TEXAS STATE UNIVERSITY

The rising STAR *of Texas*

TxDOT Report 0-7072-1
Improve Data Quality for Automated Pavement Distress Data Collection

Texas State University, Ingram School of Engineering; San Marcos, Texas
Feng Wang (PI)
Xiaohua Luo
Jueqiang Tao
Gauri Ravikant Mahajan
Ajmain Faieq

Submitted August 2024; Published November 2024

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA/TX-25/0-7072-R1		

4. Title and Subtitle	5. Report Date
Improve Data Quality for Automated Pavement Distress Data Collection	August 2024
	6. Performing Organization Code
7. Author(s)	**8. Performing Organization Report No.**
Feng Wang, Xiaohua Luo, Jueqiang Tao, Gauri Ravikant Mahajan, and Ajmain Faieq	Report 0-7072-R1
9. Performing Organization Name and Address	**10. Work Unit No. (TRAIS)**
Texas State University Ingram School of Engineering 601 University Drive San Marcos, TX 78666	**11. Contract or Grant No.** Project 0-7072
12. Sponsoring Organization Name and Address	**13. Type of Report and Period Covered**
Texas Department of Transportation Research and Technology Implementation Division P.O. Box 5080 Austin, TX 78763-5080	Technical Report 09/2020-08/2024 **14. Sponsoring Agency Code**

15. Supplementary Notes

Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration

16. Abstract

Since Fiscal Year 2017, the Texas Department of Transportation (TxDOT) has employed advanced automated and semi-automated techniques using 3D laser technology and high-resolution cameras to collect pavement condition data over 197,000 lane miles. Although this approach marks a significant technological advancement, concerns about the accuracy and precision of the collected data persist. To address these issues, TxDOT implemented a comprehensive quality management plan (QMP) that includes equipment calibration, weekly quality control verifications, and a dual-layered quality assurance (QA) process involving third-party visual assessments and internal rechecks by TxDOT personnel. This study evaluated current automated pavement condition data collection techniques through a literature review and a survey of highway agencies. It also analyzed four years of historical pavement condition data from 25 districts, examining accuracy, precision, and distress evaluations across different pavement types. Stratified sampling methods were used to improve data quality audits, ensuring representativeness in scenarios with high variability among population units. A pilot study in the San Antonio district validated the proposed QA framework, proving effective in identifying data quality issues. The findings highlighted significant challenges, such as inconsistencies and false positives, which compromise data validation. The necessity of manual validation and the establishment of robust quality assurance protocols were emphasized to ensure data reliability. The study underscored the importance of transparent benchmarks and stratified sampling methods to enhance the robustness of pavement condition data collection and management.

17. Key Words	18. Distribution Statement
Automated pavement condition data collection, data quality management, quality assurance, data accuracy and precision, sampling method, data quality threshold	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161; www.ntis.gov.

19. Security Classification (of this report)	20. Security Classification (of this page)	21. No. of Pages	22. Price
Unclassified.	Unclassified.	219	N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

DISCLAIMERS

ACKNOWLEDGMENTS

The authors would like to express appreciation to the Texas Department of Transportation for sponsoring this research. Special gratitude to the following individuals who provided support and guidance to the research team during the project study: Project Manager Jade Adediwura and all the Project Monitoring Committee members. We would also like to thank the following TxDOT personnel who provided support and technical input during the study: Jenny Li, Feng Hong, Hui Wu, Matthew Rogers, Soojun Ha, Diana Rogerio, and Aldo Madrid.

IMPROVE DATA QUALITY FOR AUTOMATED PAVEMENT DISTRESS DATA COLLECTION

Final Report
July 2024

Principal Investigator
Feng Wang, PhD, PE
Ingram School of Engineering, Texas State University

Co-Principal Investigator
Xiaohua Luo, PhD
Ingram School of Engineering, Texas State University

Research Assistant(s)
Jueqiang Tao, Gauri Ravikant Mahajan, and Ajmain Faieq
Ingram School of Engineering, Texas State University

Authors
Feng Wang, Xiaohua Luo, Jueqiang Tao, Gauri Ravikant Mahajan, and Ajmain Faieq

Sponsored by
Texas Department of Transportation

A report from
Ingram School of Engineering
Texas State University
601 University Drive
San Marcos, TX 78666

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1 Introduction

The Texas Department of Transportation (TxDOT) started to use an automated/semi-automated method utilizing 3D laser technology and high-resolution cameras to collect pavement condition data on over 197,000 lane miles since Fiscal Year (FY) 2017. However, the local experience also shows that there are still accuracy and precision issues associated with the reliability of the existing automated/semi-automated data collection methods. Therefore, TxDOT has used a quality management plan (QMP) for pavement data collection to control the quality of automated pavement data collection and processing. The QMP includes equipment calibration before data collection and weekly verification during data collection for QC. For QA of automated data, TxDOT has a data validation process that uses certified raters from an independent third party to visually evaluate about 6% of Texas roadbed miles for surface distresses and TxDOT personnel to recheck profile measurements on selected sections (TxDOT 2018a). The minimum tolerance for data quality of the surface distress audit is distress scores of 15 points in error for no more than 10% of the pavement sections. Nevertheless, with a review of the pavement condition reports over the last few years, it is quite evident that the existing data validation method has failed to reliably locate the pavement sections with low accuracy and/or precision in data. Therefore, the development of data quality assurance procedures is vital to guarantee the quality and reliability of the pavement management information system (PMIS) data.

This study developed quality assurance procedures for TxDOT in accepting automated pavement condition data. By following national quality management guides and good experience from state or provincial highway agencies in the US, the research team developed data quality assurance procedures for TxDOT by establishing data analysis/modeling capabilities, data audit sampling method, data quality tolerance values (data quality thresholds), implementation framework of data quality check, and data quality check validation procedures in PMIS. The main components of the research included the development of a cost-efficient audit sampling method and a set of data-enabled quality acceptance criteria. The cost-efficient audit sampling method can assist TxDOT in locating the pavement sections with potential data quality issues. The data-enabled quality acceptance criteria can provide TxDOT with minimum quality tolerances/thresholds and acceptance decision support to ensure the data quality of the delivered automated pavement condition data. In the project, a pilot study for a selected TxDOT District was implemented with the developed procedures to evaluate the effectiveness of the proposed data quality assurance procedures for pavement condition data quality improvement. This project enabled TxDOT to enhance the accuracy, precision, and reliability of the automated pavement condition data which would eventually help the State of Texas improve its pavement performance condition.

The investigation detailed within this study is articulated through a structured exposition across seven chapters, each dedicated to a specific facet as delineated below:

Chapter 2 delves into the deployment of automated pavement condition data collection, elucidating data quality intricacies sourced from the United States. It encompasses an analysis of the questionnaire survey results pertaining to the data quality management plan, as well as an examination of the challenges inherent in the current automated data collection framework.

Chapter 3 engages in a meticulous examination of historical pavement condition data, encompassing normality testing and an initial analysis of accuracy and precision concerning comprehensive scores and individual distress parameters.

Chapter 4 expounds upon the proposed sampling methodology devised for conducting an audit on the quality of automated pavement condition data.

Chapter 5 introduces the components essential for verifying the consistency of automated pavement condition data quality, alongside the delineation of associated thresholds. The rationale behind threshold selection is explicated, along with its application in data quality assessments. Additionally, this chapter elaborates on the operationalization of these thresholds through the use of flowcharts to implement data quality consistency checks.

Chapter 6 showcases a pilot study that employs the developed data quality indexes and thresholds in accordance with the prescribed flowcharts. To substantiate the proposed data quality assessment indexes, thresholds, and protocols, raw image inspections are undertaken to address data quality discrepancies in sections that failed the initial data quality assessment.

Chapter 7 culminates in a synthesis of the conclusions derived from the data quality analysis and proffers recommendations aimed at enhancing the quality of automated pavement condition data.

Chapter 2 Literature Review on Quality Assurance of Automated Pavement Condition Data Collection

2.1 Implementation of Automated Pavement Condition Data Collection

2.1.1 Automated Pavement Condition Data Collection

The traditional manual pavement condition survey is based on walking or traveling at a slow speed and noting the existing surface distress (Pierce and Weizel 2019). It is quite a labor-intensive and time-consuming process which is hard to cover the entire roadway length. To overcome the challenges of the manual survey, high-speed automated data collection technologies were widely adopted at network-level pavement condition data collection by many highway agencies. The automated data collection is a process of collecting pavement condition data using imaging technologies or other sensor equipment (McGhee 2004). Data and images collected through automated data collection require processing using either fully or semi-automated methods. For the semi-automated data processing, the collected image and data are processed using imaging technologies or other sensor equipment but involving significant human input during the processing and/or recording of the data (Flintsch and McGhee 2009). The semi-automated method usually processes images at workstations by personnel trained to rate visible cracks and other distresses (Pierce et al. 2013). For the fully automated data processing, the pavement condition is identified and quantified through techniques that require either no or very minimal human intervention (Flintsch and McGhee 2009). The current fully automated uses video and/or laser technology to detect and classify pavement cracking in real-time at highway speeds. Alternatively, the data collection vendors use systems to capture the pavement image first and then detect and classify the cracks using automated post-processing (Pierce and Weizel 2019).

Recently investigators have confirmed that automated data collection technologies have pushed forward the innovation of pavement performance quality assessment (Flintsch and McGhee 2009; McGhee 2004; Pierce et al. 2013; Pierce and Weizel 2019). The automated pavement condition survey has become a commonly acceptable data collection method because of its benefits of minimal impact on traffic, a significant increase in safety, more time efficiency, and the possibility of 100% network coverage. A recent survey of highway transportation agencies by the National Cooperative Highway Research Program (NCHRP) shows that 45 out of 57 responses (46 U.S. highway agencies and 11 Canadian provincial and territorial governments) are using automated data collection methods exclusively, 6 agencies using both manual and automated condition surveys, and only 6 agencies using manual pavement condition surveys (Pierce and Weizel 2019). With the wide application of automated pavement condition surveys, it is important to capture the agencies' experience in their implementation of automated pavement condition data collection.

2.1.2 Data Quality Management Program

High-quality pavement performance data can provide critical information to support decisions involving the Federal-aid program for highway pavements (FHWA 2018). To enhance the quality of the important pavement performance data, the Federal Highway Administration (FHWA) promulgated a rule, the National Performance Management Measures: Assessing Pavement Condition for the National Highway Performance Program and Bridge Condition for the National Highway Performance Program (PM2) (FHWA 2018). The rule PM2, which was effective in 2017, established ride (IRI), rutting, faulting, and cracking percent, or present serviceability rating (PSR) as the pavement condition metrics. The state highway agencies were required to collect and report these pavement condition metrics to the FHWA Highway Performance Monitoring System (HPMS) to determine the pavement performance condition in terms of good, fair, and poor per 23 CFR 490.309(c) (FHWA 2018).

To collect the pavement condition metrics accurately and report the entire highway pavement performance comparably, each state highway agency was required to develop a Data Quality Management Program (DQMP) following the requirements of FHWA and their own states according to 23 CFR 490.319(c). The DQMP is also required by the Moving Ahead for Progress in the 21st Century Act (MAP-21) and Fixing America's Surface Transportation (FAST) Act to evaluate the pavement performance for highway agencies (Simpson et al 2018). A DQMP is a document that defines the acceptable level of data quality and describes how the data collection process will ensure this level of quality in its deliverables and processes (FHWA 2018). Specifically, the DQMP includes methods and processes of five components: 1) data collection equipment calibration and certification; 2) certification process for persons performing manual data collection; 3) data quality control measures to be conducted before data collection begins and periodically during the data collection program; 4) data sampling, review and checking processes; and 5) error resolution procedures and data acceptance criteria (Simpson et al 2018). For state agencies, the DQMP aims to address the errors that occurred due to data collection equipment malfunction, unintended mistakes by operators, computer glitches, mechanical failures, and other issues that can result in poor data and the need for expensive recollection efforts (FHWA 2018). Reviewing state highway agencies' DQMPs could be an efficient way to understand how state highway agencies collect and report their pavement condition data. However, the data metrics vary by agencies. According to state highway agencies' data collection manuals, the data definitions are also unique. Therefore, the DQMPs are good resources to better understand the way that state highway agencies collect the pavement condition data and enhance the data quality.

2.1.3 Automated Data Collection Protocols and Standards

A data collection protocol/standard is a description of the procedures for consistently collecting and recording the pavement condition data in the same manner (FHWA 2018). In accordance with 23 CFR 490.309(c), the pavement condition metrics shall be collected and reported following the standardized HPMS format on an annual cycle for the Interstate roadways and on a 2-year maximum cycle for all other required sections. The HPMS format conforms to ten AASHTO (American Association of State Highway and Transportation Officials) Standards with some modifications specified in the HPMS Field Manual for IRI, cracking percent, rutting for asphalt pavements, and faulting for jointed concrete pavements (FHWA 2018). However, the automated data collection standards are not limited to the HPMS Field Manual associated with the AASHTO Standards. A previous survey shows that some state agencies also use ASTM standards in their automated data collection, especially in measuring profile, macrotexture, and analyzing precision and bias (Pierce and Weitzel 2019). Meanwhile, the Long-term Pavement Performance (LTPP) Distress Identification Manual is also adopted by a few state agencies. Some state highway agencies such as the California Department of Transportation (Caltrans), Pennsylvania Department of Transportation (PennDOT), TxDOT, etc. have their own standards for automated data collection which serve their state-level data collection, analysis, and decision making, Hence, a review of the automated data collection protocols and standards being used by state highway agencies will be included in this study.

2.2 Quality Improvement of Automated Pavement Data Collection

2.2.1 Data Quality Control

According to the AASHTO R10-06, QC includes the activities needed to adjust production processes toward achieving the desired level of quality of pavement condition data (AASHTO 2006). QC contains sampling, testing, inspection, and corrective action (where required) to maintain continuous control of a production process (FHWA 2018). The activities for QC are required by state highway agencies' DQMP and primarily implemented by the data collection team to monitor, assess, and adjust data collection processes (Chang et al 2020). The QC activities may include equipment calibration, software checks and control, verification, or blind site data collection, which are performed during data collection (Pierce and Weitzel 2019). The pavement performance indicators for QC control, verification, or blind site check mainly focus on IRI, rutting, faulting, cracking, location, etc., but the specific requirements/tolerances for the control site checks vary among state highway agencies.

2.2.2 Data Quality Assurance

After data processing and vendor's internal quality check, the pavement condition data are submitted to the agency. The agency team conducts a final data acceptance check for QA. Data acceptance criteria for QA at the agency's final data quality assessment are defined in the state highway agency's DQMP. A review of highway agencies' DQMPs shows that each state agency has its own data sampling rate and method to select samples and conduct QA. The QA criteria are in a wide range depending on state agencies' different needs. However, the major contents of QA include IRI, rutting, faulting, cracking, image, etc. If the submitted pavement condition data does not pass the data acceptance for QA, there are corrective actions for the data collection team to take to prevent erroneous data collection or data analysis procedures from being proceeded (FHWA 2017).

Even with both QC and QA procedures, the state agencies are still struggling with data quality issues during applying the automated data collection technologies. The quality of the automated data varies due to the factors in equipment, algorithms, operation procedures, human interventions, etc. The reason that caused this problem could be the deficient QC and QA during and after the data collection. Therefore, this study aims to review the successful practices and discuss issues that the state highway agencies are using in evaluating the quality of the automated pavement condition data.

2.3 Questionnaire Result of Automated Pavement Data Collection

2.3.1 Data Collection Methods

The questionnaire survey result shows that automated and semi-automated pavement data collection methods have been widely adopted by state highway agencies in the United States. Figure 2.1 summarizes the pavement condition data collection methods currently used by the agencies. As shown in the figure, 32 over the 33 agencies who responded have used automated or semi-automated data collection methods. Among these 32 agencies, 12 of them use automated or semi-automated data collection technologies for more than 10 years, 8 of them have 5 to 10 years of experience, and 5 of them have at least 1 to 4 years of experience (7 state agencies did not respond to this question). This result indicates that each state may be at a different stage of using automated/semi-automated data collection technologies. Specifically, the automated data collection in the California Department of Transportation (Caltrans) still needs manual interventions for QC/QA. Florida DOT uses fully automated Laser Crack Measurement System (LCMS) for HPMS. But for the pavement condition survey, they are still in a transition from manual distress data collection to fully automated ratings. Mississippi DOT uses manual data collection instead of automated for concrete pavement cracking evaluation, which is 3% of the lane miles. Nevada DOT and South Dakota DOT use manual data collection for distress and automated technologies for profile, rutting, and faulting. Alaska DOT uses semi-automated for patching and raveling evaluation.

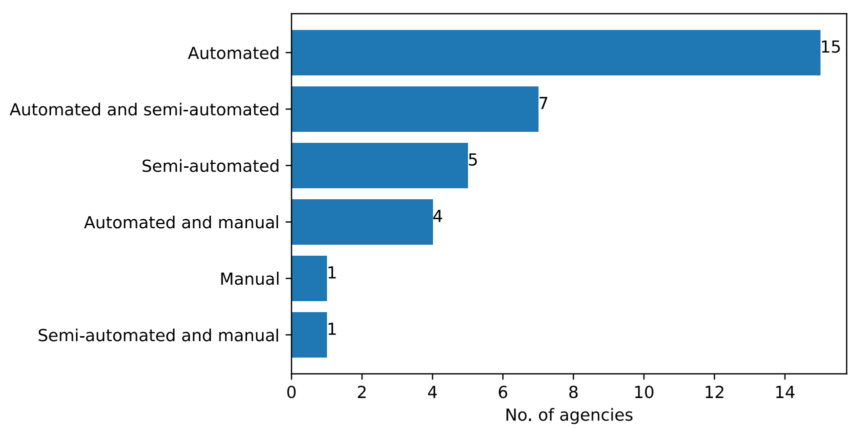


Figure 2.1 Summary of Agency Data Collection Methods (Total # of Responses = 33)

2.3.2 Data Collection Service Provider

The questionnaire survey result summarized in Figure 2.2 shows that there are three ways for state highway agencies to collect pavement condition data. First of all, 20 out of the 33 respondents contract with vendors for pavement condition surveys. Contracting with a vendor is a usual way for state highway agencies which can save a lot of time for engineers and staff. However, some state highway agencies still take additional actions to enhance the quality of vendor's services. For instance, Caltrans has a field crew to perform QC/QA. Indiana

Department of Transportation (INDOT) and Pennsylvania Department of Transportation (PennDOT) collect the project-level pavement condition data by their own staff. Meanwhile, the price of contracting with a vendor can be quite different from that of using in-house staff in conducting the data collection. Secondly, 11 out of the 33 respondents collect the data by their own staff. Some of these state agencies own data collection vehicles. For example, the Minnesota Department of Transportation (MnDOT), Maryland DOT, and Washington DOT. Thirdly, 2 out of the 33 respondents use both the vendor and staff for data collection. For example, Florida DOT collects Interstate highways using LCMS while a vendor collects non-Interstate roads.

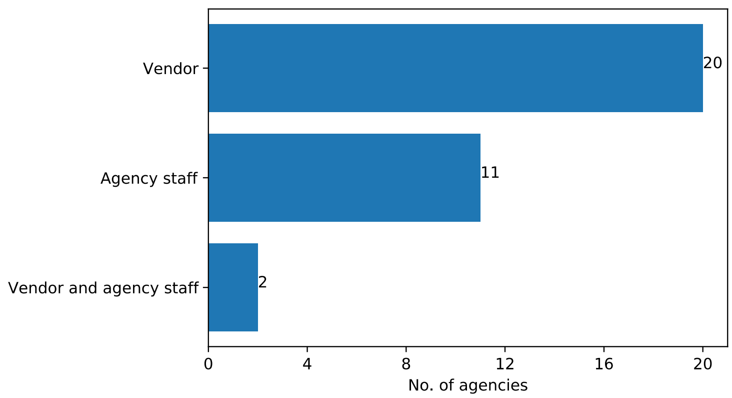


Figure 2.2 Summary of Data Collection Service Providers (Total # of Responses = 33)

2.3.3 Data Collection Protocols

Before the implementation of automated data collection, a state highway agency should specify its data collection metrics and protocol. As mentioned in the background part, the data standards and protocols vary by agencies. Although FHWA requires states to collect and report pavement condition data following the HPMS field manual, generally a state agency has more than one data collection protocol to use. The commonly used protocols include various ASTM standards, AASHTO standards, and the LTPP standard. Furthermore, a lot of agencies have standards of their own design, such as Delaware DOT, Florida DOT, Illinois DOT, MnDOT, Mississippi DOT, Nevada DOT, Nebraska DOT, Ohio DOT, Oregon DOT, South Dakota, TxDOT, Washington DOT, and Wyoming DOT. Specifically, Alabama DOT has a "Network-level Pavement Condition Data Collection Procedure," Caltrans has an "Automated Pavement Condition Survey Manual", and PennDOT has its Publication 336 "Automated Pavement Condition Survey Field Manual."

2.3.4 Data Collection Items

The data items collected by state agencies using automated/semi-automated data collection methods primarily include distress data (different kinds of cracking), roughness (IRI), rutting, and faulting according to FHWA's data report requirements. Some state highway agencies also collect additional items. For example, Arkansas DOT collects macro texture; Caltrans collects mean profile depth (MPD); Florida DOT plans to expand raveling as a separate distress category;

Louisiana Department of Transportation and Development (LAODTD) collects friction texture, macrotexture, horizontal alignment and vertical alignment data, and fill quantity; Mississippi DOT collects friction data; TxDOT collects skid number.

2.3.5 Data Collection Length and Cycle

The data collection length which is collected every cycle depends on the state's roadway network length. Figure 2.3 summarizes the survey results about the state's data collection length and frequency. Out of the 32 respondents, 26 collect the pavement condition data by roadbed miles, 4 states collect pavement data by lane miles (Caltrans, Florida DOT, Georgia DOT, and Montana DOT), and 2 states collect pavement data by centerline miles (Delaware DOT, Ohio DOT, and Tennessee DOT). The centerline mile is defined as the distance measured between the beginning point and the end point shown on the design plan regardless of the number of lanes or roadbeds, and the roadbed mile is defined as the distance along each roadbed regardless of the number of lanes. Among the 32 states that use automated or semi-automated data collections, Texas holds the biggest automated data collection network. Caltrans has the second-longest roadway length conducted with automated pavement condition data collection.

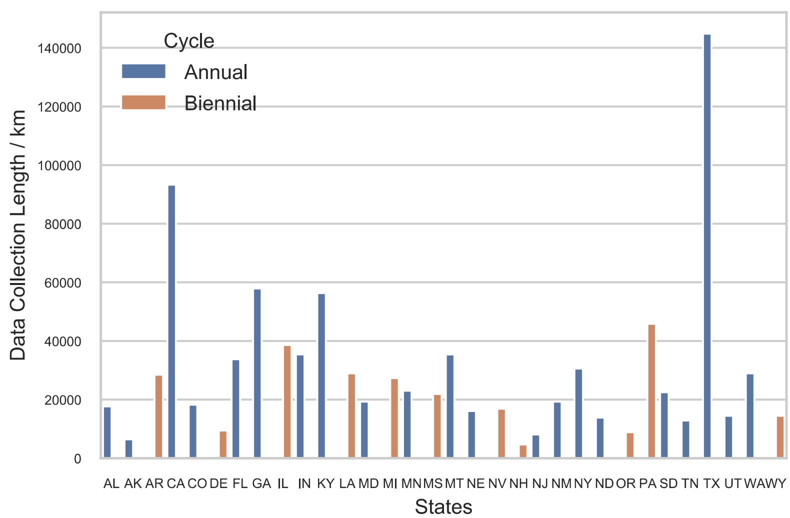


Figure 2.3 Data Collection Lengths and Cycles of State Highway Agencies (32 Responses with Automated or Semi-Automated Data Collection)

In the 2016 Field Manual, FHWA specified that the data collection frequency for Interstate System pavement is annual and for non-Interstate National Highway System (NHS) pavement is biennial (Simpson et al 2020). Both the annual data collection frequency for Interstate System pavement and the biennial data collection frequency for non-Interstate NHS require annual data reporting to HPMS making the most recently collected data replacing the data from the previous data collection cycle. To manage the state roadway network and meet FHWA's data reporting requirements, 21 of the 32 respondent states (blue bars in Fig. 3.) collect all state-maintained roads in their system annually. The rest of the 11 state highway agencies collect the Interstate or both the Interstate and non-Interstate NHS annually but collect the other state-maintained roads biennially.

2.3.6 QC/QA Processes

During a virtual interview, a senior pavement engineer from AgileAssets Inc. highlighted that "Pavement survey accuracy is really important because it concerns multi-million-dollar maintenance plan." However, the accuracy of the existing automated survey technologies can be easily affected by survey equipment. Therefore, the QC before and during the data collection and QA after the data collection are crucial to enhance the quality of the pavement condition data.

The QC activities include automated data collection equipment certification, verification, and calibration. Table 2.1 lists the QC activities taken by the 32 responding state highway agencies using automated or semi-automated data collections. The result shows that most of the state highway agencies conduct equipment certification, verification, and calibration for cracking, IRI, and rutting by vendors and staff. Some of the state highway agencies contract with an independent third party for equipment certification, but very few agencies use a third party for verification and calibration. The result also indicates that some state highway agencies only apply verification and calibration for IRI and rutting, but not for cracking.

Table 2.1 Quality control of automated pavement data collection at state highway agencies

		Vendor/contractor	**Agency staff**	**A third party**
Who does the equipment certification for	distress data (cracking)	AK, CO, DE, GA, IL, IN, KY, LA, MD, MI, NE, NY, NM, WY	AL, AR, IL, KY, MD, MN, MS, MT, NV, NH, SD	AL, CA, FL, GA, TX, WA
	roughness (IRI)	AK, AR, CO, DE, GA, IL, IN, KY, LA, MI, NE, NY, NM, WY	AR, IL, MD, MI, MN, MS, MT, ND, NV, NH, OR, PA, SD	AL, AK, CA, FL, GA, NH, NJ, TN, TX
	rutting	CO, DE, GA, IL, IN, KY, LA, MD, MI, NE, NY, NM, TN, WY	AL, AR, IL, MD, MN, MS, NV, NH, PA, SD, WA,	CA, FL, GA, TX
Who does the equipment *verification for	distress data (cracking)	AL, AK, CO, DE, GA, IL, IN, KY, LA, MD, MI, NY, NM, OR, TN, TX, WY	AR, CA, FL, IL, MD, MI, MN, MT, NV, NE, NH, PA, SD, WA	FL, NJ
	roughness (IRI)	AL, AK, CO, DE, GA, IL, IN, KY, LA, MD, MI, NY, NM, OR, TN, TX, WY	AR, CA, FL, IL, MD, MI, MN, MS, MT, ND, NV, NE, NH, NM, PA, SD, WA	FL, NJ
	rutting	AL, AK, CO, DE, GA, IL, IN, KY, LA, MD, MI, NY, NM, OR, PA, TN, TX, WY	AR, CA, FL, IL, MD, MI, MN, MS, MT, NE, NH, NJ, SD, WA,	FL
Who does the equipment	distress data (cracking)	AL, FL, GA, IL, IN, KY, LA, MI, NE, NH, NY, NJ, NM, WY	AK, AR, CA, FL, IL, MD, MI, MN, NE, NH, OR, PA, SD, TN,	FL, NJ, TX

**calibration for			WA,	
	roughness (IRI)	CO, GA, IL, IN, KY, LA, MI, NE, NH, NY, NM, WY	AK, AR, CA, IL, MD, MI, MN, MS, MT, ND, NV, NE, NH, OR, PA, SD, TN, TX	AL, NJ
	rutting	CO, FL, GA, IL, IN, KY, LA, MD, MI, NE, NH, NY, NJ, NM, WY	AL, AK, AR, CA, FL, IL, MD, MI, MN, MS, NV, NE, NH, OR, PA, SD, TN, TX, WA	FL
who does the data acceptance check		MD, NH, TN, TX	AL, AK, CA, CO, DE, FL, GA, IL, KY, MD, MI, ND, NV, NE, NH, NY, NJ, NM, OR, PA, SD, TN, TX, WA, WY	DE, NM, TX

Note. *Verification: weekly check that the inertial profiler for IRI measurements and the 3D systems for rut measurements are in good operating conditions; **calibration: comparison of data collected using an inertial profiler and skid trucks with those of a reference device (TxDOT 2018).

The QA activities are involved in the data acceptance check process which includes data allowable range check, data quality validation, and data sampling checks with a specific sampling rate and method for the automated pavement condition survey. Table 2.2 shows the QA activities taken by the 32 respondents using automated or semi-automated data collection. The result indicates that most of the state highway agencies have data allowable range checks as well as data quality validation processes for distress data, IRI, rutting, and faulting. These state highway agencies also conduct data sampling processes with different sampling rates and sampling methods. The sampling rates for distress data are mainly in the range of 0.5%-10%. In some states, the sampling rates for distress data can be 25%, 35%, and even 100%. For IRI, rutting, and faulting, more states are exercising a sampling rate of 100% of the collected network length, than the states that apply sampling rates of 0.5%-10% (except for Illinois DOT who uses a sampling rate of 50% for IRI and rutting). The most commonly used sampling method is random sampling by picking a desired sample size (% of the surveyed state network pavement sections, or population) and selecting observations from the population. Systematic sampling and stratified sampling are also used by many state highway agencies. Systematic sampling is conducted by selecting sample units or elements (pavement sections) of a population at a regular interval determined in advance. Stratified sampling is applied by dividing the sample elements (pavement sections) of a population (all the pavement sections in the state-maintained network) into subgroups or strata, and then randomly selecting elements from each of these strata. Generally, there are more similarities between elements within a stratum than elements in different strata. Different from other states, Caltrans uses cluster sampling, which is very similar to the stratified sampling, by dividing the population into multiple groups or clusters, and then selecting random elements from these clusters.

Table 2.2 Quality assurance of automated pavement condition data at state highway agencies

	Distress data (cracking)	Roughness (IRI)	Rutting	Faulting
Are there any data allowable range checks	AL, AK, CA, CO, DE, FL, IN, KY, LA, MD, MS, ND, NV, NE, NY, NJ, NM, OR, PA, SD, TN, TX, WA, WY	AK, CA, CO, DE, FL, GA, IL, IN, KY, LA, MD, MI, MS, ND, NV, NE, NH, NY, NJ, NM, OR, PA, SD, TN, TX, WA, WY	AK, CA, CO, DE, FL, GA, IL, IN, KY, LA, MD, MI, MS, NV, NE, NH, NY, NM, OR, PA, SD, TN, TX, UT, WA, WY	CA, DE, FL, IL, IN, KY, LA, MI, MS, NV, NE, NY, NM, PA, SD, TN, UT, WY
Does your agency have any data quality validation process	AL, AK, CA, CO, DE, FL, GA, IL, IN, KY, LA, MD, MI, ND, NE, NH, NY, NJ, NM, OR, PA, SD, TN, TX, UT, WA, WY	AL, AK, CA, CO, DE, FL, GA, IL, IN, KY, LA, MD, MI, MS, ND, NE, NH, NY, NJ, NM, OR, PA, SD, TN, UT, WY	AL, AK, CA, CO, DE, FL, GA, IL, IN, KY, LA, MD, MI, MS, NE, NH, NY, NM, OR, PA, SD, TN, WY	AL, DE, FL, IL, IN, KY, LA, MI, MS, NE, NY, NM, PA, SD, TN, WY
Does your agency have any data sampling process and what is the sampling rate	AL (3%), AK (5%), CA (0.5-1%), CO (1%), FL (5%), GA (5%), IL (25-35%), KY (10%), LA (5%), MD (100%), MI (1%), ND (2%), NV (10%), NE (100%), NH (25%), NY (10%), NJ (5%), PA (2.5%), SD (100%), TN (2%), TX (6%), UT (5-10%), WA (5%), WY	AK (5%), CA (0.5-5%), FL (10%), GA (5%), IL (50%), KY (100%), MD (100%), MS (100%), NE (100%), NH (100%), NY (10%), NJ (5%), PA (2.5%), SD (100%), TN (2%), UT (5-10%), WY	AK (5%), CA (0.5-5%), FL (10%), GA (5%), IL (50%), KY (100%), MD (100%), MS (100%), NE (100%), NH (100%), NY (10%), SD (100%), TN (2%), WY	FL (10%), GA (5%), KY (100%), MS (100%), NE (100%), NY (10%), PA (2.5%), SD (100%), WY
What is the sampling method	AL (stratified), AK (systematic), CA (cluster), CO (random and stratified), FL (random), GA (random), IL (random), KY (systematic), LA (random), MD (systematic), MI (stratified, random, and systematic), ND (stratified), NV (random, and systematic), NH (systematic), NY (random), PA (random), TN (systematic), TX (random), UT (stratified), WA (random), WY (random)			

Furthermore, these QA activities for data acceptance checks are mainly conducted by the agency staff which generally take much of their time. Only a few state highway agencies are working together with a vendor or a third party to conduct the data acceptance process.

One of the open questions in the questionnaire is about the data quality issues that the state highway agencies are facing. Table 2.3 summarizes some typical data quality issues and possible reasons from the responses of the state highway agencies. Eight states mentioned issues about cracking data, e.g., cracking identification/determination, cracking detection, and cracking classification. Some state agencies have data quality issues with specific pavement types, such as jointed concrete pavement (JCP). The IRI data collection has caused issues in some state agencies, especially within the urban areas. The IRI sensors are very sensitive to the traffic environment, and the reasons that cause the IRI issues could be the low vehicle speeds and frequent stops due to traffic signals. In addition, another issue that has been raised is alignment of the vendor collected data with the state referencing systems and standards. Potentially, there could be more data quality issues from the states that did not respond to the questionnaire survey.

Table 2.3 Data quality issues of state highway agencies

State highway agency	Data quality issues and possible reasons
Alabama DOT	1) Cracking data has been underreported by vendor since the beginning. It's getting better. 2) OGFC remains a challenge. The vendor may have trouble rating it.
Alaska DOT	1) Low speed IRI collection, which is likely a challenge in most states in urban areas. 2) Occasionally vendor's cracking identification misses some cracks, but that has not been a large issue overall and is normally very isolated. 3) The largest issue is probably aligning the vendor collected data to states linear referencing system for HPMS reporting.
Caltrans	1) Vendors turn over. 2) Accurate execution of automated pavement data collection is a major issue. 3) At network-level, we need to accept imperfection for localized issues; but focus on project development. 4) Accurate cracking determination appears to be the most challenging.
Colorado DOT	1) Corner Breaks are interpreted manually. 2) The vendor collected data did not align with the Long-term Pavement Performance (LTPP) definition but was corrected.
Maryland DOT	1) Data quality issues do arise, but sophisticated data quality assurance and quality control checks are in place to address them. 2) These issues arise due to the nature of the data collection procedures, personnel changes in equipment operations and data processing. 3) Continuous refinement of the processes, training of new staff, and well documented Standard Operating Procedures (SOPs) allow for effective resolution of issues.
Minnesota DOT	The biggest issue we have with automated distress classification is on JCP.
Mississippi DOT	We are aware that the pavement type is crucial in the distress classification. The contractor may have issues to classify the pavement type.
Nevada DOT	1) Certain types of distress data are less reliable because so many people are involved in the collection effort. 2) We are slowly transitioning to a more centralized approach that should make it more reliable.

New Jersey DOT	Traffic lights and traffic congestion impact the quality of the IRI data in those locations.
Oregon DOT	1) Distresses rated manually from the pavement images are more likely to have problems. 2) For concrete pavement, separating important cracks from unimportant map cracking is an issue. 3) For asphalt pavement, patching, potholes, and raveling can be an issue, especially with regard to capturing the proper severity level.
PennDOT	1) Data quality is mostly limited to right edge deterioration and left edge joint distress on asphalt pavements. 2) Due to the limitation of the imaging system to capture the full extent of the lane in some cases. Also due to limitations of the crack detection software in identifying these two distresses.
South Dakota	There have been isolated issues from time to time. Usually, an equipment malfunction has been to blame.
Tennessee DOT	Data variability. The reason could be operation issues and the quality of the downward 3-D images.
Utah DOT	1) One of the biggest headaches was matching the Location Referencing System. 2) Another was how to handle routes that were closed/under construction as well as any need to recollect data.
Wyoming DOT	Consistency of automated crack detection on JCP.

In addition, the lack of a standard for the format of automated pavement condition surveys has been another problem in QC afflicting pavement engineers for a long time. AASHTO has recently approved a new standard specification (Pavement Standard Image, or PSI) to define the 2-Dimensional and 3-Dimensional (2D/3D) pavement image data format for pavement surface condition and profile surveys. This standard provides a uniform format for automated pavement condition surveys across the country which could decrease the unit price of the automated pavement condition survey. Therefore, for state highway agencies, there are some federal regulations to specify how automated pavement survey should be conducted and how the data quality should be handled. However, for municipal governments, there is no standard for automated data collection. The requirements are quite loose as the municipal governments have no clear expectations for their data collection vendors.

2.3.7 Data Collection Cost

Cost is a big concern when the state and local agencies switch to automated data collection. Many interviewees from both the government and industry believed that the current automated data collection services are too expensive. An engineer from NCE company shared that the cost of manual data collection is as low as $15 per hour. However, the price of high-quality automated data collection could be $100-$150 per mile. VDOT spends about $100-$200 per mile for an automated pavement condition survey which includes an independent third party for QA by manually reading the image data. The cost of automated data collection is quite sensitive for the customers (agencies) such as small cities and counties. For the City of Nevada at Iowa, there were five vendors bidding for the contract of city-level automated pavement condition survey.

After an evaluation of the price and the service quality, the price of the pavement condition survey from the chosen vendor was at $105/mile. Different from other state and local agencies, MnDOT conducts automated pavement condition data collection by itself. One significant advantage is cost reduction. The current cost is approximately $40/mile for the annual survey at MnDOT. MnDOT replaces their survey vans every 5-6 years, and in an average the total data collection cost is around $55/mile.

The final contract with a data collection vendor includes a per mile based cost and a fixed price cost for the project. The unit cost of the network-level pavement condition data collection depends on the state agency's requirements on collected network length, measurement items, featured information, QC/QA, and timing. Therefore, in many cases, the price for high-quality pavement condition data is unpredictable. An engineer from Applied Pavement Technology, Inc. (AP Tech) mentioned that they adopted a couple of procedures to make sure the survey data is accurate. Each procedure would add a certain amount of cost to the total cost. If survey data is proved acceptable without manual intervention, only 10% more cost would be added. If not, an unpredictable cost may be needed to make the data acceptable to the end-user. Therefore, many engineers suggested that reducing data collection costs and data processing time be urgent needs for automated data collection.

2.4 Problems with Existing Automated Data Collection

2.4.1 Data Quality of Automated Data Collection Technologies

Most of the interviewees agreed that automated data collection is the right direction to improve the work efficiency of pavement engineers. However, the current automated pavement data collection technologies still have a lot of room for improvement, especially for the image data processing algorithms. Many pavement engineers claimed that data quality is a serious issue with the current automated data collection technologies. Some interviewees pointed out that data inconsistency and discrepancy are the main issues for state and local agencies after switching to automated data collection. Take as an example, a Pavement Management supervisor at TxDOT said "Data inconsistency and false-positive cost us extra time for data validation, and it also creates troubles for us to serve the other functional departments in TxDOT."

Inconsistency means the unexpected differences between two or more repeated runs of automated data collection at the same pavement sections. Discrepancy stands for the unaccepted differences between the true distress values and the collected data at the same pavement locations. A typical manifestation of discrepancy is false-positive which is the result of inaccurate pavement distress detection. An engineer from Roadway Asset Services (RAS) concluded that the inconsistency between different pavement condition survey systems and the inconsistency between human rating and automated systems are currently among the biggest challenges. As an example, the City of Austin used 3 vendors to collect data at different times, and the data consistency has been a big issue. The main reason is the vendors all use proprietary image data formats that literally prevent sharing and cross-check of data among vendors. Several pavement engineers mentioned that the current automated pavement survey technologies tend to raise the rate of false-positive, which has caused a significant discrepancy problem.

Meanwhile, some highway agencies are also having troubles in matching automated data with historical data that were collected manually. This data continuity issue was also mentioned by many engineers from state and local agencies in interviews. In addition, an engineer from Quality Engineering Solution Inc. (QES) mentioned that the current technologies have trouble in concrete pavement surveys for patch/sealed cracking detection, crack type classification, and crack severity quantification. The positive aspect is that the vendors all have provided timely and effective technical support services when the data quality issues were reported.

In contrast, several engineers acknowledged that they are quite satisfied with the current automated data collection technologies, especially during the Covid-19 lockdown time. These engineers also believed the data inconsistency and discrepancy issues were just normal and acceptable. Meanwhile, FHWA checks the annual report submitted by state highway agencies. Most of the annual reports are based on automated data and only a small percentage of the reports are found to have data issues.

2.4.2 Promoting Automated Data Collection Technologies

One pavement engineer with experience in state highway agency, industry, and academia shared that the current automated data collection technologies are at the entry-level to fully automated data collection (without human interruption). Another senior pavement engineer from AgileAssets Inc. commented that the current automated pavement survey is not fully automated. For instance, patches still need manual labor work in detection. More pavement engineers' feedback shows that the semi-automation of pavement survey still requires a huge amount of manual labor for pavement inspection. Therefore, the current automated and semi-automated pavement survey technologies still are not yet fully automated and have limitations.

The information gained from the interviews shows that the current data accuracies for automated pavement survey companies are around 70-80%, but 95% accuracy is expected. The engineers from survey companies insisted that the current automated technologies need to be innovated, and the artificial intelligence (AI) technologies should be applied to improve the data quality. There are some companies that started using AI technologies for automated pavement data processing. For example, deep learning algorithms have been used for automated data detection, classification, and quantification. However, the interviewees from academia pointed out that the current deep learning method being used in the automated data collection technologies still needs data pre-treatment. The lack of training data due to low availability of annotated ground truth image data and difficulties in sharing data in the public domain has caused delays in developing and using AI in the technologies. Furthermore, the current AI-driven automated pavement condition survey technologies are not able to detect all types of pavement distresses. An important reason is that the current distress definition standard is designed for human raters but not for computer visions, therefore, some of the distresses can hardly be detected or measured by the current automated technologies.

2.4.3 Implementation of QA

As mentioned above, the main issues with current automated data collection technologies are data inconsistency and discrepancy. Manual correction is needed to make the data usable. This incompetency has been brought up in many interviews with pavement management engineers at state agencies. The vendors have internal QA processes, but still could not satisfy highway agencies' data quality requirements. Several pavement engineers from state and local highway agencies would not trust the survey data before validation. Many interviewees from state highway agencies indicated that they spent a lot of staff time doing image checks for data QA after receiving the automated pavement condition data. In several states, it even took engineers months to validate the yearly pavement survey data. For instance, a district engineer at TxDOT mentioned that it is always hard to verify the data from the whole network since it would cost months of time for engineers to go over all the data. In Mississippi DOT, the IT staff and pavement engineers check the image data and the historical Pavement Management System (PMS) data and make corrections to the information in the PMS. A pavement engineer shared that the data validation in Caltrans is conducted manually by three engineers working full time. It is time-consuming and labor-intensive, and there is a lot of subjectivity too. This feedback mirrors the findings learned from the reviews of state highway agencies' DQMPs in that the most labor-intensive checks were image checks, though the manual image checks only represented a subset of the data.

After interviewing with engineers from the government and industry, we found many state and local agencies contracted with third parties to examine the survey data which was delivered from the data collection vendors. For instance, VDOT is contracting with QES. It shows that the state and local agencies are spending lots of budget just to make the data right. In addition, some state highway agencies and municipal governments separate the automated data collection, data processing, and QA as individual services contracting with different entities to conduct the pavement condition evaluation work.

More suggestions for the implementation of QA are about quantifying QA. An engineer from Applied Research Associates, Inc. shared that a threshold could be used to define the data quality for QA purposes, but the value of the threshold depends on the needs of different highway agencies.

Chapter 3. Pavement Condition Historical Data Analysis

3.1 Data Cleansing and Overview

3.1.1 Dataset from 25 Districts

Table 3.1 lists each district's pavement sections under three pavement types (ACP, CRCP, and JCP) from TxDOT's PMIS annual rating data. These data are collected from FY 2017 to FY 2021. FY 2017 is the first year that TxDOT contracted with a vendor to collect the network level pavement condition data using automated pavement data collection technologies. Pavement condition data in FY 2021 is the most recently collected pavement condition data recorded in the PMIS.

Table 3.1 PMIS annual rating dataset in different pavement types (25 districts)

RESPONSIBLE DISTRICT	CRCP	JCP	ACP	Sum	CRCP_%	JCP_%	ACP_%
01 - PARIS	815	588	35607	37010	2.202	1.589	96.209
02 - FORT WORTH	7737	183	32965	40885	18.924	0.448	80.629
03 - WICHITA FALLS	1441	253	30762	32456	4.440	0.780	94.781
04 - AMARILLO	2284	15	44432	46731	4.888	0.032	95.080
05 - LUBBOCK	2508	2	57806	60316	4.158	0.003	95.839
06 - ODESSA	29	51	40331	40411	0.072	0.126	99.802
07 – SAN ANGELO	27	37	36495	36559	0.074	0.101	99.825
08 - ABILENE	87	27	42011	42125	0.207	0.064	99.729
09 - WACO	2442	400	35888	38730	6.305	1.033	92.662
10 - TYLER	348	149	41338	41835	0.832	0.356	98.812
11 - LUFKIN	319	88	30696	31103	1.026	0.283	98.691
12 - HOUSTON	19953	1507	20522	41982	47.528	3.590	48.883
13 - YOAKUM	337	39	40421	40797	0.826	0.096	99.078
14 - AUSTIN	2283	6	37925	40214	5.677	0.015	94.308
15 - SAN ANTONIO	296	40057	10134	50487	0.586	79.341	20.072
16 - CORPUS CHRISTI	4	0	34617	34621	0.012	0.000	99.988
17 - BRYAN	511	130	35462	36103	1.415	0.360	98.225
18 - DALLAS	9189	7194	30380	46763	19.650	15.384	64.966
19 - ATLANTA	925	421	28848	30194	3.064	1.394	95.542
20 - BEAUMONT	1452	2681	23356	27489	5.282	9.753	84.965
21 - PHARR	79	26	27810	27915	0.283	0.093	99.624
22 - LAREDO	184	101	24804	25089	0.733	0.403	98.864

23 - BROWNWOOD	95	6	28476	28577	0.332	0.021	99.647
24 - EL PASO	4223	6	18758	22987	18.371	0.026	81.603
25 - CHILDRESS	651	222	26782	27655	2.354	0.803	96.843
Total	58219	54189	816626	929034	6.267	5.833	87.901

3.1.2 Overview of 25 Districts

At the third stage, the research team started to analyze the PMIS annual rating data and audit data for all 25 districts. Figure 3.1 to Figure 3.6 present an overview of the PMIS annual rating data of 25 districts based on various aspects.

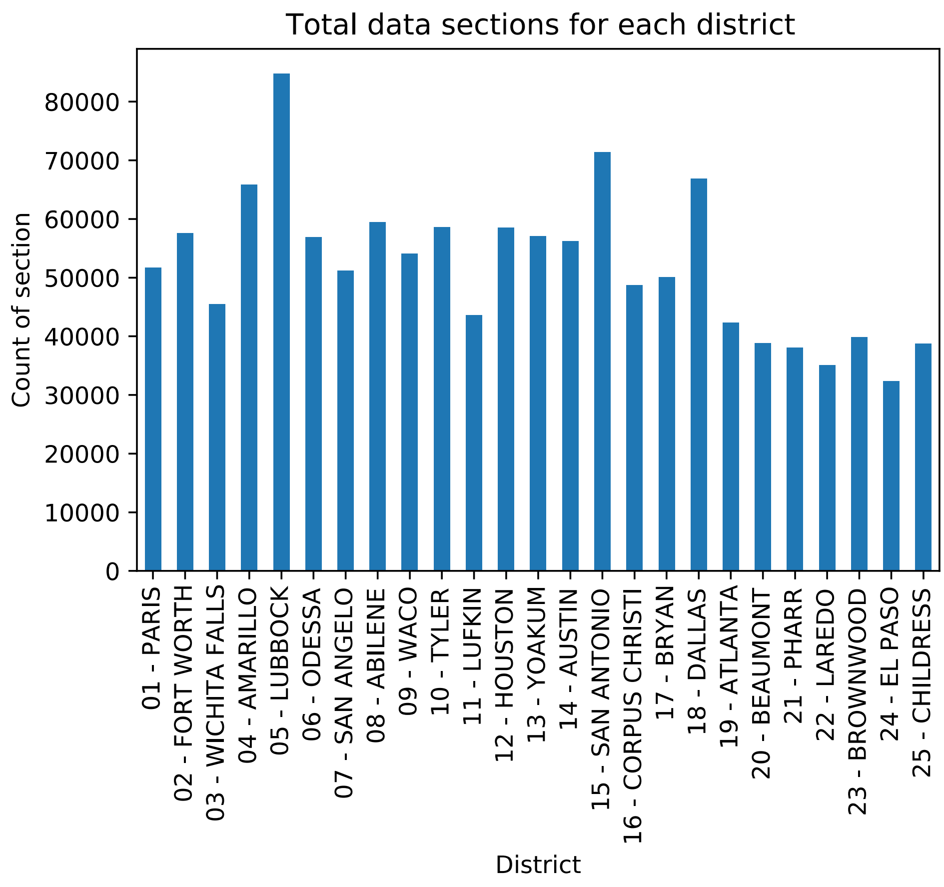


Figure 3.1 Total Data Sections for Each District

Figure 3.2 Data Collection Sections of 25 Districts

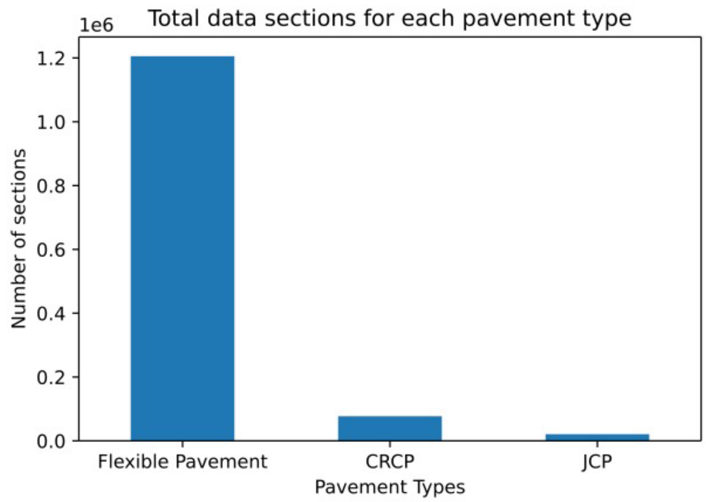


Figure 3.3 Data Collection Sections of Each Pavement Type

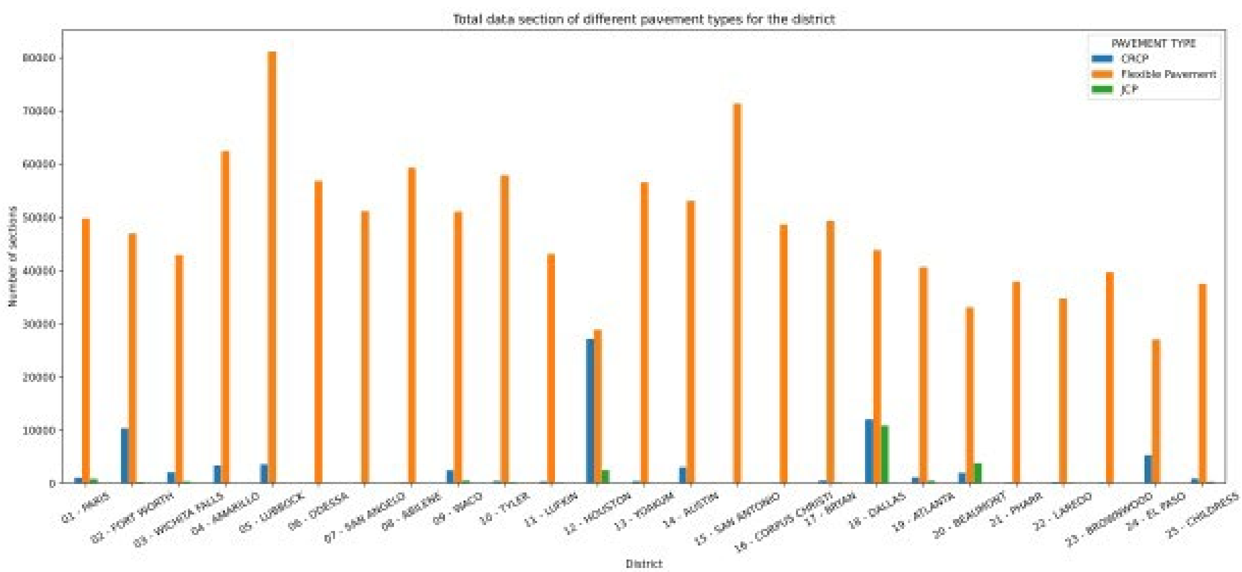


Figure 3.4 Data Collection Sections of Pavement Types For 25 Districts

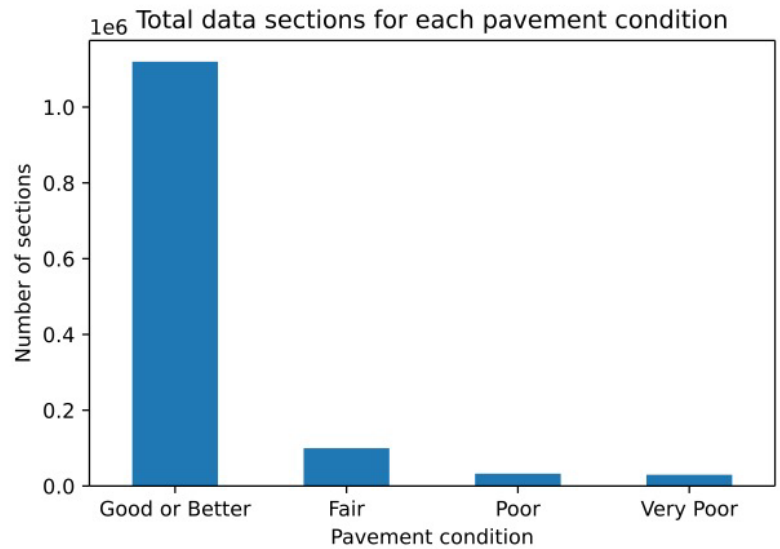


Figure 3.5 Data Collection Sections of Each Pavement Condition

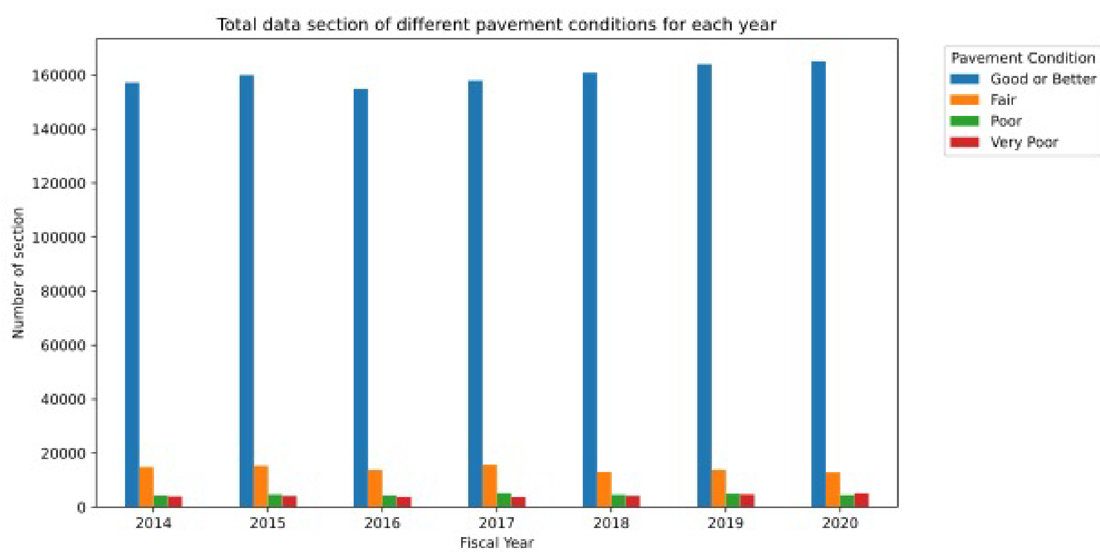


Figure 3.6 Data Collection Sections of Different Pavement Condition for Each Year

3.2 Data Normality Test

3.2.1 Normality Test Rule

In statistical analysis, many methods require a precondition of the data follow a normal distribution. Therefore, the normality test is important before the data analysis. There are two kinds of skewed distributions that are different from the normal distribution. 1) Negative skew: The left tail is longer; the mass of the distribution is concentrated on the right of the figure. The distribution is said to be left-skewed, left-tailed, or skewed to the left, and the mean is skewed to the left of a typical center of the data. 2) Positive skew: The right tail is longer; the mass of the distribution is concentrated on the left of the figure. The distribution is said to be right-skewed, right-tailed, or skewed to the right, and the mean is skewed to the right of a typical center of the data. Figure 3.7 shows the typical shapes of the normal distribution and the two skewed distributions.

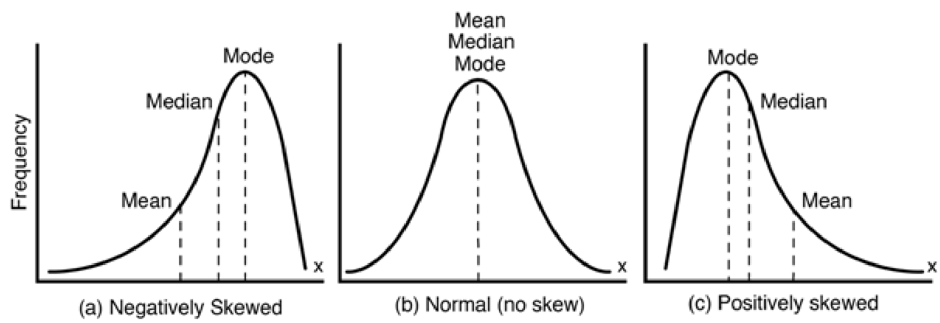
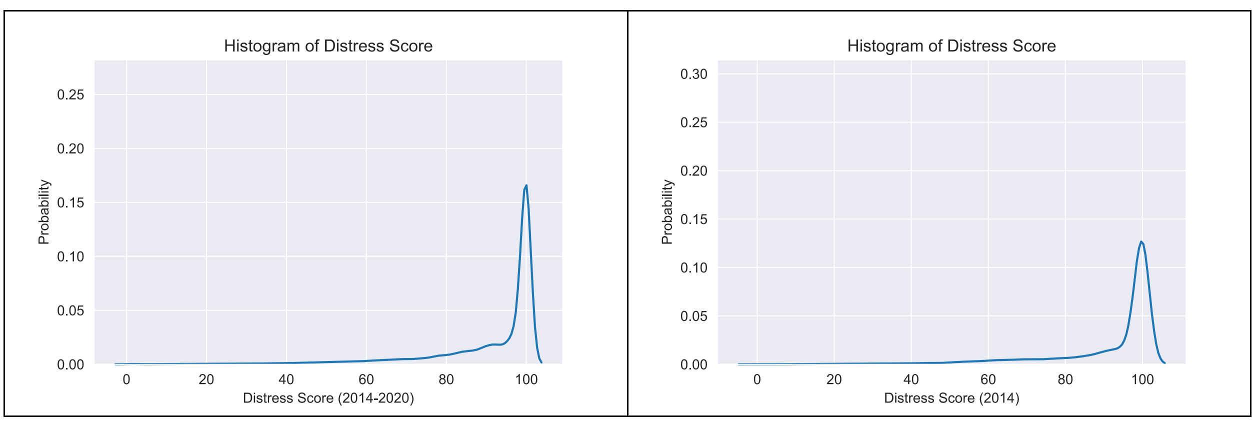


Figure 3.7 Three Kinds of Distributions

3.2.2 Normality Test

In order to find appropriate methods for data analysis, the normality test was conducted to check the PMIS annual rating data. Figure 3.8 presents the distress score histograms of all the 25 districts from FY 2014-2020. The results show that the distress scores from individual sections do not follow a normal distribution. Figure 3.9 shows every two years' distress score differences of all the 25 districts from FY 2014-2020, the results indicate that histograms of two years' distress score differences follow a normal distribution. Therefore, the changes of distress score, ride score, and condition score will be used in future study.



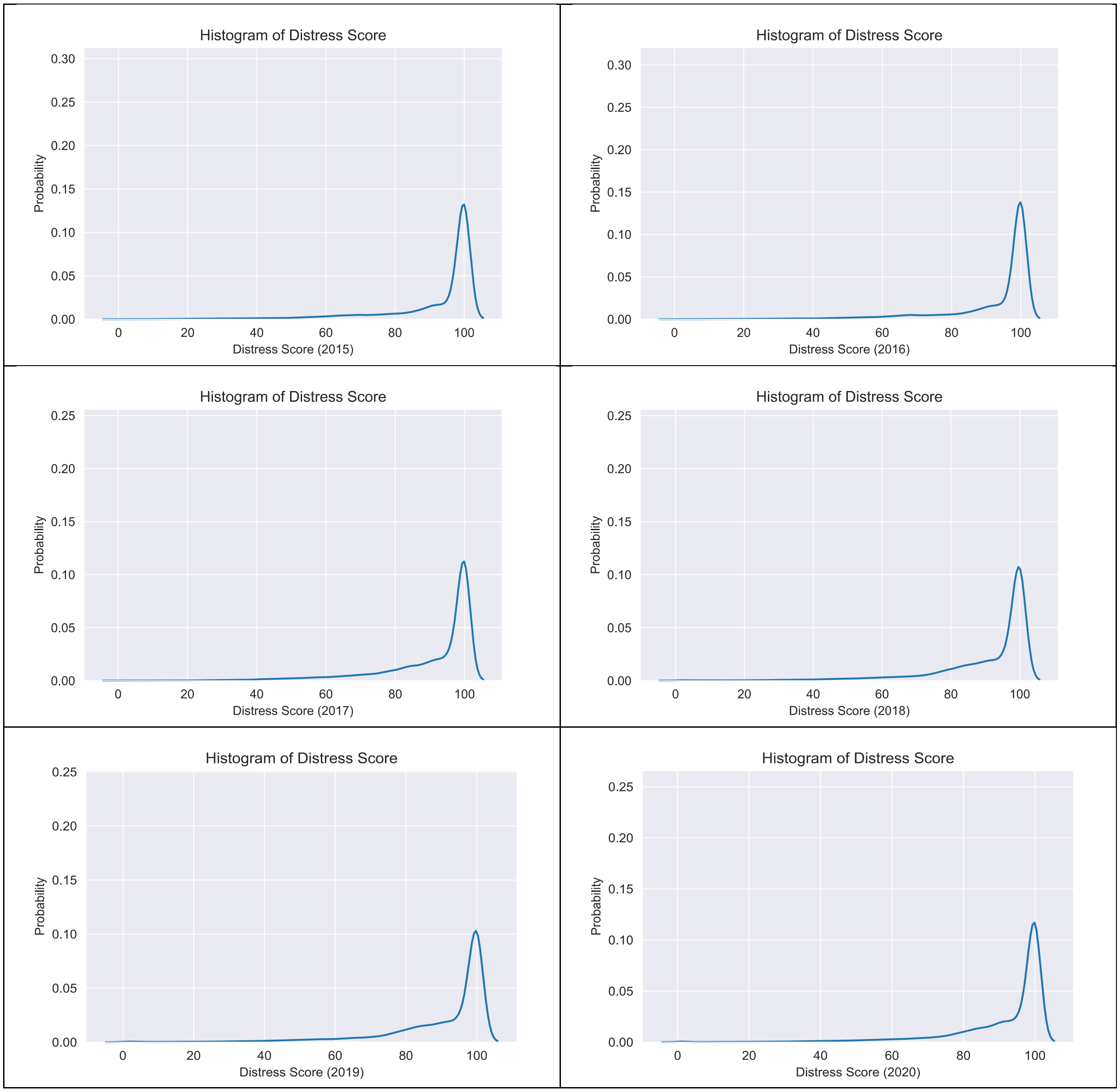


Figure 3.8 Distress Score Histograms of all the 25 Districts from FY 2014-2020

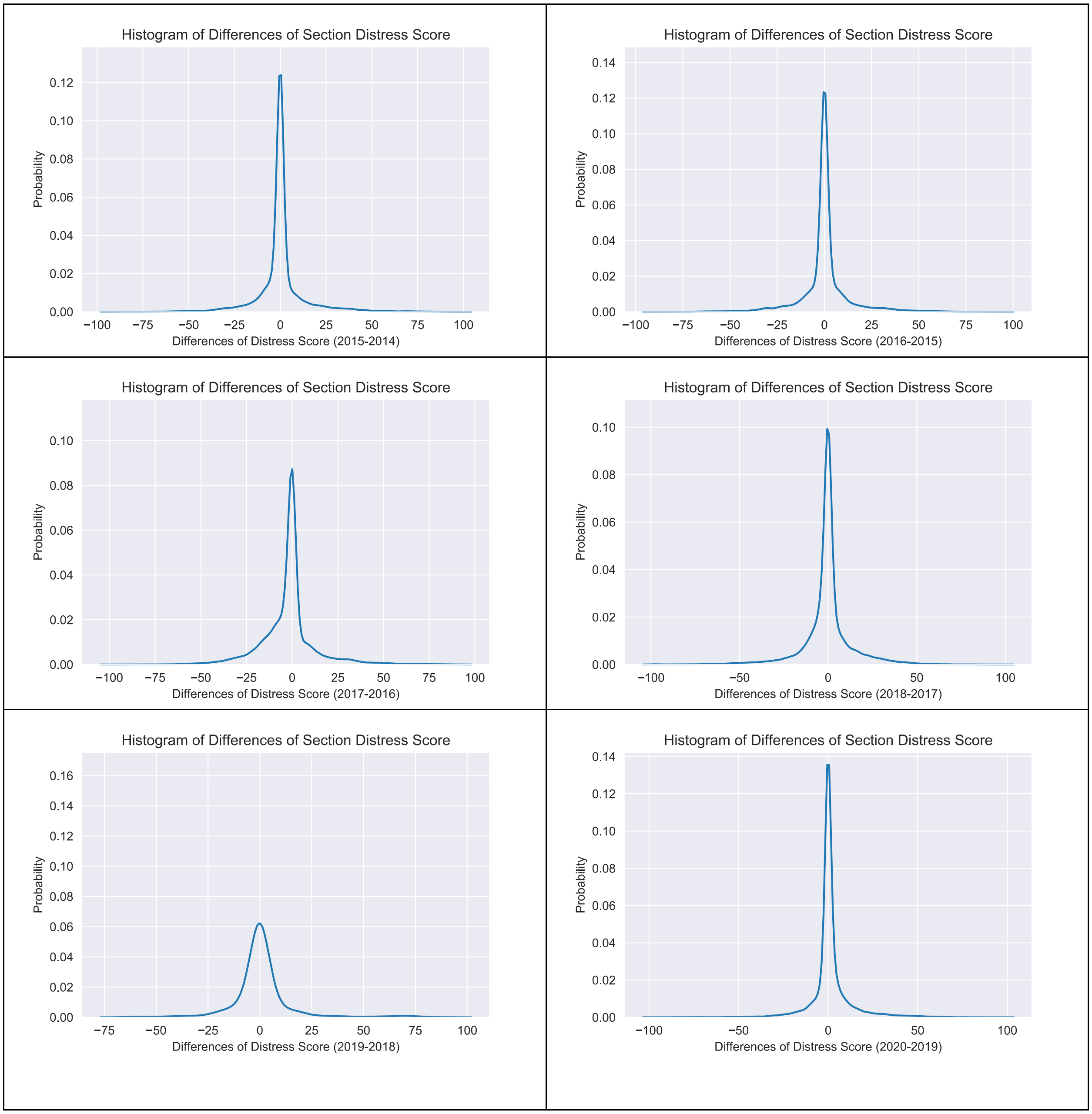


Figure 3.9 Distress Score Change Histograms of all the 25 Districts from FY 2014-2020

3.3 Automated Data Accuracy and Precision

3.3.1 Data Preparation for Accuracy and Precision Analysis

For accuracy calculation, the research team merged PMIS annual rating data with audit data by following the features of 'FISCAL YEAR', 'RESPONSIBLE DISTRICT', 'SIGNED HWY AND ROADBED ID', 'DIRECTION', 'LANE CODE', 'BEGINNING TRM NUMBER', 'BEGINNING TRM DISPLACEMENT', 'ENDING TRM NUMBER', and 'ENDING TRM DISPLACEMENT'.

The data were merged following the rule of the “merge” function in python. Figure 3.10 shows the way how the data were merged. If the right matrix finds one row that has the same values of key1 and key2 as row 0 in the left matrix, column C and column D of row 0 in the right matrix will be added to the left matrix. If the right matrix finds two rows that have the same values of key1 and key2 as row 2 in the left matrix, row 2 will be duplicated and column C and column D of rows 1 and 2 in the right matrix will be added to the left matrix.

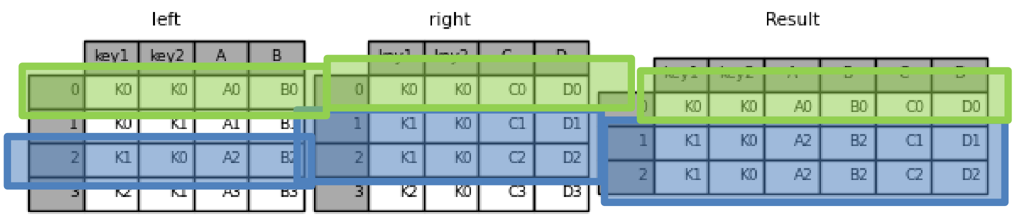


Figure 3.10 Rule of the “Merge” Function in Python

3.3.2 Methodology of Accuracy and Precision Analysis

Precision and accuracy are two ways that scientists think about error. Accuracy refers to how close a measurement is to the true or accepted value. Precision refers to how close measurements of the same item are to each other. Precision is independent of accuracy. That means it is possible to be very precise but not very accurate, and it is also possible to be accurate without being precise. The best quality scientific observations are both accurate and precise.

A classic way of demonstrating the difference between precision and accuracy is with a dartboard. Think of the bulls-eye (center) of a dartboard as the true value. The closer darts land to the bulls-eye, the more accurate they are (Figure 3.11).

- If the darts are neither close to the bulls-eye, nor close to each other, there is low accuracy and low precision.
- If all the darts land very close together, but far from the bulls-eye, there is high precision, but low accuracy.
- If the darts are all about an equal distance from and spaced equally around the bulls-eye there is mathematical accuracy because the average of the darts is in the bulls-eye. This represents data that is highly accurate, but low precise.
- If the darts land close to the bulls-eye and close together, there is both high accuracy and high precision.

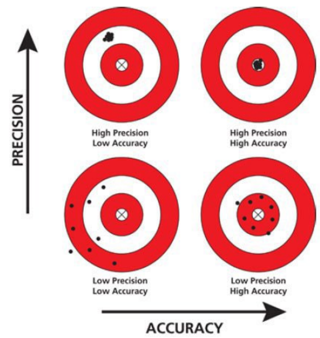


Figure 3.11 Dartboards Showing Different Accuracy and Precision Scenarios

		Predicted condition	
	Total population = P + N	Positive (PP)	Negative (PN)
Actual condition	Positive (P)	True positive (TP), hit	False negative (FN), type II error, miss, underestimation
	Negative (N)	False positive (FP), type I error, false alarm, overestimation	True negative (TN), correct rejection

Figure 3.12 Confusion Matrix Showing Accuracy and Precision

To extract more information about model performance the confusion matrix is used. The confusion matrix helps us visualize whether the model is "confused" in discriminating between the two classes. As seen in Figure 3.12, it is a 2×2 matrix. The labels of the two rows and columns are Positive and Negative to reflect the two class labels. In this example, the row labels represent the ground-truth labels, while the column labels represent the predicted labels. The accuracy and precision can be calculated using the following equations.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{2}$$

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

3.3.2 Examples of Precision and Accuracy Analysis for Automated Data

The precision and accuracy analyses of the automated pavement condition data follow the methodology introduced above. There are two scenarios in precision and accuracy analyses based on the data availability.

Scenario I: Precision

The data used in Scenario I for data precision include the PMIS annual rating data of the 25 districts from FY 2014 to FY 2020 (will include 2021 in the future). The rule to determine the confusion matrix is by comparing every two years' PMIS annual rating data of the same section. If the distress score difference of two years' PMIS annual rating data (e.g., FY2015-FY2014) in one pavement section is less than 15 points, this section can be considered as True Positive (TP). If the distress score difference of two years' PMIS annual rating data is larger than 15 points in a pavement section, this section can be considered as False Negative (FN). If the distress score difference of two years' PMIS annual rating data is smaller than -15 points in a pavement section, this section can be considered as False Positive (FP). In this definition, there is no True Negative (TN), so the TP represents both the TP and TN in the precision equation. The data used for distress score precision analysis in Scenario I are presented in Figure 3.13. The histogram of distress score difference of two years' PMIS annual rating data is shown in Figure 3.14. The precision analysis is presented in Table 3.2. The distress score precision is 90.88% in the period of FY 2014-2020.

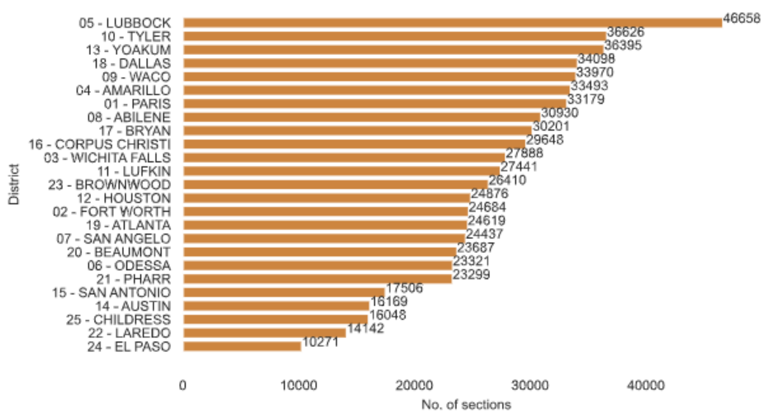


Figure 3.13 PMIS Annual Rating Data Two-Year Comparison (2014-2020)

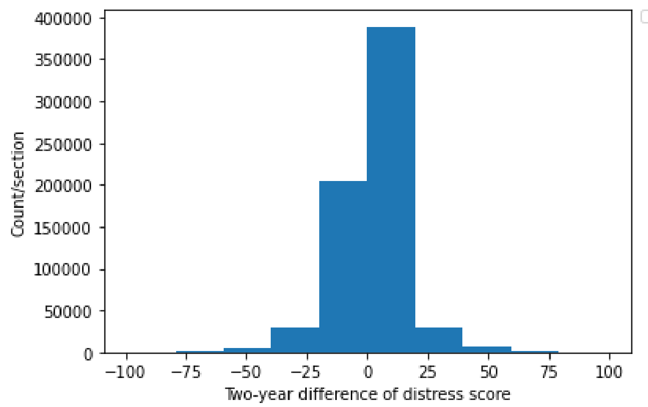


Figure 3.14 Tolerance: Two Years' Distress Score within 15 points (2014-2020)

Table 3.2 Confusion matrix of automated data precision (distress score)

Two-year comparison	P	N
P	**563,099 (TP)**	**50,389 (FN)**
N	**56,490 (FP)**	**-- (TN)**

$$\text{Precision} = \frac{TP}{TP+FP} = \frac{563099}{563099+56490} = 90.88\%$$

Scenario II:

The data used in Scenario II for accuracy include the PMIS annual rating data and audit data of the 25 districts from FY 2017 to FY 2021. The rule to determine the confusion matrix is comparing the PMIS annual rating data and audit data of the same section. If the distress score difference of the PMIS annual rating data and audit data (PMIS annual rating value – audit value) in one pavement section is less than 15 points, this section can be considered as True Positive (TP). If the distress score difference of the PMIS annual rating data is 15 points larger than that of the audit data in a pavement section, this section can be considered as False Negative (FN). If the distress score difference of the PMIS annual rating data is 15 points smaller than that of the audit data in a pavement section, this section can be considered as False Positive (FP). In this definition, there is no True Negative (TN), so the TP represents both the TP and TN in the accuracy equation. The data used for distress score accuracy analysis in Scenario II is presented in Figure 3.15. The histogram of distress score difference of PMIS annual rating data and audit data of the same section is shown in Figure 3.16. The accuracy analysis is presented in Table 3.3. The distress score accuracy is 81.76% in the period of FY 2017-2021.

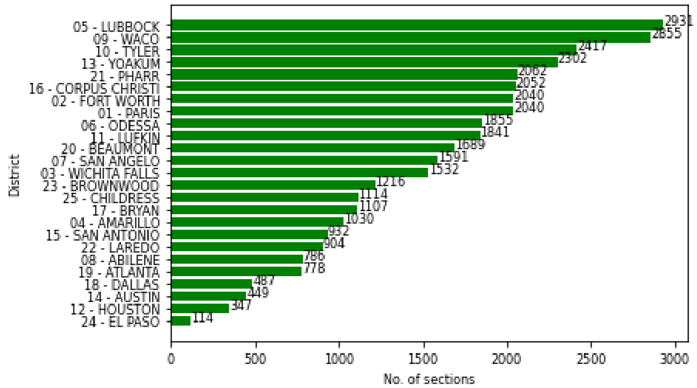


Figure 3.15 PMIS Annual Rating Data and Audit Data Comparison (2017-2021)

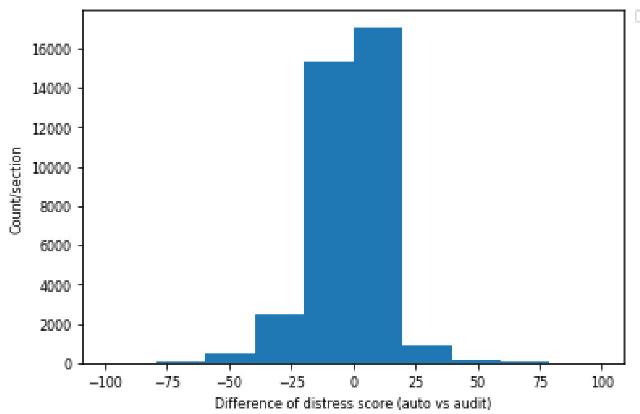


Figure 3.16 Tolerance: Auto and Audit Data Distress Score within 15 Points (2017-2021)

Table 3.3 Confusion matrix of automated data accuracy (distress score)

Two-year comparison	P	N
P	29,819 (TP)	1,419 (FN)
N	5,233 (FP)	-- (TN)

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} = \frac{29819+0}{29819+0+5233+1419} = 81.76\%$$

3.4 Individual Distress Analysis for Automated Data Collection

3.4.1 Overall Dataset for Individual Distress Analysis

The overall dataset used for individual distress analysis is the PMIS annual rating data and the audit data from FY 2017 to 2021 shown in Table 3.4.

Table 3.4 Dataset for individual distress analysis

Pavement Type	Sections after merging	Sections used for analysis
CRCP	1368	1325
JCP	322	310
ACP	34781	34533

The distress types of ACP, CRCP, and JCP included in the individual distress analysis are the following.

- ACP: alligator cracking (ACP ALLIGATOR CRACKING PCT), longitudinal cracking (ACP LONGITUDE CRACKING), transverse cracking (ACP TRANSVERSE CRACKING QTY), patching (ACP PATCHING PCT).
- CRCP: CRCP SPALLED CRACKS, CRCP PUNCHOUT, CRCP ACP PATCHES, CRCP PCC PATCHES, CRCP AVG CRACK SPACING.
- JCP: JCP FAILED JNTS CRACKS QTY, JCP FAILURES QTY.

3.4.2 Individual Distress Analysis for ACP

The individual distress analysis for ACP was conducted by comparing the PMIS annual rating data (auto data) using automated data collection technology and manual audit data. A confusion matrix was used to analyze the pavement sections with/without individual distress measurement values from both the auto and audit data, where “0” means the measurement values were 0 (no distress) and “! =0” means the measurement values were not 0 (distress appears). The confusion matrix results are shown in Table 3.5, Table 3.7, Table 3.9, and Table 3.11.

Table 3.5 Confusion matrix of ACP alligator cracking

		Auto data	
		0	!=0
Audit data	0	23981	5315
	!=0	2703	2554

Table 3.6, Table 3.8, Table 3.10, and Table 3.12 show the rates of the automatically measured individual distress by comparing with the audit data in the same data sections. This analysis includes six parameters listed in the tables below. “=” means the auto data and audit data results have the same distress measurement values which are not equal to 0. “s” means the auto data distress measurement values are larger than the values of audit data and not equal to 0, this parameter indicates the “sensitivity” of the automated data collection technology comparing with the manual data collection. “ns” means the auto data distress measurement values are smaller than the values of audit data and not equal to 0, this parameter indicates the “non-sensitivity” of the automated data collection technology comparing with the manual data

collection. “0” means both auto data and audit data distress measurement values are equal to 0. “FP” means the auto data have distress measurement values, but the audit data distress measurement values are equal to 0. “FN” is the opposite, the auto data distress measurement values are equal to 0, but the audit data have distress measurement values. Histograms of the individual distress measurements under “s”, “ns”, “FP”, and “FN” conditions are presented in Figure 3.17 to Figure 3.20.

Table 3.6 Data quality analysis of ACP alligator cracking

		Section	Rate
ACP Alligator Cracking !=0		2554	1
"="	auto = audit	674	0.01950627
s	auto > audit	638	0.01846439
ns	auto < audit	1242	0.03594478
0	auto = 0, audit = 0		0.69403525
FP	auto !=0, audit = 0		0.153821665
FN	audit !=0, auto = 0		0.07822765
FP+s	auto !=0, audit = 0, auto > audit		0.172286053
FN+ns	audit !=0, auto = 0, auto < audit		0.114172431
0 & "="	auto = audit		0.713541516

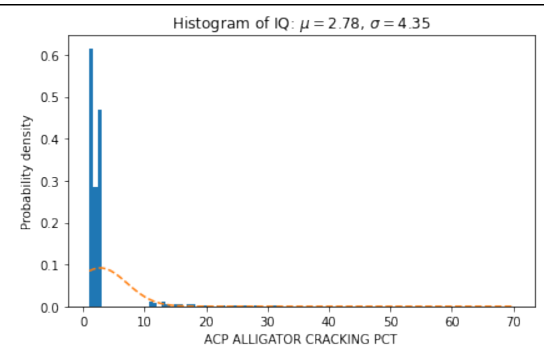
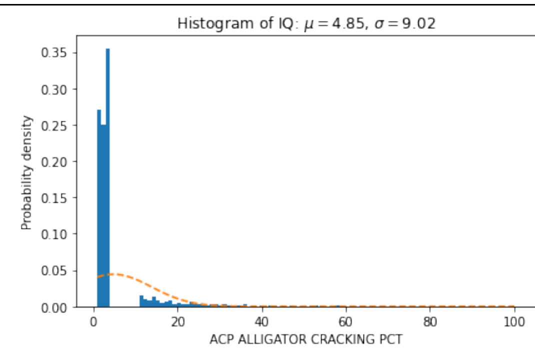
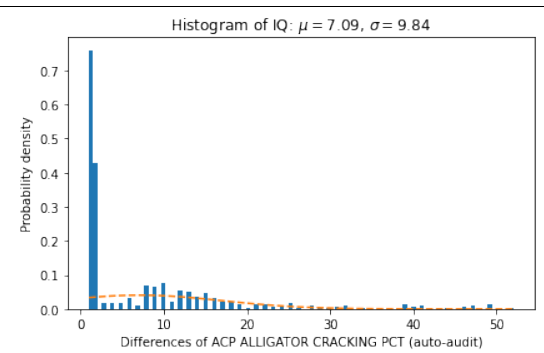
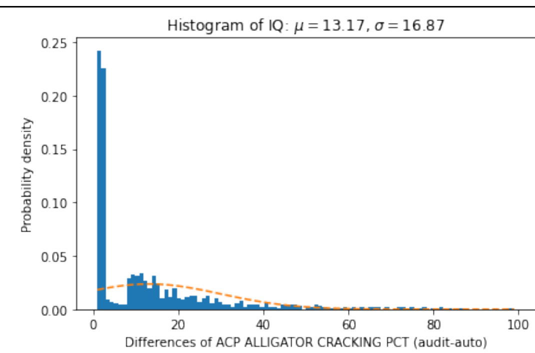
	
FP	FN
	
S	ns

Figure 3.17 Histogram Analysis of ACP Alligator Cracking

Table 3.7 Confusion matrix of ACP longitudinal cracking

		Auto data	
		0	!=0
Audit data	0	10471	11475
	!=0	1862	10745

Table 3.8 Data quality analysis of ACP longitudinal cracking

		Section	Rate
ACP Longitudinal Cracking !=0		10745	1
"="	auto = audit	1656	0.04792637
s	auto > audit	4306	0.12462015
ns	auto < audit	4783	0.13842503
0	auto = 0, audit = 0		0.303041704
FP	auto !=0, audit = 0		0.332098515
FN	audit !=0, auto = 0		0.05388823
FP+s	auto !=0, audit = 0, auto > audit		0.456718664
FN+ns	audit !=0, auto = 0, auto < audit		0.192313258
0 & "="	auto = audit		0.350968078

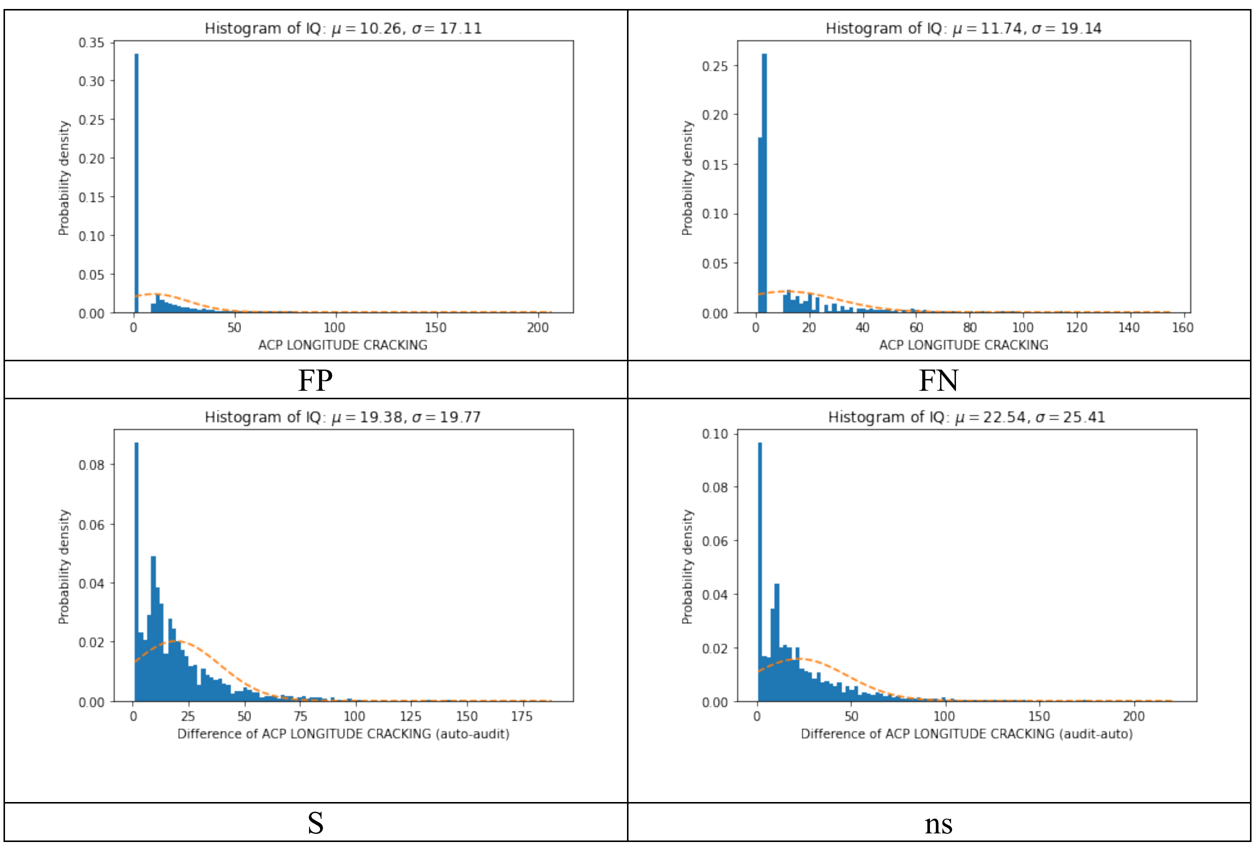


Figure 3.18 Histogram Analysis of ACP Longitudinal Cracking

Table 3.9 Confusion matrix of ACP transverse cracking

		Auto data	
		0	!=0
Audit data	0	27103	1994
	!=0	2847	2609

Table 3.10 Data quality analysis of ACP transverse cracking

		Section	Rate
ACP Transverse Cracking !=0		2609	1
"="	auto = audit	1027	0.02972246
s	auto > audit	343	0.00992678
ns	auto < audit	1239	0.03585796
0	auto = 0, audit = 0		0.784389199
FP	auto !=0, audit = 0		0.057708448
FN	audit !=0, auto = 0		0.082395161
FP+s	auto !=0, audit = 0, auto > audit		0.067635227
FN+ns	audit !=0, auto = 0, auto < audit		0.118253118
0 & "="	auto = audit		0.814111655

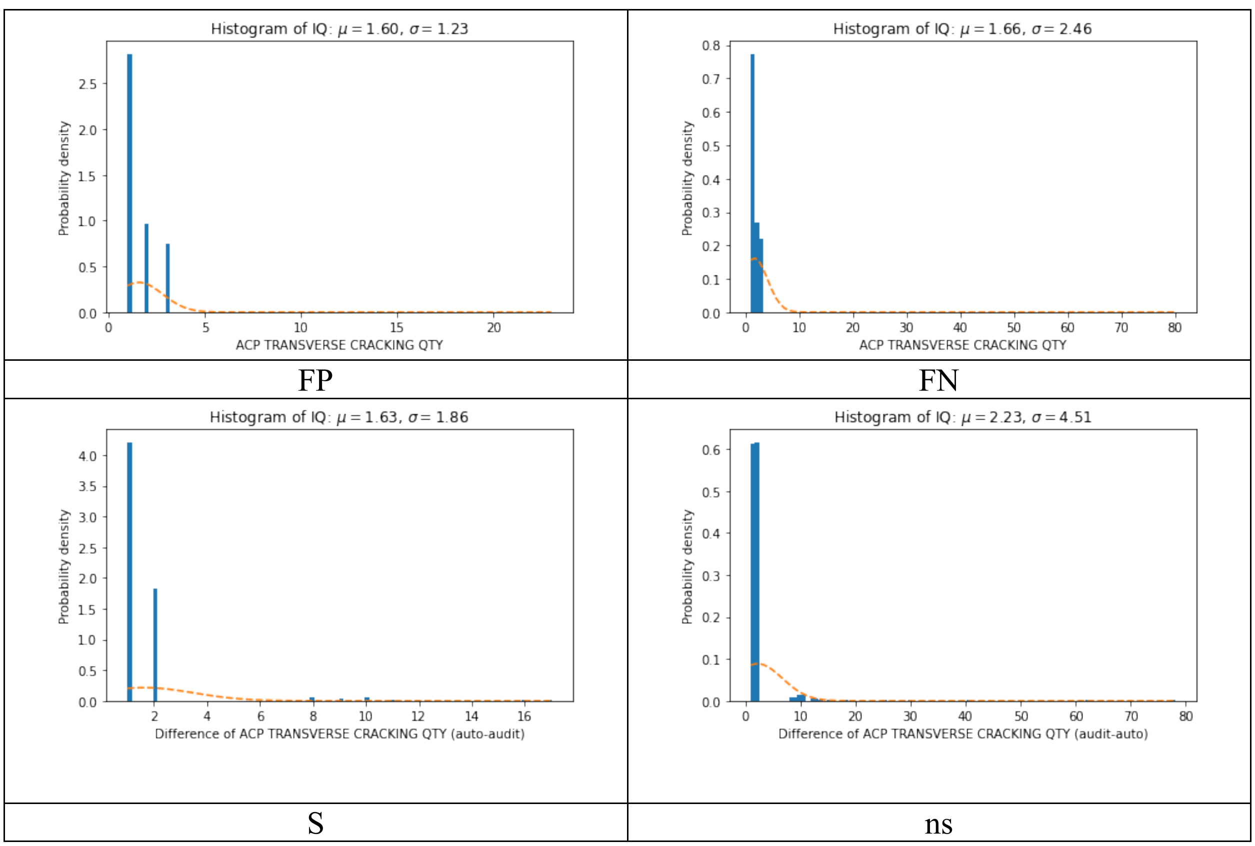


Figure 3.19 Histogram Analysis of ACP Transverse Cracking

Table 3.11 Confusion matrix of ACP patching

		Auto data	
		0	!=0
Audit data	0	30396	1175
	!=0	1201	1781

Table 3.12 Data quality analysis of ACP patching

		Section	Rate
ACP Patching !=0		1781	1
"="	auto = audit	612	0.01771192
S	auto > audit	373	0.01079501
Ns	auto < audit	796	0.02303707
0	auto = 0, audit = 0		0.879692067
FP	auto !=0, audit = 0		0.03400573
FN	audit !=0, auto = 0		0.034758198
FP+s	auto !=0, audit = 0, auto > audit		0.044800741
FN+ns	audit !=0, auto = 0, auto < audit		0.057795271
0 & "="	auto = audit		0.897403988

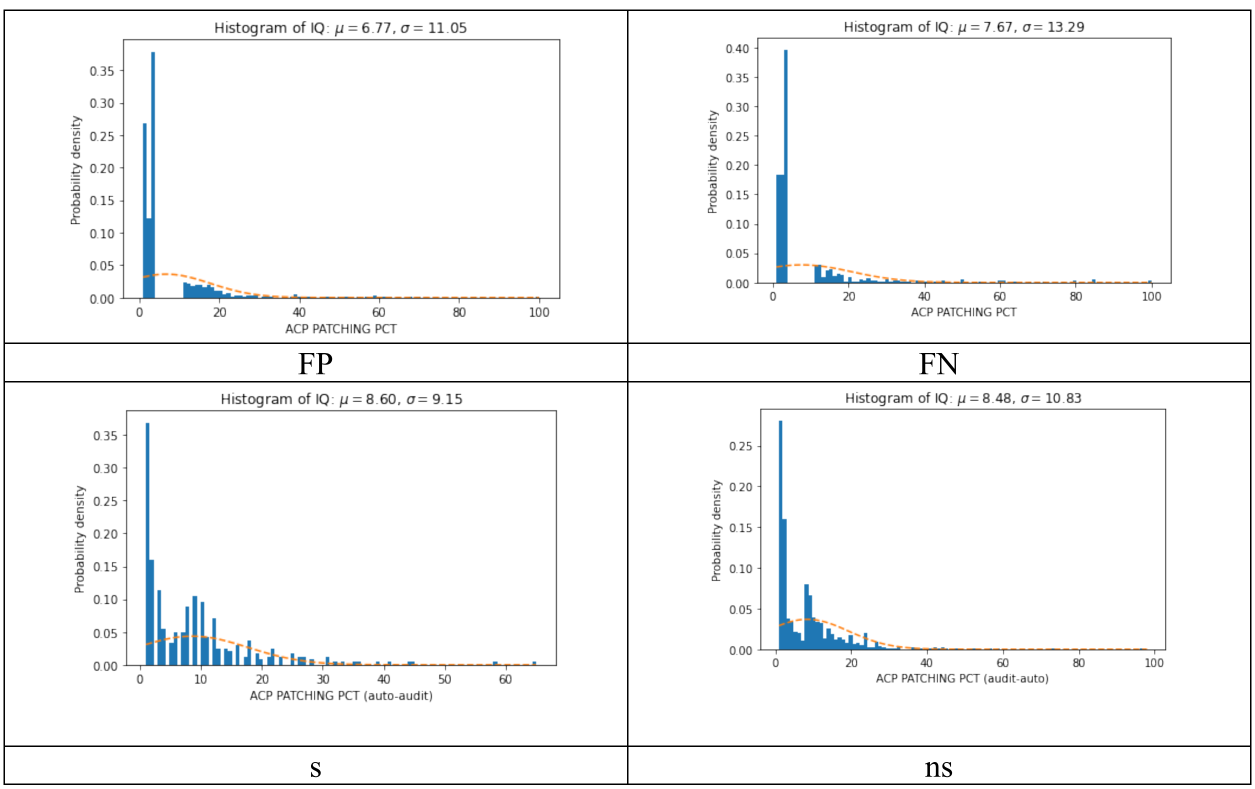


Figure 3.20 Histogram Analysis of ACP Patching

3.4.3 Individual Distress Analysis for CRCP

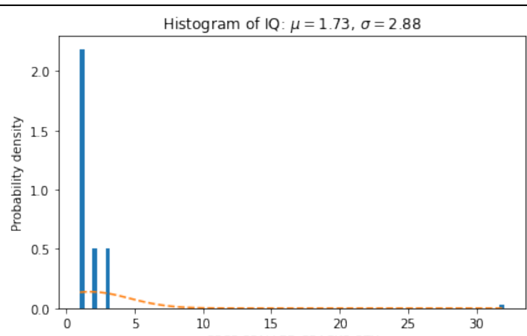
The individual distress analysis for CRCP was conducted by comparing the PMIS annual rating data (auto data) using automated data collection technology and the manual audit data, the confusion matrix results are shown in Table 3.13, Table 3.15,Table 3.17, Table 3.19, and Table 3.21. The rates of the automatically measured individual distress by comparing with the audit data in the same data sections are listed in Table 3.14, Table 3.16, Table 3.18, Table 3.20, and Table 3.22. Histograms of the individual distress measurements under “s”, “ns”, “FP”, and “FN” conditions are presented in Figure 3.21 to Figure 3.25.

Table 3.13 Confusion matrix of CRCP SPALLED CRACKS

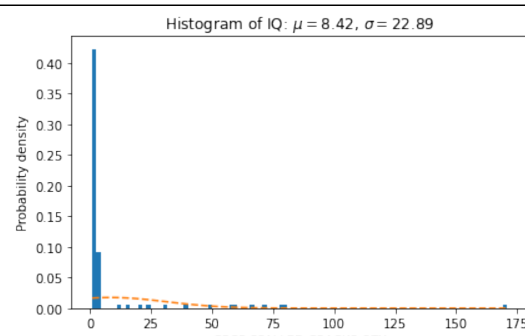
		Auto data	
		0	!=0
Audit data	0	1009	121
	!=0	110	85

Table 3.14 Data quality analysis of CRCP SPALLED CRACKS

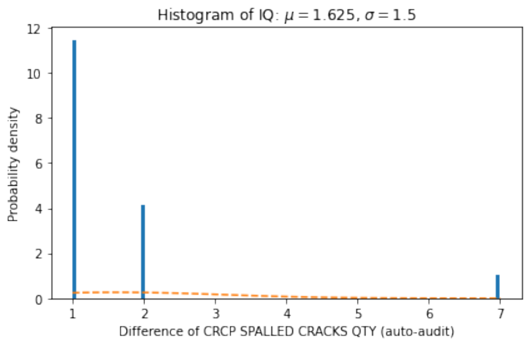
		Section	Rate
ACP Alligator Cracking !=0		85	1
"="	auto = audit	33	0.02490566
s	auto > audit	16	0.01207547
ns	auto < audit	36	0.02716981
0	auto = 0, audit = 0		0.761509434
FP	auto !=0, audit = 0		0.091320755
FN	audit !=0, auto = 0		0.083018868
FP+s	auto !=0, audit = 0, auto > audit		0.103396226
FN+ns	audit !=0, auto = 0, auto < audit		0.110188679
0 & "="	auto = audit		0.786415094

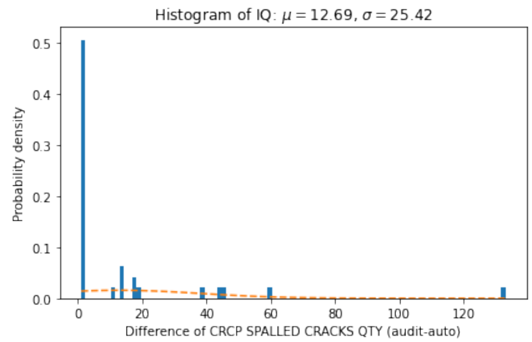


FP



FN





s	ns

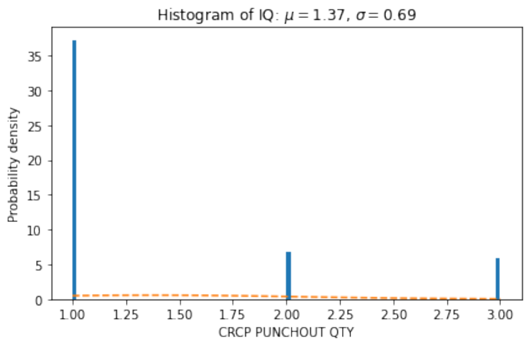
Figure 3.21 Histogram Analysis of CRCP SPALLED CRACKS

Table 3.15 Confusion matrix of CRCP PUNCHOUT

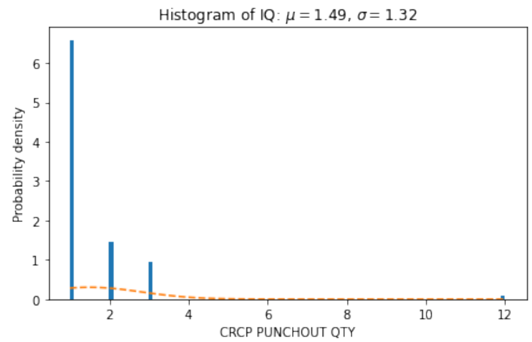
		Auto data	
		0	!=0
Audit data	0	1152	51
	!=0	87	35

Table 3.16 Data quality analysis of CRCP PUNCHOUT

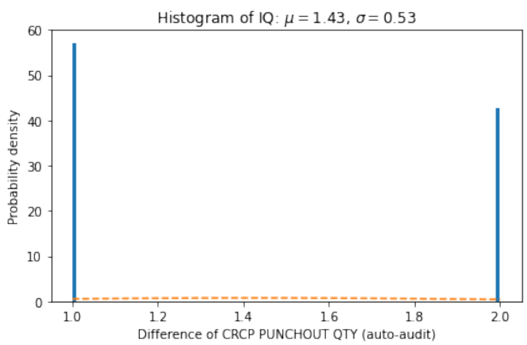
		Section	Rate
ACP Longitudinal Cracking !=0		35	1
"="	auto = audit	17	0.01283019
s	auto > audit	7	0.00528302
ns	auto < audit	11	0.00830189
0	auto = 0, audit = 0		0.869433962
FP	auto !=0, audit = 0		0.038490566
FN	audit !=0, auto = 0		0.065660377
FP+s	auto !=0, audit = 0, auto > audit		0.043773585
FN+ns	audit !=0, auto = 0, auto < audit		0.073962264
0 & "="	auto = audit		0.882264151



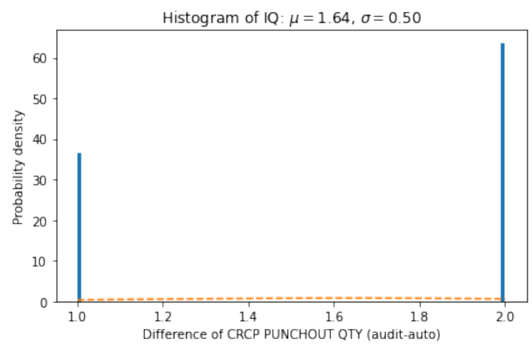
FP



FN



s



ns

Figure 3.22 Histogram Analysis of CRCP PUNCHOUT

Table 3.17 Confusion matrix of CRCP ACP PATCHES

		Auto data	
		0	!=0
Audit data	0	1201	69
	!=0	23	32

Table 3.18 Data quality analysis of CRCP ACP PATCHES

		Section	Rate
ACP Transverse Cracking !=0		32	1
"="	auto = audit	14	0.01056604
s	auto > audit	13	0.00981132
ns	auto < audit	5	0.00377358
0	auto = 0, audit = 0		0.906415094
FP	auto !=0, audit = 0		0.052075472
FN	audit !=0, auto = 0		0.017358491
FP+s	auto !=0, audit = 0, auto > audit		0.061886792

FN+ns	audit !=0, auto = 0, auto < audit	0.021132075
0 & "="	auto = audit	0.916981132

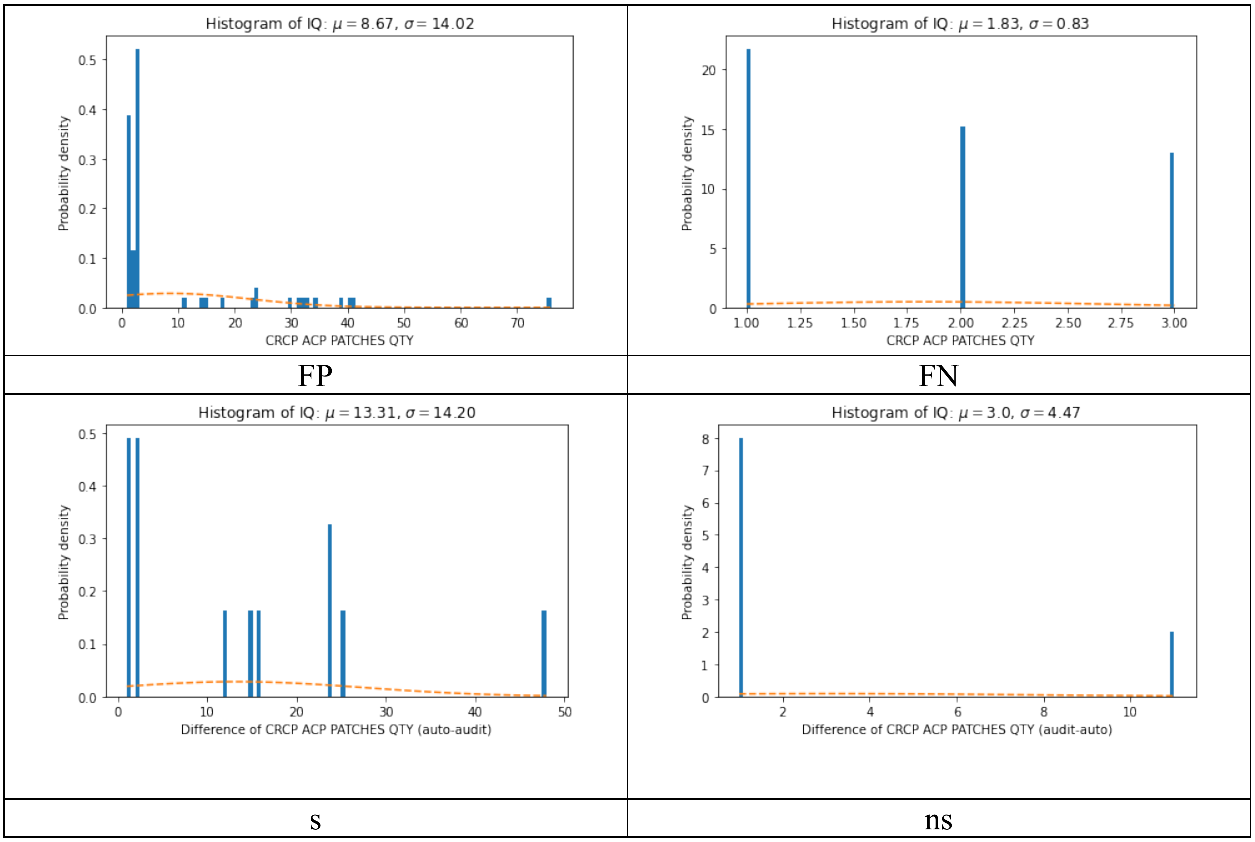


Figure 3.23 Histogram Analysis of CRCP ACP PATCHES

Table 3.19 Confusion matrix of CRCP PCC PATCHES

		Auto data	
		0	!=0
Audit data	0	963	75
	!=0	92	195

Table 3.20 Data quality analysis of CRCP PCC PATCHES

		Section	Rate
ACP Patching !=0		195	1
"="	auto = audit	97	0.07320755
s	auto > audit	56	0.04226415
ns	auto < audit	42	0.03169811
0	auto = 0, audit = 0		0.726792453
FP	auto !=0, audit = 0		0.056603774
FN	audit !=0, auto = 0		0.069433962

FP+s	auto !=0, audit = 0, auto > audit	0.098867925
FN+ns	audit !=0, auto = 0, auto < audit	0.101132075
0 & "="	auto = audit	0.8

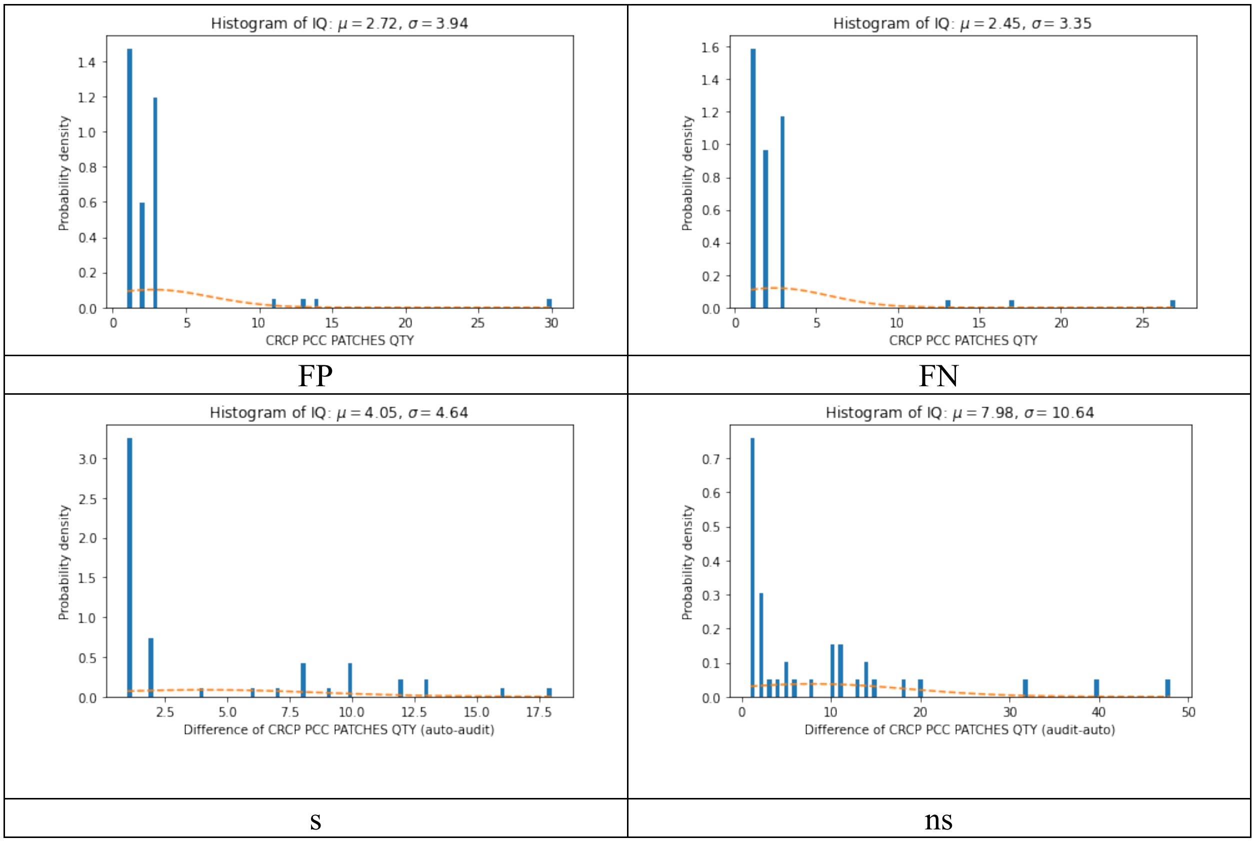


Figure 3.24 Histogram Analysis of CRCP PCC PATCHES

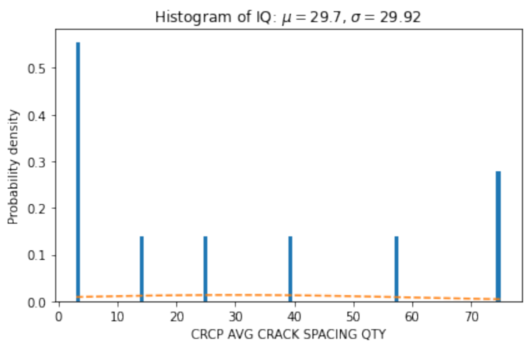
Table 3.21 Confusion matrix of CRCP AVG CRACK SPACING

		Auto data	
		0	!=0
Audit data	0	0	10
	!=0	0	1315

Table 3.22 Data quality analysis of CRCP AVG CRACK SPACING

		Section	Rate
ACP Patching !=0		1315	1
"="	auto = audit	549	0.41433962
s	auto > audit	640	0.48301887
ns	auto < audit	126	0.09509434
0	auto = 0, audit = 0		0
FP	auto !=0, audit = 0		0.00754717

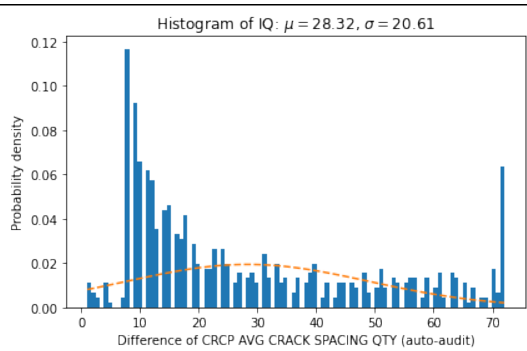
FN	audit !=0, auto = 0	0
FP+s	auto !=0, audit = 0, auto > audit	0.490566038
FN+ns	audit !=0, auto = 0, auto < audit	0.09509434
0 & "="	auto = audit	0.414339623

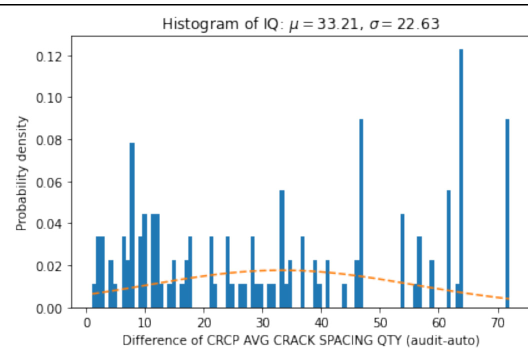


N/A

FP

FN





s

ns

Figure 3.25 Histogram Analysis of CRCP AVG CRACK SPACING

3.4.4 Individual Distress Analysis for JCP

The individual distress analysis for JCP was conducted by comparing the PMIS annual rating data (auto data) using automated data collection technology and the manual audit data, the confusion matrix results are shown in Table 3.23 and Table 3.25. The rates of the automatically measured individual distress by comparing with the audit data in the same data sections are listed in Table 3.24 and Table 3.26. Histograms of the individual distress measurements under “s”, “ns”, “FP”, and “FN” conditions are presented in Figure 3.26 and Figure 3.27.

Table 3.23 Confusion matrix of JCP FAILED JNTS CRACKS QTY

		Auto data	
		0	!=0
Audit data	0	151	53
	!=0	47	59

Table 3.24 Data quality analysis of JCP FAILED JNTS CRACKS QTY

		Section	Rate
ACP Alligator Cracking !=0		59	1
"="	auto = audit	23	0.07419355
s	auto > audit	20	0.06451613
ns	auto < audit	16	0.0516129
0	auto = 0, audit = 0		0.487096774
FP	auto !=0, audit = 0		0.170967742
FN	audit !=0, auto = 0		0.151612903
FP+s	auto !=0, audit = 0, auto > audit		0.235483871
FN+ns	audit !=0, auto = 0, auto < audit		0.203225806
0 & "="	auto = audit		0.561290323

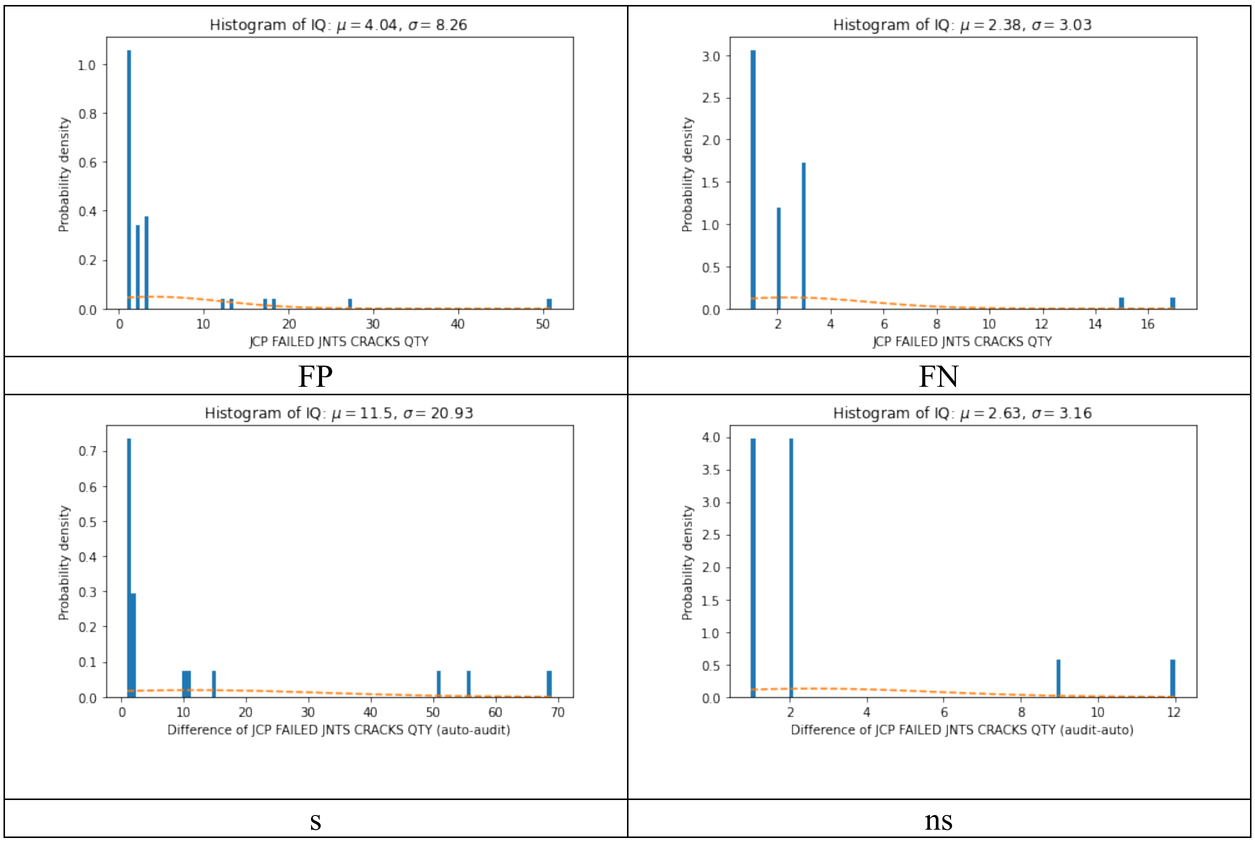


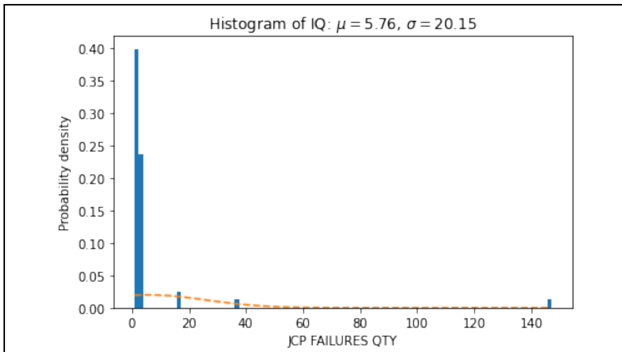
Figure 3.26 Histogram Analysis of JCP FAILED JNTS CRACKS QTY

Table 3.25 Confusion matrix of JCP FAILURES QTY

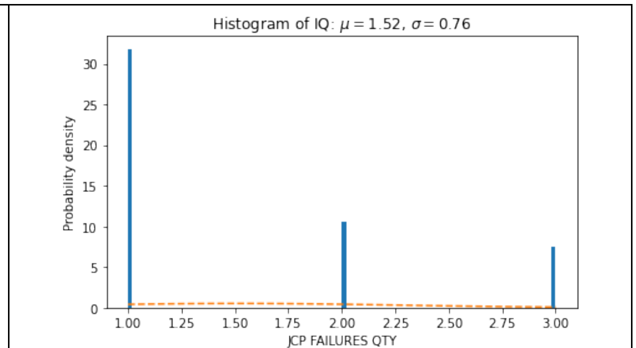
		Auto data	
		0	!=0
Audit data	0	136	55
	!=0	33	86

Table 3.26 Data quality analysis of JCP FAILURES QTY

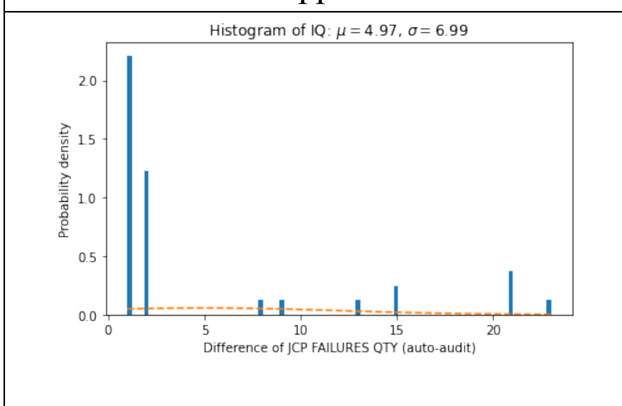
		Section	Rate
ACP Longitudinal Cracking !=0		86	1
"="	auto = audit	34	0.10967742
s	auto > audit	37	0.11935484
ns	auto < audit	15	0.0483871
0	auto = 0, audit = 0		0.438709677
FP	auto !=0, audit = 0		0.177419355
FN	audit !=0, auto = 0		0.106451613
FP+s	auto !=0, audit = 0, auto > audit		0.296774194
FN+ns	audit !=0, auto = 0, auto < audit		0.15483871
0 & "="	auto = audit		0.548387097



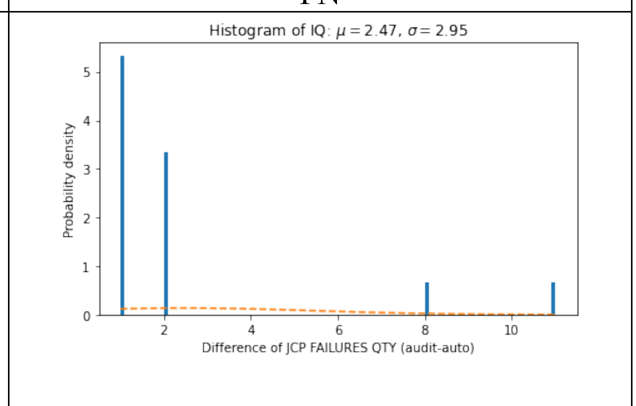
FP



FN



S



ns

Figure 3.27 Histogram Analysis of JCP FAILURES QTY

Chapter 4 Sampling Method for Data Quality Audit

To sample more units to have higher accuracy in sampling when there is high variability among units (or elements) of a population, Stratified Sampling divides the population into several strata with similarities among units or smaller variability within each stratum. The design of a stratified sample constitutes three steps: Step I Defining strata; Step II Optimal allocating sample sizes to strata; and Step III Rechecking and iterative updating sample sizes.

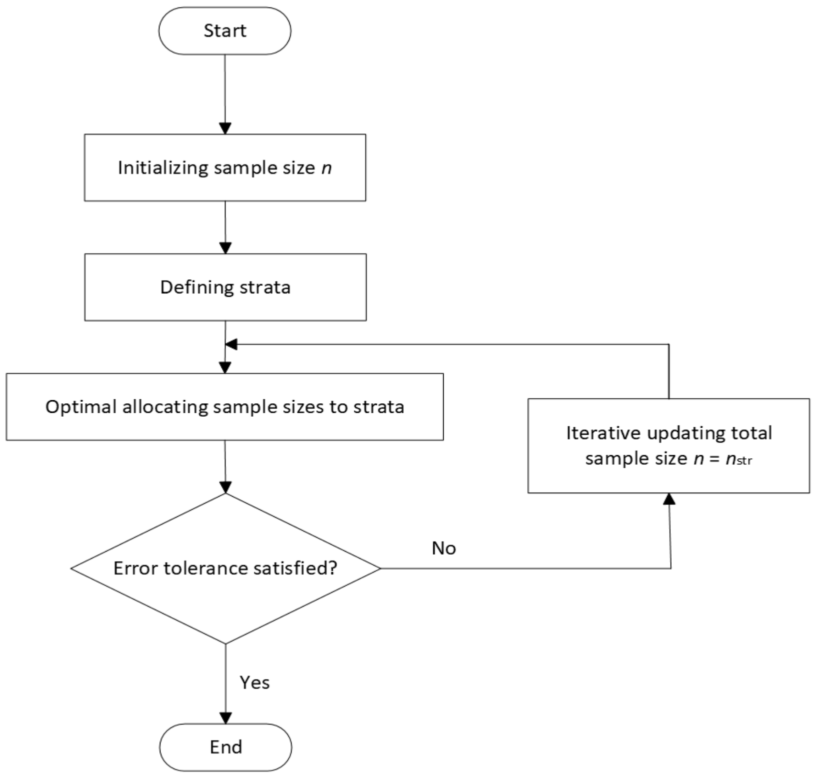


Figure 4.1 Flowchart of Stratified Sampling Method

However, the design of a stratified sample to be used for the audit (or validation) of the automated pavement condition data would be an iterative process, as shown in Figure 4.1. An initial sample size n would be used, for example, the 6% roadbed miles or a certain number of sections. In defining the strata, the population or TxDOT-maintained pavement network is to be divided into different layers or strata. Then the initial sample size is allocated to the strata using an optimal allocation method. The stratified sample is rechecked for sample accuracy or error tolerance and a new sample size is generated. With budget constraints, engineering judgment, and past good experience all considered, an appropriate new sample size n_{str} would be used to replace the initial sample size n and the sample sizes for the strata would be allocated again.

Step I: Defining Strata

Stratified sampling can provide higher precisions than a simple random sampling method.

However, the stratification needs a data-enabled preparation procedure before a stratified sampling method can be successfully implemented. The first step is defining strata for the stratified sampling method. The major considerations for setting up the potential influencing factors in defining the strata are the engineers' practical experience as well as the available data.

The statistical test model ANOVA (Analysis of Variance) and visualization comparison are used to check the distress score, condition score, and ride score of the automated collected PMIS annual rating data (auto data)

Based on statewide data from 2017 to 2021, the selected influencing factors for defining strata are 1) Pavement Type, 2) Pavement Surface Age, and 3) Highway Service Level.

1) Pavement Type: asphalt concrete pavement (ACP), Continuously Reinforced Concrete Pavement (CRCP), and Jointed Concrete Pavement (JCP)
2) Pavement Surface Age: old (surface age more than 3 years) and young (surface age less than 3 years)
3) Highway service level: Interstate Highway (IH), U.S. Highway (US), State Highway (SH), and Farm to Market Road (FM)

Therefore, the designed stratified sampling method for audit sample selection has three strata based on the statistical analysis using the currently available data.

When we divide the population of N elements into H layers or strata, with N_h units in stratum h. For stratified sampling to work, we must know the values of $N_1, N_2, ..., N_H$, and must have $N_1 + N_2 + ... + N_H = N$. If n_h observations are taken from N_h population units in the stratum h, then the total sample size of the stratified sample is $n = n_1 + n_2 + ... + n_H$.

The following notations are defined for the stratified sampling method.

$y_{hj} =$ value of the jth unit in stratum h

$$\bar{y}_{hU} = \frac{\sum_{j=1}^{N_h} y_{hj}}{N_h} = \text{population mean in stratum } h$$

$$S_h^2 = \frac{\sum_{j=1}^{N_h} \left(y_{hj} - \bar{y}_{hU} \right)^2}{N_h - 1} = \text{population variance in stratum } h$$

$$\bar{y}_h = \frac{\sum_{j=1}^{n_h} y_{hj}}{n_h} = \text{sample mean in stratum } h \text{ (estimate for } \bar{y}_{hU}\text{)}$$

$$s_h^2 = \frac{\sum_{j=1}^{n_h} \left(y_{hj} - \bar{y}_h\right)^2}{n_h - 1} = \text{sample variance in stratum } h \text{ (estimate for } S_h^2\text{)}$$

$$v = \sum_{h=1}^{H} \frac{n}{n_h} \left(\frac{N_h}{N}\right)^2 S_h^2 = \text{sample estimate for population variance } S^2$$

Step II: Optimal Allocating Sample Sizes to Strata

The calculation of the sample sizes follows a proportional allocation process in which the number of sample units for each stratum n_h is proportional to the size of the stratum N_h. By using the proportional allocation, the true percentage of each stratum in the network population can be well represented. For example, 90% of ACP, and 10% of CRCP and JCP in TxDOT's pavement network could be well represented in the stratified sample.

Stratum variance should be considered in sample size allocation for each stratum. If the variances are more or less equal across all strata, then the proportional allocation would probably be the best allocation strategy. In cases where stratum variances vary significantly from each other, An optimal allocation considering the variance differences among strata can result in smaller sampling cost or better sampling efficiency. In practice, when sampling units are quite different from each other, these units in a stratum with higher variability would be sampled with a higher sampling rate than the units of a stratum with lower variability.

Data collection cost is an important factor to be considered in allocation of sample sizes to strata in a stratified sampling method. As envisioned by researchers of this project, the total cost of audit data collection is mainly impacted by the data collection labor cost, data collection equipment cost, and data collection travel cost. The subtotal costs associated with a stratum might be different from each other based on the characteristics of the stratum.

The objective of optimal allocation is to gain the most information for the least cost. A simple cost function is given below: Let C represent the total cost, c_0 represent overhead costs such as maintaining a team for auditing, and c_h represent the cost of taking an observation in stratum h which includes data collection labor cost, data collection equipment cost, and data collection travel cost, so that

$$C = c_0 + \sum_{h=1}^{H} c_h n_h \qquad (4)$$

In the sampling method, we want to allocate observations to strata so as to minimize the estimated variance for a given total cost C, or equivalently, to minimize C for a fixed variance. Suppose that the costs $c_1, c_2, \dots, c_H$ are known. To minimize the total cost for a fixed variance, we can prove using calculus that the optimal allocation has n_h proportional to

$$\frac{N_h S_h}{\sqrt{c_h}} \tag{5}$$

for each h. Thus, the optimal sample size for stratum h is calculated in the following equation

$$n_h = \left(\frac{\frac{N_h S_h}{\sqrt{c_h}}}{\sum_{l=1}^{H} \frac{N_l S_l}{\sqrt{c_l}}} \right) n \tag{6}$$

Where, n_h is the sample size in stratum *h*;
N_h is the population of units in stratum *h*;
S_h is the standard deviation of units in stratum *h;*
c_h is the cost of collecting an audit sample unit in stratum *h;*
H is the number of strata in the population;
N_l is the population of units in stratum *l*;
S_l is the standard deviation of units in stratum *l*;
c_l is the cost of collecting an audit sample unit in stratum *l*;
n is the size of the overall stratified sample.

In the practice of implementing a stratified sampling to the pavement sections of TxDOT's network for audit of the automated pavement condition data, the stratum should be more heavily sampled if

- The stratum accounts for a larger part of the network population.
- The variance within the stratum is larger. That is to say to sample more heavily in the stratum to compensate for the heterogeneity.
- Sampling in the stratum is inexpensive.

Sometimes applying the above optimal allocation formula results in one or more of the "optimal" n_h's being larger than the population size N_h in those strata. In that case, take a sample size of N_h in those strata, and then apply the above equation to the remaining strata.

Overall, in calculating the optimal sample sizes, the percentage of each stratum in the population, data collection cost, and population variance are included as factors in allocating an optimal sample size to each stratum.

Step III: Rechecking and Iterative Updating Sample Sizes

With a random sampling method, the total sample size n (sections or miles, over TxDOT's maintained overall network denoted as N sections or miles) used for the audit is currently 6% of the roadbed miles, and the current audit tolerance e for measurement error in Distress Score is 15 points. Based on the overall network variance S^2, tolerance e for the measurement error of pavements in the network, and the Z_α value ($Z_{\alpha/2}$ for two tails) of the standard normal distribution, sample size n can be calculated using the following equation for the associated confidence interval and the confidence level $(1-\alpha)$, which is quality assurance of the pavement condition data.

$$n = \left(\frac{Z_{\alpha/2} S}{e} \right)^2 \tag{7}$$

Although an initial sample size *n* is used in implementing Steps I and II, we are able to recheck the validity of the initial use of the sample size. In fact, in allocating sample sizes to the strata, we are able to calculate the allocation rate $\frac{n}{n_h}$ at each stratum *h*. Therefore we may assess the population variance S^2 with the following formula, i.e.,

$$\nu = \sum_{h=1}^{H} \frac{n}{n_h} \left(\frac{N_h}{N} \right)^2 S_h^2 = \text{sample estimate for population variance } S^2$$

Therefore, the sample size of a stratified sample with *H* layers or strata could be recalculated as:

$$n_{str} = \left(\frac{Z_{\alpha/2} \cdot \nu}{e} \right)^2 \tag{8}$$

The validity of the new sample size n_{str} needs to be carefully evaluated before a possible adoption for implementation. It should be noted that the newly calculated sample size should be compared with the existing 6% (roadbed miles) sample size *n*, with considerations of an available budget as well as engineering judgement learned from the successful experience of pavement engineers in the Pavement Preservation Branch and TxDOT Districts, along with any other relevant limiting constraints. Once an acceptable new sample size n_{str} is agreed by the stakeholders, Step II “Optimal Allocating Sample Sizes to Strata” needs to be conducted again to reallocate the sample sizes to the sampling strata.

Chapter 5 Development of Data Quality Consistency Check Indexes and Threshold

5.1 Data Quality Consistency Check Components

5.1.1 Yearly Change of Sections with Distress Decrease

Deterioration of constructed asphalt and concrete pavement is natural. It is natural because over time the materials begin to break down and become affected by elements such as traffic loading, rain, sunlight, and chemicals that encounter the pavement surface. The index type of yearly change of sections with distress decrease is based on the characteristics of pavement deterioration. Affecting by the elements, the pavement performance should deteriorate which presents as distress occurs or distress measurements increase. If there are sections with distress disappearance or distress measurements decrease over time, these sections can be considered to have data quality issues. However, the decision of sections with data quality issues should be made by considering the engineer's experience.

Table 5.1 No. of sections with yearly distress decrease

Fiscal year	ACP	CRCP	JCP
2017	36445	1035	2543
2018	36429	1040	2552
2019	38525	1078	2306
2020	38410	1054	2297
2021	40464	1086	275

The dataset (number of sections) used for the yearly change of sections with distress decrease is presented in Table 5.1.

1) Indexes

The data quality check indexes for yearly change of sections with distress decrease or disappearance include four types of distresses in ACP, five types of distress in CRCP, and two types of distress in JCP. The data used are the PMIS annual rating data (auto data) using automated data collection technology. All the sections with maintenance in FY 2017-2020 recorded in TxDOT's database are removed from the dataset to make sure all the pavement sections are under natural deterioration.

The distress types in ACP are the percentage of wheel-path length with alligator cracking (ACP ALLIGATOR CRACKING PCT), the length in feet per station of visually observed longitudinal cracking (ACP LONGITUDE CRACKING), the number of visually observed transverse cracks per station (ACP TRANSVERSE CRACKING QTY), and the percentage of lane area with patching (ACP PATCHING PCT) in the rated lane of the data collection section.

The distress types in CRCP are spalled transverse cracks in quantity (CRCP SPALLED CRACKS QTY), the number of punchouts and failures (CRCP PUNCHOUT QTY), the number of asphalt patches (CRCP ACP PATCHES QTY), the number of visually observed concrete (PCC) patches (CRCP PCC PATCHES QTY), and the average observed pacing, in feet, between transverse cracks (CRCP AVG CRACK SPACING QTY) in the rated lane of the data collection section.

The distress types in JCP are the number of visually observed transverse spalled cracks or failed joints and cracks (JCP FAILED JNTS CRACKS QTY) and the number of visually observed failures in the rated lane of the data collection section.

2) Thresholds

The threshold for the yearly change of sections with a distress decrease or disappearance is 0 tolerance, which means every section with a distress decrease or disappearance compared to the previous year should be considered to have data quality issues. These sections should be selected and go through the QA processes by re-checking with pavement images and re-evaluating by TxDOT pavement engineers. Based on the TxDOT provided data from FY 2017-2020, the research team conducted data analysis to find the pavement sections with individual distress decrease or disappearance. The number and percentage of sections are listed in Table 5.2.

Table 5.2 Yearly changes of sections with distress decrease

Pavement type	Distress type	FY 2018 vs 2017		FY 2019 vs 2018		FY 2020 vs 2019		FY 2021 vs 2020	
		sections	%	sections	%	sections	%	sections	%
ACP	ACP ALLIGATOR CRACKING PCT	10944	30.04	230	0.60	5009	13.04	2180	5.39
	ACP LONGITUDE CRACKING	15576	42.76	648	1.68	10502	27.34	5560	13.74
	ACP TRANSVERSE CRACKING QTY	5876	16.13	86	0.22	1445	3.76	821	2.03
	ACP PATCHING PCT	2033	5.58	125	0.32	1214	3.16	1289	3.19
CRCP	CRCP SPALLED CRACKS QTY	16	1.54	17	1.58	41	3.89	50	4.60
	CRCP PUNCHOUT QTY	29	2.79	0	0.00	9	0.85	32	2.95
	CRCP ACP PATCHES QTY	9	0.02	0	0.00	21	1.99	25	2.30

	CRCP PCC PATCHES QTY	70	6.73	36	3.34	60	5.69	31	2.85
	CRCP AVG CRACK SPACING QTY	64	6.15	21	1.95	129	12.24	135	12.43
JCP	JCP FAILED JNTS CRACKS QTY	26	1.02	45	1.95	21	0.91	68	24.73
	JCP FAILURES QTY	37	1.45	22	0.95	37	1.61	43	15.64

5.1.2 Individual Distress Change

The individual distress change analysis for ACP, CRCP, and JCP was conducted and presented in Chapter 3 of this report. The data used are the PMIS annual rating data (auto data) using automated data collection technology and the manual audit data. The dataset includes the sections with maintenance. Based on the comments and suggestions of eliminating the impacts of pavement performance from regular maintenance provided in Project Update Meeting # 4, the research team reanalyzed the individual distress change using a new dataset. In the new dataset, all the sections with maintenance in FY 2017-2020 recorded in TxDOT's database are removed from PMIS annual rating data, but the manual audit data remains the same. The indexes included in the individual distress change analysis are the same distress types of ACP, CRCP, and JCP as a yearly change of sections with distress decrease, which are listed as follows.

- ACP: alligator cracking (ACP ALLIGATOR CRACKING PCT), longitudinal cracking (ACP LONGITUDE CRACKING), transverse cracking (ACP TRANSVERSE CRACKING QTY), patching (ACP PATCHING PCT).
- CRCP: CRCP SPALLED CRACKS, CRCP PUNCHOUT, CRCP ACP PATCHES, CRCP PCC PATCHES, CRCP AVG CRACK SPACING.
- JCP: JCP FAILED JNTS CRACKS QTY, JCP FAILURES QTY.

The overall dataset used for individual distress analysis is the paired PMIS annual rating data with the audit data from FY 2017 to 2021 shown in. Further, the confusion matrix was used to analyze the pavement sections with/without individual distress measurement values from both the auto and audit data, where "0" means the measurement values were 0 (no distress) and "!=0" means the measurement values were not 0 (distress appears). In the confusion matrix, TP represents the number of sections with both auto data and audit data not equal to 0, FN represents the number of sections with auto data equal to 0 and audit data not equal to 0, FP represents the number of sections with auto data not equal to 0 and audit data equal to 0, and TN represents the number of sections with auto data equal to 0 and audit data not equal to 0. The explanation of the confusion matrix is shown in Figure 5.1.

Table 5.3 Dataset for individual distress analysis

Pavement type	Sections after merging	Sections used for analysis
CRCP	1368	1325
JCP	322	310
ACP	3655	3488

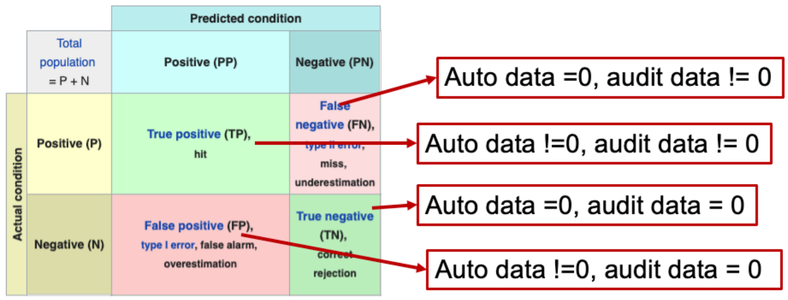


Figure 5.1 Confusion Matrix for Auto Data and Audit Data Comparison

a) Individual distress changes for ACP

The updated individual distress analysis for ACP was conducted by comparing the PMIS annual rating data (auto data) without construction sections using automated data collection technology and manual audit data. The confusion matrix results are shown in Table 5.4 to Table 5.7.

Table 5.4 Confusion matrix of ACP alligator cracking

		Auto data	
		0	!=0
Audit data	0	2563	405
	!=0	262	258

Table 5.5 Confusion matrix of ACP longitudinal cracking

		Auto data	
		0	!=0
Audit data	0	1338	707
	!=0	297	1146

Table 5.6 Confusion matrix of ACP transverse cracking

		Auto data	
		0	!=0
Audit data	0	2760	110
	!=0	368	250

Table 5.7 Confusion matrix of ACP patching

		Auto data	
		0	!=0
Audit data	0	3043	105
	!=0	112	228

Table 5.8 to Table 5.11 show the rates of the automatically measured individual distresses by comparing with the audit data in the same data sections. This analysis includes six parameters listed in the tables below. “=” means the auto data and audit data results have the same distress measurement values which are not equal to 0. “s” means the auto data distress measurement values are larger than the values of audit data and not equal to 0, this parameter indicates the “sensitivity” of the automated data collection technology comparing with the manual data collection. “ns” means the auto data distress measurement values are smaller than the values of audit data and not equal to 0, this parameter indicates the “non-sensitivity” of the automated data collection technology comparing with the manual data collection. “0” means both auto data and audit data distress measurement values are equal to 0. “FP” means the auto data have distress measurement values, but the audit data distress measurement values are equal to 0. “FN” is the opposite, the auto data distress measurement values are equal to 0, but the audit data have distress measurement values.

Table 5.8 Data quality analysis of ACP alligator cracking

		Section	Rate
ACP Alligator Cracking !=0		258	7.40%
"="	auto = audit	37	1.06%
S	auto > audit	90	2.58%
Ns	auto < audit	131	3.76%
0	auto = 0, audit = 0		73.48%
FP	auto !=0, audit = 0		11.61%
FN	audit !=0, auto = 0		7.51%
FP+s	auto !=0, audit = 0, auto > audit		14.19%
FN+ns	audit !=0, auto = 0, auto < audit		11.27%
0 & "="	auto = audit		74.54%
Accuracy	Accuracy for distress detection		80.88%

Table 5.9 Data quality analysis of ACP longitudinal cracking

		Section	Rate
ACP Longitudinal Cracking !=0		1146	32.86%
"="	auto = audit	51	1.46%
s	auto > audit	501	14.36%
ns	auto < audit	594	17.03%
0	auto = 0, audit = 0		38.36%

FP	auto !=0, audit = 0	20.27%
FN	audit !=0, auto = 0	8.51%
FP+s	auto !=0, audit = 0, auto > audit	34.63%
FN+ns	audit !=0, auto = 0, auto < audit	25.54%
0 & "="	auto = audit	39.82%
Accuracy	Accuracy for distress detection	71.22%

Table 5.10 Data quality analysis of ACP transverse cracking

		Section	Rate
ACP Transverse Cracking !=0		250	7.17%
"="	auto = audit	71	2.04%
S	auto > audit	25	0.72%
Ns	auto < audit	154	4.42%
0	auto = 0, audit = 0		79.13%
FP	auto !=0, audit = 0		3.15%
FN	audit !=0, auto = 0		10.55%
FP+s	auto !=0, audit = 0, auto > audit		3.87%
FN+ns	audit !=0, auto = 0, auto < audit		14.97%
0 & "="	auto = audit		81.16%
Accuracy	Accuracy for distress detection		86.30%

Table 5.11 Data quality analysis of ACP patching

		Section	Rate
ACP Patching !=0		228	6.54%
"="	auto = audit	27	0.77%
S	auto > audit	68	1.95%
Ns	auto < audit	133	3.81%
0	auto = 0, audit = 0		87.24%
FP	auto !=0, audit = 0		3.01%
FN	audit !=0, auto = 0		3.21%
FP+s	auto !=0, audit = 0, auto > audit		4.96%
FN+ns	audit !=0, auto = 0, auto < audit		7.02%
0 & "="	auto = audit		88.02%
Accuracy	Accuracy for distress detection		93.78%

b) Individual distress changes for CRCP

The individual distress analysis for CRCP was conducted by comparing the PMIS annual rating data (auto data) without construction sections using automated data collection technology and the manual audit data, the confusion matrix results show in Table 5.12, Table 5.14, Table 5.16, Table 5.18, and Table 5.20. The rates of the automatically measured individual distress by comparing with the audit data in the same data sections are listed in Table 5.13, Table 5.15, Table 5.17, Table 5.19, and Table 5.21.

Table 5.12 Confusion matrix of CRCP SPALLED CRACKS

		Auto data	
		0	!=0
Audit data	0	1009	121
	!=0	110	85

Table 5.13 Data quality analysis of CRCP SPALLED CRACKS

		Section	Rate
CRCP SPALLED CRACKS !=0		85	6.42%
"="	auto = audit	24	1.81%
S	auto > audit	17	1.28%
Ns	auto < audit	44	3.32%
0	auto = 0, audit = 0		76.15%
FP	auto !=0, audit = 0		9.13%
FN	audit !=0, auto = 0		8.30%
FP+s	auto !=0, audit = 0, auto > audit		10.42%
FN+ns	audit !=0, auto = 0, auto < audit		11.62%
0 & "="	auto = audit		77.96%
Accuracy	Accuracy for distress detection		82.57%

Table 5.14 Confusion matrix of CRCP PUNCHOUT

		Auto data	
		0	!=0
Audit data	0	1152	51
	!=0	87	35

Table 5.15 Data quality analysis of CRCP PUNCHOUT

		Section	Rate
CRCP PUNCHOUT !=0		35	2.57%
"="	auto = audit	16	1.21%
s	auto > audit	7	0.53%
ns	auto < audit	11	0.83%

0	auto = 0, audit = 0	86.94%
FP	auto !=0, audit = 0	3.85%
FN	audit !=0, auto = 0	6.57%
FP+s	auto !=0, audit = 0, auto > audit	4.38%
FN+ns	audit !=0, auto = 0, auto < audit	7.40%
0 & "="	auto = audit	88.15%
Accuracy	Accuracy for distress detection	89.58%

Table 5.16 Confusion matrix of CRCP ACP PATCHES

		Auto data	
		0	!=0
Audit data	0	1201	69
	!=0	23	32

Table 5.17 Data quality analysis of CRCP ACP PATCHES

		Section	Rate
CRCP ACP PATCHES !=0		32	2.42%
"="	auto = audit	8	0.60%
s	auto > audit	14	1.06%
ns	auto < audit	10	0.75%
0	auto = 0, audit = 0		90.64%
FP	auto !=0, audit = 0		5.21%
FN	audit !=0, auto = 0		1.74%
FP+s	auto !=0, audit = 0, auto > audit		6.26%
FN+ns	audit !=0, auto = 0, auto < audit		2.49%
0 & "="	auto = audit		91.25%
Accuracy	Accuracy for distress detection		93.06%

Table 5.18 Confusion matrix of CRCP PCC PATCHES

		Auto data	
		0	!=0
Audit data	0	963	75
	!=0	92	195

Table 5.19 Data quality analysis of CRCP PCC PATCHES

		Section	Rate
CRCP PCC PATCHES !=0		195	14.72%
"="	auto = audit	59	4.45%
s	auto > audit	78	5.89%

ns	auto < audit	58	4.38%
0	auto = 0, audit = 0		72.68%
FP	auto !=0, audit = 0		5.66%
FN	audit !=0, auto = 0		6.94%
FP+s	auto !=0, audit = 0, auto > audit		11.55%
FN+ns	audit !=0, auto = 0, auto < audit		11.32%
0 & "="	auto = audit		77.13%
Accuracy	Accuracy for distress detection		87.40%

Table 5.20 Confusion matrix of CRCP AVG CRACK SPACING

		Auto data	
		0	!=0
Audit data	0	0	10
	!=0	0	1315

Table 5.21 Data quality analysis of CRCP AVG CRACK SPACING

		Section	Rate
CRCP AVG CRACK SPACING !=0		1315	99.25%
"="	auto = audit	65	4.91%
s	auto > audit	1041	78.57%
ns	auto < audit	209	15.77%
0	auto = 0, audit = 0		0.00%
FP	auto !=0, audit = 0		0.75%
FN	audit !=0, auto = 0		0.00%
FP+s	auto !=0, audit = 0, auto > audit		79.32%
FN+ns	audit !=0, auto = 0, auto < audit		15.77%
0 & "="	auto = audit		4.91%
Accuracy	Accuracy for distress detection		99.25%

c) Individual distress analysis for JCP

The individual distress analysis for JCP was conducted by comparing the PMIS annual rating data (auto data) without construction sections using automated data collection technology and the manual audit data, the confusion matrix results are shown in Table 5.22 and Table 5.24. The rates of the automatically measured individual distress by comparing with the audit data in the same data sections are listed in Table 5.23 and Table 5.25.

Table 5.22 Confusion matrix of JCP FAILED JNTS CRACKS QTY

		Auto data	
		0	!=0
Audit data	0	151	53
	!=0	47	59

Table 5.23 Data quality analysis of JCP FAILED JNTS CRACKS QTY

		Section	Rate
ACP Alligator Cracking !=0		59	20.00%
"="	auto = audit	23	7.42%
s	auto > audit	20	6.45%
ns	auto < audit	19	6.13%
0	auto = 0, audit = 0		48.71%
FP	auto !=0, audit = 0		17.10%
FN	audit !=0, auto = 0		15.16%
FP+s	auto !=0, audit = 0, auto > audit		23.55%
FN+ns	audit !=0, auto = 0, auto < audit		21.29%
0 & "="	auto = audit		56.13%
Accuracy	Accuracy for distress detection		67.74%

Table 5.24 Confusion matrix of JCP FAILURES QTY

		Auto data	
		0	!=0
Audit data	0	136	55
	!=0	33	86

Table 5.25 Data quality analysis of JCP FAILURES QTY

		Section	Rate
ACP Longitudinal Cracking !=0		86	27.74%
"="	auto = audit	21	6.77%
s	auto > audit	43	13.87%
ns	auto < audit	22	7.10%
0	auto = 0, audit = 0		43.87%
FP	auto !=0, audit = 0		17.74%
FN	audit !=0, auto = 0		10.65%
FP+s	auto !=0, audit = 0, auto > audit		31.61%
FN+ns	audit !=0, auto = 0, auto < audit		17.74%
0 & "="	auto = audit		50.65%
Accuracy	Accuracy for distress detection		71.61%

5.1.3 Accuracy of the Annual Rating Data Compared with the Audit Data

The accuracy analysis of the annual rating data compared with the audit data in this section was conducted using the data retrieved from the Annual Audit Detail Rating Comparison report.

1) Accuracy analysis of Distress Score based on a 15-point Threshold

The accuracy of the distress score based on the TxDOT's currently used 15-point threshold was calculated by comparing the distress scores from the annual rating data and audit data. The accuracy can be used to evaluate how close the automated pavement condition data (annual rating data) is to the ground truth (audit data). The 15-point threshold-based accuracy of the distress score is one of the indexes in Task 5 to evaluate the quality of the automated pavement condition data and the performance of the currently used automated pavement condition data collection technology. The distress score accuracy analysis results are shown in Table 5.26.

After analyzing the data, the accuracy of the distress score for ACP in Fiscal Year (FY) 2017 is 88.824%. It shows that 88.824% of the annual rating data collected using automated data collection technology are correct compared to the audit data in FY 2017. From FY 2017 to 2019, the accuracy of distress scores for ACP increased continually. However, from FY 2019 to 2020, the distress score accuracies have a significant decrease, which means the quality of the automated data collected in FY 2019 is lower than that of FY 2020. In FY 2021, the distress score accuracy increased to 95.218% which is the highest distress score in the recent 5 years. The overall distress score accuracy of 91.472% from FY 2017 to 2021 for ACP passed the TxDOT data quality requirement of 90%. Therefore, the automated data quality based on the ACP distress score is acceptable for TxDOT.

The accuracy of the distress score for CRCP in FY 2017 is 86.195%. From FY 2017 to FY 2018, the accuracy of distress scores for CRCP increased gradually, but from FY 2018 to 2020, the distress score accuracy showed a significant decrease. It indicates that the quality of the automated data collected in FY 2019 and FY 2020 is lower than that of FY 2018. In FY 2021, the distress score accuracy increased to 90.456% which is the highest distress score in 5 years. The overall distress score accuracy of 86.744% from FY 2017 to FY 2021 for CRCP does not pass the TxDOT data quality requirement of 90%. Therefore, the automated data quality based on the CRCP distress score is not acceptable for TxDOT.

The accuracy of distress score for JCP in FY 2017 is 81.463% which is the highest distress score in 5 years. From FY 2017 to FY 2018, the distress score accuracy showed a massive decrease. It indicates that the quality of the automated data collected in FY 2018 was lower than that of FY 2017. The accuracy of distress score continued to decrease till FY 2020 which is the lowest distress score in five years. In FY 2021, the distress score accuracy increased to 76.296%. The overall distress score accuracy for JCP is 77.366% from FY 2017 to FY 2021. It does not pass the TxDOT data quality requirement of 90%. Hence, the automated data quality based on the JCP distress score is not acceptable for TxDOT.

Table 5.26 Data quality analysis of JCP FAILURES QTY

Pavement Types	Fiscal Year					
	2017	2018	2019	2020	2021	Overall
ACP	88.824%	89.749%	92.863%	90.92%	95.218%	91.472%
CRCP	86.195%	89.384%	87.217%	83.433%	90.456%	86.744%
JCP	81.463%	75.214%	76.351%	75.0%	76.296%	77.366%

2) Population analysis of Distress Score differences

a) Percentage of data within the 15-point Distress Score threshold

The percentage of data within the 15-point distress score threshold was calculated by checking the population of distress score differences and calculating the percentage of data that is within the range of -15 points and +15 points of the distress score. This data range is the requirement of TxDOT to check if the comparison of the distress scores from the annual rating data and audit data passes the data accuracy requirement. The percentage of data within 15-point Distress Score results are presented in Table 5.27.

After analyzing the data, the percentage of data between 15-point of the distress score for ACP in FY 2017 is 82.86%. It shows that 82.86% of the annual rating data collected using automated data collection technology are within the 15-point threshold compared to the audit data in FY 2017. From FY 2017 to FY 2019, the percentage of data between 15-point Distress Score increased significantly. It indicates that the data collected using automated technology shows correct data compared to the audit data. From FY 2019 to FY 2020, the percentage of data decreased significantly. The percentage of data for FY 2021 is 96.038% which is the highest percentage of data that is within 15-point Distress Score compared to the audit data. Overall, the percentage is 87.088% from FY 2017 to FY 2021.

The percentage of data between 15 points of the distress score for CRCP in FY 2017 is 73.345%. From FY 2017 to FY 2018, the percentage of data with 15-points distress score increased significantly. It indicates that the data collected using automated technology shows more correct data compared to the audit data. From FY 2018 to FY 2020, the percentage of data decreased drastically. The percentage of data for FY 2021 is 84.652% which is the highest percentage of data within 15-point distress score compared to the audit data. Overall, the percentage is 74.405% from FY 2017 to FY 2021.

The percentage of data between 15 points of the distress score for JCP in FY 2017 is 53.107%. From FY 2017 to FY 2018, the percentage of data between 15-point distress scores decreased significantly. From FY 2018 to FY 2019, the percentage of data increased, but it decreased in FY 2020. The percentage of data for FY 2021 is 57.66% which is the highest percentage of data that is within a 15-point distress score compared to the audit data. Overall, the percentage is 48.7% from FY 2017 to FY 2021.

Table 5.27 Percentage of data within15-point Distress Score

Pavement Types	Fiscal Year					
	2017	2018	2019	2020	2021	Overall
ACP	82.86%	83.204%	90.007%	84.503%	96.038%	87.088%
CRCP	73.745%	81.095%	76.353%	66.642%	84.652%	74.405%
JCP	53.107%	42.348%	46.756%	44.337%	57.66%	48.7%

b) Develop the point-based threshold based on $\mu \mp 2\sigma$

The range of distress score differences was calculated by checking the population of distress score differences and calculating the points that hold 95% of the data or within two standard deviations from the mean. This data range is generally accepted by many highway agencies to check the range of the distress score differences between the annual rating data and audit data. The distress score difference thresholds are shown in Table 5.28.

The range of distress score difference in points holding 95% of the data for ACP in FY 2017 is (-21, 23). It indicates that 95% of the data collected using automated technology (annual data) and audit data falls within the range. From FY 2017 to FY 2018, the range did not change. From FY 2018 to FY 2019, the range of distress score difference decreased and then increased in FY 2020. The range of distress score difference for FY 2021 is (-15, 15) which is the lowest Distress Score range in 5 years. Overall, the range of distress score difference is (-19, 21) from FY 2017 to FY 2021.

The range of distress score difference in points holding 95% of the data for CRCP in FY 2017 is (-29, 22). From FY 2017 to FY 2018, the distress score difference range decreased significantly. From FY 2018 to FY 2020, the range of distress score differences increased. The range for FY 2021 is (-21, 21) which is the lowest Distress Score range in 5 years. Overall, the distress score difference range is (-26, 26) from FY 2017 to FY 2021.

The range of distress score difference in points holding 95% of the data for JCP in FY 2017 is (-46, 32). From FY 2017 to FY 2018, the range decreased significantly. From FY 2018 to FY 2019, the range of distress score difference decreased and then increased in FY 2020. The range for FY 2021 is (-41, 32) which is the lowest Distress Score range in 5 years. Overall, the range of distress score difference is (-26, 26) from FY 2017 to FY 2021.

Table 5.28 Distress Score difference thresholds of the data (within 2 standard deviations)

Pavement Types	Fiscal Year					
	2017	2018	2019	2020	2021	Overall
ACP	(-21, 23)	(-21, 23)	(-17, 19)	(-19, 22)	(-15, 15)	(-19, 21)
CRCP	(-29, 22)	(-23, 22)	(-25, 25)	(-27, 34)	(-21, 21)	(-26, 26)
JCP	(-46, 32)	(-60, 40)	(-53, 37)	(-55, 46)	(-41, 32)	(-50, 37)

c) Individual distress measurement difference analysis -- $\mu \mp 2\sigma$ based threshold and accuracy for each distress type

The thresholds for each pavement distress type are calculated based on mean plus or minus two standard deviations, and the dataset used for thresholds development is from the overall data (FY 2017 to FY 2021). The thresholds are determined to check the accuracy of each distress type on the respective Fiscal Year and overall. The results are presented in Table 5.29.

After analyzing the data, the threshold for ACP patching difference between the auto data and audit data is (-9.516, 9.27). It indicates that 95% of the ACP patching differences recorded from FY 2017 to FY 2021 are within this range. Based on this threshold, the accuracy of the distress-type Patching for ACP in FY 2017 is 95.849%. It shows that 95.849% of Patching data recorded are within the difference threshold in FY 2017. From FY 2017 to 2019, the accuracy of Patching for ACP increased continually. However, from FY 2019 to 2020, the accuracy of ACP Patching has a slight decrease. In FY 2021, the accuracy increased to 97.949% which is the highest accuracy for ACP patching in the recent 5 years. The overall ACP Patching accuracy from FY 2017 to 2021 is 97.23%.

The threshold of the quantity differences for ACP Failure is (-1.55, 1.51) based on the data from FY 2017 to 2021. The accuracy of ACP Failure for FY 2017 is 99.813% which is the highest accuracy in five years. From FY 2017 to FY 2018, the accuracy of ACP Failure decreased significantly. From FY 2018 to 2019, the accuracy increased, but from FY 2019 to FY 2020, there is a massive drop in the accuracy. From FY 2020 to FY 2021, the accuracy increased to 99.246%. The overall accuracy for ACP Failure is 95.5% for FY 2017 to FY 2021.

The threshold of the quantity differences for ACP Block Cracking is (-4.81, 4.84). The accuracy of ACP Block Cracking in FY 2017 is 91.035%. From FY 2017 to FY 2018, the accuracy increased significantly which is the highest accuracy in five years. From FY 2018 to FY 2020, the accuracy decreased significantly. In FY 2020, the accuracy of ACP Block Cracking is 91.49% which is the lowest in five years. From FY 2020 to FY 2021, the accuracy increased to 95.361%. The overall accuracy for ACP Block Cracking is 95.46% for FY 2017 to FY 2021.

The threshold for ACP Alligator Cracking difference between the auto data and audit data is (-12.29, 11.51). Based on the threshold, the accuracy of ACP Alligator Cracking for FY 2017 is

93.292%. From FY 2017 to FY 2018, the accuracy of ACP Alligator Cracking decreased significantly. From FY 2018 to FY 2019, the accuracy increased significantly but from FY 2019 to FY 2020, the accuracy decreased to 95.065%. The accuracy for ACP Alligator cracking in FY 2021 is 99.943% which is the highest accuracy for ACP Alligator Cracking in five years. The overall accuracy for ACP Alligator Cracking is 95.45% for FY 2017 to FY 2021.

The threshold for ACP Longitudinal Cracking difference is (-41.99, 46.14). Based on the threshold, the accuracy of ACP Longitudinal Cracking for FY 2017 is 86.954%. From FY 2017 to FY 2020, the accuracy of ACP Longitudinal Cracking increased significantly to 97.933% which is the highest accuracy in five years. From FY 2020 to FY 2021, the accuracy decreased significantly. The overall accuracy of ACP Longitudinal Cracking from FY 2017 to FY 2021 is 95.45%.

The threshold of the quantity differences for ACP Transverse Cracking is (-3.48, 3.18). Based on the threshold, the accuracy of ACP Transverse Cracking for FY 2017 is 93.167%. From FY 2017 to FY 2018, the accuracy for ACP Transverse Cracking decreased significantly. From FY 2018 to FY 2019, the accuracy increased to 98.952%. From FY 2019 to FY 2020, ACP Transverse cracking accuracy decreased again. From 2020 to FY 2021, the accuracy increased to 99.233% which is the highest accuracy for ACP Transverse Cracking in five years. The overall accuracy for ACP Transverse Cracking from FY 2017 to FY 2021 is 95.45%

After analyzing the data of CRCP, the threshold for CRCP Spalled Cracks quantity difference between the annual data and audit data is (-11.61, 10.72). Based on the threshold, the accuracy of CRCP Spalled Cracks for FY 2017 is 100%. It shows that 100% of the data recorded are within the difference threshold in FY 2017. From FY 2017 to FY 2019, the accuracy of CRCP Spalled Cracks decreased drastically to 76.83% which is the lowest accuracy in five years. From FY 2019 to FY 2020, the accuracy increased to 99.913%. The accuracy of CRCP Spalled Cracks for FY 2020 is the highest in five years. From FY 2020 to FY 2021, the accuracy decreased significantly. The overall accuracy of CRCP Spalled Cracks from FY 2017 to FY 2021 is 95.446%.

The threshold for CRCP Punchout difference is (-1.79, 1.66). Based on the threshold, the accuracy of CRCP Punchout for FY 2017 is 98.28%. From FY 2017 to FY 2018, the accuracy of CRCP Punchout increased slightly. From FY 2018 to FY 2019, the accuracy decreased drastically to 88.148%. The accuracy of CRCP Punchout for FY 2021 is 99.821% which is the highest accuracy in five years. The overall accuracy of CRCP Punchout from FY 2017 to FY 2021 is 95.46%.

The threshold for CRCP ACP Patches difference is (-6.41, 7.03). The accuracy of CRCP ACP Patches for FY 2017 based on the threshold is 100%. From FY 2017 to FY 2019, the accuracy for CRCP ACP Patches decreased drastically to 70.014% which is the lowest accuracy in five years. From FY 2019 to FY 2021, the accuracy of CRCP ACP Patches increased significantly to 100%. The overall accuracy of CRCP ACP Patches from FY 2017 to FY 2021 is 95.437%.

The threshold for CRCP PCC Patches difference is (-6.57, 6.55). The accuracy of CRCP PCC Patches for FY 2017 is 94.573%. From FY 2017 to FY 2018, the accuracy of CRCP PCC Patches decreased significantly. From FY 2018 to FY 2019, the accuracy decreased significantly to 87.894%. From FY 2019 to FY 2021, the accuracy of CRCP PCC Patches increased significantly to 99.36% which is the highest accuracy in five years. The overall accuracy of CRCP PCC Patches from FY 2017 to FY 2021 is 95.45%.

Based on the data, the threshold for CRCP Average Crack Spacing difference is (-32.11, 61.83). The accuracy of CRCP Average Crack Spacing for FY 2017 is 99.347%. From FY 2017 to FY 2018, the accuracy of CRCP Average Crack Spacing increased significantly. From FY 2018 to FY 2021, the accuracy decreased significantly to 86.455% which is the lowest accuracy for CRCP Average Crack Spacing in five years. The overall accuracy of CRCP Average Crack Spacing from FY 2017 to FY 2021 is 95.45%

After analyzing the data of JCP, the threshold for JCP Failed Joint and Cracks quantity difference between the annual data and audit data is (-17.22, 20.73). Based on the threshold, the accuracy of JCP Failed Joint and Cracks for FY 2017 is 90.578%. It shows that 90.578% of the data recorded are within the difference threshold in FY 2017. From FY 2017 to FY 2019, the accuracy of JCP Failed Joint and Cracks increased significantly to 100%. From FY 2019 to FY 2020, the accuracy decreased to 82.342% which is the lowest accuracy in five years. From FY 2020 to FY 2021, the accuracy of JCP Failed Joint and Cracks increased significantly. The overall accuracy of JCP Failed Joint and Cracks from FY 2017 to FY 2021 is 95.471%.

The threshold for JCP Failures difference is (-28.62, 33.3). The accuracy of JCP Failures for FY 2017 is 100%. From FY 2017 to FY 2018, the accuracy of JCP Failures decreased significantly. From FY 2018 to FY 2021, the accuracy increased significantly to 100%. The overall accuracy of JCP Failures from FY 2017 to FY 2021 is 95.447%.

The threshold for JCP Shattered Slabs difference is (-12.35, 13.53). Based on the threshold, the accuracy of JCP Shattered Slabs for FY 2017 is 100%. From FY 2017 to FY 2018, the accuracy of JCP Shattered Slabs decreased drastically to 58.949% which is the lowest accuracy in five years. From FY 2018 to FY 2021, the accuracy increased to 100% and remained the same for three years. The overall accuracy of JCP Shattered Slabs from FY 2017 to FY 2021 is 95.45%.

The threshold for JCP Slabs with Longitudinal Cracks difference is (-13.73, 12.1). The accuracy of JCP Slabs with Longitudinal Cracks for FY 2017 is 99.631%. From FY 2017 to FY 2018, the accuracy of JCP Slabs with Longitudinal Cracks increased significantly to 100% which is the highest accuracy for JCP Slabs with Longitudinal Cracks. From FY 2018 to FY 2020, the accuracy decreased significantly to 84.682% which is the lowest accuracy in five years From FY 2020 to FY 2021, the accuracy of JCP Slabs with Longitudinal Cracks increased significantly. The overall accuracy of JCP Slabs with Longitudinal Cracks from FY 2017 to FY 2021 is 95.452%.

The threshold for JCP PCC Patches difference is (-14.84, 15.4). The accuracy of JCP PCC Patches for FY 2017 is 85.435%. From FY 2017 to FY 2019, the accuracy of JCP PCC Patches increased significantly to 99.96% which is the highest accuracy for JCP PCC Patches in five years. From FY 2019 to FY 2020, the accuracy decreased significantly to 97.28%. From FY 2020 to FY 2021, the accuracy of JCP PCC Patches increased significantly. The overall accuracy of JCP PCC Patches from FY 2017 to FY 2021 is 95.45%.

The threshold for JCP Apparent Joint Spacing difference is (-21.38, 23.98). The accuracy of JCP Apparent Joint Spacing for FY 2017 is 97.132%. From FY 2017 to FY 2018, the accuracy of JCP Apparent Joint Spacing decreased. From FY 2018 to FY 2019, the accuracy increased significantly to 99.17% which is the highest accuracy of JCP Apparent Joint Spacing in five years. From FY 2019 to FY 2021, the accuracy of JCP Apparent Joint Spacing decreased significantly. The overall accuracy of JCP Apparent Joint Spacing from FY 2017 to FY 2021 is 95.45%.

Table 5.29 $\mu \mp 2\sigma$ based threshold and accuracy for each distress type

Pavement Types	Distress Type	Fiscal Year						Threshold
		2017	2018	2019	2020	2021	Overall	
ACP	ACP PATCHING (PCT)	82.544%	96.268%	99.76%	96.804%	99.273%	95.45%	(-9.52,9.27)
	ACP FAILURE (QTY)	99.813%	96.114%	97.205%	88.081%	99.246 %	95.5%	(-1.55,1.51)
	ACP BLOCK CRACKING (PCT)	91.035%	99.546%	98.02%	91.49%	95.361%	95.46%	(-4.81,4.84)
	ACP ALLIGATOR CRACKING (PCT)	93.292%	90.432%	95.816%	95.065%	99.943%	95.45%	(-12.29,11.51)
	ACP LONGITUDE CRACKING (Feet)	86.954%	96.475%	97.397%	97.933%	97.726%	95.45%	(-41.99,46.14)
	ACP TRANSVERSE CRACKING (QTY)	93.167%	92.68%	98.952%	91.502%	99.233%	95.45%	(-3.48,3.18)
CRCP	CRCP SPALLED CRACKS (QTY)	100%	87.376%	76.83%	99.913%	98.402%	95.446%	(-11.61,10.72)
	CRCP PUNCHOUT (QTY)	98.28%	98.332%	99.67%	88.148%	99.821%	95.46%	(-1.79,1.66)
	CRCP ACP PATCHES (QTY)	100%	91.762%	70.014%	99.963%	100%	95.437%	(-6.41,7.03)
	CRCP PCC PATCHES (QTY)	94.573%	99.231%	87.894%	96.205%	99.36%	95.45%	(-6.57,6.55)

	CRCP AVG CRACK SPACING (QTY)	99.347%	99.662%	98.64%	88.541%	86.455%	95.45%	(-32.11,61.83)
JCP	JCP FAILED JNTS CRACKS (QTY)	90.578%	95.77%	100%	82.342%	100%	95.471%	(-17.22,20.73)
	JCP FAILURES (QTY)	100%	64.564%	97.598%	100%	100%	95.447%	(-28.62,33.3)
	JCP SHATTERED SLABS (QTY)	100%	58.949%	100%	100%	100%	95.45%	(-12.35,13.53)
	JCP SLABS WITH LONGITUDINAL CRACKS (QTY)	99.631%	100%	85.611%	84.682%	98.91%	95.452%	(-13.73,12.1)
	JCP PCC PATCHES (QTY)	85.435%	92.418%	99.96%	97.28%	99.92%	95.45%	(-14.84,15.4)
	JCP APPARENT JOINT SPACE(Feet)	97.132%	95.563%	99.17%	93.93%	88.28%	95.45%	(-21.38,23.98)

5.1.4 Precision of the Annual Rating Data

Precision is the degree to which an instrument or process will repeat the same value. The precision comparison of the annual rating data is conducted by comparing every two years' auto data in the specific same sections using automated data collection technology. Based on the confusion matrix in Figure 33, the precision is equal to *TP/(TP+FP).*

1) Indexes

The indexes included in the annual rating data analysis are 11 distress types of ACP, CRCP, and JCP, which are listed as follows.

ACP: alligator cracking (ACP ALLIGATOR CRACKING PCT), longitudinal cracking (ACP LONGITUDE CRACKING), transverse cracking (ACP TRANSVERSE CRACKING QTY), patching (ACP PATCHING PCT).

CRCP: CRCP SPALLED CRACKS, CRCP PUNCHOUT, CRCP ACP PATCHES, CRCP PCC PATCHES, CRCP AVG CRACK SPACING.

JCP: JCP FAILED JNTS CRACKS QTY, JCP FAILURES QTY.

The overall dataset used for individual distress analysis is the paired PMIS annual rating data with the audit data from FY 2017 to 2021 shown in Table 5.30.

Table 5.30 Dataset for individual distress analysis

Pavement Type	Sections used	Pairs for two years' comparison
JCP	8661	2846
CRCP	4967	1142
ACP	185356	64636

b) Thresholds

The precision analysis of the annual rating data was conducted by comparing every two years' PMIS annual rating data (auto data) without construction sections using automated data collection technology. For the precision analysis, every two years, 15% of the distress measurements were considered an acceptable bias from the automated data collection equipment. A confusion matrix was used to analyze the results of the 2-year difference in the distress measurements. In the confusion matrix, TP represents the number of sections with that 2-year difference of the distress measurements within the range of 0% to 15%, FN represents the number of sections with that 2-year difference of the distress measurements less than -15%, FP represents the number of sections with that 2-year difference of the distress measurements more than 15%, and TN represents the number of sections with that 2-year difference of the distress measurements within the range of -15% to 0%. The explanation of the confusion matrix is shown in Figure 5.2.

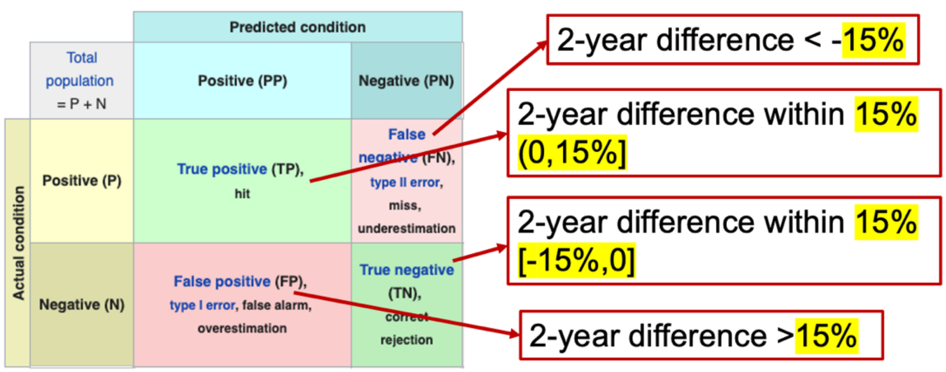


Figure 5.2 Confusion Matrix for 2-Year Difference in the Distress Measurements

a) Precision analysis of annual rating data for ACP

The updated precision analysis of annual rating data for ACP was presented in Table 5. 31 to Table 5.34.

Table 5. 31 Precision analysis of ACP alligator cracking

Result	counts	Rate
FN	710	1.10%
T-	15870	24.55%
0	37605	58.18%
T+	10012	15.49%
FP	439	0.68%
Precision:	--	99.31%

Table 5.32 Precision analysis of ACP longitudinal cracking

Result	counts	Rate
FN	22	0.03%
T-	30016	46.44%
0	10582	16.37%
T+	24004	37.14%
FP	12	0.02%
Precision:	--	99.98%

Table 5.33 Precision analysis of ACP transverse cracking

Result_1	counts	Rate
FN	6	0.01%
T-	7124	11.02%
0	53324	82.50%
T+	4182	6.47%
FP	0	0.00%
Precision:	--	100%

Table 5.34 Precision analysis of ACP patching

Result_1	counts	Rate
FN	891	1.38%
T-	3452	5.34%
0	54443	84.23%
T+	4884	7.56%
FP	966	1.49%
Precision:	--	98.48%

b) Precision analysis of annual rating data for CRCP

The updated precision analysis of annual rating data for CRCP was presented in Table 5.35 to

Table 5.39.

Table 5.35 Precision analysis of CRCP SPALLED CRACKS

Result_1	counts	Rate
FN	0	0.00%
T-	103	9.02%
0	899	78.72%
T+	140	12.26%
FP	0	0.00%
Precision:	--	100%

Table 5.36 Precision analysis of CRCP PUNCHOUT

Result_1	counts	Rate
FN	0	0.00%
T-	52	4.55%
0	1017	89.05%
T+	73	6.39%
FP	0	0.00%
Precision:	--	100%

Table 5.37 Precision analysis of CRCP ACP PATCHES

Result_1	counts	Rate
FN	0	0.00%
T-	52	4.55%
0	1053	92.21%
T+	37	3.24%
FP	0	0.00%
Precision:	--	100%

Table 5.38 Precision analysis of CRCP PCC PATCHES

Result_1	counts	Rate
FN	0	0.00%
T-	132	11.56%
0	842	73.73%
T+	168	14.71%
FP	0	0.00%
Precision:	--	100%

Table 5.39 Precision analysis of CRCP AVG CRACK SPACING

Result_1	counts	Rate
FN	99	8.67%
T-	182	15.94%
0	124	10.86%
T+	393	34.41%
FP	344	30.12%
Precision:	--	67.02%

c) Precision analysis of annual rating data for JCP

The updated precision analysis of annual rating data for JCP was presented in Table 5.40 and Table 5.41.

Table 5.40 Precision analysis of JCP FAILED JNTS CRACKS

Result_1	counts	Rate
FN	0	0.00%
T-	98	3.44%
0	2639	92.73%
T+	109	3.83%
FP	0	0.00%
Precision:	--	100%

Table 5.41 Precision analysis of JCP FAILURES

Result_1	counts	Rate
FN	0	0.00%
T-	88	3.09%
0	2643	92.87%
T+	115	4.04%
FP	0	0.00%
Precision:	--	100%

5.2 Data Quality Consistency Check Thresholds

The Performing Agency proposed three methods to determine the data quality thresholds based on the data quality consistency check components developed in TM5:

- Index 1: Yearly change of sections with distress decrease
- Index 2: Individual distress change
- Index 3: Accuracy of the annual rating data comparing with the audit data

- Index 4: Precision of the annual rating data
- Index 5: Yearly change of each section DS, CS, and RS

The traditional method that is used for threshold development is based on the two-sigma method. The two-sigma method has been used by many other state DOTs in their data quality program. The threshold development of these data quality consistency check components/indexes for each distress type applied stratified factors of highway service level (IH, US, SH, and FM).

5.2.1 Sigma method for data quality index threshold development

The research team selected sigma-based method as an alternative of data quality check indexes threshold development method. The data pre-analysis results show that the data in indexes 2, 3, 4, and 5 follow the normal distribution. Therefore, the research team first developed 2-sigma thresholds based on engineers' experience.

Index #1: No. of sections with yearly distress decrease

The threshold value of yearly change of sections with a distress value decrease or distress value disappearance is zero tolerance, which means the sections which have distress decrease or distress disappearance will be considered as having data quality issues and shall be flagged for re-evaluation. However, based on the data analysis, the zero tolerance of sections with yearly distress decrease is too tight, the research team will only recommend for internal evaluation for annual data quality check.

Index #2 Individual distress change

The individual thresholds of each distress type were determined by comparing the annual PMIS rating data using automated data collection technology and the manual audit data. The research team determined the threshold using the audit dataset retrieved from the Annual Audit Detail Rating Comparison report. The threshold of each distress type was calculated by checking the measurement differences between the annual rating data and the audit data and determining the points that hold 95% of the data or within two standard deviations of the mean.

Based on the threshold, the sections that are within and above the threshold are selected. Based on the threshold, 'FN' represents the number of sections that are outside of the threshold range in negative values of the difference in distress measurement values when comparing the annual data and audit data. 'FP' represents the number of sections that are out of the threshold range in positive values of the difference in measurement values. 'T-' represents the number of sections that are within the threshold range in negative values comparing the annual data with the audit data. 'T+' represents the number of sections that are within the threshold range in positive values comparing the annual data with the audit data. '0' represents the number of sections that exhibited no change in comparison of annual data and audit data.

The thresholds are determined following two stratified factors (Pavement type and highway service level). The pavement types are categorized as: Asphalt Concrete Pavement (ACP), Continuously Reinforced Concrete Pavement (CRCP), and Jointed Concrete Pavement (JCP). There are four types of highway service levels: Interstate Highway (IH), US Highway (US), State Highway (SH), and Farm to Market Road (FM). The pavement surface age data is not included due to unavailability of sufficient data for analysis.

Table 5.42 shows the thresholds of each distress type of ACP for a specific highway service level. The threshold is in the format of the difference in the distress measurement value comparing the annual rating data and audit data. A total of 24 thresholds have been determined.

Table 5.42 Thresholds for individual analysis for ACP

Distress Type	**Highway service level**	**No. of Sections**	**Threshold**
ACP Alligator Cracking (Percentage of Wheel path Length)	**IH**	Sections % FN 74 0.024866 T- 359 0.120632 0 2134 0.717070 T+ 378 0.127016 FP 31 0.010417	(-11.62, 10.59)
	US	Sections % FN 248 0.025585 T- 1087 0.112143 0 6800 0.701537 T+ 1468 0.151449 FP 90 0.009285	(-15.81, 14.37)
	SH	Sections % FN 315 0.028553 T- 1274 0.115482 0 7639 0.692440 T+ 1702 0.154278 FP 102 0.009246	(-14.49, 13,01)
	FM	Sections % FN 615 0.022141 T- 2572 0.092598 0 18971 0.683000 T+ 5257 0.189264 FP 361 0.012997	(-9.83, 9.6)
ACP Longitudinal Cracking (Length in Feet per station).	**IH**	Sections % FN 81 0.027227 T- 820 0.275630 0 885 0.297479 T+ 1134 0.381176 FP 55 0.018487	(-64.0, 60.59)
	US	Sections % FN 364 0.037553 T- 2048 0.211286 0 3241 0.334365 T+ 3777 0.389663 FP 263 0.027133	(-50.51, 49.91)
	SH	Sections % FN 369 0.033451 T- 2256 0.204515	(-45.69, 47.48)

		0 3465 0.314115 T+ 4649 0.421449 FP 292 0.026471	
	FM	Sections % FN 670 0.024122 T- 4557 0.164062 0 7741 0.278694 T+ 13939 0.501836 FP 869 0.031286	(-33.72, 41.34)
ACP Transverse Cracking (Quantity)	**IH**	Sections % FN 69 0.023185 T- 318 0.106855 0 2350 0.789651 T+ 210 0.070565 FP 29 0.009745	(-3.91, 3.62)
	US	Sections % FN 242 0.024966 T- 1636 0.168782 0 6956 0.717631 T+ 752 0.077582 FP 107 0.011039	(-4.04, 3.44)
	SH	Sections % FN 197 0.017857 T- 1682 0.152466 0 8143 0.738125 T+ 927 0.084028 FP 83 0.007524	(-4.91, 4.47)
	FM	Sections % FN 589 0.021205 T- 1950 0.070204 0 23248 0.836982 T+ 1695 0.061024 FP 294 0.010585	(-2.38, 2.24)
ACP Patching (Percentage of Lane Area)	**IH**	Sections % FN 45 0.015121 T- 93 0.031250 0 2693 0.904906 T+ 97 0.032594 FP 48 0.016129	(-7.29, 7.44)
	US	Sections % FN 104 0.010729 T- 381 0.039307 0 8798 0.907665 T+ 312 0.032188 FP 98 0.010110	(-9.62, 9.39)
	SH	Sections %	(-9.41, 9.04)

		FN	180	0.016316	
		T-	570	0.051668	
		0	9786	0.887056	
		T+	384	0.034808	
		FP	112	0.010152	
	FM		Sections	%	(-9.73, 9.49)
		FN	448	0.016129	
		T-	1564	0.056308	
		0	24188	0.870824	
		T+	1165	0.041943	
		FP	411	0.014797	
ACP Block Cracking (Percentage of Lane Area)	**IH**		Sections	%	(-7.71, 7.6)
		FN	10	0.003360	
		T-	6	0.002016	
		0	2931	0.984879	
		T+	17	0.005712	
		FP	12	0.004032	
	US		Sections	%	(-5.08, 5.26)
		FN	31	0.003198	
		T-	16	0.001651	
		0	9481	0.978129	
		T+	86	0.008872	
		FP	79	0.008150	
	SH		Sections	%	(-6.75, 6.68)
		FN	36	0.003263	
		T-	31	0.002810	
		0	10766	0.975888	
		T+	131	0.011875	
		FP	68	0.006164	
	FM		Sections	%	(-3, 3.03)
		FN	52	0.001872	
		T-	21	0.000756	
		0	27446	0.988119	
		T+	137	0.004932	
		FP	120	0.004320	
ACP Failure (Quantity)	**IH**		Sections	%	(-1.37, 1.28)
		FN	42	0.014113	
		T-	74	0.024866	
		0	2826	0.949597	
		T+	26	0.008737	
		FP	8	0.002688	
	US		Sections	%	(-1.39, 1.38)
		FN	45	0.004643	
		T-	102	0.010523	
		0	9464	0.976375	
		T+	47	0.004849	

		FP 35 0.003611	
	SH	Sections % FN 216 0.019579 T- 0 0.000000 0 10709 0.970722 T+ 0 0.000000 FP 107 0.009699	(-0.92, 0.88)
	FM	Sections % FN 232 0.008353 T- 626 0.022537 0 26386 0.949957 T+ 381 0.013717 FP 151 0.005436	(-1.8, 1.75)

Table 5.43 shows the thresholds of each distress type of CRCP (Continuously Reinforced Concrete Pavement) for a specific highway service level. The threshold will be in the format of the difference in the distress measurement value comparing the annual rating data and audit data. A total of 20 thresholds are determined.

Table 5.43 Thresholds for individual analysis for CRCP

Distress Type	**Highway service level**	**No. of Sections**	**Threshold**
CRCP Spalled Cracks (Quantity)	**IH**	Sections % FN 32 0.019988 T- 122 0.076202 0 1293 0.807620 T+ 136 0.084947 FP 18 0.011243	(-3.88, 3.6)
	US	Sections % FN 6 0.008982 T- 63 0.094311 0 539 0.806886 T+ 57 0.085329 FP 3 0.004491	(-10.4, 9.75)
	SH	Sections % FN 13 0.016993 T- 76 0.099346 0 618 0.807843 T+ 58 0.075817 FP 0 0.000000	(-21.56, 18.99)
	FM	Sections % FN 3 0.012658 T- 20 0.084388 0 194 0.818565 T+ 20 0.084388 FP 0 0.000000	(-2.26, 2.07)

CRCP Punchout (Quantity)	**IH**	Sections % FN 30 0.018738 T- 105 0.065584 0 1399 0.873829 T+ 51 0.031855 FP 16 0.009994	(-2.2, 1.98)
	US	Sections % FN 10 0.014970 T- 32 0.047904 0 597 0.893713 T+ 20 0.029940 FP 9 0.013473	(-1.27, 1.21)
	SH	Sections % FN 18 0.023529 T- 38 0.049673 0 665 0.869281 T+ 28 0.036601 FP 16 0.020915	(-1.47, 1.41)
	FM	Sections % FN 11 0.046414 T- 0 0.000000 0 219 0.924051 T+ 0 0.000000 FP 7 0.029536	(-0.7, 0.64)
CRCP ACP Patches (Quantity)	**IH**	Sections % FN 8 0.004997 T- 43 0.026858 0 1500 0.936914 T+ 29 0.018114 FP 21 0.013117	(-3.1, 3.26)
	US	Sections % FN 3 0.004491 T- 16 0.023952 0 617 0.923653 T+ 19 0.028443 FP 13 0.019461	(-4.45, 4.97)
	SH	Sections % FN 0 0.000000 T- 28 0.036601 0 695 0.908497 T+ 21 0.027451 FP 21 0.027451	(-11.05, 12.95)
	FM	Sections % FN 2 0.008439 T- 10 0.042194 0 221 0.932489	(-4.83, 4.68)

		T+	2	0.008439	
		FP	2	0.008439	
CRCP PCC Patches (Quantity)	**IH**		Sections	%	(-4.97, 5.33)
		FN	38	0.023735	
		T-	115	0.071830	
		0	1251	0.781387	
		T+	146	0.091193	
		FP	51	0.031855	
	US		Sections	%	(-7.91, 7.27)
		FN	11	0.016467	
		T-	73	0.109281	
		0	521	0.779940	
		T+	53	0.079341	
		FP	10	0.014970	
	SH		Sections	%	(-8.1, 7.63)
		FN	8	0.010458	
		T-	58	0.075817	
		0	636	0.831373	
		T+	59	0.077124	
		FP	4	0.005229	
	FM		Sections	%	(-6.78, 7.42)
		FN	2	0.008439	
		T-	27	0.113924	
		0	171	0.721519	
		T+	29	0.122363	
		FP	8	0.033755	
CRCP Average Crack Spacing (Average observed spacing in feet)	**IH**		Sections	%	(-31.3, 57.61)
		FN	26	0.016240	
		T-	140	0.087445	
		0	103	0.064335	
		T+	1215	0.758901	
		FP	117	0.073079	
	US		Sections	%	(-29.18, 51.48)
		FN	8	0.011976	
		T-	76	0.113772	
		0	49	0.073353	
		T+	484	0.724551	
		FP	51	0.076347	
	SH		Sections	%	(-31.94, 67.92)
		FN	10	0.013072	
		T-	90	0.117647	
		0	39	0.050980	
		T+	566	0.739869	
		FP	60	0.078431	
	FM		Sections	%	(-32.81, 86.72)
		FN	3	0.012658	

		T- 32 0.135021 0 11 0.046414 T+ 191 0.805907 FP 0 0.000000	

Table 5.44 shows the thresholds of each distress type of JCP (Jointed Concrete Pavement) for a specific highway service level. The threshold is in the format of the difference in the distress measurement value comparing the annual rating data and audit data. A total of 24 thresholds have been determined.

Table 5.44 Thresholds for individual analysis for JCP

Distress Type	**Highway service level**	**No. of Sections**	**Threshold**
JCP Failed Joint Cracks (Quantity)	**IH**	Sections % FN 0 0.000000 T- 18 0.120000 0 80 0.533333 T+ 47 0.313333 FP 5 0.033333	(-22.62, 28.65)
	US	Sections % FN 0 0.000000 T- 52 0.191176 0 130 0.477941 T+ 79 0.290441 FP 11 0.040441	(-19.48, 24.77)
	SH	Sections % FN 5 0.023923 T- 44 0.210526 0 88 0.421053 T+ 64 0.306220 FP 8 0.038278	(-10.61, 11.77)
	FM	Sections % FN 4 0.040816 T- 18 0.183673 0 57 0.581633 T+ 15 0.153061 FP 4 0.040816	(-3.51, 3.3)
JCP Failures (Quantity)	**IH**	Sections % FN 0 0.000000 T- 18 0.120000 0 72 0.480000 T+ 58 0.386667 FP 2 0.013333	(-54.07, 65.0)
	US	Sections % FN 1 0.003676 T- 45 0.165441	(-18.19, 21.86)

		0 125 0.459559 T+ 98 0.360294 FP 3 0.011029	
	SH	Sections % FN 5 0.023923 T- 35 0.167464 0 77 0.368421 T+ 86 0.411483 FP 6 0.028708	(-13.71, 16.26)
	FM	Sections % FN 1 0.010204 T- 17 0.173469 0 47 0.479592 T+ 28 0.285714 FP 5 0.051020	(-7.18, 9.65)
JCP PCC Patches (Quantity)	**IH**	Sections % FN 3 0.020000 T- 28 0.186667 0 85 0.566667 T+ 27 0.180000 FP 7 0.046667	(-12.28, 12.73)
	US	Sections % FN 4 0.014706 T- 68 0.250000 0 156 0.573529 T+ 39 0.143382 FP 5 0.018382	(-19.66, 19.22)
	SH	Sections % FN 4 0.019139 T- 25 0.119617 0 151 0.722488 T+ 22 0.105263 FP 7 0.033493	(-11.43, 12.06)
	FM	Sections % FN 0 0.000000 T- 9 0.091837 0 65 0.663265 T+ 21 0.214286 FP 3 0.030612	(-9.03, 12.32)
JCP Shattered Slabs (Quantity)	**IH**	Sections % FN 0 0.000000 T- 0 0.000000 0 140 0.933333 T+ 7 0.046667 FP 3 0.020000	(-25.45, 29.65)
	US	Sections %	(-1.39, 1.56)

		FN 0 0.000000 T- 0 0.000000 0 266 0.977941 T+ 3 0.011029 FP 3 0.011029	
	SH	Sections % FN 0 0.000000 T- 2 0.009569 0 189 0.904306 T+ 13 0.062201 FP 5 0.023923	(-4.81, 5.66)
	FM	Sections % FN 1 0.010204 T- 1 0.010204 0 93 0.948980 T+ 2 0.020408 FP 1 0.010204	(-2.63, 2.75)
JCP Slabs with Longitudinal Cracks (Quantity)	**IH**	Sections % FN 0 0.000000 T- 0 0.000000 0 140 0.933333 T+ 7 0.046667 FP 3 0.020000	(-25.45, 29.65)
	US	Sections % FN 0 0.000000 T- 0 0.000000 0 266 0.977941 T+ 3 0.011029 FP 3 0.011029	(-1.39, 1.56)
	SH	Sections % FN 0 0.000000 T- 2 0.009569 0 189 0.904306 T+ 13 0.062201 FP 5 0.023923	(-4.81, 5.66)
	FM	Sections % FN 1 0.010204 T- 1 0.010204 0 93 0.948980 T+ 2 0.020408 FP 1 0.010204	(-2.63, 2.75)
JCP Apparent Joint Space	**IH**	Sections % FN 0 0.000000 T- 17 0.113333 0 115 0.766667 T+ 13 0.086667	(-9.36, 11.1)

		FP 5 0.033333	
	US	Sections % FN 7 0.025735 T- 36 0.132353 0 180 0.661765 T+ 42 0.154412 FP 7 0.025735	(-20.88, 21.09)
	SH	Sections % FN 0 0.000000 T- 44 0.210526 0 112 0.535885 T+ 36 0.172249 FP 17 0.081340	(-28.46, 36.76)
	FM	Sections % FN 4 0.040816 T- 9 0.091837 0 68 0.693878 T+ 15 0.153061 FP 2 0.020408	(-10.02, 8.49)

Index #3 Accuracy of the annual rating data compared with the audit data

The accuracy of the annual rating data compared with the audit data in index 3 was determined using the dataset retrieved from the Annual Audit Detail Rating Comparison report. The accuracy is determined to check the similarity of the annual rating data compared to the ground truth. The accuracy is equal to (TP+TN)/(TP+TN+FP+FN).

Table 5.45 shows the accuracy thresholds of the Distress Score of the annual rating compared with the audit data. The accuracy is determined using the population of Distress Score differences and calculating the points that hold 95% of the data or within two standard deviations from the mean. This data range calculation is a frequent practice among highway agencies to check the range of Distress Score differences between the annual data and audit data. The accuracy of Distress Scores is determined for each pavement type on a specific highway service level.

In Table 5.45, for JCP IH and FM levels, the number of sections available for the accuracy threshold analysis is low. We can suggest that the DS threshold used to calculate the accuracy threshold of the Distress Score needs further evaluation due to an insufficient amount of data.

Table 5.45 Accuracy thresholds of the annual rating compared with the audit data for the Distress Score

Pavement Type	**Highway service level**	**DS Range**	**No. of Sections**	**Accuracy thresholds of Distress Score**

ACP	**IH**	(-18.77, 22.6)	Sections % FN 49 0.016465 T- 337 0.113239 0 1890 0.635081 T+ 563 0.189180 FP 137 0.046035	93.75%
	SH	(-18.9, 21.9)	Sections % FN 262 0.023747 T- 1370 0.124173 0 6893 0.624762 T+ 1984 0.179824 FP 524 0.047494	92.88%
	US	(-19.02, 22.13)	Sections % FN 177 0.018261 T- 1211 0.124936 0 6161 0.635613 T+ 1722 0.177654 FP 422 0.043537	93.82%
	FM	(-18.44, 19.28)	Sections % FN 699 0.025165 T- 4286 0.154300 0 18130 0.652698 T+ 3637 0.130936 FP 1025 0.036901	93.79%
CRCP	**IH**	(-26.47,26.71)	Sections % FN 58 0.036227 T- 189 0.118051 0 1103 0.688944 T+ 186 0.116177 FP 65 0.040600	92.32%
	SH	(-25.86, 24.92)	Sections % FN 33 0.043137 T- 84 0.109804 0 533 0.696732 T+ 93 0.121569 FP 22 0.028758	92.81%
	US	(-24.63, 24.6)	Sections % FN 25 0.037425 T- 69 0.103293 0 468 0.700599 T+ 83 0.124251 FP 23 0.034431	92.81%
	FM	(-32.77, 32.56)	Sections % FN 10 0.042194 T- 30 0.126582 0 158 0.666667	91.98%

			T+ 30 0.126582 FP 9 0.037975	
JCP	**IH**	(-59.59, 40.25)	Sections % FN 10 0.066667 T- 47 0.313333 0 74 0.493333 T+ 18 0.120000 FP 1 0.006667	92.67%
	SH	(-52.64, 41.03)	Sections % FN 8 0.038278 T- 78 0.373206 0 86 0.411483 T+ 31 0.148325 FP 6 0.028708	93.3%
	US	(-46.23, 34.29)	Sections % FN 14 0.051471 T- 82 0.301471 0 119 0.437500 T+ 55 0.202206 FP 2 0.007353	94.12%
	FM	(-41.31, 30.31)	Sections % FN 7 0.071429 T- 22 0.224490 0 55 0.561224 T+ 13 0.132653 FP 1 0.010204	91.84%

Based on the two standard deviation thresholds of each distress type on a specific highway service level determined in Table 5.43, Table 5.44 and Table 5.45, the accuracy in measurement value differences of each distress type is presented in Table 5.46, Table 5.47 and Table 5.48.

Table 5.46 shows the accuracy thresholds for each distress type of ACP. The accuracy will be calculated by comparing the measurement values of the annual rating data and audit data. For the distress type of ACP Failure at the SH level, the difference in distress measurement value of annual data compared with the audit data shows a discrepancy because most sections have no change and are within the threshold range as the difference in measurement value is zero. Some sections are out of the threshold range. The threshold must be evaluated further with more data.

Table 5.46 Accuracy thresholds of the annual rating compared with the audit data for the Distress Score

Distress Type	**Highway service level**	**No. of Sections**	**Accuracy**
ACP Alligator Cracking (Percentage of Wheel path Length)	**IH**	Sections % FN 74 0.024866 T- 359 0.120632 0 2134 0.717070	96.47%

		T+ 378 0.127016 FP 31 0.010417	
	US	Sections % FN 248 0.025585 T- 1087 0.112143 0 6800 0.701537 T+ 1468 0.151449 FP 90 0.009285	96.51%
	SH	Sections % FN 315 0.028553 T- 1274 0.115482 0 7639 0.692440 T+ 1702 0.154278 FP 102 0.009246	96.22%
	FM	Sections % FN 615 0.022141 T- 2572 0.092598 0 18971 0.683000 T+ 5257 0.189264 FP 361 0.012997	96.49%
ACP Longitudinal Cracking (Length in Feet per station).	**IH**	Sections % FN 81 0.027227 T- 820 0.275630 0 885 0.297479 T+ 1134 0.381176 FP 55 0.018487	95.43%
	US	Sections % FN 364 0.037553 T- 2048 0.211286 0 3241 0.334365 T+ 3777 0.389663 FP 263 0.027133	93.53%
	SH	Sections % FN 369 0.033451 T- 2256 0.204515 0 3465 0.314115 T+ 4649 0.421449 FP 292 0.026471	94.01%
	FM	Sections % FN 670 0.024122 T- 4557 0.164062 0 7741 0.278694 T+ 13939 0.501836 FP 869 0.031286	94.46%
ACP Transverse Cracking (Quantity)	**IH**	Sections % FN 69 0.023185	96.71%

		T- 318 0.106855 0 2350 0.789651 T+ 210 0.070565 FP 29 0.009745	
	US	Sections % FN 242 0.024966 T- 1636 0.168782 0 6956 0.717631 T+ 752 0.077582 FP 107 0.011039	96.4%
	SH	Sections % FN 197 0.017857 T- 1682 0.152466 0 8143 0.738125 T+ 927 0.084028 FP 83 0.007524	97.46%
	FM	Sections % FN 589 0.021205 T- 1950 0.070204 0 23248 0.836982 T+ 1695 0.061024 FP 294 0.010585	96.82%
ACP Patching (Percentage of Lane Area)	**IH**	Sections % FN 45 0.015121 T- 93 0.031250 0 2693 0.904906 T+ 97 0.032594 FP 48 0.016129	96.88%
	US	Sections % FN 104 0.010729 T- 381 0.039307 0 8798 0.907665 T+ 312 0.032188 FP 98 0.010110	97.92%
	SH	Sections % FN 180 0.016316 T- 570 0.051668 0 9786 0.887056 T+ 384 0.034808 FP 112 0.010152	97.35%
	FM	Sections % FN 448 0.016129 T- 1564 0.056308 0 24188 0.870824 T+ 1165 0.041943 FP 411 0.014797	96.91%

ACP Block Cracking (Percentage of Lane Area)	**IH**	Sections % FN 10 0.003360 T- 6 0.002016 0 2931 0.984879 T+ 17 0.005712 FP 12 0.004032	99.26%
	US	Sections % FN 31 0.003198 T- 16 0.001651 0 9481 0.978129 T+ 86 0.008872 FP 79 0.008150	98.87%
	SH	Sections % FN 36 0.003263 T- 31 0.002810 0 10766 0.975888 T+ 131 0.011875 FP 68 0.006164	99.06%
	FM	Sections % FN 52 0.001872 T- 21 0.000756 0 27446 0.988119 T+ 137 0.004932 FP 120 0.004320	99.38%
ACP Failure (Quantity)	**IH**	Sections % FN 42 0.014113 T- 74 0.024866 0 2826 0.949597 T+ 26 0.008737 FP 8 0.002688	98.32%
	US	Sections % FN 45 0.004643 T- 102 0.010523 0 9464 0.976375 T+ 47 0.004849 FP 35 0.003611	99.17%
	SH	Sections % FN 216 0.019579 T- 0 0.000000 0 10709 0.970722 T+ 0 0.000000 FP 107 0.009699	97.07%
	FM	Sections % FN 232 0.008353 T- 626 0.022537 0 26386 0.949957	98.62%

		T+ 381 0.013717 FP 151 0.005436	

Table 5.47 shows the accuracy thresholds of each distress type of CRCP. The accuracy will be calculated by comparing the measurement values of the annual rating data and audit data.
For the distress type of CRCP Punchout at the FM level, the quantity threshold value is a decimal number. The difference in distress measurement value of annual data compared with the audit data shows a discrepancy because most sections have no change and are within the threshold range. Some sections are out of the threshold range. The threshold must be evaluated further with more data.

Table 5.47 Accuracy thresholds of the annual rating compared with the audit data for CRCP

Distress Type	**Highway service level**	**No. of Sections**	**Accuracy**
CRCP Spalled Cracks (Quantity)	**IH**	Sections % FN 32 0.019988 T- 122 0.076202 0 1293 0.807620 T+ 136 0.084947 FP 18 0.011243	96.88%
	US	Sections % FN 6 0.008982 T- 63 0.094311 0 539 0.806886 T+ 57 0.085329 FP 3 0.004491	98.65%
	SH	Sections % FN 13 0.016993 T- 76 0.099346 0 618 0.807843 T+ 58 0.075817 FP 0 0.000000	98.3%
	FM	Sections % FN 3 0.012658 T- 20 0.084388 0 194 0.818565 T+ 20 0.084388 FP 0 0.000000	98.73%
CRCP Punchout (Quantity)	**IH**	Sections % FN 30 0.018738 T- 105 0.065584 0 1399 0.873829 T+ 51 0.031855 FP 16 0.009994	97.13%

	US	Sections % FN 10 0.014970 T- 32 0.047904 0 597 0.893713 T+ 20 0.029940 FP 9 0.013473	97.16%
	SH	Sections % FN 18 0.023529 T- 38 0.049673 0 665 0.869281 T+ 28 0.036601 FP 16 0.020915	95.56%
	FM	Sections % FN 11 0.046414 T- 0 0.000000 0 219 0.924051 T+ 0 0.000000 FP 7 0.029536	92.41%
CRCP ACP Patches (Quantity)	**IH**	Sections % FN 8 0.004997 T- 43 0.026858 0 1500 0.936914 T+ 29 0.018114 FP 21 0.013117	98.19%
	US	Sections % FN 3 0.004491 T- 16 0.023952 0 617 0.923653 T+ 19 0.028443 FP 13 0.019461	97.6%
	SH	Sections % FN 0 0.000000 T- 28 0.036601 0 695 0.908497 T+ 21 0.027451 FP 21 0.027451	97.25%
	FM	Sections % FN 2 0.008439 T- 10 0.042194 0 221 0.932489 T+ 2 0.008439 FP 2 0.008439	98.31%
CRCP PCC Patches (Quantity)	**IH**	Sections % FN 38 0.023735 T- 115 0.071830 0 1251 0.781387	94.44%

		T+	146	0.091193	
		FP	51	0.031855	
	US		Sections	%	96.86%
		FN	11	0.016467	
		T-	73	0.109281	
		0	521	0.779940	
		T+	53	0.079341	
		FP	10	0.014970	
	SH		Sections	%	98.43%
		FN	8	0.010458	
		T-	58	0.075817	
		0	636	0.831373	
		T+	59	0.077124	
		FP	4	0.005229	
	FM		Sections	%	95.78%
		FN	2	0.008439	
		T-	27	0.113924	
		0	171	0.721519	
		T+	29	0.122363	
		FP	8	0.033755	
CRCP Average Crack Spacing (Average observed spacing in feet)	**IH**		Sections	%	91.07%
		FN	26	0.016240	
		T-	140	0.087445	
		0	103	0.064335	
		T+	1215	0.758901	
		FP	117	0.073079	
	US		Sections	%	91.17%
		FN	8	0.011976	
		T-	76	0.113772	
		0	49	0.073353	
		T+	484	0.724551	
		FP	51	0.076347	
	SH		Sections	%	90.85%
		FN	10	0.013072	
		T-	90	0.117647	
		0	39	0.050980	
		T+	566	0.739869	
		FP	60	0.078431	
	FM		Sections	%	98.73%
		FN	3	0.012658	
		T-	32	0.135021	
		0	11	0.046414	
		T+	191	0.805907	
		FP	0	0.000000	

Table 5.48 shows the accuracy thresholds of each distress type of JCP. The accuracy will be

calculated by comparing the measurement values of the annual rating data and audit data.
For JCP Shattered Slabs and JCP Slabs with Longitudinal Cracks in all highway service levels, the quantity threshold value is a decimal number. The difference in distress measurement value of annual data compared with the audit data shows a discrepancy because most sections have no change and are within the threshold range. Some sections are out of the threshold range. The thresholds must be evaluated further with more data.

Table 5.48 Accuracy thresholds of the annual rating compared with the audit data for JCP

Distress Type	Highway service level	No. of Sections			Accuracy
JCP Failed Joint Cracks (Quantity)	**IH**		Sections	%	96.67%
		FN	0	0.000000	
		T-	18	0.120000	
		0	80	0.533333	
		T+	47	0.313333	
		FP	5	0.033333	
	US		Sections	%	95.96%
		FN	0	0.000000	
		T-	52	0.191176	
		0	130	0.477941	
		T+	79	0.290441	
		FP	11	0.040441	
	SH		Sections	%	93.78%
		FN	5	0.023923	
		T-	44	0.210526	
		0	88	0.421053	
		T+	64	0.306220	
		FP	8	0.038278	
	FM		Sections	%	91.84%
		FN	4	0.040816	
		T-	18	0.183673	
		0	57	0.581633	
		T+	15	0.153061	
		FP	4	0.040816	
JCP Failures (Quantity)	**IH**		Sections	%	98.67%
		FN	0	0.000000	
		T-	18	0.120000	
		0	72	0.480000	
		T+	58	0.386667	
		FP	2	0.013333	
	US		Sections	%	98.53%
		FN	1	0.003676	
		T-	45	0.165441	
		0	125	0.459559	
		T+	98	0.360294	
		FP	3	0.011029	

	SH	Sections % FN 5 0.023923 T- 35 0.167464 0 77 0.368421 T+ 86 0.411483 FP 6 0.028708	94.74%
	FM	Sections % FN 1 0.010204 T- 17 0.173469 0 47 0.479592 T+ 28 0.285714 FP 5 0.051020	93.88%
JCP PCC Patches (Quantity)	**IH**	Sections % FN 3 0.020000 T- 28 0.186667 0 85 0.566667 T+ 27 0.180000 FP 7 0.046667	93.3%
	US	Sections % FN 4 0.014706 T- 68 0.250000 0 156 0.573529 T+ 39 0.143382 FP 5 0.018382	96.69%
	SH	Sections % FN 4 0.019139 T- 25 0.119617 0 151 0.722488 T+ 22 0.105263 FP 7 0.033493	94.74%
	FM	Sections % FN 0 0.000000 T- 9 0.091837 0 65 0.663265 T+ 21 0.214286 FP 3 0.030612	96.94%
JCP Shattered Slabs (Quantity)	**IH**	Sections % FN 0 0.000000 T- 0 0.000000 0 140 0.933333 T+ 7 0.046667 FP 3 0.020000	98%
	US	Sections % FN 0 0.000000 T- 0 0.000000 0 266 0.977941	98.9%

		T+ 3 0.011029 FP 3 0.011029	
	SH	Sections % FN 0 0.000000 T- 2 0.009569 0 189 0.904306 T+ 13 0.062201 FP 5 0.023923	97.61%
	FM	Sections % FN 1 0.010204 T- 1 0.010204 0 93 0.948980 T+ 2 0.020408 FP 1 0.010204	97.96%
JCP Slabs with Longitudinal Cracks (Quantity)	**IH**	Sections % FN 0 0.000000 T- 0 0.000000 0 140 0.933333 T+ 7 0.046667 FP 3 0.020000	98%
	US	Sections % FN 0 0.000000 T- 0 0.000000 0 266 0.977941 T+ 3 0.011029 FP 3 0.011029	98.9%
	SH	Sections % FN 0 0.000000 T- 2 0.009569 0 189 0.904306 T+ 13 0.062201 FP 5 0.023923	97.61%
	FM	Sections % FN 1 0.010204 T- 1 0.010204 0 93 0.948980 T+ 2 0.020408 FP 1 0.010204	97.96%
JCP Apparent Joint Space	**IH**	Sections % FN 0 0.000000 T- 17 0.113333 0 115 0.766667 T+ 13 0.086667 FP 5 0.033333	96.67%
	US	Sections % FN 7 0.025735	94.85%

		T- 36 0.132353 0 180 0.661765 T+ 42 0.154412 FP 7 0.025735	
	SH	Sections % FN 0 0.000000 T- 44 0.210526 0 112 0.535885 T+ 36 0.172249 FP 17 0.081340	91.87%
	FM	Sections % FN 4 0.040816 T- 9 0.091837 0 68 0.693878 T+ 15 0.153061 FP 2 0.020408	93.88%

Index #4: Precision of the annual rating data

The precision of each distress type measurement value is determined by comparing two years' annual data in the same section. Precision is equal to TP/(TP+FP). The dataset used for precision analysis of the annual rating is the paired PMIS annual rating data of the same section from FY2017 to FY2021.

Table 5.49 shows the precision threshold for ACP. The precision is calculated by comparing the measurement values of two years' annual rating data in the same section. The threshold format is the difference of each distress measurement value in two years.

For ACP Block Cracking distress type at the IH level, there are no fixed threshold range for precision as there is no data available for precision analysis (marked in Table 5.49). Furthermore, for ACP Block Cracking at US level, there is not enough data available for analysis, so we cannot say that the threshold range used to calculate the precision for ACP Block Cracking at US level is correct. Therefore, the threshold range must be further evaluated with sufficient data. For ACP Failure at IH and US highway service levels, the quantity threshold value is a decimal number. The difference in distress measurement value of annual data compared with the audit data shows a discrepancy because most sections have no change and are within the threshold range as the difference in measurement value is zero. Some sections are out of the threshold range. Therefore, the precision calculated for these levels may not be usable for the analysis framework of PMIS data. The thresholds must be evaluated further with more data.

Table 5.49 Precision threshold for individual distress type for ACP

Distress Type	Highway service level	Threshold range	No. of Sections	Threshold
ACP Alligator Cracking (Percentage of	**IH**	(-5.01, 5.53)	Sections % FN 5 0.010989 T- 29 0.063736	97.11%

Wheel path Length)			0 369 0.810989 T+ 39 0.085714 FP 13 0.028571	
	US	(-8.84, 7.23)	Sections % FN 24 0.053812 T- 68 0.152466 0 308 0.690583 T+ 41 0.091928 FP 5 0.011211	98.82%
	SH	(-11.0, 10.64)	Sections % FN 179 0.020672 T- 1798 0.207645 0 5342 0.616930 T+ 1184 0.136736 FP 156 0.018016	98.16%
	FM	(-8.97, 8.45)	Sections % FN 2879 0.027966 T- 22619 0.219715 0 60277 0.585515 T+ 15018 0.145881 FP 2154 0.020923	97.85%
ACP Longitudinal Cracking (Length in Feet per station).	**IH**	(-30.58, 43.64)	Sections % FN 8 0.017582 T- 85 0.186813 0 121 0.265934 T+ 222 0.487912 FP 19 0.041758	95.75%
	US	(-47.54, 46.3)	Sections % FN 11 0.024664 T- 208 0.466368 0 43 0.096413 T+ 169 0.378924 FP 15 0.033632	96.55%
	SH	(-43.9, 42.87)	Sections % FN 275 0.031759 T- 3582 0.413674 0 1322 0.152674 T+ 3239 0.374062 FP 241 0.027832	97.13%
	FM	(-41.3, 39.05)	Sections % FN 3103 0.030142 T- 43460 0.422159 0 18665 0.181307 T+ 34888 0.338893 FP 2831 0.027500	97.16%
ACP	**IH**	(-2.12, 2.01)	Sections %	98.39%

Transverse Cracking (Quantity)			FN 21 0.046154 T- 28 0.061538 0 346 0.760440 T+ 53 0.116484 FP 7 0.015385	
	US	(-1.69, 1.5)	Sections % FN 25 0.056054 T- 58 0.130045 0 306 0.686099 T+ 45 0.100897 FP 12 0.026906	97.15%
	SH	(-2.9, 2.67)	Sections % FN 300 0.034646 T- 1081 0.124841 0 6315 0.729299 T+ 813 0.093891 FP 150 0.017323	98.21%
	FM	(-1.84, 1.64)	Sections % FN 4211 0.040905 T- 6360 0.061779 0 86677 0.841958 T+ 4058 0.039418 FP 1641 0.015940	98.34%
ACP Patching (Percentage of Lane Area)	**IH**	(-16.25, 20.67)	Sections % FN 0 0.000000 T- 17 0.037363 0 389 0.854945 T+ 27 0.059341 FP 22 0.048352	95.16%
	US	(-10.52, 10.38)	Sections % FN 12 0.026906 T- 34 0.076233 0 359 0.804933 T+ 29 0.065022 FP 12 0.026906	97.24%
	SH	(-11.55, 11.53)	Sections % FN 151 0.017439 T- 449 0.051854 0 7401 0.854718 T+ 513 0.059245 FP 145 0.016746	98.3%
	FM	(-12.3, 12.33)	Sections % FN 1746 0.016960 T- 5442 0.052862 0 87045 0.845532 T+ 6754 0.065607	98.06%

			FP 1960 0.019039	
ACP Block Cracking (Percentage of Lane Area)	**IH**	No data		
	US	(0.0, 0.0)	Sections % FN 0 0.0 0 42 1.0 FP 0 0.0	100%
	SH	(-7.45, 9.09)	Sections % FN 4 0.012821 T- 6 0.019231 0 257 0.823718 T+ 30 0.096154 FP 15 0.048077	95.13%
	FM	(-6.25, 6.99)	Sections % FN 7 0.002205 T- 14 0.004411 0 2996 0.943919 T+ 88 0.027725 FP 69 0.021739	97.82%
ACP Failure (Quantity)	**IH**	(-0.33, 0.33)	Sections % FN 5 0.010989 T- 0 0.000000 0 446 0.980220 T+ 0 0.000000 FP 4 0.008791	99.11%
	US	(-0.82, 0.81)	Sections % FN 2 0.004484 T- 0 0.000000 0 438 0.982063 T+ 0 0.000000 FP 6 0.013453	98.65%
	SH	(-1.38, 1.4)	Sections % FN 57 0.006583 T- 101 0.011664 0 8304 0.959002 T+ 129 0.014898 FP 68 0.007853	99.21%
	FM	(-1.43, 1.45)	Sections % FN 1002 0.009733 T- 1886 0.018320 0 96593 0.938279 T+ 2214 0.021506 FP 1252 0.012162	98.77%

Table 5.50 shows the precision threshold (Change in measurement values) for individual distress types of CRCP.

For each CRCP distress type at IH and US levels, the number of sections used for the precision threshold analysis is exceptionally low. Hence, the threshold ranges need further evaluation because the threshold might be incorrect due to the low available data.

Table 5.50 Precision threshold for individual distress type for CRCP

Distress Type	Highway service level	Threshold range	No. of Sections	Threshold
CRCP Spalled Cracks (Quantity)	**IH**	(-3.8, 3.53)	Sections % FN 1 0.083333 T- 2 0.166667 0 7 0.583333 T+ 1 0.083333 FP 1 0.083333	90.91%
	US	(-7.2, 4.76)	Sections % FN 1 0.111111 T- 1 0.111111 0 7 0.777778 T+ 0 0.000000 FP 0 0.000000	100%
	SH	(-5.23, 5.46)	Sections % FN 5 0.002600 T- 175 0.091004 0 1507 0.783671 T+ 231 0.120125 FP 5 0.002600	99.74%
	FM	(-2.18, 2.3)	Sections % FN 7 0.010029 T- 38 0.054441 0 589 0.843840 T+ 49 0.070201 FP 15 0.021490	97.83%
CRCP Punchout (Quantity)	**IH**	(-2.41, 2.25)	Sections % FN 1 0.083333 T- 1 0.083333 0 8 0.666667 T+ 2 0.166667 FP 0 0.000000	100%
	US	(-8.56, 6.11)	Sections % FN 1 0.111111 T- 0 0.000000 0 8 0.888889 T+ 0 0.000000 FP 0 0.000000	100%
	SH	(-1.27, 1.31)	Sections % FN 28 0.014561	98.36%

			T- 78 0.040562 0 1690 0.878835 T+ 96 0.049922 FP 31 0.016121	
	FM	(-0.82, 0.85)	Sections % FN 20 0.028653 T- 0 0.000000 0 648 0.928367 T+ 0 0.000000 FP 30 0.042980	95.58%
CRCP ACP Patches (Quantity)	**IH**	(-0.85, 0.85)	Sections % FN 1 0.083333 T- 0 0.000000 0 10 0.833333 T+ 0 0.000000 FP 1 0.083333	90.91%
	US	(-0.78, 0.56)	Sections % FN 1 0.111111 T- 0 0.000000 0 8 0.888889 T+ 0 0.000000 FP 0 0.000000	100%
	SH	(-2.29, 2.22)	Sections % FN 21 0.010920 T- 56 0.029121 0 1776 0.923557 T+ 57 0.029641 FP 13 0.006760	99.32%
	FM	(-9.42, 9.85)	Sections % FN 3 0.004298 T- 7 0.010029 0 666 0.954155 T+ 16 0.022923 FP 6 0.008596	99.14%
CRCP PCC Patches (Quantity)	**IH**	(-14.28, 19.12)	Sections % FN 0 0.000000 T- 2 0.166667 0 8 0.666667 T+ 1 0.083333 FP 1 0.083333	91.67%
	US	(-0.56, 0.78)	Sections % FN 0 0.000000 T- 0 0.000000 0 8 0.888889 T+ 0 0.000000 FP 1 0.111111	88.89%

	SH	(-5.95, 6.53)	Sections % FN 36 0.018721 T- 134 0.069683 0 1503 0.781591 T+ 196 0.101924 FP 54 0.028081	97.14%
	FM	(-6.49, 6.66)	Sections % FN 15 0.021490 T- 62 0.088825 0 539 0.772206 T+ 66 0.094556 FP 16 0.022923	97.66%
CRCP Average Crack Spacing (Average observed spacing in feet)	**IH**	(-27.13, 22.96)	Sections % FN 1 0.083333 T- 3 0.250000 0 2 0.166667 T+ 6 0.500000 FP 0 0.000000	100%
	US	(-41.74, 40.18)	Sections % FN 0 0.000000 T- 4 0.444444 0 2 0.222222 T+ 3 0.333333 FP 0 0.000000	100%
	SH	(-36.17, 46.83)	Sections % FN 52 0.027041 T- 542 0.281851 0 297 0.154446 T+ 915 0.475819 FP 117 0.060842	93.75%
	FM	(-38.88, 57.33)	Sections % FN 19 0.027221 T- 170 0.243553 0 101 0.144699 T+ 376 0.538682 FP 32 0.045845	95.29%

Table 5.51 shows the precision threshold (Change in measurement values) for individual distress types of JCP. For each JCP distress type at IH, US, and FM levels, the number of sections used for the precision threshold analysis is exceptionally low. Hence, the threshold ranges need further evaluation because the threshold may provide incorrect data.

For each JCP distress type at the IH level, there is no data available for analysis after merging. The threshold cannot be calculated for JCP at the IH level.

Table 5.51 Precision threshold for individual distress type for JCP

Distress Type	Highway service level	Threshold range	No. of Sections			Threshold
JCP Failed Joint Cracks (Quantity)	**IH**	No data				
	US	(-4.64, 5.95)		Sections	%	90.63%
			FN	0	0.00000	
			T-	9	0.28125	
			0	10	0.31250	
			T+	10	0.31250	
			FP	3	0.09375	
	SH	(-8.03, 8.38)		Sections	%	97.74%
			FN	10	0.024510	
			T-	84	0.205882	
			0	170	0.416667	
			T+	135	0.330882	
			FP	9	0.022059	
	FM	(-2.87, 3.25)		Sections	%	98.06%
			FN	4	0.037383	
			T-	20	0.186916	
			0	47	0.439252	
			T+	34	0.317757	
			FP	2	0.018692	
JCP Failures (Quantity)	**IH**	No data				
	US	(-18.62, 22.25)		Sections	%	96.88%
			FN	0	0.00000	
			T-	9	0.28125	
			0	13	0.40625	
			T+	9	0.28125	
			FP	1	0.03125	
	SH	(-13.29, 13.38)		Sections	%	98.25%
			FN	9	0.022059	
			T-	98	0.240196	
			0	173	0.424020	
			T+	121	0.296569	
			FP	7	0.017157	
	FM	(-7.45, 7.92)		Sections	%	94.17%
			FN	4	0.037383	
			T-	32	0.299065	
			0	38	0.355140	
			T+	27	0.252336	
			FP	6	0.056075	
JCP PCC Patches (Quantity)	**IH**	No data				
	US	(-36.42, 30.86)		Sections	%	96.67%
			FN	2	0.06250	
			T-	7	0.21875	

			0 14 0.43750 T+ 8 0.25000 FP 1 0.03125	
	SH	(-16.93, 16.68)	Sections % FN 10 0.024510 T- 58 0.142157 0 278 0.681373 T+ 56 0.137255 FP 6 0.014706	98.49%
	FM	(-6.79, 4.94)	Sections % FN 7 0.065421 T- 23 0.214953 0 64 0.598131 T+ 10 0.093458 FP 3 0.028037	97%
JCP Shattered Slabs (Quantity)	**IH**	No data		
	US	(-0.77, 0.83)	Sections % FN 2 0.06250 T- 0 0.00000 0 27 0.84375 T+ 0 0.00000 FP 3 0.09375	90%
	SH	(-2.85, 2.74)	Sections % FN 10 0.024510 T- 13 0.031863 0 377 0.924020 T+ 2 0.004902 FP 6 0.014706	98.49%
	FM	(-0.5, 0.52)	Sections % FN 2 0.018692 T- 0 0.000000 0 103 0.962617 T+ 0 0.000000 FP 2 0.018692	98.1%
JCP Slabs with Longitudinal Cracks (Quantity)	**IH**	No data		
	US	(-5.72, 7.66)	Sections % FN 0 0.00000 T- 3 0.09375 0 21 0.65625 T+ 7 0.21875 FP 1 0.03125	96.88%
	SH	(-7.0, 6.09)	Sections % FN 15 0.036765 T- 65 0.159314 0 264 0.647059 T+ 58 0.142157	98.47%

			FP 6 0.014706	
	FM	(-4.53, 3.92)	Sections % FN 4 0.037383 T- 14 0.130841 0 78 0.728972 T+ 10 0.093458 FP 1 0.009346	99.03%
JCP Apparent Joint Space	**IH**	No data		
	US	(-2.71, 2.31)	Sections % FN 1 0.03125 T- 15 0.46875 0 8 0.25000 T+ 6 0.18750 FP 2 0.06250	93.55%
	SH	(-6.85, 7.17)	Sections % FN 10 0.024510 T- 101 0.247549 0 150 0.367647 T+ 136 0.333333 FP 11 0.026961	97.24%
	FM	(-10.08, 11.25)	Sections % FN 2 0.018692 T- 40 0.373832 0 23 0.214953 T+ 33 0.308411 FP 9 0.084112	91.43%

Index #5 Yearly change of each section DS, CS, and RS

The threshold of yearly change of DS, CS, and RS in the same section is presented. Threshold values are determined by calculating the DS, CS, and RS differences of the same section compared with the following year determining the points that hold 95% of the data or within two standard deviations of the mean.

Table 5.52 shows the thresholds of Distress Score for yearly change of each section based on the two determining factors (pavement type and highway service level). The threshold is in the format of two years' DS differences. A total of 12 thresholds are determined in this table. In Table 81, for CRCP pavement type in IH and US levels, the number of sections used for the DS threshold range analysis is extremely low. Hence, the threshold ranges for DS need further evaluation because the threshold might be incorrect. For JCP pavement type at US and FM levels, the number of sections used for DS range analysis is extremely low. Hence, the thresholds may also be flagged for further evaluation with sufficient data. For the JCP pavement type at the IH level, there is no data available for analysis after merging. Hence, the threshold cannot be calculated for JCP at the IH level.

Table 5.52 Thresholds of DS yearly change of section

Pavement Type	Highway service level	DS range	No. of Sections	Threshold
ACP	**IH**	(-18.56, 13.55)	Sections % FN 28 0.061538 T- 175 0.384615 0 162 0.356044 T+ 82 0.180220 FP 8 0.017582	98.13%
	SH	(-21.5, 21.55)	Sections % FN 258 0.029796 T- 2664 0.307657 0 2969 0.342880 T+ 2475 0.285830 FP 293 0.033838	96.51%
	US	(-19.02, 18.8)	Sections % FN 13 0.029148 T- 170 0.381166 0 103 0.230942 T+ 145 0.325112 FP 15 0.033632	96.54%
	FM	(-22.3, 23.62)	Sections % FN 2964 0.028792 T- 30767 0.298863 0 28348 0.275365 T+ 37159 0.360953 FP 3709 0.036028	96.29%
CRCP	**IH**	(-61.18, 53.34)	Sections % FN 1 0.083333 T- 2 0.166667 0 6 0.500000 T+ 3 0.250000 FP 0 0.000000	100%
	SH	(-28.17, 26.3)	Sections % FN 90 0.046802 T- 258 0.134165 0 1308 0.680187 T+ 198 0.102964 FP 69 0.035881	96.24%
	US	(-46.63, 63.74)	Sections % FN 0 0.000000 T- 1 0.111111 0 7 0.777778 T+ 0 0.000000 FP 1 0.111111	88.89%

	FM	(-32.55, 30.72)	Sections % FN 31 0.044413 T- 90 0.128940 0 480 0.687679 T+ 75 0.107450 FP 22 0.031519	96.7%
JCP	**IH**	No data		
	SH	(-42.67, 40.8)	Sections % FN 20 0.049020 T- 96 0.235294 0 189 0.463235 T+ 86 0.210784 FP 17 0.041667	95.62%
	US	(-19.51, 15.64)	Sections % FN 2 0.06250 T- 5 0.15625 0 18 0.56250 T+ 7 0.21875 FP 0 0.00000	100%
	FM	(-45.77, 45.96)	Sections % FN 4 0.037383 T- 28 0.261682 0 38 0.355140 T+ 30 0.280374 FP 7 0.065421	93.2%

In Table 5.53, for CRCP pavement type at IH and US levels, the number of sections used for the CS threshold analysis is extremely low. Hence, the threshold ranges for CS need further evaluation because we cannot say the threshold can be used for future analysis of PMIS annual rating data of each Fiscal Year. The threshold may provide incorrect data.

For JCP pavement type at US and FM levels, the number of sections used for CS threshold analysis is extremely low. Hence, the thresholds may also be flagged for further evaluation with sufficient data. For the JCP pavement type at the IH level, there is no data available for analysis after merging. Hence, the threshold cannot be calculated for JCP at the IH level.

Table 5.53 Thresholds of CS yearly change of section

Pavement Type	**Highway service level**	**CS range**	**No. of Sections**	**Threshold**
ACP	**IH**	(-19.93, 14.28)	Sections % FN 29 0.063736 T- 180 0.395604 0 157 0.345055 T+ 84 0.184615 FP 5 0.010989	98.83%

	SH	(-27.25, 26.4)	Sections % FN 275 0.031906 T- 2820 0.327184 0 2769 0.321267 T+ 2512 0.291449 FP 243 0.028194	97.09%
	US	(-20.39, 19.93)	Sections % FN 13 0.029148 T- 175 0.392377 0 101 0.226457 T+ 140 0.313901 FP 17 0.038117	96.07%
	FM	(-26.3, 26.86)	Sections % FN 3058 0.029792 T- 32276 0.314443 0 26841 0.261493 T+ 37168 0.362102 FP 3302 0.032169	96.68%
CRCP	**IH**	(-61.97, 54.47)	Sections % FN 1 0.083333 T- 4 0.333333 0 1 0.083333 T+ 6 0.500000 FP 0 0.000000	100%
	SH	(-28.94, 27.46)	Sections % FN 92 0.047967 T- 348 0.181439 0 1048 0.546403 T+ 364 0.189781 FP 66 0.034411	96.39%
	US	(-51.27, 67.49)	Sections % FN 0 0.000000 T- 3 0.333333 0 3 0.333333 T+ 2 0.222222 FP 1 0.111111	88.89%
	FM	(-33.16, 32.26)	Sections % FN 31 0.044928 T- 163 0.236232 0 307 0.444928 T+ 167 0.242029 FP 22 0.031884	96.66%
JCP	**IH**	No data		
	SH	(-38.29, 36.11)	Sections % FN 17 0.041975 T- 135 0.333333	97.68%

			0 90 0.222222 T+ 154 0.380247 FP 9 0.022222	
	US	(-35.25, 34.81)	Sections % FN 1 0.03125 T- 10 0.31250 0 11 0.34375 T+ 9 0.28125 FP 1 0.03125	96.77%
	FM	(-38.95, 39.53)	Sections % FN 4 0.037383 T- 43 0.401869 0 13 0.121495 T+ 42 0.392523 FP 5 0.046729	95.15%

Table 5.54 shows the thresholds of Ride Score for yearly change of each section based on the two determining factors. The threshold is in the format of two years' RS differences. A total of 12 thresholds are determined in this table.

In Table 5.54, for CRCP pavement type at IH and US levels, the number of sections used for the RS range analysis is extremely low. Hence, the threshold ranges for DS need further evaluation because the threshold might be incorrect. For JCP pavement type at US and FM levels, the number of sections used for RS threshold analysis is extremely low. Hence, the thresholds may also be flagged for further evaluation with sufficient data. For the JCP pavement type at the IH level, there is no data available for analysis after merging. Therefore, the threshold cannot be calculated for JCP at IH levels.

Table 5.54 Thresholds of RS yearly change of section

Pavement Type	**Highway service level**	**RS range**	**No. of Sections**	**Threshold**
ACP	**IH**	(-0.54, 0.38)	Sections % FN 15 0.032967 T- 303 0.665934 0 28 0.061538 T+ 101 0.221978 FP 8 0.017582	98.18%
	SH	(-0.67, 0.64)	Sections % FN 192 0.022276 T- 4705 0.545887 0 298 0.034575 T+ 3183 0.369300 FP 241 0.027961	97.14%
	US	(-0.47, 0.49)	Sections % FN 7 0.015695	96.36%

			T- 208 0.466368 0 21 0.047085 T+ 194 0.434978 FP 16 0.035874	
	FM	(-0.6, 0.58)	Sections % FN 1914 0.018645 T- 53243 0.518660 0 2603 0.025357 T+ 42397 0.413005 FP 2498 0.024334	97.52%
CRCP	**IH**	(-0.77, 0.66)	Sections % FN 1 0.083333 T- 5 0.416667 0 0 0.000000 T+ 6 0.500000 FP 0 0.000000	100%
	SH	(-0.55, 0.63)	Sections % FN 19 0.009886 T- 880 0.457856 0 57 0.029657 T+ 899 0.467742 FP 67 0.034860	96.48%
	US	(-0.92, 1.3)	Sections % FN 0 0.000000 T- 4 0.444444 0 0 0.000000 T+ 4 0.444444 FP 1 0.111111	88.89%
	FM	(-0.5, 0.56)	Sections % FN 20 0.028694 T- 276 0.395983 0 27 0.038737 T+ 357 0.512195 FP 17 0.024390	97.49%
JCP	**IH**	No data		
	SH	(-0.69, 0.64)	Sections % FN 11 0.027160 T- 177 0.437037 0 18 0.044444 T+ 194 0.479012 FP 5 0.012346	98.73%
	US	(-0.93, 0.9)	Sections % FN 1 0.03125 T- 17 0.53125 0 0 0.00000 T+ 14 0.43750	100%

			FP 0 0.00000	
	FM	(-0.53, 0.6)	Sections % FN 4 0.037383 T- 43 0.401869 0 7 0.065421 T+ 47 0.439252 FP 6 0.056075	94.17%

5.3 Threshold development method selection

The research team decided to apply a sigma-based method for threshold development. To compare with the thresholds of different sigma values, the research team compared the thresholds of 1 sigma, 1.25 sigma, 1.5 sigma, and 2 sigma-based methods. By using the auto and audit data, the research team developed index #2 threshold and accuracy. Table 5.55, Table 5.57, and Table 5.59 list the measurement value thresholds based on 1 sigma, 1.25 sigma, 1.5 sigma, and 2 sigma methods of different distress types in IH, US, SH, and FM highway network levels for three pavement types (ACP, CRCP, and JCP). Table 5.56, Table 5.58, and Table 5.60 list accuracies of each distress type based on 1 sigma, 1.25 sigma, 1.5 sigma, and 2 sigma methods in IH, US, SH, and FM highway network levels for three pavement types (ACP, CRCP, and JCP).

Table 5.55 Comparison of sigma-based methods for measurement threshold development in ACP

Distress Type	Network Level	Thresholds			
		1 Sigma	**1.25 Sigma**	**1.5 Sigma**	**2 Sigma**
ACP Alligator Cracking (Percentage of Wheel path Length)	**IH**	(-6.07,5.04)	(-7.46,6.42)	(-8.85,7.81)	(-11.62,10.59)
	US	(-8.26,6.83)	(-10.15,8.71)	(-12.04,10.6)	(-15.81,14.37)
	SH	(-7.61,6.14)	(-9.33,7.85)	(-11.05,9.57)	(-14.49,13.01)
	FM	(-4.97,4.74)	(-6.19,5.95)	(-7.4,7.17)	(-9.83,9.6)
ACP Longitudinal Cracking (Length in Feet per station).	**IH**	(-32.85,29.44)	(-40.64,37.23)	(-48.43,45.02)	(-64.0,60.59)
	US	(-25.41,24.8)	(-31.68,31.08)	(-37.96,37.35)	(-50.51,49.91)
	SH	(-22.4,24.18)	(-28.22,30.01)	(-34.05,35.83)	(-45.69,47.48)
	FM	(-14.96,22.57)	(-19.65,27.27)	(-24.34,31.96)	(-33.72,41.34)
ACP Transverse Cracking (Quantity)	**IH**	(-2.03,1.74)	(-2.5,2.21)	(-2.97,2.68)	(-3.91,3.62)
	US	(-2.17,1.57)	(-2.64,2.03)	(-3.11,2.5)	(-4.04,3.44)
	SH	(-2.56,2.13)	(-3.15,2.71)	(-3.73,3.3)	(-4.91,4.47)
	FM	(-1.23,1.08)	(-1.52,1.37)	(-1.81,1.66)	(-2.38,2.24)
ACP Patching (Percentage of Lane Area)	**IH**	(-3.61,3.76)	(-4.53,4.68)	(-5.45,5.6)	(-7.29,7.44)
	US	(-4.87,4.64)	(-6.05,5.82)	(-7.24,7.01)	(-9.62,9.39)
	SH	(-4.8,4.43)	(-5.95,5.58)	(-7.1,6.73)	(-9.41,9.04)
	FM	(-4.93,4.68)	(-6.13,5.89)	(7.33 ,7.09)	(-9.73,9.49)
ACP Block Cracking (Percentage of Lane Area)	**IH**	(-3.88,3.77)	(-4.84,4.73)	(-5.8,5.69)	(-7.71,7.6)
	US	(-2.5,2.67)	(-3.14,3.32)	(-3.79,3.96)	(-5.08,5.26)
	SH	(-3.39,3.32)	(-4.23,4.16)	(-5.07,5)	(-6.75,6.68)

	FM	(-1.5,1.52)	(-1.87,1.9)	(-2.25,2.27)	(-3,3.03)
ACP Failure (Quantity)	**IH**	(-0.71,0.62)	(-0.87,0.78)	(-1.04,0.95)	(-1.37,1.28)
	US	(-0.7,0.69)	(-0.87,0.86)	(-1.05,1.03)	(-1.39,1.38)
	SH	(-0.47,0.43)	(-0.58,0.54)	(-0.69,0.66)	(-0.92,0.88)
	FM	(-0.91,0.86)	(-1.14,1.08)	(-1.36,1.31)	(-1.8,1.75)

Table 5.56 Comparison of sigma-based methods for accuracy check in ACP

Distress Type	**Network Level**	**Accuracy**			
		1 Sigma	**1.25 Sigma**	**1.5 Sigma**	**2 Sigma**
ACP Alligator Cracking (Percentage of Wheel path Length)	**IH**	93.08%	94.15%	95.09%	96.47%
	US	93.31%	94.57%	95.30%	96.51%
	SH	92.17%	93.38%	94.76%	96.22%
	FM	91.66%	93.84%	95.24%	96.49%
ACP Longitudinal Cracking (Length in Feet per station).	**IH**	86.42%	89.88%	92.67%	95.43%
	US	84.41%	87.70%	90.04%	93.53%
	SH	84.87%	88.32%	90.71%	94.01%
	FM	85.67%	89.16%	91.35%	94.46%
ACP Transverse Cracking (Quantity)	**IH**	93.45%	94.93%	94.93%	96.71%
	US	89.95%	91.89%	94.10%	96.40%
	SH	92.63%	94.63%	95.60%	97.46%
	FM	93.22%	93.22%	93.22%	96.82%
ACP Patching (Percentage of Lane Area)	**IH**	95.23%	95.67%	96.07%	96.88%
	US	95.96%	96.87%	97.20%	97.92%
	SH	94.82%	95.54%	96.17%	97.35%
	FM	94.31%	95.33%	95.98%	96.91%
ACP Block Cracking (Percentage of Lane Area)	**IH**	98.86%	99.03%	99.09%	99.26%
	US	98.50%	98.57%	98.57%	98.87%
	SH	98.66%	98.77%	98.90%	99.06%
	FM	99.14%	99.14%	99.31%	99.38%
ACP Failure (Quantity)	**IH**	94.96%	94.96%	97.45%	98.32%
	US	97.64%	97.64%	99.17	99.17%
	SH	97.07%	97.07%	97.07%	97.07%
	FM	95%	98.62%	98.62%	98.62%

Table 5.57 Comparison of sigma-based methods for measurement threshold development in CRCP

Distress Type	**Network Level**	**Thresholds**			
		1 Sigma	**1.25 Sigma**	**1.5 Sigma**	**2 Sigma**
CRCP Spalled Cracks (Quantity)	**IH**	(-2.01,1.73)	(-2.48,2.2)	(-2.95,2.66)	(-3.88, 3.6)
	US	(-5.36,4.71)	(-6.62,5.97)	(-7.88,7.23)	(-10.4, 9.75)
	SH	(-11.42,8.86)	(-13.95,11.39)	(-16.49,13.92)	(-21.56, 18.99)
	FM	(-1.18,0.99)	(-1.45,1.26)	(-1.72,1.53)	(-2.26, 2.07)
CRCP Punchout	**IH**	(-1.16,0.93)	(-1.42,1.19)	(-1.68,1.46)	(-2.2, 1.98)

(Quantity)	**US**	(-0.65,0.59)	(-0.81,0.75)	(-0.96,0.9)	(-1.27, 1.21)
	SH	(-0.75,0.69)	(-0.93,0.87)	(-1.11,1.05)	(-1.47, 1.41)
	FM	(-0.37,0.31)	(-0.45,0.39)	(-0.53,0.48)	(-0.7, 0.64)
CRCP ACP Patches (Quantity)	**IH**	(-1.51,1.67)	(-1.91,2.07)	(-2.3,2.47)	(-3.1, 3.26)
	US	(-2.09,2.62)	(-2.68,3.2)	(-3.27,3.79)	(-4.45, 4.97)
	SH	(-5.05,6.95)	(-6.55,8.45)	(-8.05,9.95)	(-11.05, 12.95)
	FM	(-2.45,2.3)	(-3.05,2.9)	(-3.64,3.49)	(-4.83, 4.68)
CRCP PCC Patches (Quantity)	**IH**	(-2.4,2.75)	(-3.04,3.4)	(-3.68,4.04)	(-4.97,5.33)
	US	(-4.12,3.48)	(-5.07,4.43)	(-6.02,5.38)	(-7.91,7.27)
	SH	(-4.16,3.7)	(-5.15,4.68)	(-6.13,5.66)	(-8.1,7.63)
	FM	(-3.23,3.87)	(-4.12,4.76)	(-5.01,5.65)	(-6.78,7.42)
CRCP Average Crack Spacing (Average observed spacing in feet)	**IH**	(-9.07,35.38)	(-14.63,40.94)	(-20.19,46.5)	(-31.3, 57.61)
	US	(-9.02,31.32)	(-14.06,36.36)	(-19.1,41.4)	(-29.18, 51.48)
	SH	(-6.97,42.95)	(-13.21,49.19)	(-19.45,55.44)	(-31.94, 67.92)
	FM	(-2.93,56.84)	(-10.4,64.31)	(-17.87,71.78)	(-32.81, 86.72)

Table 5.58 Comparison of sigma-based methods for accuracy check in CRCP

Distress Type	**Network Level**	**Accuracy**			
		1 Sigma	**1.25 Sigma**	**1.5 Sigma**	**2 Sigma**
CRCP Spalled Cracks (Quantity)	**IH**	93.63%	95.44%	95.44%	96.88%
	US	97.16%	97.46%	98.35%	98.65%
	SH	97.91%	98.04%	98.17%	98.30%
	FM	88.61%	96.20%	96.20%	98.73%
CRCP Punchout (Quantity)	**IH**	92.88%	96.07%	96.07%	97.13%

	US	89.37%	89.37%	89.37%	97.16%
	SH	86.93%	86.93%	95.56%	95.56%
	FM	92.41%	92.41%	92.41%	92.41%
CRCP ACP Patches (Quantity)	**IH**	96.81%	97.19%	97.75%	98.19%
	US	96.56%	97.16%	97.16%	97.60%
	SH	96.99%	96.99%	96.99%	97.25%
	FM	96.62%	98.31%	98.31%	98.31%

CRCP PCC Patches (Quantity)	**IH**	90.51%	92.57%	93.19%	94.44%
	US	94.61%	95.36%	95.96%	96.86%
	SH	95.56%	96.21%	97.25%	98.43%
	FM	91.14%	92.83%	94.09%	95.78%
CRCP Average Crack Spacing (Average observed spacing in feet)	**IH**	82.32%	85.01%	86.45%	91.07%
	US	85.33%	86.68%	88.02%	91.17%
	SH	78.56%	81.57%	84.44%	90.85%
	FM	63.71%	80.59%	97.47%	98.73%

Table 5.59 Comparison of sigma-based methods for measurement threshold development in JCP

Distress Type	Network Level	Thresholds			
		1 Sigma	1.25 Sigma	1.5 Sigma	2 Sigma
JCP Failed Joint Cracks (Quantity)	**IH**	(-9.8,15.83)	(-13.01,19.04)	(-16.21,22.24)	(-22.62,28.65)
	US	(-8.42,13.71)	(-11.18,16.48)	(-13.95,19.24)	(-19.48, 24.77)
	SH	(-5.02,6.17)	(-6.42,7.57)	(-7.81,8.97)	(-10.61, 11.77)
	FM	(-1.8,1.6)	(-2.23,2.03)	(-2.65,2.45)	(-3.51, 3.3)
JCP Failures (Quantity)	**IH**	(-24.3,35.24)	(-31.74,42.68)	(-39.19,50.12)	(-54.07, 65.0)
	US	(-8.18,11.85)	(-10.68,14.35)	(-13.18,16.85)	(-18.19, 21.86)
	SH	(-6.21,8.77)	(-8.09,10.64)	(-9.96,12.51)	(-13.71, 16.26)
	FM	(-2.97,5.44)	(-4.03,6.49)	(-5.08,7.55)	(-7.18, 9.65)
JCP PCC Patches (Quantity)	**IH**	(-6.03,6.48)	(-7.59,8.04)	(-9.15,9.6)	(-12.28, 12.73)
	US	(-9.94,9.5)	(-12.37,11.93)	(-14.8,14.36)	(-19.66, 19.22)
	SH	(-5.56,6.19)	(-7.02,7.65)	(-8.49,9.12)	(-11.43, 12.06)
	FM	(-3.7,6.98)	(-5.03,8.32)	(-6.36,9.65)	(-9.03, 12.32)
JCP Shattered Slabs (Quantity)	**IH**	(-11.68,15.88)	(-15.12,19.32)	(-18.57,22.77)	(-25.45, 29.65)
	US	(-0.65,0.82)	(-0.84,1.01)	(-1.02,1.19)	(-1.39, 1.56)
	SH	(-2.19,3.04)	(-2.85,3.7)	(-3.5,4.35)	(-4.81, 5.66)
	FM	(-1.28,1.41)	(-1.62,1.74)	(-1.96,2.08)	(-2.63, 2.75)
JCP Slabs with Longitudinal Cracks (Quantity)	**IH**	(-10.78,9.65)	(-13.33,12.2)	(-15.89,14.75)	(-20.99,19.86)
	US	(-5.72,4.04)	(-6.94,5.26)	(-8.16,6.48)	(-10.6,8.92)
	SH	(-5.91,4.54)	(-7.21,5.85)	(-8.52,7.15)	(-11.13, 9.76)
	FM	(-6.63,3.79)	(-7.93,5.09)	(-9.23,6.39)	(-11.83,9.0)
JCP Apparent Joint Space	**IH**	(-4.25,5.98)	(-5.53,7.26)	(-6.8,8.54)	(-9.36, 11.1)
	US	(-10.39,10.59)	(-13.01,13.22)	(-15.63,15.84)	(-20.88,21.09)
	SH	(-12.16,20.45)	(-16.23,24.53)	(-20.31,28.6)	(-28.46,36.76)
	FM	(-5.39,3.86)	(-6.55,5.02)	(-7.71,6.18)	(-10.02, 8.49)

Table 5.60 Comparison of sigma-based methods for accuracy check in JCP

Distress Type	Network Level	Accuracy			
		1 Sigma	1.25 Sigma	1.5 Sigma	2 Sigma
JCP Failed Joint Cracks (Quantity)	**IH**	95.33%	96%	96.67%	96.67%
	US	92.65%	93.75%	94.49%	95.96%
	SH	88.52%	89.95%	90.43%	93.78%
	FM	78.57%	87.76%	87.76%	91.84%
JCP Failures	**IH**	97.33%	98.67%	98.67%	98.67%

(Quantity)	**US**	95.59%	96.69%	97.43%	98.53%
	SH	90.91%	91.87%	93.78	94.74%
	FM	88.78%	90.82%	91.84%	93.88%
JCP PCC Patches (Quantity)	**IH**	86.67%	88.67%	91.33%	93.30%
	US	93.01%	94.12%	94.85%	96.69%
	SH	91.39%	92.34%	92.82%	94.74%
	FM	87.76%	91.84%	96.94%	96.94%
JCP Shattered Slabs (Quantity)	**IH**	97.33%	98%	98%	98%
	US	97.79%	98.90%	98.90%	98.90%
	SH	97.13%	97.13%	97.13%	97.61%
	FM	96.94%	96.94%	97.96%	97.96%
JCP Slabs with Longitudinal Cracks (Quantity)	**IH**	92.67%	94.67%	96%	96.67%
	US	88.60%	89.71%	91.18%	94.85%
	SH	89.47%	90.43%	92.82%	95.22%
	FM	85.71%	87.76%	89.80%	93.88%
JCP Apparent Joint Space	**IH**	90.67%	94%	94.67%	96.67%
	US	91.54%	93.01%	93.38%	94.85%
	SH	85.17%	86.60%	88.04%	91.87%
	FM	89.80%	90.82%	90.82%	93.88%

By using the auto and audit data, the research team developed index #3 distress score threshold and accuracy. Table 5.61 lists the distress score thresholds based on 1 sigma, 1.25 sigma, 1.5 sigma, and 2 sigma methods of different distress types in IH, US, SH, and FM highway network levels for three pavement types (ACP, CRCP, and JCP). Table 5.62 lists accuracies of each distress type based on 1 sigma, 1.25 sigma, 1.5 sigma, and 2 sigma methods in IH, US, SH, and FM highway network levels for three pavement types (ACP, CRCP, and JCP).

Table 5.61 Comparison of sigma-based methods for distress score threshold development

Pavement Type	**Network Level**	**DS Range**			
		1 Sigma	**1.25 Sigma**	**1.5 Sigma**	**2 Sigma**
ACP	**IH**	(-8.43,12.25)	(-11.01,14.84)	(-13.6,17.42)	(-18.77,22.6)
	SH	(-8.7,11.7)	(-11.25,14.25)	(-13.8,16.8)	(-18.9,21.9)
	US	(-8.73,11.84)	(-11.31,14.41)	(-13.88,16.99)	(-19.02,22.13)
	FM	(-9.01,9.85)	(-11.37,12.21)	(-13.73,14.57)	(-18.44,19.28)
CRCP	**IH**	(-13.17,13.42)	(-16.5,16.74)	(-19.82,20.07)	(-26.47,26.71)
	SH	(-13.17,12.22)	(-16.34,15.4)	(-19.52,18.57)	(-25.86,24.92)
	US	(-12.32,12.3)	(-15.4,15.37)	(-18.48,18.45)	(-24.63,24.6)
	FM	(-16.44,16.23)	(-20.52,20.31)	(-24.6,24.39)	(-32.77,32.56)
JCP	**IH**	(-34.63,15.29)	(-40.87,21.53)	(-47.11,27.77)	(-59.59,40.25)
	SH	(-29.23,17.61)	(-35.08,23.46)	(-40.94,29.32)	(-52.64,41.03)
	US	(-26.1,14.16)	(-31.13,19.2)	(-36.16, 24.23)	(-46.23,34.29)
	FM	(-23.4,12.4)	(-27.88,16.88)	(-32.36,21.36)	(-41.31,30.31)

Table 5.62 Comparison of sigma-based methods for distress score accuracy check

Pavement Type	Network Level	Accuracy			
		1 Sigma	1.25 Sigma	1.5 Sigma	2 Sigma
ACP	**IH**	86.59%	89.08%	90.66%	93.75%
	SH	85.52%	88.78%	90.18%	92.88%
	US	85.75%	89.13%	90.73%	88.30%
	FM	85.24%%	89.52%	91.07%	93.79%
CRCP	**IH**	85.51%	88.01%	90.07%	92.32%
	SH	85.62%	87.06%	92.81%	92.81%
	US	86.23%	88.47%	90.87%	92.81%
	FM	83.12%	87.76%	90.30%	91.98%
JCP	**IH**	84%	85.33%	86%	92.67%
	SH	84.69%	87.56%	88.52%	93.30%
	US	84.19%	87.50%	89.71%	94.12%
	FM	83.67%	84.69%	86.73%	91.84%

By comparing every two years' auto data, the research team developed index #4 two-year distress measurement difference threshold and precision. Table 5.63, Table 5.65, and Table 5.67 list the two years' measurement difference thresholds based on 1 sigma, 1.25 sigma, 1.5 sigma, and 2 sigma methods of different distress types in IH, US, SH, and FM highway network levels for three pavement types (ACP, CRCP, and JCP). Table 5.64, Table 5.66, and Table 5.68 list precisions of each distress type based on 1 sigma, 1.25 sigma, 1.5 sigma, and 2 sigma methods in IH, US, SH, and FM highway network levels for three pavement types (ACP, CRCP, and JCP).

Table 5.63 Comparison of sigma-based methods for measurement difference threshold development in ACP

Distress Type	Network Level	Threshold			
		1 Sigma	1.25 Sigma	1.5 Sigma	2 Sigma
ACP Alligator Cracking (Percentage of Wheel path Length)	**IH**	(-4.74,5.23)	(-5.98,6.48)	(-7.23, 7.72)	(-9.72， 10.22)
	US	(-5.0,5.77)	(-6.34,7.12)	(-7.69, 8.46)	(-10.38, 11.16)
	SH	(-5.41,6.11)	(-6.85,7.55)	(-8.29, 8.99)	(-11.17, 11.87)
	FM	(-4.28,5.3)	(-5.47,6.5)	(-6.67, 7.69)	(-9.07, 10.09)
	All	(-4.68,5.54)	(-5.96,6.82)	(-7.24, 8.1)	(-9.8, 10.65)
ACP Longitudinal Cracking (Length in Feet per station).	**IH**	(-22.77,28.36)	(-29.16,34.75)	(-35.55, 41.14)	(-48.33, 53.92)
	US	(-21.35,26.22)	(-27.29,32.17)	(-33.24, 38.12)	(-45.13, 50.01)
	SH	(-19.04,23.01)	(-24.3,28.26)	(-29.55, 33.52)	(-40.06, 44.03)
	FM	(-19.11,24.13)	(-24.51,29.53)	(-29.92, 34.94)	(-40.73, 45.74)
	All	(-19.88,24.72)	(-25.45,30.3)	(-31.03, 35.87)	(-42.18, 47.02)
ACP Transverse Cracking (Quantity)	**IH**	(-1.32,1.87)	(-1.72,2.27)	(-2.12, 2.67)	(-2.92, 3.47)
	US	(-1.23,1.88)	(-1.62,2.27)	(-2, 2.66)	(-2.78, 3.44)
	SH	(-1.03,1.57)	(-1.36,1.89)	(-1.69, 2.22)	(-2.34, 2.87)
	FM	(-0.8,1.11)	(-1.04,1.35)	(-1.28, 1.58)	(-1.75, 2.06)

	All	(-1.0,1.43)	(-1.3,1.73)	(-1.61, 2.04)	(-2.21, 2.65)
ACP Patching (Percentage of Lane Area)	**IH**	(-5.78,6.14)	(-7.27,7.63)	(-8.76, 9.12)	(-11.74, 12.1)
	US	(-8.78,10.16)	(-11.15,12.52)	(-13.52, 14.89)	(-18.25, 19.62)
	SH	(-6.54,6.79)	(-8.2,8.46)	(-9.87, 10.12)	(-13.2, 13.45)
	FM	(-6.16,5.94)	(-7.67,7.45)	(-9.18, 8.96)	(-12.2, 11.98)
	All	(-6.64,6.8)	(-8.32,8.48)	(-10, 10.16)	(-13.36, 13.52)
ACP Block Cracking (Percentage of Lane Area)	**IH**	(-4.07,4.61)	(-5.15, 5.7)	(-6.24, 6.78)	(-8.41, 8.95)
	US	(-13.94,10.77)	(-17.02,13.86)	(-20.11, 16.95)	(-26.29, 23.13)
	SH	(-4.46,3.23)	(-5.42,4.19)	(-6.39, 5.15)	(-8.31, 7.07)
	FM	(-5.44,4.84)	(-6.72,6.12)	(-8.01, 7.41)	(-10.57, 9.98)
	All	(-7.43,6.34)	(-9.15,8.06)	(-10.87, 9.78)	(-14.31, 13.23)
ACP Failure (Quantity)	**IH**	(-1.34,1.29)	(-1.67,1.62)	(-2.0, 1.95)	(-2.66, 2.61)
	US	(-0.82,0.8)	(-1.02,1.0)	(-1.23, 1.2)	(-1.63, 1.6)
	SH	(-0.51,0.52)	(-0.64,0.65)	(-0.77, 0.78)	(-1.03, 1.04)
	FM	(-0.71,0.69)	(-0.89,0.86)	(-1.07, 1.04)	(-1.42, 1.39)
	All	(-0.8,0.77)	(-0.99,0.97)	(-1.19, 1.16)	(-1.58, 1.56)

Table 5.64 Comparison of sigma-based methods for precision check in ACP

Distress Type	**Network Level**	**Precision**			
		1 Sigma	**1.25 Sigma**	**1.5 Sigma**	**2 Sigma**
ACP Alligator Cracking (Percentage of Wheel path Length)	**IH**	95.05%	95.89%	96.55%	97.82%
	US	95.31%	96.44%	96.91%	97.87%
	SH	95.77%	96.43%	96.86%	97.83%
	FM	93.71%	94.92%	95.75%	97.41%
	All	94.35%	95.35%	96.65%	97.05%
ACP Longitudinal Cracking (Length in Feet per station).	**IH**	89.59%	91.98%	93.75%	96.01%
	US	90.05%	92.45%	94.23%	96.53%
	SH	90.78%	92.86%	94.35%	95.59%
	FM	90.13%	92.48%	94.20%	96.48%
	All	89.93%	92.56%	94.16%	96.50%
ACP Transverse Cracking (Quantity)	**IH**	90.90%	94.38%	94.47%	96.20%
	US	88.71%	93.72%	93.72%	96.22%
	SH	91.28%	91.28%	94.87%	94.96%
	FM	94.37%	94.61%	94.61%	97.49%
	All	92.33%	92.69%	96.07%	96.13%
ACP Patching (Percentage of Lane Area)	**IH**	97.15%	97.41%	97.79%	98.23%
	US	97.88%	98.06%	98.22%	98.56%
	SH	97.55%	97.95%	98.26%	98.56%
	FM	96.81%	97.36%	97.62%	98.16%
	All	97.24%	97.69%	98.03%	98.43%
ACP Block Cracking (Percentage of Lane	**IH**	96.81%	97.87%	97.87%	97.87%
	US	96.64%	97.06%	97.35%	97.60%
	SH	96.59%	96.60%	97.38%	97.78%

Area)	**FM**	99.22%	99.48%	99.49%	99.62%
	All	98.64%	98.74%	98.92%	99.23%
ACP Failure (Quantity)	**IH**	98.96%	98.96%	98.96%	99.44%
	US	99.15%	99.66%	99.66%	99.66%
	SH	98.88%	98.88%	98.88%	99.60%
	FM	97.52%	97.52%	99.12%	99.12%
	All	98.01%	98.01%	99.27%	99.27%

Table 5.65 Comparison of sigma-based methods for measurement difference threshold development in CRCP

Distress Type	**Network Level**	**Threshold range**			
		1 Sigma	**1.25 Sigma**	**1.5 Sigma**	**2 Sigma**
CRCP Spalled Cracks (Quantity)	**IH**	(-4.28,4.05)	(-5.32,5.09)	(-6.36, 6.13)	(-8.44, 8.22)
	US	(-7.48,7.96)	(-9.41,9.89)	(-11.34, 11.82)	(-15.21, 15.68)
	SH	(-3.95,4.01)	(-4.95,5.01)	(-5.94, 6)	(-7.93, 7.99)
	FM	(-3.34,3.04)	(-4.14,3.84)	(-4.94, 4.64)	(-6.54, 6.23)
	All	(-4.64,4.53)	(-5.79,5.68)	(-6.93, 6.82)	(-9.23, 9.11)
CRCP Punchout (Quantity)	**IH**	(-0.74,0.79)	(-0.93,0.99)	(-1.13, 1.18)	(-1.51, 1.56)
	US	(-1.02,1.23)	(-1.3,1.51)	(-1.58, 1.79)	(-2.14, 2.35)
	SH	(-0.55,0.53)	(-0.68,0.66)	(-0.82, 0.8)	(-1.09, 1.07)
	FM	(-0.48,0.47)	(-0.6,0.59)	(-0.72, 0.71)	(-0.95, 0.94)
	All	(-0.74,0.79)	(-0.93,0.98)	(-1.12, 1.17)	(-1.5, 1.55)
CRCP ACP Patches (Quantity)	**IH**	(-2.62,2.76)	(-3.29,3.43)	(-3.96, 4.1)	(-5.31, 5.45)
	US	(-2.69,2.79)	(-3.37,3.47)	(-4.06, 4.16)	(-5.43, 5.53)
	SH	(-4.64,3.99)	(-5.72,5.07)	(-6.79, 6.15)	(-8.95, 8.31)
	FM	(-3.95,3.52)	(-4.88,4.46)	(-5.82, 5.39)	(-7.68, 7.26)
	All	(-3.13,3.09)	(-3.91,3.87)	(-4.69, 4.65)	(-6.25, 6.21)
CRCP PCC Patches (Quantity)	**IH**	(-2.91,2.85)	(-3.63,3.57)	(-4.35, 4.29)	(-5.78, 5.73)
	US	(-2.3,2.17)	(-2.86,2.73)	(-3.42, 3.29)	(-4.54, 4.41)
	SH	(-2.7,2.73)	(-3.38,3.41)	(-4.06, 4.09)	(-5.41, 5.45)
	FM	(-3.11,2.79)	(-3.85,3.53)	(-4.59, 4.27)	(-6.07, 5.74)
	All	(-2.84,2.76)	(-3.54,3.46)	(-4.24, 4.16)	(-5.64, 5.56)
CRCP Average Crack Spacing (Average observed spacing in feet)	**IH**	(-22.32,14.8)	(-26.96,19.44)	(-31.6, 24.08)	(-40.88, 33.36)
	US	(-17.95,14.21)	(-21.97,18.23)	(-25.98, 22.25)	(-34.02, 30.29)
	SH	(-29.17,17.03)	(-34.94,22.81)	(-40.71, 28.58)	(-52.26, 40.13)
	FM	(-30.11,16.47)	(-35.93,22.29)	(-41.75, 28.11)	(-53.39, 39.76)
	All	(-23.87,15.43)	(-28.79,20.34)	(-33.7, 25.26)	(-43.53, 35.08)

Table 5.66 Comparison of sigma-based methods for precision check in CRCP

Distress Type	**Network Level**	**Precision**			
		1 Sigma	**1.25 Sigma**	**1.5 Sigma**	**2 Sigma**
CRCP Spalled Cracks (Quantity)	**IH**	99%	99.36%	99.53%	99.70%
	US	98.55%	98.55%	98.55%	98.75%

	SH	98.41%	98.55%	98.70%	98.77%
	FM	99.20%	99.20%	99.40%	99.60%
	All	98.89%	99.15%	99.30%	99.54%
CRCP Punchout (Quantity)	**IH**	93.69%	93.69%	97.68%	97.68%
	US	97.77%	97.77%	97.77%	98.34%
	SH	96.32%	96.32%	96.32%	98.98%
	FM	94.79%	94.79%	94.79%	94.79%
	All	94.27%	94.27%	98%	94.79%
CRCP ACP Patches (Quantity)	**IH**	98.41%	98.75%	99.05%	99.17%
	US	98.23%	98.53%	98.63%	98.73%
	SH	98%	98.60%	98.60%	98.68%
	FM	98.38%	98.69%	99.20%	99.20%
	All	98.58%	98.58%	98.85%	99.12%
CRCP PCC Patches (Quantity)	**IH**	95%	96.34%	97.18%	97.68%
	US	95.27%	95.27%	97.02%	97.94%
	SH	93.95%	95.43%	96.67%	97.50%
	FM	94.07%	95.32%	97.02%	97.47%
	All	94.77%	96.17%	97.17%	97.70%
CRCP Average Crack Spacing (Average observed spacing in feet)	**IH**	90.41%	92.66%	94.67%	97.07%
	US	89.87%	92.04%	94.31%	97.27%
	SH	88.02%	90.52%	93.88%	96.47%
	FM	87.79%	90.68%	93.05%	96.45%
	All	89.82%	92.35%	94.33%	97.01%

Table 5.67 Comparison of sigma-based methods for measurement difference threshold development in JCP

Distress Type	**Network Level**	**Threshold range**			
		1 Sigma	**1.25 Sigma**	**1.5 Sigma**	**Threshold range**
JCP Failed Joint Cracks (Quantity)	**IH**	(-22.67,12.13)	(-27.02,16.48)	(-31.37, 20.83)	(-40.07, 29.53)
	US	(-6.97,7.06)	(-8.72,8.81)	(-10.48, 10.57)	(-13.98, 14.07)
	SH	(-4.65,2.91)	(-5.6,3.86)	(-6.54, 4.8)	(-8.43, 6.69)
	FM	(-15.31,13.27)	(-18.88,16.84)	(-22.45, 20.41)	(-29.6, 27.56)
	All	(-13.22,9.4)	(-16.04,12.23)	(-18.87, 15.06)	(-24.52, 20.71)
JCP Failures (Quantity)	**IH**	(-9.03,7.44)	(-11.09,9.49)	(-13.14, 11.55)	(-17.26, 15.67)
	US	(-4.98,4.52)	(-6.17,5.71)	(-7.36, 6.9)	(-9.74, 9.28)
	SH	(-6.48,6.15)	(-8.06,7.73)	(-9.64, 9.31)	(-12.8, 12.46)
	FM	(-3.88,5.07)	(-5.0,6.19)	(-6.12, 7.31)	(-8.36, 9.55)
	All	(-6.69,6.2)	(-8.3,7.81)	(-9.91, 9.43)	(-13.13, 12.65)
JCP PCC Patches (Quantity)	**IH**	(-4.45,5.27)	(-5.66,6.49)	(-6.87, 7.7)	(-9.3, 10.13)
	US	(-5.46,5.54)	(-6.84,6.91)	(-8.21, 8.29)	(-10.96, 11.04)
	SH	(-7.45,9.63)	(-9.58,11.76)	(-11.72, 13.9)	(-15.98, 18.17)
	FM	(-3.87,4.81)	(-4.95,5.89)	(-6.04, 6.98)	(-8.2, 9.15)
	All	(-6.04,7.31)	(-7.71,8.98)	(-9.38, 10.6)	(-12.71, 13.98)

JCP Shattered Slabs (Quantity)	**IH**	(-0.73,0.62)	(-0.9,0.79)	(-1.07, 0.96)	(-1.41, 1.3)
	US	(-0.33,0.29)	(-0.41,0.37)	(-0.49, 0.44)	(-0.64, 0.6)
	SH	(-1.2,1.37)	(-1.52,1.69)	(-1.85, 2.01)	(-2.49, 2.66)
	FM	(-0.8,0.72)	(-0.99,0.91)	(-1.18, 1.1)	(-1.56, 1.49)
	All	(-0.94,0.96)	(-1.18,1.2)	(-1.42, 1.44)	(-1.89, 1.92)
JCP Slabs with Longitudinal Cracks (Quantity)	**IH**	(-3.42 ,4.37)	(-4.39,5.35)	(-5.36, 6.32)	(-7.31, 8.27)
	US	(-3.76,2.23)	(-4.5,2.98)	(-5.25, 3.73)	(-6.75, 5.22)
	SH	(-3.57,4.82)	(-4.61,5.87)	(-5.66, 6.92)	(-7.76, 9.02)
	FM	(-8.18,11.77)	(-10.67,14.26)	(-13.17, 16.76)	(-18.16, 21.74)
	All	(-4.69,5.68)	(-5.99,6.98)	(-7.29, 8.28)	(-9.88, 10.87)
JCP Apparent Joint Space	**IH**	(-10.59,4.16)	(-12.43,6.0)	(-14.27, 7.85)	(-17.96,11.53)
	US	(-5.86,5.4)	(-7.27,6.8)	(-8.68, 8.21)	(-11.49, 11.03)
	SH	(-4.21,3.5)	(-5.17,4.46)	(-6.14, 5.43)	(-8.07, 7.35)
	FM	(-7.43,7.33)	(-9.27,9.17)	(-11.11, 11.02)	(-14.8, 14.7)
	All	(-7.02,4.9)	(-8.51,6.39)	(-10, 7.88)	(-12.98, 10.86)

Table 5.68 Comparison of sigma-based methods for precision check in JCP

Distress Type	**Network Level**	**Precision**			
		1 Sigma	**1.25 Sigma**	**1.5 Sigma**	**2 Sigma**
JCP Failed Joint Cracks (Quantity)	**IH**	98.91%	99.45%	100%	100%
	US	98.51%	99.25%	99.25%	99.25%
	SH	95.17%	96.23%	96.30%	97.32%
	FM	94.85%	94.85%	94.90%	95.92%
	All	98.06%	98.48%	98.62%	99.03%
JCP Failures (Quantity)	**IH**	95.72%	97.37%	98.98%	99.49%
	US	93.33%	95.28%	95.31%	96.95%
	SH	95.29%	96.01%	97.68%	98.36%
	FM	89.13%	90.82%	92.08%	96.08%
	All	94.34%	95.27%	97.23%	98.51%
JCP PCC Patches (Quantity)	**IH**	91.19%	93.26%	94.33%	96.94%
	US	92.13%	93.02%	93.80%	96.15%
	SH	93.42%	94.12%	95.13%	97.41%
	FM	92.78%	93.94%	95%	97.03%
	All	93.40%	94.37%	95.63%	97.56%
JCP Shattered Slabs (Quantity)	**IH**	97.41%	97.41%	97.47%	98.99%
	US	98.46%	--	98.46%	98.46%
	SH	96.41%	96.41%	97.71%	97.71%
	FM	97.03%	97.03%	97.71%	98.04%
	All	96.97%	97.84%	97.84%	97.84%
JCP Slabs with Longitudinal Cracks (Quantity)	**IH**	96.97%	97.47%	97.47%	97.98%
	US	93.22%	93.39%	96.88%	98.45%
	SH	91.13%	92.59%	93.98%	95.71%
	FM	93.88%	93.94%	95.96%	96.04%
	All	94.67%	95.57%	96.16%	96.99%

JCP Apparent Joint Space	**IH**	97.27%	98.91%	98.93%	98.96%
	US	99.25%	99.25%	99.25%	99.25%
	SH	94.50%	95.59%	95.99%	97%
	FM	97%	97.03%	98.04%	98.04%
	All	97.19%	97.35%	97.92%	98.48%

By comparing every two years' auto data, the research team developed index #5 two-year distress score, condition score, and ride score difference threshold and precision. Table 5.69 lists the two years' distress score difference thresholds based on 1 sigma, 1.25 sigma, 1.5 sigma, and 2 sigma methods of different distress types in IH, US, SH, and FM highway network levels for three pavement types (ACP, CRCP, and JCP). Table 5.70 lists distress score precisions of each distress type based on 1 sigma, 1.25 sigma, 1.5 sigma, and 2 sigma methods in IH, US, SH, and FM highway network levels for three pavement types (ACP, CRCP, and JCP).

Table 5.69 Comparison of sigma-based methods for distress score difference threshold development

Pavement Type	**Network Level**	**DS range**			
		1 Sigma	**1.25 Sigma**	**1.5 Sigma**	**2 Sigma**
ACP	**IH**	(-12.8,12.0)	(-15.9,15.1)	(-19, 18.19)	(-25.20, 24.39)
	SH	(-11.54,10.71)	(-14.33,13.5)	(-16.31, 15.67)	(-21.64, 21.0)
	US	(-10.98,10.34)	(-13.64,13.01)	(-17.11, 16.28)	(-22.67, 21.84)
	FM	(-12.96,10.84)	(-15.94,13.82)	(-18.92, 16.8)	(-22.87, 22.75)
	All	(-12.35,10.88)	(-15.26,13.78)	(-18.16, 16.69)	(-23.97, 22.5)
CRCP	**IH**	(-13.1,12.34)	(-16.28,15.51)	(-19.46, 18.69)	(-25.82, 25.05)
	SH	(-12.71,11.46)	(-15.73,14.48)	(-21.21, 22.06)	(-28.42, 29.27)
	US	(-14.0,14.85)	(-17.61,18.45)	(-18.75, 17.5)	(-24.79, 23.55)
	FM	(-14.13,15.33)	(-17.81,19.02)	(-21.5, 22.7)	(-28.86, 30.06)
	All	(-13.33,12.97)	(-16.62,16.25)	(19.91, 19.54)	(26.48, 26.12)
JCP	**IH**	(-21.19,29.92)	(-27.58,36.31)	(-33.97, 42.69)	(-46.75, 55.47)
	SH	(-16.42,16.67)	(-20.56,20.81)	(-31.71, 31.66)	(-42.27, 42.22)
	US	(-21.15,21.1)	(-26.43,26.38)	(-24.69, 24.94)	(-32.96, 33.22)
	FM	(-25.32,19.59)	(-30.93,25.2)	(-36.54, 30.81)	(-47.77, 42.04)
	All	(-21.18,22.75)	(-26.67,28.24)	(-32.16, 33.73)	(-43.14, 44.71)

Table 5.70 Comparison of sigma-based methods for distress score precision check

Pavement Type	**Network Level**	**Precision**			
		1 Sigma	**1.25 Sigma**	**1.5 Sigma**	**2 Sigma**
ACP	**IH**	91.19%	93.38%	94.94%	97.05%
	SH	91.19%	93.37%	94.77%	97.06%
	US	91.19%	93.30%	95.11%	97.03%
	FM	91.19%	92.90%	94.72%	97.04%
	All	91.19%	92.98%	94.81%	97.09%
CRCP	**IH**	91.19%	93.82%	95.07%	96.59%
	SH	91.19%	94.54%	93.59%	95.52%
	US	91.19%	93.10%	96.04%	96.47%

	FM	91.19%	92.39%	93.59%	96.11%
	All	91.19%	93.82%	96.80%	96.44%
JCP	**IH**	91.19%	87.57%	93.68%	95.88%
	SH	91.19%	92.97%	93.79%	95.62%
	US	91.19%	92.96%	94.62%	96.92%
	FM	91.19%	89.13%	91.40%	98.95%
	All	91.19%	90.54%	92.60%	96.24%

5.2.2 Threshold comparison and selection using FY 2021 and FY 2022 data

With the developed four sigma-based thresholds, the research team applied FY 2021 and FY 2022 auto and audit data to evaluate the pavement condition using 1.0 sigma, 1.5 sigma, and 2.0 sigma thresholds. Based on the analysis results from 1.0 sigma, 1.5 sigma, and 2.0 sigma thresholds, the research team decided to use 1.5 sigma as the recommended threshold method. The details of 1.5 sigma analysis for Index #2, Index #3, Index #4, and Index #5 are discussed subsequently.

1) 1.5 sigma analysis

Index #2

As shown in Table 5.71, CRCP Average Crack Spacing (SH, ALL), JCP Failed Joint Cracks (SH), JCP Slabs with Longitudinal Cracks (IH, US, FM), and JCP Apparent Joint Space (US, SH, FM, ALL) show an accuracy difference more than 10%. Those distresses may have issues by using the FY17-21 thresholds to locate inaccurate measurements. JCP Failed Joint Cracks (ALL), and JCP Slabs with Longitudinal Cracks (ALL) do not present a high accuracy difference, which indicates that the high difference could be caused by the limited amount of data while highway level was used as a filter. CRCP Average Crack Spacing (ALL) and JCP Apparent Joint Space (ALL) present an accuracy difference of more than 10%, which indicates that the FY17-21 thresholds of CRCP Average Crack Spacing and JCP Apparent Joint Space may have the most severe problem. The evaluation results for ACP data are good.

Table 5.71 Evaluation results on FY 2022 distress data (1.5 Sigma)

Pavement Type	Distress Type	Network Level	No. of data	Accuracy_FY1721	Accuracy_FY22	Difference
ACP	**ACP Alligator Cracking** (Percentage of Wheel path Length)	**IH**	84	0.950941	0.976344	3%
		US	343	0.952956	0.979045	3%
		SH	421	0.947607	0.984552	4%
		FM	1177	0.952369	0.979054	3%
		All	2041	0.952688	0.980607	3%
	ACP Longitudinal Cracking (Length in Feet per station).	**IH**	302	0.926411	0.929032	0%
		US	1206	0.900444	0.909357	1%
		SH	1255	0.907088	0.931395	2%
		FM	3386	0.913486	0.922716	1%
		All	6186	0.912223	0.923010	1%
	ACP Transverse Cracking	**IH**	118	0.949261	0.948387	-0%
		US	513	0.940988	0.946881	1%

	(Quantity)	**SH**	460	0.956037	0.971831	2%
		FM	793	0.932208	0.938967	1%
		All	1894	0.948742	0.952196	0%
	ACP Patching (Percentage of Lane Area)	**IH**	43	0.960685	0.948387	-1%
		US	156	0.972042	0.985867	1%
		SH	182	0.961748	0.972285	1%
		FM	650	0.959785	0.968942	1%
		All	1041	0.961740	0.970232	1%
	ACP Block Cracking (Percentage of Lane Area)	**IH**	3	0.990927	0.997849	1%
		US	36	0.985660	0.988791	0%
		SH	45	0.989032	0.990459	0%
		FM	85	0.993052	0.988805	-0%
		All	170	0.990522	0.989431	-0%
	ACP Failure (Quantity)	**IH**	13	0.974462	0.991398	2%
		US	46	0.991747	0.994639	0%
		SH	68	0.970722	0.969105	-0%
		FM	325	0.986211	0.986096	-0%
		All	459	0.988046	0.989140	0%
CRCP	**CRCP Spalled Cracks** (Quantity)	**IH**	71	0.954403	0.952941	-0%
		US	29	0.983533	0.973822	-1%
		SH	47	0.981699	0.996241	1%
		FM	7	0.962025	0.952381	-1%
		All	154	0.986561	0.988372	0%
	CRCP Punchout (Quantity)	**IH**	45	0.960650	0.967647	1%
		US	25	0.893713	0.869110	-2%
		SH	38	0.955556	0.939850	-2%
		FM	5	0.924051	0.920635	-0%
		All	113	0.963653	0.963953	0%
	CRCP ACP Patches (Quantity)	**IH**	30	0.977514	0.979412	0%
		US	12	0.971557	0.994764	2%
		SH	33	0.969935	0.996241	3%
		FM	5	0.983122	0.968254	-1%
		All	80	0.978619	0.987209	1%
	CRCP PCC Patches (Quantity)	**IH**	99	0.931918	0.941176	1%
		US	36	0.959581	0.973822	1%
		SH	53	0.972549	0.977444	0%
		FM	14	0.940928	0.984127	4%
		All	202	0.945021	0.958140	1%
	CRCP Average Crack Spacing (Average observed spacing in feet)	**IH**	312	0.864460	0.879412	1%
		US	177	0.880240	0.790576	-9%
		SH	256	0.844444	0.533835	-31%
		FM	63	0.974684	1.000000	3%
		All	808	0.857666	0.724419	-13%
JCP	**JCP Failed Joint Cracks** (Quantity)	**IH**	8	0.966667	1.000000	3%
		US	27	0.944853	0.983051	4%

		SH	26	0.904306	1.000000	10%
		FM	8	0.877551	0.857143	-2%
		All	69	0.958848	0.992424	3%
	JCP Failures (Quantity)	**IH**	6	0.986667	1.000000	1%
		US	33	0.974265	0.932203	-4%
		SH	26	0.937799	0.951220	1%
		FM	7	0.918367	0.928571	1%
		All	72	0.976680	0.969697	-1%
	JCP PCC Patches (Quantity)	**IH**	8	0.913333	0.944444	3%
		US	15	0.948529	0.983051	3%
		SH	18	0.928230	0.951220	2%
		FM	4	0.928571	1.000000	7%
		All	45	0.941015	0.977273	4%
	JCP Shattered Slabs (Quantity)	**IH**	0	0.980000	1.000000	2%
		US	2	0.988971	1.000000	1%
		SH	4	0.971292	1.000000	3%
		FM	0	0.979592	1.000000	2%
		All	6	0.990398	1.000000	1%
	JCP Slabs with Longitudinal Cracks (Quantity)	**IH**	13	0.960000	0.833333	-13%
		US	34	0.911765	0.813559	-10%
		SH	19	0.928230	0.975610	5%
		FM	4	0.897959	1.000000	10%
		All	70	0.938272	0.901515	-4%
	JCP Apparent Joint Space	**IH**	2	0.946667	1.000000	5%
		US	46	0.933824	0.457627	-48%
		SH	25	0.880383	0.609756	-27%
		FM	10	0.908163	0.357143	-55%
		All	83	0.924554	0.583333	-34%

Index #3

As shown in Table 5.72, JCP DISTRESS SCORE (US, SH, ALL) shows an accuracy difference of more than 10%. Those distresses may have issues by using the FY17-21 thresholds to locate inaccurate measurements. JCP DISTRESS SCORE (ALL) presents an accuracy difference of more than 10%, which indicates that the FY17-21 thresholds of JCP DISTRESS SCORE may have the most severe problem. The evaluation results for ACP and CRCP data are good.

Table 5.72 Evaluation results on FY 2022 DISTRESS SCORE data (1.5 Sigma)

Pavement Type	Distress Type	Network Level	No. of data	Accuracy_FY1721	Accuracy_FY22	Difference
ACP	**DISTRESS SCORE**	**IH**	152	0.906586	0.920430	1%
		US	601	0.907253	0.954678	5%
		SH	622	0.901840	0.955475	5%
		FM	1587	0.910718	0.941676	3%

		All	2984	0.907139	0.945894	4%
CRCP	**DISTRESS SCORE**	**IH**	124	0.900687	0.870588	-3%
		US	58	0.908683	0.895288	-1%
		SH	85	0.900654	0.838346	-6%
		FM	20	0.902954	0.873016	-3%
		All	287	0.899817	0.859302	-4%
JCP	**DISTRESS SCORE**	**IH**	11	0.860000	0.888889	3%
		US	30	0.897059	0.779661	-12%
		SH	28	0.885167	0.609756	-28%
		FM	5	0.867347	0.857143	-1%
		All	74	0.884774	0.750000	-13%

Index #4

As shown in Table 5.73, CRCP Average Crack Spacing (US), and JCP Failed Joint Cracks (FM) show an accuracy difference of more than 10%. Those distresses may have issues by using the FY17-21 thresholds to locate inaccurate measurements. However, CRCP Average Crack Spacing (ALL), and JCP Failed Joint Cracks (ALL) do not present a high accuracy difference, which indicates that the high difference could be caused by the limited amount of data while highway level was used as a filter.

Table 5.73 Evaluation results on FY 2022 distress data (1.5 Sigma)

Pavement Type	Distress Type	Network Level	No. of data	Precision_FY1721	Precision _FY22	Difference
ACP	**ACP Alligator Cracking** (Percentage of Wheel path Length)	**IH**	2277	0.940100	0.951302	1%
		US	5496	0.941290	0.965433	2%
		SH	5562	0.944109	0.966786	2%
		FM	8511	0.929390	0.961689	3%
		All	21846	0.941210	0.964973	2%
	ACP Longitudinal Cracking (Length in Feet per station).	**IH**	6734	0.898269	0.914914	2%
		US	16890	0.900113	0.914764	1%
		SH	14607	0.901246	0.908153	1%
		FM	22462	0.899679	0.925414	3%
		All	60693	0.901332	0.916045	1%
	ACP Transverse Cracking (Quantity)	**IH**	1603	0.929657	0.970694	4%
		US	4430	0.908316	0.962335	5%
		SH	3935	0.925346	0.963642	4%
		FM	3807	0.931261	0.967496	4%
		All	13775	0.939996	0.968385	3%
	ACP Patching (Percentage of Lane Area)	**IH**	934	0.958639	0.975596	2%
		US	2022	0.974050	0.987241	1%
		SH	2039	0.965458	0.975422	1%
		FM	3585	0.950182	0.967885	2%
		All	8580	0.959163	0.975408	2%
	ACP Failure	**IH**	469	0.975183	0.977775	0%

	(Quantity)	**US**	291	0.991261	0.995924	0%
		SH	348	0.978419	0.984589	1%
		FM	1140	0.979683	0.988477	1%
		All	2248	0.983394	0.991126	1%
CRCP	**CRCP Spalled Cracks** (Quantity)	**IH**	670	0.988924	0.990364	0%
		US	302	0.984601	0.994465	1%
		SH	457	0.979211	0.975172	-0%
		FM	75	0.983251	0.996890	1%
		All	1504	0.986820	0.988210	0%
	CRCP Punchout (Quantity)	**IH**	317	0.959054	0.979795	2%
		US	119	0.969201	0.986470	2%
		SH	174	0.918280	0.908082	-1%
		FM	59	0.896552	0.908243	1%
		All	669	0.964711	0.981569	2%
	CRCP ACP Patches (Quantity)	**IH**	353	0.982883	0.968604	-1%
		US	130	0.973051	0.969865	-0%
		SH	281	0.962007	0.969361	1%
		FM	38	0.972414	0.995334	2%
		All	802	0.976084	0.968830	-1%
	CRCP PCC Patches (Quantity)	**IH**	854	0.944286	0.907056	-4%
		US	409	0.940327	0.899139	-4%
		SH	471	0.935484	0.919176	-2%
		FM	172	0.929064	0.931571	0%
		All	1906	0.942921	0.914352	-3%
	CRCP Average Crack Spacing (Average observed spacing in feet)	**IH**	2772	0.866756	0.854523	-1%
		US	1437	0.876805	0.723247	-15%
		SH	1647	0.857348	0.782356	-7%
		FM	474	0.844335	0.892691	5%
		All	6330	0.861820	0.812305	-5%
JCP	**JCP Failed Joint Cracks** (Quantity)	**IH**	201	0.910448	0.949640	4%
		US	428	0.985185	0.955994	-3%
		SH	321	0.922581	0.926829	0%
		FM	102	0.894231	0.994286	10%
		All	1052	0.950667	0.978803	3%
	JCP Failures (Quantity)	**IH**	198	0.965174	0.920863	-4%
		US	423	0.903704	0.825493	-8%
		SH	356	0.951613	0.894309	-6%
		FM	122	0.894231	0.868571	-3%
		All	1099	0.937333	0.888404	-5%
	JCP PCC Patches (Quantity)	**IH**	120	0.910448	0.935252	2%
		US	342	0.896296	0.827011	-7%
		SH	291	0.945161	0.896341	-5%
		FM	70	0.913462	0.868571	-4%
		All	823	0.933333	0.888404	-4%
	JCP Shattered	**IH**	16	0.960199	0.953237	-1%

	Slabs (Quantity)	**US**	35	0.948148	0.946889	-0%
		SH	37	0.964516	0.975610	1%
		FM	11	0.961538	0.982857	2%
		All	99	0.965333	0.978180	1%
	JCP Slabs with Longitudinal Cracks (Quantity)	**IH**	107	0.960199	0.931655	-3%
		US	342	0.918519	0.839150	-8%
		SH	242	0.906452	0.926829	2%
		FM	101	0.931373	0.988571	6%
		All	792	0.937166	0.943267	1%
	JCP Apparent Joint Space	**IH**	192	0.920398	0.985612	7%
		US	403	0.985185	0.966616	-2%
		SH	297	0.925806	0.936992	1%
		FM	96	0.961538	0.988571	3%
		All	988	0.941333	0.973192	3%

Index #5

As shown in Table 5.74, JCP DISTRESS SCORE (US), and JCP RIDE SCORE (FM) show an accuracy difference of more than 10%. Those distresses may have issues by using the FY17-21 thresholds to locate inaccurate measurements. JCP DISTRESS SCORE (ALL), and JCP RIDE SCORE (ALL) do not present a high accuracy difference, which indicates that the high difference could be caused by the limited amount of data while highway level was used as a filter. The evaluation results for ACP and CRCP data are good.

Table 5.74 Evaluation results on FY 2022 DISTRESS SCORE, CONDITION SCORE, and RIDE SCORE data (1.5 Sigma)

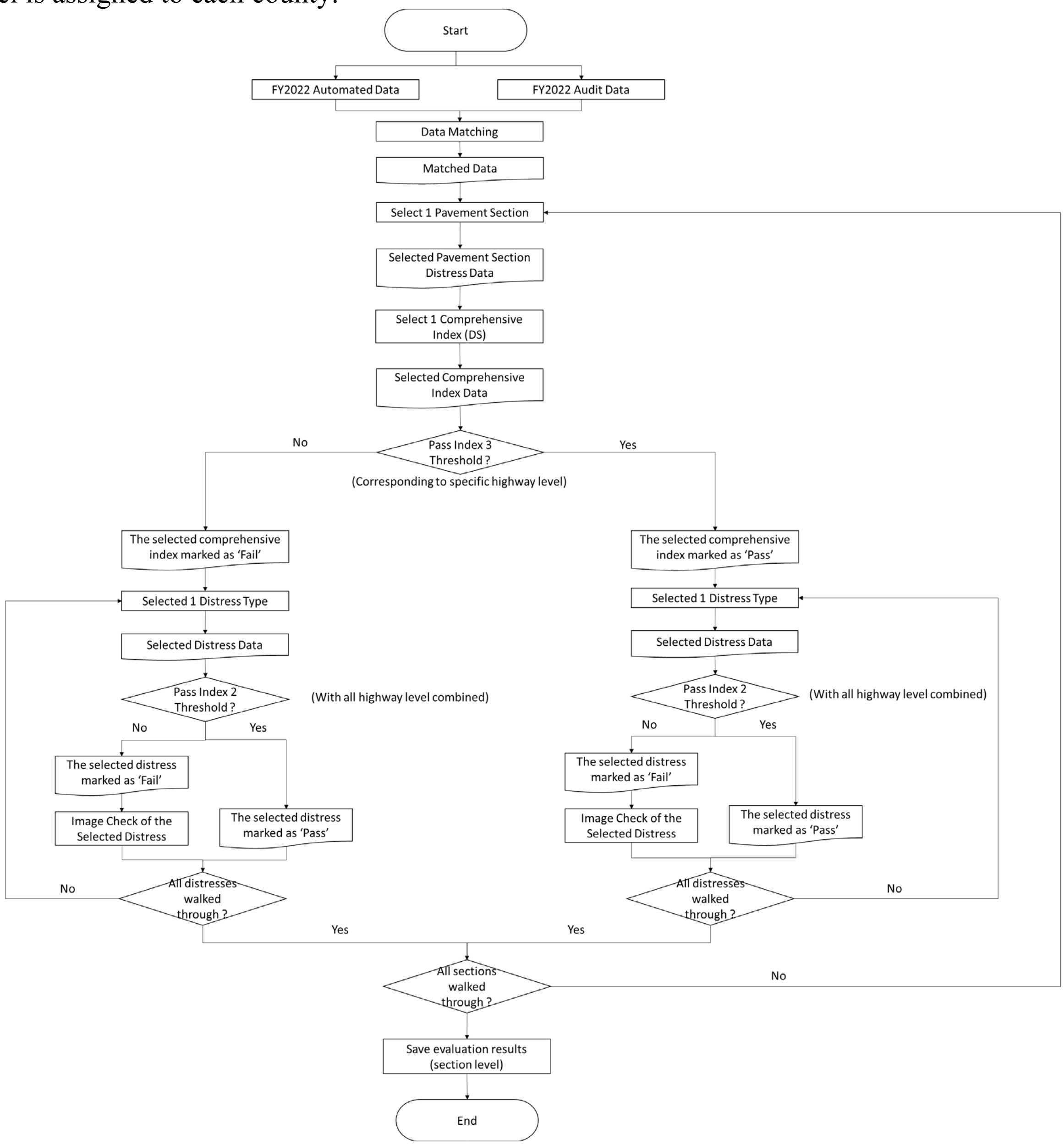
Pavement Type	Distress Type	Network Level	No. of data	Precesion_FY1721	Precision_FY22	Difference
ACP	**DISTRESS SCORE**	**IH**	5603	0.886594	0.917638	3%
		US	15545	0.894590	0.916273	2%
		SH	14455	0.892429	0.904787	1%
		FM	21625	0.885185	0.930173	4%
		All	57228	0.890426	0.922634	3%
	CONDITION SCORE	**IH**	5692	0.889868	0.919184	3%
		US	15781	0.895403	0.917560	2%
		SH	14748	0.896838	0.912747	2%
		FM	22091	0.893903	0.932611	4%
		All	58312	0.897229	0.929002	3%
	RIDE SCORE	**IH**	8606	0.921876	0.943548	2%
		US	23429	0.926621	0.921885	-0%
		SH	21729	0.934112	0.927777	-1%
		FM	32302	0.929701	0.941106	1%
		All	86066	0.929936	0.933185	0%
CRCP	**DISTRESS SCORE**	**IH**	1172	0.902333	0.883121	-2%
		US	518	0.909528	0.883764	-3%

		SH	700	0.878853	0.893291	1%
		FM	222	0.876847	0.877138	0%
		All	2612	0.896471	0.882233	-1%
	CONDITION SCORE	**IH**	1469	0.896255	0.892657	-0%
		US	713	0.905679	0.878695	-3%
		SH	937	0.876712	0.896351	2%
		FM	385	0.878698	0.899687	2%
		All	3504	0.891332	0.888286	-0%
	RIDE SCORE	**IH**	3082	0.927972	0.944963	2%
		US	1519	0.916266	0.903941	-1%
		SH	1792	0.928212	0.966173	4%
		FM	600	0.931953	0.974922	4%
		All	6993	0.927143	0.947354	2%
JCP	**DISTRESS SCORE**	**IH**	186	0.885572	0.884892	-0%
		US	411	0.911111	0.813354	-10%
		SH	353	0.877419	0.859756	-2%
		FM	118	0.817308	0.868571	5%
		All	1068	0.868000	0.865337	-0%
	CONDITION SCORE	**IH**	199	0.865672	0.888489	2%
		US	471	0.874074	0.881279	1%
		SH	414	0.883117	0.899796	2%
		FM	144	0.875000	0.914286	4%
		All	1228	0.866310	0.893058	3%
	RIDE SCORE	**IH**	264	0.935323	0.967626	3%
		US	624	0.866667	0.948250	8%
		SH	462	0.941558	0.963190	2%
		FM	168	0.884615	0.982857	10%
		All	1518	0.919786	0.958099	4%

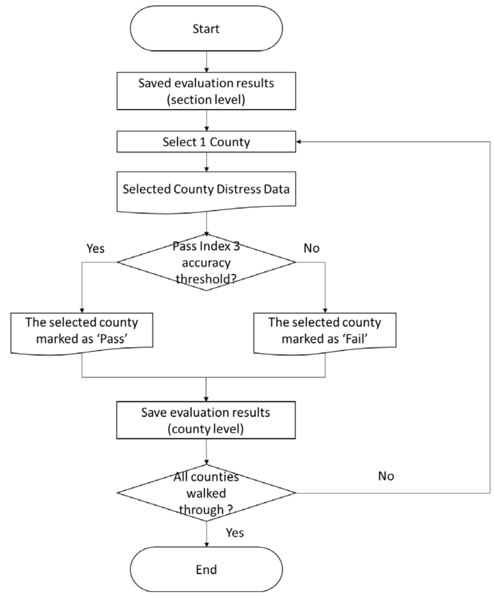
5.3 Implementation of Data Quality Consistency Check Components

5.3.1 Implementation of Data Accuracy Check for Distresses

Figure 5.4 shows the flowchart for FY2022 Data Quality Check Using the Thresholds of Index 2 and Index 3 computed from FY2017-2021. To facilitate comparisons between the automated and audit data, the measurements of both sources are first matched within the same pavement section. Subsequently, the matched data are subjected to data quality analysis after being filtered by county level, pavement section numbers, comprehensive index measurements (DS), and pavement distress measurements. The flowchart then proceeds to assess all pavement section data within each county and applies three thresholds to evaluate the selected pavement section: Index 3 threshold (corresponding to IH, SH, US, FM), Index 2 threshold (for all highway levels combined), and Index 3 accuracy threshold (for all highway levels combined). The results of the Index 3 and Index 2 threshold evaluations are stored as 'Fail' or 'Pass', where the former identifies pavement sections with potential data quality issues, and the latter helps identify which pavement distress causes the data quality issue. The Index 3 accuracy threshold (for all highway levels combined) is applied after all pavement sections in a county are evaluated. The percentage of pavement sections that pass the Index 3 threshold (corresponding to IH, SH, US, FM) is then computed and compared with the Index 3 accuracy threshold (for all highway levels combined). Finally, a 'Fail' or 'Pass' label is assigned to each county.



(a) Section level analysis

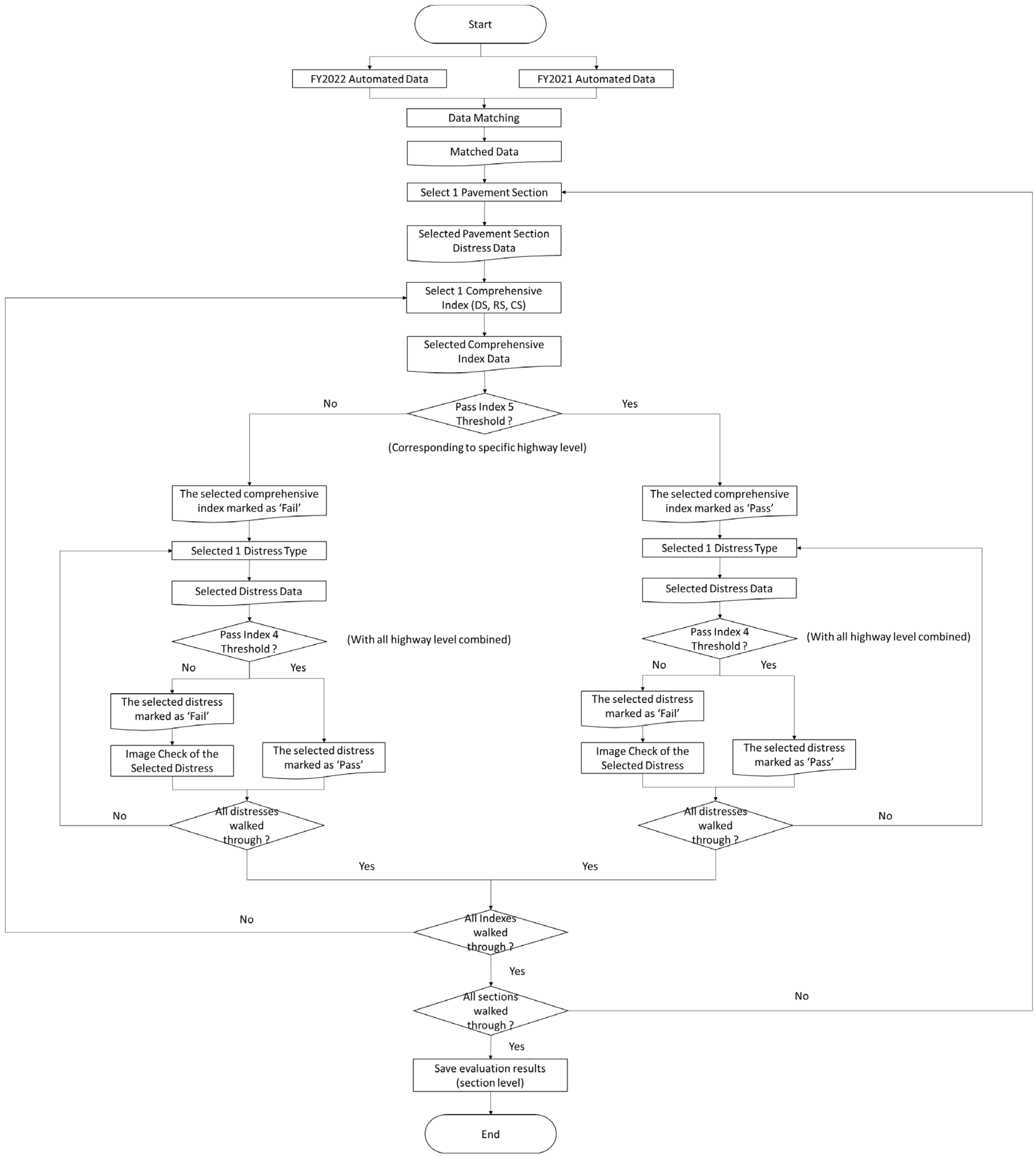


(b) County level analysis

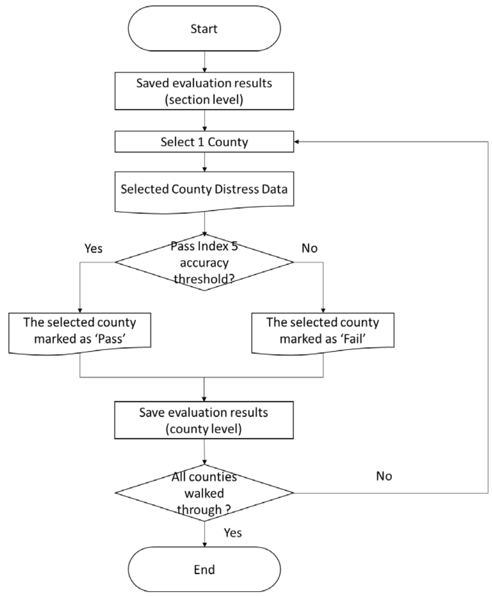
Figure 5.3 Flowchart for FY2022 Data Quality Check Using the Thresholds of Index 2 and Index 3

5.3.2 Implementation of Data Precision Check for Distresses

Figure 5.5 shows the flowchart for FY2022 Data Quality Check Using the Thresholds of Index 4 and Index 5 computed from FY2017-2021. To facilitate comparisons between the automated and audit data, the measurements of both sources are first matched within the same pavement section. Subsequently, the matched data are subjected to data quality analysis after being filtered by county level, pavement section numbers, comprehensive index measurements (DS, RS, CS), and pavement distress measurements. The flowchart then proceeds to assess all pavement section data within each county and applies three thresholds to evaluate the selected pavement section: Index 4 threshold (corresponding to IH, SH, US, FM), Index 4 threshold (for all highway levels combined), and Index 5 accuracy threshold (for all highway levels combined). The results of the Index 5 and Index 4 threshold evaluations are stored as 'Fail' or 'Pass', where the former identifies pavement sections with potential data quality issues, and the latter helps identify which pavement distress causes the data quality issue. The Index 5 accuracy threshold (for all highway levels combined) is applied after all pavement sections in a county are evaluated. The percentage of pavement sections that pass the Index 5 threshold (corresponding to IH, SH, US, FM) is then computed and compared with the Index 5 accuracy threshold (for all highway levels combined). Finally, a 'Fail' or 'Pass' label is assigned to each county.



(a) Section level analysis

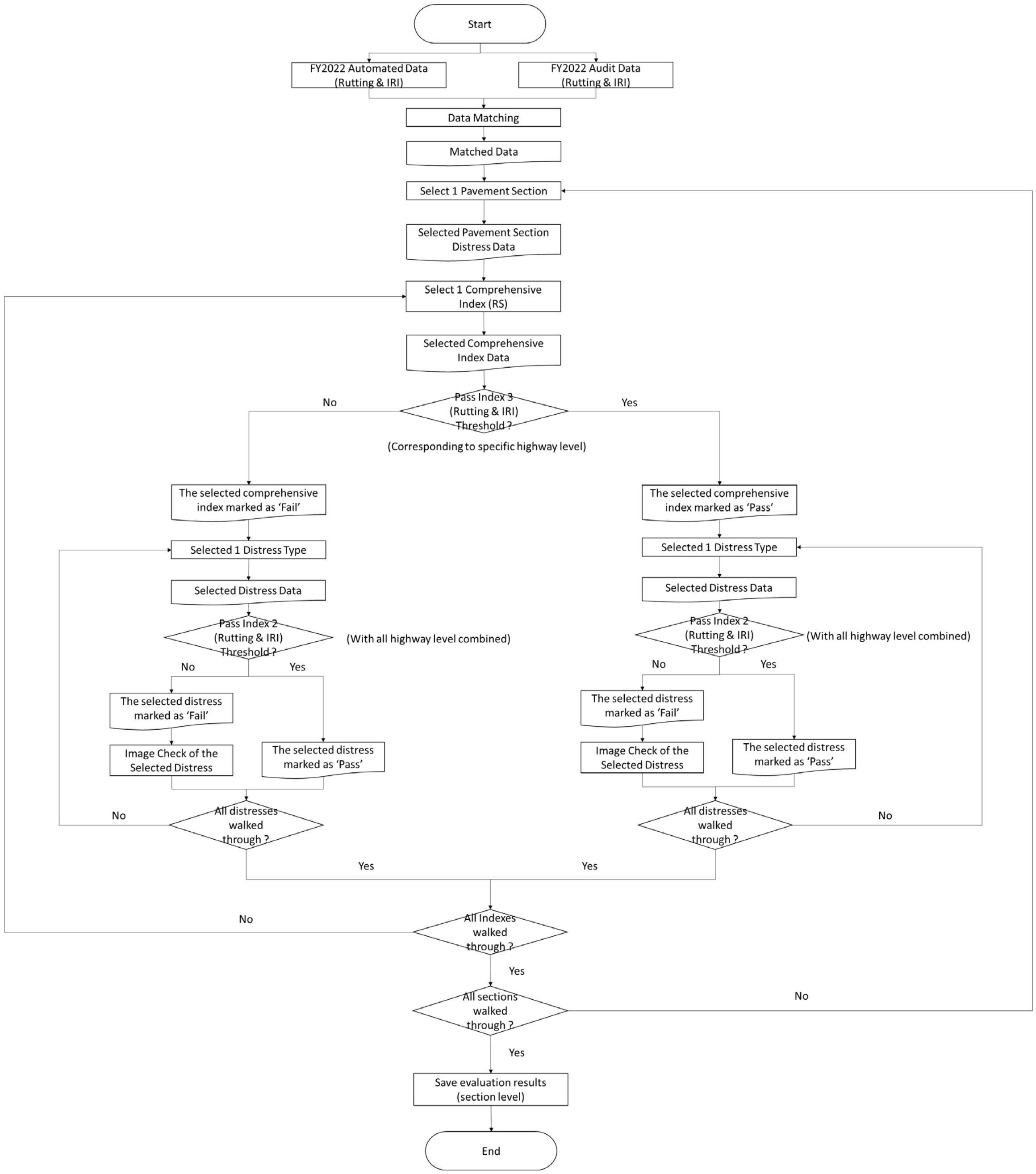


(b) County level analysis

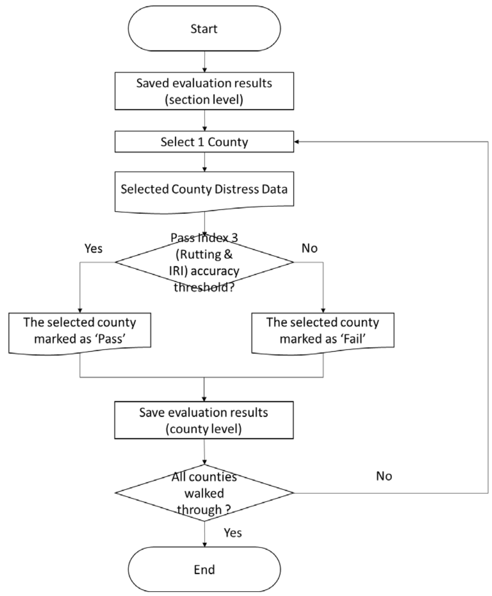
Figure 5.4 Flowchart for FY2022 Data Quality Check Using the Thresholds of Index 4 and Index 5

5.3.3 Implementation of Data Accuracy Check for Rutting and IRI

Figure 5.6 shows the flowchart for FY2022 Data Quality Check Using the Thresholds of Index 2 (Rutting & IRI) and Index 3 (Rutting & IRI) computed from FY2017-2021. To facilitate comparisons between the automated and audit data, the measurements of both sources are first matched within the same pavement section. Subsequently, the matched data are subjected to data quality analysis after being filtered by county level, pavement section numbers, comprehensive index measurements (RS), and pavement distress measurements. The flowchart then proceeds to assess all pavement section data within each county and applies three thresholds to evaluate the selected pavement section: Index 3 threshold (corresponding to IH, SH, US, FM), Index 2 threshold (for all highway levels combined), and Index 3 accuracy threshold (for all highway levels combined). The results of the Index 3 and Index 2 threshold evaluations are stored as 'Fail' or 'Pass', where the former identifies pavement sections with potential data quality issues, and the latter helps identify which pavement distress causes the data quality issue. The Index 3 accuracy threshold (for all highway levels combined) is applied after all pavement sections in a county are evaluated. The percentage of pavement sections that pass the Index 3 threshold (corresponding to IH, SH, US, FM) is then computed and compared with the Index 3 accuracy threshold (for all highway levels combined). Finally, a 'Fail' or 'Pass' label is assigned to each county.



(a) Section level analysis

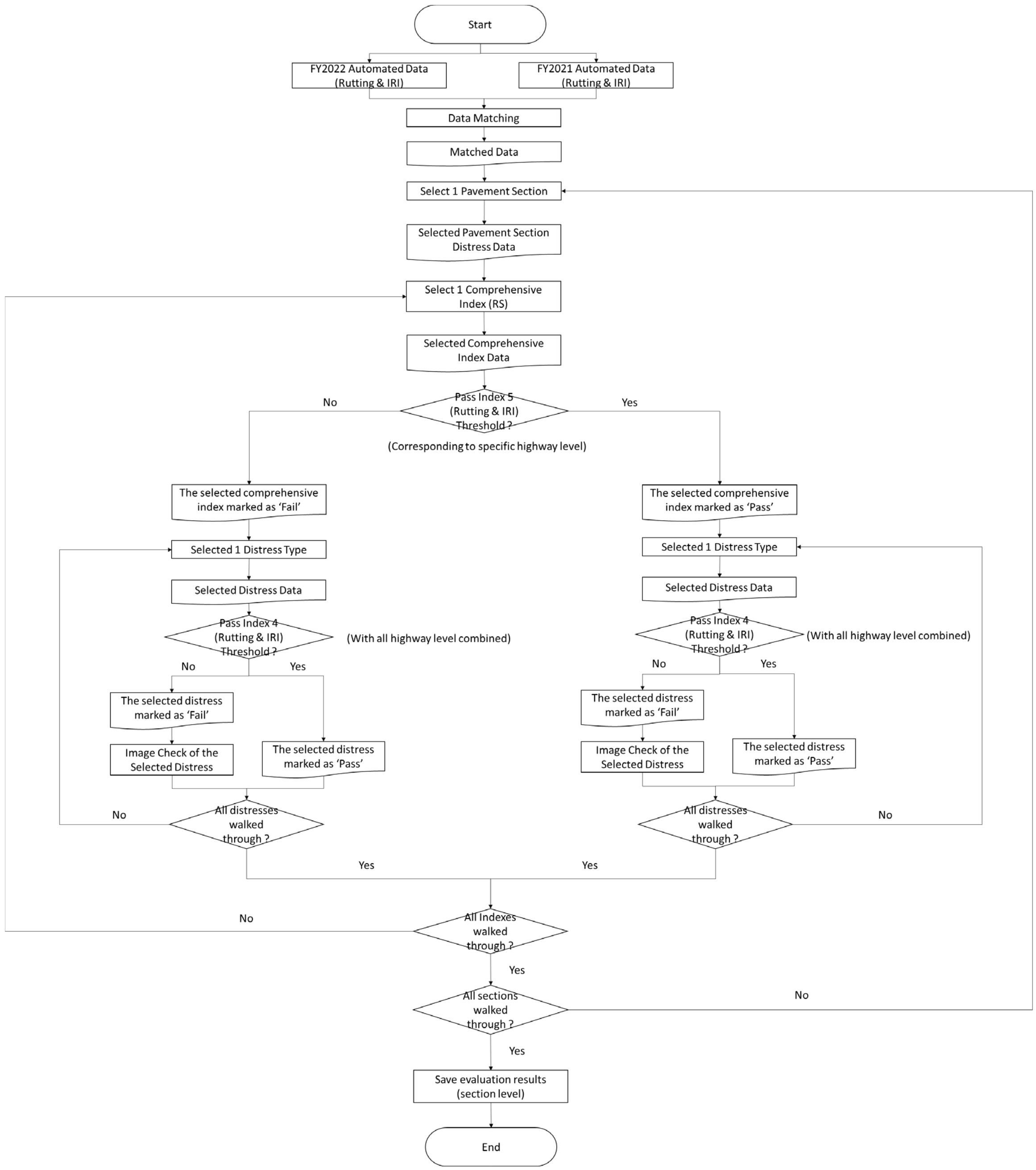


(b) County level analysis

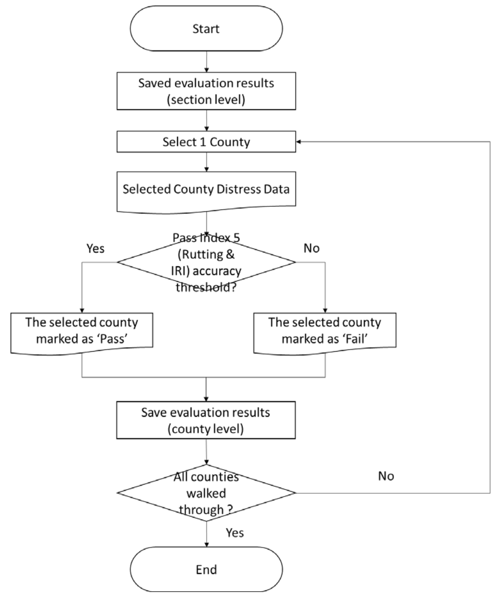
Figure 5.5 Flowchart for FY2022 Data Quality Check using the Thresholds of Index 2 (Rutting & IRI) and Index 3 (Rutting & IRI)

5.3.4 Implementation of Data Precision Check for Rutting and IRI

Figure 5.7 shows the flowchart for FY2022 Data Quality Check Using the Thresholds of Index 4 (Rutting & IRI) and Index 5 (Rutting & IRI) computed from FY2017-2021. To facilitate comparisons between the automated and audit data, the measurements of both sources are first matched within the same pavement section. Subsequently, the matched data are subjected to data quality analysis after being filtered by county level, pavement section numbers, comprehensive index measurements (RS), and pavement distress measurements. The flowchart then proceeds to assess all pavement section data within each county and applies three thresholds to evaluate the selected pavement section: Index 4 threshold (corresponding to IH, SH, US, FM), Index 4 threshold (for all highway levels combined), and Index 5 accuracy threshold (for all highway levels combined). The results of the Index 5 and Index 4 threshold evaluations are stored as 'Fail' or 'Pass', where the former identifies pavement sections with potential data quality issues, and the latter helps identify which pavement distress causes the data quality issue. The Index 5 accuracy threshold (for all highway levels combined) is applied after all pavement sections in a county are evaluated. The percentage of pavement sections that pass the Index 5 threshold (corresponding to IH, SH, US, FM) is then computed and compared with the Index 5 accuracy threshold (for all highway levels combined). Finally, a 'Fail' or 'Pass' label is assigned to each county.



(a) Section level analysis



(b) County level analysis

Figure 5.6 Flowchart for FY2022 Data Quality Check using the Thresholds of Index 4 (Rutting & IRI) and Index 5 (Rutting & IRI)

Chapter 6 Pilot Study Results of Data Quality Assurance

6.1 Data Quality Check

In this chapter, the image check result of the pilot study district San Antonio is presented to validate the data quality check procedure result of precision. The image check processes used the images from San Antonio district which were collected from FY 2021 and 2022. There are a total of three distress results per image, 1st one is the Pathway automatically processed distress result technically recorded in the PMIS database, 2nd one is the manual audit result, and 3rd one is the distress result from manually reviewing Pathway images conducted by the research team's well-trained raters. The sections selected for the image check are those with potential data quality issues from the pilot study using the data quality validation procedures. The image check result in this chapter is presented by each distress types and Fiscal Year. Based on the data availability, only typical distress types of alligator cracking, longitudinal cracking, transverse cracking, patching, and failure from asphalt pavement are included.

6.1.1 Data Quality Check for Index #2

Table 6.1 presents the data quality check results for San Antonio on Index 2. For ACP, the table reveals the accuracy of ACP Block Cracking, ACP Alligator Cracking, ACP Transverse Cracking, and ACP Failure are in good data quality. ACP Patching shows slightly lower accuracy compared to the Index 2 threshold, which suggests a minor data quality concern. ACP Longitude Cracking shows severe lower accuracy compared to the Index 2 threshold, which may suggest a strong data quality concern. The primary data quality issue identified in this analysis is FN. This finding implies that the annual data might tend to underestimate the severity of distress on the pavement.

For CRCP, the table reveals the accuracy of CRCP Spalled Cracks, CRCP Punchout, CRCP ACP Patches, CRCP PCC Patches, and CRCP AVG Crack Spacing are in good data quality. CRCP PCC Patches show slightly lower accuracy compared to the Index 2 threshold, which suggests a minor data quality concern. CRCP AVG Crack Spacing shows severe lower accuracy compared to the Index 2 threshold, which may suggest a strong data quality concern. The primary data quality issue identified in this analysis is FP. This finding implies that the annual data might tend to overestimate the severity of distress on the pavement.

For JCP, only a few of the pavement section data was involved in the pilot study for Index 2. Consequently, the evaluation results may not reflect the actual condition.

Table 6.1 Evaluation results on FY 2022 DISTRESS SCORE, CONDITION SCORE, and RIDE SCORE data (1.5 Sigma)

Pavement Type	Distress Type	Network Level	No. of Sections	Accuracy	Threshold
ACP	ACP PATCHING PCT	All	sections 96.17% FN 20 0.0259067 T- 24 0.0310881	95.60%	96.17%

			0 684 0.8860104 T+ 30 0.0388601 FP 14 0.0181347		
	ACP FAILURE QTY	All	sections 98.80% FN 1 0.0012953 T- 11 0.0142487 0 752 0.9740933 T+ 7 0.0090674 FP 1 0.0012953	99.74%	98.80%
	ACP BLOCK CRACKING PCT	All	sections 99.05% FN 0 0.0000000 T- 0 0.0000000 0 758 0.9818653 T+ 8 0.0103627 FP 6 0.0077720	99.22%	99.05%
	ACP ALLIGATOR CRACKING PCT	All	sections 95.27% FN 23 0.0297927 T- 78 0.1010363 0 569 0.7370466 T+ 94 0.1217617 FP 8 0.0103627	95.98%	95.27%
	ACP LONGITUDE CRACKING	All	sections 91.22% FN 65 0.0841969 T- 198 0.2564767 0 246 0.3186528 T+ 219 0.2836788 FP 44 0.0569948	85.88%	91.22%
	ACP TRANSVERSE CRACKING QTY	All	sections 94.87% FN 29 0.0375648 T- 114 0.1476684 0 595 0.7707254 T+ 31 0.0401554 FP 3 0.0038860	95.85%	94.87%
CRCP	CRCP SPALLED CRACKS QTY	All	sections % FN 0 0.0000000 T- 0 0.0000000 0 11 0.8461538 T+ 2 0.1538462 FP 0 0.0000000	100.00%	98.65%
	CRCP PUNCHOUT QTY	All	sections % FN 0 0.0000000 T- 0 0.0000000 0 13 1.0000000 T+ 0 0.0000000 FP 0 0.0000000	100.00%	96.37%

	CRCP ACP PATCHES QTY	All	 FN T- 0 T+ FP	sections 0 0 13 0 0	% 0.0000000 0.0000000 1.0000000 0.0000000 0.0000000	100.00%	97.86%
	CRCP PCC PATCHES QTY	All	 FN T- 0 T+ FP	sections 1 0 12 0 0	% 0.0769231 0.0000000 0.9230769 0.0000000 0.0000000	92.31%	94.50%
	CRCP AVG CRACK SPACING QTY	All	 FN T- 0 T+ FP	sections 0 0 1 7 5	% 0.0000000 0.0000000 0.0769231 0.5384615 0.3846154	61.54%	85.77%
JCP	JCP FAILED JNTS CRACKS QTY	All	 FN T- 0 T+ FP	sections 0 0 0 2 0	% 0.0000000 0.0000000 0.0000000 1.0000000 0.0000000	100.00%	95.88%
	JCP FAILURES QTY	All	 FN T- 0 T+ FP	sections 0 1 0 0 1	% 0.0000000 0.5000000 0.0000000 0.0000000 0.5000000	50.00%	97.67%
	JCP SHATTERED SLABS QTY	All	 FN T- 0 T+ FP	sections 0 0 2 0 0	% 0.0000000 0.0000000 1.0000000 0.0000000 0.0000000	100.00%	99.04%
	JCP SLABS WITH LONGITUDINAL CRACKS	All	 FN T- 0 T+ FP	sections 0 0 2 0 0	% 0.0000000 0.0000000 1.0000000 0.0000000 0.0000000	100.00%	93.83%
	JCP PCC PATCHES QTY	All	 FN T-	sections 0 1	% 0.0000000 0.5000000	100.00%	94.10%

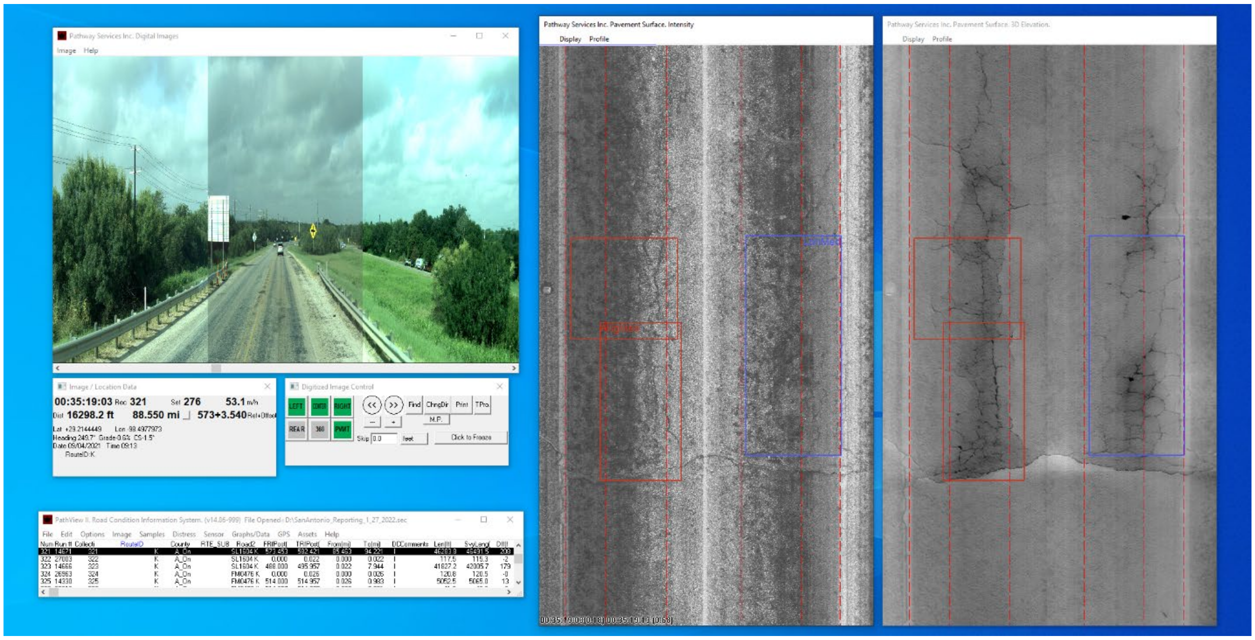
			0	0	0.0000000		
			T+	1	0.5000000		
			FP	0	0.0000000		
	JCP APPARENT JOINT SPACE	All		sections	%	0.00%	92.46%
			FN	2	1.0000000		
			T-	0	0.0000000		
			0	0	0.0000000		
			T+	0	0.0000000		
			FP	0	0.0000000		

A thorough examination was conducted for the pavement sections flagged for data quality issues. The findings of this visual check are summarized in Table 6.2. Upon reviewing the results, it becomes evident that a total of 105 pavement sections exhibit data quality issues as per the automated measurements. Additionally, 105 pavement sections were identified as having data quality issues based on the audit measurements. Remarkably, among these sections, 52 pavement sections were found to have data quality issues according to both the automated and audit measurements.

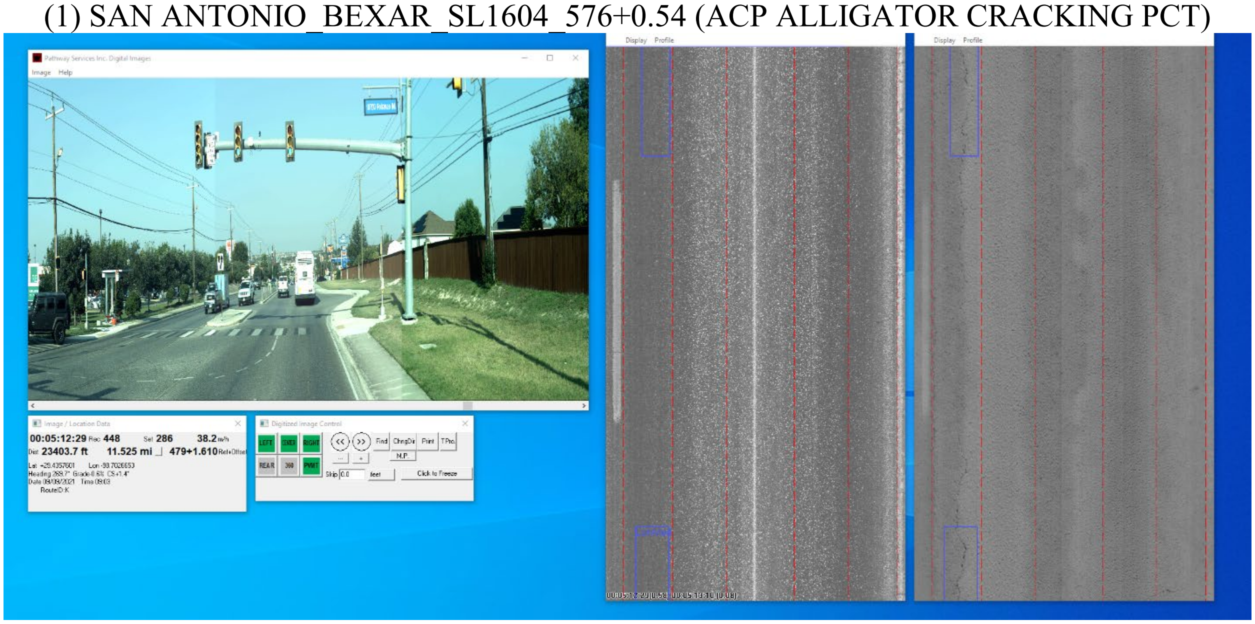
Table 6.2 Visual check results of pavement sections with data quality issues for San Antonio

Pavement Type	Distress Type	Pavement sections with quality issues		
		Automated	Audit	Both
ACP	ACP ALLIGATOR CRACKING PCT	13	9	3
	ACP FAILURE QTY	0	1	0
	ACP LONGITUDE CRACKING	70	56	40
	ACP PATCHING PCT	8	21	0
	ACP TRANSVERSE CRACKING QTY	12	16	9
	ACP BLOCK CRACKING PCT	2	2	0
Total		105	105	52

Figure 6.1 visually presents the outcomes of the automated data detection/identification process. It becomes apparent that certain discrepancies exist in the automated analysis. Specifically, FP cracks are observed in the detection of longitudinal cracking, while FN cracks are apparent in the identification of alligator cracking, transverse cracking, block cracking, and patching. Regarding failures, the audit rating incorrectly identifies overlay as failures. This misclassification requires attention and correction.

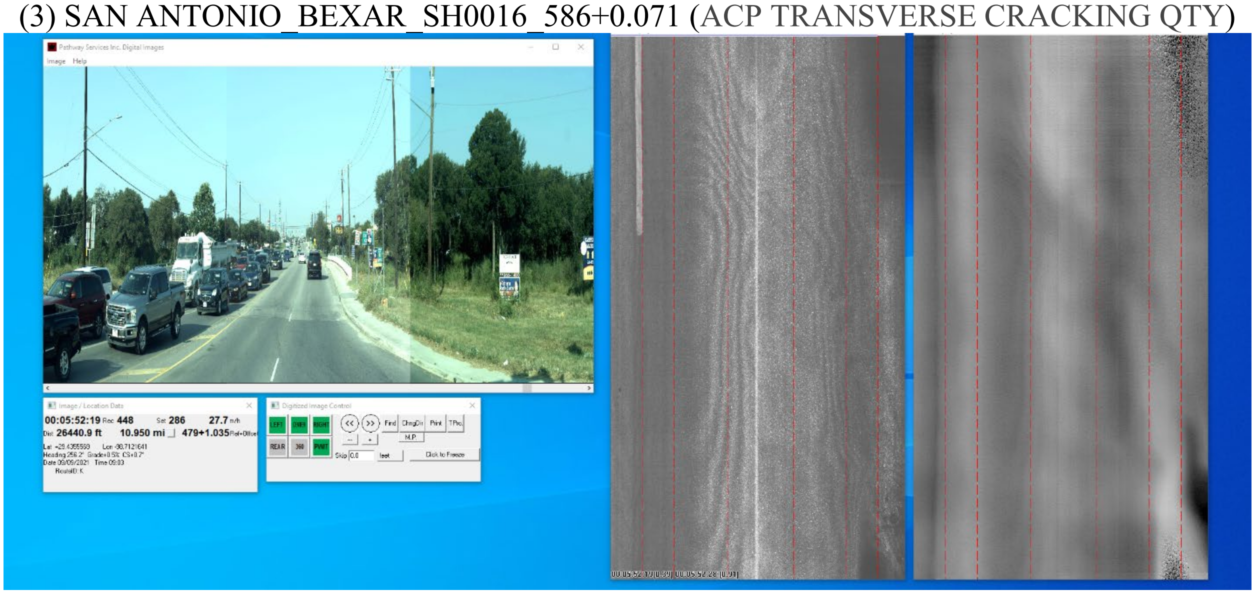


(1) SAN ANTONIO_BEXAR_SL1604_576+0.54 (ACP ALLIGATOR CRACKING PCT)

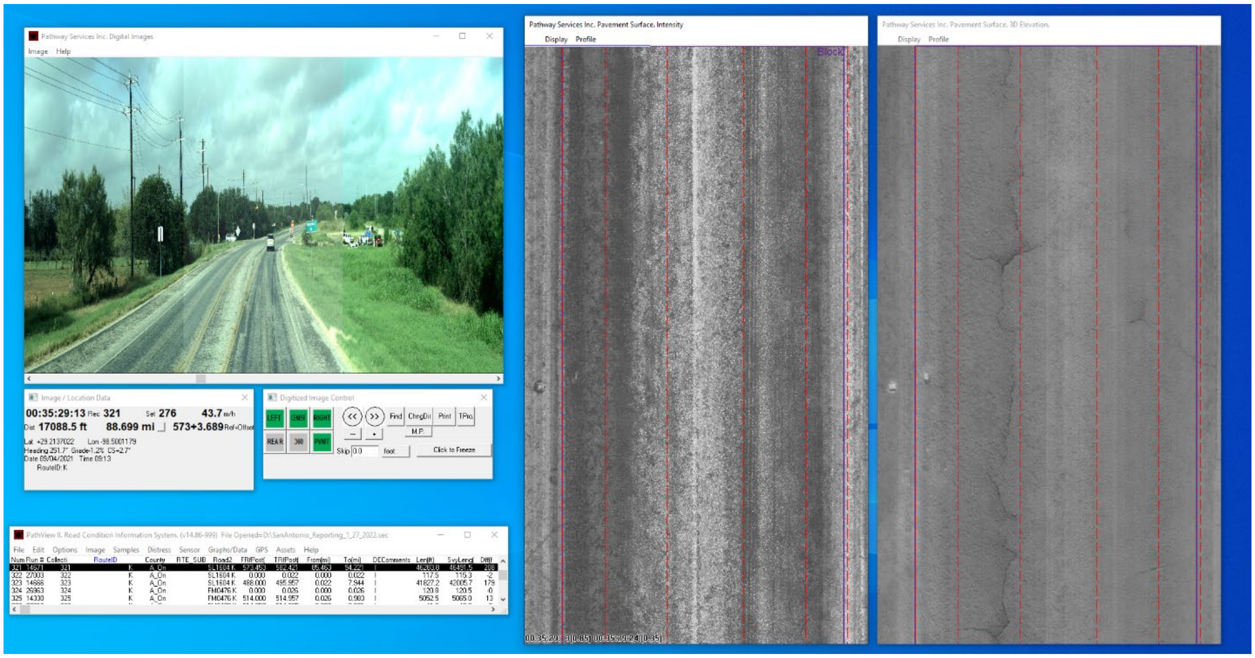


(2) SAN ANTONIO_BEXAR_FM1957_480+0.61 (ACP LONGITUDE CRACKING)

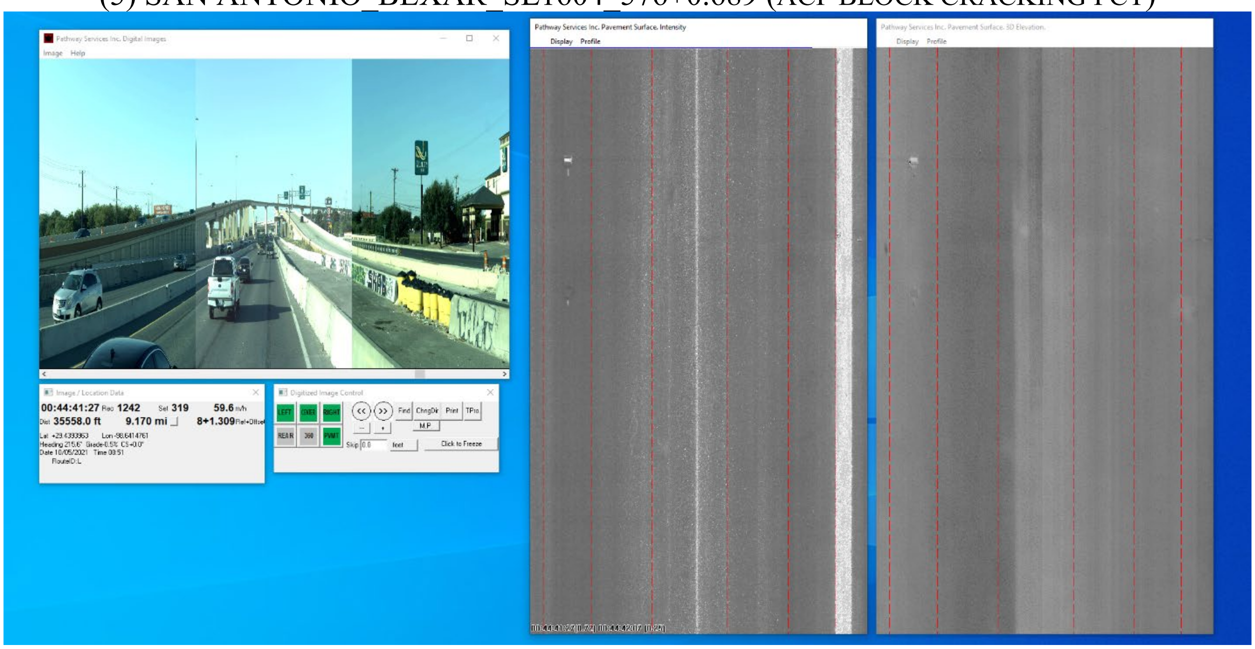
(3) SAN ANTONIO_BEXAR_SH0016_586+0.071 (ACP TRANSVERSE CRACKING QTY)



(4) SAN ANTONIO_BEXAR_IH0010_480+0.035 (ACP PATCHING PCT)



(5) SAN ANTONIO_BEXAR_SL1604_576+0.689 (ACP BLOCK CRACKING PCT)



(6) SAN ANTONIO_BEXAR_IH0410_9+0.309 (ACP FAILURE QTY)

Figure 6.1 Visualization of the Distress Detection/Identification Errors

6.1.2 Data Quality Check for Index #3

Table 6.3 presents the data quality check results for San Antonio on Index 3. For ACP, the table reveals the accuracy of Distress Score under SH network level, and US network level, indicating that they exhibit good data quality. Distress Score under FM network level shows slightly lower accuracy compared to the Index 3 threshold, which suggests a minor data quality concern. Distress Score under IH network level shows severe lower accuracy compared to the Index 3 threshold, which may suggest a strong data quality concern. The primary data quality issue identified in this analysis is FP. This finding implies that the annual data might tend to overestimate the severity of the Distress Score.

For CRCP, the table reveals the accuracy of the Distress Score under the IH network level, indicating that they exhibit good data quality. No CRCP Distress Score data under FM network level, SH network level, and US network level were included in the Index 3 polit study in FY 2022.

For JCP, only a few of the Distress Score data were involved in the pilot study for Index 3. Consequently, the evaluation results may not reflect the actual condition.

Table 6.3 Index 3 Data quality check for San Antonio

Pavement Type	Distress Type	Network Level	No. of Sections			Accuracy	Threshold
ACP	Distress Score	FM		sections	%	88.53%	91.07%
			FN	15	0.0537634		
			T-	38	0.1362007		
			0	181	0.6487455		
			T+	28	0.1003584		
			FP	17	0.0609319		
	Distress Score	SH		sections	%	92.74%	90.18%
			FN	7	0.0299145		
			T-	18	0.0769231		
			0	162	0.6923077		
			T+	37	0.1581197		
			FP	10	0.0427350		
	Distress Score	US		sections	%	96.61%	90.73%
			FN	0	0.0000000		
			T-	11	0.0932203		
			0	86	0.7288136		
			T+	17	0.1440678		
			FP	4	0.0338983		
	Distress Score	IH		sections	%	85.96%	90.66%
			FN	4	0.0350877		
			T-	14	0.1228070		
			0	64	0.5614035		
			T+	20	0.1754386		
			FP	12	0.1052632		
CRCP	Distress Score	FM		sections	%	Nan	90.30%
			FN	Nan	Nan		
			T-	Nan	Nan		
			0	Nan	Nan		
			T+	Nan	Nan		
			FP	Nan	Nan		

	Distress Score	SH	sections % FN Nan Nan T- Nan Nan 0 Nan Nan T+ Nan Nan FP Nan Nan	Nan	90.07%
	Distress Score	US	sections % FN Nan Nan T- Nan Nan 0 Nan Nan T+ Nan Nan FP Nan Nan	Nan	90.87%
	Distress Score	IH	sections % FN 0 0 T- 0 0 0 12 0.9230769 T+ 0 0 FP 1 0.0769231	92.30%	90.07%
JCP	Distress Score	FM	sections % FN 1 0.5000000 T- 0 0.0000000 0 1 0.5000000 T+ 0 0.0000000 FP 0 0.0000000	50.00%	86.73%
	Distress Score	SH	sections % FN Nan Nan T- Nan Nan 0 Nan Nan T+ Nan Nan FP Nan Nan	Nan	88.52%
	Distress Score	US	sections % FN Nan Nan T- Nan Nan 0 Nan Nan T+ Nan Nan FP Nan Nan	Nan	89.71%
	Distress Score	IH	sections % FN Nan Nan T- Nan Nan 0 Nan Nan T+ Nan Nan FP Nan Nan	Nan	86.00%

6.1.3 County-level analysis for Index 2 & Index 3

Table 6.4 presents the county-level analysis results for San Antonio. For ACP, there are 7 counties from the San Antonio district were selected for the data audit. The table reveals the accuracy of Comal, Guadalupe, Kerr, and McMullen, indicating that they exhibit good data quality. Wilson shows slightly lower accuracy compared to the Index 3 threshold, which suggests a minor data quality concern. Bandera and Bexar show severe lower accuracy compared to the Index 3 threshold, which may suggest a strong data quality concern. The primary data quality issue identified in this analysis is FP. This finding implies that the annual data might tend to slightly underestimate the severity of overall pavement condition.

For CRCP, there are 1 county from San Antonio district were selected for the data audit. The table reveals the accuracy of Bexar, indicating that they exhibit good data quality.

For JCP, only a few of the Distress Score data was involved in the pilot study for Index 2 and Index 3. Consequently, the evaluation results may not reflect the actual condition.

Table 6.4 County-level data quality check for San Antonio

Distress Type	County	No. of Sections			Accuracy	Threshold	Decisions
ACP	10 - BANDERA		sections	%	78.43%	90.71%	Fail
		FN	3	0.0697674			
		T-	6	0.1395349			
		0	27	0.6279070			
		T+	7	0.1627907			
		FP	0	0.0000000			
	15 - BEXAR		sections	%	82.28%	90.71%	Fail
		FN	8	0.0269360			
		T-	34	0.1144781			
		0	182	0.6127946			
		T+	44	0.1481481			
		FP	29	0.0976431			
	46 - COMAL		sections	%	94.74%	90.71%	Pass
		FN	0	0.0000000			
		T-	3	0.0394737			
		0	53	0.6973684			
		T+	16	0.2105263			
		FP	4	0.0526316			
	95 - GUADALUPE		sections	%	91.07%	90.71%	Pass
		FN	4	0.0357143			
		T-	17	0.1517857			
		0	70	0.6250000			
		T+	15	0.1339286			
		FP	6	0.0535714			
	133 - KERR		sections	%	93.18%	90.71%	Pass

		FN	3	0.0340909			
		T-	3	0.0340909			
		0	69	0.7840909			
		T+	10	0.1136364			
		FP	3	0.0340909			
	162 - MCMULLEN		sections	%	100.00%	90.71%	Pass
		FN	0	0.0000000			
		T-	2	0.0500000			
		0	35	0.8750000			
		T+	3	0.0750000			
		FP	0	0.0000000			
	247 - WILSON		sections	%	89.89%	90.71%	Fail
		FN	8	0.0898876			
		T-	16	0.1797753			
		0	57	0.6404494			
		T+	7	0.0786517			
		FP	1	0.0112360			
CRCP	15 - BEXAR		sections	%	92.31%	89.98%	Pass
		FN	0	0.0000000			
		T-	0	0.0000000			
		0	12	0.9230769			
		T+	0	0.0000000			
		FP	1	0.0769231			
JCP	46 - COMAL		sections	%	50.00%	88.48%	Fail
		FN	1	0.5000000			
		T-	0	0.0000000			
		0	1	0.5000000			
		T+	0	0.0000000			
		FP	0	0.0000000			

6.1.4 Data Quality Check for Index #4

Table 6.5 presents the data quality check results for San Antonio on Index 4. For ACP, the table reveals the accuracy of ACP Patching, ACP Failure, ACP Alligator Cracking, ACP Longitude Cracking, and ACP Transverse Cracking, indicating that they exhibit good data quality. There are similar amounts of pavement sections with FP and FN. This finding implies that both unexpected decreasing and increasing in distress measurements could be equally observed.

For CRCP, the table reveals the accuracy of CRCP Spalled Cracks, CRCP Punchout, CRCP ACP Patches, CRCP PCC Patches, and CRCP Avg Crack Spacing, indicating that they exhibit good data quality.

For JCP, only a few of the pavement section data were involved in the pilot study for Index 4. Consequently, the evaluation results may not reflect the actual condition.

Table 6.5 Index 4 data quality check for San Antonio

Pavement Type	Distress Type	Network Level	No. of Sections			Precision	Threshold
ACP	ACP PATCHING PCT	All		sections	%	98.71%	95.92%
			FN	200	0.0204351		
			T-	435	0.0442253		
			0	8716	0.8859292		
			T+	361	0.0367019		
			FP	124	0.0127084		
	ACP FAILURE QTY	All		sections	%	98.72%	98.34%
			FN	71	0.0072184		
			T-	123	0.0125051		
			0	9388	0.954453		
			T+	129	0.0131151		
			FP	125	0.0127084		
	ACP BLOCK CRACKING PCT	All		sections	%	Nan	95.68%
			FN	Nan	Nan		
			T-	Nan	Nan		
			0	Nan	Nan		
			T+	Nan	Nan		
			FP	Nan	Nan		
	ACP ALLIGATOR CRACKING PCT	All		sections	%	98.05%	94.12%
			FN	277	0.0285685		
			T-	1318	0.1339976		
			0	6700	0.6807645		
			T+	1355	0.1377593		
			FP	186	0.0189101		
	ACP LONGITUDE CRACKING	All		sections	%	95.00%	90.13%
			FN	381	0.0390403		
			T-	2457	0.2497967		
			0	3311	0.3362139		
			T+	3214	0.3267588		
			FP	473	0.0481903		
	ACP TRANSVERSE CRACKING QTY	All		sections	%	99.03%	94.00%
			FN	98	0.0099634		
			T-	228	0.0231802		
			0	8707	0.8852176		
			T+	709	0.0720821		
			FP	94	0.0095567		
CRCP	CRCP SPALLED CRACKS QTY	All		sections	%	100.00%	98.68%
			FN	0	0.0000000		
			T-	0	0.0000000		
			0	79	0.9753086		
			T+	2	0.0246914		
			FP	0	0.0000000		

	CRCP PUNCHOUT QTY	All	sections % FN 1 0.0123457 T- 0 0.0000000 0 80 0.9876543 T+ 0 0.0000000 FP 0 0.0000000	100.00%	96.47%
	CRCP ACP PATCHES QTY	All	sections % FN 1 0.0123457 T- 1 0.0123457 0 79 0.9753086 T+ 0 0.0000000 FP 0 0.0000000	100.00%	97.61%
	CRCP PCC PATCHES QTY	All	sections % FN 0 0.0000000 T- 1 0.0123457 0 71 0.8765432 T+ 8 0.0987654 FP 1 0.0123457	98.75%	94.29%
	CRCP AVG CRACK SPACING QTY	All	sections % FN 20 0.2469136 T- 22 0.2716049 0 17 0.2098765 T+ 19 0.2345679 FP 3 0.0370370	92.31%	86.18%
JCP	JCP FAILED JNTS CRACKS QTY	All	sections % FN 0 0.0000000 T- 1 0.2500000 0 1 0.2500000 T+ 2 0.5000000 FP 0 0.0000000	100.00%	95.07%
	JCP FAILURES QTY	All	sections % FN 0 0.0000000 T- 2 0.5000000 0 1 0.2500000 T+ 0 0.0000000 FP 1 0.2500000	50.00%	93.73%
	JCP SHATTERED SLABS QTY	All	sections % FN 1 0.2500000 T- 0 0.0000000 0 3 0.7500000 T+ 0 0.0000000 FP 0 0.0000000	100.00%	96.53%
	JCP SLABS WITH	All	sections % FN 0 0.0000000 T- 0 0.0000000	100.00%	93.72%

	LONGITUDINAL CRACKS		0 4 1.0000000 T+ 0 0.0000000 FP 0 0.0000000		
	JCP PCC PATCHES QTY	All	sections % FN 0 0.0000000 T- 2 0.5000000 0 1 0.2500000 T+ 1 0.2500000 FP 0 0.0000000	100.00%	93.33%
	JCP APPARENT JOINT SPACE	All	sections % FN 0 0.0000000 T- 1 0.2500000 0 1 0.2500000 T+ 2 0.5000000 FP 0 0.0000000	100.00%	94.13%

6.1.5 Data Quality Check for Index #5

Table 6.6 presents the ACP data quality check results for San Antonio on Index 5. For Distress Score, the table reveals the precession of Distress Score under FM network level, and US network level, indicating that they exhibit good data quality. Distress Score under SH network level and IH network level show slightly lower precision compared to the Index 5 threshold, which suggests a minor data quality concern. The primary data quality issue identified in this analysis is FP. This finding implies that the dominant pavement sections with data quality issues show an unexpected increase in distress measurements.

For Condition Score, the table reveals the accuracy of Condition Score under the FM network level, indicating that they exhibit good data quality. Condition Score under SH network level, US network level, and IH network level show slightly lower precision compared to the Index 5 threshold, which suggests a minor data quality concern. The primary data quality issue identified in this analysis is FP. This finding implies that the dominant pavement sections with data quality issues show an unexpected increase in distress measurements.

For Ride Score, Ride Score under FM network level shows slightly lower precision compared to the Index 5 threshold, which suggests a minor data quality concern. Ride Score under SH network level, US network level, and IH network level show severe lower precision compared to the Index 5 threshold, which may suggest a strong data quality concern. The primary data quality issue identified in this analysis is FP. This finding implies that the dominant pavement sections with data quality issues show an unexpected increase in distress measurements.

Table 6.6 Index 5 Data quality check for San Antonio (ACP)

Distress Type	Network Level	No. of Sections	Precision	Threshold
Distress Score	FM	sections % FN 175 0.0443487	89.98%	88.52%

		T-	1705	0.4320831		
		0	757	0.1918398		
		T+	1102	0.2792701		
		FP	207	0.0524582		
	SH		sections	%	85.56%	89.24%
		FN	109	0.0459334		
		T-	872	0.3674673		
		0	642	0.2705436		
		T+	549	0.2313527		
		FP	201	0.0847029		
	US		sections	%	89.59%	89.46%
		FN	50	0.0462535		
		T-	426	0.3940796		
		0	271	0.2506938		
		T+	271	0.2506938		
		FP	63	0.0582794		
	IH		sections	%	87.58%	88.66%
		FN	125	0.0513136		
		T-	830	0.3407225		
		0	706	0.2898194		
		T+	591	0.2426108		
		FP	184	0.0755337		
CONDITION SCORE	FM		sections	%	89.80%	89.39%
		FN	180	0.0457433		
		T-	1706	0.4335451		
		0	662	0.1682338		
		T+	1178	0.2993647		
		FP	209	0.0531131		
	SH		sections	%	86.15%	89.68%
		FN	105	0.0443038		
		T-	879	0.3708861		
		0	588	0.2481013		
		T+	606	0.2556962		
		FP	192	0.0810127		
	US		sections	%	87.89%	89.54%
		FN	51	0.0471785		
		T-	419	0.3876041		
		0	260	0.2405180		
		T+	277	0.2562442		
		FP	74	0.0684551		
	IH		sections	%	88.04%	88.99%
		FN	126	0.0517241		
		T-	863	0.3542693		
		0	651	0.2672414		
		T+	623	0.2557471		

		FP	173	0.0710181		
RIDE SCORE	FM		sections	%	88.20%	92.97%
		FN	105	0.0266836		
		T-	1940	0.4930114		
		0	101	0.0256671		
		T+	1566	0.3979670		
		FP	223	0.0566709		
	SH		sections	%	86.24%	93.41%
		FN	56	0.0236287		
		T-	1108	0.4675105		
		0	61	0.0257384		
		T+	979	0.4130802		
		FP	166	0.0700422		
	US		sections	%	82.23%	92.66%
		FN	23	0.0212766		
		T-	546	0.5050879		
		0	22	0.0203515		
		T+	399	0.3691027		
		FP	91	0.0841813		
	IH		sections	%	81.86%	92.19%
		FN	98	0.0402299		
		T-	993	0.4076355		
		0	38	0.0155993		
		T+	1063	0.4363711		
		FP	244	0.1001642		

Table 6.7 presents the CRCP data quality check results for San Antonio on Index 5. Only a few of the pavement section data were involved in the pilot study for Index 5. Consequently, the evaluation results may not reflect the actual condition.

Table 6.7 Data quality check for San Antonio (CRCP)

Distress Type	Network Level	No. of Sections			Precision	Threshold
Distress Score	FM		sections	%	Nan	87.68%
		FN	Nan	Nan		
		T-	Nan	Nan		
		0	Nan	Nan		
		T+	Nan	Nan		
		FP	Nan	Nan		
	SH		sections	%	85.71%	87.88%
		FN	0	0.0000000		
		T-	0	0.0000000		
		0	6	0.8571429		
		T+	0	0.0000000		
		FP	1	0.1428571		
	US		sections	%	100.00%	90.95%

		FN	0	0.0000000		
		T-	0	0.0000000		
		0	5	1.0000000		
		T+	0	0.0000000		
		FP	0	0.0000000		
	IH		sections	%	98.33%	90.23%
		FN	1	0.0144928		
		T-	8	0.1159420		
		0	57	0.8260870		
		T+	2	0.0289855		
		FP	1	0.0144928		
CONDITION SCORE	FM		sections	%	Nan	87.87%
		FN	Nan	Nan		
		T-	Nan	Nan		
		0	Nan	Nan		
		T+	Nan	Nan		
		FP	Nan	Nan		
	SH		sections	%	80.00%	87.67%
		FN	0	0.0000000		
		T-	2	0.2857143		
		0	4	0.5714286		
		T+	0	0.0000000		
		FP	1	0.1428571		
	US		sections	%	80.00%	90.57%
		FN	0	0.0000000		
		T-	0	0.0000000		
		0	3	0.6000000		
		T+	1	0.2000000		
		FP	1	0.2000000		
	IH		sections	%	94.34%	89.63%
		FN	1	0.0144928		
		T-	15	0.2173913		
		0	46	0.6666667		
		T+	4	0.0579710		
		FP	3	0.0434783		
RIDE SCORE	FM		sections	%	Nan	93.20%
		FN	Nan	Nan		
		T-	Nan	Nan		
		0	Nan	Nan		
		T+	Nan	Nan		
		FP	Nan	Nan		
	SH		sections	%	100.00%	92.82%
		FN	0	0.0000000		
		T-	7	1.0000000		
		0	0	0.0000000		

		T+	0	0.0000000		
		FP	0	0.0000000		
	US		sections	%	50.00%	91.63%
		FN	0	0.0000000		
		T-	1	0.2000000		
		0	0	0.0000000		
		T+	2	0.4000000		
		FP	2	0.4000000		
	IH		sections	%	90.48%	92.80%
		FN	1	0.0144928		
		T-	26	0.3768116		
		0	8	0.1159420		
		T+	30	0.4347826		
		FP	4	0.0579710		

Table 6.8 presents the JCP data quality check results for San Antonio on Index 5. Only a few of the pavement section data were involved in the pilot study for Index 5. Consequently, the evaluation results may not reflect the actual condition.

Table 6.8 Index 5 Data quality check for San Antonio (JCP)

Distress Type	Network Level	No. of Sections			Precision	Threshold
Distress Score	FM		sections	%	100.00%	81.73%
		FN	1	0.2500000		
		T-	0	0.0000000		
		0	2	0.5000000		
		T+	1	0.2500000		
		FP	0	0.0000000		
	SH		sections	%	Nan	87.74%
		FN	Nan	Nan		
		T-	Nan	Nan		
		0	Nan	Nan		
		T+	Nan	Nan		
		FP	Nan	Nan		
	US		sections	%	Nan	91.11%
		FN	Nan	Nan		
		T-	Nan	Nan		
		0	Nan	Nan		
		T+	Nan	Nan		
		FP	Nan	Nan		
	IH		sections	%	Nan	88.56%
		FN	Nan	Nan		
		T-	Nan	Nan		
		0	Nan	Nan		
		T+	Nan	Nan		
		FP	Nan	Nan		

CONDITION SCORE	FM	sections % FN 1 0.2500000 T- 2 0.5000000 0 0 0.0000000 T+ 0 0.0000000 FP 1 0.2500000	50.00%	87.50%
	SH	sections % FN Nan Nan T- Nan Nan 0 Nan Nan T+ Nan Nan FP Nan Nan	Nan	88.31%
	US	sections % FN Nan Nan T- Nan Nan 0 Nan Nan T+ Nan Nan FP Nan Nan	Nan	87.41%
	IH	sections % FN Nan Nan T- Nan Nan 0 Nan Nan T+ Nan Nan FP Nan Nan	Nan	86.57%
RIDE SCORE	FM	sections % FN 0 0.0000000 T- 3 0.7500000 0 0 0.0000000 T+ 0 0.0000000 FP 1 0.2500000	75.00%	88.46%
	SH	sections % FN Nan Nan T- Nan Nan 0 Nan Nan T+ Nan Nan FP Nan Nan	Nan	94.16%
	US	sections % FN Nan Nan T- Nan Nan 0 Nan Nan T+ Nan Nan FP Nan Nan	Nan	86.67%
	IH	sections % FN Nan Nan T- Nan Nan	Nan	93.53%

		0 Nan Nan T+ Nan Nan FP Nan Nan			

6.1.6 County-level analysis for Index 4 & Index 5

Table 6.9 presents the county-level analysis results of the Distress Score for San Antonio. For ACP, there are 15 counties from the San Antonio district were selected for the data audit. The table reveals the accuracy of Guadalupe, Wilson, Comal, Uvalde, Bandera, Kerr, Gillespie, and La Salle, indicating that they exhibit good data quality. Bexar, Atascosa, Mcmullen, and Kendall show slightly lower accuracy compared to the Index 5 threshold, which suggests a minor data quality concern. Frio, Medina, and Caldwell show severe lower accuracy compared to the Index 5 threshold, which may suggest a strong data quality concern. The primary data quality issue identified in this analysis is FP. This finding implies that the dominant pavement sections with data quality issues show an unexpected increase in Distress Score measurements.

For CRCP, there are 3 counties from the San Antonio district were selected for the data audit. The table reveals the accuracy of Bexar, Kendall, and Uvalde, indicating that they exhibit good data quality. There are similar amounts of pavement sections with FP and FN. This finding implies that both unexpected decreasing and increasing in Distress Score measurements could be equally observed. However, the number of Distress Score data involved in the pilot study for Index 4 and Index 5 may not be sufficient. Consequently, the evaluation results may not reflect the actual condition.

For JCP, only a few of the Distress Score data was involved in the pilot study for Index 4 and Index 5. Consequently, the evaluation results may not reflect the actual condition.

Table 6.9 County-level data quality check for San Antonio (Distress Score)

Distress Type	**County**	**No. of Sections**	**Accuracy**	**Threshold**	**Determination**
ACP	95 - GUADALUPE	sections % FN 39 0.0421622 T- 421 0.4551351 0 164 0.1772973 T+ 250 0.2702703 FP 51 0.0551351	90.27%	89.04%	Pass
	247 - WILSON	sections % FN 24 0.0345324 T- 272 0.3913669 0 200 0.2877698 T+ 161 0.2316547 FP 38 0.0546763	91.08%	89.04%	Pass
	46 - COMAL	sections % FN 17 0.0268562	92.10%	89.04%	Pass

		T-	208	0.3285940			
		0	186	0.2938389			
		T+	189	0.2985782			
		FP	33	0.0521327			
	15 - BEXAR		sections	%	88.47%	89.04%	Fail
		FN	122	0.0473419			
		T-	980	0.3802872			
		0	772	0.2995731			
		T+	528	0.2048894			
		FP	175	0.0679084			
	232 - UVALDE		sections	%	93.34%	89.04%	Pass
		FN	12	0.0169972			
		T-	238	0.3371105			
		0	171	0.2422096			
		T+	250	0.3541076			
		FP	35	0.0495751			
	83 - FRIO		sections	%	82.31%	89.04%	Fail
		FN	53	0.0694626			
		T-	251	0.3289646			
		0	176	0.2306684			
		T+	201	0.2634338			
		FP	82	0.1074705			
	7 - ATASCOSA		sections	%	85.35%	89.04%	Fail
		FN	74	0.0747475			
		T-	451	0.4555556			
		0	224	0.2262626			
		T+	170	0.1717172			
		FP	71	0.0717172			
	163 - MEDINA		sections	%	82.73%	89.04%	Fail
		FN	60	0.0803213			
		T-	325	0.4350736			
		0	120	0.1606426			
		T+	173	0.2315930			
		FP	69	0.0923695			
	10 - BANDERA		sections	%	96.10%	89.04%	Pass
		FN	6	0.0155844			
		T-	144	0.3740260			
		0	91	0.2363636			
		T+	135	0.3506494			
		FP	9	0.0233766			
	162 - MCMULLEN		sections	%	87.18%	89.04%	Fail
		FN	14	0.0448718			
		T-	128	0.4102564			
		0	72	0.2307692			
		T+	72	0.2307692			

		FP	26	0.0833333			
	131 - KENDALL		sections	%	86.59%	89.04%	Fail
		FN	24	0.0585366			
		T-	154	0.3756098			
		0	55	0.1341463			
		T+	146	0.3560976			
		FP	31	0.0756098			
	133 - KERR		sections	%	93.12%	89.04%	Pass
		FN	14	0.0214067			
		T-	254	0.3883792			
		0	125	0.1911315			
		T+	230	0.3516820			
		FP	31	0.0474006			
	28 - CALDWELL		sections	%	63.64%	89.04%	Fail
		FN	0	0.0000000			
		T-	0	0.0000000			
		0	6	0.5454545			
		T+	1	0.0909091			
		FP	4	0.3636364			
	87 - GILLESPIE		sections	%	100.00%	89.04%	Pass
		FN	0	0.0000000			
		T-	4	0.1666667			
		0	14	0.5833333			
		T+	6	0.2500000			
		FP	0	0.0000000			
	142 - LA SALLE		sections	%	100.00%	89.04%	Pass
		FN	0	0.0000000			
		T-	3	0.7500000			
		0	0	0.0000000			
		T+	1	0.2500000			
		FP	0	0.0000000			
CRCP	15 - BEXAR		sections	%	96.20%	89.65%	Pass
		FN	1	0.0126582			
		T-	8	0.1012658			
		0	66	0.8354430			
		T+	2	0.0253165			
		FP	2	0.0253165			
	131 - KENDALL		sections	%	100.00%	89.65%	Pass
		FN	0	0.0000000			
		T-	0	0.0000000			
		0	1	1.0000000			
		T+	0	0.0000000			
		FP	0	0.0000000			
	232 - UVALDE		sections	%	100.00%	89.65%	Pass
		FN	0	0.0000000			

		T-	0	0.0000000			
		0	1	1.0000000			
		T+	0	0.0000000			
		FP	0	0.0000000			
JCP	46 - COMAL		sections	%	75.00%	86.80%	Fail
		FN	1	0.2500000			
		T-	0	0.0000000			
		0	2	0.5000000			
		T+	1	0.2500000			
		FP	0	0.0000000			

Table 6.10 presents the county-level analysis results of Condition Score for San Antonio. For ACP, there are 15 counties from the San Antonio district were selected for the data audit. The table reveals the accuracy of Guadalupe, Wilson, Comal, Uvalde, Bandera, Kerr, Gillespie, and La Salle, indicating that they exhibit good data quality. Bexar, Atascosa, Mcmullen, and Kendall show slightly lower accuracy compared to the Index 5 threshold, which suggests a minor data quality concern. Frio, Medina, and Caldwell show severe lower accuracy compared to the Index 5 threshold, which may suggest a strong data quality concern. The primary data quality issue identified in this analysis is FP. This finding implies that the dominant pavement sections with data quality issues show an unexpected increase in Condition Score measurements.

For CRCP, there are 3 counties from the San Antonio district were selected for the data audit. The table reveals the accuracy of Bexar, Kendall, and Uvalde, indicating that they exhibit good data quality. There are similar amounts of pavement sections with FP and FN. This finding implies that both unexpected decreasing and increasing on Condition Score measurements could be equally observed. However, the number of Condition Score data involved in the pilot study for Index 4 and Index 5 may not be sufficient. Consequently, the evaluation results may not reflect the actual condition.

For JCP, only a few of the Condition Score data was involved in the pilot study for Index 4 and Index 5. Consequently, the evaluation results may not reflect the actual condition.

Table 6.10 County-level data quality check for San Antonio (Condition Score)

Distress Type	County	No. of Sections			Accuracy	Threshold	Determination
ACP	95 - GUADALUPE		sections	%	88.11%	89.72%	Fail
		FN	39	0.0422535			
		T-	422	0.4572048			
		0	149	0.1614301			
		T+	244	0.2643554			
		FP	69	0.0747562			
	247 - WILSON		sections	%	90.36%	89.72%	Pass
		FN	29	0.0417867			
		T-	269	0.3876081			
		0	193	0.2780980			
		T+	166	0.2391931			

		FP	37	0.0533141			
	46 - COMAL		sections	%	88.15%	89.72%	Fail
		FN	30	0.0474684			
		T-	198	0.3132911			
		0	170	0.2689873			
		T+	190	0.3006329			
		FP	44	0.0696203			
	15 - BEXAR		sections	%	88.86%	89.72%	Fail
		FN	113	0.0438665			
		T-	1020	0.3959627			
		0	689	0.2674689			
		T+	581	0.2255435			
		FP	173	0.0671584			
	232 - UVALDE		sections	%	93.06%	89.72%	Pass
		FN	18	0.0255319			
		T-	237	0.3361702			
		0	153	0.2170213			
		T+	267	0.3787234			
		FP	30	0.0425532			
	83 - FRIO		sections	%	84.40%	89.72%	Fail
		FN	47	0.0615990			
		T-	239	0.3132372			
		0	173	0.2267366			
		T+	232	0.3040629			
		FP	72	0.0943644			
	7 - ATASCOSA		sections	%	87.27%	89.72%	Fail
		FN	51	0.0517241			
		T-	446	0.4523327			
		0	201	0.2038540			
		T+	217	0.2200811			
		FP	71	0.0720081			
	163 - MEDINA		sections	%	84.20%	89.72%	Fail
		FN	60	0.0803213			
		T-	340	0.4551539			
		0	111	0.1485944			
		T+	178	0.2382865			
		FP	58	0.0776439			
	10 - BANDERA		sections	%	93.77%	89.72%	Pass
		FN	17	0.0442708			
		T-	158	0.4114583			
		0	74	0.1927083			
		T+	129	0.3359375			
		FP	6	0.0156250			
	162 - MCMULLEN		sections	%	87.18%	89.72%	Fail
		FN	14	0.0448718			

		T- 115 0.3685897 0 66 0.2115385 T+ 91 0.2916667 FP 26 0.0833333			
	131 - KENDALL	sections % FN 28 0.0684597 T- 148 0.3618582 0 55 0.1344743 T+ 148 0.3618582 FP 30 0.0733496	85.61%	89.72%	Fail
	133 - KERR	sections % FN 16 0.0245399 T- 269 0.4125767 0 107 0.1641104 T+ 231 0.3542945 FP 29 0.0444785	92.81%	89.72%	Pass
	28 - CALDWELL	sections % FN 0 0.0000000 T- 0 0.0000000 0 6 0.5454545 T+ 2 0.1818182 FP 3 0.2727273	72.73%	89.72%	Fail
	87 - GILLESPIE	sections % FN 0 0.0000000 T- 4 0.1666667 0 14 0.5833333 T+ 6 0.2500000 FP 0 0.0000000	100.00%	89.72%	Pass
	142 - LA SALLE	sections % FN 0 0.0000000 T- 2 0.5000000 0 0 0.0000000 T+ 2 0.5000000 FP 0 0.0000000	100.00%	89.72%	Pass
CRCP	15 - BEXAR	sections % FN 1 0.0126582 T- 16 0.2025316 0 52 0.6582278 T+ 5 0.0632911 FP 5 0.0632911	92.41%	89.13%	Pass
	131 - KENDALL	sections % FN 0 0.0000000 T- 0 0.0000000 0 1 1.0000000 T+ 0 0.0000000	100.00%	89.13%	Pass

		FP 0 0.0000000			
	232 - UVALDE	sections % FN 0 0.0000000 T- 1 1.0000000 0 0 0.0000000 T+ 0 0.0000000 FP 0 0.0000000	100.00%	89.13%	Pass
JCP	46 - COMAL	sections % FN 1 0.2500000 T- 2 0.5000000 0 0 0.0000000 T+ 0 0.0000000 FP 1 0.2500000	50.00%	86.63%	Fail

Table 6.11 presents the county-level analysis results of Ride Score for San Antonio. For ACP, there are 15 counties from the San Antonio district were selected for the data audit. The table reveals the accuracy of Atascosa and La Salle, indicating that they exhibit good data quality. Guadalupe, Bexar, Frio, Medina, and Bandera show slightly lower accuracy compared to the Index 5 threshold, which suggests a minor data quality concern. Wilson, Comal, Uvalde, Mcmullen, Kendall, Kerr, Caldwell, and Gillespie show severe lower accuracy compared to the Index 5 threshold, which may suggest a strong data quality concern. The primary data quality issue identified in this analysis is FP. This finding implies that the dominant pavement sections with data quality issues show an unexpected increase in Ride Score measurements.

For CRCP, there are 3 counties from the San Antonio district were selected for the data audit. The table reveals the accuracy of Kendall and Uvalde, indicating that they exhibit good data quality. Bexar shows slightly lower accuracy compared to the Index 5 threshold, which suggests a minor data quality concern. The primary data quality issue identified in this analysis is FP. This finding implies that the dominant pavement sections with data quality issues show an unexpected increase in Ride Score measurements. However, the number of Ride Score data involved in the pilot study for Index 4 and Index 5 may not be sufficient. Consequently, the evaluation results may not reflect the actual condition.

For JCP, only a few of the Ride Score data was involved in the pilot study for Index 4 and Index 5. Consequently, the evaluation results may not reflect the actual condition.

Table 6.11 County-level data quality check for San Antonio (Ride Score)

Distress Type	County	No. of Sections	Accuracy	Threshold	Determination
ACP	95 - GUADALUPE	sections % FN 34 0.0368364 T- 490 0.5308776 0 23 0.0249187 T+ 341 0.3694475 FP 35 0.0379198	92.32%	92.99%	Fail

	247 - WILSON		sections	%	86.47%	92.99%	Fail
		FN	17	0.0244957			
		T-	279	0.4020173			
		0	24	0.0345821			
		T+	298	0.4293948			
		FP	76	0.1095101			
	46 - COMAL		sections	%	83.41%	92.99%	Fail
		FN	35	0.0553797			
		T-	270	0.4272152			
		0	15	0.0237342			
		T+	243	0.3844937			
		FP	69	0.1091772			
	15 - BEXAR		sections	%	92.05%	92.99%	Fail
		FN	70	0.0271739			
		T-	1191	0.4623447			
		0	60	0.0232919			
		T+	1121	0.4351708			
		FP	134	0.0520186			
	232 - UVALDE		sections	%	86.97%	92.99%	Fail
		FN	18	0.0255319			
		T-	377	0.5347518			
		0	23	0.0326241			
		T+	214	0.3035461			
		FP	73	0.1035461			
	83 - FRIO		sections	%	88.34%	92.99%	Fail
		FN	15	0.0196592			
		T-	264	0.3460026			
		0	8	0.0104849			
		T+	402	0.5268676			
		FP	74	0.0969856			
	7 - ATASCOSA		sections	%	95.66%	92.99%	Pass
		FN	10	0.0101420			
		T-	464	0.4705882			
		0	22	0.0223124			
		T+	461	0.4675456			
		FP	29	0.0294118			
	163 - MEDINA		sections	%	89.83%	92.99%	Fail
		FN	34	0.0455154			
		T-	418	0.5595716			
		0	15	0.0200803			
		T+	238	0.3186078			
		FP	42	0.0562249			
	10 - BANDERA		sections	%	90.91%	92.99%	Fail
		FN	12	0.0312500			
		T-	235	0.6119792			

		0	7	0.0182292			
		T+	108	0.2812500			
		FP	22	0.0572917			
	162 - MCMULLEN		sections	%	85.90%	92.99%	Fail
		FN	8	0.0256410			
		T-	96	0.3076923			
		0	3	0.0096154			
		T+	169	0.5416667			
		FP	36	0.1153846			
	131 - KENDALL		sections	%	85.85%	92.99%	Fail
		FN	20	0.0488998			
		T-	201	0.4914425			
		0	10	0.0244499			
		T+	141	0.3447433			
		FP	37	0.0904645			
	133 - KERR		sections	%	85.93%	92.99%	Fail
		FN	9	0.0138037			
		T-	302	0.4631902			
		0	12	0.0184049			
		T+	248	0.3803681			
		FP	81	0.1242331			
	28 - CALDWELL		sections	%	54.55%	92.99%	Fail
		FN	0	0.0000000			
		T-	0	0.0000000			
		0	0	0.0000000			
		T+	6	0.5454545			
		FP	5	0.4545455			
	87 - GILLESPIE		sections	%	54.17%	92.99%	Fail
		FN	0	0.0000000			
		T-	0	0.0000000			
		0	0	0.0000000			
		T+	13	0.5416667			
		FP	11	0.4583333			
	142 - LA SALLE		sections	%	100.00%	92.99%	Pass
		FN	0	0.0000000			
		T-	0	0.0000000			
		0	0	0.0000000			
		T+	4	1.0000000			
		FP	0	0.0000000			
CRCP	15 - BEXAR		sections	%	91.14%	92.71%	Fail
		FN	1	0.0126582			
		T-	33	0.4177215			
		0	7	0.0886076			
		T+	32	0.4050633			
		FP	6	0.0759494			

	131 - KENDALL	sections % FN 0 0.0000000 T- 0 0.0000000 0 1 1.0000000 T+ 0 0.0000000 FP 0 0.0000000	100.00%	92.71%	Pass
	232 - UVALDE	sections % FN 0 0.0000000 T- 1 1.0000000 0 0 0.0000000 T+ 0 0.0000000 FP 0 0.0000000	100.00%	92.71%	Pass
JCP	46 - COMAL	sections % FN 0 0.0000000 T- 3 0.7500000 0 0 0.0000000 T+ 0 0.0000000 FP 1 0.2500000	75.00%	91.98%	Fail

6.1.7 Index 2 (Rutting & IRI)

Table 6.12 presents the data quality check results for San Antonio on Index 2. For ACP, the table reveals the accuracy of ACP Rut Auto Severe, IRI Left Score, IRI Right Score, and IRI Average Score, indicating that they exhibit good data quality. ACP Rut Auto Deep shows slightly lower accuracy compared to the Index 2 threshold, which suggests a minor data quality concern. ACP Rut Auto Shallow shows severe lower accuracy compared to the Index 2 threshold, which may suggest a strong data quality concern. The primary data quality issue identified in this analysis is FP. This finding implies that the annual data might tend to overestimate the severity of Rutting & IRI in the pavement.

No CRCP and JCP section data was delivered in the pilot study for Index 2.

Table 6.12 Index 2 (Rutting & IRI) Data quality check for San Antonio

Pavement Type	Distress Type	Network Level	No. of Sections	Accuracy	Threshold
ACP	ACP RUT AUTO SHALLOW AVG PCT	All	sections % FN 0 0.0000000 T- 4 0.0065681 0 42 0.0689655 T+ 292 0.4794745 FP 271 0.4449918	55.50%	75.86%
	ACP RUT AUTO DEEP AVG PCT	All	sections % FN 0 0.0000000 T- 0 0.0000000	92.45%	94.64%

			0 382 0.6272578 T+ 181 0.2972085 FP 46 0.0755337		
	ACP RUT AUTO SEVERE AVG PCT	All	sections % FN 0 0.0000000 T- 0 0.0000000 0 602 0.9885057 T+ 6 0.0098522 FP 1 0.0016420	99.84%	99.86%
	IRI LEFT SCORE (IN/MILE)	All	sections % FN 8 0.0131363 T- 417 0.6847291 0 25 0.0410509 T+ 147 0.2413793 FP 8 0.0131363	96.72%	93.80%
	IRI RIGHT SCORE (IN/MILE)	All	sections % FN 12 0.0197044 T- 398 0.6535304 0 19 0.0311987 T+ 163 0.2676519 FP 12 0.0197044	95.24%	93.19%
	IRI AVERAGE SCORE (IN/MILE)	All	sections % FN 10 0.0164204 T- 418 0.6863711 0 26 0.0426929 T+ 138 0.2266010 FP 11 0.0180624	95.57%	93.45%

6.1.8 Index 3 (Rutting & IRI)

Table 6.13 presents the data quality check results for San Antonio on Index 3. For ACP, the table reveals the accuracy of the Distress Score under FM network level, SH network level, and US network level, indicating that they exhibit good data quality. Distress Score under IH network level shows severe lower accuracy compared to the Index 3 threshold, which may suggest a strong data quality concern. There are similar amounts of pavement sections with FP and FN. This finding implies that both unexpected decreasing and increasing in Distress Score measurements could be equally observed.

No CRCP and JCP section data was delivered in the pilot study for Index 3.

Table 6.13 Index 3 (Rutting & IRI) Data quality check for San Antonio

Pavement Type	Distress Type	Network Level	No. of Sections			Accuracy	Threshold
ACP	Distress Score	FM		sections	%	94.12%	92.73%
			FN	13	0.0332481		
			T-	80	0.2046036		
			0	2	0.0051151		
			T+	286	0.7314578		
			FP	10	0.0255754		
	Distress Score	SH		sections	%	94.03%	92.57%
			FN	4	0.0298507		
			T-	35	0.2611940		
			0	0	0.0000000		
			T+	91	0.6791045		
			FP	4	0.0298507		
	Distress Score	US		sections	%	96.20%	94.13%
			FN	3	0.0379747		
			T-	33	0.4177215		
			0	1	0.0126582		
			T+	42	0.5316456		
			FP	0	0.0000000		
	Distress Score	IH		sections	%	80.00%	92.85%
			FN	0	0.0000000		
			T-	1	0.2000000		
			0	0	0.0000000		
			T+	3	0.6000000		
			FP	1	0.2000000		

6.1.9 County-level Index 2 & Index 3 (Rutting & IRI)

Table 6.14 presents the county-level analysis results for San Antonio. For ACP, 7 counties from the Dallas district were selected for the data audit. The table reveals the accuracy of Medina and Guadalupe, indicating that they exhibit good data quality. Bandera, Uvalde, Comal, Kendall, and Bexar show severe lower accuracy compared to the Index 2 and Index 3 thresholds, which may suggest a strong data quality concern. There are similar amounts of pavement sections with FP and FN. This finding implies that both unexpected decreasing and increasing on Distress Score measurements could be equally observed. There are similar amounts of pavement sections with FP and FN. This finding implies that both unexpected decreasing and increasing on Distress Score measurements could be equally observed. However, the number of pavement sections involved in the pilot study for Index 2 and Index 3 may not be sufficient for some of the counties. Consequently, the evaluation results may not reflect the actual condition.

No CRCP and JCP sections from Dallas were included in the Index 2 and Index 3 pilot study in FY 2022.

Table 6.14 County-level data quality check for San Antonio

Distress Type	County	No. of Sections			Accuracy	Threshold	Determination
ACP	10 - BANDERA		sections	%	33.33%	93.33%	Fail
		FN	1	0.3333333			
		T-	0	0.0000000			
		0	0	0.0000000			
		T+	1	0.3333333			
		FP	1	0.3333333			
	163 - MEDINA		sections	%	95.02%	93.33%	Pass
		FN	17	0.0292096			
		T-	142	0.2439863			
		0	3	0.0051546			
		T+	408	0.7010309			
		FP	12	0.0206186			
	232 - UVALDE		sections	%	0.00%	93.33%	Fail
		FN	0	0.0000000			
		T-	0	0.0000000			
		0	0	0.0000000			
		T+	0	0.0000000			
		FP	1	1.0000000			
	46 - COMAL		sections	%	83.33%	93.33%	Fail
		FN	1	0.1666667			
		T-	1	0.1666667			
		0	0	0.0000000			
		T+	4	0.6666667			
		FP	0	0.0000000			
	95 - GUADALUPE		sections	%	100.00%	93.33%	Pass
		FN	0	0.0000000			
		T-	4	0.5714286			
		0	0	0.0000000			
		T+	3	0.4285714			
		FP	0	0.0000000			
	131 - KENDALL		sections	%	80.00%	93.33%	Fail
		FN	0	0.0000000			
		T-	1	0.2000000			
		0	0	0.0000000			
		T+	3	0.6000000			
		FP	1	0.2000000			
	15 - BEXAR		sections	%	80.00%	93.33%	Fail
		FN	1	0.2000000			
		T-	1	0.2000000			
		0	0	0.0000000			
		T+	3	0.6000000			
		FP	0	0.0000000			

6.1.10 Index 4 (Rutting & IRI)

Table 6.15 presents the data quality check results for San Antonio on Index 4. For ACP, the table reveals the precision of ACP Rut Auto Severe, IRI Left Score, IRI Right Score, and IRI Average Score, indicating that they exhibit good data quality. ACP Rut Auto Deep shows slightly lower accuracy compared to the Index 4 threshold, which suggests a minor data quality concern. ACP Rut Auto Shallow shows severe lower accuracy compared to the Index 4 threshold, which may suggest a strong data quality concern. There are similar amounts of pavement sections with FP and FN. This finding implies that both unexpected decreases and increases in Rutting and IRI measurements could be equally observed.

For CRCP, the table reveals the precision of the IRI Left Score, IRI Right Score, and IRI Average Score, indicating that they exhibit good data quality. There are similar amounts of pavement sections with FP and FN. This finding implies that both unexpected decreasing and increasing in Rutting and IRI measurements could be equally observed. However, the number of pavement sections involved in the pilot study for Index 4 may not be sufficient. Consequently, the evaluation results may not reflect the actual condition.

For JCP, only a few of the pavement section data were involved in the pilot study for Index 4. Consequently, the evaluation results may not reflect the actual condition.

Table 6.15 Index 4 (Rutting & IRI) data quality check for San Antonio

Pavement Type	Distress Type	Network Level	No. of Sections			Precision	Threshold
ACP	ACP RUT AUTO SHALLOW AVG PCT	All		sections	%	82.17%	90.95%
			FN	786	0.0799105		
			T-	2173	0.2209231		
			0	1602	0.1628711		
			T+	4049	0.4116511		
			FP	1226	0.1246442		
	ACP RUT AUTO DEEP AVG PCT	All		sections	%	95.94%	96.58%
			FN	288	0.0292802		
			T-	975	0.0991257		
			0	6893	0.7007930		
			T+	1332	0.1354209		
			FP	348	0.0353802		
	ACP RUT AUTO SEVERE AVG PCT	All		sections	%	99.37%	93.33%
			FN	54	0.0054900		
			T-	0	0.0000000		
			0	9720	0.9882066		
			T+	0	0.0000000		
			FP	62	0.0063034		
	IRI LEFT	All		sections	%	97.69%	95.07%

	SCORE (IN/MILE)		FN 169 0.0172063 T- 4375 0.4454286 0 452 0.0460191 T+ 4704 0.4789249 FP 122 0.0124211		
	IRI RIGHT SCORE (IN/MILE)	All	sections % FN 414 0.0421503 T- 4466 0.4546935 0 369 0.0375687 T+ 4330 0.4408471 FP 243 0.0247404	95.08%	93.83%
	IRI AVERAGE SCORE (IN/MILE)	All	sections % FN 280 0.0285074 T- 4378 0.4457341 0 458 0.0466300 T+ 4565 0.4647730 FP 141 0.0143555	97.27%	94.53%
CRCP	IRI LEFT SCORE (IN/MILE)	All	sections % FN 0 0.0000000 T- 38 0.4691358 0 6 0.0740741 T+ 36 0.4444444 FP 1 0.0123457	97.67%	94.36%
	IRI RIGHT SCORE (IN/MILE)	All	sections % FN 3 0.0370370 T- 37 0.4567901 0 5 0.0617284 T+ 35 0.4320988 FP 1 0.0123457	97.56%	95.74%
	IRI AVERAGE SCORE (IN/MILE)	All	sections % FN 2 0.0246914 T- 32 0.3950617 0 8 0.0987654 T+ 38 0.4691358 FP 1 0.0123457	97.87%	95.13%
JCP	IRI LEFT SCORE (IN/MILE)	All	sections % FN 0 0.0000000 T- 1 0.2500000 0 0 0.0000000 T+ 3 0.7500000 FP 0 0.0000000	100.00%	98.25%
	IRI RIGHT SCORE (IN/MILE)	All	sections % FN 0 0.0000000 T- 1 0.2500000 0 0 0.0000000	66.67%	95.87%

			T+ 2 0.5000000 FP 1 0.2500000		
	IRI AVERAGE SCORE (IN/MILE)	All	sections % FN 0 0.0000000 T- 1 0.2500000 0 0 0.0000000 T+ 2 0.5000000 FP 1 0.2500000	66.67%	97.44%

6.1.11 County-level analysis for Index 4 & Index 5 (Rutting & IRI)

Table 6.16 presents the county-level analysis results of Ride Score for San Antonio. For ACP, there are 15 counties from the San Antonio district were selected for the data audit. The table reveals the accuracy of Atascosa and La Salle, indicating that they exhibit good data quality. Guadalupe, Bexar, Frio, Medina, and Bandera show slightly lower accuracy compared to the Index 5 threshold, which suggests a minor data quality concern. Wilson, Comal, Uvalde, McMullen, Kendall, Kerr, Caldwell, and Gillespie show severe lower accuracy compared to the Index 5 threshold, which may suggest a strong data quality concern. The primary data quality issue identified in this analysis is FP. This finding implies that the dominant pavement sections with data quality issues show an unexpected increase in Ride Score measurements.

For CRCP, there are 3 counties from the San Antonio district were selected for the data audit. The table reveals the accuracy of Kendall and Uvalde, indicating that they exhibit good data quality. Bexar shows slightly lower accuracy compared to the Index 5 threshold, which suggests a minor data quality concern. The primary data quality issue identified in this analysis is FP. This finding implies that the dominant pavement sections with data quality issues show an unexpected increase in Ride Score measurements. However, the number of Ride Score data involved in the pilot study for Index 4 and Index 5 may not be sufficient. Consequently, the evaluation results may not reflect the actual condition.

For JCP, only a few of the Ride Score data was involved in the pilot study for Index 4 and Index 5. Consequently, the evaluation results may not reflect the actual condition.

Table 6.16 County-level data quality check for San Antonio (Ride Score)

Distress Type	County	No. of Sections	Accuracy	Threshold	Determination
ACP	95 - GUADALUPE	sections % FN 34 0.0368364 T- 490 0.5308776 0 23 0.0249187 T+ 341 0.3694475 FP 35 0.0379198	92.32%	92.99%	Fail
	247 - WILSON	sections % FN 17 0.0244957	86.47%	92.99%	Fail

		T-	279	0.4020173			
		0	24	0.0345821			
		T+	298	0.4293948			
		FP	76	0.1095101			
	46 - COMAL		sections	%	83.41%	92.99%	Fail
		FN	35	0.0553797			
		T-	270	0.4272152			
		0	15	0.0237342			
		T+	243	0.3844937			
		FP	69	0.1091772			
	15 - BEXAR		sections	%	92.05%	92.99%	Fail
		FN	70	0.0271739			
		T-	1191	0.4623447			
		0	60	0.0232919			
		T+	1121	0.4351708			
		FP	134	0.0520186			
	232 - UVALDE		sections	%	86.97%	92.99%	Fail
		FN	18	0.0255319			
		T-	377	0.5347518			
		0	23	0.0326241			
		T+	214	0.3035461			
		FP	73	0.1035461			
	83 - FRIO		sections	%	88.34%	92.99%	Fail
		FN	15	0.0196592			
		T-	264	0.3460026			
		0	8	0.0104849			
		T+	402	0.5268676			
		FP	74	0.0969856			
	7 - ATASCOSA		sections	%	95.66%	92.99%	Pass
		FN	10	0.0101420			
		T-	464	0.4705882			
		0	22	0.0223124			
		T+	461	0.4675456			
		FP	29	0.0294118			
	163 - MEDINA		sections	%	89.83%	92.99%	Fail
		FN	34	0.0455154			
		T-	418	0.5595716			
		0	15	0.0200803			
		T+	238	0.3186078			
		FP	42	0.0562249			
	10 - BANDERA		sections	%	90.91%	92.99%	Fail
		FN	12	0.0312500			
		T-	235	0.6119792			
		0	7	0.0182292			
		T+	108	0.2812500			

		FP	22	0.0572917			
	162 - MCMULLEN		sections	%	85.90%	92.99%	Fail
		FN	8	0.0256410			
		T-	96	0.3076923			
		0	3	0.0096154			
		T+	169	0.5416667			
		FP	36	0.1153846			
	131 - KENDALL		sections	%	85.85%	92.99%	Fail
		FN	20	0.0488998			
		T-	201	0.4914425			
		0	10	0.0244499			
		T+	141	0.3447433			
		FP	37	0.0904645			
	133 - KERR		sections	%	85.93%	92.99%	Fail
		FN	9	0.0138037			
		T-	302	0.4631902			
		0	12	0.0184049			
		T+	248	0.3803681			
		FP	81	0.1242331			
	28 - CALDWELL		sections	%	54.55%	92.99%	Fail
		FN	0	0.0000000			
		T-	0	0.0000000			
		0	0	0.0000000			
		T+	6	0.5454545			
		FP	5	0.4545455			
	87 - GILLESPIE		sections	%	54.17%	92.99%	Fail
		FN	0	0.0000000			
		T-	0	0.0000000			
		0	0	0.0000000			
		T+	13	0.5416667			
		FP	11	0.4583333			
	142 - LA SALLE		sections	%	100.00%	92.99%	Pass
		FN	0	0.0000000			
		T-	0	0.0000000			
		0	0	0.0000000			
		T+	4	1.0000000			
		FP	0	0.0000000			
CRCP	15 - BEXAR		sections	%	91.14%	92.71%	Fail
		FN	1	0.0126582			
		T-	33	0.4177215			
		0	7	0.0886076			
		T+	32	0.4050633			
		FP	6	0.0759494			
	131 - KENDALL		sections	%	100.00%	92.71%	Pass
		FN	0	0.0000000			

		T- 0 0.0000000 0 1 1.0000000 T+ 0 0.0000000 FP 0 0.0000000			
	232 - UVALDE	sections % FN 0 0.0000000 T- 1 1.0000000 0 0 0.0000000 T+ 0 0.0000000 FP 0 0.0000000	100.00%	92.71%	Pass
JCP	46 - COMAL	sections % FN 0 0.0000000 T- 3 0.7500000 0 0 0.0000000 T+ 0 0.0000000 FP 1 0.2500000	75.00%	91.98%	Fail

Further, the image check was performed for the road sections in San Antonio for FY 2021 and FY 2022 to check the precision. The details about image checks for various distresses such as ACP alligator cracking, ACP longitudinal cracking, ACP transverse cracking, ACP patching, and ACP failure are discussed in the subsequent sections.

6.2 Image Check Result for Accuracy

6.2.1 Alligator Cracking Image Check Result for Accuracy

ACP Alligator Cracking from FY 2022

A) Observation: None correct (3 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Audit	Software	Manual observation
15-Bexar	IH0037 R	133	0.5	133	1	1	19	11	14
15-Bexar	SL1604 K	576	0.5	576	1	8	18	8	13
46-Comal	FM0306 R	524	0	524	0.5	1	11	1	0

B) Observation: Audit correct (12 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Audit	Software	Manual observation
10-Bandera	SH0173 K	484	0	484	0.5	15	29	0	29

15-Bexar	FM1346 K	508	0	508	0.5	3	14	3	14
15-Bexar	FM1346 K	508	0.5	508	1	6	33	6	33
15-Bexar	FM1346 K	508	1	508	1.5	1	17	1	17
15-Bexar	IH0035 R	138	0	138	0.5	11	0	11	0

C) Observation: Both the Annual Rating and Software are correct (3 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Audit	Software	Manual observation
15-Bexar	FM1535 R	490	0	490	1	0	19	0	0
15-Bexar	SH0016 L	580	1.5	580	2	0	13	0	0
15-Bexar	SH0016 L	586	0.5	586	1	0	20	0	0

Summary:

The alligator cracking road surface images data for FY 2022 was obtained for various counties in San Antonio district of Texas. The image analysis is performed for the road sections using Path view software and manual observation. Throughout the section image analysis, manual observation is assumed to be correct. Additionally, several significant findings are noted concerning the ACP alligator cracking distress rating.

The distress ratings are classified into four categories: a) Annual rating, b) Audit rating, c) Software detection, and d) Manual observation. Based on the mentioned categories, there are three cases, which are enlisted as follows:

1) None of the four distress rating categories are correct (Annual rating ≠ Audit rating ≠ Software detection ≠ Manual observation)
2) Audit rating is correct (Audit = Manual observation)
3) Both annual rating and software detection are correct (Annual rating = Software detection = Manual observation)

The following is a detailed discussion of the three cases:

1) **None of the 4 distress rating categories are correct (Annual rating ≠ Audit rating ≠ Software detection ≠ Manual observation)**

It has been observed that the distress rating varies for each category of observation based on an image analysis. This is mostly because the manual observation rating differs from the other three categories. For some sections, the annual rating and software detected less distress than manual observation. This was primarily caused by the software's inadequacy to correctly detect the alligator cracking on the images of the pavement surface. In this situation, the manual observation rating exceeded the software and annual rating results. Additionally, audit rating reported more alligator cracking than the other three categories. This might be caused due to incorrect distress identification, mistaking longitudinal cracking, or raveling for alligator cracking.

2) **Audit rating is correct (Audit rating = Manual observation)**

Based on the image analysis, it has been observed that the ACP alligator cracking rating was the same for both audit rating and manual observation. From this, it can be concluded that the audit rating is correct since it has accounted for all the ACP alligator cracking for that road section.

3) **Both annual rating and software detection are correct (Annual rating = Software detection = Manual observation)**

Based on the image analysis, the annual rating and software detection of ACP alligator cracking was found to be the same as manual observation. From this, it can be concluded that the annual rating and software detection are correct since both have detected all the ACP alligator cracking that was identified manually for that road section.

Overall, for most of the sections with the alligator cracking, the audit rating and manual observation distress ratings are the same. Also, it is to be noted that the manual rating could be lesser than the other three categories due to the poor quality of images for the road sections. Furthermore, there is a need to improve the software capabilities to precisely detect the distress types.

6.2.2 Longitudinal Cracking Image Check Result for Accuracy

ACP Longitudinal Cracking from FY 2022

A) Observation: None correct (33 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Audit	Software	Manual observation
15-Bexar	FM0078 K	500	1.5	500	2	49	105	84	95
15-Bexar	FM1303 K	498	1.5	498	2	60	9	60	12
15-Bexar	FM1535 R	490	0.3	490	0.5	91	152	43	55
15-Bexar	IH0410 R	21	0.5	21	0.8	70	208	18	26
15-Bexar	PA1502 L	490	1.5	490	2	59	3	12	20

B) Observation: Audit correct (29 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Audit	Software	Manual observation
15-Bexar	IH0037 R	135	0.5	135	1.1	5	39	5	39
15-Bexar	IH0037 L	136	0.5	136	1	6	59	6	59
15-Bexar	IH0410 R	12	0	12	0.5	4	71	4	71
15-	IH0410	41	0	41	0.5	8	67	8	67

Bexar	R								
15-Bexar	IH0410 L	11	0	11	0.5	9	48	9	48

C) Observation: Annual rating correct (1 section)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Audit	Software	Manual observation
10-Bandera	SH0173 K	484	0.5	484	1	3	44	0	3

D) Observation: Software correct (6 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Audit	Software	Manual observation
15-Bexar	FM1303 K	498	0	498	0.5	130	56	85	85
15-Bexar	FM1535 K	486	0	486	0.5	4	55	0	0
15-Bexar	FM1957 K	480	1	480	1.5	110	174	110	110
15-Bexar	SH0218 K	502	0.6	502	1	52	8	27	27
15-Bexar	SH0218 L	502	0.4	502	0.6	65	0	38	38

E) Observation: Both the Annual Rating and Software are correct (16 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Audit	Software	Manual observation
15-Bexar	FM1957 K	480	0	480	0.5	52	95	52	52
15-Bexar	IH0410 R	26	5	26	5.5	0	41	0	0
15-Bexar	IH0035 R	150	0	150	0.5	66	111	66	66
15-Bexar	SH0016 L	580	1.5	580	2	79	153	79	79
15-Bexar	SH0016 L	582	0	582	0.5	9	50	9	9

Summary:

The distress ratings are classified into four categories: a) Annual rating, b) Audit rating, c) Software detection, and d) Manual observation. Based on the mentioned categories, there are five cases, which are enlisted as follows:

1) None of the four distress rating categories are correct (Annual rating ≠ Audit rating ≠ Software detection ≠ Manual observation)
2) Audit rating is correct (Audit = Manual observation)

3) Annual rating is correct (Annual rating = Manual observation)
4) Software detection is correct (Software detection = Manual observation)
5) Both annual rating and software detection are correct (Annual rating = Software detection = Manual observation)

The following is a detailed discussion of the five cases:

1) None of the 4 distress rating categories are correct (Annual rating ≠ Audit rating ≠ Software detection ≠ Manual observation)

It has been observed that the distress rating for longitudinal cracking varies for each category of observation based on an image analysis. This is mostly because the manual observation rating differs from the other three categories. For majority of the sections, the audit rating reported more distress than manual observation. Additionally, the annual rating and software reported more longitudinal cracking as compared to manual observation. This was mostly caused by software and annual ratings that incorrectly identified distress, mistaking lane strip, road markings, crack sealing (maintenance activities) or raveling for the longitudinal cracking. On another note, in some instances, longitudinal cracking was not reported both by the software as well as the annual rating. This was primarily caused by the software's inadequacy to correctly detect the longitudinal cracking on the images of the pavement surface. In this situation, the manual observation rating exceeded the software and annual ratings results.

2) Audit rating is correct (Audit rating = Manual observation)

Based on the image analysis, it has been observed that the ACP longitudinal cracking rating was the same for both audit rating and manual observation. From this, it can be concluded that the audit rating is correct since it has accounted for all the ACP longitudinal cracking for that road section. In this case, the annual rating and software reported less distress ratings due to their inadequacy to detect all the longitudinal cracks on the pavement surface.

3) Annual rating is correct (Annual rating = Manual observation)

Based on the image analysis, only for one section it has been observed that the longitudinal cracking rating was same for both annual rating and manual observation. From this, it can be concluded that the annual rating is correct since it has accounted for all the ACP longitudinal cracking for that road section.

4) Software detection is correct (Software detection = Manual observation)

Based on the image analysis, it has been observed that the ACP longitudinal cracking rating was the same for both software detection and manual observation. From this, it can be concluded that the software rating is correct since it has detected all the locations of longitudinal cracking for that road section. But this was mainly observed for the road sections with zero distress.

5) Both annual rating and software detection are correct (Annual rating = Software detection = Manual observation)

Based on the image analysis for year 2022, the annual rating and software detection of ACP longitudinal cracking was found to be the same as manual observation. From this, it can be concluded that the annual rating and software detection are correct since both detected all the longitudinal cracking that was identified manually for that road section. However, in this case, the audit reported more longitudinal cracking than the other three rating categories.

6.2.3 Transverse Cracking Image Check Result for Accuracy

ACP Transverse Cracking from FY 2022

A) Observation: None correct (3 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Audit	Software	Manual observation
15-Bexar	IH0037R	135	0	135	0.5	3	10	3	1
15-Bexar	IH0037R	49	0.6	49	1	0	5	0	1
15-Bexar	SH0016L	580	1	580	1.5	2	5	2	3

B) Observation: Audit correct (3 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Audit	Software	Manual observation
15-Bexar	SL0353K	506	0	506	0.2	4	1	1	1
15-Bexar	SL0353K	506	0.2	506	0.7	2	5	5	5
15-Bexar	SL1604K	580	1	580	1.5	4	1	1	1

C) Observation: Software correct (4 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Audit	Software	Manual observation
10-Bandera	SH0016K	560	1.5	560	2	0	3	2	2
15-Bexar	IH0035R	150	0	150	0.5	1	4	2	2
15-Bexar	SH0016L	586	0	586	0.5	0	5	1	1
95-Guadalupe	FM0464L	520	0	520	0.5	0	8	1	1
95-Guadalupe	UA0090K	520	0	520	0.5	2	7	3	3

D) Observation: Both the Annual Rating and Software are correct (7 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Audit	Software	Manual observation
247-Wilson	SH0097 K	538	1.5	538	2	0	38	0	0
15-Bexar	IH0410 R	50	0	50	0.5	2	7	2	2
15-Bexar	IH0410 R	50	0.5	50	1	1	7	1	1
15-Bexar	SH0016 L	580	1.5	580	2	4	10	4	4
15-Bexar	SH0016 L	586	0.5	586	1	1	9	1	1

Summary:
Like ACP alligator cracking and longitudinal cracking, the distress ratings for transverse cracking are also divided into four categories: a) Annual rating, b) Audit rating, c) Software detection, and d) Manual observation. Also, based on the mentioned categories, there are four cases same as that for alligator and longitudinal cracking. The detailed discussion of the four cases is as follows:

1) **None of the 4 distress rating categories are correct (Annual rating ≠ Audit rating ≠ Software detection ≠ Manual observation)**

It has been observed that the distress rating varies for each category of observation based on an image analysis. This is mostly because the manual observation rating differs from the other three categories. For some sections, the annual rating reported more transverse cracking than manual observation. This was mostly caused by incorrect distress identification, mistaking sealed cracks (maintenance activities) or raveling for the transverse cracking. On another note, for one section, the software could not detect the distress on the pavement surface images and hence the ratings based on software was zero due to its inadequacy to correctly detect the transverse cracking.

2) **Audit rating is correct (Audit rating = Manual observation)**

Based on the image analysis, it has been observed that the ACP transverse cracking rating was the same for both audit rating and manual observation. From this, it can be concluded that the audit rating is correct since it has accounted for all the ACP transverse cracking for that road section.

3) **Software detection is correct (Software detection = Manual observation)**

In this case, it has been observed that the transverse cracking rating was the same for both software detection and manual observation. From this, it can be concluded that the software rating is correct since it has detected all the locations of longitudinal cracking for that road section.

4) **Both annual rating and software detection are correct (Annual rating = Software detection = Manual observation)**

Based on the image analysis for the year 2022, the annual rating and software detection of ACP transverse cracking were found to be the same as manual observation. From this, it can be concluded that the annual rating and software detection are correct since both detected all the transverse cracking that was identified manually for the road sections.

6.2.4 Patching Image Check Result for Accuracy

ACP Patching from FY 2022

A) Observation: Both Audit and software are correct (7 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Audit	Software	Manual observation
247-Wilson	SH0123K	518	0	518	0.5	27	0	0	0
247-Wilson	SH0123K	518	0.5	518	1	21	0	0	0
247-Wilson	US0087K	716	0.5	716	1	13	0	0	0
95-Guadalupe	IH0010L	594	0.2	594	0.7	49	0	0	0

B) Observation: Annual rating correct (4 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Audit	Software	Manual observation
247-Wilson	FM2505K	530	0	530	0.5	10	0	0	10
247-Wilson	FM2505K	530	0.5	530	1	22	0	0	22
247-Wilson	FM2579K	518	1	518	1.5	11	0	0	11
247-Wilson	FM2579K	518	1.5	518	2	14	0	0	14

Summary:
The distress ratings for ACP patching are divided into four categories: a) Annual rating, b) Audit rating, c) Software detection, and d) Manual observation. Likewise, based on the stated categories, there are two cases, which are discussed as follows:

1) **Both Audit rating and Software detection rating are correct (Audit rating = Software detection = Manual observation)**

Based on the image analysis, it has been observed that the ACP patching ratings from both audit rating and software detection are the same as that of the manual observation ratings. From this, it can be concluded that the audit and software ratings are correct since it has reported all the ACP patching locations on the road sections wherein suitable maintenance activities have been performed.

2) **Annual is correct (Annual rating = Manual observation)**

This case has been observed for the majority of the pavement sections wherein the ACP patching rating was the same for annual rating and manual observation. Additionally, in this case, the audit and software were inadequate to detect the patching on the pavement surface.

In summary, the distress ratings for pavement conditions are evaluated through annual ratings, audit ratings, software detection, and manual observation methods. The analysis indicates that annual ratings, audit ratings, and software detection can either underestimate or overestimate ACP (alligator, longitudinal, and transverse) cracking and patching when compared to manual observations. This discrepancy might arise from difficulties in accurately identifying the type of distress or from misinterpretations among the various distress types. Additionally, overestimation can result from the incorrect identification of lane stripes and raveling as cracking.
Moreover, the ratings generated by software detection are not consistently aligned with those from annual ratings, audit ratings, and manual observations. These inaccuracies in distress identification can significantly impact maintenance and rehabilitation strategies, leading to incorrect budget allocations over the pavement service lives.

6.3 Image Check Result for Precision

6.3.1 Alligator Cracking Image Check Result for Precision

1. ACP Alligator Cracking from FY 2022

A) Observation: None correct (174 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
95 - Guadalupe	BS0123B K (899)	492	0	492	0.152	11	0	4
15 - Bexar	US0281 L (940)	520	0	520	0.5	11	14	5
163 – Medina	FM1250 K (595)	450	0.5	450	1	51	53	48
7 - Atascosa	SH0173 K (463)	544	0	544	0.5	2	2	9
83 – Frio	SS0581 K (234)	614	0.5	614	1	14	14	44

B) Observation: Annual rating (correct) (15 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15 - Bexar	IH0035 R (463)	142	0	142	0.5	24	20	24
95 - Guadalupe	SH0080 K (78)	490	1.5	490	1.98	22	34	22

15 - Bexar	US0087 R (199)	704	0.5	704	1	18	25	18
133 - Kerr	SH0027 K (822)	482	0.5	482	1	9	14	9
163 - Medina	IH0035 A (442)	128	0	128	0.5	3	5	3

C) **Observation:** Software (correct) (33 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15 – Bexar	US0090 X (580)	556	0	556	0.209	5	0	0
95 – Guadalupe	IH0010 L (897)	613	0	613	0.5	8	12	12
163 – Medina	FM0462 K (597)	500	0.5	500	1	39	6	6
15 – Bexar	IH0410 A (1344)	46	0	46	0.5	2	4	4
15 – BEXAR	FM1516 K (1269)	500	0.5	500	1	13	16	16

D) **Observation:** Both the Annual Rating and Software are correct (83 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
95 – Guadalupe	FM1339 K (1048)	526	0.5	526	1	0	0	0
15 – Bexar	IH0035 R (412)	141	0.5	141	1.093	23	23	23
162 - Mcmullen	SH0016 K (49)	658	0.5	658	1	0	0	0
15 - Bexar	IH0410 X (1354)	31	0.678	31	1.203	24	24	24
83 – Frio	FM0140 K (480)	462	1	462	1.5	1	1	1

E) **Observation:** Annual Rating > Software and Manual observation (170 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
95 - Guadalupe	FM0725 K (995)	482	0	482	0.5	13	1	3
15 - Bexar	IH0010 A (1340)	580	0	580	0.5	25	20	7
95 - Guadalupe	IH0010 A (84)	612	0	612	0.5	15	5	3

163 - Medina	FM0462 K (599)	488	0	488	0.5	24	3	7
163 - Medina	FM0462 K (599)	486	1.543	486	1.984	32	5	3

F) Observation: Software > Annual Rating and Manual observation (38 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
163 - Medina	FM1250 K (595)	450	0	450	0.5	46	95	46
46 - Comal	FM2673 K (886)	506	0	506	0.5	7	9	4
83 – Frio	IH0035 X (402)	88	0	88	0.5	12	14	2
15 - Bexar	IH0410 A (1344)	44	0	44	0.5	9	21	4
15 - Bexar	FM0078 K (1274)	504	0.5	504	1	21	24	11

G) Observation: Manual observation > Software and Annual Rating (80 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
95 - Guadalupe	FM0466 K (189)	544	0	544	0.103	3	3	10
15 - Bexar	IH0035 R (412)	135	0	135	0.5	10	14	32
163 - Medina	FM0462 K (596)	506	1.5	506	1.937	11	13	21
15 - Bexar	FM1346 K (151)	504	0	504	0.5	11	15	20
163 - Medina	FM0462 K (599)	488	1	488	1.5	1	4	8

The road surface image data (for 2021 and 2022) was obtained for various counties in the San Antonio district of Texas. The image analysis is performed for the road sections using Path view software and manual observation. Throughout the section on image analysis, manual observation is assumed to be correct. Additionally, several significant findings are noted concerning the ACP alligator cracking distress rating.

The distress ratings are classified into three categories: a) Annual rating, b) Software detection, and c) Manual observation. Based on the mentioned categories, there are seven cases, which are enlisted as follows:

1) None of the three distress rating categories are correct (Annual rating ≠ Software detection ≠ Manual observation)
2) Annual rating is correct (Annual rating = Manual observation)

3) Software detection is correct (Software detection = Manual observation)
4) Both annual rating and software detection are correct (Annual rating = Software detection = Manual observation)
5) Annual Rating > Software and Manual observation
6) Software > Annual Rating and Manual observation
7) Manual observation > Software and Annual Rating

The following is a detailed discussion of the seven cases:

1) None of the 3 distress rating categories are correct (Annual rating ≠ Software detection ≠ Manual observation)

It has been observed that the distress rating varies for each category of observation based on an image analysis. This is mostly because the manual observation rating differs from the other two categories. For some sections in the year 2022, the annual rating and software detected more distress than manual observation. This was mostly caused by software and annual ratings that incorrectly identified distress, mistaking longitudinal cracking, or raveling for alligator cracking. Also, in some instances, less alligator cracking was reported by the software and annual rating. This was primarily caused by the software's inadequacy to correctly detect the alligator cracking on the images of the pavement surface. In this situation, the manual observation rating exceeded the software and annual ratings results.

2) **Annual rating is correct (Annual rating = Manual observation)**

Based on the image analysis, it has been observed that the ACP alligator cracking rating was the same for both annual rating and manual observation. From this, it can be concluded that the annual rating is correct since it has accounted for all the ACP alligator cracking for that road section.

3) **Software detection is correct (Software detection = Manual observation)**

Based on the image analysis for the year 2022, it has been observed that the ACP alligator cracking rating was the same for both software detection and manual observation. From this, it can be concluded that the software rating is correct since it has detected all the ACP alligator cracking for that road section. However, for the year 2021, the software detection was correct most of the time when there was zero distress present for the road section.

4) Both annual rating and software detection are correct (Annual rating = Software detection = Manual observation)

Based on the image analysis, the annual rating and software detection of ACP alligator cracking was found to be the same as manual observation. From this, it can be concluded that the annual rating and software detection are correct since both have detected all the ACP alligator cracking that was identified manually for that road section.

5) Annual Rating > Software and Manual observation

This case is a sub-type of case 1 where none of the annual rating and software detection are correct. In this scenario, the annual rating for ACP alligator cracking was more than both software and manual observation. The reason for this difference may be due to incorrect distress identification or misinterpretation among distress types.

6) **Software > Annual Rating and Manual observation**

This case is also a sub-type of case 1 where none of the annual rating and software detection are correct. In this case, the software detects more locations for ACP alligator cracking because the software could not differentiate between the ACP alligator cracking and raveling distresses. Additionally, the software also detects longitudinal or transverse cracking as alligator cracks. This leads to higher distress ratings by the software as compared to annual ratings and manual observation.

7) **Manual observation > Software and Annual Rating**

This case is also a sub-type of case 1 where none of the annual rating and software detection are correct. In this case, the ACP alligator cracking was accurately identified manually at all locations. However, all the locations were not detected by the software and annual rating for the precise alligator cracking distress identification.

Overall, for most of the sections especially with the longitudinal cracking, there is a difference among annual rating, software detection, and manual observation distress ratings. It is also to be noted that the manual rating could be lesser due to the poor quality of images for the road sections. Furthermore, there is a need to improve the software capabilities to precisely detect the distress types.

2. ACP Alligator Cracking from FY 2021

ACP alligator cracking of FY 2021 has a similar result as FY 2021, the results are listed below.

A) **Observation:** None correct (111 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
163 - Medina	FM0462 K (849)	494	0	494	0.5	4	0	11
15 - Bexar	US0090 X (1103)	566	0	566	0.5	9	0	1
163 - Medina	FM0462 K (849)	490	1.5	490	1.979	0	0	4
95 - Guadalupe	IH0010 A (1463)	612	0	612	0.5	72	0	1
163 - Medina	FM0462 K	486	0.043	486	0.543	59	0	15

B) Observation: Annual rating (correct) (10 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
163 – Medina	FM0462 K (849)	492	0	492	0.5	2	0	2
95 – Guadalupe	US0090 K (1461)	662	0	662	0.5	8	0	8
15 – Bexar	FM1560 K (510)	494	0	494	0.5	5	0	5
163 – Medina	FM1250 K (1441)	450	0	450	0.5	30	0	30
131 – Kendall	RM1376 K (925)	476	1	476	1.5	8	0	8

C) Observation: Software (correct) (164 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
83 – Frio	IH0035 X (787)	87	0.5	87	1	1	0	0
95 – Guadalupe	SH0080 K (1409)	492	1.5	492	1.995	11	0	0
15 – Bexar	SL0353 K (1443)	504	1.5	504	1.954	25	0	0
15 – Bexar	FM1346 K (1152)	504	1.5	504	2.002	21	0	0
95 – Guadalupe	FM0078 K (157)	522	1	522	1.5	10	0	0

D) Observation: Both the Annual Rating and Software are correct (32 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15 – Bexar	IH0035 R (800)	139	0	139	0.5	0	0	0
247 – Wilson	US0181 R (358)	530	1.261	530	2.027	0	0	0
46 – Comal	US0281 L (760)	504	1	504	1.5	0	0	0
232 – Uvalde	RM2690 K (1298)	506	0.5	506	1	0	0	0
15 – Bexar	IH0035 X (1039)	134	0.5	134	1.035	0	0	0

E) Observation: Annual Rating > Software and Manual observation (129 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15 - Bexar	IH0035 L (770)	142	0.5	142	1.048	17	0	0
133 – Kerr	SH0027 K (950)	482	0.5	482	1	24	0	6
162 – Mcmullen	SH0016 K (76)	658	0.5	658	1	9	3	1
15 – Bexar	IH0410 A (1167)	31	0.98	31	1.203	50	0	31
15 – Bexar	SH0151 L (1313)	480	1	480	1.5	30	23	3

F) **Observation:** Software > Annual Rating and Manual observation (2 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15 – Bexar	SH0211 K (833)	486	0	486	0.5	10	14	0
7 – Atascosa	IH0037 R (29)	114	0	114	0.5	6	22	9

G) **Observation:** Manual observation > Software and Annual Rating (32 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
163 – Medina	FM0462 K (849)	494	0	494	0.5	4	0	11
163 – Medina	FM0462 K (848)	500	0.5	500	1	0	0	4
95 – Guadalupe	SH0123 K (1089)	476	0	476	0.5	2	0	5
83 – Frio	SS0581 K (1052)	616	0	616	0.5	20	0	43
15 – Bexar	IH0037 A (390)	128	0	128	0.5	1	0	6

6.3.2 Longitudinal Cracking Image Check Result for Precision

1. ACP Longitudinal Cracking from FY 2022

A) **Observation:** None correct (165 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15 – Bexar	FM0078 K (1273)	504	0	504	0.5	63	69	79
46 – Comal	FM0306 K (974)	504	0	504	0.5	13	20	49
15 - Bexar	US0281 X (1180)	522	0	522	0.478	1	1	6
163 – Medina	IH0035 X (421)	130	0	130	0.542	3	3	0
15 - Bexar	US0181 L (309)	514	0	514	0.5	12	6	8

B) Observation: Annual rating (correct) (2 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
46-Comal	FM0306 K (975)	500	0	500	0.5	1	2	1
15-Bexar	IH0410 A (1341)	34	0	34	0.5	46	50	46

C) Observation: Software (correct) (34 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15 - Bexar	IH0010 X (1088)	589	0	589	0.5	8	12	12
15 - Bexar	IH0035 R (1333)	152	0	152	0.5	77	71	71
95 - Guadalupe	IH0010 A (84)	612	0	612	0.5	7	2	2
131 – Kendall	IH0010 X (859)	524	0	524	0.578	55	4	4
15 - Bexar	US0181 R (296)	516	0	516	0.202	88	0	0

D) Observation: Both the Annual Rating and Software are correct (37 sections)

County	Highway And Roadbed	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation

	ID							
95 - Guadalupe	IH0010 L (1068)	600	0	600	0.5	0	0	0
15 - Bexar	US0281 A (1172)	524	0	524	0.5	70	70	70
15 - Bexar	US0181 R (296)	508	0	508	0.5	39	39	39
163 – Medina	IH0035 L (368)	126	0.5	126	1.01	0	0	0
15 - Bexar	US0090 R (1232)	564	0	564	0.5	19	19	19

E) **Observation:** Annual Rating > Software and Manual observation (84 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15 - Bexar	IH0035 R (412)	141	0	141	0.5	143	124	60
15 - Bexar	IH0035 R (413)	146	0	146	0.5	75	70	48
15 - Bexar	IH0035 L (365)	141	0	141	0.5	78	40	18
95 - Guadalupe	IH0010 A (77)	628	0	628	0.267	44	38	21
15 - Bexar	IH0010 A (1340)	579	0	579	0.5	172	143	71

F) **Observation:** Software > Annual Rating and Manual observation (13 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
95 - Guadalupe	US0090 K (912)	662	0	662	0.5	147	150	25
46 - Comal	US0281 L (941)	510	0	510	0.5	144	149	41
15 - Bexar	IH0010 L (1069)	594	0	594	0.184	93	109	67
15 - Bexar	IH0035 L (365)	139	0	139	0.5	166	190	28
15 - Bexar	IH0035 R (412)	137	0	137	0.5	150	175	89

G) **Observation:** Manual observation > Software and Annual Rating (13 sections)

County	Highway And Roadbed	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation

	ID							
83 - Frio	IH0035 R (407)	117	0	117	0.5	70	74	78
15 - Bexar	IH0035 R (413)	143	0	143	0.5	66	66	91
15 - Bexar	IH0035 L (364)	142	0	142	0.5	1	49	62
95 - Guadalupe	IH0010 L (1067)	605	0	605	0.5	15	15	75
95 - Guadalupe	IH0010 R (1078)	605	0	605	0.5	7	19	49

The road surface images data (for 2021 and 2022) was obtained for various counties in San Antonio district of Texas. The image analysis is performed for the road sections using Path view software and manual observation. Throughout the section image analysis, manual observation is assumed to be correct. Additionally, several significant findings are noted concerning the rating of ACP longitudinal cracking.

Like the ACP alligator cracking, the longitudinal cracking distress ratings are also classified into three categories: a) Annual rating, b) Software detection, and c) Manual observation. Based on the mentioned categories, there are seven cases, which are enlisted as follows:

1) None of the three distress rating categories are correct (Annual rating ≠ Software detection ≠ Manual observation)
2) Annual rating is correct (Annual rating = Manual observation)
3) Software detection is correct (Software detection = Manual observation)
4) Both annual rating and software detection are correct (Annual rating = Software detection = Manual observation)
5) Annual Rating > Software and Manual observation
6) Software > Annual Rating and Manual observation
7) Manual observation > Software and Annual Rating

The following is a detailed discussion of the seven cases:

1) **None of the 3 distress rating categories are correct (Annual rating ≠ Software detection ≠ Manual observation)**

It has been observed that the distress rating varies for each category of observation based on an image analysis. This is mostly because the manual observation rating differs from the other two categories. For the majority of the sections for the year 2022, the annual rating and software detected more distress than manual observation. This was mostly caused by software and annual ratings that incorrectly identified distress, mistaking lane strip, road markings, crack sealing (maintenance activities) or raveling for the longitudinal cracking. On another note, at some instances, longitudinal cracking was not reported both by the software as well as the annual rating. This was primarily caused by the software's inadequacy to correctly detect the longitudinal cracking on the images of the pavement surface. In this situation, the manual observation rating exceeded the software and annual ratings results. For the year 2021, the software could not detect the distress on the section images and hence the ratings based on software were always zero.

2) **Annual rating is correct (Annual rating = Manual observation)**

Based on the image analysis, for very few sections it has been observed that the longitudinal cracking rating was the same for both annual rating and manual observation. From this, it can be concluded that the annual rating is correct since it has accounted for all the ACP longitudinal cracking for that road section.

3) **Software detection is correct (Software detection = Manual observation)**

Based on the image analysis, it has been observed that the ACP longitudinal cracking rating was the same for both software detection and manual observation. From this, it can be concluded that the software rating is correct since it has detected all the locations of longitudinal cracking for that road section. But this was mainly observed for the road sections with zero distress.

4) **Both annual rating and software detection are correct (Annual rating = Software detection = Manual observation)**

Based on the image analysis for the year 2022, the annual rating and software detection of ACP longitudinal cracking were found to be the same as manual observation. From this, it can be concluded that the annual rating and software detection are correct since both detected all the longitudinal cracking that was identified manually for that road section. This was true for 2021 road sections only with zero distress.

5) **Annual Rating > Software and Manual observation**

This case is a sub-type of case 1 where none of the annual rating and software detection are correct. In this scenario, the annual rating for longitudinal cracking was more than both software and manual observation. The reason for this difference may be due to incorrect distress identification or misinterpretation among distress types.

6) **Software > Annual Rating and Manual observation**

This case is also a sub-type of case 1 where none of the annual rating and software detection are correct. In this case, the software detects more locations for ACP longitudinal cracking because the software could not differentiate among the ACP longitudinal cracking, lane strips, road markings, and raveling. Additionally, the software also detects alligator or transverse cracking as longitudinal cracks. This leads to higher distress ratings by the software as compared to annual rating and manual observation.

7) **Manual observation > Software and Annual Rating**

This case is also a sub-type of case 1 where none of the annual rating and software detection are correct. In this case, the ACP longitudinal cracking was accurately identified manually for all locations. However, all these locations were not detected by the software and annual rating for the precise longitudinal cracking distress identification.

2. ACP Longitudinal Cracking from FY 2021

Longitudinal cracking from FY 2021 image check results are shown below.

A) **Observation:** None correct (95 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15 - Bexar	IH0010 X (1088)	589	0	589	0.5	40	0	11
15 - Bexar	IH0035 R (412)	141	0	141	0.5	65	4	10
15 - Bexar	IH0035 L (365)	140	0	140	0.5	237	0	3
15 - Bexar	IH0035 L (364)	142	0	142	0.5	96	0	50
95 - Guadalupe	IH0010 A (1083)	610	0	610	0.292	23	0	5

B) Observation: Software (correct) (117 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
83 - Frio	IH0035 R (795)	117	0	117	0.5	30	0	0
15 - Bexar	IH0035 L (771)	136	0	136	0.5	258	0	0
95 - Guadalupe	IH0010 L (462)	600	0	600	0.5	36	0	0
232 - Uvalde	US0090 K (414)	488	0	488	0.5	111	0	0
15 - Bexar	US0281 R	528	0	528	0.5	51	0	0

C) Observation: Both the Annual Rating and Software are correct (21 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15 – Bexar	US0281 L (36)	526	0	526	0.5	0	0	0
133 – Kerr	IH0010 R (600)	517	0	517	0.5	0	0	0
95 – Guadalupe	IH0010 R (460)	608	0	608	0.5	0	0	0
133 – Kerr	IH0010 R (599)	511	0	511	0.5	0	0	0
232 – Uvalde	RM2690 K (1298)	504	1.5	504	1.978	0	0	0

D) Observation: Annual Rating > Software and Manual observation (183 sections)

County	Highway And Roadbed	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation

	ID							
46 - Comal	FM0306 K (820)	504	0	504	0.5	47	0	19
163 - Medina	IH0035 L (774)	125	0	125	0.5	84	0	11
15 - Bexar	US0090 L (469)	570	0	570	0.5	67	0	11
15 - Bexar	US0090 X (1103)	560	0	560	0.5	30	0	2
46 - Comal	US0281 L (759)	512	0	512	0.5	106	3	18

E) Observation: Manual observation > Software and Annual Rating (13 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
131 - Kendall	IH0010 X (894)	524	0	524	0.578	2	0	12
15-Bexar	IH0010 X (1169)	579	0.5	579	1.034	0	0	12
95-Guadalupe	IH0010 R (460)	608	0.5	608	0.998	0	0	1
15-Bexar	FM0078 K (154)	504	0.5	501	1	24	4	48
15-Bexar	IH0010 A (546)	567	0.5	567	1.011	9	0	29

6.3.3 Transverse Cracking Image Check Result for Precision

1. ACP Transverse Cracking from FY 2022

A) Observation: None correct (48 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15-Bexar	US0090 L (1416)	570	0	570	0.5	5	5	4
95-Guadalupe	IH0010 R (892)	612	0	612	0.5	5	5	3
15-Bexar	IH0010 X (1087)	591	0	591	0.5	3	4	0
15-Bexar	IH0010 X (1087)	590	0	590	0.5	3	5	1
95-Guadalupe	UA0090 K	0	0	0	0.096	2	7	4

B) Observation: Annual rating (correct) (12 sections)

County	Highway	Beginning	Beginning	Ending	Ending TRM	Annual	Software	Manual

	And Roadbed ID	TRM Number	TRM Displacement	TRM Number	Displacement	rating		observation
15-Bexar	US0181 L (309)	512	0	512	0.5	3	4	3
15-Bexar	US0281 A (1172)	524	0	524	0.5	3	4	3
15-Bexar	IH0010 A (1340)	579	0	579	0.5	3	5	3
15-Bexar	IH0035 L (364)	144	0	144	0.5	1	4	1
15-Bexar	US0087 R (200)	706	0	706	0.5	0	1	0

C) **Observation:** Software (correct) (12 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
46-Comal	FM0725 K (1000)	0	0	0	0.009	13	0	0
15-Bexar	IH0035 L (364)	142	0	142	0.5	0	1	1
95-Guadalupe	FM1979 K (918)	532	0	532	0.211	1	0	0
15-Bexar	SS0122 K (1428)	506	0	506	0.5	1	2	2
15-Bexar	IH0410 L	4	0	4	0.5	3	6	6

D) **Observation:** Both the Annual Rating and Software are correct (31 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
46-Comal	BS0046CR (1008)	0	0	0	0.037	0	0	0
232-Uvalde	FM0030 K (688)	0	0	0	0.042	0	0	0
15-Bexar	US0281 X (1179)	524	0	524	0.5	4	4	4
15-Bexar	US0090 X (1224)	568	0	568	0.348	0	0	0
15-Bexar	US0090 R	572	0	572	0.5	6	6	6

E) **Observation:** Annual Rating > Software and Manual observation (9 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
46-Comal	FM0725 K	0	0	0	0.009	13	0	0

	(1000)							
15-Bexar	IH0010 X (1337)	580	0	580	0.153	6	5	3
95-Guadalupe	FM1979 K (918)	532	0	532	0.211	1	0	0
15-Bexar	IH0010 A (1208)	566	0.688	566	0.973	5	4	4
15-Bexar	SL1604 A (1115)	538	0.932	538	1.432	5	2	3

F) **Observation:** Software > Annual Rating and Manual observation (19 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15-Bexar	US0181 L (309)	512	0	512	0.5	3	4	3
15-Bexar	IH0010 A (1340)	579	0	579	0.5	3	5	3
95-Guadalupe	IH0010 L (897)	613	0	613	0.5	5	6	4
15-Bexar	IH0035 L (364)	144	0	144	0.5	1	4	1
15-Bexar	IH0035 R (413)	147	0	147	0.5	0	2	1

G) **Observation:** Manual observation > Annual Rating and Software (1 section)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15-Bexar	IH0010 X	593	0	593	0.592	5	5	9

Like ACP alligator cracking and longitudinal cracking, the distress ratings for transverse cracking are also divided into three categories: a) Annual rating, b) Software detection, and c) Manual observation. Also, based on the mentioned categories, there are seven cases same as that for alligator and longitudinal cracking. The detailed discussion of the seven cases is as follows:

1) **None of the 3 distress rating categories are correct (Annual rating ≠ Software detection ≠ Manual observation)**

It has been observed that the distress rating varies for each category of observation based on an image analysis. This is mostly because the manual observation rating differs from the other two categories. For the majority of the sections for the year 2022, the annual rating and software detected more transverse cracking than manual observation. This was mostly caused by software and annual ratings that incorrectly identified distress, mistaking sealed cracks (maintenance activities) or raveling for the transverse cracking. On another note, for the year 2021, the software could not detect the distress on the pavement surface images and hence the ratings based on the software were always zero due to its inadequacy to correctly detect the transverse cracking.

2) **Annual rating is correct (Annual rating = Manual observation)**

Based on the image analysis, for very few sections it has been observed that the transverse cracking rating was the same for both annual rating and manual observation. From this, it can be concluded that the annual rating is correct since it has reported all the ACP transverse cracking for the road sections.

3) **Software detection is correct (Software detection = Manual observation)**

In this case, it has been observed that the transverse cracking rating was the same for both software detection and manual observation. From this, it can be concluded that the software rating is correct since it has detected all the locations of longitudinal cracking for that road section. However, this condition has been mainly reflected in many road sections with zero distress in the year 2021.

4) **Both annual rating and software detection are correct (Annual rating = Software detection = Manual observation)**

Based on the image analysis for the year 2022, the annual rating and software detection of ACP transverse cracking was found to be the same as manual observation. From this, it can be concluded that the annual rating and software detection are correct since both detected all the transverse cracking that was identified manually for the road sections. This condition was true for 2021 road sections only with zero transverse cracks.

5) **Annual Rating > Software and Manual observation**

This case is a sub-type of case 1 where none of the annual rating and software detection are correct. In this situation, the annual rating for transverse cracking was more than both software and manual observation. This was observed for many sections (2021) and the reason for this difference may be due to incorrect distress identification or misinterpretation among distress types.

6) **Software > Annual Rating and Manual observation**

This case is also a sub-type of case 1 where none of the annual rating and software detection are correct. In this case, the software detects more locations for ACP transverse cracking because the software could not differentiate between the ACP transverse cracking and raveling for some road sections. This leads to higher distress ratings by the software as compared to annual rating and manual observation.

7) **Manual observation > Software and Annual Rating**

This scenario is a subtype of case 1, where both the annual rating and software detection are incorrect. In this instance, ACP transverse cracking was manually identified accurately at all locations. However, the software and annual rating failed to detect the transverse cracking distress at these locations.

2. ACP Transverse Cracking from FY 2021

Transverse cracking image check results are listed below.

A) **Observation:** None correct (26 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15-Bexar	US0090 X (1224)	568	0	568	0.348	5	5	0
95-Guadalupe	IH0010 L (897)	613	0	613	0.5	7	0	1
15-Bexar	IH0035 L (364)	143	0	143	0.5	3	0	2
15-Bexar	IH0010 A (1208)	567	0	567	0.5	0	0	2
15-Bexar	SS0122 K (1428)	506	0	506	0.5	3	0	5

B) **Observation:** Annual rating (correct) (9 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15-Bexar	US0090 L (1416)	570	0	570	0.5	2	0	2
131 - Kendall	IH0010 R (1381)	523	0	523	0.5	2	0	2
95-Guadalupe	UA0090 K (173)	0	0	0	0.096	4	0	4
15-Bexar	SL0353 K (1253)	502	0	502	0.5	1	0	1
15-Bexar	US0181 L (309)	512	0.5	512	1	1	0	1

C) **Observation:** Software (correct) (54 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
46-Comal	BS0046CR (1008)	0	0	0	0.037	2	0	0
232-Uvalde	FM0030K (688)	0	0	0	0.042	2	0	0
46-Comal	FM0725 K (1000)	0	0	0	0.009	6	0	0
15-Bexar	US0281 X (1179)	524	0	524	0.5	1	0	0

15-Bexar	US0281 X (1180)	522	0	522	0.478	2	0	0

D) Observation: Both the Annual Rating and Software are correct (13 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15-Bexar	US0181 L (309)	512	0	512	0.5	0	0	0
15-Bexar	IH0010 A (1340)	579	0	579	0.5	0	0	0
15-Bexar	IH0035 R (412)	141	0	141	0.5	0	0	0
15-Bexar	IH0010 X (1337)	580	0	580	0.153	0	0	0
15-Bexar	IH0010 X (1337)	579	0	579	0.5	0	0	0

E) Observation: Annual Rating > Software and Manual observation (62 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
46-Comal	BS0046CR (1008)	0	0	0	0.037	2	0	0
46-Comal	FM0725 K (1000)	0	0	0	0.009	6	0	0
15-Bexar	US0281 K (1180)	522	0	522	0.478	2	0	0
95-Guadalupe	IH0010 R (892)	612	0	612	0.5	2	0	0
95-Guadalupe	IH0010 L (897)	613	0	613	0.5	7	0	1

F) Observation: Software > Annual Rating and Manual observation (0 sections)

G) Observation: Manual observation > Annual Rating and Software (15 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15-Bexar	US0090 R (1417)	568	1.951	568	2.39	0	0	2
15-Bexar	SL0013 K (1439)	492	1.944	492	1.97	0	0	3
15-Bexar	FM2696 L (932)	484	1.578	484	1.916	0	0	1
15-Bexar	SS0122 K (1428)	504	1	504	1.5	4	0	6

15-Bexar	IH0010 A (1208)	566	0.688	566	0.973	0	0	3

6.3.4 Patching Image Check Result for Precision

1. ACP Patching from FY 2022

A) Observation: None correct (14 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15-Bexar	FM0078 K (1273)	508	0	508	0.5	6	0	1
15-Bexar	IH0035 L (364)	143	0	143	0.5	13	0	21
7-Atascosa	FM0791 K (229)	514	0	514	0.5	22	0	18
83-Frio	FM2779 K (474)	530	0	530	0.5	14	0	33
7- Atascosa	FM2924 K (224)	514	0	514	0.5	2	0	6

B) Observation: Annual rating (correct) (8 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
46-Comal	FM2779 K (474)	528	1	528	1.5	22	0	22
232-Uvalde	US0090 K (732)	484	1	484	1.5	14	0	14
15-Bexar	IH0035 R (413)	146	0.5	146	1.006	37	0	37
83-Frio	FM2779 K (474)	528	0.5	528	1	17	0	17
247-Wilson	FM1681 K (18)	530	0	530	0.5	2	0	2

C) Observation: Software (correct) (46 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
95-Guadalupe	IH0010 R (1077)	600	0	600	0.5	20	0	0
15-Bexar	IH0035 R (413)	145	0	145	0.5	2	0	0
15-Bexar	FM1628 K (158)	502	0	502	0.5	13	0	0

247-Wilson	FM3161 K (101)	526	0	526	0.5	28	0	0
162-Mcmullen	FM1962 K (28)	496	0	496	0.5	1	0	0

D) Observation: Both the Annual Rating and Software are correct (82 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
95-Guadalupe	Fm0020 k (1017)	530	0	530	0.5	0	0	0
10-Bandera	FM0470 K (692)	446	0	446	0.5	0	0	0
247-Wilson	FM0537 K (68)	522	0	522	0.5	0	0	0
46-Comal	FM0311 K (876)	502	0	502	0.5	0	0	0
15-Bexar	US0090 X (1224)	568	0	568	0.348	0	0	0

E) Observation: Annual rating > Software and Manual observation (51 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15-Bexar	FM0078 K (1273)	508	0	508	0.5	6	0	1
95-Guadalupe	IH0010 R (1077)	600	0	600	0.5	20	0	0
15-Bexar	IH0035 R (413)	145	0	145	0.5	2	0	0
15-Bexar	FM1628 K (158)	502	0	502	0.5	13	0	0
7-Atascosa	FM0791 K (229)	514	0	514	0.5	22	0	18

F) Observation: Software > Annual Rating and Manual observation (0 section)

G) Observation: Manual observation > Annual Rating and Software (5 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15-Bexar	IH0035 L (364)	143	0	143	0.5	13	0	21
7-Atascosa	FM0791 K (228)	510	0	510	0.5	13	0	18
83-Frio	FM2779 K (474)	530	0	530	0.5	14	0	33

7-Atascosa	FM2924 K (224)	514	0	514	0.5	2	0	6
247-Wilson	FM1344 K (221)	528	0	528	1.5	10	0	14

Similar to ACP cracking (alligator, longitudinal, and transverse), the distress ratings for ACP patching are divided into three categories: a) Annual rating, b) Software detection, and c) Manual observation. Likewise, based on the stated categories, there are seven cases same as that for ACP cracking. The detailed discussion of the seven cases is as follows:

1) **None of the 3 distress rating categories are correct (Annual rating ≠ Software detection ≠ Manual observation)**

It has been observed that the distress rating varies for each category of observation based on an image analysis. This is mostly because the manual observation rating differs from the other two categories. For the years 2021 and 2022, the software could not detect the distress on the section images and hence the ratings based on the software were always zero due to its inadequacy to correctly detect the patching on the pavement surface.

2) **Annual rating is correct (Annual rating = Manual observation)**

Based on the image analysis, for a very few sections, it has been observed that the ACP patching rating was the same for both annual rating and manual observation. From this, it can be concluded that the annual rating is correct since it has reported all the ACP patching locations on the road sections wherein suitable maintenance activities have been performed.

3) **Software detection is correct (Software detection = Manual observation)**

This case has been observed for the majority of the pavement sections wherein the ACP patching rating was the same for both software detection and manual observation. However, this condition has been mainly reflected for many road sections with zero patching in both years. The software was inadequate to detect the patching on the pavement surface.

4) **Both annual rating and software detection are correct (Annual rating = Software detection = Manual observation)**

Based on the image analysis, the annual rating and software detection of ACP patching were found to be the same as manual observation. This was true for 2021 and 2022 road sections only with zero patching.

5) **Annual Rating > Software and Manual observation**

This case is a sub-type of case 1 where none of the annual rating and software detection are correct. In this situation, the annual rating for ACP patching was more than both software and manual observation. This was observed for many sections (2021 and 2022) and the reason for this difference could be due to incorrect distress identification or misinterpretation among distress types.

6) **Software > Annual Rating and Manual observation**

This case has not been observed for ACP patching for both 2021 and 2022 pavement sections.

7) Manual observation > Software and Annual Rating

This scenario is a subtype of case 1, where both the annual rating and software detection are incorrect. In this instance, ACP patching was manually identified accurately at all locations. However, the software and annual rating failed to detect the patching at these locations. This case was observed only for five sections for 2021 and 2022.

2. ACP Patching from FY 2021

The ACP patching image check results are presented below.

A) **Observation:** None correct (18 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
10-Bandera	FM0470 K (1030)	446	0	446	0.5	18	0	9
247-Wilson	FM0537 K (365)	522	0	522	0.5	18	0	1
15-Bexar	IH0035 L (770)	143	0	143	0.5	0	0	1
247-Wilson	FM1681 K (1165)	530	0	530	0.5	14	0	1
83-Frio	FM1582 K (189)	554	0	554	0.5	11	0	2

B) **Observation:** Annual rating (correct) (3 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
247-Wilson	FM1344 K (351)	532	0	532	0.5	16	0	16
131-Kendall	FM3351 K (828)	468	0.5	468	1	3	0	3
46-Comal	FM1863 K (969)	496	1.5	496	1.9	10	0	10

C) **Observation:** Software (correct) (89 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
95-Guadalupe	FM0020 K (1396)	530	0	530	0.5	18	0	0
15-Bexar	FM0078 K (154)	508	0	508	0.5	16	0	0

95-Guadalupe	FM0020 K (1397)	540	0	540	0.261	12	0	0
46-Comal	FM0311 K (1340)	502	0	502	0.5	11	0	0
15-Bexar	US0090 X (1311)	568	0	568	0.348	21	0	0

D) **Observation:** Both the Annual Rating and Software are correct (39 Sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
95-Guadalupe	IH0010 R (459)	600	0	600	0.5	0	0	0
15-Bexar	FM1628 K (1144)	502	0	502	0.5	0	0	0
7-Atascosa	FM0791 K (250)	514	0	514	0.5	0	0	0
7-Atascosa	FM0791 K (250)	510	0	510	0.5	0	0	0
247-Wilson	FM3161 K (347)	526	0	526	0.5	0	0	0

E) **Observation:** Annual rating > Software and Manual observation (104 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
95-Guadalupe	FM0020 K (1396)	530	0	530	0.5	18	0	0
15-Bexar	FM0078 K (154)	508	0	508	0.5	16	0	0
95-Guadalupe	FM0020 K (1397)	540	0	540	0.261	12	0	0
46-Comal	FM0311 K (1340)	502	0	502	0.5	11	0	0
15-Bexar	US0090 X (1311)	568	0	568	0.348	21	0	0

F) **Observation:** Software > Annual Rating and Manual observation (0 sections)

G) **Observation:** Manual observation > Annual Rating and Software (5 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
10-Bandera	FM0470 K (1029)	436	0	436	0.5	10	0	18
15-Bexar	IH0035 L (770)	143	0	143	0.5	0	0	1

232-Uvalde	RM0187 K (1019)	520	0	520	0.5	17	0	26
232-Uvalde	US0090 K (414)	484	1	484	1.5	0	0	6
83-Frio	FM2779 K (1202)	528	1	528	1.5	0	0	8

6.3.5 Failure Check Result for Precision

1. ACP Failure from FY 2022

A) Observation: None correct (32 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
7-Atascosa	FM0140 K (272)	490	0	490	0.37	8	0	2
163-Medina	FM0462 K (596)	508	0	508	0.5	3	0	8
7-Atascosa	IH0035 A (443)	132	0	132	0.739	3	1	5
163-Medina	FM1796 K (701)	448	0	448	0.5	8	0	13
232-Uvalde	FM1052 K (753)	514	0	514	0.5	0	0	1

B) Observation: Annual rating (correct) (6 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
7-Atascosa	FM1099 K (227)	508	0	508	0.5	4	0	4
163-Medina	FM2200 K (535)	506	0	506	0.5	3	0	3
15-Bexar	IH0037 X (1423)	127	0	127	0.5	3	0	3
7-Atascosa	FM2924 K (224)	514	0.5	514	1	13	0	13
163-Medina	FM0462 K (597)	500	0.5	500	1	4	0	4

C) Observation: Software (correct) (33 sections)

County	Highway And Roadbed	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation

	ID							
95-Guadalupe	IH0010 R (892)	613	0	613	0.5	5	0	0
95-Guadalupe	IH0010 A (84)	612	0	612	0.5	9	0	0
15-Bexar	IH0035 A (444)	136	0	136	0.5	2	0	0
15-Bexar	FM1518 K (206)	494	0	494	0.5	2	0	0
131-Kendall	FM1621 K (825)	470	0	470	0.5	1	0	0

D) Observation: Both the Annual Rating and Software are correct (40 Sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15-Bexar	FM0327 K (212)	508	0	508	0.5	0	0	0
163-Medina	IH0035 A (442)	129	0	129	0.5	0	0	0
15-Bexar	IH0035 A (444)	137	0	137	0.5	0	0	0
7-Atascosa	IH0037 R (1319)	106	0	106	0.5	0	0	0
83-Frio	IH0035 X (401)	93	0	93	0.5	0	0	0

E) Observation: Annual rating > Software and Manual observation (45 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
7-Atascosa	FM0140 K (272)	490	0	490	0.37	8	0	2
95-Guadalupe	IH0010 R (892)	613	0	613	0.5	5	0	0
95-Guadalupe	IH0010 A (84)	612	0	612	0.5	9	0	0
163-Medina	IH0035 A (440)	122	0	122	0.79	9	0	1
163-Medina	FM2200 K (535)	510	0	510	0.5	12	0	5

F) Observation: Software > Annual Rating and Manual observation (0 sections)

G) Observation: Manual observation > Annual Rating and Software (30 sections)

County	Highway And Roadbed	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation

	ID							
163-Medina	FM0462 K (596)	508	0	508	0.5	3	0	8
7-Atascosa	IH0035 A (443)	132	0	132	0.739	3	1	5
163-Medina	FM1796 K (701)	448	0	448	0.5	8	0	13
232-Uvalde	FM1052 K (753)	514	0	514	0.5	0	0	1
163-Medina	FM0462 K (596)	504	0	504	0.5	2	0	16

Similar to ACP cracking (alligator, longitudinal, and transverse) and patching, the rating for ACP failure quantity is divided into three categories: a) Annual rating, b) Software detection, and c) Manual observation. Based on these listed categories, there are seven cases same as that for ACP cracking and patching. The detailed discussion of the seven cases is as follows:

1) **None of the 3 distress rating categories are correct (Annual rating ≠ Software detection ≠ Manual observation)**

It has been observed that the distress rating varies for each category of observation based on an image analysis. This is mostly because the manual observation rating differs from the other two categories. For the years 2021 and 2022, the software could not detect the failure quantity on the section images and hence the ratings based on software were always zero (except only for one section) due to its inadequacy to correctly detect the failure quantity on the pavement surface.

2) **Annual rating is correct (Annual rating = Manual observation)**

Based on the image analysis, for a very few sections, it has been observed that the ACP failure rating was the same for both annual rating and manual observation. From this, it can be concluded that the annual rating is correct since it has reported all the ACP failure quantity on the road sections.

3) **Software detection is correct (Software detection = Manual observation)**

This case has been observed for the pavement sections wherein the ACP failure rating was the same for both software detection and manual observation. However, this condition has been mainly reflected for many road sections with zero failure in both years. The software was incapable to detect the failure quantity on the pavement surface.

4) **Both annual rating and software detection are correct (Annual rating = Software detection = Manual observation)**

Based on the image analysis, the annual rating and software detection of ACP failure were found to be the same as manual observation. This was true for 2021 and 2022 road sections only with zero failure.

5) **Annual Rating > Software and Manual observation**

This case is a sub-type of case 1 where none of the annual rating and software detection are correct. In this situation, the annual rating for ACP failure was more than both software and manual observation. This was observed for many sections (2021 and 2022) and the reason for this difference may be due to incorrect distress identification.

6) Software > Annual Rating and Manual observation

This case has not been observed for ACP failure for both 2021 and 2022 pavement sections.

7) Manual observation > Software and Annual Rating

This scenario is a subtype of case 1, where both the annual rating and software detection are incorrect. In this case, ACP failure was manually identified accurately at all locations. However, the annual rating failed to detect all the failure locations. Also, the software could not detect the failure locations on the pavement surface images for both years. This situation leads to underestimation of the failure quantity by the annual rating and software.

2. ACP Failure from FY 2021

The ACP failure image check results are listed below.

A) Observation: None correct (20 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
163-Medina	FM0462 K (849)	492	0	492	0.5	0	0	2
15-Bexar	FM0327 K (1370)	508	0	508	0.5	3	0	1
95-Guadalupe	IH0010 R (1466)	613	0	613	0.5	0	0	2
15-Bexar	IH0035 A (1040)	137	0	137	0.5	2	0	1
131-Kendall	FM1621 K (899)	470	0	470	0.5	3	0	5

B) Observation: Annual rating (correct) (9 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
163-Medina	FM1796 K (1436)	448	0	448	0.5	10	0	10
163-Medina	FM2200 K (1434)	506	0	506	0.5	1	0	1
83-Frio	IH0035 X	83	0	83	0.5	2	0	2
10-Bandera	RM1077 K (1252)	480	0	480	0.5	2	0	2
95-Guadalupe	BS0123B K (1469)	488	0	488	1	3	0	3

C) Observation: Software (correct) (22 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
7-Atascosa	FM0140 K (207)	490	0	490	0.37	4	0	0
163-Medina	IH0035 A (1236)	129	0	129	0.5	2	0	0
7-Atascosa	FM1099 K (253)	508	0	508	0.5	1	0	0
95-Guadalupe	FM1150 K (1398)	540	0	540	0.5	4	0	0
7-Atascosa	FM2924 K (255)	514	0	514	0.5	5	0	0

D) Observation: Both the Annual Rating and Software are correct (40 Sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
163-Medina	FM0462 K (847)	508	0	508	0.5	0	0	0
95-Guadalupe	IH0010 A (1463)	612	0	612	0.5	0	0	0
7-Atascosa	IH0035 A	132	0	132	0.739	0	0	0
163-Medina	IH0035 A	122	0	122	0.79	0	0	0
15-Bexar	FM1518 K (746)	494	0	494	0.5	0	0	0

E) Observation: Annual rating > Software and Manual observation (30 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
15-Bexar	FM0327 K (1370)	508	0	508	0.5	3	0	1
163-Medina	IH0035 A (1236)	129	0	129	0.5	2	0	0
15-Bexar	IH0035 A (1040)	137	0	137	0.5	2	0	1
7-Atascosa	FM0140 K (207)	490	0	490	0.37	4	0	0
7-Atascosa	FM1099 K (253)	508	0	508	0.5	1	0	0

F) Observation: Software > Annual Rating and Manual observation (0 sections)

G) Observation: Manual observation > Annual Rating and Software (14 sections)

County	Highway And Roadbed ID	Beginning TRM Number	Beginning TRM Displacement	Ending TRM Number	Ending TRM Displacement	Annual rating	Software	Manual observation
163-Medina	FM0462 K (849)	492	0	492	0.5	0	0	2
163-Medina	FM0462 K (847)	504	0	504	0.5	0	0	13
95-Guadalupe	IH0010 R (1466)	613	0	613	0.5	0	0	2
131-Kendall	FM1621 K (899)	470	0	470	0.5	3	0	5
163-Medina	FM1250 K (1441)	450	0	450	0.5	1	0	12

In summary, the distress ratings for pavement conditions are evaluated through annual ratings, software detection, and manual observation methods. The analysis indicates that annual ratings and software detection can either underestimate or overestimate ACP (alligator, longitudinal, and transverse) cracking and failure when compared to manual observations. This discrepancy might arise from difficulties in accurately identifying the type of distress or from misinterpretations among the various distress types. Additionally, overestimation can result from the incorrect identification of lane stripes and raveling as cracking.

Moreover, the ratings generated by software detection are not consistently aligned with those from annual ratings and manual observations. These inaccuracies in distress identification can significantly impact maintenance and rehabilitation strategies, leading to incorrect budget allocations over the pavement service lives.

Chapter 7 Conclusions and Recommendations

7.1 Conclusions

This study aimed to understand the data quality of automated pavement condition data by evaluating the accuracy and reliability of automated pavement condition data. To implement the data quality evaluation, the research team selected data quality check indexes, developed data quality thresholds, designed implementation procedures, and validated the data quality check results using raw images. The primary focus was on developing data quality thresholds and designing data quality consistency check procedures to locate the pavement data with potential data quality issues and then correct the pavement condition to improve the accuracy and precision of the automated pavement condition data.

Based on the analysis conducted, the following conclusions are drawn:

1. **Problems with Existing Automated Pavement Condition Data Collection**: the literature review and questionnaire survey indicate that automated pavement condition data collection is the commonly accepted data collection method that employed by highway agencies. Automated pavement data collection technologies encounter data quality issues such as inconsistency, discrepancy, and false positives, impacting data validation and vendor collaboration. Despite not being fully automated, these technologies rely heavily on manual labor for inspection, achieving accuracies of 70-80% instead of the desired 95%. While AI applications like deep learning algorithms show potential, challenges like data pre-treatment and limited ground truth data hinder progress. Manual validation remains essential, with state agencies investing resources and time in this process, often outsourcing to third-party contractors. Implementing measurable QA procedures and establishing data quality benchmarks are recommended to enhance operational efficacy. Addressing these obstacles is imperative for optimizing the effectiveness of automated data collection technologies in pavement engineering.

2. **A comprehensive analysis of historical pavement condition data**: the historical data analysis focuses on 25 districts and three pavement types (ACP, CRCP, and JCP) spanning FY 2017 to FY 2021. Utilizing automated data collection technologies, the dataset offers a detailed breakdown of pavement sections per district. The normality test on distress scores across five districts indicated deviation from a normal distribution, though the distress score differences over two years followed a normal distribution, hinting at the potential for future analysis. Through merging PMIS annual ratings with audit data, accuracy and precision analyses were conducted, defining accuracy as proximity to the true value and precision as measurement consistency. The chapter elucidates the application of confusion matrices for model performance visualization and metric calculation, illustrating the significance of accuracy and precision in interpreting pavement condition data for informed decision-making in pavement management.

3. **Stratified Sampling Method** highlights the importance of using this method for data quality audits, especially in cases of high variability among population units. The process involves defining strata based on key factors, allocating sample sizes optimally to each stratum to minimize costs and maximize information gain, and iteratively updating sample sizes as needed. By carefully considering population characteristics, variance, and cost implications, researchers can design a stratified sampling method that accurately represents the population and enhances the accuracy and reliability of audit results in pavement condition assessments.

4. **Threshold Development of Data Consistency Check:** the accuracy check was designed using automated data comparing with audit data of distresses and DS in the same fiscal year. Similarly, precision analysis was conducted by examining the disparities in individual distress and comprehensive scores across two consecutive years. The establishment of thresholds was achieved by scrutinizing the variation in distress measurements between annual ratings and audits/two consecutive years, with a focus on identifying data points encompassing 87% of the dataset or falling within 1.5 standard deviations from the mean. Procedures were devised for executing the data quality check process based on the established thresholds.

5. **Pilot Study of Data Quality Assurance**: The San Antonio district was chosen as the site for a pilot study on pavement condition data quality assurance, which employed the established thresholds and devised procedures. A raw image inspection of the roadway sections was performed to validate the suggested data quality thresholds and procedures. The findings indicate that the proposed data quality assurance approach effectively identifies sections with potential data quality concerns.

7.2 Recommendations

Based on the conclusions derived from the questionnaire survey, historical pavement condition data analysis, sampling method development, and implementation of data quality assurance, the following recommendations are proposed to enhance the quality of automated pavement condition data collection.

1. **Development of a Unified Rating System**: Establishing a unified rating system that integrates both automated and manual observations can help minimize discrepancies. This system should incorporate the strengths of both methods to provide a more accurate and consistent evaluation of pavement conditions.
2. **Regular Audits and Validation Checks**: Conducting regular audits and validation checks of the collected data is crucial. Implementing a systematic approach to auditing can help identify and rectify data quality issues promptly. This includes using the developed cost-efficient audit sampling method and data-enabled quality acceptance criteria.
3. **Continuous Improvement of Data Quality Management Plans**: Updating and continuously improving the data quality management plans based on the latest research

and technological advancements is vital. This ensures that the data collection methods remain effective and reliable over time.

By implementing these recommendations, TxDOT can significantly improve the accuracy, precision, and reliability of automated pavement condition data. This will ultimately lead to better maintenance and rehabilitation strategies, optimal resource allocation, and improved pavement performance across the state.

These conclusions and recommendations form a comprehensive strategy to address the identified challenges and enhance the overall effectiveness of automated pavement condition data collection in Texas.

References

Pierce, L.M., and Weitzel, N.D. *NCHRP Synthesis 531: Automated Pavement Condition Surveys, Report No. Project 20-05, Topic 49-15.* National Cooperative Highway Research Program. Transportation Research Board, Washington, D.C., 2019.

McGhee, K.H. NCHRP Synthesis 334: Automated Pavement Distress Collection Techniques, Report No. *Project 20-5, Topic 34-04. National Cooperative Highway Research Program. Transportation Research B*oard, Washington, D.C., 2004.

Flintsch, G. W., and McGhee, K.H. *Quality Management of Pavement Condition Data Collection. Vol. 401.* Transportation Research Board, Washington, D.C., 2009.

Pierce, L. M., McGovern, G and Zimmerman, K. *Practical Guide for Quality Management of Pavement Condition Data Collection*. Federal Highway Administration, Washington, D.C.,2013.

FHWA (Federal Highway Administration). *Guidelines for Development and Approval of State Data Quality Management Programs*. Washington, D.C., 2018.

Simpson, A., Rada, G., Bryce, J., Serigos, P., Visintine, B, and Groeger, J. *Interstate Highway Pavement Sampling Data Quality Management Plan, Report No. FHWA-HIF-18-032*. Federal Highway Administration, U.S. Department of Transportation. Washington, D.C., 2018.

AASHTO (American Association of State Highway and Transportation Officials). *Standard Practice for Definition of Terms Related to Quality and Statistics as Used in Highway Construction*. AASHTO R 10-06. American Association of State Highway and Transportation Officials, Washington, D.C., 2006.

Chang, George, Amanda Gilliland, Gonzalo R. Rada, Pedro A. Serigos, Amy L. Simpson, and Sareh Kouchaki. *Successful Practices for Quality Management of Pavement Surface Condition Data Collection and Analysis Phase I: Task 2-Document of Successful Practices*. FHWA-RC-20-0007. Federal Highway Administration, Washington, D.C., 2020.

FHWA (Federal Highway Administration). *National Performance Management Measures; Assessing Pavement Condition for the National Highway Performance Program and Bridge Condition for the National Highway Performance Program*. Washington, D.C., 2017.

Simpson, A., P. Serigos, S. Kouchaki, G. Rada, and Groeger, J. *Interstate Highway Pavement Sampling: Final Phase 2 Report. Report No. FHWA-HIF-20-085*. Federal Highway Administration, U.S. Department of Transportation. Washington, D.C., 2020.

TxDOT (Texas Department of Transportation). *Quality Management Plan for Pavement Data Collection, Texas Department of Transportation*, Maintenance Division, Pavement Preservation Branch, Austin, TX, 2018.

Georgia Department of Transportation (GDOT), *Long-Term Pavement Performance Monitoring and Management Plan*, 2020.

Arkansas Department of Transportation (ARDOT), *Pavement Performance Data Quality Management Program*, 2018.

California Department of Transportation (Caltrans), *Data Quality Management Plan*, 2018.

Delaware Department of Transportation (DelDOT), *Pavement Distress Data Quality Management Plan*, 2020.

Florida Department of Transportation (FDOT), *Data Quality Management Plan for Pavement Condition Data Collection*, 2021.

Illinois Department of Transportation (IDOT), *National Performance Management Measures Data Quality Management Program*, 2018.

Indiana Department of Transportation (INDOT), *Pavement Condition Data Quality Management Plan*, 2018.

Alabama Department of Transportation (ALDOT), *Network-Level Pavement Condition Data Collection Procedure*, 2020.

Alaska Department of Transportation & Public Facilities (AK DOT&PF), *State of Alaska Highway Distress Data Collection Quality Management Plan*, 2018.

Kentucky Transportation Cabinet (KYTC), *Data Quality Management Plan*.

Louisiana Department of Transportation and Development (LADOTD), *Quality Control by Data Collection Vendor*.

Minnesota Department of Transportation (MnDOT), *Pavement Data Quality Management Plan*, 2018.

Mississippi Department of Transportation (MDOT), *Development of a Pavement Management Manual and Data Quality Plan*, 2017.

Nebraska Department of Transportation (NDOT), *Data Quality Management Program for the Collection of Pavement Condition,* 2018.

New Hampshire Department of Transportation (NHDOT), *Pavement Data Quality Management Program*, 2020.

New Jersey Department of Transportation (NJDOT), Asset Management Plan, *Pavement Data Quality Management Programs*, 2019.

New York State Department of Transportation (NYSDOT), *Quality Management Plan for Pavement Condition Data Collection*, 2018.

North Dakota Department of Transportation (NDDOT), *Network-Level Pavement Condition Data Collection Quality Management Plan*, 2018.

Oregon Department of Transportation (ODOT), *Data Quality Management Plan for Pavement Condition*, 2018.

Pennsylvania Department of Transportation (PennDOT), *Data Quality Management Program*, 2018.

South Dakota Department of Transportation (SDDOT), *Quality Management Plan for Network Level Pavement Condition Data Collection*, 2018.

Tennessee Department of Transportation (TDOT), *Guidelines on Quality Management of Pavement Condition Data*, 2018.

Washington State Department of Transportation (WSDOT), *Pavement Condition Data Collection Quality Management Plan*, 2018.

Wyoming Department of Transportation (WYDOT), *Network-Level Pavement Condition Data Collection Quality Management Plan*, 2019.

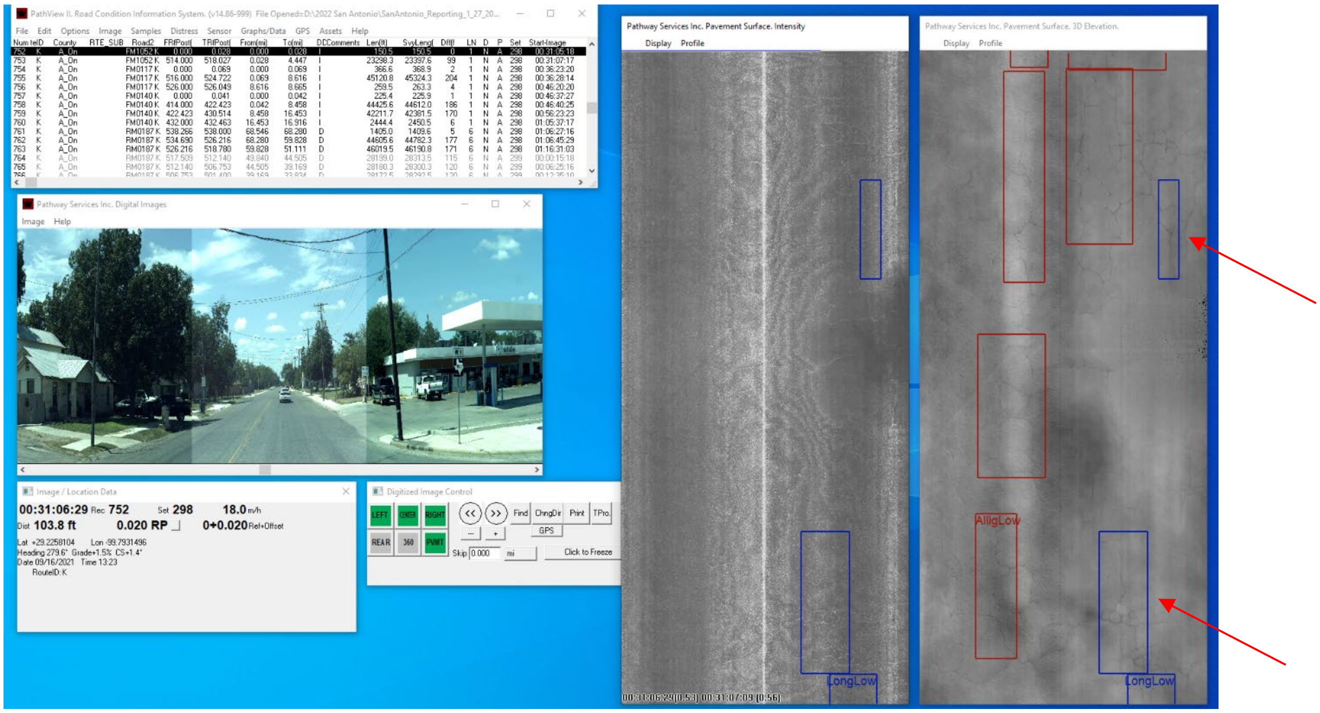
Maryland Department of Transportation (MDOT), *Pavement Data Quality Management Program*, 2018.

Colorado Department of Transportation (CDOT), *Pavement Data Quality Management Program*, 2018.

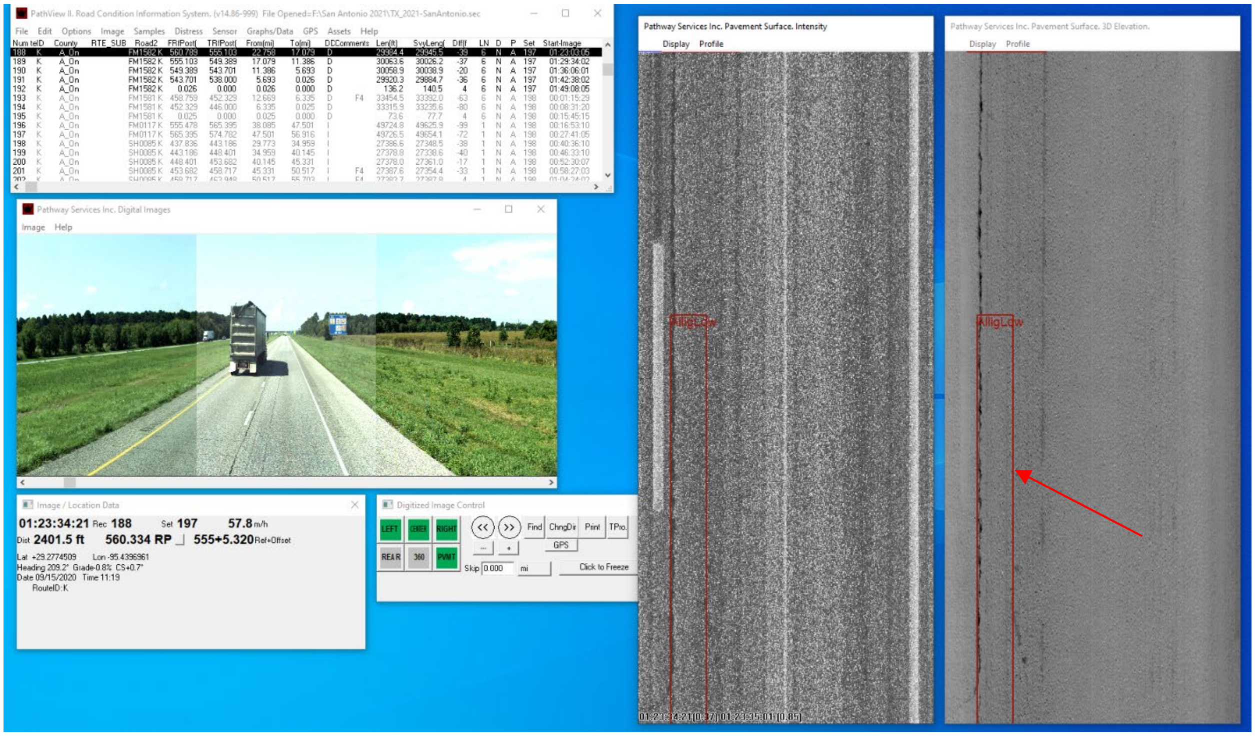
Appendix: Image Examples

There are differences in the distress detection by the software based on the following cases:

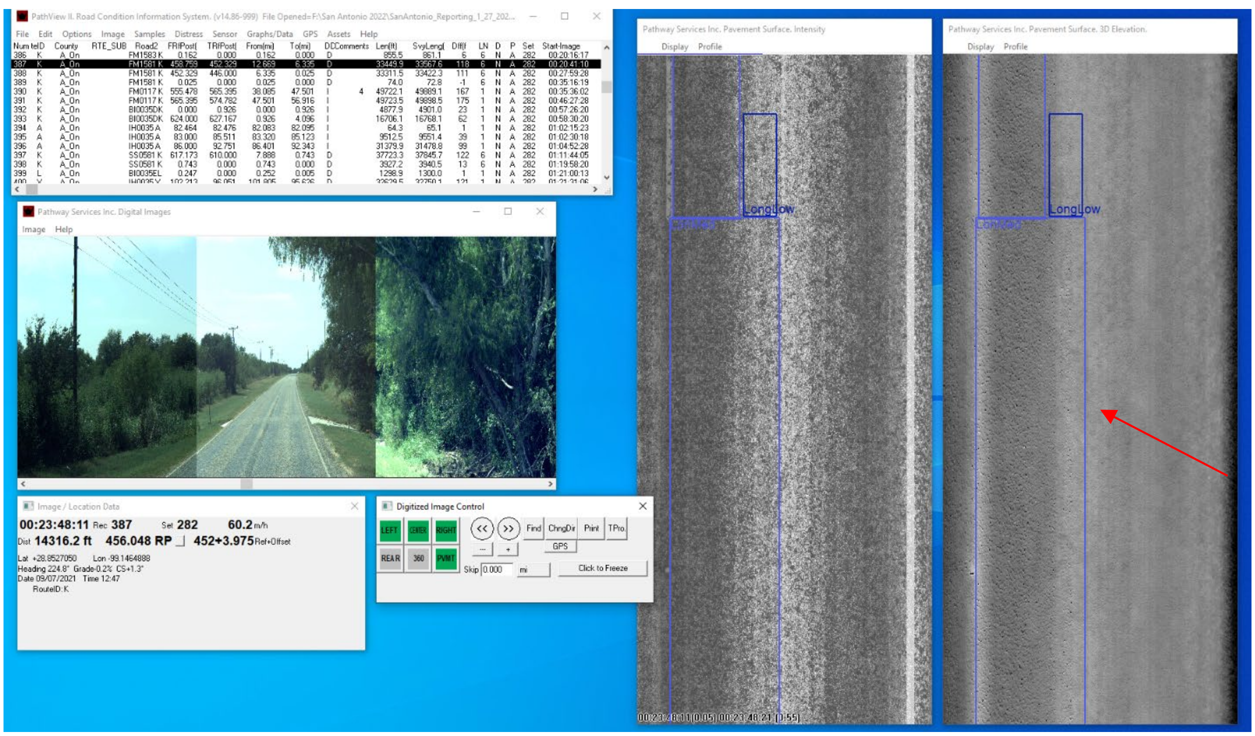
Case 1: Software identified alligator cracking as longitudinal cracking (blue box)



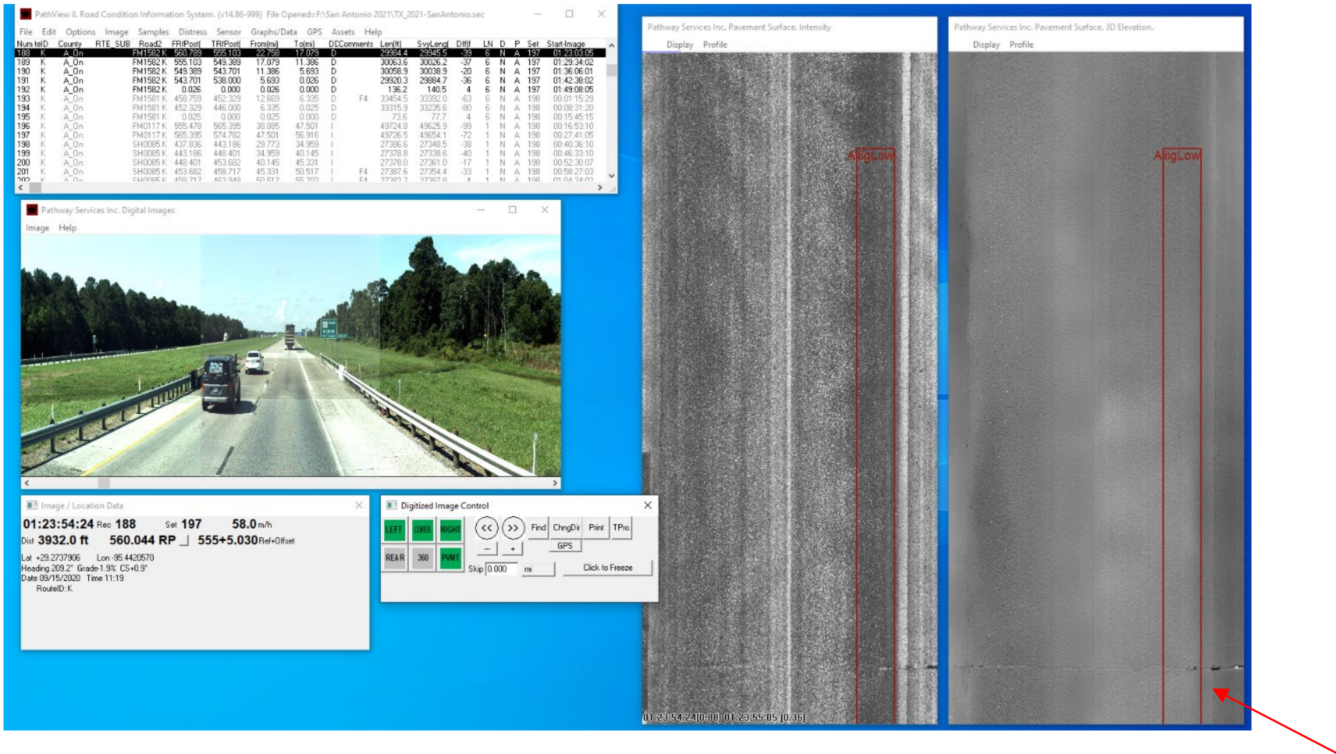
Case 2: Software identified longitudinal cracking as alligator cracking (red box)



Case 3: Software identified raveling as longitudinal cracking (blue box)

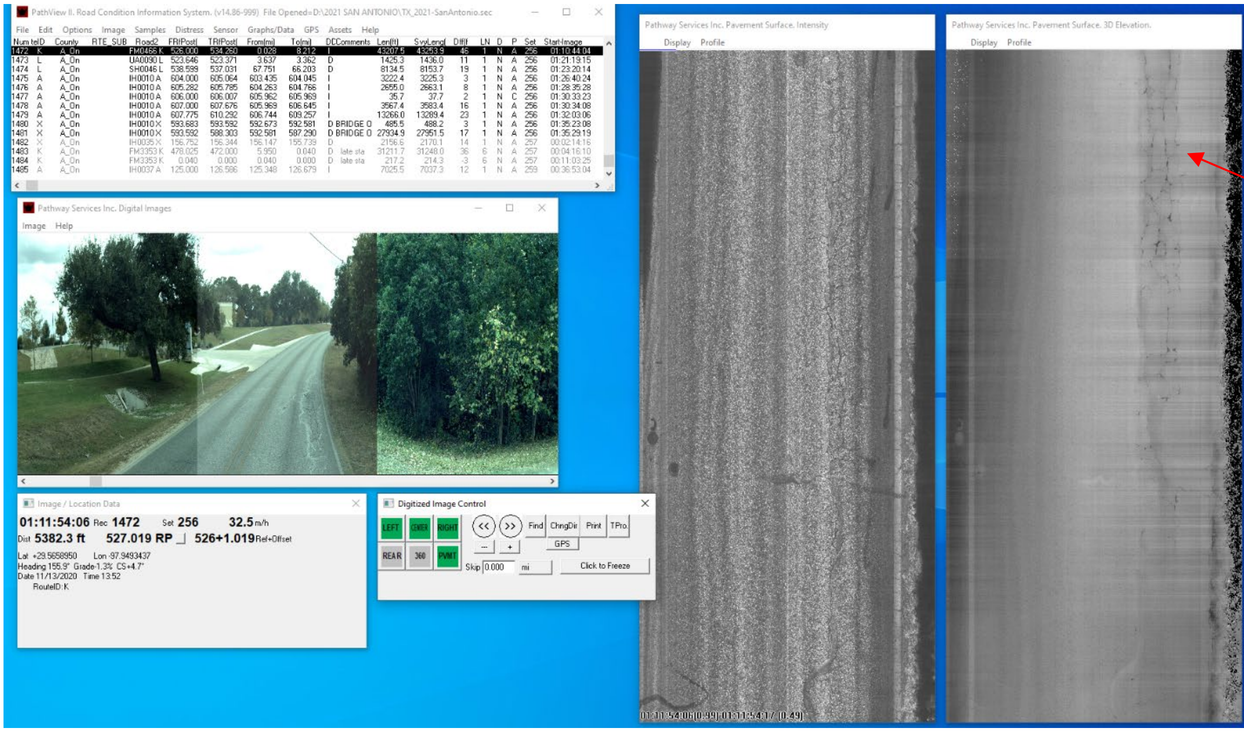


Case 4: Software identified transverse cracking as alligator cracking (red box)

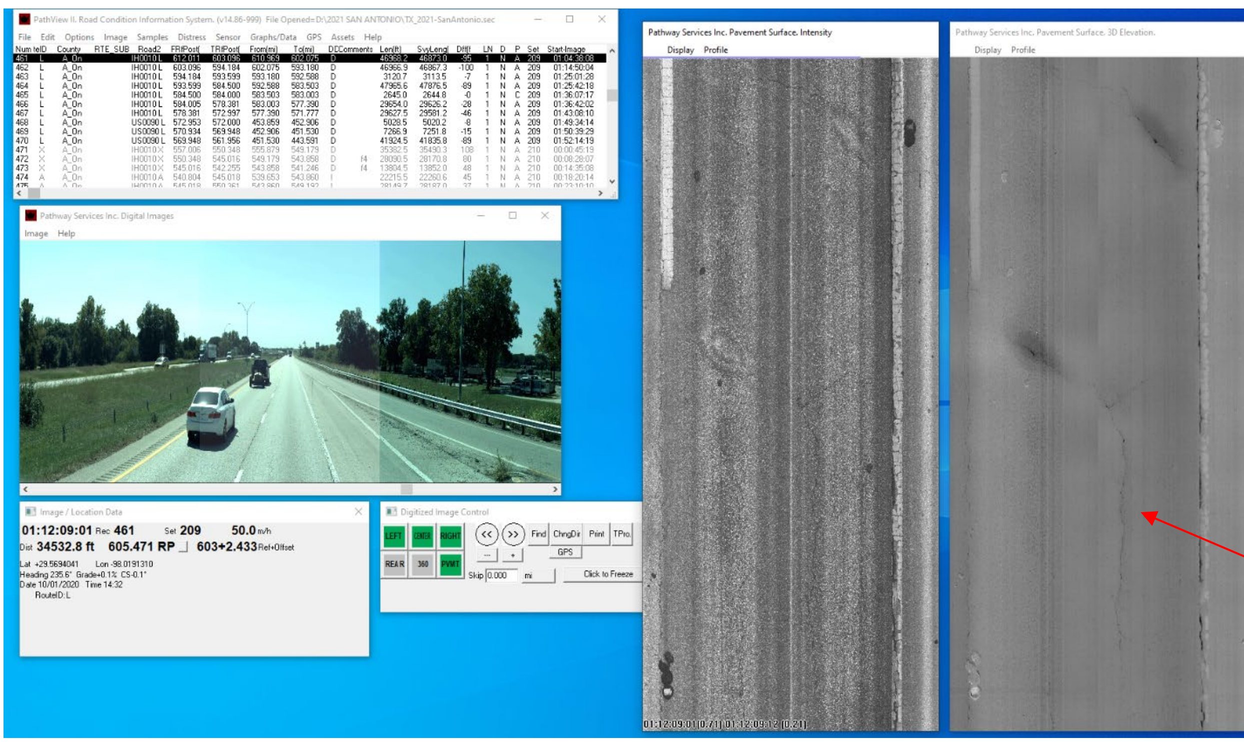


Case 5: Software identified transverse cracking as longitudinal cracking (blue box)

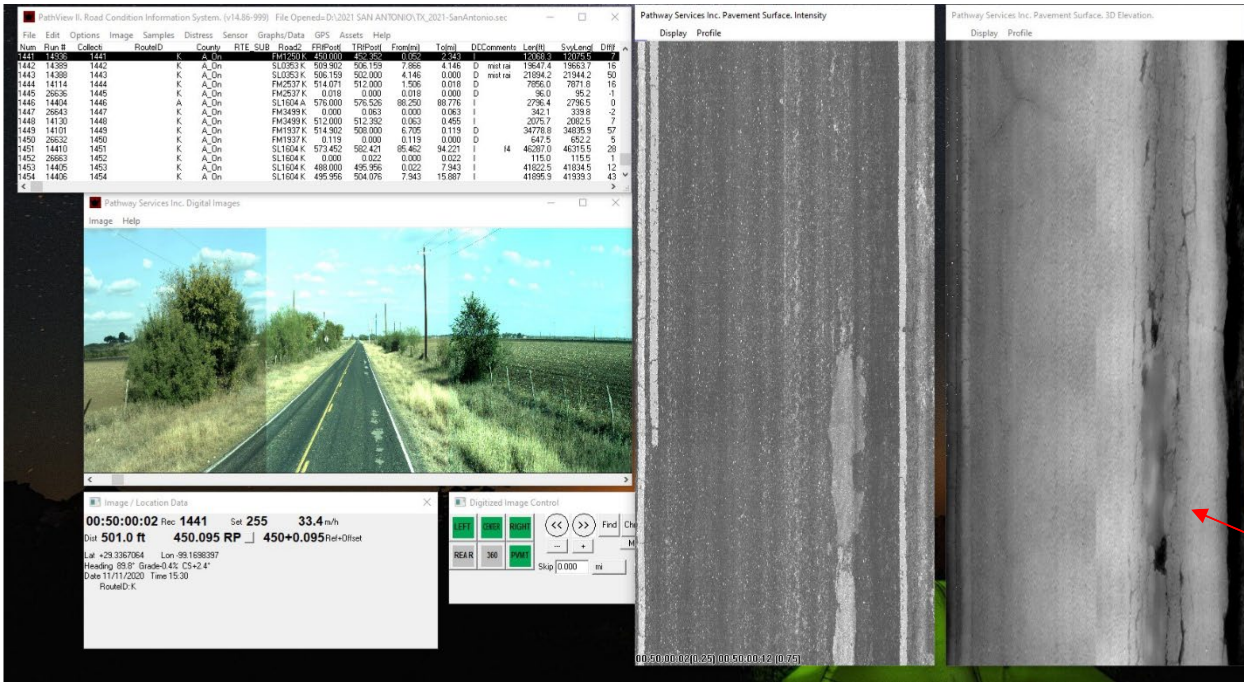
Case 6: Software failed to detect alligator cracking



Case 7: Software failed to detect longitudinal cracking



Case 8: Software failed to detect ACP failure



Pathway Services Inc. Digital Images
Image / Location Data
00:50:00:02 Rec 1441 Set 255 33.4 m/h
Dist 501.0 ft 450.095 RP 450+0.095 Ref+Offset
Digitized Image Control
LEFT
RIGHT
REAR
360
PVMT
Find
Pathway Services Inc. Pavement Surface, Intensity
Display Profile
Pathway Services Inc. Pavement Surface, 3D Elevation
Display Profile