Improving Data Quality of Automated Pavement Condition Data Collection – Summary of State of the Practices of Transportation Agencies and Views of Professionals

Xiaohua Luo[1]; Haitao Gong[2]; Jueqiang Tao[3]; Feng Wang[4]; Jana Minifie[5]; Xin Qiu[6]

Abstract: Automated or semi-automated pavement condition data collection is replacing its predecessor manual data collection in many state and local highway agencies due to the advantages of reducing labor, time, and cost. However, the practical experience from highway agencies indicates that there are still data quality issues with the pavement condition data collected using the existing image and sensor based data collection technologies. This study aims to investigate the implementation experience and issues of automated or semi-automated pavement condition surveys. An online questionnaire survey was conducted along with scheduled virtual/phone interviews to gather relevant information about state of the practice and state of the art from government, industry, and academia. Open questions about the data quality and quality control & quality assurance (QC/QA) were used to receive first-hand inputs from highway agencies and pavement experts. The study has compiled the following observations: 1) a uniform data collection protocol for automated data collection is an urgent need for highway agencies; 2) the current QA has too much human intervention; 3) the cost ranging $100-$200 per mile is a big burden for state and local agencies; 4) the main issues of data quality are presented as data inconsistency and discrepancy; 5) a higher accuracy is expected if the image processing algorithms are improved using artificial intelligence technologies; and 6) the existing automated data collection is not available for project-level data collection.

Keywords: Automated or semi-automated data collection; Pavement condition data; Data quality; State and local agencies

[1] Postdoctoral Research Associate, Ingram School of Engineering, Texas State University, 601 University Dr., San Marcos, TX 78666. Email: x_l33@txstate.edu
[2] PhD candidate, Material Science Engineering and Commercialization (MSEC), Texas State University, 601 University Dr., San Marcos, TX 78666. Email: h_g153@txstate.edu
[3] PhD candidate, Material Science Engineering and Commercialization (MSEC), Texas State University, 601 University Dr., San Marcos, TX 78666. Email: j_t552@txstate.edu
[4] Associate Professor, Ingram School of Engineering, Texas State University, 601 University Dr., San Marcos, TX 78666 (corresponding author). Email: f_w34@txstate.edu
[5] Professor, Dept. of Management, Texas State University, 601 University Dr., San Marcos, TX 78666. Email: jm13@txstate.edu
[6] Professor, Dept. of Road and Traffic Engineering, Zhejiang Normal University, 688 Yinbin Rd., Jinhua, Zhejiang 321004. Email: xqiu@zjnu.cn

Introduction

In the past two decades, one of the greatest improvements in pavement management is using 3-Dimensional (3D) camera and laser sensor based automated or semi-automated technologies for pavement condition data collection. Updated existing technology has sensor system with the ability of obtaining 1 mm resolution 3D pavement image data. These pavement image data are at full lane coverage in all 3 directions which are collected at highway speed up to 100 km/h (Wang et al. 2015). In recent years, researchers have been working on algorithms of automated pavement cracking detection and analysis on 3D pavement image data using deep learning and neural networks (Tsai and Chatterjee 2017; Tsai et al. 2017; Zhang et al. 2017; Zhang et al. 2019; Hsieh and Tsai 2020; Tsai et al. 2021; Yang et al. 2021). In the last two decades, many other efforts were cumulatively devoted to driving the technological innovation of the automated pavement condition data collection (McGhee 2004; Flintsch and McGhee 2009; Pierce et al. 2013; Pierce and Weizel 2019; Chang et al. 2020). With these dedicated and continued efforts, the automated or semi-automated technologies for pavement condition data collection have become more and more sophisticated. Compared with the traditional manual data collection method, automated or semi-automated pavement condition surveys can provide safer, faster, and more cost-effective data collection operations and more convenient and efficient data services. Therefore, automated or semi-automated technologies have been widely adopted by state and local highway agencies for pavement condition data collection.

However, problems have arisen with the use of these technologies in pavement data collection. Many highway agencies have reported data quality issues with the automated pavement condition data (Pierce and Weizel 2019). Highway agencies have taken various quality management measures for automated data collection, which include monitoring of quality control (QC) requirements on the vendor side and implementation of quality assurance (QA) procedures on the agency side to improve data quality. For instance, the Virginia Department of Transportation (VDOT) contracts with a third party to validate and verify 10% of the collected pavement condition data (Flintsch and McGhee 2009). Texas Department of Transportation (TxDOT) conducts a quality assurance audit that uses TxDOT in-house staff and a third-party contractor to visually evaluate about 6% of roadbed miles for surface distresses and ride quality (TxDOT 2016). There is a lot of valuable experience and lessons already learned by the state and local agencies in implementing automated data collection in pavement management, but there is a lack of an appropriate platform to share these experience and lessons. Most of the reported experience about QC/QA of the automated pavement condition data collection focuses on state highway agencies (Flintsch and McGhee 2009; Pierce et al. 2013; Pierce and Weizel 2019; Chang et al. 2020). There are still some studies focusing on the issues that local agencies are having. The data collection vendors and industry are the ones who conduct data collection for the agencies and face the problems of data acquisition. They should be consulted regarding their perspectives on data collection technology and data quality improvement. It is necessary to conduct a comprehensive study that combines the experience of implementing automated or semi-automated technologies in pavement condition surveys by researchers and industry for both state and local highway agencies.

The main objective of this study is to collect and present the experience of U.S. highway agencies along with the industry and researchers in implementing automated or semi-automated pavement condition surveys. The study focuses on data collection methods, service providers, protocols, requirements, and costs. To fulfill this objective, the study was conducted by using an online questionnaire survey along with virtual/phone interviews. The questionnaire served to collect current policies and practices of state highway agencies regarding automated data collection and data quality management. The virtual/phone interviews were designed to cover various relevant topics in implementing and conducting automated pavement condition surveys and possible ways for improving the data quality.

Background

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 a labor-intensive and time-consuming process making it difficult 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 involve 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 is using 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 the automated data collection technologies have pushed forward the innovation of pavement performance quality assessment (McGhee 2004; Flintsch and McGhee 2009; Pierce et al. 2013; Pierce and Weizel 2019). The automated pavement condition survey has become a commonly acceptable data collection method because: 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 are using both manual and automated condition surveys, and the other 6 agencies are using manual pavement condition surveys (Pierce and Weizel 2019). With the wide application of automated pavement condition surveys in state DOTs, the agencies are the end-users of automated pavement condition data collection technologies and the collected pavement condition data. It is important to capture the agencies' experience in their implementation of automated pavement condition data collection.

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 quality 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. The DQMPs are good resources to better understand the way that state highway agencies collect the pavement condition data and enhance the data quality.

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). The Long-term Pavement Performance (LTPP) Distress Identification Manual is also adopted by a few state agencies. Some state highway agencies, such as California Department of Transportation (Caltrans) and Pennsylvania Department of Transportation (PennDOT), have their own standards for automated data collection which serve their state-level data collection, analysis, and decision making. A review of the automated data collection protocols and standards being used by state highway agencies is included in this study.

Quality Improvement of Automated Pavement Data Collection

With the wide application of automated pavement data collection technologies, there is a major concern that the quality of the automatically collected pavement condition data varies with the differences of equipment, algorithms, operation procedures, and human interventions. The AFH20 Quality Assurance Management Committee of the Transportation Research Board (TRB) proposed and got approved of an NCHRP Synthesis Study titled "Agency Inspection and Monitoring of Quality Control (QC) Plans for use in administering Quality Assurance Specifications." A major consideration of the AFH20 committee is to interview some of the states that have good practices for QC plans as well as those that do not have requirements to get a clearer picture of the state of the practice (TRB 2020). TxDOT funded a research project titled "Improve Data Quality for Automated Pavement Distress Data Collection" to address the data accuracy and precision issues associated with the reliability of the existing automated and semi-automated data collection methods. TxDOT also desires to establish data acceptance and QA guidelines, procedures, or specifications for automated and semi-automated pavement condition surveys that could be used to improve data quality management practices for contracting pavement condition data collection (TxDOT 2020). The highway agencies are making efforts to manage the quality of automated pavement data, including monitoring of QC requirements on the vendor side and implementation of QA procedures on the agency side (Pierce and Weitzel 2019).

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 check and control, verification, or blind site data collection, which are performed during data collections (Pierce and Weitzel 2019). The pavement performance indicators for QC, verification, or blind site check mainly focus on IRI, rutting, faulting, cracking, and location, but the specific requirements/tolerances for the control site checks vary among state highway agencies.

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. The major contents of QA include IRI, rutting, faulting, cracking, and images. If the submitted pavement condition data did not pass the data acceptance check 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 when applying automated data collection technologies. The quality of the automated data varies due to the factors in equipment, algorithms, operation procedures, and human interventions. The reason that caused this problem could be the deficient QC and QA during and after the data collection. This study aims to review the successful practices and discuss the issues in the automated pavement condition data collection.

Methodology

The methods used in this study include an online questionnaire survey and virtual/phone interviews. The questionnaire named "Automated Pavement Condition Data Collection - Data Quality Control and Quality Assurance (QC/QA) Questionnaire" was designed to support a TxDOT research project titled "Improve Data Quality for Automated Pavement Distress Data Collection." The questionnaire was drafted in December 2020. Five TxDOT pavement engineers reviewed, commented, and suggested the original draft. There are five sections (i.e., data collection, DQMP for quality control, DQMP for quality assurance, open question for data quality issues, and DQMP standard sharing request) with a total of thirty-five questions in the finalized version of the questionnaire. The questionnaire was created using Qualtrics Surveys software and distributed to fifty-two State Pavement Management Engineers (including Washington D.C. and Puerto Rico) on April 1st, 2021 through TxDOT's email system.

The interviews were conducted through the National Science Foundation's Innovation Corps (I-Corps) program. According to the COVID-19 pandemic, all the interviews were conducted virtually using online communication platforms (e.g., Zoom and Microsoft Teams) and phone calls. The interviewees were pavement experts from the government, industry, and academia with practical experience in automated pavement condition data collection. The questions being asked during the interviews vary according to the positions and responsibilities of the interviewees. For the interviewees from the government, the questions mainly focused on the methods used, challenges faced, data quality issues experienced, and the price consideration for pavement condition surveys. The questions for the interviewees in the industry focused on the development of technologies, efficiency of data collection, quality control methods, and marketing experience. The questions for the academic researchers concentrated on technical experience in research and development, along with challenges and innovation trends of the technologies for automated detection, classification, and quantification of pavement distresses. The interviews were conducted during the I-Corps program from January 12th to February 23rd, 2021.

Results and Discussion

A tremendous amount of information was collected by the questionnaire survey and interview responses. By the end of April 2021, there were thirty-seven responses to the online questionnaire received from thirty-three state highway agencies. Among them, twenty-nine highway agencies also shared their DQMP standards. In addition, 101 pavement experts were interviewed including 77 by virtual meetings and 24 by phone calls.

The aggregate results are presented by category. Due to the space limit, only the results falling in the following categories are presented. It is believed that, at the very least, it provides a fairly current picture of the automated pavement condition data collection community from these highway agencies, companies, and researchers. It should be noted that the reported results and associated analysis only reflect the opinions of the respondents in the online survey and interviews.

Practice of Pavement Condition Data Collection

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. Fig. 1. summarizes the pavement condition data collection methods currently used by the agencies. 32 of 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 Caltrans still needs manual interventions for QC/QA. Florida DOT uses fully automated Laser Crack Measurement System (LCMS) for HPMS. 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.

Data Collection Service Provider

The questionnaire survey result summarized in Fig. 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. First, 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 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, such as 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.

Implementation of Pavement Condition Data Collection

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. 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 have standards of their own design.

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 reporting 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 and vertical alignment data, and fill quantity; Mississippi DOT collects friction data; TxDOT collects skid number.

Data Collection Length and Cycle

The data collection length which is collected every cycle depends on the state's roadway network length. Fig. 3. summarizes the survey results about the state's data collection length and frequency. 26 states collect the pavement condition data by roadbed miles, 4 states collect pavement data by lane miles, and 2 states collect pavement data by centerline miles. 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. 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.

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.

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. 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 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 third parties 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.

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 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%. The sampling rates for distress data can also be 25%, 35%, and even 100%. For IRI, rutting, and faulting, most states are exercising a sampling rate of 100% of the collected network length, and a few states 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 stratified sampling, by dividing the population into multiple groups or clusters, and then selecting random elements from these clusters.

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 3 summarizes some typical data quality issues and possible reasons from the responses of state highway agencies. Eight states mentioned issues about cracking data, such as 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 in the urban areas. The IRI sensors are very sensitive to 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.

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. For state highway agencies, there are some federal regulations to specify how automated pavement surveys should be conducted and how the data quality should be handled. 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.

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 could be 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 $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 on 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 are accurate. Each procedure would add a certain amount of cost to the total cost. If survey data is proved as 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 agency. Therefore, many engineers suggested that reducing data collection costs and data processing time are urgent needs for automated data collection.

Problems with Existing Automated Data Collection

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. 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 also create 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 measurements 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 that the vendors all use proprietary image data formats that literally prevent sharing and cross-checking 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. An engineer from Quality Engineering Solution Inc. (QES) mentioned that the current technologies have issues 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 that the data inconsistency and discrepancy issues are 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.

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 the 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 on 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 an accuracy of 95% 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, the 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 automated data collection technologies still needs data pre-treatment. The lack of training data due to the low availability of annotated ground truth image data and difficulties in sharing data in the public domain have caused delays in developing and using AI in the technologies. 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.

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 problem 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. Four pavement engineers from state and local highway agencies pointed out that they 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 four 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 work together to 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. QA 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.

Many state and local agencies contracted with third parties to examine the survey data which were 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 budgets just to make the data right. 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.

Extend Automated Data Collection to Project Level

A few agencies mentioned in the questionnaire survey that they are using automated technologies for some project-level data collection. Many pavement engineers from both the industry and government also highlighted in the interviews that extending the automated pavement condition data collection technologies to project-level data collection is necessary. However, the current technologies are still not fully ready. An engineer from the City of Austin, who worked with a vendor for automated pavement data collection, shared the experience that the current automated data collection is not yet available for project-level data collection. There were many issues with IRI data collected in the city network. The engineer who worked on PMS data spent a lot of time on QA. Another engineer revealed that the project-level data collection in TxDOT San Antonio District is more than just IRI data, and it also requires Falling Weight Deflectometer (FWD) and Ground Penetrating Radar (GPR) to collect structure data conducted by TxDOT staff. MnDOT is struggling with traffic control and seasonal limitations in using GPR and FWD for project-level pavement condition surveys. An engineer at TxDOT Houston District mentioned that the inaccurate GPS reference is another important issue that limited the use of automated data collection at project-level. Pavement engineers from Virginia DOT, TxDOT Dallas and Pharr Districts, NCE, and StreetSaver Inc. also provided the same feedback.

The interview result shows that although there is a tremendous need to extend the automated pavement condition data collection from network-level to project-level, the current automated data collection technologies still have some issues that need to be addressed before being applied as the primary method for project-level data collection. The first issue is that the data accuracy and precision used for project-level design model calibration are typically higher than that used for network-level performance trends (Chang et al 2020). The second issue is that the current automated pavement condition data collection technologies cannot provide all the data items (e.g., structure data) needed for project-level decision making.

Conclusions

The automated pavement condition survey is essential as it saves great amounts of resources (time and cost) for the customers who need pavement condition data. In addition to the state highway agencies, municipal governments also rely on automated pavement survey services. This study employed a questionnaire survey to investigate the implementation of automated or semi-automated pavement surveys and to summarize the QC/QA practices that are conducted by state and local highway agencies. The study also conducted 101 virtual or phone interviews to learn the practical insights about the above issues that the government, industry, and academia are perceiving in conducting automated or semi-automated data collection. Based on the survey questionnaire and interviews, the following findings are observed:

- Most of the state and local highway agencies conduct automated or semi-automated pavement data collection. A lot of state highway agencies have more than ten years of experience in using automated or semi-automated technologies. Contracting with a vendor is a prevailing way to conduct a pavement condition survey.
- There is no uniform data collection protocol for automated or semi-automated pavement data collection. ASTM standards, AASHTO standards, and the LTPP standard are the commonly used standards, but state highway agencies also have standards of their own design.
- The data collection items and frequency vary among state highway agencies. Most of the states collect state maintained network-level pavement condition data annually.
- The QC of the automated or semi-automated pavement condition data collection is typically conducted by vendors and agency staff. QA activities are mainly conducted by the agency staff which take them plenty of time. Random, stratified, and systematic sampling methods with a specific sampling rate of the roadway network length are used for state agencies' QA purposes. However, the current QA process has huge demands on staff time of agencies. An innovation of the QA process could help promote automated pavement condition surveys.
- The existing high cost of automated or semi-automated pavement condition data collection with a range of $100-$200 per mile is a big concern for state and local agencies. State highway agencies and local agencies may have different cost expectations on automated pavement condition surveys.

- The main issue of the automated or semi-automated pavement condition data collection is data quality which presents as data inconsistency and discrepancy. The agencies spend plenty of time of their own or contract with an individual third party to address the data quality issues and make the data usable. Meanwhile, vendors provide timely and effective technical support services to help the agencies and/or the third party address the data quality issues when the issues are reported.
- Although the existing automated pavement condition data collection technologies are widely adopted for network-level pavement surveys, there are data inconsistency and discrepancy problems that must be corrected through a time-demanding QA process. These inconsistency and discrepancy problems are partially due to the immature data collection technologies as well as the vendors' use of proprietary image data formats that prevent sharing and cross-checking of data for agencies to work with different vendors at different times.
- Extending the automated data collection to project-level is a tremendous need for pavement engineers, but the current technologies are still immature in data accuracy and precision for project-level use. The AI concepts and models are expected to be more researched and utilized to further improve and optimize the image processing capabilities and to be applied to the continued advancement of automated pavement data collection technologies.

Data Availability Statement

No data, models, or code were generated or used during the study.

Acknowledgments

This research work has received funding support from the Texas Department of Transportation (TxDOT Project 0-7072) and the National Science Foundation (NSF Award No. 2051192). The authors are grateful to Dr. Jenny Li and Dr. Feng Hong of TxDOT for their tireless help to plan and implement the questionnaire survey. The authors would like to thank the 33 highway agencies for providing their data inputs to the questionnaire survey and the 101 pavement engineers and researchers for sharing their valuable knowledge and insights.

References

Wang, K.C., Q.J. Li, G. Yang, Y. Zhan, and Y. Qiu. 2015. "Network level pavement evaluation with 1 mm 3D survey system." *Journal of traffic and transportation engineering (English edition)*, 2(6), 391-398. https://doi.org/10.1016/j.jtte.2015.10.005

Yang, G., K.C. Wang, J.Q. Li, Y. Fei, Y. Liu, K.C. Mahboub, and A.A. Zhang. 2021. "Automatic Pavement Type Recognition for Image-Based Pavement Condition Survey Using Convolutional Neural Network." *J. Comput. Civ. Eng.* 35(1): 04020060. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000944

Zhang, A., K.C. Wang, B. Li, E. Yang, X. Dai, Y. Peng, Y. Fei, Y. Liu, J.Q. Li, and C. Chen. 2017. "Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network." *Computer-Aided Civil and Infrastructure Engineering*, 32(10), 805-819. https://doi.org/10.1111/mice.12297

Zhang, A., K.C. Wang, Y. Fei, Y. Liu, C. Chen, G. Yang, J.Q. Li, E. Yang, and S. Qiu. 2019. "Automated pixel-level pavement crack detection on 3D asphalt surfaces with a recurrent neural network." *Computer-Aided Civil and Infrastructure Engineering*, 34(3), 213-229. https://doi.org/10.1111/mice.12409

Tsai, Y.C., Y. Zhao, B. Pop-Stefanov, and A. Chatterjee. 2021. "Automatically detect and classify asphalt pavement raveling severity using 3D technology and machine learning." *Int. J. Pavement Res. Technol.*, 14(4), 487-495. https://doi.org/10.1007/s42947-020-0138-5

Hsieh, Y.A., and Y.J. Tsai. 2020. "Machine learning for crack detection: review and model performance comparison." *J. Comput. Civ. Eng.* 34(5): 04020038. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000918

Tsai, Y.C., and A. Chatterjee. 2017. "Comprehensive, quantitative crack detection algorithm performance evaluation system." *J. Comput. Civ. Eng.* 31(5): 04017047. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000696

Tsai, Y.C., A. Chatterjee, and C. Jiang. 2017. “Challenges and lessons from the successful implementation of automated road condition surveys on a large highway system.” *25th European Signal Processing Conference (EUSIPCO)*, 2031-2035. IEEE. Kos, Greece. https://doi.org/10.23919/EUSIPCO.2017.8081566

McGhee, K.H. 2004. *NCHRP Synthesis 334: Automated Pavement Distress Collection Techniques, Report No. Project 20-5, Topic 34-04.* National Cooperative Highway Research Program. Transportation Research Board, Washington, D.C.

Flintsch, G. W., and K.H. McGhee. 2009. *Quality Management of Pavement Condition Data Collection. Vol. 401.* Transportation Research Board, Washington, D.C.

Pierce, L. M., G. McGovern, and K. Zimmerman. 2013. *Practical Guide for Quality Management of Pavement Condition Data Collection*. Federal Highway Administration, Washington, D.C.

Pierce, L.M., and N.D. Weitzel. 2019. *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.

Chang, George, Amanda Gilliland, Gonzalo R. Rada, Pedro A. Serigos, Amy L. Simpson, and Sareh Kouchaki. 2020. *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.

TxDOT (Texas Department of Transportation). 2016. *Texas Transportation Asset Management Plan*. Texas Department of Transportation, Maintenance Division, Austin, TX.

FHWA (Federal Highway Administration). 2018. *Guidelines for Development and Approval of State Data Quality Management Programs*. Washington, D.C.

FHWA (Federal Highway Administration). 2017. *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.

Simpson, A., G. Rada, J. Bryce, P. Serigos, B. Visintine, and J. Groeger. 2018. *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.

TRB (Transportation Research Board). 2020. Proposed NCHRP Synthesis Study "Agency Inspection and Monitoring of Quality Control Plans for use in administering Quality Assurance Specifications", draft in process and discussions, Quality Assurance Management Committee (AFH20), TRB, Washington D.C.

TxDOT (Texas Department of Transportation) 2020. FY 2021 Research project Statement 21-047 "Improve Data Quality for Automated Pavement Distress Data Collection", Texas Department of Transportation, Research and Technology Implementation Division (RTI), Austin, TX.

AASHTO (American Association of State Highway and Transportation Officials). 2006. *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.

Simpson, A., P. Serigos, S. Kouchaki, G. Rada, and J. Groeger. 2020. *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.

TxDOT (Texas Department of Transportation). 2018. Quality Management Plan for Pavement Data Collection, Texas Department of Transportation, Maintenance Division, Pavement Preservation Branch, Austin, TX.

List of figure captions

Fig. 1. Summary of agency data collection methods (total # of responses = 33)

Fig. 2. Summary of data collection service providers (total # of responses = 33)

Fig. 3. Data collection lengths and cycles of state highway agencies (32 responses with automated or semi-automated data collection)

Tables

Table 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
Who does the equipment certification for 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
Who does the equipment certification for 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
Who does the equipment *verification for 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

Who does the equipment *verification for 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 **calibration for 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, WA,	FL, NJ, TX
Who does the equipment **calibration for 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
Who does the equipment **calibration for 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).

Table 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	AK (5%), CA (0.5-5%), FL (10%), GA (5%), IL (50%), KY (100%), MD (100%), MS (100%), NE (100%), NH (100%), NY	AK (5%), CA (0.5-5%), FL (10%), GA (5%), IL (50%), KY (100%), MD (100%), MS (100%), NE (100%), NH (100%), NY	FL (10%), GA (5%), KY (100%), MS (100%), NE (100%), NY (10%), PA (2.5%), SD (100%), WY

	(5%), PA (2.5%), SD (100%), TN (2%), TX (6%), UT (5-10%), WA (5%), WY	(10%), NJ (5%), PA (2.5%), SD (100%), TN (2%), UT (5-10%), WY	(10%), SD (100%), TN (2%), 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)		

Note: the numbers in the brackets mean the sampling rates of the state DOTs.

Table 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 on 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 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 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.

602