indiana Design Manual (2013). Indiana Department of Transportation, Indianapolis, Indiana. http://www.in.gov/indot/design_manual/design_manual_2013.htm

INDOT Standard Specifications (2020). https://www.in.gov/dot/div/contracts/standards/book/index.html.

Initiatives. NCHRP 20-07, National Cooperative Highway Research Program, 2005.

Irfan, M., Khurshid, M. B., Bai, Q., Labi, S., and Morin, T. L. (2012). "Establishing optimal project-level strategies for pavement maintenance and rehabilitation–A framework and case study." *Engineering Optimization*, 44(5), 565-589.

Jacobs, T. L. (1992). "Optimal long-term scheduling of bridge deck replacement and rehabilitation." *Journal of Transportation Engineering*, 118(2), 312-322.

Ji, Y. G., Nantung, T., and Tompkins, B. (2015). "Evaluation for UTBWC on SR-11 as pavement preservation treatment: a case study." *International Journal of Pavement Research and Technology*, 8(4), 267-275.

Jorge, D., and Ferreira, A. (2012). "Road network pavement maintenance optimization using the HDM-4 pavement performance prediction models." *International Journal of Pavement Engineering*, 13(1), 39-51.

Kargah-Ostadi, N., Zhou, Y., and Rahman, T. (2019). “Developing performance prediction models for pavement management systems in local governments in absence of age data.” *Transportation Research Record*, 2673(3), 334-341.

Keane, P. L., and Wu, L. (1985). “*An integrated decision-making methodology for optimal maintenance strategies.*” Department of General Engineering Technical Report, University of Illinois, Urban-Champaign, IL.

Keane, P.L. and Keane, M.I. (1985). “*An integrated decision-making methodology for optimal maintenance strategies.*” MS design project, Department of General Engineering, University of Illinois at Urban-Champaign.

Khan, M. U., Mesbah, M., Ferreira, L., and Williams, D. J. (2014). “Development of road deterioration models incorporating flooding for optimum maintenance and rehabilitation strategies.” *Road and Transport Research: A Journal of Australian and New Zealand Research and Practice*, 23(1), 3-24.

Khattak, M. J., Landry, C., Veazey, J., and Zhang, Z. (2013). “Rigid and composite pavement index-based performance models for network pavement management system in the state of Louisiana.” *International Journal of Pavement Engineering*, 14(7), 612-628.

Khattak, M.J., Nur, M.A., Bhuyan, M.R.U.K., and Gaspard, K. (2014). “International Roughness Index models for HMA overlay treatment of flexible and composite pavements.” *International Journal of Pavement Engineering*, 15(4), 334-344.

Khurshid, M. B. (2010). “*A framework for establishing optimal performance thresholds for highway asset interventions.*” Ph.D., Purdue University, United States - Indiana.

Khurshid, M. B., Irfan, M., Ahmed, A., and Labi, S. (2014). “Multidimensional benefit–cost evaluation of asphaltic concrete overlays of rigid pavements.” *Structure and Infrastructure Engineering*, 10(6), 792–810.

Khurshid, M. B., Irfan, M., and Labi, S. (2011). “An analysis of the cost-effectiveness of rigid pavement rehabilitation treatments.” *Structure and Infrastructure Engineering*, 7(9), 715–727.

Kim, S.H. and Kim, N. (2006). “Development of performance prediction models in flexible pavement using regression analysis method.” *KSCE Journal of Civil Engineering*, 10(2), 91–96.

Kobayashi, K., Do, M., and Han, D. (2010). “Estimation of Markovian transition probabilities for pavement deterioration forecasting.” *KSCE Journal of Civil Engineering*, 14 (3), 343-351.

Kobayashi, K., Kaito, K., and Lethanh, N. (2012). “A Bayesian estimation method to improve deterioration prediction for infrastructure system with Markov chain model.” *International Journal of Architecture, Engineering and Construction*, 1(1), 1-13.

Kulkarni, R. B., Miller, D., Ingram, R. M., Wong, C. W., and Lorenz, J. (2004). “Need-based project prioritization: alternative to cost-benefit analysis.” *Journal of Transportation Engineering*, 130(2), 150-158.

Labi, S., Faiz, A., Saeed, T. U., Alabi, B. N. T., and Woldemariam, W. (2019). “Connectivity, accessibility, and mobility relationships in the context of low-volume road networks.” *Transportation Research Record*, 0361198119854091.

Labi, S., Hwee, K. S., Lamptey, G., and Nunoo, C. (2006). “*Long-term benefits of microsurfacing applications in Indiana–methodology and case study*.” In Proceedings of the 85th Annual Meeting of the Transportation Research Board.

Labi, S., and Sinha, K. C. (2003). “Measures of short-term effectiveness of highway pavement maintenance.” *Journal of Transportation Engineering*, 129(6), 673-683.

Lamptey, G., Labi, S., and Li, Z. (2008). “Decision support for optimal scheduling of highway pavement preventive maintenance within resurfacing cycle.” *Decision Support Systems*, 46(1), 376–387.

Lamptey, G., Singh, L., Labi, S., and Sinha, K. C. (2010). “Systematic framework for incorporating safety in network-level transportation planning and programming.” *Journal of Transportation Engineering*, 136(5), 436-447.

La Torre, F., Domenichini, L., and Darter, M. I. (1998). “*Roughness prediction model based on the artificial neural network approach.*” Proc., 4th Int. Conf. on Managing Pavements, Transportation Research Board, Washington, DC.

Lee, J., and Shields, T. (2010). “*Treatment guidelines for pavement preservation.*”

Lethanh, N. and Adey, B.T. (2012). “A Hidden Markov model for modeling pavement deterioration under incomplete monitoring data.” *World Academy of Science, Engineering and Technology*, 6(1), 46-53.

Lethanh, N. and Adey, B.T. (2013). “Use of exponential hidden Markov models for modeling pavement deterioration.” *International Journal of Pavement Engineering*, 14(7), 645-654.

Lethanh, N., Kaito, K., and Kobayashi, K. (2015). “Infrastructure deterioration prediction with a Poisson hidden Markov model on time series data.” *Journal of Infrastructure Systems*, 21(3), 0401405.

Li, M., and Wang, H. (2019). “Development of ANN-GA program for backcalculation of pavement moduli under FWD testing with viscoelastic and nonlinear parameters.” *International Journal of Pavement Engineering*, 20(4), 490-498.

Li, N. (1997). “*Development of a probabilistic based integrated pavement management system.*” Ph.D. Thesis, WaterLoo, Ontario, Canada.

Li, N., Xie, W.C., and Haas, R. (1996). “Reliability-based processing of Markov chains for modeling pavement network deterioration.” *Transportation Research Record: Journal of the Transportation Research Board*, (1524), 203-213.

Li, Y., and Madanat, S. (2002). “A steady-state solution for the optimal pavement resurfacing problem.” *Transportation Research Part A: Policy and Practice*, 36(6), 525–535.

Li, Z. (2005). “*A probabilistic and adaptive approach to modeling performance of pavement infrastructure.*” PhD Thesis, University of Texas, Austin, Tx.

Li, Z., Madanu, S., Zhou, B., Wang, Y., and Abbas, M. (2010). “A heuristic approach for selecting highway investment alternatives.” *Computer‐Aided Civil and Infrastructure Engineering*, 25(6), 427-439.

Li, Z., and Puyan, M. (2006). “A stochastic optimization model for highway project selection and programming under budget uncertainty.” *Applications of Advanced Technology in Transportation (2006)*, Proceedings

Los Angeles County Department of Public Works, Pavement Management Section, https://dpw.lacounty.gov/gmed/lacroads/TreatmentMicrosurfacing.aspx

Lou, Z., Gunaratne, M., Lu, J. J., and Dietrich, B. (2001). “Application of neural network model to forecast short-term pavement crack condition: Florida case study.” *Journal of Infrastructure Systems*, 7(4), 166-171.

Luo, Z. (2013). “Pavement performance modeling with an auto-regression approach.” *International Journal of Pavement Engineering*, 14(1), 85-94.

MacLeod, D. R., and Walsh, R. (1998). “*Markov modeling-a case study.*” In Proc., 4th International Conference on Managing Pavements.

Madanat, S. M., Karlaftis, M. G., and McCarthy, M. G. (1997). “Probabilistic infrastructure deterioration models with panel data.” *Journal of Infrastructure Systems*, 3(1), 4-9.

Madanat, S. M., Nakat, Z. E., and Sathaye, N. (2005). “*Development of empirical mechanistic pavement performance models using data from the Washington state PMS database.*” University of California, Davis and Berkeley.

Madanat, S., Bulusu, S., and Mahmoud, A. (1995a). “Estimation of infrastructure distress initiation and progression models.” *Journal of Infrastructure Systems*, 1(3), 146-150.

Madanat, S., Mishalani, R., and Ibrahim, W. H. W. (1995b). “Estimation of infrastructure transition probabilities from condition rating data.” *Journal of Infrastructure Systems*, 1(2), 120-125.

Madanat, S., Park, S., and Kuhn, K. (2006). “Adaptive optimization and systematic probing of infrastructure system maintenance policies under model uncertainty.” *Journal of Infrastructure Systems*, 12(3), 192–198.

Mamlouk, M. S., and Zaniewski, J. P. (1998). “*Pavement preventive maintenance: description, effectiveness and treatments.*” Flexible Pavement Rehabilitation and Maintenance, ASTM STP 1348, West Conshohocken, Pa., 121-135

Mandiartha, P., Duffield, C.F., Thompson, R G., and Wigan, M. (2010). “*Stochastic based models for low volume road (LVR) network management.*” In ARRB Conference, 24th, 2010ARRB Group Limited, Melbourne.

Martin, T., and Kadar, P. 2012. “*The scoping and development of probabilistic road deterioration models.*” Austroads: Sydney (Austroads Publication No. AP-T201-12).

Mauch, M., and Madanat, S. (2001). “Semiparametric hazard rate models of reinforced concrete bridge deck deterioration.” *Journal of Infrastructure Systems*, 7(2), 49-57.

Mazari, M., and Rodriguez, D.D. (2016). “Prediction of pavement roughness using a hybrid gene expression programming-neural network technique.” *Journal of Traffic and Transportation Engineering (English Edition)*, 3(5), 448–455.

Medury, A., and Madanat, S. (2014). “Simultaneous network optimization approach for pavement management systems.” *Journal of Infrastructure Systems*, 20(3), 04014010.

Meneses, S., Ferreira, A., and Collop, A. (2013). “Multi-objective decision-aid tool for pavement management.” *Proceedings of the Institution of Civil Engineers - Transport*, 166(2), 79–94.

Midland Asphalt Materials Inc. 2018, http://www.midlandasphalt.com/pavement-preservation/novachip/

Mills, L. N., Attoh-Okine, N. O., and McNeil, S. (2011). "Hierarchical Markov Chain Monte Carlo simulation for modeling transverse cracks in highway pavements." *Journal of Transportation Engineering*, 138(6), 700-705.

Mishalani, R. G., and Madanat, S. M. (2002). "Computation of infrastructure transition probabilities using stochastic duration models." *Journal of Infrastructure Systems*, 8(4), 139-148.

Mohammadi, A., Amador-Jimenez, L., and Elsaid, F. (2019). "Simplified pavement performance modeling with only two-time series observations: a case study of Montreal Island." *Journal of Transportation Engineering, Part B: Pavements*, 145(4), 05019004.

Montgomery, S. R., Labi, S., and Haddock, J. E. (2018). "Load transfer restoration with diamond grinding on rigid pavements: Short-term and long-term effectiveness." *Journal of Transportation Engineering, Part B: Pavements*, 144(1), 04017019.

Morcous, G., and Lounis, Z. (2005). "Maintenance optimization of infrastructure networks using genetic algorithms." *Automation in construction*, 14(1), 129-142.

Moreira, A. V., Tinoco, J., Oliveira, J. R. and Santos, A. (2018). "An application of Markov chains to predict the evolution of performance indicators based on pavement historical data." *International Journal of Pavement Engineering*, 19(10), 937-948.

Morian, D. A., Zhao, Y., Arellano, J., and Hall, D. E. (2005). "Asphalt pavement rehabilitation treatment: Analysis of 20 years of performance." *Transportation Research Record*, 1905(1), 36-43.

National Highway Construction Cost Index (NHCCI) 2.0 (2019). Retrieved from https://www.fhwa.dot.gov/policy/otps/nhcci/pt1.cfm

NCHRP 1-37A Design Guide (2002). 2002 Design Guide – Design of New and Rehabilitated Pavement Structures, Draft Final Report, Part 1 – Introduction and Part 2 – Design Inputs, Prepared for the National Cooperative Highway Research Program by ERES Division of ARA.

Nesbit, D. M., Sparks, G. A., and Neudorf, R. D. (1993). "A semi-Markov formulation of the pavement maintenance optimization problem." *Canadian Journal of Civil Engineering*, 20(3), 436–447.\

Nunez, M. M., and Shahin, M. Y. (1986). “Pavement condition data analysis and modeling.” *Transportation Research Record*, 1070, 125-132.

Odoki, J. B., and Kerali, H. G. R. 2000. “*Analytical framework and model descriptions. Highway Development and Management (HDM-4) Volume 4.*” The World Road Association. Washington DC.

Office of Management and Budget. (2018). “Budget Of The U.S. Government, 2019.”

Ong, G. P. (Raymond), Nantung, T., and Sinha, K. (2010). “*Indiana pavement preservation program.*” JTRP Technical Reports.

Ortiz-García, J. J., Costello, S. B., and Snaith, M. S. (2006). “Derivation of transition probability matrices for pavement deterioration modeling.” *Journal of Transportation Engineering*, 132(2), 141–161.

Osorio-Lird, A., Chamorro, A., Videla, C., Tighe, S., and Torres-Machi, C. (2018). “Application of Markov chains and Monte Carlo simulations for developing pavement performance models for urban network management.” *Structure and Infrastructure Engineering*, 14(9), 1169-1181.

Ouyang, Y., and Madanat, S. (2004). “Optimal scheduling of rehabilitation activities for multiple pavement facilities: exact and approximate solutions.” *Transportation Research Part A: Policy and Practice*, 38(5), 347–365.

Panthi, K. (2009). “*A methodological framework for modeling pavement maintenance costs for projects with performance-based contracts.*” Ph.D. Thesis, Florida International University, Florida, USA.

Pérez-Acebo, H., Bejan, S., and Gonzalo-Orden, H. (2018). “Transition probability matrices for flexible pavement deterioration models with half-year cycle time.” *International Journal of Civil Engineering*, 16(9), 1045-1056.

Pérez-Acebo, H., Mindra, N., Railean, A., and Rojí, E. (2019). “Rigid pavement performance models by means of Markov Chains with half-year step time.” *International Journal of Pavement Engineering*, 20(7), 830-843.

Peshkin, D. G, and Hoerner, T. E. (2005). “*Pavement preservation: practices, research plans, and initiatives.*” Final Report, APTech Project No. 03-126-RR1, Prepared for National Cooperative Highway Research Program, Transportation Research Board, National Research Council.

Peshkin, D. G., Hoerner, T. E., and Zimmerman, K. A. (2004). "*Optimal timing of pavement preventive maintenance treatment applications.*" NCHRP Report 523, Transportation Research Board, Washington, DC.

Peshkin, D., Smith, K. L., Wolters, A., Krstulovich, J., Moulthrop, J., and Alvarado, C. (2011). "*Preservation approaches for high-traffic-volume roadways.*" SHRP2 Report S2-R26-RR-1. The Second Strategic Highway Research Program, Washington, DC.

Plati, C., Georgiou, P., and Papavasiliou, V. (2016). "Simulating pavement structural condition using artificial neural networks." *Structure and Infrastructure Engineering*, 12(9), 1127-1136.

Plati, C., Georgouli, K., Cliatt, B., and Loizos, A. (2017). "Incorporation of GPR data into genetic algorithms for assessing recycled pavements." *Construction and Building Materials*, 154, 1263-1271.

Porras-Alvarado, J. D., Zhang, Z., Salazar, L. G. L., and Chi, E. C. (2014). "*Probabilistic approach to modeling pavement performance using IRI data.*" In 93rd Annual Meeting of the Transportation Research Board, Washington, DC.

Prozzi, J., and Madanat, S. (2000). "Using duration models to analyze experimental pavement failure data." *Transportation Research Record: Journal of the Transportation Research Board*, (1699), 87-94.

Prozzi, J.A., and Madanat, S.M. (2003). "Incremental nonlinear model for predicting pavement serviceability." *Journal of Transportation Engineering*, 129(6), 635-641.

Pulugurta, H. (2007). "*Development of pavement condition forecasting models.*" Ph.D. Thesis, The University of Toledo, Ohio, United States.

Pulugurta, H., Shao, Q., and Chou, Y.J. (2009). "Pavement condition prediction using Markov process." *Journal of Statistics and Management Systems*, 12(5), 853-871.

Qiao, Y., Labi, S., and Fricker, J. D. (2019). "Hazard-based duration models for predicting actual duration of highway projects using nonparametric and parametric survival analysis." *Journal of Management in Engineering*, 35(6), 04019024.

Rajagopal, A. (2006). "*Developing pavement performance prediction models and decision trees for the city of Cincinnati. Ohio Department of Transportation.*" Technical Report, FHWA/OH-2006/14.

Ramaswamy, R. and Ben-Akiva, M. (1990). "Estimation of highway pavement deterioration from in-service pavement data." *Transportation Research Record*, 1272, 96-106.

Ranjith, S., Setunge, S., Gravina, R. and Venkatesan, S. (2013). "Deterioration prediction of timber bridge elements using the Markov chain." *Journal of Performance of Constructed Facilities*, 27(3), 319-325.

Reddy, M. A., Reddy, K. S., and Pandey, B. B. (2004). "Selection of genetic algorithm parameters for backcalculation of pavement moduli." *International Journal of Pavement Engineering*, 5(2), 81-90.

Roadresource (2019). Data retrieved from: https://roadresource.org/

Rose, S., Mathew, B.S., Isaac, K.P., and Abhaya, A.S. (2018). "Risk based probabilistic pavement deterioration prediction models for low volume roads." *International Journal of Pavement Engineering*, 19(1), 88-97.

Saeed, T. U., Qiao, Y., Chen, S., Alqadhi, S., Zhang, Z., Labi, S., and Sinha, K. C. (2017). "*Effects of bridge surface and pavement maintenance activities on asset rating*." Joint Transportation Research Program, Publication No. FHWA/IN/JTRP-2017/19. West Lafayette, IN: Purdue University.

Saeed, T. U., Qiao, Y., Chen, S., Gkritza, K., and Labi, S. (2017). "Methodology for probabilistic modeling of highway bridge infrastructure condition: Accounting for improvement effectiveness and incorporating random effects." *Journal of Infrastructure Systems*, 23(4), 04017030.

Silva, F., Van Dam, T., Bulleit, W., and Ylitalo, R. (2000). "Proposed pavement performance models for local government agencies in Michigan." *Transportation Research Record*, 1699, 81-86.

Sandra, A. K. and Sarkar, A. K. (2013). "Development of a model for estimating International Roughness Index from pavement distresses." *International Journal of Pavement Engineering*, 14(8), 715-724.

Santos, J., and Ferreira, A. (2013). "Life-cycle cost analysis system for pavement management at project level." *International Journal of Pavement Engineering*, 14(1), 71-84.

Santos, J., Ferreira, A., and Flintsch, G. (2019). "An adaptive hybrid genetic algorithm for pavement management." *International Journal of Pavement Engineering*, 20(3), 266-286.

Sarwar, M.T. and Anastasopoulos, P.C. (2016). “Three-stage least squares analysis of postrehabilitation pavement performance.” *Transportation Research Record*, 2589(1), 97-109.

Sarwar, M.T. and Anastasopoulos, P.C. (2017). “The effect of long-term non-invasive pavement deterioration on accident injury-severity rates: A seemingly unrelated and multivariate equations approach.” *Analytic methods in accident research*, 13, 1-15.

Shafahi, Y., and Hakhamaneshi, R. (2009). “Application of a maintenance management model for Iranian railways based on the Markov chain and probabilistic dynamic programming.” *Scientia Iranica, Transaction A: Civil Engineering*, 16(1), 87-97.

Sindi, W., and Agbelie, B. (2020). “Assignments of pavement treatment options: Genetic Algorithms versus Mixed-Integer Programming.” *Journal of Transportation Engineering, Part B: Pavements*, 146(2), 04020008.

State of California Department of Transportation. (2008). “*Maintenance Technical Advisory Guide.*” Flexible Pavement Preservation Second Edition, Common Distresses on Flexible Pavements, Caltrans Division of Maintenance.

Storn, R. (1996). “*On the usage of differential evolution for function optimization.*” In Proceedings of North American Fuzzy Information Processing (pp. 519-523). IEEE.

Sulistyowati, S., and Soehartono, S. (2019). “*Performance modeling of Kaligawe road in Semarang using Markov chains.*” NEO TEKNIKA, 5(1).

Tabatabaee, N. and Ziyadi, M. (2013). “Bayesian approach to updating Markov-based models for predicting pavement performance.” *Transportation Research Record*, 2366(1), 34-42.

Taylor, C. K. (2013). “*An introduction to abstract algebra.*” bookboon.com, 1st edition, ISBN 978-87-403-0367-4

Terzi, S. (2007). “Modeling the pavement serviceability ratio of flexible highway pavements by artificial neural networks.” *Construction and Building Materials*, 21(3), 590-593.

Thomas, O., and Sobanjo, J. (2012). “Comparison of Markov chain and Semi-Markov models for crack deterioration on flexible pavements.” *Journal of Infrastructure Systems*, 19(2), 186-195.

Thomas, O., and Sobanjo, J. (2016). “Semi-Markov models for the deterioration of bridge elements.” *Journal of Infrastructure Systems*, 22(3), 04016010.

“Trading Economics, United States.” (2019). <https://tradingeconomics.com/united-states/gdp> data retrieved on March 11, 2020.

Tsuda, Y., Kaito, K., Aoki, K., and Kobayashi, K. (2006). “Estimating Markovian transition probabilities for bridge deterioration forecasting.” *Structural Engineering / Earthquake Engineering, Japan Society of Civil Engineering*, 23(2), 241-256.

Viking Construction Inc. (2017), http://www.vciss.com/micro.htm

Walls III, J., and Smith, M. R. 1998. “*Life-cycle cost analysis in pavement design-interim technical bulletin.*” Rep. FHWA-SA-98-079, US DOT, Washington, DC.

Wang, K.C., Zaniewski, J., and Way, G. (1994). “Probabilistic behavior of pavements.” *Journal of Transportation Engineering*, 120(3), 358-375.

Washington, S.P., Karlaftis, M.G., and Mannering, F. (2011). “*Statistical and econometric methods for transportation data analysis.*” Second Edition. Chapman and Hall/CRC, Boca Raton FL.

Wang, F., Zhang, Z., and Machemehl, R. (2003). “Decision-making problem for managing pavement maintenance and rehabilitation projects.” *Transp. Res. Rec.*, 1853, 21–28.

Wang, K. C. P., Zaniewski, J., and Way, G. (1994). “Probabilistic behavior of pavements.” *Journal of Transportation Engineering, ASCE*, 120(3), 358–375.

Wilde, W. J., Thompson, L., and Wood, T. J. (2014). “*Cost-effective pavement preservation solutions for the real world.*” (No. MN/RC 2014-33) Department of Transportation, Research Services & Library.

Wong, W. G., He, G. P., and Luk, S. T. (2003). “*Development of road management systems in China.*” Proceedings of the Institution of Civil Engineers - Transport, 156(4), 179–188.

Wright, K. (2016). “*America’s roads and bridges are in disrepair—find out where your state ranks*.” AlterNet.

Wu, Z., and Flintsch, G. W. (2009). “Pavement preservation optimization considering multiple objectives and budget variability.” *Journal of Transportation Engineering*, 135(5), 305–315.

Xu, G., Bai, L., Sun, Z., Nowaczyk, T., Shive, C., and Wilcoxson, J. (2015). “*Pavement deterioration prediction model and project selection for Kentucky highways.*” 9th International Conference on Managing Pavement Assets (ICMPA9), Washington, DC.

Yamany, M. S., and Abraham, D. M. (2020a). “*Prediction of pavement performance using non-homogeneous Markov models: incorporating the impact of preventive maintenance.*” Transportation Research Board 99th Annual Meeting, Washington D.C.

Yamany, M. S., and Abraham, D. M. (2020b). “Hybrid approach to incorporate preventive maintenance effectiveness into probabilistic pavement performance models.” *Journal of Transportation Engineering, Part B: Pavements.*

Yamany, M. S., Abraham, D. M., and Labi S. (2019a). “*Probabilistic modeling of pavement performance using Markov chains: a state-of-the-art review*.” ASCE International Conference on Transportation & Development, Virginia, USA

Yamany. M.S., Abraham, D.M., and Labi, S. (2020a). “Comparative analysis of Markovian methodologies for modeling infrastructure system performance.” *Journal of Infrastructure Systems*.

Yamany, M. S., and Elwakil, E. (2019). “Modelling of cast-in-place concrete tunnel liners condition.” *Structure and Infrastructure Engineering*. Doi: 10.1080/15732479.2019.1687529

Yamany, M. S., and Elwakil, E. (2020). “*Condition modeling of cast-in-place concrete liners for highway tunnels across the United States.*” Transportation Research Board 99th Annual Meeting, Washington D.C.

Yamany, M. S., Saeed, T. U., and Volovski, M. (2019b), “*Performance prediction of interstate flexible pavement across the midwestern United States: random-parameter regression vs. artificial neural network.*” Transportation Research Board 98th Annual Meeting, Washington D.C.

Yamany, M. S., Saeed, T. U., Volovski, M., and Ahmed, A. (2020b), “Characterizing the performance of interstate flexible pavements using artificial neural networks and random parameters regression.” *Journal of Infrastructure Systems*, 26(2), 04020010. Doi: 10.1061/(ASCE)IS.1943-555X.0000542

Yang, J., Gunaratne, M., Lu, J.J., and Dietrich, B. (2005). “Use of recurrent Markov chains for modeling the crack performance of flexible pavements.” *Journal of Transportation Engineering*, 131(11), 861-872.

Yang, J., Lu, J., Gunaratne, M., and Xiang, Q. (2003). "*Forecasting overall pavement condition with Neural Networks: Application on Florida highway network.*" Transportation Research Board, Washington, DC.

Yang, J., Lu, J., Gunaratne, M., and Dietrich, B. (2006). "Modeling crack deterioration of flexible pavements: comparison of recurrent Markov chains and artificial neural networks." *Transportation research record: Journal of the Transportation Research Board*, 1974(1), 18-25.

Yang, Y. N., Kumaraswamy, M. M., Pam, H. J., and Xie, H. M. (2013). "Integrating semiparametric and parametric models in survival analysis of bridge element deterioration." *Journal of Infrastructure Systems*, 19(2), 176-185.

Zaghloul, S., Helali, K., and Bekheet, W. (2006). "*Development and implementation of Arizona Department of Transportation (ADOT) Pavement Management System (PMS).*" (No. FHWA-AZ-06-494). Arizona. Dept. of Transportation.

Zhao, Z., and Guo, Z. (2013). "*Prediction of asphalt pavement International Roughness Index (IRI) by combined approach of empirical regression and Markov.*" In ICTE 2013: Safety, Speediness, Intelligence, Low-Carbon, Innovation (pp. 2106-2113).

Ziari, H., Sobhani, J., Ayoubinejad, J., and Hartmann, T. (2016). "Prediction of IRI in short and long terms for flexible pavements: ANN and GMDH methods." *International Journal of Pavement Engineering*, 17(9), 776-788.

# APPENDIX

## A.1. Questionnaire Survey 1

**Preventive Maintenance of Pavements**

I am Mohamed Yamany, a Ph.D. candidate in the Lyles School of Civil Engineering, Purdue University. For my doctoral studies, I am analyzing the effect of preventive maintenance on pavement condition. This survey aims at collecting **data** of the **preventive maintenance treatments** implemented on **Interstate Asphalt Pavements.**

Thank you for participating in this survey.

Q1. Name of State Transportation Agency (STA):

____________________________________________________

____________________________________________________

____________________________________________________

____________________________________________________

Q2. Please enter each preventive maintenance treatment used by your agency for Interstate Flexible Pavements, and at what pavement ages (from 1 to 35) each treatment is implemented.

| Treatments | Pavement Age in years | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 |
| 1: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 5: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 6: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Q3. What are the criteria for the choice of preventive maintenance treatments and their timings? Please list each criterion (e.g., cost, effectiveness, worst first, etc.) and its score from 1 to 5; where 1 represents the least important criterion and 5 is the most important criterion?

| **Criteria** | **Score** | | | | |
|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** |
| **1:** | | | | | |
| **2:** | | | | | |
| **3:** | | | | | |
| **4:** | | | | | |
| **5 :** | | | | | |

Q4. Which pavement condition indicator does your agency use?

| Y/N | **Pavement Condition Indicator** |
|---|---|
| | IRI |
| | PCI |
| | PSI |
| | Other, please enter________________________________________ |

Q5. Based on the chosen condition indicator, how much improvement in pavement condition do you think would occur (a) immediately (within 1 year) after implementing the treatment, and (b) over the pavement lifetime? For example, crack sealing is typically expected to decrease pavement IRI by 5 in/mile or 5% immediately after treatment implementation and decrease the rate of deterioration over pavement lifetime by 10%.

| **Treatments** | **How much improvement (value or % of improvement)?** | |
|---|---|---|
| | **Immediate improvement** | **Over pavement lifetime improvement** |
| 1: | | |
| 2: | | |
| 3: | | |
| 4: | | |
| 5: | | |
| 6: | | |

Q6. What is the cost per lane mile of each treatment for agency?

| **Treatments** | **Cost $ per lane mile** |
|---|---|
| 1: | |
| 2: | |
| 3: | |
| 4: | |
| 5: | |
| 6: | |

Q7. On average, what is the production rate of implementing each treatment?

| **Treatments** | **Linear or square foot per day (approximate)** |
|---|---|
| 1: | |
| 2: | |
| 3: | |
| 4: | |
| 5: | |
| 6: | |

Q8. If you are interested in engaging further with the researcher on this study and for receiving a copy of the final results, please fill in the following information.

Name:

__________________________________________________________________

State Transportation Agency (STA):

__________________________________________________________________

Role in STA:

__________________________________________________________________

Postal Address:

__________________________________________________________________

Email Address:

__________________________________________________________________

Phone Number:

__________________________________________________________________

## A.2. Questionnaire Survey 2

# <u>Prediction of Pavement Condition</u>

### Validation by Subject Matter Experts

This research aims at predicting pavement condition when the impact of preventive maintenance is considered in pavement condition prediction models.

Data for the period 1989 to 2016 were retrieved from the Long-Term Pavement Performance (LTPP) for:

- Road/pavement group: Flexible (black-topped roads that include asphalt and composite) Interstate roads
- Location: Midwestern states (Indiana, Illinois, Wisconsin, Michigan, Ohio, Minnesota, Iowa and Missouri)
- Pavement condition is expressed in the International Roughness Index (IRI, in/mi.)

<u>The range of the IRI values is:</u>

- From less than 60 in/mi. as a newly constructed pavement
- To more than 100 in/mi.

<u>Pavement condition was categorized into 5 states based on the value of the IRI:</u>

- State 1 if IRI is less than or equal to 60 in/mi.
- State 2 if IRI is greater than 60 in/mi. or less than or equal to 70 in/mi.
- State 3 if IRI is greater than 70 in/mi. or less than or equal to 80 in/mi.
- State 4 if IRI is greater than 80 in/mi. or less than or equal to 100 in/mi.
- State 5 if IRI is greater than 100 in/mi.

Past research studies related to pavement condition prediction have a common assumption: pavement condition deteriorates over pavement age and does not improve, i.e., they do not consider the effect of preventive maintenance treatments. My research study relaxes the common assumption of past research studies by considering the impact of preventive maintenance in pavement condition prediction model.

Through my research study, I was able to develop a research framework to estimate the actual times and types of preventive maintenance treatments that could have been implemented to pavement surfaces. Also, this research framework was built to incorporate the estimated times and types of preventive maintenance treatments with the typical pavement condition data (pavement condition indicator (IRI), traffic and climate factors). Finally, I developed a pavement condition prediction model that considers the effect of preventive maintenance on pavement performance.

To validate my developed model, I developed another model based on the aforementioned common assumption of past research studies, and then compared its results with that of my model.

The following pages contain graphs (Figures 1 - 5), each one displays the probability of pavement being in a specific condition at a particular pavement age. These graphs were developed under the following assumptions:

(1) Annual Average Freezing Index (representing the climate condition effect) = 772 °F days, which is the mean value obtained from the collected data,

(2) Annual Average Daily Truck Traffic (representing the traffic loading effect) = 2169 trucks, which is the mean value obtained from the collect data,

(3) Truck growth rate is assumed to be 2.8%,

(4) Application of micro-surfacing at any year from pavement age of 4 to 10 year with a probability distribution calculated from the developed research framework,

(5) Application of Ultra-Thin Bonded Wearing Course (UTBWC) at any year from pavement age of 11 to 15 year with a probability distribution calculated from the developed research framework, and

(6) Application of HMA thin overlay at any year from pavement age of 16 to 20 year with a probability distribution calculated from the developed research framework

Please review Figures 1 to 5 (probability curves) and give comments as requested.

**Pavements in state 1 (IRI ≤ 60 in/mi.)**

The probability curve represents the likelihood of pavement being in state 1 (IRI ≤ 60 in/mi.) over its life. My study indicates that at 35 years the pavement is expected to be in poor condition).

There are two curves: the green curve (S1c) represents the probabilities resulting from my research study, while the black dashed curve (S1p) represents the probabilities resulting from past research studies.

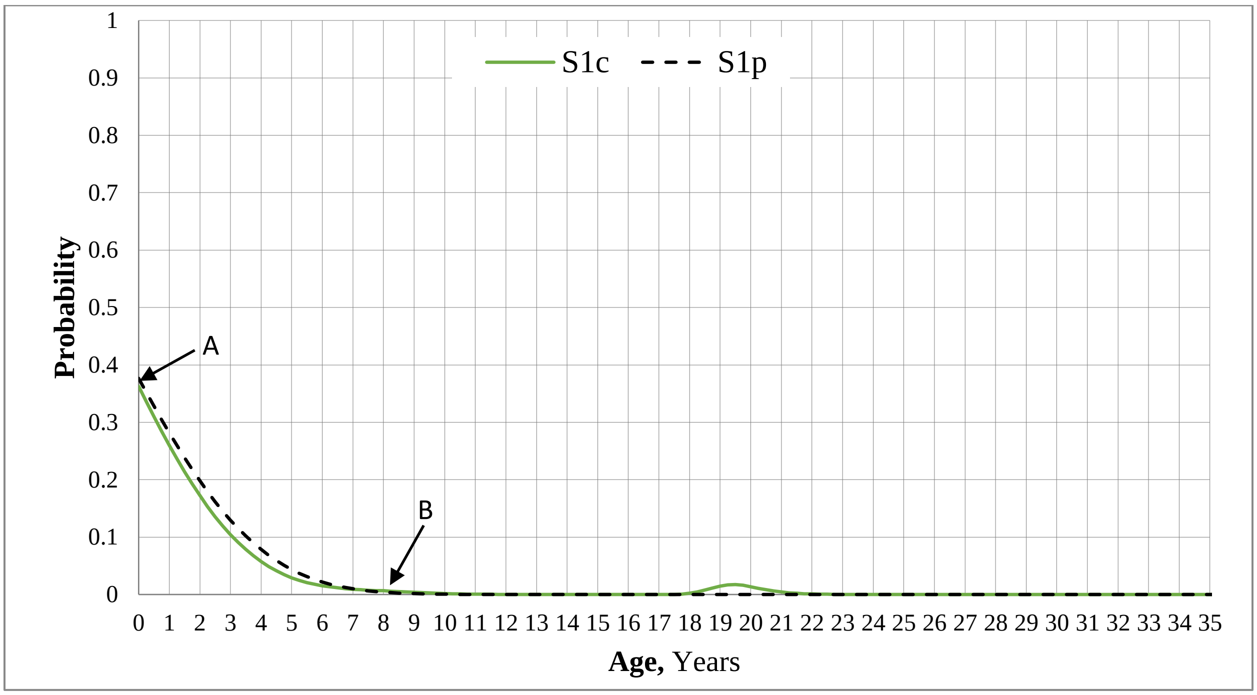



Figure A.2.1. Probability curves of pavements being in condition state 1

The probability curve reads as follows: at pavement age of zero (right after construction or major rehabilitation), there is about 38% of probability that the pavement IRI is less than or equal to 60 in/mi (see point A). At the age of 8 years there is a zero probability that the pavement IRI is less than or equal to 60 in/mi. (see point B).

Since both curves (green and black dashed) are similar, do you agree with their decreasing trend?

☐ Yes – Please provide reason

☐ No – Please provide reason

Do you agree with the values of the probabilities shown in Figure 1?

- ☐ Yes – Please provide reason

- ☐ No – Please provide reason

**Pavements in State 2 (60<IRI ≤ 70 in/mi.)**

The probability curve represents the likelihood of pavement being in state 2 (60<IRI ≤ 70 in/mi.) over its age (from 0 to 35 years).

There are two curves in Figure 2: the green curve (S2c) represents the probabilities resulting from my research study, while the black dashed curve (S2p) represents the probabilities resulting from past research studies.

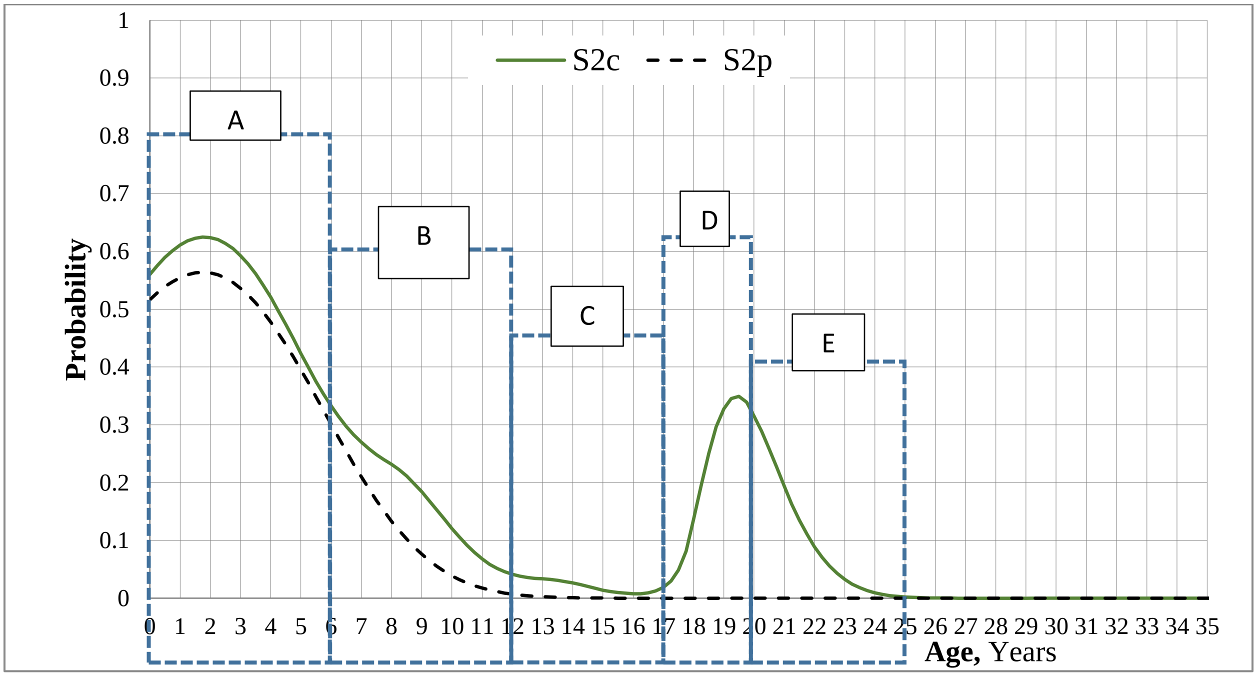



Figure A.2.2. Probability curves of pavements being in condition state 2

Note: The probability here is the likelihood of pavement IRI ranges from 60 to 70 in/mi. over time.

Both probability curves (green and black dashed) have similar trends from zero to 6 years, but after the age of 6 years the black dashed curve keeps going down with a specific rate while the green curve trend can be described as follows:

- The probability decreases with different rates from the age of 6 to 12 years due to effect of the expected application of preventive maintenance treatments such as micro-surfacing or UTBWC.

- The probability begins to increase back again at the age of 17 years due to the effect of the expected application of preventive maintenance treatments such as HMA thin overlay.
- The probability is about 35% at the age of 20 years when preventive maintenance treatments such as HMA thin overlay is most likely implemented.

On a scale from 0 to 3; where 0 refers to "strongly disagree" or "never seen in practice", 1 being "disagree", 2 being "agree", and 3 "being strongly agree" or "most similar to practice", to what extent do you agree with the trends of the green and black-dashed curves over the intervals (A, B, C, D, and E) of pavement age shown in Figure 2?

| **Intervals (years)** | **A** (0-6) | **B** (6-12) | **C** (12-17) | **D** (17-20) | **E** (20-25) |
|---|---|---|---|---|---|
| Green Curve | | | | | |
| Black-dashed Curve | | | | | |

Comments:

**Pavements in State 3 (70<IRI ≤ 80 in/mi.)**

The probability curve represents the likelihood of pavement being in state 3 (70<IRI ≤ 80 in/mi.) over its age (from 0 to 35 years).

There are two curves in Figure 3: the green curve (S3c) represents the probabilities resulting from my research study, while the black dashed curve (S3p) represents the probabilities resulting from past research studies.

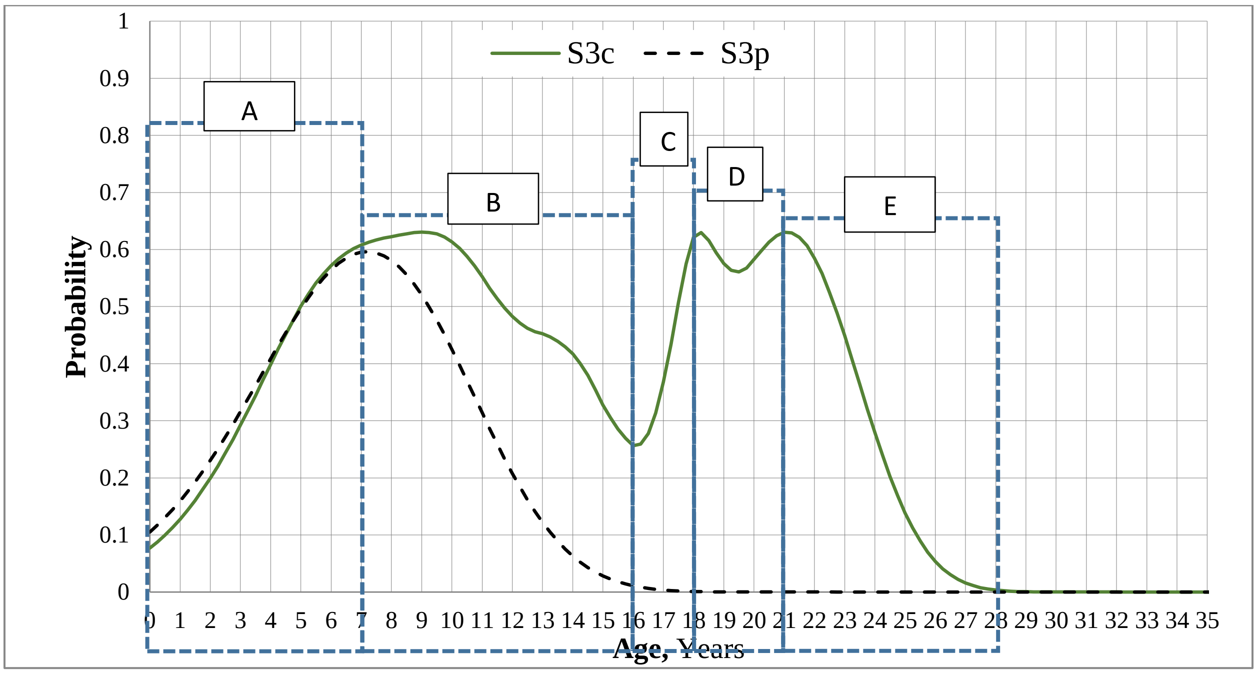



Figure A.2.3. Probability curves of pavements being in condition state 3

On a scale from 0 to 3; where 0 being "strongly disagree" or "never seen in practice", 1 being "disagree", 2 "being agree", and 3 being "strongly agree" or "most similar to practice", to what extent do you agree with the trends of the green and black-dashed curves over the intervals (A, B, C, D, and E) of pavement age shown in Figure 3?

| **Intervals (years)** | **A** (0-7) | **B** (7-16) | **C** (16-18) | **D** (18-21) | **E** (21-28) |
|---|---|---|---|---|---|
| Green Curve | | | | | |
| Black-dashed Curve | | | | | |

Comments:

**Pavements in State 4 (80<IRI ≤ 100 in/mi.)**

The probability curve represents the likelihood of pavement being in state 4 (80<IRI ≤ 100 in/mi.) over its age (from 0 to 35 years).

There are two curves in Figure A. 2. 4: the green curve (S4c) represents the probabilities resulting from my research study, while the black dashed curve (S4p) represents the probabilities resulting from past research studies.

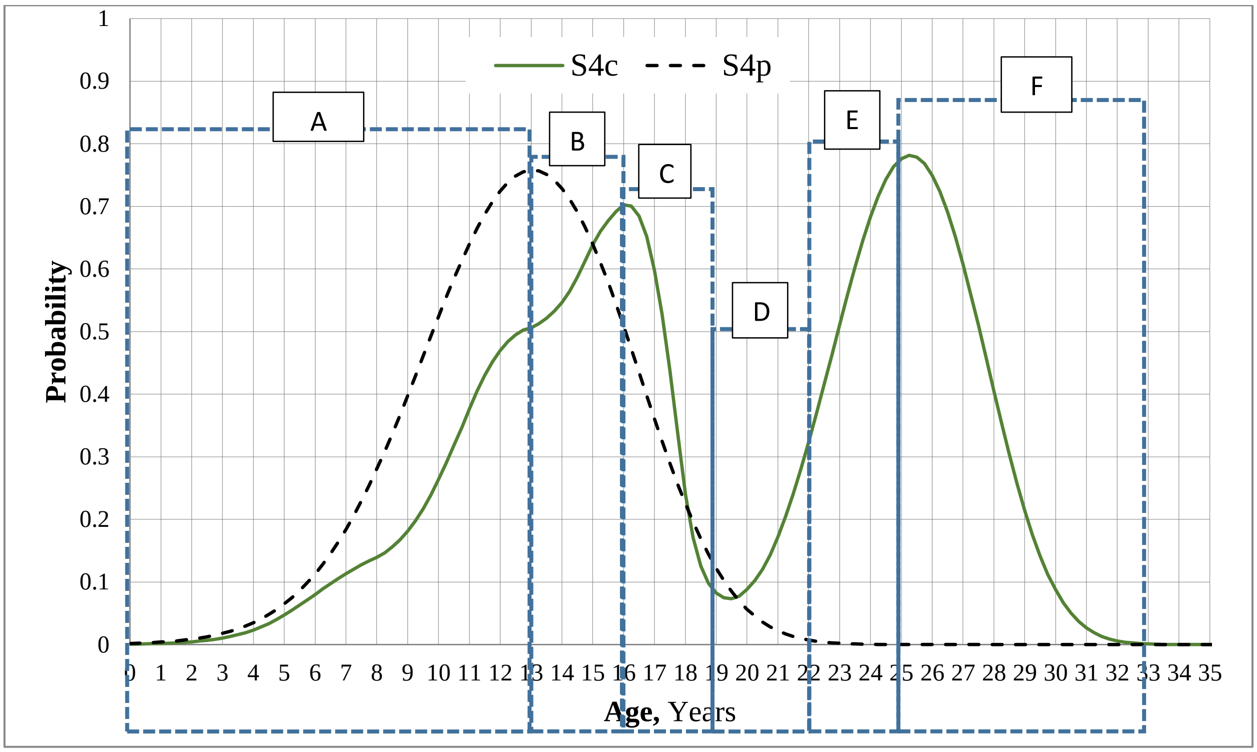



Figure A.2.4. Probability curves of pavements being in condition state 4

On a scale from 0 to 3; where 0 being "strongly disagree" or "never seen in practice", 1 being "disagree", 2 being "agree", and 3 being "strongly agree" or "most similar to practice", to what extent you are agree with the trends of the green and black-dashed curves over the intervals (A, B, C, D, E, and F) of pavement age shown in Figure 4?

| **Intervals (years)** | **A** (0-13) | **B** (13-16) | **C** (16-19) | **D** (19-22) | **E** (22-25) | **F** (25-32) |
|---|---|---|---|---|---|---|
| Green Curve | | | | | | |
| Black-dashed Curve | | | | | | |

Comments:

**Pavements in State 5 (IRI > 100 in/mi.)**

The probability curve represents the likelihood of pavement being in state 5 (IRI > 100 in/mi.) over its age (from 0 to 35 years).

There are two curves in Figure 5: the green curve (S5c) represents the probabilities resulting from my research study, while the black dashed curve (S5p) represents the probabilities resulting from past research studies.

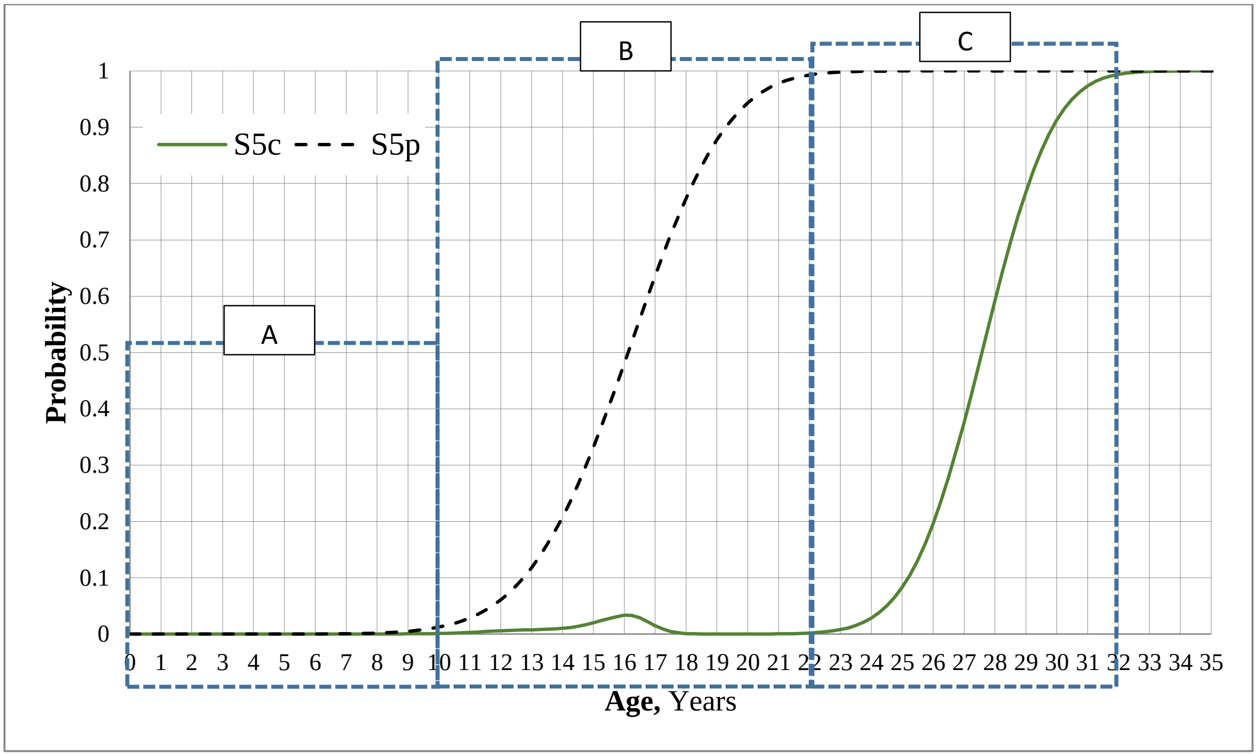



Figure A.2.5. Probability curves of pavements being in condition state 5

On a scale from 0 to 3; where 0 being "strongly disagree" or "never seen in practice", 1 being "disagree", 2 being "agree", and 3 being "strongly agree" or "most similar to practice", to what extent do you agree with the trends of the green and black-dashed curves over the intervals (A, B and C) of pavement age shown in Figure?

| **Intervals (years)** | **A** (0-10) | **B** (10-22) | **C** (22-32) |
|---|---|---|---|
| Green Curve | | | |
| Black-dashed Curve | | | |

Comments:

## A.3. Questionnaire Survey 3

## The Effect of Traffic Load on Pavement Condition

Pavement condition is expressed in the International Roughness Index (IRI, in/mi.).
Traffic load is expressed in the Annual Average Daily Truck Traffic (AADTT).

Figure 1 shows the effect of the cumulative AADTT (annual average daily truck multiplied by pavement age) on the probability of pavement being in any condition state such as state 1 (IRI less than 60 in/mi.). It shows the effect of increasing the cumulative AADTT by 10,000, which is equivalent to approximately 5 years of 2000 trucks per year.

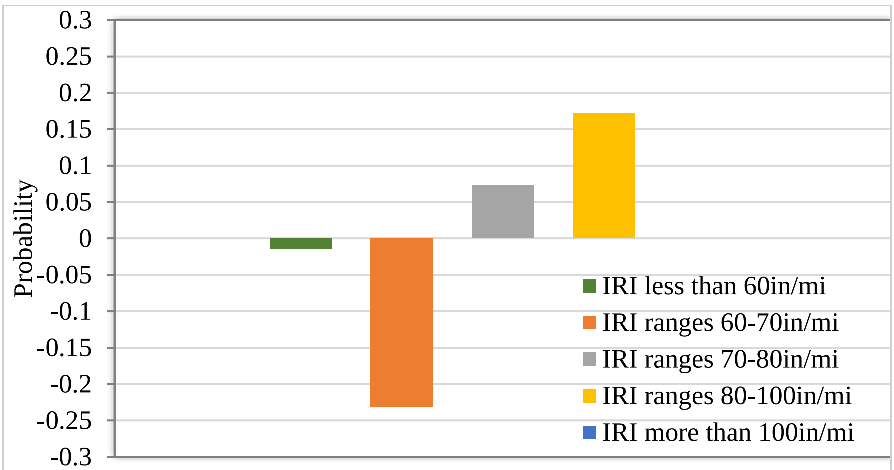



Figure A.3.1. Change in the probability of pavements in any state due to an increase in AADTT

Figure 1 shows that increasing the cumulative AADTT by 10,000 could cause the following:

- Decrease in the probability that pavement IRI is less than 60 in/mi by about 1%
- Decrease in the probability that pavement IRI ranges from 60 to 70 in/mi by about 22%
- Increases in the probability that pavement IRI ranges from 70 to 80 in/mi. by about 7%
- Increase in the probability that pavement IRI ranges from 80 to 100 by 17%

On a scale from 0 to 3; 0 being “strongly disagree” or “never seen in practice”, 1 being “disagree”, 2 being “agree”, and 3 being “strongly agree” or “most similar to practice”, to what extent do you agree with the effect of cumulative AADTT on each state of pavement condition?

| Condition state | Effect | Score (0-3) | Comment |
|---|---|---|---|
| IRI ≤ 60in/mi | -1% | | |
| 60<IRI ≤ 70in/mi | -22% | | |
| 70<IRI ≤ 80in/mi | +7% | | |
| 80<IRI ≤ 100in/mi | +17% | | |
| IRI > 100in/mi | + | | |

## **The Effect of Weather Condition on Pavement Condition**

Weather condition is expressed in the Annual Average Freezing Index (AAFI), °F days. AAFI is the cumulative number of degree-days below 32 °F during the year.

Figure 2 shows the difference in the probability of pavement being in each condition state between pavements in two regions of different AAFI, e.g., Minnesota and Indiana. The average AAFI values in 2017 were found to be: 1387 (°F days) in Minnesota, and 232 (°F days) in Indiana. So, the difference in the AAFI between these two states is about 1000 (°F days). Figure 2 shows the effect of 5,000 cumulative AAFI (5 years of 1000 (°F days) difference in AAFI) on the probability of pavement in any condition state.

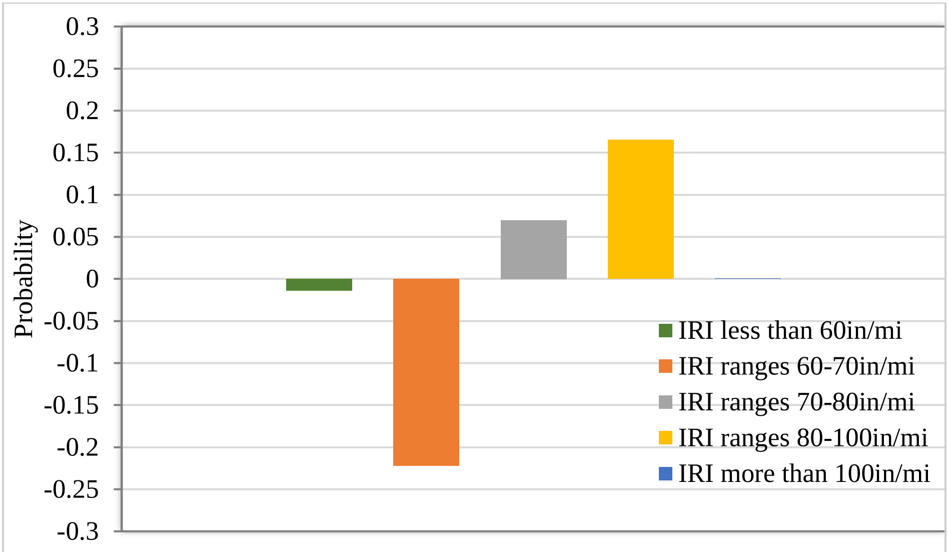



Figure A.3.2 Change in the probability of pavements in any state due to change in AAFI

For instance, Figure 2 indicates that after 5 years of 1000 (°F days) difference in the AAFI between two regions, the probability of pavements in the colder region having an IRI of 60 to 70in/mi. (i.e., being in State 1) is 22% less than that in the warmer region.

On a scale from 0 to 3; 0 being “strongly disagree” or “never seen in practice”, 1 being disagree, 2 being “agree”, and 3 being “strongly agree” or “most similar to practice” , to what extent do you agree with the effect of the cumulative AAFI on each state of pavement condition?

| Condition state | Effect | Score (0-3) | Comment |
|---|---|---|---|
| IRI ≤ 60in/mi | -1% | | |
| 60<IRI ≤ 70in/mi | -22% | | |
| 70<IRI ≤ 80in/mi | +7% | | |
| 80<IRI ≤ 100in/mi | +16% | | |
| IRI > 100in/mi | + | | |

# The Effect of Micro-surfacing Application on Pavement Condition

Figure 3 shows the effect of applying micro-surfacing to pavement surface on the probability of pavement being in any condition state such as state 1 (IRI less than 60 in/mi).

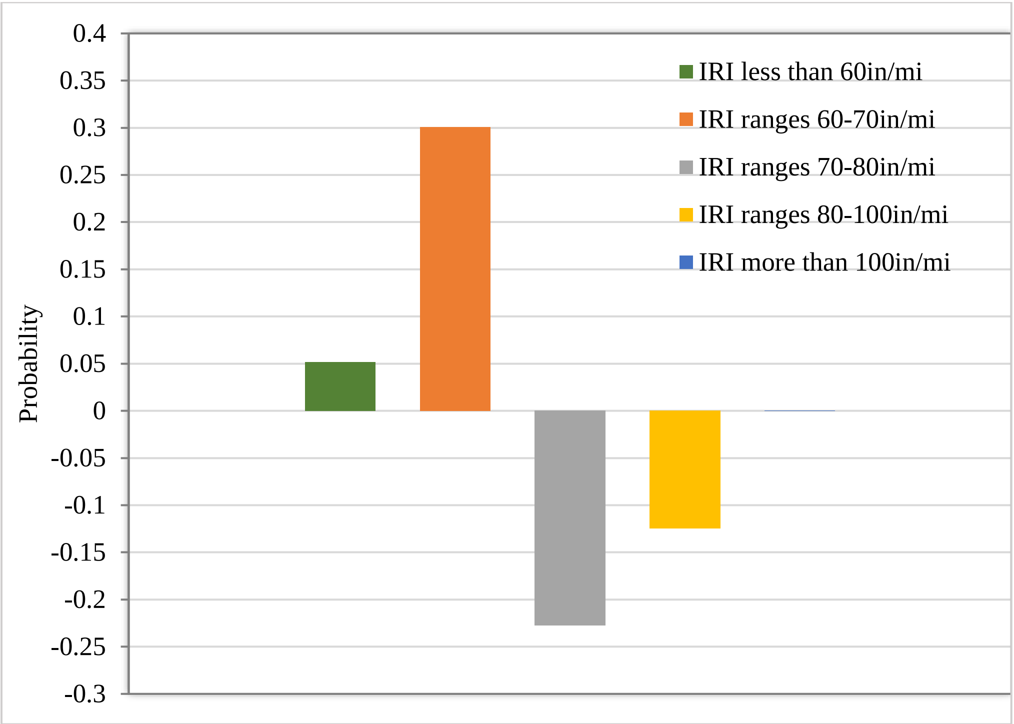



Figure A.3.3. Change in the probability of pavements in any state due to micro-surfacing application

On a scale from 0 to 3; 0 being “strongly disagree” or “never seen in practice”, 1 being “disagree”, 2 being “agree”, and 3 being “strongly agree” or “most similar to practice”, to what extent do you agree with the effect of applying micro-surfacing on each state of pavement condition?

| Condition state | Effect | Score (0-3) | Comment |
|---|---|---|---|
| IRI ≤ 60in/mi | +5% | | |
| 60<IRI ≤ 70in/mi | +30% | | |
| 70<IRI ≤ 80in/mi | -22% | | |
| 80<IRI ≤ 100in/mi | -12% | | |
| IRI > 100in/mi | - | | |

## The Effect of Ultra-Thin Bonded Wearing Course Application on Pavement Condition

Figure 4 shows the effect of applying ultra-thin bonded wearing course (UTBWC) to pavement surface on the probability of pavement being in any condition state such as state 1 (IRI less than 60 in/mi.).

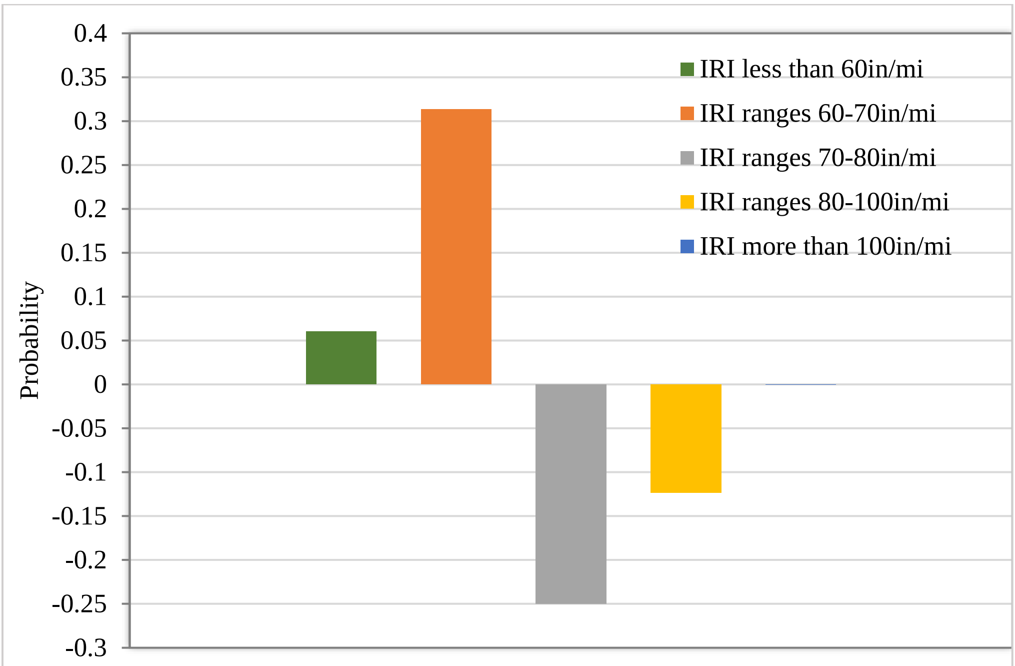



Figure A.3.4. Change in the probability of pavements in any state due to UTBWC application

On a scale from 0 to 3; 0 being “strongly disagree” or “never seen in practice”, 1 being “disagree”, 2 being “agree”, and 3 being “strongly agree” or “most similar to practice, to what extent do you agree with the effect of applying ultra-thin bonded wearing course on each state of pavement condition?

| Condition state | Effect | Score (0-3) | Comment |
|---|---|---|---|
| IRI ≤ 60in/mi | +6% | | |
| 60<IRI ≤ 70in/mi | +31% | | |
| 70<IRI ≤ 80in/mi | -25% | | |
| 80<IRI ≤ 100in/mi | -12% | | |
| IRI > 100in/mi | - | | |

## The Effect of HMA Thin Overlay Application on Pavement Condition

Figure 5 shows the effect of applying HMA thin overlay to pavement surface on the probability of pavement being in any condition state such as state 1 (IRI less than 60 in/mi).

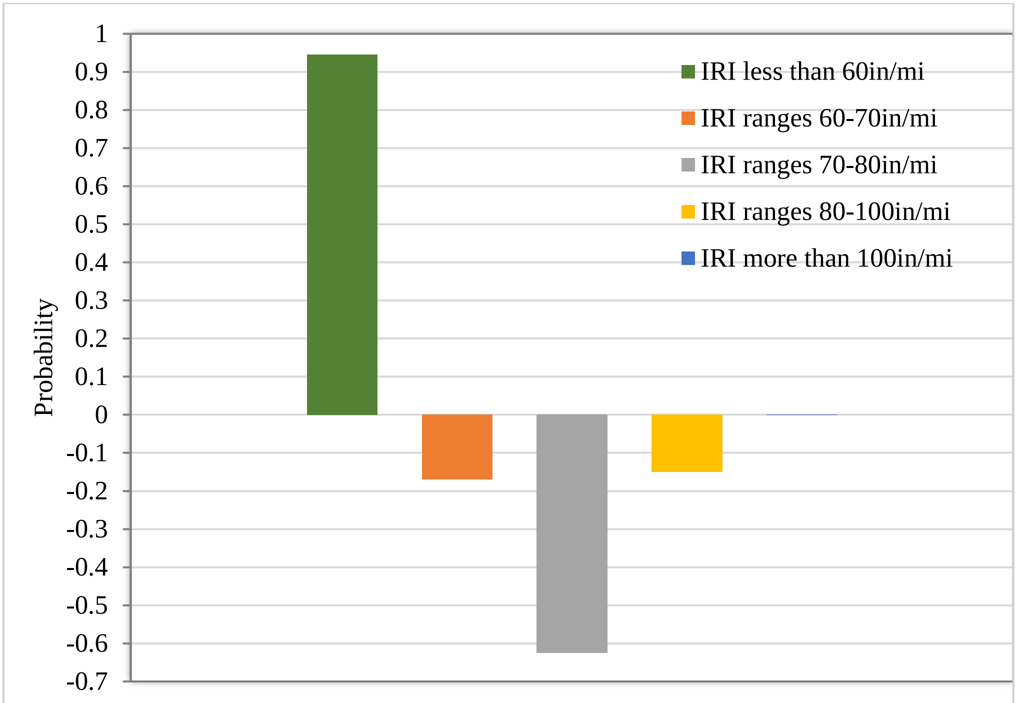



Figure A.3.5. Change in the probability of pavements in any state due to thin overlay application

On a scale from 0 to 3; 0 being “strongly disagree” or “never seen in practice”, 1 being disagree, 2 being “agree”, and 3 being “strongly agree” or “most similar to practice”, to what extent do you agree with the effect of applying HMA thin overlay on each state of pavement condition?

| Condition state | Effect | Score (0-3) | Comment |
|---|---|---|---|
| IRI ≤ 60in/mi | +95% | | |
| 60<IRI ≤ 70in/mi | -17% | | |
| 70<IRI ≤ 80in/mi | -63% | | |
| 80<IRI ≤ 100in/mi | -15% | | |
| IRI > 100in/mi | - | | |