

DEVELOPMENT OF DISTRESS AND PERFORMANCE MODELS OF COMPOSITE  
PAVEMENTS FOR PAVEMENT MANAGEMENT

by

Xiazhi Fang

A dissertation submitted to the faculty of  
The University of North Carolina at Charlotte  
in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy in  
Infrastructure and Environmental Systems

Charlotte

2017

Approved by:

---

Dr. Don Chen

---

Dr. Thomas Nicholas

---

Dr. John Hildreth

---

Dr. Jy S. Wu

---

Dr. Churlzu Lim



## ABSTRACT

XIAZHI FANG. Development of distress and performance models of composite pavements for the North Carolina Pavement Management System. (Under the direction of DR. DON CHEN)

Roadway systems in the United States have become huge assets that need massive resources to maintain and operate. To meet the long-term performance goal, government agencies developed pavement management systems (PMSs) to help them manage roadway assets effectively with limited resources. Currently, some PMSs in the United States have been designed for two types of pavements: asphalt and concrete. The composite pavement, another pavement type, which is the result of concrete pavement rehabilitations and constructed with an asphalt surface layer over a concrete base, was treated as asphalt. However, the literature review indicates that compared to asphalt pavements, composite pavements perform differently and have different dominant distresses. In addition, as the amount of composite pavements increases, it is necessary to investigate them independently to incorporate more accurate information into the PMS. Therefore, the goal of this research is to improve and to expand the PMS with an additional pavement type: composite pavements. To achieve this goal, the PMS managed by the North Carolina Department of Transportation (NCDOT) was used as a case study, and several objectives were accomplished in this research: 1) to identify composite pavements and generate the raw data based on the construction history; 2) to clean the raw data and mitigate errors using statistical methods and engineers' experiences; 3) to develop nonlinear models to describe dominant distresses and pavement performances; 4) to propose quantile regression (QR) models to predict pavement performances; and 5) to investigate the pavement treatment effectiveness by exploring performance index jumps.

Based on findings of this research, it was concluded that the automated data were more consistent with engineers' experience and revealed more information than the windshield data; longitudinal cracking and transverse cracking were found to be the dominant distresses in composite pavements, followed by alligator cracking and raveling; Interstate composite pavements deteriorated faster than both US and NC composite pavements, and NC composite pavements had the slowest deterioration rate; QR models can be used as a new prediction method of pavement performances at both the project and the network levels; in general the "Resurfacing" treatment was more effective than the "Chip Seal" treatment; and The average service life of asphalt and composite pavements were similar, but composite pavements have a smaller variation of service lives than that of asphalt pavements.

It was recommended that the automated data should be used in future PMS related research projects, due to its better data quality, and because of the robust performance of QR models at both network and project levels, QR models should be incorporated in the future PMS.

In summary, this research expanded the existing NCDOT PMS with composite pavements, proposed systematic methods to improve the quality of performance data, enriched the diversity of prediction models by exploring potentials of QR models, and investigated the effectiveness of pavement treatments. Essentially, transportation agencies can use the findings of this research to make informative investment decisions.

## TABLE OF CONTENTS

LIST OF FIGURES .....	viii
LIST OF TABLES .....	xi
CHAPTER 1 INTRODUCTION .....	1
1.1 Background of the PMS .....	1
1.2 Research Objectives .....	2
1.3 Contributions of this Study .....	4
1.4 Organization of the Dissertation .....	5
CHAPTER 2 LITERATURE REVIEW .....	8
2.1 Introduction of Composite Pavements .....	8
2.2.1 PMS Components .....	10
2.2.2 Pavement Distresses .....	14
2.2.3 Pavement Performance .....	15
2.3 Pavement Performance Models .....	16
2.3.1 Model Types .....	17
2.3.2 Performance Model Development Considerations .....	18
2.3.3 Deterministic Pavement Performance Models .....	19
2.3.4 Probabilistic Pavement Performance Models .....	20
2.4 Quantile Regression .....	22
2.4.1 Introduction .....	22
2.4.2 Applications of QR in Other Fields .....	25
CHAPTER 3 DATA PREPARATION .....	27
3.1 Overview of the Research Methodology .....	27

3.2 Data Sources.....	29
3.2.1 The Windshield and Automated Performance Data .....	30
3.2.2 Differences in Data Distributions .....	32
3.3 Data Merging Process .....	36
3.4 Data Cleansing .....	40
<b>CHAPTER 4 DEVELOPMENT OF NONLINEAR DISTRESS AND PERFORMANCE MODELS.....</b>	<b>45</b>
4.1 Distress Models .....	45
4.1.1 Calculation of Distress Index .....	45
4.1.2 Development of Distress Models .....	46
4.2 Performance Models .....	57
4.2.1 Calculation of Pavement Condition Rating (PCR).....	57
4.2.2 Development of Performance Models .....	58
<b>CHAPTER 5 DEVELOPMENT OF QUANTILE REGRESSION MODELS .....</b>	<b>63</b>
5.1 Model Development and Prediction at the Project Level .....	63
5.3 Model Development and Application at the Network Level .....	72
<b>CHAPTER 6 TREATMENT EFFECTIVENESS .....</b>	<b>77</b>
6.1 Definitions of performance jumps.....	77
6.2 Analyses of Descriptive statistics and Visualizations .....	79
6.3 Comparison of Treatment Effectiveness .....	85
6.4 Comparison of pavement performance between Asphalt and composite pavements.....	90
<b>CHAPTER 7 CONCLUSIONS, LIMITATIONS AND RECOMMENDATIONS .....</b>	<b>93</b>
7.1 Conclusions .....	93
7.2 Limitations of research.....	95

7.3 Recommendations .....	96
BIBLIOGRAPHY .....	97
APPENDIX A QUANTILE REGRESSION MODELS FOR PAVEMENT FAMILIES AT THE PROJECT LEVEL .....	104
APPENDIX B QUANTILE REGRESSION MODEL FOR PAVEMENT CLASSIFICATIONS AT THE NETWORK LEVEL .....	107

## LIST OF FIGURES

Figure 1: General structure of the PMS .....	13
Figure 2: Research methodology .....	28
Figure 3: Alligator cracking in the windshield data.....	33
Figure 4: Alligator cracking in the automated data .....	34
Figure 5: Raveling in the windshield data .....	35
Figure 6: Raveling in the automated data .....	35
Figure 7: The flow chart of data merging process .....	36
Figure 8: Comparing mile posts to merge two data sets.....	38
Figure 9: A data error example .....	41
Figure 10: An example of performance jumps .....	41
Figure 11: Sample sizes of pavement families .....	43
Figure 12: The boxplot of transverse cracking index in US_0-5K.....	43
Figure 13: Boxplots of distresses of US_0-5K (i) .....	46
Figure 14: Boxplots of distresses of US_0-5K (ii) .....	47
Figure 15: Distress models for transverse cracking.....	49
Figure 16: Distress models for longitudinal cracking.....	50
Figure 17: Distress models for alligator cracking.....	50
Figure 18: Distress models for raveling.....	51
Figure 19: Transverse cracking models for pavement families .....	53
Figure 20: Longitudinal Cracking for pavement families.....	53
Figure 21: Alligator cracking for pavement families.....	54
Figure 22: Raveling for pavement families .....	54



Figure 23: Distress models for Interstate .....	55
Figure 24: Distress models for US_0-5K.....	55
Figure 25: Distress models for US_5K plus .....	56
Figure 26: Distress models for NC .....	56
Figure 27: Performance model of Interstate .....	58
Figure 28: Performance model of US_0-5K plus .....	59
Figure 29: Performance model of US_5K plus.....	59
Figure 30: Performance model of NC.....	60
Figure 31: Performance models for pavement families.....	61
Figure 32: Process of QR model development and prediction at the project level.....	64
Figure 33: A hypothetical example of using QR models at the project level .....	69
Figure 34: QR model for US_0-5K plus at the project level .....	69
Figure 35: A hypothetical example of using QR models at the network level .....	73
Figure 36: Interstate QR model at the network level.....	74
Figure 37: US 0-5K QR model at the network level .....	74
Figure 38: US 5K plus QR model at the network level .....	75
Figure 39: NC QR model at the network level .....	75
Figure 40: Calculation of a jump .....	79
Figure 41: The histogram of jumps.....	80
Figure 42: The distribution of pre-treatment PCRs .....	81
Figure 43: The accumulative curve of the pre-treatment PCR .....	82
Figure 44: The histogram of pre-treatment ages.....	83
Figure 45: Accumulative curve of pre-treatment ag2 .....	84

Figure 46: Parallel coordinate plot.....	84
Figure 47: The boxplot of treatment effectiveness based on treatment types.....	86
Figure 48: The boxplot of treatment effectiveness based on pavement families.....	87
Figure 49: The boxplots of jumps and real jumps .....	90

## LIST OF TABLES

Table 1: Comparison of the windshield and the automated data in North Carolina.....	31
Table 2: Parameters of performance models .....	60
Table 3: Parameters of random effect performance models .....	62
Table 4: In-sample and out-sample errors .....	70
Table 5: Prediction errors (out-sample) with different intervals .....	71
Table 6: Descriptive statistics for jumps.....	80
Table 7: Descriptive statistics for pre-treatment PCRs.....	81
Table 8: Descriptive statistics for pre-treatment ages.....	83
Table 9: Intercepts of performance curves.....	89
Table 10: The result of t-test for “real_age” .....	91

## CHAPTER 1 INTRODUCTION

Currently, some pavement management systems (PMSs) in the United States have been designed for two types of pavements: asphalt and concrete. The goal of this research is to improve and to extend the PMS with an additional pavement type: composite pavements.

### 1.1 Background of the PMS

The PMS is an organized procedure constructed upon statistical and mathematical methodologies to effectively and economically manage and maintain roadways. The concept of pavement management began in the 1970s when the demand to supervise a tremendous amount of roadways increased significantly. According to the United States Department of Transportation (USDOT), early in 1980, the estimated roadway lane-miles had reached eight million miles, which was more than 90% of the total roadway lane-miles that have been constructed by 2013 (USDOT, 2013).

In 1980, the first national workshop on pavement management was held to disseminate the concept of pavement management and to help states comprehend the goals of pavement management. In this workshop, PMS was first defined as “a system which involves the identification of optimum strategies at various management levels and maintain pavements at an adequate level of serviceability. These include, but are not limited to, systematic procedures for scheduling maintenance and rehabilitation activities based on optimization of benefits and minimization of costs” (FHWA, 1997). Since then,

an increasing number of organizations have collaboratively promoted the implementation of PMS, such as the Federal Highway Administration (FHWA), the Transportation Research Board (TRB), the American Association of State Highway and Transportation Officials (AASHTO), and the National Highway Institute (NHI). Due to these efforts, more states have developed their own PMSs to assist the decision-making processes and to manage their roadways.

Roadway systems in the United States have become huge assets that need massive resources to maintain and to operate. By the 1990s, over one trillion dollars had been invested to the nation's highway and bridge systems, and over 62 billion was devoted annually to physical preservations and operational improvements (FHWA, 1997). In the 21st century, highway officials are facing the challenges of managing "an ever-expanding, still-evolving, yet aging highway network" (FHWA, 1998). To address these challenges, it is essential for government agencies to keep improving their PMSs to help them effectively manage roadway assets with limited resources.

## 1.2 Research Objectives

Some state PMSs were developed for two types of pavements: asphalt and concrete. Composite pavements have been historically considered as asphalt by these PMSs because they have an asphalt surface layer. However, according to the literature review, composite pavements perform differently than asphalt pavements and have different dominant distresses. In addition, as the amount of composite pavements increases, it is necessary to investigate them independently to incorporate more accurate information into the PMS.

Therefore, the goal of this research is to improve and to extend the PMS with an additional pavement type: composite pavements. In this study, the North Carolina Department of Transportation (NCDOT) PMS was used as a case study, and composite pavements refer to pavements built with a relatively thin asphalt surface layer over a concrete base, typically they are results of concrete pavement rehabilitations. To achieve the goal, research expects to:

- (1) Identify composite pavements and generate the raw data according to the construction history. Since composite pavements have been considered as asphalt pavements by NCDOT, performance records of composite pavements can be extracted from the asphalt database by investigating when original asphalt pavements were converted to composite.
- (2) Clean the data and mitigate errors using statistical methods and engineers' experiences. High quality data is fundamental for a successful pavement management analysis. Since the pavement condition data was collected in the field, it is inevitable that the raw data contains certain subjective errors or measuring mistakes.
- (3) Develop nonlinear models to describe dominant distresses and pavement performances. Due to the different pavement structures, the processes of distress development and performance deterioration of composite pavements can be different from conventional asphalt and concrete pavements. A better

understanding of these processes will be very important to the preparation of future maintenance and rehabilitation strategies.

- (4) Propose Quantile Regression (QR) models to predict pavement performances. In addition to developing nonlinear models, this study proposes QR models to predict pavement performances. It aims to establish a new form of pavement performance models and to improve the accuracy of predictions.
- (5) Investigate the effectiveness of pavement treatments by exploring performance condition jumps. Pavement performances are improved after different types of treatments. To quantify the improvement, condition jumps are calculated for each type of treatment.

### 1.3 Contributions of this Study

This research will benefit the pavement management agencies through the improvement and extension of the PMS. Contributions of this research include:

- (1) Extended the current PMS with a new type of pavement, composite pavements, and developed a new set of distress and performance models for composite pavements. These performance models will allow more accurate descriptions of pavement performance, assist engineers to select more appropriate treatments, and help engineers make more reasonable investment decisions.
- (2) Explored the potentials of QR models in managing pavements. The literature review did not reveal any previous studies related to similar applications. Compared to existing deterministic pavement performance models, QR models have several

advantages due to their nonparametric characteristics, including the ability to reduce noises from outliers, and the capability to present multi-dimensional information. Compared to existing probabilistic pavement performance models, QR models require less computational resources and can handle continuous performance ratings.

- (3) Confirmed the NCDOT maintenance thresholds. The average PCR value and the average age when pavement sections have been treated are 61.9 and 12.7, respectively. These findings are consistent with NCDOT maintenance practices.

#### 1.4 Organization of the Dissertation

This dissertation includes literature review, data preparation, nonlinear distress and performance model development, QR model development, treatment effectiveness, and conclusions, limitations and recommendations.

Chapter 1 introduces the overall background of the research. The research goal and objectives are presented.

Chapter 2 provides a literature review for the research. The background of composite pavements is introduced. The PMS is defined in terms of its structures and significant components. Additionally, the currently existing performance models, deterministic and probabilistic, are reviewed. Lastly, the definition of QR models is introduced, along with its advantages and applications.



Chapter 3 presents the sources of data used in this research. It introduces the research methodology, including data preparation, model development, and the investigation of treatment effectiveness. In addition, available data sources are described along with their characteristics. Data sets are generated based on appropriate statistical methods and experiential judgements. Detail approaches to identify composite pavement sections and to clean the raw data are elaborated.

Chapter 4 presents the development of nonlinear distress and performance models. The sigmoidal model form is used to develop both distress and performance models. Dominant distresses of composite pavements are discussed. Representative pavement sections are selected using several screening conditions in order to obtain more reliable and precise data sets, which are then used to develop distress and performance models.

Chapter 5 describes the development of QR models. In this research, QR models are developed to predict pavement performances at both project and network levels. At the project level, the accuracy of predictions are evaluated using the Mean Absolute Percentage Error (MAPE). At the network level, QR models are developed presenting five curves at different distribution locations using representative pavement sections.

In Chapter 6, the effectiveness of pavement treatments are evaluated. Performance jumps resulting from treatments are calculated to quantify the improvement in the pavement condition. The treatment effectiveness of different treatment types is discussed using data visualization and descriptive statistics.

Chapter 7 summarizes conclusions, discusses limitations of this research, and provides recommendations for future studies.

## CHAPTER 2 LITERATURE REVIEW

This chapter reviews composite pavements, PMSs, pavement distresses, pavement performance models, and quantile regression (QR) models. Two different types of performance models are discussed, deterministic and probabilistic, followed by considerations and assumptions of these models. At the end, the algorithm and the effectiveness of QR models are reviewed.

### 2.1 Introduction of Composite Pavements

Determined by the Committee on Composite Pavement Design of the Highway Research Board, the official definition of composite pavements is: “A structure comprising multiple, structurally significant, layers of different, sometimes heterogeneous composition. Two layers or more must employ dissimilar, manufactured binding agents” (Smith, 1963).

In the United States, the majority of existing composite pavements are the result of concrete pavement rehabilitations that construct hot mixed asphalt (HMA) layers on top of concrete bases (Flintsch et al., 2008; FHWA, 2016). New composite roadways have also been constructed since the 1950s by various states and local highway agencies, such as the states of New Jersey and Washington and the cities of New York, Washington, D. C., and Columbus, Ohio (Rao, 2012). Worldwide, composite pavements have been built in the last few decades, especially in European countries. These countries have been constructing composite pavements on major roadways for several decades and have achieved substantial

experience (Hassan et al., 2008; Rao, 2013). Countries such as Germany, France, and Spain have built 30% to 50% of their main road networks using long-life, semi-rigid (composite) structures (Thogersen et al., 2004; Flintsch et al., 2008).

In 2006, Merrill et al. (2006) reported that composite pavements constructed in the U.K., the Netherlands, and Hungary performed satisfactorily in terms of rutting, cracking, and deflection. Additionally, compared to asphalt pavements, composite pavements tended to have longer lives. Similarly, in the United State, the FHWA Zero Maintenance Pavement Study Composite identified composite pavements as one of the most promising low-maintenance pavements (Darter & Barenberg, 1976; Rao et al., 2003). More advantages were discovered by Rao et al. (2012). In their study, they listed various situations when composite pavements were the optimal solution and summarized those advantages as lower life-cycle cost, rapid renewal, sustainable concrete pavement treatment solution, noise reduction with flexible surface layer, and combination of the structural capacity of concrete and the functional characteristics of asphalt surfacing.

Similarly, in 2008, Flintsch et al. (2008) indicated that composite pavements were able to reduce both structural and functional problems that typical flexible or rigid pavements possess. Additionally, based on the result of the deterministic agent-cost life cycle cost analysis, they concluded that a composite pavement with cement-treated base is a cost-effective alternative for a typical interstate highway, and a composite pavement with a continuously reinforced concrete pavement base may be a cost-effective option for highways with very high traffic volumes.

## 2.2 Introduction of the PMS

Based on the different pavement management applications, a PMS can be classified into two levels: network and project. The PMS at the network level concentrates on the selection of optimal maintenance and rehabilitation strategies, and the effective allocation of constraint resources. At the project level, the PMS focuses on specific pavement projects for their designs, life-cycle cost analyses, and selections of optimal alternative treatments.

### 2.2.1 PMS Components

A typical PMS consists of three subsystems: information, analysis, and implementation (Hudson, et al., 1979).

#### 1. The information subsystem

The information subsystem plays a fundamental role in a PMS, which includes the information of pavement inventories, performance conditions, treatment histories, traffic loads, and costs. Pavement condition data usually includes different types of distresses, roughness, rutting, skid resistance, and structural capacity (Vitillo, n. d.). The time intervals and methods of data collection vary among different states and different types of pavements. For example, In North Carolina, performance data for Interstate pavements are collected annually, whereas US and NC highways are surveyed every two years. Prior to 2010, pavement condition data was collected using the windshield method, and the automated data collection method has been used since then.

Windshield data, also known as manual data, are collected manually by raters driving a vehicle at a low-speed (15-20 miles per hour) or by walking on the pavements.

The windshield survey has several issues, such as the high rating subjectivity, the intensive labor force, and the safety concerns (Flinstch & Bryant, 2006; Wang et al., 2010; NCDOT, 2011). As technology develops, automated data collection methods, which use laser sensors, cameras, or three-D laser scanning to capture distress information, have been invented and widely adopted (Ong et al., 2010). Data collected using these methods are referred to as automated data.

In 2007, Chang Albitres et al. conducted a thorough literature review of automated pavement surveys and concluded that compared to the windshield data, the automated data had the following characteristics: 1) it included higher performance index values, 2) it was difficult to recognize distresses at low-severity level during automated surveys, 3) it had a large variance in repeating survey data, and 4) the datasets provided by different vendors could be significantly different. In their study, the ground truth data was created by calculating the averages of three windshield data sets that were collected by different raters for the same roadway sections. Then, the automated data was compared to the ground truth data to evaluate the qualities of the data sets. They also concluded that it was easier for automated surveys to capture transverse and alligator cracking distress, than to capture distresses such as raveling and rutting.

In 2010, Wang et al. introduced the new laser-imaging technology for automated surveys. This technology was able to collect data at a high accuracy level by taking the advantage of “shadow-free images”. They indicated that windshield data had larger variations due to “complex pavement condition, varying data collection methods, rater

inconsistencies, interrater uniformity, time and transcription, referencing, and data entry.” Also they concluded that automated surveys were able to obtain data as precise as windshield surveys and overcame the inconsistency and the inefficiency issues that the windshield data had.

## 2. The analysis subsystem

The analysis subsystem provides a variety of methods to interpret pavement performances and to identify cost-effective treatments and strategies. Treatments are actions to correct specific pavement deficiencies, whereas strategies are combinations of treatments to maintain pavement networks at an acceptable condition level over a reasonable time period. With an increasing extent of analyses, these methods can be grouped as pavement condition analyses, priority assessment models, and network optimization models. The main functions of pavement condition analyses are to aggregate different types of distresses into a composite index and to develop models to describe distress development and performance deterioration processes. Priority assessment models are used to generate a prioritized list of projects, which can be completed in two steps. The first step determines the maintenance strategy for an individual project based on the prediction of future conditions, and the second step develops the priority list according to the cost-benefit analysis results or other measures of cost effectiveness. Network optimization models are used in a similar approach as the priority assessment model, but the main difference is that the subject of network optimization models is at the network

level, and the purpose is to identify an optimally balanced strategy to meet budget and policy constraints.

### 3. The implementation subsystem

The implementation subsystem is used to verify the reliability of a PMS, including the effectiveness of the information subsystem, the analysis subsystem, and the strategies generated based on those efforts.

Figure 1 summarizes the structure of the PMS.

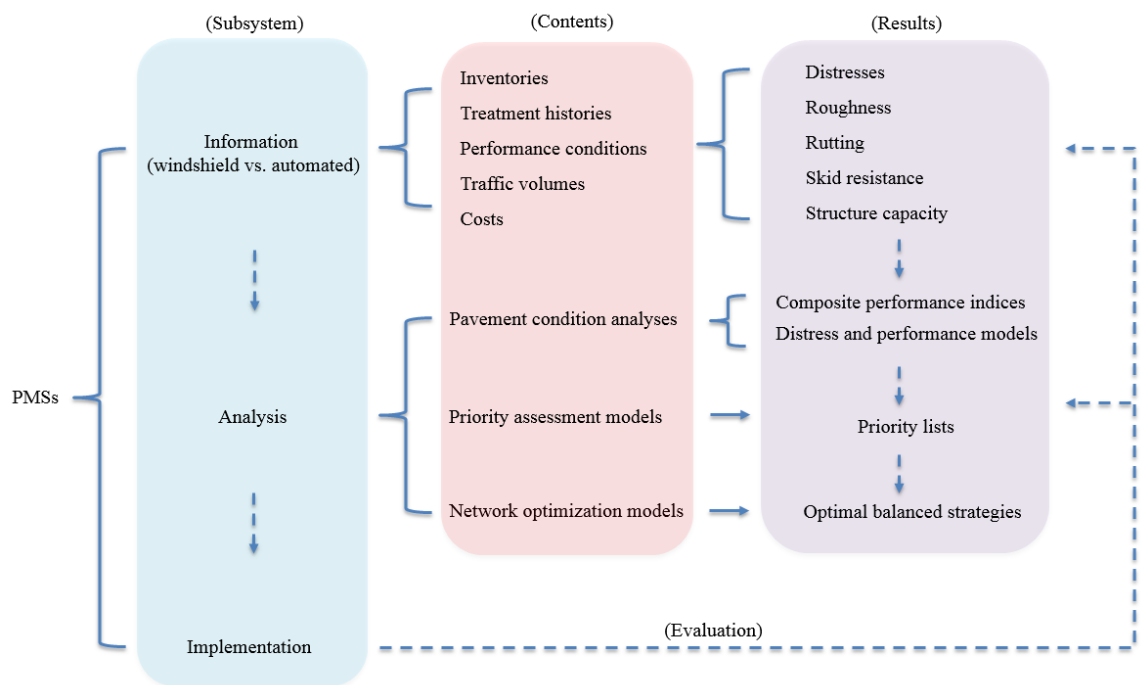


Figure 1: General structure of the PMS



### 2.2.2 Pavement Distresses

Distresses in composite pavements are similar to those in flexible pavements due to the same materials being used as the top layer, and all of these distresses could potentially affect the performance and the structural capacity of composite pavements (Von Quintus et al., 1979; Flintsch et al., 2008). Several studies indicated that reflective cracking was the major distress type for composite pavements (Von Quintus, 1979; NCHRP, 2004). Reflective cracks were caused by the relative horizontal and vertical movements of cracks or joins in the underlying rigid layer (Flintsch et al., 2008). In addition to reflective cracking, due to the similarity of distresses between composite and asphalt pavements, other distresses found in asphalt pavements, such as transverse cracking, longitudinal cracking, raveling, alligator cracking, and rutting, also occur in composite pavements. Different states have slightly different definitions of distresses, which affect the interpretation of the distress data in the analysis. In this study, the definitions of distresses that used by the NCDOT are provided below.

Cracks that are perpendicular to the pavement centerlines and are not over the joins of underlying concrete layers are defined as transverse cracks (NCDOT, 2011). Transverse cracking is also referred to as thermal cracking because it is generally caused by a sharp temperature drop or repeated temperature cycling. Longitudinal cracks are cracks that are parallel to the pavement centerlines and are not over the joins of underlying concrete layers. In North Carolina, cracks that are developed outside the wheel paths are categorized as longitudinal cracks. Cracks that are developed inside the wheel paths are categorized as

alligator cracks at the low-severity level. Therefore, longitudinal cracking in North Carolina are considered as one of non-load related distresses. Another non-load related distress is raveling, which is the result of the loss of aggregate particles in the asphalt layers due to the lost bond between aggregates and asphalt binders. Water collected around raveling can cause safety hazards and severe traffic accidents.

Alligator cracking is the primary structural distress in asphalt pavements and is categorized as load related distress, according to NCDOT. The early stage of alligator cracking is shown as longitudinal cracking in the wheel paths, and then, these cracks progress into a network of interconnecting cracks, which look like the skin of the alligator (Miller & Bellinger, 2003). The main cause of alligator cracking is heavy traffic loads applied to the pavements. Similarly, rutting, another load related distress, is also caused by heavy traffic loads, which appears as longitudinal depressions in the wheel paths.

### 2.2.3 Pavement Performance

Pavement performance is defined as the ability of a pavement to serve the public as designed (Yoder and Witczak, 1975). Many factors, such as traffic volumes, pavement designs (material and layer thickness), maintenance levels, and environmental associated factors, can impact pavement performance after construction or treatments. Pavement performance can be evaluated by distresses, structural capacities, friction measurements, and roughness (Gramling, 1994). The Pavement Condition Rating (PCR), one of indices designed to represent the overall pavement performance, is calculated by combining effects of various types of distresses, their severities, and sizes (Highway Preservation Systems,

2001). PCR has been widely used by various state agencies but may be calculated using different formulas. In North Carolina, PCR ranges from 0 to 100, is a composite performance index, and is calculated by considering all types of distresses that carry different weights.

In addition to PCR, the Present Serviceability Index (PSI), the Pavement Condition Index (PCI), and the Pavement Quality Index (PQI) are the other commonly used performance indices. PSI uses pavement roughness to estimate pavement performance (Yoder & Witczak, 1975; Roberts et al., 1991) and was developed for the American Association of State Highway Officials (AASHO) Road Test to measure ride quality using the longitudinal profile variation data (Sun, 2001). PCI was developed by the US Army Corps of Engineers, and it is calculated based on the amounts and types of distresses in pavements. PCI is a numerical value that is between 0 and 100. Different from the three indices discussed above, PQI uses both the present ride-ability of pavements and the future pavement deterioration to estimate pavement performance. It is an ordinal measurement, which ranges from 2 to 10 with 2 representing the poorest performance and 10 representing the best (Lashlee, 2004).

### 2.3 Pavement Performance Models

Pavement performance models are the fundamental component of the analysis subsystem. They are used to predict future performances and to help agencies generate the priority list of pavements that need treatments. Eventually, treatment strategies can be developed using the priority list, the importance of listed pavement functions and the

estimated costs of treatments. Accurate pavement performance models can result in more effective and efficient pavement management strategies (Prozzi & Madanat, 2004).

### 2.3.1 Model Types

Performance models can be categorized into three types based on the development methods: mechanistic, empirical, and mechanistic-empirical. Mechanistic models are developed using physical principles, such as soil mechanistic theories, mechanical properties of pavement materials, and multilayer structural conditions (Li, 2005). Empirical models are developed by statistical techniques with explanatory variables. Mechanistic-empirical models incorporate the advantages of both abovementioned methods. In fact, there is no absolute boundary between mechanistic and empirical methods because all mechanistic models include some empirical elements, while empirical models also reflect some mechanistic principles (Li, 2005). In practice, empirical and mechanistic-empirical methods have been most commonly used by agencies due to the large amount of available data (Li, 2005; Lytton, 1987).

Performance models can also be classified into two categories based on the results: deterministic and probabilistic. The deterministic model provides a single predicted value as the pavement performance prediction, whereas the probabilistic one offers multiple predicted values with corresponding probabilities. Typically, probabilistic models are more useful for the PMS at the network level, while deterministic models are more valuable at the project level (Lytton, 1987).

### 2.3.2 Performance Model Development Considerations

Several considerations most frequently discussed by various researchers are introduced in this section. They are latent performance, heterogeneity, and serial dependence (Chun & Durango-Cohen, 2008).

Latent performance refers to the fact that condition ratings are not able to accurately represent the true deterioration of infrastructure facilities. Deterioration itself is unobservable, while the only observable characteristics of deterioration are distresses that resulted from deterioration processes (Ben-Akiva & Ramaswamy, 1993). Therefore, condition ratings could not be directly linked to the facility deterioration process (Madanat et al., 1995). Since it is impossible to observe the deterioration process, all researchers still use condition ratings as the indicator of pavement performance. Latent performance was used by researchers (Chun & Durango-Cohen, 2008; Li, 2005) to develop probabilistic models.

Heterogeneity refers to the assumption that the deterioration process of each infrastructure facility is different because of some “unobservable factors and facility-specific characteristics” (Madanat et al., 1995). The uncorrected heterogeneity could cause a biased estimation of coefficients in the model. Most panel data sets, especially for those with large amounts, suffer from this issue (Heckman, 1981a; Prozzi & Madanat, 2004). Panel data is a type of data set that is collected over time from many sections. Since it is impossible to generate an individual model for each roadway section, pavement

performance is considered probabilistic because of the existence of heterogeneity (Abaza et al, 2001).

### 2.3.3 Deterministic Pavement Performance Models

Deterministic pavement performance models are models that provide a single prediction value of pavement performance. The original performance model was developed using the nonlinear model to capture the lost value in PSI for flexible pavements by AASHO. The explanatory variables in this model include cumulative traffic loads and the structural designs of pavements (HRB, 1962). In 1984, based on the AASHO model, Garcia-Diaz and Riggins (1984) proposed a methodology that used an S-shaped curve to capture the changes of deterioration rates as the traffic level accumulates. The AASHO and various S-shaped curve models have limitations, such as the failure of considering uncertainties in pavement behaviors under uncertain traffic and weather conditions, difficulties in identifying the comprehensive substantial factors and parameters, measure errors, and subjective obstacles in collecting condition data (Li et al 1996). However, these models are still widely used by state agencies due to their simplicity for implementation.

To improve the original AASHO model, in 2004, Prozzi and Madanat (2004) proposed a joint estimation method that combined the experimental and the field data to include the realistic long-term traffic impact on pavement performances. In addition, they added a more detailed specification of each parameter and variable for the original model, incorporated a new environmental variable and adopted random effects in their model to account for the heterogeneity.

#### 2.3.4 Probabilistic Pavement Performance Models

Probabilistic pavement performance models predict multiple pavement performance rating values with corresponding probabilities. These models can reveal the uncertainties of pavement conditions, which were caused by unobserved explanatory variables, presence of measurement errors, and inherent stochasticity of the deterioration processes (Madanat et al., 1995). The Markov transition process was the most popular probabilistic model to be used for predicting infrastructure performance in the 1990s, which forecasts the transition probabilities of the performance of an infrastructure deteriorating from one condition status to another (Madanat et al., 1995; Li et al, 1996).

Based on different statistical assumptions, transition probability matrices (TPM) used in the Markov process can be classified into two different types: homogeneous and nonhomogeneous. Homogeneous TPM has two assumptions. Firstly, pavements at different condition states have the same probabilities of transferring to a lower state. Secondly, the present condition state is related only to the preceding state, which means the resulting Markov model has no memory of the entire condition history (Golabi et al., 1982; Hass, 1997; Chen & Mastin, 2014). Different from homogeneous TPM, nonhomogeneous TPM assumes that deterioration probabilities at different states are different, which can capture additional information (especially the changes in traffic and weather factors). However, compared to homogeneous TPM, nonhomogeneous TPM can produce less accurate results (Li et al, 1996).

The Markov transition methodology discussed above is state-based, which predicts the probabilities that facilities deteriorate from one condition state to another during a data-collection cycle (typically, one or two years). Later, the time-based TPM was developed by DeStefano and Grivas (1998) and Mishalani and Madanat (2002). Compared to state-based models, time-based models can be used to develop the probability density function of the time that the facility takes to leave a particular condition state for a lower state (Mishalani and Madanat, 2002). Even though these models overcome some limitations of state-based models, time-based models require more accurate and frequent observations of performance data during the deterioration period, which is not easy to obtain (Li's dissertation (2015)).

In addition to the Markov transition process, econometric models have also been widely used to calculate transition probabilities. Madanat et al. (1995) used ordered logistic analysis to develop incremental models for each discrete condition state. The models explicitly interpreted the deterioration process as the function of the exogenous variables and, in turn, were used to calculate the transition probabilities for the Markov transition process. In their research, discrete condition ratings were used to reduce computational complexities, and the incremental model was used to capture the nonstationary structure of deterioration process and predict changes in conditions over time.

In 1997, Madanat et al. (1997) incorporated a random-effect factor in their logistic model to control the heterogeneity existing in infrastructure facilities, which cannot be captured by the existing explanatory variables. The study revealed that the random-effect



logistic model was better than the ordinary model in terms of accuracy of predicted results. Additionally, they also concluded that the deterioration process was dependent on infrastructures' condition histories (Madanat & Karlaftis, 1997).

In summary, pavement performance models, including the deterministic and the probabilistic performance models, play a significant role in PMSs. Even though deterministic models cannot account for uncertainties in predictions as probabilities models do, deterministic models are still frequently used by transportation agencies because of their simplicity and convenience in understanding and implementation into PMSs (Wolters & Zimmerman, 2010; Li et al., 2006).

## 2.4 Quantile Regression

Quantile regression (QR) was first introduced by Koenker and Basset (1978) as an extension of conventional regression models. In the following sections, basic concepts of QR are introduced, and the applications of QR in multiple fields are also reviewed.

### 2.4.1 Introduction

For the traditional regression, model parameters are calculated using the least sum squared error. The traditional simple linear regression model can be written as:

$$y = \beta_0 + \beta_1 x + e$$

where  $y$  is the dependent variable;  $x$  is the independent variable;  $\beta_0$  is the estimated intercept;  $\beta_1$  is the estimated coefficient for  $x$ ; and  $e$  is the error. The least sum squared error is defined as:

$$\min \sum_{i=1}^n (y_i - \beta_1 x_i - \beta_0)^2$$

Different from the traditional regression, a simple QR model can be written as:

$$Q_\theta(y|x) = \beta_0(\theta) + \beta_1(\theta)x + e$$

where  $\theta$  is a specific conditional quantile of interest;  $y$  is the dependent variable;  $x$  is the independent variable;  $\beta_0$  and  $\beta_1$  are the intercept and the estimated coefficient of  $x$ , respectively, for  $y$  at  $\theta^{th}$  percentile; and  $e$  is the error.

The  $\theta^{th}$  QR model is defined as any solution to minimize the following terms (Koenher & d'Orey, 1978):

$$\min \sum_{i=1}^n \rho_\theta(y_i - \beta_1(\theta)x_i - \beta_0(\theta))$$

where  $\rho_\theta(\cdot)$  is defined as:

$$\rho_\theta(u) = \begin{cases} (\theta - 1)u, & u < 0 \\ \theta u & u > 0 \end{cases}$$

The main difference between the traditional regression and QR is that the traditional regression focuses exclusively on a specific location of the dependent variable distribution, while QR studies different locations of a dependent variable distribution. QR offers “a global view” on the relationships between the independent variable and explanatory variables (Davino et al., 2013). Furthermore, as discussed by Davino et al. (2013), QR can

address different types of distributions and eliminate the dependence on normality assumptions.

In Davino et al.'s book (2013), 10 models with different error patterns, including homogeneous and heterogeneous errors, were examined using QR. The results indicated that if the error pattern followed the normal distribution, slopes estimated by QR at different quantiles were all the same as slopes estimated by Ordinary Least Squares (OLS). If the error in the dependent variable was heterogeneous, the slopes estimated by QR at different quantiles were different. In this case, using OLS regression analysis cannot provide a complete picture of the relationship between variables, since it only focuses on changes at means. If the error was homogeneous but not in normal distribution, the slopes across different quantiles were the same, but the intercepts at each quantile were different (Davino et al., 2013). In other words, QR regression can reveal more information than the traditional regression does by studying relationships between explanatory variables and the dependent variable at different quantiles.

Another advantage of QR is its nonparametric characteristic. That means assumptions of error distributions are not needed. Therefore, QR is able to model different types of error distributions and can provide more realistic and robust outcomes. In addition, QR is less sensitive to outliers compared to the traditional regression due to its ability to study dependent variables at multiple quantiles (Galvao, 2011). QR estimates can also be used to obtain prediction intervals for the conditional distribution of the dependent variable.

Due to the nonparametric nature of QR, such prediction interval outperforms the ordinary prediction interval (Zhou & Portnoy, 1996; Davino et al., 2013).

#### 2.4.2 Applications of QR in Other Fields

QR has been widely used in different fields, such as economics, energy and biology, but not in pavement management yet. Applications of QR in different fields are reviewed below.

Eide and Showalter (1998) investigated the effect of school quality on student performance (test score) using QR and indicated results were different compared to those obtained from traditional regression. They concluded that the effect of school quality was not influential on the average of test scores, but might have different effects on the test scores at different conditional quantiles. McClain and Rex (2001) used QR to investigate the relationship between dissolved oxygen levels and the maximum size of deep-sea organisms. They concluded that the dissolved oxygen concentration may influence the maximum size of deep-sea organisms. Dimelis and Louri (2002) used QR to investigate the influence of a firm's foreign involvement on the productivity distribution, and concluded that foreign involvement was especially influential to the lower productivity local firms. Machodo and Mata (2005) used QR to decompose the changes in wage distribution. The results indicated that education had a greater influence on wages of individuals who were at the top of the wage distribution than those of at the bottom of the wage distribution. Coad and Rao (2008) used QR to confirm the importance of innovation

for a firm regarding its growth and concluded that innovation had significant positive influence on the firms' growth.

In addition, QR has been used to generate probabilistic forecasting, especially in the energy industry. Bremnes (2004) and Haque et al. (2014) used QR to generate probabilistic forecasts of wind power. Hong et al. (2014) used QR to forecast long-term probabilistic energy load. Indicated in Bremnes's study (2004), the advantages of QR included "no distribution assumption and flexible inclusion of predictive information." These advantages are essential to help decision-makers make an optimal economic decision.

## CHAPTER 3 DATA PREPARATION

This chapter summarizes the methodology used to generate the data set, develop distress and performance models, and investigate treatment effectiveness of composite pavements using the NCDOT PMS as a case study. This chapter also elaborates on several steps that are needed to extract composite pavement data from the asphalt data set, including choosing the appropriate data sources, identifying the composite pavement sections, and cleansing the data.

### 3.1 Overview of the Research Methodology

Distress and performance models for composite pavements requires the development of database containing all necessary information from the existing asphalt performance and the construction data sets. Since composite pavements have been historically classified as asphalt pavements by the NCDOT, their condition data are included in the asphalt data set. As the information resides in multiple data sources, the two existing data sets were merged according to roadway spatial relationships. After data sets are merged, the resulting data set needs to be cleaned using statistical methods and field experiences. Following this procedure, the windshield and the automated data sets for composite pavements are generated.

In this study, the windshield and the automated data sets for composite pavements were compared in terms of data attributes, precisions and distributions using descriptive statistics and graphs. Based on the data comparison result, it was decided to use the

automated data to develop pavement distress and performance models because the pattern of automated data tends to be more consistent with engineers' experience, and to use the windshield data to investigate treatment effectiveness because it has a longer history.

Pavement distress models were developed using non-linear regressions (S-shaped curves). Pavement performance Models were developed using traditional non-linear regressions and quantile regression (QR) models.

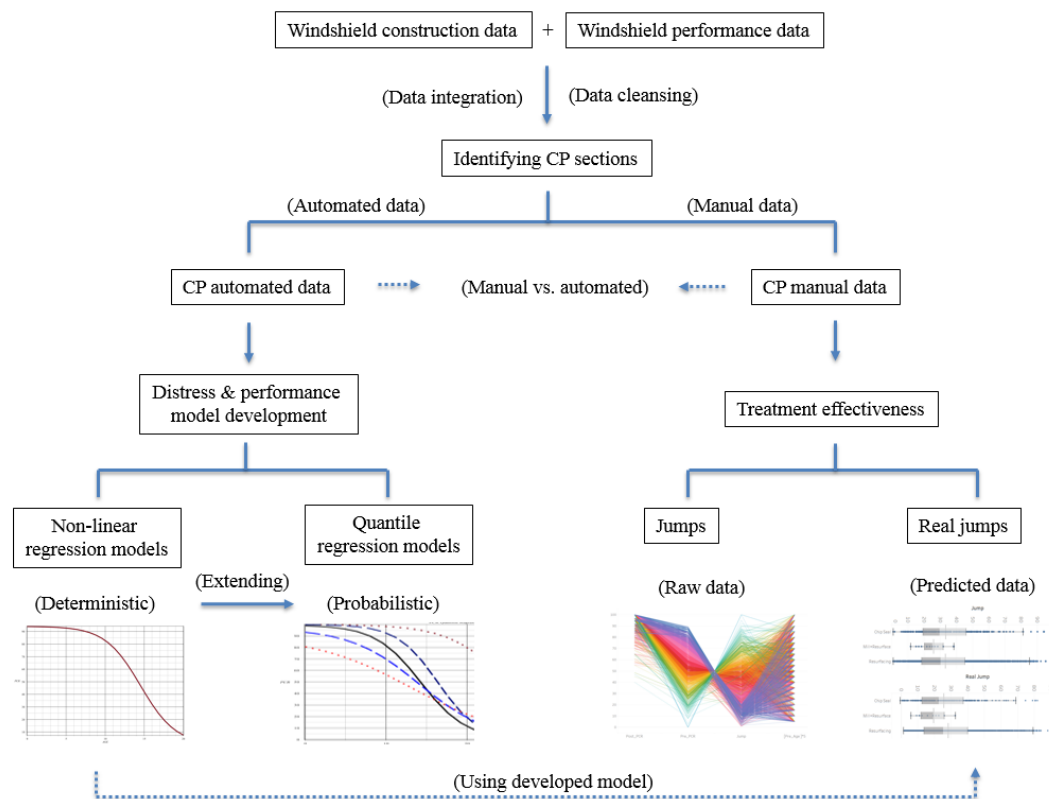


Figure 2: Research methodology

Using the windshield data, treatment effectiveness was quantified by the differences between the pre-treatment and the post-treatment PCRs, which was referred to

performance jumps in this study. Two different jumps were investigated: jumps and real jumps, which were calculated as the difference between observed pre-treatment PCR and post-treatment PCR, and observed pre-treatment PCR and predicted real post-treatment PCR, respectively. Figure 2 shows the work flow used in this research.

### 3.2 Data Sources

The construction and the performance data sets used by this research were provided by the NCDOT. The construction data set contains the historical treatments applied to the NC roadway system from the early 1900s to 2015. In this data set, treated pavement sections are identified by the county name, the route number, the mileposts, and the treatment year. Other information, such as the treatment types and materials, is also included.

The performance data set is created from roadway survey results and consists of performance ratings of different types of distresses. In the performance data set, pavement sections are identified by the county name, the route number, the mileposts, and the effective year which indicates the year when the data was collected. Based on data collection methods, two types of performance data have been generated and maintained by NCDOT: the windshield and the automated data. The NCDOT had launched windshield roadway surveys since the early 1980s, and the available windshield data are from 1982 to 2010. In 2011, the NCDOT started using the automated technique to collect distress data. The automated data used in this research are from 2013 to 2015. The following sections



review characteristics of windshield and automated data and discuss the differences between these two types of data, in terms of data formats and data patterns.

### 3.2.1 The Windshield and Automated Performance Data

In North Carolina, the windshield data has a much longer history than the automated data. However, the automated data contain more types of distresses than the windshield data. For example, in the automated data, patching has been categorized into wheel path and non-wheel path patching, and longitudinal cracking and reflective cracking have been added to the automated data. The measurements of distresses are also different. In the windshield data, alligator cracking is the only distress that is measured using continuous interval ratings, while the other distresses are rated with discrete ordinal ratings, such as “None”, “Light”, “Moderate”, and “High”. Table 1 summarizes the differences in terms of distress measurements, data processing methods and analysis processes between the windshield and the automated data (Chen et al., 2014).

Table 1: Comparison of the windshield and the automated data in North Carolina

Data	Distresses	Data Type	Severity Level	Unit	Analysis Process			
					1	2	3	4
Windshield	Alligator cracking	Interval/ continuous	4	Percentage	No	Yes	No	No (already calculated in raw data)
	Transverse cracking	Ordinal/ discrete	4	N/A	Assigning discrete values	No	Yes	
	Rutting							
	Raveling							
	Oxidation							
	Patching							
	Bleeding							
Automated	Alligator cracking	Interval/ continuous	3	Square feet	Yes	Yes	No	Used with weight factors
	Raveling		3			Yes		
	Wheelpath patching		1			No		
	Non-wheelpath patching		1			No		
	Transverse cracking		3	Yes				
	Reflective cracking		3	Yes				
	Longitudinal cracking		2	Yes				
	Longitudinal lane joint		2	Yes				
	Bleeding		2	Yes		Not included		
	Delamination		1	Square feet				No

Note: In the “Analysis Process” column, “1” represents Data Normalization; “2” represents Distress Composite Index; “3” represents Numeric Data Transformation; and “4” represents Calculation of Performance Composite Index (Chen et al., 2014).

In 2014, Chen et al. (2014) developed distress and performance models using both the windshield data and the automated data for NCDOT. Since the windshield data used in their study included interval and ordinal distress ratings, two types of statistical models were developed: sigmoidal models for interval data and piecewise linear models for ordinal data. Sigmoidal models described the development of distresses over pavement age, while piecewise linear models consisted of several straight lines to indicate the changes in different severity levels. For the automated data, the sigmoidal model form was used for all distress models as well as the performance model, because the data were all continuous with a range from 0 to 100 (Chen et al., 2014).

### 3.2.2 Differences in Data Distributions

In this research, pavement sections were categorized into four pavement families based on their pavement classifications and traffic volumes: Interstate, US\_0-5K, US\_5K plus, NC\_0-5K, and NC\_5K plus. As an example, the US\_0-5K family represents US roadway sections with an Annual Average Daily Traffic (AADT) between 0 and 5,000. Alligator cracking and raveling of US\_0-5K were used as an example to compare the distributions of the windshield and the automated data. Figures 3 and 4 show boxplots of alligator cracking of the windshield data and the automated data, respectively. The x axis represents the pavement age, and the y axis represents alligator cracking index. The horizontal bar within each box represents the average index value at each age. In Figure 3, the windshield data at the age of 0 were reasonable, because outliers were removed during the data integration process. However, Figure 4 shows that the automated data at age of 0

were fairly low, and the distribution of the data points after age of 13 repeated the previous pattern. This indicates that the automated data also has quality issues, such as measuring errors and pavement age not appropriately reset. Based on the boxplots, from age of 1 to 12, alligator cracking in the automated data showed an obvious deterioration trend. In the windshield data, after age of 9, the average alligator cracking index values were almost the same, which contradicted engineers' experience that the amount of alligator cracking increased as pavement sections became aged.

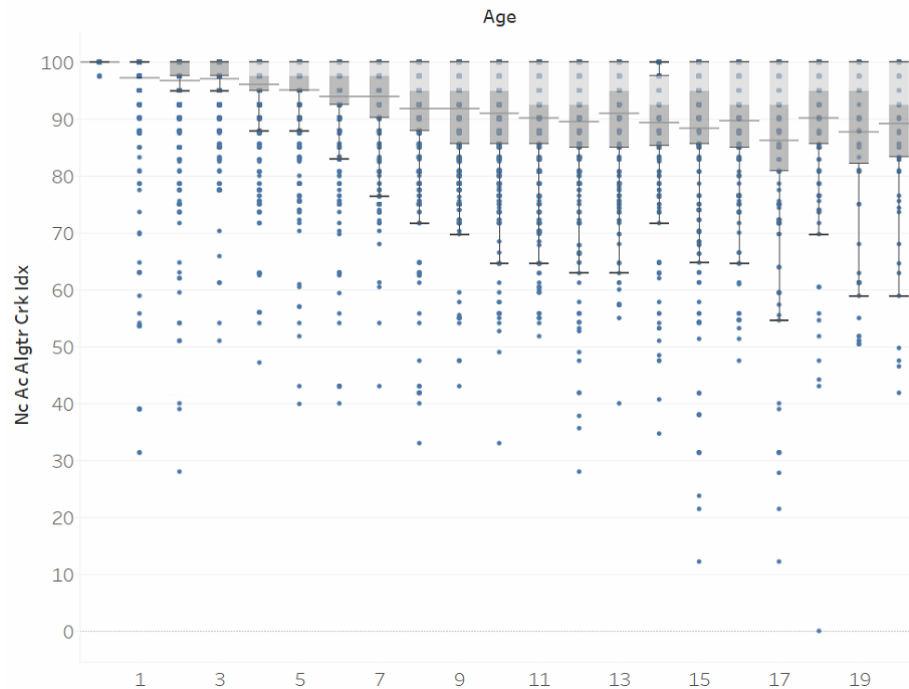


Figure 3: Alligator cracking in the windshield data

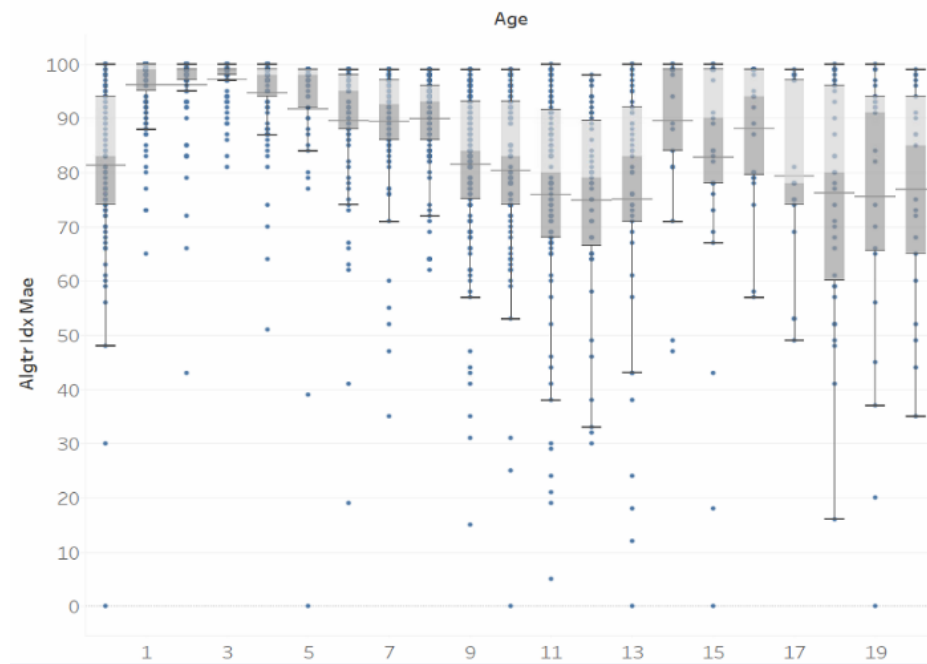


Figure 4: Alligator cracking in the automated data

It was also observed that the distresses collected in the windshield data were not as precise as those in the automated data. For example, Figure 5 shows the average windshield raveling index values at each age. Since the raveling index values were originally ordinal ratings and were converted into discrete ratings during the data processing procedure, real conditions of raveling cannot be represented precisely. Therefore, it was hard to show the deterioration process of raveling using the windshield data. However, in the boxplot developed using the automated data, the deterioration trend was obvious, as shown in Figure 6.

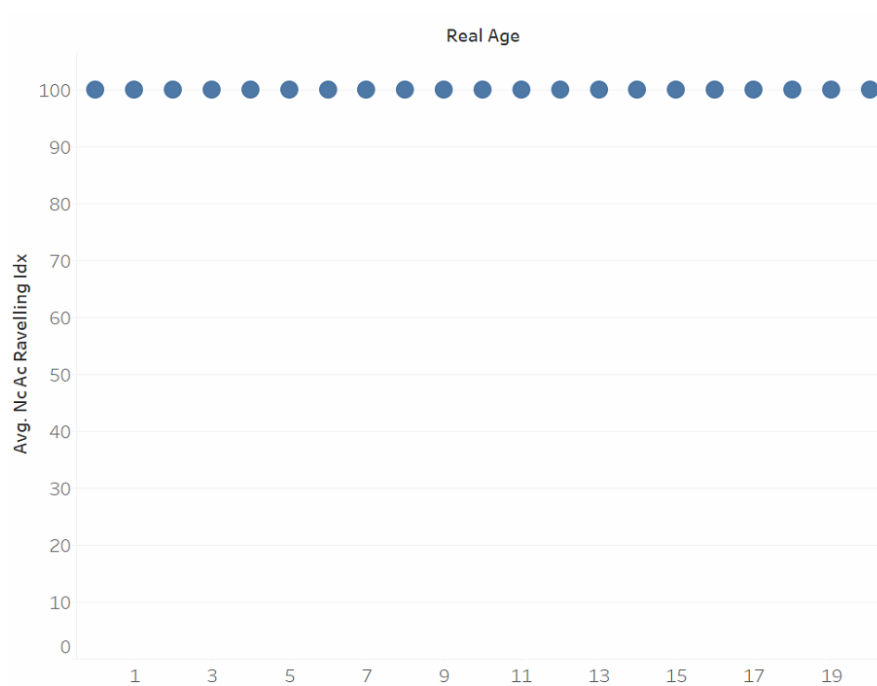


Figure 5: Raveling in the windshield data

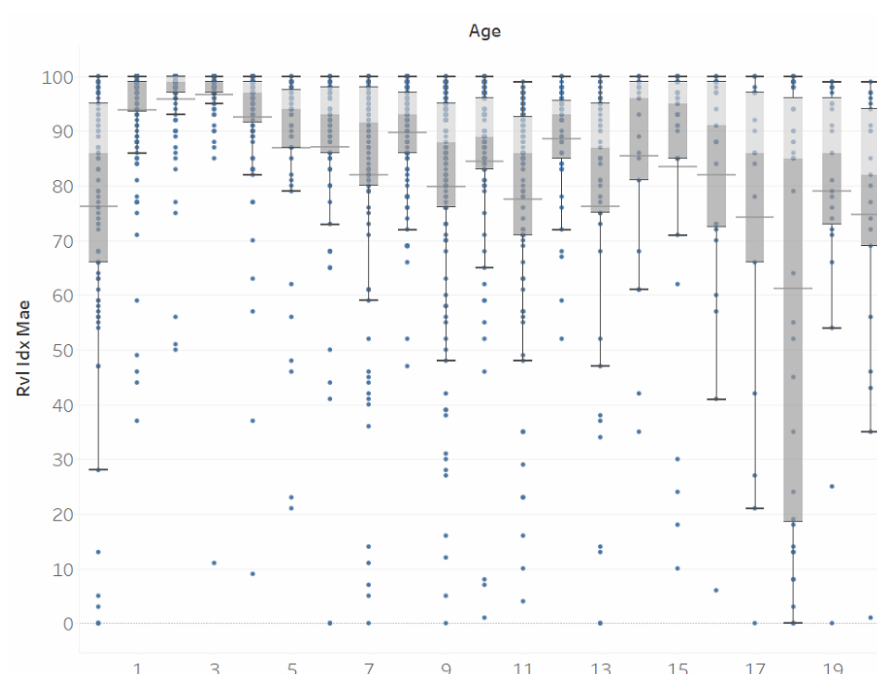


Figure 6: Raveling in the automated data

In summary, the windshield data and the automated data are different not only in measurement methods and distress types, but also in data patterns. According to the comparison between the windshield and the automated data using alligator cracking and raveling of US\_0-5K as an example, it was decided to use the windshield data to identify composite pavement sections and to use the automated data to develop distress and performance models.

### 3.3 Data Merging Process

To identify composite pavement sections from the automated data, several steps were performed as shown in the flow chart below (Figure 7). In this flow chart, CP indicates composite pavements. This process was conducted using SAS (SAS, 2016).

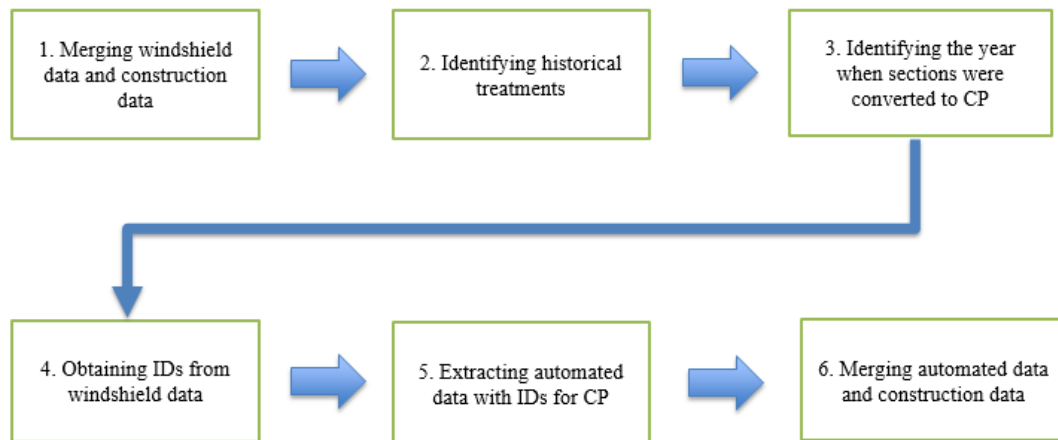


Figure 7: The flow chart of data merging process

*Step One: Merging the windshield performance data and the construction data*

The windshield performance data (from 1982 to 2010) and the construction data (from 1900 to 2015) were merged using section identifications, including the county name, the route number, the mileposts, and the effective year. Since the mileposts of the same pavement section in these two data sets were not always the same, a threshold of 50% of the length between the starting and ending mileposts was used to merge these two data sets.

Figure 8 shows the algorithm of the merging process. In this figure, the blue bar represents the roadway section that has been constructed. This information is included in the construction data. The green bars represent the roadway pavement performance surveying sections with four possible location scenarios. This information is included in the windshield performance data. The first scenario shows that the construction section covers the entire performance section, and these two data sets can easily be merged. The second and the third scenarios show that the performance section crosses the boundaries of the construction section. If the length of the performance section has more than 50% covered by the construction section, then these two data sets can be merged. The fourth scenario shows that the performance section covers the entire construction section. If the construction section is longer than 50% of the performance section, then these two data sets can be merged as well.



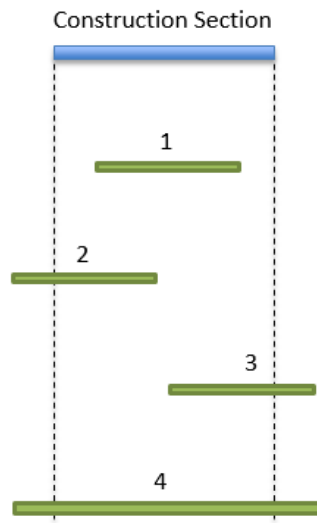


Figure 8: Comparing mile posts to merge two data sets

*Step Two: Identifying historical treatments for each pavement section*

After Step One, each pavement section and its historical treatments are linked together. Based on the information obtained from the construction data, the materials used for each historical treatment, either asphalt or concrete, can be identified.

*Step Three: Identifying the year when pavement sections converted to composite pavements*

For a pavement section, if its top layer treatment materials, from the earliest to the latest, changed from concrete to asphalt, then the section can be identified as a composite pavement section after the first asphalt treatment. In addition, the year of the latest treatment before the performance data was collected was considered as the “reborn” year of a pavement section, and was used to calculate the age of this pavement section.

*Step Four: Obtaining identifications (IDs) from the composite windshield data*

After previous steps, the windshield performance data of composite pavements from 1982 to 2010 were obtained from the original asphalt performance dataset. IDs of composite pavement sections, i.e., county names, route names, and mileposts, were extracted. Two assumptions were made in this step. They are: 1) once a concrete pavement section was converted to a composite section, it would stay as a composite section, and 2) composite pavement sections in the automated data (from 2013 to 2015) would use the same IDs as the windshield data.

*Step Five: Extracting automated performance data using IDs*

The IDs obtained from Step Four were used to identify composite pavements from the automated data. This process was completed by locating IDs obtained in the previous step in the automated performance data and comparing corresponding starting and ending mileposts.

*Step Six: Merging automated performance data with the construction data and calculating the age*

After Step Five, the automated performance data set of composite pavements was obtained. However, the latest treatment in the windshield data, which was used to calculate pavement age, was not necessarily the latest treatment in the automated data because pavement section might be treated after the windshield data was collected. Therefore, these extracted automated performance data was merged again with the construction data. Then

the year in which the latest treatment was obtained was used to calculate the age of a composite pavement.

### 3.4 Data Cleansing

To improve the quality of the automated data for composite pavements, data cleansing was implemented in the merging process based on statistical methods and NCDOT engineers' field experience. Outliers of the performance indices at each age in the automated data set were removed using interquartile ranges. Errors in the windshield performance data set were discovered and fixed based on field experience. For instance, a low performance index value at early ages and a high performance index value at late ages were considered as outliers and thus removed. Data cleansing processes were conducted using SAS.

Special attention should be given to the age of the windshield performance data identified in Step Three. In this step, if the performance data was collected in the same year when it was treated but right before the treatment, the performance rating would be fairly low but the pavement age would be reset to 0. An example is shown in Figure 9 (the "layer\_year" column indicates the corresponding latest treatment year of the performance data). In this example, the first treatment was performed in 1994, and the second was in 2007. The condition data (PCR of 48.4) in the year of 2007 should have an age of 13 instead of 0. To address this issue, all the performance data at age of 0, with PCR values less than 95, were removed. This was because if a pavement section was just treated, it should have a PCR value close to 100.

COUNTY	route1	OFFSET_FROM	OFFSET_TO	EFF_YEAR	PCR	layer_year	age
60	10000077	0	2.029	1994	100	1994	0
60	10000077	0	2.029	1995	95	1994	1
60	10000077	0	2.029	1996	95	1994	2
60	10000077	0	2.029	1997	95	1994	3
60	10000077	0	2.029	2000	80	1994	6
60	10000077	0	2.029	2001	71.7	1994	7
60	10000077	0	2.029	2002	68.4	1994	8
60	10000077	0	2.029	2003	60.1	1994	9
60	10000077	0	2.029	2004	55.9	1994	10
60	10000077	0	2.029	2005	55.9	1994	11
60	10000077	0	2.029	2006	48.4	1994	12
60	10000077	0	2.029	2007	48.4	2007	0
60	10000077	0	2.029	2008	100	2007	1
60	10000077	0	2.029	2009	100	2007	2

Figure 9: A data error example

Based on the examination of merged data set developed in Step Three, some sections had obvious jumps in the PCR value but no associated treatments were found, as shown in Figure 10. To address this issues, these PCR values after jumps were assigned a new age of 0 and were considered as the results of treatments that have not been recorded in the construction data.

COUNTY	route1	OFFSET_FROM	OFFSET_TO	EFF_YEAR	PCR	layer_year	age
60	10000077	10.397	12.06	1996	86.7	1996	0
60	10000077	10.397	12.06	1997	86.7	1996	1
60	10000077	10.397	12.06	1998	86.7	1996	2
60	10000077	10.397	12.06	2000	86.7	1996	4
60	10000077	10.397	12.06	2001	86.7	1996	5
60	10000077	10.397	12.06	2002	86.7	1996	6
60	10000077	10.397	12.06	2003	83.4	1996	7
60	10000077	10.397	12.06	2004	100	1996	8
60	10000077	10.397	12.06	2005	96.7	1996	9
60	10000077	10.397	12.06	2006	96.7	1996	10
60	10000077	10.397	12.06	2007	83.4	1996	11
60	10000077	10.397	12.06	2008	100	1996	12
60	10000077	10.397	12.06	2009	100	1996	13

Figure 10: An example of performance jumps

After the initial data cleansing process was completed, a further cleansing was performed for these extracted automated data using the interquartile range (IQR).

For each pavement family, pavement performance data was sorted by age, and outliers were removed at each age using the IQR method. The IQR is defined below:

$$IQR = Q_3 - Q_1$$

$$Bottom\ Boundary = Q_1 - 1.5 * IQR$$

$$Upper\ Boundary = Q_3 + 1.5 * IQR$$

where  $Q_1$  is the 25<sup>th</sup> percentile,  $Q_3$  is the 75<sup>th</sup> percentile, and IQR is the interquartile range. Data at each age beyond the corresponding bottom and upper boundaries were considered as outliers and removed. Roadway sections with age greater than 20 were excluded from further analyses according to NCDOT engineers' recommendation. Sample sizes of the cleaned data, pavement classifications and AATD values are shown in Figure 11.

Using these cleaned data, boxplots were developed to visualize the data in terms of distributions, percentiles, median (represented by the short horizontal line in the box), and mean values (indicated by the diamond in the box). As an example, the boxplot of transverse cracking index of US\_0-5K is shown in Figure 12.

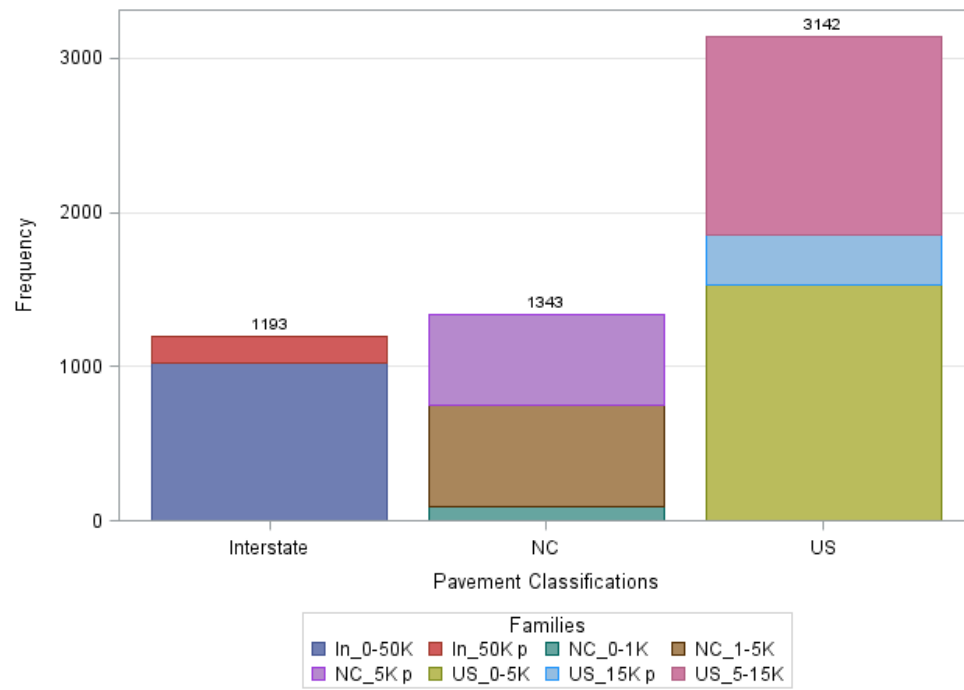


Figure 11: Sample sizes of pavement families

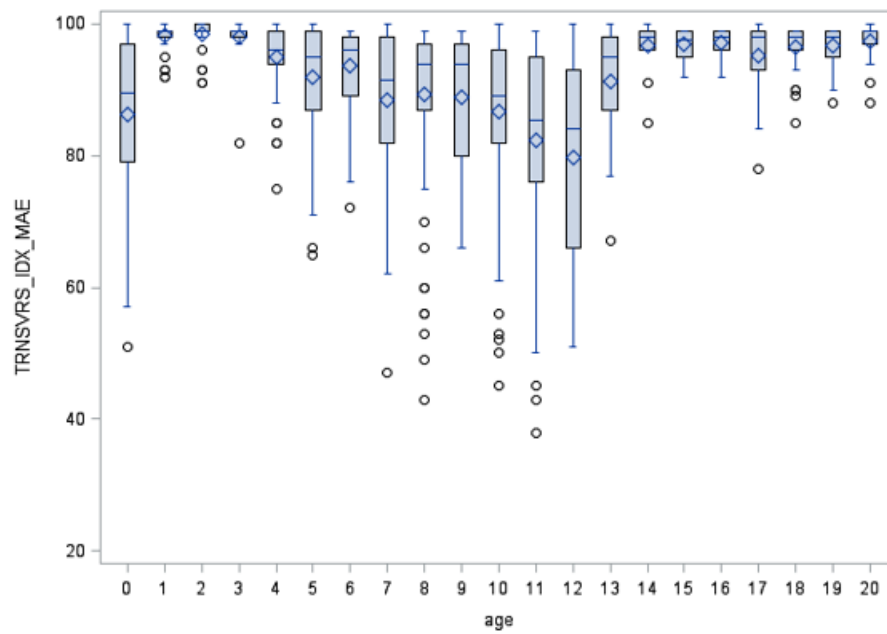


Figure 12: The boxplot of transverse cracking index in US\_0-5K

From this boxplot, it can be noticed that the average value of transverse cracking indices decreased over time and then increased at the age of 13, and the distribution of data after the age of 13 repeated the previous pattern. Similar situations were observed in other roadway families. This was possibly caused by pavement ages not being appropriately reset for some pavement sections. To address this issue, performance data after age of 12 were not used to develop distress and performance models.

It was also observed that the average index value at age of 0 was fairly low, which was not expected based on field experience. Right after a treatment, there should be no transverse cracking observed in pavements, and the transverse cracking index should be close to 100. This issue was probably caused by the fact that the performance data was collected in the same year when it was treated but right before the treatment, as shown in Figure 9. Similar situations were observed in other roadway families as well. Therefore, data at the age of 0 was also not used to develop the models. The cleansing processes for other distress indices and PCR were conducted using the same algorithm described above.

## CHAPTER 4 DEVELOPMENT OF NONLINEAR DISTRESS AND PERFORMANCE MODELS

In this chapter, the procedures of developing distress and performance models are described, the dominant distresses of composite pavements are discussed.

### 4.1 Distress Models

#### 4.1.1 Calculation of Distress Index

In the automated data, distress data were recorded based on distress types and amounts of distress at different severity levels. To develop distress and performance models, it is necessary to calculate a distress index for each distress, and the value of this index can represent different severity ratings of the distress for each roadway section. These distress indices are composite indices and can be calculated using the Maximum Allowable Extent (MAE) functions (Chen et al., 2014).

Two steps were involved in the distress index calculations. The first step was to normalize the raw data into percentages at each severity level. Then, a composite distress index value was obtained using the MAE spreadsheet provided by NCDOT. The algorithm and details of the distress index calculation was presented by Chen et al. (2014). The calculated distress index values were used for developing distress models and later the performance models.



#### 4.1.2 Development of Distress Models

Using automated data sets of five pavement families, boxplots were generated for eight types of distresses, including transverse cracking, longitudinal cracking, raveling, alligator cracking, rutting, longitudinal lane joint, wheel path patching, and non-wheel path patching. The boxplots of distresses of US\_0-5K are shown in Figures 13 and 14, as an example. The x axis represents age, and the y axis represents corresponding distress index values.

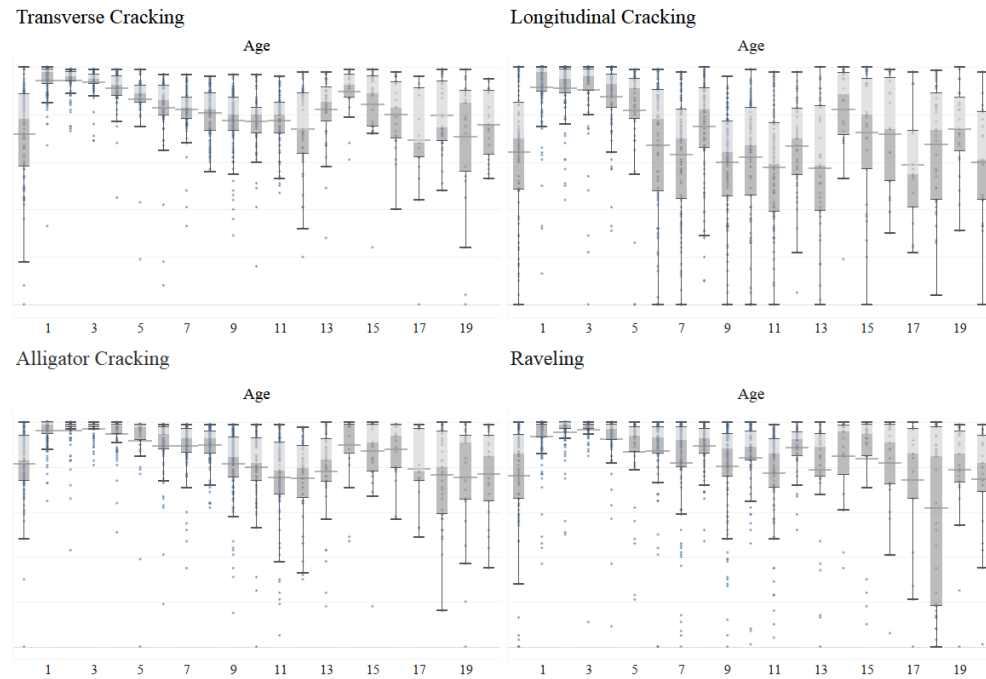


Figure 13: Boxplots of distresses of US\_0-5K (i)

From these boxplots, it can be observed that longitudinal cracking and transverse cracking were the main types of distresses occurring in composite pavements, followed by alligator cracking and raveling. Alligator cracking and raveling occurred frequently in aged

pavements. Rutting, longitudinal lane joint, wheel path patching, and non-wheel path patch did not occur frequently in the first 20-year period. Similar patterns were observed from other pavement families.

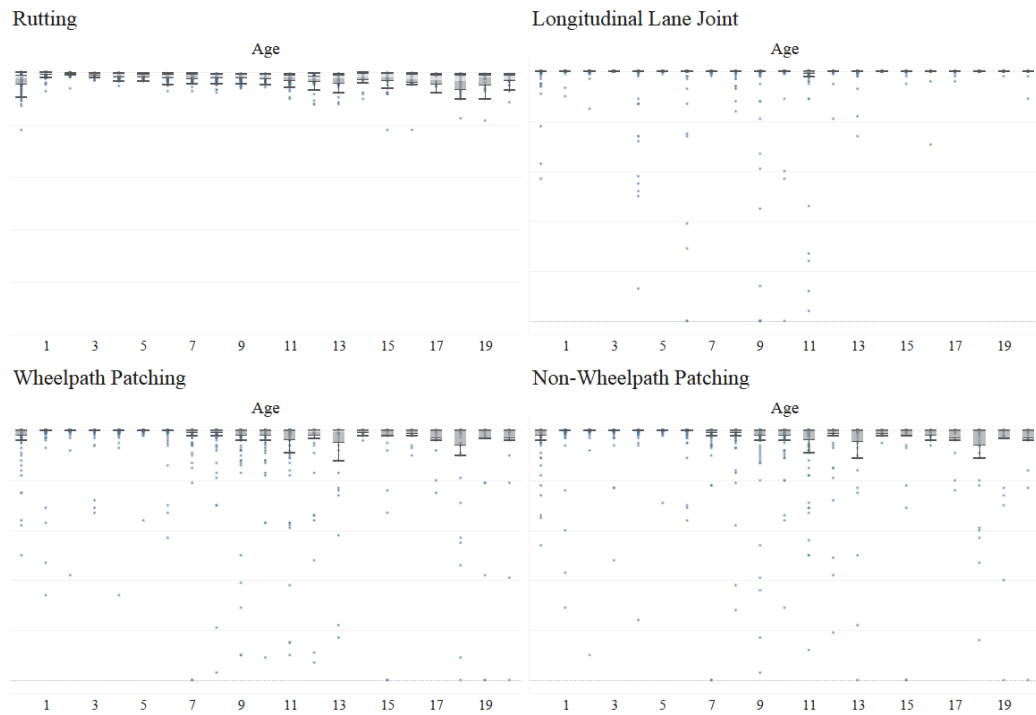


Figure 14: Boxplots of distresses of US\_0-5K (ii)

These boxplots also show that for transverse cracking, longitudinal cracking, alligator cracking and raveling, the average distress index values decreased and then increased at around age of 12, and the average values at age of 0 were not reasonable. As discussed in section 3.2.3., it was decided that distress models were developed using the data with age from 1 to 12.

The sigmoidal equation (Chen et al., 2014; Chen & Mastin, 2015) used to develop distress models for different pavement families can be written as

$$y = \frac{a}{1 + e^{\frac{-x+b}{c}}} \quad (1)$$

where  $y$  is PCR;  $x$  is pavement age;  $a$ ,  $b$ , and  $c$  are model parameters.

The sigmoidal equation can be further converted into a linear equation (Equation 2) (Chen et al., 2014), as follows:

$$-\frac{x-b}{c} = \ln(e^{\ln a - \ln y} - 1)$$

Let  $Y = \ln(e^{\ln a - \ln y} - 1)$ , then  $-\frac{x-b}{c} = Y$ .

Therefore,

$$Y = -\frac{1}{c}x + \frac{b}{c} \quad (2)$$

To estimate the parameters of the sigmoidal model, the value of parameter  $a$  was assigned to 100, since the initial index value at age of 0 should be close to 100. With the assigned initial value of  $a$ , the initial estimates of  $b$  and  $c$  can be easily calculated by running the simple linear regression analysis using Equation 2.

Distress Models for transverse cracking, longitudinal cracking, alligator cracking and raveling were generated for five pavement families. The resulting curves were shown in Figures 15 through 18.

Figure 15 shows the transverse cracking curves of five pavement families. It can be noticed that Interstate routes performed better than others, which means Interstate had the least amount of transverse cracking. In other words, Interstate was more resistant to the

weather change than the other families because transverse cracking is a non-load related distress.

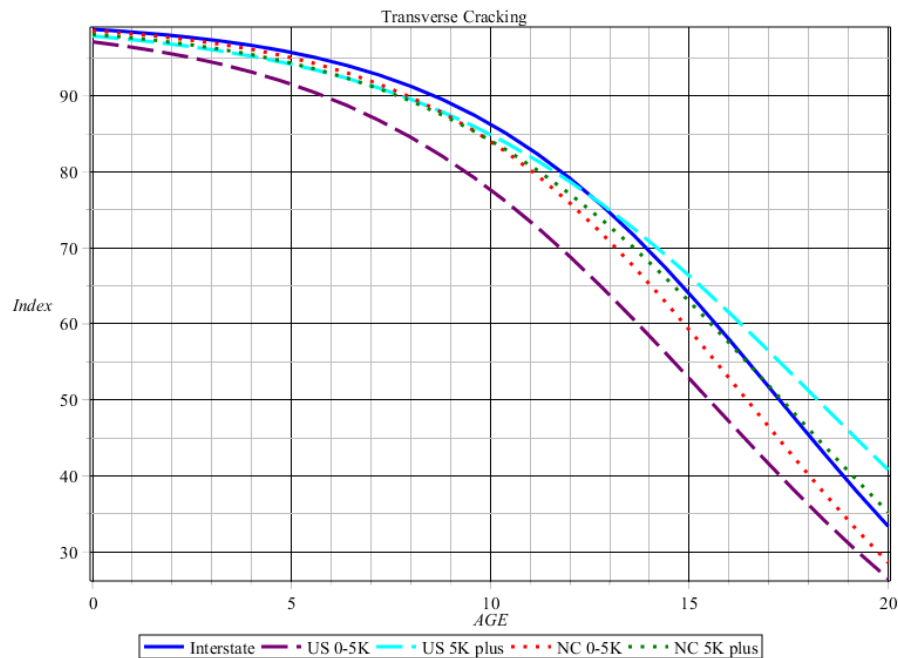


Figure 15: Distress models for transverse cracking

Figure 16 shows the longitudinal cracking curves of five pavement families. Interstate performed worse than others, followed by US\_0-5K and US\_5K plus. NC\_0-5K and NC\_5K plus performed better than Interstate and US routes. Figure 17 shows the alligator cracking curves of five pavement families. It can be noticed that the deterioration process of all pavement families had a similar trend. Figure 18 shows the raveling curves of five pavement families. All five pavement families performed similarly during the first 10-year period.

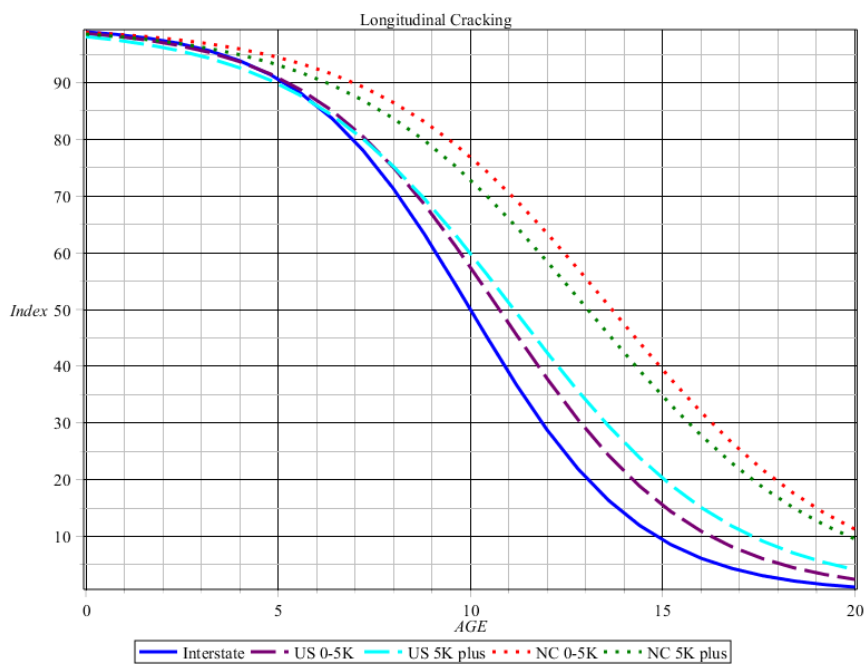


Figure 16: Distress models for longitudinal cracking

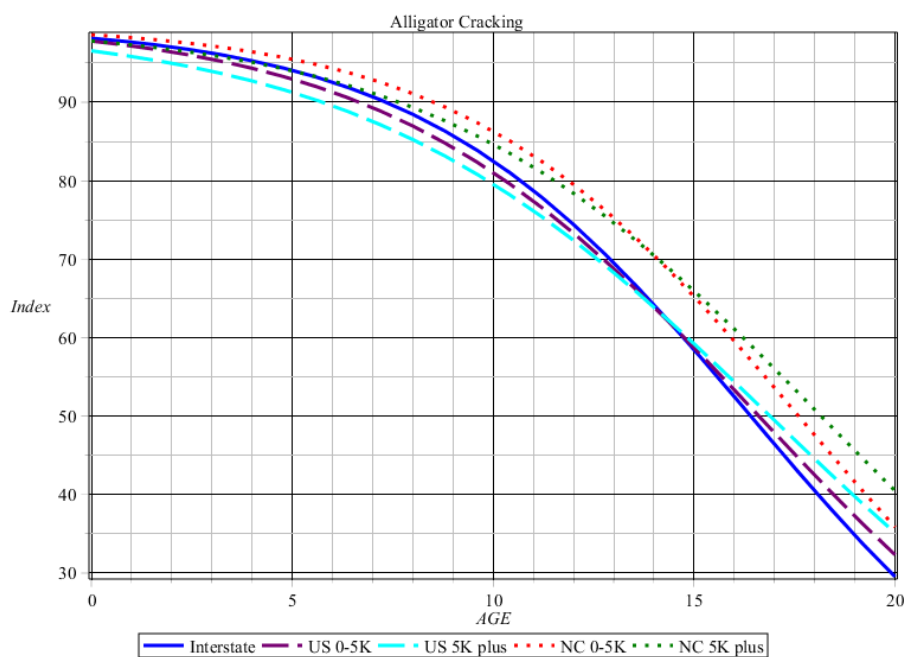


Figure 17: Distress models for alligator cracking

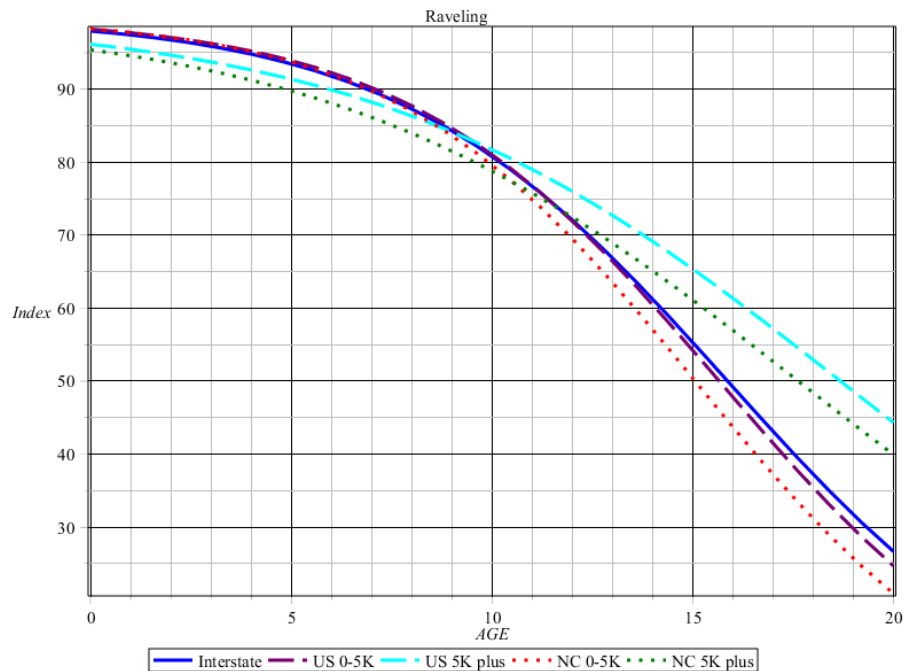


Figure 18: Distress models for raveling

An additional set of distress models was also developed using the representative sections. Since the automated data set only has three years of data, which is 2013, 2014, and 2015, every pavement section has a maximum of three pavement performance observations. There are a large number of sections that have only one or two observations. With this small amounts of observations, it is very challenging to identify outliers based on changes in PCR values. Therefore, it was decided to use selected pavement sections that have good data to develop an additional set of distress models. These models are expected to be more precise and reliable than if all sections are used. The following assumptions were used to select the representative sections:

- Age should be less than or equal to 13 to avoid the repeated distribution pattern;
- Pavement sections should have three observations with consecutive ages; and
- PCR values decrease over time.

Due to the small sample sizes of representative sections, five pavement families were combined into four: Interstate, US\_0-5K, US\_5K plus and NC. Figure 19 through 22 show distress models for transverse cracking, longitudinal cracking, alligator cracking and raveling, and the curves were color-coded by pavement. Similar conclusions can be drawn from these distress models, but more distinguished patterns of the same distress for different pavement families were revealed. Figure 23 through 26 show distress models for pavement families, Interstate, US\_0-5K, US\_5K plus, and NC, and the colors were coded by distress types.

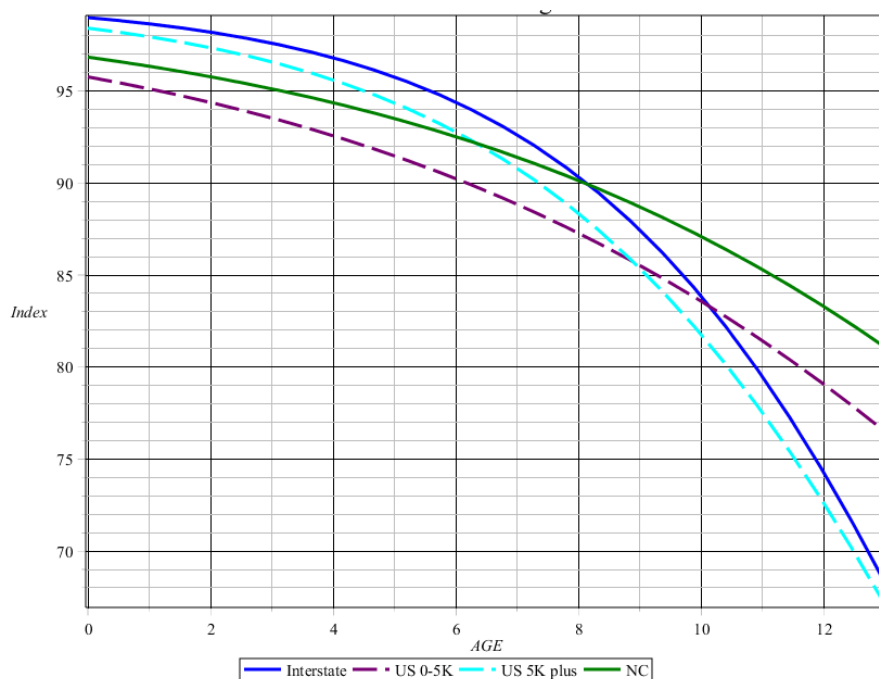


Figure 19: Transverse cracking models for pavement families

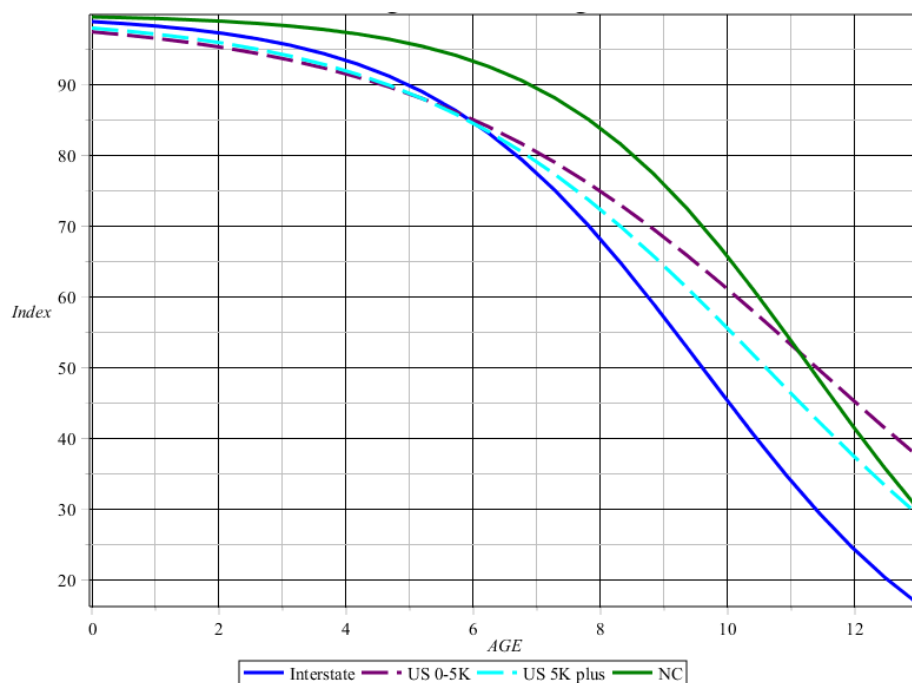


Figure 20: Longitudinal Cracking for pavement families



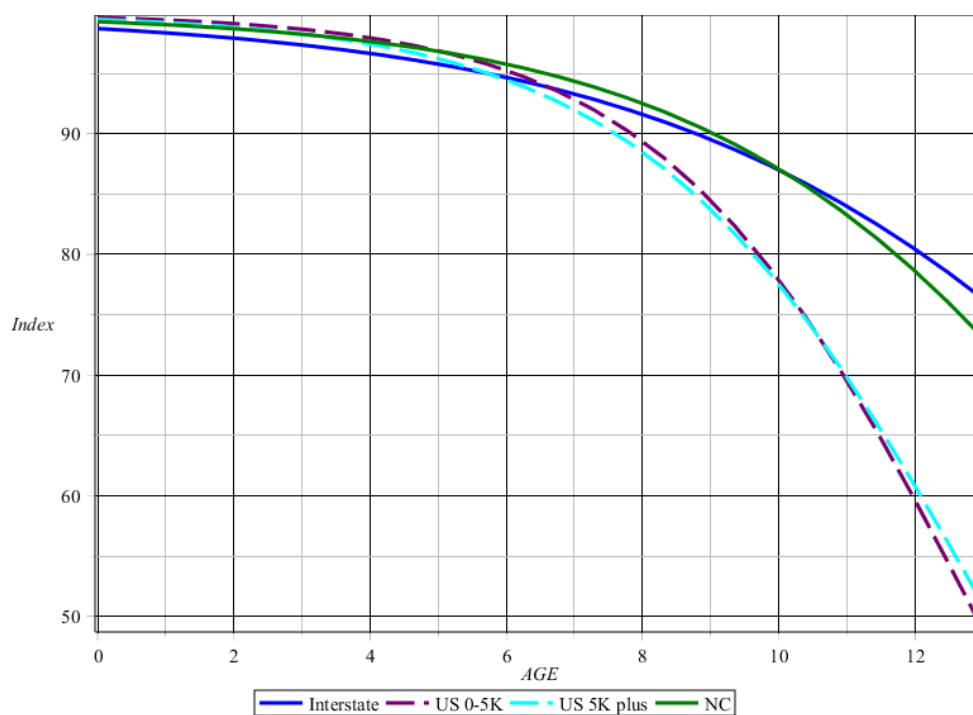


Figure 21: Alligator cracking for pavement families

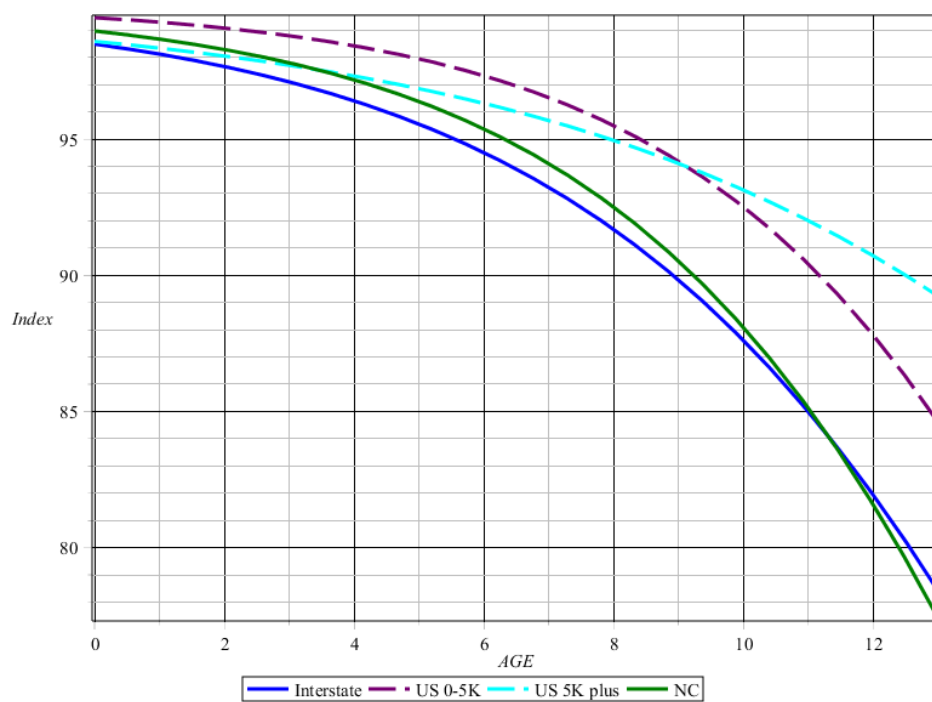


Figure 22: Raveling for pavement families

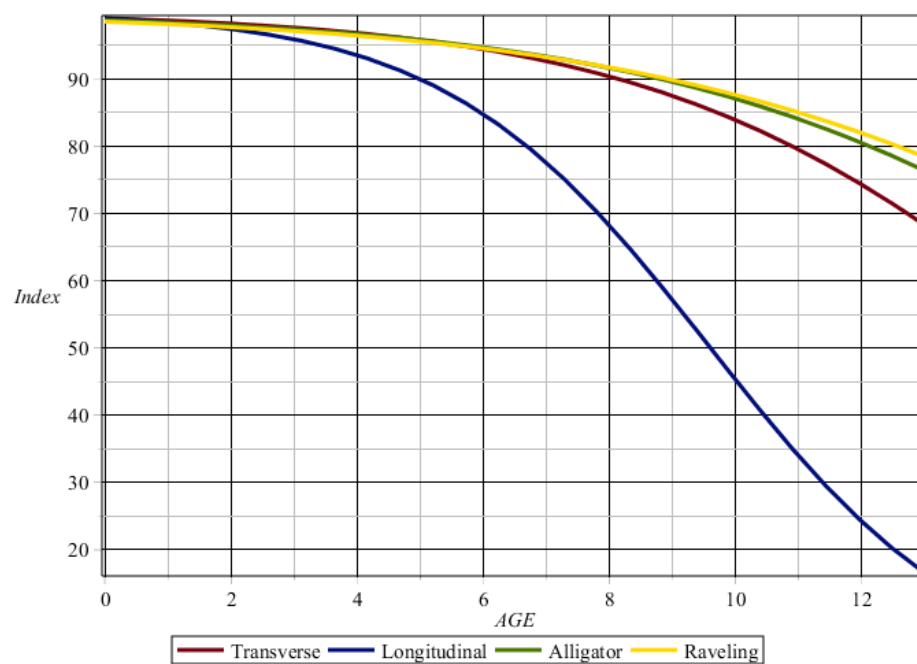


Figure 23: Distress models for Interstate

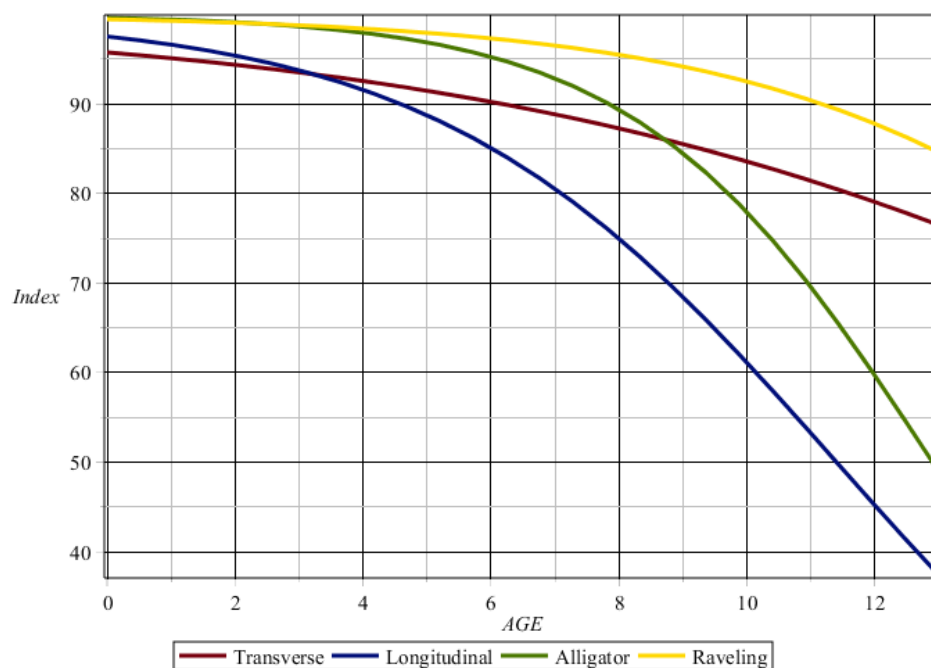


Figure 24: Distress models for US\_0-5K

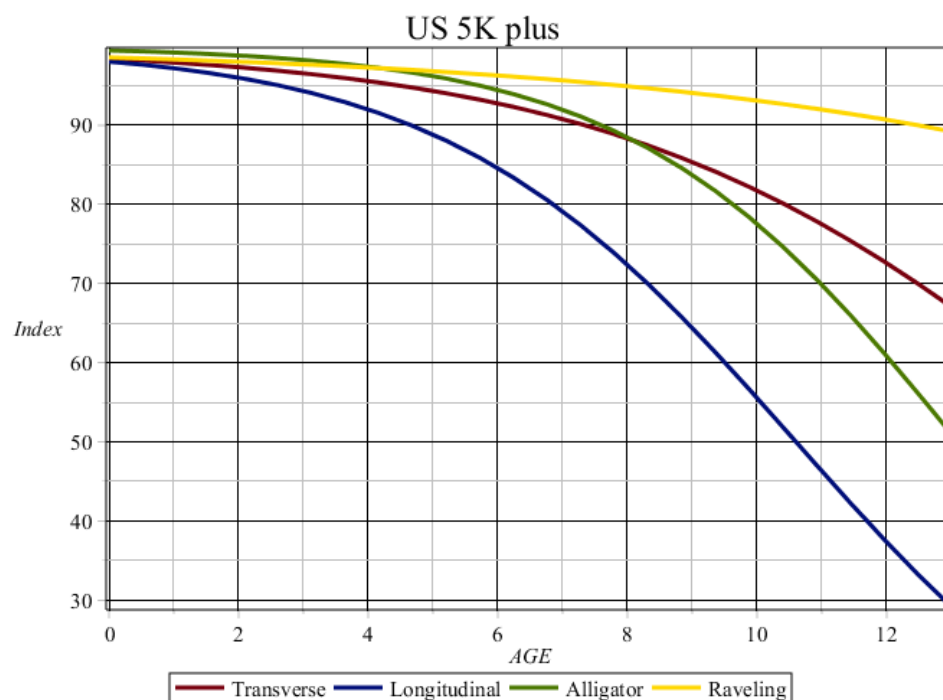


Figure 25: Distress models for US\_5K plus

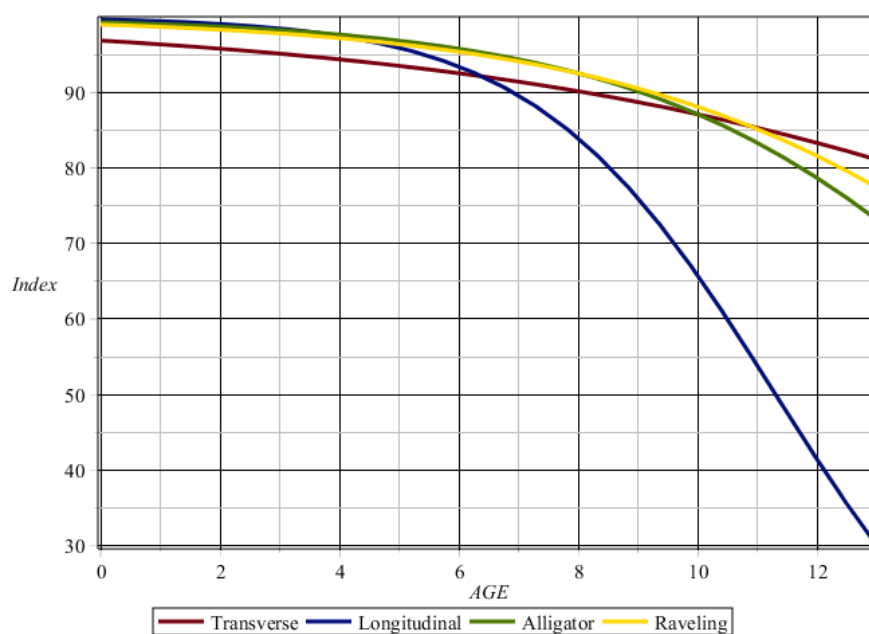


Figure 26: Distress models for NC

Based on all these distress models, it can be concluded that generally, longitudinal and transverse cracking are the dominate distresses for composite pavements, followed by alligator cracking and raveling. However, dominant distresses for different pavement families may be different, and also, dominant distresses for the same pavement family may change over time. For example, for US\_5K plus, longitudinal cracking and alligator cracking, instead of transverse cracking, were the dominant distresses after age of 8.

## 4.2 Performance Models

### 4.2.1 Calculation of Pavement Condition Rating (PCR)

As a composite index of pavement performance, PCR was calculated from distress index values of different distress types and corresponding weight factors. The weight factors were determined by a study conducted by Chen et al. (2014). Since distresses were categorized into load-related (LDR) and non-load related (NDR) in this research, PCR value was the smaller value of LDR and NDR index values. The equations of calculating LDR, NDR, and PCR are shown below:

- $LDR = 0.532 * (\text{Alligator Cracking Index}) + 0.152 * (\text{Wheel Patching}) + 0.089 * (\text{Non-Wheel Patching Index}) + 0.228 * (\text{Rutting Index})$
- $NDR = 0.425 * (\text{Transverse Cracking Index}) + 0.225 * (\text{Longitudinal Cracking Index}) + 0.175 * (\text{Longitudinal Lane Joint Index}) + 0.175 * (\text{Raveling Index})$
- $PCR = \min (NLR, LDR)$

#### 4.2.2 Development of Performance Models

The same sigmoidal equation (Equation 1) was used to develop pavement performance models. Representative pavement sections were selected based on the same criteria used for distress models and used to develop performance models. Figures 27 through 30 show scatterplots of representative sections with estimated performance trend lines for Interstate, US and NC, respectively.

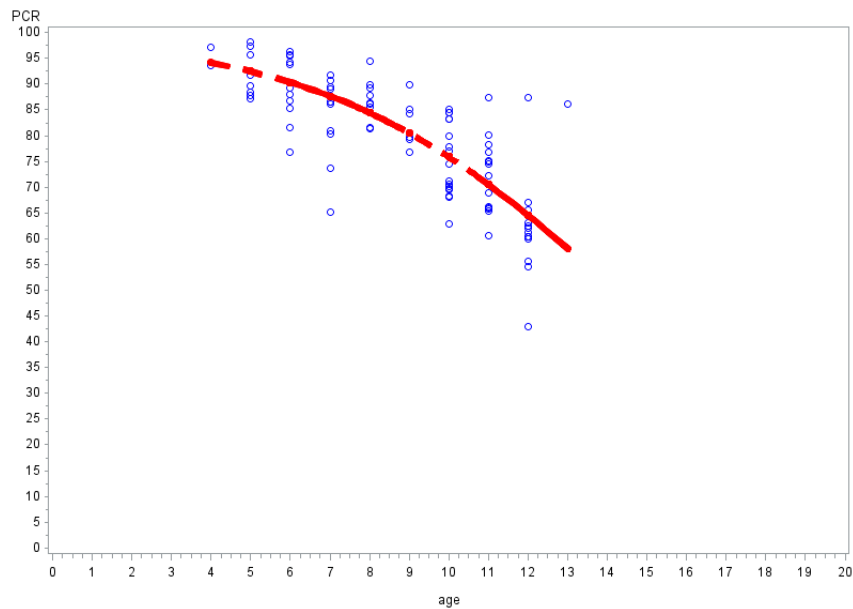


Figure 27: Performance model of Interstate

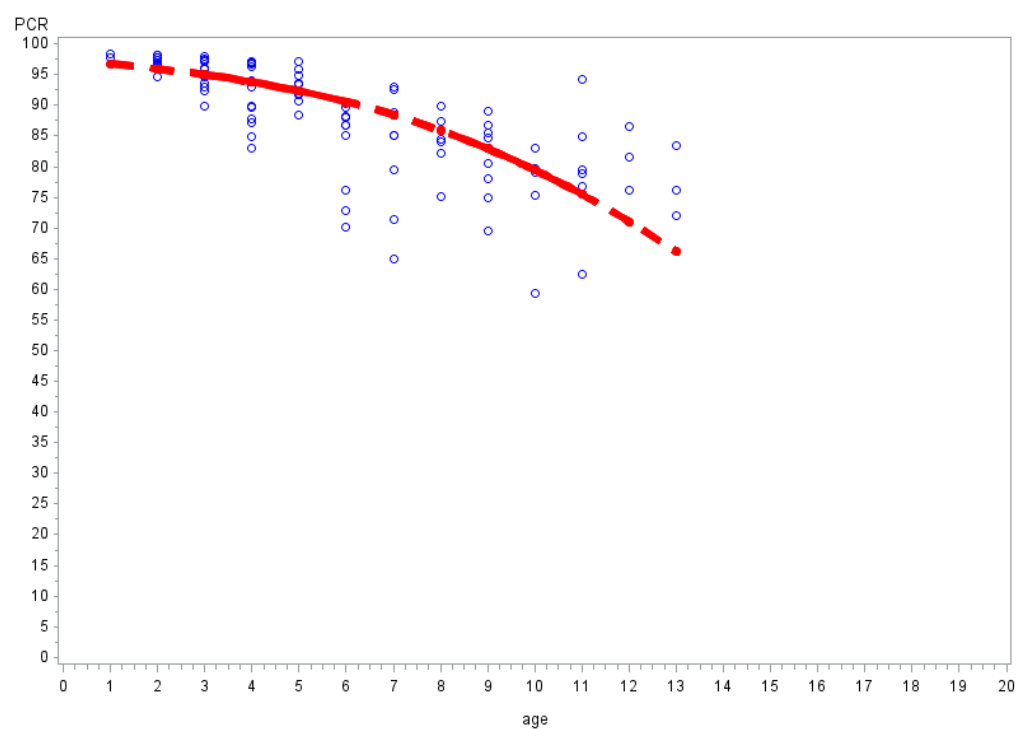


Figure 28: Performance model of US\_0-5K plus

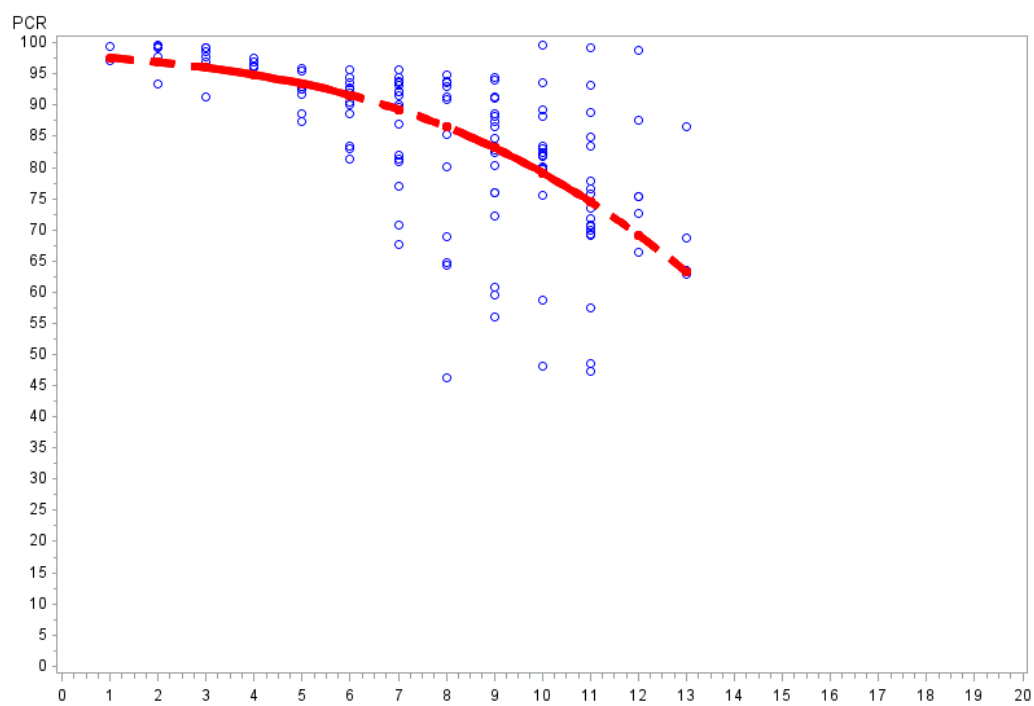


Figure 29: Performance model of US\_5K plus

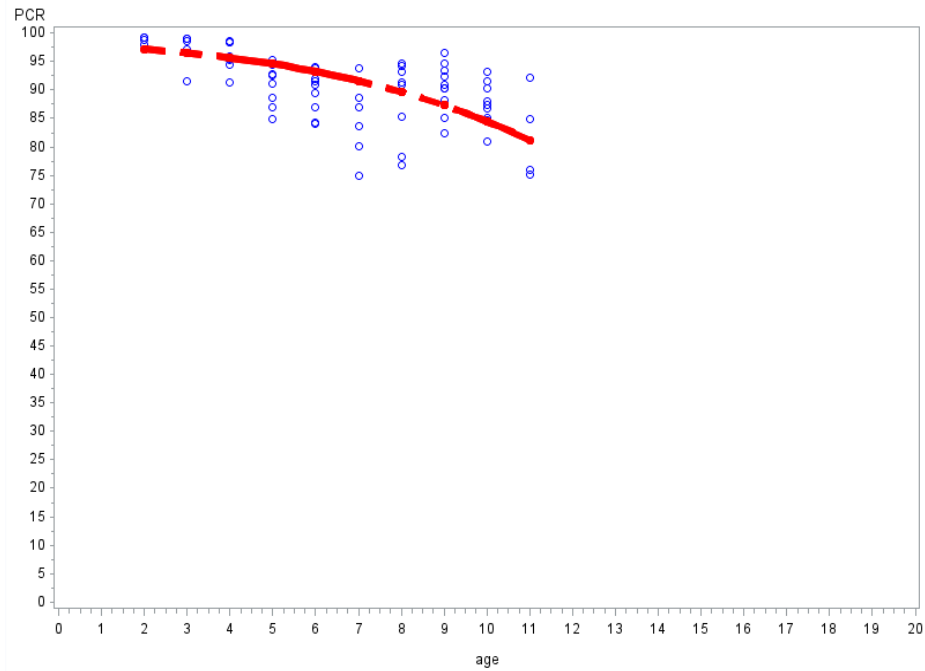


Figure 30: Performance model of NC

These scatterplots show that different pavement families have different performance deterioration trends (represented by red curve lines), and Interstate routes show the most aggressive deterioration trend. Estimated parameters, i.e.  $a$ ,  $b$ , and  $c$ , of these performance models were summarized in Table 2.

Table 2: Parameters of performance models

Family	$a$	$b$	$c$
Interstate	100	14.2	-3.7
US_0-5K	100	15.9	-4.4
US_5K plus	100	15.1	-3.8
NC	100	17.3	-4.3

To compare the performance of pavements in these four families, four model curves are plotted in one chart and shown in Figure 31 below:

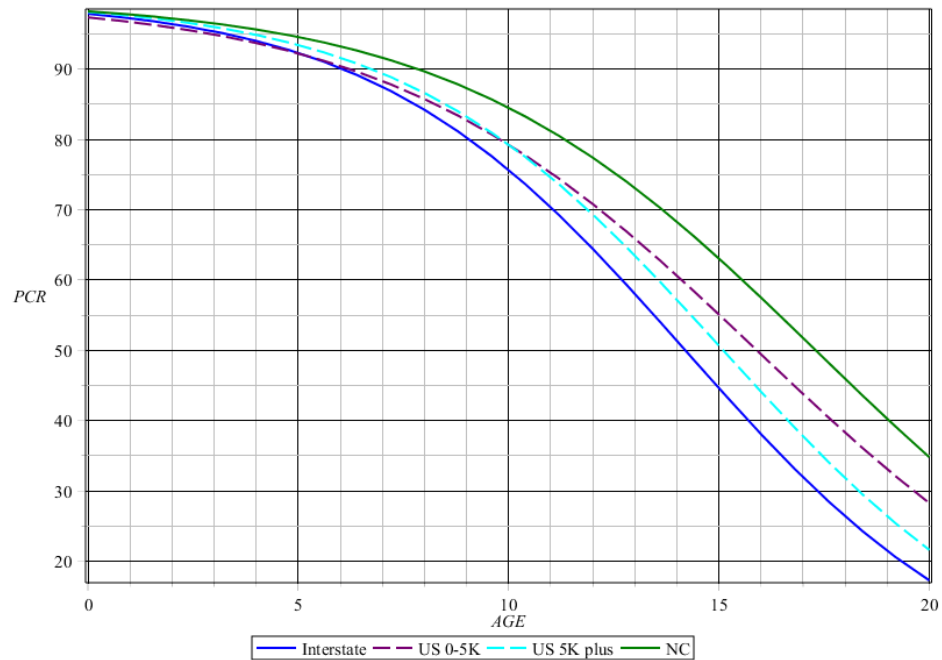


Figure 31: Performance models for pavement families

Figure 31 shows that in all the curves the PCR values reach 60 at ages around 12.7, 14.0, 13.5 and 15.5 for Interstate, US\_0-5K, US\_5K plus and NC, respectively. This indicates that NC CP sections performed better than US CP sections, and US CP sections performed better than Interstate CP sections. This was probably caused by different traffic volumes undertaken by these CP sections. Interstate CP sections carried the largest amounts of traffic among three roadway families, and thus deteriorated faster than the other two roadway families.



Random effect performance models were also developed based on the sigmoidal model, and the model parameters are included in Table 3. A comparison between the nonlinear performance models and the random effect performance models indicated that these two groups of curves were quite different, except the Interstate family. However, based on the random effect models, similar conclusions can be drawn according to relative deterioration rates. Due to the simplicity, sigmoidal models were emphasized in this study and recommended for future implementation.

Table 3: Parameters of random effect performance models

Family	$a$	$b$	$c$
Interstate	100	14.5	-4.0
US_0-5K	100	20.2	-7.0
US_5K plus	100	18.5	-6.3
NC	100	21.5	-6.4

## CHAPTER 5 DEVELOPMENT OF QUANTILE REGRESSION MODELS

In this chapter, the development of quantile regression (QR) pavement performance models is presented, and the application of developed QR models is discussed at both project and network levels.

### 5.1 Model Development and Prediction at the Project Level

At the project level, QR models was used to predict the performance of a specific pavement section. Due to the existence of large variations and heterogeneity in pavement performance data, pavement sections within the same pavement family have quite different deterioration process. However, since the performance of a roadway section is highly related to its previous condition (Chu & Duango-Cohen, 2007), pavement performance can be predicted using historical data by selecting the appropriate conditional distribution curve. The first section below provides the general description about the algorithm and the process of developing QR performance models at the project level. The following section discusses the accuracy of predictions.

At the project level, QR models were developed for each pavement family, i.e., Interstate, US\_0-5k, US\_5K plus, and NC. The performance data used to develop the network level QR models was the same one that was used to develop nonlinear distress and performance models (Chapter 4). The same performance data was expanded to include roadway sections that have two observations in order to develop the project level QR

models because the sample size of the existing data set was too small to develop a robust QR model.

Figure 32 shows the process of model development and performance prediction at the project level. There are six steps in this process: data preparation, data splitting, error measurement, model development, prediction and model selection, and improvement.

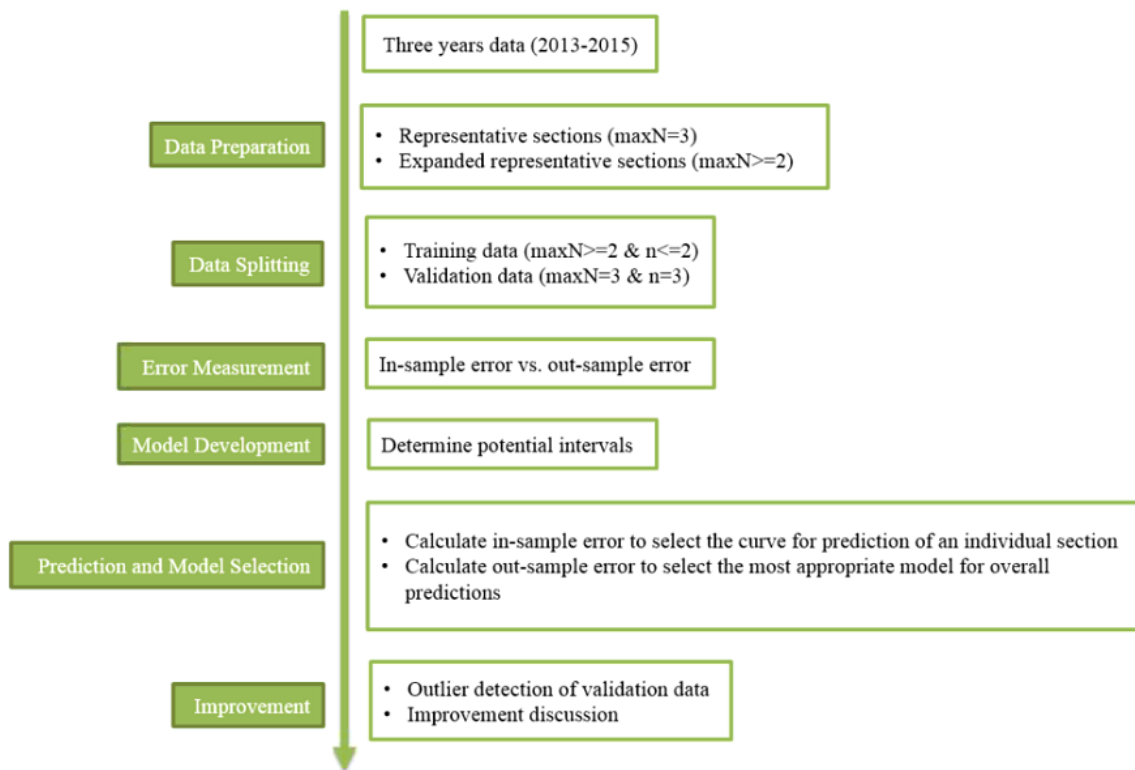


Figure 32: Process of QR model development and prediction at the project level

#### *Step One: Data preparation*

In this step, the expanded representative roadway sections were selected based on the following assumptions:

- Ages should be less than or equal to 13 to avoid the repeated distribution pattern;
- Pavement sections should have two or three observations with consecutive ages; and
- PCR values decrease over time.

These assumptions are slightly different than those of in Chapter 4 in that, instead of having three observations, pavement sections that have two observations were also included in order to obtain larger sample sizes.

#### *Step Two: Data splitting*

In this step, the expanded data were split into training (in-sample) and validation (out-sample) data sets. The training data set was used to develop QR models, and the validation data set was used to select the most appropriate QR model.

#### *Step Three: Error measurement*

In-sample error is also known as the fitting error, which is calculated using the in-sample data and indicates how well the model fits the data. Out-sample error is also known as the prediction error, which is calculated using the out-sample data and used to select the most appropriate prediction model.

In this research, errors were measured by Mean Absolute Percentage Error (MAPE). The MAPE equation can be written as:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|y - \hat{y}|}{y} \quad (3)$$

where  $y$  is the observation;  $\hat{y}$  is the predicted value;  $n$  is the total number of predictions.

The lower the MAPE, the better performance the model has.

MAPE is not an appropriate indicator when it is used to measure the error when the actual observation is close to 0. This was not the case in this study because typically all the PCR values were greater than 40 in the first 12 years. Besides MAPE, the Mean Squared Error (MSE) and the Mean Absolute Error (MAE) can also be used for measuring model errors. MSE can be calculated as:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2 \quad (4)$$

MAE can be calculated as:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y - \hat{y}| \quad (5)$$

These three equations indicate that MSE and MAE are determined by the squared errors and absolute errors, respectively, and they can range from 0 to an infinite value. Whereas MAPE is determined by the standardized errors. It ranges from 0 to 1, and is unit-less. These make MAPE a much better error indicator than MSE and MAE. Therefore, MAPE was selected to measure the model errors in this study.

#### *Step Four: Model development*

The sigmoidal performance model (Equation 1) was used as the base model to develop QR models that have multiple curves with the same percentile interval. Different potential percentile intervals were used to develop multiple QR models. For example, if the 5 percentile interval is used, then there will be 19 resulting curves for each family. Each curve represents a different percentile.

#### *Step Five: Prediction and model selection*

In this research, in-sample error was used to select the most appropriate curve within a QR model to predict the performance of each individual section. Then out-sample error was used to select the most appropriate QR model for overall predictions and to evaluate the accuracy of predictions.

To predict performances using a QR model, it is important to ensure each section in the validation data set has enough historical data (at least two known historical observations). To satisfy this requirement, more constraints were applied to data splitting: all the third observation of roadway sections that have three observations were put into the validation data set, and the rest of the data were put into the training data set.

#### *Step Six: Model improvement*

After out-sample errors were calculated, outliers of the validation data were examined using different MAPE (out-sample) thresholds, i.e. 0.1, 0.15, 0.2, and 0.3, and

the overall MAPE without outliers was calculated. By investigating the outliers, suggestions for improving the future model were discussed.

A hypothetical example is provided in Figure 33 to illustrate the procedure to predict pavement performance using QR models. The QR model in this figure has five curves. The black dots represent hypothetical historical data points of a pavement section of interest, and the red dots represent its predicted future performance index values. From Figure 33, it can be observed that among all five curves, the 50<sup>th</sup> percentile curve is the most appropriate curve to predict the performance trend. To verify this observation, in-sample MAPE calculated based on these black dots is used as a more precise measurement to select the most appropriate curve for each pavement section. This selected curve is then used to predict future performance index values (red dots). Lastly, the out-sample MAPE is calculated based on the predicted values (red dots) and real performance index values (which are not shown in this figure) of all sections, which is used to evaluate the overall prediction performance of a QR model.

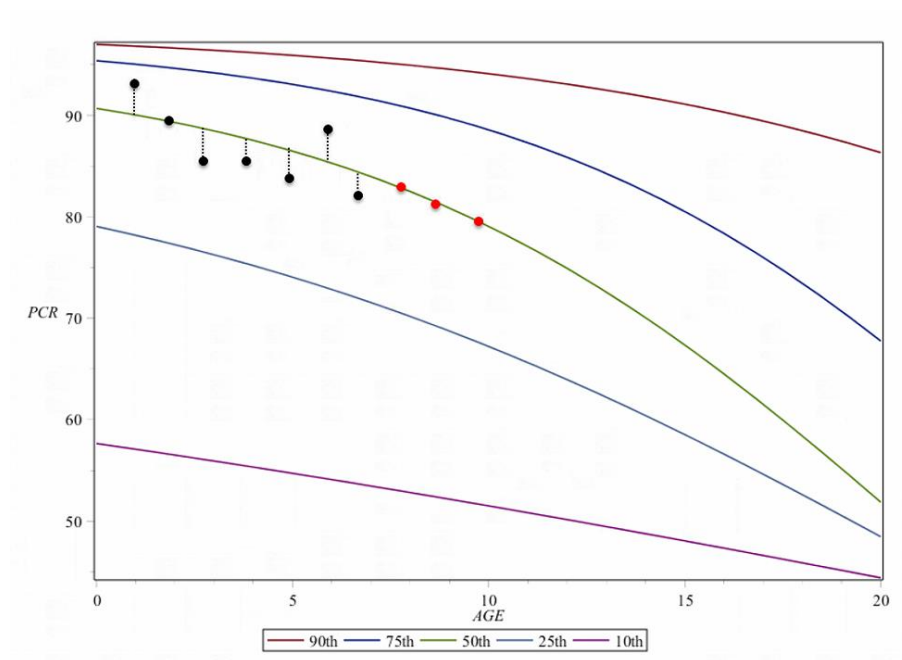


Figure 33: A hypothetical example of using QR models at the project level

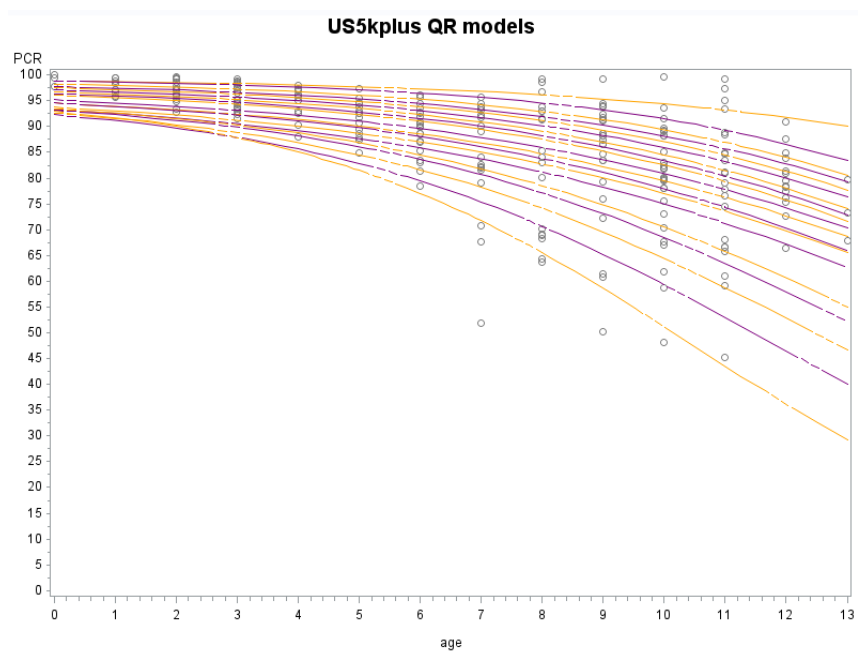


Figure 34: QR model for US\_0-5K plus at the project level



Figure 34 shows the QR model for US\_0-5k plus. This model were developed using a percentile interval of 5, thus 19 QR curves were obtained. Figures for other families can be found in Appendix A. As shown in Figure 34, at each age, there are 19 different performance predictions, indicated by 19 different curves. Two colors of yellow and purple were used to distinguish adjacent curves. The in-sample MAPE of the validation section's first two observations (which were included in training data set) were calculated. The curve that provides the smallest MAPE value is selected as the curve to predict the third observations.

Table 4 shows results of in-sample and out-sample MAPE values for each pavement family. It can be observed that in-sample errors are always smaller than out-sample errors, because the model was generated using the training data by minimizing the in-sample error. The total number of predictions is 130 with the overall MAPE (out-sample) value of 8.00%, which indicates that generally, the predicted value is about 8.00% different from the real value.

Table 4: In-sample and out-sample errors

Family	Training data set		Validation data set	
	Sample size	MAPE (In-sample)	Sample size	MAPE (Out-sample)
Interstate	120	3.26%	31	11.14%
US_0-5K	120	1.67%	32	8.24%
US_5K plus	224	2.50%	43	8.41%
NC	88	2.45%	24	2.76%

To obtain a better QR model, two additional percentile intervals of 10 and 2.5 were examined, and therefore, a total of three models were generated using percentile intervals of 10, 5 and 2.5, respectively. Outliers of the validation data were defined with four different thresholds, the number of outliers and the corresponding MAPE values without outliers were calculated, as shown in Table 5. For example, in the second row of Table 5, the threshold of “MAPE>0.30” indicates that predicted values that have MAPE (out-sample error) greater than 0.3 were considered as an outlier, and the model with the 5 percentile interval had 4 outliers. After removing outliers, the overall MAPE value is 6.52%.

Table 5: Prediction errors (out-sample) with different intervals

Threshold	10		5		2.5	
	Number of outliers	Error (out-sample)	Number of outliers	Error (out-sample)	Number of outliers	Error (out-sample)
(All)	0 (130)	8.58%	0 (130)	8.00%	0 (130)	7.91%
MAPE>0.30	4	7.14%	4	6.52%	4	6.40%
MAPE>0.20	12	5.96%	13	5.15%	12	5.14%
MAPE>0.15	23	4.81%	20	4.39%	20	4.28%
MAPE>0.10	38	3.68%	34	3.34%	35	3.13%

Table 4 shows that as the percentile interval decreases, the corresponding MAPE values decreases, which means the smaller the percentile interval, the more accurate the prediction. However, there was no obvious improvement as the percentile interval changed from 5 to 2.5. Therefore, to choose a simpler model, the percentile interval of 5 was still used in this research. For the model with a 5 percentile interval, there are 34 prediction out

of 130 sections have MAPE greater than 10%. Excluding these extreme large MAPE values, the overall MAPE is 3.34%, which is calculated based on 73.8% of all predictions. This indicated the QR model is robust.

As a result, each pavement family has a QR model with a percentile interval of 5. Future performance can be predicted using the corresponding pavement family's QR model. The prediction process is the same as what was discussed in Step Five in the previous section: a pavement section's historical data are used to select the most appropriate curve residing in a pavement family's QR model, and the future performance of this section is predicted using the selected curve.

Poor predictions with MAPE values greater than 20% were examined. It was found that those poor predictions were caused by the sudden drop of the third PCR value (which was included in the validation data set), and this drop could not be captured from the previous two PCR values (which were included in the training data set). In the future, the prediction can be improved by using performance data with a longer history.

### 5.3 Model Development and Application at the Network Level

At the network level, the probability that pavement performance falling into a certain performance range can be obtained by investigating the upper and the bottom quantile boundaries. For instance, using a hypothetical example as shown in Figure 35, the probability of pavement performance that lies between the boundaries of 25th and 75th percentile curves is 0.5. Similarly, the probability of pavement performance that lies between the boundaries of 10th and 90th percentile curves is 0.8. More probabilities of

pavement performance locations can be found in Figure 35. This information captured by the QR can help decision-makers to develop more reliable strategies by considering the uncertainties of pavement performances.

Different than the project level where 19 QR curves were generated to achieve a high prediction accuracy, at the network level, five curves are sufficient because the purpose is to provide the probability of achieving a pavement performance goal to pavement managers. In this study, representative sections that have three observations were used to develop QR models at the network level. The models for four pavement families are shown in Figures 36 through 39. The same QR models overlaid with scatterplots can be found in Appendix B.

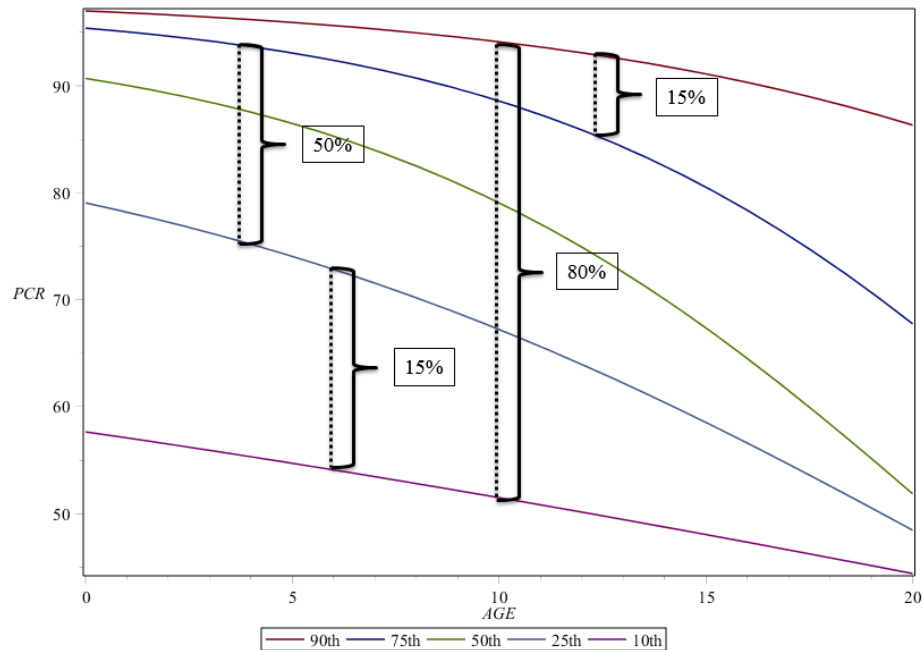


Figure 35: A hypothetical example of using QR models at the network level

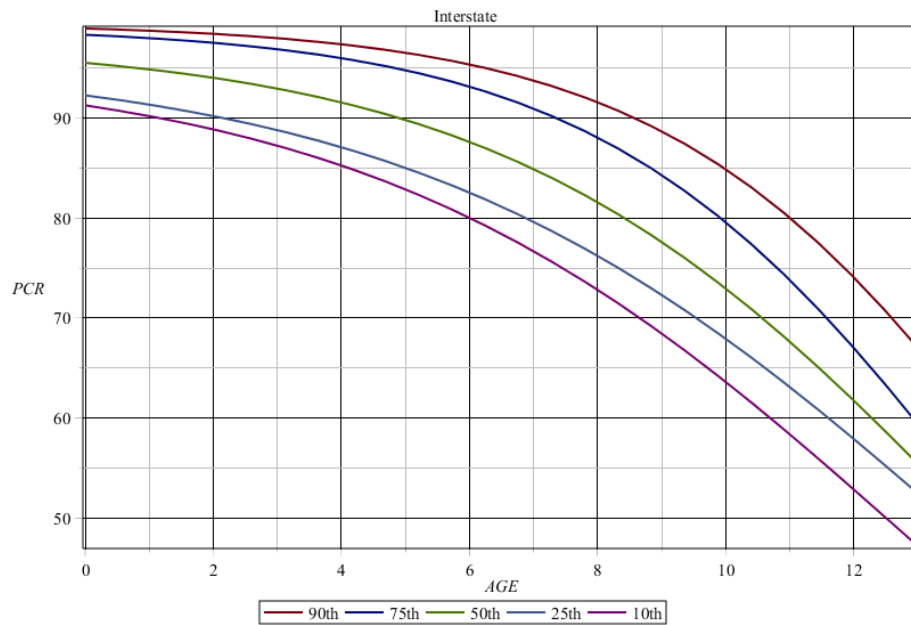


Figure 36: Interstate QR model at the network level

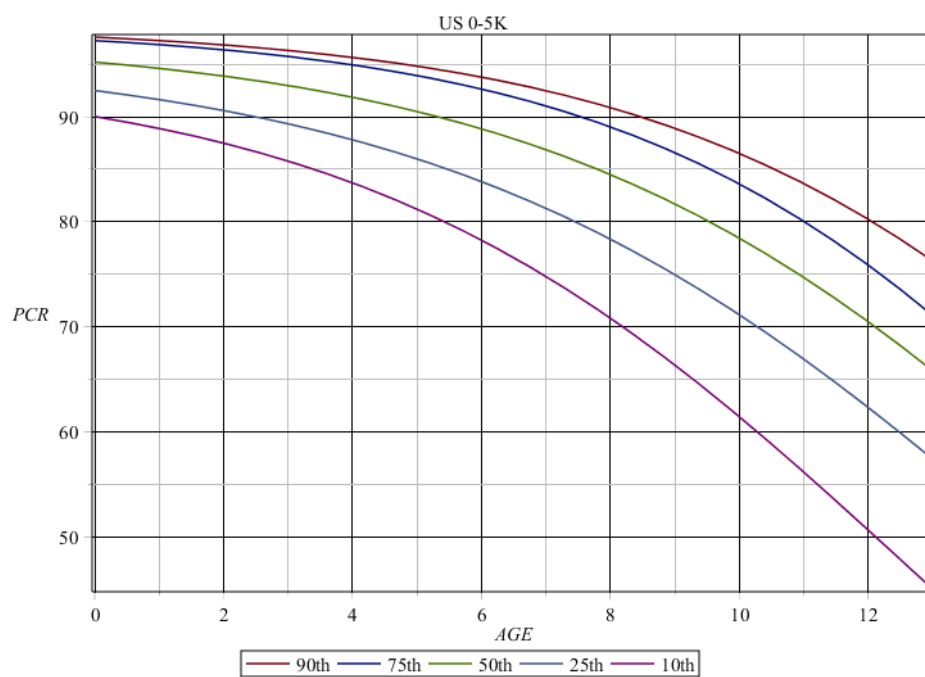


Figure 37: US 0-5K QR model at the network level

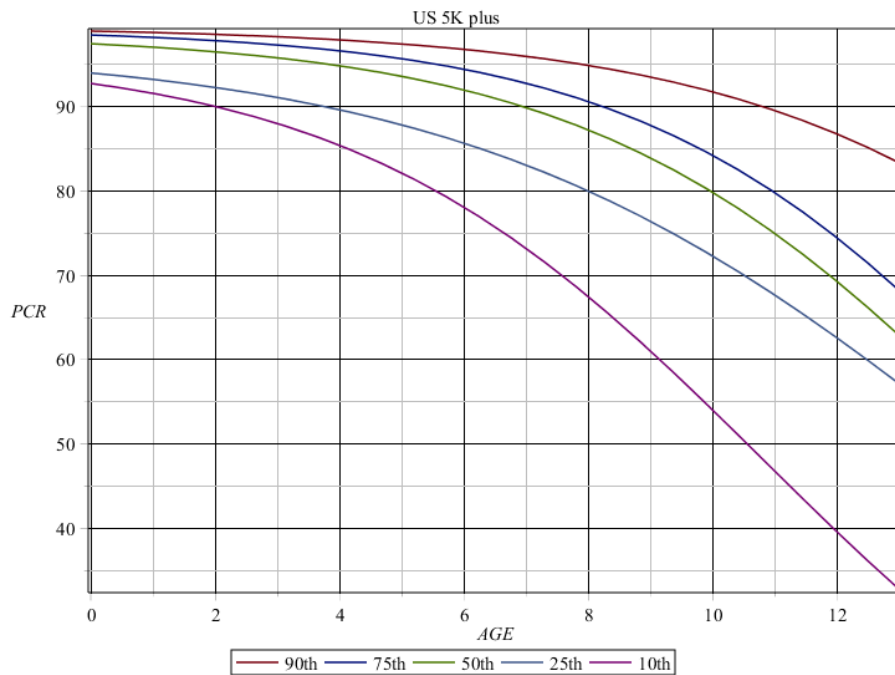


Figure 38: US 5K plus QR model at the network level

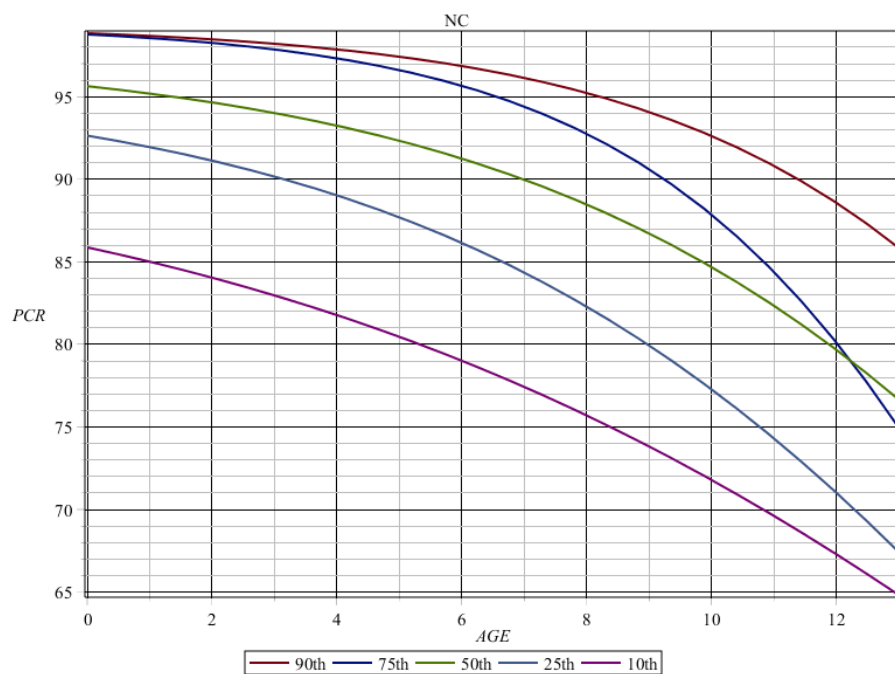


Figure 39: NC QR model at the network level

In these figures, the green curve represents the 50<sup>th</sup> percentile curve, which means 50% of the sections have performance index values above that curve, and 50% of the sections are below that curve. Over time, the performance ranges of US pavement families increase, while Interstate and NC routes have relatively constant ranges. Overall, NC routes remain in good conditions longer than the other two pavement classifications. Theoretically, the 75<sup>th</sup> percentile curve in the NC QR model (Figure 31) should not cross the 50<sup>th</sup> percentile curve. However, these two curves cross each other at age of 12 because of the insufficient data at older ages.

## CHAPTER 6 TREATMENT EFFECTIVENESS

To maintain the large amount of existing composite pavements with limited budget, transportation agencies need to seek effective ways to manage pavement networks. Calculating the net performance jump for each treatment can quantify the improvement in the pavement condition and the effectiveness of treatments. This chapter consists of definitions of performance jumps, analyses of descriptive statistics and visualizations, and comparisons of treatment effectiveness.

### 6.1 Definitions of performance jumps

In this section, the windshield composite pavement data, which was extracted from the original windshield asphalt data, was used for calculating jumps due to the long treatment history. Since the windshield data includes a sufficient amount of data, five pavement families were investigated: Interstate, US\_0-5K, US\_5K plus, and NC.

In this study, The PCR rating right before the treatment is defined as the pre-treatment PCR, which indicates the pavement condition that triggered the treatment. the PCR rating right after the treatment is defined as the post-treatment PCR, which indicates the pavement condition after the treatment. The difference between the pre- and the post-treatment PCRs represents the treatment effectiveness (Equation 3), which is also referred to as a jump in the following discussions.

$$\text{Jump} = \text{Post\_PCR} - \text{Pre\_PCR} \quad (3)$$



where Post\_PCR represents the post-treatment PCR, Pre\_PCR represents the pre-treatment PCR.

An example is shown in Figure 40. The “real\_n” column stands for the number of treatments that have been applied to the pavement section since it was converted to a composite pavement section. When the “real\_n” equals 0, it means after the section becomes composite, it has not been treated yet. Similarly, when the “real\_n” increases by one unit, it indicates that the section was treated sometime between the years of the “Pre\_PCR” and the “Post\_PCR”.

In this example, this particular section has PCR records from the years of 1982 to 1998. During this period, the “real\_n” increased from 0 (in 1986) to 1 (in 1998). Therefore, in this case, the PCR values of 79.2 (in 1986) and 95 (in 1998) are the pre- and the post-treatment PCRs, respectively. The difference is calculated as the jump that represents treatment effectiveness, which is 15.8.

In Figure 40, the “real\_age” is the pre-treatment age that indicates the time from the previous treatment to the current treatment. For example, in Figure 32, the “real\_age” value of 13 (in 1986) is the pre-treatment age, which indicates that this section was treated at the age of 13, meaning the time from the previous treatment (in 1973) to the current treatment (in 1987) was 13 years. The “Work Code” column includes various treatment types, such as “AC Construction/Reconstruction”, “Chip Seal”, and “Resurfacing”.

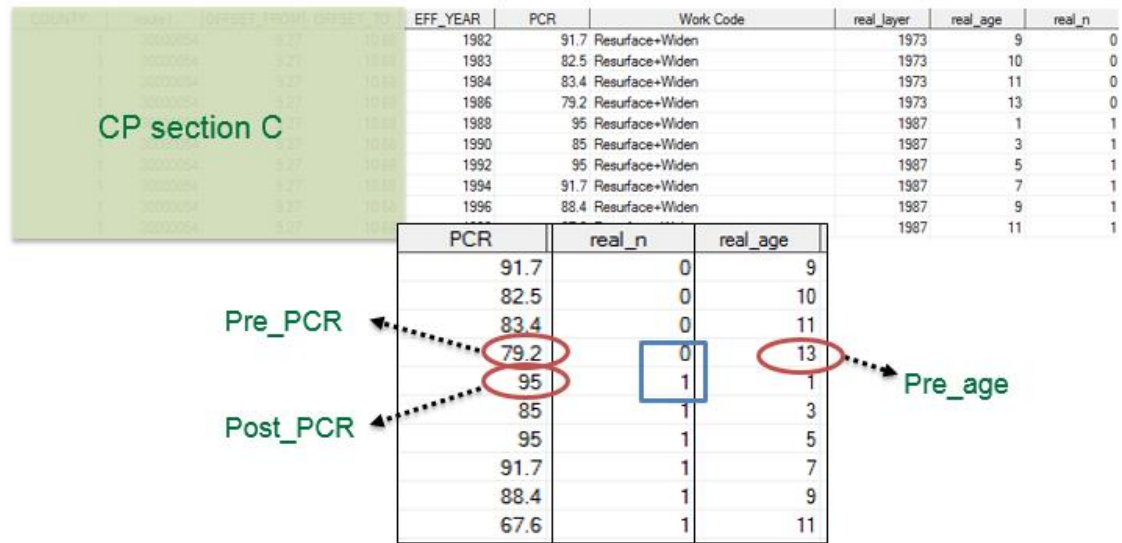


Figure 40: Calculation of a jump

## 6.2 Analyses of Descriptive statistics and Visualizations

All jumps were calculated using Equation 3 and were used in the following analyses. Due to errors discussed in Chapter 2, outliers also existed in calculated jumps. To cleanse the data, jumps that were less than 0 or had a pre-treatment PCR greater than 90 were removed. Figure 41 shows the distribution of jumps with an interval of five. It can be observed from this histogram that jumps ranging from 15 to 20 had the highest frequency. Table 6 shows the average value of jumps is 34.86, which means on average a treatment can increase the PCR value of a roadway section by 34.86.

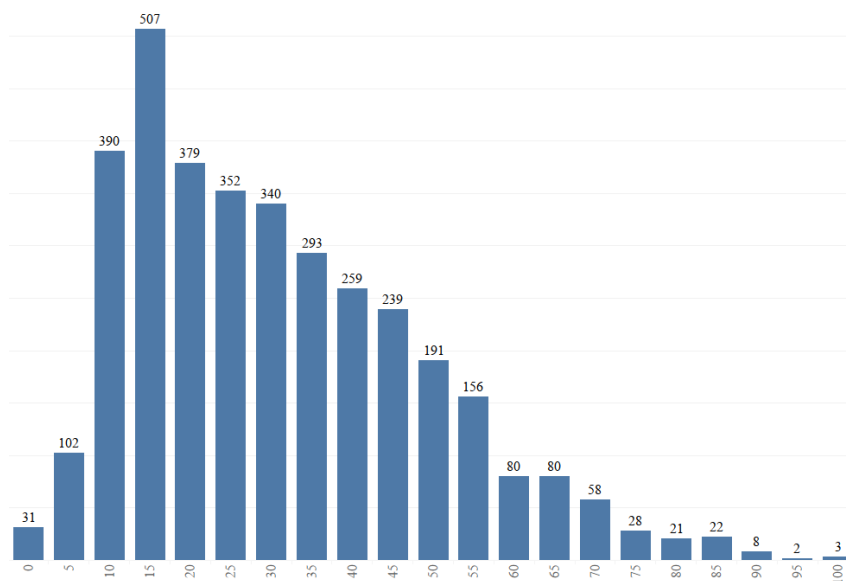


Figure 41: The histogram of jumps

Table 6: Descriptive statistics for jumps

Analysis Variable : jump				
N	Mean	Std Dev	Minimum	Maximum
3295	34.8594234	17.4623073	10.4000000	100.0000000

To explore pavement conditions that triggered treatments, descriptive statistics and the distribution of the pre-treatment PCR were obtained and are shown in Table 6 and Figure 42. Table 7 indicates that the average pre-treatment PCR was 61.9. This means that on average, treatments were applied when the pre-treatment PCR value reached 61.9. The accumulative curve of pre-treatment PCR was developed, as shown in Figure 43. The curve indicates that the 25<sup>th</sup> percentile of the pre-treatment PCR is around 50, the median is around 65, and the 75<sup>th</sup> percentile is around 75. This means that 25 percent of pavements

were treated when their pre-treatment PCR values were less than 50, 50 percent of the pavements were treated when their pre-treatment PCR values were less than 65, and 75 percent of pavements were treated when their pre-treatment PCR values were less than 75. The median pre-treatment PCR value of 65 is consistent with NCDOT maintenance practices.

Table 7: Descriptive statistics for pre-treatment PCRs

Analysis Variable : pre_pcr				
N	Mean	Std Dev	Minimum	Maximum
3295	61.9080728	17.4233949	0	89.2000000

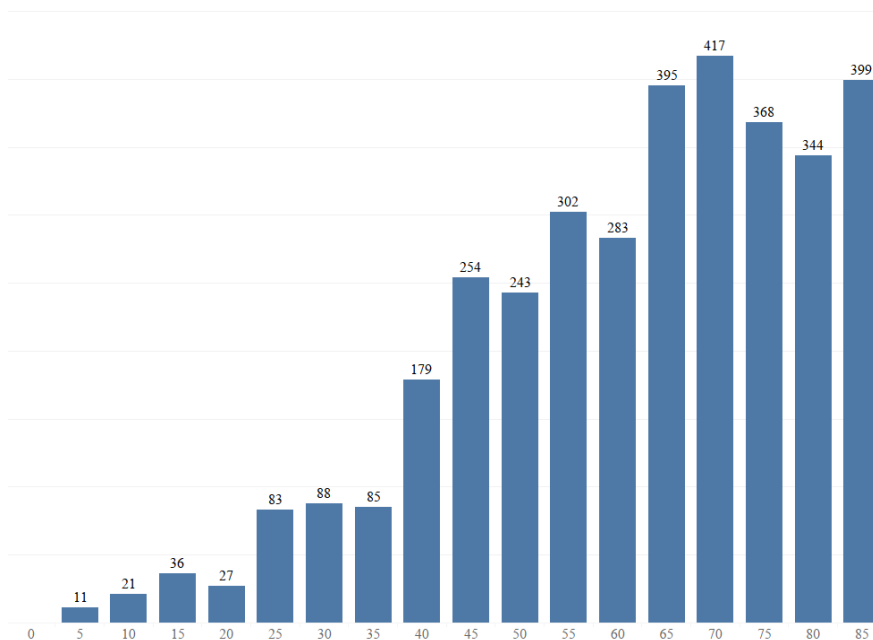


Figure 42: The distribution of pre-treatment PCRs

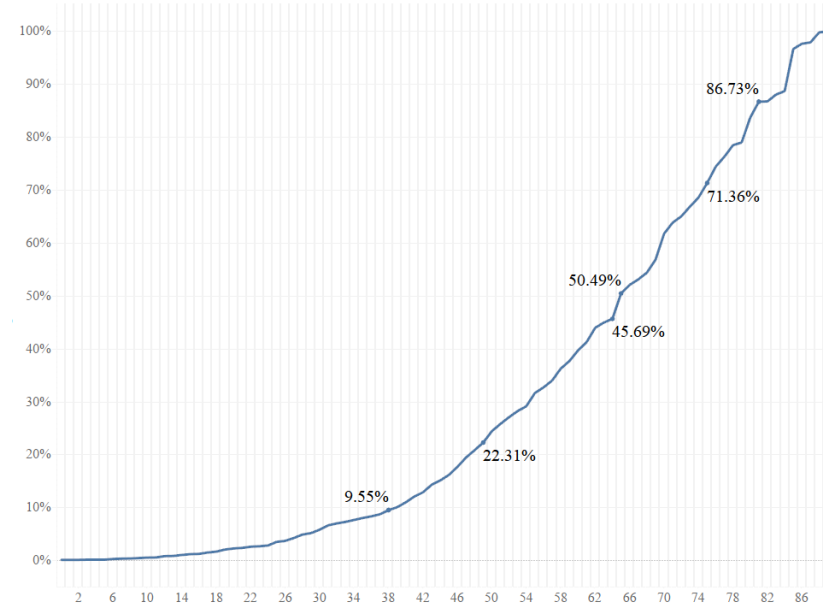


Figure 43: The accumulative curve of the pre-treatment PCR

Similarly, to investigate ages at which pavements were treated, descriptive statistics and the histogram of the pre-treatment age were also obtained and are shown in Table 8 and Figure 44. The mean value of the pre-treatment age was about 13 years, and the most frequent pre-treatment age was 13 as well. That means on average, pavements were treated when they were at age of 13. Furthermore, based on the accumulative curve shown in Figure 45, the pre-treatment age has the 25th percentile of 7, the median of 11, and the 75th percentile of 15. This means that 25 percent of the pavements were treated before the age of 7, 50 percent of the pavements were treated before the age of 11, and 75 percent of the pavements were treated before the age of 15. The median pre-treatment age of 11 is also consistent with NCDOT maintenance practices.

Table 8: Descriptive statistics for pre-treatment ages

Analysis Variable : pre_age				
N	Mean	Std Dev	Minimum	Maximum
3295	12.7690440	7.7190191	1.0000000	63.0000000

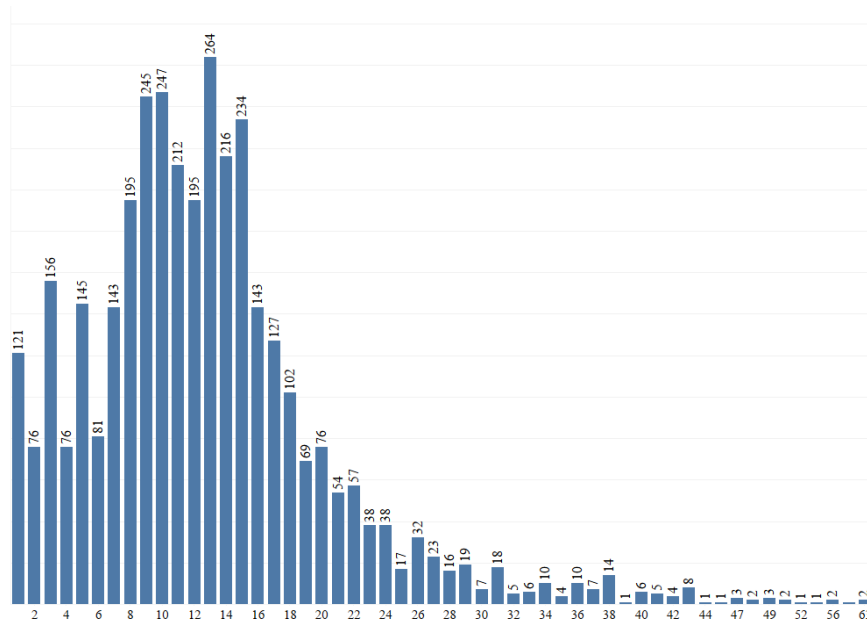


Figure 44: The histogram of pre-treatment ages

To investigate the relationships among the pre-treatment PCR, the post-treatment PCR, the jump, and the pre-treatment age, a visualization graph, called the parallel coordinate, was developed, as shown in Figure 46. In this figure, four vertical axis represent the post-treatment PCR, the pre-treatment PCR, the jump, and the pre-treatment age (multiplied by 5 for the visual clarity purpose), respectively. These four values of the same roadway section were connected with a line using the same color. In this visualization, colors were assigned based on the pre-treatment PCR values with an interval value of five.

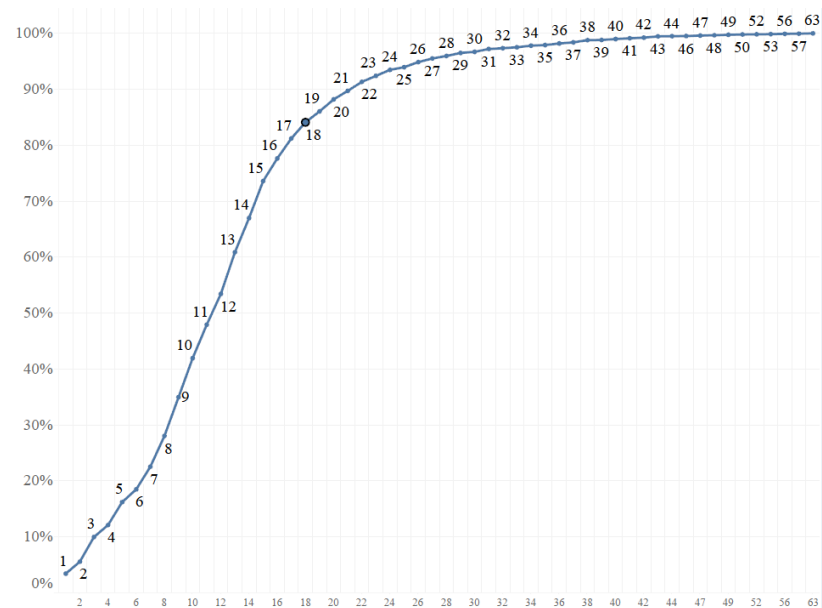


Figure 45: Accumulative curve of pre-treatment ag2

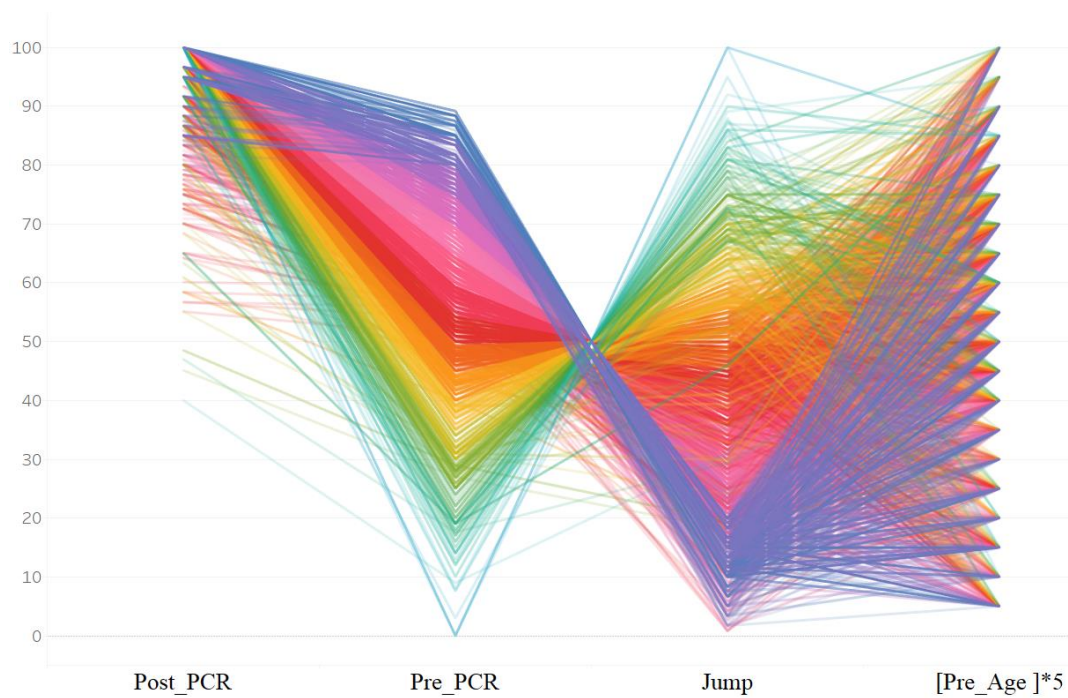


Figure 46: Parallel coordinate plot

It can be noticed that the majority of treatments achieved high Post\_PCR values ranging from 90 to 100 regardless the magnitudes of the Pre\_PCR values. Even though some Post\_PCR values were less than 90, they were still greater than the corresponding Pre\_PCR values because all the connecting lines were higher at the Post\_PCR axis. It can also be noticed that the lines connecting the jumps and the Pre\_PCR values had an “X” pattern, indicating a negative relationship between these two values. This was because right after a treatment no distresses should be observed in the pavement, and the Post\_PCR should remain close to 100. Therefore, if a pavement was treated when it had a lower pre-treatment PCR, the jump in PCR after the treatment would be greater than those that were treated with a higher pre-treatment PCR. On the axis of “pre\_age\*5”, pre-treatment PCR values of 60 (represented by red lines) were spread out evenly over the entire age range, which indicates that in general a roadway section was treated when its pre-treatment PCR reached 60, regardless of age.

### 6.3 Comparison of Treatment Effectiveness

To further investigate treatment effectiveness, comparisons among different treatment types and pavement families were conducted. Figures 47 and 48 show the boxplots of the post-treatment PCR, the pre-treatment PCR, the jump, and the treatment age, which were categorized by treatment types and pavement families. Figure 47 shows that the sample sizes of treatment “Mill+Resurface” were small, and the other two different treatments, “Chip Seal” and “Resurface”, had similar average values of Pre\_PCR, post\_PCR, jumps, and age. Figure 48 shows that Interstate routes had a greater jump, which



means the average treatment effectiveness of Interstate routes was greater than other pavement families.

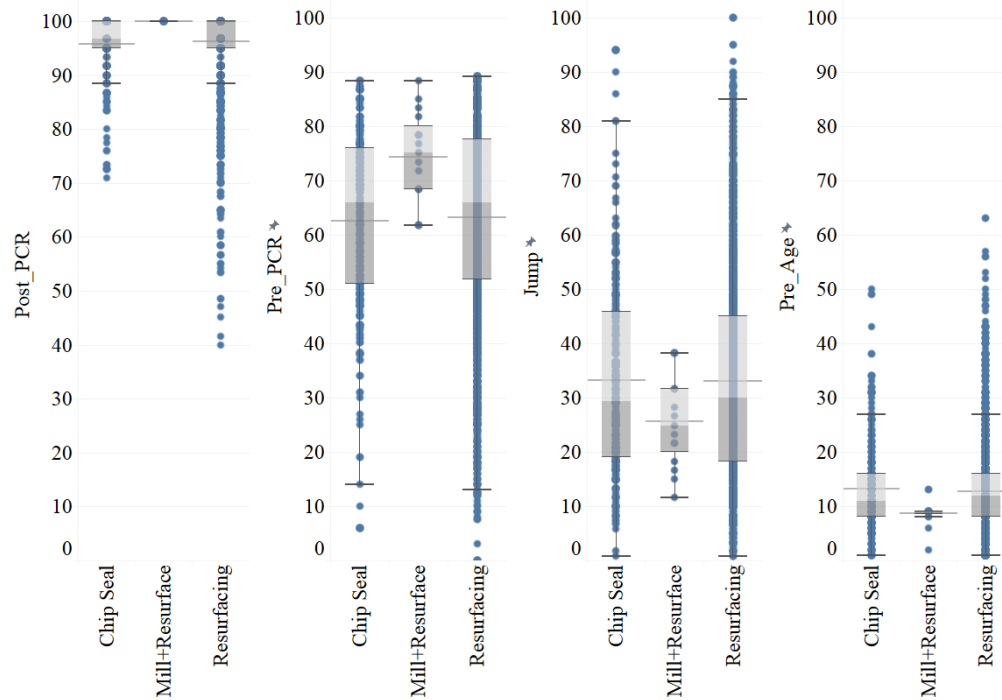


Figure 47: The boxplot of treatment effectiveness based on treatment types

The similarities among pavement families and treatment types were caused by the fact that all post-treatment PCR values were set to 100 regardless of families and treatment types. However, the real post-treatment PCR values could be different than 100 and varied based on treatment types. Right after treatments, even though no distresses should be observed, after a short period of time (one or two years), performance conditions could deteriorate differently in different pavement sections that were treated with different methods.

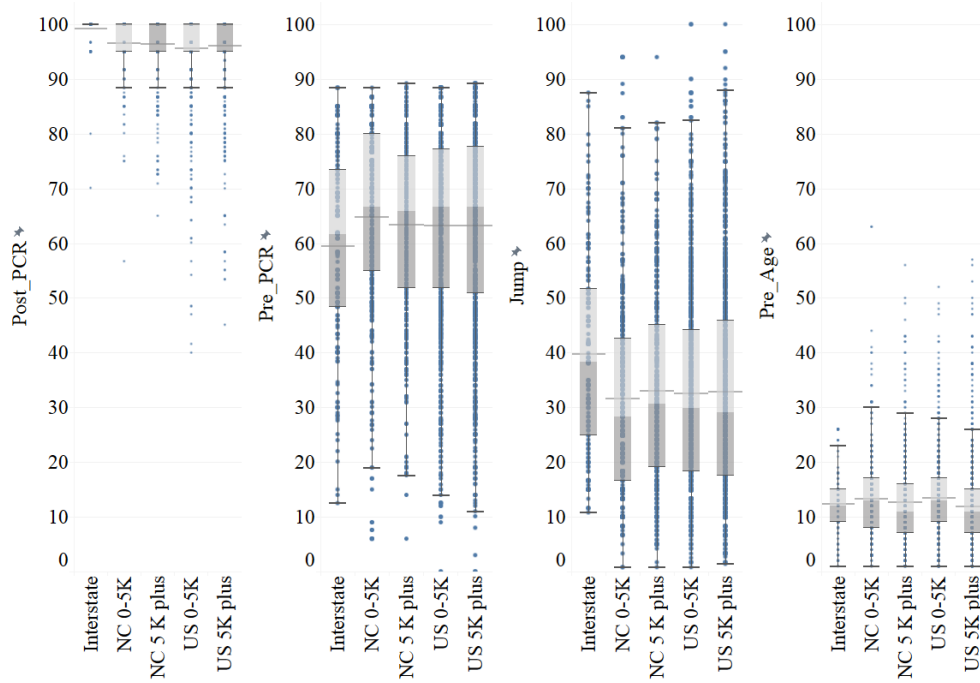


Figure 48: The boxplot of treatment effectiveness based on pavement families

To estimate the real post-treatment PCR values, the same sigmoidal model form (Equation 1) was used.

$$y = \frac{a}{1 + e^{\frac{-x+b}{c}}} \quad (1)$$

where  $y$  is PCR;  $x$  is age;  $a$ ,  $b$ , and  $c$  are model parameters.

The sigmoidal equation can be further converted to a linear equation (Equation 2) (Chen et al., 2014).

$$-\frac{x-b}{c} = \ln(e^{\ln a - \ln y} - 1)$$

Let  $Y = \ln(e^{\ln a - \ln y} - 1)$ , then  $-\frac{x-b}{c} = Y$ .

Therefore,

$$Y = -\frac{1}{c}x + \frac{b}{c} \quad (2)$$

To estimate the parameters of the sigmoidal model, the initial value of parameter  $a$  was assigned to 100. With the assigned initial value of  $a$ , the parameters of  $b$  and  $c$  can be easily estimated by running the simple linear regression analysis using Equation 2. To estimate real post-treatment PCR, the nonlinear regression model (Equation 1) was developed by setting the parameters of  $a$  and  $b$  changeable and fixing the parameter  $c$  that was obtained from the simple linear regression analysis. The intercept of the resulting model is the real post-treatment PCR. The windshield historical data of composite pavements, from age of 1 to age of 12, were used to estimate the intercept (real post-treatment PCR or initial PCR at age of 0).

Table 9 shows the calculated intercepts for pavement families and treatment types. The results indicate that for “Mill+Resurface”, only the Interstate family had an adequate sample size, and that its real post-treatment PCR was slightly greater than that of the “resurface” treatment. Because of the small sample sizes of “Mill+Resurface” for the other families, a comparison of treatment effectiveness was conducted between the treatments of “Chip Seal” and “Resurface”. Table 9 indicates that four out of five families had greater real post-treatment PCR values for “Resurface” than that of “Chip Seal”. It can be concluded that generally the “Resurface” treatment is more effective than “Chip Seal”.

Table 9: Intercepts of performance curves

Family	Treatment	Intercept	Sample Size
Interstate	Chip Seal	95.05	107
	Resurface	93.70	2,128
	Mill+Resurface	94.13	258
US0-5K	Chip Seal	88.88	651
	Resurface	91.76	5,818
	Mill+Resurface	N/A	7
US5Kplus	Chip Seal	88.62	716
	Resurface	91.01	6,819
	Mill+Resurface	N/A	21
NC0-5K	Chip Seal	86.52	304
	Resurface	90.67	1,957
	Mill+Resurface	N/A	0
NC5Kplus	Chip Seal	87.73	778
	Resurface	90.50	2,165
	Mill+Resurface	N/A	13

The differences between real post-treatment PCR values and pre-treatment PCR values were calculated to represent the real treatment effectiveness. This difference is referred to as the real jump in this study. Figure 49 shows the boxplots of jumps and real jumps, categorized by treatment types and pavement families. It was noticed that all average values of real jumps were less than jumps. Overall, “Resurface” was a more effective treatment than “Chip Seal”. In addition, Interstate routes gained larger jumps than other roadway families.

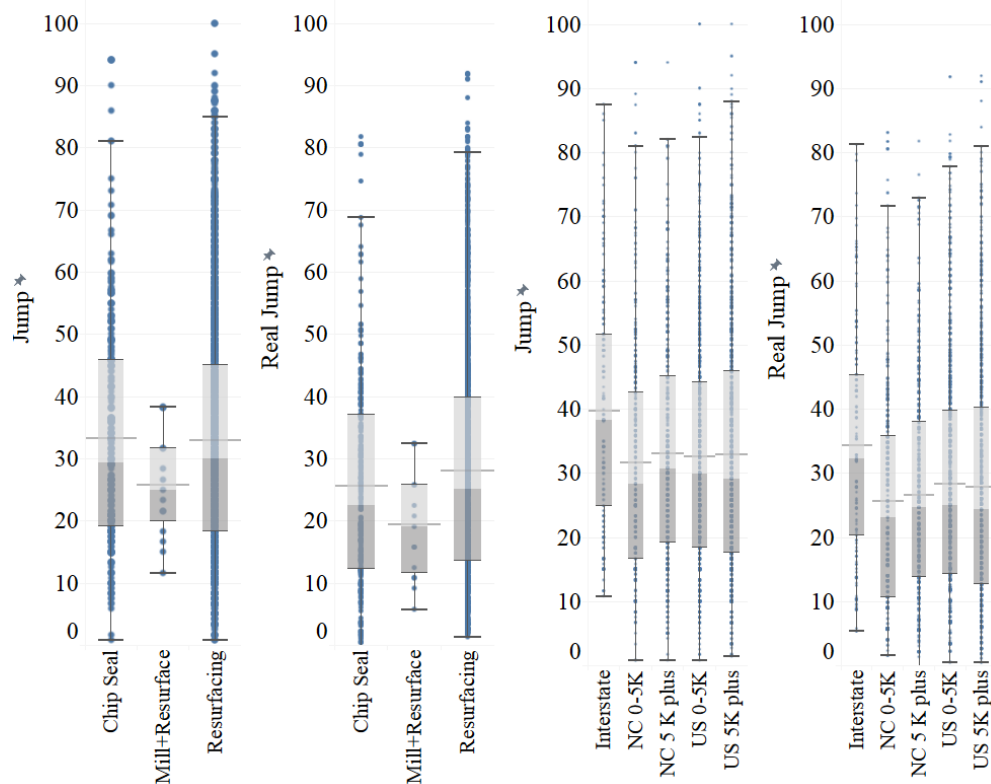


Figure 49: The boxplots of jumps and real jumps

#### 6.4 Comparison of pavement performance between Asphalt and composite pavements

The literature review indicated that asphalt pavements perform differently than composite pavements. This statement was validated in this section. A performance indicator, pavements' service lives, was used to compare the performance of asphalt and composite pavements. In this comparison, "real\_age" is the number of years from the first treatment (which is after the pavements were converted to composite pavements) to the second, representing the service life of a roadway section. Since only two groups of "real\_age", asphalt and composite pavements, need to be compared under each pavement family, a t-test of average service lives and their variations was used to investigate whether

the service lives were significantly different from each other. The test result is shown in Table 10. In this table, ASP represents Asphalt Pavements, and CP represents Composite Pavements.

Table 10: The result of t-test for “real\_age”

Pavement families	Number of records		Mean value		Mean p-value	Standard deviation		Std Dev p-value
	ASP	CP	ASP	CP		ASP	CP	
Interstate	78	53	8.51	8.98	0.589	4.90	4.82	0.9148
US0-5K	231	236	15.18	15.46	0.759	11.17	7.69	<.0001
US5Kplus	472	298	11.08	12.36	0.433	9.75	8.13	0.0002
NC0-5K	945	124	16.24	15.98	0.746	8.85	7.50	0.0208
NC5Kplus	229	115	14.50	14.28	0.821	9.59	7.41	0.0016

Given a confidence level of 0.05, in the “Mean p-value” column, all families have p-values much greater than 0.05, which means for all pavement families, the average service life of asphalt pavements was not significantly different from that of composite pavements. Among all pavement families, it can be concluded that Interstate highways were treated more frequently than other pavements.

In the “Std Dev p-value” column, four out five families have p-values less than 0.05, which means in most pavement families, the variation in the service lives of composite pavements was significantly different from that of asphalt pavements. In addition, in these four pavement families, the variation of the service lives of composite pavement is less than that of asphalt pavements. This indicates that the quality of composite

pavements is more consistent than that of asphalt pavements, and the pavement performance for composite pavements is more predictable. These advantages of composite pavements will benefit the resource allocation and the treatment scheduling.

In summary, according to the t-test results, no significant difference between the average service lives of asphalt and composite pavements was found. This means that the average time of composite pavements can last from the first treatment to the second is similar to that of asphalt pavements. However, for composite pavements, the service lives tended to have a smaller variation than asphalt pavements, which means after the first treatment, the majority of composite pavements were treated in a relatively shorter time range compared to asphalt pavements.

## CHAPTER 7 CONCLUSIONS, LIMITATIONS AND RECOMMENDATIONS

This chapter summarizes conclusions drawn from the findings, discusses limitations of this research, and provides recommendations for future studies.

### 7.1 Conclusions

This research was conducted to expand the PMS with composite pavements, propose systematic methods to improve the quality of performance data, enrich the diversity of prediction models by exploring potentials of QR models, and investigate the effectiveness of pavement treatments.

The quality of pavement performance data collected using the windshield method and the automated method was systematically investigated. It was concluded that the automated data were more precise and contained more information than the windshield data.

Various types of distresses in composite pavements were studied across pavement families. Generally, longitudinal cracking and transverse cracking were found to be the dominant distresses in composite pavements, followed by alligator cracking and raveling. However, dominant distresses for different pavement families may be different, and also, dominant distresses for the same pavement family may change over time.

Representative roadway sections were selected in order to develop robust performance models. Sigmoidal models were developed to describe the general



deterioration trend for pavement performance. From the developed performance curves, it was observed that Interstate composite pavements deteriorated faster than both US and NC composite pavements, and NC composite pavements had the slowest deterioration. This is probably because the performance of composite pavements is sensitive to traffic volumes. The higher the traffic volume, the faster the pavement deteriorates.

Quantile regression (QR) models was proposed as a new prediction method of pavement performance at both project and the network levels. At the project level, the prediction capability of QR models was evaluated using MAPE. The results indicated that QR models can predict pavement performance reasonable well. At the network level, QR models were proved to be able to provide the probability of pavement performance values, which is essential for decision makers to develop informative pavement management strategies.

Treatment history of composite pavements was examined using the windshield data. it was concluded that 25 percent of pavements were treated when their pre-treatment PCR values were less than 50, 50 percent of the pavements were treated when their pre-treatment PCR values were less than 65, and 75 percent of pavements were treated when their pre-treatment PCR values were less than 75. It was also concluded that 25 percent of the pavements were treated before the age of 7, 50 percent of the pavements were treated before the age of 11, and 75 percent of the pavements were treated before the age of 15. The median pre-treatment PCR value of 65 and pre-treatment age of 11 are consistent with NCDOT maintenance practices.

The effectiveness of treatments was investigated using performance jumps in the raw data and estimated performance jumps using nonlinear models. It was concluded that in general the “Resurfacing” treatment was more effective than the “Chip Seal” treatment. The average service life of asphalt and composite pavements were similar. However, for composite pavements, the service lives tended to have a smaller variation than asphalt pavements. This indicates that the quality of composite pavements is more consistent than that of asphalt pavements, and the pavement performance for composite pavements is more predictable. These advantages of composite pavements will benefit the resource allocation and the treatment scheduling.

## 7.2 Limitations of research

In this study, composite pavements only refer to pavements with a relatively thin asphalt surface layer over a concrete base, which are the results of concrete pavement rehabilitations. Performance of newly constructed composite pavements is beyond the scope of this research.

In this study, selected representative roadway sections were used to develop distress and performance models. Since the original automated data has only three years of data, the maximum number of performance observations of each roadway section is three. With this small amount of data, the developed models may not be able to provide an accurate performance prediction for each individual roadway section. However, the framework established by this study can be used for future studies when more data become available.

### 7.3 Recommendations

Several recommendations are provided in this section for future studies, as follows:

- (1) It is recommended that the automated data should be used in future PMS related research projects, due to its better data quality. Treatment effectiveness of composite pavements can be re-conducted when adequate automated data is available in the future. In addition, pavement performance prediction using the proposed QR models can be improved using automated data that has sufficient data.
- (2) It is recommended that classification or clustering methodologies can be used to identify the significant factors in order to group pavement sections. In this study, pavements were categorized into groups based on either pavement classifications, traffic volumes, or treatment types. Due to the limitations of available data, no studies were conducted to determine the most appropriate factor that should be used to group the pavement sections. This could be the reason why no obvious differences in performance and treatment effectiveness were found among some pavement families.
- (3) It is recommended to use a different performance indicator other than service lives, such as PCR, to compare performance of asphalt and composite pavements. This way more differences in performance of two pavement types might be discovered.
- (4) Because of the accurate predictions and the suitable applications of QR models at both network and project levels, it is recommended that QR model can be incorporated in the future PMS.

## BIBLIOGRAPHY

Abaza, K. A., Ashur, S. A., Abu-Eisheh, S. A., & Rabay'a, A. (2001). "Macroscopic optimum system for management of pavement rehabilitation." *Journal of transportation engineering*, 127(6), 493-500.

Abrevaya, J., & Dahl, C. M. (2008). "The effects of birth inputs on birthweight: evidence from quantile estimation on panel data." *Journal of Business & Economic Statistics*, 26(4), 379-397.

Abrevaya, J., & Dahl, C. M. (2008). "The effects of birth inputs on birthweight: evidence from quantile estimation on panel data." *Journal of Business & Economic Statistics*, 26(4), 379-397.

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

Bremnes, J. B. (2004). "Probabilistic wind power forecasts using local quantile regression." *Wind Energy*, 7(1), 47-54.

Canay, I. A. (2011). "A simple approach to quantile regression for panel data." *The Econometrics Journal*, 14(3), 368-386.

Chang Albitres, C., Smith, R., & Pendleton, O. (2007). "Comparison of automated pavement distress data collection procedures for local agencies in San Francisco bay area, California." *Transportation Research Record: Journal of the Transportation Research Board*, (1990), 119-126.

Chen, D., Cavalline, T. L., and Ogunro, V. O. (2014). "Development and validation of pavement deterioration models and analysis weight factors for the NCDOT pavement management system (phase I: windshield survey data)." Rep. No. FHWA/NC/2011-01, Federal Highway Administration (FHWA), Washington, DC.

Chen, D., Hildreth, J., Nicholas, T., and Dye, M. (2014). "Development and validation of pavement deterioration models and analysis weight factors for the NCDOT pavement management system (phase II: automated data)." Rep. No. FHWA/NC/2011-01, Federal Highway Administration (FHWA), Washington, DC.

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

Chu, C. Y., & Durango-Cohen, P. L. (2008). "Estimation of dynamic performance models for transportation infrastructure using panel data." *Transportation Research Part B: Methodological*, 42(1), 57-81.

Coad, A., & Rao, R. (2008). "Innovation and firm growth in high-tech sectors: A quantile regression approach." *Research policy*, 37(4), 633-648.

Darter, M. I., & Barenberg, E. J. (1976). *Zero-maintenance pavement: Results of field studies on the performance requirements and capabilities of conventional pavement systems* (No. FHWA-RD-76-105 Intrm Rpt.).

Davino, C., Furno, M., & Vistocco, D. (2013). *Quantile regression: theory and applications*. John Wiley & Sons.

DeStefano, P. D., & Grivas, D. A. (1998). "Method for estimating transition probability in bridge deterioration models." *Journal of Infrastructure Systems*, 4(2), 56-62.

Dimelis, S., & Louri, H. (2002). "Foreign ownership and production efficiency: a quantile regression analysis." *Oxford Economic Papers*, 54(3), 449-469.

Eide, E., & Showalter, M. H. (1998). "The effect of school quality on student performance: A quantile regression approach." *Economics letters*, 58(3), 345-350.

Federal Highway Administration (1996). "The National Highway System: a commitment to America's future." <http://www.fhwa.dot.gov/publications/publicroads/96spring/p96sp2.cfm> (Jun.10, 2016).

Federal Highway Administration (1997). "Pavement Management System – past, present, and future, public roads: 80 years old, but the best is yet to come." <http://www.fhwa.dot.gov/publications/publicroads/98julaug/pavement.cfm> (Jun.10, 2016).

Federal Highway Administration (1998). "Pavement preservation: a road map for the future." <http://www.fhwa.dot.gov/infrastructure/asstmgmt/roadmap.pdf> (Jun. 10, 2016).

Federal Highway Administration (2016). "Guidance to designing and constructing long-life composite pavement systems that provide durable, sustainable aggregate surfaces at a potentially lower cost than conventional pavement." [https://www.fhwa.dot.gov/goshrp2/Solutions/All/R21/New\\_Composite\\_Pavement\\_Systems](https://www.fhwa.dot.gov/goshrp2/Solutions/All/R21/New_Composite_Pavement_Systems) (Jun. 6, 2016).

Federal Highway Administration (2016). "What is the National Highway System?" [http://www.fhwa.dot.gov/planning/national\\_highway\\_system/](http://www.fhwa.dot.gov/planning/national_highway_system/) (Jun. 10, 2016).

Flintsch, G. W., Diefenderfer, B. K., & Nunez, O. (2008). "Composite pavement systems: Synthesis of design and construction practices."

Galvao, A. F. (2011). "Quantile regression for dynamic panel data with fixed effects." *Journal of Econometrics*, 164(1), 142-157.

Garcia-Diaz, A., & Riggins, M. (1984). "Serviceability and distress methodology for predicting pavement performance." *Transportation Research Record*, 997, 56-61.

Geraci, M., & Bottai, M. (2007). "Quantile regression for longitudinal data using the asymmetric Laplace distribution." *Biostatistics*, 8(1), 140-154.

Golabi, K., Kulkarni, R. B., & Way, G. B. (1982). "A statewide pavement management system." *Interfaces*, 12(6), 5-21.

Gramling, W.L., (1994). Synthesis of Highway Practice 203: Current Practices in Determining Pavement Condition. Transportation Research Board, National Research Council. Washington, D.C.

Gujarati, D. N. (2009). *Basic econometrics*. Tata McGraw-Hill Education.

Haas, R., & Transportation Association of Canada. (1997). *Pavement design and management guide*. Ottawa: Transportation Association of Canada.

Hao, L., & Naiman, D. Q. (2007). *Quantile regression* (No. 149). Sage.

Haque, A. U., Nehrir, M. H., & Mandal, P. (2014). "A hybrid intelligent model for deterministic and quantile regression approach for probabilistic wind power forecasting." *IEEE Transactions on Power Systems*, 29(4), 1663-1672.

HASSAN, K. E., Nicholls, J. C., Harding, H. M., & Nunn, M. E. (2008). *Durability of continuously reinforced concrete surfaced with asphalt*.

Heckman, J. J. (1981a). "Statistical models for discrete panel data." Structural analysis of discrete data with econometric applications, C. F. Manski and D. McFadden, eds., MIT Press, Cambridge, Mass., 114-178.

Highway Preservation Systems, Ltd. (2001). "Pavement condition evaluation manual." <<http://www.camineros.com/docs/cam014.pdf>> (Nov. 4, 2014).

Highway Research Board. (1962). "The AASHO road test." *Special Rep. No. 61A-E*, National Academy of Science, National Research Council, Washington, D.C.

- Hong, T., Wilson, J., & Xie, J. (2014). "Long term probabilistic load forecasting and normalization with hourly information." *IEEE Transactions on Smart Grid*, 5(1), 456-462.
- Hudson, W.R., Hass, R., and Pedigo, R.D. (1979). "Pavement Management System Development." National Cooperative Highway Research Program Report 215.
- Jiang, Y., Saito, M., & Sinha, K. C. (1988). *Bridge performance prediction model using the Markov chain* (No. 1180).
- Koenker, R., & Machado, J. A. (1999). "Goodness of fit and related inference processes for quantile regression." *Journal of the american statistical association*, 94(448), 1296-1310.
- Lashlee, J. T., Chaney, D., and Bearden, K. (2004). "Pavement management system report." Public Works Department Planning and Design Division.
- Li, N., Xie, W. C., & 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., Cheetham, A., Zaghoul, S., Helali, K., & Bekheet, W. (2006). "Enhancement of Arizona pavement management system for construction and maintenance activities." *Transportation Research Record: Journal of the Transportation Research Board*, (1974), 26-36.
- Li, Z. (2005). "A probabilistic and adaptive approach to modeling performance of pavement infrastructure."
- Lytton, R. L. (1987, November). "Concepts of pavement performance prediction and modeling." In *Proc., 2nd North American Conference on Managing Pavements* (Vol. 2).
- Machado, J. A., & Mata, J. (2005). "Counterfactual decomposition of changes in wage distributions using quantile regression." *Journal of applied Econometrics*, 20(4), 445-465.
- Madanat, S. M., Karlaftis, M. G., & McCarthy, P. S. (1997). "Probabilistic infrastructure deterioration models with panel data." *Journal of infrastructure systems*, 3(1), 4-9.
- Madanat, S., & Ibrahim, W. H. W. (1995). "Poisson regression models of infrastructure transition probabilities." *Journal of Transportation Engineering*, 121(3), 267-272.
- Madanat, S., Mishalani, R., & Ibrahim, W. H. W. (1995). "Estimation of infrastructure transition probabilities from condition rating data." *Journal of infrastructure systems*, 1(2), 120-125.

McClain, C., & Rex, M. (2001). "The relationship between dissolved oxygen concentration and maximum size in deep-sea turrid gastropods: an application of quantile regression." *Marine Biology*, 139(4), 681-685.

Merrill, D., Van Dommelen, A., & Gaspar, L. (2006). "A review of practical experience throughout Europe on deterioration in fully-flexible and semi-rigid long-life pavements." *International Journal of Pavement Engineering*, 7(2), 101-109.

Miller, J. S., & Bellinger, W. Y. (2003). *Distress identification windshield for the long-term pavement performance program* (No. FHWA-RD-03-031).

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

Olidis, C., & Hein, D. (2004). "Guide for the Mechanistic-Empirical Design of New and Rehabilitated Pavement Structures Materials Characterization Is your Agency Ready?" In *annual conference of the Transportation Association of Canada*.

Pavement management system. <<http://mail.tku.edu.tw/yinghaur/lee/pms/nhi-pms/NHI-PMS98.pdf>> (Jun. 24, 2016).

Powell, D. (2010). "Unconditional quantile regression for panel data with exogenous or endogenous regressors." *Available at SSRN 1722954*.

Prozzi, J. A., & Madanat, S. M. (2004). "Development of pavement performance models by combining experimental and field data." *Journal of Infrastructure Systems*, 10(1), 9-22.

Rao, S., Darter, M., Tompkins, D., Vancura, M., Khazanovich, L., Signore, J., Coleri, E., Wu, R.Z., and Harvey, J. (2013). "Composite Pavement Systems: Volume 1: HMA/PCC Composite Pavements." SHRP 2 report, S2-R21-RR-2. Washington, D.C.: Transportation Research Board, 2013.

Rao, S., Darter, M., Tompkins, D., Vancura, M., Khazanovich, L., Signore, J., & Vandenbossche, J. (2012). *Composite Pavement Systems Volume 1: HMA/PCC Pavements* (No. SHRP 2 Renewal Project R21).

Roberts, F. L., Kandhal, P.S., Lee, D. Y., and Kennedy, T. W. (1991). "Hot mix asphalt materials, mixture design, and construction." NAPA Education Foundation, Lanham, Md.

SAS Institute. <[http://www.sas.com/en\\_us/home.html](http://www.sas.com/en_us/home.html)> (Dec. 2, 2016).



Smith, R. E., Palmeri, R. P., Darter, M. I., & Lytton, R. L. (1984). *Pavement Overlay Design Procedures and Assumptions, Vol. 3: Guide for Designing Overlay*. FHWA-RD-85-008. Federal Highway Administration, Washington, DC.

Smith, R. J. (1963). "Definition of composite pavement structures." *Highway Research Record*, 37, 1-4.

Sun, L. (2001). "Developing spectrum-based models for international roughness index and present serviceability index." *Journal of transportation engineering*, 127(6), 463-470.

Thoegersen, F., Busch, C., & Henrichsen, A. (2004). "Mechanistic design of semi-rigid pavements-An incremental approach". *Mechanistic Design of Semi-Rigid Pavements*, (138).

United States Department of Transportation (2013). "Table 1-6: estimated U.S. roadway lane-miles by functional system." [http://www.rita.dot.gov/bts/sites/rita.dot.gov/bts/files/publications/national\\_transportation\\_statistics/html/table\\_01\\_06.html](http://www.rita.dot.gov/bts/sites/rita.dot.gov/bts/files/publications/national_transportation_statistics/html/table_01_06.html) (Jun. 10, 2016).

Vitillo, N., "Pavement Management System Overview." <http://www.state.nj.us/transportation/eng/pavement/pdf/PMSOverviews0709.pdf> (Jun. 23, 2016).

Von Quintus, H. L., Finn, F. N., Hudson, W. R., & Roberts, F. L. (1979). *Flexible and composite structures for premium pavements: Volume 1 and 2*. FHWA-RD-81-154 FHWA-RD-81-155. Federal Highway Administration, Washington, DC.

Wang, K. C., Hou, Z., & Williams, S. (2010). "Precision test of cracking surveys with the automated distress analyzer." *Journal of Transportation Engineering*, 137(8), 571-579.

Wang, K. C., Zaniewski, J., & Way, G. (1994). "Probabilistic behavior of pavements." *Journal of transportation Engineering*, 120(3), 358-375.

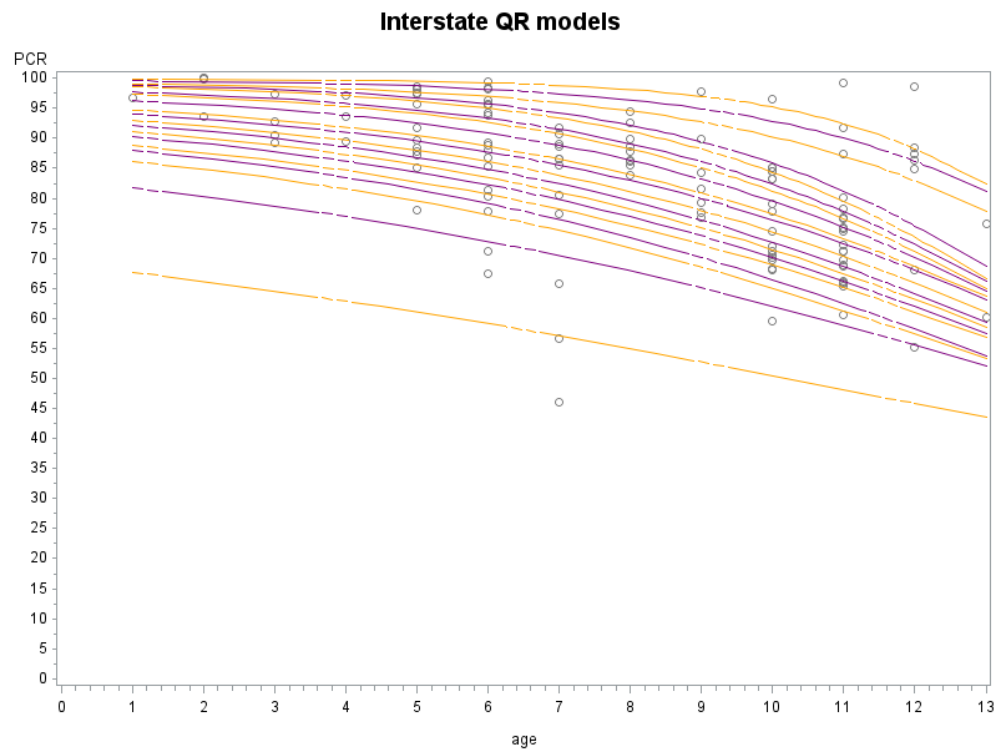
Wolters, A. S., & Zimmerman, K. A. (2010). *Research of Current Practices in Pavement Performance Modeling* (No. FHWA-PA-2010-007-080307).

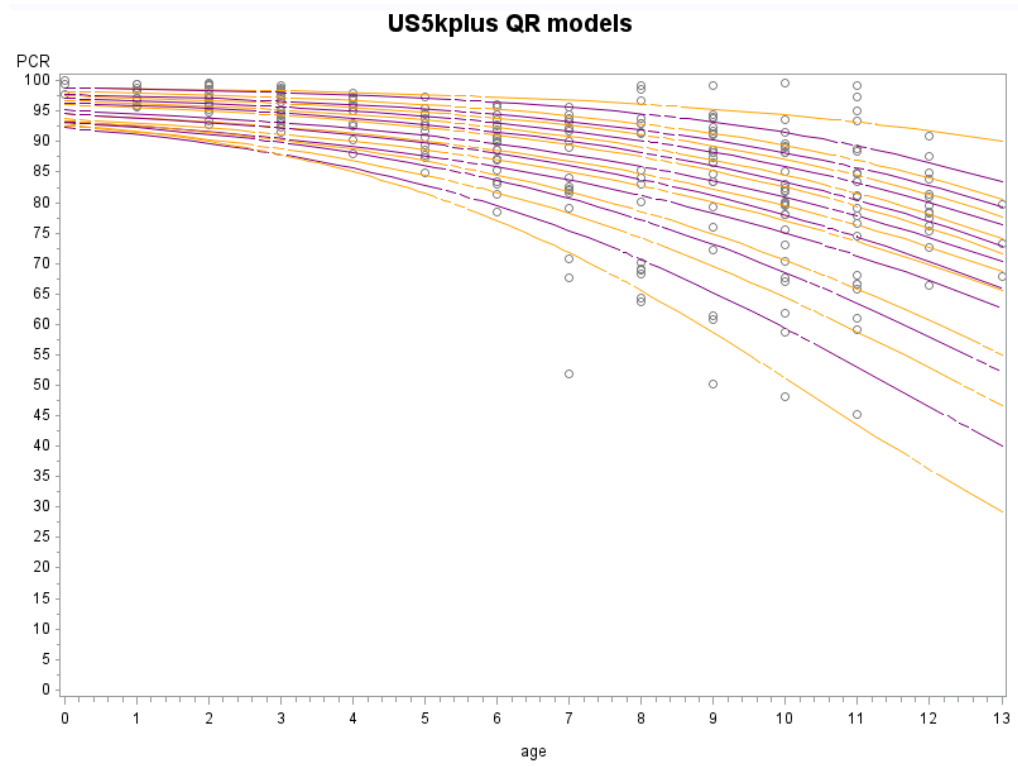
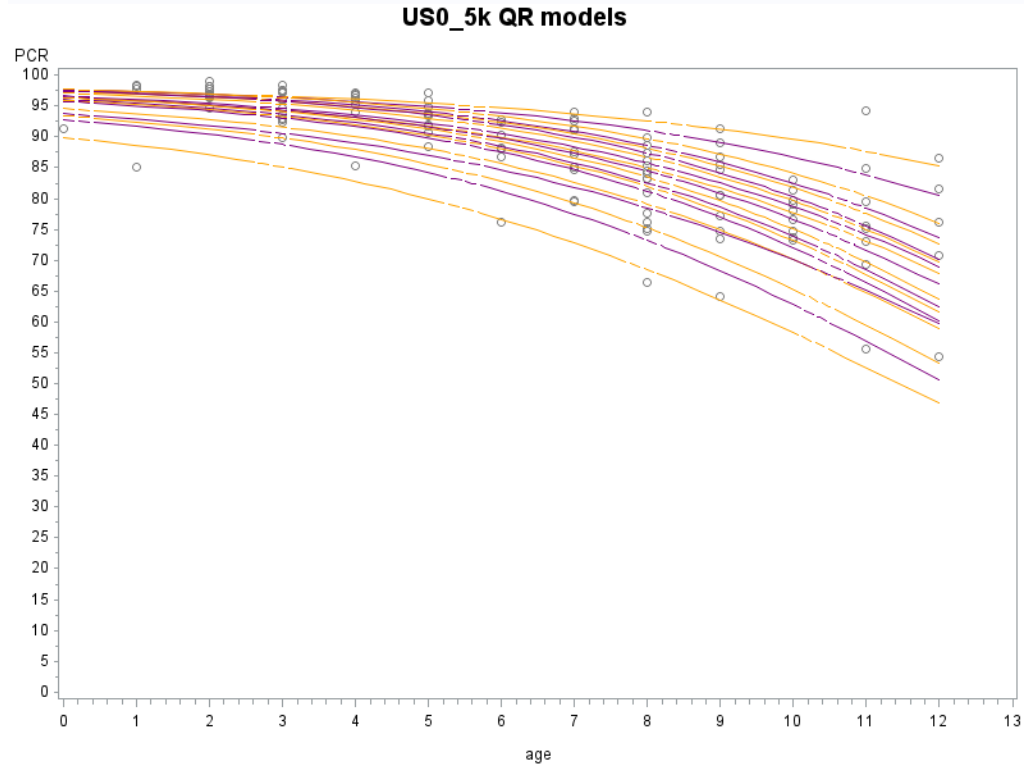
Yoder, E. J., & Witczak, M. W. (1975). *Principles of pavement design*. John Wiley & Sons. North Carolina Department of Transportation (2011). "NCDOT digital imagery distress evaluation handbook."

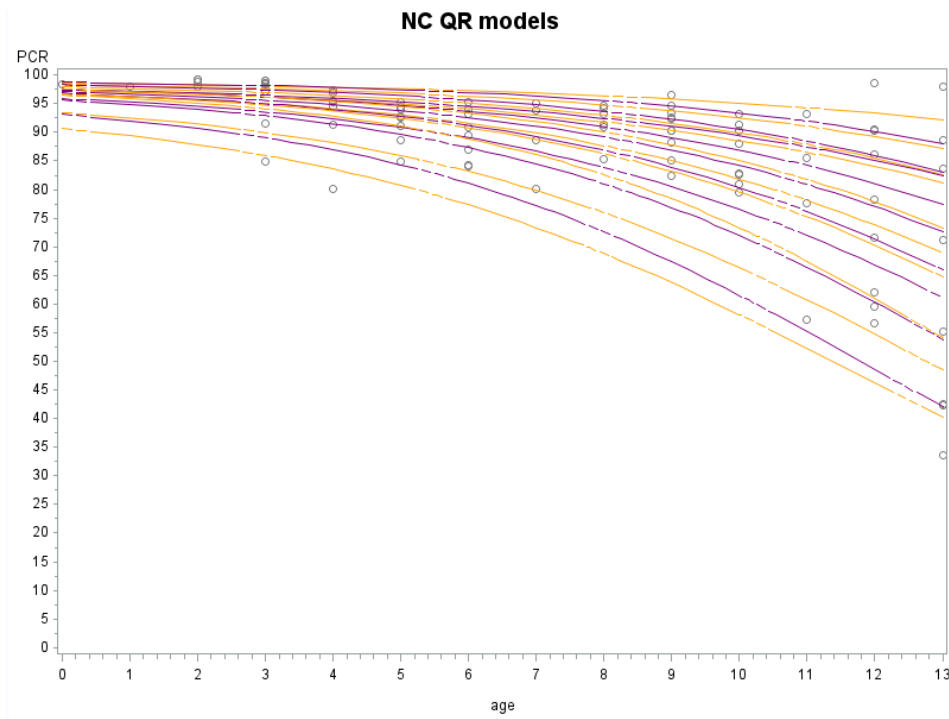
Yoder, E. J., and Witczak, M. W. (1975). *Principles of Pavement Design*, 2nd Ed. Wiley, New York

Zhou, K. Q., & Portnoy, S. L. (1996). "Direct use of regression quantiles to construct confidence sets in linear models." *The Annals of Statistics*, 24(1), 287-306.

## APPENDIX A QUANTILE REGRESSION MODELS FOR PAVEMENT FAMILIES AT THE PROJECT LEVEL







## APPENDIX B QUANTILE REGRESSION MODEL FOR PAVEMENT CLASSIFICATIONS AT THE NETWORK LEVEL

