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, M.ASCE[2]; Jueqiang Tao[3]; Feng Wang, M.ASCE[4]; Jana Minifie[5]; and Xin Qiu[6]

Abstract: Automated or semi-automated pavement condition data collection is replacing manual data collection in many state and local highway agencies due to its advantages of reducing labor, time, and cost. However, the practical experience of highway agencies indicates that there are still data quality issues with the pavement condition data collected using existing image and sensor-based data collection technologies. This study aims to investigate the implementation experiences and issues of automated or semi-automated pavement condition surveys. An online questionnaire survey was conducted, along with scheduled virtual/phone interviews to gather information from government, industry, and academia about the state of the practice and state of the art. 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) Highway agencies urgently need a uniform data collection protocol for automated data collection; (2) the current QA requires too much human intervention; (3) cost ($100–$200 per mile) is a significant burden for state and local agencies; (4) the main issues regarding data quality are data inconsistencies and discrepancies; (5) agencies expect a greater accuracy once the image processing algorithms are improved using artificial intelligence technologies; and (6) existing automated data collection methods are not available for project-level data collection. **DOI: 10.1061/JPEODX.0000392.**

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

Introduction

In the past two decades, one of the greatest improvements in pavement management is the use of three dimensional (3D) camera and laser sensor-based automated or semi-automated technologies for pavement condition data collection. Updated existing technology has sensor systems able to obtain 1 mm resolution 3D pavement image data. These data provide full lane coverage in all three directions, and can be collected at highway speeds up to 100 km/h (Wang et al. 2015). In recent years, researchers have been creating algorithms for 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, 2021; Zhang et al. 2017, 2019; Hsieh and Tsai 2020; Yang et al. 2021). In the last two decades, many other efforts were devoted to driving technological innovation in the field of automated pavement condition data collection (McGhee 2004; Flintsch and McGhee 2009; Pierce et al. 2013; Pierce and Weitzel 2019; Chang et al. 2020). With these dedicated and continuing efforts, 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 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, and many highway agencies have reported data quality issues (Pierce and Weitzel 2019). Agencies have taken various quality management measures to improve automated data collection, including monitoring quality control (QC) requirements on the vendor side and implementing quality assurance (QA) procedures on the agency side. 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). The Texas Department of Transportation (TxDOT) conducts a quality assurance audit that uses in-house staff and a third-party contractor to visually evaluate about 6% of roadbed miles for surface distress and ride quality (TxDOT 2016). State and local agencies have acquired much valuable experience and many lessons learned by implementing automated data collection, but there is a lack of an appropriate platform to share these experience and lessons. Most of the reported experience about quality control & quality assurance (QC/QA) of the automated pavement condition data collection focuses on state highway agencies (Flintsch and McGhee 2009; Pierce et al. 2013; Pierce and Weitzel 2019; Chang et al. 2020), although some studies focus on the issues that local agencies are

[1]Postdoctoral Research Associate, Ingram School of Engineering, Texas State Univ., 601 University Dr., San Marcos, TX 78666. ORCID: https://orcid.org/0000-0002-0686-450X. Email: x_l33@txstate.edu

[2]Ph.D. Candidate, Material Science Engineering and Commercialization (MSEC), Texas State Univ., 601 University Dr., San Marcos, TX 78666. Email: h_g153@txstate.edu

[3]Ph.D. Candidate, MSEC, Texas State Univ., 601 University Dr., San Marcos, TX 78666. Email: j_t552@txstate.edu

[4]Associate Professor, Ingram School of Engineering, Texas State Univ., 601 University Dr., San Marcos, TX 78666 (corresponding author). ORCID: https://orcid.org/0000-0002-1528-9711. Email: f_w34@txstate.edu

[5]Professor, Dept. of Management, Texas State Univ., 601 University Dr., San Marcos, TX 78666. Email: jm13@txstate.edu

[6]Professor, Dept. of Road and Traffic Engineering, Zhejiang Normal Univ., 688 Yinbin Rd., Jinhua, Zhejiang 321004. Email: xqiu@zjnu.cn

Note. This manuscript was submitted on November 15, 2021; approved on May 5, 2022; published online on July 5, 2022. Discussion period open until December 5, 2022; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Transportation Engineering, Part B: Pavements***, © ASCE, ISSN 2573-5438.

having. Data collection vendors perform data collection for the agencies and directly face the problems of data acquisition; thus, their perspectives on data collection technology and data quality improvement are important. It is necessary to conduct a comprehensive study that combines the experiences 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 present the experiences of US highway agencies, 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, 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 related to 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 surface distresses (Pierce and Weitzel 2019). It is a labor-intensive and time-consuming process that is difficult to cover the entire roadway length. To overcome the challenges of the manual survey, many highway agencies have adopted high-speed automated data collection technologies for network-level pavement condition data collection. 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. In semi-automated data processing, the collected images 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 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). Some fully automated systems use video and/or laser technology to detect and classify pavement cracking in real-time, and at highway speeds. Alternatively, some data collection vendors use systems to capture the pavement image first, and then detect and classify the cracks using automated post-processing (Pierce and Weitzel 2019).

Recently, investigators have confirmed that the automated data collection technologies have advanced pavement performance quality assessments (McGhee 2004; Flintsch and McGhee 2009; Pierce et al. 2013; Pierce and Weitzel 2019). The automated pavement condition survey has become a commonly acceptable data collection method because it has a minimal impact on traffic, while providing a significant increase in safety, increased 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 responders (46 US highway agencies and 11 Canadian provincial and territorial governments) are using automated data collection methods exclusively. Six of the other 12 agencies are using both manual and automated condition surveys, and the final six agencies are using manual pavement condition surveys (Pierce and Weitzel 2019). With the wide adoption of automated pavement condition surveys by state DOTs, the agencies are the end-users of automated pavement condition data collection technologies, and of the collected pavement condition data. It is thus important to capture those agencies' experiences in the 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 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). Rule PM2, which became effective in 2017, surface roughness [International Roughness Index (IRI)], rutting, faulting, and cracking percent, along with present serviceability rating (PSR) as the pavement condition metrics. The state highway agencies are required to collect and report these pavement condition metrics to the FHWA's Highway Performance Monitoring System (HPMS) as either good, fair, or poor, per 23 CFR 490.309(c) (FHWA 2018).

To collect 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 (Simpson et al. 2018). The 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 processes and deliverables (FHWA 2018). The DQMP has 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 data collection; (4) data sampling, review, and checking processes; and (5) error resolution procedures and data acceptance criteria (Simpson et al. 2018). The DQMP aims to address the errors that occur 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 agency's DQMPs can be an efficient way to understand how state highway agencies collect, enhance and report their pavement condition data. However, the data metrics vary by agency, and according to state highway agencies' data collection manuals, different data definitions are used.

Automated Data Collection Protocols and Standards

A data collection protocol/standard is a description of the procedures used to ensure consistency in the collection and recording of pavement condition data (FHWA 2018). In accordance with 23 CFR 490.309(c), pavement condition metrics are to be collected and reported following the standardized HPMS format on an annual cycle for Interstate highways, and on a two-year maximum cycle for all other roads. 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. A previous survey showed that some state agencies also use ASTM standards in their automated data collection, especially in measuring profile and macrotexture, and analyzing precision and bias (Pierce and Weitzel 2019). The Long-Term Pavement Performance Distress Identification Manual (LTPP) has also been adopted by a few state agencies. Some state highway agencies, such as the California Department of Transportation (Caltrans) and Pennsylvania Department of Transportation (PennDOT), have their own standards for automated data collection that 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 adoption of automated pavement data collection technologies, there is much concern that the quality of the automatically collected pavement condition data varies with differences in equipment, algorithms, operation procedures, and human interventions. The AFH20 Quality Assurance Management Committee of the Transportation Research Board (TRB) directed a NCHRP Synthesis Study titled "Agency Inspection and Monitoring of Quality Control (QC) Plans for Use in Administering Quality Assurance Specifications." A major objective of the AFH20 committee was to interview states with good practices for QC plans, as well as those that do not have such 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 existing automated and semi-automated data collection methods. TxDOT also wanted to establish data acceptance and QA guidelines, procedures, and 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). Various other highway agencies are also making efforts to manage the quality of automated pavement data, including monitoring of vendor QC requirements and agency QA procedures (Pierce and Weitzel 2019).

Data Quality Control

According to the AASHTO R10-06 (AASHTO 2006), QC includes the activities needed to adjust production processes to achieve the desired level of quality of pavement condition data. QC includes sampling, testing, inspection, and corrective action (where required), to maintain continuous control of a production process (FHWA 2018). Specific QC activities are required by each state's DQMP and are primarily implemented by the data collection team to monitor, assess, and adjust data collection processes (Chang et al. 2020). QC activities may include equipment calibration, software check and control, verification, and blind site data collection, which are performed during data collection (Pierce and Weitzel 2019). The pavement performance indicators for QC, verification, and 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 internal quality check, the pavement condition data are submitted to the agency. The agency team then 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, the data collection team takes corrective actions to prevent erroneous data collection or data analysis procedures from being processed (FHWA 2017).

Even with existing 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 previously-identified variances in equipment, algorithms, operation procedures, and human interventions. The reasons that cause this problem could be deficient QC and QA during and after data collection. This study aims to identify and review successful practices, and discuss the outstanding issues in the automated pavement condition data collection.

Methodology

The methods used in this study include an online questionnaire 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 on, and provided suggestions to improve the original draft. The finalized version of the questionnaire has five sections (data collection, DQMP for quality control, DQMP for quality assurance, open question for data quality issues, and DQMP standard sharing request) with thirty-five questions. The questionnaire was created using Qualtrics Surveys software and distributed to 52 pavement management engineers (including Washington, DC and Puerto Rico) on April 1, 2021 through the TxDOT email system.

The interviews were conducted pursuant to the National Science Foundation's Innovation Corps (I-Corps) program. Because of the COVID-19 pandemic, all interviews were conducted virtually, via 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 asked during the interviews varied according to the positions and responsibilities of the interviewees. For government interviewees, the questions mainly focused on the methods used, challenges faced, data quality issues experienced, and price considerations. For industry interviewees, the questions focused on the development of technologies, efficiency of data collection, quality control methods, and marketing experience. For academic researchers, the questions focused on technical experience in research and development, along with challenges and innovation trends of the technologies. The interviews were conducted during the I-Corps program period, from January 12 to February 23, 2021.

Results and Discussion

A tremendous amount of information was collected by the questionnaire survey and interview responses. By the end of April 2021, 37 responses to the online questionnaire were received from 33 state highway agencies. Twenty-nine agencies also shared their

DQMP standards. In addition, 101 pavement experts were interviewed, 77 by virtual meetings and 24 by phone calls. The aggregate results were organized by category. Due to space limits, only those results falling into the following categories are presented, which provide a fairly current picture of the automated pavement condition data collection community. 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 result shows that automated and semi-automated pavement data collection methods have been widely adopted by state highway agencies. Fig. 1 summarizes the pavement condition data collection methods currently in use. Of the 33 agencies that responded, 32 have used automated or semi-automated data collection methods. Among these 32 agencies, 12 have used automated or semi-automated data collection technologies for more than 10 years; eight have five to 10 years of experience; and five have one to four years of experience (seven agencies did not respond to this aspect of the question). This result indicates that different states may be at different stages of using automated/semi-automated data collection technologies. For instance, the automated data collection in Caltrans still needs manual intervention for QC/QA. Florida DOT uses a fully automated laser crack measurement system (LCMS) for HPMS, while for the pavement condition survey, it is still in a transition from manual distress data collection to fully automated ratings. Mississippi DOT uses manual data collection for concrete pavement cracking evaluation, which constitutes 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 a semi-automated system for patching and raveling evaluation.

Data Collection Service Provider

The questionnaire survey results summarized in Fig. 2 show that there are three ways in which state highway agencies collect pavement condition data:

- Of the 33 respondents, 20 contract with vendors for pavement condition surveys. Contracting with a vendor is a common

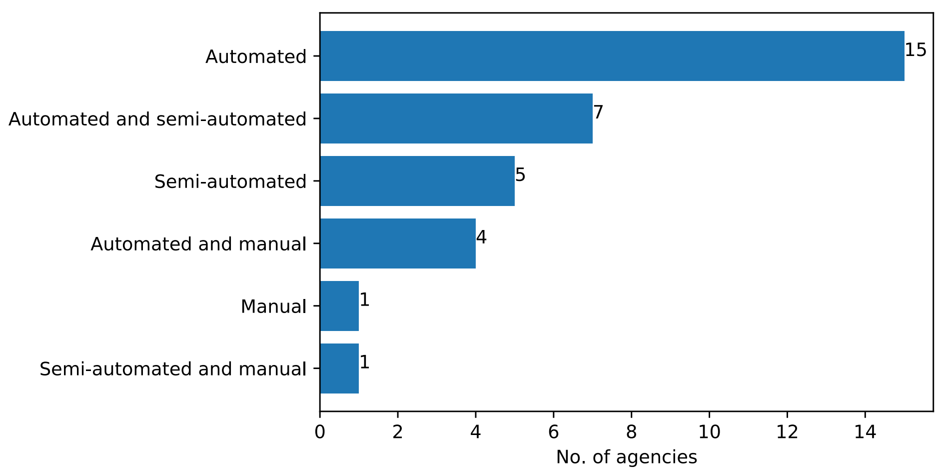

Fig. 1. Summary of agency data collection methods (33 responses).

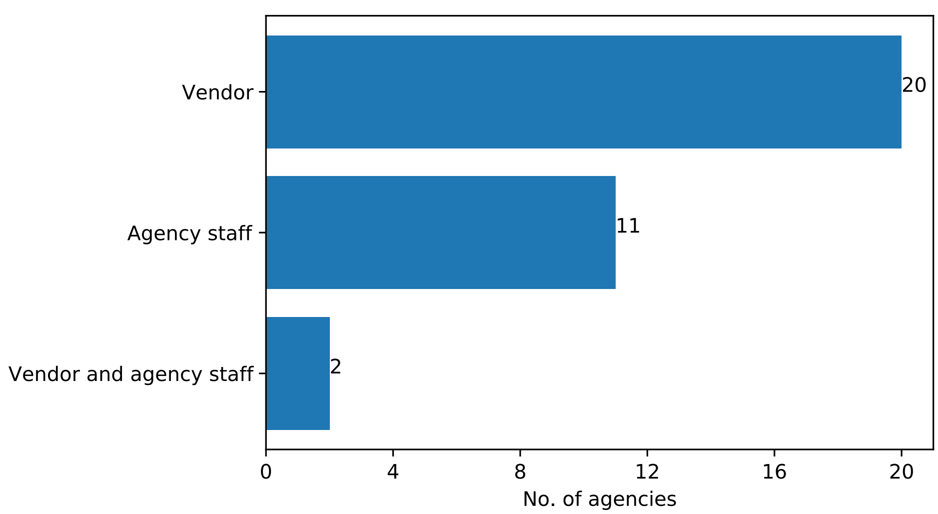

Fig. 2. Summary of data collection service providers (33 responses).

approach for state highway agencies and can save a lot of time for engineers and staff, but the cost of contracting with a vendor can be much different from that of using in-house staff. Some state highway agencies take additional actions to enhance the quality of the vendor's services. For instance, Caltrans has a field crew to perform QC/QA. The Indiana Department of Transportation (INDOT) and PennDOT collect project-level pavement condition data using their own staff.

- Of the 33 respondents, 11 collect data using their own staff. Some of these state agencies, such as the Minnesota Department of Transportation (MnDOT), Maryland DOT, and Washington DOT, own data collection vehicles.
- The final two respondents use both vendors and staff for data collection. For example, Florida DOT collects Interstate highway data using its own LCMS, while a vendor collects data from non-Interstate roads.

Implementation of Pavement Condition Data Collection

Data Collection Protocols

Before implementing automated data collection, a state highway agency should specify its data collection metrics and protocols. As mentioned in the background, data standards and protocols vary by agency. Although FHWA requires states to collect and report pavement condition data following the HPMS field manual, generally a state agency can use more than one data collection protocol. The commonly used protocols include various ASTM standards, AASHTO standards, and the LTPP standard. Delaware, Florida, Illinois, Minnesota, Mississippi, Nevada, Nebraska, Ohio, Oregon, South Dakota, Texas, Washington, and Wyoming 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 agencies collect additional items. For example, Arkansas collects macro texture; Caltrans collects mean profile depth (MPD); Florida plans to expand raveling as a separate distress category; the Louisiana Department of Transportation and Development (LAODTD) collects friction texture, macrotexture, horizontal and vertical alignment data, and fill quantity; Mississippi collects friction data; and TxDOT collects skid numbers.

Data Collection Length and Cycle

The data collection length depends on the state's roadway network length. Fig. 3 summarizes the survey results about each state's data collection length and frequency. Of the 32 respondent states, 26 collect pavement condition data by roadbed miles, four collect pavement data by lane miles, and two 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.) Of the 32 states that use automated or semi-automated data collection, Texas has the longest automated data collection network, and Caltrans has the second-longest.

In its 2016 Field Manual, FHWA specified that the data collection frequency for the 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 so that the most recently collected data replaces the data from the prior data collection cycle. To manage the state roadway network and meet FHWA's data reporting requirements, 21 of the 32 respondent states (solid bars in Fig. 3) collect data on all state-maintained roads annually. The other 11 state highway agencies collect data annually on the Interstate, or both the Interstate and non-Interstate NHS, but collect data biennially on the other state-maintained roads.

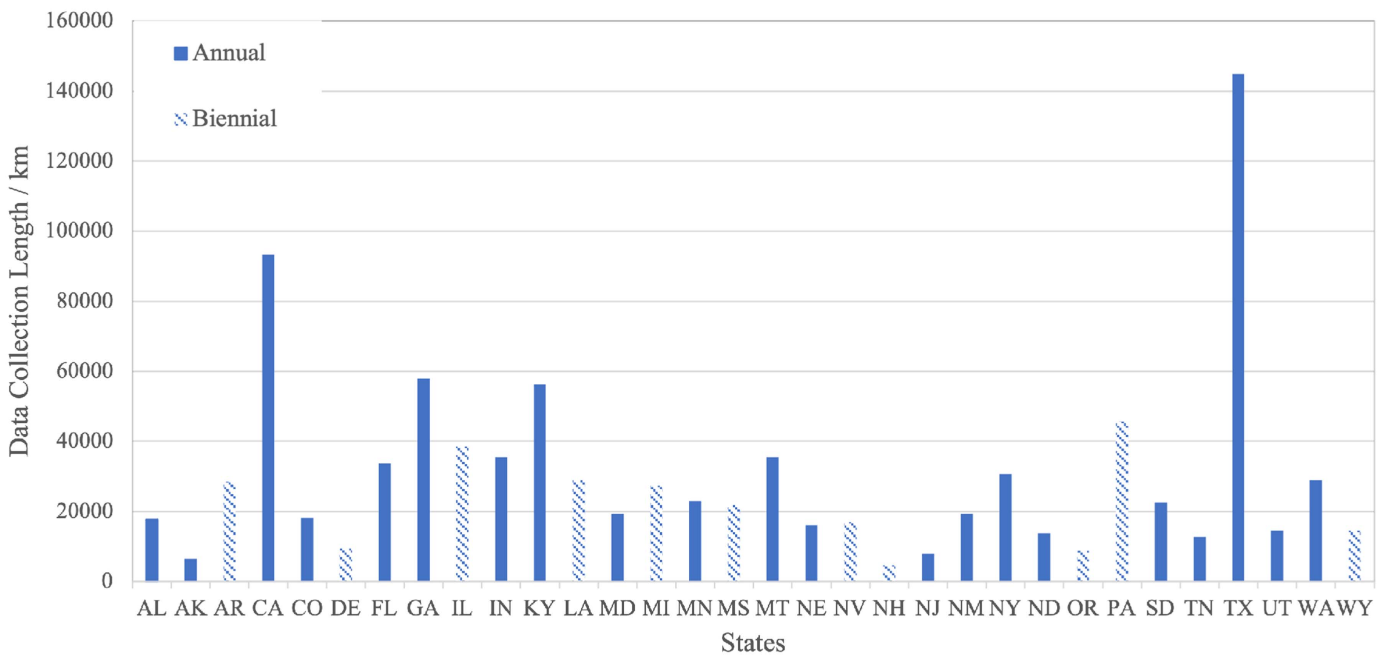

Fig. 3. Data collection lengths and cycles of state highway agencies (32 responses with automated or semi-automated data collection).

J. Transp. Eng., Part B: Pavements, 2022, 148(3): 04022042

QC/QA Processes

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

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 results show 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 results also indicate that some state highway agencies only apply verification and calibration for IRI and rutting, and not for cracking.

QA activities are involved in the data acceptance check process, which includes data allowable range check, data quality validation, data sampling checks with a specific sampling rate, and methods for the automated pavement condition survey. Table 2 provides the QA activities undertaken 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 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 sampling 100% of the collected network length, while a few states apply sampling rates of 0.5%–10% (except for Illinois, which 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 (i.e., percent 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 separating the sample elements (pavement sections) of a population (all pavement sections in the state-maintained network) into subgroups or strata, and then randomly selecting elements from each stratum. Generally, there are more similarities among elements within a stratum than among elements in different strata. Contrary to 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, and they generally take a considerable amount of time. Only a few state highway agencies are working with a vendor or other third party to perform the data acceptance checks.

One of the open questions in the questionnaire concerns the data quality issues that the state highway agencies are facing. Table 3 summarizes the responses of state highway agencies to some typical data quality issues and their possible causes. 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). IRI data

Table 1. Quality control of automated pavement data collection at state highway agencies

Quality control items	Vendor/contractor	Agency staff	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 condition; and 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

Quality assurance items	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 if so 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)			

Note: Numbers in parentheses are the sampling rates of the state DOTs.

collection has caused issues in some state agencies, especially in urban areas. The IRI sensors are very sensitive to the traffic environment, and the low vehicle speeds and frequent stops due to traffic signals in urban areas can cause issues with IRI data collection. Another issue that has been raised is alignment of the vendor collected data with the state referencing systems and standards.

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

Data Collection Cost

Cost is a big concern when state and local agencies switch to automated data collection. Many interviewees from both the government and industry believe that the current automated data collection services are too expensive. An engineer from National Construction Enterprises shared that the cost of manual data collection could be as low as \$15 per hour, while 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 that includes an independent third party for QA that manually reads the image data. The cost of automated data collection is quite significant to customers (agencies) such as small cities and counties. The City of Nevada (Iowa) had five vendors bidding for its automated pavement condition survey. After an evaluation of price and service quality, the price of the pavement condition survey from the chosen vendor was \$105 per mile. Unlike other state and local agencies, MnDOT conducts automated pavement condition data collection itself. One significant advantage is cost reduction—The current cost is approximately \$40 per mile for the annual survey. MnDOT replaces their survey vans every 5–6 years, and on average the total data collection cost is around \$55 per mile.

The final contract with a data collection vendor typically includes a fixed price, plus a per-mile charge. The unit cost of network-level pavement condition data collection depends on the state agency's requirements regarding 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. For instance, an engineer from Applied Pavement Technology (AP Tech) stated that they could adopt various procedures to ensure the survey data are accurate, and each procedure would add a certain amount to the total cost. Thus, if survey data proved acceptable without manual intervention, a 10% surcharge would be added. If not, an additional charge, unpredictable at the time of contract formation, might be needed to make the data acceptable to the agency. For reasons such as this, many engineers advised 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 can improve the work efficiency of pavement engineers. However, the

Table 3. Data quality issues of state highway agencies

Agency	Data quality issues and possible reasons
Alabama DOT	(1) Cracking data has been underreported by vendor since the beginning. It is 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 probably is aligning the vendor collected data to state's linear referencing system for HPMS reporting.
Caltrans	(1) Vendor 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 and focus on project development. (4) Accurate cracking determination appears to be the most challenging.
Colorado DOT	(1) Corner breaks are interpreted manually. (2) Vendor-collected data did not align with the 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 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 classifying 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) Limitations of the imaging system to capture the full extent of the lane in some cases. Also 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 need to recollect data.
Wyoming DOT	Consistency of automated crack detection on JCP.

current automated pavement data collection technologies have a lot of room for improvement, especially regarding data processing algorithms. Pavement engineers claimed that data quality is a serious issue with current automated data collection technologies. Some interviewees noted that data inconsistencies and discrepancies are the main issues after state and local agencies switch to automated data collection. Take as an example a Pavement Management supervisor at TxDOT, who 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 differences between two or more runs of automated data collection at the same pavement locations. Discrepancy means differences between the true distress values and the collected measurements at the same pavement locations. A typical manifestation of discrepancy is a false positive. Several pavement engineers mentioned that current automated pavement survey technologies tend to raise the rate of false-positive, which has caused a significant discrepancy problem. An engineer from Roadway Asset Services (RAS) concluded that the inconsistencies between different pavement condition survey systems, as well as the inconsistencies between human ratings and automated systems, are currently among the biggest challenges. As an example, the City of Austin used three vendors to collect data at different times, and the resulting data inconsistencies have been a big issue. The main reason is that the vendors all used proprietary image data formats that prevent sharing and cross-checking of data.

Meanwhile, some highway agencies are also having trouble matching automated data with historical data that were collected manually. This data continuity issue was raised by many engineers from state and local agencies in interviews. An engineer from Quality Engineering Solutions (QES) said that current technologies have issues in concrete pavement surveys for patched/sealed cracking detection, crack type classification, and crack severity quantification. He also noted, however, that the vendors all have provided timely and effective technical support services when data quality issues were reported.

In contrast, several engineers reported that they are quite satisfied with current automated data collection technologies, especially during the Covid-19 lockdown; these engineers believed that data inconsistency and discrepancy issues are normal and acceptable. This conclusion is supported by the FHWA experience: Most of the annual reports submitted by state highway agencies to the FHWA are based on automated data, and only a small percentage of the reports are found to have data issues.

Promoting Automated Data Collection Technologies

Current automated and semi-automated pavement survey technologies are not fully automated, and have limitations. One pavement engineer with experience in government, industry, and academia shared that the current automated data collection technologies are far removed from fully automated (without human interruption) data collection. A senior pavement engineer from AgileAssets commented that its current pavement survey is not fully automated; for instance, patches still need manual detection. Feedback from other pavement engineers confirms that semi-automated pavement surveys still require a huge amount of manual labor.

The information gained from the interviews shows that while an accuracy of 95% is expected, current data accuracies for automated pavement surveys are around 70%–80%. The engineers from survey companies insisted that current automated technologies need to be improved, and that artificial intelligence (AI) technologies should be applied to improve the data quality. Some companies have started using AI technologies. 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 still needs data pre-treatment. The lack of training data due to the low availability of annotated ground truth image data, and difficulties with sharing data in the public domain, have caused delays in developing and using AI. Current AI-driven automated pavement condition survey technologies cannot detect all types of pavement distresses. An important reason for this is that the current distress definition standard is designed for human raters, and not for computers, so some distresses cannot adequately be detected or measured by current automated technologies.

Implementation of QA

As mentioned previously, the main issues with current automated data collection technologies are data inconsistencies and discrepancies, for which manual correction is needed to make the data usable. This problem was 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 agencies said that they would not trust the survey data without validation. Many interviewees indicated that they spent substantial staff time doing image checks for data QA. In four states, it took engineers months to validate the yearly pavement survey data. For instance, a district engineer at TxDOT mentioned that it is hard to verify the data from the whole network because it would take engineers months to review all the data. In Mississippi, 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 data validation in Caltrans is conducted manually by three engineers working full-time. QA is time-consuming and labor-intensive, and involves much subjectivity. This feedback mirrors the findings from the reviews of agency 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 (e.g., VDOT contracts with QES) to examine the survey data provided by the data collection vendors, which is expensive. Some highway agencies and municipal governments treat automated data collection, data processing, and QA as individual services, and contract with different entities to conduct the pavement condition evaluation work.

More suggestions regarding the implementation of QA concern quantifying QA. An engineer from Applied Research Associates suggested that a threshold could be used to define data quality for QA purposes, which can vary depending 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 said that extending the automated pavement condition data collection technologies to project-level data collection is necessary. However, current technologies are not fully ready for project-level data collection, as was noted by an engineer from the City of Austin who worked with a vendor for automated pavement data collection. The engineer noted that there were many issues with IRI data collected in the city network. The engineer, who worked on PMS data, spent much time on QA. Another engineer revealed that the project-level data collected in the TxDOT San Antonio District is more than just IRI data, and thus also requires TxDOT staff to utilize a falling weight deflectometer (FWD) and ground penetrating radar (GPR) to collect structure data. 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 inaccurate GPS referencing is another important issue that limits the use of automated data collection at the project level. Pavement engineers from Virginia DOT, TxDOT Dallas and Pharr Districts, NCE, and StreetSaver provided the same feedback.

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

Conclusions

Automated pavement condition surveying is essential, as it saves much time and cost for customers who need pavement condition data (state highway agencies and municipal governments). 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 obtain practical insights about the issues that government, industry, and academia perceive about automated or semi-automated data collection. Based on the survey questionnaire and interviews, the following findings are observed:

- Most state and local highway agencies conduct automated or semi-automated pavement data collection. Many state highway agencies have more than ten years of experience in using automated or semi-automated technologies. Contracting with a vendor is the 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 use standards of their own design.
- Data collection items and frequency vary among state highway agencies, although most states collect state maintained network-level pavement condition data annually.
- QC of automated or semi-automated pavement condition data is typically conducted by vendors and agency staff. QA activities are mainly conducted by agency staff, and takes a substantial amount 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 makes huge demands on agency staff time.

Innovation in the QA process is needed to promote automated pavement condition surveys.

- The high cost of automated or semi-automated pavement condition data collection ($100–$200) per mile is a major 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 in automated or semi-automated pavement condition data collection is data quality, which presents as data inconsistencies and discrepancies. Agencies both contract with third parties and use internal staff to address the data quality issues and make the data usable. Vendors provide timely and effective technical support services to help the agencies and/or third parties address data quality issues.
- Although existing automated pavement condition data collection technologies are widely adopted for network-level pavement surveys, data inconsistency and discrepancy problems must be corrected through an intensive QA process. These inconsistency and discrepancy problems are due to immature data collection technologies, as well as to vendors' use of proprietary image data formats that prevent sharing and cross-checking of data.
- Extending automated data collection to the project-level is a tremendous need for pavement engineers, but current technologies are still too immature in data accuracy and precision for project-level use. AI concepts and models are expected to be more widely used to further improve and optimize image processing 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

AASHTO. 2006. *Standard practice for definition of terms related to quality and statistics as used in highway construction*. AASHTO R 10-06. Washington, DC: AASHTO.

Chang, G., A. Gilliland, G. R. Rada, P. A. Serigos, A. L. Simpson, and S. 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. Washington, DC: Federal Highway Administration.

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, DC: FHWA.

FHWA (Federal Highway Administration). 2018. *Guidelines for development and approval of state data quality management programs*. Washington, DC: FHWA.

Flintsch, G. W., and K. H. McGhee. 2009. Vol. 401 of *Quality management of pavement condition data collection*. Washington, DC: Transportation Research Board.

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.

McGhee, K. H. 2004. *NCHRP Synthesis 334: Automated pavement distress collection techniques, Report No. Project 20-5, Topic 34-04*. Washington, DC: Transportation Research Board.

Pierce, L. M., G. McGovern, and K. Zimmerman. 2013. *Practical guide for quality management of pavement condition data collection*. Washington, DC: Federal Highway Administration.

Pierce, L. M., and N. D. Weitzel. 2019. *NCHRP synthesis 531: Automated pavement condition surveys. Report No. Project 20-05, Topic 49-15*. Washington, DC: National Cooperative Highway Research Program. Transportation Research Board.

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*. Washington, DC: Federal Highway Administration.

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*. Washington, DC: Federal Highway Administration.

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*. Washington, DC: TRB.

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." In *Proc., 25th European Signal Processing Conf. (EUSIPCO), 2031-2035*. Washington, DC: IEEE.

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.

TxDOT (Texas Department of Transportation). 2016. *Texas transportation asset management plan*. Austin, TX: TxDOT.

TxDOT (Texas Department of Transportation). 2018. *Quality management plan for pavement data collection*. Austin, TX: TxDOT.

TxDOT (Texas Department of Transportation). 2020. *Improve data quality for automated pavement distress data collection*. Austin, TX: TxDOT.

Wang, K. C. P., Q. J. Li, G. Yang, Y. Zhan, and Y. Qiu. 2015. "Network level pavement evaluation with 1 mm 3D survey system." *J. Traffic Transp. Eng.* 2 (6): 391–398. https://doi.org/10.1016/j.jtte.2015.10.005.

Yang, G., K. C. P. 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, 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." *Comput.-Aided Civ. Infrastruct. Eng.* 34 (3): 213–229. https://doi.org/10.1111/mice.12409.

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." *Comput.-Aided Civ. Infrastruct. Eng.* 32 (10): 805–819. https://doi.org/10.1111/mice.12297.