Update and Documentation of MDOT Warranty Process and Distress Thresholds

FHWA Technical Report Documentation Page

1. Report No. FHWA/MDOT-RD-18-273	2. **Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle Update and Documentation of MDOT Warranty Process and Distress Thresholds		**5. Report Date** September 20, 2018 **6. Performing Organization Code**
7. Author(s) Feng Wang, Farshad Amini, Xiaohua Luo, and Jueqiang Tao		**8. Performing Organization Report No.**
9. Performing Organization Name and Address Jackson State University PO Box 18839 1400 J.R. Lynch Street, Jackson, MS 39217		**10. Work Unit No. (TRAIS)** **11. Contract or Grant No.** 107453/101000
12. Sponsoring Agency Name and Address Mississippi Department of Transportation PO Box 1850 Jackson, MS 39215-1850		**13. Type Report and Period Covered** 08/26/2016 to 12/31/2018 **14. Sponsoring Agency Code**
15. Supplementary Notes		

16. Abstract
Pavement warranty is an innovative contracting procedure, which has been adopted in Mississippi to enhance pavement performance and protect the investment in pavement construction since 2000. MDOT uses converted deduct points to monitor/evaluate the distresses and the distress thresholds are in the form of deduct points, which are not as straightforward as use of direct quantities. Moreover, the validity and applicability of continual use of these conversion equations have become problematic with time, because the equations were empirically developed in 1990s reflecting the data, experience, and technologies at that time. Currently, 100% collection is now possible and can be used in lieu of the sampling scheme in MDOT. Therefore, the objectives of this study are to research the performance of the current pavement warranty program, and the possible ways of updating new thresholds for the warranty program. The survival analyses were conducted to compare the pavement performance for warranty contracting versus the general non-warranty contracting in MS. The bootstrapping method was utilized to process the skewed historical distress measurement data into normal distributions. The upper bound of confidence interval or standard deviations of the normalized distress measurement data was determined as an alternative method to rebuild the threshold for each distress type. The confusion matrix was administrated to consider a smooth transition of updating the thresholds from the current deduct point based system to a new measurement based system. A Visual Basic based macro program was developed in MS Excel to implement the new thresholds. Eventually, the classification accuracy index, the one sample t-test, and the Levene's test were employed to validate the automated data collection technology currently used for the non-warranty pavements maintained by the state. The results have shown that the average performance of warranty pavements is superior to that of the non-warranty pavements in Mississippi. The thresholds based on the 1.5 σ method with the normalized non-warranty distress data could be an alternative to rebuilding the distress thresholds. The current automated data collection technology is not accurate enough for project-level pavement distress surveys required by the warranty pavements; therefore, project level analysis will continue to be done by staff augmentation and QC/QA of the automated method.

17. Key Words Pavement warranty, pavement distress, deduct point, threshold value, pavement performance, Pavement Management System (PMS), statistic method		**18. Distribution Statement** Unclassified	
19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 138	**22. Price**

Disclaimer

Jackson State University and the Mississippi Department of Transportation do not endorse service providers, products, or manufacturers. Trade names or manufacturers' names appear herein solely because they are considered essential to the purpose of this report.

The contents of this report do not necessarily reflect the views and policies of the sponsor agency.

MDOT Statement of Nondiscrimination

The Mississippi Department of Transportation (MDOT) operates its programs and services without regard to race, color, national origin, sex, age, or disability in accordance with Title VI of the Civil Rights Act of 1964, as amended and related statutes and implementing authorities.

Author Acknowledgments

The authors would like to thank the Mississippi Department of Transportation (MDOT) for providing research funding to this study. The authors are grateful to the support and assistance from MDOT engineers Ms. Cynthia Smith, Ms. Marta Charria, Mr. James Watkins (retired), Mr. Rhea Vincent, Mr. Alex Middleton, Mr. Alan Hatch, and Mr. Randy Battey (retired).

The project was a joint effort that pools expertise from the Department of Civil and Environmental Engineering, and the Institute for Multimodal Transportation (IMTrans) at Jackson State University (JSU), Jackson, Mississippi. JSU PhD student Ms. Xiaohua Luo contributed in conducting literature reviews, selecting statistical method, analyzing pavement distress data, evaluating performance of warranty and non-warranty pavements, developing measurement based warranty thresholds, and drafting the research report. Jueqiang Tao, visiting graduate student from the Zhejiang Normal University, Zhejiang, China, helped conduct literature reviews and develop new measurement based warranty thresholds, developed the macro program, and drafted the validation of the automated pavement cracking measurement part of the research report. Dr. Farshad Amini, Professor and Chair of the Department of Civil and Environmental Engineering, helped in developing the proposal, reviewing the final report, and budget management for the project. Dr. Feng Wang, Associate Professor of the Department of Civil and Environmental Engineering and Director of IMTrans, helped in drafting the proposal, fulfilling the research tasks, monitoring and reporting the progresses, and revising the final report for the project.

Table of Contents

List of Tables

List of Figures

CHAPTER 1. INTRODUCTION

1.1 Problem Statement

Pavement warranty is an innovative contracting procedure adopted by state transportation agencies. Many states view implementing warranties as a way to protect their investment in pavement construction. The major benefit of pavement warranty is enhanced pavement performance. However, establishing warranty criteria and associated distress thresholds to monitor the performance of warranty pavements (maintained projects) is an apparently challenging issue that must be carefully studied. Currently Mississippi Department of Transportation (MDOT) uses converted deduct points to monitor/evaluate the distresses and the distress thresholds are in the form of deduct points, while many other states use distress indicators and thresholds directly from measurements of pavement distresses or densities of distresses. Also, MDOT uses 500-ft samples, which are based on what the original deduct points method used. However, 100% collection is now possible and can be used in lieu of sampling. A recent survey study on pavement warranty programs in the US and Canada revealed that: (1) Although composite indexes are still used for condition evaluation of warranty pavements, the individual distress thresholds are the major criteria used by state DOTs to manage their warranty projects; and (2) Unlike the practice in Mississippi that uses deduct point based distress thresholds, all the other states that responded with existing warranty programs set their threshold limits based on the maximum allowed quantity of the distress measurement for the distress type.

This research study will use MDOT's pavement management system (PMS) data to statistically develop a new specification of distress thresholds based on direct measurements of pavement distresses or densities of pavement distresses for MDOT's pavement warranty program. Based on the investigation, the research will make suggestions for MDOT to adopt new features of the warranty specification on distress thresholds and on how MDOT's PMS database can be employed to monitor the performance of maintained projects in a more efficient manner. This research will analytically develop new distress thresholds based on direct measurements of pavement distresses or distress densities using the state of art and state of practice pavement evaluation technologies for MDOT's pavement warranty program. The reviewing of literature has provided solid understanding on the state of the practice methodologies of developing pavement distress assessment methods, pavement performance indicators and associated thresholds. The possibility of including ride quality (roughness) into the current MDOT's pavement warranty program and issuing an appropriate distress threshold will be investigated. The new performance thresholds will be determined using statistical models and data mining for appropriate ranges of acceptable pavement performance, with a special consideration for smooth transition from the old deduct point based system to the new system. The development of the new performance thresholds will be conducted by statistically analyzing the non-warranty pavement data saved in MDOT's PMS and especially the distress survey data collected annually for the warranty projects. Analytical capabilities such as statistical distributions and histograms of distress measurements, bootstrapping and data cleansing for correction of skewness and outliers, survival modeling and performance comparisons between warranty versus non-warranty pavements over service life, development of new distress thresholds based on confidence interval and cumulative distribution function methods, and correlation analyses of distress quantities measured using automated versus semi-manual methods, will be developed to establish the new warranty specification on distress thresholds. Confusion matrix for a smooth transition from the old system to the new system will be conducted to check the validity and applicability of the new specification of distress thresholds. In the confusion matrix analysis, the cases in which the old thresholds were exceeded will be revisited by analyzing the annual warranty survey reports rechecked with the newly developed distress thresholds, and necessary adjustments will be made to remain consistency. A macro program in Visual Basic and the tutorial narratives are developed to convey the analytical capabilities embedded in MS Excel, allowing the MDOT pavement managers and engineers to make needed adjustments in implementing new measurement based distress thresholds. As a result of the research study, this report will evaluate the performance of the existing pavement warranty program, introduce a specification of the updated warranty process and develop a training course module for implementing the new warranty program.

1.2 Background

Pavement warranty is regarded as one of the most prominent programs for protecting investment in pavement constructions and preventing early failures. It has been widely used in Europe to enhance the quality and to reduce the life-cycle costs of pavements. Potential benefits of implementing pavement warranty include decreased construction oversight, enhanced pavement quality, flexibility in pavement type and material selection, and the potential for the industry to use its knowledge more productively, primarily because of the shift from a passive to an active attitude toward quality (Qi et al. 2013). In 1999, a survey by the Ohio Department of Transportation indicated that only 12 states had introduced performance-based pavement warranty with a total of about 37 projects. However, in the following few years, the number, scope, and period of warranty projects have significantly increased. For example, Wisconsin considered 10 warranty projects in 2003; a recent project in New Mexico extended the warranty period to 20 years. Pavement warranty has been applied to different types of projects, such as new constructions, overlays on asphalt concrete or Portland cement concrete (PCC) pavements, and rehabilitations (Wang et al. 2005).

The most important technical challenge in implementing pavement warranty is to establish warranty specifications and to evaluate the performance of the pavements under warranty. The purpose of the warranty specification is to establish responsibilities, expectations, and consistency in the department's implementation and administration of warranty requirements in the maintained projects. Pavement warranty specifications usually include descriptions of the length of warranty period, the types and threshold levels of distresses or performance/distress indicators, the methods to measure the distresses or performance indicators, and the corresponding remedial actions. Proper selection of warranty items and establishment of their threshold levels constitute the most important component in the implementation of a pavement warranty. The most common distresses in warranty specifications are rutting and various types of cracking. Some states also include ride quality as a warranty item. Sometimes, friction, patching area, potholes, raveling, flushing, distortions, and disintegrated area can be in warranty specifications as well. It should be noted that each state developed its own warranty items on the basis of its local experience. For example, ride quality information is not included in some states' warranty items, whereas Indiana, one of the states that have ride information in warranty items, believes it is the best indicator to represent the performance of the pavement (El Gendy et al. 2012).

The threshold values vary state by state because in many cases, distress thresholds are determined by local experience with pavement performance in every individual state. It has been common practice for each state to review and statistically analyze the performance data stored in its PMS database to establish the threshold levels that are acceptable by the local contractors. A PMS database includes a tremendous amount of information about both warranty and non-warranty pavements through periodic data collection and analysis, which should be utilized directly in the development of an appropriate range of acceptable pavement performance in terms of distress thresholds. Currently MDOT uses deduct points converted from pavement distress measurements and the distress thresholds are each maximum deduct point allowable for a distress type, while many other states use distress indicators and thresholds directly from measurements of pavement distresses or densities of distresses (Qi et al. 2015). A recent survey study conducted by the authors on pavement warranty programs in the US and Canada has revealed that: (1) Although composite indexes are still used for condition evaluation of network of pavements, the individual distress thresholds are the major criteria used by state DOTs to manage their warranty projects; and (2) Unlike the practice in Mississippi that uses deduct point based threshold levels, all the other states that responded with existing warranty programs set their threshold limits based on the maximum allowed quantity of the distress measurement for the distress type (Qi et al. 2012). A quick retrospect of the history of the pavement evaluation methods used in Mississippi would help in better understanding the problem.

MDOT pavement distress assessment and evaluation methods have evolved over the last decades. Beginning in the early 1990s, pavement distresses were measured manually from video images and subsequently converted to deduct points using the deduct point curve equations which were empirically finalized in 1995 for MDOT's pavement management system (George 1995). In this method, two 500-ft samples per mile were human/visual-rated. The conditions of the pavements were evaluated by using deduct points converted from measured pavement distresses for individual distress types and an empirically developed composite pavement condition rating (PCR) that combines multiple distress types. In recent years, new technology has emerged in the data collection arena, and MDOT has changed from the human-rated sampling method to 100% automated distress survey at the network-level since 2010. The network-level pavement distress data are collected biannually for all state-maintained pavements by pavement data collection contractors. The concept of distress density (computed from distress measurement and affected area) has been employed with the automated data collection, and the PCR equations have been revised to accommodate the density concept. The conditions of pavements for the network-level surveys are evaluated by using deduct point equations for individual distress types and the revised distress density based PCR model (Smith 2009).

The MDOT pavement warranty program was started in the year 2000 and with the accumulation of about more than two decades, MDOT has awarded a total of 21 projects across the state to 9 different contractors to date, among which several are still within the warranty period. Of the total 21 projects, 18 are Hot Mix Asphalt (HMA) pavements with two having a 5-year warranty and sixteen having a 7-year warranty period, while 3 are Jointed Reinforced Concrete Pavement (JRCP) pavements with two in a 5-year warranty and one in a 10-year warranty (Qi et al. 2012). The distress assessment and evaluation methods specified for the warranty projects are the same as the methods that MDOT had used before 2010 (with manual data collection and deduct point equations), primarily because (1) the warranty projects were all contracted before 2010 and these contracts are still active; and (2) the manual distress survey method was regarded more appropriate for the project level (Qi et al. 2012, Battey and Watkins 2009).

A deduct point curve equation is actually a regression model that relates a pavement engineer's perception of loss of points in the rating of pavement performance to the measured distress of the pavement for the particular distress type. One of the major advantages of using the deduct points and the conversion curves is that the conversions can change every individual distress type from its original measurement unit and value range to the unified performance rating scale. After the conversions, different types of distresses become comparable on the same rating scale and could be added together to form the so-called composite PCR. However, the major disadvantages of using the deduct points in the pavement warranty program are: (1) The conversions of distress measurements into deduct points using the regressed deduct point equations can reduce the accuracy of the objective distress measurements by introducing subjective and random errors. (2) Since the deduct point equations were empirically developed in 1990s reflecting the data, experience, and technologies at that time, therefore, the thresholds need to be revisited to reflect new technologies and methods. (3) The composite PCR model that combines multiple pavement distresses along with the deduct point conversions on which the PCR model is based would be more needed for the reporting of pavement condition at the network level, than for the management of warranty projects which is indeed in the project level. The above disadvantages may explain the extreme unpopularity of the deduct point usage in pavement warranty programs in the US and Canada (Qi et al. 2015).

MDOT currently collects and rates the warranty pavement conditions in yearly surveys using in-house profilers, video logging, and staff time. The surveys still use two human/visual-rated 500-ft samples per mile and the old deduct point curves. Table 1 and Table 2 list the current distress threshold in deduct points for each of the 20 warrantied distress types for asphalt and concrete pavements (Qi et al. 2012, Battey and Watkins 2009). Apparently ride quality or roughness is not included in the current MDOT pavement warranty program as a distress indicator, although it has been widely used in other states. Furthermore, there exist the needs for MDOT not only to transition the project-level data collection for warranty projects to 100% automated distress surveys, from 500-ft sampling to 100% mileage surveying, but also to update from the deduct point based old warranty specification of distress thresholds to a more straightforward method of using direct distress measurements based or distress densities based threshold specification.

Table 1 Warranty Thresholds and Remedial Actions for Asphalt Pavements Used by Mississippi DOT

Distress Type	Threshold Level (deduct points)	Remedial Action
Alligator Cracking	10.0	Remove and replace distressed layers, the area to be equal to 150% of the distressed area to a depth not to exceed the warranty pavement
Block Cracking	3.0	Remove and replace distressed layers, the area to be equal to 110% of the distressed area to a depth not to exceed the warranty pavement
Reflection Cracking	9.0	Seal cracks according to the current Department SOP.
Edge Cracking	3.0	Remove and replace the distressed layers, the area to be equal to 110% of the distressed area
Longitudinal Cracking	4.0	Remove and replace distressed layers to a depth not to exceed the warranty pavement
Transverse Cracking	3.0	Remove and replace distressed layers to a depth not to exceed the warranty pavement
Potholes	5.0	Remove and replace distressed layers, the area to be equal to 150% of the distressed area to a depth not to exceed the warranty pavement
Rutting	5.0	Remove and replace the surface layer
Raveling/Segregation	0.2	Apply a chip seal or a partial depth repair
Surface Bleeding	0.4	Remove and replace surface distressed area mixture full depth
Friction	35	Milling, surface treatment, or overlay to correct inadequacy

Table 2 Warranty Thresholds, and Remedial Actions for Concrete Pavements Used by Mississippi DOT

Distress Type	Threshold (deduct points)	Remedial Action
Corner Breaks	4.3	Saw and square affected area; place dowels on transverse joints
Faulting of Transverse Joints	2.7	Diamond Grind - ensure positive drainage
Joint Seal Damage	1.66	Seal according to current MDOT policy
Longitudinal Cracking single crack	1.4	Stitch and Seal according to current MDOT policy
Transverse Cracking, single crack	1.97	Retrofit 3 dowels per wheel path; seal entire crack
Multiple cracks involved	3.5	According to current MDOT policy
Spalling of Longitudinal Joints	1.15	Clean (hydro-blast, sandblast or other) and fill
Spalling of Transverse Joints	4.4	Clean (hydro-blast, sandblast or other) and fill
Map Cracking & Scaling	1.77	Thin overlay with material that has good adhesion to concrete

1.3 Study Objectives

This research initiative will be identified as the Update and Documentation of MDOT Warranty Process and Distress Thresholds. The major objectives of this research effort are as follows:

1. Literature review about analyzing and establishing pavement performance of warranty projects in Mississippi and distress thresholds for MDOT's pavement warranty specification;

2. Development of statistical models for performance evaluation of warranty and non-warranty pavements in Mississippi;

3. Development of direct distress measurements based or distress densities based MDOT specification on distress thresholds;

4. Investigation of the viability of including ride quality or roughness (rutting and IRI) into MDOT's pavement warranty performance indicators;

5. Study of the validity, applicability, and consistency of applying the new specification of distress thresholds;

6. Documentation of the new warranty program and development of a training course module for MDOT.

The purpose of Objective 1 is to conduct a literature review to search for state of the art/practice methodologies of developing distress thresholds. Federal guidelines and experiences from pioneering states in pavement warranty programs will be understood. Statistical models will be established for the development of new distress thresholds.

The purpose of Objective 2 is to develop statistical models to evaluate the performance of warranty pavements, as well as that of the non-warranty pavements based on different characteristics of the two pavement programs in Mississippi. For the warranty program the failure criteria are the maximum allowed deduct point based thresholds, while for the non-warranty program, there is not an existing definition for the terminal pavement condition or no clearly defined (targeted) failure condition has existed for the pavements without a warranty.

The purpose of Objective 3 is to develop pavement distress threshold quantities and severities based on direct measurements of pavement distresses or densities of distresses, and ultimately to establish MDOT's new specification on pavement distress thresholds using pavement survey data in MDOT's PMS database.

The purpose of Objective 4 is to evaluate the possibility of adding ride quality or roughness (rutting and IRI) as a potential warranty performance indicator to the current pavement warranty program, and to develop an appropriate distress threshold for the newly added performance indicator.

The purpose of Objective 5 is to develop analytical capabilities of comparing the current deduct point based thresholds with direct distress measurements based or distress densities based thresholds to identify and understand possible differences between the two systems of distress thresholds. The objective is also to apply the newly developed distress thresholds to historical cases in which old thresholds were exceeded, to ensure validity, applicability, and consistency of applying the new specification of distress thresholds.

The purpose of Objective 6 is to develop training materials or modules about the newly developed pavement warranty process which will be used for education and training of MDOT engineers in implementation stage to be familiar with the new warranty specifications.

The proposed research entails joint efforts that pool expertise from MDOT's research, maintenance, construction and other related Divisions, and the Department of Civil and Environmental Engineering at Jackson State University. The successful completion of the proposed study will enhance MDOT’s pavement management functionalities by updating the current pavement warranty process with more objective and reasonable distress thresholds for the warranted distress items of the pavement warranty program in Mississippi.

1.4 Organization

The remainder of the report is organized as follows: Chapter 2 presents a summary of the literature reviews in the areas of pavement warranty implementation experiences, warranty threshold setting guidelines and methods, and statistical methods for data analysis and processing. Chapter 3 compares the performance of the warranty versus non-warranty pavements over service time in Mississippi. Chapter 4 describes the data used in the study. Chapter 5 is focused on the methodology employed in the research. Chapter 6 develops and identifies the measurement based new distress thresholds. Chapter 7 of introduces the implementation and adjustment of the new thresholds. Chapter 8 aims to validate the applicability of using automated measurements in place of the current semi-manual rating measurements in collection of pavement distress data for warranty projects. The conclusions and future research recommendations are summarized in Chapter 9.

CHAPTER 2. LITERATURE REVIEW

Pavement warranty is an innovative contracting procedure adopted by state transportation agencies to protect investment in pavement construction and maintenance. There are mainly three components in the implementation of warranty projects: 1) pavement distress indicators; 2) pavement distress thresholds; and 3) pavement distress remedial action. It is recommended that the pavement condition indicators for warranty projects are friction, IRI, rutting depth, and cracks. Determining the threshold criteria for the pavement distress is dependent on the availability of the historical distress data and especially based on realistic and attainable PMS data. The remedial action for addressing a pavement distress is determined by the agency to guarantee the continuing performance condition of warranty pavements (FHWA 2017). The pavement distress indicators are the main factors to evaluate the pavement performance, and the pavement distress thresholds are the key criteria to make the maintenance decisions of warranty pavements. The warranty thresholds could simply be pavement distress measurements or pavement condition index (PCI) in some states, while in Mississippi they are pavement deteriorations in deduct points, which are converted using empirical equations from the pavement distress measurements (George 1995), and the distress thresholds are each maximum deduct point allowable for a distress type (Qi et al. 2015).

The following sections summarize the pavement warranty programs of relevant states reported in the literature, as well as previous studies on pavement warranty conducted by researchers, and the documents of the current pavement warranty practice and related studies in Mississippi. There is also explanation on the methods applied in the development of distress thresholds in warranty programs. Eventually, there are sections that illustrate the data analysis and processing methods considered in this study to set up and evaluate the distress thresholds.

2.1 Pavement Warranties in the United States

Warranty contracts, where the contractor guarantees the product and assumes responsibility for repair and maintenance for a designated period, are either being used or being critically examined for use by many states. Recently, with the publication of Federal Highway Administration's (FHWA) final rule on warranty clauses, the warranty contracts are used in areas such as bridge painting, pavement making, and freeway management among states in the US. Later on, the FHWA is interested in advancing short and long-term performance warranties for Hot Mix Asphalt (HMA) and Portland Cement Concrete (PCC) pavements on pavement preservation treatments (Hughes 1996).

The use of warranties on roadway construction projects has seen gradual popularity since the FHWA developed warranty guidance documents in the 1990s in the US (FHWA 2017). A warranty process model was developed and refined based on data collected from state highway agencies, as well as from studying individual specifications, programs, and projects (Anderson and Russell 2001). A pavement performance warranty program is administrated to specify the minimum performance conditions of a particular pavement over a specified period of time and to define the contractor’s remedial responsibility in case of premature failures (Qi et al. 2013). It is an innovative contracting procedure widely used in Europe and adopted by state transportation agencies in the US to enhance the quality and to reduce the life-cycle costs of pavements (Anderson and Russell 2001).

Table 3 Pavement Warranty Projects by States (by 2012)

Agency	Pavement Type	No. of Projects	Warranty Period (years)
Indiana	HMA	4	5
	Pavement preservation (micro surfacing)	30	3
	JPCP	3	5
Illinois	HMA overlay	3	5
	Full Depth HMA (20- yr. design)	3	5
	Full Depth HMA (30- yr. design)	7	5
	JPCP (20- yr. design)	3	5
	JPCP (30- yr. design)	4	5
	CRCP (30- yr. design)	8	5
	CRCP (40- yr. design)	1	5
Louisiana	HMA	2	3
	JPCP	1	3
Mississippi	HMA	2	5
	HMA	16	7
	JRCP	2	5
	JRCP	1	10
Pennsylvania	HMA	8	5
Wisconsin	HMA	199	5
	HMA	8	3
	HMA	3	7
	Dowel Bar Retrofit	10	3
	JPCP	16	5

To identify and understand the pavement warranty practices and specifications in the US, a comprehensive survey study on pavement warranty was conducted by the research team in the previous project “Evaluation of MDOT’s Distress Thresholds for Maintained Pavement Projects.” The survey questionnaire was developed in February 2011 using the contact list of the Research Advisory Committee (RAC) of the American Association of State Highway and Transportation Officials (AASHTO). There are five respondents considered relevant to Mississippi with respect to pavement type and warranty period among all the participated state DOTs in the survey. The result listed in Table 3 reported that they had warranty projects on

HMA pavement, and the average warranty period required was 5 years for HMA pavements. Four out of six state DOTs stated that they applied warranties on JPCP (Jointed Plain Concrete Pavement) pavement projects, and the average warranty period required was 5 years. In addition, Indiana DOT has applied warranties to 30 micro surfacing projects; MDOT has used warranty contracting on 18 HMA and 3 JRCP projects; In Wisconsin, warranties were applied to 10 Dowel Bar Retrofit projects.

The survey questionnaire results show that pavement warranty has been applied to different types of HMA and PCC pavements. In addition to the often applications in pavement preservation or rehabilitation projects, it was reported in the survey that warranties were also implemented for the full depth in HMA and PCC pavement construction projects. The warranty period was usually between 3 to 7 years, and the most commonly used warranty term was 5 years, while no warranty period of more than 10 years was reported in the survey.

Table 4 Pavement Warranty Items Specified by State DOTs

Agency	Ride Quality/roughness	Physical Distresses	Structural Capacity	Safety
British Columbia	√	√	√	
Florida	√	√		
Illinois	√	√		
Indiana	√	√		√
Louisiana	√	√		
Mississippi		√		√
Nova Scotia		√		
Pennsylvania	√	√		√
Wisconsin	√	√		

In projects under warranty program, the contractor takes a specific set of duties and associated risks, and will take the resulting cost in case of a premature problem during the warranty period. The risk is transferred from highway agencies to the contractors to various degrees. The most significant technical challenge in implementing a pavement warranty is to establish warranty specifications and to evaluate the performance of the pavements under warranty (Wang et al. 2005). The purpose of the warranty specification is to establish responsibilities, expectations, and consistency in the department's implementation and administration of warranty requirements for the warranted projects. Pavement warranty specifications usually include descriptions of the length of warranty period, the types and threshold levels of distresses or performance/distress indicators, the methods to measure the distresses, and the corresponding remedial actions (FHWA 2017). Proper selection of warranty items and establishment of their threshold levels constitute the most important components in the implementation of a pavement warranty. It should be noted that each state develops its own warranty items on the basis of its local experience, as presented in the survey questionnaire results in Table 4. For example, ride quality information is not included in some states' warranty items, whereas Indiana, one of the states that have ride information in warranty items, believes it is the best indicator to represent the performance of the pavement (Qi et al. 2012). A comprehensive literature review shows that although warranty items vary from state to state, most of the reviewed states include rutting, roughness, and various types of cracking as pavement condition indicators under warranty for both asphalt and concrete warranted pavements, whereas raveling and bleeding are specifically for asphalt pavements, and faulting and spalling for concrete pavements (Qi et al. 2018).

2.2 Implementation of Warranty Program in Mississippi

MDOT implemented its pavement warranty program in 2000 with the purpose to foster innovation and improve pavement quality. So far, there have been 21 pavement performance warranty projects with warranty periods from 5 to 10 years, where the pavement distress is surveyed annually, and the performance of these pavements is monitored in the entire duration of the warranty period. The main type of pavement condition under warranty for pavement warranty projects in Mississippi is physical distress. In addition, ride quality data measured in International Roughness Index (IRI), in. /mi., is collected for construction acceptance of the pavement warranty projects in Mississippi.

The distresses are defined and measured according to the Long-Term Pavement Performance (LTPP) Distress Identification Manual (DIM) (Miller and Bellinger 2003). A deduct point value of the distress is then calculated by entering the appropriate empirical deduct point curve with the measured distress level and interpolating against the observed distress severity level. For each distress type, there is an associated threshold value. If the calculated deduct point exceeds the threshold level for the distress type, the contractor is required to perform the remedial action listed in Table 1 and Table 2.

The warranty pavement is surveyed annually, and a minimum of two 500 ft. sections are sampled per mile to collect pavement roughness and physical distress data. Roughness, rutting in asphalt pavement, and faulting in concrete pavements are collected on 100% of the sampled sections and stored electronically. Pavement distresses are evaluated by a video graphic technique. Video logging for distress identification is conducted in both directions regardless of whether the road is divided or undivided. For multilane divided roads, the distress data is usually collected in the outside lanes, but MDOT also reserves the right to collect data in all lanes if needed. In addition to the severity level, MDOT survey records distress extent information as well (MDOT 2015).

MDOT uses the High Speed Inertial Profiler to collect roughness data and transverse faulting on warranty pavements. Scanning laser rut measurement technology has been applied to rutting data collection since spring 2009 to replace the 3-point laser rut measurement. Surface distress data of the sampled sections are collected by a video graphic technique. A distress evaluation is then performed on the video images. Trained personnel in the office review the video pictures of the road surface and categorize and document the distress severity and extent. If a distress threshold is exceeded, the contractor will be notified of the results within 30 days of the annual survey and have 45 days to perform any required remedial action. If 30% or more of the project segments require a remedial action, then the entire project will receive that action. Should the contractor contest the results, the conflict resolution team would have to pass a judgment employing the simple majority rules. Apparently ride quality or roughness is not included in the current MDOT pavement warranty program as a distress indicator, although it has been widely used in other states.

Figure 1 Surface Conditions of Warranty vs. Non-warranty Sections of US Highway 49 in Simpson County.

A preliminary study was performed to compare the performance of the HMA overlay warranty project in US-49 in Simpson County and its control road segment in terms of PCR and distress index after 4 years of service. Figure 1 shows that over the four-year period, the pavement condition of the warranty project maintained a higher condition score with a slight trend of deterioration at the end of the period. On the other hand, the pavement condition of the non-warranty project deteriorated in an early stage at an accelerated rate.

To examine the amount of the increase in investment to achieve the warranty project objectives, the costs of two pairs of HMA warranty and non-warranty projects in Mississippi were analyzed and compared. It shows that compared to the non-warranty projects, the warranty projects exhibited 15.40% and 33.35% increases respectively in unit cost per mile of the freeway as well as 11.86% and 53.65% increases in unit bid cost per mile of HMA material, respectively (Qi et al. 2013).

2.3 Distress Thresholds in Warranty Programs

In developing a pavement warranty program, the agency needs to establish the pavement distress types to be included in the contract, the threshold values of the distresses, and the remedial actions to be taken to recover the pavement from the distresses if any of the thresholds is exceeded. The proper selection of warranty items and establishment of their threshold levels constitute the most important components in the implementation of a pavement warranty program. In the pavement warranty specification, the specified threshold value of a distress type is indicated as the tolerable distress level for the particular distress type as an indicator of the pavement performance in warranty. Warranty provisions may define zero-tolerance thresholds for specific distress types and severity levels, meaning that the existence of any sign of distress requires remedial action. They may also define maximum allowable tolerances for thresholds, which if exceeded trigger remedial actions. If the requiring-remedial program segments reach a specific percentage (e.g., 30% in Mississippi), the entire segment has to receive the remedial action. When the contractor contests the results, the conflict resolution team will have to pass a judgment by employing the simple majority rule (Qi et al. 2012). Recently, thresholds have been typically measured by visual inspection, laser profiling, or individual measurements (Scott III et al. 2011). Agencies establish warranty thresholds depending on the availability of the historical PMS data and statistical methods. Actually, the threshold values vary state by state because the thresholds are determined by local experience with pavement performance in every individual state. To date, various methods have been developed to define the warranty distress thresholds. For instance, some states, such as Florida, Illinois, and Indiana, directly determine the thresholds based on specific values of pavement distress measurements or distress densities, while Pennsylvania and Louisiana use varied percentages of segment for some distress types (Qi et al. 2015, Battaglia 2009). Table 5 to Table 8 list typical warranty thresholds used by selected experienced states implementing pavement warranty contracting.

Table 5 Distress Thresholds for Asphalt Pavements (Part 1)

Distress Type	Distress Thresholds by State			
	Louisiana	British Columbia	Indiana	Illinois
Alligator Cracking	N/A	N/A	N/A	50 ft^2 moderate, or any high severity
Block Cracking	N/A	N/A	N/A	100 ft^2 moderate, or any high severity
Reflection Cracking	N/A	N/A	N/A	N/A
Edge Cracking	50 linear ft. total length with crack width > 0.25 in; or > 100 linear ft. total length	N/A	N/A	10 ft. high severity
Longitudinal Cracking	50 linear ft. total length with crack width > 0.25 in; or > 200 linear ft. total length	N/A	0 ft, severity 2	10 ft. moderate, any high severity
Transverse Cracking	N/A	N/A	0 ft, severity 1	10 ft. moderate, any high severity
IRI	N/A	IRI > 2 m/km	90 in./mi.	110 in./mi.
Rutting	0.35 in. averaged in any 50 foot length in any wheel path	Visible rutting after 1 year	0.25 in.	0.30 in.
Potholes	Any occurrence	Any occurrence	6 in^2	Any occurrence
Surface Bleeding	10ft^2	Any occurrence	N/A	500 ft^2 moderate, any high severity
Friction	N/A	N/A	25	N/A
Raveling/ Segregation	10ft^2	Any occurrence	N/A	500 ft^2 moderate, any high severity
Others	Shoving: any occurrence Fatigue cracking:10 ft^2	N/A	N/A	N/A

Table 6 Distress Thresholds for Asphalt Pavements (Part 2)

Distress Type	Distress Thresholds by State		
	Florida	Pennsylvania	Wisconsin
Alligator Cracking	> 30 ft. length of cracking per 0.1 mi. segment for cracks >1/8 in.	Medium: average crack width > hairline and ≤ 0.25 in. High: average crack width > 0.25 in.	≥ 50 ft^2 in a segment of medium (M) or higher
Block Cracking	> 30 ft. length of cracking per 0.1 mi. segment for cracks >1/8 in.	Anything >0%	≥ 50 ft^2 in a segment of level low (L) or higher
Reflection Cracking	> 30 ft. length of cracking per 0.1 mi. segment for cracks >1/8 in.	N/A	N/A
Edge Cracking	> 30 ft. length of cracking per 0.1 mi. segment for cracks >1/8 in.	Anything >0%	≥ 50 linear ft. of the segment length
Longitudinal Cracking	> 30 ft. length of cracking per 0.1 mi. segment for cracks >1/8 in.	Anything >0%	≥ 50 linear ft. in a segment
Transverse Cracking	> 30 ft. length of cracking per 0.1 mi. segment for cracks >1/8 in.	Low: average crack width > hairline and ≤ 0.25 in. Medium: average crack width > 0.25 in. and ≤ 0.5 in. High: average crack width > 0.5 in.	≥ 100 linear ft. in a segment of level low (L) or higher
IRI	Ride Number, any 0.1mi. segment with RN <3.5	N/A	N/A
Rutting	0.25 in.	> 9.5 mm (3/8 in.)	≥ 0.375 in. in depth
Potholes	Any potholes	Anything > 0%	Any occurrence
Surface Bleeding	Width ≥ 1 ft. and ≥ 10 ft. long	N/A	≥ 50 ft^2 in a segment of level medium (M) or higher
Friction	N/A	N/A	N/A
Raveling/ Segregation	Raveling length≥ 10ft	Medium: surface is rough and pitted, may have loose particles. High: Surface is very rough and highly pitted.	≥ 50 ft^2 in a segment of level medium (M) or higher
Other distress	N/A	Miscellaneous cracking/ Edge deterioration/ Left edge joint: Low, Medium, or High	Longitudinal & transverse distortion (includes bumps & sags, corrugation, depression, swell and other distortions) ≥ 50 ft^2 in a segment of level medium (M) with ½ in. of vertical distortion or higher

Table 7 Distress Thresholds for Concrete Pavements (Part 1)

Distress Type	Distress Thresholds by State			
	Louisiana	Indiana	Illinois	Florida
Corner Breaks	Any occurrence	N/A	Any moderate or high	N/A
Faulting of Transverse Joints	0.25 in. maximum; 0.125 in. minimum	0.25 in.	N/A	N/A
Joint Seal Damage	Any occurrence	12 ft. cumulative total	N/A	N/A
Longitudinal Cracking	Any occurrence	0 ft., severity 2	10 ft. moderate, or any high	4 cracks in any lane mile > 1/8 in. or any crack > 3/8 in.
Multiple Cracks Involved	N/A	N/A	N/A	4 cracks in any lane mile > 1/8 in. or any crack > 3/8 in.
Transverse Cracking	Any occurrence	0 ft., severity 1	10 ft. moderate, or any high	4 cracks in any lane mile > 1/8 in. or any crack > 3/8 in.
Spalling of Longitudinal Joints	> 2 in. width	N/A	10 ft. moderate, or any high	N/A
Spalling of Transverse Joints	> 2 in. width	N/A	10 ft. moderate, or any high	N/A
Map Cracking & Scaling	N/A	N/A	50 ft^2	4 cracks in any lane mile > 1/8 in. or any crack > 3/8 in.
IRI	N/A	90 in./mi.	150 in/mi	Ride Number, any 0.1 mile segment with RN <3.5
Other Distresses	Tire texture (tire gauge): 0.125 in. mean texture depth; Macrotexture (sand patch): 20% maximum loss over warranty period; Lane-to-AC Shoulder Separation: Any occurrence; Popouts: Any occurrence; Spalled Areas: Areas > 25 in^2 and/or with depth > 1 in.	N/A	Punchouts in CRCP – any moderate or high severity	Spalling is divided between "in wheel path" and "outside wheel path". In wheel path, four areas in any lane mile exceeding 1 inch in width and 6 inches in length OR any single area exceeding 3 inches in width. For areas outside wheel path, four areas in any lane mile exceeding 1.5 inches in width and 12 inches in length OR any single area exceeding 3 inches in width. Shattered slabs –cracking patterns that divide the slab into three or more segments require full slab replacement.

Table 8 Distress Thresholds for Concrete Pavements (Part 2)

Distress Type	Distress Thresholds by State	
	Wisconsin	Pennsylvania
Corner Breaks	N/A	N/A
Faulting of Trans. Joints	≥ 3 faulted joints or cracks per station with faulting > ¼ in.	Medium: Absolute value of elevation difference is ≥ 0.25 in. and < 0.5 in. High: Absolute value of elevation difference is ≥ 0.5 in.
Joint Seal Damage	N/A	N/A
Long. Cracking	N/A	Medium: Average crack width ≤ 0.25 in. wide, spalling ≥ 2.0 in. wide for ≤ 50% length. High: Average crack width > 0.25 in. wide, spalling 2.0 in. wide for > 50% length.
Multiple Cracks Involved	Broken panels	N/A
Trans. Cracking	≥ 8 broken panels	Medium: Average crack width > hairline & ≤ 0.25 in. wide, Spalling ≥ 2.0 in. wide for ≤ 50% length, or faulting ≥ 0.25 in. and < 0.50 in. High: Average crack width > 0.25 in. wide, spalling ≥ 2.0 in. wide for > 50% length, or faulting ≥ 0.5 in.
Spalling of Long. Joints	Any distress > 2 in. in width or any faulting less than ½ in. at the longitudinal joint within a 0.1-mi. segment. Or, Faulted longitudinal joint (> ½ in.).	Medium: Average spalled width ≥ 3.0 in. and < 6.0 in. for an accumulated spalled length of at least 25 ft. High: Average spalled width ≥ 6.0 in. for an accumulated spalled length of at least 25 ft.
Spalling of Trans. Joints	Distress ≥ 2 in. in width in the wheel paths on 5 joints or cracks in any one 0.1 mile segment.	Medium: > 2.0 in wide for ≤ 50% of joint length. High: > 2.0 in wide for > 50% of joint length
Map Cracking & Scaling	10% of surface	N/A
IRI	N/A	N/A
Other Distresses	Slab Breakup: Cracks or slabs broken into 2 pieces. >4 cracked slabs per segment (0.1 mi.) at 3 years of age and >8 slabs per segment at 5 years of age. A slab is defined as a section of pavement bounded on the ends by joints and on the sides by a centerline joint and/or the edge of pavement; Or ≥ 1 slabs broken into ≥ 3 pieces; Or no distressed patches. Any patch present must be in good condition and performing satisfactorily. Surface Distress: Distress is present on > 0.5% and < 10% of the surface area on any one 0.1-mi. segment, if surface distress is < 1 inch in depth.	Broken slab: Low: At least 4 pieces in a 20-ft. length with average width ≤ hairline in the outside wheelpath, no faulting and IRI ≤ 100 in./mi.; Medium: At least 4 pieces in a 20-ft. length with an average crack width > hairline and ≤ 0.25 in., may have faulting > 0.25 in. and ≤ 0.50 in. or IRI > 100 and ≤ 200 in./mi.; High: At least 4 pieces in a 20-ft. length with an average crack width > 0.25 in., may have faulting > 0.5 in. or IRI > 200 in./mi.

Currently MDOT sets the threshold for each distress type as the maximum allowable deduct points converted from the corresponding pavement distress measurement. The distress measurements are conducted every year on warranted pavements by a team from the Research Division. The deduct point conversion curve equations were developed in 1995 for MDOT's PMS (George 1995), which are actually regression models that relate pavement engineers' perception of pavement performance to measured distresses for particular distress types. However, the regressed deduct point equations can be confusing to use for warranty projects. Further, the old deduct point curves and distress thresholds may not be suitable to evaluate the current warranty pavements due to the emergence of new pavement technology and new pavement materials (Qi et al. 2012). Therefore, the use of direct distress measurements based or distress densities based thresholds would be easier, more straightforward, and more appropriate for contractors and MDOT than the use of the current deduct points based thresholds.

Another potential issue associated with threshold-setting methods referred in National Cooperative Highway Research Program (NCHRP) report 699 is regarding whether to specify a distinct threshold or a graduated scale or performance curve during the warranty period. Montana DOT argued that the distinct minimum performance thresholds would be more appropriate than the graduated thresholds over the warranty period, because the performance curve might be modified once the remedial action is performed (Scott III et al. 2011). Therefore, the objective distress measurement based distinct thresholds would be the appropriate option for setting distress thresholds for warranty pavement performance in this study.

2.4 Statistical Methods for Warranty Programs

Because of the varying operating environments and the difference in specifications of pavement warranty programs among states, it is still difficult to arrive in the establishment of a generally accepted procedure or method to evaluate the effectiveness of pavement warranty programs. Therefore, one of the most important technical challenges in pavement warranty implementation is to find an appropriate evaluation method, which serves to examine the performance comparison between warranty and non-warranty pavements.

Survival analysis is a branch of statistics for assessing the expected duration time until the occurrence of a specific failure event. As a typical time-related model, the survival analysis method gives, based on the history of condition data, the probability of the survival event of a living or nonliving system up to a terminal point. It was first adopted in pavement performance modeling in the 1930s (Winfrey 1969). The Highway Design and Maintenance Standards Study (HDM) first used it for the World Bank to predict the initiation of fatigue cracking (Paterson and Chesher 1986). Afterward, the survival analysis was used to reanalyze the American Association of State Highway Officials (AASHO) road test data created in the 1960s, and the result showed that the survival model was more appealing than the original AASHO formulations (Prozzi and Madanat 2000). The Illinois Department of Transportation (IDOT) applied the survival method to analyze the failure probability of longevity highway pavements with the increasing of pavement age and cumulative equivalent single axle loads (ESALs) using the historical data recorded in the Illinois Pavement Feedback System (IPFS) database, moreover, the analysis was conducted on various bare pavement types, overlays in categories by thickness, and overlaid pavement types (Gharaibeh and Darter 2002, Gharaibeh and Darter 2003, Gharaibeh and Shirazi

2009). Survival analysis has also been adopted to reflect pavement transverse cracking initiation time with Weibull hazard functions (Wang et al. 2005) and to identify the factors affecting the cracking initiation of resurfaced asphalt pavements (Dong and Huang 2014). In this study, the survival analysis based evaluation method is utilized to compute the survival probabilities of warranty and non-warranty pavements during different lengths of service time to show the superior performance of the warranty pavements.

Due to the superior performance of warranty pavements and the disadvantages of the existing deduct point thresholds, it is necessary to develop new thresholds to update and improve the warranty program in Mississippi. In order to analytically develop new warranty thresholds based on direct measurements of pavement distresses or distress densities for MDOT's pavement warranty program, statistical analysis models and data mining methods were applied to determine the new distress thresholds for appropriate ranges of acceptable pavement performance. NCHRP Report 699 has proposed that standard deviation of the historical non-warranty performance data is an alternative way to establish acceptable pavement performance thresholds (Scott III et al. 2011). In order to better understand the administration of the standard deviation method in developing the pavement performance thresholds, the following example presented in Figure 2 provides an approach for establishing baseline IRI thresholds using PMS data in Indiana DOT.

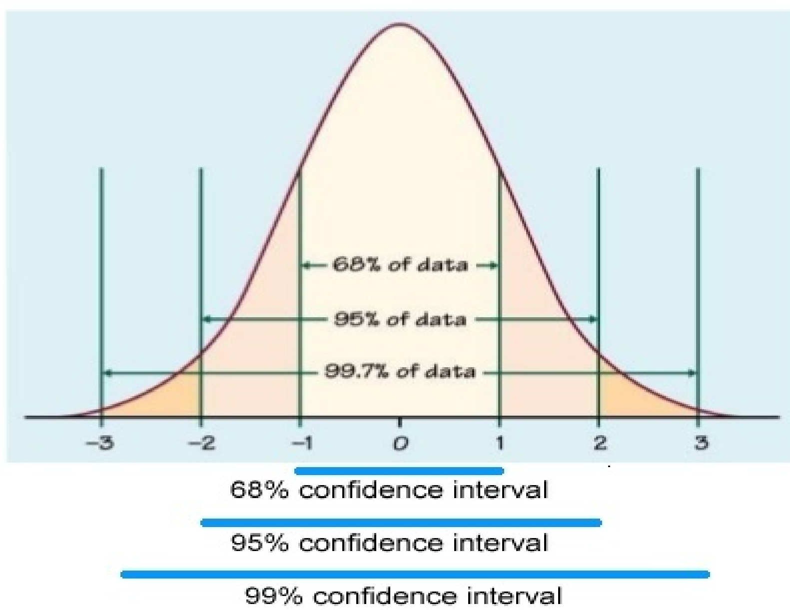


Figure 2 Confidence Interval of Standard Deviations in Standard Normal Distribution

In this example, the PMS data were taken from an Indiana DOT 10-year HMA pavements using high-speed data collection methods. Since the length of data collection segments of warranty pavements is shorter than that of the road sections collected in PMS, to develop the warranty threshold, the PMS data were reprocessed into 0.1 mile segments to compile the IRI dataset. Then, the reprocessed dataset was checked for normality to determine the statistical distribution and to calculate the standard deviation of the IRI dataset. In the standard deviation (denoted as symbol σ) method as shown in Figure 2, the data was assumed to be normally distributed. In a two-side test, approximately 68% of the data would fall within 1.0 σ, and 95% of the population would fall within 2.0 σ around the population mean. As a starting point, the DOT may set the the threshold at 2.0 σ (with 2.5% of measured sections would exceed the threshold) to reduce the failure risk to the agency for a 10-year warranted pavement. However, based on the practice experience or improved consistency, the DOT's found that setting the threshold at 2.0 σ would be too loose. Therefore, it was suggested that the threshold for warranty pavement should be tightened to between 1.0 σ and 2.0 σ (Scott III et al. 2011).

As presented in the example, the objective distress measurement thresholds could be determined based on the standard deviation of the historical non-warranty performance data. Hence, the accurate historical performance data are vital in the development of the distress measurement based thresholds. Once the network inventory and pavement condition data have been collected, a database can be established to store and use the information. Although a manual filing system may be possible for a small network, the efficiency and cost effectiveness of storing data on a computer makes an automated database the most practical alternative, especially with the size and complexity of a state pavement network. Therefore, the recent historical performance data are collected and stored using automatically or semi-automatically method among states.

Based on the available data collection and storage methods, there are many reasons that may contribute to inaccuracy in the results of surveyed data, such as outliers, system errors, data missing, and data deficiencies. Moreover, the experience of the authors shows that the historical distress data recorded in a PMS may be much skewed and do not follow a normal distribution (Luo et al. 2018). The bootstrapping method, which is usually used on skewed data with non-normal or unknown distributions (Archilla 2006, Li et al. 2010, Romanoschi and Metcalf 2000), is adopted to process the skewed historical distress measurement data into normal distributions. The nonparametric bootstrapping method was presented as an alternative approach to conduct the economic cost analysis (Desgagne et al. 1998). It has surged in popularity for analysis of pavement structure life to determine the confidence intervals for probability of failure at a given time (Romanoschi and Metcalf 2000). In the study, bootstrapping was a procedure that involved random re-sampling of the existing data with replacement. In sampling with replacement, every data element was returned to the data set or sample after sampling, so that the observed data size could be more than a given sample. In this way the sample error could be reduced and the distribution could approximately follow a normal distribution.

With the bootstrapping approach, the probability distribution function (PDF) and cumulative distribution function (CDF) curves could be constructed at each percentile of the skewed non-warranted pavement distress data. The PDF and CDF curves could represent the characteristics of the non-warranty historical distress data and could also be applied to set up the measurement based threshold value based on the defined standard deviation for each individual distress type.

In order to achieve a smooth transition from the old deduct point based system to the new distress measurement based system, the confusion matrix method was applied to indicate the deviation between the deduct point based and new measurement based thresholds. The confusion matrix is an error matrix with two dimensions ("actual" and "predicted") that is used to visualize the summarized results of test data for which the true and false values are known. It is widely used in the performance evaluation of models in the field of machine learning and specifically the problem of statistical classification (Stehman 1997, Deng et al. 2016). In this study, the confusion matrix was used as an indicator to present the accuracy of the maintenance decision making by a new threshold being identical with that of the existing threshold.

2.5 Summary

This section summarizes the information about the practicing experience of selected pavement warranty programs in the US, with the specifications in the length of warranty periods, the types and threshold levels of distresses or performance/distress indicators, the methods to measure the distresses, and the corresponding remedial actions. The warranty program became popular in the US from 1990's in order to enhance the quality and to reduce the life-cycle costs of pavements. However, the pavement types involved in warranty, the number of warranty projects, the warranty period, and the warranty performance indicators vary among states because each state has different circumstances and requirements.

In an attempt to foster innovation and improve pavement quality, the state of Mississippi initiated the warranty program in 2000. Up to now, 21 pavement performance warranty projects were let with warranty periods from 5 to 10 years. The performance of the warranty pavement is surveyed annually with a video logging process. The distress evaluation is then performed by trained personnel in the office. The initial comparison between the distress conditions of the warranty versus non-warranty pavements showed better results of the warranty pavements in Mississippi.

The warranty program in Mississippi is quite special in that it uses deduct point based distress thresholds instead of direct distress measurement based thresholds. However, there have been questions regarding the conversion of distress measurements into deduct points using the empirically regressed conversion equations, which are not straightforward or easy to use. Hence, it is considered more appropriate to develop measurement based warranty thresholds by applying survival analysis to evaluate the performance of warranty and non-warranty pavements, adopting bootstrapping method to process the skewed PMS data, to develop thresholds based on normalized data distributions, using standard deviation, PDF, and/or CDF, and utilizing the optimized confusion matrix to have smooth transition from the existing system to the new system.

CHAPTER 3. EVALUATION OF WARRANTY AND NON-WARRANTY PAVEMENT PERFORMANCE IN MISSISSIPPI

The MDOT pavement warranty program was started in the year 2000 and with the accumulation of more than a decade of condition survey data. To date, MDOT has awarded a total of 21 warranty projects for both asphalt and concrete pavements in different warranty periods. The main type of pavement condition under warranty for pavement warranty projects in Mississippi is physical distress. In addition, ride quality data measured in IRI (in. /mi.), are collected for construction acceptance of projects in Mississippi. The distress assessment and evaluation methods specified for the warranty projects are the same as the methods that MDOT had used before 2010 (with manual data collection and deduct point equations). It is primarily because the warranty projects were all contracted before 2010 and these contracts are still active, and the manual distress survey method was regarded more appropriate for the project level (Qi et al. 2012, Battey and Watkins 2009). Therefore, one of the most important technical challenges in pavement warranty implementation is to find an appropriate evaluation method, which serves to examine the performance comparison between warranty and non-warranty pavements. This chapter attempts to establish a survival analysis based evaluation method and compare the performance of warranty versus non-warranty pavements using the field surveyed pavement distress data in Mississippi. Due to the Mississippi circumstance of using deduct point based thresholds in warranty contracting, two failure criteria, deduct point based and deduct point back-calculated, were employed to compare the performance of warranty and non-warranty pavements.

3.1 Initial Employment of Pavement Performance Evaluation Method for Warranty

In order to evaluate the performance of warranty versus non-warranty pavements, previous studies usually used distress indicators (e.g., rutting, cracking, etc.) in statistical tests such as pairwise comparisons and two-sample t-tests (Qi et al. 2013, West et al. 2011). Using linear regression, the Wisconsin DOT found that the warranted pavements performed better than the non-warranty pavements, with lower median Pavement Distress Index (PDI) and IRI values after 12 years in service (Battaglia 2009). Indiana DOT created CDF of IRI and rut depth as the criteria, and the results indicated that the warranted pavement performed longer and more cost-effectively compared with their counterparts (Sadeghi et al. 2016). Nonlinear regression models were also developed to estimate the time specific model parameters of general pavement performance combining experimental and field data (Prozzi and Madanat 2004). According to the comprehensive review of previous studies, it has been revealed that the majority of the methods used were devoted to evaluating the conditions of warranty versus non-warranty pavements by general statistical tests based comparisons for specific time points, but few efforts have been made to compare the possible differences in pavement performance over a continuous period of service time. Apparently, the time factor plays an important role in the performance evaluation of pavements, especially for pavement warranty program in which pavement distresses are monitored closely and pavement service quality is considered for the entire duration of the warranty period. Therefore, based on the literature review one research challenge is to explore an appropriate method that considers the influence of time factor on the evaluation of the pavement performance.

3.2 Deduct Point Method in Mississippi

Since the 1990s, MDOT pavement distress data were manually measured and stored in MDOT's PMS, and then converted to deduct points using deduct point curve equations, which were related to pavement engineers' perception of losing points in the rating of pavement performance based on the measured distress. Currently the thresholds for each distress type are deduct points converted from the corresponding pavement distress measurements which were conducted every year on warranted pavements. The deduct point conversion regression models which are developed in 1995 related on pavement engineers' perception of pavement performance for particular distress types (George 1995). The major advantage of using the deduct points and the conversion curves was that the conversions could change every individual distress type from its original measurement unit and value range to an unified performance rating scale (George 1995), so that different types of distresses became comparable on the same rating scale (100 points for a perfect pavement condition) and could be added together to form the composite PCR. However, the PCR model that combines multiple pavement distresses based on the deduct point conversions is more relevant for reporting conditions on the network level, rather than for the management of warranty projects which involve analysis at the project level.

With the start of the MDOT pavement warranty program in 2000, the distress measurements, the deduct point conversion equations, and the thresholds for individual distress were set up empirically (Qi et al. 2012, Battey and Watkins 2009). These equations and thresholds are developed based on the manually collected data. Later, the deduct point conversion equations for individual distress measurements were revised to reflect the distress density based pavement measurements and data collection technologies for pavement management generally. But it did not include the warranty projects because the warranty contracts were already in place. The current distress thresholds in deduct points for each warrantied distress type for asphalt and concrete pavements (Qi et al. 2012, Battey and Watkins 2009) are listed in Table 1 and Table 2, respectively.

3.3 Survival Analysis Method in Pavement

Survival analysis concerns analyzing the expected duration of time until the occurrence of an failing event (Dong and Huang 2014). In this study, the survival model was employed to identify the pavement performance over the service time up to the designated failures of asphalt pavements. In order to determine the failure time, a failure threshold should be defined for each distress. When the distress exceeds this threshold, the pavement is considered to have failed, thus the time that a pavement can survive without a distress failure can be identified (Wang et al. 2005).

In modeling the survival analysis, T is denoted as the survival time of the pavement structure in terms of the response variable or distress type, from the initial status along the service time *t*. *T* is determined by the actual time of failures of individual pavement sections in the sample, and $T \geq 0$. When the deteriorating density or rate $f(t)$ is given as a function of time *t*, $f(u)du$ is the probability of function f at t=u, the probability of having the survival time could be denoted as a duration function (a cumulative distribution function) $F(t)$, which is given by

$$F(t) = P(t \geq T) = \int_0^t f(u)\,du \qquad \textbf{3-1}$$

The duration function $F(t)$ represents the probability $P(t \geq T)$ that the pavement cannot survive (be failed) after *t* time of service. Based on the duration function $F(t)$, the survival function could be acquired. If $S(t)$ represents the probability of surviving beyond time *t* or that the pavement survives the past time *t*, it can be easy to see that

$$S(t) = P(T > t) = 1 - F(t) \qquad \textbf{3-2}$$

As t ranges from 0 to ∞, the survival function has the following properties

1) It is non-increasing, or monotonically decreasing;
2) At time $t = 0$, $S(t) = 1$, which means the probability of surviving past time 0 is 1, because no failure has yet occurred; and
3) At time $t = \infty$, $S(t) = S(\infty) = 0$, which means the survival curve goes to 0 as time approaches to infinity.

One important method employed to build the survival model is the hazard function $h(t)$, which measures the risk of failure of an entity at time t. The definition of the function is shown in equation 3-4. Actually, the hazard function measures the conditional probability of survival of an infinitesimally short time period between *t* and $t + \Delta t$, given that the pavement has survived up to time t. It is the instantaneous rate at which an event occurs, provided that no previous failure has happened (Dong and Huang 2014).

$$h(t) = \frac{f(t)}{S(t)} \tag{3-3}$$

In particular, by definition

$$f(t) = \lim_{\Delta t \to 0} \frac{F(t+\Delta t) - F(t)}{\Delta t} \tag{3-4}$$

The hazard function can be expressed as

$$\begin{aligned} h(t) &= \lim_{\Delta t \to 0} \frac{F(t+\Delta t) - F(t)}{\Delta t \Delta S(t)} = \lim_{\Delta t \to 0} \frac{P\left(t < T \le t + \Delta t\right)}{\Delta t \Delta S(t)} \\ &= \lim_{\Delta t \to 0} \frac{P\left(t < T \le t + \Delta t \middle| T > t\right)}{\Delta t} \end{aligned} \tag{3-5}$$

where T is survival time, and $\Delta t > 0$. In this study, $S(t)$ was estimated using the indicator functions associated with individual pavement sections.

The survival model is widely used in analyses of fatigue cracking deterioration, the effectiveness of maintenance treatment, and the pavement friction degradation (Wang et al. 2005, Li et al. 2017, Dong and Huang 2015). In this study, it assumed that every entity in the sample follows the same survival function (with no covariates or other individual differences), and the non-parametric Kaplan-Meier estimator would be the empirical survival function. Let φ_i be the indicator function for the event outcome of an entity, i denote an entity (a pavement section in the study), and n be the number of pavement sections. The indicator function φ_i takes the value 0 if the pavement section i is considered to have failed and 1 otherwise. By averaging the outcomes for all pavement sections in the sample, the Kaplan-Meier estimator $\hat{S}(t)$ is simply the proportion of pavement sections alive in the sample at time t, and can be regarded as a point estimate of the survival function $S(t)$ at time t (Rodrıguez 2005), i.e.,

$$\hat{S}(t) = \frac{1}{n} \sum_{i=1}^{n} \varphi_i \left\{ t \ge T \right\} \tag{3-6}$$

In order to fit the Kaplan-Meier survival probability curve to model the data in more detail, exponential, Weibull, Gamma, and Log Normal distributions could be used to estimate the survival distribution with parametric assumptions. In this study, Weibull and Log Normal distributions were employed as built-in distributions to fit the Kaplan-Meier survival probability curve using the open-source statistical software R 3.3.3. By specifying a parametric form for $S(t)$, it would be easy to compute for selected quantiles of the distribution, estimate the expected failure time, and derive a smooth function for estimating $S(t)$.

3.4 Effective Estimation of Warranties in Mississippi

3.4.1 Determination of Terminal Condition

MDOT network-level pavement distress data were manually measured and stored in MDOT's PMS from the early 1990s to 2010, and then converted to deduct points using deduct point curve equations. The warranty program began in 2000; therefore, deduct point based thresholds were used as the failure conditions in the survival analyses for the warranty pavements, and the failure conditions could be retrieved directly from the distress reports. A few years after 2000, MDOT began to move toward automated distress surveys for the pavements not in a warrant, although a prioritization factor and a repair decision tree have been used to decide maintenance for non-warranty pavements. There is no direct distress threshold has been set up for any terminal distress condition of the non-warranty pavements in Mississippi. Therefore, only with the deduct point based thresholds, it is difficult to evaluate the performance of warranty and non-warranty pavements on the same basis. In order to solve this problem, the deduct point based thresholds for alligator cracking, longitudinal cracking, and transverse cracking were used as the failure conditions in the survival analyses for the warranty pavements, and the failure conditions could be retrieved directly from the distress reports. For non-warranty pavements, the empirical conversion equations were used to convert from the deduct point based thresholds from the warranty pavements to direct distress measurement formatted terminal values.

The main items of pavement condition used for warranty projects in Mississippi are physical distresses. For each distress type, there is an equation to convert the distress measurement into deduct points. In this study, 10 asphalt warranty pavement projects each with a 7-year warranty period were selected as a data sample from the annual maintained HMA pavement distress reports in Mississippi for the warranty pavements including distress types of alligator cracking, longitudinal cracking, transverse cracking, rutting, and IRI. The deduct point based thresholds for alligator cracking, longitudinal cracking, transverse cracking, and rutting were used as the failure conditions in the survival analyses for the warranty pavements. In addition, although IRI has not been included as a warranty item in MDOT's current pavement warranty program, it is also collected regularly for construction quality acceptance and performance monitoring. As considered as a warranty item in other states (e.g., Florida, Indiana, and Illinois) and its threshold ranges from 90 to 133 in./mi., the IRI was included as a warranty item in this study and its threshold was set at 125 in./mi. based on a previous research effort (Qi et al. 2018).

The non-warranty asphalt pavement distress data collected from 2000 to 2014 were retrieved from MDOT's PMS to assess the average performance of the non-warranty pavements in Mississippi, the distress types included alligator cracking, longitudinal cracking, and transverse cracking. Table 9 lists the data sizes of the asphalt warranty and non-warranty pavements in survival analysis. In order to define the terminal condition for non-warranty pavements, the empirical conversion equations were used to convert from the deduct point based thresholds set up for the warranty pavements to direct distress measurement formatted terminal values. The procedure steps are as follows: 1) 10 warranty projects of asphalt pavements were selected to calculate percentages of samples at low, medium, and high severities respectively for the employed pavement distresses; 2) each deduct point based threshold value was broken down into three severity levels in proportion to the percentages; 3) the deduct points of three severity levels were used to calculate the distress amount in each severity level by the appropriate empirical conversion equations; and 4) a total distress measurement threshold value could be obtained by adding the calculated distress quantities in three severity levels.

The base percentages of distress measurements in deduct points for the warranty pavements at low, medium, and high severities were 49%, 45%, 6%; 58%, 35%, 7%; and 98.6%, 1.2%, 0.2% calculated from alligator cracking, longitudinal cracking, and transverse cracking respectively. Pavement with a larger percentage of high level severity distresses are generally more damaged than the pavements with a smaller percentage of high severity even at the same total lengths of cracking. Two scenarios of sensitivity analysis (SA I and SA II) for failure thresholds were set up by using the empirical percentages of (60%, 30%, 10%) and (75%, 20%, 5%) in two sets of distress severity levels respectively, which could address whether or not and how the distress severity distribution would affect the survival analysis results. Therefore, the total distress threshold values could be recalculated based on the two sets of percentages. In this study, the deduct point and back-calculation methods were employed to define the terminal failure conditions for the warranty and non-warranty pavements.

Table 9 Data Sizes of Warranty and Non-warranty Pavement for Survival Analysis

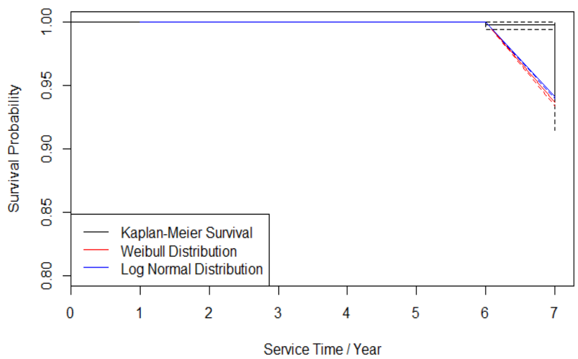
Warranty Pavements				Non-warranty Pavements	
Location	Data Size /segments	Location	Data Size /segments	Distress Type	Data Size /section
Bolivar US61	319	Grenada Yalobusha I55	736	Alligator Cracking	33797
Carroll Montgomery I55	412	Lincoln I55	284	Longitudinal Cracking	40699
Coahoma US61	381	Montgomery Carroll Grenada I55	504	Transverse Cracking	11790
Desoto MS302 Swinnea Rd to US78	604	Simpson US49 Blain	128	—	—
Desoto MS302 US51 to Swinnea Rd	301	Toyota Rd MS-780	191	—	—

3.4.2 Survival Analysis of the Warranty and Non-warranty Pavements

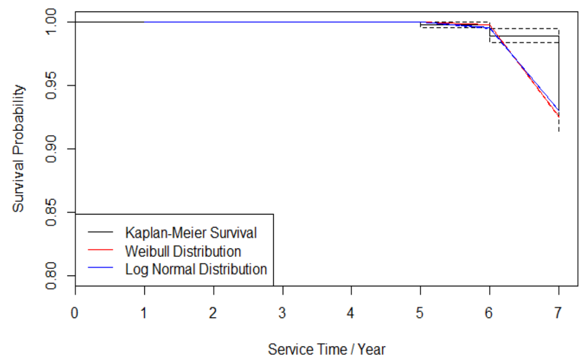
The survival analysis was conducted by running the noncommercial statistical software R 3.3.3. No assumption of distribution was needed for the warranty and non-warranty pavements distress data, and due to its applicability to unknown distributions, the Kaplan-Meier estimator was adopted to non-parametrically analyze the survival probabilities of the pavements. Meanwhile, several distribution models were fitted to generate survival probability function curves in the survival analyses (Jackson 2016), and eventually the Weibull and Log Normal time-to-event models along with their model parameters that best fitted the Kaplan-Meier curves maximizing the log-likelihood were selected as most appropriate survival probability fitting models.

3.4.2.1 Survival Analysis of Warranty Pavements

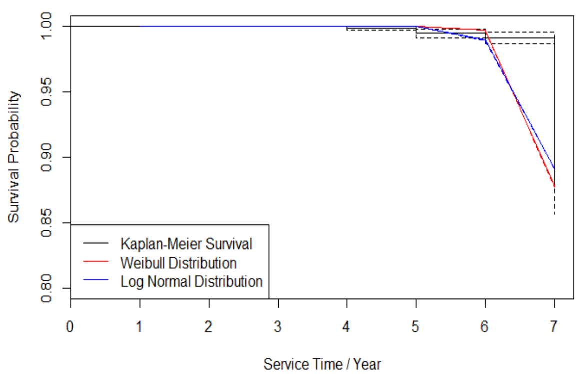
The Weibull and Log Normal distributions have seen frequent applications in reliability analysis with lifetime (or failure time) data. Survival probability fitting curves obtained by Weibull and Log Normal distributions were used to assess the deterioration of pavement performance.



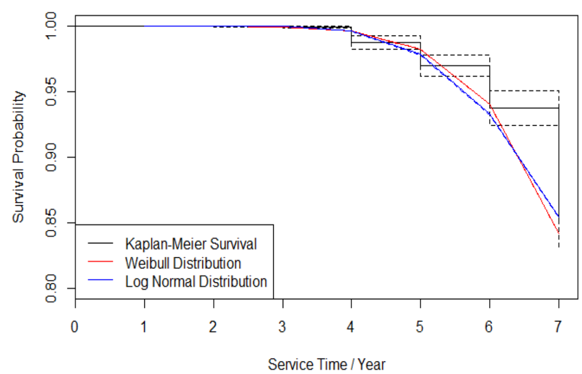
(a) Alligator Cracking



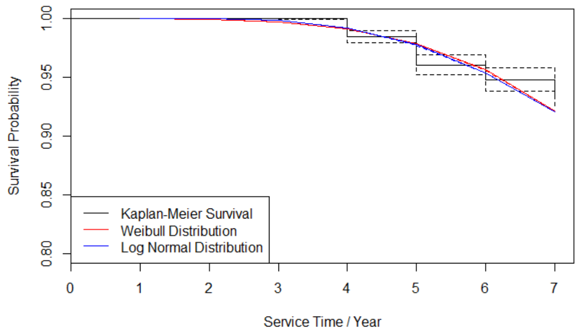
(b) Longitudinal Cracking



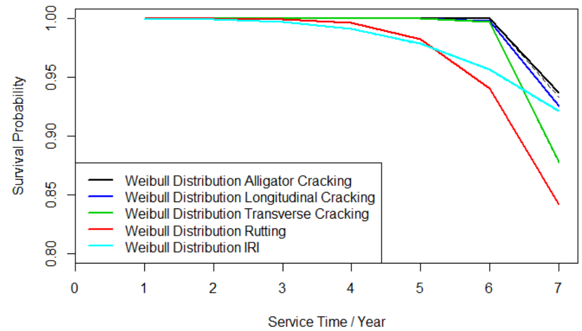
(c) Transverse Cracking



(d) Rutting



(e) IRI



(f) All

Figure 3 Survival Probabilities and Fitting Curves of Warranty Pavements

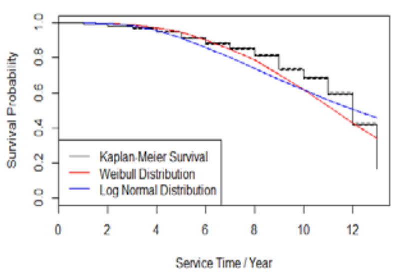
Figure 3 (a)-(e) presents the Kaplan-Meier survival curves and fitting curves of Weibull and Log Normal distributions of different distresses for warranty pavements. It is noticeable that the Weibull distribution model can generate better fitting curves obtaining maximum log-likelihood and therefore Figure 3(f) collects the Weibull distribution fitting curves for all the distress types.

As shown in Figure 3, the survival probabilities of warranty pavements are equal to 1 for all distresses during the initial service time, which indicates that there is only a small possibility of failure at the initial time. After the initial time, the survival probability curves start to fall slightly over time. According to Figure 3 (a)-(c), it is observed that the failures of alligator cracking, longitudinal cracking, and transverse cracking start to accelerate at around the 5th or 6th year of service time, while the rutting and IRI performances start to deteriorate at the 3rd or 4th year as shown in Figure 3 (d)-(e). This phenomenon is presented comparatively in Figure 3 (f) with all the fitting curves of the Weibull distribution model included in the same figure. Moreover, as seen in Figure 3(f), the order of the deteriorating speeds of these distresses from the highest to the lowest would be rutting, transverse cracking, IRI, longitudinal cracking, and alligator cracking, and the performance deterioration of the alligator cracking occurs the latest.

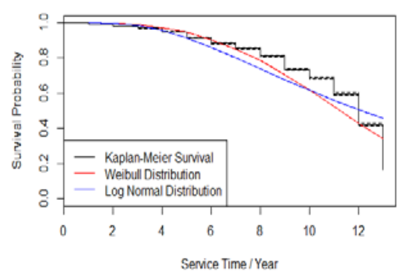
3.4.2.2 Survival Analysis of Non-warranty Pavements

Figure 4 shows the survival probability curves of different distresses for the non-warranty pavements in the base case and two scenarios of SA I and SA II with different distress severity distributions.

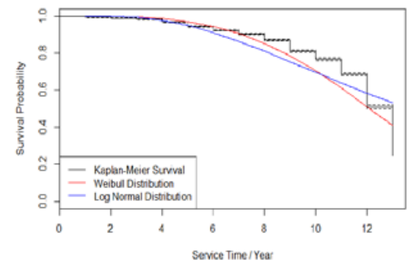
The survival probabilities of all the three types of cracking decrease continuously along with service time for the base case of distress severity distribution, and the two scenarios with varied severity distributions of distresses in SA I, and SA II. In addition, the fitting curves of the Weibull distribution model are more accurate than those of the Log Normal distribution model, because the Weibull distribution model appears to fit the Kaplan-Meier survival curve better than the Log Normal model. Therefore, the Weibull distribution model was employed to present the survival probability fitting curves of different distresses for the non-warranty pavements.



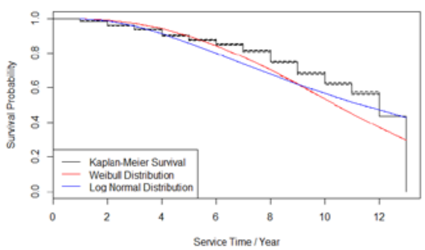
(a) Alligator Cracking



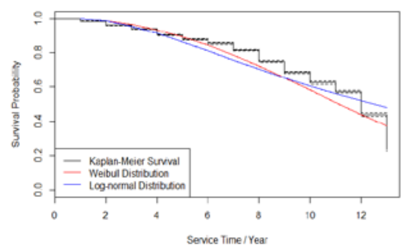
(b) Alligator Cracking (SA I)



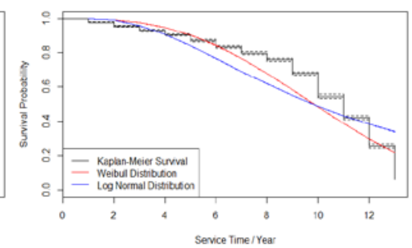
(c) Alligator Cracking (SA II)



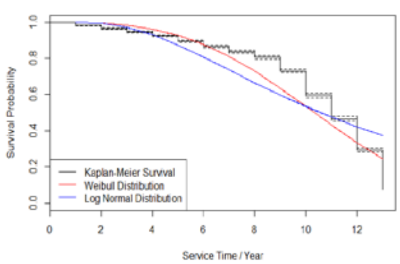
(d) Longitudinal Cracking



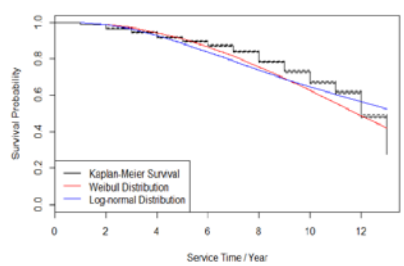
(e) Longitudinal Cracking (SA I)



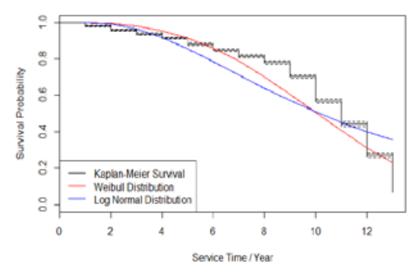
(f) Longitudinal Cracking (SA II)



(g) Transverse Cracking



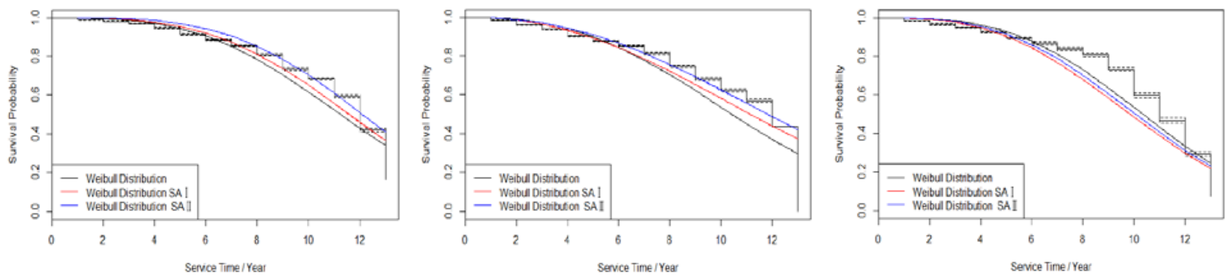
(h) Transverse Cracking (SA I)



(i) Transverse Cracking (SA II)

Figure 4 Survival Probabilities and Fitting Curves of Non-warranty Pavements

Figure 5 shows the survival probability curves for the base case and fitting curves of the Weibull distribution for the base, SA I, and SA II scenarios of three types of cracking distresses for the non-warranty pavements. The fitting curves of the Weibull distribution for the base, SA I, and SA II scenarios have similar deterioration trends, and the survival probabilities decrease along with service time. However, it shows that the most serious distress types in base condition and SA II are transverse cracking, while longitudinal cracking is the most deteriorated distress type for SA I. Therefore, the analytical results indicate that the choice of different failure criteria of distresses and the distress severity distribution of the non-warranty pavements do have a significant impact on the results of the survival analyses.

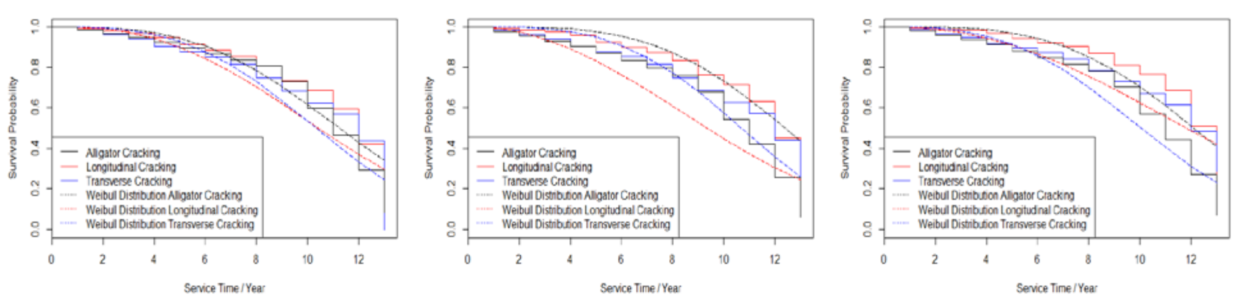


(a) Alligator Cracking (b) Longitudinal Cracking (c) Transverse Cracking

Figure 5 Sensitivity Analyses of Distress Severity Distributions for Non-warranty Pavements in Survival Probability and Weibull Fitting Curves

3.4.2.3 Comparison of Warranty and Non-warranty Pavements

The information displayed in Figure 6 is the combination of Kaplan-Meier survival curves and Weibull distribution fitting curves of different distresses for the non-warranty pavements. As shown in Figure 6, the results indicate that the survival probabilities of the three types of cracking decrease over time at increasing deceleration rates. Different from the situations in warranty pavements, the non-warranty pavements gain cracking distresses at an early stage. The non-warranty pavements start to deteriorate only after about 2 years of service time in the three types of cracking, which confirms the superiority of the warranty pavements to the non-warranty pavements. Furthermore, transverse cracking has the lowest survival probability in the three cracking types. As observed before, the comparison of survival analysis results in the three scenarios of base condition, SA I, and SA II shows that out of the three types of cracking, alligator cracking has the highest survival probability for all the time; the fitting curves in Figure 6 (a) and (c) show that the transverse cracking performance is the worst during the service time in the base condition and SA II; the fitting curves in Figure 6 (b) show that longitudinal cracking performance is the worst in SA I. The main reason would be that different failure thresholds of distresses may affect the results of the survival analyses, which could also be concluded from the analyses of sensitivity analyses of base condition, SA I, and SA II. Therefore, it is vitally important to define the terminal condition of different distresses for a meaningful survival analysis.



(a) Base condition (b) SA I (c) SA II

Figure 6 Survival Probabilities and Fitting Curves for Distress Types, Threshold Types, and Severity Distribution Scenarios of Non-warranty Pavements

Comparisons of Figure 3 (a)-(c) and Figure 4 (a), (d), and (g) show that the survival probabilities of the warranty pavements are much higher than those of the non-warranty pavements during the same length of service time. Although the service periods of the warranty and non-warranty pavements are different in the datasets, the fitting trend curves still indicate that the average performance of warranty pavements is better than that of the non-warranty pavements. Furthermore, it can be observed that with the warranty period of 7 years in the dataset, there would be only few or even no failures of the three types of cracking in the initial 4 years of service in the warranty pavements. Therefore, the pavement warranty program could be an effective way of maintaining the performance of pavements. Moreover, in warranty pavements the most serious distress is rutting, while in the non-warranty pavements, the most serious distress would be longitudinal cracking.

3.5 Summary Remarks

This section uses survival analysis method to evaluate the performance of warranty pavements, as well as that of the non-warranty pavements. The research serves to emphasize that: 1) the survival analysis is an effective way to assess the performance of warranty and non-warranty pavements, especially for the skewed non-warranty distress data; 2) the average performance of warranty pavements is superior to that of the non-warranty pavements; 3) the beginning of the distress failures for the warranty and non-warranty pavements could be checked easily in the survival curves; and 4) the most serious distress types in warranty and non-warranty pavements could be illustrated in the comparison of the survival performance lines.

The survival analysis demonstrates the numerical assessments of the relative deterioration trends of pavement performance during service time of the warranty and non-warranty projects in Mississippi using the survival probability concept. The results show that the deteriorating speeds of non-warranty pavements are obviously faster than the warranty pavements, especially at the beginning of the warranty pavement service time. Moreover, the decreasing rates of the survival probability curves of the warranty pavements are slower than those of the non-warranty pavements for individual distresses. These results indicate that the performance of the warranty pavements is better than the non-warranty pavements in Mississippi.

As the analysis has revealed, it is important to have the same distress criteria defined for the terminal conditions for the warranty and non-warranty pavements in order to apply the survival analysis method to compare the performance of warranty versus non-warranty pavements. It seems rather unique for Mississippi to use the deduct points based warranty distress thresholds and the empirical conversion equations to evaluate the performance of the warranty pavements in the current warranty practice while the non-warranty pavements are evaluated by direct distress measurements. There are differences in the data collection method and technology between the warranty and non-warranty pavements in Mississippi. However, Mississippi may not be alone in having these discrepancies and inconsistence in the management of warranty and non-warranty pavements. Therefore, it would be very possible and practical for new evaluation methods and/or technologies to be adopted for the pavements with a contracting warranty, which could make the evaluation of the pavement performance easier both for the warranty and non-warranty pavements. Further, based on the survival concept, the failure probability assessments of the warranty and non-warranty pavements during service time have shown that the performance of the warranty pavements is significantly better than that of the non-warranty pavements at the same service time level (Luo et al. 2018).

CHAPTER 4. DATA DESCRIPTION

In order to update the distress threshold values for the warranty projects in Mississippi, distress survey data from both warranty and non-warranty pavements were used in the study. The pavement distress survey data applied in this study included warranty and non-warranty pavement data over various service time periods. This section describes the data sources, data items, data preparation, data screening, and data processing and visualization.

4.1 Data Sources

The pavement condition survey of the state-maintained highway system is conducted every two years by dividing the road system into homogeneous pavement analysis sections for all state-maintained pavements to monitor the overall status of the state's roadways in Mississippi. The contractor collects both condition and distress data. Condition data includes the following: IRI, PCR, roughness rating, rut depth, faulting, and texture. The distress data includes cracking, potholes, patching, punch-outs, and joint deterioration.

Up to 2008, the longitudinal profile was collected using a South Dakota profiler that uses laser sensors for measurement. The data collection vehicle was mounted with five video cameras to capture images of the shoulders, wheel paths, and the right-of-way. The images were then digitized into frames for analysis purposes with one frame representing approximately 50 feet. The data collection vehicle was also equipped with a GPS receiver to log coordinate data. Distress evaluation was then performed on the digitized image using human raters. The distress data were quantified in terms of severity and extent on 500-ft samples within each analysis section. A random sampling of approximately 20 percent was used for the distress evaluation. Starting with the 2010 survey, MDOT began to have data collection vendors automatically collect 100% of the lane rather than samples. Roughness, rutting, faulting, and texture measurements are collected on 100 percent of the state-maintained system.

Rutting and IRI are collected as an average value per section and percentages of values in category low, medium and high for every section. The category of rutting is defined 1/16 to 1/8 inch as low, 1/8 to 1/4 inch as medium, and larger than 1/4 inch as high. The rutting extent is calculated by length of rutting in one category divided by the total probable rutting. Total probable rutting is equal to the length of the sample/section. For example, if a 6000 feet section has 3000 ft. low rutting, the percentage of low rutting is 3000/6000 which equals 50 percent. For the IRI, the category is classified 0 to 150 in/mi as low, 150 to 300 in/mi as medium, and larger than 300 in/mi as high.

Table 10 Datasets of Non-warranty Projects

Distress Number	Distress Type	# of Sections at Severity Level			Pavement Type
		Low (0)	Medium (1)	High (2)	
0	Longitudinal Cracking	1148	534	207	Concrete
1	Transverse Cracking	1155	1205	453	Concrete
5	Faulting of Transverse Joints	0	0	0	Concrete
7	Corner Break	509	325	154	Concrete
9	Joint Seal Deterioration	395	291	161	Concrete
10	Spalling-Longitudinal	414	152	79	Concrete
11	Spalling-Transverse	755	457	310	Concrete
12	Map Cracking	42	9	5	Concrete
19	Alligator Cracking	24595	14629	1696	Asphalt
20	Block Cracking	14141	9616	1073	Asphalt
21	Edge Cracking	7023	5138	2019	Asphalt
22	Longitudinal Cracking	33723	20319	2970	Asphalt
23	Transverse Cracking	34961	24046	5461	Asphalt
27	Potholes	2512	0	0	Asphalt
29	Bleeding	3750	2055	841	Asphalt
30	Reflective Cracking-Transverse	5814	5710	1185	Asphalt

Table 11 Datasets of Non-warranty Projects at Survey Years

Distress Number	# of Sections at Survey Year								Pavement Type
	2000	2002	2004	2006	2008	2010	2012	2014	
0	245	163	154	161	123	330	303	410	Concrete
1	597	263	241	257	290	365	332	468	Concrete
5	0	0	0	0	0	0	0	0	Concrete
7	110	112	115	132	80	147	179	113	Concrete
9	9	11	21	34	46	241	212	273	Concrete
10	76	72	76	99	14	35	144	129	Concrete
11	116	110	111	113	101	194	352	425	Concrete
12	21	2	7	2	4	5	11	4	Concrete
19	4765	4436	4488	5858	5308	5096	6678	4291	Asphalt
20	1059	1116	1528	2202	1818	6563	6707	3837	Asphalt
21	908	372	399	769	1280	3105	7202	145	Asphalt
22	4452	4362	5152	5978	8202	9101	8554	11211	Asphalt
23	6167	5979	6273	6783	8253	9285	8865	12863	Asphalt
27	620	28	30	106	89	739	165	735	Asphalt
29	685	330	424	304	175	1661	3025	42	Asphalt
30	1695	1341	1628	1581	1573	1233	2511	1147	Asphalt

Based on the data recorded in MDOT's PMS database, the distress data collected from asphalt (COMP, FLEX, and OFLEX) and concrete (CRCP, JCP, and JRCP) pavements were used as non-warranty historical pavement data in the analysis. The period of the distress data was from 2000 to 2014, because 2000 was the warranty program's starting time in Mississippi and 2014 was the latest time with completed network-level distress survey. Based on the available data recorded in MDOT's PMS, the rutting and IRI data in network-level pavements collected from 2000 to 2016 were used in this research. Table 10 lists the dataset of road sections for each severity level of distresses used to set up the distress threshold values. Table 11 presents the survey year of road sections for each individual distress type of the non-warranty pavements. The dataset sizes for the concrete pavements are much smaller than those of the asphalt pavements as shown in these tables. The datasets of the rutting and IRI for both asphalt and concrete pavements from 2000 to 2016 are shown in Table 12 and Table 13.

Table 12 Datasets for Rutting and IRI of Non-warranty Asphalt Pavement Projects

Section Survey Year	# of Sections at Severity Level			Average Value (# of sections)
	Low	Medium	High	
2000	4869	4869	4869	4869
2002	5019	5019	5019	5019
2004	5226	5226	5226	5226
2006	5266	5266	5266	5266
2008	5472	5472	5472	5472
2010	5520	5520	5520	5520
2012	5553	5553	5553	5553
2014	5614	5614	5614	5614
2016	5691	5691	5691	5691

Table 13 Datasets for Rutting and IRI of Non-warranty Concrete Pavement Projects

Section Survey Year	# of Sections at Severity Level			Average Value (# of Sections)
	Low	Medium	High	
2000	319	319	319	319
2002	279	279	279	279
2004	254	254	254	254
2006	246	246	246	246
2008	241	241	241	241
2010	239	239	239	239
2012	250	250	250	250
2014	201	201	201	201
2016	203	203	203	203

The warranty pavement data were retrieved from the annual maintained HMA and PCC pavement distress reports of the warranty projects in Mississippi. The projects were surveyed every year during the warranty periods. MDOT continued to use manual rating and sampling on the in-house-collected warranty projects since this was better suited to project-level rating and since most warranty contracts were in place prior to 2010. Two random one-tenth mile segments in each mile were selected for pavement distress evaluation before 2017, while 100% of each mile (totally 10 one-tenth mile segments) were evaluated after 2017 in Mississippi. To date, MDOT has awarded a total of 21 warranty projects for both asphalt and concrete pavements in different warranty periods. In the study, typical distress measurements of pavement warranty projects were selected as warranty pavement data. All warranty projects used in this research were completed before 2017; therefore, two random one-tenth mile segments in each mile were selected to present the pavement condition of the warranty pavements. Each surveyed pavement segment recorded for a distress type was considered as one sample element, except for the empty record items. The warranty pavement segments with maintenance decisions illustrated in Table 14 and Table 15 (not including the zero record items), were employed to evaluate the rationality of distress threshold values. Similarly to the non-warranty pavement datasets, the sizes of datasets for the warranty concrete pavements are much smaller than those of the warranty asphalt pavements as shown in the tables.

Table 14 Datasets of Warranty Projects

Warranty Project	Remedial Action (# of Segments)		Pavement Type
	NO (0)	YES (1)	
BolivarUS61	510	103	Asphalt
Carroll I-55	55	74	Concrete
Carroll/Montgomery I-55	261	13	Asphalt
Covington US-84	697	228	Asphalt
Desoto I-55	11	0	Concrete
Desoto MS-302 (US-51 to Swinnea Rd)	105	34	Asphalt
Desoto/Tunica I-69 MS-304	38	0	Concrete
Grenada/Yalobusha I-55	557	57	Asphalt
Lauderdale MS-19 NB	50	0	Asphalt
Lauderdale MS-19 SB	125	8	Asphalt
Lincoln I-55	68	2	Asphalt
Montgomery US-82	184	27	Asphalt
Montgomery/Carroll/Grenada I-55	249	10	Asphalt
Newton/Neshoba MS-19	325	46	Asphalt
Pontotoc MS-9	192	10	Asphalt
Simpson US-49 Superior	74	0	Asphalt
Toyota Rd MS-780	56	0	Asphalt

Table 15 Datasets of Warranty Projects for Distress Types

Distress Number	Distress Description	Remedial Action (# of Segments)		Pavement Type
		NO (0)	YES (1)	
0	Longitudinal Cracking	3	9	Concrete
1	Transverse Cracking	4	7	Concrete
5	Faulting of Transverse Joints	49	0	Concrete
7	Corner Breaks	30	13	Concrete
10	Spalling of Longitudinal Joint	5	3	Concrete
11	Spalling of Transverse Joint	13	42	Concrete
19	Alligator Cracking	231	65	Asphalt
20	Block Cracking	108	141	Asphalt
22	Longitudinal Cracking	1452	32	Asphalt
23	Transverse Cracking	1003	237	Asphalt
27	Potholes	0	7	Asphalt
200	Rutting	659	56	Asphalt

4.2 Distress Items

The pavement condition and distress data for warranty projects considered in this study include IRI, rutting, and the distress types that are listed in Table 16. In current warranty contracting, rutting is employed as a distress indicator in warranty asphalt pavements and is recorded and classified as percentages in low, medium, and high severity levels. There is a deduct point based threshold for rutting that is converted by plugging in the percentages at low, medium, and high severity categories to empirical equations. Although the IRI was not considered as a distress indicator in previous warranty contracts, the average IRI and average rutting depth for each segment (0.1 mile) were surveyed to evaluate the pavement condition for the warranty pavements. From the state of practice of other states and the accumulated past data collections, the rutting depth and IRI data are analyzed in this study to research the possible way of including the IRI as indicators in the new warranty program.

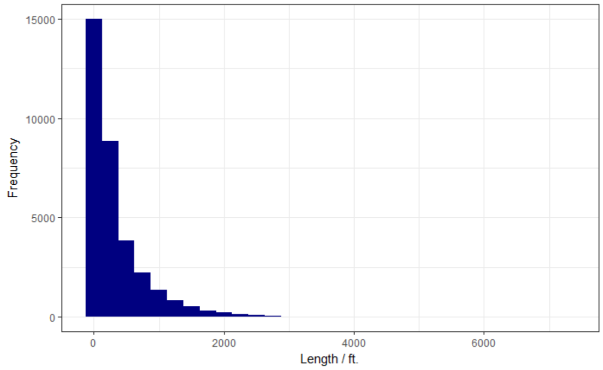
The warranty pavement distress measurements stored in the annual maintained pavement distress reports were collected in 500 ft. long segments for each distress type. For comparison purposes, each of the 500 ft. long pavement sample segments within every pavement analysis section was considered as a sample element. Likewise, the distress measurements stored in the PMS database were converted to consistent units and organized by distress type, lane, analysis section, and pavement survey time. Then the section based distress measurements were divided by their respective lengths of sections and multiplied by segment length of 500 ft. to arrive at segment based distress measurements that are equivalent to the warranty sampling element. Therefore, every 500 ft. length segment was considered as one sample element for both warranty and non-warranty pavements in this study.

Table 16 Distress Types in Warranty Projects

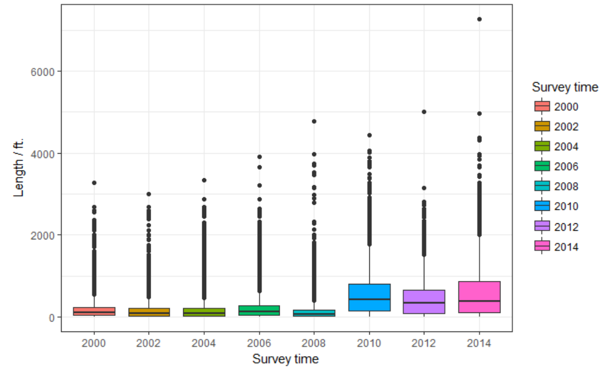
Pavement Type	Distress #	Distress Type and Measurement Unit	Pavement Type	Distress #	Distress Type and Measurement Unit
Asphalt	19	Alligator Cracking/sq.ft.	Concrete	0	Longitudinal Cracking/ft.
	20	Block Cracking/sq.ft.		1	Transverse Cracking/ft.
	21	Edge Cracking/ft.		5	Faulting of Transverse Joints/no.
	22	Longitudinal Cracking/ft.		7	Corner Breaks/no.
	23	Transverse Cracking/ft.		9	Joint Seal Deterioration/no.
	27	Potholes/no.		10	Spalling of Longitudinal Joint/in.
	29	Bleeding/sq.ft.		11	Spalling of Transverse Joint/in.
	30	Reflection Cracking/ft.		12	Map Cracking & Scaling/sq.ft.
	200	Rutting/percentage			Average IRI/(in/mi)
		Average Rutting Depth/in			
		Average IRI/(in/mi)			

4.3 Data Screening

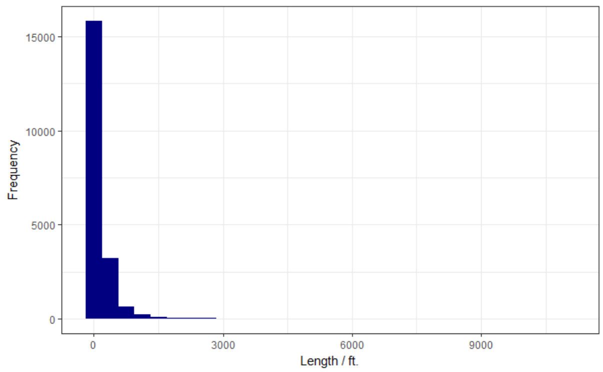
Prior to the data analysis, the first step was to cleanse the dataset to remove all the null and unusable records in the PMS database and warranty reports. Faulting of transverse joints, corner breaks, joint seal deterioration, and map cracking & scaling in concrete non-warranty pavements and potholes in asphalt non-warranty pavements did not have valid data or the values were all zeros. Therefore, these distress types could not be used for the statistical analyses and comparisons. In this study, the warranty thresholds were set up based on the average experience of the pavement network in Mississippi; hence, the outliers of the distress measurements in the PMS database needed to be cleansed out. Due to rater subjectivity (earlier) and (later) automated distress detection not being perfect, the distress measurement raw data recorded in the PMS contains outliers. For instance, the histograms of longitudinal cracking for asphalt pavement in Figure 7 are apparently skewed, and the boxplots ordered by survey time show a lot of outliers in each service year.



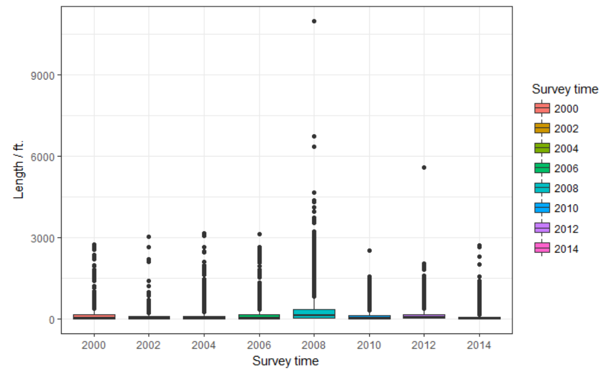
(a) Histogram at Low Severity Level



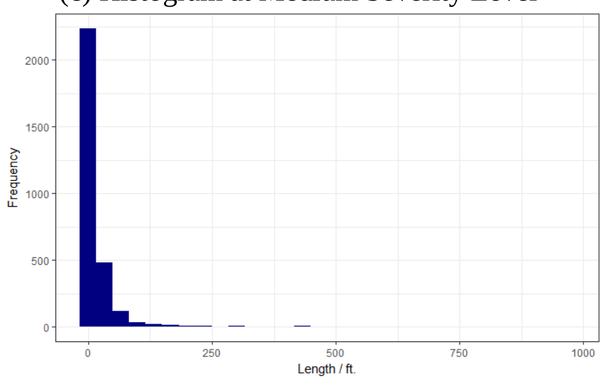
(b) Boxplot at Low Severity Level



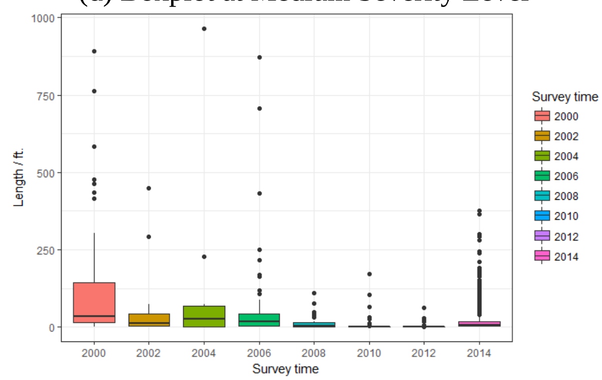
(c) Histogram at Medium Severity Level



(d) Boxplot at Medium Severity Level



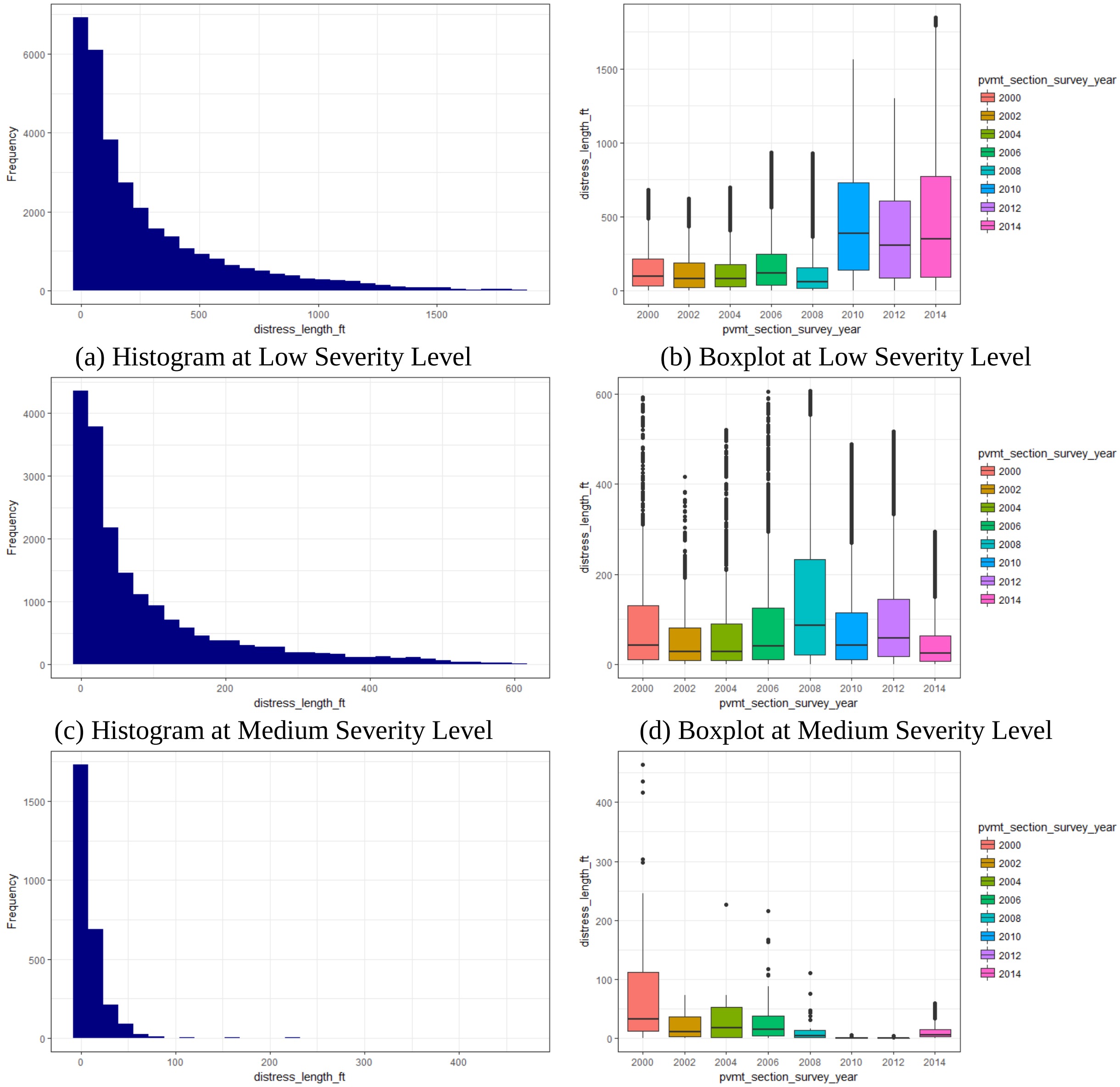
(e) Histogram at High Severity Level



(f) Boxplot at High Severity Level

Figure 7 Histograms and Boxplots of Longitudinal Cracking Raw Data for Non-warranty Asphalt Pavements

Therefore, the data between the 95% confidence intervals were selected to reduce outliers and improve the accuracy of statistical estimates of the characteristics of the distresses. The cleansed results for non-warranty longitudinal cracking in Figure 8 show that the tails of histograms are reduced, and the medians are closer to the middles of the boxplots, which implies that the 95% confidence interval data censoring is effective.



(a) Histogram at Low Severity Level

(b) Boxplot at Low Severity Level

(c) Histogram at Medium Severity Level

(d) Boxplot at Medium Severity Level

(e) Histogram at High Severity Level

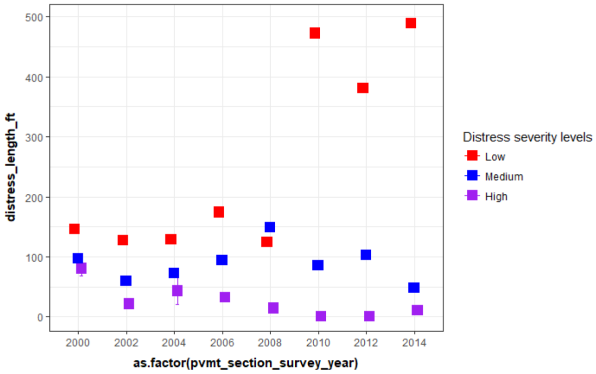
(f) Boxplot at High Severity Level

Figure 8 Histograms and Boxplots of Longitudinal Cracking Cleansed Data for Non-warranty Asphalt Pavements

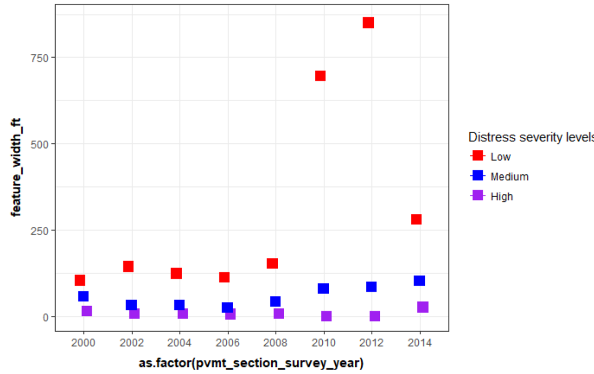
4.4 Statistical Analysis and Visualization

4.4.1 Overview of Distresses Data

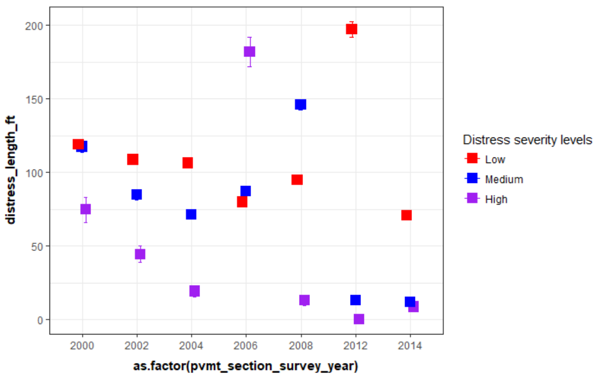
To investigate the non-warranty pavement distress characteristic over time, distress data of each individual distress type at different severity levels (outlier reduced by 95% confidence interval) were categorized into groups of different service years. Using the available distress data for both asphalt and concrete pavements of non-warranty contracts, survey data of every two years from 2000 through 2014 were categorized into 8 groups for each severity level of a distress type. Figure 9 and Figure 10 present the boxplots with error bars of each individual distress at three



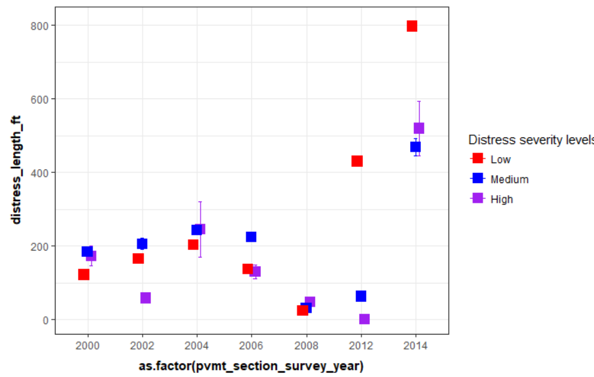
(a) Longitudinal Cracking



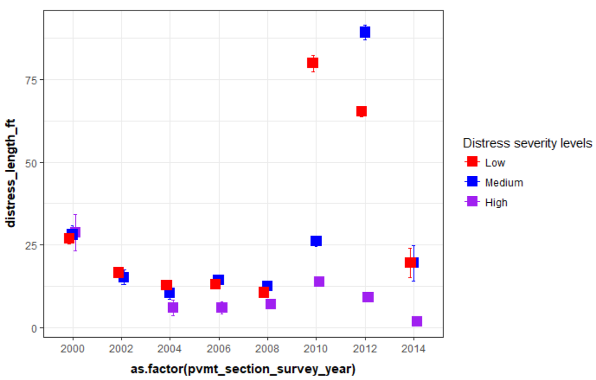
(b) Transverse Cracking



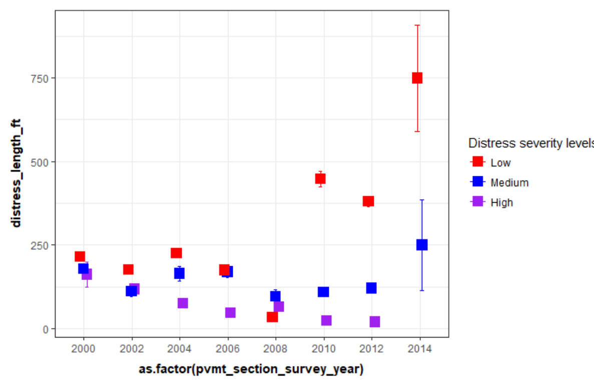
(c) Alligator Cracking



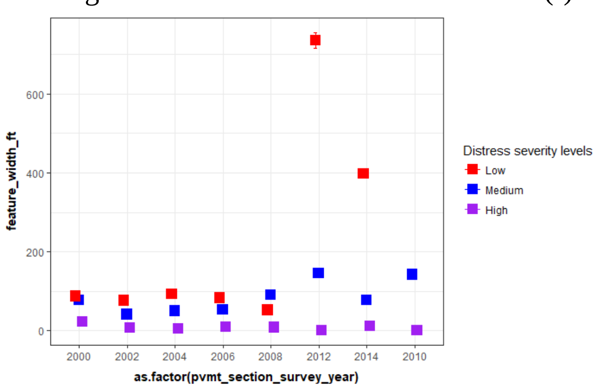
(d) Block Cracking



(e) Edge Cracking



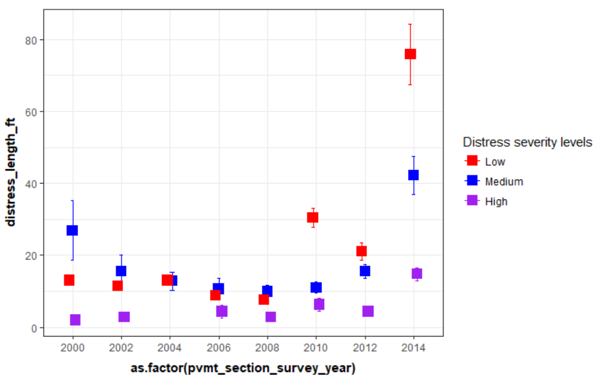
(f) Bleeding



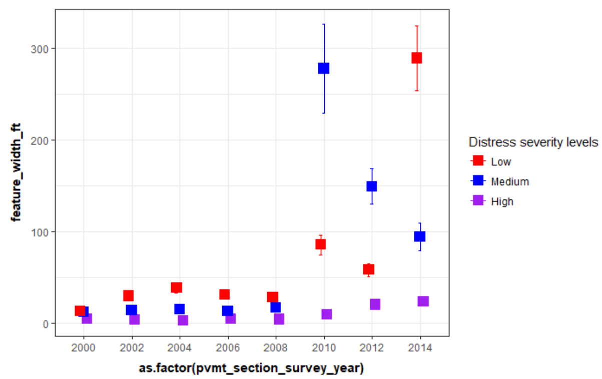
(g) Reflection Cracking

Figure 9 Boxplots of Distresses with Error Bars for Asphalt Pavements

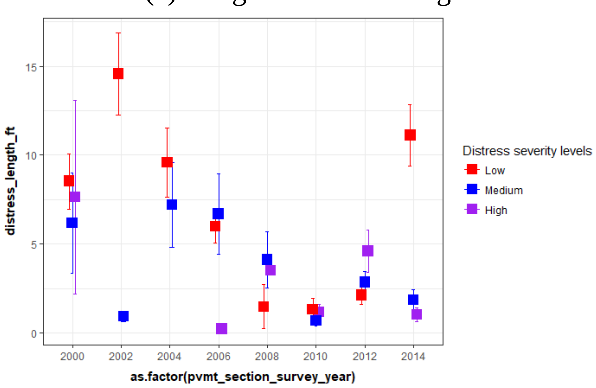
severity levels for asphalt and concrete pavements. The figures show that the measurements at low severity level are highest, and that at high severity level are lowest in almost every service year for each distress, except for few service years. The measurements have an obvious increase starting at 2010 for all of the distresses in asphalt and concrete pavements, especially at low severity level. Because MDOT switched vendor companies around that time, the data collection results may be affected by the change of the data vendors. Also, the 2010 survey was the first wherein the vendor used automated distress detection technology as opposed to two 500-ft samples per mile used previously.



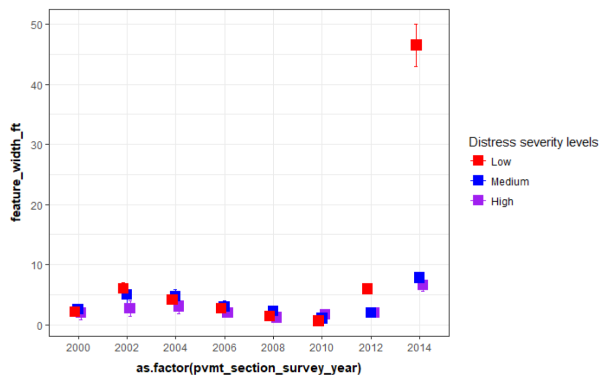
(a) Longitudinal Cracking



(b) Transverse Cracking



(c) Spalling of Longitudinal Joint



(d) Spalling of Transverse Joint

Figure 10 Boxplots of Distresses with Error Bars for Concrete Pavements

4.4.2 Distresses in Asphalt Pavements

After data screening, there were valid non-warranty data for longitudinal cracking, transverse cracking, alligator cracking, block cracking, edge cracking, bleeding, and reflection cracking to be used for the asphalt pavements.

(1) Longitudinal Cracking in Asphalt Pavements

The datasets of longitudinal cracking in non-warranty asphalt pavements are ordered by survey year for each severity level in Table 17. The table shows the numbers of road sections contained in longitudinal cracking at three severity levels in service years. There are the highest number of sections in longitudinal cracking at the low severity level with fine cracking and/or few interconnecting cracks. The statistical descriptive figures presented in Figure 8 indicate that the measurement data at the three severity levels are all right-skewed. The numbers of longitudinal cracking occurring in warranty projects with maintenance decisions “Yes, need remedial action” and “No, does not need remedial action” are listed in Table 18. It shows that most of the data records (segments in warranty) do not need maintenance as they do not exceed the deduct point threshold.

Table 17 Longitudinal Cracking Datasets in Non-warranty Asphalt Pavements

Section Survey Year	Severity Level		
	Low (0)	Medium (1)	High (2)
2000	3473	904	75
2002	3598	740	24
2004	3680	1461	11
2006	4201	1711	66
2008	4510	3650	42
2010	5100	3754	247
2012	4239	4134	181
2014	4922	3965	2324

Table 18 Longitudinal Cracking Datasets in Warranty Asphalt Pavements

Warranty Project	Remedial Action	
	NO (0)	YES (1)
Bolivar US61	179	22
Carroll/Montgomery I-55	74	0
Covington US-84	390	0
Desoto MS-302 (US-51 to Swinnea Rd)	26	0
Grenada/Yalobusha I-55	92	0
Lauderdale MS-19 NB	24	0
Lauderdale MS-19 SB	78	0
Lincoln I-55	21	0
Montgomery US-82	114	0
Montgomery/Carroll/Grenada I-55	77	0
Newton/Neshoba MS-19	218	0
Pontotoc MS-9	100	10
Simpson US-49	52	0
Toyota Rd MS-780	7	0

(2) Transverse Cracking in Asphalt Pavements

Table 19 lists the numbers of road sections contained in the transverse cracking of the non-warranty asphalt pavements at three severity levels over service years. Most of the sections containing transverse cracking are at the low and medium severity levels. The statistical descriptive figures presented in Figure 11 indicate that the measurement data at the three severity levels are right-skewed. The numbers of transverse cracking occurring in the warranty projects with maintenance decisions are listed in Table 20. It shows that the most numbers of segments in in the warranty projects does not exceed the deduct point threshold for transverse cracking.

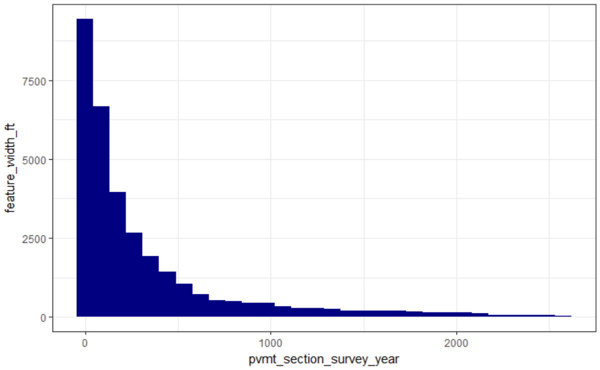
Table 19 Transverse Cracking Datasets in Non-warranty Asphalt Pavements

Section Survey Year	Severity Level		
	Low (0)	Medium (1)	High (2)
2000	3836	2068	263
2002	3936	1911	132
2004	4026	2204	43
2006	4330	2286	167
2008	4658	3531	64
2010	5090	3559	636
2012	4201	3967	697
2014	4884	4520	3459

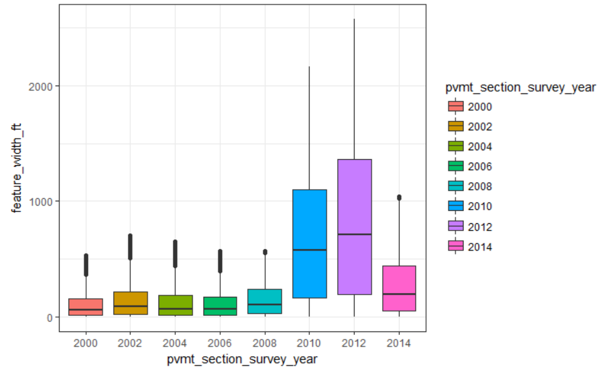
Table 20 Transverse Cracking Datasets in Warranty Asphalt Pavements

Warranty Project	Remedial Action	
	NO (0)	YES (1)
Bolivar-US61	160	18
Carroll/Montgomery I-55	159	6
Covington US-84	111	110
Desoto MS-302 (US-51 to Swinnea Rd)	13	0
Grenada/Yalobusha I-55	158	42
Lauderdale MS-19 NB	7	0
Lauderdale MS-19 SB	38	4
Lincoln I-55	25	2
Montgomery US-82	24	1
Montgomery/Carroll/Grenada I-55	131	8
Newton/Neshoba MS-19	85	46
Pontotoc MS-9	59	0

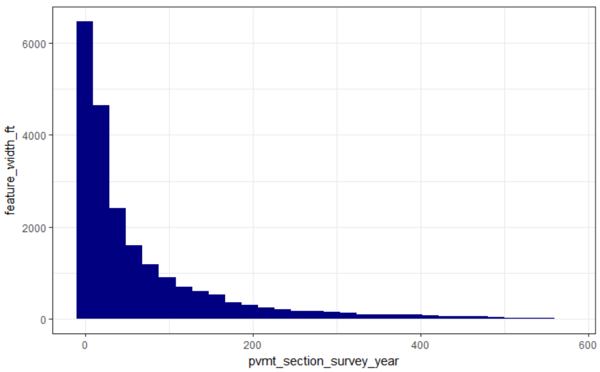
Simpson US-49	22	0
Toyota Rd MS-780	11	0



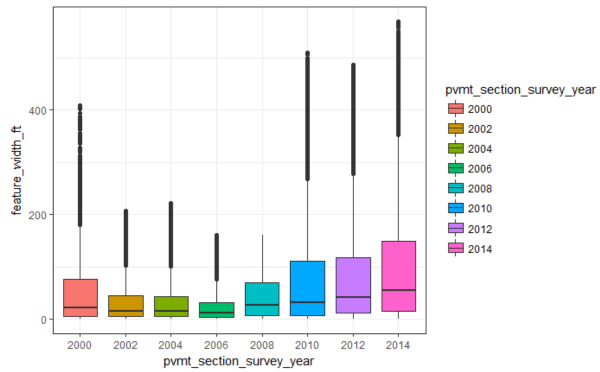
(a) Histogram at Low Severity Level



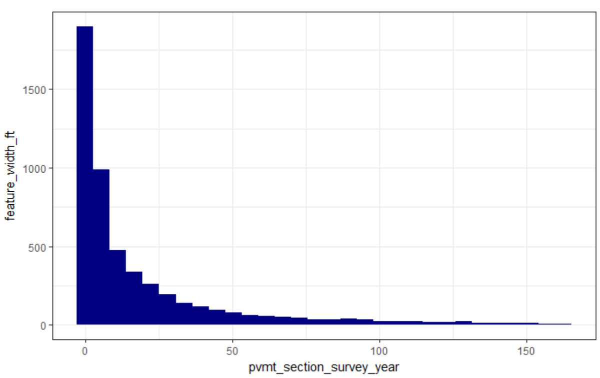
(b) Boxplot at Low Severity Level



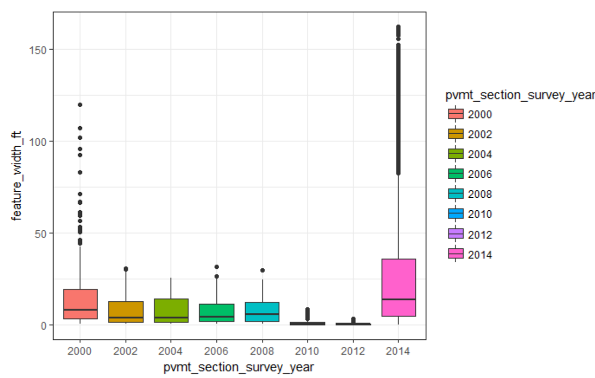
(c) Histogram at Medium Severity Level



(d) Boxplot at Medium Severity Level



(e) Histogram at High Severity Level



(f) Boxplot at High Severity Level

Figure 11 Histograms and Boxplots of Transverse Cracking in Non-warranty Asphalt Pavements

(3) Alligator Cracking in Asphalt Pavements

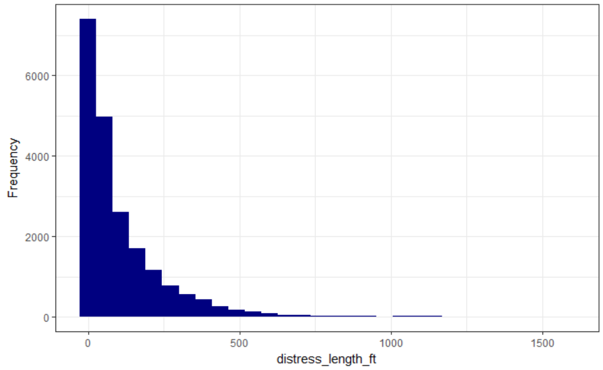
Table 21 lists the numbers of road sections in alligator cracking in non-warranty asphalt pavements at three severity levels over service years. Most sections containing alligator cracking are at low and medium severity levels. The low severity level cracking is of any forms of fine longitudinal cracks with few interconnected cracks found in the wheel paths. The medium severity level cracking consists of wider longitudinal cracks with interconnected cracks found in the wheel paths. This cracking also starts to show the alligator pattern and other deteriorations, and the cracking depths become apparent. The statistical descriptive figures presented in Figure 12 indicate that the measurement data at the three severity levels are all right-skewed. The numbers of warranty projects to have alligator cracking occurring with maintenance decisions are listed in Table 22. It shows that most of the transverse cracking appearing in warranty projects does not exceed the deduct point threshold.

Table 21 Alligator Cracking Datasets in Non-warranty Asphalt Pavements

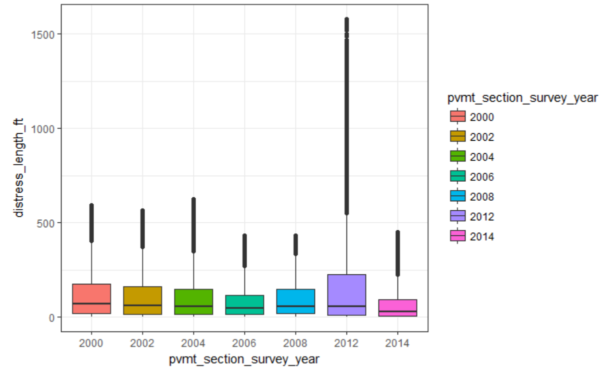
Section Survey Year	Severity Level		
	Low (0)	Medium (1)	High (2)
2000	3338	1273	154
2002	3183	1141	112
2004	3157	1289	42
2006	3568	1627	663
2008	2857	2403	48
2010	2664	2399	33
2012	3452	3175	51
2014	2376	1322	593

Table 22 Alligator Cracking Datasets in Warranty Asphalt Pavements

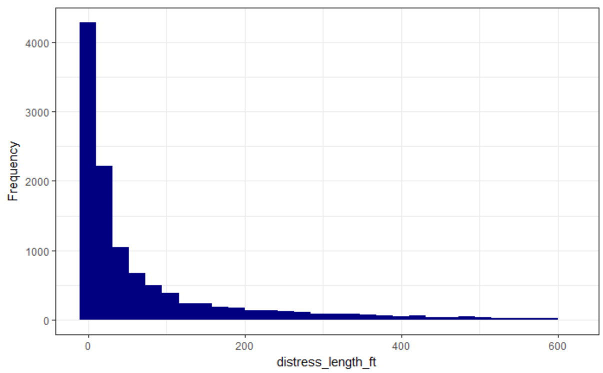
Warranty Project	Remedial Action	
	NO (0)	YES (1)
Bolivar US61	76	12
Carroll/Montgomery I-55	2	0
Covington US-84	85	23
Desoto MS-302 (US-51 to Swinnea Rd)	13	4
Grenada/Yalobusha I-55	4	0
Lauderdale MS-19 SB	9	4
Lincoln I-55	3	0
Montgomery US-82	27	20
Montgomery/Carroll/Grenada I-55	9	2
Pontotoc MS-9	3	0



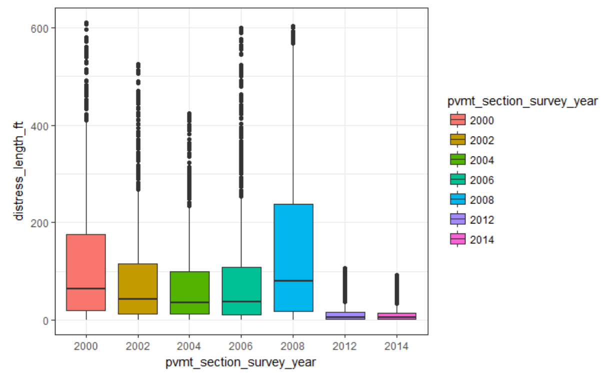
(a) Histogram at Low Severity Level



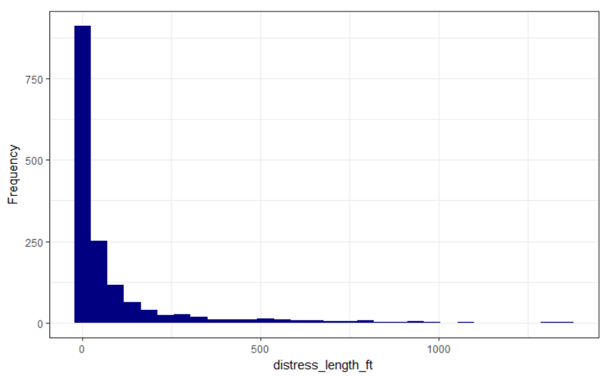
(b) Boxplot at Low Severity Level



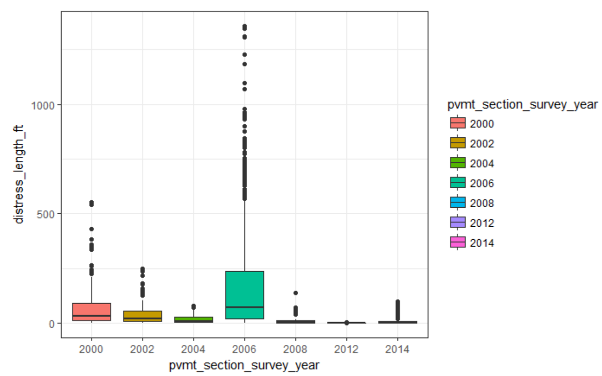
(c) Histogram at Medium Severity Level



(d) Boxplot at Medium Severity Level



(e) Histogram at High Severity Level

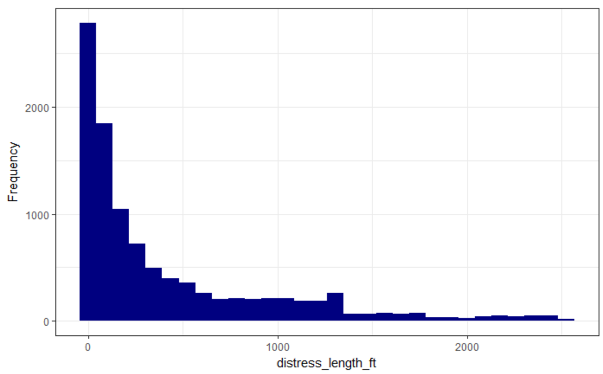


(f) Boxplot at High Severity Level

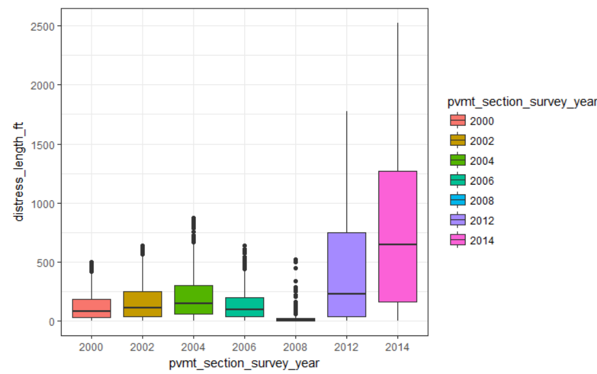
Figure 12 Histograms and Boxplots of Alligator Cracking in Non-warranty Asphalt Pavements

(4) Block Cracking in Asphalt Pavements

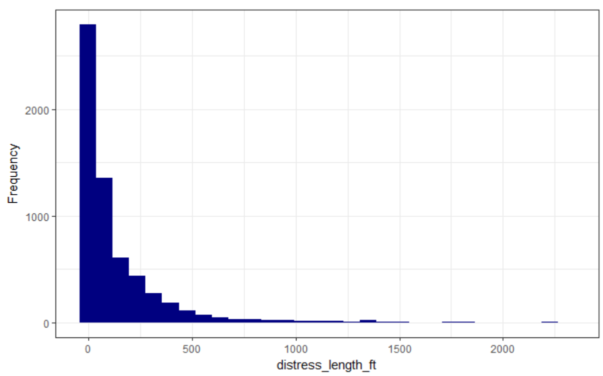
Table 23 lists the numbers of road sections with block cracking at three severity levels over service years in the non-warranty asphalt pavements. Most sections containing block cracking are at low and medium severity levels. The statistical descriptive figures presented in Figure 13 indicate that the measurement data at the three severity levels are all right-skewed. The numbers of warranty projects with block cracking occurring with maintenance decisions are listed in Table 24. It shows that about half the warranty projects with block cracking exceed the deduct point threshold, but the dataset size of the block cracking in the warranty projects seems not large enough for a reliable statistical analysis. Block cracking is no longer quantified in the network-level surveys done every two years due to difficulty of automated distress detection technology's ability to discern the block pattern.



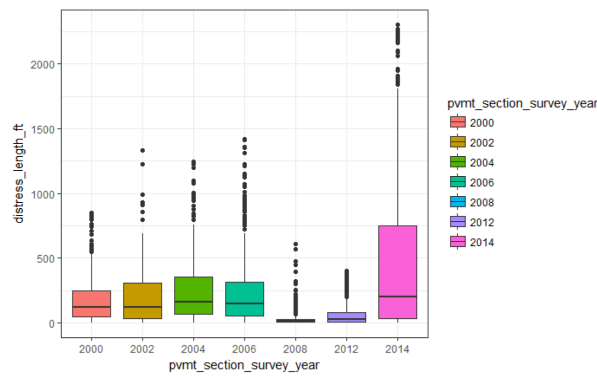
(a) Histogram at Low Severity Level



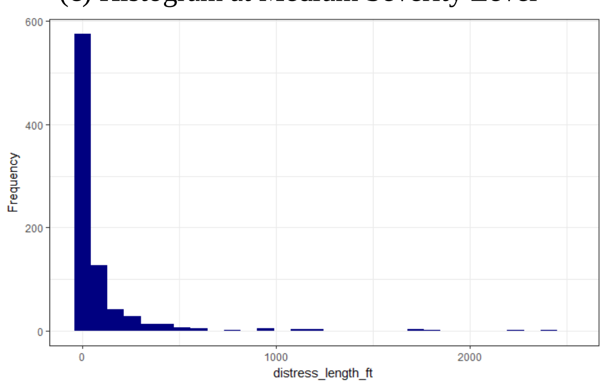
(b) Boxplot at Low Severity Level



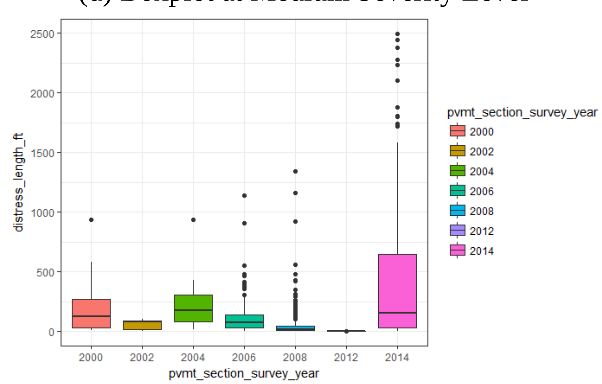
(c) Histogram at Medium Severity Level



(d) Boxplot at Medium Severity Level



(e) Histogram at High Severity Level



(f) Boxplot at High Severity Level

Figure 13 Histograms and Boxplots of Block Cracking in Non-warranty Asphalt Pavements

Table 23 Block Cracking Datasets in Non-warranty Asphalt Pavements

Section Survey Year	Severity Level		
	Low (0)	Medium (1)	High (2)
2000	591	415	53
2002	875	230	11
2004	1074	441	13
2006	1146	954	102
2008	817	570	431
2010	3253	3114	196
2012	3306	3231	170
2014	3079	661	97

Table 24 Block Cracking Datasets in Warranty Asphalt Pavements

Warranty Project	Remedial Action	
	NO (0)	YES (1)
Bolivar US61	58	48
Covington US-84	40	89
Lincoln I-55	3	0
Montgomery US-82	6	4
Montgomery/Carroll/Grenada I-55	1	0

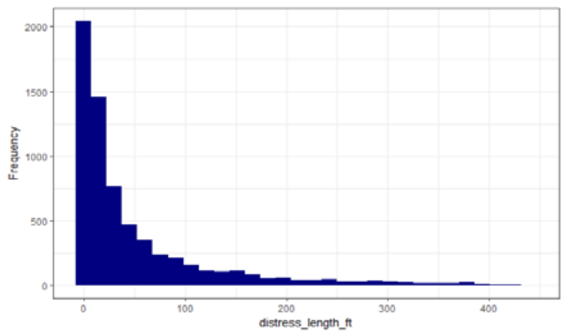
(5) Edge Cracking in Asphalt Pavements

Table 25 lists the numbers of road sections with edge cracking at three severity levels in the non-warranty asphalt pavements over service years. Out of the three severity levels, the low severity level has the highest number of sections. The histogram figures presented in Figure 14 indicates that the measurement data at all the three severity levels are right-skewed. There is no edge cracking data occurring on the warranty projects.

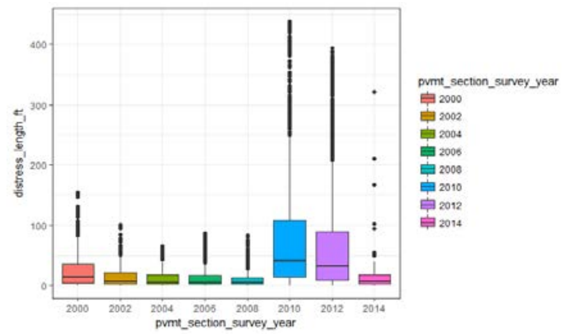
Table 25 Edge Cracking Datasets in Non-warranty Asphalt Pavements

Section Survey Year	Severity Level		
	Low (0)	Medium (1)	High (2)
2000	630	229	49
2002	315	56	1
2004	289	100	10

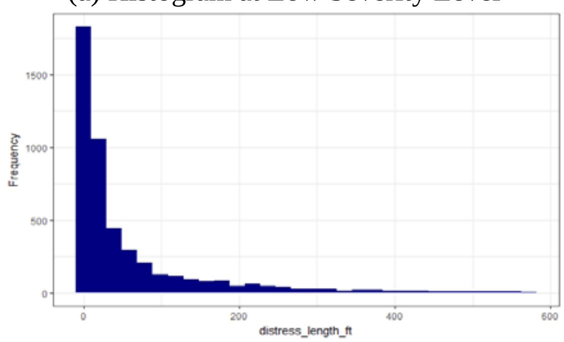
2006	564	173	32
2008	508	661	111
2010	1502	780	823
2012	3112	3109	981
2014	103	30	12



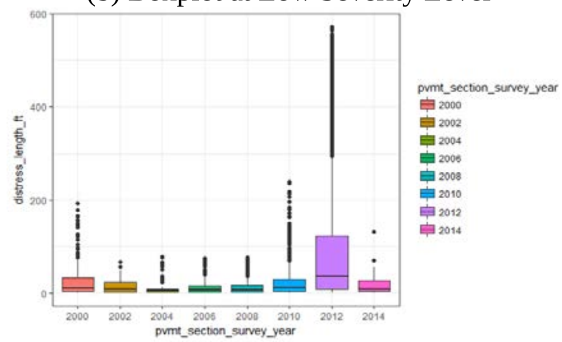
(a) Histogram at Low Severity Level



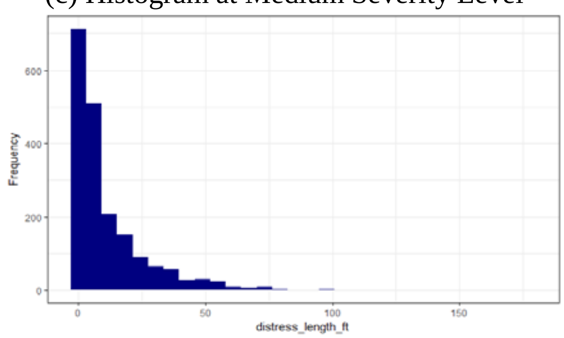
(b) Boxplot at Low Severity Level



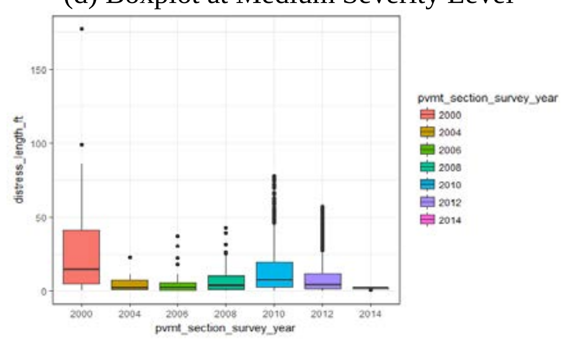
(c) Histogram at Medium Severity Level



(d) Boxplot at Medium Severity Level



(e) Histogram at High Severity Level



(f) Boxplot at High Severity Level

Figure 14 Histograms and Boxplots of Edge Cracking in Non-warranty Asphalt Pavements

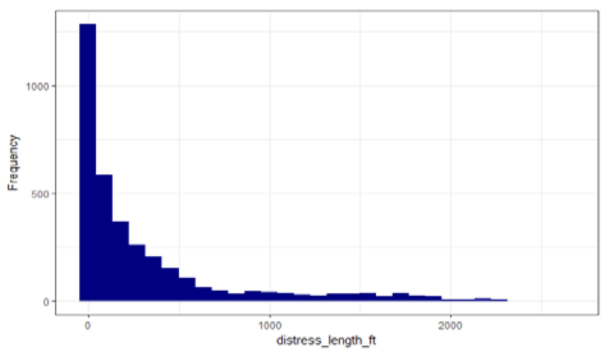
(6) Bleeding in Asphalt Pavements

Table 26 Bleeding Datasets in Non-warranty Asphalt Pavements

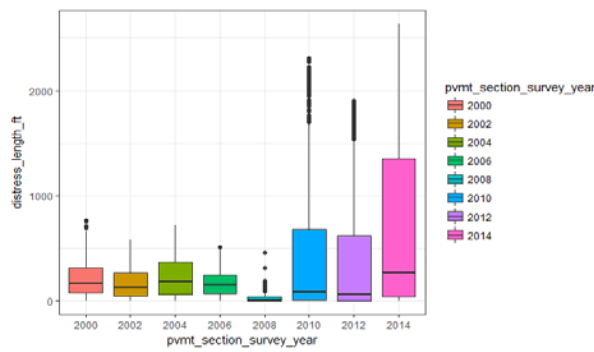
Section Survey Year	Severity Level		
	Low (0)	Medium (1)	High (2)
2000	546	123	16
2002	297	31	2
2004	370	52	2

2006	230	63	11
2008	123	50	2
2010	811	517	333
2012	1338	1213	474
2014	35	6	1

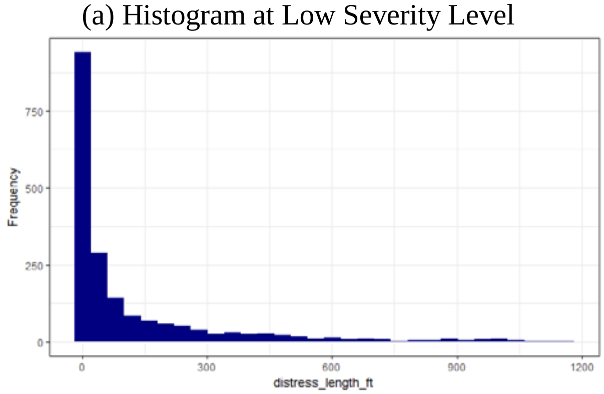
Table 26 lists the numbers of road sections with bleeding at three severity levels over service years in the non-warranty asphalt pavements. The low severity level has the highestnumber of sections. The histogram figures presented in Figure 15 indicate that the measurement data at the three severity levels are all right-skewed. There is no bleeding data or occurring for the warranty projects.



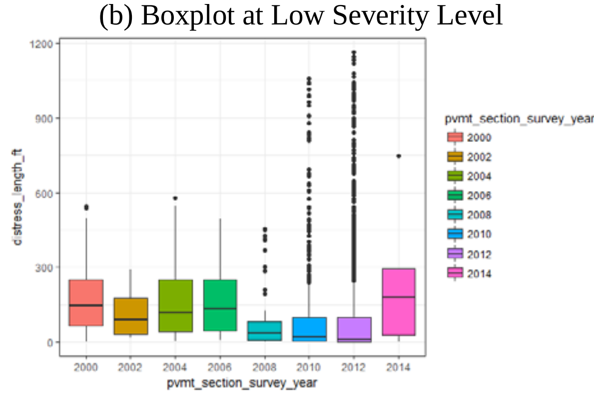
(a) Histogram at Low Severity Level



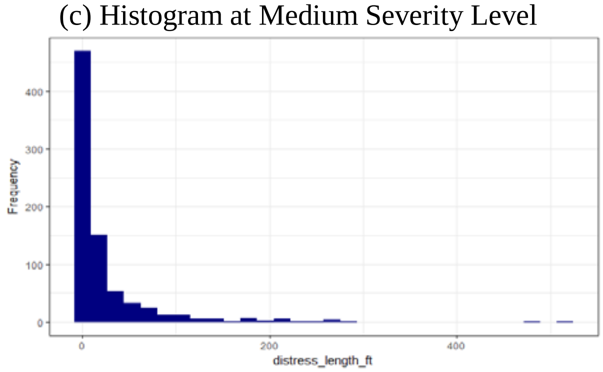
(b) Boxplot at Low Severity Level



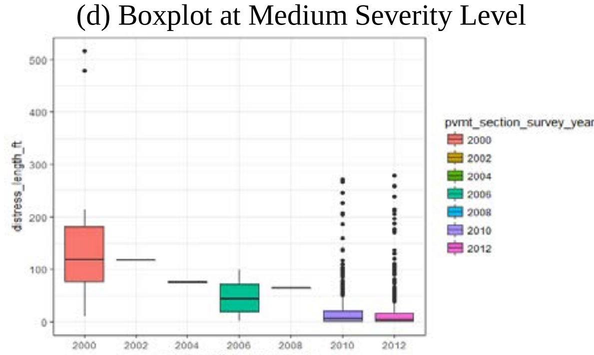
(c) Histogram at Medium Severity Level



(d) Boxplot at Medium Severity Level



(e) Histogram at High Severity Level

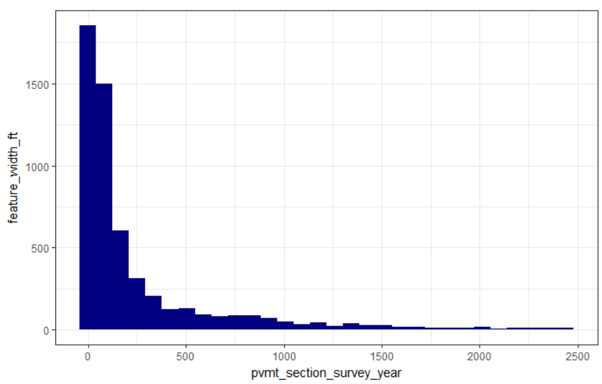


(f) Boxplot at High Severity Level

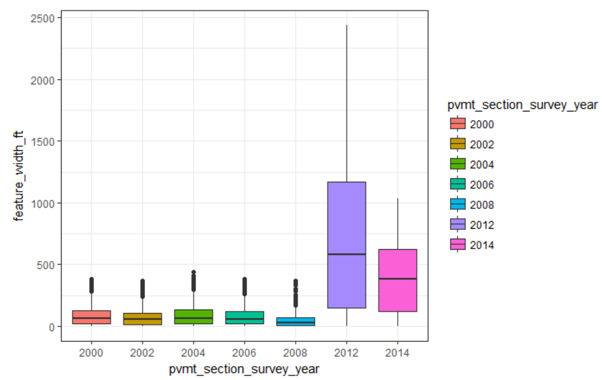
Figure 15 Histograms and Boxplots of Bleeding in Non-warranty Asphalt Pavements

(7) Reflection Cracking in Asphalt Pavements

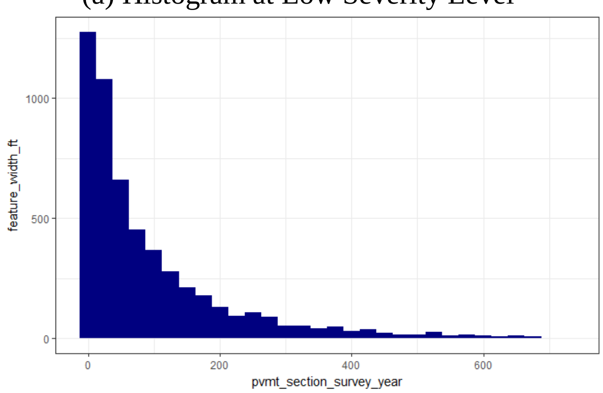
Table 27 lists the numbers of road sections with reflection cracking at three severity levels in the non-warranty asphalt pavements over their service years. Out of the three severity levels the low and medium severity levels have the highest number of sections. The statistical descriptive figures presented in Figure 16 indicate that the measurement data at the three severity levels are all right-skewed. There is no reflection cracking data or occurring for the warranty projects. Reflection cracking is no longer quantified during the every-two-year network-level survey due to the difficulty of automated data collection's ability to distinguish reflective cracks from transverse cracks. Indeed, it was also difficult for human manual raters to distinguish reflective cracking on a video screen.



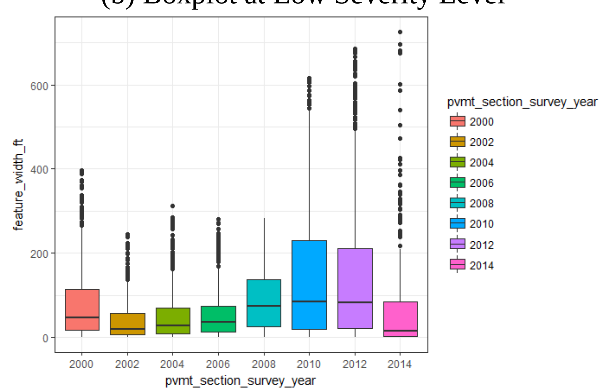
(a) Histogram at Low Severity Level



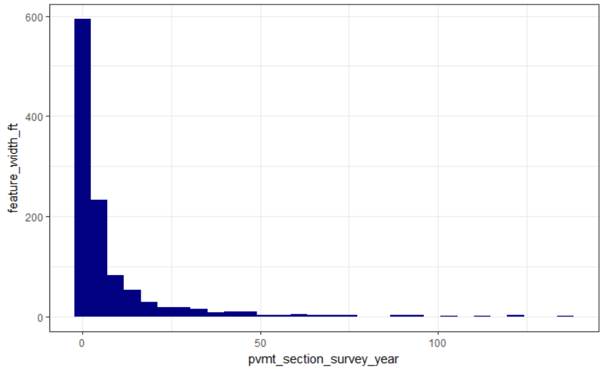
(b) Boxplot at Low Severity Level



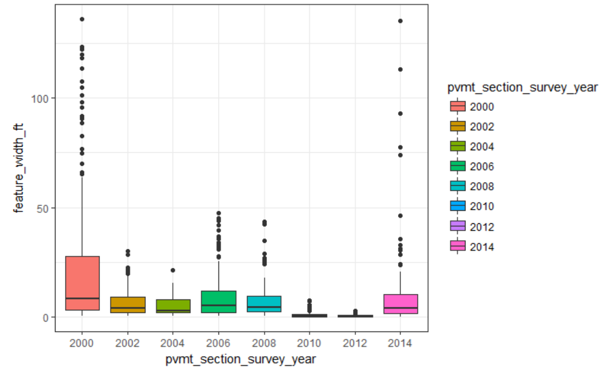
(c) Histogram at Medium Severity Level



(d) Boxplot at Medium Severity Level



(e) Histogram at High Severity Level



(f) Boxplot at High Severity Level

Figure 16 Histograms and Boxplots of Reflection Cracking in Non-warranty Asphalt Pavements

Table 27 Reflection Cracking Datasets in Non-warranty Asphalt Pavements

Section Survey Year	Severity Level		
	Low (0)	Medium (1)	High (2)
2000	807	635	253
2002	747	528	66
2004	896	696	36
2006	782	620	179
2008	676	817	80
2010	0	1011	222
2012	1172	1093	246
2014	734	310	103

4.4.3 Distresses in Concrete Pavements

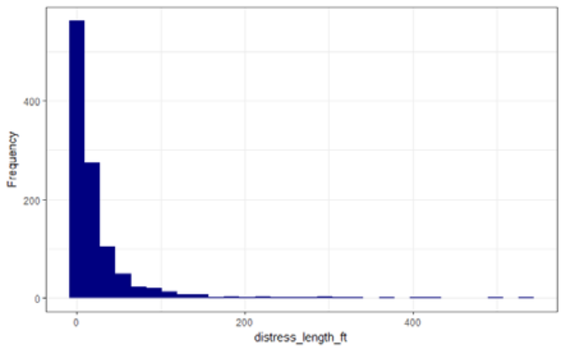
After data screening, there were valid data of longitudinal cracking, transverse cracking, spalling of longitudinal joint, and spalling of transverse joint in the non-warranty concrete pavements to perform a statistical analysis.

(1) Longitudinal Cracking in Concrete Pavements

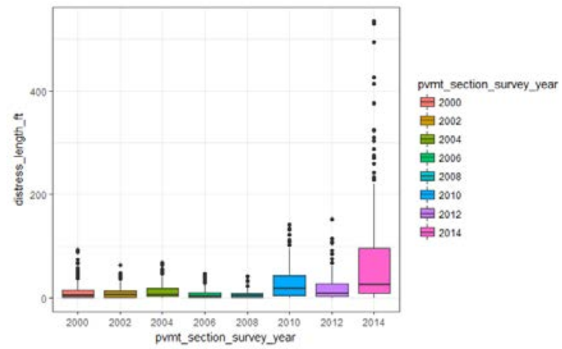
Table 28 lists the numbers of road sections with longitudinal cracking at three severity levels over service years in the non-warranty concrete pavements. Out of the three severity levels, the low severity level has the highest numbers of sections. The statistical descriptive figures presented in Figure 17 indicate that the measurement data at the three severity levels are right-skewed, and the measurement values are smaller than the measurements in the asphalt pavements. There has been some longitudinal cracking occurring in the warranty projects with maintenance decisions, but the data size was too small to be included in a statistical study.

Table 28 Longitudinal Cracking Datasets in Non-warranty Concrete Pavements

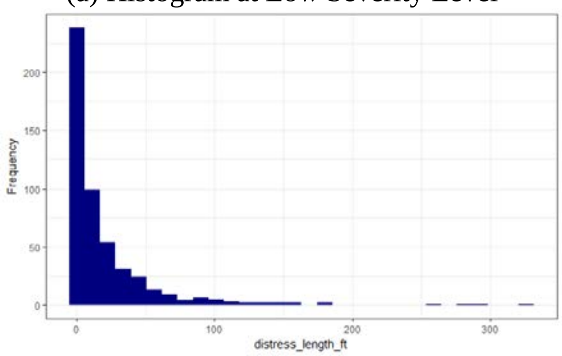
Section Survey Year	Severity Level		
	Low (0)	Medium (1)	High (2)
2000	204	39	2
2002	143	18	2
2004	130	24	0
2006	125	30	6
2008	63	52	8
2010	168	113	49
2012	136	127	40
2014	179	131	100



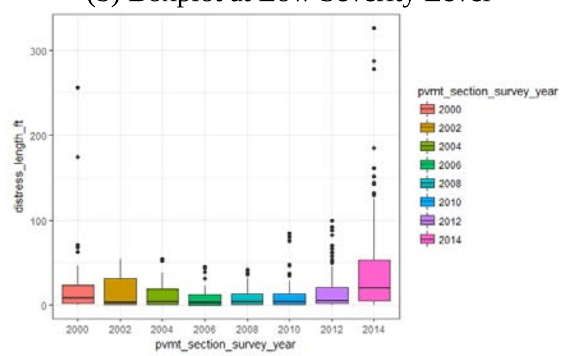
(a) Histogram at Low Severity Level



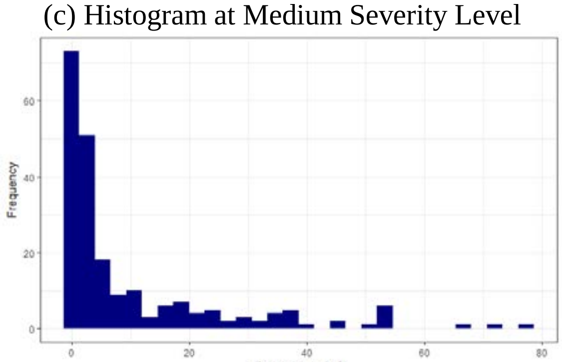
(b) Boxplot at Low Severity Level



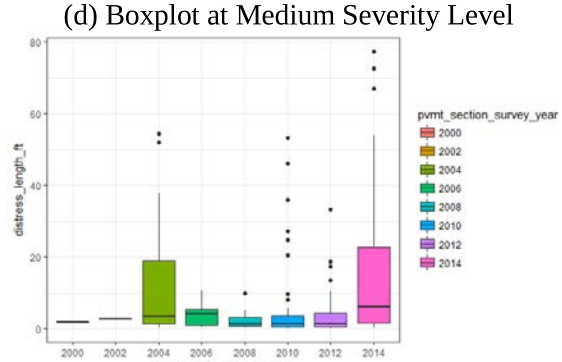
(c) Histogram at Medium Severity Level



(d) Boxplot at Medium Severity Level



(e) Histogram at High Severity Level

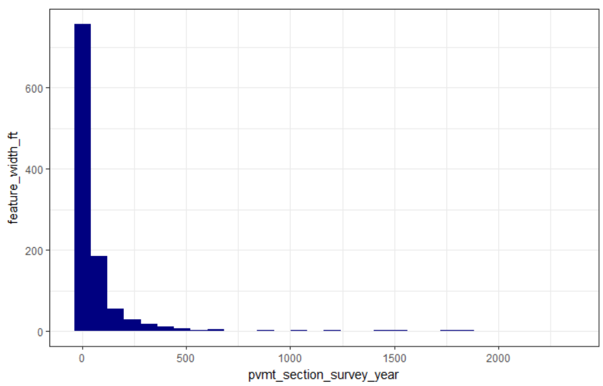


(f) Boxplot at High Severity Level

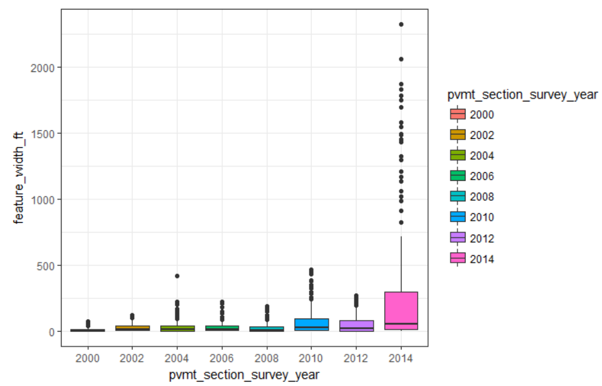
Figure 17 Histograms and Boxplots of Longitudinal Cracking in Non-warranty Concrete

(2) Transverse Cracking in Concrete Pavements

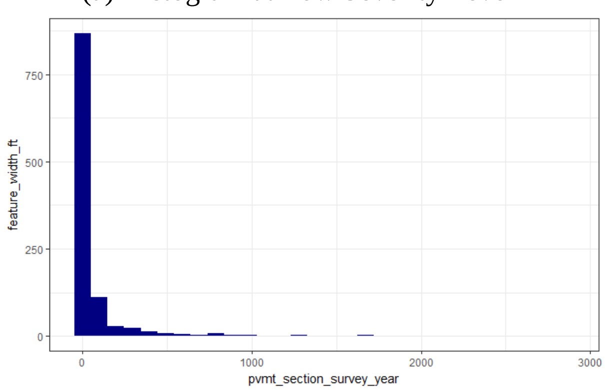
Table 29 lists the numbers of road sections with transverse cracking at three severity levels over service years in the non-warranty concrete pavements. The low and medium severity levels have the highest number of sections. The statistical descriptive figures presented in Figure 18 indicate that the measurement data at the three severity levels are right-skewed, and the measurement values are smaller than the measurements in the asphalt pavements. Although there has been some transverse cracking occurring in the warranty projects with maintenance decisions, the data size was too small to be included in the study.



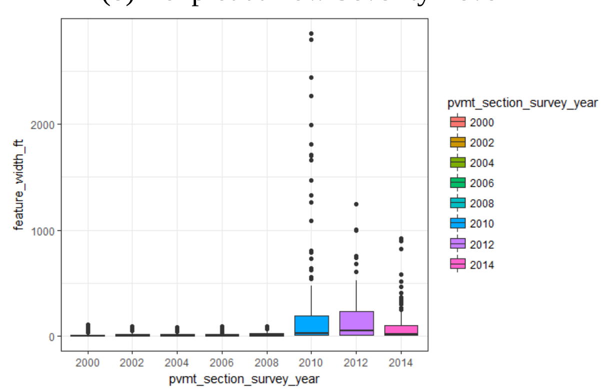
(a) Histogram at Low Severity Level



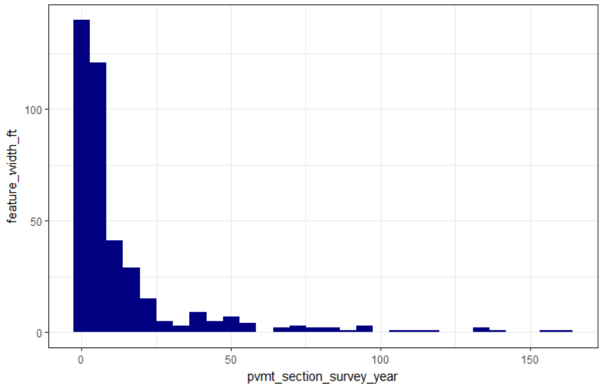
(b) Boxplot at Low Severity Level



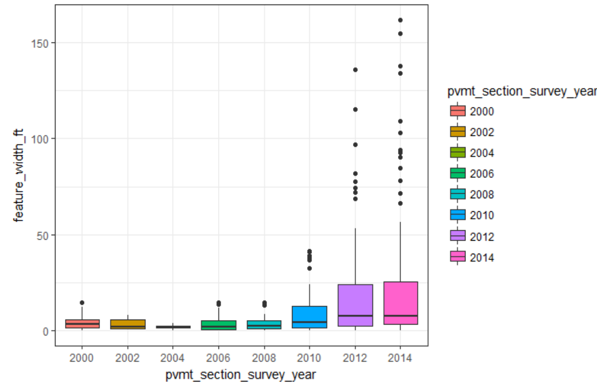
(c) Histogram at Medium Severity Level



(d) Boxplot at Medium Severity Level



(e) Histogram at High Severity Level



(f) Boxplot at High Severity Level

Figure 18 Histograms and Boxplots of Transverse Cracking in Non-warranty Concrete Pavements

Table 29 Transverse Cracking Datasets in Non-warranty Concrete Pavements

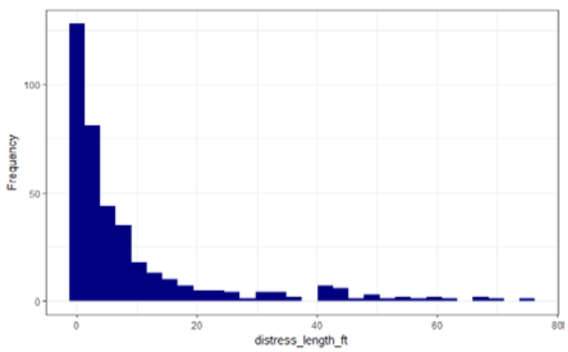
Section Survey Year	Severity Level		
	Low (0)	Medium (1)	High (2)
2000	248	292	57
2002	125	130	8
2004	133	103	5
2006	125	109	23
2008	86	141	63
2010	131	146	88
2012	111	137	84
2014	196	147	125

(3) Spalling of Longitudinal Joint in Concrete Pavements

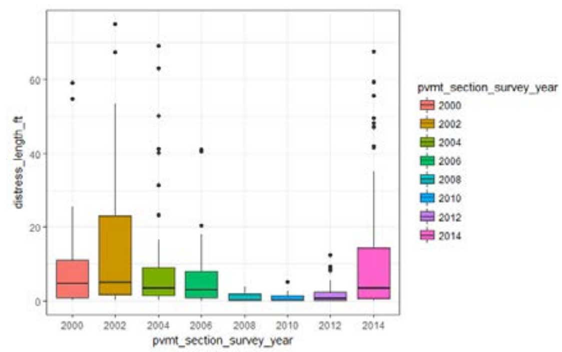
Table 30 lists the numbers of road sections with spalling of longitudinal joint at three severity levels over service years in the non-warranty concrete pavements. Out of the three severity levels, the low severity level has the highest number of sections. The statistical descriptive figures presented in Figure 19 indicate that the measurement data at the three severity levels are all right-skewed. There has been no data or occurring for spalling of longitudinal joint in the warranty projects.

Table 30 Spalling of Longitudinal Joint Datasets in Non-warranty Concrete Pavements

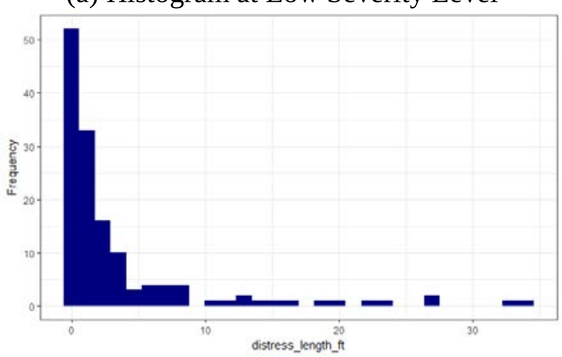
Section Survey Year	Severity Level		
	Low (0)	Medium (1)	High (2)
2000	61	11	4
2002	69	3	0
2004	63	12	1
2006	76	19	4
2008	3	9	2
2010	9	13	13
2012	41	61	42
2014	92	24	13



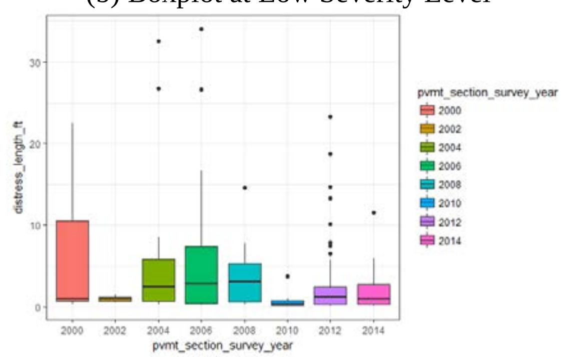
(a) Histogram at Low Severity Level



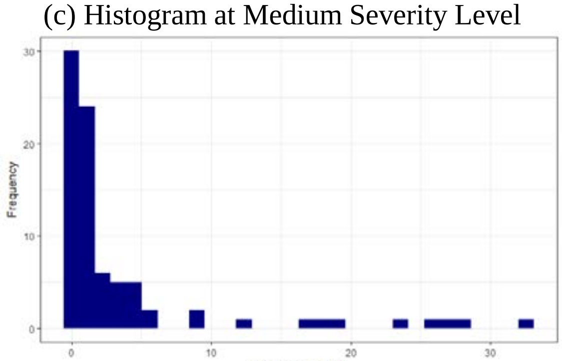
(b) Boxplot at Low Severity Level



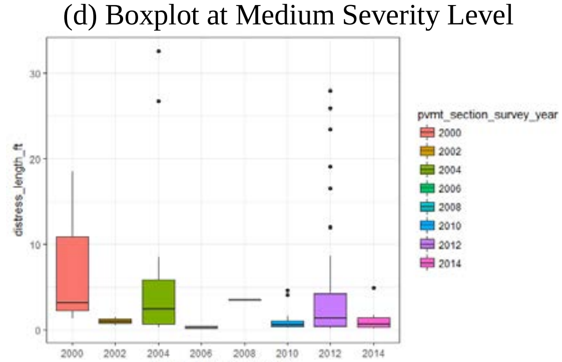
(c) Histogram at Medium Severity Level



(d) Boxplot at Medium Severity Level



(e) Histogram at High Severity Level



(f) Boxplot at High Severity Level

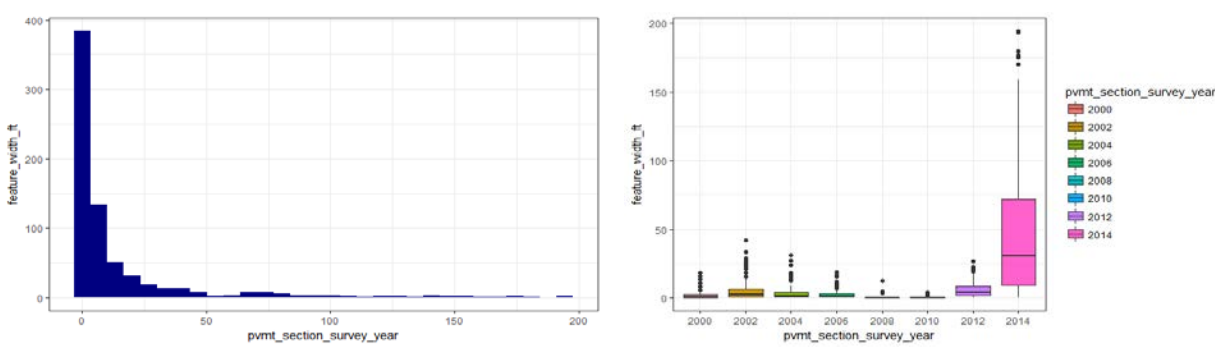
Figure 19 Histograms and Boxplots of Spalling of Longitudinal Joint in Non-warranty Concrete Pavements

(4) Spalling of Transverse Joint in Concrete Pavements

Table 31 lists the numbers of road sections with spalling of transverse joint at three severity levels over service years in the non-warranty concrete pavements. Out of the three severity levels, the low severity level has the highest number of sections with spalling of transverse joint. The statistical descriptive figures presented in Figure 20 indicate that the measurement data at the three severity levels are all right-skewed. There has been no data or record of occurring for spalling of transverse joint in the warranty projects.

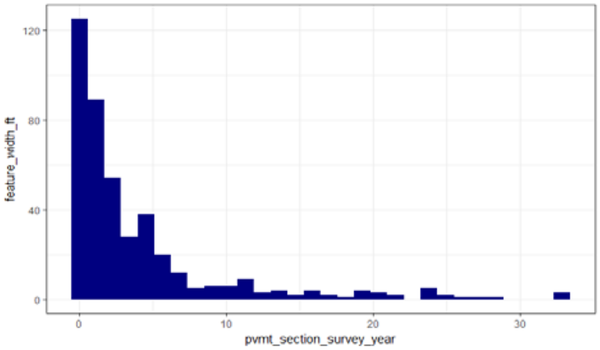
Table 31 Spalling of Transverse Joint Datasets in Non-warranty Concrete Pavements

Section Survey Year	Severity Level		
	Low (0)	Medium (1)	High (2)
2000	86	23	7
2002	88	16	6
2004	80	28	3
2006	73	27	13
2008	18	38	45
2010	78	61	55
2012	141	125	86
2014	191	139	95

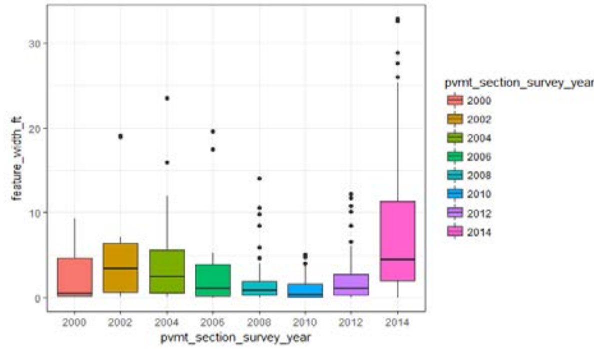


(a) Histogram at Low Severity Level

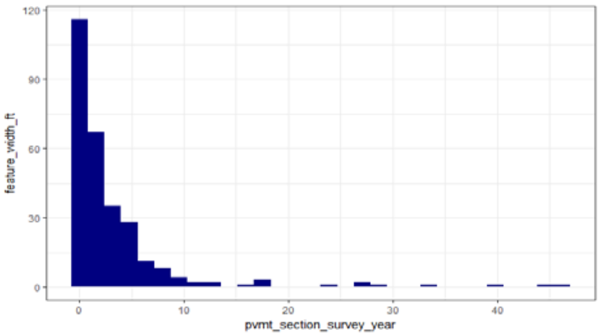
(b) Boxplot at Low Severity Level



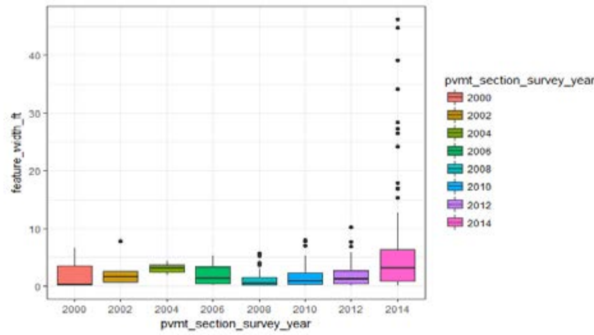
(c) Histogram at Medium Severity Level



(d) Boxplot at Medium Severity Level



(e) Histogram at High Severity Level



(f) Boxplot at High Severity Level

Figure 20 Histograms and Boxplots of Spalling of Transverse Joint in Non-warranty Concrete Pavements

4.4.4 Rutting and IRI

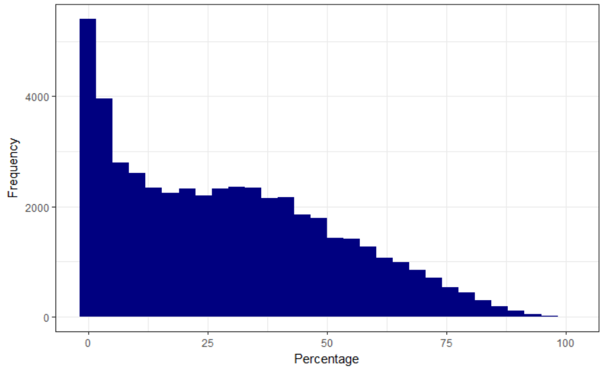
After data screening, usable valid data for rutting depth in percentage, average rutting depth, IRI in percentage, and average IRI in both asphalt and concrete pavements were available for statistical analysis. Table 12 and Table 13 list the numbers of road sections recorded in the PMS to present network pavement condition for rutting and IRI in Mississippi.

(1) Rutting Depth

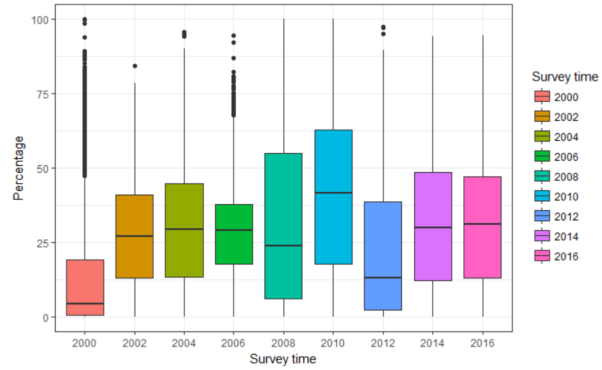
Different from the distress data, the rutting data contains average rut depth values and percentages falling in categories of low, medium, and high severities. The classification rule of the category is presented in "Data Sources" Section. For the condition data, each selected section is required to collect the average values and percentages for rutting, IRI, and faulting. The statistical descriptive figures for rutting percentages in asphalt non-warranty pavements are presented in Figure 21, which indicate that the percentages at three categories are right-skewed, and the percentages at the high category are the smallest. The rutting percentage records as distress in the warranty projects with maintenance decisions are listed in Table 32. It shows that most of the rutting depths in the warranty projects did not exceed the deduct point threshold.

Table 32 Rutting Datasets in Warranty Asphalt Pavements

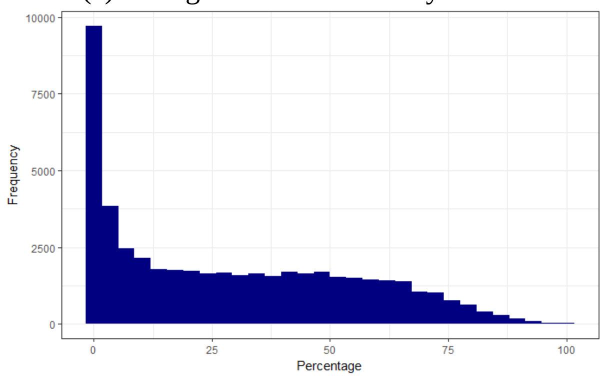
Warranty Project	Remedial Action	
	NO (0)	YES (1)
Bolivar US61	37	3
Carroll/Montgomery I-55	26	0
Covington US-84	71	6
Desoto MS-302 (US-51 to Swinnea Rd)	53	30
Grenada/Yalobusha I-55	303	15
Lauderdale MS-19 NB	19	0
Lincoln I-55	16	0
Montgomery US-82	13	2
Montgomery/Carroll/Grenada I-55	31	0
NewtonNeshoba MS-19	22	0
Pontotoc MS-9	30	0
Toyota Rd MS-780	38	0



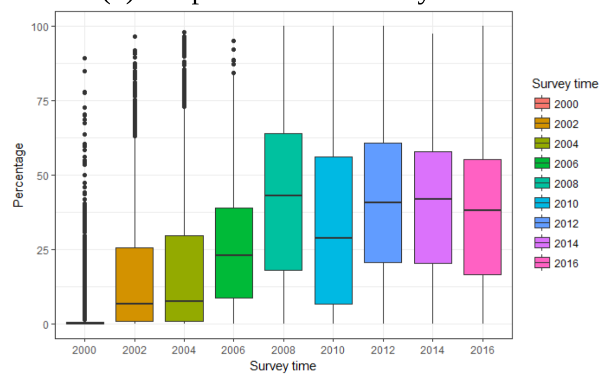
(a) Histogram at Low Severity Level



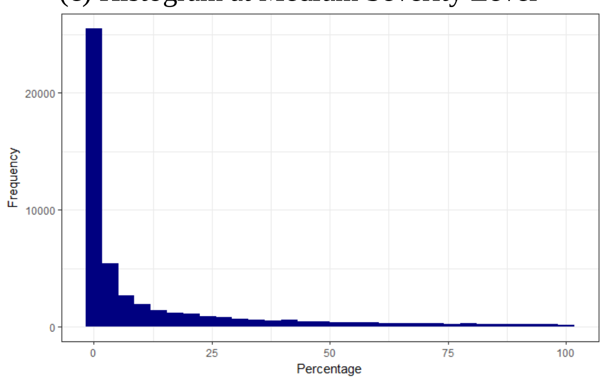
(b) Boxplot at Low Severity Level



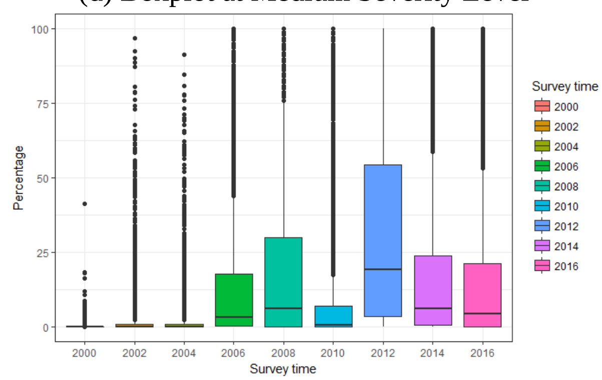
(c) Histogram at Medium Severity Level



(d) Boxplot at Medium Severity Level



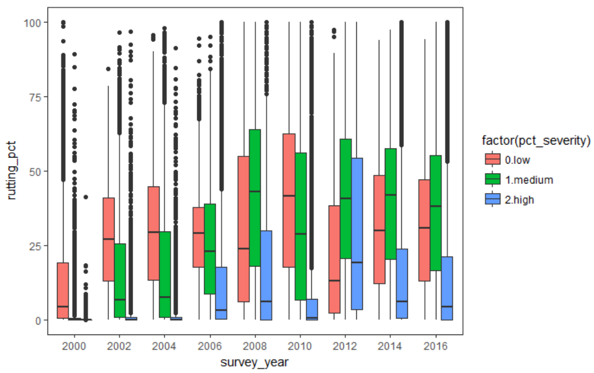
(e) Histogram at High Severity Level



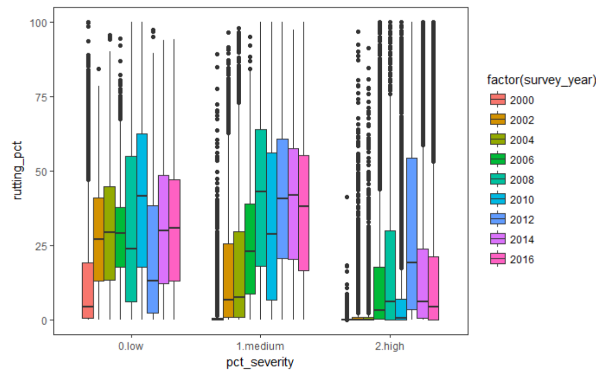
(f) Boxplot at High Severity Level

Figure 21 Histograms and Boxplots of Rutting in Non-warranty Asphalt Pavements

The boxplots of the rutting data for the asphalt pavements ordered by severity categories and service years are presented in Figure 22. Figure 22 (a) shows that the percentage of rutting for asphalt pavement at categories of low, medium, and high was apparently low before the service year 2008, and the main category at that time was low, which means majority of the rutting depth in the network was lower than 1/8 inch. However, after that time, the percentage of all 3 categories increased noticeably, and the main category turned to medium (except service year 2010). It indicates that the rutting depth in the network became serious, with larger percentage of the rutting depth getting larger than 1/8 inch. Each category over service years is compared in Figure 22 (b), and the results show that most of the rutting occurring in Mississippi's network is in categories of low and medium. The percentage of rutting in low and medium increases over service time; moreover, the percentage in medium category with rutting depth larger than 1/8 inch but smaller than 1/4 inch increases more rapidly.



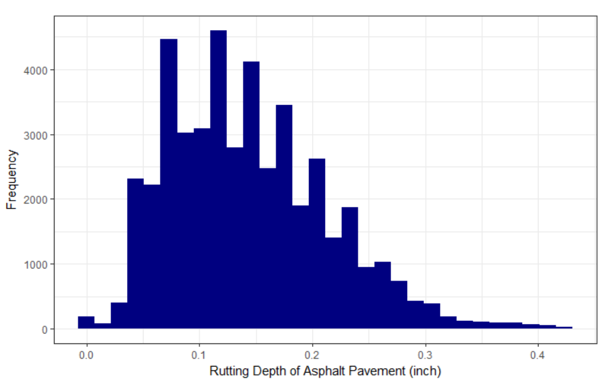
(a) Percentage of Rutting Depth in Categories



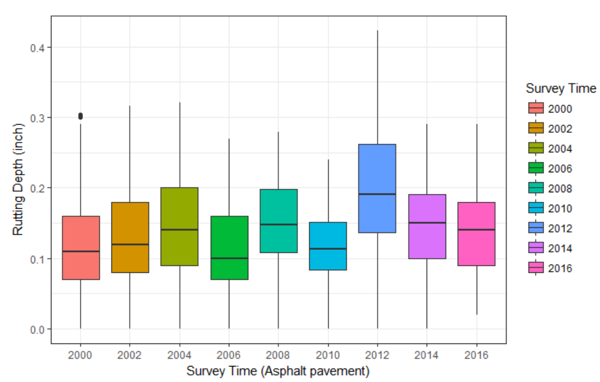
(b) Percentage of Rutting Depth in Different Years

Figure 22 Boxplots of Rutting Depth Percentages in Non-warranty Asphalt Pavement

The average rutting depth is another pavement condition indicator recorded in PMS for asphalt pavements, and the statistical description of the indicator is presented in Figure 23. Figure 23 (a) indicate that the distributions of the average rutting depths are right-skewed in asphalt pavements. Figure 23 (b) show that the average rutting depth increases slightly over time, except for service year 2012.



(a) Histogram of Average Rutting Depth



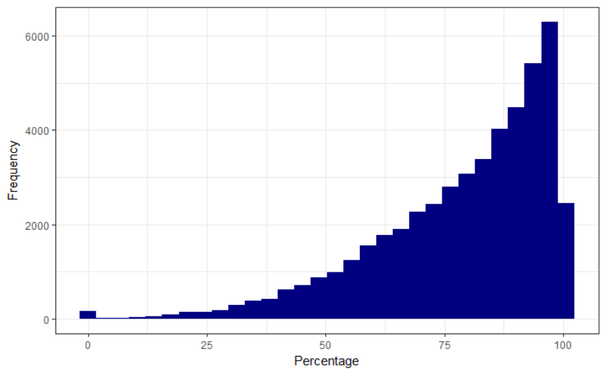
(b) Boxplot Average Rutting Depth

Figure 23 Histograms and Boxplots of Rutting Depth in Non-warranty Asphalt Pavements

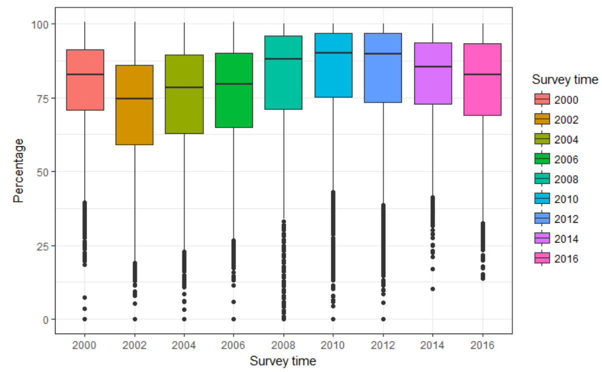
(2) Average IRI

The IRI data included average values and percentages in severity categories of low, medium, and high in non-warranty pavements. The classification rule of the category is presented in "Data

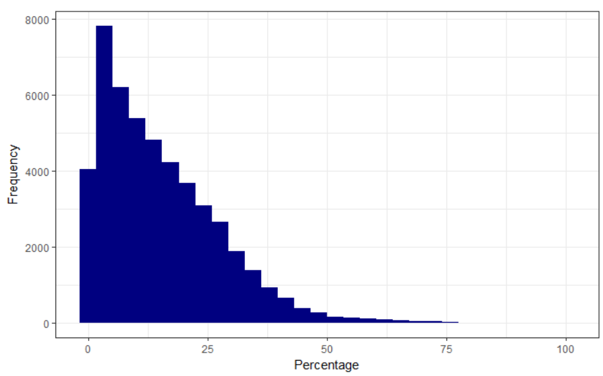
Sources” Section. In the warranty projects, the IRI was not considered as a warranty item in contracting but recorded as condition data (only average IRI) in every surveyed segment. The statistical descriptive figures for IRI percentages in asphalt and concrete non-warranty pavements are presented in Figure 24 and Figure 25, which indicate that the percentages of IRI in the low severity category are left-skewed for asphalt and concrete pavements, while in medium and high categories right-skewed. Compared with the percentage in low and medium categories, the percentage values in the high category are quite small for both the asphalt and concrete pavements.



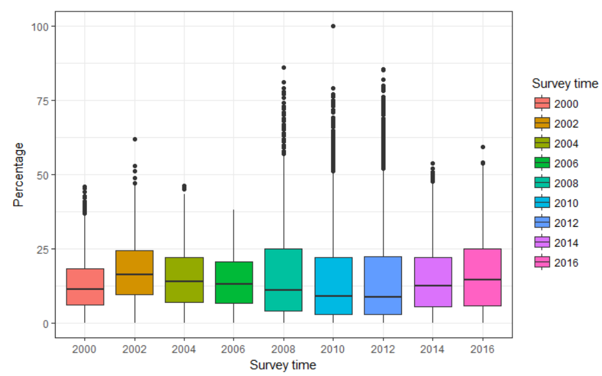
(a) Histogram at Low Severity Level



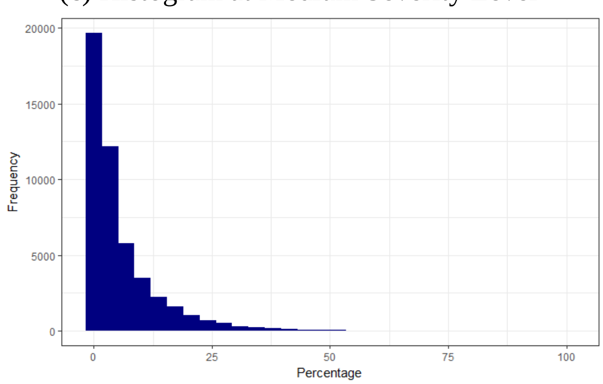
(b) Boxplot at Low Severity Level



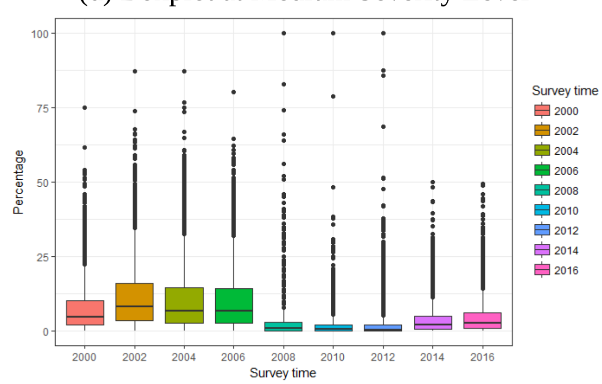
(c) Histogram at Medium Severity Level



(d) Boxplot at Medium Severity Level

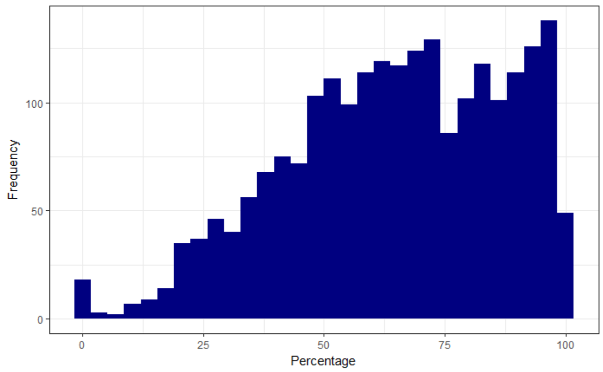


(e) Histogram at High Severity Level

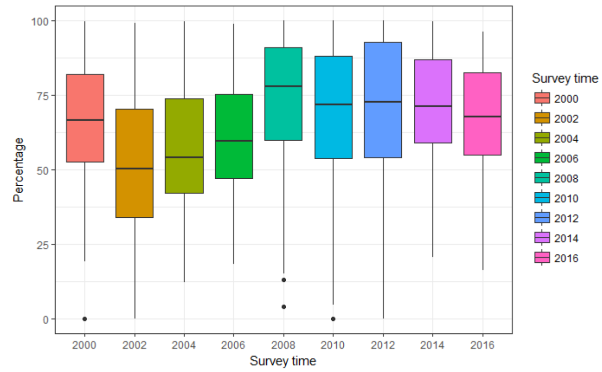


(f) Boxplot at High Severity Level

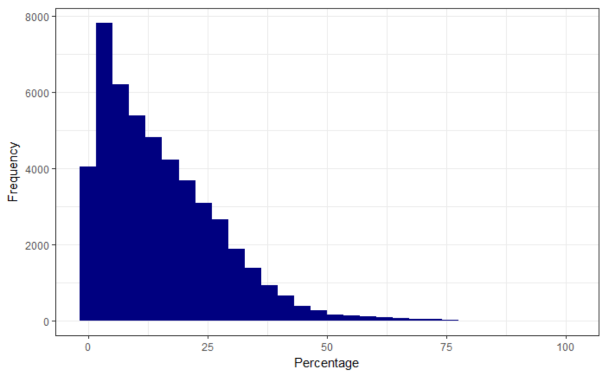
Figure 24 Histograms and Boxplots of IRI in Non-warranty Asphalt Pavements



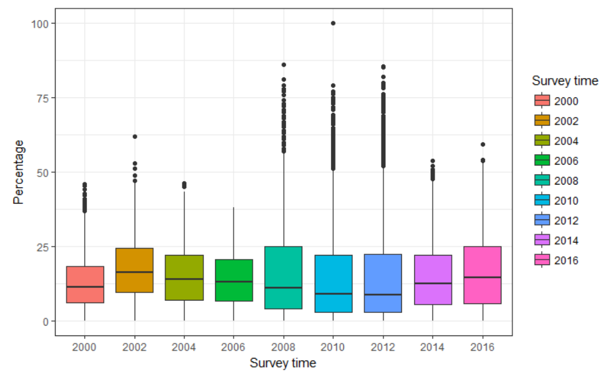
(a) Histogram at Low Severity Level



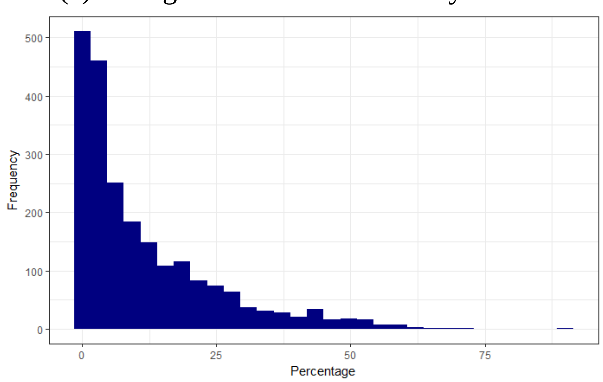
(b) Boxplot at Low Severity Level



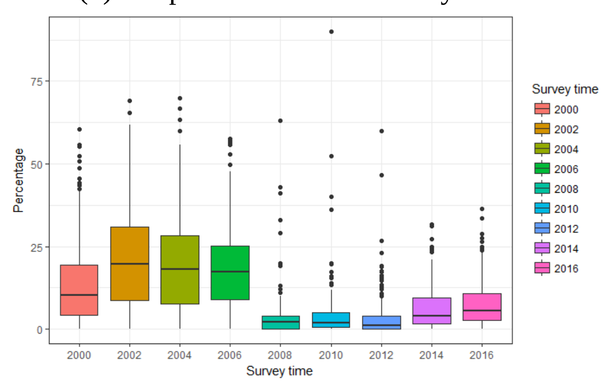
(c) Histogram at Medium Severity Level



(d) Boxplot at Medium Severity Level



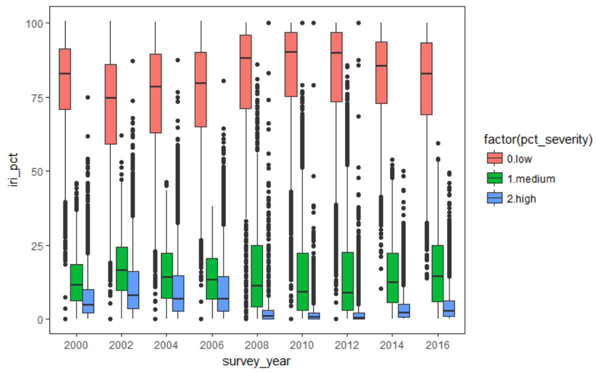
(e) Histogram at High Severity Level



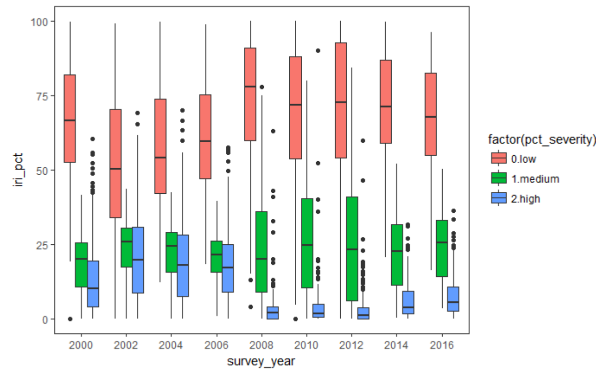
(f) Boxplot at High Severity Level

Figure 25 Histograms and Boxplots of IRI in Non-warranty Concrete Pavements

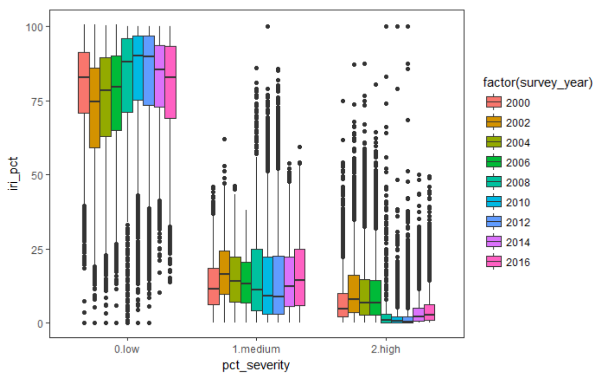
The boxplots of IRI for the asphalt and concrete pavements ordered by the three severity categories and service years are presented in Figure 26. It is obvious that the average IRI in the low category consists of the majority average IRI in the network over all the service years in both asphalt and concrete pavements, which means the average IRI is mostly lower than 150 in/mi. Figure 26 (c) and (d) show that the average IRI in the three severity categories increases slightly over the service time in the network, and there is only one rapid rise appearing after service year 2008 in concrete pavements. This may be due to 2010 being the first 100% automated survey as opposed to having used two 500-ft samples per mile in previous surveys.



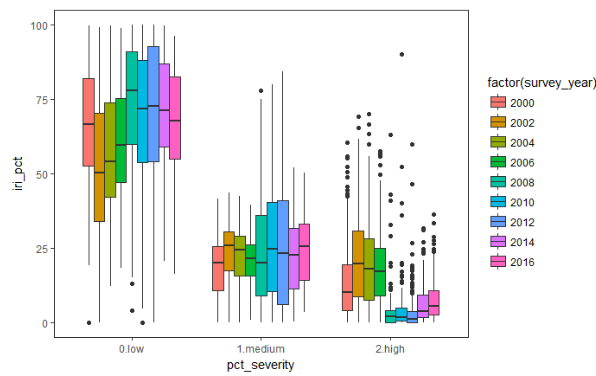
(a) Percentage of IRI in Categories (Asphalt Pavement)



(b) Percentage of IRI in Categories (Concrete Pavement)



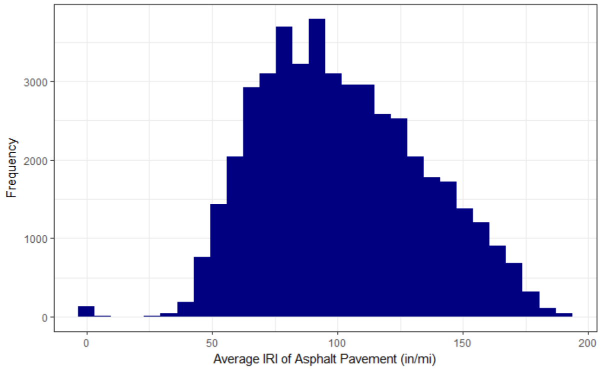
(c) Percentage of IRI in Different Years (Asphalt Pavement)



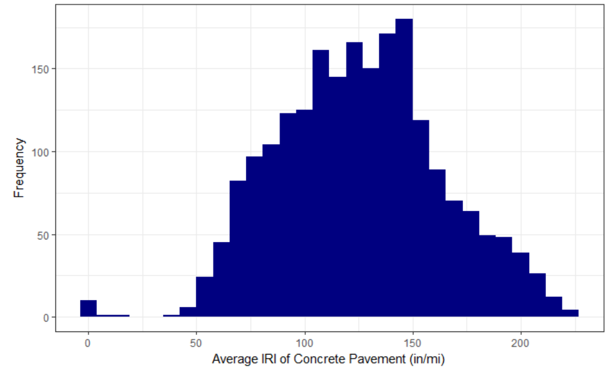
(d) Percentage of IRI in Different Years (Concrete Pavement)

Figure 26 Boxplots of IRI Percentage in Non-warranty Pavements

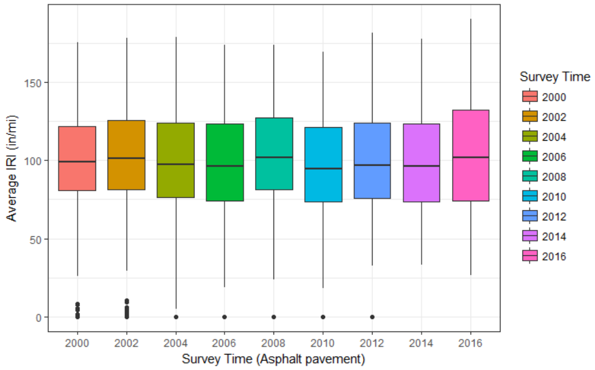
The average IRI has also been tracked in PMS as a pavement condition indicator for both the non-warranty asphalt and concrete pavements, and the statistical descriptions of the IRI's in the non-warranty pavements are presented in Figure 27. Figure 27 (a) and (b) indicate that the distribution of the average IRI seems approximately following normal distributions in both the asphalt and concrete pavements. As expected, the average IRI in the concrete pavements is larger than that in the asphalt pavements as shown in Figure 27 (e) and (f). Figure 27 (c)-(e) show that the average IRI keeps in a certain range over time in the asphalt pavements, while it has increased slightly in the concrete pavements after service year 2008.



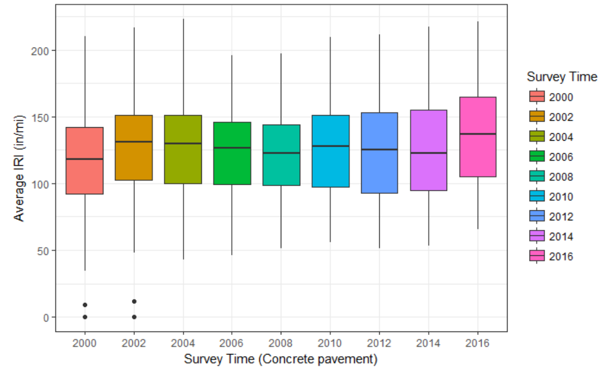
(a) Histogram of Average IRI (Asphalt Pavement)



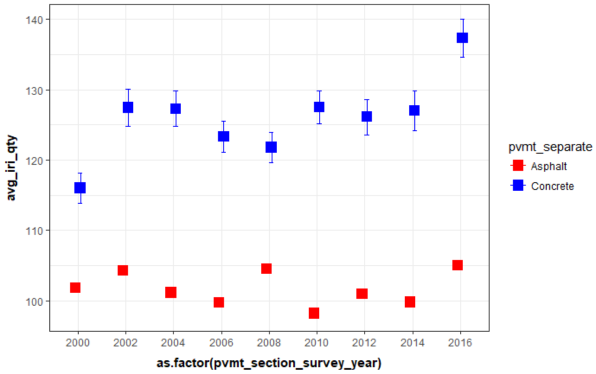
(b) Histogram of Average IRI (Concrete Pavement)



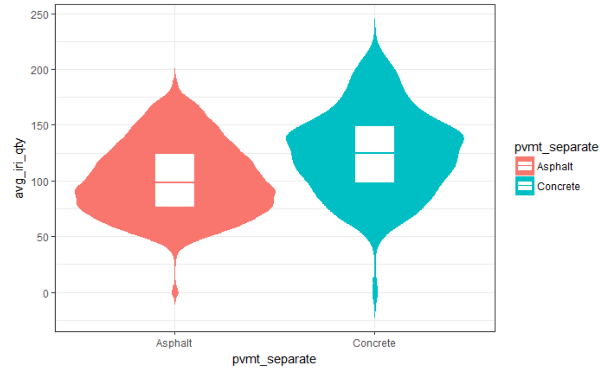
(c) Boxplot Average IRI (Asphalt Pavement)



(d) Boxplot Average IRI (Concrete Pavement)



(e) Boxplot with Error Bar in Different Years



(f) Violin Plot

Figure 27 Histograms, Boxplots and Violin Plots of IRI Averages in Non-warranty Pavements

CHAPTER 5. METHODOLOGY

This chapter presents the methodology employed in the study to possibly develop the new distress thresholds for MDOT's pavement warranty program. The survival analysis and models are introduced in the "Literature Review" and "Evaluation of Warranty and Non-warranty Pavement Performance in Mississippi" chapters, and will not be included in this chapter. In this chapter, first, to follow NCHRP Report 699, the standard deviation or confidence interval method is employed to align the cumulative percentiles of the non-warranty pavement distress data (assuming normal distributions) with the desired distress thresholds for the pavement warranty program. Then, to deal with the skewedness in realistic datasets and for processing and cleansing the raw data, the bootstrapping method is used to overcome the bias/error inducing irregularity and uncertainty in data distributions and to regenerate the characterizing PDF and CDF curves for different distress types. Finally, the confusion matrix model is introduced to evaluate the equivalency and smoothness of the maintenance decision making by using the new measurement based thresholds in the proposed warranty program and by using the deduct point based thresholds in the existing warranty program.

5.1 Standard Deviation

NCHRP Report 699 provides a statistical method to develop acceptable pavement distress thresholds based on the historical performance data. The method assumes the pavement distress data follow normal distributions, and then suggests setting the threshold at 2.0 σ or two times standard deviations above the mean (where only 5% of measured sections would exceed the threshold based on one tail test) to reduce the risk to the contractor for a 10-year warranted pavement (Scott III et al. 2011). According to the additional experience or improved consistency, the threshold is supported to be further tightened (to between 1.0 σ and 2.0 σ). This method was implemented to create historical pavement performance data distributions which satisfied the performance requirements in Canada (Smith et al. 2016). Actually, the pavement performance data do not always follow normal distributions in practice, but some of them are quite skewed, and normalization would be a big issue in the use of this method. In another way of interpretation, the threshold could be based on the best-fit probability distribution or probability density function, and setting the threshold at the upper bound of the confidence interval that would have confidence probabilities of 68%, to 95% for the data to fall within the interval with the threshold at the upper bound of the interval from 1.0 σ to 2.0 σ respectively in distance to the sample mean (Scott III et al. 2011, Smith et al. 2016).

5.2 Bootstrapping Resampling

Bootstrapping, a common resampling method, is a general approach of statistical inference with mild modeling assumptions and broad applicability for raw data with complex or even unknown distributions. It is a nonparametric method that relies on large amounts of computation rather than mathematical analysis and distributional assumptions (Efron and Tibshirani 1994). The bootstrapping method involves random re-sampling of the existing raw data with replacements to provide the statistical characteristics of a complex or unknown statistical distribution, and a bootstrapping sample is taken from the original population. In this study, the bootstrapping technique was used to re-sample the skewed non-warranty pavement distress measurement data to regenerate a bootstrapped distribution (Delucchi and Bostrom 2004).

In resampling of a one-sample situation, let $X = [X_1, X_2, \cdots, X_n]$ be an independent and identically distributed (i.i.d.) random sample of size *n*, and let $x = [x_1, x_2, \cdots, x_n]$ denote the observed realization from an underlying population with unknown distribution function F,

$$X_i = x_i, \quad X_i \sim F \quad i = 1, 2, \cdots, n.$$

Define $R(X, F)$ as the actual random variable, which is the function of X and F. Then the observed sample $x = [x_1, x_2, \cdots, x_n]$ could be employed to estimate the distribution characteristics of $R(X, F)$ (Efron 1979). Assume $\theta = \theta(F)$ is some parameter of interest such as the mean, medium, correlation, or standard deviation of F. $\hat{\theta} = \tau(X)$ is an estimator of the parameter θ of F ,with τ denoting some function, that can be used to estimate θ from the data. In this setting, it is the deviation of $\hat{\theta}$ from θ that is of our primary interest.

The general idea of the bootstrapping method is to approximate F from the data and use the approximation $\hat{F}$, instead of F. In practice, an approximation of the estimator distribution is obtained by drawing N samples of size n from the true distribution F (with replacement, if F is discrete). Two main broad areas, to obtain the approximation $\hat{F}$, are the parametric and non-parametric bootstraps relied on a known distribution function and an unknown distribution function respectively. In the study, the distributions of the observations are obviously skewed and unknown to us. Thus, the non-parametric bootstrapping is the alternative resampling method. Assume a data set $x = [x_1, x_2, \cdots, x_n]$ is available, the implementation of the non-parametric bootstrapping resampling method is presented as follows (Chernick and LaBudde 2014) (as shown in Figure 28).

Step 1. Choose the number of bootstrap re-samples N; frequently $N \in [1000, 2000]$.

Step 2. Sample a new data set x^* of size n from x with replacement.

Step 3. Estimate θ from x^* by calling the estimator $\hat{\theta}_i^*$, for $i = 1, 2, \cdots, N$. Store $\hat{\theta}_i^*$.

Step 4. Repeat step 2 and step 3 for *N* times.

Step 5. Use the empirical distribution of $\left[\hat{\theta}_1^*, \hat{\theta}_2^*, \cdots, \hat{\theta}_N^*\right]$ as an approximation of the true distribution of $\hat{\theta}$ (use mean as the distribution function in this case).

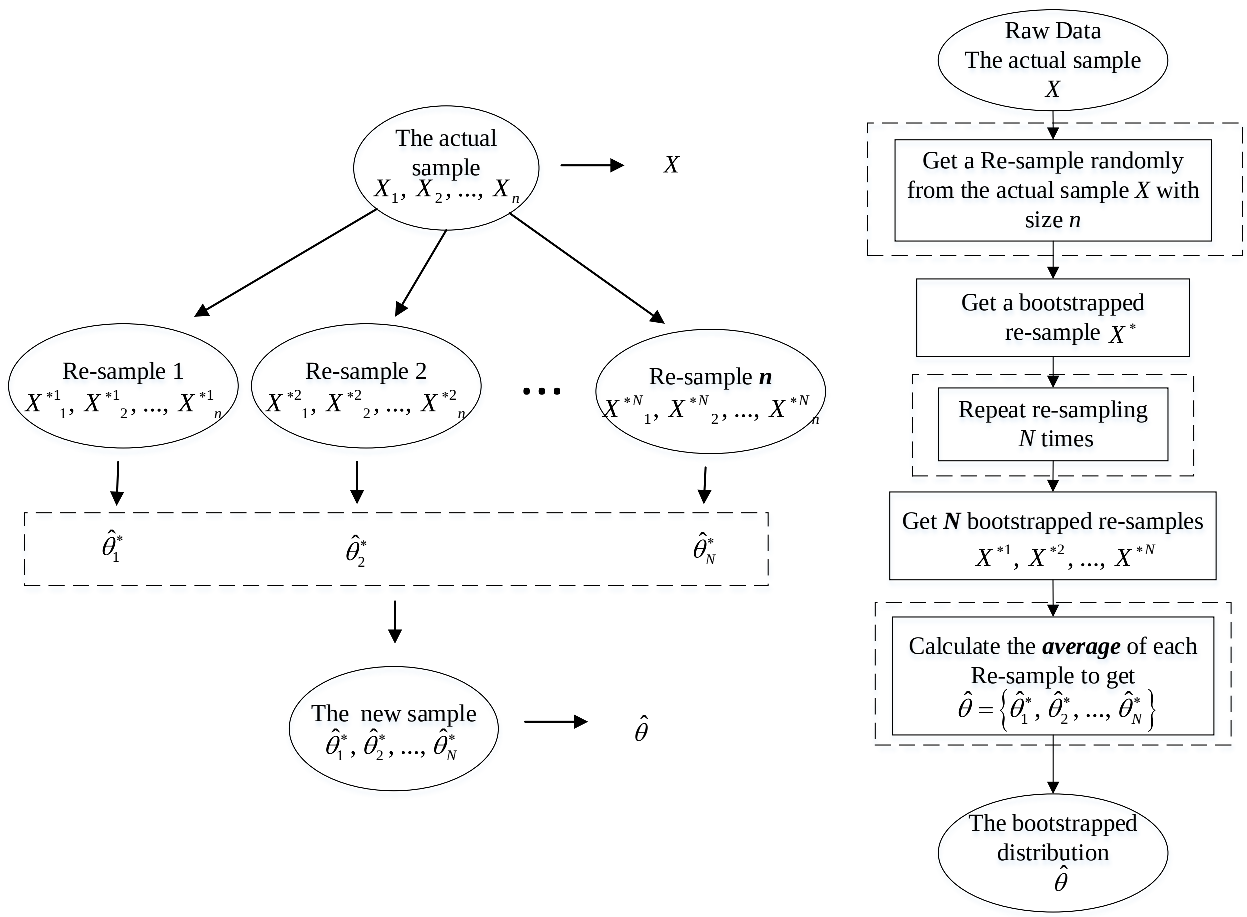


Figure 28 Basic Idea of Bootstrapping Method

5.4 Confusion Matrix and Accuracy

A confusion matrix, also known as an error matrix, is a table that is used to visualize the performance of a model in the field of machine learning and specifically the problem of statistical classification (Stehman 1997). It is a kind of contingency table, with two dimensions ("actual" and "predicted"), to summarize the results of test data for which the true and false values are known. In the confusion matrix, each column represents the instances in a predicted class while each row represents the instances in an actual class. The two columns have two possible classes "YES" and "NO," respectively, which means the pavement is predicted to be failed or not failed, while the two rows have two possible classes “YES” and “NO,” respectively, which means the pavement has actually failed or not failed (Powers 2011). In predictive analytics, the confusion matrix for binary classifier is a table that reports the number of true positives, true negatives, false positives, and false negatives.

Table 33 Confusion Matrix

	Predicted: NO	Predicted: YES
Actual: NO	TN	FP
Actual: YES	FN	TP

The values in the confusion matrix are whole numbers (the number of the record road segment for each distress at the specific confusion matrix categories, not rate) in Table 33:

1) n is the size of the total dataset in this confusion matrix.
2) TP is the true positives, which the predicted is “YES” and the actual is also “YES”.
3) TN is the true negatives, which the predicted is “NO”, and the actual is also “NO”.
4) FP is the false positives, which the predicted is “YES”, but the actual is “NO” (also known as a "Type I error.")
5) FN is the false negatives, which the predicted is “NO”, but the actual is “YES” (also known as a "Type II error.")

The confusion matrix is frequently employed to organize and display correct classifications (accuracy) have been proposed for summarizing the information contained in the confusion matrix (Stehman 1997). The equation of the accuracy is determined as:

$$Accuracy\ (ACC) = \frac{\sum True\ positive + \sum True\ negative}{\sum Total\ population} = \frac{TP + TN}{TP + FP + FN + TN} \quad \textbf{5-1}$$

CHAPTER 6. DEVELOPMENT OF NEW THRESHOLDS

This chapter presents the data processing and analysis results of the study following the methodology described in the previous chapter. The results are divided into two parts: the development of warranty distress thresholds, and the comparison and selection of warranty thresholds. In the first part, by employing the longitudinal cracking as an example, the non-warranty distress measurement data were processed by using the bootstrapping method to normalize the data distributions, followed by basic statistical processing to plot the PDF and CDF curves of the normalized distress distributions. Then the best-fit hypothesis was applied to create the example best-fit thresholds for longitudinal cracking, and the confusion matrix was utilized to evaluate the accuracy of the example best-fit thresholds. Next, the standard deviation method was applied to develop multiple threshold values (from 1.0 σ, 1.1 σ, …, 2.0 σ) of other four distress types to find the appropriate value for the number of standard deviations in the development of the distress thresholds. Eventually, the accuracy of the best-fit thresholds for five typical distress types were evaluated to search for the best options for measurement based thresholds with the least deviations. In the second part, the best-fit thresholds of five distress types and the 1.0 σ and 1.5 σ associated thresholds of all the distress types (with sufficient and valid data) were presented to compare the applicability of the two methods (i.e., the best-fit and the standard deviation methods). Finally, the two methods were employed to determine the thresholds of IRI to possibly include IRI in the new warranty program.

6.1 Development of Warranty Distress Thresholds

In this study, two methods were adopted to set up the distress measurement based thresholds. One is based on the PDF of the non-warranty data in PMS. The other method, which is called the best-fit method, relies on both the bootstrapping normalized distribution of the non-warranty data in PMS and warranty annual maintained pavement distress reports.

6.1.1 Standard Deviation Method

From the PDFs, the probabilities of 68% and 86.6% of the data would fall within the ranges of ± 1.0 σ and ± 1.5 σ, respectively, around the sample means. The bootstrapping normalized CDFs of distress measurements at three severity levels, (as illustrated in

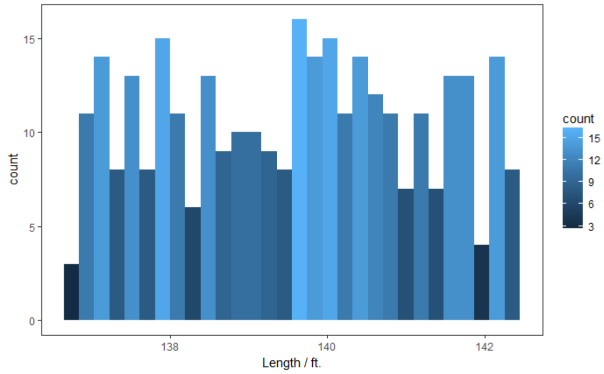
Figure 30 (b), (d), and (f) for longitudinal cracking,) were created to develop pavement distress threshold values with 1.0 σ and 1.5 σ, respectively. As an example for how to use the CDF curve (as shown in Figure 30 (b), (d), or (f)), the first step is to find the percentile probability at the y or vertical axis, then follow the CDF curve to determine its corresponding distress amount at the x or horizontal axis. The located distress amount is the corresponding distress threshold with a distance in terms of the number of standard deviations from the mean for the distress type at the specific severity level. For example, one time standard deviation above the sample mean or the symbol 1.0 σ in a normal distribution corresponds with the cumulative percentile value at 68% on the CDF curve. (Using Figure 30 (b) as an example, on the vertical or y axis locate 68% and draw a horizontal line to intersect with the CDF curve and from the intersection point draw a vertical line downward to arrive at the value of 287.4 ft on the x or horizontal axis. That way, the threshold value for longitudinal cracking at low severity level, based on the one time standard deviation method or the 1.0 σ method would be determined as 287.4 ft.)

6.1.1.1 Non-parametric Bootstrap Simulation--Longitudinal Cracking

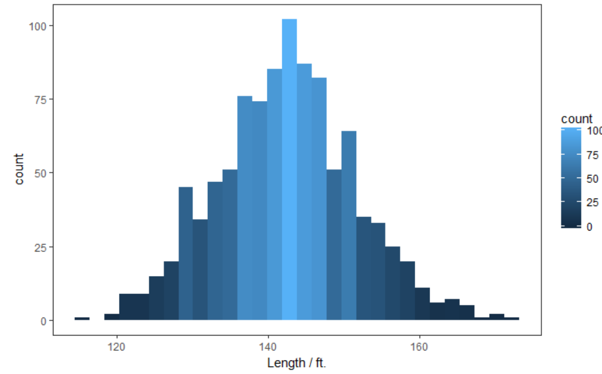
Most of the non-warranted pavement physical distresses do not follow a normal distribution. Therefore, the bootstrapping algorithm was used to generalize every percentile (from 1% to 100%) of the data for each distress type when no alternative approach was sufficiently accurate. In sampling with replacement, every sampled data point was returned to the dataset after sampling, and the resampled data size could be larger than the original one. In this way, the sample error could be reduced. Eventually, the CDF curves of the non-warranted pavement distress data in every percentile were constructed. For instance, in order to reduce the sample error of the longitudinal cracking at low severity level (sample size: 33,723), the bias-corrected percentile non-parametric bootstrap simulation procedure was used to construct the selected percentile intervals (from 0% to 100%, a total of 100 data samples). The bootstrapping was done by running the noncommercial statistical software R 3.3.3.

The data processing consists of the following steps: 1) Retrieve data points from MDOT PMS database within percentiles between 0% and 1% of the longitudinal cracking at low severity level as a raw bootstrap sample X. Set the number of bootstrap re-samples, N, at N = 1000. 2) Form a new data set x^* of size n by sampling from X with replacement, with n = 338 (1/100 of raw data sample size). 3) Compute the statistic mean θ from x^* by calling the estimator $\hat{\theta}_i^*$, for $i = 1, 2, \cdots, N$. Store the calculated $\hat{\theta}_i^*$. 4) Repeat step 2 and step 3 for N times. 5) The empirical distribution of $\left[\hat{\theta}_1^*, \hat{\theta}_2^*, \cdots, \hat{\theta}_N^*\right]$ is the approximation of the true distribution of $\hat{\theta}$. 6) Repeat the 5 steps (from step 1 to step 5) for 99 times for the rest of the percentile intervals.

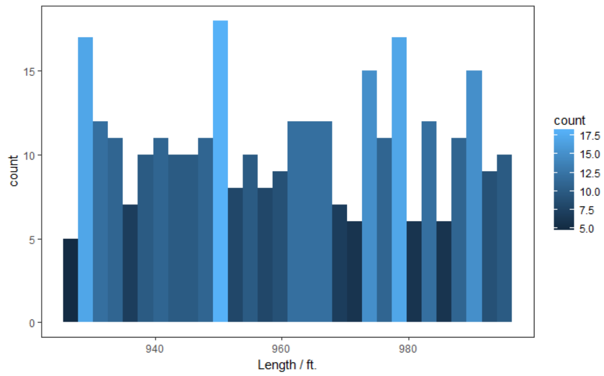
As shown in Figure 29, the histograms of the raw data at percentile intervals 49%-50% and 94%-95% and the bootstrapping resampled data are selected to illustrate the effectiveness of the bootstrapping resampling method. The raw data in Figure 29 (a) and (c) are discrete and do not follow a normal distribution. In contrast, the bootstrapping resampled datasets in Figure 29 (b) and (d) can approximately follow a normal distribution. Therefore, the bootstrapping method is effective in the normalization of the discrete sample data with unknown distributions.



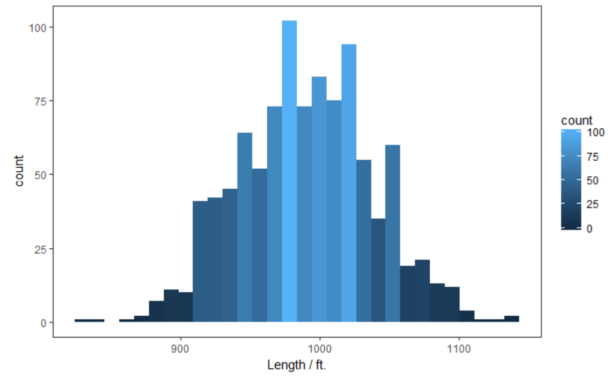
(a) Percentile 49%-50% Raw Data



(b) Percentile 49%-50% Resampled Data



(c) Percentile 94%-95% Raw Data

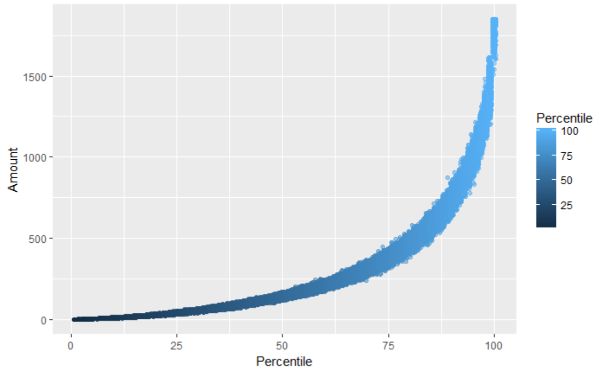


(d) Percentile 94%-95% Resampled Data

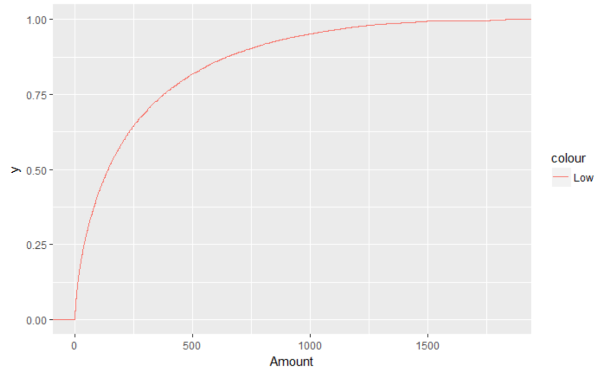
Figure 29 Histograms of Percentile Intervals of Longitudinal Cracking at Low Severity Level before and after Bootstrapping

6.1.1.2 CDF Curves of Bootstrapped Distress Data

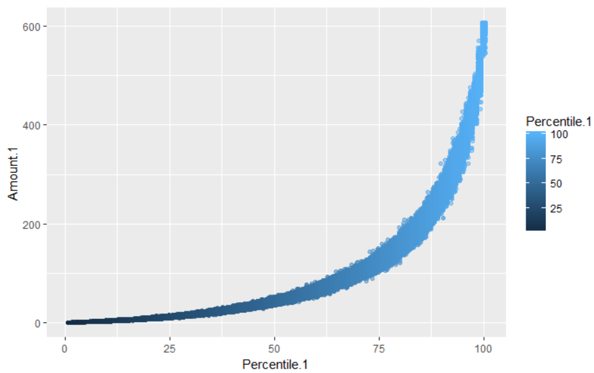
The non-parametric bootstrap simulation procedure presented in the previous section was followed to retrieve and resample distress data in each percentile interval of the transverse cracking, alligator cracking, block cracking, edge cracking, bleeding, and reflection cracking at three severity levels. The selected bootstrapped CDF curves of longitudinal cracking in asphalt non-warranty pavements are shown in Figure 30. Figure 30 (a), (c), and (e) plot the resampled distress amounts of longitudinal cracking with cumulative percentile at low, medium, and high severity levels, respectively. Figure 30 (b), (d), and (f) show the CDFs of the longitudinal cracking at three severity levels drawn from the bootstrap normalized data. At a severity level, the distress length at any percentile could be found in the CDF curves.



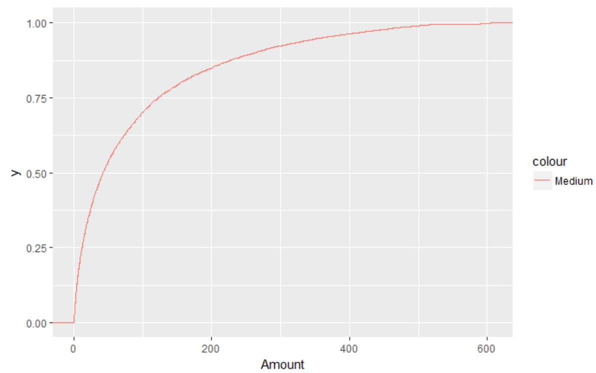
(a) Distress Amount of Low Severity Level



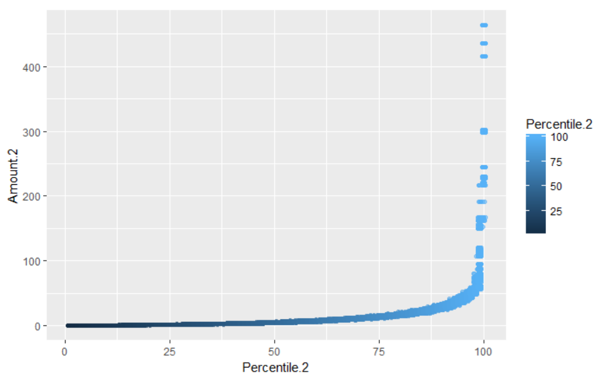
(b) CDF of Low Severity Level



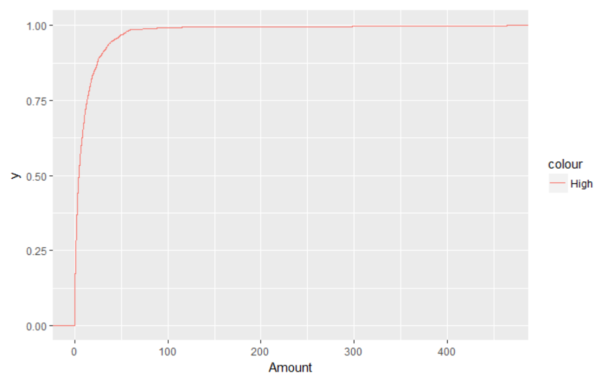
(c) Distress Amount of Medium Severity Level



(d) CDF of Medium Severity Level



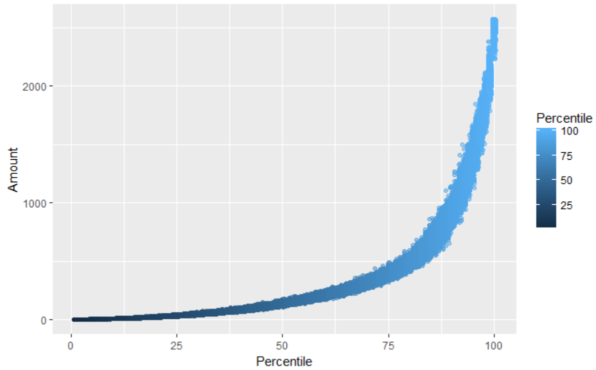
(e) Distress Amount of High Severity Level



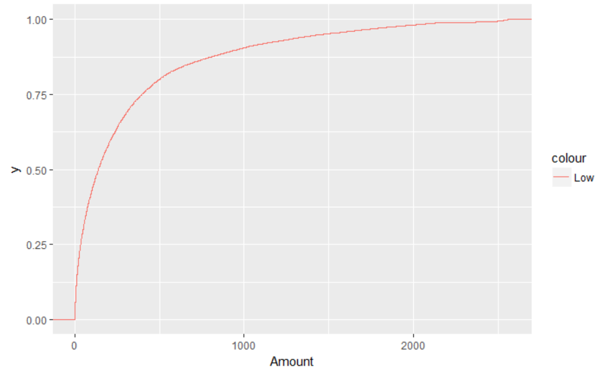
(f) CDF of High Severity Level

Figure 30 Cumulative Distribution Curves of Longitudinal Cracking in Asphalt Pavements

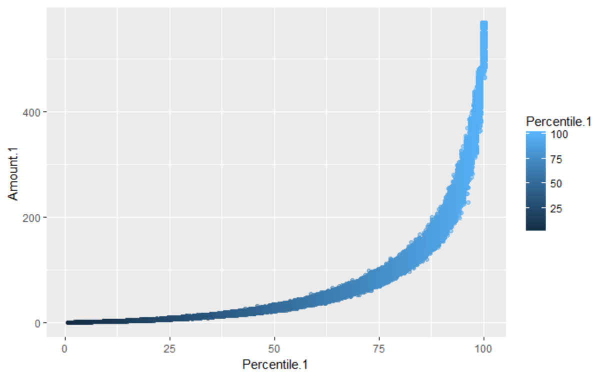
The plots of the resampled distress amount and the CDF curves at three severity levels representing transverse cracking, alligator cracking, block cracking, edge cracking, bleeding, and reflection cracking in asphalt non-warranty pavements are presented in Figure 31 through Figure 36. For the concrete pavements, the plots and CDF curves for longitudinal cracking, transverse cracking, spalling of longitudinal joint, and spalling of transverse joint are displayed in Figure 37 through Figure 40.



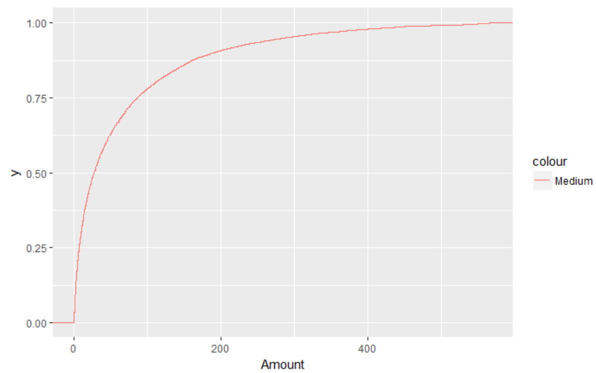
(a) Distress Amount of Low Severity Level



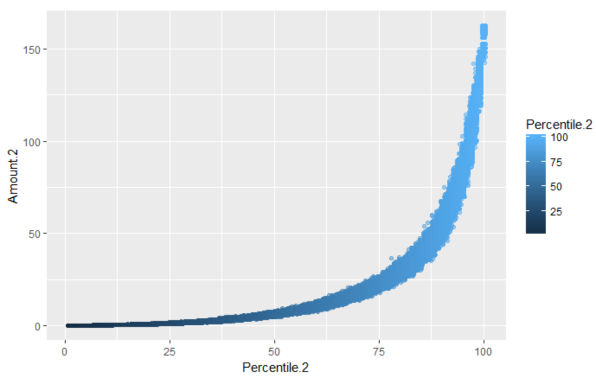
(b) CDF of Low Severity Level



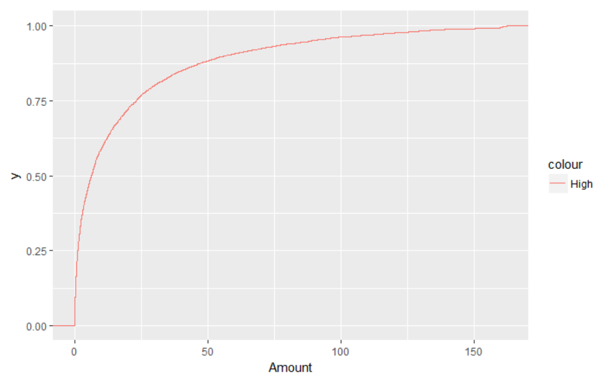
(c) Distress Amount of Medium Severity Level



(d) CDF of Medium Severity Level

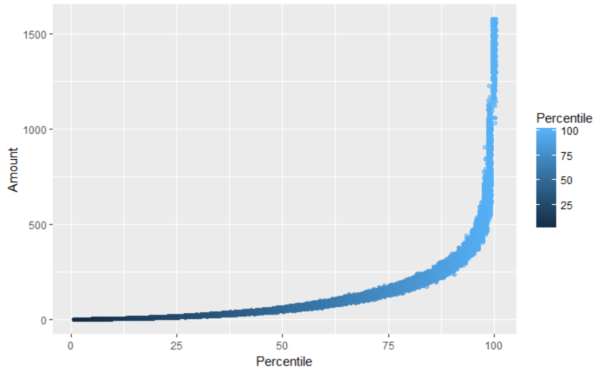


(e) Distress Amount of High Severity Level

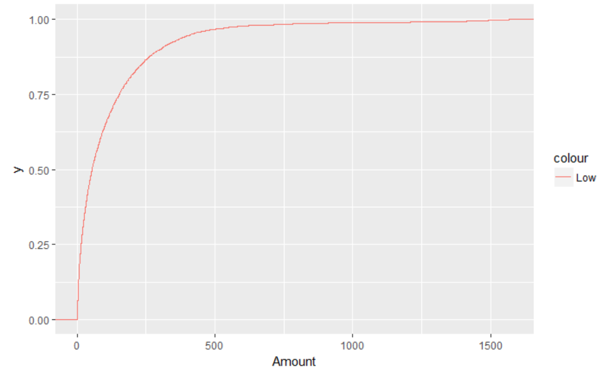


(f) CDF of High Severity Level

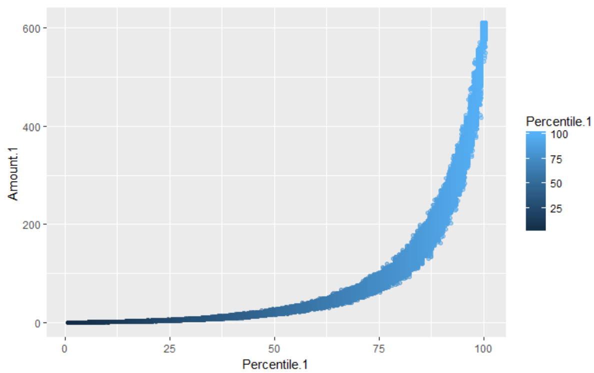
Figure 31 Cumulative Distribution Curves of Transverse Cracking in Asphalt Pavements



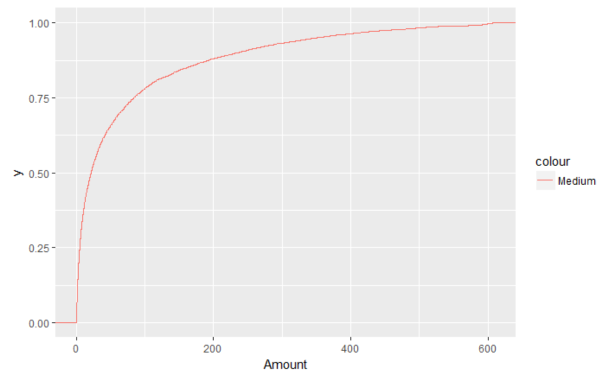
(a) Distress Amount of Low Severity Level



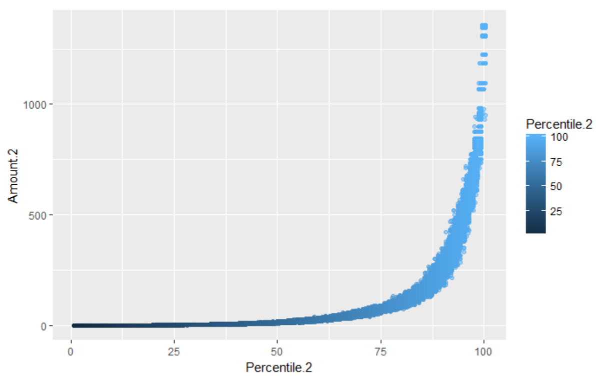
(b) CDF of Low Severity Level



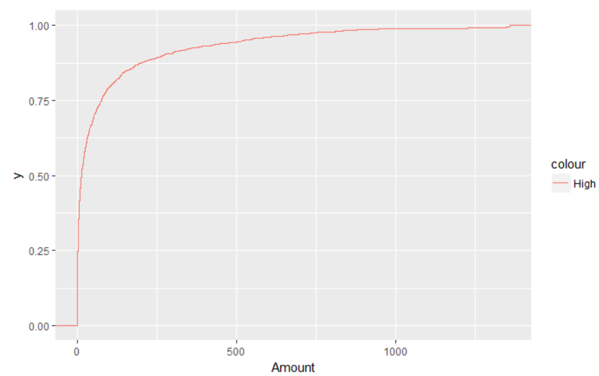
(c) Distress Amount of Medium Severity Level



(d) CDF of Medium Severity Level

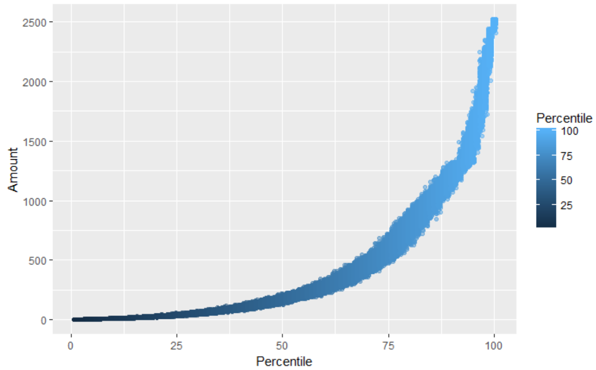


(e) Distress Amount of High Severity Level

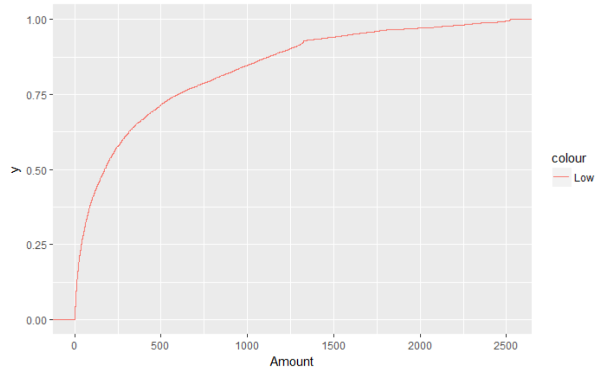


(f) CDF of High Severity Level

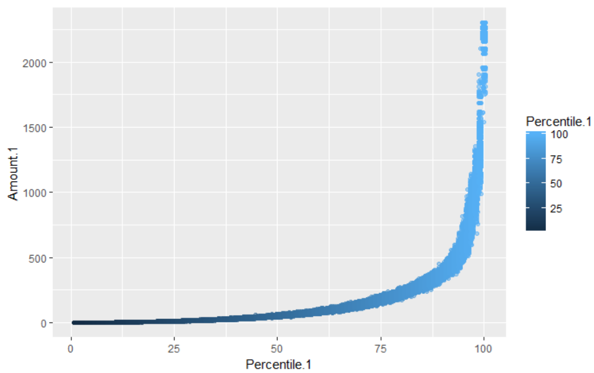
Figure 32 Cumulative Distribution Curves of Alligator Cracking in Asphalt Pavements



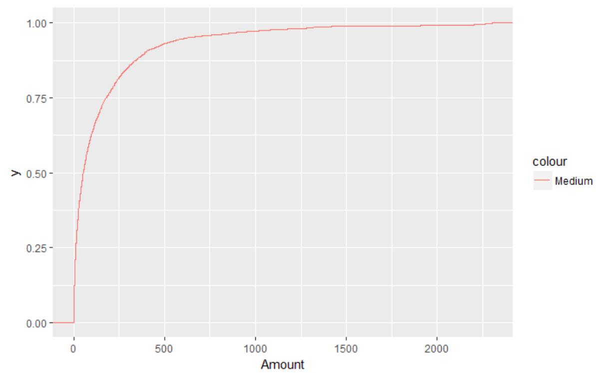
(a) Distress Amount of Low Severity Level



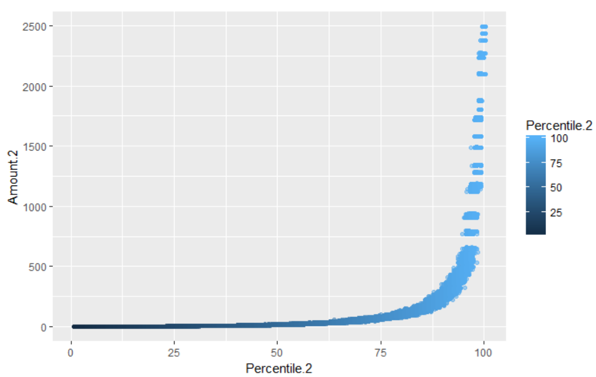
(b) CDF of Low Severity Level



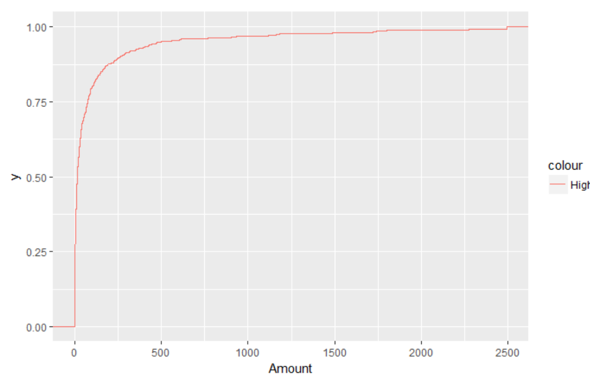
(c) Distress Amount of Medium Severity Level



(d) CDF of Medium Severity Level

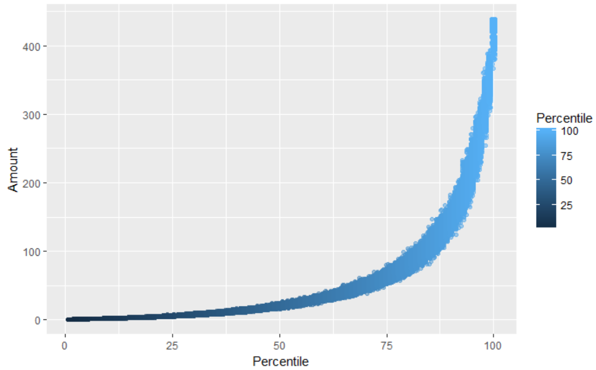


(e) Distress Amount of High Severity Level

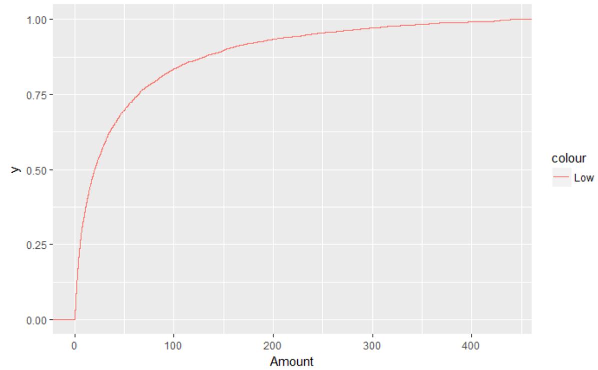


(f) CDF of High Severity Level

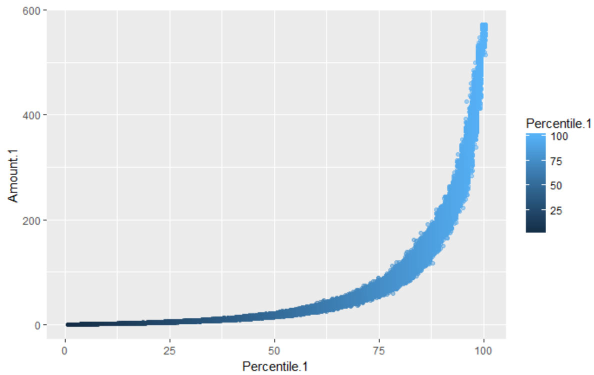
Figure 33 Cumulative Distribution Curves of Block Cracking in Asphalt Pavements



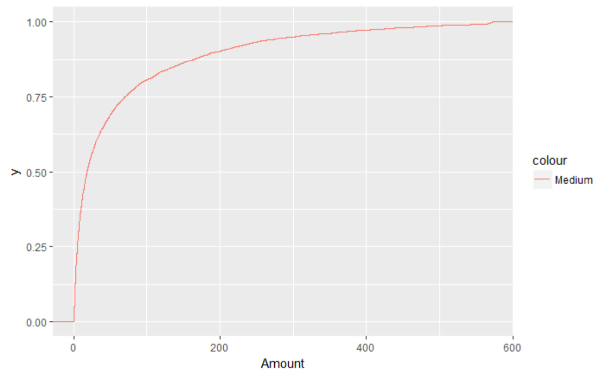
(a) Distress Amount of Low Severity Level



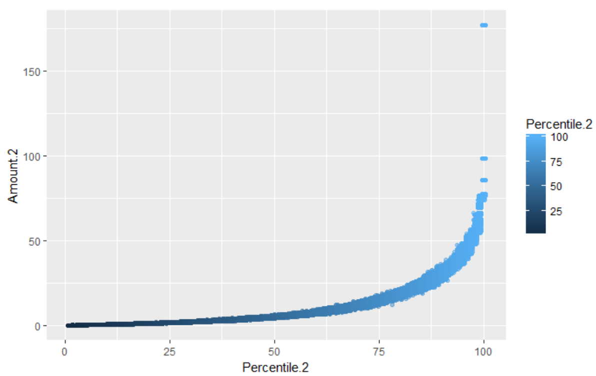
(b) CDF of Low Severity Level



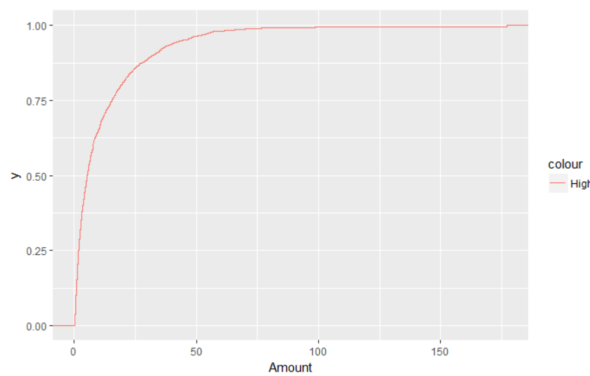
(c) Distress Amount of Medium Severity Level



(d) CDF of Medium Severity Level

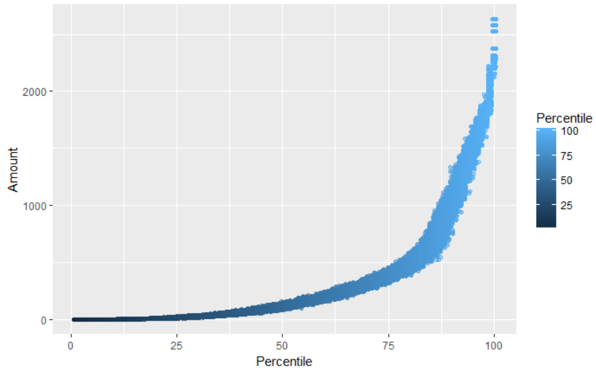


(e) Distress Amount of High Severity Level

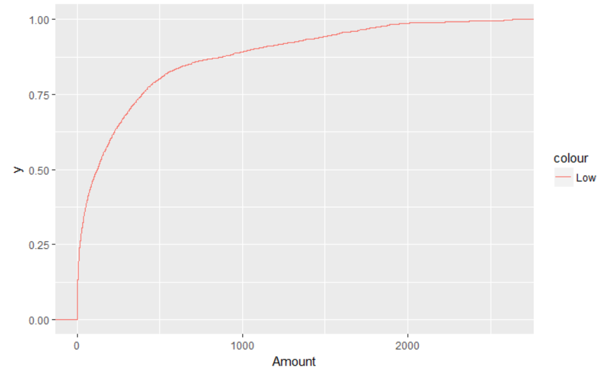


(f) CDF of High Severity Level

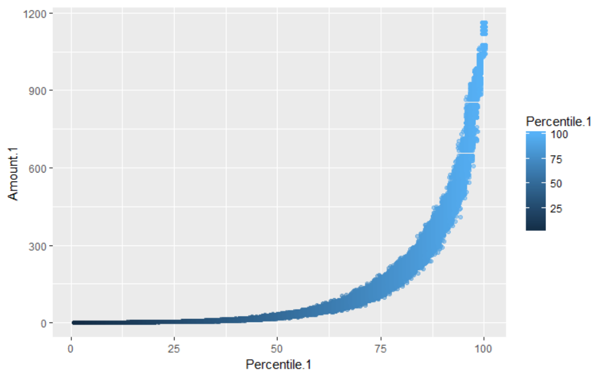
Figure 34 Cumulative Distribution Curves of Edge Cracking in Asphalt Pavements



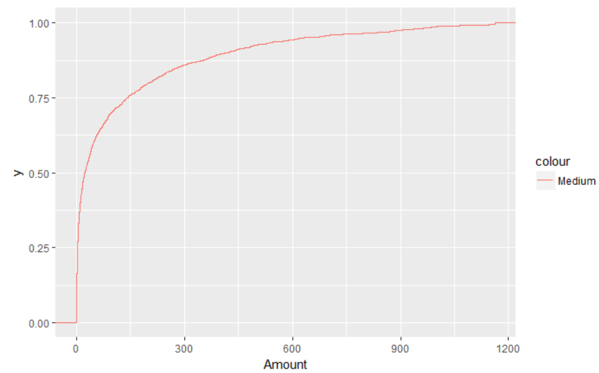
(a) Distress Amount of Low Severity Level



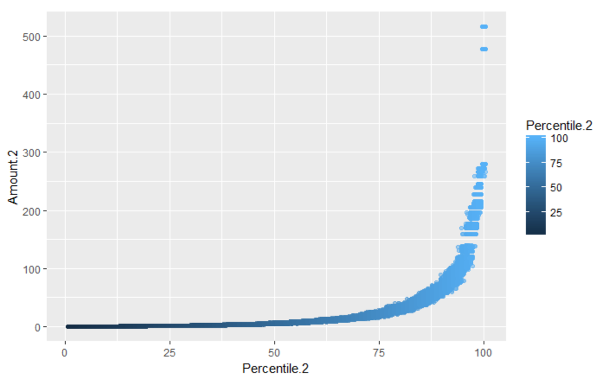
(b) CDF of Low Severity Level



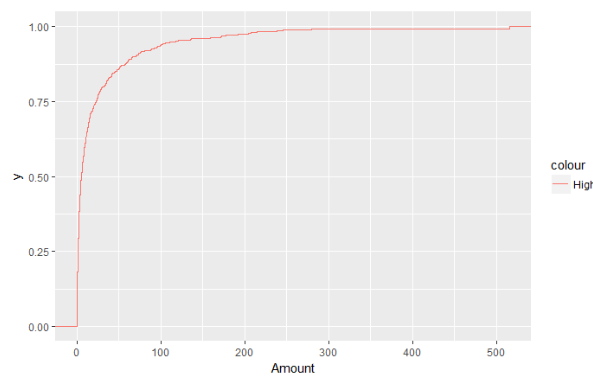
(c) Distress Amount of Medium Severity Level



(d) CDF of Medium Severity Level

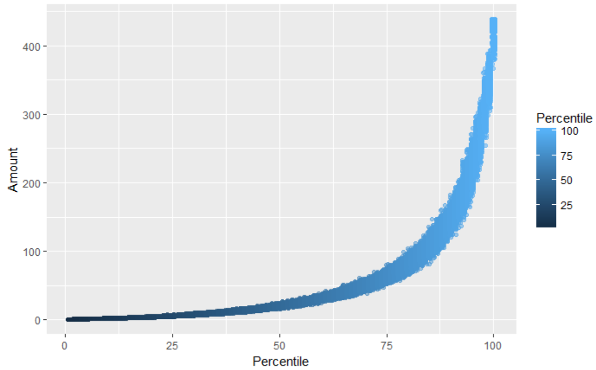


(e) Distress Amount of High Severity Level

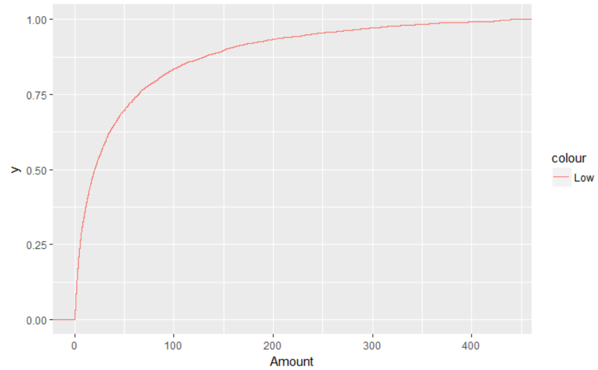


(f) CDF of High Severity Level

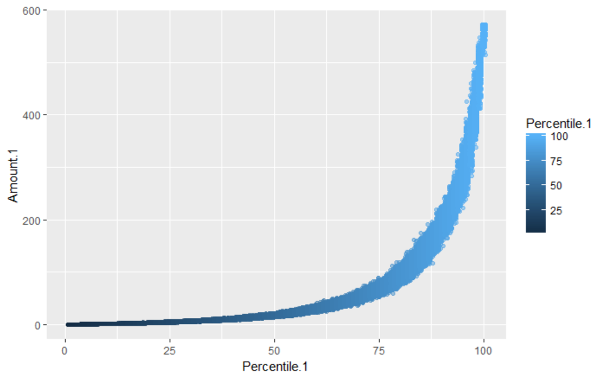
Figure 35 Cumulative Distribution Curves of Bleeding in Asphalt Pavements



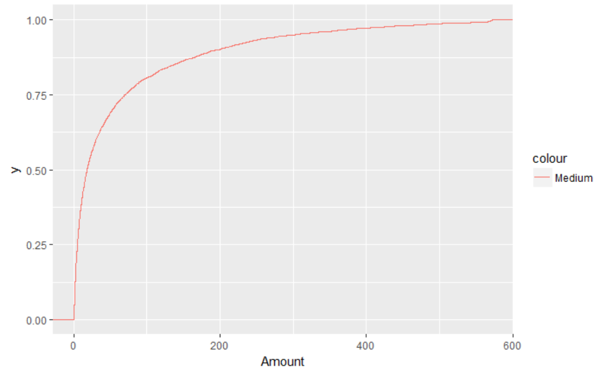
(a) Distress Amount of Low Severity Level



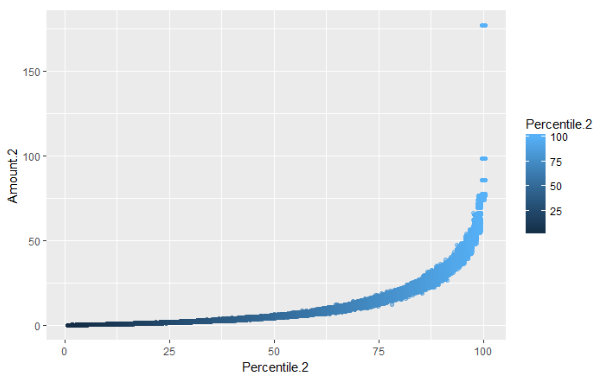
(b) CDF of Low Severity Level



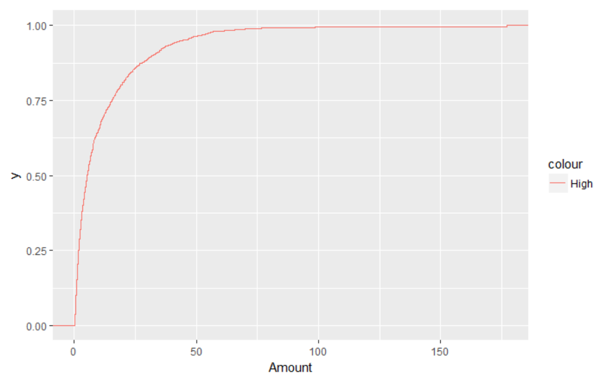
(c) Distress Amount of Medium Severity Level



(d) CDF of Medium Severity Level

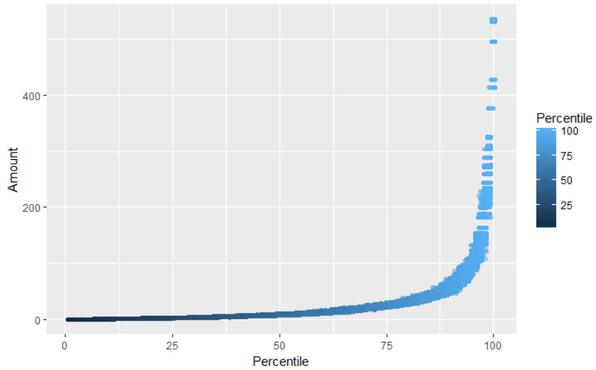


(e) Distress Amount of High Severity Level

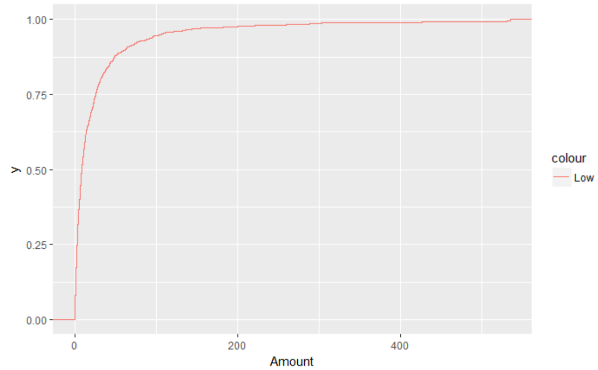


(f) CDF of High Severity Level

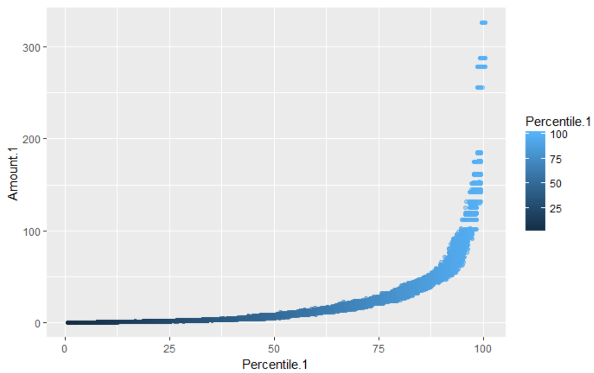
Figure 36 Cumulative Distribution Curves of Reflection Cracking in Asphalt Pavements



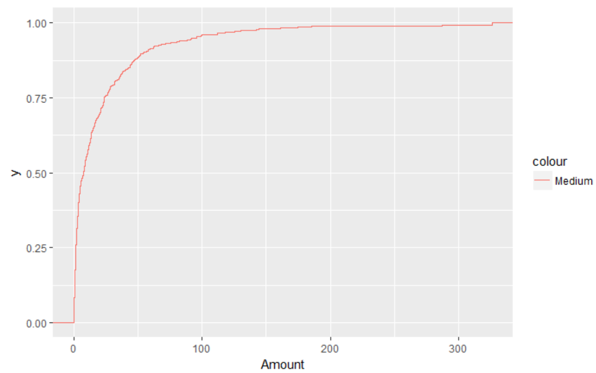
(a) Distress Amount of Low Severity Level



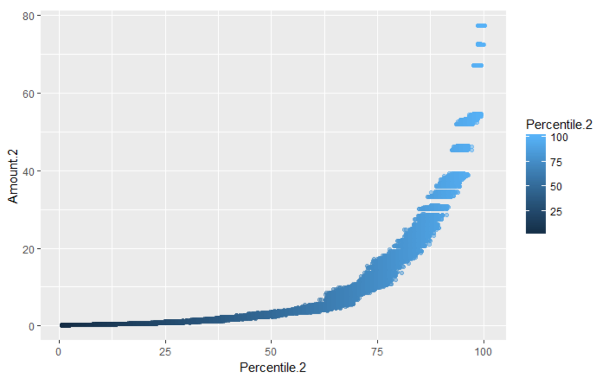
(b) CDF of Low Severity Level



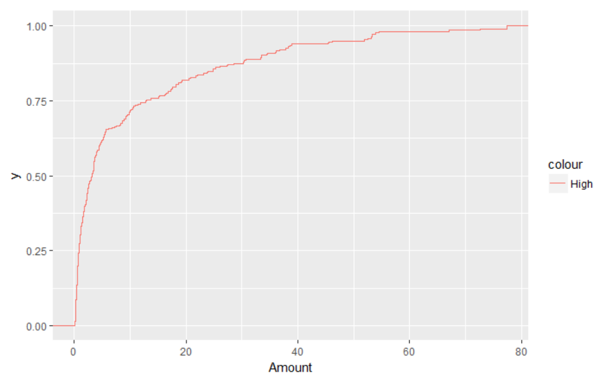
(c) Distress Amount of Medium Severity Level



(d) CDF of Medium Severity Level

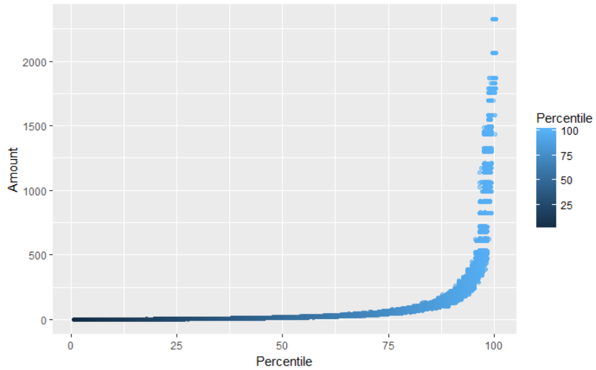


(e) Distress Amount of High Severity Level

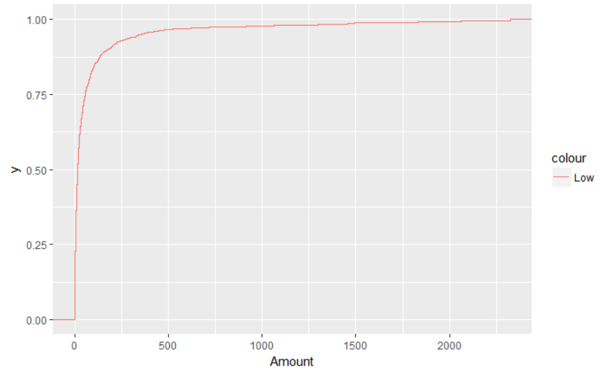


(f) CDF of High Severity Level

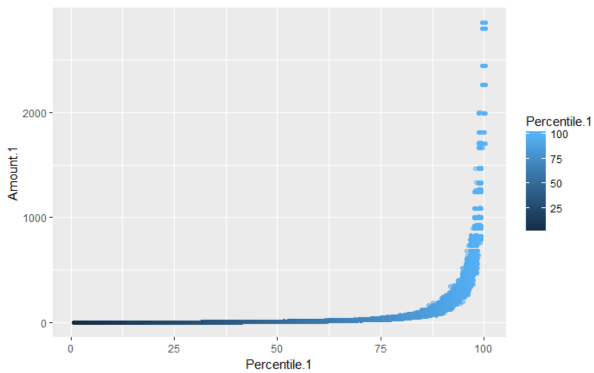
Figure 37 Cumulative Distribution Curves of Longitudinal Cracking in Concrete Pavements



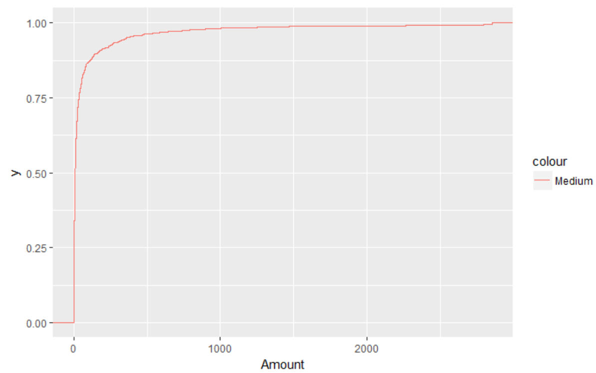
(a) Distress Amount of Low Severity Level



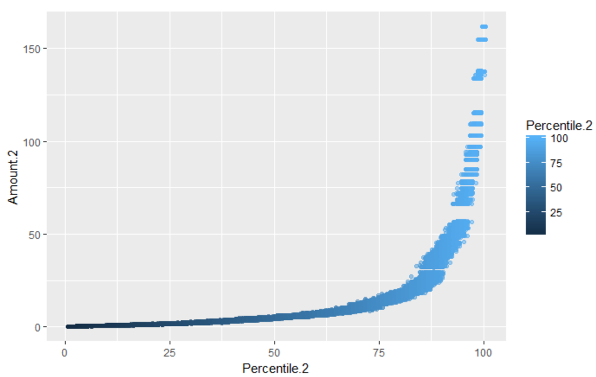
(b) CDF of Low Severity Level



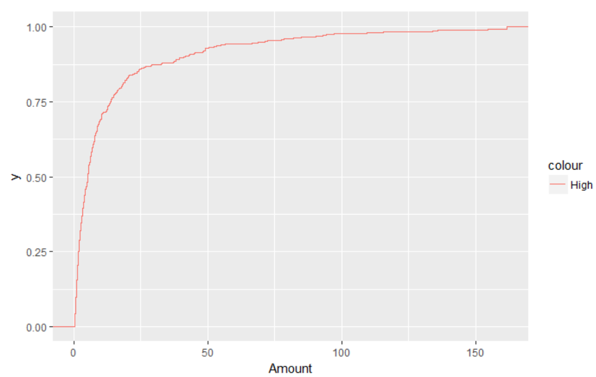
(c) Distress Amount of Medium Severity Level



(d) CDF of Medium Severity Level

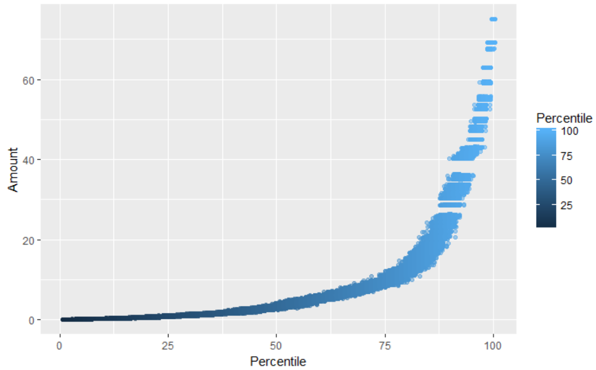


(e) Distress Amount of High Severity Level

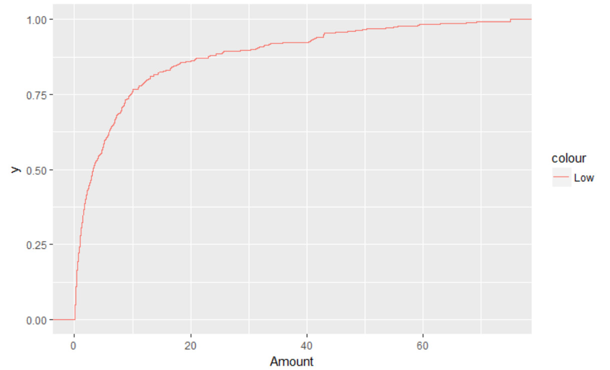


(f) CDF of High Severity Level

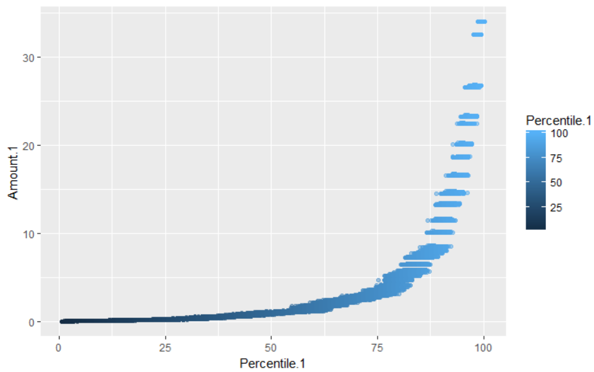
Figure 38 Cumulative Distribution Curves of Transverse Cracking in Concrete Pavements



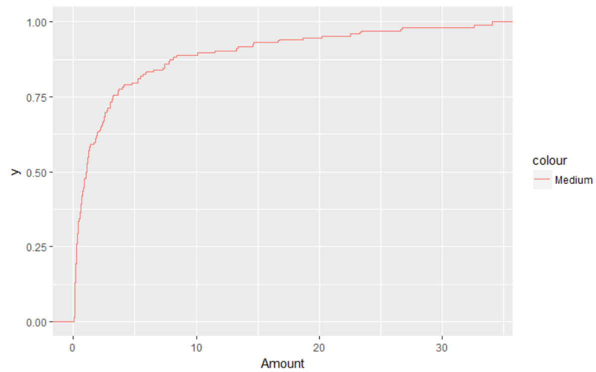
(a) Distress Amount of Low Severity Level



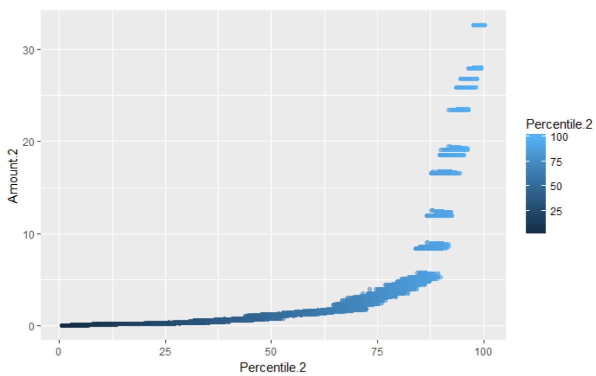
(b) CDF of Low Severity Level



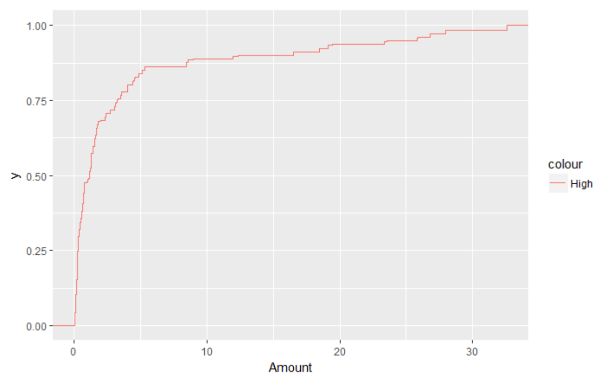
(c) Distress Amount of Medium Severity Level



(d) CDF of Medium Severity Level

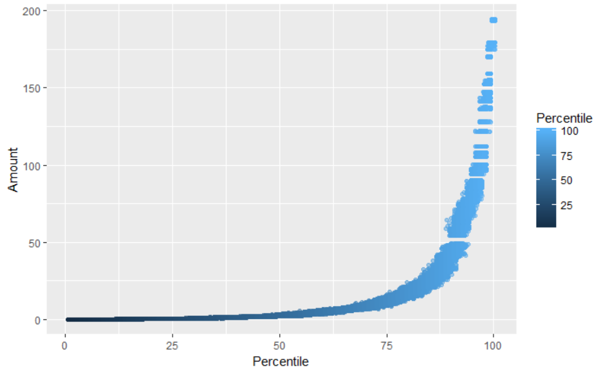


(e) Distress Amount of High Severity Level

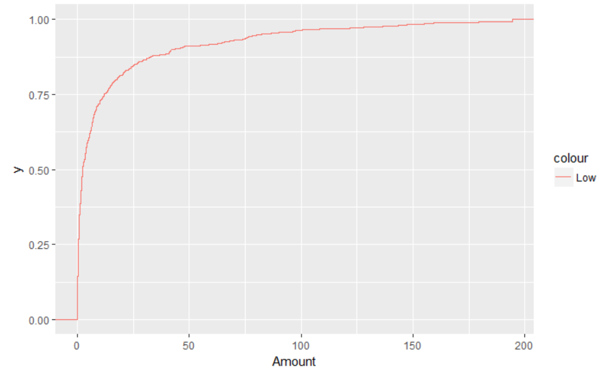


(f) CDF of High Severity Level

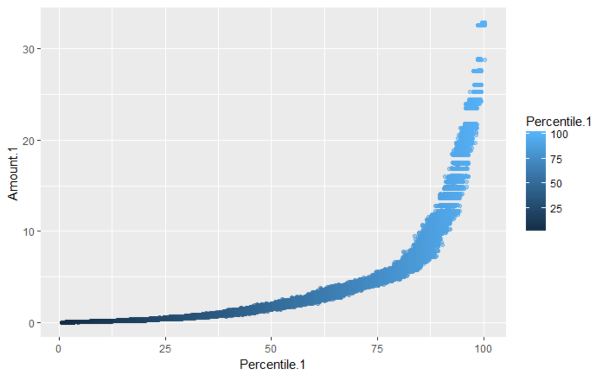
Figure 39 Cumulative Distribution Curves of Spalling of Longitudinal Joint in Concrete Pavements



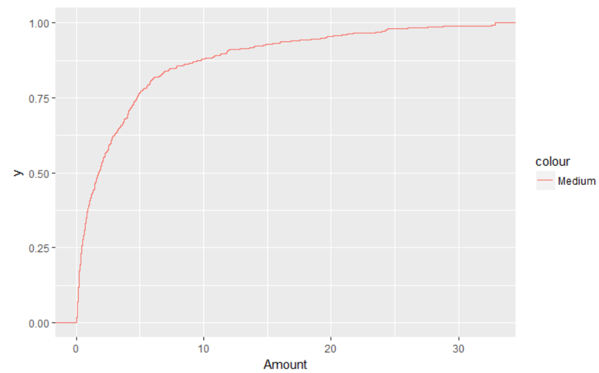
(a) Distress Amount of Low Severity Level



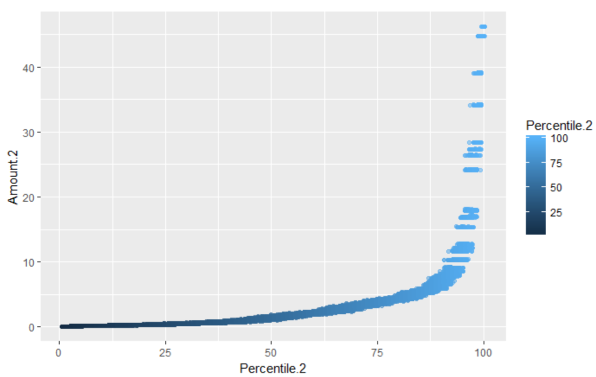
(b) CDF of Low Severity Level



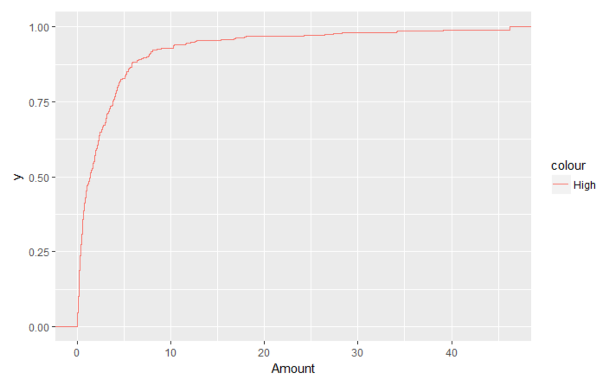
(c) Distress Amount of Medium Severity Level



(d) CDF of Medium Severity Level

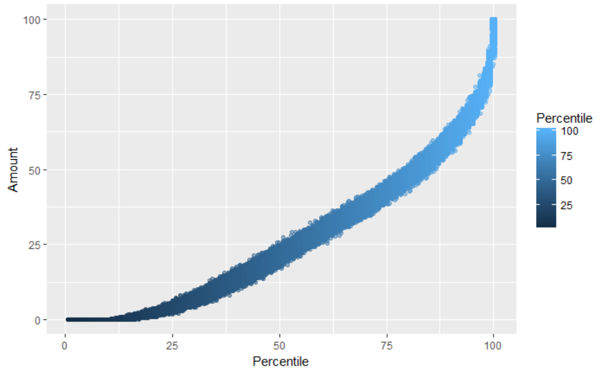


(e) Distress Amount of High Severity Level

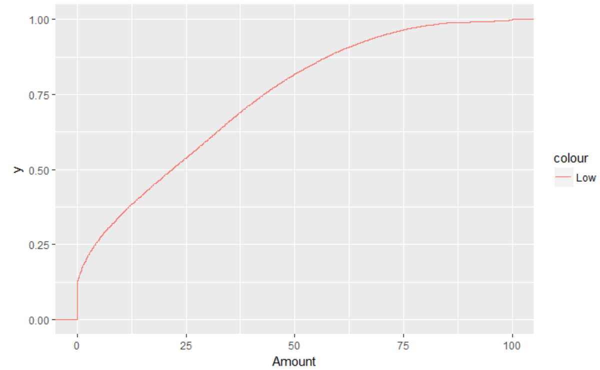


(f) CDF of High Severity Level

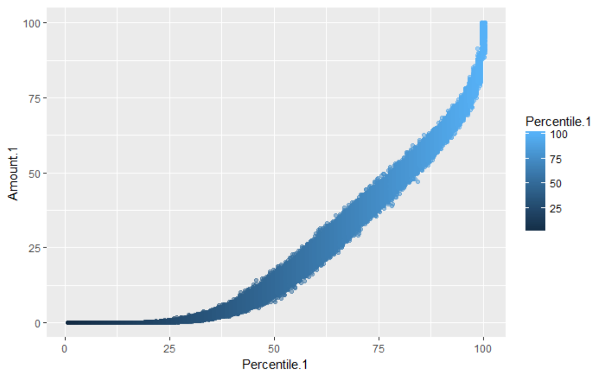
Figure 40 Cumulative Distribution Curves of Spalling of Transverse Joint in Concrete Pavements



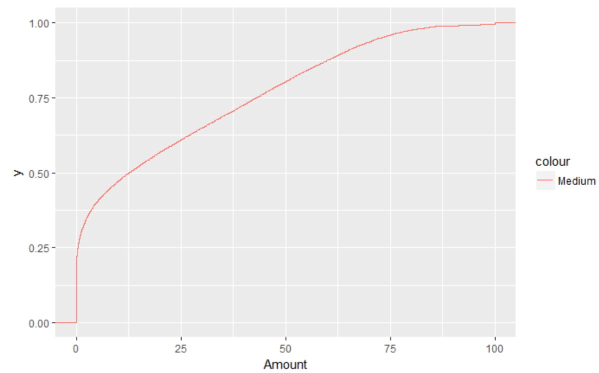
(a) Distress Amount of Low Severity Level



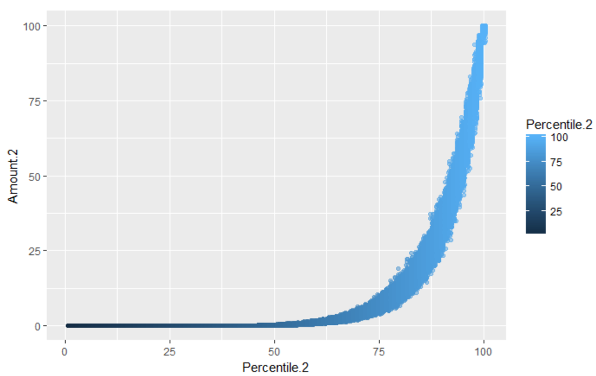
(b) CDF of Low Severity Level



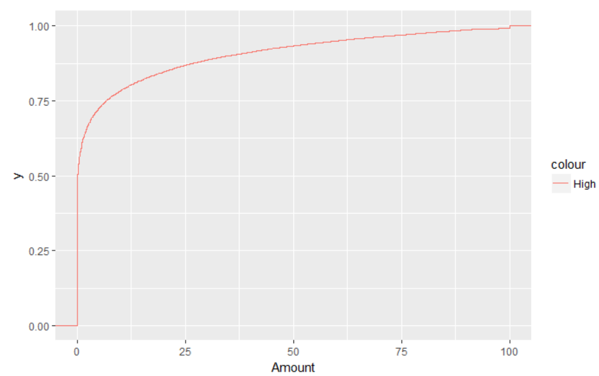
(c) Distress Amount of Medium Severity Level



(d) CDF of Medium Severity Level

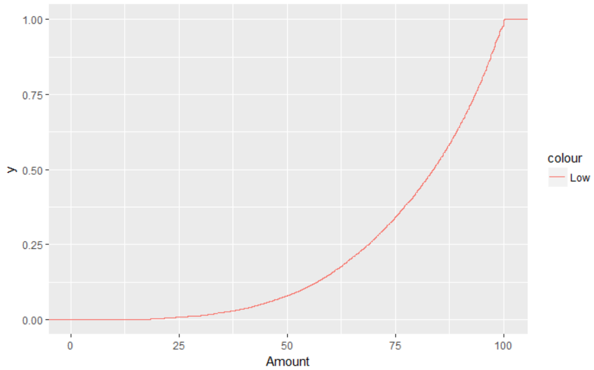


(e) Distress Amount of High Severity Level

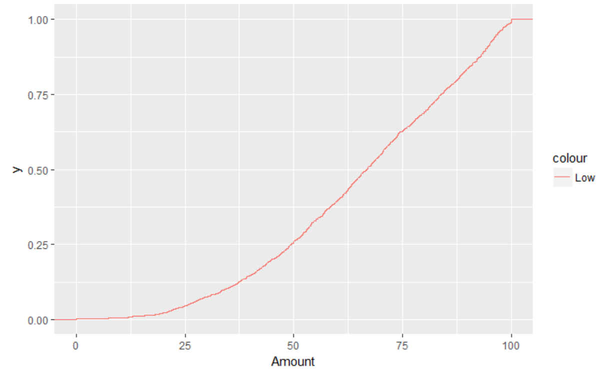


(f) CDF of High Severity Level

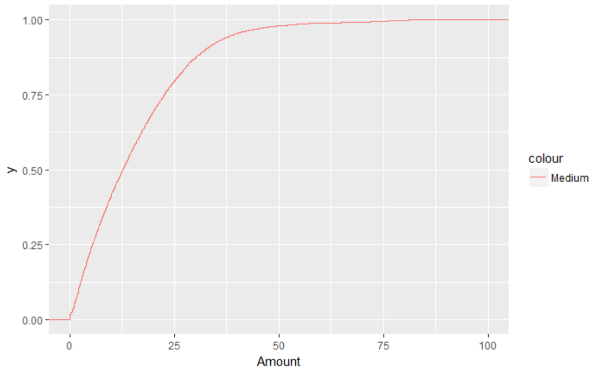
Figure 41 Cumulative Distribution Curves of Rutting Percentages in Asphalt Pavements



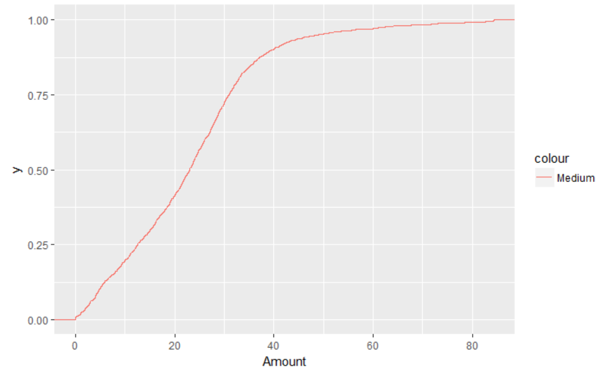
(a) Low Severity Level in Asphalt Pavement



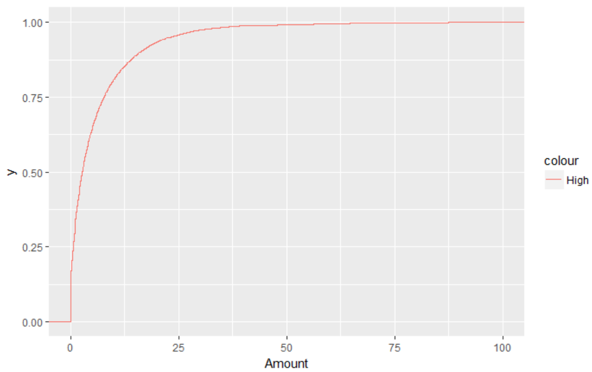
(b) Low Severity Level in Concrete Pavement



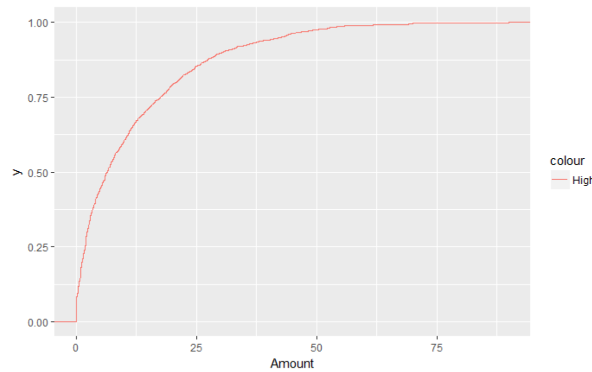
(c) Medium Severity Level in Asphalt Pavement



(d) Medium Severity Level in Concrete Pavement



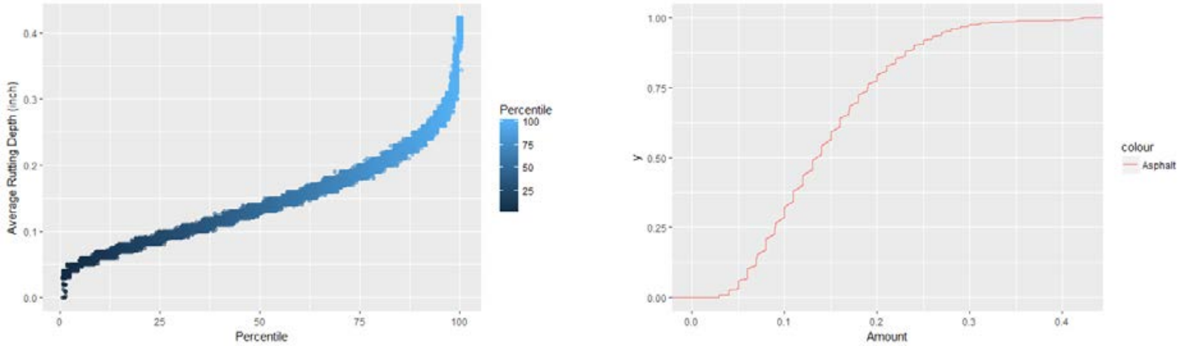
(e) High Severity Level in Asphalt Pavement



(f) High Severity Level in Concrete Pavement

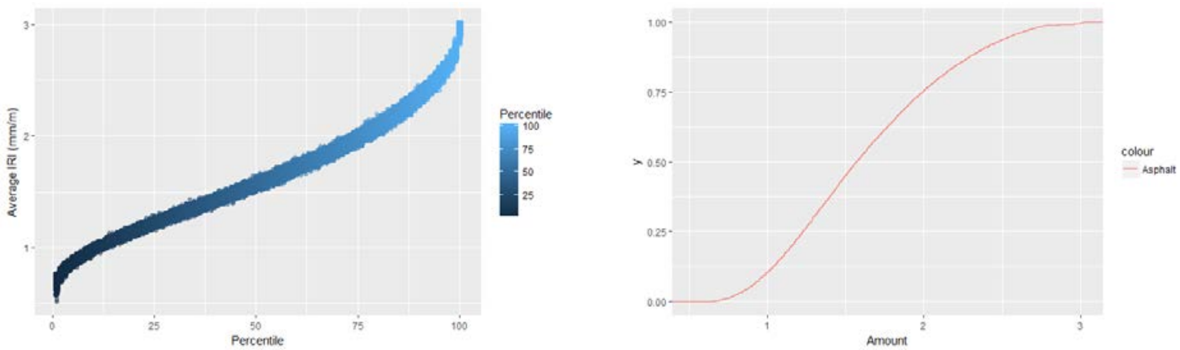
Figure 42 CDF's of IRI Percentages in Asphalt and Concrete Pavements

Figure 41 presents the resampled plots and CDF curves of rutting in percentages for low, medium, and high categories in asphalt pavements. Figure 42 performs the resampled CDF curves of IRI in percentages for low, medium, and high categories in asphalt and concrete pavements. Figure 43 and Figure 44 show the plots of the resampled distributions and CDF curves for average rutting depth and average IRI in non-warranty pavements.

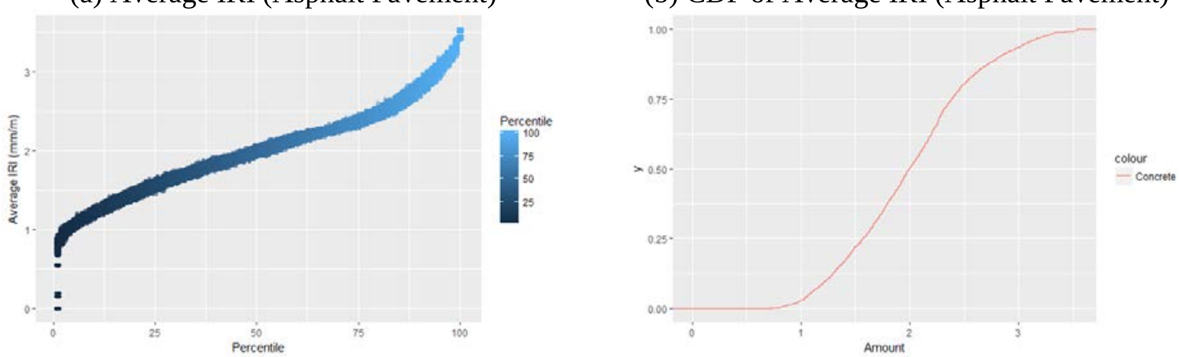


(a) Average Rutting Depth (Asphalt Pavement) (b) CDF of Average Rutting Depth (Asphalt Pavement)

Figure 43 Cumulative Distributions of Average Rutting Depth



(a) Average IRI (Asphalt Pavement) (b) CDF of Average IRI (Asphalt Pavement)



(c) Average IRI (Concrete Pavement) (d) CDF of Average IRI (Concrete Pavement)

Figure 44 Cumulative Distributions of Average IRI

6.1.2 Best-Fit Method

6.1.2.1 Best-Fit Consideration in Updating Warranty Program

The other method relies on both the bootstrapping normalized distribution of the non-warranty data in PMS and the warranty annual maintained pavement distress reports. The invalid records and outliers of the non-warranty distress data were cleansed using the bootstrapping method. The normalized non-warranty distress data and the CDF distributions over the 100 percentile intervals from 0-100% were generated to compute the sample means associated with the percentile intervals at three severity levels for each distress type. Thus, there are 100 sets of means for each distress type at three severity levels. Each of the 100 sets of sample means is associated with a percentile and includes three specific distress threshold values for the individual distress type at the three severity levels respectively.

The distress survey data and remedial maintenance decisions for each distress type at three severity levels recorded in the warranty annual maintained pavement distress reports were employed to check with each of the 100 sets of distress thresholds. In accordance with each threshold, a maintenance decision of "YES" or "NO" would be made for the distress measurement recorded for the pavement segment to determine whether or not the pavement segment would need a remedial action for the distress type at a certain severity level. Comparing the "predicted" maintenance decisions with the "actual" maintenance decisions recorded in the warranty project reports, an accuracy of the number of consistent maintenance decisions could be obtained to evaluate the "equivalency" of the new threshold to the existing threshold. The "equivalency" between the existing and the new thresholds is important for the smooth or "seamless" transition in updating the pavement warranty program of a state highway agency.

6.1.2.2 Confusion Matrix Analysis -- Longitudinal Cracking

Taking longitudinal cracking as an example, the best-fit procedure includes the following steps:

1) Compute the mean values of the empirical samples bootstrapped from the raw longitudinal cracking data in a generic percentile interval of 49% to 50%. The set includes 145.05 ft, 43.49 ft, and 4.67 ft per 1/10 mile for low, medium, and high severity levels, respectively.

2) Assume the total crack lengths of 145.05 ft, 43.49 ft, and 4.67 ft per 1/10 mile as new distress measurement based thresholds for longitudinal cracking at three severity levels for the warranty pavements.

3) Use the threshold values of 145.05 ft, 43.49 ft, and 4.67 ft per 1/10 mile to compare with the historical longitudinal cracking data at three severity levels recorded for warranty pavement segments in the annual warranty project maintenance reports. Make new maintenance decisions with the new threshold criteria, e.g., the segment maintenance decision is "YES" when at least one severity level maintenance decision is "YES." As shown in Table 34, there are 1,484 pavement segments recorded in warranty projects with longitudinal cracking. By applying the set of thresholds, 698 of the segments would have maintenance decision of "YES," while 786 "NO."

4) Develop a confusion matrix with the new maintenance decisions versus the actual maintenance decisions made by using the existing thresholds. The calculated accuracy is 0.551, which means 55.1% of the segments that would have the same maintenance decisions by using the new thresholds as the actual decisions with the existing deduct point based thresholds.

5) Repeat Steps 1 through 4 to calculate the confusion matrix accuracies for all percentile intervals over the 0-100% percentile range of the distress data to find the best-fit percentile interval with the highest accuracy.

Table 34 Confusion Matrix of Longitudinal Cracking Threshold (49%-50%)

n = 1484	Predicted: NO	Predicted: YES
Actual: NO	786	666
Actual: YES	0	32
Accuracy	55.1%	

6.1.3 Selection of Standard Deviation Multiplier (1.0-2.0)

As the standard deviation method assumes the pavement distress data follow a normal distribution, a state DOT is suggested to set the threshold of a distress type at 2.0 σ (where only 5% of measured sections would exceed the threshold) to reduce the risk to the contractor for a 10-year warranted pavement (Scott III et al. 2011). According to additional experience or for improved consistency, the threshold is suggested to be further tightened (to between 1.0 σ and 2.0 σ). However, the specific standard deviation multiplier (between 1.0 and 2.0) is still unknown. In order to answer this question, this study calculated 11 groups of distress measurement thresholds based on 11 different standard deviation multipliers (i.e., 1.0, 1.1, …, to 2.0) for longitudinal cracking, transverse cracking, alligator cracking, block cracking, and rutting in asphalt pavements. Confusion matrix method was applied to evaluate the deviations of the deduct point based thresholds and the 11 groups of measurement based thresholds for each individual distress type. The results for longitudinal cracking, transverse cracking, alligator cracking, and block cracking are presented in Table 35 to Table 38, and the results for rutting in percentages and average rutting depth are shown in Table 39 and Table 40.

As shown in Table 35, the degree of similarity in the maintenance decision making between the deduct point based thresholds and the 11 groups of measurement based thresholds increase with the increase in the standard deviation multiplier initially, and slightly decrease afterward. The highest accuracy appears in the multiplier range from 1.4 to 1.8, where the percentiles of the normalized distribution are from 84% to 93%. The highest accuracy could reach 99.1%, which means 99.1% of the recorded segments would be given the same maintenance decision with the measurement based thresholds as the deduct point based thresholds at three severity levels.

Table 35 Standard Deviation Multipliers and Thresholds for Longitudinal Cracking in Asphalt Pavements

Multiplier	Percentile	Amount / (ft. per 1/10 mile)			Accuracy
		Low	Medium	High	

1.0	0.68	287.37	92.32	10.21	0.821
1.1	0.73	348.98	112.10	12.25	0.890
1.2	0.77	410.72	134.60	14.83	0.937
1.3	0.81	478.13	159.95	17.33	0.966
1.4	0.84	548.71	188.66	20.12	0.984
1.5	0.87	620.46	216.78	23.78	0.989
1.6	0.89	699.84	248.08	26.55	0.991
1.7	0.91	781.63	278.33	30.77	0.989
1.8	0.93	860.86	311.40	34.77	0.987
1.9	0.94	942.08	343.14	39.22	0.982
2.0	0.95	1024.00	373.23	43.86	0.982

Table 36 Standard Deviation Multipliers and Thresholds for Transverse Cracking in Asphalt Pavements

Multiplier	Percentile	Amount / (ft. per 1/10 mile)			Accuracy
		Low	Medium	High	
1.0	0.68	294.63	63.24	116.39	0.936
1.1	0.73	361.63	78.49	20.63	0.879
1.2	0.77	435.95	95.56	25.43	0.840
1.3	0.81	516.50	114.26	30.85	0.816
1.4	0.84	622.60	134.90	37.25	0.810
1.5	0.87	757.68	155.08	44.59	0.809
1.6	0.89	906.44	178.82	53.01	0.809
1.7	0.91	1039.12	205.90	62.49	0.809
1.8	0.93	1213.76	235.25	73.28	0.809
1.9	0.94	1365.42	269.66	83.58	0.809
2.0	0.95	1545.98	301.72	92.68	0.809

The thresholds for transverse cracking in asphalt pavements in Table 36 show that the degree of similarity in the maintenance decisions made by the two kinds of thresholds continually decreases with the increase in standard deviation multiplier. The highest accuracy appears in the standard deviation multiplier of 1.0, where the percentile of the normalized distribution is 68%. The highest accuracy could reach 93.6%, which means 93.6% of the recorded segments would be given the same maintenance decision with the measurement based thresholds as the deduct point based thresholds at three severity levels.

The results of alligator cracking in Table 37 are similar to the results of longitudinal cracking. The accuracy increases with the increase in standard deviation multiplier first, and then slightly decreases afterward. The highest accuracy appears in the range of 1.6 σ to 1.9 σ, where the percentiles of the normalized distribution are from 89% to 94%. At most, 89.5% of the recorded segments would be given the same maintenance decision with the measurement based thresholds as the deduct point based thresholds at three severity levels.

Table 37 Standard Deviation Multipliers and Thresholds for Alligator Cracking in Asphalt Pavements

Multiplier	Percentile	Amount / (sq.ft. per 1/10 mile)			Accuracy
		Low	Medium	High	
1.0	0.68	115.97	56.83	47.68	0.666
1.1	0.73	140.65	74.46	65.29	0.720
1.2	0.77	164.34	94.02	85.30	0.764
1.3	0.81	192.31	116.03	112.65	0.804
1.4	0.84	220.57	147.65	141.45	0.824
1.5	0.87	250.13	180.72	183.62	0.841
1.6	0.89	283.89	217.71	245.75	0.878
1.7	0.91	318.93	254.34	301.94	0.878
1.8	0.93	357.12	290.13	387.49	0.895
1.9	0.94	391.58	332.01	483.42	0.889
2.0	0.95	430.75	365.84	550.15	0.872

Table 38 Standard Deviation Multipliers and Thresholds for Block Cracking in Asphalt Pavements

Multiplier	Percentile	Amount / (sq.ft. per 1/10 mile)			Accuracy
		Low	Medium	High	
1.0	0.68	422.28	126.59	45.25	0.819
1.1	0.73	535.19	159.15	68.24	0.855
1.2	0.77	671.27	199.83	83.95	0.892
1.3	0.81	826.78	237.87	108.99	0.944
1.4	0.84	964.43	278.05	142.51	0.984
1.5	0.87	1082.93	324.66	177.83	0.932
1.6	0.89	1190.48	374.51	241.86	0.904
1.7	0.91	1288.88	424.73	293.74	0.876
1.8	0.93	1326.67	491.14	383.46	0.867
1.9	0.94	1527.35	573.33	466.45	0.807
2.0	0.95	1672.56	681.83	579.14	0.775

The results of block cracking in Table 38 is similar to the results of longitudinal cracking and alligator cracking. The accuracy increases with the increase in standard deviation multiplier first, and slightly decreases afterward. The highest accuracy appears in the multiplier range of 1.3 to 1.5, where the percentiles of the normalized distribution are from 81% to 87%. At most, 98.4% of the recorded segments would be given the same maintenance decision with the measurement based thresholds as the deduct point based thresholds at three severity levels.

Table 39 Standard Deviation Multipliers and Thresholds for Rutting Percentage in Asphalt

Pavements

Multiplier	Percentile	Amount / percentage			Accuracy
		Low	Medium	High	
1.0	0.68	39.89	43.05	7.99	0.936
1.1	0.73	43.47	47.75	11.88	0.933
1.2	0.77	47.12	52.00	16.78	0.930
1.3	0.81	50.82	55.86	22.30	0.930
1.4	0.84	54.39	59.59	28.58	0.926
1.5	0.87	57.82	62.84	35.72	0.924
1.6	0.89	61.01	65.78	42.98	0.924
1.7	0.91	64.11	68.60	51.00	0.923
1.8	0.93	67.22	71.01	58.17	0.922
1.9	0.94	70.00	73.80	65.30	0.922
2.0	0.95	72.58	76.00	71.94	0.922

Table 40 Standard Deviation Multipliers and Thresholds for Average Rutting Depth in Asphalt Pavements

Multiplier	Percentile	Average Rutting Depth (inch per 1/10 mile)	Accuracy
1.0	0.68	0.17	0.862
1.1	0.73	0.18	0.909
1.2	0.77	0.20	0.937
1.3	0.81	0.21	0.945
1.4	0.84	0.22	0.945
1.5	0.87	0.23	0.941
1.6	0.89	0.24	0.931
1.7	0.91	0.25	0.929
1.8	0.93	0.26	0.926
1.9	0.94	0.27	0.924
2.0	0.95	0.28	0.923

Both the rutting in percentages and the average rutting depth are considered as potential distress items in the warranty program. The accuracy results for rutting in percentage are presented in Table 39. The accuracy continually decreases with the increase in standard deviation multiplier. The highest accuracy appears in the standard deviation of 1.0 σ, where the percentile of the normalized distribution for percentage is 68%. The highest accuracy could reach 93.6%, which means 93.6% of the recorded segments would be given the same maintenance decision with the measurement based thresholds as the deduct point based thresholds at three severity levels. The results for average rutting depth are shown in Table 40. The accuracy increases with the increase in standard deviation multiplier first, and slightly decreases afterward. The highest accuracy appears in the standard deviation range from 1.2 σ to 1.5 σ, where the percentiles of the normalized distribution are from 77% to 87%. At most, 94.5% of the recorded segments would be given the same maintenance decision with the measurement based thresholds as the deduct point based thresholds at three severity levels.

Overall, the standard deviation threshold and accuracy results of the longitudinal cracking, alligator cracking, block cracking, and average rutting depth share the similar characteristics that the accuracy increases with the increase of the standard deviation multiplier initially, and gradually decreases. In addition, the highest accuracy appears in the range of 1.3 to 1.7, and the distress measurement thresholds at 1.5 σ could be the most representing for the four warranty distress items. However, the transverse cracking and rutting percentage may differ from the other distress items in terms of the best multiplier value for the accuracy of maintenance decisions. Hence, it would be a reasonable option to set the new measurement based distress thresholds by adding 1.5 σ to the sample means of the distress measurement data. However, the distress thresholds set by adding 1.0 σ to the sample means of the distress measurement data were also administrated to compare with the 1.5 σ thresholds in this study.

6.1.4 Best-Fit Accuracy

Following the above 5 steps, the curves of accuracy versus the distress percentile for longitudinal cracking, transverse cracking, alligator cracking, and block cracking are shown from Figure 45 to Figure 48. The curves for rutting in percentages and average rutting depth are presented in Figure 49 and Figure 50. The figures display that as the percentile increases, the accuracy increases first, then decreases, and hence, there is a peak with the highest accuracy in every distress type. The results imply that there exists a best-fit percentile and distress thresholds associated with the best-fit percentile in every distress type. Adopting the distress thresholds at the best-fit percentile would allow pavement engineers to use the new thresholds to make the most equivalent maintenance decisions to the actual decisions in the warranty reports for the distress type. Therefore, the percentile at the highest accuracy can be employed to find the best-fit measurement based thresholds for each distress type at the three severity levels.

(1) Accuracy of Threshold – Longitudinal Cracking

The curve of the accuracy of maintenance decisions along with distribution percentile (equivalent to the standard deviation multiplier) of threshold for longitudinal cracking is shown in Figure 45. With the increase of the percentile, the accuracy increases, and the highest accuracy, 99.12%, occurs at the percentile 89%, which is also listed in Table 35. The comparison result of maintenance decisions between the deduct point based thresholds and best-fit measurement based thresholds is listed in the confusion matrix in Table 41. The result shows that there are 1,484 valid segment records in the selected warranty projects. 21 of them make decisions of “YES” (remedial actions needed) with both of the deduct point based thresholds and measurement based thresholds; 1,450 of them make decisions of “NO” (remedial actions not needed) with both of the two sets of thresholds; 2 of them make decisions of “YES” with the deduct point based thresholds, but that of “NO” with the measurement based thresholds; and 11 of them make decisions of “NO” with the deduct point based thresholds, but that of “YES” with the measurement based thresholds. Hence, there are 1,471 records that make the same decisions with both of the two sets of thresholds, while 13 of them make the different decisions. The percentage of the same decisions in the total records is 99.12%, which is identified as accuracy value in the confusion matrix.

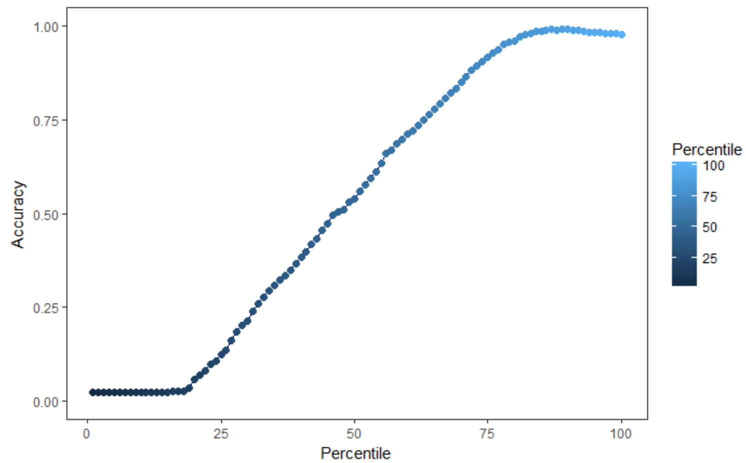


Figure 45 Accuracy of Longitudinal Cracking Thresholds in Asphalt Pavements

Table 41 Confusion Matrix of Longitudinal Cracking Threshold with Highest Accuracy

n = 1484	Predicted: NO	Predicted: YES
Actual: NO	1450	11
Actual: YES	2	21
Accuracy	99.1%	

(2) Accuracy of Threshold – Transverse Cracking

The curve of the accuracy of maintenance decisions along with distribution percentile of transverse cracking threshold is shown in Figure 46. With the increase of the percentile, the accuracy increases, and the highest accuracy, 99.44%, occurs at the percentile of 63%. The comparison result of maintenance decisions between the deduct point based thresholds and best-fit measurement based thresholds is listed in the confusion matrix in Table 42. The result shows that there are 1,240 valid segment records in the selected warranty projects. 236 of them make decisions of “YES” (remedial actions needed) with both of the deduct point based thresholds and measurement based thresholds; 997 of them make decisions of “NO” (remedial actions not needed” with both of the two sets of thresholds; 6 of them make decisions of “YES” with the deduct point based thresholds, but that of “NO” with the measurement based thresholds; and 1 of them makes decision of “NO” with the deduct point based thresholds, but that of “YES” with the measurement based thresholds. Hence, there are 1,233 records that make same decisions with both of the two sets of thresholds, while 7 of them make different decisions. The percentage of the same decisions in the total records is 99.4%, which is identified as accuracy value in the confusion matrix.

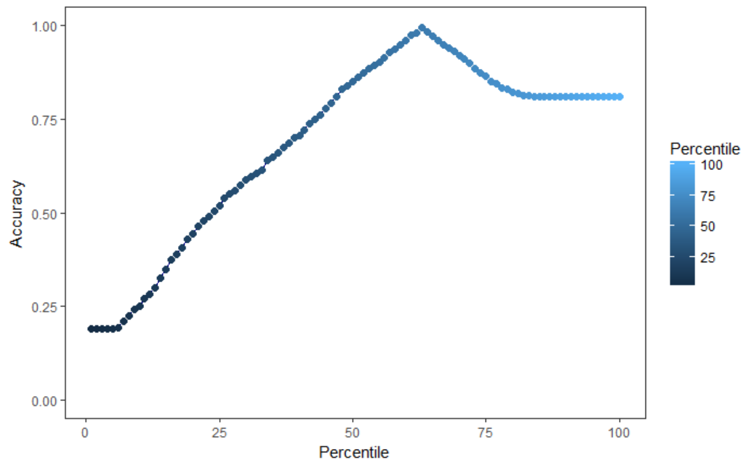


Figure 46 Accuracy of Transverse Cracking Thresholds in Asphalt Pavements

Table 42 Confusion Matrix of Transverse Cracking Threshold with Highest Accuracy

n = 1240	Predicted: NO	Predicted: YES
Actual: NO	997	1
Actual: YES	6	236
Accuracy	99.4%	

(3) Accuracy of Threshold – Alligator Cracking

The curve of the accuracy of maintenance decisions along with distribution percentile of alligator cracking threshold is shown in Figure 47. The highest accuracy, 89.5%, occurs at the percentile of 93%. The comparison result of maintenance decisions between the deduct point based thresholds and best-fit measurement based thresholds is listed in the confusion matrix in Table 43. The result shows that there are 296 valid segment records in the selected warranty projects; 35 of them make decisions of "YES" (remedial actions needed) with both of the deduct point based thresholds and measurement based thresholds; 230 of them make decisions of "NO" (remedial actions not needed) with both of the two sets of thresholds; 1 of them makes decision of "YES" with the deduct point based thresholds, but that of "NO" with the measurement based thresholds; and 30 of them make decisions of "NO" with the deduct point based thresholds, but that of "YES" with the measurement based thresholds. Hence, there are 265 records that make same decisions with both of the two sets of thresholds, while 31 of them make different decisions. The percentage of the same decisions is 89.5%, which is identified as accuracy in the confusion matrix.

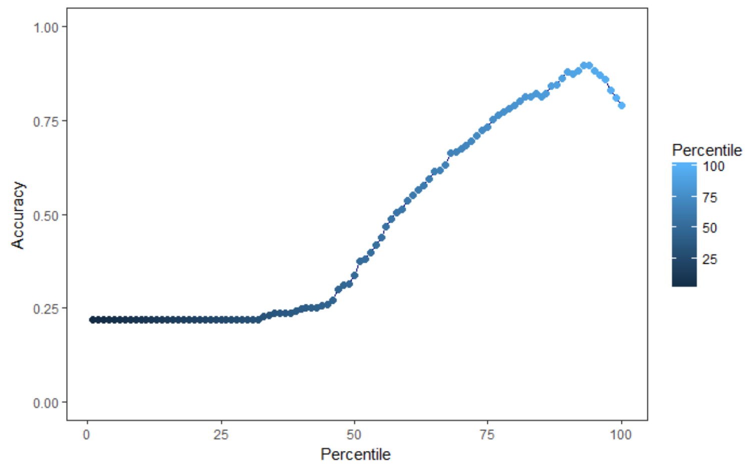


Figure 47 Accuracy of Alligator Cracking Thresholds in Asphalt Pavements

Table 43 Confusion Matrix of Alligator Cracking Threshold with Highest Accuracy

n = 296	Predicted: NO	Predicted: YES
Actual: NO	230	30
Actual: YES	1	35
Accuracy	89.53%	

(4) Accuracy of Threshold – Block Cracking

The curve of the accuracy of maintenance decisions along with distribution percentile of block cracking threshold is shown in Figure 48. The highest accuracy, 98.4%, occurs at the percentile of 84%. The comparison result of maintenance decisions between the deduct point based thresholds and best-fit measurement based thresholds is listed in the confusion matrix in Table 44. The result shows that there are 249 valid segment records in the selected warranty projects; 139 of them make decisions of “YES” (remedial actions needed) with both of the deduct point based thresholds and measurement based thresholds; 106 of them make decisions of “NO” (remedial actions not needed) with both of the two sets of thresholds; 2 of them make decisions of “YES” with the deduct point based thresholds, but that of “NO” with the measurement based thresholds; and 2 of them make decisions of “NO” with the deduct point based thresholds, but that of “YES” with the measurement based thresholds. Hence, there are 245 records that make same decisions by both of the two sets of thresholds, while 4 of them make different decisions. The percentage of same decisions is 98.4%, which is identified as accuracy in confusion matrix.

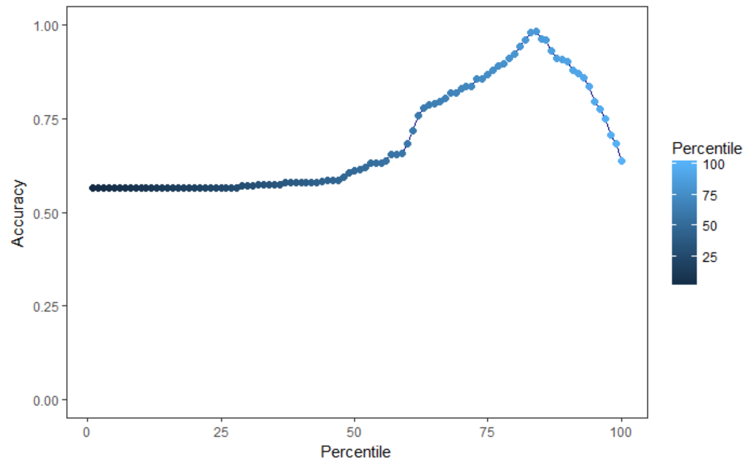


Figure 48 Accuracy of Block Cracking Thresholds in Asphalt Pavements

Table 44 Confusion Matrix of Block Cracking Threshold with Highest Accuracy

n = 249	Predicted: NO	Predicted: YES
Actual: NO	106	2
Actual: YES	2	139
Accuracy	98.4%	

(5) Accuracy of Threshold – Rutting in Percentage

The curve of the accuracy of maintenance decisions along with distribution percentile of thresholds for rutting in percentages is shown in Figure 49. The highest accuracy, 99.4%, occurs at the percentage 33%. The comparison result of maintenance decisions between the deduct point based thresholds and best-fit measurement based thresholds is listed in the confusion matrix in Table 45. The result shows that there are 715 valid segment records in the selected warranty projects; 55 of them make decisions of "YES" (remedial actions needed) with both of the deduct point based thresholds and measurement based thresholds; 656 of them make decisions of "NO" (remedial actions not needed) with both of the two sets of thresholds; 1 of them makes decision of "YES" with the deduct point based thresholds, but that of "NO" with the measurement based thresholds; and 3 of them make decisions of "NO" with the deduct point based thresholds, but that of "YES" with the measurement based thresholds. Hence, there are 711 records that make same decisions with both of the two sets of thresholds, while 4 of them make different decisions. The percentage of same decisions is 99.4%, which is identified as accuracy in the confusion matrix.

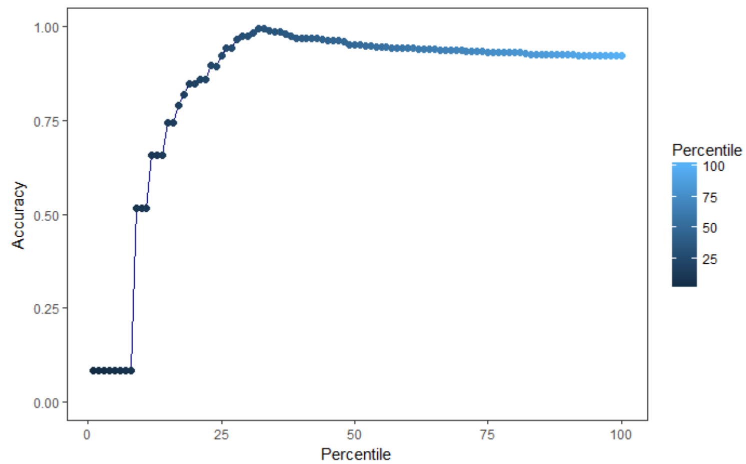


Figure 49 Accuracy of Rutting Percentage Thresholds in Asphalt Pavement

Table 45 Confusion Matrix of Rutting Percentage Threshold with Highest Accuracy

n = 715	Predicted: NO	Predicted: YES
Actual: NO	656	3
Actual: YES	1	55
Accuracy	99.4%	

(6) Accuracy of Threshold – Average Rutting Depth

The curve of the accuracy of maintenance decisions along with distribution percentile of thresholds for average rutting depth is shown in Figure 50. The highest accuracy, 94.8%, occurs at the percentage 83%. The comparison result of maintenance decisions between the deduct point based thresholds and best-fit measurement based thresholds is listed in the confusion matrix in Table 46. The result shows that there are 715 valid segment records in the selected warranty projects; 21 of them make decisions of “YES” (remedial actions needed) with both of the deduct point based thresholds and measurement based thresholds; 657 of them make decisions of “NO” (remedial actions not needed) with both of the two sets of thresholds; 35 of them make decisions of “YES” with the deduct point based thresholds, but that of “NO” with the measurement based thresholds; and 2 of them make decisions of “NO” with the deduct point based thresholds, but that of “YES” with the measurement based thresholds. Hence, there are 678 records that make same decisions with both of the two sets of thresholds, while 37 of them make different decisions. The percentage of same decisions is 94.8%, which is identified as accuracy in the confusion matrix.

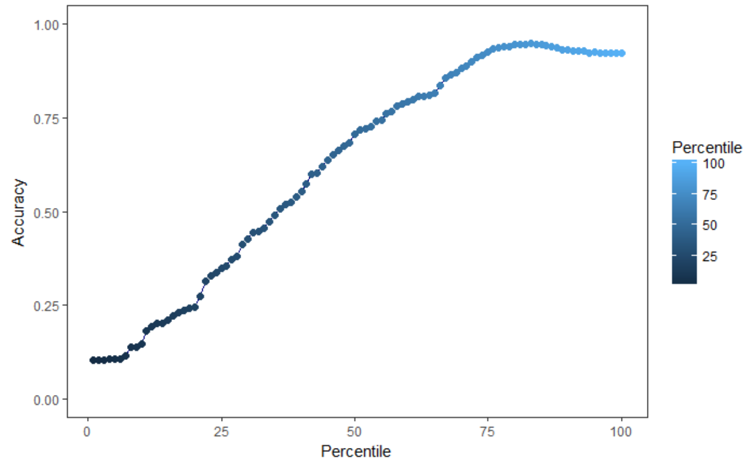


Figure 50 Accuracy of Average Rutting Depth Thresholds in Asphalt Pavement

Table 46 Confusion Matrix of Average Rutting Depth Threshold with Highest Accuracy

n = 715	Predicted: NO	Predicted: YES
Actual: NO	657	2
Actual: YES	35	21
Accuracy	94.8%	

6.2 Selection of Warranty Threshold Method

6.2.1 Thresholds Based on Best-Fit Method

The distress thresholds of the best-fit method with the existing practice for the asphalt pavements are listed in Table 47. The best-fit thresholds appear at different percentiles of the four distress types, which are at 93%, 84%, 89%, and 63% for alligator cracking, block cracking, longitudinal cracking, and transverse cracking, respectively. The highest accuracies also reach to different values for the four distress types. The accuracies of the distress types are all larger than 80%, and the accuracies of the longitudinal cracking and transverse cracking reach the values of larger than 99%. In addition, the accuracy of block cracking is at high percentage value of 98.4%. However, the results of alligator cracking and block cracking may not be as accurate as those of longitudinal cracking and transverse cracking. The data sizes of alligator cracking and block cracking in the warranty reports are relatively small with only 296 and 249 records, respectively. Obviously, using the best-fit method to set up the distress measurement threshold requires enough valid data from the warranty reports.

Table 47 Thresholds Fitting with Existing Practice for Asphalt Pavements

Distress #	Distress Type	Distress level	# of data in warranty	Percentile/%	Distress amount/(1/10 mile)	Accuracy
19	Alligator Cracking/sq.ft.	low	296	93	357.12	0.895
		medium			290.13	
		high			387.49	
20	Block Cracking/sq.ft.	low	249	84	964.43	0.984
		medium			278.05	
		high			142.51	
22	Longitudinal Cracking/ft.	low	1484	89	699.84	0.991
		medium			248.08	
		high			26.55	
23	Transverse Cracking/ft.	low	1240	63	237.27	0.994
		medium			49.26	
		high			12.32	
200	Rutting/percentage	low	715	33	13.76	0.994
		medium			8.50	
		high			0.004	
200	Rutting/average rutting depth (in)	--	715	83	0.21	0.948

In the warranty reports, both the rutting percentages in low, medium, and high categories and the average rutting depth are surveyed in every road segment. Therefore, there are two kinds of analyses for rutting, one with rutting in percentage, and the other with average rutting depth. For rutting in percentage, the existing maintenance decisions recorded in the warranty reports are made by comparing with the deduct point based thresholds and the deduct point values calculated from the surveyed rutting in percentages using the empirical equations. The predicting maintenance decisions are made by comparing with the measurement based thresholds at the best percentile of the normalized rutting percentages in the three severity categories of the non-warranty pavements and the surveyed rutting in percentages recorded in the warranty reports. In the average rutting depth, the applied maintenance decisions are the same as the rutting in percentage, while the predicted maintenance decisions are different. The predicted maintenance decisions for the average rutting depths are made by comparing the measurement based thresholds at the best-fit percentile of the normalized average rutting depth of the non-warranty pavements with the surveyed average rutting depth for every road segments recorded in the warranty reports. Under this circumstance, although the data sizes are the same, the best-fit distress thresholds for rutting are different, and the accuracy differs as well. For rutting in percentages, the highest accuracy is 99.4% when setting the measurement based thresholds at percentile 33% of normalized rutting percentages in three categories of the non-warranty pavements. The highest accuracy in the average rutting depth analysis is 94.8% with the measurement based thresholds set at percentile 83% of the normalized average rutting depth of the non-warranty pavements.

6.2.2 Thresholds Based on Standard Deviation Method

Based on the method of adding 1.0 σ or 1.5 σ to the sample means of the normalized distress measurement data, the distress measurement based thresholds for asphalt pavements are listed in Table 48 and Table 49, and similarly the thresholds for concrete pavements are shown in Table 50 and Table 51. In Table 48, the thresholds were developed by adding 1.0 σ to the sample means of the normalized distress measurement data for all of the nine distress types at three severity levels (except for the pothole measured as the number of counts). However, the accuracy of comparing the existing maintenance decisions and predicted maintenance decisions using the thresholds listed in Table 48 was only calculated for alligator cracking, block cracking, longitudinal cracking, transverse cracking, rutting in percentages, and average rutting depth. It was because there were not valid data recorded or all the values were zeros in the warranty reports for asphalt pavements. The situation also occurred in adding 1.5 σ to the sample means of the normalized distress measurement data creating the measurement based thresholds for asphalt pavements in Table 49. For the concrete pavement thresholds based on adding 1.0 σ to the sample means of the normalized distress measurement data presented in Table 50, not all of the thresholds in the eight distress types could be calculated. It was because there were not enough non-warranty data or the values were all zeros for faulting of transverse joints, corner breaks, joint seal deterioration, and map cracking in the PMS. In addition, the accuracy analyses could not be administrated due to insufficient or invalid data recorded in the warranty reports for concrete pavements. This situation also occurred to the 1.5 σ based thresholds of concrete pavements in Table 51.

Table 48 Standard Deviation (1.0 σ) Thresholds for Asphalt Pavements

Distress #	Distress Type	Distress level	# of data in warranty	Distress amount/(1/10 mile)	Accuracy
19	Alligator Cracking/sq.ft.	low	296	115.97	0.666
		medium		56.83	
		high		47.68	
20	Block Cracking/sq.ft.	low	249	422.28	0.819
		medium		126.59	
		high		45.25	
21	Edge Cracking/ft.	low	0	45.78	NA
		medium		48.17	
		high		10.94	
22	Longitudinal Cracking/ft.	low	1484	287.37	0.821
		medium		92.32	
		high		10.21	
23	Transverse Cracking/ft.	low	1240	294.63	0.936
		medium		63.24	
		high		16.39	
27	Potholes/no.		0	16.00	NA
29	Bleeding/sq.ft.	low	0	291.75	NA
		medium		86.00	
		high		14.30	
30	Reflection Cracking	low	0	173.96	NA
		medium		98.66	
		high		5.07	
200	Rutting/percentage	low	715	39.89	0.936
		medium		43.05	
		high		7.99	
200	Rutting/average rutting depth (in)		715	0.17	0.862

Table 49 Standard Deviation (1.5 σ) Thresholds for Asphalt Pavements

Distress #	Distress Type	Distress level	# of data in warranty	Distress amount/(1/10 mile)	Accuracy
19	Alligator Cracking/sq.ft.	low	296	250.13	0.841
		medium		180.72	
		high		183.62	
20	Block Cracking/sq.ft.	low	249	1082.93	0.932
		medium		324.66	
		high		177.83	
21	Edge Cracking/ft.	low	0	123.15	NA
		medium		152.50	
		high		26.19	
22	Longitudinal Cracking/ft.	low	1484	620.46	0.989
		medium		216.78	
		high		23.78	
23	Transverse Cracking/ft.	low	1240	757.68	0.809
		medium		155.08	
		high		44.59	
27	Potholes/no.		0	16.00	NA
29	Bleeding/sq.ft.	low	0	783.62	NA
		medium		319.91	
		high		52.07	
30	Reflection Cracking	low	0	619.06	NA
		medium		214.92	
		high		18.47	
200	Rutting/percentage	low	715	57.82	0.924
		medium		62.84	
		high		35.72	
200	Rutting/average rutting depth (in)		715	0.23	0.941

Table 50 Standard Deviation (1.0 σ) Thresholds for Concrete Pavements

Distress #	Distress Type	Distress level	Distress amount/(1/10 mile)
0	Longitudinal Cracking/ft.	low	66.12
		medium	35.58
		high	14.41
1	Transverse Cracking/ft.	low	37.95
		medium	60.16
		high	36.84
5	Faulting of Transverse Joints	low	NA
		medium	
		high	
7	Corner Breaks/no.	low	NA
		medium	
		high	
9	Joint Seal Deterioration/no.	low	NA
		medium	
		high	
10	Spalling of Longitudinal Joint/in.	low	31.39
		medium	22.73
		high	9.39
11	Spalling of Transverse Joint/in.	low	15.78
		medium	9.36
		high	7.74
12	Map Cracking & Scaling/sq.ft.	low	60.86
		medium	69.07
		high	30.08

Table 51 Standard Deviation (1.5 σ) Thresholds for Concrete Pavements

Distress #	Distress Type	Distress level	Distress amount/(1/10 mile)
0	Longitudinal Cracking/ft.	low	95.76
		medium	52.72
		high	19.17
1	Transverse Cracking/ft.	low	124.87
		medium	98.65
		high	53.66
5	Faulting of Transverse Joints	low	NA
		medium	
		high	
7	Corner Breaks/no.	low	NA
		medium	
		high	
9	Joint Seal Deterioration/no.	low	NA
		medium	
		high	
10	Spalling of Longitudinal Joint/in.	low	147.33
		medium	94.00
		high	26.90
11	Spalling of Transverse Joint/in.	low	23.67
		medium	11.98
		high	10.74
12	Map Cracking & Scaling/sq.ft.	low	91.81
		medium	93.92
		high	30.08

6.2.3 Comparison of Thresholds

If the new distress thresholds are determined based on the method of adding 1.0 σ or 1.5 σ to the sample means of the normalized distress measurement data as suggested by NCHRP Report 699, maintenance decisions with the new thresholds could be checked against the distress data and the actual maintenance decisions in the annual warranty maintenance reports, and the accuracy of making same decisions by using the new and old thresholds are calculated and listed in Table 52. The comparisons show that the accuracy of the 1.5 σ thresholds for alligator cracking, block cracking, longitudinal cracking, and average rutting depth is higher than that of the 1.0 σ thresholds. The comparison result may differ in transverse cracking and rutting in percentage, but overall, the 1.5 σ thresholds should be more appropriate than the 1.0 σ thresholds if a smooth transition of the existing warranty practice is to be considered.

The accuracy of maintenance decisions using the new thresholds is clearly influenced by the percentile interval of the distress distribution as shown in Figure 45 to Figure 50. The peaks of the accuracy curves represent the highest accuracy with the assumed best-fit sets of distress thresholds for the four distress types. The highest accuracy and its corresponding percentile intervals of the best-fit thresholds are also listed in Table 52. The accuracy of the best-fit thresholds is slightly higher than that of the 1.5 σ thresholds for alligator cracking, block cracking, longitudinal cracking, rutting in percentage, and average rutting depth, while the accuracy of the best-fit thresholds is much higher than that of the 1.5 σ thresholds for transverse cracking.

For the alligator cracking and longitudinal cracking, the percentiles of the best-fit method (93% and 89%) are higher than those of the 1.0 σ and 1.5 σ methods (68% and 86.6%). The percentiles of the best-fit method in block cracking and average rutting depth are 84% and 83%, which are between the 1.0 σ and 1.5 σ methods. And the percentiles of the best-fit method for transverse cracking and rutting in percentage are lower than the 1.0 σ method, which only cover 63% and 33% of the distribution, respectively.

Table 52 Accuracy of Different Methods for Asphalt Pavements

Distress #	Distress Type	68% (1.0 σ)	86.6% (1.5 σ)	Best Fitted	
		Accuracy	Accuracy	Accuracy	Percentile
19	Alligator Cracking/sq.ft.	0.666	0.841	0.895	93%
20	Block Cracking/sq.ft.	0.819	0.932	0.984	84%
22	Longitudinal Cracking/ft.	0.821	0.989	0.991	89%
23	Transverse Cracking/ft.	0.936	0.809	0.994	63%
200	Rutting/percentage	0.940	0.924	0.994	33%
200	Rutting/average rutting depth (in)	0.862	0.941	0.948	83%

It has been reported that some state agencies plan to set the warranty distress thresholds at 2.0σ, which could help reduce the risk of failures to the contractors for a 10-year warranty pavement project. The additional experience or improved consistency suggested that the thresholds should be tightened (to between 1.0 σ and 2.0 σ) or the warranty should be extended (Scott III et al. 2011). The comparison of the two methods in this study has shown that the 1.5σ method would be a practical approach to rebuilding the new warranty thresholds in Mississippi based on the historical warranty and non-warranty pavement data.

6.2.4 Consideration of IRI Thresholds

The IRI data of both the IRI in percentages of severity levels and the average IRI for asphalt pavements and concrete pavements have been collected in the non-warranty pavements in biennial surveys, while only average IRI was recorded as condition data but without a threshold in the existing warranty program in Mississippi. Therefore, the standard deviation method was also applied to set up the IRI thresholds based on the 1.0 σ and 1.5 σ procedures, and the results are presented in Table 53. Compared with the IRI in percentage thresholds in asphalt and concrete pavements, both the 1.0 σ and 1.5 σ based thresholds in asphalt pavements are larger than the thresholds in concrete pavements at low severity level. The results are consistent to the observation that concrete pavements are likely to be rougher than asphalt pavements in driving.

Table 53 Standard Deviation (1.0 and 1.5 σ) Thresholds for IRI (Percentage and Average)

Distress Types	Pavement type	Distress level	Amount (1.0 σ)	Amount (1.5 σ)
IRI (percentage)	Asphalt	low	91.11	96.86
		medium	19.38	29.39
		high	5.80	13.35
	Concrete	low	79.16	92.1
		medium	28.68	36.55
		high	13.11	26.48
Average IRI (mm/m)	Asphalt	--	1.86 (114 in./mi.)	2.27 (140 in./mi.)
	Concrete	--	2.27 (140 in./mi.)	2.68 (164 in./mi.)

Table 54 Thresholds for IRI in FHWA Website

Performance Indicators	Pavement Type	Evaluation Section	Threshold Values	Transformed Threshold Values
IRI	Asphalt	100 m (325 ft)	1.3 m/km (80 in./mi.)	2.08 mm/m (128 in./mi.)
	Concrete	100 m (325 ft)	1.3 m/km (80 in./mi.)	2.08 mm/m (128 in./mi.)

The FHWA suggests the IRI be included as a performance indicator in a warranty program and provides recommended IRI thresholds for both asphalt and concrete pavements as shown in Table 54. The IRI threshold values are determined by using the PMS data, while the selected section length is 100 m (328 ft) or 0.06 mile. In order to compare the threshold value on the same basis, the section length is extended to 160 m (0.1 mile). The results show that the recommeneded IRI threshold values are at 2.08 mm/m in both asphalt and concrete pavements. Compared with the standard deviation based thresholds in Table 53, it shows that the recommended IRI thresholds are looser than the 1.0 σ thresholds but tighter than the 1.5 σ thresholds.

CHAPTER 7. IMPLEMENTATION OF NEW THRESHOLDS

With special consideration for a smooth transition from the old deduct point based system to the new direct measurement based system, three groups of the new thresholds based on maximum allowed quantity of the direct measurement for each associated distress type were developed in Chapter 6. The thresholds based on the 1.5 σ method with the normalized non-warranty distress data were suggested to be an alternative to rebuild the distress thresholds for the pavement warranty program. However, the thresholds based on 1.0 σ method and the best-fit method were also suggeted as other two threshold options. Therefore, the three distress threshold options were applied to set up the maintenance decision making model based on the Visual Basic programming language in the Microsoft environment.

7.1 Introduction of the Maintenance Decision Software

The maintenance decision model was developed in Visual Basic for Applications (VBA), which is an implementation of Microsoft's event-driven programming language Visual Basic 6 (VB6) and its associated integrated development environment (IDE). The intermediate code is then executed by a virtual machine (hosted by the host application). Codes written in VBA are compiled to Microsoft P-Code (packed code), a proprietary intermediate language, which dwells in the host applications (Access, Excel, Word, Outlook, and PowerPoint) stored as a separate stream in component object model (COM) Structured Storage files (e.g., .doc or .xls) independent of the document streams. VBA enables building user-defined functions, automating processes and accessing Windows API and other low-level functionality through dynamic-link libraries. Different from the VB6, VBA code is normally run within a host application, rather than as a standalone program, and it can control one application from another using object linking and embedding (OLE) automation. VBA is built into most Microsoft Office applications, including Office for Mac OS X (except version 2008), and other Microsoft applications, including Microsoft MapPoint and Microsoft Visio. VBA is also implemented, at least partially, in applications published by companies other than Microsoft, including ArcGIS, AutoCAD, CorelDraw, LibreOffice, Reflection, SolidWorks, and WordPerfect.

In this research, the pavement warranty maintenance decision software was developed in VBA in Microsoft Excel, which can automatically create a Microsoft Excel report from Microsoft Excel data. The main procedures to make maintenance decision with the new measurement based distress threshold are 1) pavement type identification; 2) distress type and threshold option selection; 3) data procession; and 4) maintenance decision report development.

7.2 Administration of Maintenance Decision with New Thresholds

7.2.1 Data Requirement

In order to administer the maintenance decision software, the warranty survey data should be loaded to the software following the format as the examples in Table 55 and Table 56 with default remedial action at “No”. The survey data include the beginning log mile, ending log mile, road direction, lane position, average rutting depth, average IRI, and distress amount at each severity level for all the distress types. The format of the pavement surface distress data in the tables follows the data format in the annual maintained pavement distress report, and all the distress data should be summarized into segment (0.1 mi.) based. The distress type, severity level and amount should be entered into the first, second and third column separately. For example, the engineer should fill in the form with the warranty project survey data in 0.1-mile segment as in Table 55 (the recorded data in the table are assumed distress values).

Table 55 Asphalt Pavement Survey Record Form

Beginning Log Mile:	20.439		
Ending Log Mile:	20.534		
Direction:	N		
Lane:	Outside		
Average Rut Depth (in):	11		
Average IRI (in/mi):	11111		
Distresses	**Severity**	**Amount**	**Remedial Action**
Alligator Cracking	Low	1000	
	Medium	11110	
	High	0	
		Total:	N
Bleeding	Low	1111	
	Medium	0	
	High	0	
		Total:	N
Block Cracking	Low	111110	
	Medium	0	
	High	0	
		Total:	N
Edge Cracking	Low	11111	
	Medium	0	
	High	0	
		Total:	N
Longitudinal Cracking	Low	11111	
	Medium	0	
	High	0	
		Total:	N
Potholes	Low	11111	
		Total:	N
Raveling & Weathering	Low	11111	
	Medium	0	
	High	0	
		Total:	N
Reflection Cracking	Low	11111	
	Medium	0	
	High	0	
		Total:	N
Rutting	Low	11111	
	Medium	0	
	High	0	
		Total:	N
Transverse Cracking	Low	11111	
	Medium	0	
	High	0	
		Total:	N

Table 56 Concrete Pavement Survey Record Form

Beginning Log Mile:	0.617		
Ending Log Mile:	0.712		
Direction:	N		
Lane:	Outside		
Average Faulting (in):	0		
Average IRI (in/mi):	66.042		
Distresses	**Severity**	**Amount**	**Remedial Action**
Broken Slabs	Low	0	
		Total:	N
Corner Breaks	Low	1111	
	Medium	111	
	High	1110	
		Total:	N
Faulting of Trans. Joints	Medium	111	
	High	1111	
		Total:	N
Joint Seal Damage	Low	111	
	Medium	1111	
	High	1111	
		Total:	N
Longitudinal Cracking	Low	111	**N**
	Medium	111	
	High	1111	
		Total:	N
Longitudinal Cracking (Single JCP)	Low	0	
	Medium	0	
	High	0	
		Total:	N
Map Cracking & Scaling	Low	11111	
	Medium	111	
	High	111	
		Total:	N
Spall. of Long. Joints	Low	11	
	Medium	111	
	High	111	
		Total:	N

Spall. of Trans. Joints	Low	111	
	Medium	111	
	High	111	
		Total:	N
Transverse Cracking	Low	111	
	Medium	111	
	High	111	
		Total:	N
Transverse Cracking (Single JCP)	Low	0	
	Medium	0	
	High	0	
		Total:	N

7.2.2 Procedures of Maintenance Decision

The software was developed based on the Microsoft Excel, which is named as "MDOT Pavement Warranty Software." The user can double-click to open the software and fill in the form with the survey data, then the maintenance decision software can be run with the following procedures.

(1) Pavement Type Identification

Figure 51 presents the MainForm or main interface of the software. There are two buttons in the MainForm, "Asphalt Pavement" and "Concrete Pavement", which are used to select the pavement type. Before the data processing, the engineer should choose the pavement type based on the pavement warranty projects. For example, the engineer chooses the warranty pavement type as 'Asphalt Pavement'.

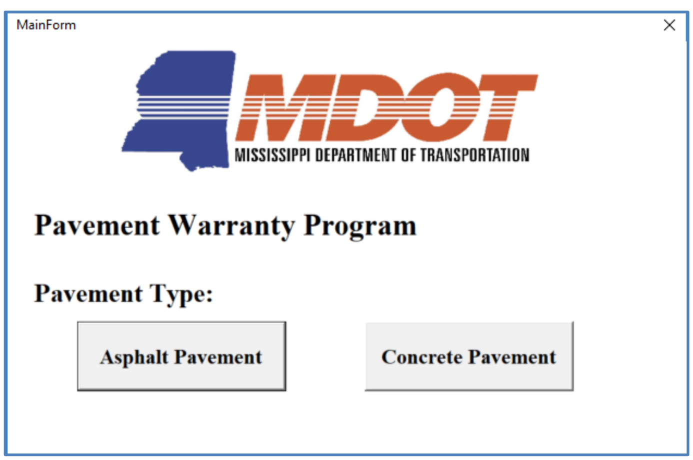


Figure 51 MainForm of the Software

(2) Distress Type and Threshold Option Selection

After the selection of the warranty pavement type as asphalt pavement, the distress types and thresholds information window will appear shown as Figure 52. In this window, the engineer can select a distress type from a drop-down menu shown in Figure 53. After selection of a distress type, the engineer can also select the threshold option performed as in Figure 54. The 1.0 σ, 1.5 σ, and best-fit based thresholds are built in this software as threshold options. Next, the summary of the numerical information about the specific distress type and severity level could be presented in the blank window, for example, the longitudinal cracking measurement thresholds at low, medium, and high severity levels with the option of 1.5 σ are presented in Figure 55.

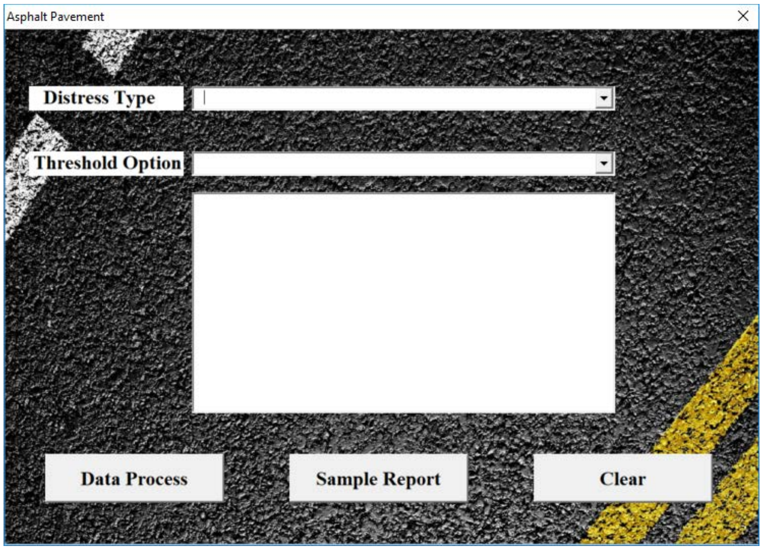


Figure 52 Distress Types and Thresholds Information Window

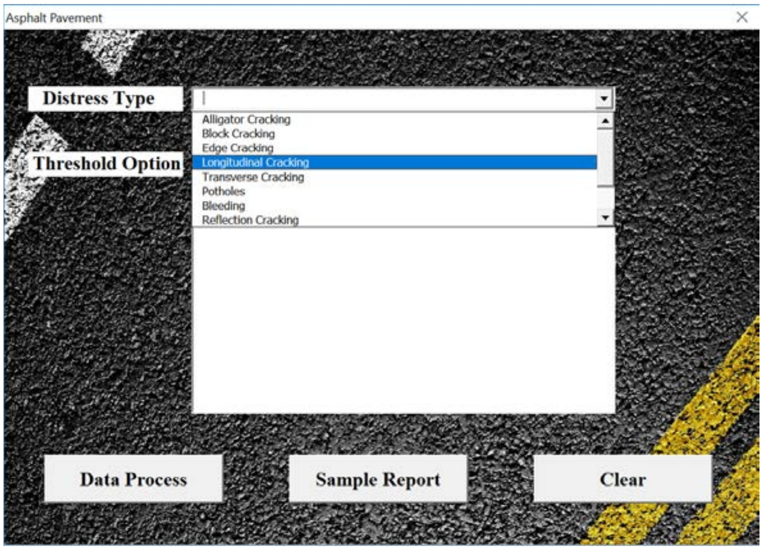


Figure 53 Selection of Distress Types

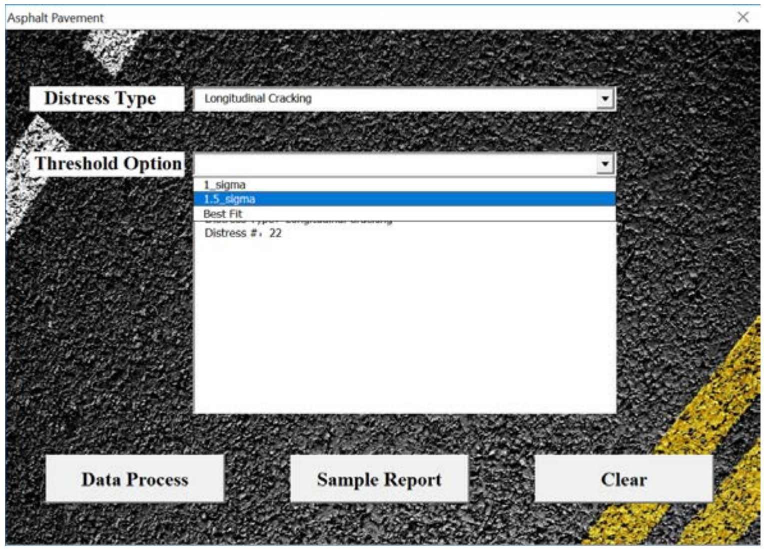


Figure 54 Selection of Threshold Options

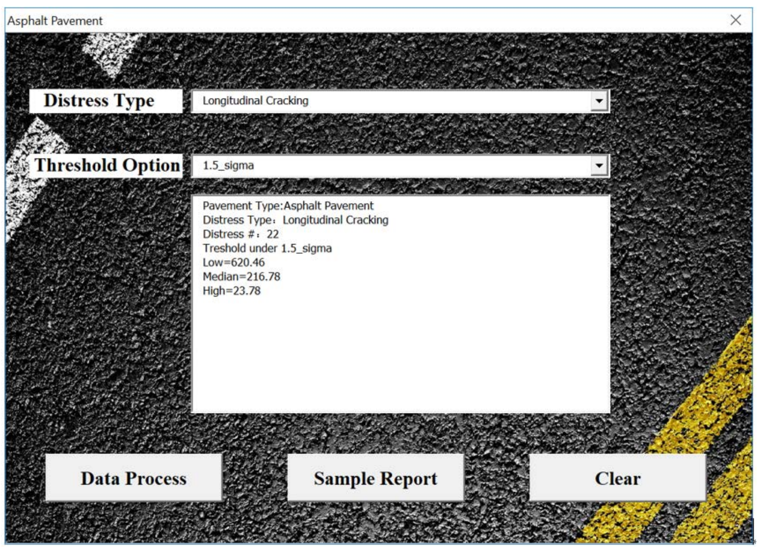


Figure 55 Visualization of the Distress Threshold Information

(3) Distress Evaluation

After an information check, click the “Data Process” button, then the software will run to make maintenance decision with the 1.5 σ thresholds for all of the distress data recorded in Table 55. A brief report will appear in the new window showing the number of sections for each pavement distress that exceeds the thresholds in Figure 56.

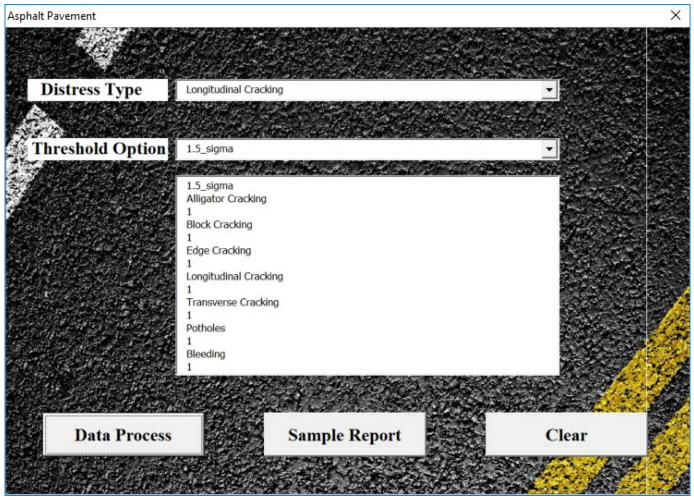


Figure 56 Brief Report for Maintenance Decision

(4) Maintenance Decision Report Development

A detailed report could be illustrated in Excel after clicking the “Sample Report” button. For example, in Table 57, the sample report shows the segments with direction, lane, distress types, the issued severity level, distress amount, and the evaluated threshold level that need remedial actions. Moreover, the maintenance decisions are also performed in the initial data record form as shown in Table 58. The distress that exceeds the thresholds would be in red and the Remedial Action will be changed to “Y” (means “YES, need remedial action”) from the initial “N” (means “NO, does not need remedial action”). The distress report in Excel can also be saved as .pdf format using the “Save As” button in Excel.

Table 57 Sample Report for Maintenance Decision in Excel

Beginning Mile	Ending Mile	Direction	Lane	Distress Type	Severity level	Amount	Threshold Level
20.439	20.439	20.439	20.439	Average Rut Depth		11	1.5_sigma
20.439	20.439	20.439	20.439	Average IRI		11111	1.5_sigma
20.439	20.439	20.439	20.439	Bleeding	Low	1111	1.5_sigma
20.439	20.439	20.439	20.439	Block Cracking	Low	111110	1.5_sigma
20.439	20.439	20.439	20.439	Edge Cracking	Low	11111	1.5_sigma
20.439	20.439	20.439	20.439	Longitudinal Cracking	Low	11111	1.5_sigma
20.439	20.439	20.439	20.439	Potholes	Low	11111	1.5_sigma
20.439	20.439	20.439	20.439	Reflection Cracking	Low	11111	1.5_sigma
20.439	20.439	20.439	20.439	Rutting	Low	11111	1.5_sigma
20.439	20.439	20.439	20.439	Transverse Cracking	Low	11111	1.5_sigma

Table 58 Survey Record Form with Maintenance Decision

Beginning Log Mile:	20.439		
Ending Log Mile:	20.534		
Direction:	N		
Lane:	Outside		
Average Rut Depth (in):	11		
Average IRI (in/mi):	11111		
Distresses	Severity	Amount	Remedial Action
Alligator Cracking	Low	1000	
	Medium	11110	
	High	0	
		Total:	Y
Bleeding	Low	1111	
	Medium	0	
	High	0	
		Total:	Y
Block Cracking	Low	111110	
	Medium	0	
	High	0	
		Total:	Y
Edge Cracking	Low	11111	
	Medium	0	
	High	0	
		Total:	Y
Longitudinal Cracking	Low	11111	
	Medium	0	
	High	0	
		Total:	Y
Potholes	Low	11111	
		Total:	Y
Raveling & Weathering	Low	11111	
	Medium	0	
	High	0	
		Total:	N
Reflection Cracking	Low	11111	
	Medium	0	
	High	0	
		Total:	Y
Rutting	Low	11111	
	Medium	0	
	High	0	
		Total:	Y
Transverse Cracking	Low	11111	
	Medium	0	
	High	0	
		Total:	Y

7.3 Adjustment of the Threshold Options

If the engineer wants to select other threshold options, the "Clear" button can be used to eliminate the current evaluation result, followed by using the "Threshold Option" to select other threshold options, for example, the best-fit threshold option is selected in Figure 57 and the distress measurement thresholds are listed in the window, accordingly.

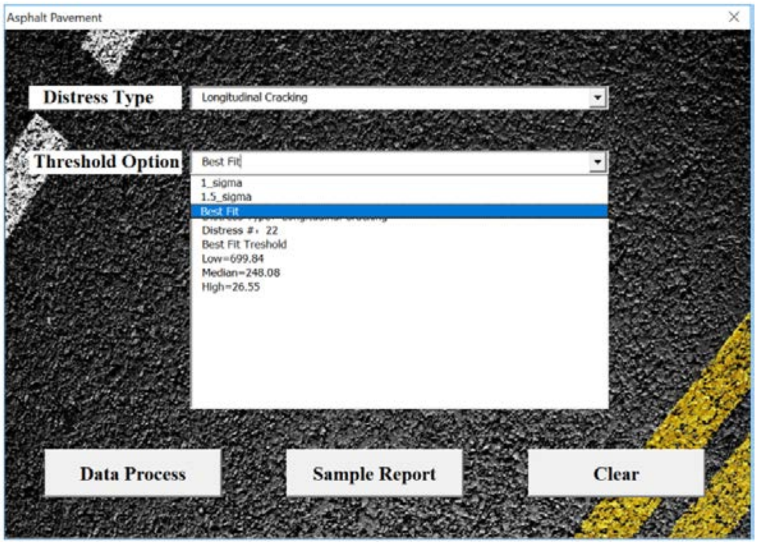


Figure 57 Adjustment of Threshold Options

However, based on the data limitation in warranty projects, for the best-fit based thresholds, only alligator cracking, block cracking, longitudinal cracking, transverse cracking, rutting in percentage, and average rutting depth in asphalt pavement are available. Other distress types in both asphalt and concrete pavements follow 1.5 σ based thresholds, because the 1.5 σ thresholds are suggested by the authors as the new distress measurement based thresholds in the warranty program in Mississippi. The availability of the threshold levels (1.0 σ, 1.5 σ, and best-fit) in each distress type for both asphalt and concrete pavements is listed in Table 59 and Table 60.

Table 59 Threshold Levels in Distress Types for Asphalt Pavements

Distress Type	1.0_Sigma	1.5_Sigma	Best-fit
Average Rut Depth	√	√	√
Average IRI	√	√	
Alligator Cracking	√	√	√
Bleeding	√	√	
Block Cracking	√	√	√
Edge Cracking	√	√	
Longitudinal Cracking	√	√	√
Potholes	√	√	
Raveling & Weathering			
Reflection Cracking	√	√	
Rutting	√	√	√

Table 60 Threshold Levels in Distress Types for Concrete Pavements

Distress Type	1.0_Sigma	1.5_Sigma	Best-fit
Average Faulting			
Average IRI	√	√	
Broken Slabs			
Corner Breaks			
Faulting of Trans. Joints			
Joint Seal Damage			
Longitudinal Cracking	√	√	
Longitudinal Cracking (Single JCP)			
Map Cracking & Scaling	√	√	
Spall. of Long. Joints	√	√	
Spall. of Trans. Joints	√	√	
Transverse Cracking	√	√	
Transverse Cracking (Single JCP)			

CHAPTER 8. VALIDATION OF AUTOMATED PAVEMENT CRACKING MEASUREMENTS

In order to statistically verify the applicability of using automated measurements in place of manual rating measurements to collect pavement cracking data, the classification accuracy index was first employed as an evaluation criterion to explore the precision of crack detection under different section lengths. Second, the one sample t-test was conducted to investigate the accuracy of the automated method, and then the correlation between the measurement errors and pavement conditions was analyzed. Finally, the Levene’s test was employed to evaluate the precision of the automated method based on the variance of the automated and semi-automated measurement results.

8.1 Distress Data Collection Technology

The study result in the previous report indicates that the roughness, rutting, cracking, and joint faulting are the most commonly surveyed performance data for warranty projects (Qi et al. 2012). Both manual/semi-manual and automated data collection technologies used recently are reported in the comprehensive survey as shown in Table 61.

Table 61 Distress Data Items and Data Collection Technologies

Distress Type	Data Collection Technology	BC*	FL	IL	IN	LA	MS	NS*	PA	WI
Cracking	Manual	√	√	√	√	√		√	√	√
	Film Video					√	√			√
	Digital Image			√		√		√	√	√
Rut-Depth	Three Sensor		√		√			√		
	Scanning Laser	√		√		√	√			√
	Five Sensor					√				√
Joint-Faulting	Ultrasonic						√	√		
	Laser			√		√				√
	Handheld fault meter				√	√	√		√	

*Note: BC = British Columbia; NS = Nova Scotia

Eight states employ the wholly manual method or a revised version, the so-called “semi-manual” method involving videotaping/digitalizing of road surface images in the field and manual data collection by watching the playback of the videotape or digital images in the office. Automated technology is well developed and widely used for rutting data collection. All states reported gathering rutting data with sensors or laser technologies except for Pennsylvania. The Pennsylvania DOT uses a manual method with 12-ft straight edges and is intended to switch to scanning lasers in 2012. As for joint faulting measurement, ultrasonic, laser, and handheld fault meter technologies were all reported in the survey. For instance, Indiana measures joint faulting with a handheld fault meter manually. Louisiana uses laser assessment for preliminary analysis of joint-faulting. In the event of excessive distress, on-site assessment is conducted using a handheld fault meter.

8.2 Description and Comparison of Data Collection Technologies

Pavement distress data recorded in the PMS is always used for pavement condition evaluation, performance prediction, and maintenance and rehabilitation (M&R) activities. Unreliable pavement distress data could result in inaccurate or even completely erroneous condition evaluation which could negatively affect the relevant decisions. A quantitative assessment of the impact of systematic and random errors in pavement condition data on PMS outputs was conducted to revealed that the measurement error would lead to an underestimated pavement condition and result in unreasonable pavement M&R prioritization and budget (Saliminejad and Gharaibeh 2013). From evaluating the influence of measurement error of pavement crack length and width on pavement performance evaluation and maintenance planning, it was discovered that it would significantly affect the reliability of pavement maintenance decisions (Jia et al. 2016). Therefore, it is critical for a PMS to provide accurate pavement distress data through appropriate and reliable pavement condition surveys.

Walking, windshield and semi-automated survey are three common pavement distress collecting methods which are depending on human labor. The difference between the three processes is that walking and windshield survey are conducted in the field, while the semi-automated method collects distress data on recorded pavement surface images. Although many manual distress ratings for agencies are executed according to well-defined criteria, a certain amount of subjectivity and the experience of the raters have some influence on the ratings (Tan and Cheng 2014). In addition, with the acceleration of pavement degradation and construction, the requirement of timeliness and comprehensiveness of PMS pavement distress data is elevating. Such distress survey methods which are based on visual rating are in low collection efficiency and fail to meet the requirement. Consequently, many researchers begin to study the fully automated method and try to replace manual rating with the fully automated image recognition based surveying (Salman et al. 2013, Pereira and Pereira 2015, Lins and Givigi 2016). Due to the rapid development of computer and image processing technology in the last two decades, the fully automated pavement distress survey method has gradually matured. Moreover, its high collection efficiency and safety make automated pavement surveying acceptable for some pavement distress collections. However, challenges quickly reveal themselves. The fully automated distress survey process consists of image acquisition, image processing and data output which is much more complex than the visual rating. Each step may lead to new measurement errors and result in a deviation of the measurement results from the actual condition or ground truth. In 2008, NCHRP conducted a survey of pavement distress collection methods adopted by 65 US research institutions, and the results showed that most research institutions used the semi-automated method to collect distress data and only a few of them employed the fully automated method (McGhee 2004). The reliability of pavement distress data collected by the fully-automated system is one of the biggest obstructions which hinders the states to select it for distress survey.

To ensure the reliability of the automated method, quality control (QC) and quality assurance (QA) approaches are used by each state. QC is mainly completed by vendors, and the purpose is conducting an equipment test to ensure that the fully automated pavement survey equipment satisfies the industry requirements. QA is examined by users to control the accuracy and precision of the collected data. Its examination process involves checking the consistency of equipment performance to ensure the equipment meeting the vendors' commitments and assessing the magnitude of equipment errors to ensure the reliability of PMS. Therefore, the minimum measurement error was set up and adopted in many states to check the reliability of the fully automated data in each data collection method. However, because of the varying pavement circumstances and specifications of pavement condition evaluation among states, it is still difficult to arrive in the establishment of a generally accepted standard to evaluate the effectiveness of the fully automated system. Therefore, it has been of great interest to statistically evaluate the reliability of the automated data collected on maintained highways.

8.3 Pavement Distress Collection in Mississippi

MDOT collects pavement condition data on state maintained highways in every two years. Before 2010, pavement surface distresses such as cracks were manually rated from pavement distress images captured by vehicle mounted video-logging cameras, and a random sampling of approximately 20 percent of the highway mileage was used for the distress evaluation. After 2010, MDOT replaced the semi-manual data collection method with a fully automated surveying technology to cover 100% mileage of pavement investigation in network-level.

By the end of 2017, there were approximately 27,500 lane-miles of state-maintained pavements in Mississippi. MDOT collected and managed the pavement distress data on 13,800 lane-miles (the rightmost through lane in each direction on divided highways, north- and eastbound lanes on undivided), which were more than 50% of the total state-maintained mileage, and the total number of pavement distress data records had reached 25,223,783 involving 3,145 construction projects. These data recorded in PMS were also widely used in various aspects of pavement engineering, such as providing support for the construction of pavement performance model and making M&R decisions for nearly 11,000 projects.

MDOT owns a fairly big amount of automated distress data in its PMS for pavement design, construction and maintenance purposes. Therefore, any inaccurate distress data may bring adverse impacts on decision-making for transportation agency officials, and lead to an increase of unreasonable pavement-related cost. Therefore, it is necessary to check the automated pavement distress data against the distress data collected for the same pavement sections and understand the equivalency and reliability of using the automated technology to replace manual or semi-manual data collection practice in project level in Mississippi. Actually, MDOT is also participating in the distress data quality pooled fund study (TPF-5(299)) with FHWA and other states in an attempt to arrive at distress “ground truth.”

8.4 Automated Data Validation Results and Analyses

8.4.1 Description of Data Validation Sample

The pavement distress data collected in 2014 and 2016 from 22.871 miles of asphalt pavement sections near Meridian, Mississippi, were selected as a sample in the data validation, which is listed in Table 62. The fully automated data were collected by the automated distress identification software AutoCrack and AutoClass. The semi-automated data selected as the reference values in this study was collected through visual detection by experienced distress raters of MDOT. The raw pavement images used for distress surveys were collected by the PathRunner which was equipped with a Global Position System (GPS) and newly developed and calibrated high-resolution cameras. The section number, beginning and ending locations of each section and distress positions were determined by GPS coordinates to ensure that the semi-automated and fully-automated data were correctly matched. In addition, the pavement longitudinal and transverse profiles were recorded by the on-board South Dakota Profiler for the purpose of measuring IRI, rutting and textures. The pavement distress information, such as distress type, severity level, distress quantity measurement and spatial coordinates, was recorded and summarized in every 550-ft section (MDOT 2015, MDOT2016).

Table 62 Data Validation Sample

Project Name	Direction	Survey Sections
Lauderdale MS-19 NB	North	72
Lauderdale MS-19 SB	South	52
Neshoba SR-19	South	122
Newton SR-19	North	50
Pontotoc SR-9	North	216
Total		512

8.4.2 Effectiveness Analysis of Crack Detection

Pavement image collection, pre-processing, crack detection, crack quantification and crack classification are the five main steps in the fully automated distress collection procedure in the Pathway system. Crack detection is the most complicated and critical step although the applicability of the pre-processing algorithms and the quality of collected distress images could affect the detection accuracy which could further impact the reliability of the measurement value. Therefore, it is important to assess the fully automated crack detection.

Two kinds of errors may occur during the distress detection process. One is the extra reported cracks that do not exist on the pavement surface, and the other is the underreported cracks. The two errors can be quantified using the classification evaluation indicators false positive and missed crack (false negative) which are defined in Equation 8-1 and Equation 8-2. Moreover, two modified precision evaluation indexes "precision" and "recall" are calculated by Equation 8-3 and Equation 8-4.

$$False\ Positive = \frac{\sum_{i=1}^{n} Sec_i[Reference = 0 \,\&\, Auto > 0]}{\sum_{i=1}^{n} Sec_i} \times 100\% \quad \textbf{8-1}$$

$$Missed\ Crack = \frac{\sum_{i=1}^{n} Sec_i[Reference > 0 \,\&\, Auto = 0]}{\sum_{i=1}^{n} Sec_i} \times 100\% \quad \textbf{8-2}$$

$$Precision = \frac{\sum_{i=1}^{n} Sec_i[Reference = 0 \,\&\, Auto > 0]}{\sum_{i=1}^{n} Sec_i[Reference = 0]} \times 100\% \quad \textbf{8-3}$$

$$Recall = \frac{\sum_{i=1}^{n} Sec_i[Reference > 0 \,\&\, Auto = 0]}{\sum_{i=1}^{n} Sec_i[Reference > 0]} \times 100\% \quad \textbf{8-4}$$

Where *n* is the total number of pavement sections considered in the research; *i* is the section number; *Sec*[] is the indicator function that equals to 1 if the criteria is satisfied; *Reference* and *Auto* are the summarized cracking measurement values calculated by the semi-automated method and fully-automated method, respectively.

Table 63 Evaluation Results of Crack Detection Errors

Distress Type	False Positive	Missed Crack	Precision	Recall
Longitudinal Cracking	54.7% (280)	6.1% (31)	78.4%	20.0%
Transverse Cracking	36.9% (189)	4.7% (24)	62.0%	11.6%

Table 63 presents the results of the quantification of 4 evaluation indexes for two different distress types. The results show that in the identifying of the longitudinal and transverse cracking, there are 280 and 189 sections (with 54.7% and 36.9% of total sections) reporting as false positive cracking for longitudinal cracking and transverse cracking, and with precisions of 78.4% and 62.0%. Furtherly, 31 and 24 sections exist in missed cracking for longitudinal cracking and transverse cracking, total 6.1% and 4.7% of all the sections, and with recall of 20.0% and 11.6% in distress sections. The automated crack detection algorithm seems to report higher percentages of false positive and precision compared to the percentages of missed crack and recall. It seems that the automated method could be easily disturbed by pavement texture or other features and mistakenly identify them as cracking. Moreover, the calculation results of missed crack and recall show that the fully automated system presents a relatively low rate of underreported cracking. In addition, it is noticeable that the fully automated algorithm has a slightly higher accuracy in identifying transverse cracking than that for the longitudinal cracking. At this point in time, there must be human QC done by the vendor and QA done by MDOT to ensure more accuracy.

Due to the unavailability of the crack map in this study, the crack information which is summarized in each section was utilized to calculate the evaluation indexes. However, the numerical value of an index would be affected by the section lengths. To understand the effect of section length on evaluation results, the scaling factors were included in the evaluation process. 2 and 0.5 were the two scale factors applied in this study, which represented doubling the section length and reducing the section length by half, respectively. The results are presented in Table 64.

Table 64 Evaluation Results of Crack Detection Errors in Different Section Lengths

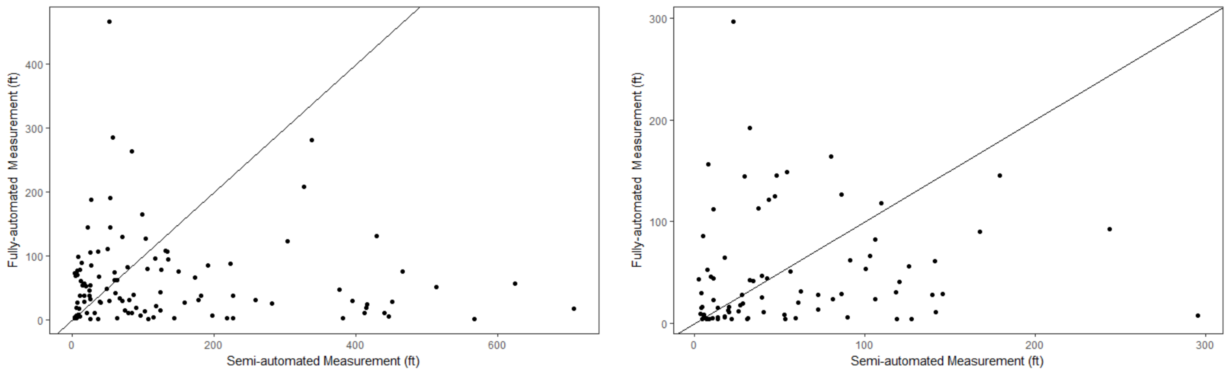
Distress Type	Scale Factor	False Positive	Missed Crack	Precision	Recall
Longitudinal Cracking	0.5	54.1%[1]	7.4%[1]	70.6%[3]	31.9%[1]
	2	49.0%[3]	4.3%[3]	84.6%[1]	10.1%[3]
Transverse Cracking	0.5	50.1%[1]	5.0%[1]	60.4%[3]	28.8%[1]
	2	49.8%[2]	3.1%[3]	78.0%[1]	8.5%[3]

Note: 1, 2, 3 represent the order of numerical values from high to low.

The result shows that, as the section length increases, the false positive and missed cracks are changing irregularly. The numerical values may increase or decrease, resulting in the evaluation results of false positive and missed crack inconsistency. The two indicators of precision and recall show a regular trend with the change of the section lengths that the precision increases and the recall decreases when the section length increases. This highlights the importance of QC/QA in the process.

8.4.3 Accuracy Analysis of Crack Measurement

The accuracy and precision analyses considered the summarized cracking measurement values of the subsections for which both the automated surveying and the rater detection were applied. In attempt to show the consistency between the fully automated and semi-automated data collection methods, Figure 58 presents the scatter plots of measurements for longitudinal and transverse cracking. The ordinate represents the fully automated measurement and the horizontal axis represents the semi-automated reference value. The points near the identity line represent the test result with high accuracy, while the points away from the identity line represent the opposite situation. The points above the identity line imply that the automated method underestimates the actual pavement distress condition, while the points under the identity line indicate an overestimation of the actual pavement distress condition.



(a) Longitudinal cracking (b) Transverse cracking

Figure 58 Semi vs. Fully Automated Measurements.

As shown in Figure 58, when the crack lengths of pavement surface are small, most of the points are close to the identity line. However, with the increase of the cracking lengths, the magnitudes of distance of the points off the identity line increase and more measurement results appear below the identity line than those above the line. The result illustrates that the accuracy of the fully automated method decreases with the deterioration of pavement condition, and the automated measurement results tend to underestimate the actual distress condition. However, the fully automated method has a lower detection accuracy for the obvious cracks on the pavement surface. Through repeated inspections of the original pavement surface images, it is found that the distress data analyzed in the research is all from the newly constructed pavement surface, and the length of the cracking is fairly small. Therefore, the fully automated method can easily produce the false negative, and a great number of pavement cracks cannot be completely detected.

In order to further study the relationship between measurement accuracy and pavement condition, the measurement error was calculated for each subsection as the difference between the fully automated and the semi-automated measurements. The Pearson correlation coefficient was utilized to quantify the correlation of the measurement error and crack length (Pearson 1895). Pearson correlation coefficient is a number between -1 and 1 which can be explained as the linear extent of two variables. The correlation degrees represented by the absolute value of the correlation coefficient are as follows: interval [0, 0.2] indicates extremely weak correlation or uncorrelated; interval (0.2, 0.4] indicates weak correlation; interval (0.4, 0.6] indicates moderate correlation; interval (0.6, 0.8] indicates strong correlation; interval (0.8, 1] indicates extremely strong correlation. The Pearson correlation coefficient is calculated by Equation 8-5.

$$Pearson(x, y) = \frac{\sum (x - \overline{x})(y - \overline{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2 \cdot \sum_{i=1}^{n} (y_i - \overline{y})^2}} \quad \textbf{8-5}$$

Where *x* and *y* are two independent variables. $\overline{x}$ and $\overline{y}$ are the mean values of the two independent variables; *n* is the total number of pairs of data points; *i* is the *i*-th data point.

Table 65 Pearson Correlation Test Results for Longitudinal/Transverse Cracking

Distress Type	Degree of Freedom	T value	P value	Correlation Coefficient
Longitudinal Cracking	107	-23.758	$<2.2\times10^{-16}$	-0.91
Transverse Cracking	80	-8.356	1.617×10^{-12}	-0.68

Table 66 One-Sample T-Test Results for Longitudinal/Transverse Cracking

Crack Type	Variation Source	Automated Data
Longitudinal Cracking	Average Value	-74.165
	Sample Numbers	109
	Degree of Freedom	108
	T Value	-4.395
	P Value	2.601×10^{-5}
	Significant Difference	Yes
Transverse Cracking	Average Value	-8.417
	Sample Numbers	82
	Degree of Freedom	81
	T Value	-1.018
	P Value	0.3177
	Significant Difference	No

The correlation analysis results show in Table 65. The Pearson correlation coefficients of longitudinal cracking and transverse cracking are -0.91 and -0.68 respectively. The P values are much smaller than 0.05, which would be an indication of a significantly strong correlation between the measurement error and crack length. The results show that with more deterioration in pavement cracking, the automated measurement error will increase, which may result in an overestimation of pavement cracking condition. In addition, a good linear correlation also indicates that the measurement error of the fully automated method can be calculated by a regression equation according to the surveyed pavement condition. Therefore, while the manual and semi-automated methods consume much human labor, the reliability of the state of practice of fully automated method is limited and cannot be promoted rapidly under the current scientific and technological circumstance. It is a feasible way to improve the accuracy of measurement results by conducting a manual correction according to the correlation between the measurement error and the pavement condition. Moreover, the measurement error of the fully automated method is not a constant value and is in a large variation interval, which indicates that the precision of the automated method may not be satisfactory enough in current data collection.

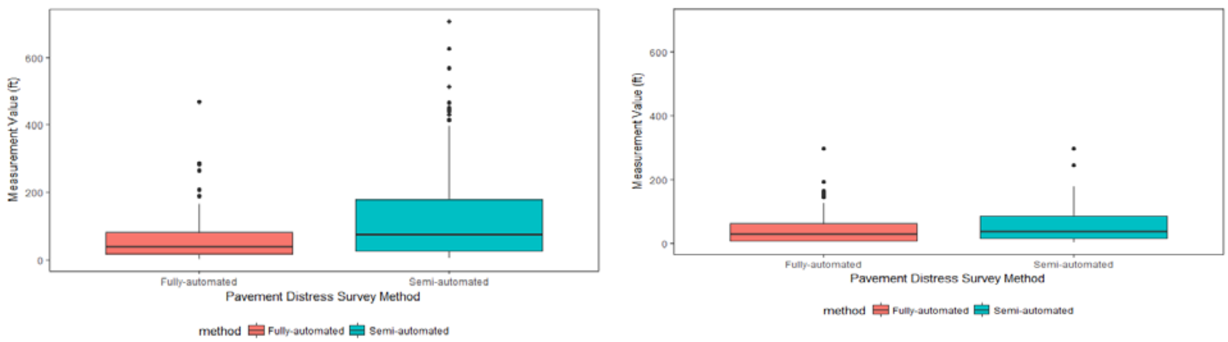
To better understand if there is a significant bias of fully automated measurements, a one-sample t-test was conducted to examine the mean of measurement error. The alternative hypothesis is that measurement error can exist, but the null hypothesis to be tested is that the mean value of measurement error is 0. The one-sample t-test results for the different survey data under 95% confidence level are listed in Table 66.

The test results show that the fully automated method significantly underestimates the actual longitudinal cracking, while it presents a great accuracy on transverse cracking measurements. In addition, the negative average measurement value would be an indication that the fully automated system has a trend to underestimate the actual distress condition. It should be noted that the data examined here excludes the section with false cracks and missed cracks. When this data is considered in the test, the distress condition would be overestimated due to the high false positive rate of fully automated method.

8.4.4 Precision Analysis of Measurements

Figure 59 presents boxplots of the fully automated and semi-automated measurements for longitudinal and transverse cracking to perform the precision of the two data collection methods. The median (the solid line in each colored box) is employed to represent the average situation of each pavement distress in the selected pavement sections using the automated or semi-automatic method. The upper and lower bounds in height of each box are Inter Quartile Range (IQR) which represent the precision of the data collection methods.

When the accuracy of the automated method is low while the precision is high, the medians of the two methods are different, and the boxplot is in different horizontal positions, but the IQR of the two methods is similar, and the shapes of boxplots are similar. When the accuracy of the automated method is low and the precision is low as well, the value of medians and IQR's of the two methods are different. Therefore, the shapes and horizontal positions of the boxplots are different. Based on the information shown in the figures, for both of the two pavement cracking types, a lower IQR and median are observed for the automated method compared with the referenced semi-automatic method, which indicates that the two methods have different accuracy, and the automated method underestimates the actual pavement distress length, and the precision of the automated distress data is lower. A potential reason for the fact that the IQR of the full automated method is lower that of the semi-automated method is that the measurement errors of the automated method are highly correlated with the pavement distress condition. When the pavement condition is good, the automated method tends to overestimate the actual pavement distress. However, when the pavement condition is poor, the automated method tends to underestimate the actual pavement distress, which leads to a more concentrated distress measurement value. In contrast, the transverse cracking measurement seems to have a better precision than the longitudinal cracking.



(a) Longitudinal cracking (b) Transverse cracking

Figure 59 Boxplots of Semi and Fully Automated Measurements.

Levene's test could be used to assess the equality of variances for a variable calculated for two or more groups, and it is less sensitive than the other homogeneity of variance test to departures from normality (Levene 1961). Hence, the Levene's test was further utilized to qualitatively evaluate the precision of the fully automated method. The statistic of Levene's test is defined as follows:

$$W = \frac{(N-k)}{(k-1)} \cdot \frac{\sum_{i=1}^{k} N_i (Z_{i\cdot} - Z_{\cdot\cdot})}{\sum_{i=1}^{k} \sum_{j=1}^{N_i} (Z_{ij} - Z_{i\cdot})^2} \qquad \textbf{8-6}$$

Where k is the number of different groups to which the sampled cases belong; N_i is the number of cases in the i^{th} group; N is the total number of cases in all groups; $Z_{ij} = |Y_{ij} - \tilde{Y}_i|$, $\tilde{Y}_i$ is the median of the i^{th} group and Y_{ij} is the value of the measured variable for the j^{th} case from the i^{th} group; $Z_{i.}$ = the mean of the Z_{ij} for group i; $Z_{..}$ is the mean of all Z_{ij}.

The test statistic W is approximately F-distributed. The alternative hypothesis is that the measurement variances are different for groups, and the null hypothesis is that all of the groups have similar population variances. With a 95% confidence, the Levene's test results for the two types of pavement cracking are listed in Table 67. The W values of the longitudinal cracking and transverse cracking are 20.073 and 0.3103. The test results show that the fully automated system has a higher precision on transverse cracking evaluation, but the precision in longitudinal cracking measurement is lower. There are two reasonable interpretations for this phenomenon. One is that the length of the transverse cracking is constrained by the width of the pavement lane and is actually much shorter than the longitudinal cracking, making the measurement error relatively small, which causes a relatively small measurement error distribution range. The other is that the error of the longitudinal cracking measurement value has a higher correlation with the pavement surface condition, which results in a larger distribution range of the measurement errors.

Table 67 Levene's Test Results for Longitudinal/Transverse Cracking

Distress Type	W Value	Degree of Freedom		P Value	Significant Difference
		k-1	N-k		
Longitudinal Cracking	20.073	1	216	1.208×10^{-5}	Yes
Transverse Cracking	0.3103	1	162	0.5782	No

8.5 Summary of Data Validation

This chapter probes into the effectiveness of crack detection by utilizing the classification accuracy index, and analyses of the accuracy and precision of the fully automated method by conducting a statistical analysis of measurement error and measurement deviation, and the relationship between the measurement error and pavement condition. The crack detection accuracy analysis results indicate that the fully automated method cannot accurately detect the pavement cracking which exists on the pavement surface, and much more distress would be detected due to the high false positive error in the fully automated method. The statistical analysis results of the accuracy for the fully automated cracking measurements illustrate that the fully automated method cannot accurately evaluate the length of longitudinal cracking, but it is available to accurately evaluate the length of transverse cracking. The measurement error is highly correlated with the pavement condition, and with the degradation of pavement condition, the measurement error is significantly increasing. A high false positive error and incomplete cracking detection may cause a high measurement error in poor pavement condition when the inconspicuous cracking exists on the pavement surface. The statistical analysis results of the variance between the fully automated data and referenced semi-automated data show that the fully automated method can precisely evaluate the transverse cracking measurement, while it cannot precisely evaluate the longitudinal cracking. It may be because of that the variation range of transverse cracking measurement value is much smaller than that of the longitudinal cracking and the longitudinal cracking measurement error has a higher correlation coefficient with the surveyed pavement condition. It is worth noting that the data collection contractors who collect survey data every two years do QA on the automated distress to improve accuracy.

Based on the research results in this chapter and also the reviews of the literature including a previous report by the authors, the research team does not suggest that automated survey technology be predominantly accepted by the MDOT to serve the functions of surface distress condition data collection for the warranty pavements. Even though automated technologies are the emerging trend for pavement surface distress data acquisition, they are not sophisticated enough to be used at project level, especially when warranty clauses are involved.

CHAPTER 9. SUMMARY AND RECOMMENDATION

9.1 Introduction

The pavement warranty program is used to specify the minimum performance conditions (i.e., distress thresholds) of a particular pavement over a specified period of time (i.e., warranty period) and to define the contractor's remedial responsibility in case of premature failures. For state highway transportation agencies, it is regarded as one of the most prominent programs for protecting investment in pavement constructions and preventing early failures. The contractor is responsible for restoring the pavement condition at no cost to the highway agency when the thresholds of warranty items are exceeded during the warranty period. It is an innovation for the contractor to reallocate performance risk, which could ensure the quality of the constructed products, enhance the pavement service performance, reduce the amount of state highway agency resources required on maintenance projects, and decrease the life-cycle cost of projects (Anderson and Russell 2001). Based on these characteristics, more states in the US and other countries are attempting to conduct the pavement warranty program.

However, because of the varying operating environments and the difference in specifications of pavement warranty programs among states, it is difficult to arrive in the establishment of a generally accepted procedure or method to evaluate the effectiveness of pavement warranty programs. Therefore, one of the most important technical challenges in pavement warranty implementation is to explore an appropriate method for the establishment of the desirable evaluation, which serves to examine the performance comparison between warranty and non-warranty pavements. Hence, this research utilized a new survival analysis based evaluation method and compared the performance of warranty versus non-warranty pavements with the use of the pavement distress survey data saved in the MDOT's PMS database. It serves to provide a feasible solution for the agencies and contractors to evaluate the effectiveness of pavement warranty programs.

In recent years, with the development of new technology in the data collection arena, MDOT has changed from the human-rated sampling method to 100% automated distress survey at the network-level for the entire state-maintained system, whereas the warranted pavement projects are still being surveyed using the semi-annual (video logging in the field and in-house manual rating) technologies (Luo et al. 2018). However, the conversion of distress measurements into deduct points using the empirically regressed deduct point equations reduces the accuracy of the objective distress measurements by increasing chances of subjective and random errors. Also, the deduct point equations were empirically developed in 1990s reflecting the data, experience, and technologies at that time, thus, the validity and applicability of continual use of these deduct point equations have become problematic with time (Qi et al. 2015). Therefore, the research study developed new distress thresholds based on direct measurements of pavement distresses or distress densities for MDOT's pavement warranty program and based on which a maintenance decision software to implement the new thresholds was developed.

MDOT owns a fairly big amount of automated distress data in its PMS for pavement design, construction and maintenance purposes. Inaccurate distress data will bring adverse impacts on decision-making for transportation agency officials, and lead to the increase of unreasonable pavement-related cost. In order to make a comprehensive evaluation of the quality of the automated pavement distress data in the PMS, the research probed into the accuracy of crack detection by utilizing the classification accuracy index, evaluated the accuracy and precision of the fully automated method by conducting a statistical analysis of measurement error and measurement deviation, and analyzed the relationship between the measurement error and pavement condition.

Overall, this study evaluated the performance of the warranty pavements versus the non-warranty pavements in Mississippi. The study also developed appropriate measurement based distress threshold values to update the existing deduct point based thresholds for MDOT. First, a comprehensive literature review was performed to review the pavement warranty practice and related studies in Mississippi and other states. Then, survival analyses were conducted to compare the pavement performance for warranty contracting versus the general non-warranty contracting. Next, bootstrapping method was employed to normalize the skewed non-warranty distress data to draw the PDF and CDF curves for each distress type. Further, the corresponding percentiles of each distress threshold value were determined from the CDF curves of the pavement distress data based on the standard deviation method and best-fit method, and then used to evaluate the appropriateness of the threshold levels. The new thresholds were implemented using a Microsoft Excel environmental Visual Basic based software developed by the research team. Eventually, the classification accuracy index, the one sample t-test, and the Levene's test were employed to validate the currently available automated data collection technology.

9.2 Summary of Findings

The study selected the warranty pavement distress data from the annual maintained pavement distress reports and non-warranty pavement distress data from 2000 to 2014 from PMS database. In addition, the selected PMS data was reprocessed into 0.1-mile segments to compile the non-warranty distress data sections. Then the pavement performance evaluation, rebuilding of the measurement based thresholds, implementation of the new thresholds in software, and validation of automated data collection technology were conducted in this research study. Main findings are illustrated as follows:

1) Evaluation of Warranty and Non-warranty Pavement Performance in Mississippi

- During the initial service time, the survival probability of the warranty pavements is equal to 1, then survival probability curves fall slightly over time;

- The order of the deteriorating speeds for the warranty pavements in this research is listed from the highest to the lowest: rutting, transverse cracking, IRI, longitudinal cracking, and alligator cracking;
- All the distresses in the non-warranty pavements decrease continuously during the service periods in three threshold scenarios;
- The average performance of warranty pavements is better than that of the non-warranty pavements.

2) Development of New Thresholds for Warranty Pavements

- The non-warranty pavement distress data is skewed; thus, the bootstrapping method could be employed to normalize the data.
- The best-fit method could realize the smoothest transition from the deduct point based thresholds to the measurement based thresholds. However, only some of the distress types with enough warranty and non-warranty data could be applied with the best-fit method to develop thresholds.
- The 1.5 σ based thresholds could be an alternative for the new measurement based thresholds due to its higher accuracy than other standard deviation based thresholds such as the 1.0 σ method.
- The rutting and IRI could be warranty items in warranty program to monitor the ride quality and roughness of the pavement condition.

3) Validation of Automated Data Collection Technology

- The crack detection algorithm in the study cannot correctly always distinguish the distress area and non-distress area on the pavements.
- There is a gap between the automated measurement and the actual value of longitudinal cracking, and the gap is highly correlated with pavement distress condition.
- The transverse cracking data collected by the automated method have higher precision than that of longitudinal cracking.

9.3 Final Remarks

Jackson State University has completed MDOT Research Project 107453/101000, “Update and Documentation of MDOT Warranty Process and Distress Thresholds.” During this research, the following achievements were attained:

- The pavement performance of warranty program versus non-warranty program in Mississippi was evaluated;

- Statistical data processing for skewed non-warranty distress data recorded in PMS was administrated;
- Two sets of distress thresholds for asphalt and concrete warranty pavement were developed using two kind of statistical methods;
- The distress thresholds were compared and selected based on existing warranty maintenance decisions;
- The new measurement based distress thresholds were implemented in maintenance decision software based on the Microsoft Excel© environment;
- Automated and semi-automated data collection technologies were validated; and
- The project report could be used as training materials for implementation of new warranty thresholds for MDOT pavement warranty program.

As a result of the study, based on statistical analyses of the warranty and non-warranty pavement distress data in Mississippi, the research team has reached the following recommendations.

- The average performance of warranty pavements is superior to that of the non-warranty pavements in Mississippi. Survival analysis is an effective way to assess the performance of warranty and non-warranty pavements. The beginning of the distress failures for the warranty and non-warranty pavements could be checked easily in the survival curves; moreover, the most serious distress types in warranty and non-warranty pavements could be illustrated in the comparison of the survival performance curves.
- The application of bootstrapping method is effective in the normalization of the skewed non-warranty distress data. Considering the need for smooth transition from the existing deduct point based system to the new measurement based system and distress data availability, the thresholds based on the 1.5 σ method with the normalized non-warranty distress data could be an alternative to rebuilding the distress thresholds for the pavement warranty program in Mississippi.
- The fully automated method cannot accurately detect the pavement cracking at project level. The fully automated method cannot accurately evaluate the length of longitudinal cracking, but it is available to accurately evaluate the length of transverse cracking. Therefore, the automated technique is currently not sophisticated enough for project-level pavement distress surveys.

References

Anderson, S., and Russell, J. (2001). "NCHRP Report 451: Guidelines for warranty, multi-parameter, and best value contracting." *Transportation Research Board*Washington, DC.

Archilla, A. R. (2006). "Use of superpave gyratory compaction data for rutting prediction." *J. Transp. Eng.*, 132(9), 734-741.

Battaglia, I. K. (2009). "Pavement warranty program in Wisconsin: 12-year evaluation." *Transportation Research Board*Washington, DC.

Battey, R., and Watkins, J. (2009). "Current status of MDOT's maintained pavement program." *Proc., Presentation at Louisiana Transportation Conference. Baton Rouge, LA.*

Chernick, M. R., and LaBudde, R. A. (2014). *An introduction to bootstrap methods with applications to R*, John Wiley & Sons.

Delucchi, K. L., and Bostrom, A. (2004). "Methods for analysis of skewed data distributions in psychiatric clinical studies: working with many zero values." *American Journal of Psychiatry*, 161(7), 1159-1168.

Deng, X., Liu, Q., Deng, Y., and Mahadevan, S. (2016). "An improved method to construct basic probability assignment based on the confusion matrix for classification problem." *Information Sciences*, 340, 250-261.

Desgagne, A., Castilloux, A.-M., Angers, J.-F., and Le Lorier, J. (1998). "The use of the bootstrap statistical method for the pharmacoeconomic cost analysis of skewed data." *Pharmacoeconomics*, 13(5), 487-497.

Dong, Q., and Huang, B. (2014). "Evaluation of influence factors on crack initiation of LTPP resurfaced-asphalt pavements using parametric survival analysis." *J. Perform. Constr. Facil.*, 28(2), 412-421.

Dong, Q., and Huang, B. (2015). "Failure probability of resurfaced preventive maintenance treatments: investigation into long-term pavement performance program." *Transportation Research Record: Journal of the Transportation Research Board*(2481), 65-74.

Efron, B. (1979). "Bootstrap methods: another look at the jackknife." *The annals of Statistics*, 1-26.

Efron, B., and Tibshirani, R. J. (1994). *An introduction to the bootstrap*, CRC press.

El Gendy, A., Qi, Y., and Wang, F. (2012). "Current state of the pavement warranty in the United States and Canada." *Proc., Transportation Research Board 91st Annual Meeting*.

FHWA (2017). "Pavement Warranties." <https://www.fhwa.dot.gov/pavement/warranty/>. (accessed in June, 2017).

George, K. (1995). "Pavement Management System-Phase II." Final Report, Report No. MS-RD-95-100, *Mississippi Department of Transportation*, Jackson, Mississippi.

Gharaibeh, N. G., and Darter, M. I. (2002). "Longevity of highway pavement in Illinois---2000 Update." Illinois Cooperative Highway Research Report No. IHR-R24, *University of Illinois at Urbana-Champaign*.

Gharaibeh, N., and Darter, M. (2003). "Probabilistic analysis of highway pavement life for Illinois." *Transportation Research Record: Journal of the Transportation Research Board*(1823), 111-120.

Gharaibeh, N. G., and Shirazi, H. (2009). "Risk-based model for pricing highway infrastructure warranties." *J. Infrastruct. Syst.*, 15(4), 378-382.

Hughes, C. (1996). "State of the practice of warranty specifications in the United States." Contract, 010-940.

Jackson, C. H. (2016). "Flexsurv: a platform for parametric survival modeling in R." *Journal of Statistical Software*, 70(8), 1-33.

Jia, X., Huang, B., Dong, Q., Zhu, D., and Maxwell, J. (2016). "Influence of pavement condition data variability on network-level maintenance decision." *Transportation Research Record: Journal of the Transportation Research Board*(2589), 20-31.

Levene, H. (1961). "Robust tests for equality of variances." *Contributions to probability and statistics. Essays in honor of Harold Hotelling*, 279-292.

Li, L., Guler, S. I., and Donnell, E. T. (2017). "Pavement friction degradation based on Pennsylvania field test data." *Transportation Research Record: Journal of the Transportation Research Board*(2639), 11-19.

Li, Z., Shao, Q., Xu, Z., and Cai, X. (2010). "Analysis of parameter uncertainty in semi-distributed hydrological models using bootstrap method: A case study of SWAT model applied to Yingluoxia watershed in northwest China." *Journal of Hydrology*, 385(1), 76-83.

Lins, R. G., and Givigi, S. N. (2016). "Automatic crack detection and measurement based on image analysis." *IEEE Trans. Instrum. Meas.*, 65(3), 583-590.

Luo, X., Wang, F., Wang, N., Qiu, X., and Tao, J. (2018). "Survival analysis of warranty and nonwarranty pavement performance using pavement management system data." *Transportation Research Board 97th Annual Meeting*Washington DC, United States.

McGhee, K. H. (2004). *Automated pavement distress collection techniques*, Transportation Research Board.

MDOT (2015). "*Pavement Management Manual*", Draft by Mississippi Department of Transportation.

MDOT (2016). "*MS DOT Distress Classification Guide*", Mississippi Department of Transportation and Pathway Service Inc.

Miller, J.S. and Bellinger, W. Y. (2003). *Distress identification manual for the Long-term pavement performance program* (Fourth Revised Edition). FHWA-RD-03-031, Federal Highway Administration, McLean, Virginia.

Paterson, W., and Chesher, A. D. (1986). "On predicting pavement surface distress with empirical models of failure times." *Transp. Res. Rec.*(1095).

Pearson, K. (1895). "Note on regression and inheritance in the case of two parents." *Proc. R. Soc. London*, 58, 240-242.

Pereira, F. C., and Pereira, C. E. (2015). "Embedded image processing systems for automatic recognition of cracks using UAVs." *IFAC-PapersOnLine*, 48(10), 16-21.

Powers, D. M. (2011). "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation."

Prozzi, J., and Madanat, S. (2000). "Using duration models to analyze experimental pavement failure data." *Transp. Res. Rec.*(1699), 87-94.

Prozzi, J., and Madanat, S. (2004). "Development of pavement performance models by combining experimental and field data." *J. Infrastruct. Syst.*, 10(1), 9-22.

Qi, Y., El Gendy, A., and Wang, F. (2012). "Evaluation of MDOT's distress thresholds for maintained pavement projects." Final Project Report No. SS# 221, *Mississippi Department of Transportation*, Jackson, Mississippi.

Qi, Y., Wang, F., El Gendy, A., and Li, Y. (2013). "Evaluation of the effectiveness of Mississippi's pavement warranty program." *Transp. Res. Rec.*(2366), 98-109.

Qi, Y., Wang, F., El Gendy, A., Li, Y., and Peng, B. (2015). "Evaluation of Mississippi’s pavement warranty thresholds using pavement management system data." *Transportation Research Board*Washington, D.C.

Qi, Y., Wang, F., El Gendy, A., and Luo, X. (2018). "Evaluation of pavement warranty thresholds based on survey study and analysis of pavement management system data." *Advances in Transportation Studies*. No. XLV .

Rodrıguez, G. (2005). "Non-parametric estimation in survival models."

Romanoschi, S., and Metcalf, J. (2000). "Evaluation of probability distribution function for the life of pavement structures." *Transportation Research Record: Journal of the Transportation Research Board*(1730), 91-98.

Sadeghi, L., McDaniel, R., and Haddock, J. E. (2016). "Effectiveness of warranted asphalt pavements in Indiana." *Transp. Res. Rec.*(2573), 69-75.

Saliminejad, S., and Gharaibeh, N. G. (2013). "Impact of error in pavement condition data on the output of network-level pavement management systems." *Transp. Res. Rec.*, 2366(1), 110-119.

Salman, M., Mathavan, S., Kamal, K., and Rahman, M. (2013). "Pavement crack detection using the Gabor filter." *Proc., Intelligent Transportation Systems-(ITSC), 2013 16th International IEEE Conference on*, IEEE, 2039-2044.

Scott III, S., T. Ferragut, M. Syrnick, and S. Anderson (2011). "Guidelines for the use of pavement warranties on highway construction projects." NCHRP Report No. 699, *Transportation Research Board*.

Smith, C. (2009). "Project Report: Implementation of automated distress rating for pavement condition surveys ", MDOT Research Division. Jackson, MS.

Smith, J. D., Lee, S., and Kazmierowski, T. J. (2016). "Use of pavement management system data to enhance pavement performance specifications in Canada." *Transportation Research Record: Journal of the Transportation Research Board*(2573), 60-68.

Stehman, S. V. (1997). "Selecting and interpreting measures of thematic classification accuracy." *RSEnv*, 62(1), 77-89.

Tan, S. G., and Cheng, D. (2014). "Quality assurance of performance data for pavement management systems." *Design, Analysis, and Asphalt Material Characterization for Road and Airfield Pavements*, 163-169.

Wang, L., Park, J., and Hill, S. (2005). "Use of pavement management system data to monitor performance of pavements under warranty." *Transp. Res. Rec.*(1940), 21-31.

Wang, Y., Mahboub, K. C., and Hancher, D. E. (2005). "Survival analysis of fatigue cracking for flexible pavements based on long-term pavement performance data." *J. Transp. Eng.*, 131(8), 608-616.

West, R., Michael, J., Turochy, R., and Maghsoodloo, S. (2011). "Use of data from specific pavement studies experiment 5 in the long-term pavement performance program to compare virgin and recycled asphalt pavements." *Transp. Res. Rec.*(2208), 82-89.

Winfrey, R. (1969). "Economic analysis for highways." *Transportation Research Board*, Washington, D.C.