GEORGIA DOT RESEARCH PROJECT 12-10

FINAL REPORT

A REMOTE SENSING AND GIS-ENABLED HIGHWAY ASSET MANAGEMENT SYSTEM PHASE 2

OFFICE OF PERFORMANCE-BASED MANAGEMENT RESEARCH
15 KENNEDY DRIVE
FOREST PARK, GA 30297-2534

<table>
<tr><td colspan="2">1.Report No.:
FHWA-GA-17-1210</td><td>2. Government Accession No.:</td><td colspan="2">3. Recipient's Catalog No.:</td></tr>
<tr><td colspan="3" rowspan="2">4. Title and Subtitle:
A Remote Sensing and GIS-enabled Asset Management System (RS-GAMS) Phase 2</td><td colspan="2">5. Report Date: February 2018</td></tr>
<tr><td colspan="2">6. Performing Organization Code:</td></tr>
<tr><td colspan="3">7. Author(s): Yichang (James) Tsai; Zhaohua Wang</td><td colspan="2">8. Performing Organ. Report No.:</td></tr>
<tr><td colspan="3" rowspan="2">9. Performing Organization Name and Address:
Georgia Institute of Technology
790 Atlantic Drive
Atlanta, GA 30332-0355</td><td colspan="2">10. Work Unit No.:</td></tr>
<tr><td colspan="2">11. Contract or Grant No.:
PI# 0010896</td></tr>
<tr><td colspan="3" rowspan="2">12. Sponsoring Agency Name and Address:
Georgia Department of Transportation
Office of Research
15 Kennedy Drive
Forest Park, GA 30297-2534</td><td colspan="2">13. Type of Report and Period Covered:
Final; May 2012 – May 2017</td></tr>
<tr><td colspan="2">14. Sponsoring Agency Code:</td></tr>
<tr><td colspan="5">15. Supplementary Notes:
Prepared in cooperation with the U.S. Department of Transportation, Federal Highway Administration.</td></tr>
<tr><td colspan="5">16. Abstract:
The objective of this project is to validate the use of commercial remote sensing and spatial information (CRS&SI) technologies, including emerging 3D line laser imaging technology, mobile light detection and ranging (LiDAR), image processing algorithms, and GPS/GIS technologies, to improve transportation asset data collection, condition assessment, and management. The research focuses include the validation of automatic asphalt pavement crack classification, concrete pavement distress detection, pavement marking retroreflectivity condition assessment using mobile LiDAR, and the long-term monitoring of pavement conditions in terms of cracking and rutting. The automatic classification of two major types of cracks, load cracking and block cracking, were comprehensively validated and show very promising results. The detection of concrete pavement distresses, including cracking, faulting, spalling, and shoulder joint distress, were validated using experimental tests conducted on I-516 and I-16. The preliminary study on pavement marking retroreflectivity condition assessment using mobile LiDAR has shown very promising results, though more large-scale tests are needed to include more marking materials. The implementation and applications of the research outcomes have been presented to demonstrate the benefit of this research project. Finally, conclusions are made and recommendations are suggested.</td></tr>
<tr><td colspan="3">17. Key Words:
3D Laser; Mobile LiDAR; Cracking; Faulting; Marking</td><td colspan="2">18. Distribution Statement: No Restriction</td></tr>
<tr><td>19. Security Classification (of this report):
Unclassified</td><td>20. Security classification (of this page): Unclassified</td><td>21. Number of Pages: 271</td><td colspan="2">22. Price:</td></tr>
</table>

GDOT Research Project No. 12-10

Final Report

A REMOTE SENSING AND GIS-ENABLED HIGHWAY ASSET MANAGEMENT SYSTEM PHASE 2

By

Yichang (James) Tsai, Ph.D., P.E.

Zhaohua Wang, Ph.D., P.E.

Georgia Institute of Technology

Contract with

Georgia Department of Transportation

In cooperation with

U.S. Department of Transportation

Federal Highway Administration

February 2018

Acknowledgements

The work described in this report was supported by the Georgia Department of Transportation (GDOT) through Research Project 12-10 and the U. S. Department of Transportation Office of the Assistant Secretary for Research and Technology (USDOT/OST-R). We would like to thank the technical advisory committee, including Mr. Butch Wlaschin, Mr. Thomas Van, Mr. Mark Swanlund, and Mr. Mike Moravec from the Federal Highway Administration; Mr. Bouzid Choubane and Mr. Bruce Dietrich from Florida DOT; Ms. Angela Alexander, Mr. Gregory Mayo, and Mr. Eric Pitts from Georgia DOT; Mr. Rick Miller from Kansas DOT; and Dr. Judith B Corley-Lay from North Carolina DOT, for their valuable contributions to this project. The strong support provided by Georgia DOT, including Mr. Buddy Gratton, Mr. Gerald Ross, Mr. Russell McMurry, Mr. David Crim, Ms. Georgene Geary, Ms. Jane Smith, Ms. Meg Pirkle, Mr. Timothy Dale Brantley, Ms. Ernay Robinson, Mr. Jim Leben, Mr. Terry Rutledge, Mr. Ritchie Swindell, Mr. Rick Deaver, Mr. David Jared, Mr. Binh Bui, and Ms. Gretel Sims is highly appreciated. We would like to thank the following state DOTs and companies for their technical support: Connecticut DOT, Florida DOT, Kansas DOT, Louisiana DOTD, Mississippi DOT, Oklahoma DOT, South Carolina DOT, Texas DOT, INO/Pavemetrics, and Trimble Geo3D. We would like to thank the research team members, Dr. James Lai, Dr. Anthony Yezzi, Ms. Yiching Wu, Dr. Vivek Kaul, Dr. Feng Li, Dr. Chengbo Ai, Mr. Zachary Lewis, Mr. Thibaut Dusanter, Dr. Chenglong Jiang, Dr. Chieh Wang, and Mr. Geoffrey Price. Mr. Jon Lindsay's editing of this report is also appreciated. Especially, we would like to thank Mr. Caesar Singh, Mr. Vasanth Ganesan, and Mr. Binh Bui for their assistance in managing this project.

Table of Contents

List of Figures

List of Tables

Executive Summary

The U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology (USDOT/OST-R) and the Georgia Department of Transportation (GDOT) co-sponsored this research project to validate the application of commercial remote sensing and spatial information (CRS&SI) technology. GDOT supported the validation of intelligent transportation asset inventory, including asphalt pavement crack classification, concrete pavement distress detection, and pavement marking retroreflectivity condition assessment using an intelligent Remote Sensing and GIS-based Asset Management System (RS-GAMS). USDOT and Georgia Institute of Technology (Georgia Tech), through cost sharing, sponsored the integration and calibration of CRS&SI technology, the Georgia Tech Sensing Vehicle (GTSV) that can be operated non-destructively at highway speed.

1. Research Focuses

The following research focuses address GDOT's and other transportation agencies' needs regarding the use of CRS&SI technologies to improve the data collection and condition assessment of the two types of transportation assets: pavements and pavement markings.

- Automatic asphalt pavement crack classification
- Concrete pavement distress detection
- Pavement marking condition assessment
- Long-term monitoring of pavement distresses, including cracking and rutting

2. Research Outcome and Major Findings

The following conclusions are based on the four research focuses:

- **Automatic asphalt pavement crack classification**

  This research validated the performance of an automatic crack classification algorithm that uses 3D line laser imaging (named "3D laser" hereafter) data and was previously developed by the PI. The algorithm is based on a multi-scale crack fundamental element (CFE) model and uses crack detection results, i.e. crack maps, as inputs, which have been validated in RS-GAMS Phase 1. The classification of the two most commonly occurring cracks, load cracking and block cracking, defined in GDOT's pavement distress survey manual, Pavement Condition Evaluation System (PACES), were implemented, tested, and validated. The promising results demonstrate that the algorithm is capable of transforming raw sensing data and detected crack maps into useful decision-support information, including crack types, severity levels, and extents.

  In the test on 1,069 pavement 3D laser images (each image covers, approximately, a 5-meter long and 4-meter wide section of pavement), GDOT pavement engineers visually reviewed each image and established the ground truth. Based on the comparison between the ground truth and the automatically classified results, the algorithm showed an accuracy of 92.2% when classifying load cracking at four severity levels and 98.1% when classifying block cracking at three severity levels. Another test was conducted on ten 100-ft test sections that were selected on State Route (SR) 236, SR 275, and SR 67 in Georgia. In each test section, GDOT

pavement engineers visually identified the crack types, severity levels, and extents in the field that were used as ground truth. Among ten test sections, four were surveyed by accurate measurements using a measuring wheel, while the other six sections were surveyed by visual estimation following GDOT's current survey practices. Then, comparison was made on deducts derived from the automatic crack classification and the field visual survey. For the wheel-measured sections, the average absolute difference of total deducts was 3.25 out of 100 (a pavement rating is between 0 and 100), and for the visually-estimated sections, the average absolute difference was 5 out of 100. Both differences were within the error tolerance based on GDOT's current practice (5 out of 100).

The validation results show that the use of 3D laser data and the corresponding algorithms could improve the productivity and efficiency of collecting pavement distress information. Moreover, the fine-grained sensing data also opened the opportunity to improve existing pavement management by adding more detailed decision-support information that could not be acquired previously.

- **Concrete pavement distress detection**

  This study validated the detection and measurements of various concrete pavement distresses (including cracking, faulting, spalling, and shoulder joint distresses) using pavement 3D laser data. The validation results demonstrate the potential for using 3D laser data for automatically detecting distresses in concrete pavements. The test sites were selected on interstate highways I-16 and I-516.

The validation of concrete pavement cracking detection using 3D laser data showed acceptable performance. The automatic crack detection results were compared to the manually digitized ground truth using a buffered Hausdorff scoring method that was developed in RS-GAMS Phase 1. The results showed that detection of cracks on I-516 (mainly longitudinal cracks) is accurate and robust; however, the detection of cracks on I-16 (mainly transverse cracks) is not as good as on I-516. The larger data acquisition interval in the driving direction, which was about 5 mm, might be the reason that some transverse cracks cannot be captured by 3D laser data. In comparison, the transverse resolution is about 1 mm, which can better capture the longitudinal cracks. Limited to the laser data resolution, hairline cracks (thinner than 2 mm) are still challenging for automatic detection.

The validation of concrete joint faulting measurement showed that it is very feasible to use 3D laser data for collecting faulting data at highway speed. Using the regression-based method, the automatic faulting measurements were consistent with manually measured ground truth using the Georgia Faultmeter in both the well-controlled lab test and the field test.

The accuracy of automatic spalling detection varied for different sizes of spalling. Spalling with widths greater than 90 mm can be successfully detected; the detection accuracy was reduced, but still acceptable for widths between 50 and 90 mm, while spalling was hard to detect when the width was less than 50 mm. Though some small spallings were not successfully detected, they can be clearly observed on the 3D laser data. Thus, the automatic detection algorithm could be further improved to handle such cases.

Since there is no dedicated application that is commercially available for shoulder joint distress detection, we explored the feasibility of using an automatic spalling detection algorithm to detect shoulder joint distress. The larger extent and depth of shoulder joint distress make them distinctive on laser range data and easier to detect. On the selected representative cases, the automatic detection results were visually consistent with field observation. However, it should be noted that due to the transverse coverage of the current pavement surface laser data (about 4 meters), the shoulder area might be missed when the vehicle wanders. In addition, a specific shoulder joint distress detection algorithm is needed to further ensure an accurate and robust detection.

- **Pavement marking retroreflectivity condition assessment**

  This research was to establish the correlation between the retroreflectivity measured by handheld retroreflectometer and the retro-intensity acquired from a LiDAR point cloud. Establishing a reliable correlation is the key step for assessing pavement marking retroreflectivity conditions using a mobile LiDAR. In this preliminary study, thermoplastic and waterborne paint were selected, which are the most commonly used pavement marking materials. Test sites were selected on Ferst Drive, Hemphill Avenue, and 17th Street on/near the Georgia Tech campus.

  It was discovered that the retro-intensity values acquired from mobile LiDAR are not sensitive to ambient temperatures and have an average standard deviation less than 0.0041. The retro-intensity acquired from mobile LiDAR held good repeatability on

the tested thermoplastic and waterborne materials with an average standard deviation of 0.0044.

There was an exponential correlation between retroreflectivity and retro-intensity with an R-square of 0.9525 for thermoplastic and 0.9267 for waterborne paint. The correlation between retroreflectivity and retro-intensity might be sensitive to different bead formulas of the pavement marking material. Separate correlation curves might be needed not only for different pavement marking material category, e.g. thermoplastic, waterborne, etc., but also for different bead formulas in the same material category. Based on the correlation results, a preliminary retro-intensity threshold corresponding to the minimum retroreflectivity (100 mcd/m^2/lux) defined in the MUTCD could be defined as 0.4263, with a 95% confidence interval ranging from 0.4035 to 0.4505 for thermoplastic and 0.3521, with a 95% confidence interval ranging from 0.2973 to 0.4264 for waterborne materials. Using the established correlations, a mobile LiDAR-based pavement marking retroreflectivity condition assessment method can be further developed.

- **Long-term monitoring of crack deterioration**

  This research project studied pavement deterioration behavior using long-term monitoring 3D laser data over time. A detailed level representation of crack information, (including length, width, orientation/direction, position/location, intensity, pattern, etc., which were derived by using 3D laser data, crack detection algorithms, and the multi-scale CFE model) was applied to study the temporal crack deterioration behavior. The results showed that it could potentially be used to support

the following applications: 1) the fundamental study of pavement mechanistic and pavement deterioration behavior, 2) validation of current pavement design methods and development of new design concepts and methods, 3) determination of adequate treatment methods and strategies based on pavement distress characteristics and their deterioration behavior, 4) development of accurate and reliable forecasting models, and 5) development of cost-effective pavement management operations/practices, such as intelligent crack sealing planning. Compared to the frequently used lab simulation data or well-controlled experimental data (e.g. AASHO road test), the in-service pavement data used in this study represents true crack deterioration mechanisms and enables the ability to explicitly connect pavement condition deterioration with the real-world factors that cause it. While it is not feasible to comprehensively cover the entire field of research, this study is believed to be a transformative, concrete, first step in changing the way researchers have approached sensing-based infrastructure condition monitoring and risk assessment.

- **Long-term monitoring of rutting deterioration**

  This research characterized the 3D rut shape and analyzed its deterioration behaviors using multiple scales and long-term pavement 3D laser data. Rut parameters, including transverse profile-based parameters, longitudinal parameters, and temporal parameters, were defined and proposed. Descriptive statistics and 2D and 3D visualizations were used to analyze the deterioration behaviors of rutting.

  Transverse profile-based rut parameters, including rut depth, rut cross-sectional area, percent deformation, and total absolute distortion, show good correlation among each

other, providing consistent information about the conditions of ruts. Longitudinal parameters also showed similar trends as the aforementioned parameters, indicating that ruts not only grow in depth and area but also grow longitudinally in length and volume. Temporal parameters provide a direct means to quantify the deterioration of rutting. These parameters can be very useful, especially when the rate of deterioration is high. The multi-scale analysis results show that seasonal variations can have appreciable effect on the deterioration of ruts. Temporal parameters derived from short periods (e.g., a few months), can be affected by seasonal variations. This finding suggests that, for different deterioration analysis applications, different analysis periods should be applied in order to obtain consistent results. The comparison among three different routes showed that traffic and roadway characteristics also play an important role in the deterioration of ruts. Other factors, such as the age and design of the pavement, can also contribute to the actual deterioration behavior of ruts. 2D and 3D visualization of ruts at the individual level shows the importance and benefits of having registered long-term 3D pavement data. Details of how ruts deteriorate, e.g., the development of dual-wheel shaped ruts and the growth in rut length and volume, were able to be identified in the selected sections. This information can further support the diagnosis of rutting and inform data-driven maintenance decisions.

3. Recommendations for Future Research

The following suggest the future research and implementation:

- With the promising results from the automatic asphalt pavement crack classification for load cracking and block cracking, it is recommended that the automatic classification be extended to other types of distresses as defined in GDOT pavement distress manual. In addition, the algorithms can be easily extended to other crack survey protocols used by different state highway agencies because of the flexibility provided by the crack CFE model.
- The validation results for automatic concrete pavement faulting measurements showed very good consistency with manual measurements using a Georgia Faultmeter. A large-scale pilot study with a state DOT, e.g., GDOT, is suggested to automate the network-level faulting measurements. This can significantly improve the productivity, data accuracy, and data coverage.
- The concrete pavement crack detection shows promising results. However, it is difficult to detect hairline, transverse cracks due to the relatively coarser data resolution in the driving direction using the current 3D laser device. Thus, to capture hairline cracks, the data capture frequency and resolution of a 3D laser device needs to be further improved. In addition, to automate the crack evaluation for concrete pavements, automatic crack classification algorithms, which can be based on the work we have done for asphalt pavements, need to be developed.
- New algorithms need to be developed because the automatic detection for concrete spalling doesn't work well on those with widths less than 50 mm. In addition, new algorithms are needed for automatic shoulder joint distress detection.
- The pavement marking validation results indicate that pavement marking retroreflectivity conditions could be measured and evaluated using mobile LiDAR

that can be operated at highway speed. However, the testing samples in this research project are limited. It is suggested that large-scale testing on more marking materials be conducted before it is implemented.

Chapter 1 Introduction

1. Background and Research Need

Roadway asset inventory and condition assessment are critical for making data-driven, informed decisions on cost-effective management of roadway assets. However, the commonly used manual data collection is labor-intensive, time-consuming, costly, dangerous, and error-prone. There is an urgent need to improve the reliability and productivity of roadway asset inventory and condition assessment. The Commercial Remote Sensing and Spatial Information (CRS&SI) technologies, including 3D line laser imaging (named "3D laser" hereafter), airborne and terrestrial Light Detection And Ranging (LiDAR), GPS/GIS, Inertial Measurement Unit (IMU), and image/signal processing, have dramatically advanced in the past few years. There is great potential to validate the applications of these technologies to bring new capabilities to roadway asset management. For this purpose, a research project, entitled "A Remote Sensing and GIS-enabled Asset Management System (RS-GAMS) Phase 2," was sponsored by the U. S. Department of Transportation Office of the Assistant Secretary for Research and Technology (USDOT/OST-R) and the Georgia Department of Transportation (GDOT). This project proposes integrating the aforementioned CRS&SI technologies into an intelligent sensing system, Georgia Tech Sensing Vehicle (GTSV), to bring new capabilities to roadway asset inventory, condition assessment, and management. In the project, GDOT supported the validation of asphalt pavement crack classification, concrete distress detection, pavement marking retroreflectivity condition evaluation, and long-term pavement condition monitoring; the USDOT and the Georgia Institute of

Technology (Georgia Tech) shared the cost and sponsored the integration and calibration of CRS&SI technology, i.e., GTSV that can be operated non-destructively at highway speed.

The technology developed in this project is an extension of the outcomes from RS-GAMS Phase 1 (RP 10-08), which was sponsored by the USDOT/OST-R and GDOT. FIGURE 1.1 shows the architecture of RS-GAMS. It includes the components of the sensing system, data processing and collection, data integration and management, and decision support. New capabilities and functions for data collection, condition assessment, and decision support can then be developed under this system framework. RS-GAMS Phase 1 has established the preliminary framework and also uses two assets (sign and asphalt pavement), as shown in the light grey color in FIGURE 1.1, to demonstrate its capability.

With the refinement of the integrated sensing system, new applications were developed and validated in the Phase 2 study, as shown in the dark grey color in FIGURE 1.1. The applications of RS-GAMS in Phase 1 focused on the development and validation of CRS&SI technology on: (1) asphalt pavement rutting identification and measurement, (2) asphalt pavement crack detection and the performance evaluation of different pavement crack detection algorithms, (3) sign asset inventory, and (4) sign retroreflectivity condition assessment.

RS-GAMS Phase 2 uses the same framework developed in Phase 1, but extends it to additional roadway assets, including concrete pavements, pavement markings, cross slopes, curves, and pavement widths that are important for transportation agencies'

engineering practices. RS-GAMS Phase 2, sponsored by GDOT, focuses on the following applications and their validations: a) asphalt pavement distress classification, b) concrete pavement distress detection, c) pavement marking retroreflectivity condition assessment, and d) long-term monitoring on pavement conditions and analysis.

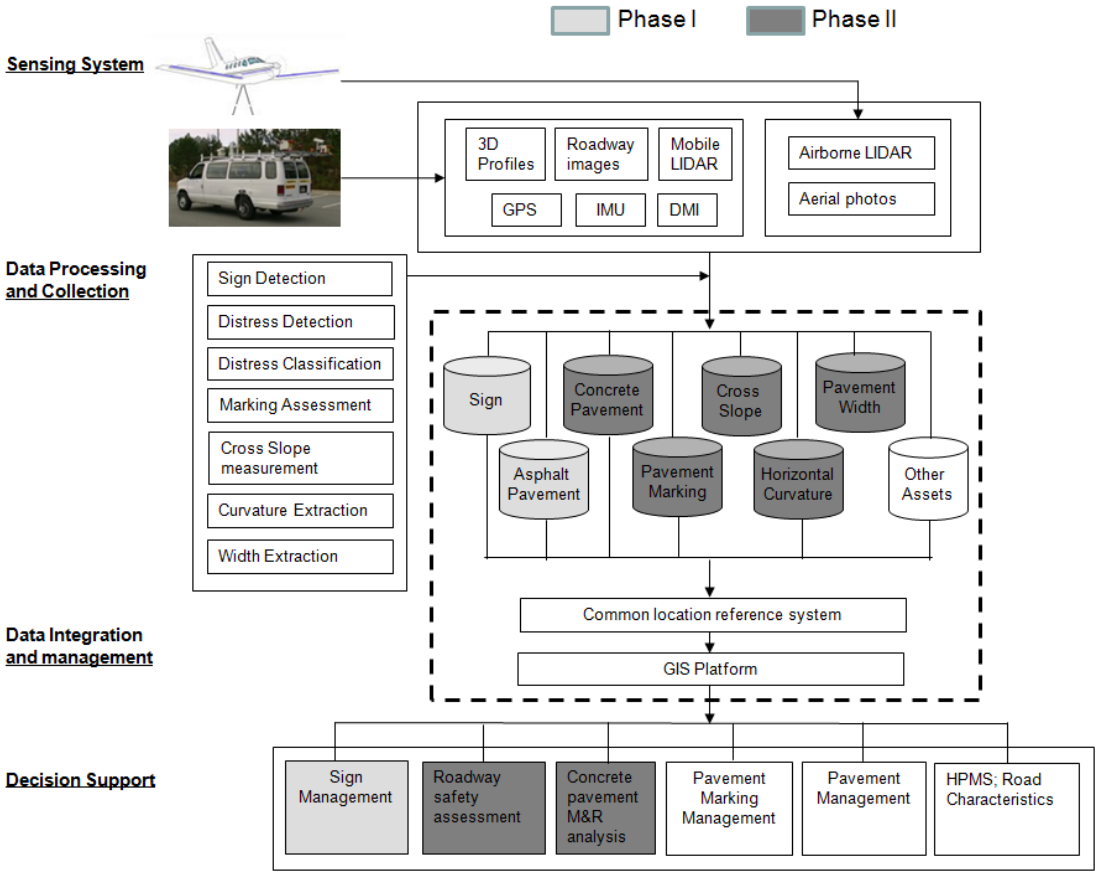


FIGURE 1.1: RS-GAMS architecture

2. Research Objective

The objective of this project is to integrate CRS&SI technologies, including emerging 3D laser, signal/image processing, and GPS/GIS technologies, into an intelligent sensing system to bring new capabilities to roadway asset inventory, condition assessment, and

management. The developed sensing system can be used to improve the reliability and productivity of pavement distress detection/classification, pavement marking condition assessment, and inventory of pavement cross slopes, roadway curvatures, and pavement widths. Project development steps are to (1) refine and calibrate the integrated sensing system, (2) test and validate the sensing system using real-world data, and (3) quantify the research benefits.

3. Research Tasks

The following research tasks address GDOT's and other transportation agencies' needs regarding the use of CRS&SI technologies to improve data collection, condition assessment, and management of two types of transportation assets: pavement and marking.

- **Validate asphalt pavement crack classification**

  This research task validated the performance of an automatic crack classification algorithm (previously developed by the PI) using pavement surface laser data. The algorithm is based on a multi-scale crack fundamental element (CFE) model and uses the crack detection results, which have been comprehensively validated in RS-GAMS Phase 1, as inputs. The classification of two commonly occurring types of cracking, load cracking and block cracking, defined in GDOT's pavement distress survey manual, Pavement Condition Evaluation System (PACES), were implemented, tested, and validated.

- **Validate concrete pavement distress detection**

  This research task validated the detection and measurements of various concrete pavement distresses, including cracking, faulting, spalling, and shoulder joint distress, using pavement surface laser data. The test sites were selected on two interstate highways: I-16 and I-516.

- **Feasibility study of using mobile LiDAR for pavement marking retroreflectivity condition assessment**

  This study focuses on evaluating the feasibility of using mobile LiDAR to perform pavement marking retroreflectivity condition assessment. Thermoplastic and waterborne paint, the most commonly used pavement marking materials, were selected to conduct the feasibility study. Test sites were selected on Ferst Drive, Hemphill Avenue, and 17th Street on/near the Georgia Tech campus.

- **Long-term monitoring of pavement distresses, including cracking and rutting**

  This task includes the continuous long-term pavement condition monitoring after the completion of the US DOT project and after the completion of the tasks listed above. This task is crucial for applying the analyzed 3D laser profile data for the actual long-term pavement performance, including cracking and rutting. The activities of this task include continuous monitoring of the pavement performance, data processing and analysis, and continuous refinement of the developed sensing technology.

4. Report Organization

This report summarizes the results of the research project co-sponsored by USDOT and GDOT. Though the results of validating data collection have been incorporated into the report for USDOT in consideration of completeness, the study of the long-term crack and rutting deterioration behavior were exclusively contained in this report for GDOT.

This report is organized into eight chapters. Chapter 1 summarizes the research background, need, and approaches. Chapter 2 presents the validation results for asphalt pavement crack classification. Chapter 3 presents validation results for concrete pavement distress detection. Chapter 4 presents the feasibility study of pavement marking retroreflectivity condition assessment. Chapter 5 presents the long-term deterioration of asphalt pavement cracks. Chapter 6 presents the deterioration of asphalt pavement rutting using long-term monitoring 3D laser data. Chapter 7 presents the implementation of the research results. Chapter 8 summarizes the conclusions and makes recommendations for future research.

Chapter 2 Validation of Automatic Asphalt Pavement Crack Classification

This chapter presents the research results for validating the automatic crack classification for asphalt pavements.

1. Introduction

Pavement surface distress evaluation is an essential component of a pavement management system (PMS). Cracking, as one of the most common types of pavement distresses, is caused by constant traffic loading, asphalt aging, environmental impact, and/or improper structural design. Progressive cracking can weaken pavement structures because it allows water and other foreign objects into the base and accelerates pavement deterioration. The proper treatment of pavement cracks at the optimal timing is important for cost-effective pavement maintenance. Many transportation agencies, including GDOT, have invested major resources in their pavement condition survey and evaluation procedures to enhance their decision-making capabilities.

Traditionally, the collection of pavement crack data is usually done by visual inspection in the field. For example, in GDOT's practice, pavement surveyors manually identify the presence, types, and severity levels of cracking, and then estimate/measure and record the extent of each crack type and severity level. Such a task is dangerous, subjective, costly, time-consuming, and labor-intensive. Therefore, automatic pavement cracking evaluation is gaining attention among transportation agencies and researchers. To transform the manual practice into an automatic procedure, at least two steps should be taken. Crack

detection is the first step in identifying the presence of pavement cracking from the collected pavement data and generate a crack map. Crack classification is the second step to automatically identify the types and severity levels from the detected crack map. Automatic crack detection using the emerging 3D laser technology has been validated in the RS-GAMS Phase 1 study. This study focuses on the validation of crack classification.

An automatic crack classification and quantification method has been previously developed by the PI following the GDOT PACES distress protocol. The objective of this research is to conduct a validation of the automatic crack classification and quantification using the crack map detected by using 3D laser data. This chapter is organized as follows. After an introduction in Section 1, Section 2 summarizes the major findings through the literature review, including both current pavement distress protocols used in different federal and state transportation agencies and existing automatic crack classification algorithms. Section 3 presents the basic concept of the automatic crack classification and quantification method previously developed by the PI, followed by a comprehensive validation in Section 4. The GDOT PACES distress protocol is used as an example to demonstrate the performance of automatic asphalt crack classification. Section 5 presents a case study that applies the automatic crack classification and quantification method on the interstate highways, which are usually high-traffic-volume roadways and are challenging for the traditional manual survey; it visualizes the network-level pavement condition to support decision-making. In the end, Section 6 summarizes the major findings.

2. Summary of Literature Review

The literature review in this chapter covers two aspects. First, the current pavement cracking survey practices from different federal and state transportation agencies are reviewed to better understand the objective of the automatic crack classification task and identify the similarities and differences between protocols. Then, the existing automatic crack classification algorithms are reviewed to understand the current status of automatic algorithm development and identify the gap between the current status and the desired crack classification objective. The major findings from the perspective of agency's current pavement distress protocols are summarized as follows:

- **Complexity**: Transportation agencies usually have complicated and diverse crack definitions in their survey practice. The real-world distress protocols from transportation agencies usually involve human identification of complex crack patterns. Furthermore, for most protocols, multiple severity levels are defined under the same crack type, which also have an important impact on the pavement condition evaluation.
- **Subjectivity**: Though the detailed crack definitions are described in the protocols, the crack patterns on the real pavement are still diverse. Even for experienced human surveyors, different engineers may provide different crack survey results on the same roadway section. An automatic crack classification and quantification method will overcome this issue by providing consistent results.
- **Diversity**: Different protocols are developed for different purposes. For example, the GDOT PACES protocol has a causal-based crack definition so that the crack survey results can directly contribute to GDOT's maintenance operations; the

Long-term Pavement Performance (LTPP) protocol is a research-oriented data collection and focuses more on how to precisely record every single crack on the road. The data can be used for Mechanistic-Empirical Pavement Design Guide (MEPDG) calibration and validation. Different purposes lead to diverse crack definitions. The diversity between protocols makes it difficult to transform the automatic crack classification algorithm from one protocol to another

The major findings from the perspective of existing automatic crack classification algorithms are summarized as follows:

- **Preliminary outcomes**: In terms of classification, most of the existing studies provide the following crack type outcomes: longitudinal, transverse, diagonal, alligator, and block, etc. Such an outcome format simplified the classification problem but limited the real-world implementation. It is hard to establish the correlation between these preliminary outcomes and the actual crack definitions used by transportation agencies.
- **Limited crack characteristics**: Crack orientation and crack amount are two major characteristics that have been studied for automatic classification purposes. However, manual survey protocols require more characteristics, such as crack location, crack intersection, and possible polygon patterns.
- **No severity levels**: The classification of different crack severity levels is explored in the existing literature.
- **Lack of flexibility**: Most existing algorithms take the entire image as an input. The crack classification will be less effective when multiple crack types appear on

the same image, and this creates a significant challenge for accurate crack extent quantification.

- **Limited performance**: Most studies only validate their methods on a small set of pavement images. Some large-scale studies have shown that current automatic crack classification and quantification survey results usually have a poor correlation with manual survey results.

Overall, an automatic crack classification and quantification method remains a challenge from the perspective of practical implementation. Targeting the gap between the current status of automatic crack classification and the real-world crack survey protocols, the following section will present a multi-scale crack representation approach for crack classification.

3. Multi-Scale Crack Representation for Crack Classification

Crack pattern, together with other crack properties, is crucial for differentiating crack types and severity levels in transportation agencies' pavement survey practices. The PI proposed a multi-scale CFE model in a previous study (Tsai & Jiang, 2012). This model topologically provides rich crack properties at three different scales (fundamental crack properties, aggregated crack properties, and clustered CFE geometrical properties) to support the development of an automatic crack classification method. It also standardizes crack performance measures for different transportation agencies and effectively deals with the diversity among different protocols. FIGURE 2.1 shows crack properties at three scales defined in the CFE model.

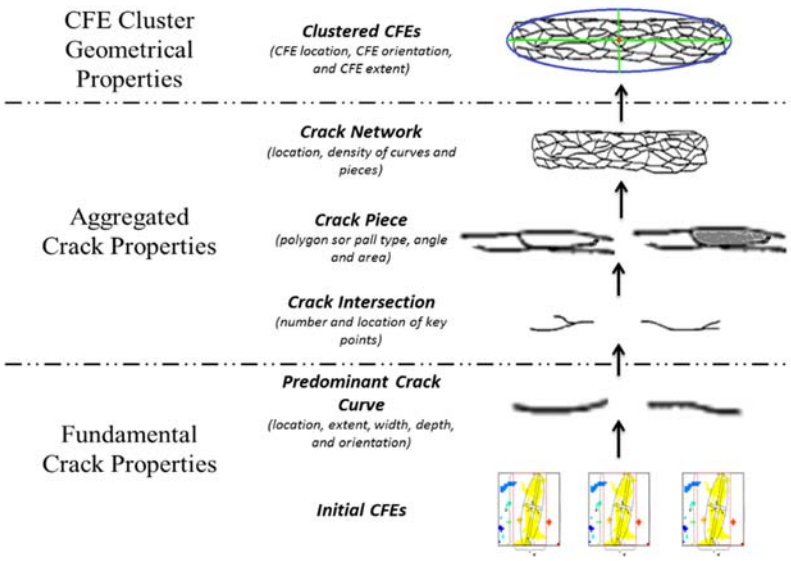


FIGURE 2.1: Multi-scale crack properties from CFE model (Tsai & Jiang, 2012)

The input for the multi-scale CFE model is the automated detected crack map. This crack map can be derived from either 2D or 3D data. Fundamental crack properties focus on each crack segment and describe the fundamental and physical properties of the cracks, including crack width, depth, length, etc.; aggregated crack properties focus more on crack patterns inside the CFE and represent how cracks interact with each other, such as crack intersection, crack polygon, crack density, etc.; clustered CFE geometrical properties treat each CFE as a whole and describe its overall properties, including the CFE center, orientation, length, and width. From the bottom, the model represents the physical characteristics of pavement cracks; from the top, it tends to mimic the pavement engineers' manual evaluation procedure in the field (from the macro to the micro level observation). When experienced pavement engineers conduct a condition survey, they do not usually measure the crack width and depth first; instead, they first identify a group of cracks that should be clustered together as one element (CFE); then, they look at the

crack pattern inside the CFE, and, finally, they measure the physical and fundamental crack properties. By clearly defining three scales of crack representation, the model can better incorporate both fundamental crack properties and human judgment. The method's features are explained in the following aspects:

- Consistency: The crack properties extracted through this model are independent from different pavement distress protocols. These crack properties, such as crack length, crack width, etc., are directly derived from the pavement data and the detected crack map and are not influenced by agencies' protocols or survey practices. This consistency is critical as a standardized crack performance measure.
- Flexibility: These properties can be easily transformed between different protocols to develop a corresponding crack classification method through certain rules and criteria (as shown in FIGURE 2.2).

Based on this multi-scale crack representation concept, the PI further developed a crack classification and quantification method following the GDOT PACES distress protocol. The method itself doesn't require specific data format or crack detection algorithms, but the accurate crack detection using 3D pavement surface data is expected to provide better input.

4. Validation of the Automatic Crack Classification Method

4.1 Validation objective

The objective of this study is to validate the feasibility of automatic crack classification and quantification on asphalt pavements. To do so, we first validated the accuracy of

automatic crack classification method on a large and diverse set of pavement images and then compared the results from the automatic crack classification and quantification with the field survey results conducted by experienced pavement engineers. The GDOT PACES distress protocol was selected as an example to conduct the validation. Load cracking and block cracking (in GDOT, transverse cracking is also categorized as block cracking. For clarity's sake, we use block/transverse (B/T), hereafter), as two predominant asphalt pavement distresses in Georgia, are the focus of this study. The crack classification method can be easily extended to other crack-related distresses and protocols.

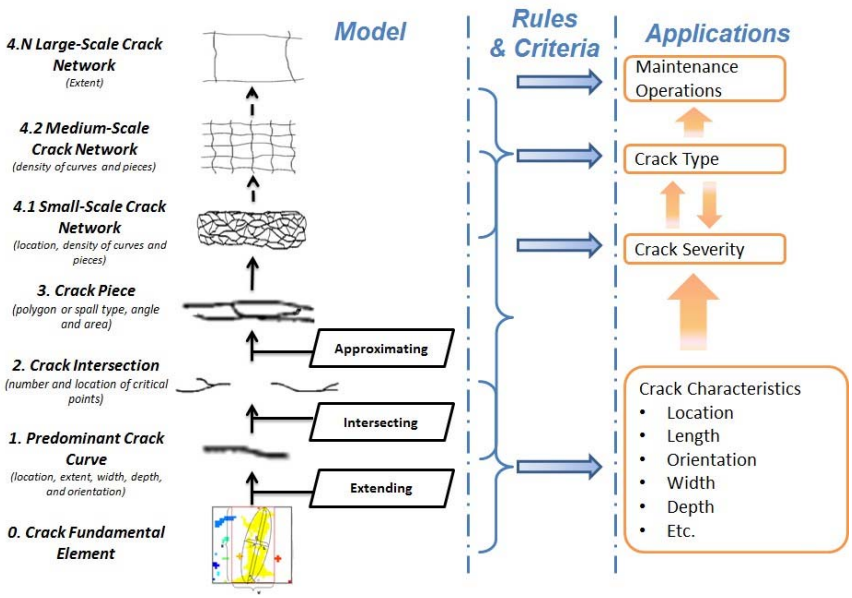


FIGURE 2.2: Crack classification using extracted crack properties

4.2 Experimental design

Two series of experimental tests are conducted to fulfill the validation objective:

- An image-based validation test is conducted on State Route 236 / Lavista Road in Atlanta, Georgia, to validate the accuracy of crack classification. With help from GDOT pavement engineers, each pavement image was visually reviewed, and the presence and severity level of load cracking and B/T cracking manually labeled. In this test, we validated and conducted an image-by-image comparison between the automatic crack classification results and manual labels.
- A site-based validation test was conducted to compare the results from automatic and manual field survey. It validated the accuracy of both crack classification and quantification. Following GDOT's current pavement surface condition survey practice, ten 100-ft pavement sections were selected on SR 236, SR 275, and SR 67. Manual field surveys for load cracking and B/T cracking were conducted by GDOT liaison engineers. On the other hand, automatic crack evaluation also provides crack types, severity levels, and extents on these selected sections. The deduct values caused by load cracking and B/T cracking were calculated and compared between automatic and manual crack surveys.

In this study, GTSV was used to collect the 3D laser data, and the crack detection was conducted using the analyzer software along with the system. The dimension of each pavement image is five meters long in the driving direction and approximately four meters wide in the transverse direction. A large and diverse dataset was prepared to cover diverse crack characteristics.

- Crack type and severity levels: load cracking and B/T cracking were the focus of this test, and the experimental data covered all four severity levels for load

cracking and three severity levels for B/T cracking. A total of 2,335 load cracking images were prepared, including 881 images with no load cracking, 1,145 images with Severity Level 1, 158 images with Severity Level 2, 136 images with Severity Level 3, and 15 images with Severity Level 4. A total of 1,224 B/T cracking images were prepared, including 97 images with no B/T cracking, 1006 images with Severity Level 1, 108 images with Severity Level 2, and 13 images with Severity Level 3. The crack type and severity level were labeled through manual review.

- Combination of multiple crack types: in the current manual field survey, the combination of multiple crack types mostly introduced the inconsistency into the survey results. A large portion of experimental data were pavement images that contained both load cracking and B/T cracking in order to validate the performance of automatic crack classification on these cases.
- Crack location: the cracks were randomly located in the left wheel path, right wheel path, and non-wheel path regions in the experimental data. The correct extraction of crack location was the basis for load and B/T cracking classification.
- Crack pattern: as the severity level increased, the crack patterns in the experimental data changed gradually from a single crack line to intersected crack networks. The capability of interpreting complex and diverse crack patterns is another important aspect to be validated for the automatic crack classification.

The following subsections will present the validation tests and results in detail.

4.3 Image-based validation

An image-based validation was first conducted on State Route 236 / Lavista Road in Atlanta, Georgia. The selected project, which has an excessive amount of load cracking and B/T cracking, starts from Milepost 0 (Atlanta city limit) to about Milepost 6.8 (I-285 bridge). The experimental data included 2,335 images for load cracking and 1,224 images for B/T cracking. The data from each crack type and severity level were randomly separated into two sets: 70% of the data were used for model training and calibration, while the rest, 30%, were used for testing.

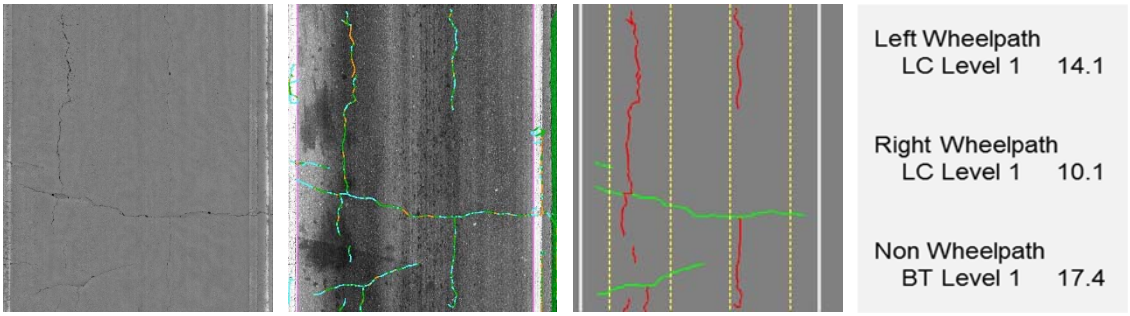
Performance of load cracking classification

For load cracking, the training set consisted of 619 images of no load cracking, 798 images of Severity Level 1, 108 images of Severity Level 2, 99 images of Severity Level 3, and 11 images of Severity Level 4. The algorithm performance on the test set is shown in Table 2.1.

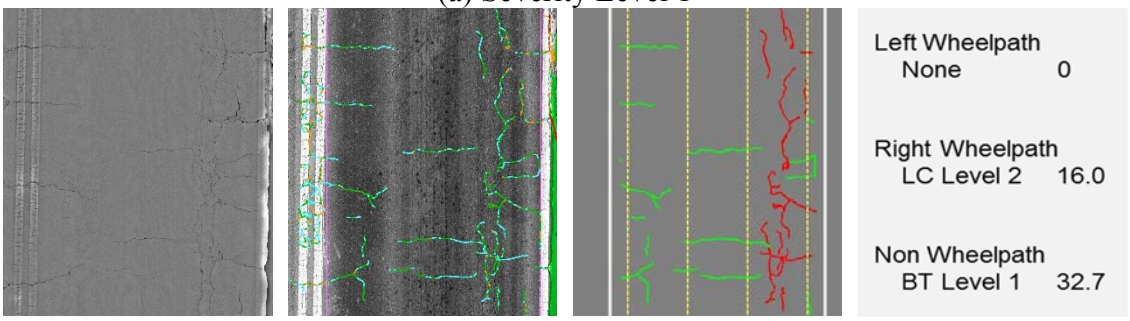
As shown in Table 2.1, a total of 700 test images are selected. The algorithm has overall high classification accuracy at about 92.2%. From the perspective of recall (i.e. the ratio of correctly classified cases to total actual cases), a larger portion of Severity Level 2 was not correctly classified compared to other severity levels. From the perspective of precision (i.e. the ratio of correctly classified cases to total classified cases), the classification for Severity Level 2 was quite low, which was mainly due to the clear difference between the sample sizes. Some representative cases for load cracking and their automatic classification and quantification outcomes are shown in FIGURE 2.3.

Table 2.1: Performance of Load Cracking Classification

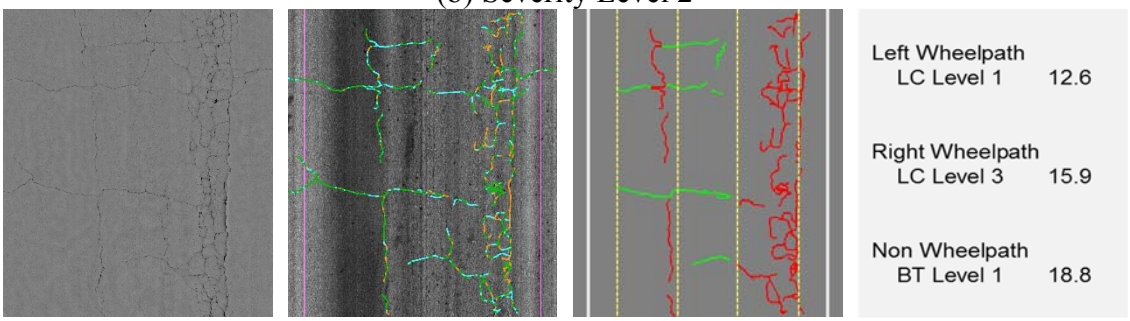
		Classified Severity Level						
		None	Level 1	Level 2	Level 3	Level 4	Total	Recall (%)
Actual Severity Level	None	247	15	0	0	0	262	94.3
	Level 1	10	317	20	0	0	347	91.4
	Level 2	0	6	42	2	0	50	84.0
	Level 3	0	0	2	35	0	37	94.6
	Level 4	0	0	0	0	4	4	100.0
	Total	257	338	64	37	4	700	
	Precision (%)	96.1	93.8	65.6	94.6	100.0		92.2



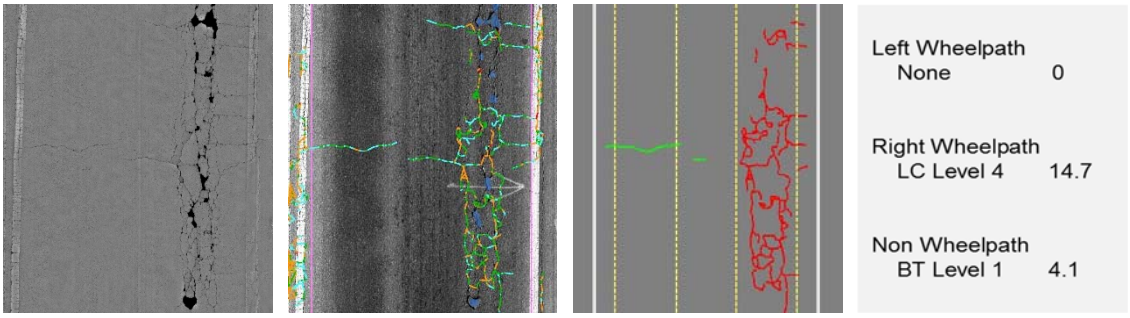
(a) Severity Level 1



(b) Severity Level 2



(c) Severity Level 3



(d) Severity Level 4

FIGURE 2.3: Representative load cracking evaluation outcomes, from left to right: range image, crack map on intensity image, load cracking (red) and B/T cracking (green), and evaluation outcomes (unit: foot)

Performance of B/T cracking classification

For B/T cracking, the training set consisted of 68 images of no B/T cracking, 703 images of Severity Level 1, 74 images of Severity Level 2, and 10 images of Severity Level 3. The algorithm performance on the test set is shown in Table 2.2.

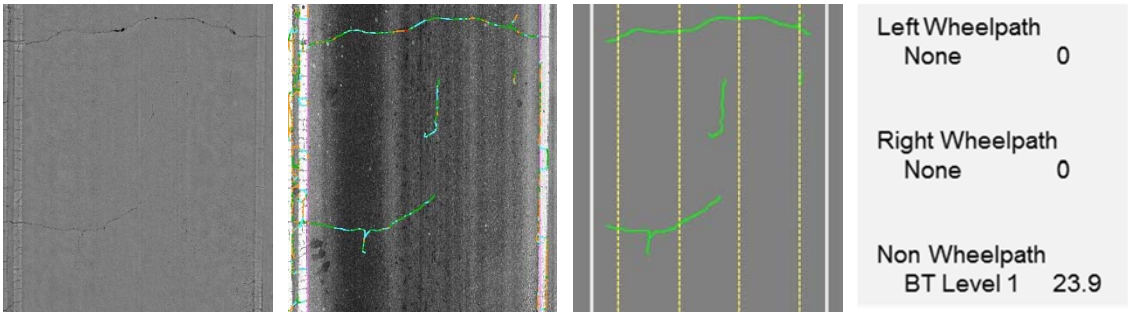
As shown in Table 2.2, 369 test images are selected. The algorithm has, overall, high classification accuracy at about 97.2%. The results are also promising from the perspectives of both precision and recall. One possible reason is that the three target classes have quite distinctive differences in the crack properties. More data are still needed to further provide a more robust classification to Severity Levels 2 and 3. Some representative cases for B/T cracking and their automatic classification and quantification outcomes are shown in FIGURE 2.4.

Table 2.2: Performance of B/T Cracking Classification

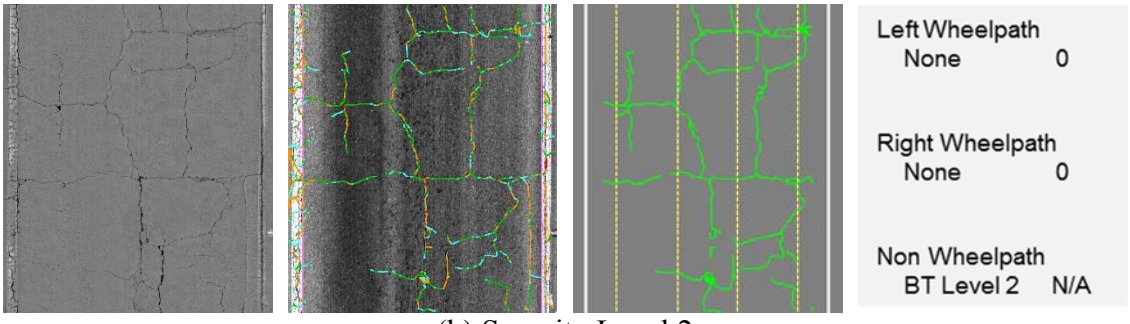
		Classified Severity Level					
		None	Level 1	Level 2	Level 3	Total	Recall (%)
Actual Severity Level	None	27	2	0	0	29	93.1
	Level 1	1	298	4	0	303	98.3
	Level 2	0	0	31	3	34	91.2
	Level 3	0	0	0	3	3	100.0
	Total	28	300	35	6	369	
	Precision (%)	96.4	99.3	88.6	50.0		97.2

4.4 Site-based validation

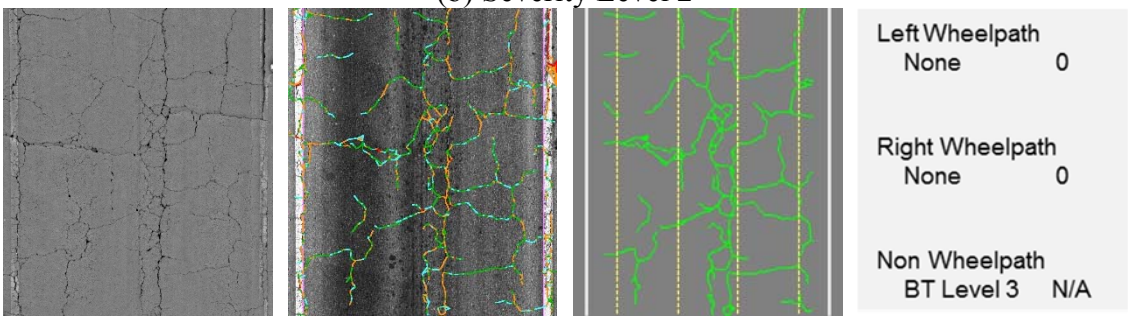
Section validations were further conducted to compare the results from automatic and manual crack surveys. Following GDOT's current crack survey practice, ten 100-ft pavement sections were selected on SR 236, SR 275, and SR 67. Manual field surveys for load cracking and B/T cracking were conducted. On the other hand, the automatic crack classification and quantification method provided crack types, severity levels, and extents. The deduct values were calculated and compared between a manual survey and an automatic crack evaluation, and the results are presented as follows:



(a) Severity Level 1



(b) Severity Level 2



(c) Severity Level 3

FIGURE 2.4: Representative B/T cracking evaluation outcomes, from left to right: range image, crack map on intensity image, load cracking (red) and B/T cracking (green), and evaluation outcomes (unit: foot)

Validation on wheel-measured sections

On four selected sites, GDOT pavement engineers visually identify the crack types and severity levels in the field, and the crack extents were measured using a measuring wheel (which is time consuming) in order to reduce the potential bias through the manual survey. The experimental results are shown in Table 2.3. The columns on the left are field crack measurement results and their corresponding deducts, and the columns on the right are automatic results. Based on the experimental results, the overall deducts given by automatic crack evaluation were close to those in the field measurement and range image inspection. For the four selected sites, the average absolute difference of the overall deduct between automatic crack evaluation and wheel measurement is 3.25, which is

within the error tolerance (5-10 deduct points) in GDOT's current survey practice. The differences were mainly caused by 1) load cracking at Severity Level 1 is partially captured or detected, which leads to a lower deduct value in automatic evaluation results (e.g. SR 236 Site #3), and 2) a slight measurement difference on high severity levels (e.g. load cracking Severity Level 4 at SR 236 Site #2) results in a big difference on the deduct points. Both reasons will be further illustrated in the following subsection.

Table 2.3: Section Validation with Wheel Measurement

(a) SR 236 Site #1

	Wheel Measurement		**Automatic Evaluation**	
	Extent(%)	Deduct	Extent(%)	Deduct
Load Lvl 1	56	15	48	15
B/T Lvl 1	100	18	100	18
Overall		33		33

(b) SR 236 Site #2

	Wheel Measurement		**Automatic Evaluation**	
	Extent(%)	Deduct	Extent(%)	Deduct
Load Lvl 1	30	10	25	9
Load Lvl 2	7	9	7	9
Load Lvl 4	11	29	7	22
B/T Lvl 1	99	18	100	18
Overall		47		40

(c) SR 236 Site #3

	Wheel Measurement		**Automatic Evaluation**	
	Extent(%)	Deduct	Extent(%)	Deduct
Load Lvl 1	41	13	27	9
Load Lvl 2	2	2	0	0
B/T Lvl 1	100	18	100	18
Overall		31		27

(d) SR 275 Site #1

	Wheel Measurement		Automatic Evaluation	
	Extent(%)	Deduct	Extent(%)	Deduct
Load Lvl 1	71	15	47	15
Load Lvl 3	9	19	8	17
B/T Lvl 1	78	18	80	18
Overall		37		35

Validation on visually-estimated sections

Beside the four sites where crack extents were measured in detail using a measuring wheel, six other sites were also evaluated in this study. On these sites, GDOT pavement engineers visually identified the crack types and severity levels, and they visually estimated the crack extents in the field (following GDOT's current practices). The experimental results are shown in Section 4.5.

On most sites, the overall deducts given by automatic crack evaluation were still close to those in the field estimation and range image inspection. For SR 275 Sites #2 and #3, the hairline cracks in the field were not captured or were only partially captured by the 3D line laser imaging system, which significantly impacted the overall deduct. For these six sites, the average absolute difference of the overall deduct between automatic crack evaluation and visual estimation was 5 out of 100, which is within the error tolerance of GDOT's current survey practice.

4.5 Limitations of the crack classification and quantification

Through the validation, the major issues of crack classification and quantification using 3D pavement data can be categorized as follows:

System resolution of current 3D line laser imaging technology

The resolution of the 3D line laser imaging system used in this study was 1 mm in the transverse direction and 5 mm in driving direction. Our previous study shows that the system has the capability to capture cracks wider than 2 mm, but has only limited performance when dealing with cracks around 1mm wide (hairline cracks). Most hairline cracks were only partially captured or were completely missing on the range image (as shown in FIGURE 2.5), which leads to certain differences between automatic evaluation outcomes and manual survey results.

Table 2.4: Section Validation with Visual Estimation

(a) SR 275 Site #2

	Visual Estimation		**Automatic Evaluation**	
	Extent(%)	Deduct	Extent(%)	Deduct
Load Lvl 1	90	15	0	0
B/T Lvl 1	10	4	0	0
Overall		19		0

(b) SR 275 Site #3

	Visual Estimation		**Automatic Evaluation**	
	Extent(%)	Deduct	Extent(%)	Deduct
Load Lvl 1	12	7	0	0
B/T Lvl 1	30	7	25	6
Overall		14		6

(c) SR 275 Site #4

	Visual Estimation		**Automatic Evaluation**	
	Extent(%)	Deduct	Extent(%)	Deduct
Load Lvl 1	20	8	14	7
B/T Lvl 1	60	11	45	9
Overall		19		16

(d) SR 275 Site #5

	Visual Estimation		Automatic Evaluation	
	Extent(%)	Deduct	Extent(%)	Deduct
Load Lvl 1	100	15	57	15
B/T Lvl 1	60	11	57	11
Overall		26		26

(e) SR 67 Site #1

	Visual Estimation		Automatic Evaluation	
	Extent(%)	Deduct	Extent(%)	Deduct
Load Lvl 1	100	15	51	15
B/T Lvl 1	100	18	81	18
Overall		33		33

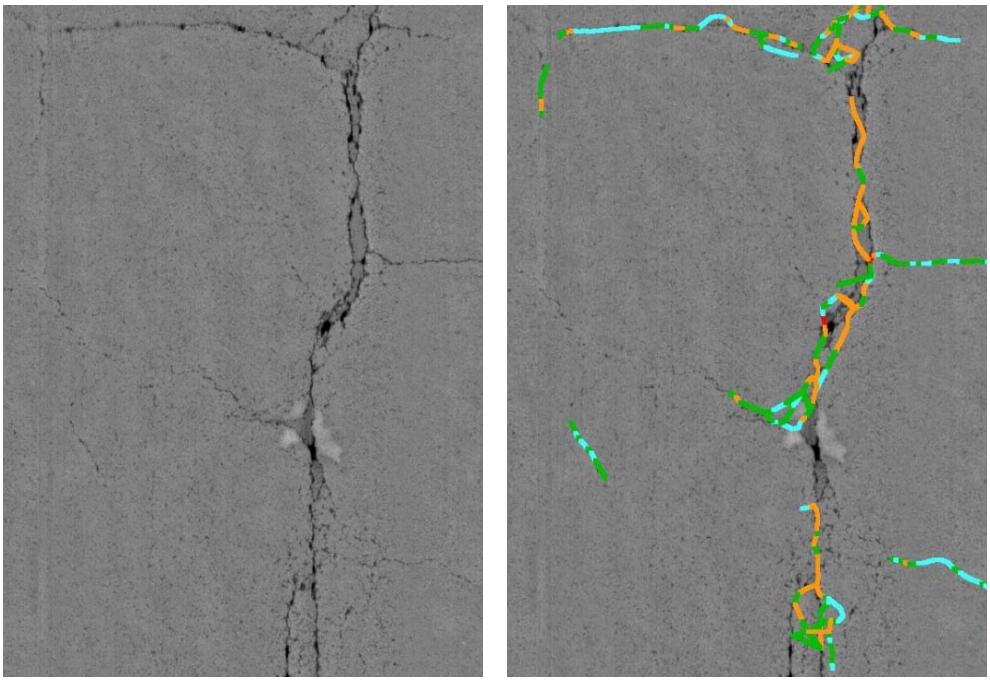
(f) SR 67 Site #2

	Visual Estimation		Automatic Evaluation	
	Extent(%)	Deduct	Extent(%)	Deduct
Load Lvl 1	70	15	88	15
Load Lvl 2	15	15	0	0
B/T Lvl 1	20	6	15	5
Overall		21		21

FIGURE 2.5: Limitation of current technology on hairline cracking (images from SR236 Site #3)

Limitation of automatic crack detection

Furthermore, the automatic crack detection algorithm cannot be guaranteed to provide a precise crack map on every image. As shown in FIGURE 2.6, some cracks were only partially detected, and the pavement areas with clearly double-crack lines were detected as only a single crack line. The limitation of automatic crack detection also impacts the performance of the subsequent crack classification and measurement.



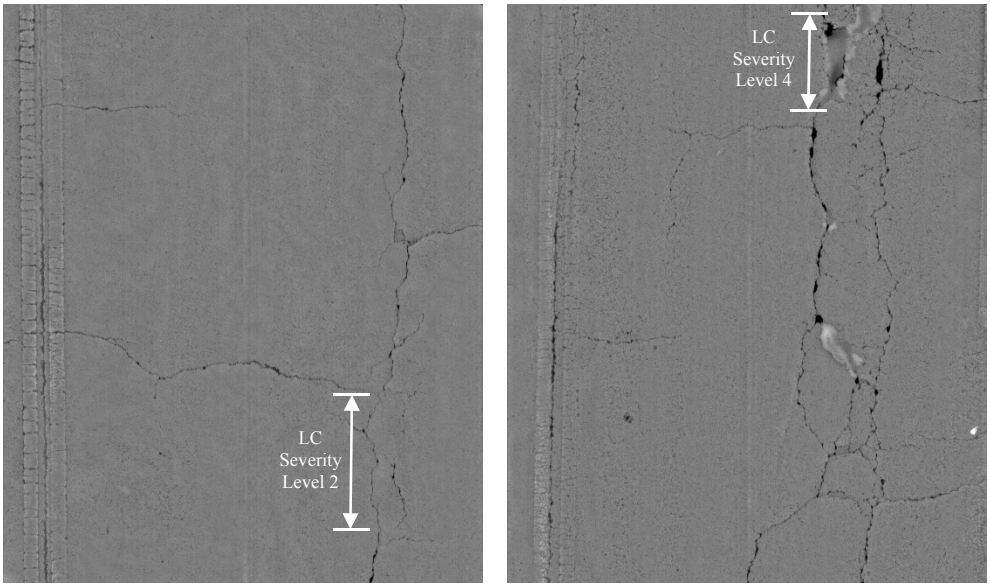
(a) Range image (b) Detected crack map

FIGURE 2.6: Limitation of crack detection algorithms (images from SR 236 Site #2)

Measurement biases through crack quantification

In this study, each wheelpath was assumed to contain only one load cracking severity level; we intentionally chose the predominant load cracking severity level when multiple severity levels occurred in the same wheelpath. As shown in FIGURE 2.7(a), although a short segment of load cracking Severity Level 2 was identified on the bottom of the image, the whole wheelpath was still measured and recorded as Severity Level 1. Similarly, the short segment of load cracking at Severity Level 4 in FIGURE 2.7(b) was disregarded, which might have had a larger impact on the deduct calculation, since high severity levels usually correspond to a much higher deduct value. However, since this automatic crack evaluation was conducted with 100-percent coverage, such biases were

assumed to counterbalance each other, and, therefore, they didn't have a significant impact on the overall condition assessment.



(a) SR236 Site #3 (b) SR236 Site #2

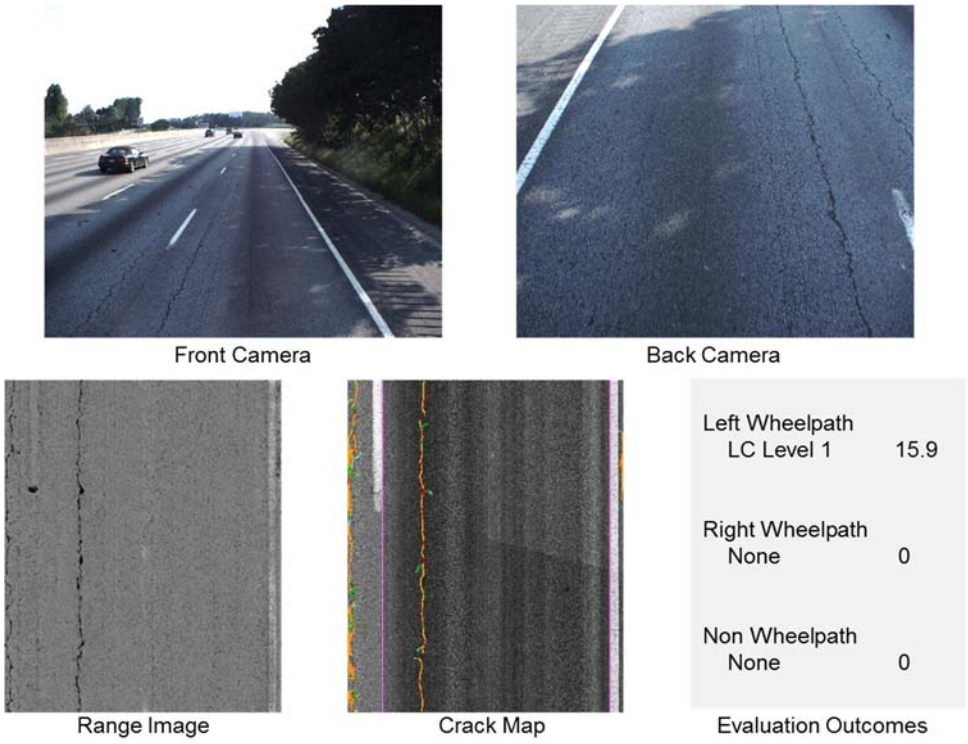
FIGURE 2.7: Assumptions of crack quantification in this study

5. Case Study on Interstate Highway I-85

We have demonstrated the performance of crack classification and quantification through a series of image-based and site-based field validation tests in previous sections. This section presents an outreach study by applying it on the interstate highways, which are usually high-traffic-volume roads and are challenging for traditional manual surveys. The selected test site is I-85 southbound between Exit 104 and Exit 85 (about 19 miles). According to the information provided by GDOT, the road segment inside the I-285 perimeter was in poor condition and would be resurfaced soon.

The pavement data was collected on the test site and the crack classification and quantification was conducted automatically. FIGURE 2.8 shows some representative

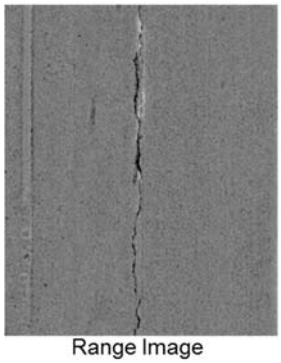
samples and the corresponding automatic crack evaluation outcomes. Due to the high traffic volume, it was difficult to conduct a field manual survey to establish the ground truth; therefore, the video-log images from front and back camera were provided for comparison purposes.



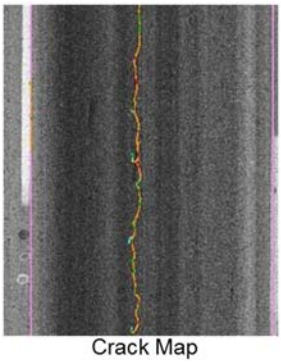
(a) Load cracking

Front Camera

Back Camera



Range Image

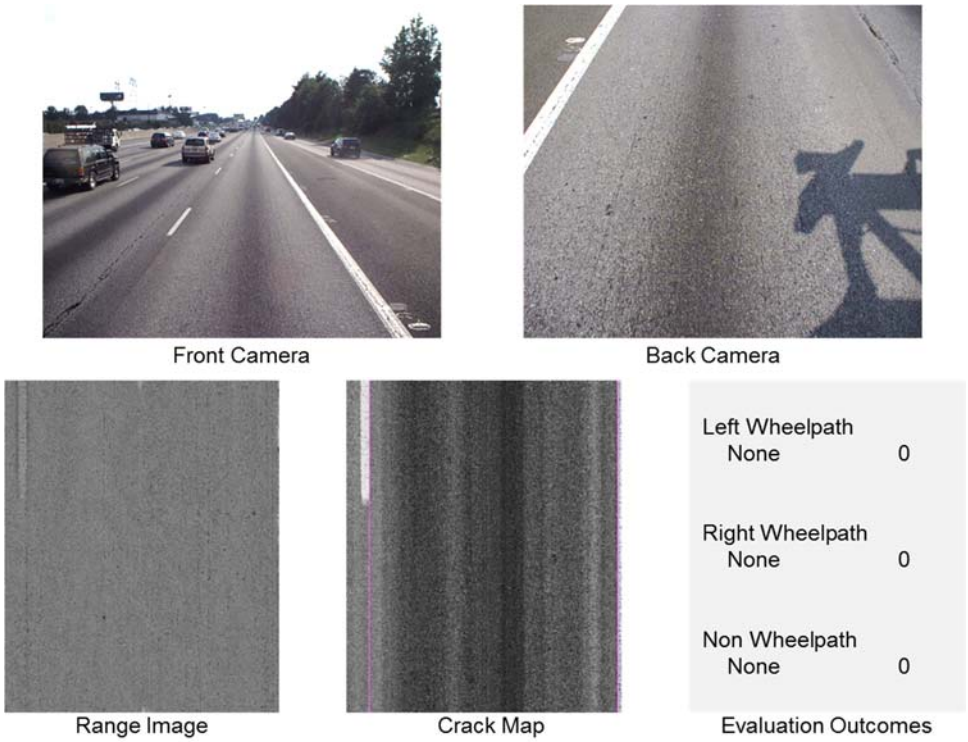


Crack Map

Left Wheelpath	
None	0
Right Wheelpath	
None	0
Non Wheelpath	
BT Level 1	18.8

Evaluation Outcomes

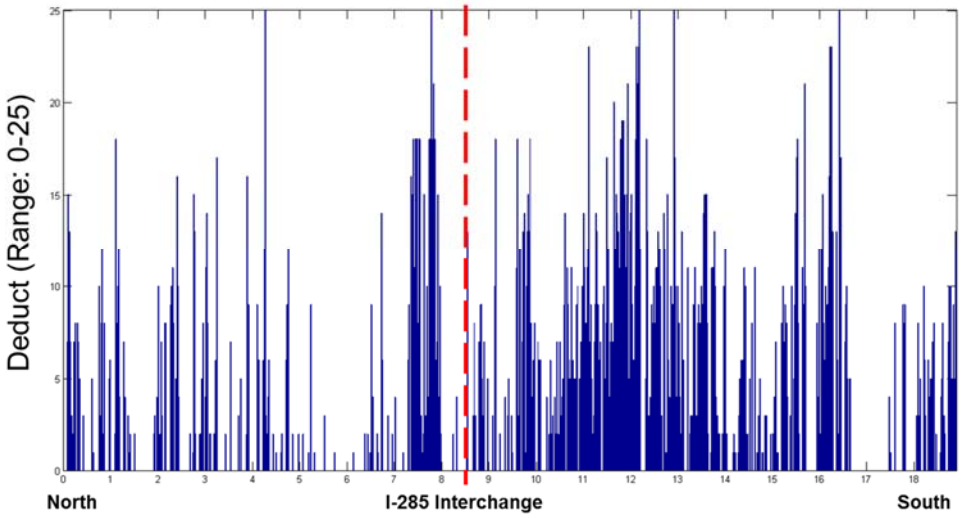
(b) B/T cracking



(c) Good pavement

FIGURE 2.8: Representative samples on the selected interstate test site

Based on the automatic crack evaluation outcomes, the deduct distribution of load and B/T cracking can be generated and visualized on a GIS map (as shown in FIGURE 2.9). The pavement condition inside the I-285 perimeter is clearly worse than that of outside I-285. The automatic crack evaluation using 3D pavement data is promising for transforming the sensing data and detected crack map into decision support information; it is especially beneficial on the high-volume-traffic roads, such as interstate highways, where traditional manual surveys are difficult to conduct.



(a) Deduct distribution

(b) GIS visualization

FIGURE 2.9: Overall crack condition on the selected interstate test site

6. Summary

Emerging 3D laser technology demonstrated its great potential on pavement condition surveys. Automatic crack detection using the 3D pavement surface data has been

validated to provide a more accurate crack map. With the detailed preservation of crack characteristics, it is expected to better support automatic crack classification and quantification. This study validates the performance of a previously developed crack classification and quantification method. The major findings are summarized as follows:

First, automatic crack classification and quantification is feasible on asphalt pavements. In this study, we validated crack classification and quantification following GDOT's PACES distress protocol. Through the validate tests, the automatic method demonstrated high accuracy in classifying load cracking, B/T cracking, and their severity levels, and the crack quantification results are close to manual field surveys conducted by experienced GDOT liaison engineers.

- An image-based validation was first conducted to validate the accuracy of crack classification. Actual pavement data were collected, manually reviewed, and labeled by GDOT pavement engineers to establish the ground truth. The automatic crack classification method shows an accuracy of 92.2% on classifying load cracking and its severity levels and 97.2% on classifying B/T cracking and its severity levels.
- A site-based validation was then conducted to compare the results from automatic and manual crack surveys. Ten different 100-ft. sections are selected on SR236, SR275, and SR67 following GDOT's current crack survey practices. For the four wheel-measured sections, the average absolute difference between automatic crack classification and quantification results and manual survey results is 3.25 out of 100, and for the six visually-estimated sections, the average absolute

difference is 5 out of 100. Both differences are within the error tolerance of GDOT's current survey practices.

Second, the automatic crack classification provides most robust classification results of low severity level cracks. The classification accuracy of load cracking and B/T cracking Severity Level 1 is constantly over 90% with the largest portion of experimental data. The load cracking and B/T cracking on interstate highways in Georgia are mostly at Severity Level 1. The automatic crack classification and quantification method is very promising to be implemented on interstate highways at this stage.

Third, the automatic crack classification and quantification, especially for quantification, is influenced by the capability of a data acquisition system and crack detection algorithm.

- Through the site-based validation, it is observed that for the sections where manual survey results are significantly different from the automatic results, most of the cracks are hairline cracks. The resolution of the 3D laser system used in this study is 1 mm in the transverse direction and 5 mm in driving direction. Its performance with hairline cracks (about 1mm wide) was less robust.
- The performance of transverse crack quantification is fair. The data acquisition interval at driving direction was 5mm to collect data at highway speed. Under this interval, some transverse cracks may be only partially detected, so their extents are underestimated through the crack quantification. With the current system, the vehicle needs to drive at less than 20km per hour to collect 3D pavement surface data at a 1mm interval.

For future research, additional experimental data should be added, especially for the high severity level cracks, to further improve the classification accuracy on these cases. Besides load cracking and B/T cracking, other crack-related distresses, such as edge cracking, reflective cracking, etc., should be incorporated, as well. The classification of these cracks may involve historical and structural data of the pavements.

For the outreach of this study, we will first test the automatic crack classification and quantification on Interstate 285 near Atlanta, and then extend the study to the entire interstate highway system in Georgia. Since interstate highways are usually challenging for a field survey due to the high traffic volume, the automatic survey results will be a good complement to GDOT's current pavement survey practices. At the next stage, these methods will be further tested and implemented on the state routes in Georgia. The method can also be extended to classify and quantify cracks for different distress protocols, e.g. FDOT's flexible pavement survey.

Beyond crack classification, the concept of multi-scale crack representation using a CFE can be extended to developed national consistent crack measures to meet the need of MAP-21. State DOTs have invested major resources to collect and maintain their legacy data over decades for pavement management and are not willing to change their distress protocol. This concept provides the opportunity to flexibly transform between their own distress protocols and national, consistent measures. In addition, 3D pavement surface data, automatic crack detection algorithms, and multi-scale crack representation from the CFE model can be used to develop next-generation pavement preservation planning tools, such as sensor-based, intelligent crack sealing planning tools.

References

Tsai, Y., Jiang C. (2012). “Crack Detection Using High-resolution 3D Line Laser Imaging Technology”, 7th RILEM International Conference on Cracking in Pavements, Delft, Netherlands, June 20, 2012.

Chapter 3 Validation of Automatic Concrete Pavement Distress Detection

1. Introduction

Pavement surface distress measurement is an essential part of a PMS for determining cost-effective maintenance and rehabilitation strategies. Visual surveys conducted by engineers in the field are still the most widely used means to inspect and evaluate pavements, although such evaluations involve high degrees of subjectivity, hazardous exposure, and low production rates. Consequently, automated distress identification is gaining wide popularity among transportation agencies.

For the past two decades, using a 2D intensity-based imaging system has been the major means most state DOTs have collected data. The intensity-based data acquisition method makes it sensitive to lighting effects. In general, because of the intensity-based data acquisition method, the performance of distress detection algorithms is severely hampered in the presence of shadows, lighting effects, non-uniform crack widths, and poor intensity contrast between cracks and surrounding pavement surfaces. The challenge persists in spite of all the research work that has been carried out to improve image acquisition techniques by minimizing the lighting defects (Kaul et al., 2010). However, it is difficult to achieve consistent crack detection under different ambient lighting conditions when using natural light for illumination (Xu, 2005). Some illumination devices, such as LED lighting, are used to provide constant lighting that prevents the impact of shadows (Xu, 2005; Xu, 2007). However, the beam width of the LED lighting is 0.5 inch, which is not thin enough to provide sufficient depth resolution. The shallow

cracks and/or thin cracks, which have low intensity contrast with surrounding pavement, are sometimes difficult to detect. Many algorithms are able to perform well only in an image data set that has images that are not too different from each other. Otherwise, manual inputs are required to adjust the input parameters so that the algorithms can perform reasonably. Although 3D stereovision has been studied recently, it is not operational. Therefore, full automation of pavement distress detection has remained a challenge, especially for accurate and reliable detection (Kaul et al., 2010).

3D laser technology includes measuring the range of an object using the projected lighting and the triangulation computation. Detailed pavement surface laser data can be collected at highway speed with adequate resolution in the x, y, and z dimensions. In recent years, this has gained great attention from researchers, industries, and transportation agencies. The 3D laser technology for pavement crack detection and rutting measurement has, for the first time, been comprehensively validated in the lab and in the field in RS-GAMS Phase 1.

This chapter is organized as follows. After introduction in Sections 1, the following four sections will present the comprehensive validation on automatic concrete distress detection using pavement surface laser data, including cracking in Section 2, faulting in Section 3, spalling in Section 4, and shoulder joint distress in Section 5. Section 6 will summarize the major findings in this study.

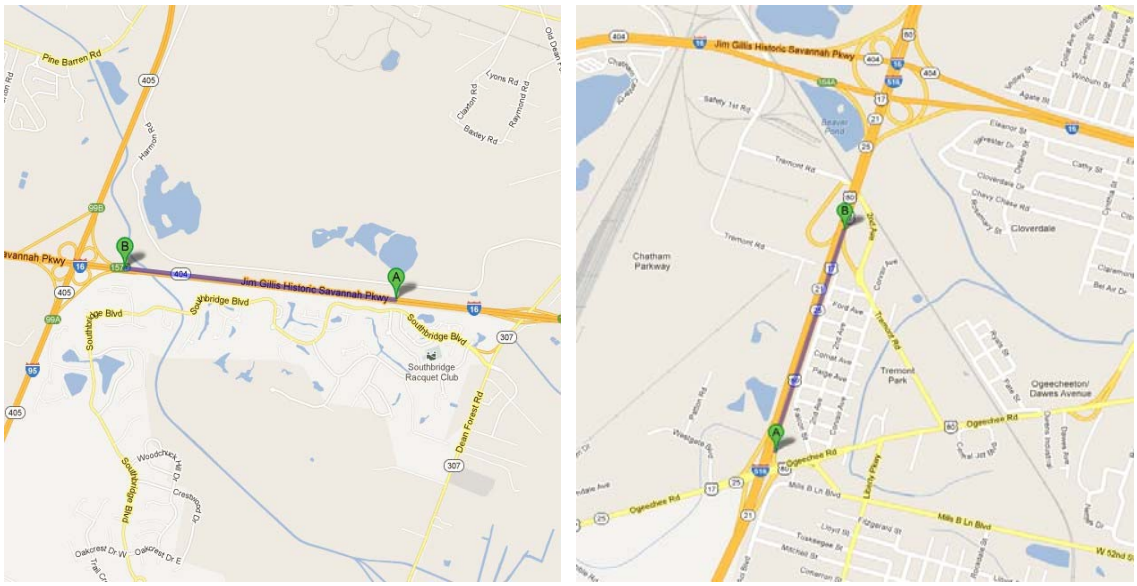
2. Validation of Automatic Concrete Pavement Crack Detection

Automatic crack detection can be conducted based on 3D laser data, and its performance on asphalt pavement has been validated in RS-GAMS Phase 1. Compared to asphalt

pavement, crack detection on concrete pavement has some unique challenges: first, transverse joints on jointed plain concrete pavement (JPCP) may impact the performance of automatic crack detection, especially for the joints with poor conditions; second, parallel grooves (transverse or longitudinal) on diamond-grooved concrete surfaces may lead to potential false positive crack detection. The section will quantitatively validate the accuracy of automatic crack detection on concrete pavement. In addition, some representative cases, including hairline cracks, misdetection as joints, and false positive detection caused by pavement damage, will be presented regarding the potential issue of concrete pavement crack detection.

2.1 Experimental data

The experimental data of the validation test were collected on interstate highways I-16 and I-516 near Savannah, Georgia. Both highways are concrete pavement on the selected test sites (as shown in FIGURE 3.1).



(a) I-16 test site (b) I-516 test site

FIGURE 3.1: Selected test sites on concrete crack detection

On the I-16 test site, we selected a total of 27 slabs westbound from MP159 to MP157; this roadway has representative cracking patterns, most of which are transverse cracking / broken slabs. Each of these slabs is marked with a unique ID for future reference (as shown in FIGURE 3.2), and the distresses on these slabs are drawn roughly for comparison purposes. On the I-516 test site, we selected 15 slabs northbound from MP3 to MP4. Similarly, each of these slabs was marked with a unique ID. Digital photos of the selected testing sites were taken from the road shoulder to track the detailed distress condition and provided certain reference through the validation.

FIGURE 3.2: Manually labeling the selected concrete slab for crack detection validation

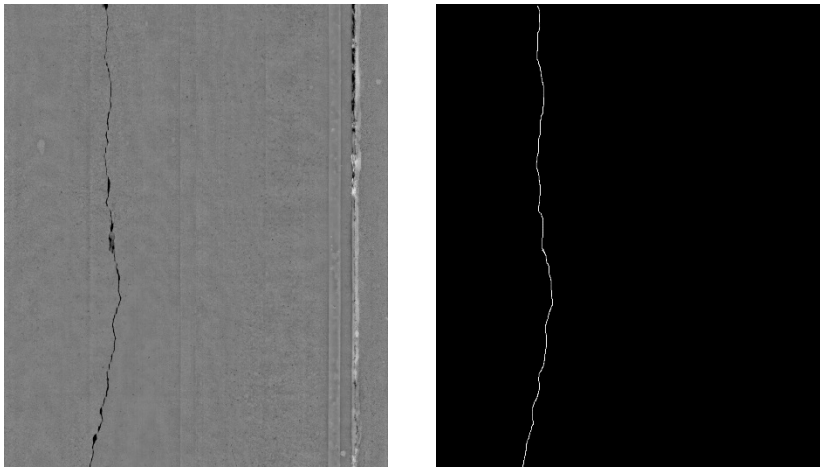
3.2 Experimental design

The following procedures are conducted to quantitatively evaluate the performance of automatic crack detection on concrete pavement:

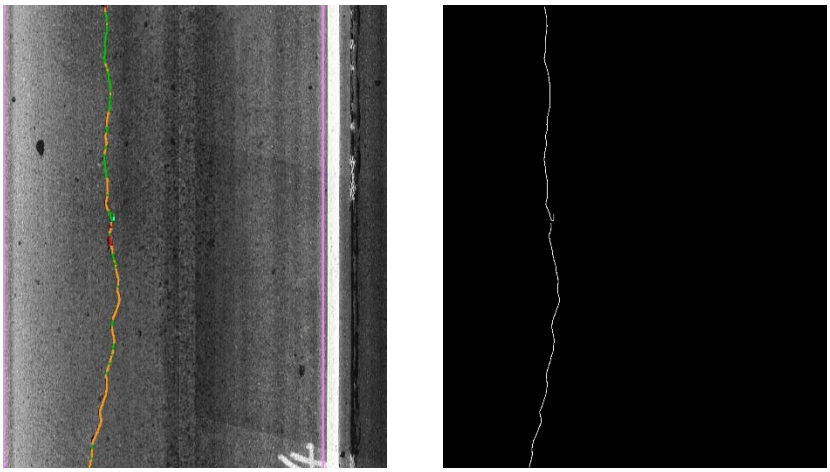
- First, the ground truth was manually digitized and extracted from the 3D laser data. The data was presented in the form of a range image (FIGURE 3.3(a)). Based on the

visual inspection of a range image, the cracking positions were manually digitized and converted to a binary ground truth crack map (FIGURE 3.3(b)).

- Then, the crack map results were generated using an automatic crack detection algorithm. The crack map can be overlaid on an intensity image or a range image (FIGURE 3.3(c)), and different colors of the detected crack line represent different crack widths. There are two typical ways to convert data into a binary crack map:
  - The first way is to manually digitize the intensity or range image with a crack map overlay (similar to the procedure used to establish ground truth).
  - The second way is to interpret the XML file as the results of distress detection and reconstruct a binary crack map (FIGURE 3.3(d)).
- Finally, the two binary crack maps were compared – ground truth and automatic detection. The buffered Hausdorff scoring method was employed to conduct an objective and quantitative evaluation.



(a) Range Image (b) Manually Digitized Ground Truth



(c) Automatic Crack Detection (d) Binary Automatic Crack Detection

FIGURE 3.3: Illustration of evaluation procedure

2.2 Introduction to buffered Hausdorff scoring method

The buffered Hausdorff scoring method was proposed in our previous study to evaluate the performance of asphalt crack detection (Kaul et. al., 2010; Tsai et. al., 2010). It incorporates the strengths of both mean square error and Hausdorff distance by modifying the Hausdorff distance metric. The Hausdorff distance is among the most popular distance measures and measures the distance between two curves; it is a metric. It has been extensively used in literature (Beauchemin et al., 1998; Wang, 2002). For any two sets of points $A = a_1, a_2, ..., a_n$ and $B = b_1, b_2, ..., b_m$,

$$H(A, B) = \max(h(A, B), h(B, A))$$

Where

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|$$

$h(A,B)$ is the greatest of all the small distances from points of A to B and is the greatest of all the small distances from points of B to A. FIGURE 3.4 illustrates this distance measurement effectively.

The value of the Hausdorff distance is large, even if one crack pixel in the segmented image is far from the ground truth image crack pixels. Seeing this limitation of the Hausdorff distance metric, a new metric was developed that does not suffer from the defects of the Hausdorff distance. The intuitive development of this measure is described next. A better distance measure than the Hausdorff distance is the modified Hausdorff distance given by $MH(A,B)$:

$$MH(A,B) = \max(h_1(A,B), h_1(B,A))$$

Where

$$h_1(A,B) = \frac{1}{m}\sum_{a\in A}\min_{b\in B}\|a-b\|$$

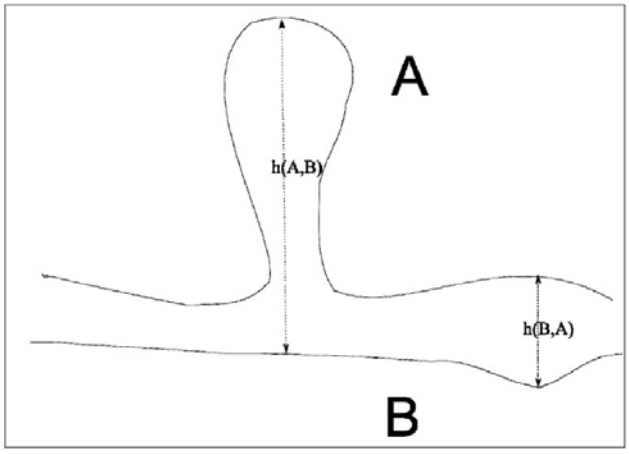


FIGURE 3.4: Illustration of Hausdorff distance

After initially using the modified Hausdorff distance measure for our image comparison, we felt that there was one more possible improvement. Once a crack pixel in the automatically segmented image falls substantially away from the closest pixel in the ground truth image, it no longer makes sense to heavily penalize this distance. Wrong detections beyond a certain distance should be penalized equally. This leads to a new distance measure, the buffered Hausdorff distance measure given by $BH(A,B)$.

$$BH(A,B) = \max(h_2(A,B), h_2(B,A))$$

Where

$$h_2(A,B) = \frac{1}{m}\sum_{a\in A} \underset{L}{sat} \min_{b\in B} \|a-b\|$$

Here, $\underset{L}{sat}$ indicates that when the distance of the crack pixel to the closest crack pixel in the other image exceeds a saturation value L, we use a constant value of L for the distance. The buffer L was chosen to be 50 in this validation, which fits the selected image resolution of 1,040 × 1,250 based on our previous sensitivity study. FIGURE 3.5 illustrates the buffered Hausdorff distance measure. The sample values of the buffered distance have a very intuitive meaning, too. The buffered distance can be interpreted as the average Euclidean distance between the crack pixels in the ground truth image and the segmented images. To compare other scoring methods with this buffered distance, a scaled scoring measure was derived as given below:

$$\text{Buffered distance score} = 100 - \frac{BH(A,B)}{L} \times 100$$

The buffered distance effectively measures the performance of the segmentation methods and generates a score that corresponds with the qualitative performance of visual inspection. Using the buffered Hausdorff scoring method, the experimental results on the I-516 and I-16 test sites are presented in the following subsections.

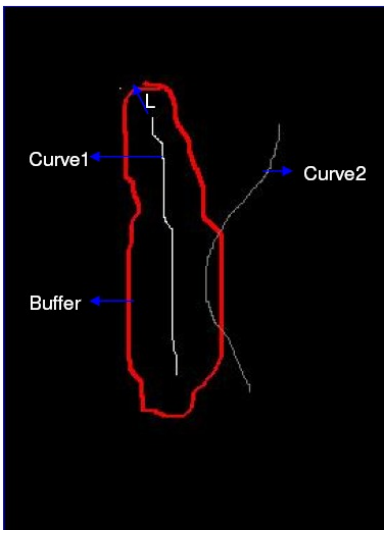


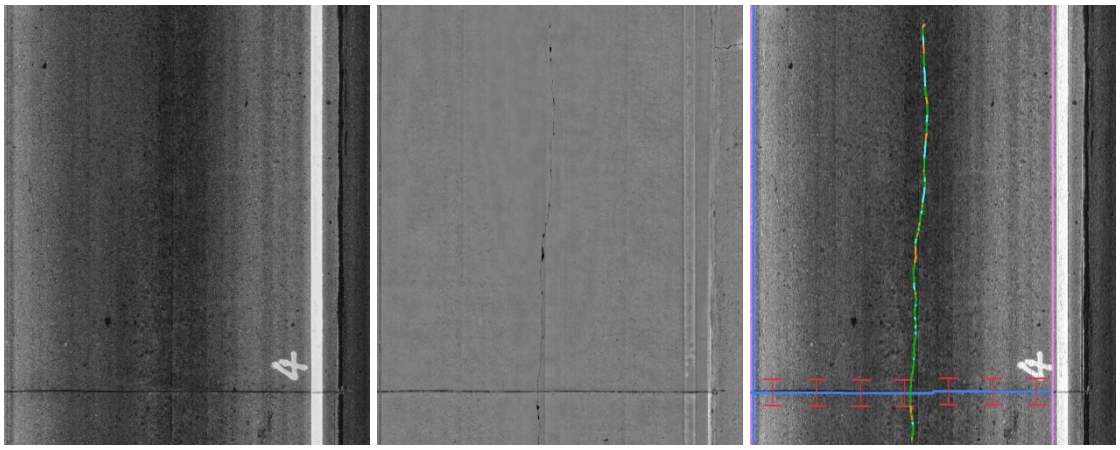
FIGURE 3.5: Illustration of Buffered Hausdorff distance measure

2.3 Validation on I-516 test site

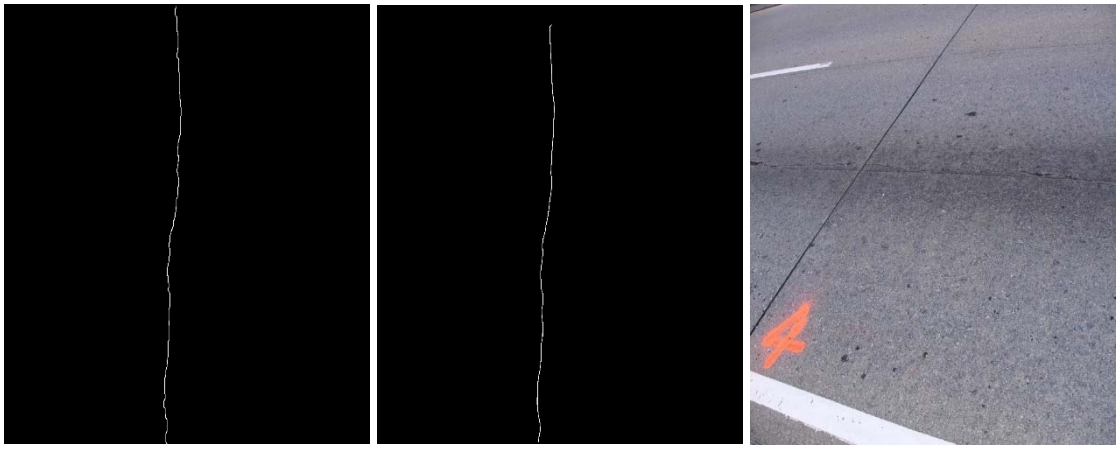
This subsection presents the experimental results on I-516. The majority of cracks on this test site are longitudinal cracking. 15 slabs with longitudinal cracking were selected. Since the normal length of the slab is larger than the collected pavement image in the driving direction, some slabs were divided into two images in the experimental test, where the consecutive image is represented with the same ID and a single quote (e.g. Slab #1').

FIGURE 3.6 and FIGURE 3.7 are two examples demonstrating the performance of automatic crack detection on this test site. FIGURE 3.6 shows longitudinal cracking on a concrete pavement with normal crack width and no spalling. FIGURE 3.6 (a) and (b)

shows the intensity and range images collected from the 3D laser system. FIGURE 3.6 (c) shows automatic crack detection results; it can be observed that, besides the cracks, the longitudinal and transverse joints for this JPCP are also extracted, which are labeled as straight blue lines. FIGURE 3.6 (d) shows the ground truth, which is manually digitized from the range image, and FIGURE 3.6 (e) shows the binary crack map generated from the automatic detection results. FIGURE 3.6 (f) is a digital photo taken from the road shoulder on this specific slab. Based on visual evaluation, the crack detection results are very close to the manually digitized ground truth. The buffered Hausdorff scoring method also gives a high score of 96.59 on this image, which indicates an accurate detection. Also, the transverse and longitudinal joints are successfully differentiated from the cracks. Similarly, FIGURE 3.7 shows longitudinal cracking with a large crack width and some spalling. The automatic crack detection shows an accurate outcome, as well, in this case.

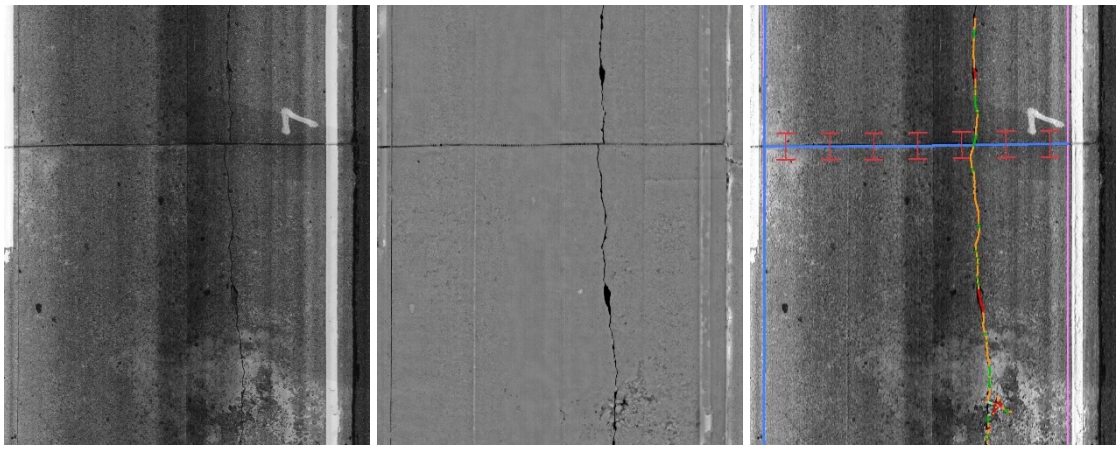


(a) Intensity Image (b) Range Image (c) Automatic Crack Detection

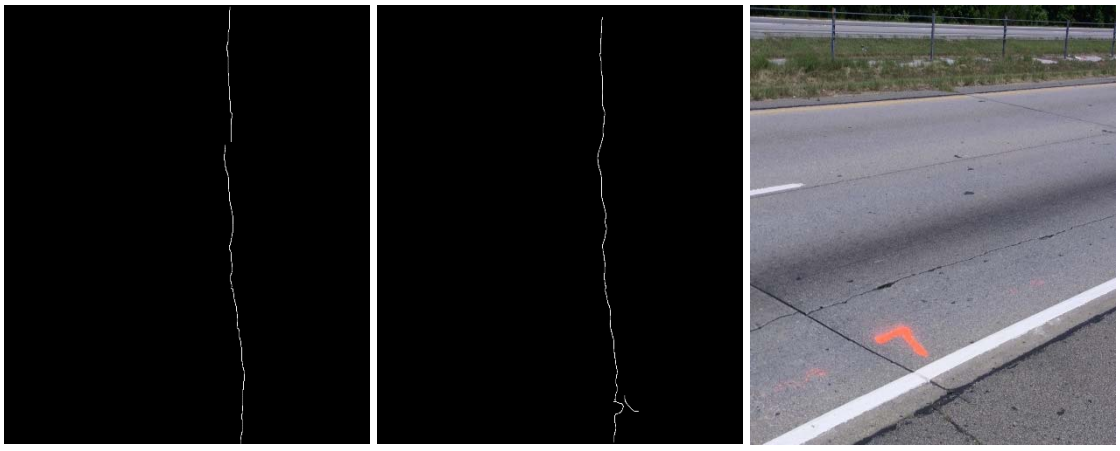


(d) Ground Truth I Binary Detection Results (f) Field Photo

FIGURE 3.6: Crack detection on Slab #4 on I-516 test site (Score: 96.5859)



(a) Intensity Image (b) Range Image (c) Automatic Crack Detection



(d) Ground Truth (e) Binary Detection Results (f) Field Photo

FIGURE 3.7: Crack detection on Slab #7 on I-516 test site (Score: 93.4601)

The overall performance on this test site is summarized in Table 3.1. The automatic crack detection using pavement surface laser data shows accurate and robust detection results on most of the images, having an overall average score of 85.89. The cases with poor performance were mainly caused by misclassification between cracks and construction joints, which will be further explained in the following subsections.

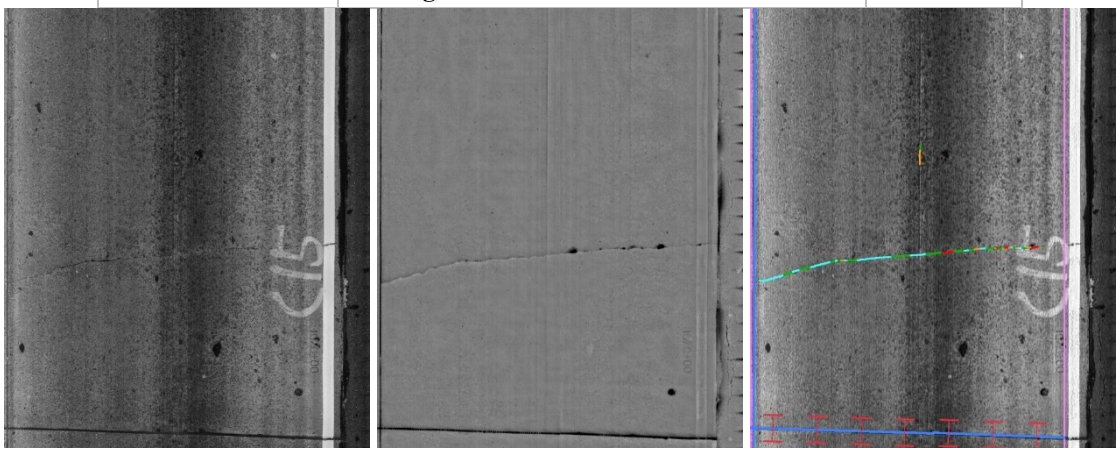
2.4 Validation on I-16 test site

This subsection presents the experimental results on I-16. The majority of the cracks on this test site are transverse cracking / broken slab. A total of 27 slabs were selected. Similarly, some slabs are divided into two images through the data collection, where the consecutive image is represented with the same ID and a single quote (e.g. Slab #C1').

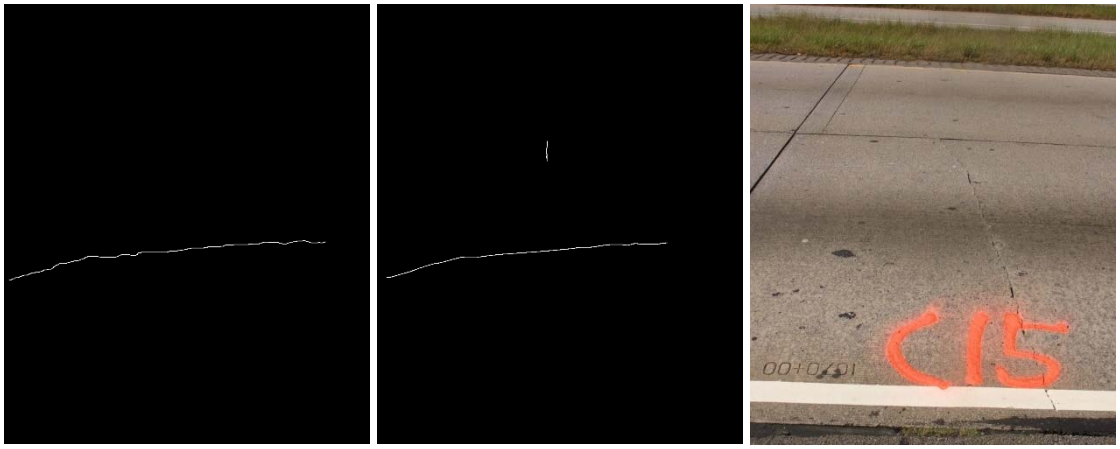
FIGURE 3.8 and FIGURE 3.9 are two examples that demonstrate the performance of automatic crack detection on this test site. FIGURE 3.8 shows transverse cracking on concrete pavement with normal crack width and slight spalling. Based on visual evaluation, the crack detection results are close to the manually digitized ground truth. The buffered Hausdorff scoring method also gives a high score of 90.45 on this image, which indicates accurate detection. Similarly, FIGURE 3.9 is a broken slab with a large crack width and severe spalling. This introduces some uncertainty into the crack detection, since some spalling is large enough to be detected separately as spalling (as shown in the blue area on the image). Based on visual inspection, the overall pattern of detected cracks is similar to the one in the ground truth, although some false positives and false negatives are observed. The buffered Hausdorff score on this image is 85.41.

Table 3.1: Performance Evaluation of Crack Detection on I-516 Test Site

Slab ID	Description	Score
1	Longitudinal cracking, wide with spalling	97.4526
1'	Longitudinal cracking, normal width	95.167
2	Longitudinal cracking, normal width	97.595
3	Longitudinal cracking, wide with spalling	98.0972
3'	Longitudinal cracking, wide with spalling	98.6694
4	Longitudinal cracking, normal width	96.5859
4'	Longitudinal cracking, normal width	81.3702
5	Longitudinal cracking, normal width	93.2204
6	Longitudinal cracking, normal width	89.6559
7	Longitudinal cracking, wide with spalling	93.4601
7'	Longitudinal cracking, wide with spalling	41.8065
8	Longitudinal cracking, wide with spalling	86.8667
8'	Longitudinal cracking, wide with spalling	46.8266
9	Longitudinal cracking, wide with spalling	90.9242
9'	Longitudinal cracking, wide with spalling	97.2637
10	Longitudinal cracking, wide with spalling	96.8542
11	Longitudinal cracking, normal width	78.0038
11'	Longitudinal and transverse cracking, normal width	63.1563
12	Longitudinal cracking, wide with spalling	87.9587
13	Longitudinal cracking, wide with spalling	54.7187
14	Longitudinal cracking, wide with spalling	96.9098
15	Longitudinal cracking, wide with spalling	95.0416
15'	Longitudinal cracking, normal width	97.8843
AVG	**Average Score on the I-516 site**	**85.8908**

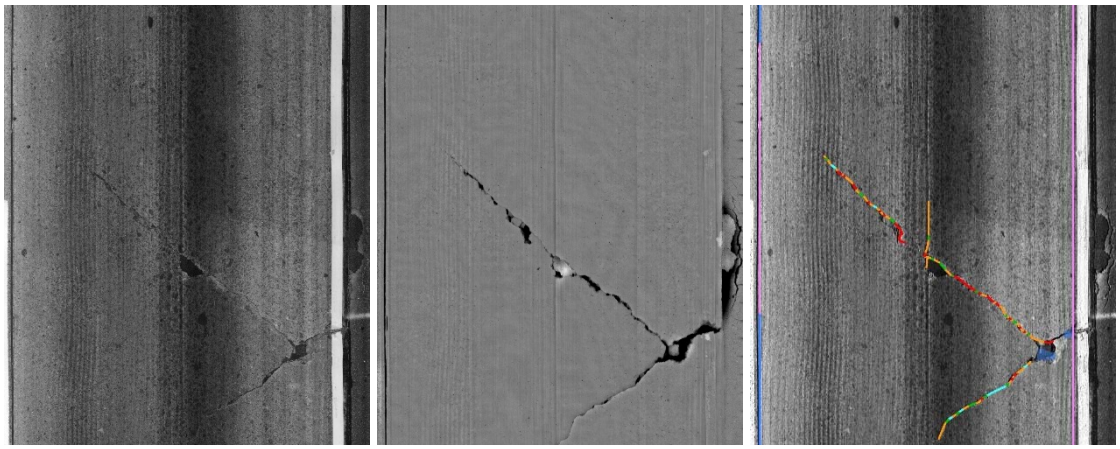


(a) Intensity Image (b) Range Image (c) Automatic Crack Detection



(d) Ground Truth (e) Binary Detection Results (f) Field Photo

FIGURE 3.8: Crack detection on Slab #C15 on I-16 test site (Score: 90.4512)



(a) Intensity Image (b) Range Image (c) Automatic Crack Detection

(d) Ground Truth (e) Binary Detection Results (f) Field Photo

FIGURE 3.9: Crack detection on slab #C1' on I-16 test site (Score: 85.4087)

The overall performance on this test site is summarized in Table 3.2. The performance of automatic crack detection on this site is not as good as on the I-516 test site, having an overall average score of 64.95. This is probably due to three reasons: 1) compared to longitudinal cracking, transverse cracking has a higher probability of being misclassified as construction joints; 2) many hairline cracks are observed on this site, which are difficult for the current system to detect due to the resolution limit; and 3) many severe spallings along the cracks interfere with the crack detection performance. More details on these failed / poor-performance cases are presented in the following subsection.

2.5 Potential issues of concrete pavement crack detection

Through the performance evaluation of automatic crack detection on concrete pavement, several major issues were identified, including hairline cracks, misdetection as joints, and false positives caused by pavement damage. Each of these issues is discussed below with a representative example to analyze the potential cause and future solution.

Table 3.2: Performance Evaluation of Crack Detection on I-16 Test Site

Slab ID	Description	Score
C1	Transverse cracking, normal width	75.3972
C1'	Broken slab, wide with spalling	85.4087
C2	Transverse cracking, normal width	59.8938
C3	Transverse cracking, normal width	84.6954
C4	Broken slab, wide with spalling	54.3036
C8	Broken slab, wide with spalling	47.1257
C10	Transverse cracking, normal width	89.279
C12	Transverse cracking, hairline	46.8648
C13	Replaced slab, hairline cracking	39.5158
C14	Broken slab, wide with spalling	50.5526
C15	Transverse cracking, normal width	90.4512
C16	Transverse cracking, normal width	47.7935
C17	Transverse cracking, hairline	45.1996

C18	Transverse cracking, normal width	92.6896
C20	Broken slab, wide with spalling	54.871
C22	Replaced slab, hairline cracking	59.9153
C24	Replaced slab, longitudinal cracking	37.121
C26	Longitudinal cracking, wide with spalling	94.0759
C27	Longitudinal cracking, wide with spalling	78.9262
AVG	**Average Score on the I-16 site**	**64.9516**

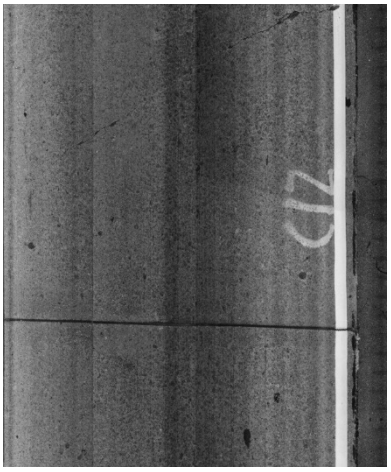
Hairline cracks

Similar to asphalt pavement, hairline cracks on concrete pavement present a major challenge for crack detection. The data acquisition interval of the current 3D line laser imaging system is 1 mm in the transverse direction and 5 mm in the driving direction. In this experimental test, it wasn't practical to manually measure the crack width at the test site due to the heavy traffic on the interstate highway. Based on our previous experience on asphalt pavement, cracks with widths below 2 mm are difficult to detect. FIGURE 3.10 shows a representative case of hairline cracking. The middle part of the transverse crack on the image is hairline. It can be seen that that portion is too thin to be observed, even from the high-resolution digital photos taken from the shoulder. The automatic detection can only extract part of the entire crack line in this case, which leads to a poor buffered Hausdorff score. The hairline cracking issue is better observed on the I-16 test site, resulting in an overall, relatively low score on that site; also, the larger data acquisition interval in the driving direction makes it more difficult to capture transverse hairline cracking.

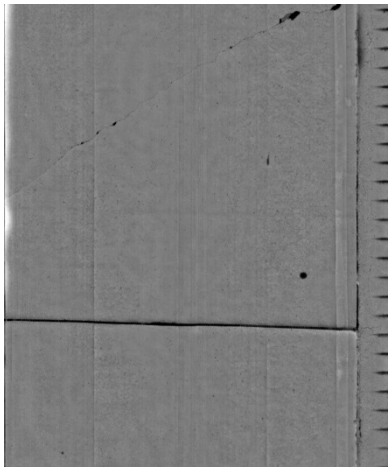
Misclassification as Joints

For concrete pavement, especially for joint plain concrete pavement, automatic crack detection needs to be conducted together with joint detection. Concrete joints have very similar appearance and characteristics with cracks; without a separate joint detection, most of the joints will be captured as concrete cracking, as well. Although the joint detection algorithm performs relatively accurately (presented in the faulting measurement validation section), conducting joint detection simultaneously with crack detection introduces the potential risk that some cracks may be misclassified as joints at the first stage.

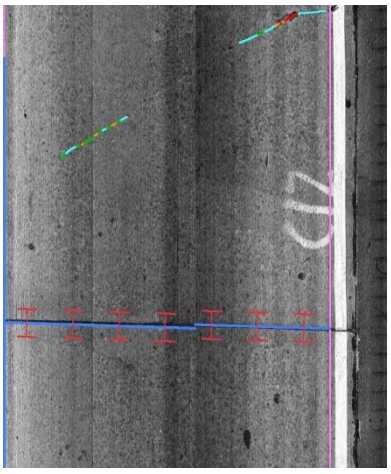
Through the experimental tests, several such cases are observed for both transverse cracking and longitudinal cracking. FIGURE 3.11 shows an example of transverse cracking; the left side of the crack is detected successfully, while the right side is misclassified as a joint. FIGURE 3.12 shows an example of longitudinal cracking; the top side of the crack is detected successfully, while the bottom side is misclassified as a joint. Based on the experimental results, these false detection cases usually happen under the following two situations: 1) cracks appear to be approximate straight lines and 2) some spalling occurs along the crack and interfere with the crack detection.



(a) Intensity Image



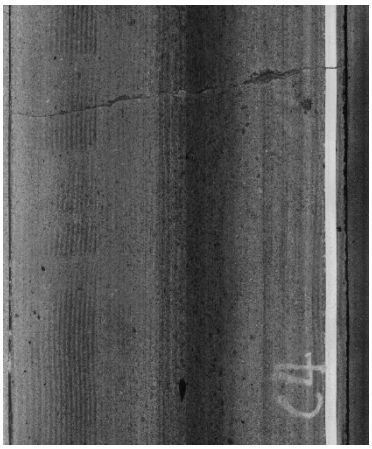
(b) Range Image

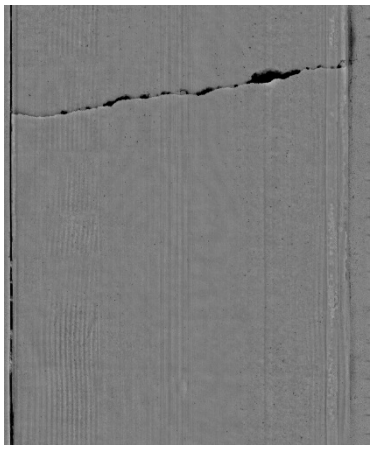


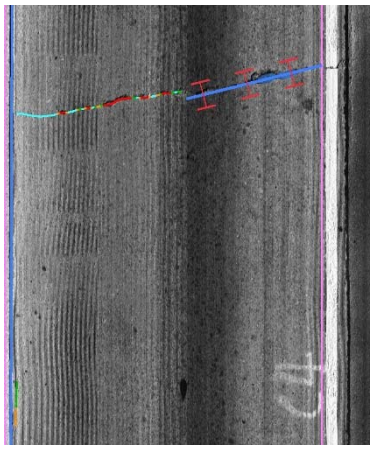
(c) Automatic Crack Detection

(d) Field Photo

FIGURE 3.10: Crack detection on hairline cracking

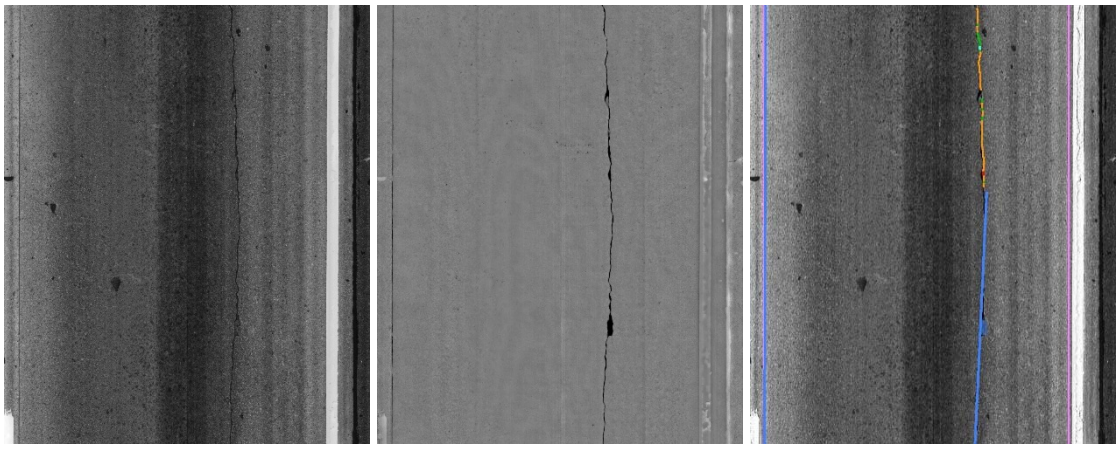






(a) Intensity Image (b) Range Image (c) Automatic Detection Results

FIGURE 3.11: Misclassification of transverse cracking as concrete joint



(a) Intensity Image (b) Range Image (c) Automatic Detection Results

FIGURE 3.12: Misclassification of longitudinal cracking as concrete joint

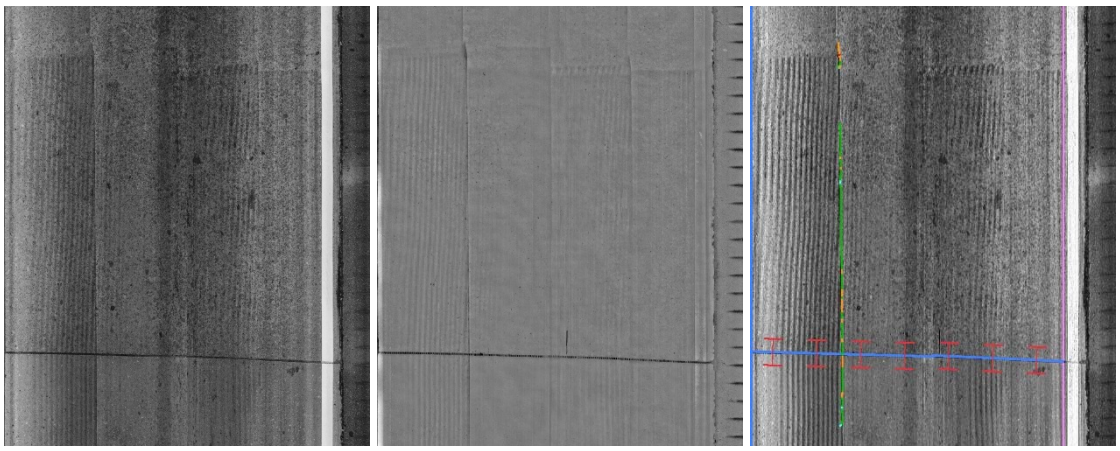
To remove these false detections, the following additional criteria can be considered as a follow-up step after crack and joint detection. First, the location of detected joints is evaluated, especially for longitudinal joints. If a longitudinal joint is located in the middle of the lane, there is a high probability that it is a false joint detection. Second, the connectivity between joints and neighboring cracks is evaluated. If a joint is directly connected to cracks, either longitudinal or transverse, it may be a false joint detection, as well. These further modifications will improve the crack and joint detection accuracy in future implementation.

False Positive Detection

Similar to asphalt pavement, false positive detections still exist in concrete pavement. This is mainly due to the data acquisition mechanism of the 3D line laser imaging system. Since pavement surface laser data is purely based on the elevation of pavement surface, any pavement damage that result in an elevation change will lead to potential risk on crack detection. One example is shown in FIGURE 3.13. We can observe a clear, straight

dent along the driving direction on the range image (FIGURE 3.13(b)), whose appearance is quite similar to cracks; automatic crack detection also extracts those out as cracking (FIGURE 3.13(c)). However, this dent is most likely to have been caused by a flat tire of heavy trucks**,** which is a false positive detection in this case.

These false positives are the issue that we have to face under the 3D data acquisition mechanism. To further remove these, the characteristics of these false detections need to be studied and classified, and then a specific filter can then be designed to remove them either at the data preprocessing stage or during the post processing of the detection results.



(a) Intensity Image (b) Range Image (c) Automatic Detection Results

FIGURE 3.13: Example of false positive detection

2.6 Summary

For automatic concrete pavement crack detection using a 3D laser data, the major findings are summarized as follows:

First, the 3D laser system shows overall acceptable performance for automatic crack detection on concrete pavement. Two test sites were selected on interstate highways I-

516 and I-16 to quantitative evaluate the crack detection performance. Automatic crack detection results were compared with the manually digitized ground truth using the buffered Hausdorff scoring method. The automatic crack detection shows quite accurate and robust results on the I-516 test site, which mainly consists of longitudinal cracking. The crack detection performance on I-16 test site was not as good; this is mainly due to the fair amount of hairline cracking and interference due to severe spalling and joint detection. Through the discussion of false detection cases in this chapter, there is still the potential to further improve the performance of these cases. In summary, from the crack evaluation perspective, the emerging 3D laser system demonstrates its capability to support further automatic concrete pavement condition evaluation.

Second, hairline cracks are still quite challenging for automatic detection. The data acquisition interval of the current 3D laser (1 mm in the transverse direction and 5 mm in the driving direction) makes it difficult to detect hairline cracks thinner than 2 mm. Also, the fact that there are more false negatives detected in transverse cracking can be explained by the large data acquisition interval in the driving direction.

Third, as on asphalt pavement, pavement damage caused by flat tires or trucks still has an impact on crack detection on concrete pavement. A dent on pavement has a very similar appearance as cracking and may lead to a false positive detection. The characteristics of these pavement damages need to be studied and classified, and a specific filter can then be designed to remove them, either at data preprocessing stage or during the post processing of the detection results.

3. Validation of Concrete Pavement Faulting Measurement

Faulting is the differential vertical displacement of the slab edge across a transverse joint caused by inadequate load transfer, differential deflection at the joint, inadequate base support, or sub-base erosion (Jung et al., 2008). The difference in elevation affects vehicle ride quality, accelerates vehicle damage, and leads to distresses, such as corner breaks and blowups; thus, faulting has a major effect on vehicle operation costs and pavement life-cycle costs (FHWA, 2006). Faulting is an important performance indicator for jointed concrete pavements and the criteria for pavement restoration decisions.

Faulting has traditionally been collected by manual methods. The Georgia Faultmeter, designed by GDOT, is one of the most popular hand-held devices for performing faulting measurement and is used by many state highway agencies, including GDOT (2004) and the Minnesota Department of Transportation (MnDOT) (Burnham, 2003). In this method, a surveyor finds a gap in traffic and sets the faultmeter at a single spot along a designated joint to measure faulting. GDOT requires the meter to be set approximately 15 cm (6 in) from the pavement edge marking in the outside lane. The surveyor pushes the button to acquire a faulting measurement and records it manually. This manual operation is labor-intensive, time-consuming, costly, and dangerous to workers and drivers. It also limits the faulting measurement to only the sampled joints, not all of the joints. GDOT conducts faulting measurement on every eighth joint (GDOT, 2004). Therefore, alternative methods for effectively and safely collecting faulting data are much needed. In addition, state highway agencies are now required to collect faulting data under the new Highway Performance Monitoring System (HPMS) reassessment (OHIP & FHWA 2008). This strongly motivates state DOTs to look for cost-effective means to collect faulting data.

Recently, some state highway agencies, such as the Florida Department of Transportation (FDOT) and the Mississippi Department of Transportation (MDOT), have collected faulting data using single laser profile data that were collected for the IRI (MaGhee, 2004; Nazef et al., 2009). With the single profile data, it is important to have an adequate sampling interval to locate the joint. Nazef et al. reported that a 1.7 cm (0.68 in) sampling interval yields a 95% joint detection rate using the algorithm developed by the FDOT. The chance of missing joints increases as the sampling interval increases.

With the advances of sensing technology, the 3D laser technique creates the potential to use not one single profile but the entire pavement surface to measure the faulting across concrete joints. FIGURE 3.14 illustrates the basic principle of capturing line laser data along the joints while the vehicle is driving at highway speed, and FIGURE 3.15 gives an example of the collected pavement surface laser data at the location of a concrete joint. It can be observed that the elevation change of the pavement surface is clearly captured by the laser data, which makes it possible to develop automatic faulting measurement algorithms.

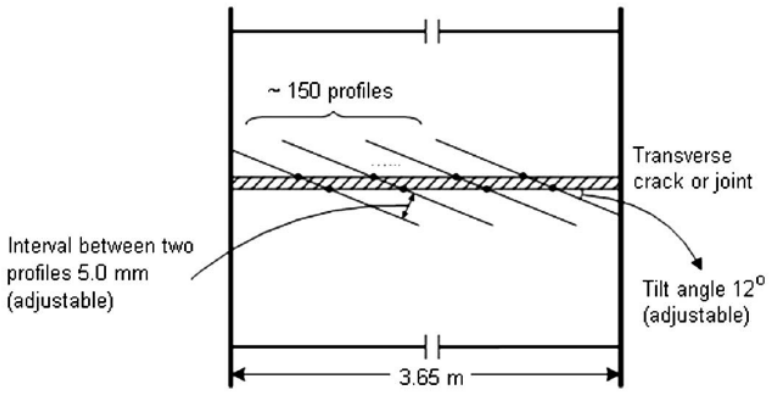
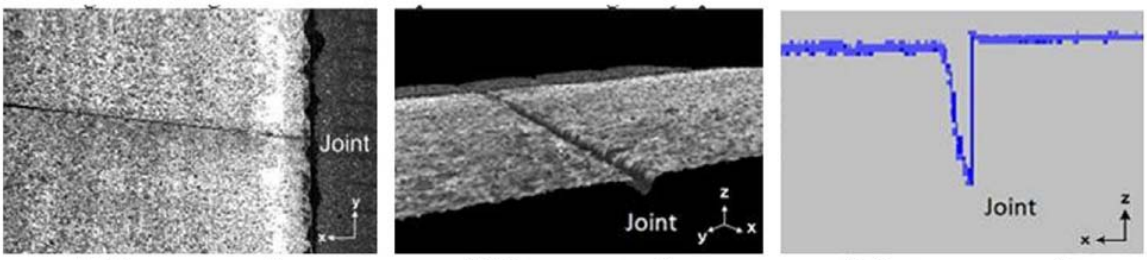


FIGURE 3.14: Illustration of the alignment of the 3D continuous laser profiles



(a) 2D pavement surface (b) 3D pavement surface (c) One transverse profile

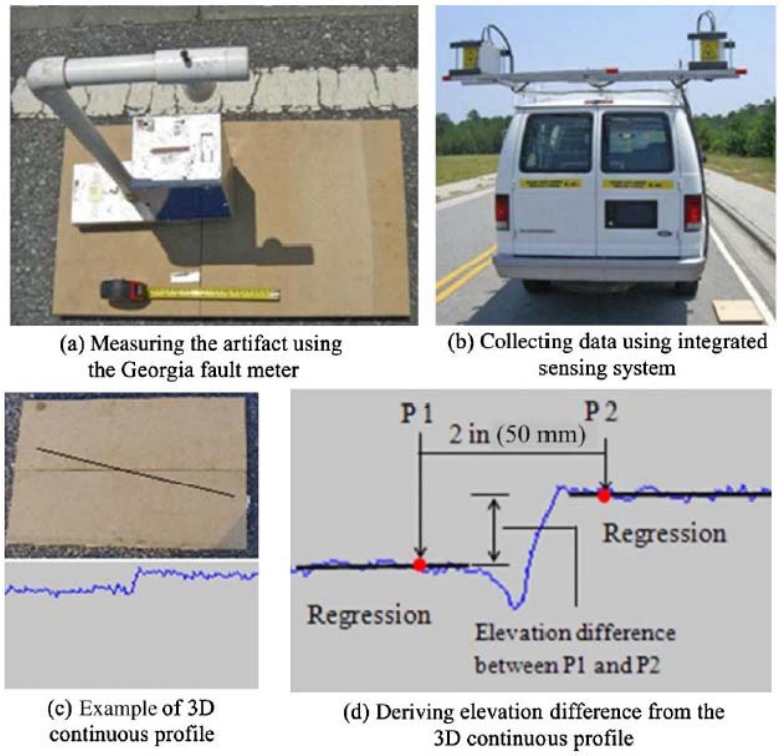
FIGURE 3.15: Example of pavement surface laser data for faulting measurement

To validate the capability of using the pavement surface laser data on faulting measurement and to examine the performance of corresponding automatic faulting measuring methods, a comprehensive validation test was conducted in our research. Using the same pavement surface laser data, two different methods are evaluated in this section, including a regression-based faulting measurement (Tsai, et al., 2012) and the concrete joint module provided in the commercial software. The test results, analysis, and major findings are presented in this report to provide an overall performance evaluation and, also, to make suggestions for future improvement.

3.1 Validation using the regression-based faulting detection method

For the regression-based method, a controlled in-lab experimental test was first conducted to examine the detection accuracy of the method; then, a field test was conducted on I-16 to evaluate the accuracy and the feasibility of operating the integrated sensing system at highway speeds.

Controlled In-lab Experimental Result and Analysis



(a) Measuring the artifact using the Georgia fault meter

(b) Collecting data using integrated sensing system

(c) Example of 3D continuous profile

(d) Deriving elevation difference from the 3D continuous profile

FIGURE 3.16: Controlled test

In-lab tests were designed to test the regression-based method with the faulting in different ranges in a well-controlled environment, as it is difficult to locate a testing section of an appropriate length that can cover a full range of faulting. The artifacts, made of two wood panels creating two flat surfaces with known elevation differences, were used to create the elevation differences ranging from 0.8 to 15 mm (1/32 to 19/32 in). The purpose was to mimic the range of faulting measured by the Georgia Faultmeter. The artifacts were set level on a fairly flat road on the Georgia Tech campus to ensure a consistent elevation difference between any two points on the two panels. The known elevation differences were also confirmed using the Georgia Faultmeter on the test site,

as shown in FIGURE 3.16. The integrated sensing system was then used to collect the pavement surface laser data with a 1mm (0.04 in) resolution in the transverse direction and a 5mm (0.2 in) interval between two profiles in the travel direction at low speed, as shown in FIGURE 3.16(b). With a 12° tilt angle, approximately 15 profiles were collected along the 38 cm (15 in) wide wood panel. FIGURE 3.16(c) shows one of the 3D continuous pavement profiles that can be used to derive the elevation difference. The elevations of the two flat surfaces can be established by applying regression to the points representing the surface, as shown in FIGURE 3.16(d). Following the footprint of the Georgia Faultmeter, the elevations of two measurement points (P1 and P2 in FIGURE 3.16(d)), separated by 50 mm (2 in.), were estimated using the regression lines, and then the elevation difference could be calculated.

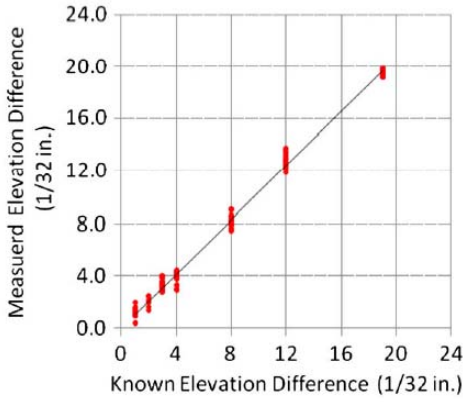


FIGURE 3.17: Known & derived elevation differences

The accuracy of the derived elevation differences was then evaluated by comparing the derived and the known elevation differences. FIGURE 3.17 shows the derived elevation differences were close to the known elevation differences with a small variation. The comparison results in Table 3.3 show that the average absolute differences are within 1

mm (0.04 in), and the variances among different profiles were fairly small (less than 1 mm) and consistent across various elevation differences. The regression-based method slightly overestimated the elevation differences. The 0.8 mm (1/32 in) elevation difference may not be detected reliably because of the 0.5 mm resolution in the vertical direction. The results of the controlled field tests have demonstrated that the regression-based method can achieve an accuracy of less than 1 mm with small variances among different profiles.

Table 3.3: Summary Result of Elevation Statistics

Known elevation difference (in.)	Known elevation difference (mm)	Sample size	Mean (mm)	Average difference (mm)	Average absolute difference (mm)	Standard deviation (mm)	Coefficient of variance
1/32	0.8	20	1.0	0.2	0.2	0.3	0.27
2/32	1.6	6	1.6	0.0	0.3	0.3	0.21
3/32	2.4	19	2.7	0.3	0.3	0.3	0.12
4/32	3.2	13	3.1	−0.1	0.3	0.4	0.13
8/32	6.4	15	6.5	0.1	0.3	0.4	0.06
12/32	9.5	20	10.1	0.6	0.6	0.4	0.04
19/32	15.1	20	15.5	0.4	0.4	0.2	0.01

Field Experimental Result and Analysis

Field tests were also conducted on I-16 to evaluate the accuracy and repeatability of the regression-based method and the feasibility of operating the integrated sensing system at highway speeds. A 450-ft test section covering 15 joints on eastbound I-16 between milepost 154 and 155 was selected because of the sampled faulting reported on the section by GDOT's engineers. The slabs are 9 m (30 ft) long and 3.65 m (12 ft) wide. The 15 slabs were first labeled and marked with a sequential number and a point where the faulting was measured on the basis of GDOT's faulting measurement practice (GDOT, 2004). The marked point was approximately 15 cm (6 in) from the pavement edge marking. The faulting was measured three times at each joint to establish the ground

truth. The integrated sensing system was then used to collect faulting data at two different highway speeds, 100 and 80 km/h (62.5 and 50 mph). The joint can be captured by the profile with 1 mm (0.04 in) resolution in the transverse direction. Three runs were repeated at each speed to evaluate the repeatability of derived faulting measurements and the feasibility of operating the integrated sensing system at highway speeds. The 3D continuous pavement profile data at the marked point where the faulting was measured by the Georgia Faultmeter was identified manually, and faulting was derived on the basis of the selected profile data using the same method described in the controlled field test section.

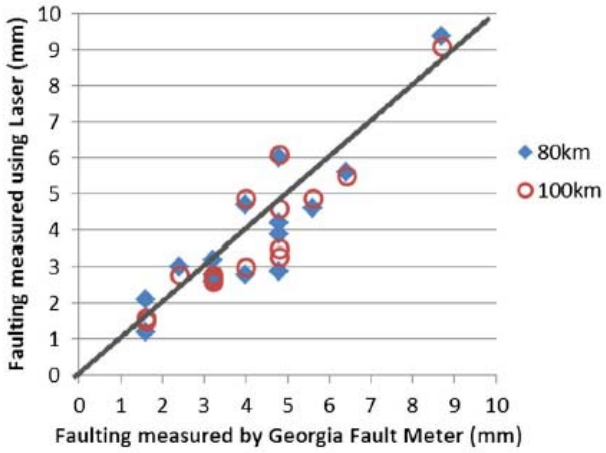


FIGURE 3.18: Measured vs. derived faulting measurements

Faulting measurements of the 15 joints derived from the 3D continuous pavement profile data were compared with those measured using the Georgia Faultmeter to evaluate the accuracy and repeatability of the regression-based method at highway speed. FIGURE 3.18 shows that the derived faulting measurements are fairly close to the ones measured using the Georgia Faultmeter, having a maximum difference of less than 2 mm. Table 3.4 summarizes the derived faulting measurements on the 15 joints collected on I-16 at

different speeds. The derived faulting measurements range from 1.2 to 9.4 mm (0.05 to 0.37 in), and 13 out of 15 joints have a difference of less than 1 mm; the largest difference is 1.9 mm.

Table 3.4: Statistics of Derived Faulting Measurement on 15 Slabs by Speeds

Slab	Georgia fault meter (mm)	80 km/h Mean (mm)	80 km/h Difference (mm)	80 km/h Standard deviation (mm)	100 km/h Mean (mm)	100 km/h Difference (mm)	100 km/h Standard deviation (mm)	80–100 Mean difference (mm)
1	4.8	3.9	–0.9	0.2	4.6	–0.2	0.3	–0.6
2	1.6	1.2	–0.4	0.3	1.6	0.0	0.2	–0.5
3	1.6	2.1	0.5	0.5	1.5	–0.1	0.3	0.6
4	4.8	6.0	1.2	0.3	6.1	1.3	0.9	–0.1
5	4.8	4.2	–0.6	0.1	3.3	–1.5	0.1	0.9
6	5.6	4.6	–1.0	0.1	4.9	–0.7	0.3	–0.3
7	6.4	5.6	–0.8	0.4	5.5	–0.9	0.1	0.1
8	3.2	2.6	–0.6	0.6	2.7	–0.5	0.2	–0.1
9	3.2	2.8	–0.4	0.8	2.8	–0.4	0.4	0.0
10	3.2	3.2	0.0	0.4	2.6	–0.6	0.3	0.5
11	4.0	4.7	0.7	0.4	4.9	0.9	0.3	–0.2
12	4.0	2.8	–1.2	0.6	3.0	–1.0	0.3	–0.2
13	2.4	3.0	0.6	0.9	2.8	0.4	0.1	0.1
14	4.8	2.9	–1.9	0.6	3.5	–1.3	0.8	–0.5
15	8.7	9.4	0.7	0.5	9.1	0.4	0.1	0.3

The repeatability of the derived faulting measurements collected in three runs at each speed was evaluated next. FIGURE 3.19 and FIGURE 3.20 show the derived faulting measurements from three runs at 80 and 100 km/h, respectively. As shown in FIGURE 3.19 and FIGURE 3.20, there is no significant difference among the different runs. The standard deviations are within 1 mm, as shown in Table 3.4. The maximum differences among three runs were also reviewed. For the data collected at 100 km/h, 13 out of 15 joints (87%) have a maximum difference of less than 1 mm (0.04 in.). Results indicate that the derived faulting measurements can achieve a desirable repeatability among different runs at the same speed. Finally, the derived faulting measurements were compared at different speeds. FIGURE 3.21 shows the derived faulting measurements are close at different speeds, and no significant difference or bias can be observed. The differences are within 1 mm (0.04 in), as shown in Table 3.4. On the basis of these

analyses, the proposed method can achieve a desirable repeatability among different runs and at different speeds, and it is feasible to operate the integrated sensing system at highway speed (e.g., 100 km/h) for collecting faulting data.

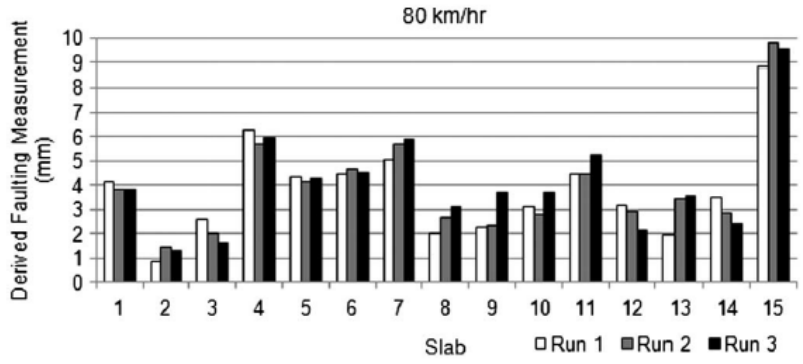


FIGURE 3.19: Derived faulting measurements at 80 km/h

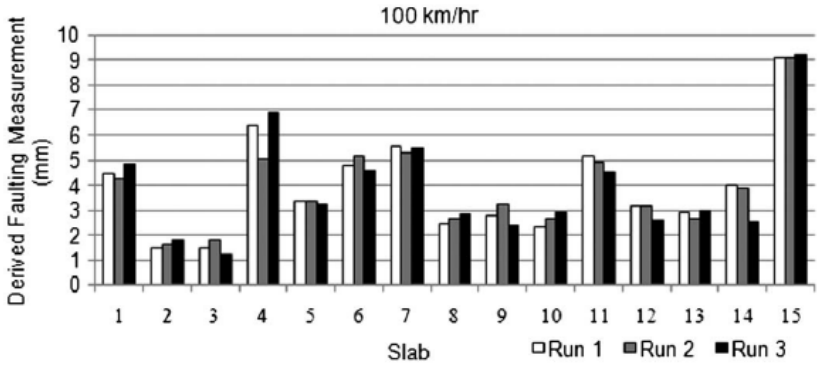


FIGURE 3.20: **Derived faulting measurements at 100 km/h**

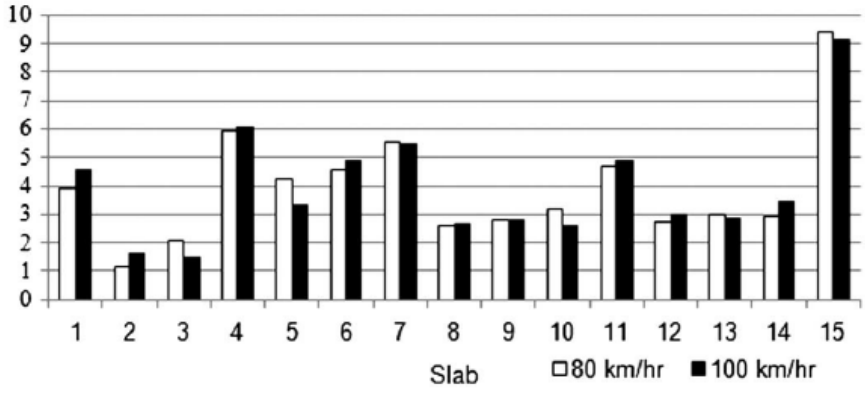


FIGURE 3.21: Comparison of faulting measurements at different speeds

On the basis of lab tests and field tests conducted on I-16, the preliminary results have demonstrated that it is feasible to collect faulting data with desirable accuracy and repeatability when using the 3D continuous profiles acquired by the integrated sensing system at highway speeds.

3.2 Validation using the concrete joint module in the commercial software

Transverse Joint Detection Module Test

To test the accuracy of the software for detecting transverse joints, we conducted a validation test on a dataset of 941 images. An example of the expected detection result is shown as FIGURE 3.22.

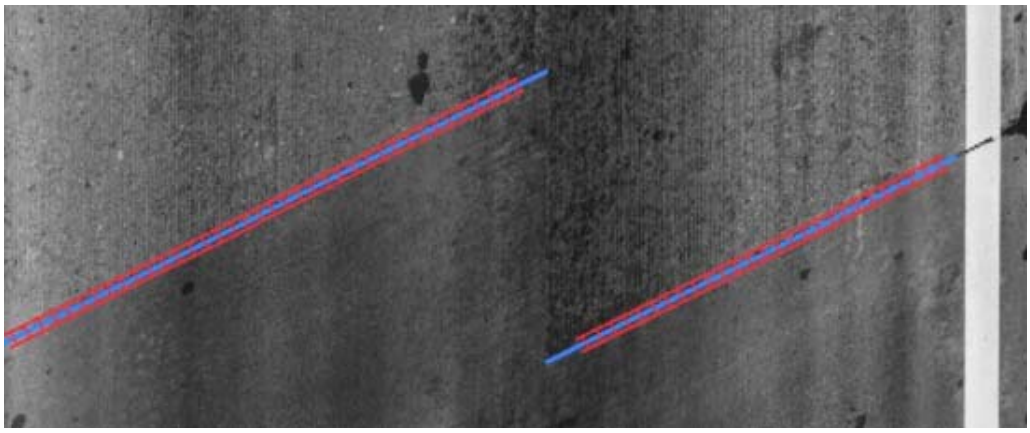


FIGURE 3.22: Example of correctly detected cases

The blue lines in the image indicate the location of the joint, while the red bars indicate the measuring points for faulting; the information will be utilized in the next section. It can be seen that for the expected outcome, the correct detection needs to be on both sides of the road, and the length of the detected lines (the blue ones) should match the actual transverse joint length.

Within these images, 5 images with transverse joints were not detected correctly (Type II error); in 4 images, other distresses, such as cracks, are mistakenly detected as joints (Type I error). Considering both Type I and Type II errors, the joint detection module still achieved an accuracy of 99.5%.

To find the reasons for the incorrect detection, we further examined the error cases. Four false-positive cases were big transverse cracks sharing the typical characteristics of joints (shown in FIGURE 3.23):

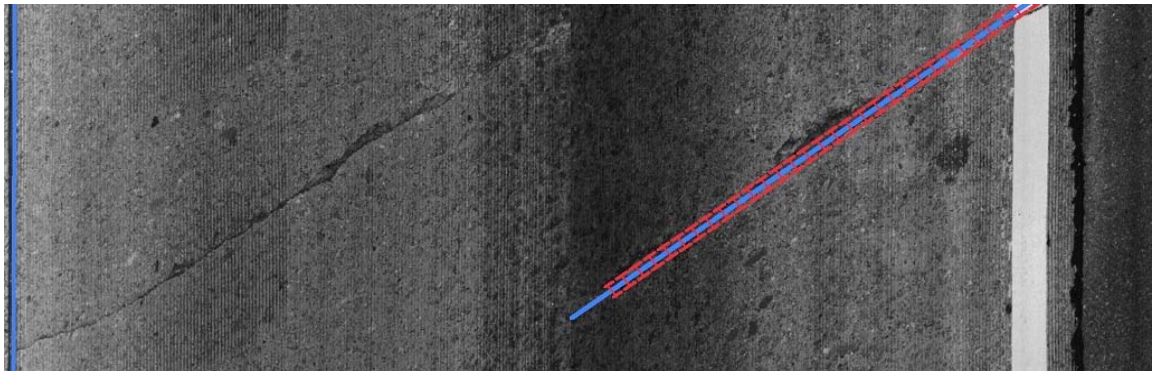


FIGURE 3.23: Example of false-positive cases

This type of case is, naturally, difficult to eliminate because these cracks have almost all the unique characteristic of joints. However, it is very rare for the crack to maintain a consistent width and straightness in both sides of the pavements. So, unlike transverse joints, normally these cases would only occur in one side of the pavement. As a result, it is easy to find them by comparing the detection results with the other side of the pavement.

Five false-negative cases have the detection results only on one side, as shown in FIGURE 3.24:

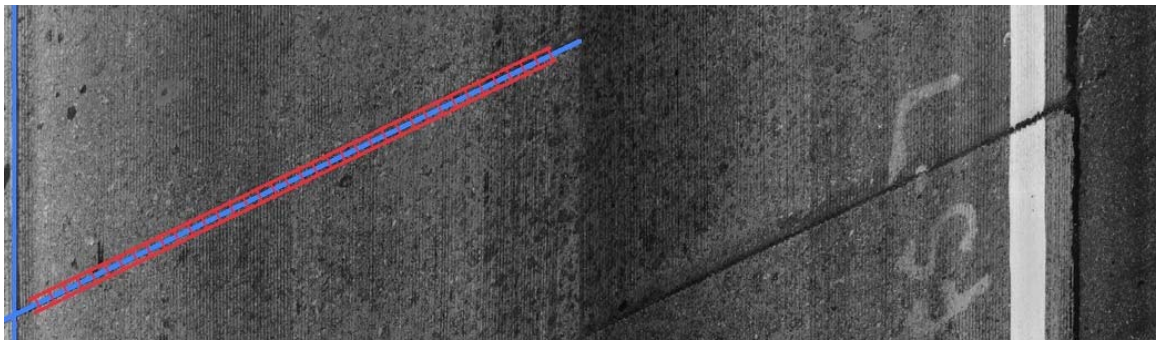


FIGURE 3.24: Example of false-negative cases

Most of these cases contain joints with a portion outside the image, which may cause problems in detection. However, it is not common for the false-negative cases to occur on both sides. By examining the detection result on the other side of the pavement, the impact of false-positive cases can be much reduced.

Some other special cases were also tested. There are 5 images containing joints with asphalt patches (FIGURE 3.25), and all such cases are detected correctly, which indicates patching may not influence the detection process.

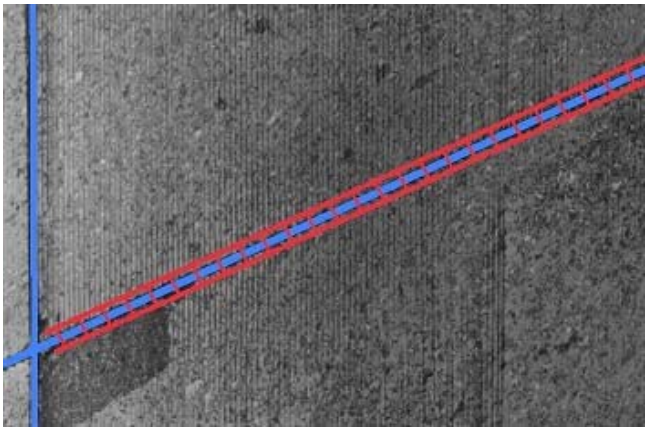


FIGURE 3.25: Example of joints with patch

Faulting Measuring Module Test

Data Collection and Ground Truth Establishment

For faulting measurement, we collected data on a section of I-16. The location is shown in the map in FIGURE 3.26. Field measurement was conducted to establish the ground truth for this validation test. Three places on the route were selected for manual ground-truth data collection; each of the three places contained 10 continuous concrete slabs. One place is around EB MP159 (Site #1), and the other two spots are on WB MP159 (Site #2) and MP156 (Site #3). All those joints measure are marked with color painting. The Georgia Faultmeter, which has already been introduced in the literature review, was used as the measuring tool, as shown in FIGURE 3.27.

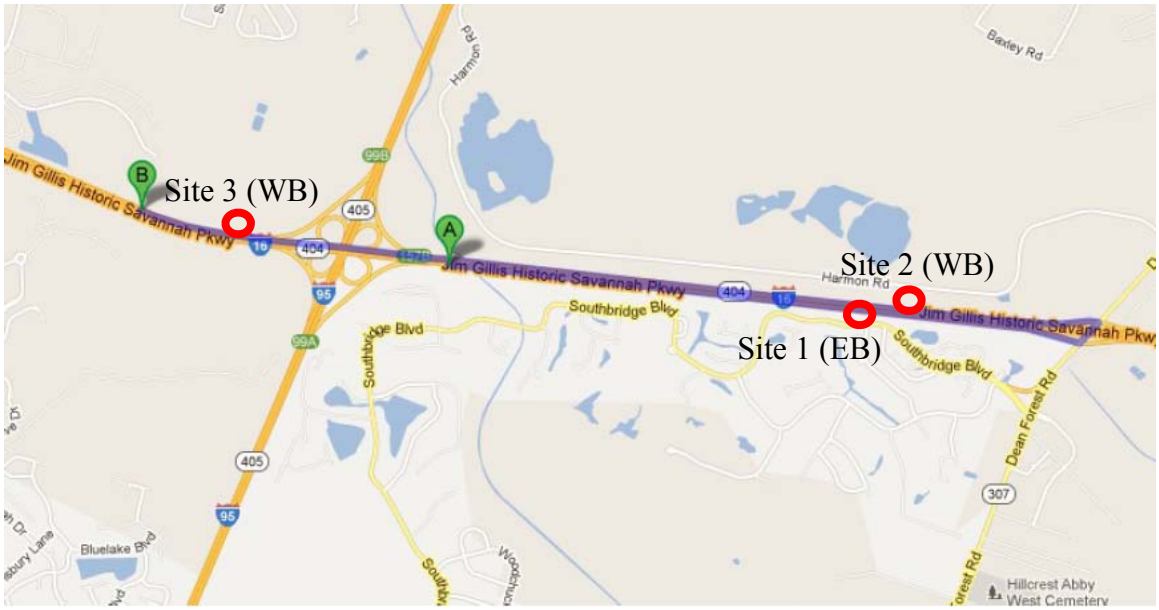


FIGURE 3.26: Locations of data collection

FIGURE 3.27: *Left*: Georgia Faultmeter; *Right*: Manual measurement on the roadway

To reduce random errors, multiple runs were conducted for the same pavement section. The marked measuring location example is shown in FIGURE 3-28, with the white dot painted on the joint indicating the measuring point:

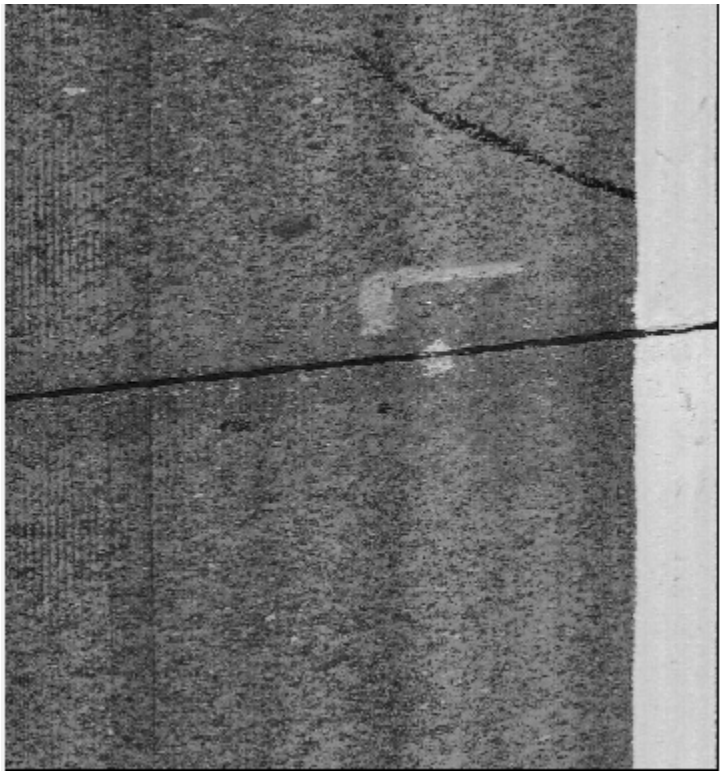


FIGURE 3-28 Marked measuring point in pavement surface laser data

Testing Results

The tests consisted of several major steps. First, a proper configuration was needed for the software to achieve good performance. In this test, we added a further averaging procedure, which calculated the average detected elevation differences among all points within the marked area to reduce possible random errors. After that, a comprehensive analysis and interpretation of the final result were made to arrive at the conclusion.

a) Parameter Configuration

The faulting detection module of the software has several adjustable parameters, which would influence the detection performance. The most important parameters are illustrated in FIGURE 3.29:

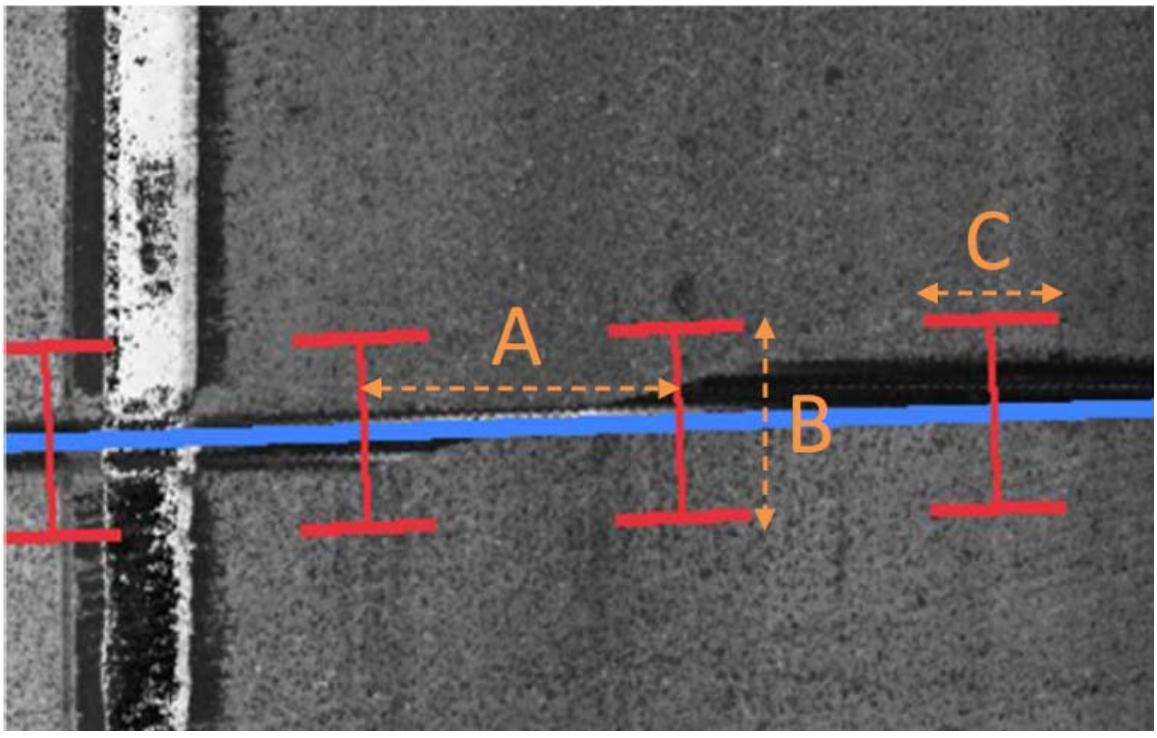


FIGURE 3.29: Faulting measurement parameters

The mechanism of the detection module obtains the elevations of both ends of the red bars in FIGURE 3.29, and then compares their differences in elevation. A is the distance between consecutive measuring points along the joint; B is the distance between two measuring ends; and C is the averaging window length. Our averaging program requires the distance A to be 1 mm, which will be elaborated upon later. As the averaging process would be done in the next step, we set the averaging window size, which is the distance C in FIGURE 3.29, as 1 mm, too. So, the only variable left adjustable is the distance B. To find out the optimal value of distance B, we tried multiple settings, including 50 mm, 100 mm, 200 mm, and 300 mm. The test results are discussed below:

b) Initial Result Averaging

As is shown in FIGURE 3-28, the measuring point is actually an area in the image, so to increase the precision, we set the distance A to a very small value (1 mm) and programmed to get the average evaluation difference within the area as the final outcome. Normally, there will be nearly 40 points within the marked area. We first find these points by the X coordinate range (horizontal), and then extract them from the XML file generated by the software.

c) Result Analysis

The final result after the averaging process is discussed in this section. From the test results, we calculated the mean and standard deviation values of the error, which is defined as the absolute difference between the software-measured value and the ground truth for different settings. We also calculated the 95% confidence upper limit of the error for every setting - basically this means the error has a probability of 95% lying below this limit.

The Site 1 data has the largest error, with more than 2 mm at 95% of the upper limit for all the settings; in the other two sites, this value can be controlled within 2 mm for some settings. However, even for Site 1, the average error value is only slightly higher than 1 mm; in the other two sites, the average errors are both below 1 mm. In reality, the faulting cases we have interests in normally have elevation differences larger than 5 mm, so a mean error of approximately 1 mm and maximum of 2 mm variance for small portion cases is acceptable. Through performance comparison among different settings of B distance, it can be seen that the 50 mm setting has the worst accuracy with the highest mean and deviation values of error in all three sites; the other three groups are much better. The performance of 100 mm, 200 mm, and 300 mm are slightly different. At Site 1 and Site 2, 100 mm has the best performance, while 300 mm has the worst, approximately 1 mm and 0.5 mm lower in the 95% upper limit, respectively; in Site 3, the 300 mm setting performs the best, while 100mm performs the worst, yet is only 0.3 mm lower in the 95% upper limit. So, 100 mm has overall slightly better performance than 300 mm. Meanwhile, the 200 mm setting has consistently good performance, which can also be a recommended setting.

Table 3.5: Test Result for Site 1 (Error = |Measured Value – Ground Truth|)

Site 1	Field Test(Ground Truth)				LCMS Results			
Sample number	Round1	Round2	Round3	Average / mm	50mm	100mm	200mm	300mm
1	0.79	0.79	2.38	1.32	0.23	0.49	1.07	2.23
2	3.18	1.59	1.59	2.12	0.41	0.54	-0.03	0.41
3	2.38	2.38	3.18	2.65	-1.62	0.62	1.51	0.21
4	2.38	3.18	3.97	3.18	0.61	1.65	2.61	2.33
5	1.59	0.79	1.59	1.32	0.83	0.48	2.40	1.00
6	2.38	3.18	3.18	2.91	5.34	1.83	1.70	2.05
7	1.59	1.59	0.79	1.32	-1.43	0.40	0.63	0.94
8	5.56	5.56	4.76	5.29	2.48	4.18	3.23	2.55
9	3.97	3.97	4.76	4.23	0.72	2.20	1.19	1.23

10	1.59	1.59	1.59	1.59	-0.11	0.74	1.09	1.02
Extra	3.18	3.18	1.59	2.65	0.86	1.30	2.37	2.56
Error Mean	/				2.28	1.29	1.18	1.26
Error Std dev.	/				1.03	0.43	0.85	0.99
95% Conf. Upper Limit	/				4.29	2.13	2.85	3.20

Table 3.6: Test Result for Site 2 (Error = |Measured Value – Ground Truth|)

Site 2	Field Test(Ground Truth)				LCMS Results			
Sample number	Round1	Round2	Round3	Average / mm	50mm	100mm	200mm	300mm
1	11.11	11.11	11.91	11.38	/	/	/	/
2	0.79	0.79	0.79	0.79	0.49	-0.021	1.438	1.057
3	4.76	4.76	5.56	5.03	6.28	4.718	4.088	3.708
4	3.97	3.18	3.18	3.44	8.06	3.588	5.491	5.464
5	7.14	6.35	5.56	6.35	2.29	7.087	7.204	8.107
6	3.18	3.18	3.18	3.18	0.047	1.54	2.776	2.7
7	3.97	3.18	2.38	3.18	10.69	2.123	3.521	4.457
8	3.18	3.18	3.18	3.18	1.46	2.645	2.669	3.648
9	5.56	5.56	5.56	5.56	2.4	4.734	4.917	4.709
10	9.53	9.53	9.53	9.53	7.27	10.826	10.013	9.154
Error Mean	/				3.11	0.82	0.76	0.98
Error Std dev.	/				2.02	0.44	0.49	0.61
95% Conf. Upper Limit	/				7.06387	1.6815	1.7269	2.1686

Table 3.7: Test Result for Site 3 (Error = |Measured Value – Ground Truth|)

Site 3	Field Test(Ground Truth)				LCMS Results			
Sample number	Round1	Round2	Round3	Average / mm	50mm	100mm	200mm	300mm
1	5.56	5.56	5.56	5.56	6.005	4.481	3.838	4.138
2	7.14	7.94	7.94	7.67	5.398	8.745	8.216	7.284
3	3.97	3.97	3.97	3.97	5.598	3.718	4.42	4.19
4	3.97	3.97	3.97	3.97	/	/	/	/
5	7.14	7.94	7.94	7.67	6.686	7.988	8.016	7.564
6	3.97	3.97	3.97	3.97	5.684	5.275	4.443	3.45
7	4.76	3.97	4.76	4.50	1.88	4.837	5.018	5.738
8	5.56	6.35	6.35	6.09	6.307	7.854	7.239	5.749
9	4.76	3.18	3.97	3.97	4.449	4.642	4.707	4.844
10	6.35	7.14	6.35	6.61	/	/	/	/
Error Mean	/				1.30	0.85	0.74	0.64
Error Std dev.	/				0.84	0.51	0.44	0.45
95% Conf. Upper Limit	/				2.94047	1.8517	1.5989	1.5282

FIGURE 3.30, FIGURE 3.31and FIGURE 3.32 visualize the detection results. It's fairly clear that the 100 mm, 200 mm and 300 mm settings all outperform the 50 mm one.

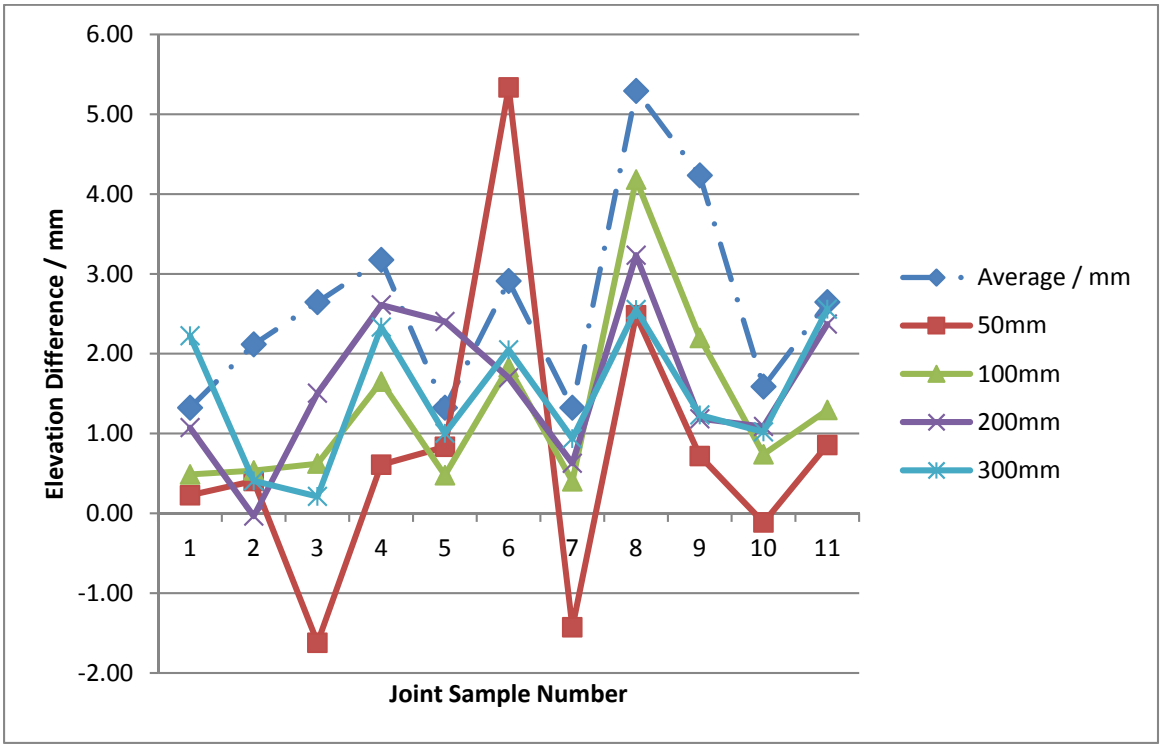


FIGURE 3.30: Site 1 result

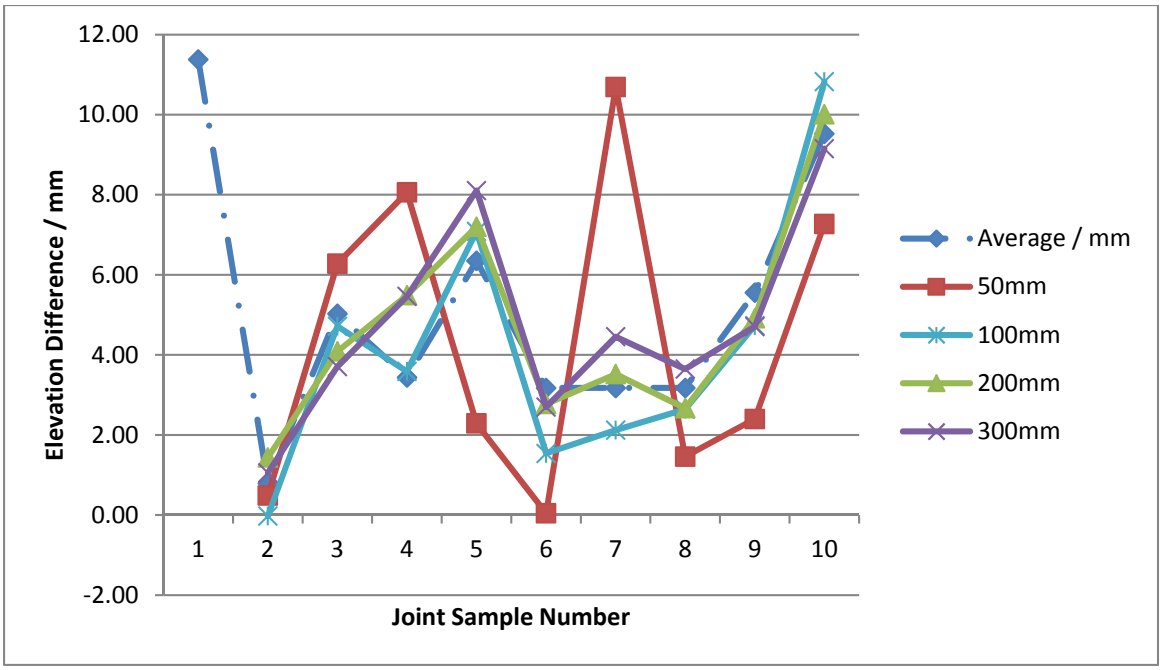


FIGURE 3.31: Site 2 result

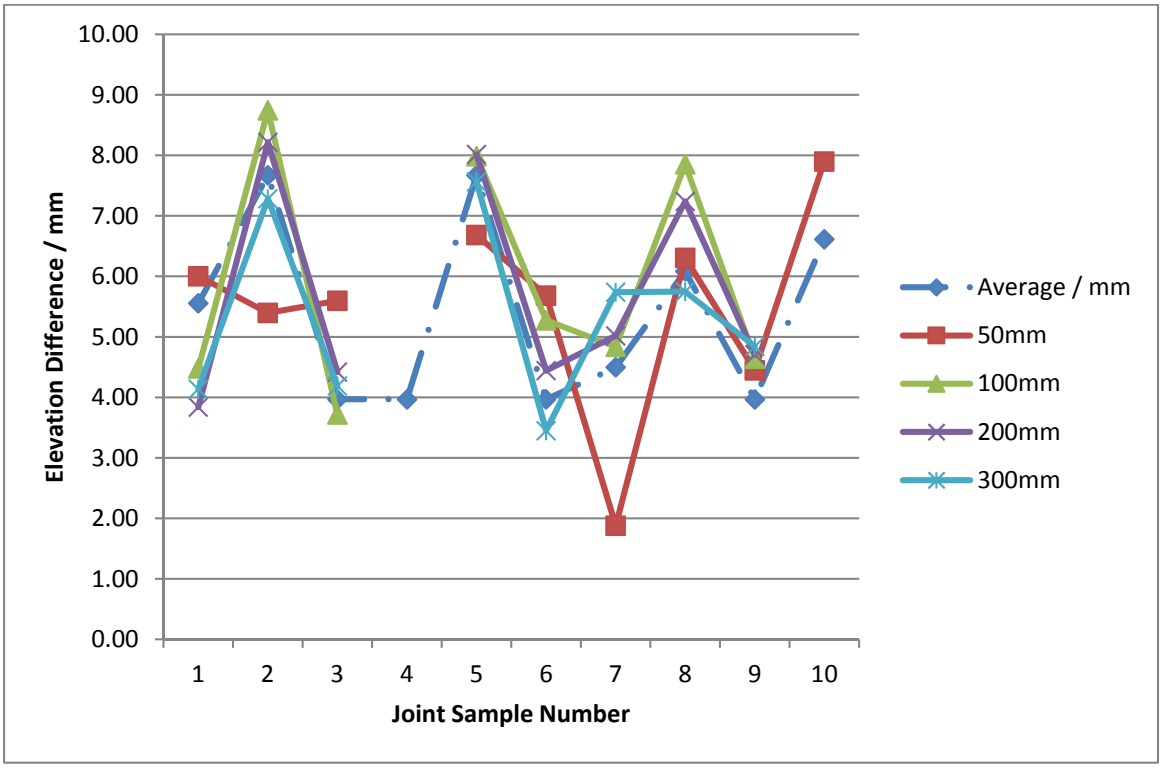


FIGURE 3.32: Site 3 result

In this test, most cases have been successfully measured, yet three cases in which joints have been successfully detected failed to provide faulting measurement information. Results with the symbol '/' indicate failure to provide faulting information. Most of these problems occurred in joints with a portion outside the image and with a marked location close to the image edge, as shown in FIGURE 3.33.

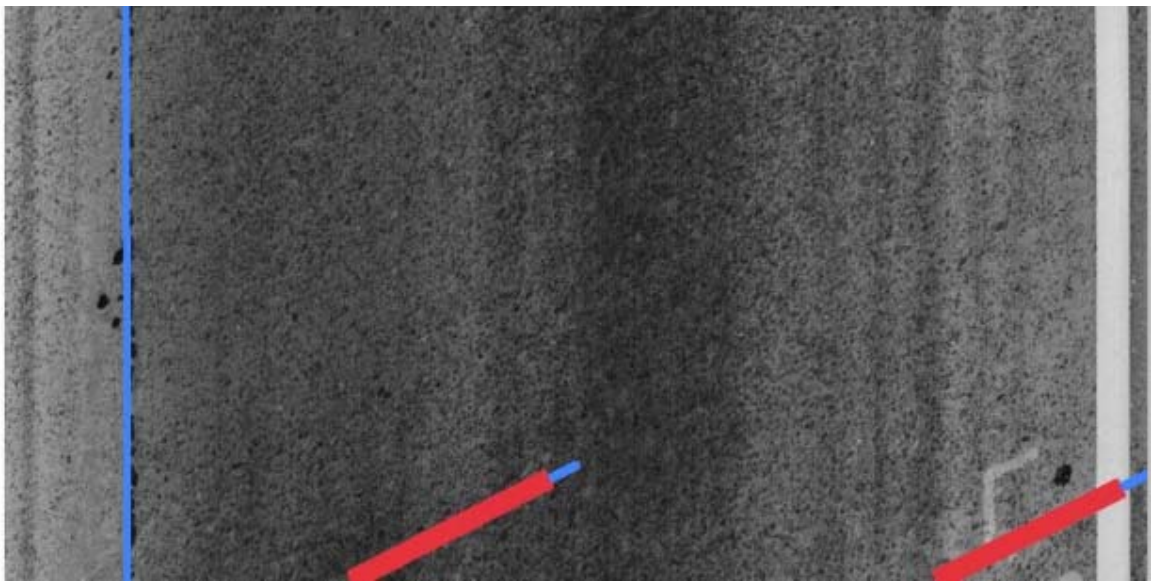


FIGURE 3.33: Example of joints with excluded part

In conclusion, with proper settings, the pavement surface laser data, along with the faulting detection software, can measure the elevation difference across transverse joints with an average error of approximately 1 mm, and less than 1 mm variance in most cases; the 95% error limits can be controlled to within 2 mm in two of the three sites tested, and slightly over 2 mm in the third site. Per the requirement of the project needs, the accuracy provided is acceptable. According to the testing results, the recommended setting of the distance between two measuring points can be either 100 mm or 200 mm. If possible, multiple tests using different settings are recommended to reduce false detection cases.

3.3 Summary

To validate if the 3D pavement surface data acquired by the new technology could provide sufficient accuracy, two detection methods are tested in the study. For the regression-based detection method, both controlled in-lab tests and field tests were conducted for performance evaluation. The controlled test shows less than 0.6 mm mean error with small variances in multiple cases; the field test, which collects data at highway speed, shows less than a 1.5 mm mean error with a standard deviation smaller than 1 mm at 80 km/h and 100 km/h. So, though the speed of data collecting can affect detection accuracy, the variation is not significant and won't affect the consequent analysis.

For the built-in faulting detection software module test, in which the data-collecting vehicle drove at approximately 96.6 km/h (60 mile/h), different software configurations were tested to achieve the optimal accuracy. It was found that with 100 mm or 200 mm as the measuring distance setting, the average error can be controlled around 1 mm with less than 1 mm variance. Future research could focus on improving handling missed cases in which joints are not completely inside the images.

Judging from these validation test results, it is concluded that with the proper detection algorithm and software configuration, collecting faulting data by using the integrated sensing system at highway speeds to collect 3D continuous profile data is feasible and produces data with desirable accuracy and repeatability. Though the regression-based method and the built-in software occasionally produce incorrectly detected cases, the overall detection accuracy is acceptable, and through post-processing, it is easy to reduce the impact of such cases.

4. Validation of Concrete Pavement Spalling Detection

Spalling is a common type of concrete pavement distress that usually occurs along the construction joints between concrete slabs. According to the LTPP's definition, spalling refers to cracking, breaking, chipping, or fraying of slab edges within 0.3 m from the face of transverse joints. GDOT defines spalling as the deterioration of concrete at the joint, which usually starts with the breaking or flaking off of the concrete at the joint; as they worsen, raveling and pop-outs occur (as shown in FIGURE 3.34). Due to its random spatial distribution pattern on the concrete pavement, it is difficult to design a sampling strategy for a manual spalling survey. Therefore, automatic spalling detection will help improve the cost-effectiveness of concrete pavement maintenance and enhance roadway safety.

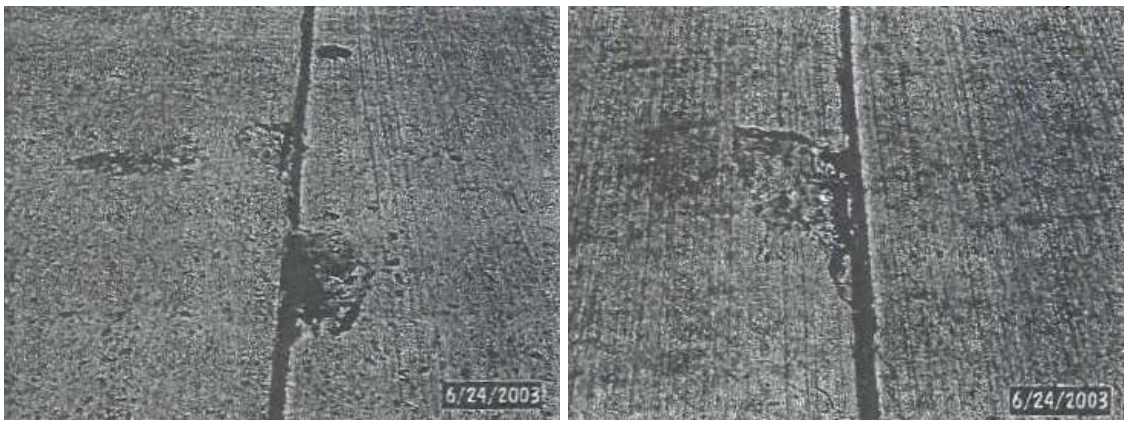


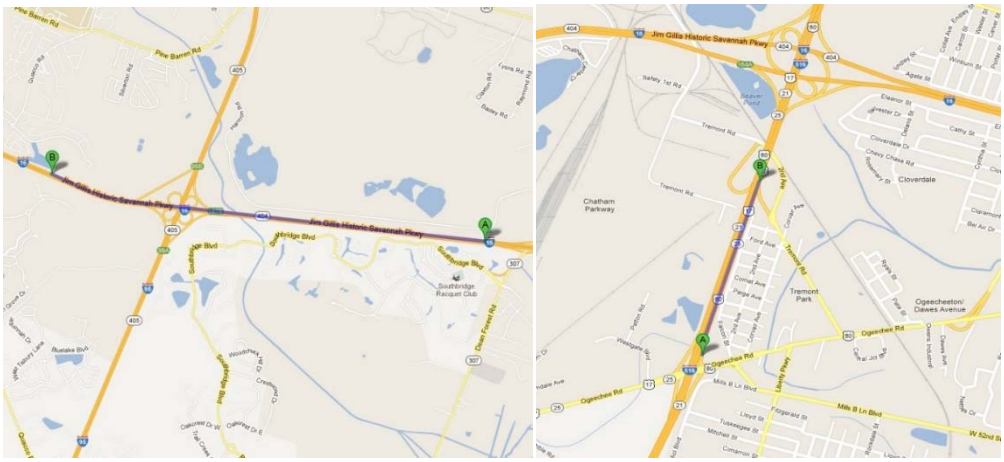
FIGURE 3.34: Example of spalling at concrete joint (GDOT 2003)

With the advances in sensing technology, an automatic spalling detection method can now be developed using 3D laser pavement data. The section will validate the accuracy of automatic spalling detection on concrete pavement. In addition, some representative false negative cases (e.g. asphalt-filled spalling and small/shallow spalling) and false

positive cases (e.g. crack spalling, shoulder-joint distress, and concrete coring) will be presented to reveal the potential issue of concrete pavement spalling detection.

4.1 Experimental data

The experimental data of the validation test were collected on interstate highways I-16 and I-516 near Savannah, Georgia. Both highways are concrete pavement at the selected test sites (as shown in FIGURE 3.35).



(a) I-16 test site (b) I-516 test site

FIGURE 3.35: Selected test sites on concrete spalling detection

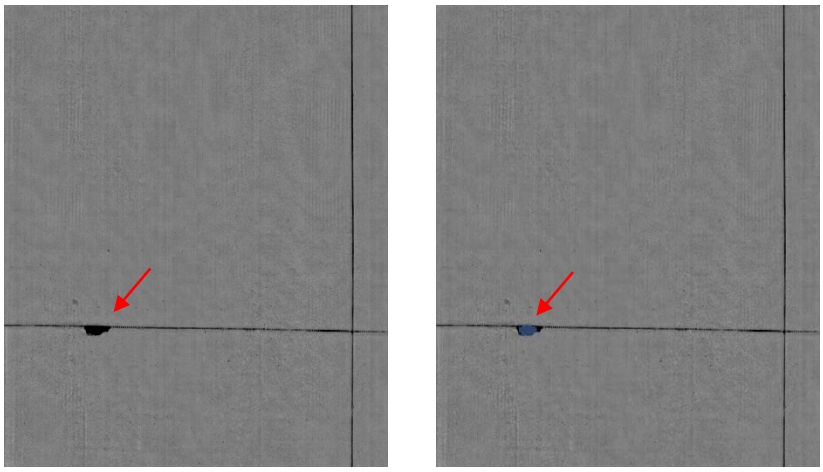
4.2 Experimental design

The following procedures were conducted to evaluate the performance of automatic spalling detection on concrete pavement:

- First, the ground truth was established by manually reviewing the collected pavement surface laser data and identifying the location of spalling. Since the distribution of spalling is sparse on actual concrete pavement, it is not practical to manually record the ground truth along the road. However, spalling has a distinctive elevation change on concrete pavement, and a 3D range image is sufficient to support the ground truth established for validation purposes (as shown in FIGURE 3.36(a)). Based on the

visual review of collected range images from the entire test site, the spalling location is manually identified.

- Then, the automatic spalling detection algorithm is applied. The spalling detection results can be overlaid on an intensity image or range image (FIGURE 3.36(b)).
- Finally, the ground truth and automatic detection results are compared. The number of correctly-detected/missed spalling are counted to calculate the overall detection precision.



(a) Ground Truth (b) Automatic Spalling Detection Result

FIGURE 3.36: Illustration of spalling evaluation procedure

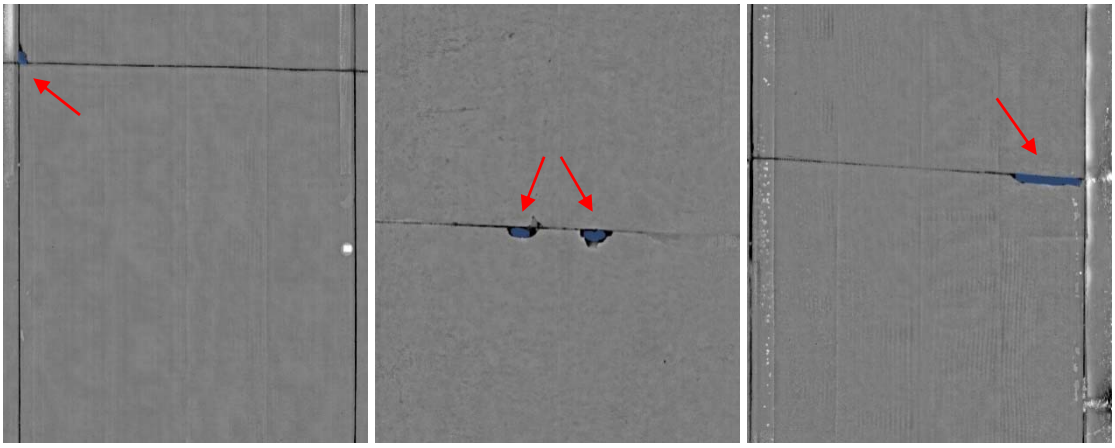
4.3 Validation on I-516 and I-16 test sites

Based on the manual review, a total of 86 spalling locations were identified on the selected test sites. Of these, 65 spallings were correctly detected, while 21 spalling locations were missed by the automatic detection algorithm, which provides a detection precision of 75.6%. Some representative cases are presented below.

Correct Detection Cases

FIGURE 3.37 shows three representative examples of correctly detected spalling. It can be observed that the location and shape of spalling do not have a significant impact on the

performance of the automatic spalling detection algorithm. The automatic detection can also extract multiple spalling locations from the same pavement image.



(a) Lane Line (b) Middle (c) Edge Line

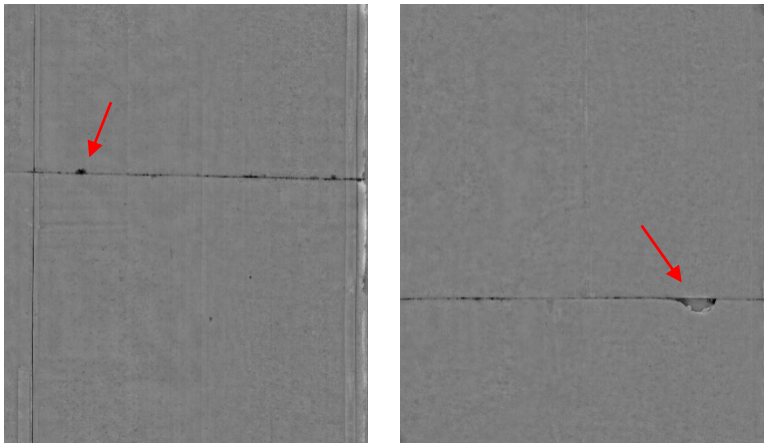
FIGURE 3.37: Examples of correctly detected spalling

False Negative Cases

After a careful review of the 21 spalling locations that were missed by the automatic detection algorithm, these false negative cases can generally be classified into two different types:

- Undersized spalling: as shown in FIGURE 3.38(a), the size of the spalling is too small to be captured by the automatic detection algorithm. Besides the area, the depth of the spalling is another factor that may influence detection performance. In some cases, the spalling is too shallow and not distinct enough on the range image. The minimum size of spalling to be detected can be adjusted as a parameter in the automatic algorithm; however, a lower limit that is too small may introduce additional false positives into the spalling detection. On the other hand, GDOT only requires recording spalling sites whose areas are larger than 1.5’’x6’’. Through a rough measurement on the range image, the length of the spalling in FIGURE 3.38(a) is about 2.4 inches (around 60mm), which does not have to be recorded.

- Asphalt-patched spalling: as shown in FIGURE 3.38(b), some spalling sites have already been patched with asphalt in previous maintenance. Although some asphalt patches may not be in good condition, they still reduce the distinctiveness of spalling on concrete pavement and introduce the potential for false positives in spalling detection.



(a) Undersized Spalling (b) Asphalt-patched Spalling

FIGURE 3.38: Examples of false negative cases of spalling detection

False Positive Cases

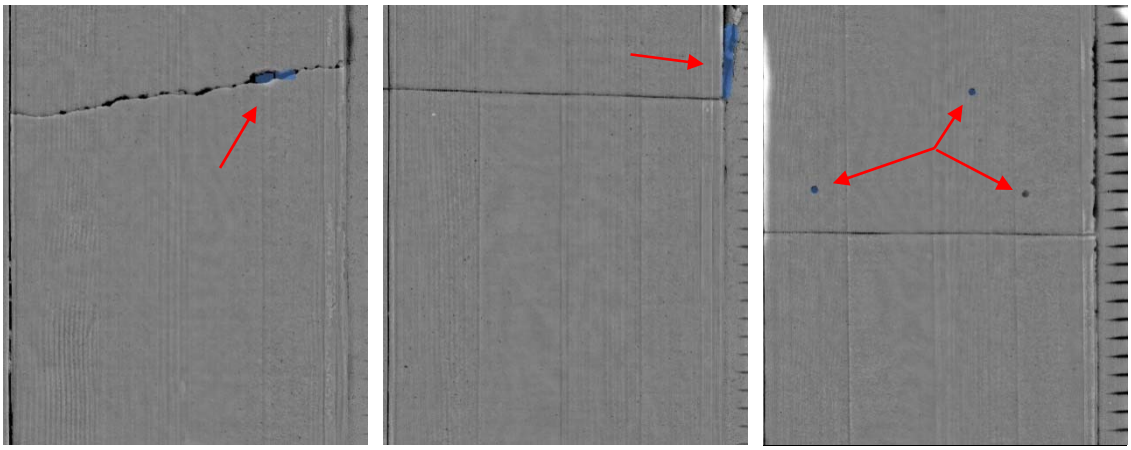
Based on the review of the spalling detection results on the entire test site, some false positive cases are also identified. These cases can generally be classified into three different types:

- Spalling on severe cracking / broken slab: As mentioned in the previous section, spalling can occur along both transverse and longitudinal cracking on concrete pavement, which usually indicates a high severity level of cracking. The spalling detection algorithm itself cannot differentiate spalling along joints or along cracks and, therefore, introduces some false positive cases through the validation test (as shown in FIGURE 3.39 (a)). Since we have already validated the joint detection algorithm on concrete pavement, these false positives can easily be removed as a

follow-up step after spalling detection by comparing the location of detected spalling and joints.

- Shoulder-joint distress: This type of distress is caused by vertical movement in concrete slabs; it usually happens on the shoulder area right after a transverse joint. It has some characteristics similar to spalling and is also extracted by the spalling detection algorithm (as shown in FIGURE 3.39(b)). Although these results are false positive for spalling detection, it does reveal the possibility of using the spalling detection algorithm to detect and measure this type of shoulder-joint distress. More details will be presented in the next section. Similarly, false positives can easily be removed by comparing the location of detected spalling and transverse joints.
- Coring: The quality control of concrete construction usually requires core drilling on the concrete surface. The deep coring has significant elevation changes and, therefore, is captured as false positives in some cases (as shown in FIGURE 3.39(c)). These false positives can also be removed based on location information.

Overall, although quite a few false positive cases are observed through the validation test, almost all of them can be eliminated by comparison with the detected joint location; therefore, these false positive cases should not have a significant influence on the real-world implementation of the automatic spalling detection algorithm.

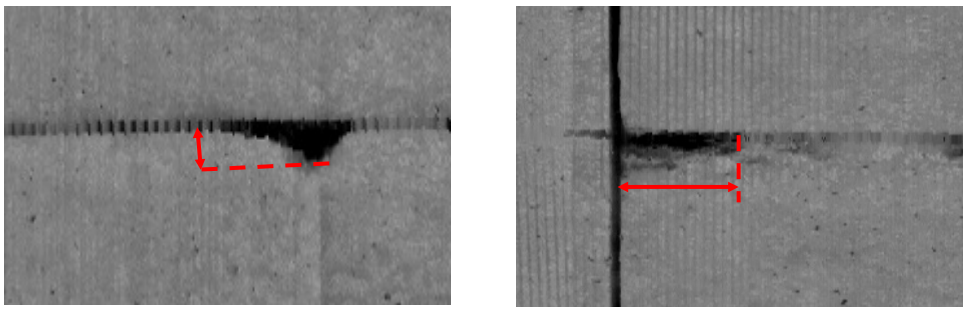


(a) Cracking (b) Shoulder-Joint Distress (c) Coring

FIGURE 3.39: Examples of false positive cases of spalling detection

Detection Accuracy vs. Spalling Size

To find the relationship between the spalling size and detection accuracy, we compared the missed cases and the correctly detected cases in terms of spalling size. According to the definition in the LTPP manual, we treat the distance from the furthest edge of the spalling to the joint as the spalling size. So, for longitudinal spallings, the size is in the transverse direction, and for transverse spallings, it is in the longitudinal direction. An example is shown in FIGURE 3.40 (a). In some cases, the spalling is located in the intersection between longitudinal and transverse joints; we use the longer distance to either joint in two directions as the spalling size, shown as FIGURE 3.40 (b).



(a) Spalling on a single joint (b) Spalling in the intersection

FIGURE 3.40: Examples of spalling size measurement

The detection results are reorganized according to the measured spalling size, and the obtained frequency for each size group is shown in FIGURE 3.41 and Table 3.8. It can be seen that the software can hardly detect spalling sizes smaller than 50 mm; the software can detect some spalling sizes larger than 50 mm but smaller than 90 mm, but the accuracy is not very good (59%); for spalling sizes larger than 90 mm, the detection accuracy becomes much better (90%). One of the reasons for these results could be that the software module was originally designed for pothole detection; though the spalling

has characteristics similar to potholes, spalling tends to be much smaller. So, the pothole detection module may have some internal filtering mechanism that eliminates small candidates and reduces computation time, which won't affect pothole detection but could cause false negative spallings. To examine such an assumption, we developed our own detection algorithm that successfully detected most of the cases missed by the software module and using the same dataset. The test indicates that the 3D data have the capability to provide desired accuracy for spalling detection, and in cases of small-sized spallings, future algorithm modification may help improve the detection accuracy.

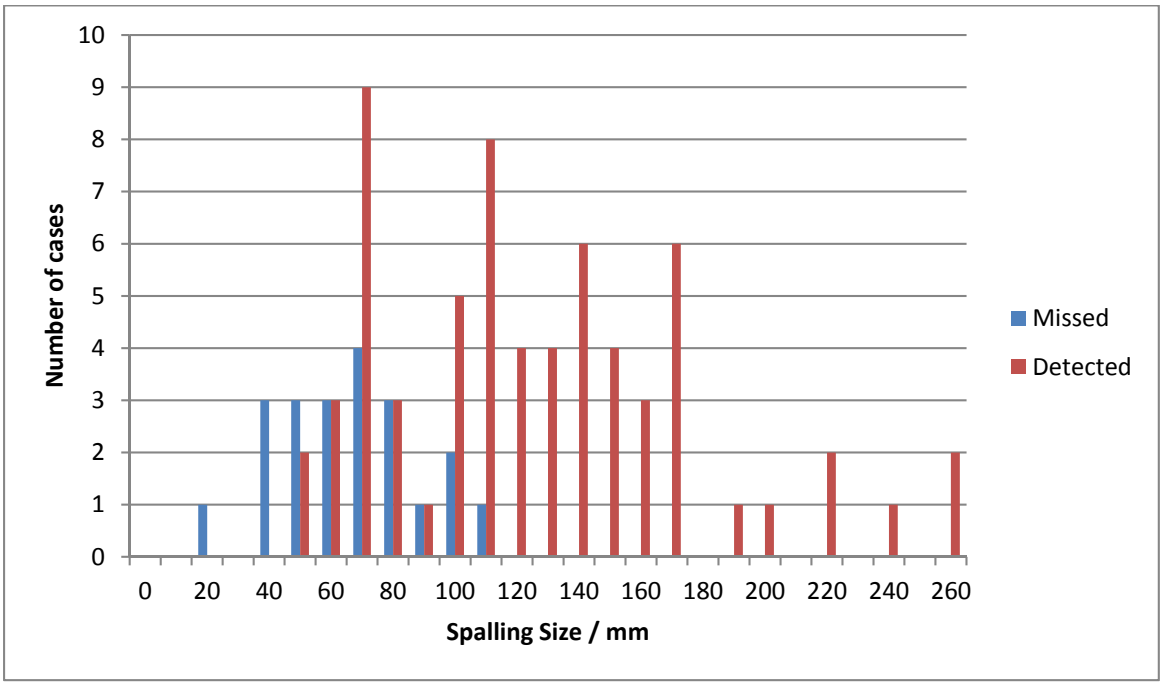


FIGURE 3.41: Frequency vs. spalling size

Table 3.8: Software Detection Accuracy for Different Spalling Sizes

Spalling size / mm	≤50	50-90	>90	Total
Missed cases	7	11	3	21
Detected cases	2	16	47	65
Accuracy	22%	59%	90%	76%

4.4 Summary

For automatic concrete pavement spalling detection using a 3D laser data, the major findings are summarized as follows:

First, the 3D laser technology shows overall acceptable performance for automatic spalling detection on concrete pavement. Two test sites were selected on interstate highways I-516 and I-16 to evaluate the spalling detection performance. The ground truth was established by manually reviewing the collected range images. The automatic spalling detection shows accurate results: 65 out of 86 spalling locations were detected, which provides a detection precision of 75.6%. The spalling detection is also robust to different spalling locations and shapes.

Second, two types of representative false negative cases are observed, including undersized spalling and asphalt-patched spalling. Undersized spalling may not influence the detection performance in real-world implementation, since GDOT's current manual survey practice already has a minimum requirement for the size of spalling to be recorded. Asphalt-patched spalling is expected to impact the detection accuracy; however, the percentage of this type of cases is minor in the selected test sites.

Third, false positive detections are also observed through the review of the experimental results, including three major types: crack spalling, shoulder-joint distress, and coring. Most of these false positives can be removed by comparing their locations with detected transverse joints.

Fourth, the detection accuracy of the built-in spalling detection software varies for locations with different sizes of spalling. Spalling less than 50 mm wide can hardly be detected; between 50 and 90 mm wide, the detection accuracy is better, but not satisfactory; for spalling locations more than 90 mm, the accuracy becomes quite good. Such results may due to the internal filtering mechanism of the algorithm to reduce computation time. Our own detection algorithm has successfully detected most cases missed by the software, indicating the capability of the data to provide the desired accuracy.

5. Validation of Concrete Pavement Shoulder Joint Distress Detection

Shoulder joint distress is another type of concrete pavement distress that occurs along the pavement shoulder. GDOT defines shoulder joint distresses as a depression in the shoulder that is caused by the vertical movement in the concrete slabs under load, which may cause material to be pumped out at the joint (as shown in FIGURE 3.42). Since this distress takes the form of large spalling locations along the roadway edge, this section will explore the possibility of an automatic spalling detection algorithm that can identify the location and extent of shoulder joint distresses using 3D pavement data.

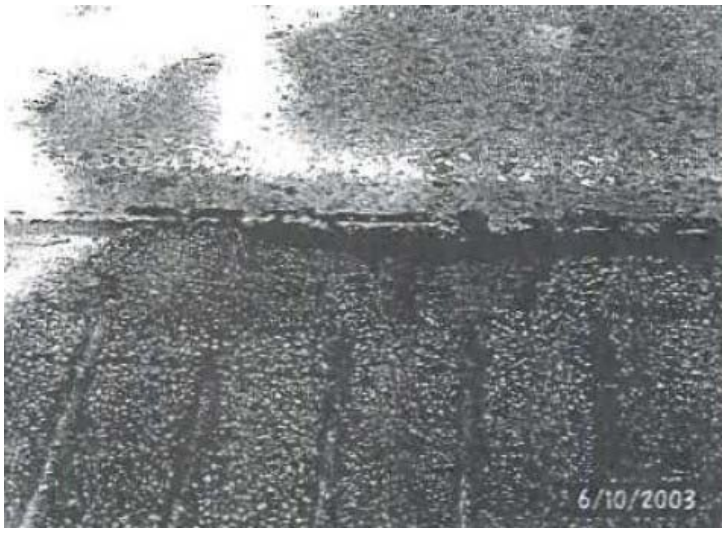


FIGURE 3.42: Example of shoulder joint distress (GDOT 2003)

5.1 Experimental data

The experimental data of the validation test were collected on interstate highway I-16 near Savannah, Georgia, with concrete pavement on the selected test site (as shown in FIGURE 3.43).

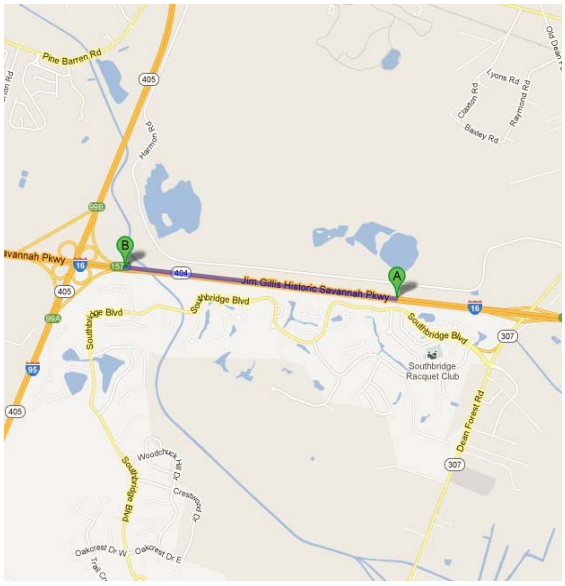


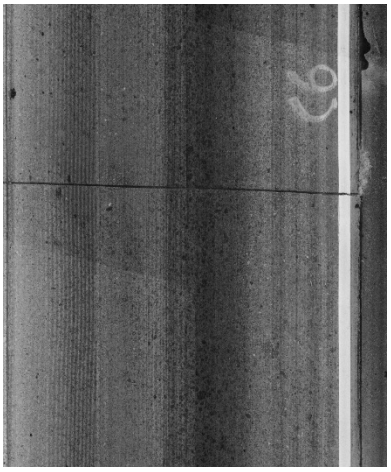
FIGURE 3.43: Selected I-16 test site on concrete shoulder joint distress detection

5.2 Experimental design

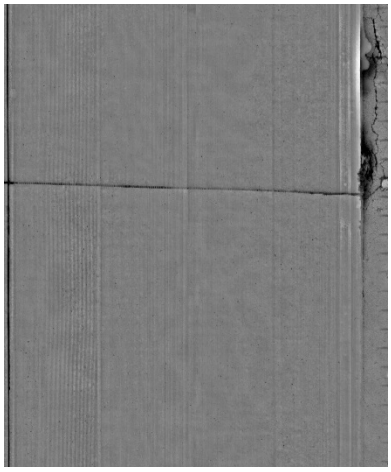
Unlike for cracking, faulting, and spalling, it is difficult to provide a quantitative ground truth for shoulder joint distresses due to their continuity and large extent. Therefore, we manually selected several slabs at the test site. Each of these slabs was marked with a unique ID, and the shoulder joint distresses along these slabs were roughly drawn on a distress map for comparison purposes. Digital photos of the selected slabs were taken from the road shoulder to track the detailed distress conditions and provide certain references through the validation. Then, the automatic spalling detection algorithm was applied on these slabs to see if it could capture the location and extent of shoulder joint distresses. The detection results were visually compared to photos and distress maps taken from the field survey.

5.3 Validation on I-16 test site

FIGURE 3.44 to FIGURE 3.47 show the experimental results of four representative cases. Images (a) and (b) show the intensity and range images of the selected location; Image (c) shows the detected shoulder joint distress (blue area); and Image (d) shows the digital photo at the same location. Based on the experimental results, it can be observed that although it is difficult to identify them in intensity images, shoulder joint distresses appear quite distinctively in range images. Having characteristics similar to spalling (elevation drop), their larger extents and areas actually simplify the automatic detection. Based on visual judgment, the detection results are consistent with field observations, including distress maps and digital photos.



(a) Intensity Image

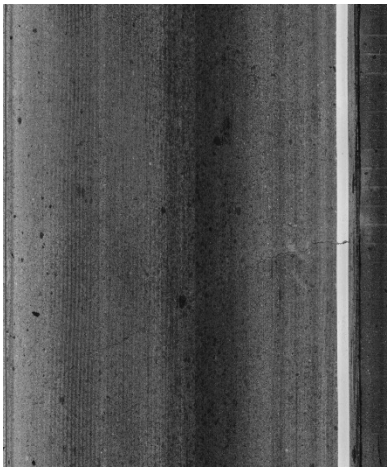


(b) Range Image

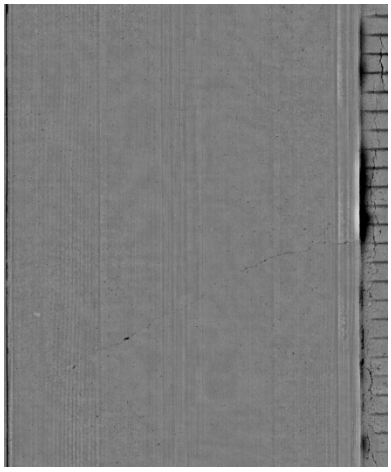
(c) Automatic Distress Detection

(d) Field Photo

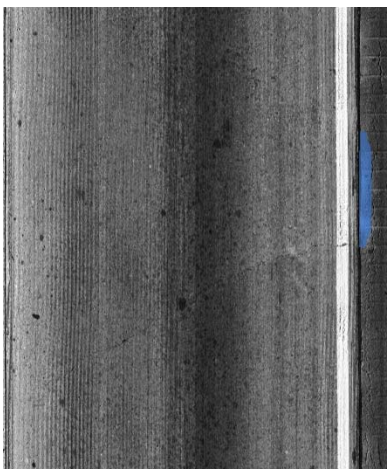
FIGURE 3.44: Shoulder-joint distress detection on Slab #C6



(a) Intensity Image



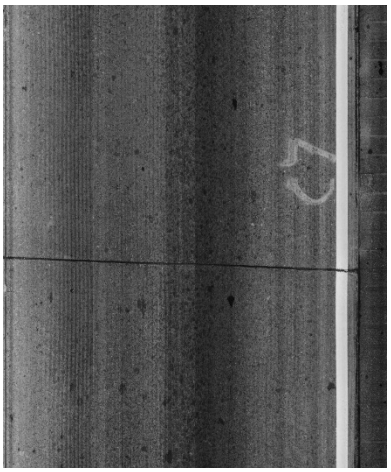
(b) Range Image



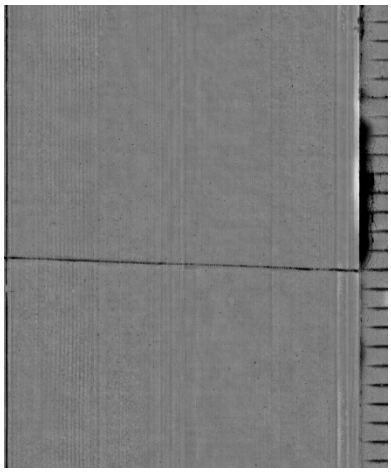
(c) Automatic Distress Detection

(d) Field Photo

FIGURE 3.45: Shoulder-joint distress detection on Slab #C6'



(a) Intensity Image

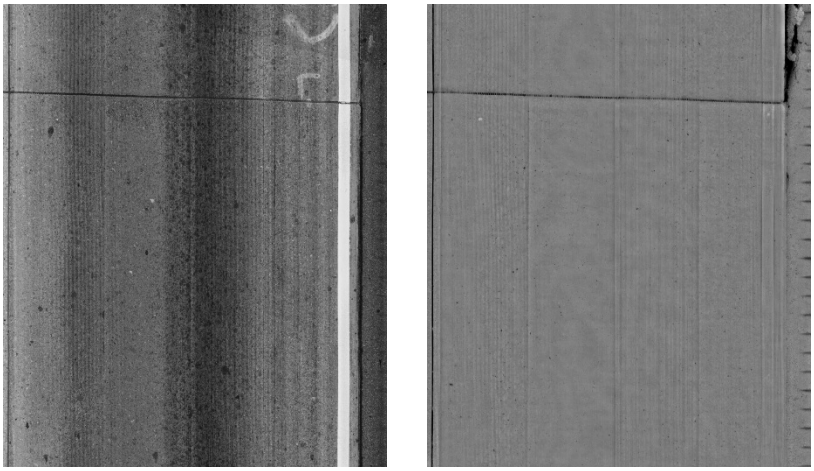


(b) Range Image

(c) Automatic Distress Detection

(d) Field Photo

FIGURE 3.46: Shoulder-joint distress detection on Slab #C7



(a) Intensity Image (b) Range Image

(c) Automatic Distress Detection (d) Field Photo

FIGURE 3.47: Shoulder-joint distress detection on Slab #C9

5.4 Summary

For automatic concrete pavement shoulder joint distress detection using a 3D laser data, the major findings are summarized as follows:

First, the 3D laser technology has shown the potential to capture the shoulder joint distresses. Because of the distinctive elevation change of such distresses on the pavement shoulder, they can also be automatically identified following principles similar to spalling

detection. The experimental results on the test sites are visually consistent with field observation. Furthermore, the automatic detection results are associated with the quantitative area and depth of the distress, which provides an objective and consistent way of making severity level judgments and extent reporting.

Second, it should be noted that due to the transverse coverage of the current 3D line laser-imaging system (about 4 meters), it is not guaranteed that the shoulder area will be captured while the vehicle is driving in the outside lane. Therefore, if this specific type of distress is targeted, it is suggested that data collection along the edge line be conducted; in this case, users should be aware that a narrow strip near the lane line might not be captured by the system.

Third, current shoulder joint distress detection is conducted using the automatic spalling detection algorithm, since both distresses share some similar characteristics. However, it is still suggested that a separate detection algorithm be developed to: 1) narrow the region of interest to the shoulder area to speed up the detection process and reduce the number of potential false positives and 2) deal with significant shoulder drop-off.

6. Summary

Emerging 3D laser technology has demonstrated its great potential to accomplish automatic distress detection on concrete pavement. A series of comprehensive validation tests were conducted on selected test sites on interstate highways I-16 and I-516. Four different types of concrete pavement distresses were validated, and the major findings are summarized as follows:

For concrete cracking, the 3D laser technology shows overall acceptable performance when performing automatic crack detection. Automatic crack detection results were compared with manually digitized ground truth data through a buffered Hausdorff scoring method. The automatic crack detection shows quite accurate and robust results on longitudinal cracks on the I-516 site; however, the crack detection performance on transverse cracks on the I-16 site was not as good. The false negative detection on transverse cracking can be explained by the larger data acquisition interval in the driving direction. As in asphalt pavement, the hairline cracks (thinner than 2 mm) are still challenging for automatic detection. Also, severe spalling and joint detection impact the crack detection performance, but only in a limited number of cases.

For concrete faulting, it is feasible to collect faulting measurement at highway speed using a 3D laser system. Using the regression-based method, the automatic faulting measurements are quite consistent with manually measured ground truths using the Georgia Faultmeter in both a controlled lab test and a field test. With proper parameter configuration, the concrete joint module in the commercial software can also provide acceptable faulting measurement results.

For concrete spalling, the automatic detection accuracy using the commercial software varies for different sizes. Spalling with widths larger than 90 mm can be successfully detected; between 50 and 90 mm wide, the detection accuracy drops but is still acceptable; spallings locations less than 50 mm wide can hardly be detected. Undersized spalling and asphalt-patched spalling are two common types of false negatives. Although some small spalling locations were not successfully detected, they can be clearly observed on a range image. Therefore, it is believed that there is still room to further

improve the automatic spalling detection algorithm. Some false positives were observed in the experimental test, including crack spalling, shoulder-joint distress, and coring. Most of these false positives can easily be removed by comparing their locations with detected transverse joints.

For shoulder joint distresses, this study explores the feasibility of using the automatic spalling detection algorithm on shoulder joint distress detection. The larger extent and depth of shoulder joint distresses make them quite distinctive in range images and relatively straightforward to detect. On the selected representative cases, the automatic detection results are visually consistent with field observations (including distress maps and digital photos). However, it should be noted that, due to the transverse coverage of the current 3D laser system (about 4 meters), it is not guaranteed that the shoulder area will be captured while the vehicle is driving in the outside lane. Also, developing a specific shoulder joint distress detection algorithm to further ensure accurate and robust detection is recommended.

References

Ahmed, M. F. M., and Haas. C. T. (2010). "The Potential of Low Cost Close Range Photogrammetry Towards Unified Automatic Pavement Distress Surveying." *Transportation Research Board Annual Meeting,* Washington, D. C.

Alekseychuk, O. (2006). "Detection of Crack-Like Indications in Digital Radiography by Global Optimisation of a Probabilistic Estimation Function." PhD Thesis, BAM-Dissertationsreihe, Band 18.

EI-Korchi, T., Gennert, M. A., Ward M. O., Wittels. N. (1991). "Lighting Design for Automated Pavement Surface Distress Evaluation." *Transportation Research Record:Journal of the Transportation Research Board*, 1311, 144-148.

Hou, Z., Wang K. C. P., Gong. W. (2007). "Experimentation of 3d Pavement Imaging through Stereovision." *Proc., International Conference on Transportation Engineering*, 376-381.

Kaul, V., Tsai Y., Mersereau. R. M. (2010). "A Quantitative Performance Evaluation of Pavement Distress Segmentation Algorithms." *Transportation Research Record: Journal of the Transportation Research Board*, 2153, 106-113.

Laurent, J., Lefebvre D., Samson. E. (2008). "Development of a New 3D Transverse Laser Profiling System for the Automatic Measurement of Road Cracks." *SURF 2008.*

Lee, H., and Kim. J. (2006). "Analysis of Errors in Ground Truth Indicators for Evaluating the Accuracy of Automated Pavement Surface Image Collection and Analysis System for Asset Management." *Journal of ASTM International*, 3(5), 1-15.

McGhee, K. H. (2004). "Automated Pavement Distress Collection Techniques: A Synthesis of Highway Practice." *NCHRP Synthesis 334,* Washington, D.C.

Nazef, A., Mraz, A., Gunaratne M., Choubane. B. (206). "Experimental Evaluation of a Pavement Imaging System: Florida Department of Transportation's Multipurpose

Survey Vehicle." *Transportation Research Record: Journal of the Transportation Research Board*, 1974, 97-106.

Raman, M., Hossain, M., Miller, R., Cumberledge, G., Lee, H., Kang. K., (2004). "Assessment of Image-Based data Collection and the AASHTO Provisional Standard for Cracking on Asphalt-Surfaced Pavements." *Transportation Research Record: Journal of Transportation Research Board*, 1889, 166-125.

Tsai, Y., Kaul, V., Mersereau, R. M. (2010). "A Critical Assessment of Pavement Distress Segmentation Methods Performance Evaluation of Pavement Distress Segmentation Methods." *Journal of Transportation Engineering*, 136 (1), 11-19.

Wang, K. C. P. (2000). "Designs and Implementations of Automated Systems for Pavement Surface Distress Survey." *Journal of Infrastructure Systems*, 6(1), 24-32.

Wang, K. C. P., and Gong. W. (2002). "Real-Time Automated Survey of Pavement Surface Distress." *Proc., International Conference on Applications of Advanced Technologies in Transportation Engineering*, 465-472.

Wang, K. C. P., and Tee. W. (2002). "Understanding of Pavement Profiling and Validation of an Implementation." *Proc., International Conference on Traffic and Transportation Studies*, 817-824.

Wang, K. C. P. (2004). "Challenges and Feasibility for Comprehensive Automated Survey of Pavement Conditions." *Proc., International Conference on Applications of Advanced Technologies in Transportation Engineering*, 531-536.

Wang, K. C. P., and Gong. W. (2005). "Real-Time Automated Survey System of Pavement Cracking in Parallel Environment." *Journal of Infrastructure Systems*, 11(3), 154-164.

Wang, K. C. P., Gong W., Hou. Z. (2008). "Automated Cracking Survey". *Proc., RILEM International Conference on Cracking in Pavements*, 881-889.

Xu, B. (2005). "Artificial Lighting for the Automated Pavement Distress Rating System." Austin, TX.

Xu, B. (2007). "Summary of Implementation of an Artificial Lighting System for Automated Visual Distress Rating System." Austin, TX.

Chapter 4 Feasibility Study of Pavement Marking Retroreflectivity Condition

1. Introduction

Pavement marking is one of the most important traffic control devices and is closely related to highway safety. It provides critical delineation information for road users. Pavement marking retroreflectivity is the most important feature of pavement marking to maintain its vital functionality during nighttime. FHWA has proposed a set of requirements as a mandate for the minimum pavement marking retroreflectivity and defined the requirements in the Manual on Uniform Traffic Control Devices (MUTCD, 2009). Public transportation agencies, including state DOTs, must design and implement their programs for pavement marking management to maintain retroreflectivity at or above the minimum level.

Traditionally, two manual methods are commonly used by public transportation agencies to collect pavement marking retroreflectivity condition data: nighttime inspection and retroreflectometer measurement. For nighttime inspection, a windshield survey is carried out at night by field engineers driving along the roadways and recording the road sections with poor retroreflectivity. Although nighttime inspections can be conducted at driving speed, the condition assessment results are subjective and inconsistent. A retroreflectometer measurement is conducted by making sampled retroreflectometer readings along the roadway. Although retroreflectometer readings can be consistent, as defined by the American Society for Testing and Materials (ASTM), field engineers need to conduct the measurement while being physically on the road. Such an operation is not

only labor-intensive and time-consuming, but, more importantly, they may be dangerous because engineers will be exposed to traffic. Therefore, there is a need for a safe and cost-effective pavement marking retroreflectivity condition assessment method that produces reliable and consistent results so that public transportation agencies can implement effective pavement marking management plans.

In recent years, many emerging sensing technologies have become technically mature and commercially available, e.g., computer vision, mobile LiDAR, scanning laser profiler, etc. It is important to explore feasible technologies and technology applications that can support a mobile pavement marking retroreflectivity condition assessment. Because a LiDAR system uses a principle similar to the one a retroreflectometer uses to measure retroreflectivity and because USDOT RS-GAMS Phase 1 has demonstrated the potential for applying LiDAR technology to the condition assessment of traffic sign retroreflectivity, a LiDAR system holds the potential to be applied to pavement marking condition assessment. Therefore, the objective of this study is to explore the feasibility of measuring a pavement marking's retroreflectivity condition using LiDAR technology.

Section 1 presents the background and identifies the objective of this study. Section 2 presents a literature review regarding the current pavement marking retroreflectivity condition practice and the previous studies on mobile assessment. Section 3 presents the research method. Section 4 presents the experimental test and the preliminary results. Section 5 presents the findings and recommendations for future research.

2. Literature Review

2.1 Significance of Pavement Marking

Pavement markings are beneficial to drivers for their significance in keeping drivers on track and ensuring driving safety by conveying continuous information of "the intended travel path for short-range operations and the roadway alignment for long-range delineation" to drivers (Carlson et al., 2009). Recognizing the significance of pavement marking on driving safety, traffic agencies invest millions of dollars in maintaining the quality of pavement marking in the United States each year. Retroreflectivity is a widely accepted performance measure of pavement markings. To maintain the retroreflectivity of pavement marking at a serviceable level within a reasonable budget, researchers have devoted many efforts to set up minimum criteria for pavement marking retroreflectivity. Based on these research findings, FHWA has published minimum criteria of retroreflectivity through the new MUTCD and has required that all public traffic agencies design their own programs to maintain the retroreflectivity of pavement markings at or above the minimum level.

Engineers have been evaluating the effect of pavement marking on the improvement of driving safety ever since the 1940s. Most studies have concluded that pavement markings do improve driver/roadway safety. Centerlines, for example, as a major pavement longitudinal marking, have been proven effective in keeping drivers from shifting from the road's centerline to the right side of the road, as well as motivating them to drive, remarkably, slower (Taragin, 1947). As for edge lines, a before-and-after study in Connecticut has proven that the presence of edge lines can reduce accidents. Besides,

centerlines also have some influence on slowing driving speed (Williston, 1960). A similar study (Basile, 1962) in Kansas has observed a reduction in fatalities on two-lane, rural highways but no significant changes in the total number of injuries with the presence of edge pavement markings. However, these very early studies all adopted the simplest research methods, and no details of the research processes are available. These methods have limitations; for example, they cannot exclude the impact of other factors, such as weather, on roadway safety. Apparently, these studies failed to consider the differences in the retroreflectivity conditions of pavement markings in their analysis. A more recent experimental study (Tsyganov et al., 2006) adopted a before-and-after study approach to quantifying the effect of edge lines on improving safety on the rural two-lane highways in Texas. In the study, a statistical crash analysis found that "edge-line treatments on rural two-lane roadways may reduce accident frequency up to 26% and the highest safety impacts occur on curved segments of roadways with lane widths of 9 to 10 ft." In addition, the researchers also conducted stationary traffic observation and driving tests under different circumstances. However, an increase in speed by an average of 5 mph, or 9%, on both straight and curved highway segments, instead of a speed drop, has been observed in the tests. Even though there are some conflicts about the impact of the pavement marking on the driving speed, these studies all agree with the conclusion that the presence of pavement markings improves roadway safety, especially when the pavement marking conditions are particularly poor. Smadi et al. (2008) studied the crash records and established spatial correlations with the pavement marking retroreflectivity using data from the entire Iowa primary road system. Although the results from the entire dataset didn't show that lower pavement marking retroreflectivity correlates to a higher

crash rate, the results with only lower retroreflectivity (≤ 200 mcd/m^2/lux) indicates a statistical significant negative correlation with the crash rate (i.e., when the pavement marking retroreflectivity is lower than 200 mcd/m^2/lux, crash rates increase as the pavement marking retroreflectivity decreases).

2.2 Research on Minimum Retroreflectivity Level

Research has shown that pavement marking will work effectively as long as the pavement marking meets minimum criteria, but there is no proof that shows higher retroreflectivity would improve roadway safety. Therefore, it's necessary to determine the minimum level of acceptable retroreflectivity. In the 1990s, FHWA sponsored the development of a computer model, named Computer-Aided Road-Marking Visibility Evaluator (CARVE), to establish a set of criteria for minimum level of retroreflectivity based on the posted speed limits and the presence of retroreflective, raised pavement markers. Then, FHWA included the research results as part of an internal report and recommended criteria for minimum retroreflectivity. Many other institutes and DOTs have, also, developed general recommendations for the minimum criteria of retroreflectivity. Although these recommendations vary, they all ranged from 90 to 130 mcd/m^2/lux (Debaillon et al., 2008). With the application of new materials and technologies in pavement markings, these criteria are no longer applicable. The University of Iowa improved the CARVE to the Target Visibility Predictor in order to keep pace with new pavement marking technology. A study (Debaillon et al., 2008) adopted this new system and incorporated many new inputs to establish updated criteria. In this research, the researchers also included pavement surface type (e.g. asphalt or concrete) and vehicle type (passenger or freight). From this study emerged a set of

minimum pavement marking retroreflectivity recommendations for typical conditions on U.S. highways. However, this research only employs a deterministic model (i.e., TarVIP) based on illumination physics and vehicle/roadway geometry. None of the results were validated using actual field retroreflectivity measurements or drivers' perceptions. Therefore, further research is necessary to validate the results. Recently, the new MUTCD has published a recommendation for minimum retroreflectivity but does not require forceful implementation because of the budget concerns of DOTs.

The presence of pavement markings is essential for driving safety, and it is significant that the pavement markings should be at or above a minimum level (i.e., at least visible to the drivers). When retroreflectivity is low, driving risk can increase significantly. In order to maintain the visibility of pavement markings or keep the retroreflectivity above the minimum criteria, methods for quantitative measurement of retroreflectivity are needed by transportation agencies for pavement marking management. The following section reviews different retroreflectivity measurement methods.

2.3 Retroreflectivity Measurement Methods

Visual Inspection

Visual inspection is a measurement practice in which evaluators stand by the roadside or drive along the road to visually rate the condition of pavement marking retroreflectivity. Choosing to drive along the road could be a very efficient method that imposes little impact on the traffic. In addition, visual inspection is also very inexpensive because no special devices are needed, and only two engineers are needed for the whole evaluation process, one for driving and one for evaluating. However, human eyes are not reliable because, when they encounter different background contrasts, the brightness of an object

will look very different. Thus, one evaluator's rating results usually differ from another's because of expertise, experience, and subjectivity. Research conducted by the Texas Department of Transportation (TxDOT) (Benz et al., 2009) tried to correlate the visual inspection results with data collected by a handheld retroreflectometer and found that there was a correlation between the average rating of several evaluators (8 different evaluators in the research), but the correlation was sometimes inconsistent. However, when individual ratings were analyzed, significant biases from the handheld retroreflectometer readings appeared; even training seems to be ineffective in improving accuracy. Factors other than retroreflectivity, including facing angle, lighting condition, etc., may also impact the visual condition assessment (FHWA, 2007). In conclusion, even though visual inspection has advantages, it has the drawback of being qualitative and subjective and, therefore, cannot be used as a standardized method.

Handheld Retroreflectometer

Using handheld retroreflectometers provides a method for quantitative measurement of pavement marking retroreflectivity. A handheld retroreflectometer could be placed on a pavement marking and readings of the pavement marking retroreflectivity at each spot can be taken. Standard procedures are available for conducting measurement with a handheld retroreflectometer, and the operation can be handled with little training. With the standard operation, both the repeatability and reproducibility can be controlled. The price of a handheld retroreflectometer is usually $12,000 to $25,000 per unit depending on the model, but it is much less expensive than a mobile device ($80,000 without a van) (Benz et al., 2009), which will be discussed later. However, there are disadvantages to using a handheld retroreflectometer. The measurement operation may require lane

closures and unsafe situations because the operators are exposed to traffic. Most importantly, a large number of samples can be expensive to acquire and process. Since measurement of retroreflectivity with a handheld retroreflectometer is a standard operation, a brief introduction of available ASTM standards is provided in the following sections.

ASTM E1710-11 has proposed standard procedures for retroreflectivity measurement under dry conditions with a portable retroreflectometer at the standard "30 meter geometry" (FIGURE 4.1). In addition, this standard has also described a standard configuration of the retroreflectometer that all products should follow.

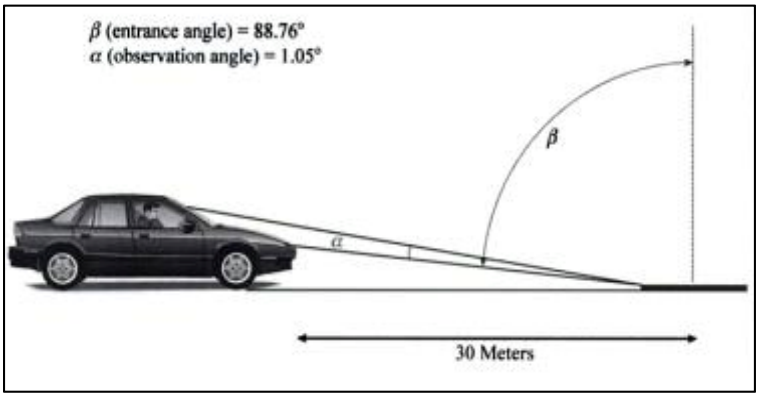


FIGURE 4.1: Standard geometry (Bernstein. 2000)

ASTM E2176-08, on the other hand, has provided standard procedures for retroreflectivity measurement under continuous wetting conditions with a retroreflectometer. This condition simulates a scenario in which it is raining (FIGURE 4.2).

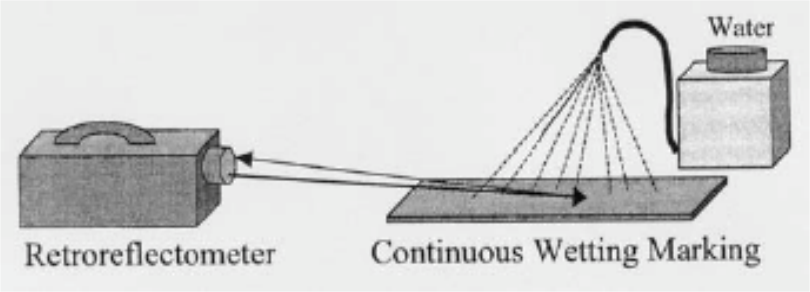


FIGURE 4.2: Illustration of continuous wetting measurement (ASTM E2176)

ASTM E2177-11 is available as standard procedures (i.e. wet recovery) for measuring retroreflectivity under wet conditions, which is similar to the scenario in which it has rained before measurement (FIGURE 4.3). However, it was claimed that this method does not necessarily reflect the condition of pavement marking after rain has fallen. A portable or mobile retroreflectometer could be used but should comply with the requirements for the apparatus described in ASTM E1710-11.

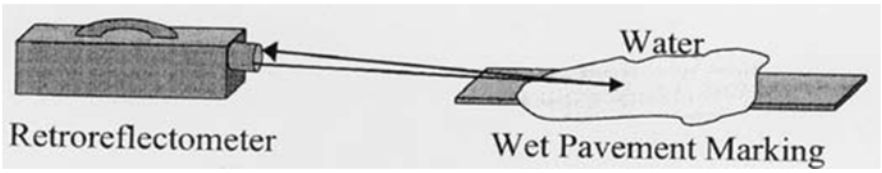


FIGURE 4.3 Illustration of wet condition (ASTM E2177)

Data collected following these standard procedures usually have good repeatability and reproducibility in practice. However, these standards have only described the application of a handheld retroreflectometer. Although a mobile retroreflectometer is also mentioned in ASTM E2177-11, it is not described in detail. In the following section, a few available mobile methods will be discussed.

Mobile Retroreflectometer

Using a mobile retroreflectometer provides a means and a method for network-level data collection. Usually, a mobile retroreflectometer system consists of a van, a mobile retroreflectometer, and a computer system for data storage. Retroreflectometers have some very attractive advantages. Mobile retroreflectometers are able to collect network-level data, and they can be operated at highway speed and do not disrupt the traffic. However, the cost of a mobile retroreflectometer is about four times that of a handheld retroreflectometer. Besides, the mobile retroreflectometers require more training and maintenance. In addition, the readings of retroreflectometers can be inconsistent for many reasons. For example, a study (Benz et al., 2009) found that the mobile retroreflectometer is very sensitive to two factors and could cause inconsistent data. The sensitivity analysis found that distance from the pavement marking to the sensors and the position across the measurement window have the most significant impact on the mobile retroreflectivity measurement. Unlike the handheld retroreflectometer, no standard operating procedures are available for mobile retroreflectometers. Thus, false calibration and operation can result in very inconsistent data and, ultimately, raise concerns about mobile retroreflectometers among the state DOTs.

LaserLux, as an example of mobile retroreflectometer available for DOTs since the 1990s, has been evaluated by the TxDOT (Benz et al., 2009). It was proven that many factors, such as the measuring of geometry, temperature, and pavement profile, had significant influence on the accuracy of the collected data. Even though there was a motion compensation system incorporated into the LaserLux, the errors still could not be removed. Similarly, other currently available mobile retroreflectometers all either have problems in accuracy or need a complex set-up and calibration process (Lee, 2011). The

National Cooperative Highway Research Program (NCHRP) Innovations Deserving Exploratory Analysis (IDEA) program (Lee, 2011) has recently proposed an advanced system to measure marking retroreflectivity at highway speed. The system under research has a tracking system to keep the measurement geometry automatically and a calibration system to simplify the process for calibration. Researchers claim that experimental tests have shown the accuracy of the system, but it still needs further evaluation and validation in the field.

Currently available mobile retroreflectometers are still not fully validated for implementation; thus, evaluation and improvement efforts are still ongoing. Besides, it is also worthwhile to explore the potential application of emerging technologies.

2.4 LiDAR Technology

LiDAR is an optical remote sensing technology that can measure the distance to or from other properties of a target by illuminating the target with light, often using pulses from a laser (Cracknell et al., 1991). The retro-value in LiDAR is the percentage of the redirected energy from the target divided by the emitted energy from LiDAR. The principle is consistent with FHWA's definition of retroreflectivity. A LiDAR-equipped vehicle has been used to collect pavement inventory data, but it has not been specifically designed for retroreflectivity measurement. However, since it applies a similar principle as a retroreflectometer, it would be promising to study the feasibility of measuring pavement marking retroreflectivity using LiDAR technology.

2.5 Summary

Pavement marking plays an important role in driving safety. The retroreflectivity condition is identified as the most important condition for pavement marking and draws major concerns from public transportation agencies because poor retroreflectivity has been proven to be strongly correlated with nighttime crash rates, especially under undesirable conditions, such as rain. Traditionally, state DOTs conduct manual condition assessment using visual inspection and handheld retroreflectometers. Evaluation by visual inspection is subjective and usually inconsistent. Assessment by handheld retroreflectometer is usually accurate and consistent with proper operation. However, it is so labor-intensive and time-consuming that it cannot fulfill the need for a continuous condition assessment covering the full network. It may, also, require extra resources for traffic control, and it may expose field engineers to open traffic. A good alternative is the mobile retroreflectometer, but available devices, such as the LaserLux, require a very complex process of calibration before each data collection. Any improper operation or setup could cause faulty data to be collected. In addition, many factors have proven to significantly impact the accuracy and consistency of mobile retroreflectometers. There is still a need to explore alternative mobile pavement marking retroreflectivity condition assessment methods. As the LiDAR system has already been used for traffic sign inventory and retroreflectivity condition assessment by utilizing the retro-intensity value that can measure the ratio of light redirected from an object to the light emitted from the laser, an alternative mobile method is possible for pavement marking retroreflectivity condition assessment.

3. Research Method

The objective of this study is to evaluate the feasibility of conducting pavement marking retroreflectivity condition assessment using LiDAR technology by establishing the preliminary correlation between the retroreflectivity readings measured by handheld retroreflectometers and the LiDAR retro-intensity values. The proposed research method is composed of four steps: test site selection, data collection, data preprocessing, and data analysis. FIGURE 4.4 shows the detailed procedures of this research.

- Test site selection: To establish a reliable correlation, the pavement marking retroreflectivity data should consist of complete coverage of different conditions. Thermoplastic was selected as the focused material in this study, as it is popularly used on state routes and interstates in Georgia and many other states. To include the complete coverage of different conditions, a nighttime visual inspection was first conducted to purposely include road sections containing different retroreflectivity conditions. Second, a preliminary retroreflectivity measurement using handheld retroreflectometer (at large measurement interval, e.g. 50 ft.) was conducted to further narrow down the road sections with different conditions. Third, detailed retroreflectivity measurements using a retroreflectometer (at small measurement intervals, e.g. 1 ft.) were conducted to collect the comprehensive pavement marking condition within each candidate section.

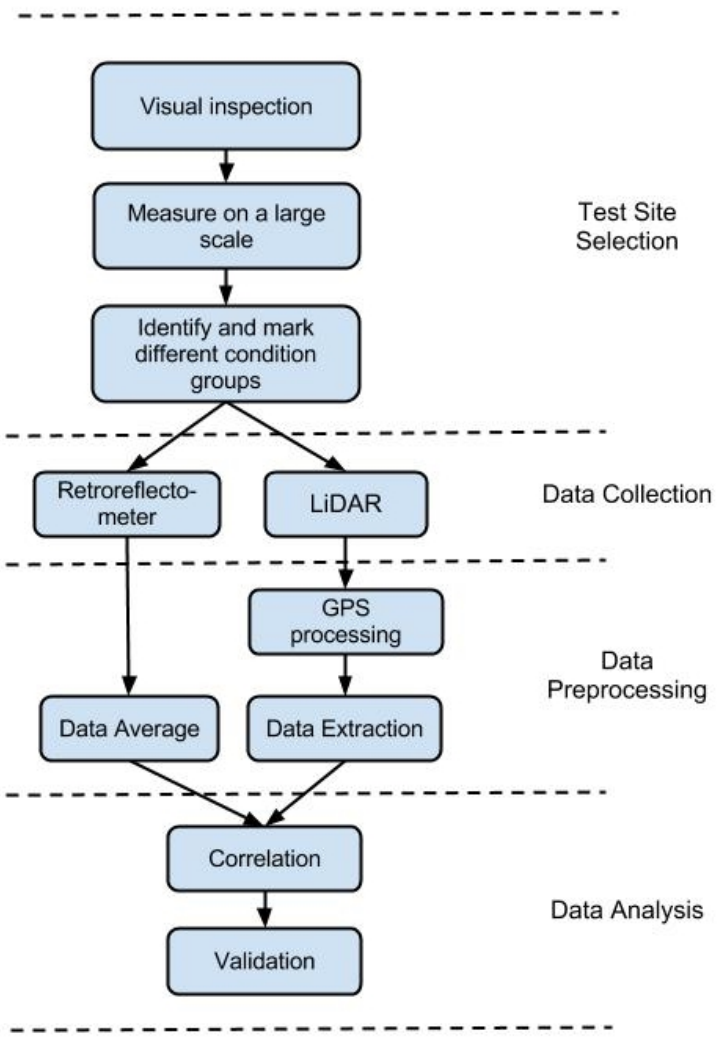


FIGURE 4.4 Flowchart of establishing correlation retroreflectivity and retro-intensity

- Data collection: To fully cover the selected road sections, three runs of LiDAR data were collected using the GTSV. The three runs of data were also used to assess the repeatability of the measurement. The corresponding retroreflectivity measurement using a retroreflectometer was also conducted right after the LiDAR data collection.
- Data preprocessing: The retroreflectometer readings are first averaged to obtain the retroreflectivity of each sample spot. The GPS of the LiDAR data are first

post-processed to obtain a higher GPS accuracy; then, the retro-intensity data are manually extracted from the LiDAR data using Trimble Analyst Software.

- Data analysis: Once both retroreflectivity readings and the corresponding retro-intensity values are collected, the correlation can be established based on their spatial association. The detailed procedures are discussed in the experimental test section.

4. Experimental Tests

Prior testing results show that pavement markings that are made of various materials have different retroreflectivity values. This research focuses mainly on thermoplastic and waterborne paint, the most popular pavement marking materials. After obtaining the LiDAR retro-intensity data and the retroreflectivity is measured at the same field locations, we conducted a regression analysis to explore the correlation between the LiDAR retro-intensity values and the retroreflectivity. With the correlation, critical values that differentiate acceptable pavement markings from unacceptable ones will be established for retro-intensity and used to assess the pavement markings' condition.

4.1 Data Collection

Data collection that covers retroreflectivity conditions from extremely bad to brand new involves the selection of a test site using the visual inspection method. Through this method, Ferst Drive (between State Street and Techwood Drive on the Georgia Tech campus) was selected for thermoplastic pavement marking testing. Hemphill Avenue and 17th Street were selected for waterborne paint pavement marking testing. To examine the detailed trend on the selected test sections, preliminary retroreflectivity measurements

were conducted using a *StripeMaster II* handheld retroreflectometer to get an idea of the distribution of retroreflectivity at the selected sites. Based on the trend, more detailed test sections were selected so that the retroreflectivity conditions were distributed evenly from bad to brand new. For thermoplastic and waterborne paint pavement markings, details about the test section selection are introduced as follows:

Thermoplastic

Thermoplastic pavement markings on Ferst Drive between State Street and Techwood Drive on Georgia Tech's Atlanta campus were installed by Georgia Tech facility management personnel. The thermoplastic pavement markings are of various conditions because of uneven aging and deterioration rates. FIGURE 4.5 shows the trend of the retroreflectivity along the selected road segment.

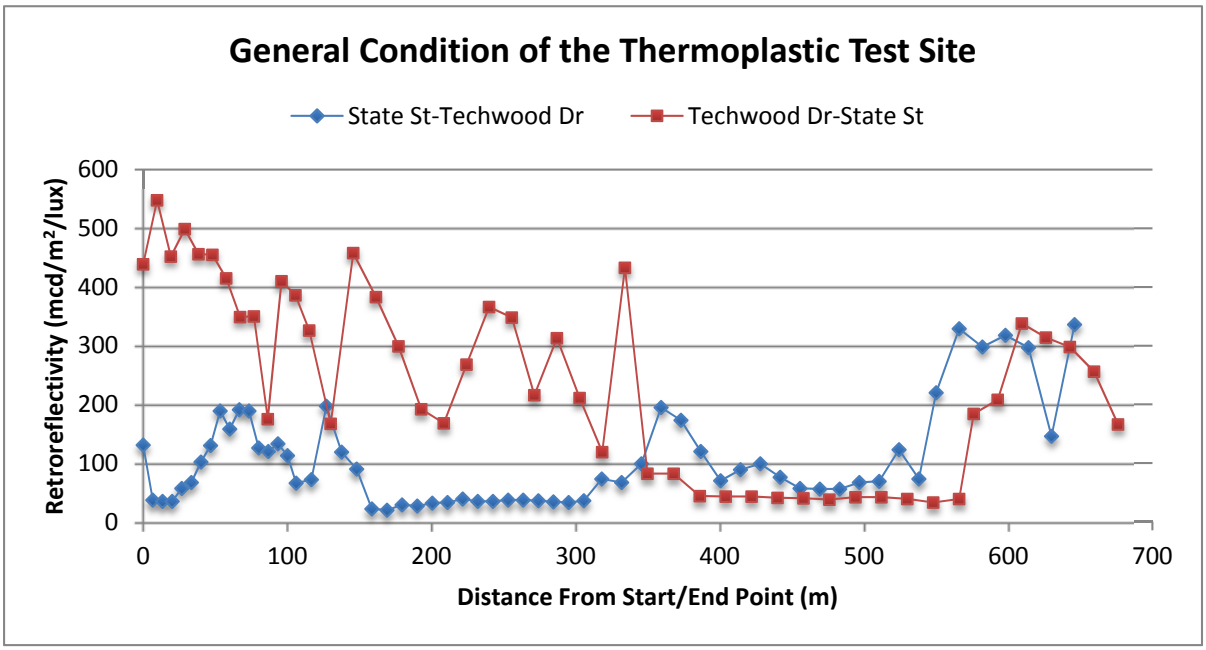


FIGURE 4.5: Pavement marking retroreflectivity on Ferst Dr. on Georgia Tech campus

It can be observed that the selected thermoplastic test site consists of varying retroreflectivity conditions ranging from 30 mcd/m^2/lux to more than 600 mcd/m^2/lux. Based on the information of the general condition in FIGURE 4.5, twelve test locations covering different ranges of retroreflectivity were selected. Each condition group was 10-ft in the longitudinal direction and contained 10 samples spaced one foot apart. To establish a reliable correlation, the retroreflectivity condition within each 1-ft section was maintained homogeneously. The detailed range for each condition group is shown in Table 4.1.

Table 4.1: Retroreflectivity of Thermoplastic Test Sections

Range (mcd/m^2/lux)	No. of Test Sections
0-50	3,10
50-100	1,4,5
100-200	2
200-300	7,12
300-400	6,11
400-500	9
500-600	10

The exact locations of different condition groups were spray-painted on the pavement to ensure that the location measurement from different times or from different sensors were consistent. After marking the samples in the field shown in FIGURE 4.6, the retroreflectivity for each sample was collected three times with a handheld retroreflectometer in compliance with the ASTM E1710-11 standard, and then the entire test site was surveyed by the GTSV three times.

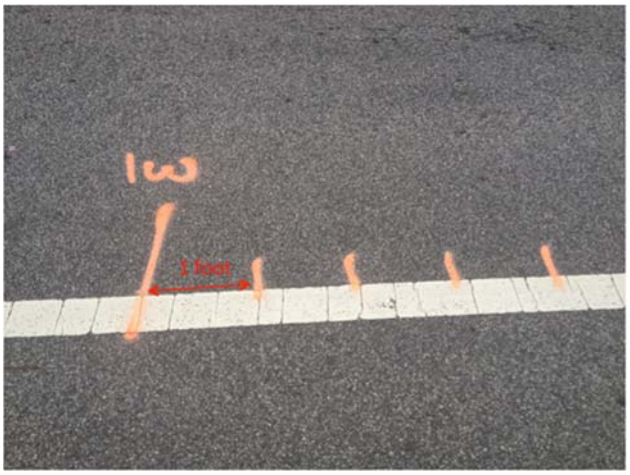


FIGURE 4.6: An example of selected thermoplastic test section

Waterborne Paint

Hemphill Avenue and 17th Street, shown in FIGURE 4.7, were identified as test sites for the road segments that were installed with waterborne paint markings under various conditions. We used an approach similar to the approach we had used to determine the test sections for thermoplastic markings. We first conducted a preliminary rough measurement of the pavement markings. FIGURE 4.8 shows the results.

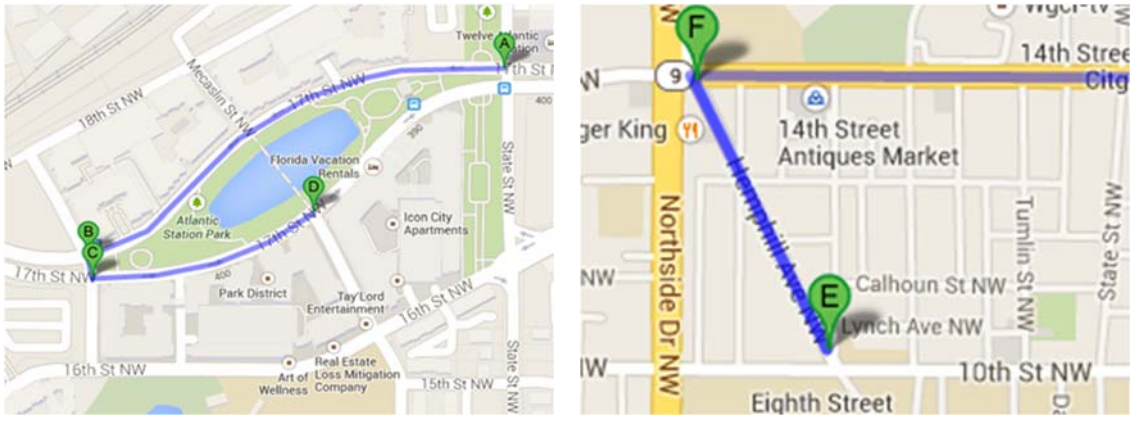


FIGURE 4.7 Test sites of waterborne paint markings

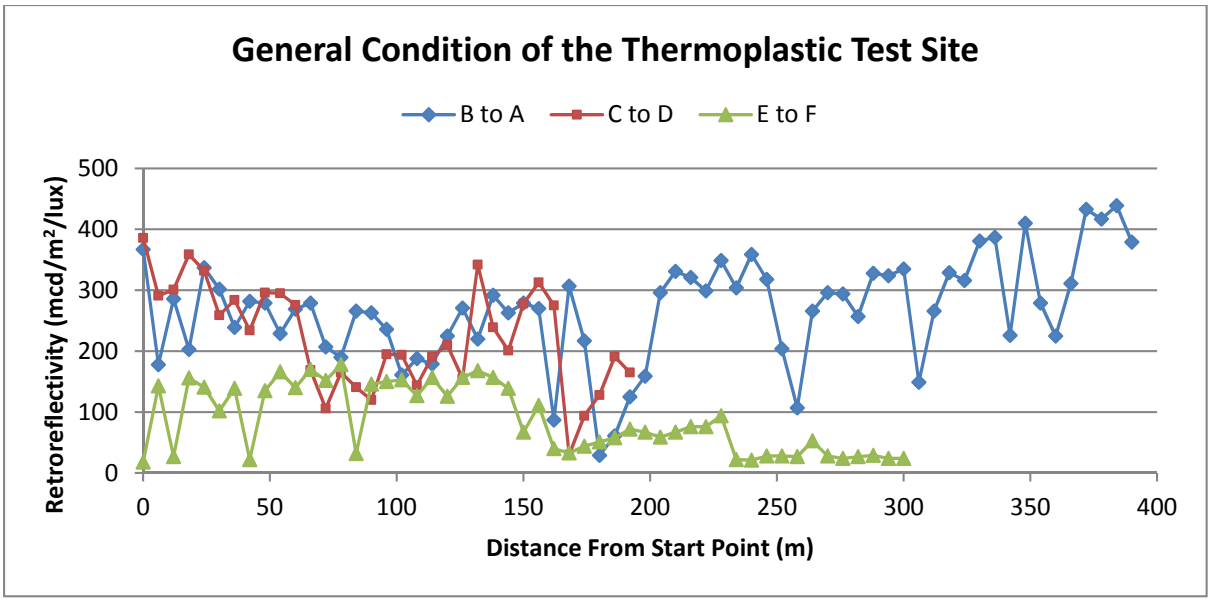


FIGURE 4.8 Pavement marking retroreflectivity condition on Hemphill Avenue and 17th Street

It can be observed that the selected test sites contain pavement markings with various conditions from 0 to 450 mcd/m^2/lux. Based on the trends on the road segments, 16 test sections were selected; eight sections were on Hemphill Avenue, and eight sections were on 17th Street. The markings selected on Hemphill Avenue are indicated by dashed lines, and each dashed line is segmented into samples 1 foot apart. On 17th Street, the test sections were also sampled one foot apart with 10 samples in each section. Table 4.2 shows the detailed distribution of conditions of the test sections.

Table 4.2: Retroreflectivity of Thermoplastic Test Sections

Range (mcd/m^2/lux)	No. of Test Sections
0-50	1, 2
50-100	3, 4, 7 ,8
100-150	5, 6, 11
150-250	12, 13
250-350	9,10,14,15
>350	16

4.2 Data Extraction

Trimble Trident 3D Analyst software was used to extract the data. All the samples were extracted into a Geographic Information System (GIS) layer through the video log images that were synchronized with the mobile LIDAR, shown as the green line in FIGURE 4.9. Then, the layer was mapped onto the calibrated LiDAR point-cloud, shown in FIGURE 4.10. At each sample location, retro-intensity values of the LiDAR points were manually extracted, shown in FIGURE 4.10. As the LiDAR device used in this study (i.e. Riegl LMS-Q120i) can acquire 10,000 points per second, at least three retro-intensity values could be obtained to associate with each sample section (i.e., a 1-ft section in the longitudinal direction). For example, the yellow rectangle shown in FIGURE 4.11 represents the actual 1-foot test section. Although there are ten LiDAR points cast over this section, only the middle point fully cast on the pavement marking can be used for establishing correlation(the highlighted point in red circles shown in FIGURE 4.11). Consequently, for each 1-foot section, approximately three retro-intensity values are associated. For the points that partially cast over the pavement marking section, the retro-intensity values might not reflect the true retroreflectivity condition because the energy was redirected by both the pavement and pavement marking.

FIGURE 4.9: Create a New Layer to Locate Sample Site in Videolog Image

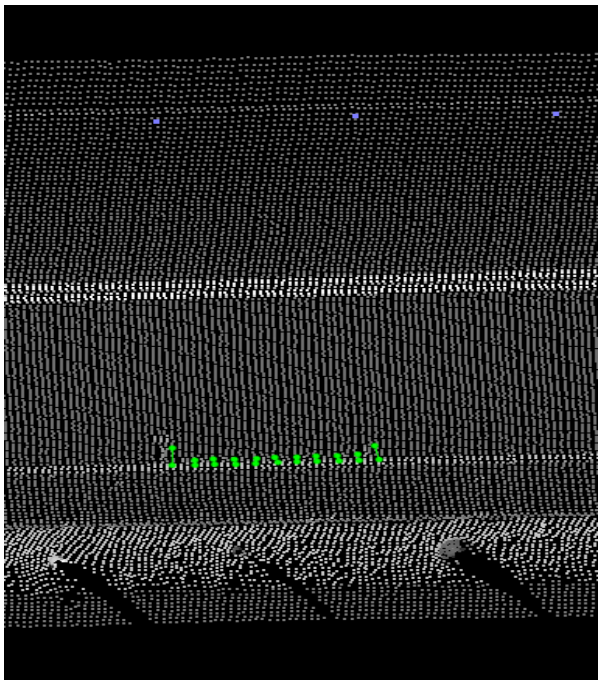


FIGURE 4.10: New Layer Mapped onto Point-cloud

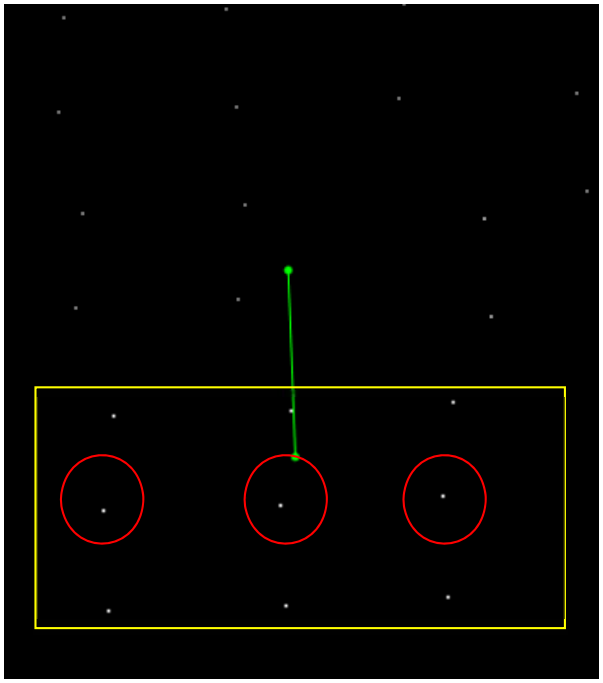


FIGURE 4.11: Selected Points Included in the Analysis

4.3 Data Analysis

Handheld Retroreflectometer Measurement

To show the general condition of the pavement markings in each test section, FIGURE 4.12 and FIGURE 4.13 illustrate the average retroreflectivity and standard deviations. FIGURE 4.12 shows that thermoplastic sections with average retroreflectivity above 200 mcd/m^2/lux, where the standard deviations at the sections with high retroreflectivity, are significantly larger than the standard deviations of the sections with low retroreflectivity. One possible explanation for the large standard deviation is because of various traffic, weather, and sunshine conditions; the condition of the pavement marking deteriorates at a dynamic rate at different locations, so the condition of the pavement markings is not uniform across the 10-foot test section. However, when the retroreflectivity of the

pavement marking is below 100, the condition of the entire section is uniformly below the standard.

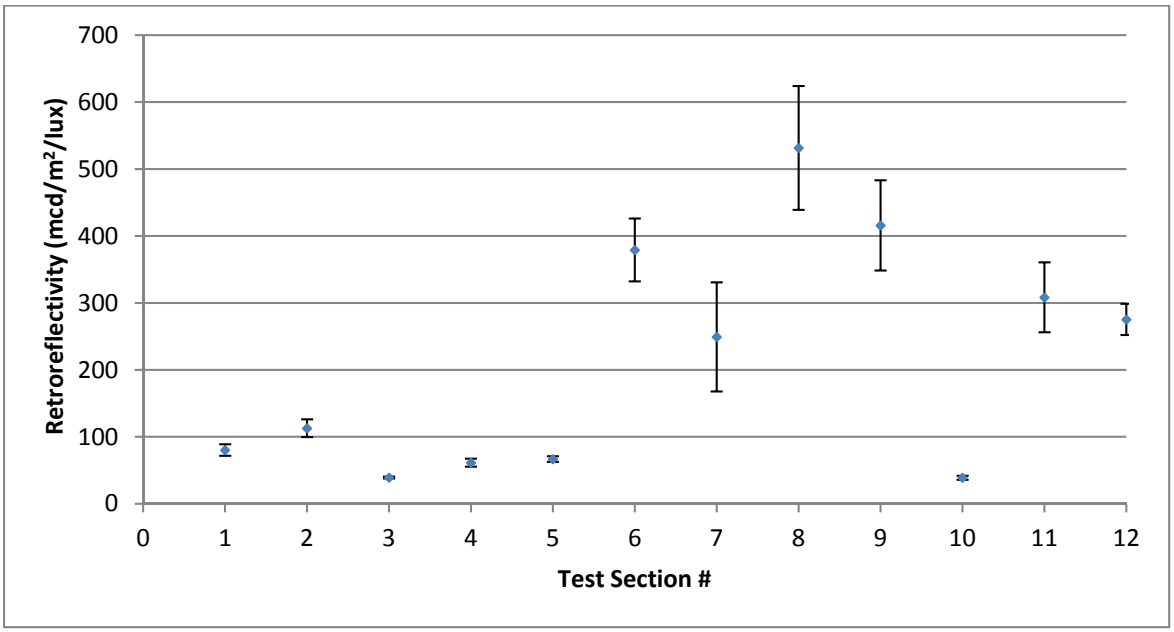


FIGURE 4.12: Average retroreflectivity on test sections thermoplastic markings

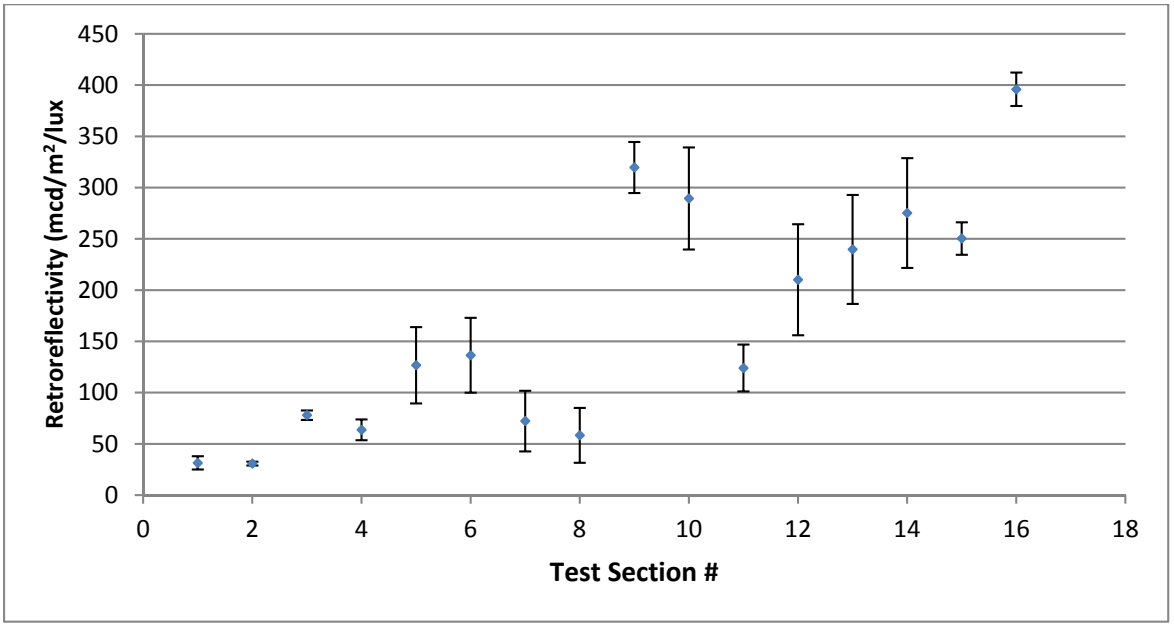


FIGURE 4.13: Average retroreflectivity on test sections waterborne paint markings

From FIGURE 4.13, it can be determined that waterborne paint shows much lower retroreflectivity than thermoplastic because even the retroreflectivity of the new waterborne paint is only around 400 mcd/m^2/lux. In addition, a similar trend can also be observed for waterborne paint; that is, the retroreflectivity above 200 mcd/m^2/lux shows the largest variation.

LiDAR Measurements

Repeatability

To assess the pavement marking condition consistently, the LiDAR system should be able to measure retro-intensity with good repeatability. To evaluate repeatability, we plotted the retro-intensity values of all of the thermoplastic test sections for the three runs in FIGURE 4.14.

The three runs by the GTSV overlapped quite well except for some of the sections with large variations, such as Sections 2 and 7. Closer examination reveals that the three runs have a maximum standard deviation of 0.0154 and an average standard deviation of 0.0044, which still holds good repeatability over all the sections. The retro-intensity of three runs for waterborne paint shows similar results.

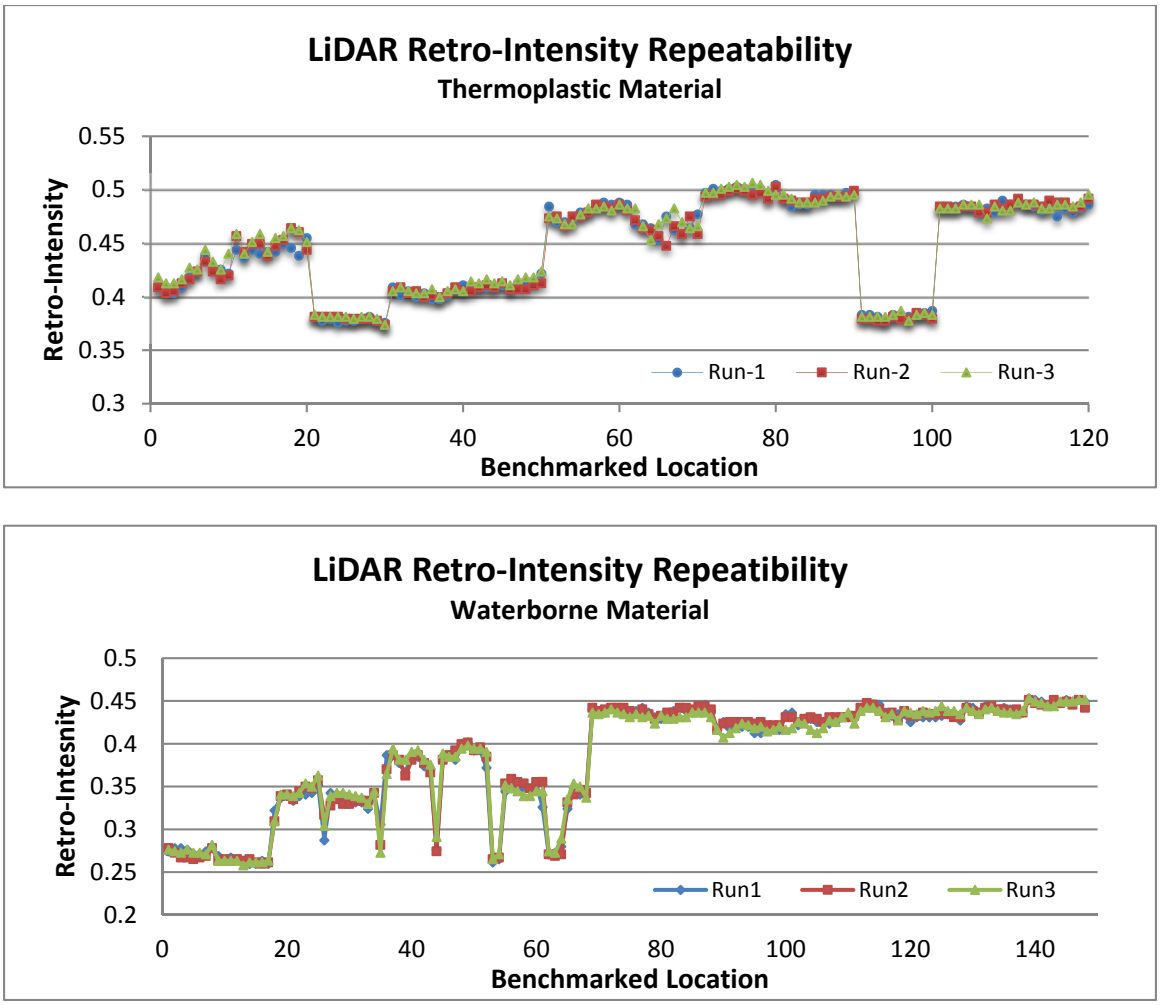


FIGURE 4.14: LiDAR retro-intensity repeatability

Sensitivity to Temperature

To assess the sensitivity of the LiDAR retro-intensity values to the temperature, a field test was conducted on I-16, which is near the Georgia Tech Savannah campus. The ambient temperature was collected at each hour for six consecutive hours (i.e. 9:00am – 3:00pm). Two locations with different retro-intensity values were selected to measure the LiDAR retro-intensity values, including a pavement marking section with high retro-intensity value and a concrete pavement surface section with low retro-intensity.

FIGURE 4.15 shows the results of the sensitivity. Both the pavement marking section (with high retro-intensity values) and the non-pavement marking section (with low retro-intensity values) show only a small variance (0.0041 and 0.0039, respectively) as the temperature changed between 88.3°F to 119.5°F.

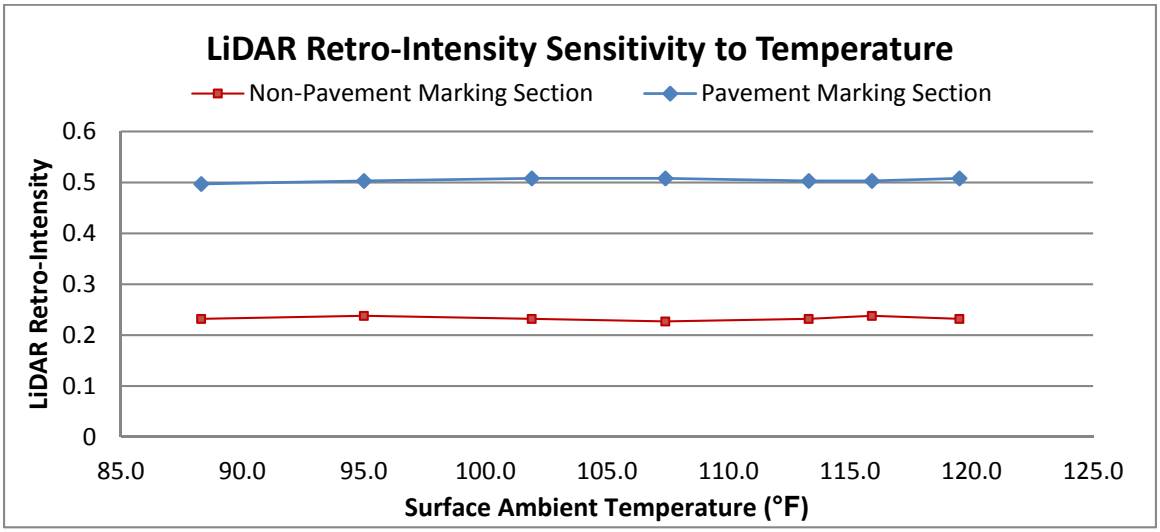
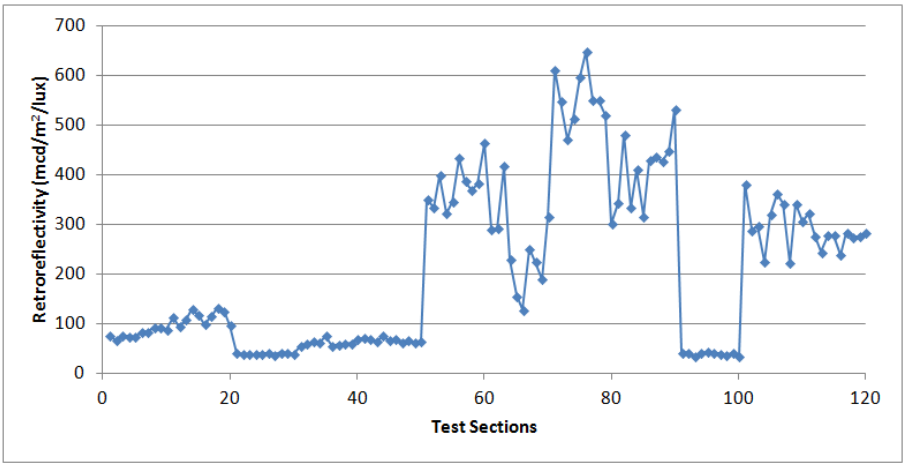


FIGURE 4.15: LiDAR retro-intensity sensitivity to ambient temperature

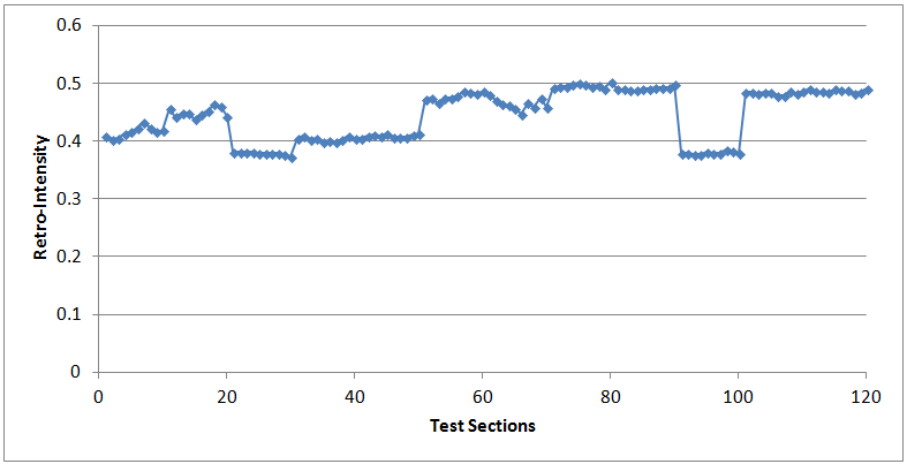
Based on the results derived from the sensitivity study, unlike the mobile pavement marking retroreflectivity measurement device, the LiDAR retro-intensity values are not sensitive to the ambient temperature. For example, the Florida Department of Transportation reported that "*changes in temperature were observed to have a significant impact on measured retroreflectivity (using LaserLux)*" (Fletcher et al., 2007). The insensitivity to temperature is a very important feature of the mobile LiDAR for establishing a reliable correlation between the retro-intensity and retroreflectivity in the following sections. It is also critical to ensure that the LiDAR-based pavement marking retroreflectivity condition assessment can produce a consistent outcome under different temperature conditions in practice once the correlation is established.

Correlation Establishment

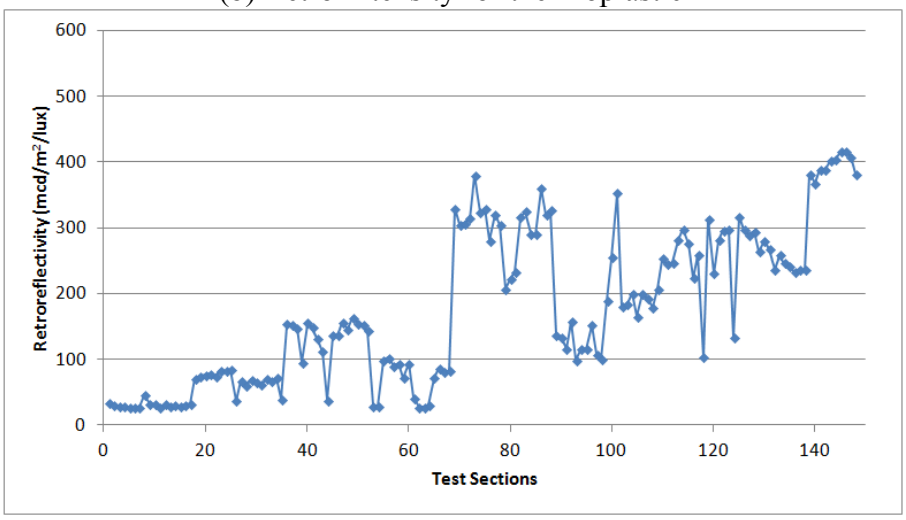
FIGURE 4.16 (a) to (d) shows the collected retroreflectivity and retro-intensity data for thermoplastic and waterborne paint materials. It can be determined that the retro-intensity values correlate well with retroreflectivity from condition group to condition group. Within the test sections, especially those with large retroreflectivity (over 200 mcd/m^2/lux), the retroreflectometer readings show a bigger variation due to the randomness of individual measurements at each single location, while the LiDAR retro-intensity readings are continuously measured using the GTSV and shows more consistent values with smaller variations.



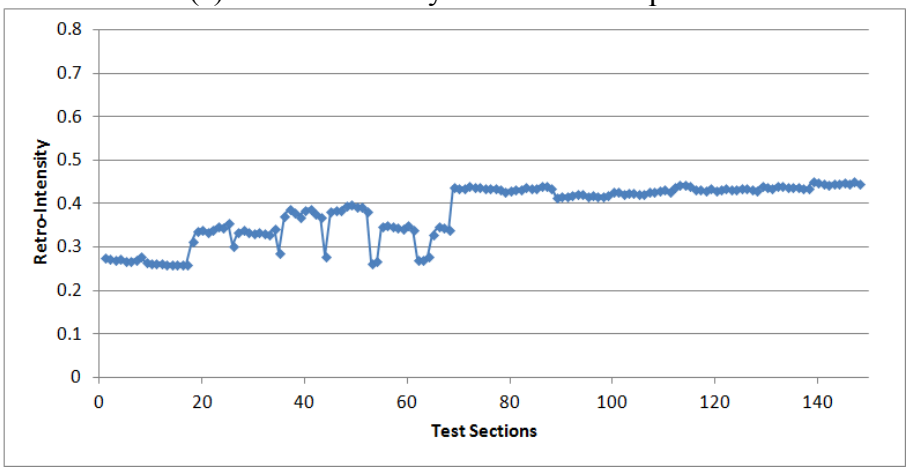
(a) Retroreflectivity for thermoplastic



(b) Retro-intensity for thermoplastic



(c) Retroreflectivity for waterborne paint



(b) Retro-intensity for waterborne paint

FIGURE 4.16: Retroreflectivity and retro-intensity for thermoplastic and waterborne paint materials

Regression using an exponential function was conducted with data from the 12 thermoplastic test sections and 16 waterborne paint test sections. The results of the regression analysis are shown in FIGURE 4.17 and FIGURE 4.18. These results do not indicate any deterministic models between the retroreflectivity and the retro-intensity values from LiDAR, although high R^2 values are observed, which indicates strong correlations.

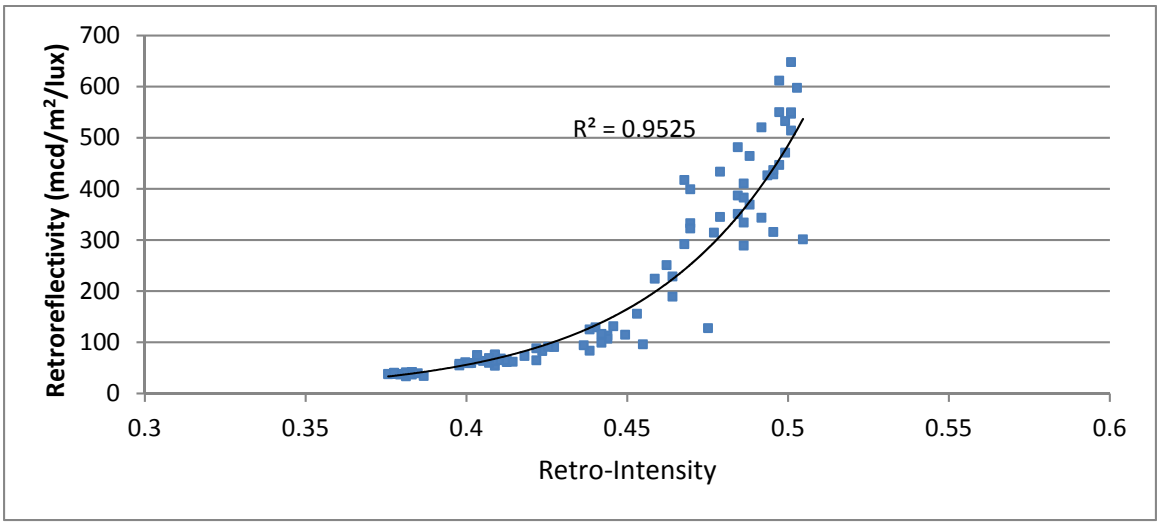


FIGURE 4.17: Correlation between retroreflectivity and LiDAR retro-intensity of thermoplastic material

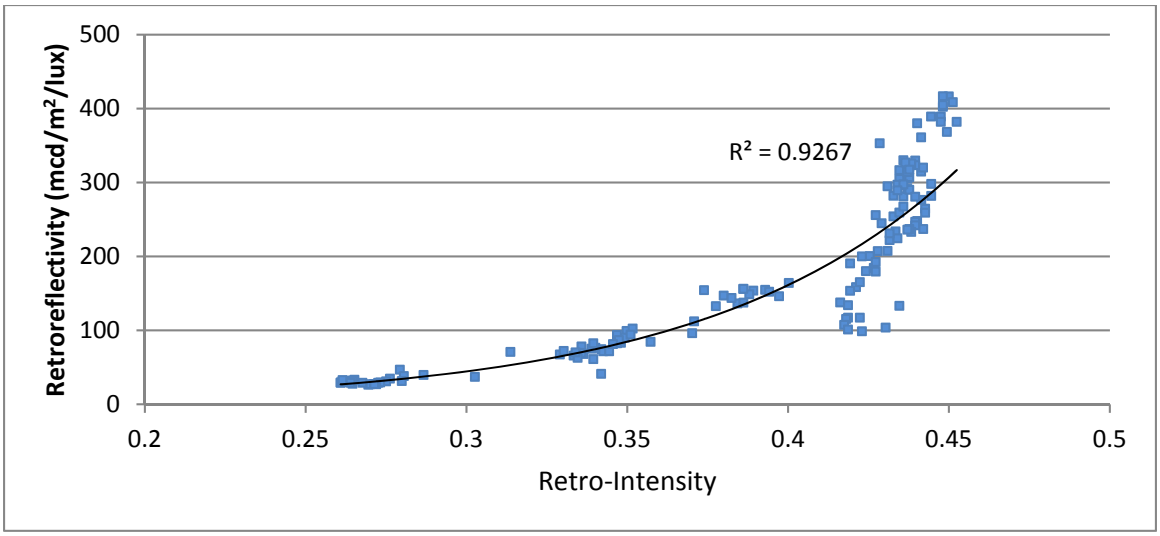


FIGURE 4.18: Correlation between retroreflectivity and retro-intensity of waterborne paint material

FIGURE 4.17 and FIGURE 4.18 show clear exponential relationships between the handheld retroreflectometer measurement and the LiDAR retro-intensity values for both thermoplastic and waterborne paint materials. The R-squares are as high as 0.9525 and 0.9267, respectively. Even though the R-square for the waterborne paint is as high as 0.9267, the middle-ranged test sections (150-250 mcd/m^2/lux for waterborne paint) show an obvious deviation from the general trend. Instead of lump summing all the samples to form a single correlation, further investigations were conducted to study the cause of such a deviation. FIGURE 4.19 shows the deviation of the two selected locations for waterborne material; the blue color represents the data collected on the Hemphill Street, whereas the red color represents the data collected on 17th Street. Although both locations have waterborne paint-based pavement markings, it is suspected that the different bead formula (i.e. bead shape and diameter distribution) could have an impact of the correlation results.

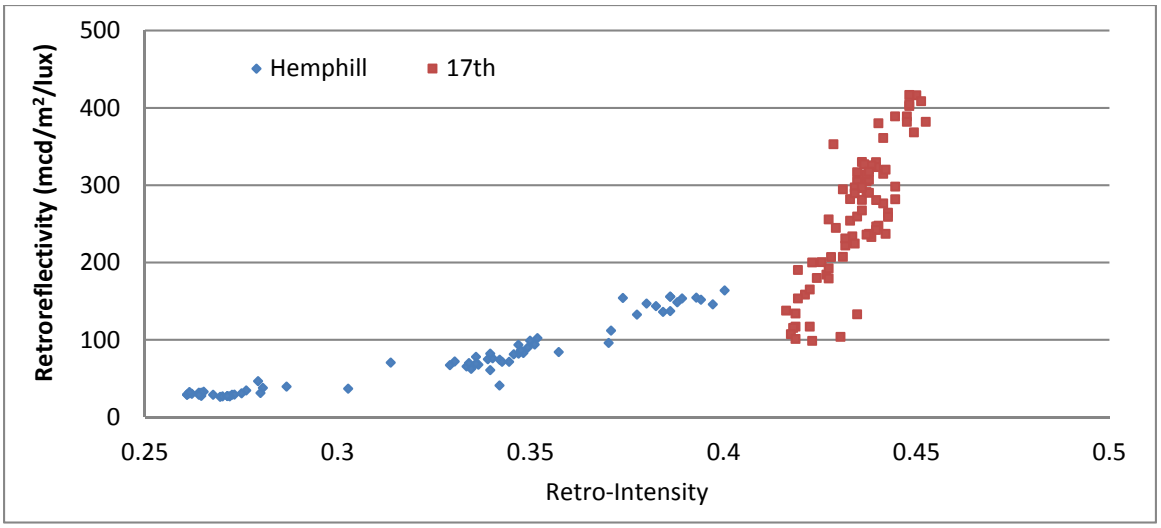
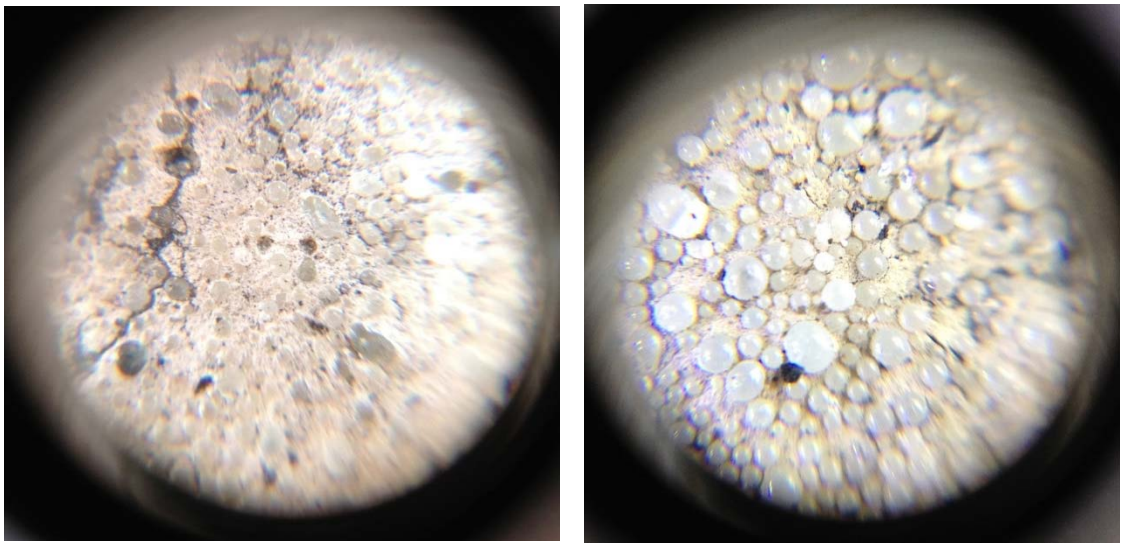


FIGURE 4.19: Different retro-intensity responses from different road sections

The test sections on Hemphill Street and 17th Street were revisited to capture the microscopic images of the pavement markings. FIGURE 4.20 shows two examples of the bead observed on Hemphill Street and 17th Street, respectively. There is some bead formula difference observed in both of the sections. Based on the observation of the microscopic images, it is suspected that Hemphill Street used Type I bead only, while 17th Street used Type I and Type IV bead mix with a 4:1 or 3:1 ratio. The roundness of the Type IV glass beads collected on 17th Street is better than the ones on Hemphill Street. In addition, the glass bead density treated on 17th Street is much higher than Hemphill Street, which is a good explanation of the general retroreflectivity differences. There is a need to further investigate the impact of different bead formulas to establish the correlation between the retro-intensity values from mobile LiDAR and the retroreflectivity from the retroreflectometer.



(a) (b)

FIGURE 4.20: Observation of microscopic glass beads on a) Hemphill Ave. and b) 17th Street

Although the derived correlations still require further validation, these correlations are promising for developing an automated pavement marking retroreflectivity condition assessment in the future. For example, assuming the derived correlation curve between the LiDAR retro-intensity and retroreflectivity were validated, if the critical value of retroreflectivity is 100mcd/m^2/lux for separating pavement marking retroreflectivity condition, the condition assessment of pavement marking using the corresponding critical retro-intensity value could be conducted. The detailed analysis results in Table 4.3 show that if the experimental tests are performed on thermoplastic independently many times, 95% of the mean retro-intensity will fall between 0.4035 to 0.4505 when the mean retroreflectivity for a sample spot is 100mcd/m^2/lux. If the experimental tests are conducted on waterborne paint, the 95% confidence interval will be 0.2973 to 0.4264 with a mean retro intensity of 0.3561 for the retroreflectivity of 100mcd/m^2/lux. These

critical retro-intensity values can be used to determine the threshold separating good pavement markings from poor ones.

Table 4.3: Critical Retro-intensity Value

Material	Retroreflectivity	Retro-Intensity	Confidence Level	Confidence Interval	
Thermoplastic	100mcd/m^2/lux	0.4263	95%	0.4035	0.4505
Waterborne paint	100mcd/m^2/lux	0.3561	95%	0.2973	0.4264

5. Conclusions and Recommendations

This study focuses on exploring the feasibility of conducting a pavement marking retroreflectivity condition assessment using LiDAR technology by establishing the correlation between the retroreflectivity values measured by handheld retroreflectometer and the retro-intensity value acquired from a LiDAR point cloud. In this study, as the most commonly used materials of DOT maintained highway and local roads, thermoplastic and waterborne paint, were used to establish the correlation. The same procedure can be followed for establishing the correlation with other pavement marking materials. For thermoplastic, Ferst Drive on the Georgia Tech campus was selected as the test site to conduct the data collection; it consisted of 12 10-foot test sample sections and 120 individual 1-foot testing sample sections. The retroreflectivity from the selected thermoplastic test sections ranged from around 30mcd/m^2/lux to about 600mcd/m^2/lux and covered the typical range for thermoplastic material from newly built to being completely deteriorated. For waterborne paint, Hemphill Avenue and 17th Street were selected as the test site for data collection. On Hemphill Avenue, eight dash lines and on 17th Street, eight 10-foot test sections were measured with the *StripeMaster II* Retroreflectometer. The retroreflectivity from the selected waterborne paint test sections

ranged from around 30mcd/m^2/lux to about 400mcd/m^2/lux. The following summarizes the findings of this study:

- Retro-intensity values acquired from mobile LiDAR are not sensitive to ambient temperatures, having an average standard deviation of less than 0.0041. The retro-intensity acquired from mobile LiDAR has good repeatability on the tested thermoplastic and waterborne materials with an average standard deviation of 0.0044.
- There is an exponential correlation between retroreflectivity and retro-intensity with an R-square of 0.9525 for thermoplastic and 0.9267 for waterborne paint.
- The correlation between retroreflectivity and retro-intensity might be sensitive to different bead formulas of the pavement marking material. Separate correlation curves might be needed not only for different pavement marking material category, e.g. thermoplastic, waterborne, etc., but also for different bead formulas under the same material category, e.g. different bead formulas, etc.
- Based on the correlation results, the preliminary retro-intensity threshold corresponding to the minimum retroreflectivity (100mcd/m^2/lux) defined in the MUTCD can be determined as 0.4263, with a 95% confidence interval ranging from 0.4035 to 0.4505 for thermoplastic and 0.3521, with a 95% confidence interval ranging from 0.2973 to 0.4264. Using the established correlations, a mobile LiDAR-based pavement marking retroreflectivity condition assessment method can be further developed.

To achieve such a mobile method, the following recommendations for future research are suggested:

1) To validate the preliminary results using additional data collected with different bead formulas.
2) To validate the established correlations for thermoplastic and waterborne pavement markings with additional data collected by both the GTSV and a retroreflectometer.
3) To extend the experimental test to other pavement markings materials (e.g., tape, polyuria) following similar procedures proposed in this study.
4) To validate the critical retro-intensity values that correspond to the minimum pavement marking retroreflectivity standards required by the MUTCD and transportation agencies by comparing with both a handheld retroreflectometer and nighttime visual inspection.
5) To develop an automatic method for extracting pavement marking retro-intensity data from the LiDAR point cloud to streamline the condition assessment as proposed in this study.

References

Bahar, G., Masliah, M., Erwin, T., Tan, E., and Hauer, E. (2006). "Pavement Marking Materials and Markers: Real-World Relationship between Retroreflectivity and Safety over Time." NCHRP Web Document, Vol. 92, 206p.

Basile, A. J. (1962). "Effect of Pavement Edge Markings on Traffic Accidents in Kansas." Highway Research Board Bulletin.

Benz, R.J., Pike, A.M., Kuchangi, S.P., and Brackett, Q. (2009). "Serviceable Pavement Marking Retroreflectivity Levels: Technical Report." Texas Transportation

Institute, Texas Department of Transportation, Federal Highway Administration, 150p.

Bernstein, R. (2000). "Summary of Evaluation Findings for 30-Meter Handheld and Mobile Pavement Marking Retroreflectometer." Highway Innovative Technology Evaluation Center (HITEC), Washington, DC USA.

Carlson, P.J., Park, E.S., and Andersen, C. K. (2009). "Benefits of pavement markings: a renewed perspective based on recent and ongoing research." Transportation Research Record: Journal of the Transportation Research Board, Vol. 2107, pp. 59-68.

Cracknell, A.P., and Hayes, L. (1991). "Introduction to remote sensing." CRC press.

Debaillon, C., Carlson, P.J., Hawkins, H.G., He, Y., Schnell, T., and Aktan, F. (2008). "Review and Development of Recommended Minimum Pavement Marking Retroreflectivity Levels." Transportation Research Record: Journal of the Transportation Research Board, Vol. 2055, pp. 71-77.

Fletcher, J., Philpott, A., Choubane, B., and Holzschuher, C. (2007). "Characterization and Mitigation of Temperature Sensitivity within Mobile Retroreflectometer Unit." Transportation Research Record: Journal of the Transportation Research Board, Vol. 2015, pp. 91-102.

Harrigan, E.T. (2006). "Pavement Marking Materials and Markers: Testing the Relationship between Retroreflectivity and Safety." NCHRP Research Results Digest, Vol. 305, 6p.

Holzschuher, C., and Philpott, A. (2006). "Mobile Retroreflectivity Unit (MRU) Surveying of Maintenance Rating Program (MRP) Sites". Florida Department of Transportation, State Materials Office.

Lee, T. (2011). "Advanced Methods for Mobile Retroreflectivity Measurement of Pavement Marking." NCHRP-IDEA Program Project Final Report, 19p.

Smadi, O., Souleyrette, R.R., Ormand, D., and Hawkins, N.R. (2008). "Pavement Marking Retroreflectivity: Analysis of Safety Effectiveness." Transportation Research Record: Journal of the Transportation Research Board, Vol. 2056, pp. 17-24.

Taragin, A. (1947). "The Effect on Driver Behavior on Center Lines on Two-Lane Roads." Highway Research Board Proceedings.

Tsyganov, A.R., Machemehl, R.R., Warrenchuk, N. M., and Wang, Y. (2006). "Before-After Comparison of Edge Line Effects on Rural Two-Lane Highway." University of Texas, Austin, Texas Department of Transportation, Federal Highway Administration.

Williston, R. M. (1960). "Effect of Pavement Edge Markings on Operator Behavior." Highway Research Board Bulletin.

Chapter 5 Long-term Monitoring of Crack Deterioration - A Pilot Study on SR 26

1. Introduction

Cracking is one of the predominant types of pavement distresses, mainly caused by vehicle repetitive load and pavement weathering. Pavement crack characteristics, including crack position (wheelpath or non-wheelpath), orientation (transverse or longitudinal), length, width, density, etc., and deterioration behaviors (e.g. crack growth in length and width and changes in branch pattern) are essential for effective pavement management. Modeling and forecasting the pavement crack conditions of a specific pavement section provides valuable information to transportation agencies for proper planning of maintenance and rehabilitation activities, budget estimation, and resource allocation. Acquiring critical insight into pavement cracking deterioration behaviors is also crucial to support the validation of the MEPDG.

A considerable number of efforts have been made to model pavement crack deterioration behaviors in literature. Many existing studies (Prozzi & Madanat, 2000; Hu, et at., 2012) used lab simulation data or controlled experimental test data, such as the American Association of State Highway Officials (AASHO) Road Test, which was an accelerated loading experiment (HRB, 1962). These studies suffered from several problems, including the fact that they were developed only for specific climate conditions and sub-grade materials, and they used predetermined values for some parameters (Reger, et al., 2013). Therefore, it is difficult to use controlled experimental outcomes to fully model the actual pavement crack deterioration behavior due to the complicated roadway

conditions in the real-world environment. On the other hand, data from actual in-service pavement sections subjected to the combined actions of highway traffic and environmental conditions are the most representative of the actual deterioration process (Prozzi & Madanat, 2004). Some studies used the data from state departments of transportations' (DOT) annual pavement condition surveys, such as Florida DOT's crack survey data (Yang, et al., 1974; Nasseri, et al., 2009; Thomas & Sobanjo, 2013) and Washington state's PMS (Reger, et al., 2013). Loizos and Karlaftis (2005) used a data set collected from in-service pavements in 15 European countries. However, these data collected in the field, tend to have a higher number of measurement errors than data collected in controlled experiments, as the data from the field are usually collected through visual roadside estimations or windshield surveys. Furthermore, these data are mostly composite condition ratings and are reported on a per-mile or higher basis, which loses the critical granularity of detailed crack deterioration behaviors.

The advances in sensor technology makes it possible to collect high-quality pavement surface data with 3D laser systems. These data have been proven robust in crack detection (Tsai & Li, 2012) and enable the extraction of detailed information: crack length, width, orientation, etc. (Tsai, et al., 2014). This has created a valuable opportunity for accurately and reliably collecting detailed pavement crack data in the field to study and gain insight into pavement characteristics/patterns and crack behaviors in a real-world environment. The objectives of this study are 1) to evaluate the feasibility of using detailed crack data gathered in the field to support the study of multi-scale crack deterioration behaviors (e.g. individual single/branched cracks, and clustered sections, such as 100-ft., 1-mile, etc.) to identify the vulnerable and robust pavement sections/spots

with high and low deterioration rates; and 2) to identify the potential issues for the analysis of long-term monitoring data. Georgia State Route 26 was selected as a pilot site because it has had 3D pavement surface data collected on it for over the past four years. The results are presented and discussed in the subsequent contexts.

2. Experimental Setup

FIGURE 5.1 shows the selected project for this pilot study on Georgia State Route 26 / U.S. 80 between Milepost 5.5 and Milepost 11.5, which has, primarily, flexible pavements with dense graded asphalt surface as the top course. According to GDOT's traffic count data, the selected project had an AADT of 24,020, and a truck percentage of 11.94% in 2013. It was resurfaced in 2004, when the surface course and part of the asphalt base course were replaced, but constant deterioration has been observed ever since.

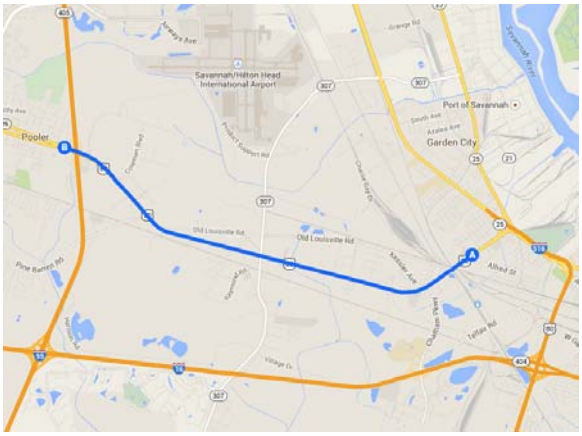
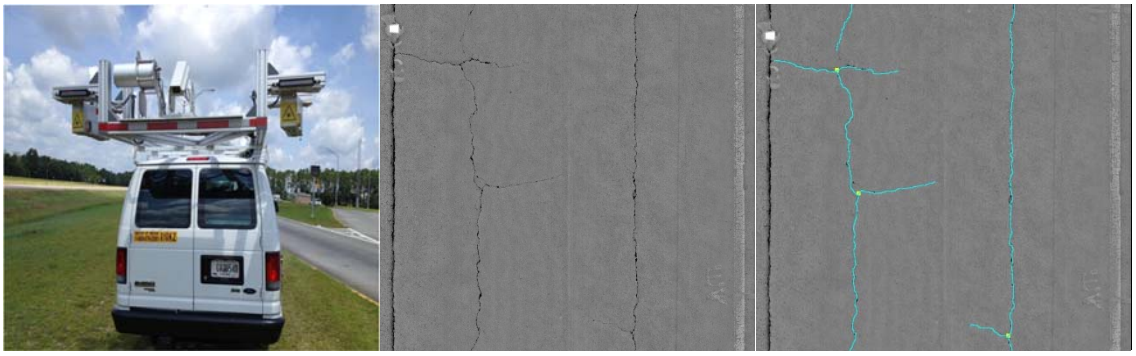


FIGURE 5.1: Selected project for crack deterioration analysis on GA State Route 26

In this pilot study, the pavement surface data collected using GTSV on five different timestamps from October 2011 to December 2013 was used. After the data acquisition,

the pavement surface data are represented in the format of range images, where each pixel on the image describes the elevation of the pavement surface. The crack detection is then conducted using the algorithm proposed in our previous study (Kaul, et al., 2012). FIGURE 5.2 shows the 3D laser imaging system and an example of pavement range image and the corresponding detected crack map.



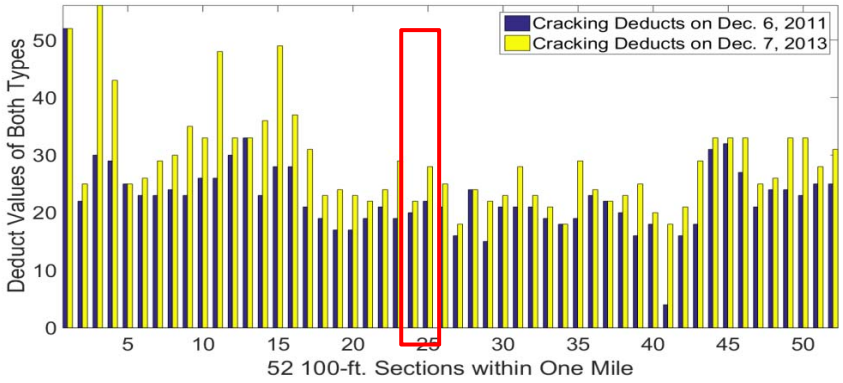
(a) 3D Laser Imaging System (b) Range Image (c) Detected Crack Map

FIGURE 5.2: Illustration of data acquisition and crack detection

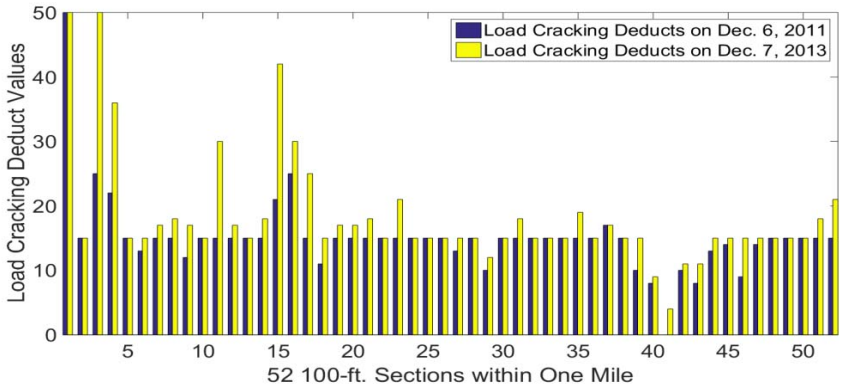
3. Analysis at One-Mile Level

The one-mile segment between Milepoints 7 and 8 is selected here as an example for demonstration purposes. According to the COPACES database, which records the annual manual survey results conducted by GDOT pavement engineers, the overall composite condition rating of this specific segment drops from 60 to 53 over the same period of time. However, such a composite rating is clearly insufficient to reveal where and how the pavement deterioration happens in detail. With the proposed crack classification method (Tsai, et al., 2014), we can now compare the changes between the two timestamps side by side on different crack types, including their severity levels.

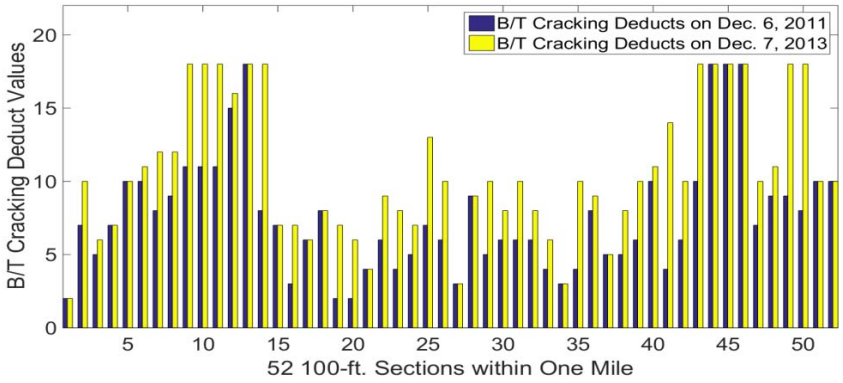
FIGURE 5.3 shows the comparison between crack deduct values between Dec. 6, 2011, and Dec. 7, 2013, across all 52 100-ft. sections on the selected one-mile segment. Each 100-ft. section is represented by two consecutive bars; the blue one on the left is the deduct value for Dec. 6, 2011, and the yellow one on the right is the deduct value for Dec. 7, 2013. FIGURE 5.3 (a) shows the crack-related deduct values composed from both load cracking and block/transverse (B/T) cracking, which are the two predominant crack distresses in Georgia. According to FIGURE 5.3 (a), the crack condition in some sections deteriorated significantly, especially at the beginning of this mile, while it remained almost unchanged in some other sections. FIGURE 5.3 (b) and FIGURE 5.3 (c) show the deterioration behaviors for two crack types, load cracking, and B/T cracking, respectively. It can be observed that in most 100-ft. sections, no significant deterioration was observed on the load cracking, while the deduct values of B/T cracking changed quite a bit. This further reveals that the crack deterioration on most of this mile is thermal related rather than load related.



(a) Comparison on both crack types combined



(b) Comparison on load cracking



(c) Comparison on B/T cracking

FIGURE 5.3: Crack condition comparison on the one-mile segment between Dec. 6, 2011 and Dec. 7, 2013

4. Analysis at Detailed Level Using a Crack Fundamental Element Model

Using crack types and the corresponding deduct values is one alternative to interpreting the crack condition on the roads. Since the detected crack maps are generated at multiple timestamps, the changes of more fundamental crack properties can be investigated.

Crack pattern, together with other crack characteristics, is crucial for differentiating crack types and severity levels in transportation agencies' pavement survey practices. Tsai, et al. (2014) proposed a multi-scale CFE model to not only topologically represent crack

patterns, but also to provide rich crack properties at three different scales (fundamental crack properties, aggregated crack properties, and CFE-clustered geometrical properties) to support the crack condition analysis. As introduced in Chapter 2, FIGURE 2.1 briefly illustrates the logic of multi-scale crack properties extraction.

Using this logic, different crack properties, such as crack length, junctions, and polygons, etc., can be extracted from the detected crack maps, and the changes of these properties can be quantitatively evaluated between different timestamps. A 200-ft. section near Milepoint 7.5 in the negative direction is selected (as highlighted in FIGURE 5.3), and the changes of different properties within this selected section are presented in the subsequent contents.

5.1 Crack Length

Crack length is one of the most basic properties used to describe a crack pattern. Many state DOTs use crack length to report the extents for different crack types. In this study, we measure the arc lengths of the detected crack curves within the selected sections. FIGURE 5.4 shows the change in total crack length, which is the sum of the lengths of all detected crack curves. It can be observed that the crack deterioration led to a 30% increase in crack length over the two-year period.

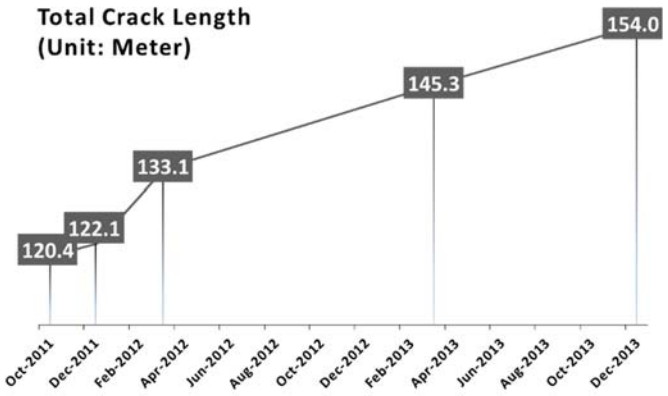


FIGURE 5.4: Change in total crack length across five different timestamps on the selected 200-ft. section near milepoint 7.5

The analysis in crack length can also be linked with other crack properties, such as crack orientations. FIGURE 5.5 shows the comparison between crack growth along the longitudinal direction (within 30 degrees of the driving direction) and other directions. It shows that the crack growth along other directions is more significant than the one along the longitudinal direction, which indicates that long and continuous longitudinal cracks already existed in both wheelpaths, and the crack deterioration was more of a branching out from these cracks rather than a further extension of them. This is also consistent with the more significant deterioration with B/T cracking (as shown in FIGURE 5.3).

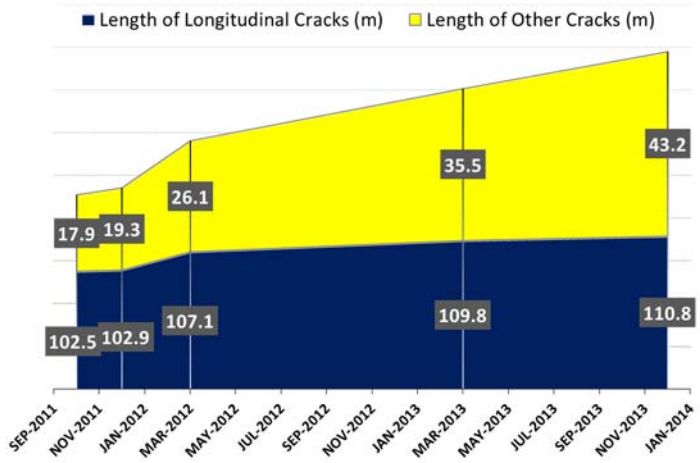
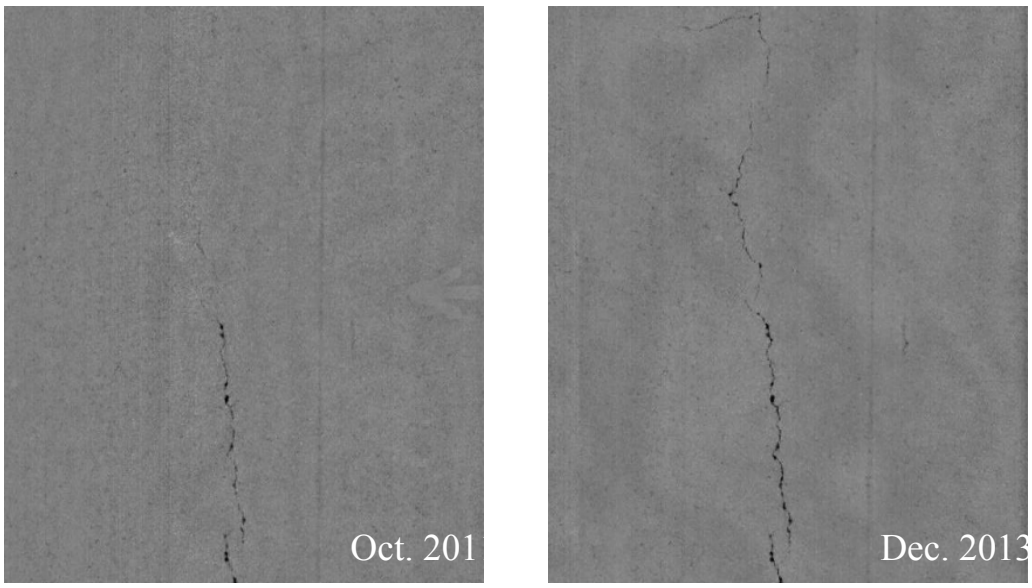


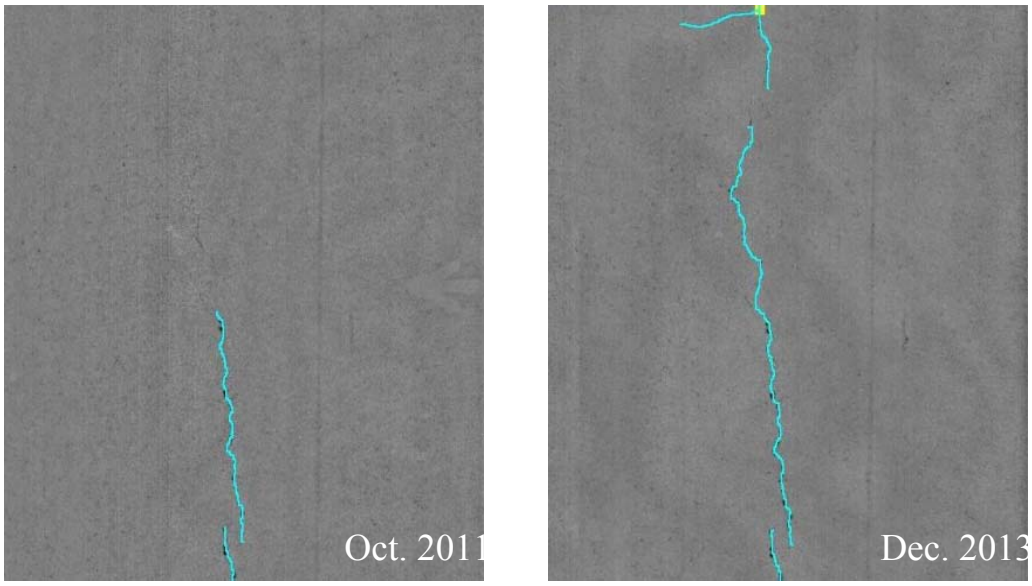
FIGURE 5.5: Crack deterioration comparison between longitudinal and other directions on the selected 200-ft. section near milepoint 7.5

The precise locations of these deterioration behaviors can be identified by comparing the detected crack maps. FIGURE 5.6 and FIGURE 5.7 present two examples of crack growth along longitudinal and transverse directions correspondingly.

Moreover, the analysis in different crack properties can be location-based. Many state DOTs emphasize the crack condition survey inside wheel paths, as they are expected to deteriorate more rapidly than the outside areas. FIGURE 5.8 shows the change of crack length inside and outside the wheel paths across five different timestamps. It can be observed that the majority of cracks lie inside the wheel paths, and crack deterioration mainly happens inside the wheel paths, as well.

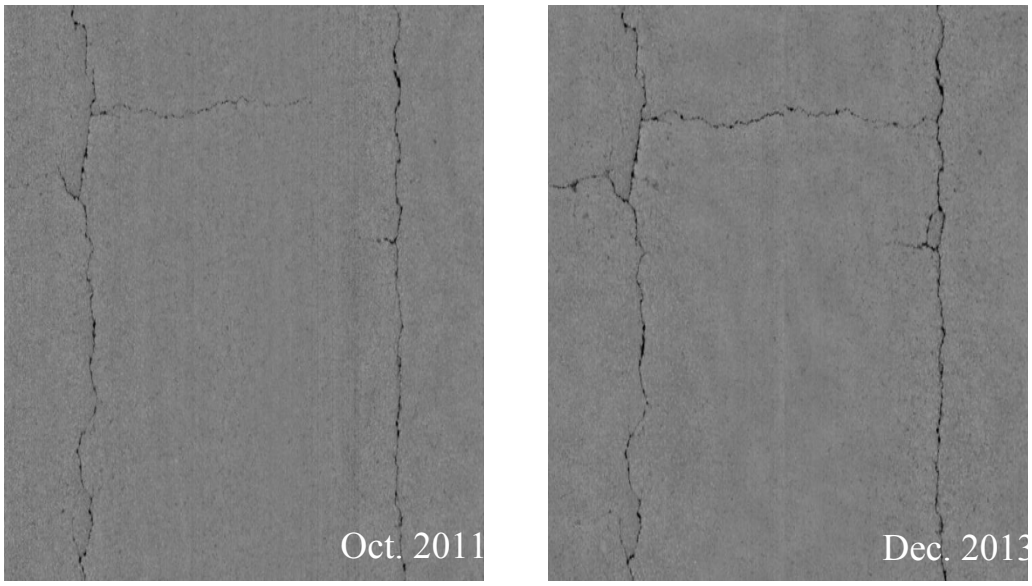


(a) Comparison between range images

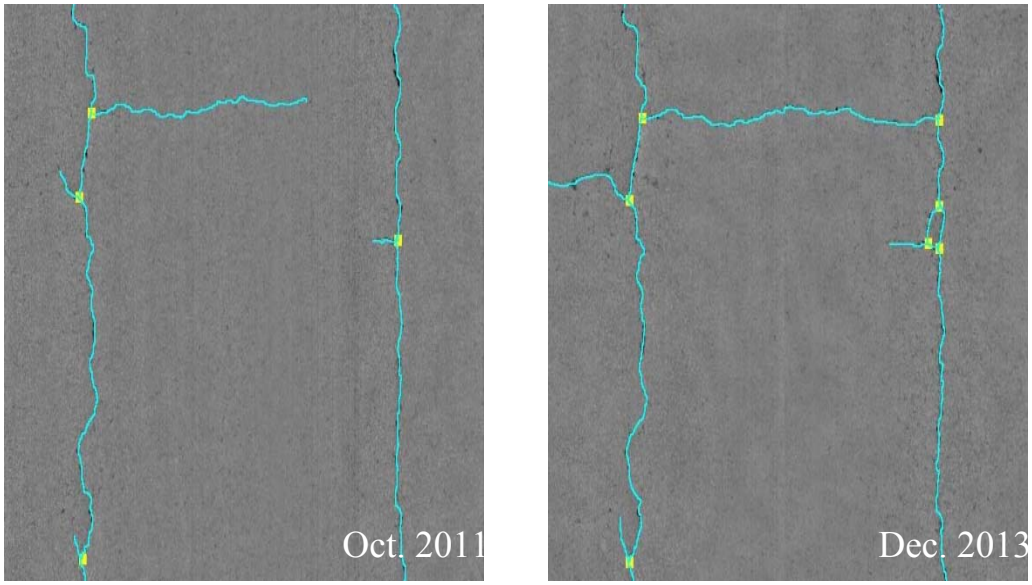


(b) Comparison between detected crack maps

FIGURE 5.6: Example of crack deterioration along the longitudinal direction on the selected section near milepoint 7.5

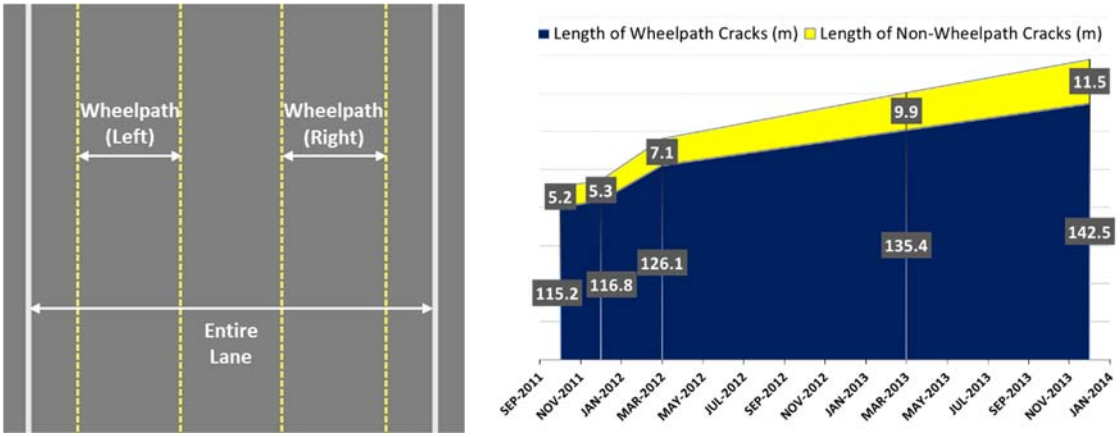


(a) Comparison between range images



(b) Comparison between detected crack maps

FIGURE 5.7: Example of crack deterioration along the transverse direction



(a) Illustration of wheelpath (b) Change of crack length

FIGURE 5.8: Crack deterioration comparison inside and outside the wheelpaths across five different timestamps on the selected 200-ft. section near milepoint 7.5

5.2 Crack Intersection

Crack intersections locate the crossing points of the predominant crack curves. The number of crack intersections indicates the complexity of the crack pattern and the severity of the crack condition. Crack intersections are usually introduced by the branching out or connection of existing crack curves. As shown in FIGURE 5.9, the number of crack intersections has increased dramatically over the two-year period, which confirms that the branching out of existing cracks was the major crack deterioration behavior during that period of time. FIGURE 5.10 presents an example in which crack branching out is clearly observed.

5.3 Crack Polygon

Many crack curves intersect with each other and form an enclosed piece-like area, which is referred to as a crack polygon. The number and extent of each crack polygon is an important property to differentiate crack types and severity levels. For example, the presence of an extensive number of small crack polygons is usually referred as “alligator

cracking," which indicates high severity load-related cracking and potential structural failure. In the study, the number of crack polygons within the selected sections is counted across different timestamps (as shown in FIGURE 5.11). The crack polygons started to appear in the second half of the given period, but the number was still limited. A further investigation on the detected crack maps shows that these distributed polygons are isolated, so their existence doesn't significantly impact the crack type and severity level within the selected sections. FIGURE 5.12 shows an example of how the crack polygons are formed through crack deterioration.

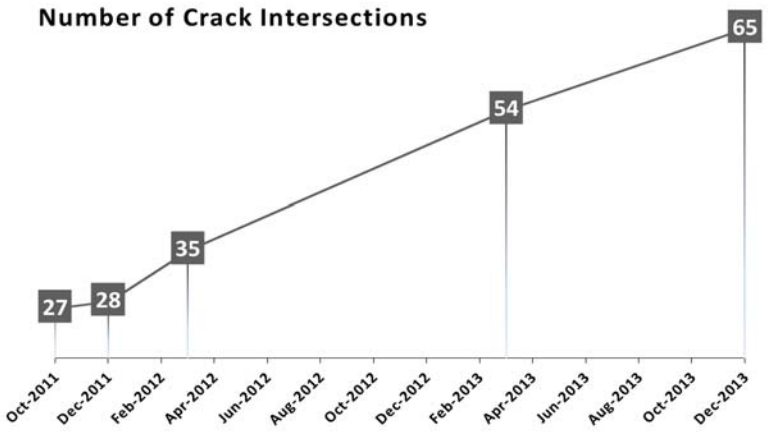
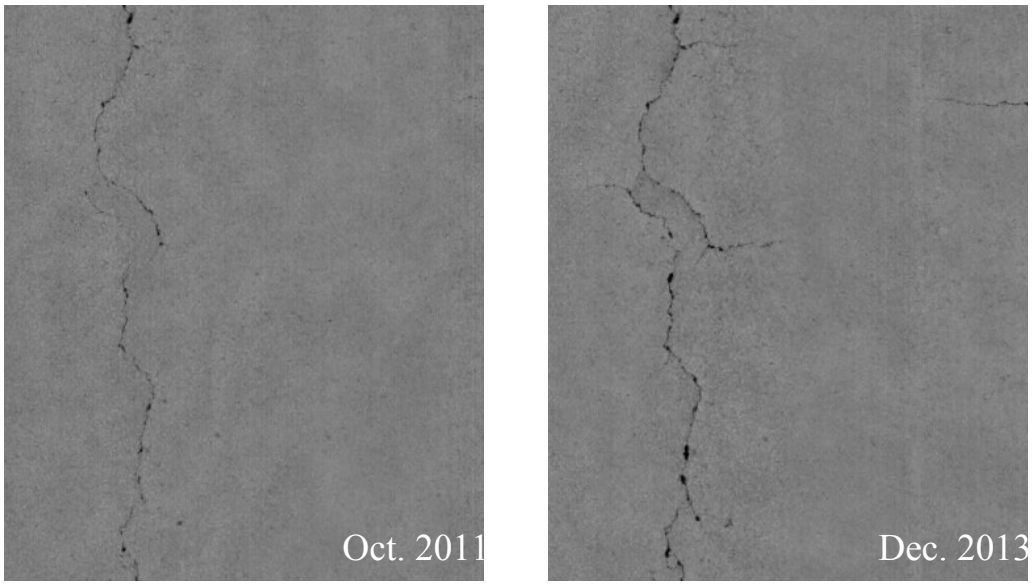
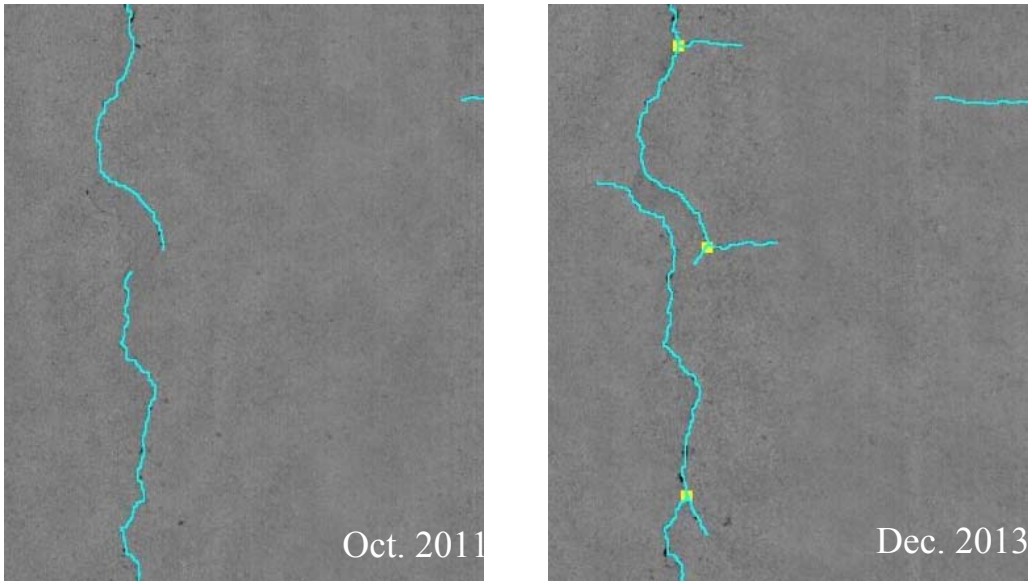


FIGURE 5.9: Change in number of crack intersections across five different timestamps on the selected 200-ft. section near milepoint 7.5



(a) Comparison between range images



(b) Comparison between detected crack maps

FIGURE 5.10: Example of crack branching out (crack intersections are marked as yellow dots)

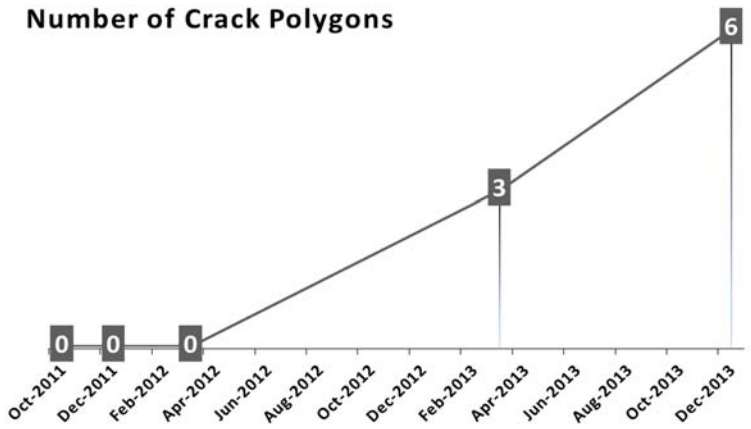


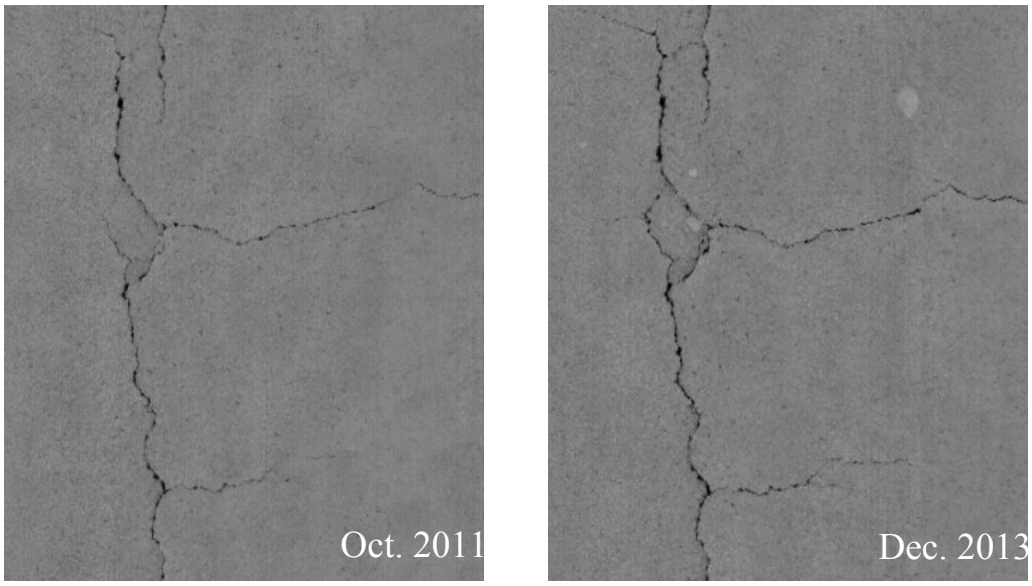
FIGURE 5.11: Change in number of crack polygons across five different timestamps on the selected 200-ft. section near milepoint 7.5

6. Issues Related to Crack Deterioration Analysis Using Long-term Monitoring Data

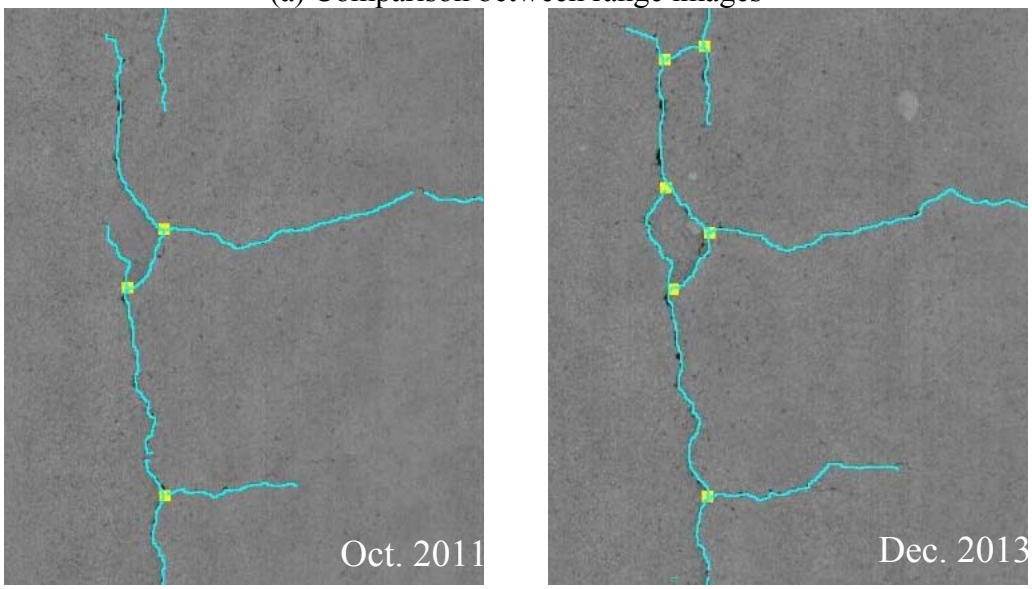
The existing issues and potential challenges, which are expected to benefit future research in this field, are identified based on the experience from this study and discussed below.

First, a method for data registration between multiple timestamps needs to be developed. In this study, multiple timestamps are manually registered together based on the markers on the road. As shown in FIGURE 5.13, in the longitudinal direction, the data from different timestamps are aligned together using the pre-labeled markers on the edge of the road. This requires these markers to be painted to label the terminus of the survey section before data collection. It becomes difficult to conduct the study at a larger scale (e.g. in a state-wide survey). Therefore, using an automatic data registration method based on GPS

and existing pavement features, such as cracks and pavement markers, is recommended for future research.



(a) Comparison between range images



(b) Comparison between detected crack maps

FIGURE 5.12: Illustration of forming crack polygons

Second, the data variability needs to be thoroughly evaluated. On one hand, the survey areas are not always consistent between multiple timestamps, and data completeness is not always guaranteed. Most pavement data acquisition systems are able to cover one lane width only if the vehicle is driven in the middle of the lane. As shown in FIGURE 5.13, in the transverse direction, due to the inconsistent driving behaviors in different surveys, the coverage cannot be guaranteed to be the same. The survey in Mar. 2012

covered the entire lane width, while a small part near the left lane marking was missed in the survey in Dec. 2011. To ensure a fair comparison, the survey area needs to be shrunk in the transverse direction. The issue may become more significantly difficult for large-scale research that covers significant amounts of time. Therefore, standardized data collection procedures need to be developed. On the other hand, the inherent data variability needs to be studied so that it is not confused with the external crack deterioration behaviors.

Third, research is needed to understand diverse internal and external factors of crack deterioration behavior. The study presented in this section focuses on monitoring and quantifying the changes. The next stage utilizes these data in the pavement management applications, including forecasting deterioration behaviors and enhancing pavement design. It requires large-scale research and, also, depends on the data availability in many other aspects, such as traffic volume, environmental factors, design, construction, existing maintenance, etc.

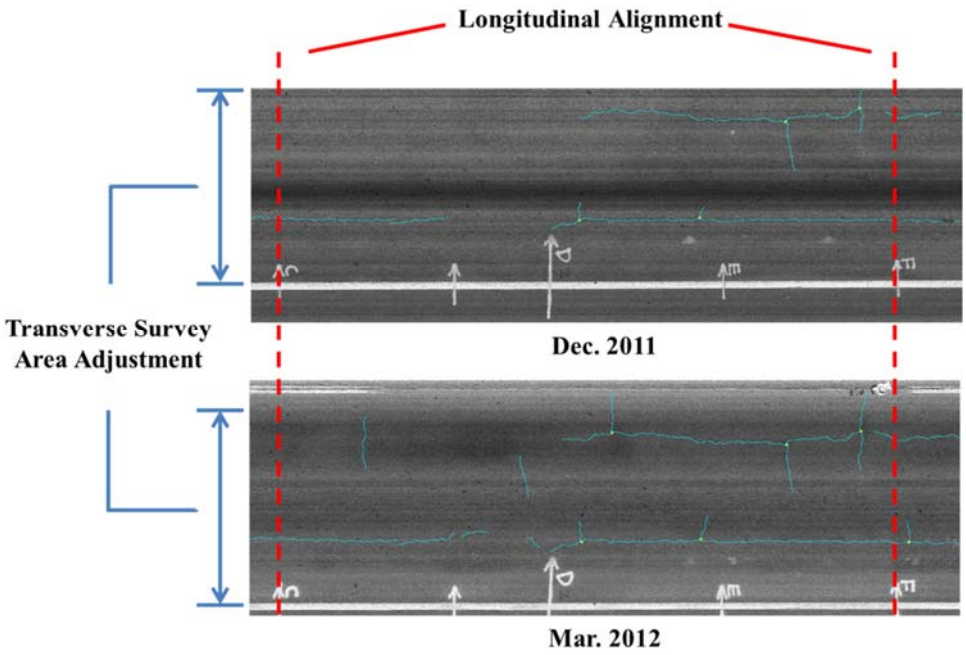


FIGURE 5.13: Illustration of manual registration

7. Summary

Pavement cracks continuously deteriorate over time. This chapter presents a potential method of monitoring pavement crack deterioration behaviors from temporal domains using high-quality, 3D pavement surface data. It also presents a detailed level representation of crack information (including length, width, orientation/direction, position/location, intensity, pattern, etc.), derived using 3D laser imaging technology, crack detection algorithms, and the multi-scale CFE model, could potentially be used to support 1) fundamental study of pavement mechanistic and pavement deterioration behavior, 2) validation of current pavement design methods and development of new design concepts and methods, 3) determination of adequate treatment methods and strategies based on pavement distress characteristics and their deterioration behavior, 4) development of accurate and reliable forecasting models, and 5) development of cost-effective pavement management operations/practices, such as intelligent crack sealing

planning. Compared to frequently used lab simulation data or controlled experimental data (e.g. AASHO road test), the in-service pavement data used in this study represent truer crack deterioration mechanisms and enable the ability to explicitly connect pavement condition deterioration with the real-world factors that cause it. While it is not feasible to comprehensively cover the entire field of research, this study is a concrete first step and is believed to be transformative in changing the way researchers have approached sensing based infrastructure condition monitoring and risk assessment.

References

Prozzi, J. A. and Madanat, S. (2000) "Using Duration Models to Analyze Experimental Pavement Failure Data," Transportation Research Record, No. 1699, pp. 87-94.

Hu, S., Zhou, F., Scullion, T., and Leidy, J. (2012) "Calibrating and Validating Overlay Tester–Based Fatigue Cracking Model with Data from National Center for Asphalt Technology," Transportation Research Record, No. 2296, pp. 57-68.

HRB. (1962) "The AASHO Road Test," Highway Research Board, National Academy of Sciences - National Research Council, Washington, D.C.

Reger, D., Christofa, E., Guler, I., and Madanat, S. (2013) "Estimation of Pavement Crack Initiation Models by Combining Experimental and Field Data," Journal of Infrastructure Systems, Vol. 19, No. 4, pp. 434-441.

Prozzi, J. A. and Madanat, S. (2004) "Development of Pavement Performance Models by Combining Experimental and Field Data," Journal of Infrastructure Systems, Vol. 10, No. 1, pp. 9-22.

Yang, J., Lu, J. J., Gunaratne, M., and Dietrich, B. (2006) "Modeling Crack Deterioration of Flexible Pavements: Comparison of Recurrent Markov Chains and Artificial Neural Networks," Transportation Research Record, No. 1974, pp. 18-25.

Nasseri, S., Gunaratne, M., Yang, J., and Nazef, A. (2009) "Application of Improved Crack Prediction Methodology in Florida's Highway Network," No. 2093, pp. 67-75.

Thomas, O. and Sobanjo, J. (2013) "Comparison of Markov Chain and Semi-Markov Models for Crack Deterioration on Flexible Pavements," Journal of Infrastructure Systems, Vol. 19, No. 2, pp. 186-195.

Loizos, A. and Karlaftis, M. G. (2005) "Prediction of Pavement Crack Initiation from In-Service Pavements: A Duration Model Approach," Transportation Research Record, No. 1940, pp. 38-42.

Tsai, Y. and Li, F. (2012) "Critical Assessment of Detecting Asphalt Pavement Cracks under Different Lighting and Low Intensity Contrast Conditions Using Emerging 3D Laser Technology," Journal of Transportation Engineering, Vol. 138, No. 5, pp. 649–656.

Tsai, Y., Jiang, C., and Huang, Y. (2014) "Multiscale Crack Fundamental Element Model for Real-World Pavement Crack Classification," Journal of Computing in Civil Engineering, Vol. 28, No. 4.

Kaul, V., Yezzi, A., and Tsai, Y. (2012) “Detecting curves with unknown endpoints and arbitrary topology using minimal paths,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 34, No. 10, pp. 1952-1965.

Chapter 6 Long-term Monitoring of Rutting Deterioration

In this chapter, 3D rut shapes and their deterioration at multiple scales, including project, segment, and individual rut-levels, are analyzed using (1) descriptive statistics of spatial parameters and the proposed temporal parameters; and (2) 2D and 3D visualization and image subtraction. In this study, a pavement project is defined as a section of pavement, typically a few miles long, which has a consistent pavement type, design, maintenance activity, etc. A pavement segment is a road section that is typically 1 mile or less in length. An individual rut refers to a road section that is smaller than 1 mile in length.

For project and segment-level deterioration, boxplots are generated to show the distribution of rut parameters at different timestamps. A boxplot is a practical tool that depicts the overall distribution of data (McGill et al. 1978). As depicted in FIGURE 6.1, a boxplot consists of a box with three horizontal bars and the whiskers (vertical bars). The box represents the middle fifty percentiles of the data (its range is defined as the interquartile range: IQR), and the horizontal bars at the top, middle, and bottom of the box denote the third quartile, the median, and the first quartile, respectively. The top whisker extends from the top of the box to the farthest data point that is within 1.5 times of the IQR from the top of the box (i.e., the upper extreme), and the bottom whisker extends from the bottom of the box in a similar manner. Any data points beyond the extent of the whiskers are considered outliers.

For individual rut level deterioration, 2D and 3D visualization, as well as image subtraction techniques that are similar to the ones shown in the previous chapter, are used to show the detailed 3D rut shapes.

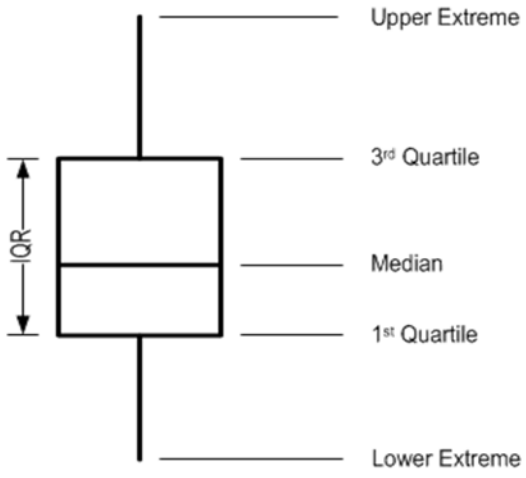


FIGURE 6.1: Illustration of a boxplot

1. Characterization of 3D Rut Shape and Its Deterioration

In this study, rut parameters defined in the literature, as well as some new parameters, are used to characterize a 3D rut shape and to quantify the change in the 3D shape. The following list shows the spatial and temporal rut parameters used in the analysis:

- Profile-based parameters:
  - LRD: PP69 Left rut depth (mm)
  - RRD: PP69 Right rut depth (mm)
  - PP69 Left rut width (mm)
  - RRW: PP69 Right rut width (mm)
  - LCA: PP69 Left cross-sectional area (mm^2)
  - RCA: PP69 Right cross-sectional area (mm^2)
  - TPA: Total positive area (mm^2)
  - TNA: Total negative area (mm^2)
  - DR: Distortion ratio

- TD: Total distortion (mm^2)
- TAD: Total absolute distortion (mm^2)
- PD: PP69 Percent deformation (%)

- Longitudinal parameters:
  - LRL: Left rut length (m)
  - RLR: Right Left rut length (m)
  - LRV: Left rut volume (m^3)
  - RRV: Right rut volume (m^3)

- Temporal parameters:
  - TEA: Total elevated area per year (mm^2/yr)
  - TDA: Total depressed area per year (mm^2/yr)
  - MED: Mean elevated distance per year (mm/yr)
  - MDD: Mean depressed distance per year (mm/yr)

Among these parameters, temporal parameters and TAD are first proposed in this research. The proposed parameters are described below.

1.1 Profile-based Parameters

Total Absolute Distortion

FIGURE 6.2 illustrates the positive and negative areas of a pavement profile, which could be sensitive to the location of the reference line. Therefore, in this study, we propose the total absolute distortion (TAD) as an additional indicator to describe the level of distortion. As defined in the following equation, TAD is the sum of absolute positive areas and absolute negative areas. TAD accounts for the distortion on both sides of the

reference line by summation, which naturally makes it less sensitive to the original share of the pavement surface.

$$TAD = |TPA| + |TNA|$$

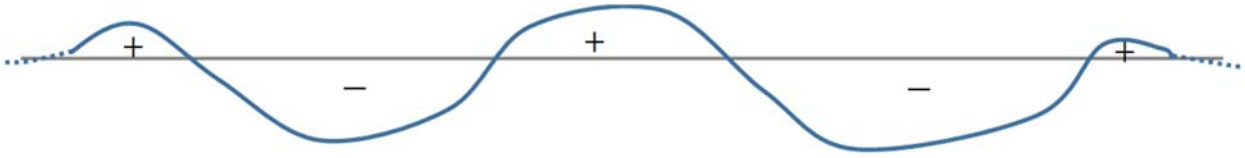


FIGURE 6.2: Illustration of positive and negative areas

Longitudinal Parameters

Rut Length

Rut length can be defined as the longitudinal extension of an individual rut measured in distance units, such as meter or foot. Rut length has been used by transportation agencies as a means to identify the percentage or extent of rut within a section length. Some other agencies use rut length as one of the indicators to define and monitor rutting. For example, the Transfund New Zealand measures rut by the length of a wheel path where rut depth exceeds 30 mm (Transfund New Zealand, 1997).

To calculate rut length, the minimum dimensions of a rut needs to be clearly defined. After taking current state DOT practices and AASHTO provisional standards into consideration, we define the minimum dimensions (i.e., depth, width, and length) of a rut to be 3 mm, 300 mm, and 30 m (1/8 in, 1 ft, and 10ft). In addition, the minimum distance between two ruts needs to be at least 6 m (20 ft), as suggested by Li (2012). If the

longitudinal length of the non-rutting area between two ruts is less than 6 m, the two ruts are considered as one, and the total length of this rut is the sum of the lengths of both ruts and the non-rutting area. Any pavement sections failing to meet these requirements, even with measured rutting, are considered localized deformation with a rut length of zero. Rut lengths in the left wheel path (denoted as LRL) are calculated separately from those in the right wheel path (denoted as RRL).

Rut Volume

Rut volume has been explored previously in the literature (Li 2012); however, it has not yet gained much attention among transportation agencies. Similar to the way rut length is calculated, rut volume is only calculated for ruts that meet the minimum dimensional and spacing requirements stated previously. Rut volume can be defined as the integral of rut areas along the length of the rut, as shown in the equation below. Rut volume in the left wheelpath (denoted as LRV) and the right wheelpath (denoted as RRV) are calculated separately.

$$RV_k = \sum_{i=1}^{n} lA_i, \forall\, i \in k$$

where

RV_k = rut volume of the kth rut;

l = longitudinal distance interval between two consecutive measurements;

A_i = the ith cross-sectional area; and

n = total number of measurements within the boundaries of rut k.

1.2 Temporal Parameters

To understand how rut deteriorates, one can compare the rut parameters of the same segment at different timestamps through statistical means (e.g., mean, median, percentiles, boxplot, etc.). With the spatiotemporally registered data, rut deterioration can be further quantified through direct comparison of the 3D rut shapes between different timestamps. In this study, the temporal parameters discussed below are proposed to measure the actual rate of change in 3D rut shapes (transverse profiles) as a more direct means to quantify rut deterioration.

Total Elevated and Depressed Areas Per Year

As depicted in FIGURE 6.3, the total elevated area per year (TEA) is the elevating rate of the total elevated area per year. Elevated areas are defined as the areas enclosed by two profiles where the elevation of the second timestamp profile (red profile) exceeds the elevation of the first timestamp profile (blue profile). Total depressed area per year (TDA), on the other hand, is the depressing rate of the sum of the depressed areas, where the elevation of the second timestamp profile is lower than the elevation of the first timestamp profile. Equations below define the total elevated and depressed area mathematically.

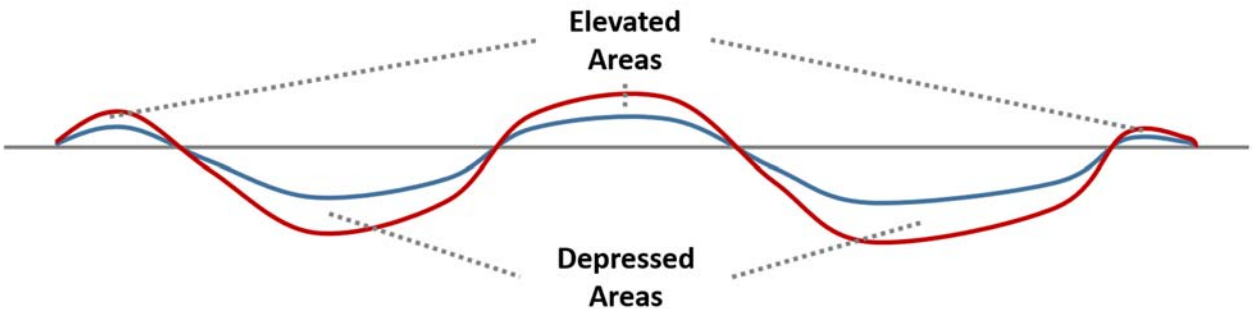


FIGURE 6.3: Illustration of temporal parameters

$$TEA = \frac{365}{t} \sum_{j=1}^{J} d_j, \forall j \in i$$

$$TDA = \frac{365}{t} \sum_{j=1}^{K} d_k, \forall k \in i$$

$$d_i = z_{2_i} - z_{1_i}, i = 1, \dots, n$$

where

t = number of days between Timestamps 1 & 2;

d_j = the jth point with positive elevation change;

d_k = the kth point with negative elevation change;

d_k = the signed elevation change of the ith point from profile z_1 to profile z_2;

z_1 = profile elevations at Timestamp 1;

z_2 = profile elevations at Timestamp 2;

J = total number of elevated points;

K = total number of depressed points; and

n = total number of points per profile, $J + K \leq n$.

Mean Elevated and Depressed Distances Per Year

In addition to the change in cross-sectional profile areas, change in elevation can also be calculated. In this study, the mean elevated distance per year (MED) is defined as the annual rate of in the average positive elevation from the first timestamp to the second timestamp. As defined in the equations below, the mean elevated distance is the rate at which the average elevation changes among all points that have a positive elevation change. Mean depressed distance per year, or MDD, can be calculated in the same manner.

$$MED = \frac{TEA}{J}$$

$$MDD = \frac{TDA}{K}$$

2. Project-Level Rut Deterioration

For project-level analysis, long-term 3D pavement data collected on Georgia State Route 26 between Mileposts 5.5 and 11.5 were used. This road section consists of two 6-mile projects, one in the eastbound direction and the other in the westbound direction. Long-term 3D pavement data were collected at 7 timestamps between March 2012 and February 2016.

2.1 Deterioration of Profile-based Parameters

FIGURE 6.4 shows the boxplots of all profile-based rut parameters of SR 26 eastbound Mileposts 5.5 to 11.5 at multiple timestamps. From this figure, it is noticed that rutting within this project slowly became more severe. This result echoes GDOT's annual survey results in which the pavement condition evaluation system (PACES) rating of this project gradually decreased over the analysis period. FIGURE 6.5 shows the boxplots of all profile-based rut parameters of SR 26 westbound, Mileposts 5.5 to 11.5, at multiple timestamps.

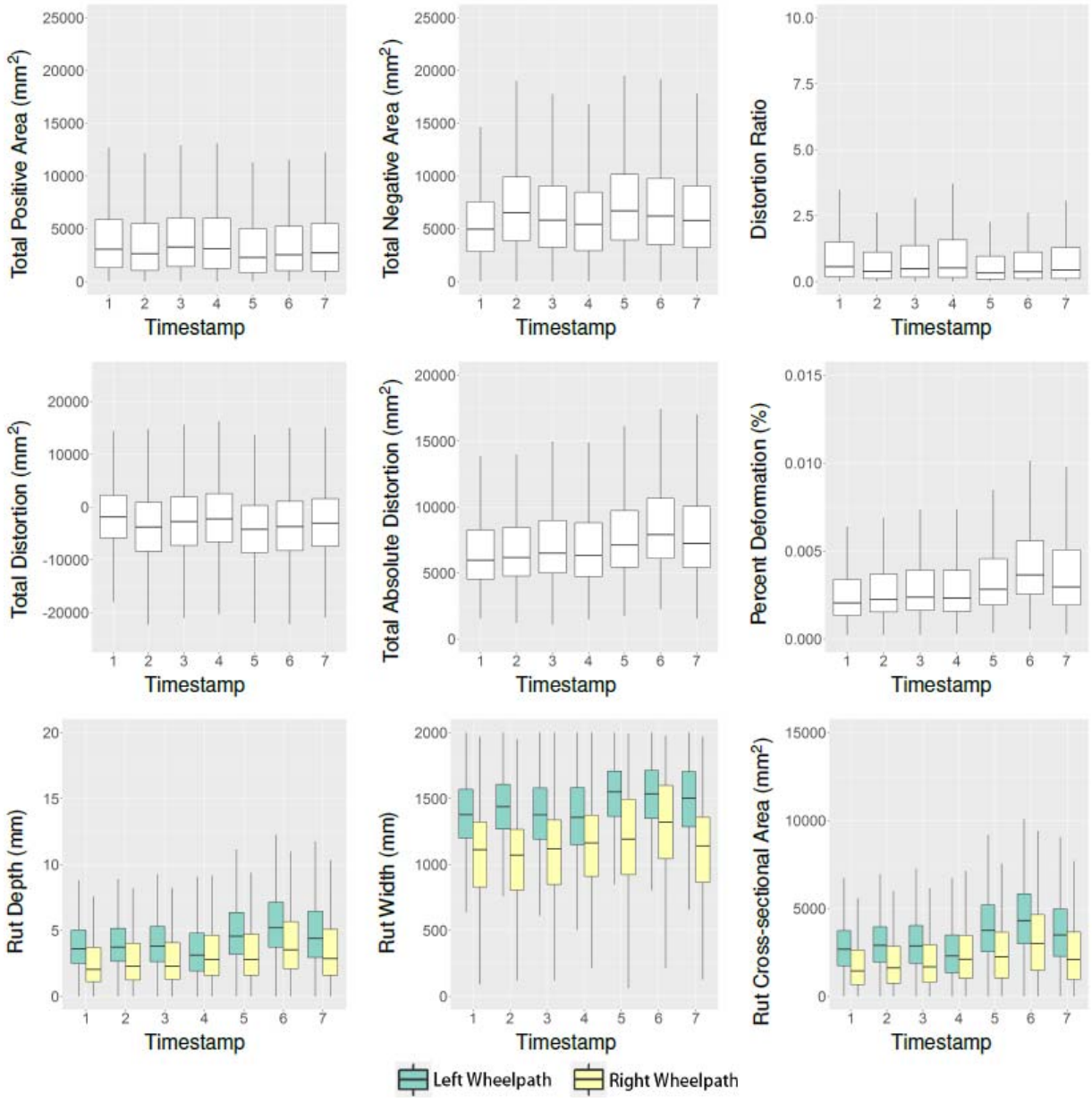


FIGURE 6.4: Deterioration of profile-based parameters at project-level: SR 26 Eastbound MP 5.5-11.5

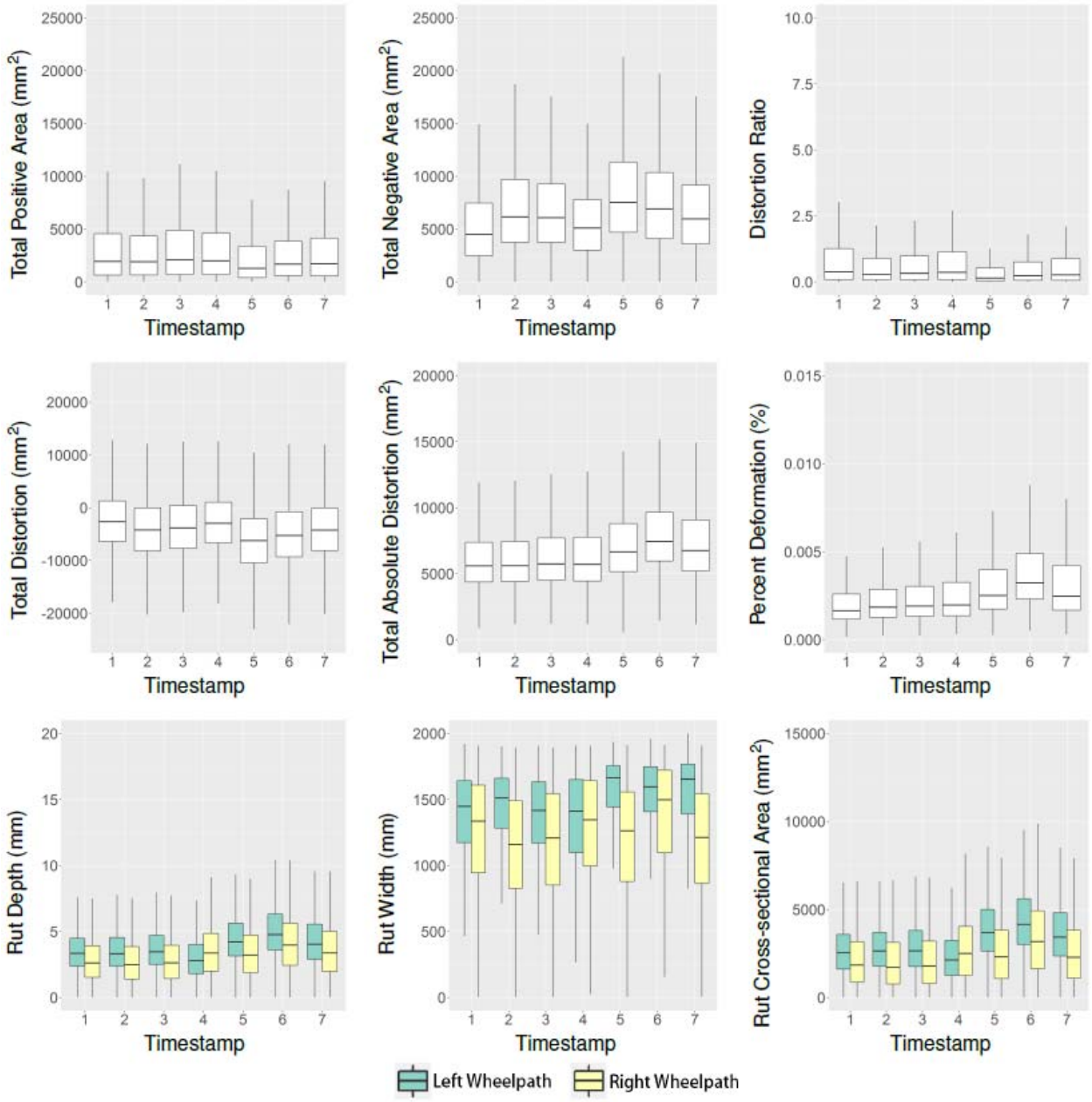


FIGURE 6.5: Deterioration of profile-based parameters at project-level: SR 26 Westbound MP 11.5-5.5

Most profile-based parameters did not change dramatically over the course of these 7 timestamps. Several profile-based parameters, including rut depth, rut cross-sectional area, percent deformation, and total absolute distortion, showed similar trends that increased gradually. Interestingly, while these parameters showed positive trends at Timestamps 1, 2, 3, 5, and 6, their trends slowed down or became negative at timestamps 4 and 7. A possible explanation of this observation can be the seasonal variability, i.e., rut deterioration slows or stops in low- temperature seasons, as depicted in FIGURE 6.6 (White et al. 2002). Note that Timestamps 4 and 7 were in December and February, respectively, and Timestamps 1, 2, 3, 5, and 6 were in summer and spring.

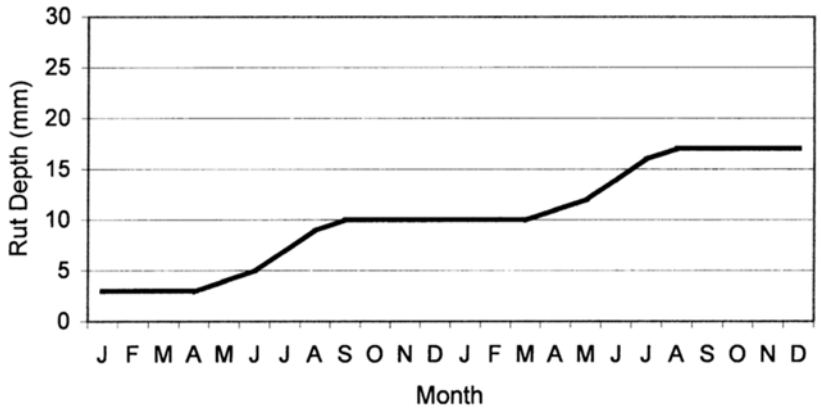


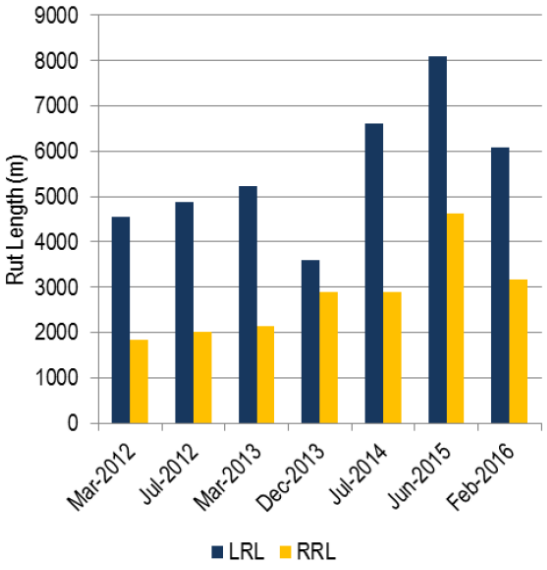
FIGURE 6.6: Potential seasonal effects on rutting (White et al. 2002)

Other profile-based parameters, such as total positive area, total negative area, distortion ratio, and total distortion, nevertheless, did not have a clear, common trend. Overall, the value of the total negative area was larger than that of the positive area, indicating that the majority of this project had larger depression than elevation. Consequently, the total distortion, which is defined as the total positive area minus the total negative area, tends to be negative throughout the analysis period.

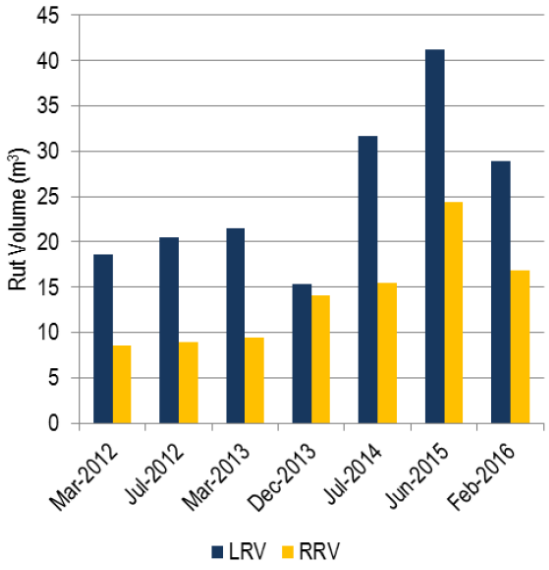
Similar findings are also observed in the other 6-mile project in the Westbound direction on SR 26, which is reasonable because the two projects share very similar, if not the same, pavement design, traffic volume, characteristics, and climates. Overall, the Eastbound direction had more severe rutting than the Westbound direction. Results of the Westbound project are shown in FIGURE 6.5.

2.2 Deterioration of Longitudinal Parameters

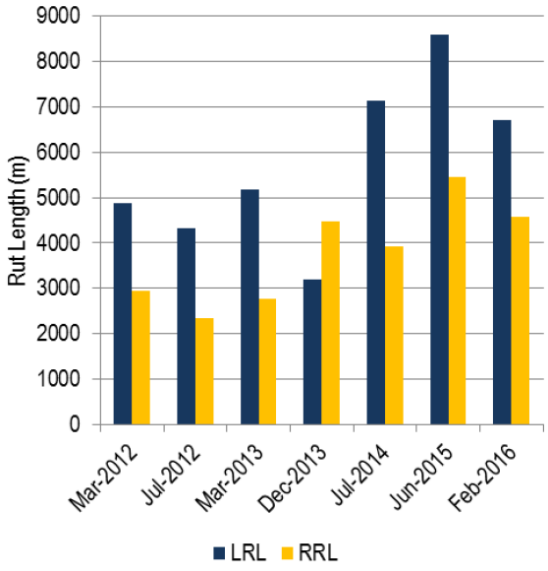
Deterioration of project-level longitudinal parameters, including the total length and volume of ruts, are summarized in this section. As depicted in FIGURE 6.7, a high correlation between rut length and rut volume can be observed. This phenomenon is expected, since both rut length and rut volume are only calculated at locations where there is rutting. In addition, these two parameters generally follow similar trends as the ones observed in a few profile-based parameters, including rut depth, rut cross-sectional area, percent deformation, and total absolute distortion. This result indicates that the deterioration of rut shape is multi-dimensional, i.e., rut deteriorates in depth, length, cross-sectional area, and volume. Seasonal variation seems to have an effect on the changes observed in longitudinal parameters, too. While positive trends in length and volume can be observed during spring and summer seasons, growth in rut length and volume became moderate and sometimes negative in winter. Further confirmation of this observed behavior can be performed in the field to verify the effect of seasonal variation on rut shapes.



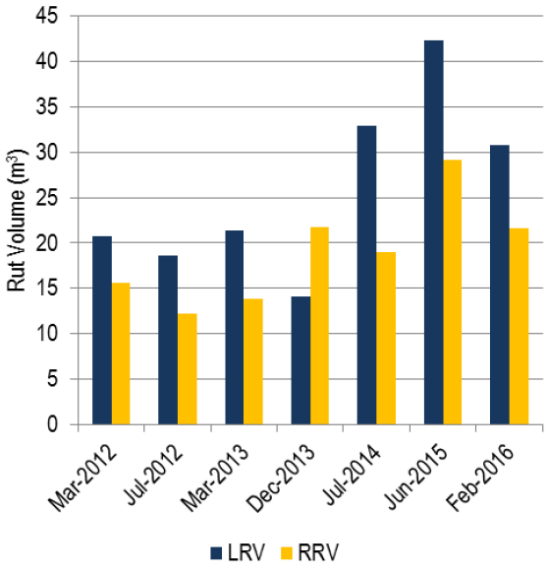
(a) Rut Length (SR26 Eastbound)



(b) Rut Volume (SR26 Eastbound)



(c) Rut Length (SR26 Westbound)



(d) Rut Volume (SR26 Westbound)

FIGURE 6.7: Deterioration of longitudinal parameters at project-level

2.3 Deterioration of Temporal Parameters

FIGURE 6.8 and FIGURE 6.9 show the deterioration of ruts in temporal rut parameters, including the mean elevated and depressed distances, and the total elevated and depressed areas. As defined in the previous section, temporal parameters at each timestamp were obtained by calculating their rates of change per year.

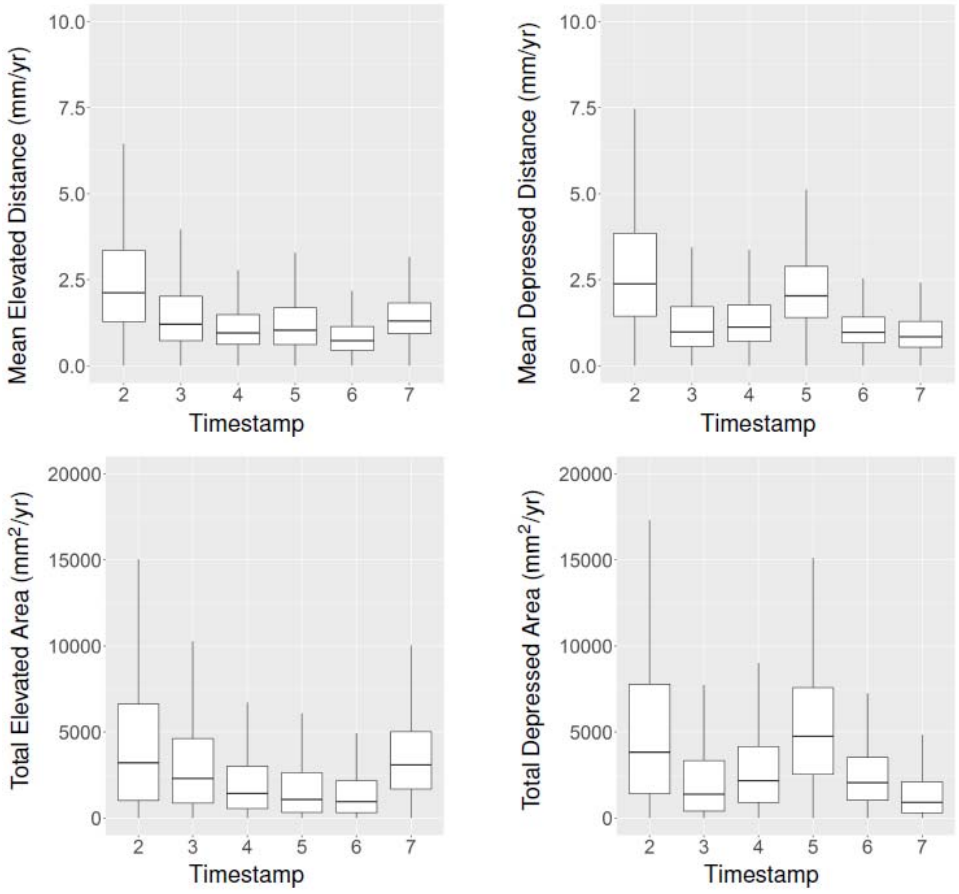


FIGURE 6.8: Deterioration of temporal parameters at project-level: SR 26 Eastbound

These two figures show that the two projects shared similar trends in all four parameters. It is noticeable that while the deterioration rates were temperate (e.g., < 1.5 mm/yr) at most timestamps, the depressed rates in depth and in area were higher at Timestamps 2

and 5. This finding can be associated with the seasonal variation as discussed above. As depicted in Figure in FIGURE 6.6, the deterioration of rut depth tends to be more rapid in late spring and summer, including April, May, June, July, and August, and it slows down during the late fall and winter season (i.e., October to February).

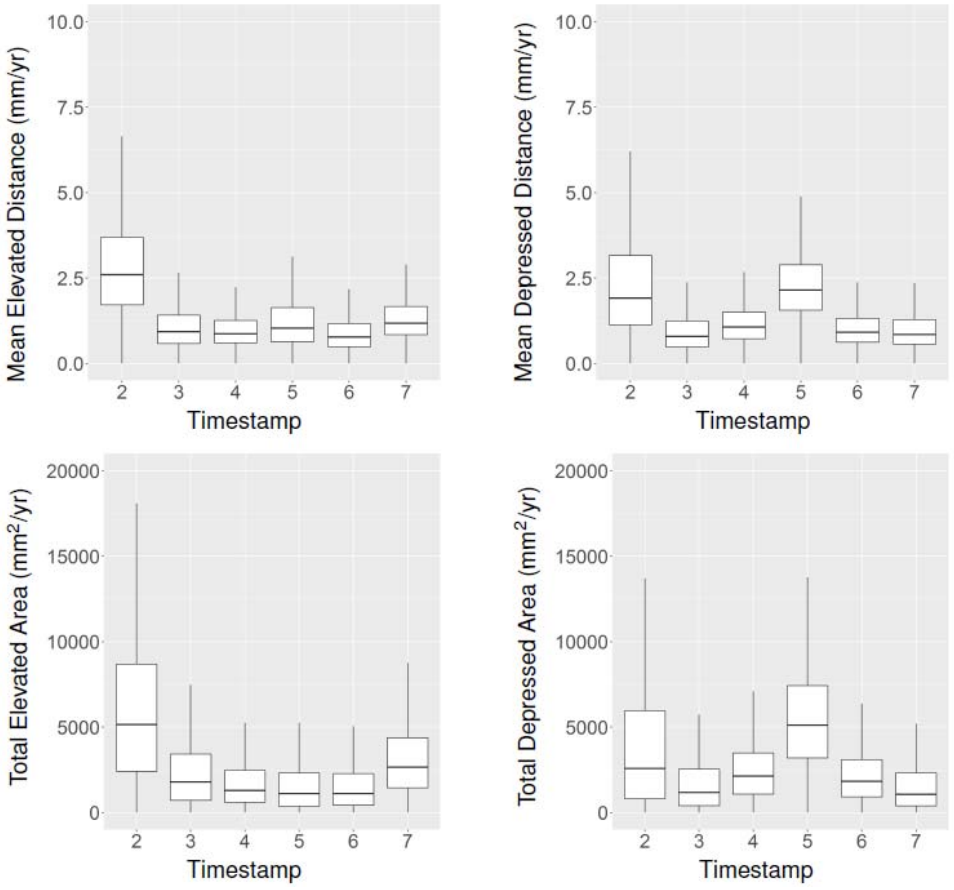


FIGURE 6.9: Deterioration of Temporal Parameters at Project-level: SR26 Westbound

By comparing these two timestamps with their respective previous timestamps, it is noted that the season transitioned from winter or early spring to summer. Given the fact that a rut deteriorates more rapidly in spring and summer, these two timestamps are, therefore, expected to have higher deterioration rates.

Interestingly, although Timestamp 6 occurred in the summer, deterioration rates at this timestamp were not as pronounced as those at Timestamps 2 and 5. This is because deterioration rates at both Timestamps 2 and 5 were calculated over shorter periods (i.e., 4 months and 7 months) in spring and summer, whereas the rates at Timestamp 6 were calculated over an 11-month period that spanned all seasons. Since the deterioration rate is typically higher in summer, when propagating a rate that is observed over a few summer months to an annual rate, the projected annual rate is expected to be higher. On the contrary, if the period between two timestamps is longer (e.g., close to a year or more than a year), the deterioration rate is more likely to be "balanced out" over the course of the period. As a result, deterioration rates at Timestamp 6 were not as noticeable as those at Timestamps 2 and 5.

3. Segment-Level Rut Deterioration

Rut deterioration in the two projects shown in the previous section, together with two other segments (one on SR 275 and the other on I-95), are further analyzed on a one-mile segment basis. Similar to the project-level analysis, deterioration of rutting in these 14 segments were analyzed using descriptive statistics of various spatial and temporal rut parameters, and the results are summarized below.

3.1 Deterioration of Wheelpath-Specific, Profile-based Parameters

FIGURE 6.10 and FIGURE 6.11 show the deterioration of wheelpath-specific rut parameters, including rut depth, rut width, and rut cross-sectional area. By comparing the deterioration of these rut parameters, several findings that are consistent with the results

of project-level analysis can be summarized. First, seasonal variation seems to have effect on all segments, as rut parameters tend to grow more slowly or negatively at Timestamps 4 and 7 and faster at the other timestamps. Second, it is evident that rut depth and rut cross-sectional area share similar trends, indicating these two parameters have higher correlation. Rut width, on the other hand, shows large variation between 0 mm and 2,000 mm (which are essentially the possible minimum and maximum width) and does not show clear correlation with other parameters.

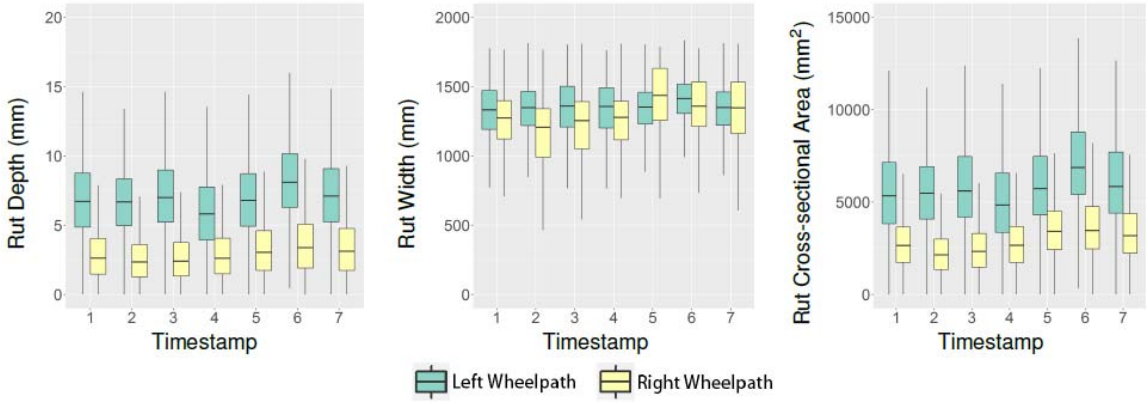


FIGURE 6.10: Deterioration of Wheelpath-specific Profile-based Parameters at Segment-level: SR 275 Northbound MP 0-1

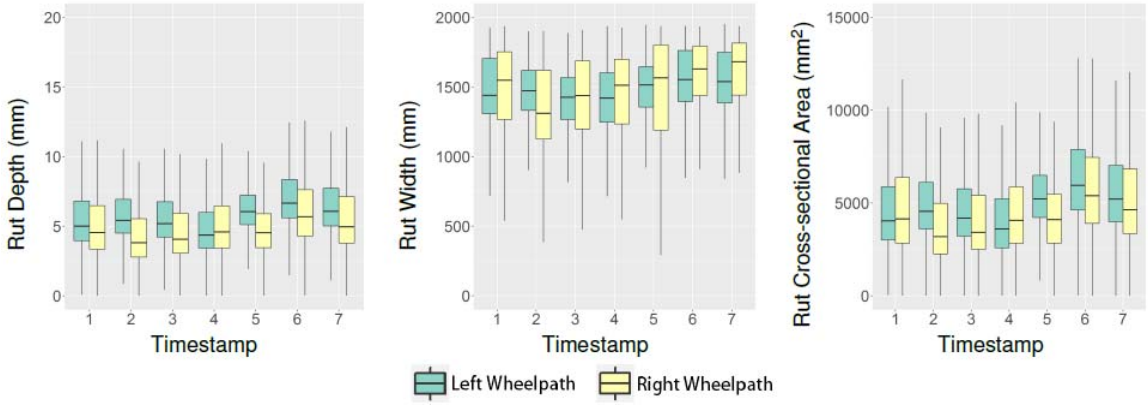


FIGURE 6.11: Deterioration of wheelpath-specific profile-based parameters at segment-level: I-95 Southbound MP 101-100

By comparing consecutive segments on SR26's two projects, as depicted in FIGURE 6.12 and FIGURE 6.13, it is noted that rut conditions vary noticeably among these segments. An overall observation is that segments in the eastbound direction of SR 26 had more severe ruts than those on the westbound direction. Segments between Mileposts 10.5 to 11.5 in both directions showed extremely severe rutting with the rut depth larger than 1 in. A possible explanation of the severe rutting in these two segments is that it has a lot of truck traffic traveling from and to the Savannah Port. There is an intersection at around Milepost 11.5 that is also at the beginning and ending of a transition slope to a bridge. As a result, excessive stop-and-go heavy truck traffic may be the cause of severe rutting in these two segments.

Furthermore, when comparing the deterioration of rut parameters in segments from different routes, it is observed that the general trend of rut deterioration was slower on SR 275 and I-95 than on several segments of SR 26. This finding indicates that rut deterioration can potentially be affected by traffic characteristics and different pavement design.

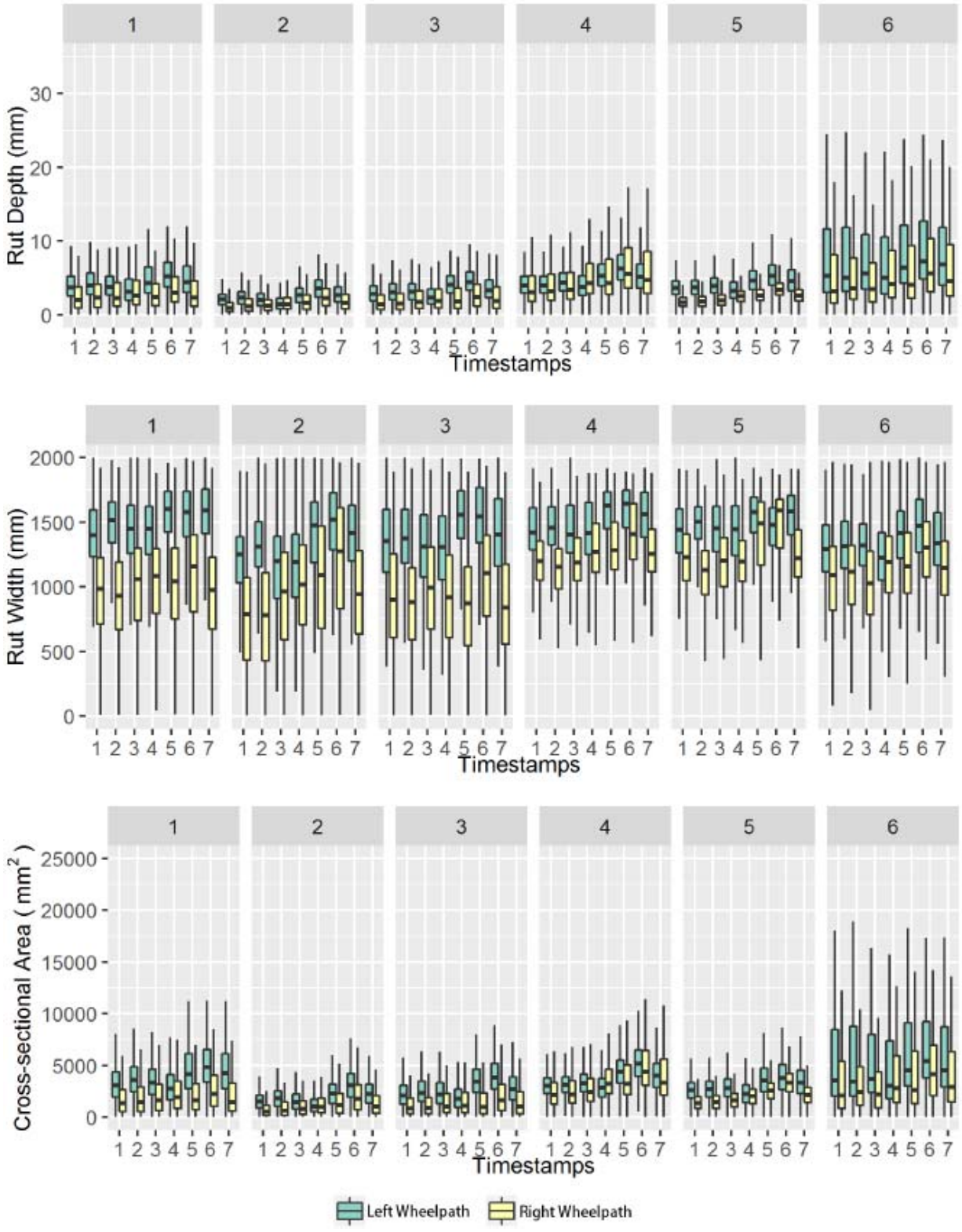


FIGURE 6.12: Deterioration of wheelpath-specific profile-based parameters at segment-level: SR 26 Eastbound MP 5.5-11.5

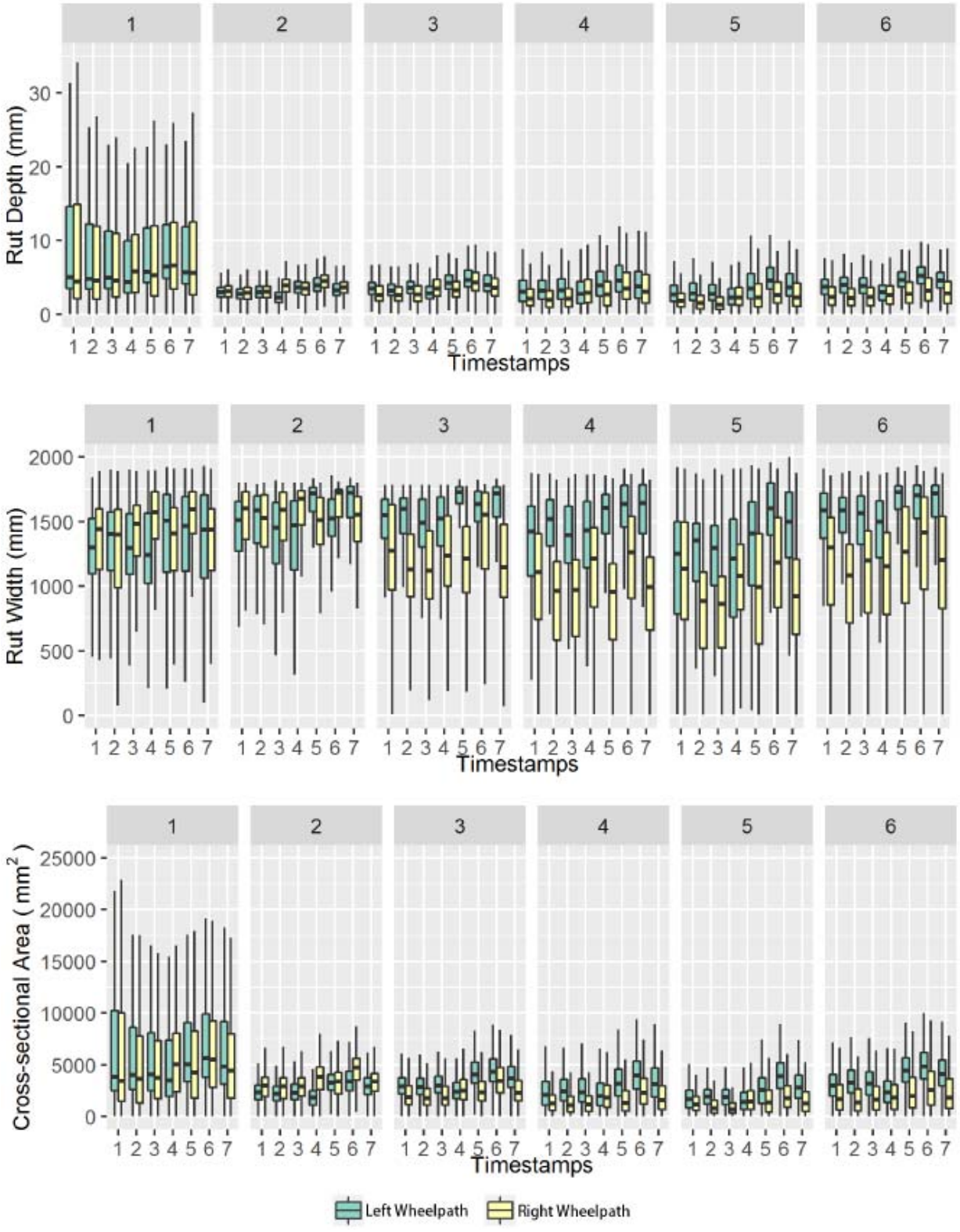


FIGURE 6.13: Deterioration of wheelpath-specific profile-based parameters at segment-level: SR 26 Westbound MP 11.5-5.5

3.2 Deterioration of Other Profile-based Parameters

Deterioration of other profile-based rut parameters, including total positive and negative areas, total distortion, total absolute distortion, distortion ratio, and percent deformation, are further analyzed in this section. Results of the deterioration of these parameters are shown in FIGURE 6.14 to FIGURE 6.15. In FIGURE 6.14, it is observed that the deterioration of other profile-based parameters was not noticeable in the SR 275 segment throughout the 7 timestamps. In FIGURE 6.15, on the other hand, a gradual deterioration can be observed in parameters such as TAD and PD.

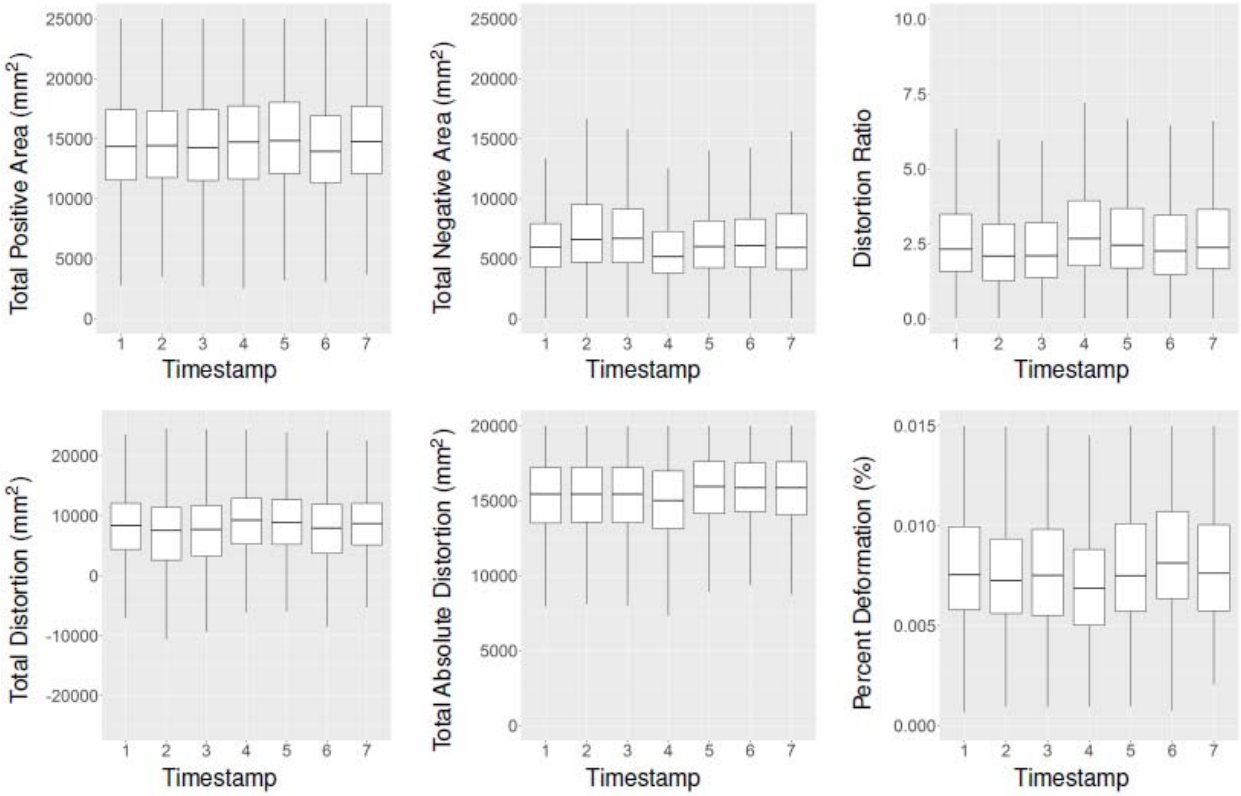


FIGURE 6.14: Deterioration of other profile-based parameters at segment-level: SR 275 MP 0-1

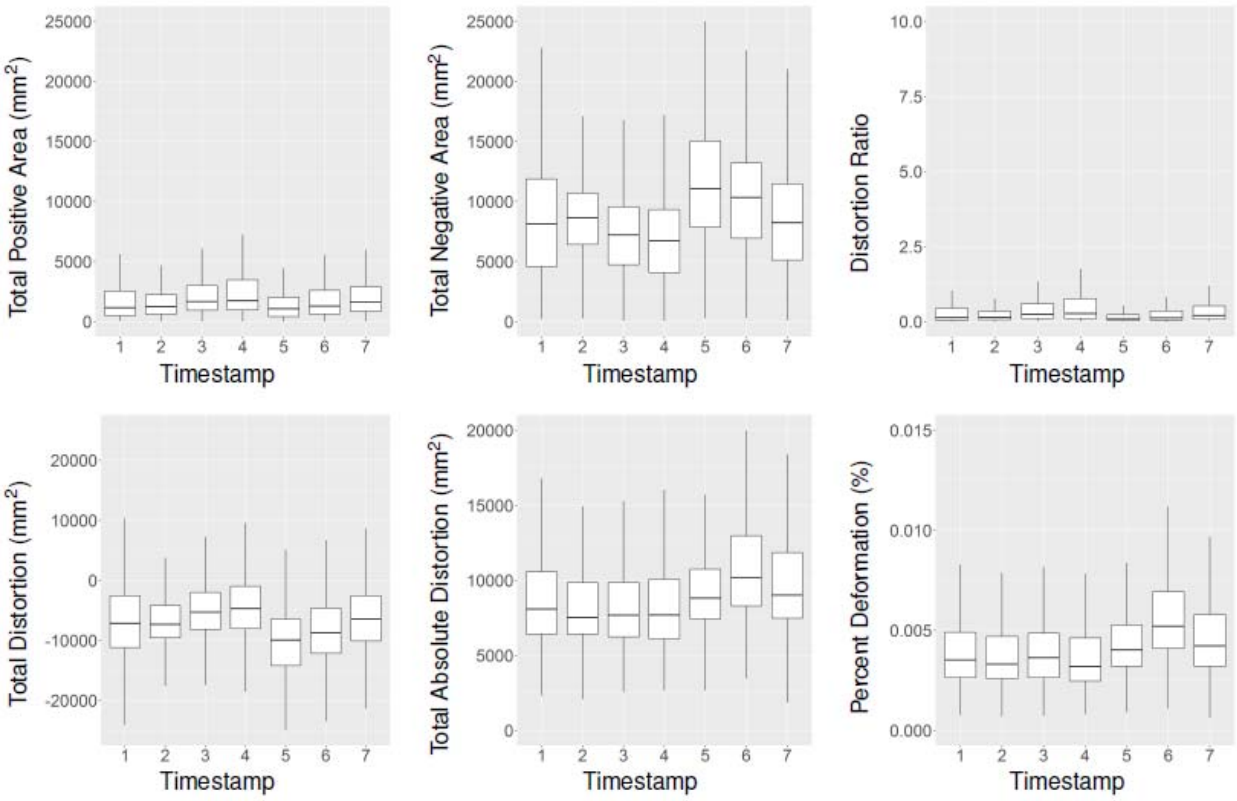


FIGURE 6.15: Deterioration of other profile-based parameters at segment-level: I-95 MP 101-100

It is noted that the rut parameters derived based on the positive and negative areas, including TPA, TNA, TD, and DR, show larger variations in their trends over the 7 timestamps in most of the segments, unlike the trends of TAD and PD, which are more consistent and similar to the trends of rut depth and cross-sectional area. Since the positive and negative areas are defined by calculating the areas defined by the transverse profile and an imaginary reference line that connects two edges of the lane, variation can be introduced into these parameters when the geometric relationship between the profile and the reference line slightly changes. Parameters such as TD and DR, therefore, cannot directly reveal true physical shape and conditions of the rut; instead, they represent the relative relationship between TPA and TNA.

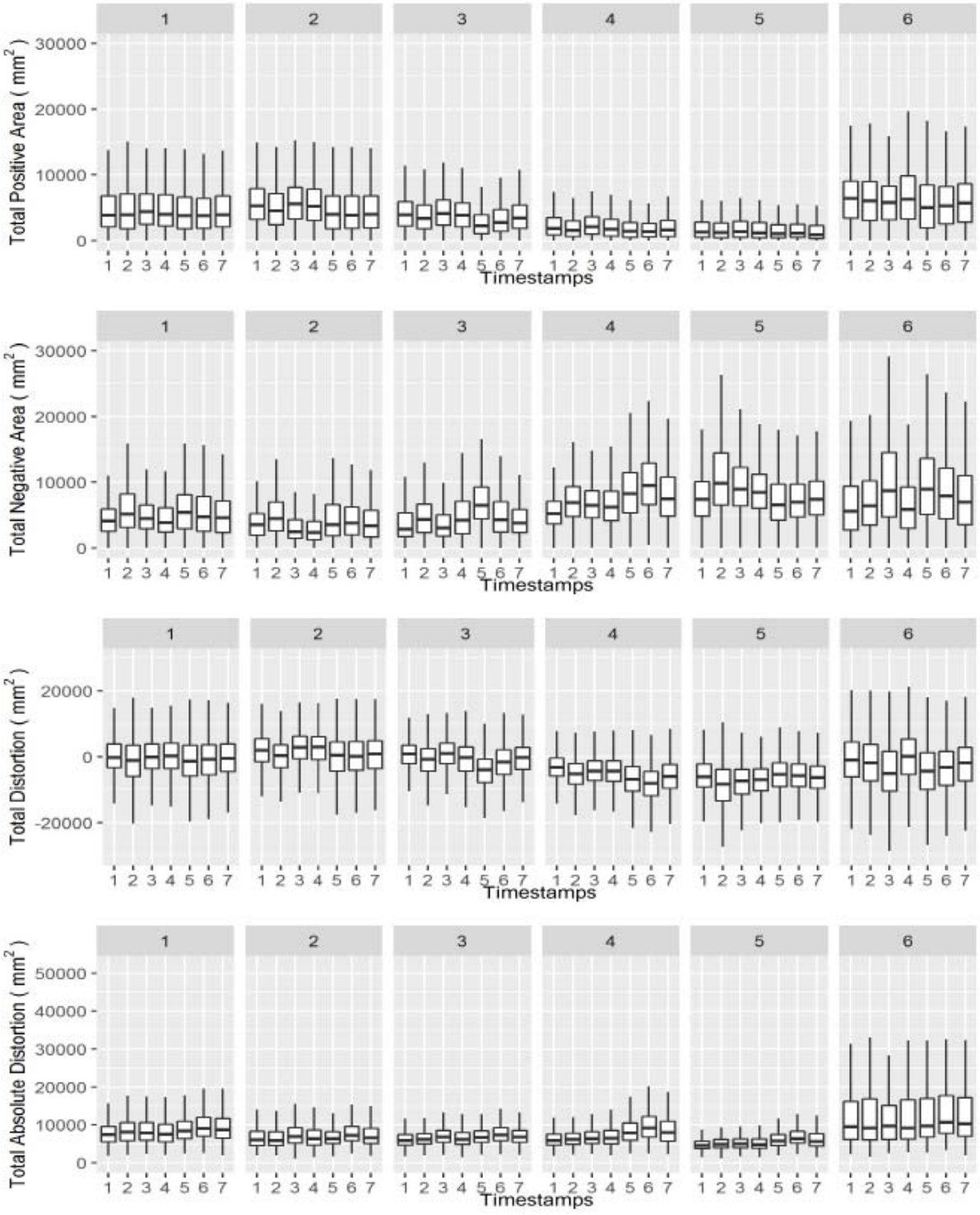


FIGURE 6.16: Deterioration of profile-area-based parameters at segment-level: SR 26 Eastbound MP 5.5-11.5

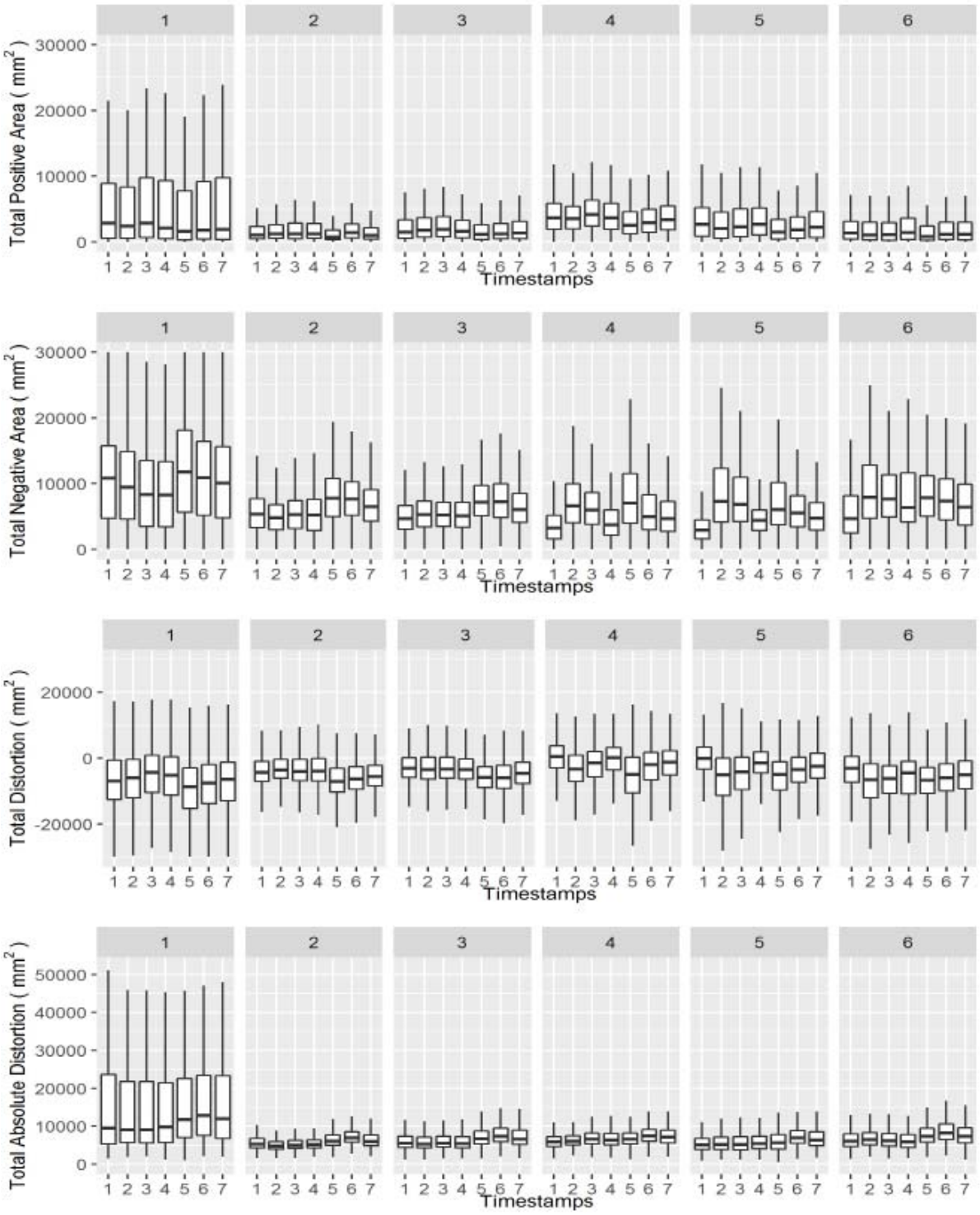


FIGURE 6.17: Deterioration of profile-area-based parameters at segment-level: SR 26 Westbound MP 11.5-5.5

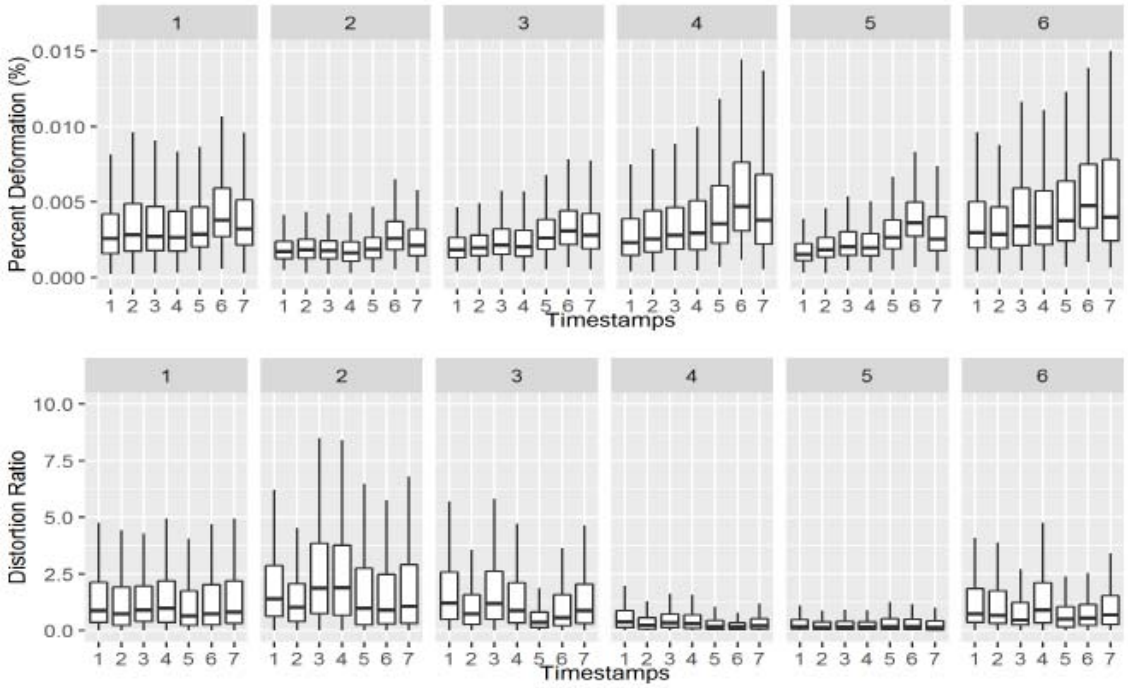


FIGURE 6.18: Deterioration of other profile-based parameters at segment-level: SR 26 Eastbound MP 5.5-11.5

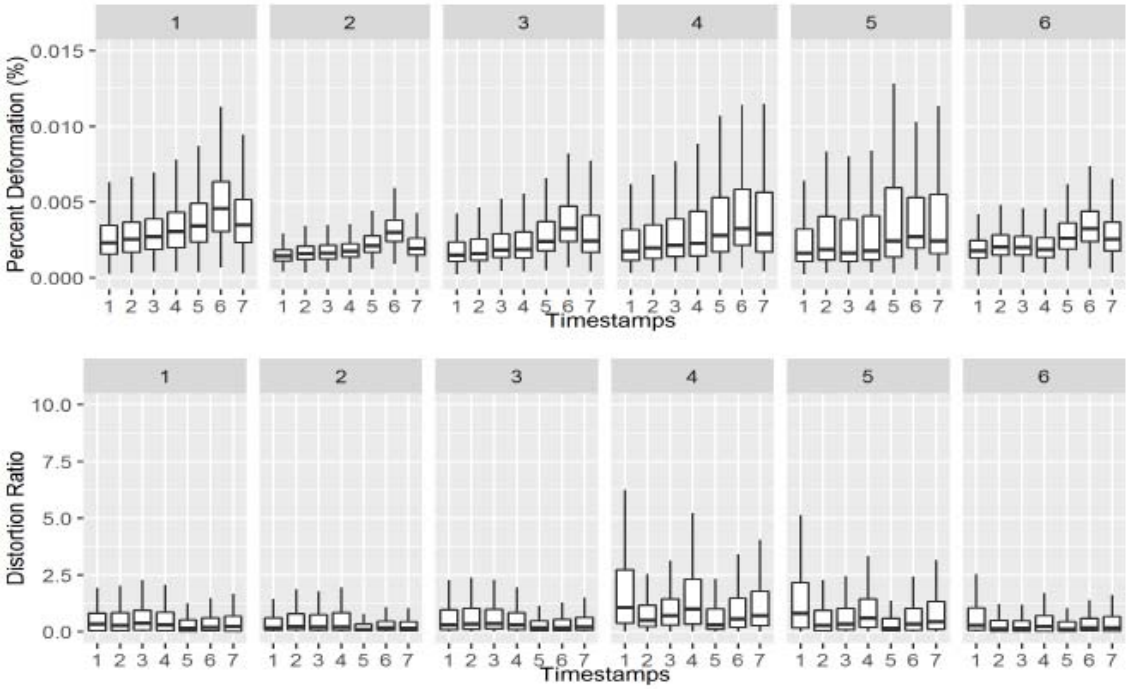


FIGURE 6.19: Deterioration of other profile-based parameters at segment-level: SR 26 Westbound MP 11.5-5.5

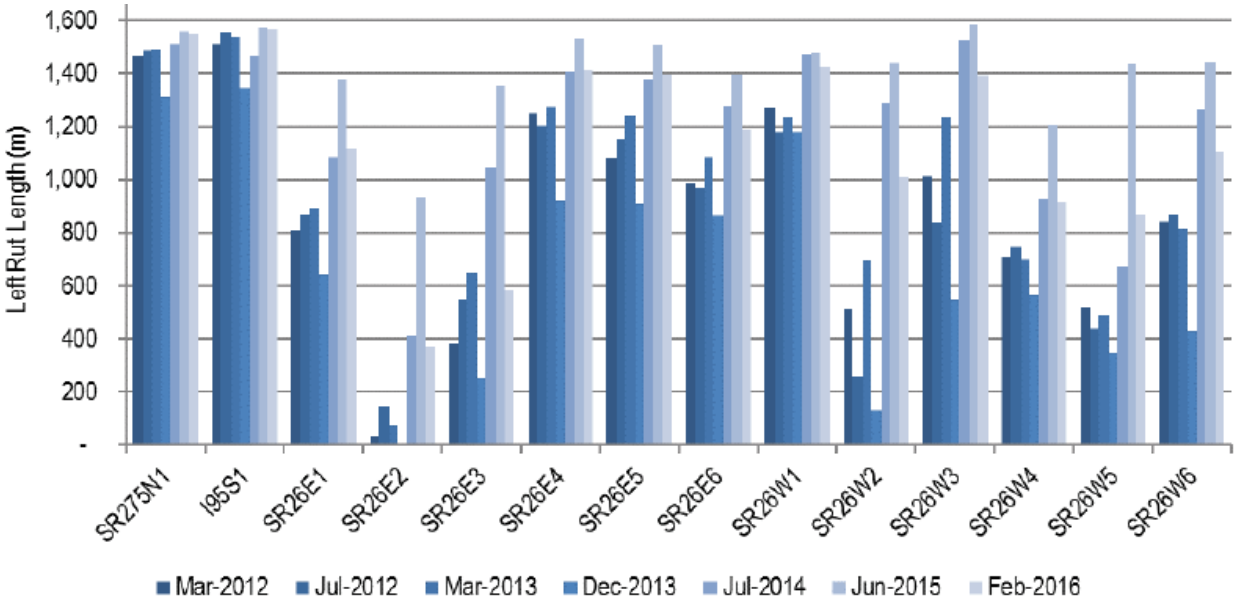
3.3 Deterioration of Longitudinal Parameters

FIGURE 6.20 and FIGURE 6.21 show the deterioration of longitudinal parameters at the segment level. In these figures, each bar represents the length (volume) of rutting in a segment at a given timestamp. By comparing the length and volume values of the same segment across multiple timestamps, a positive trend is generally observed. This indicates that most of the segments deteriorate in length and volume. When comparing the parameters between left and right wheelpaths, it is noted that the left wheelpath, overall, had more severe rutting. A similar relationship between wheelpaths can also be found in other wheelpath-specific parameters, such as rut depth and cross-sectional area. This finding, however, is inconsistent with the general expectation that rutting in the right wheelpath is typically more severe than in the left wheelpath, since traffic load is usually higher on the right wheelpath under the effect of cross slope.

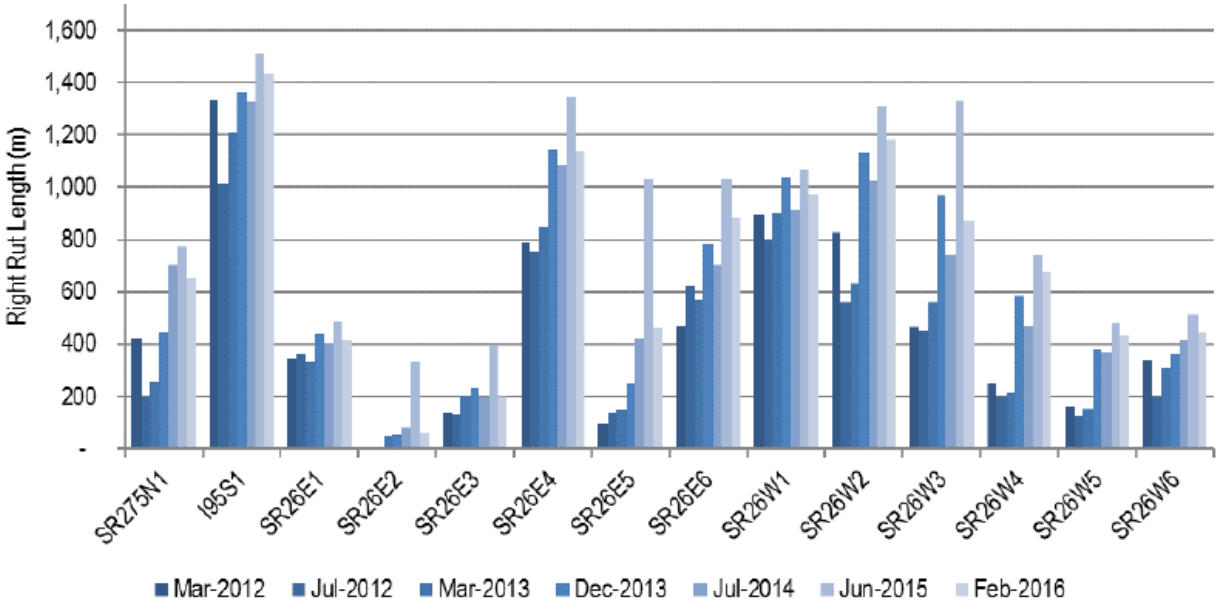
3.4 Deterioration of Temporal Parameters

FIGURE 6.22 to FIGURE 6.25 show the deterioration of temporal rut parameters. The effect of seasonal variation, as discussed in previous sections, is also evident in the trends of temporal parameters at the segment level. Note that when compared to other segments, SR 275 has, overall, smaller temporal parameter values (in FIGURE 6.22), indicating that rutting in this segment deteriorated more slowly than the other segments. I-95 and SR 26 segments, on the other hand, showed fairly similar trends and patterns. It is noted that all segments on SR 26 had very similar temporal parameter values at the same timestamp; this result indicates that although segments in the same project may have different severity levels of rutting, they tend to deteriorate at the same rate (as in this case). One possible explanation of this finding is that since all segments of these two

pavement projects shared very similar pavement design, age, and traffic characteristics, they deteriorated in a similar manner.

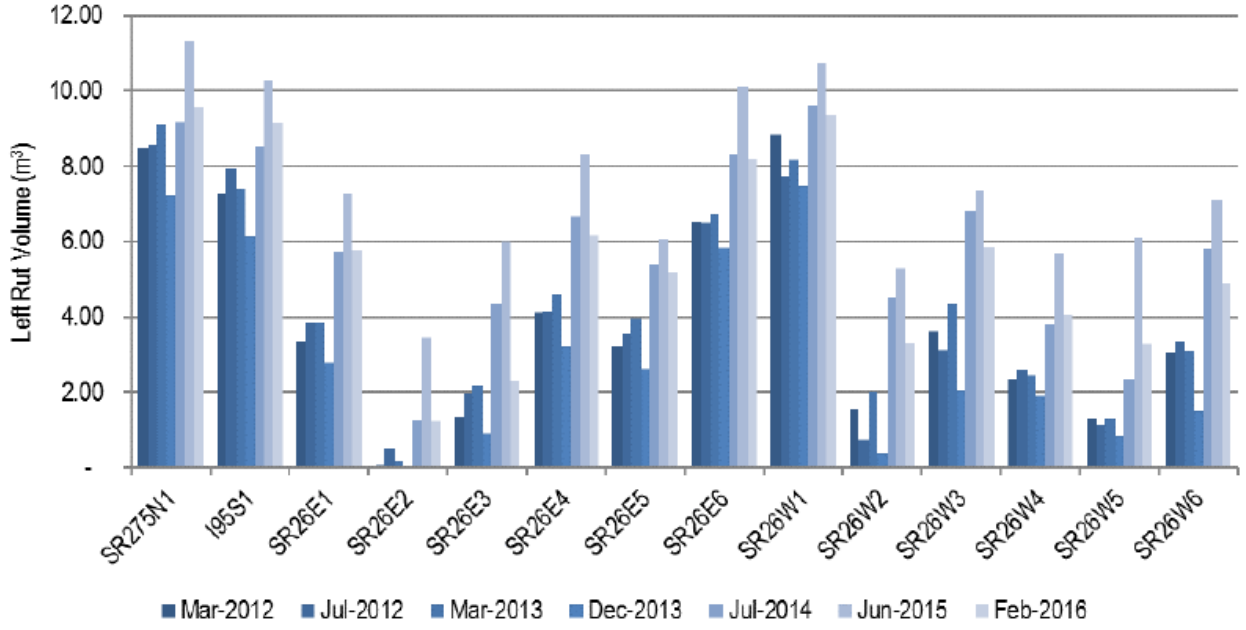


(a) Left Rut Length

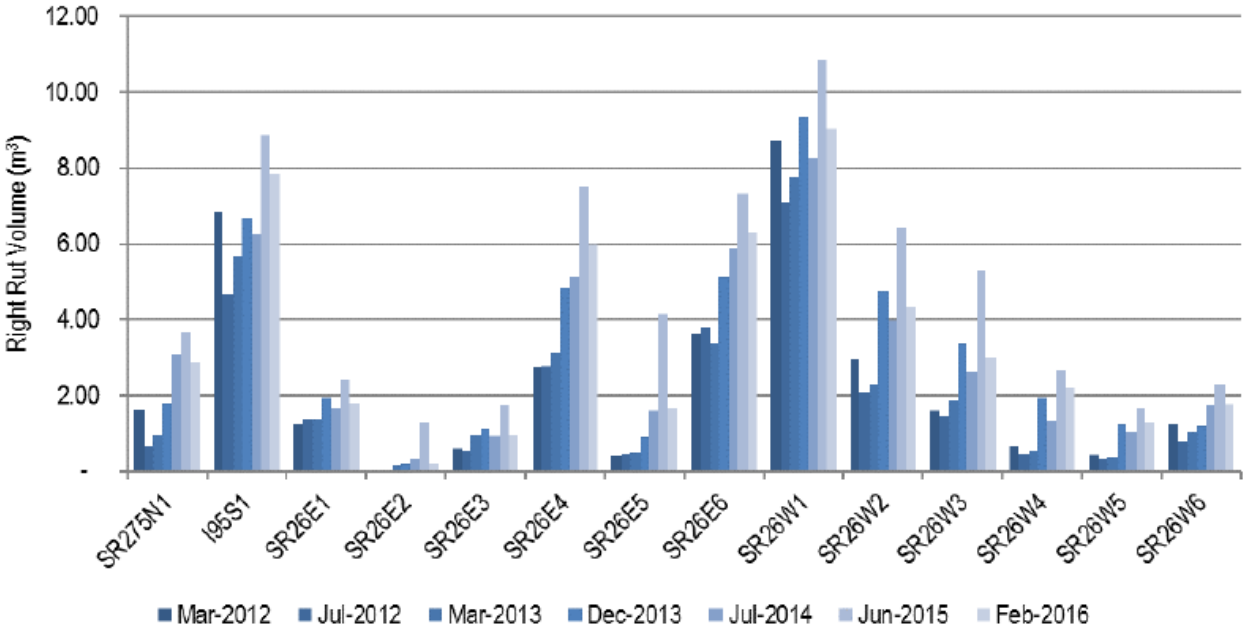


(b) Right Rut Length

FIGURE 6.20: Deterioration of rut length at segment-level



(a) Left Rut Volume



(b) Right Rut Volume

FIGURE 6.21: Deterioration of rut volume at segment-level

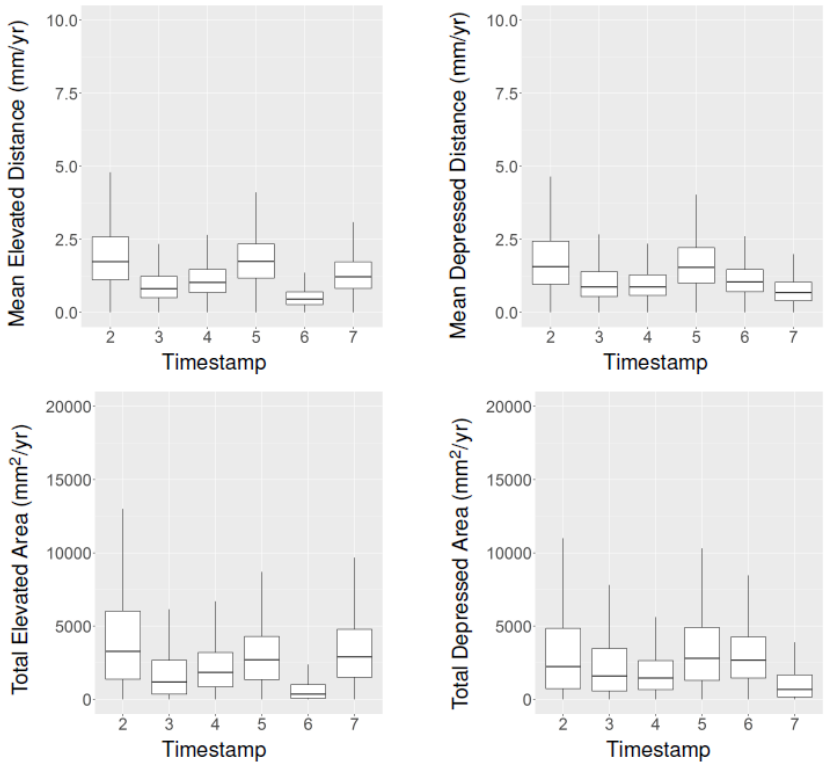


FIGURE 6.22: Deterioration of temporal parameters at segment-level: SR 275

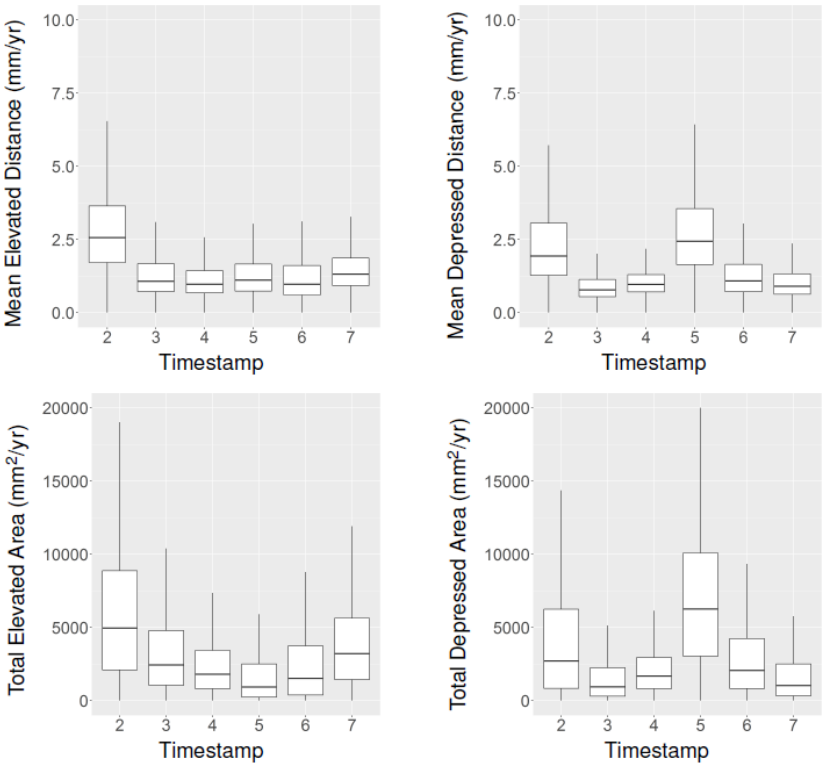


FIGURE 6.23: Deterioration of temporal parameters at segment-level: I-95

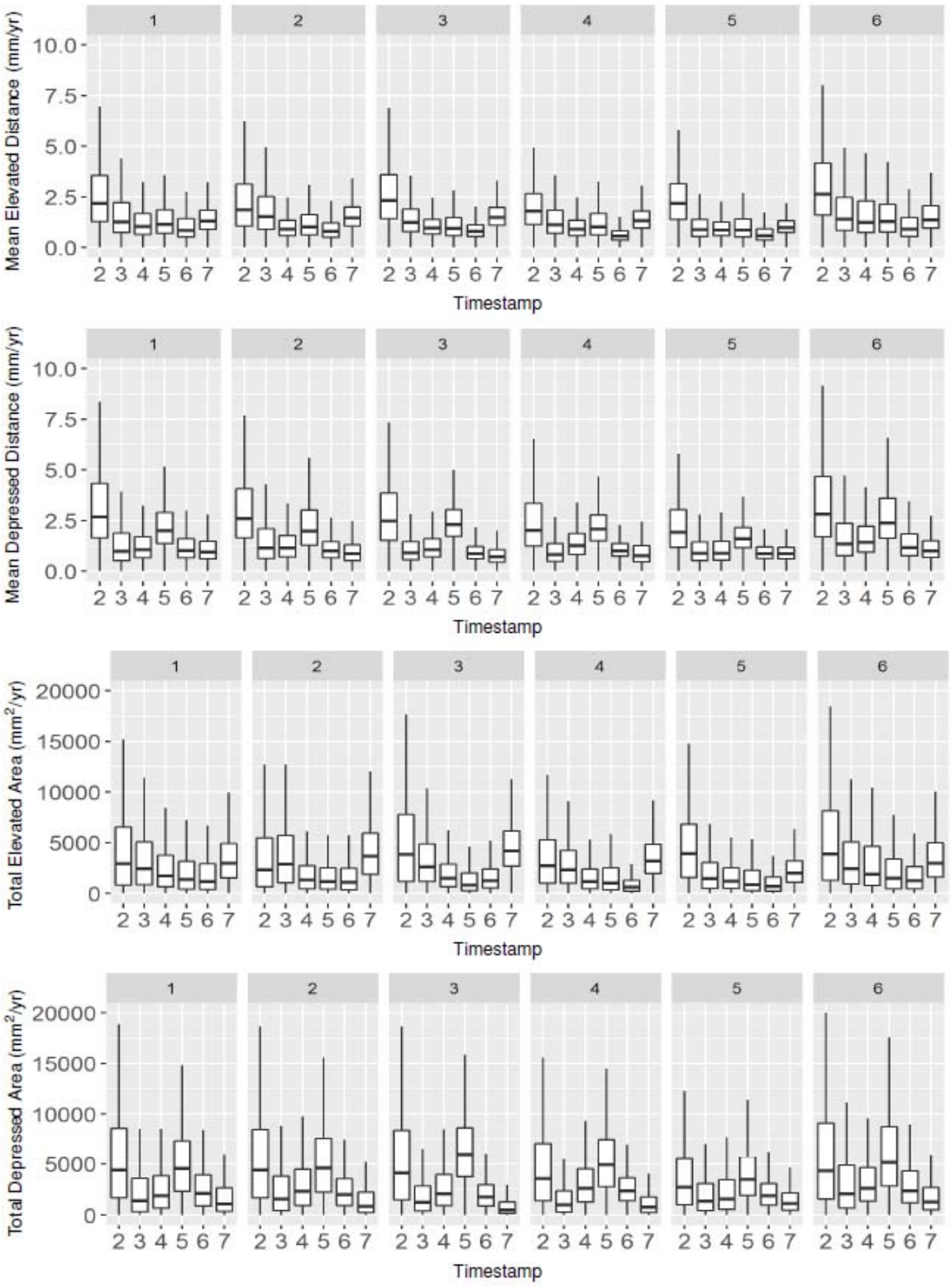


FIGURE 6.24: Deterioration of temporal parameters at segment-level: SR 26 Eastbound MP 5.5-11.5

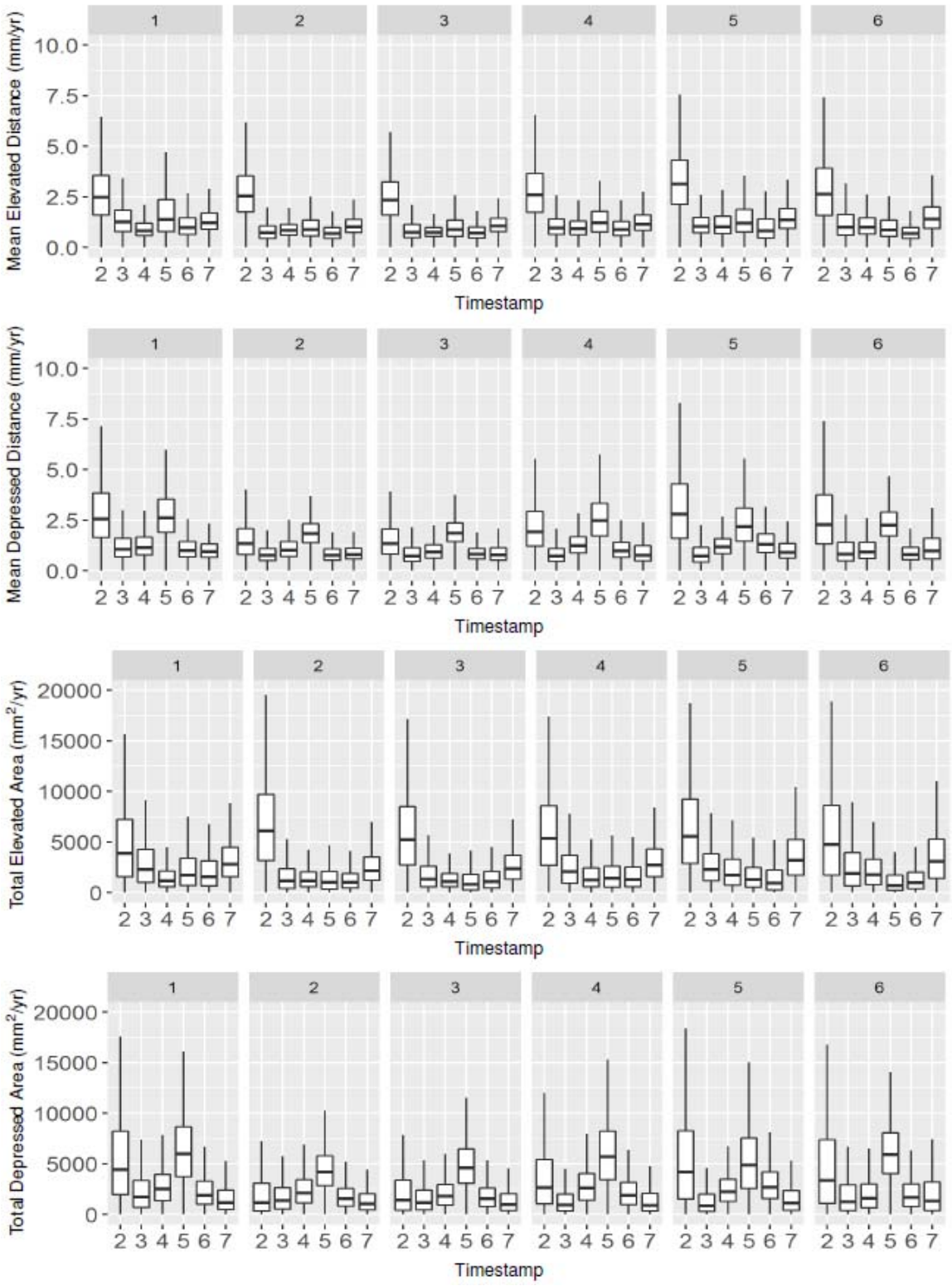


FIGURE 6.25: Deterioration of temporal parameters at segment-level: SR 26 Westbound MP 11.5-5.5

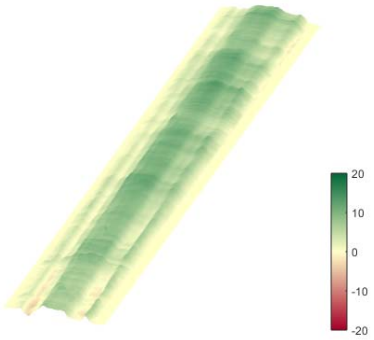
4. Individual-Rut-Level Rut Deterioration

The above sections demonstrate how 3D sensing technology and the proposed data registration method can be used as statistical means that advances the understanding of rut deterioration at larger scales. In fact, the strength of the proposed method prevails when used to study the detailed deterioration of rutting at a finer level. In this section, a few examples are selected to demonstrate how the proposed method can be used to quantify rut deterioration at the individual rut-level. Acknowledging the possible effect of seasonal variation on the shape of rutting, only summer timestamps (i.e., 7/13/2012, 7/18/2014, and 6/15/2015) were selected for the individual rut-level analysis in this section.

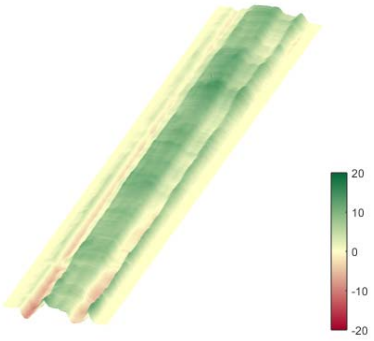
4.1 Rut Deterioration of a Selected Section on SR 26

FIGURE 6.26 and FIGURE 6.27 illustrate the 3D rut shapes of a 25-meter pavement section on SR 26 and the exact change in shape between different timestamps. By comparing the 3D shapes of these timestamps, it is clear that ruts in both wheelpaths deteriorated faster when close to the beginning section (bottom left of each figure). This finding was also observed in the field at the beginning section, which was at an intersection where heavy truck traffic frequently decelerates and accelerates. From the 3D shape change, as depicted in FIGURE 6.26 (d) and FIGURE 6.26 (e), it is clear that the change was more severe between the first two timestamps. This is reasonable because the first two timestamps were two years apart, whereas the last two timestamps were only one year apart. From the 2D deterioration maps in FIGURE 6.27 (d) and FIGURE 6.27 (e), it is evident that the deterioration in this section can be associated with heavy truck traffic because of the dual-wheel shapes observed in both wheelpaths. The

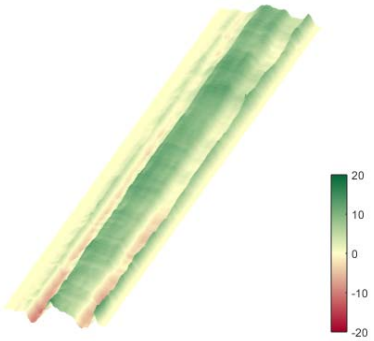
exact elevation and cross-sectional area changes can be further calculated and represented using the temporal parameters proposed in this study.



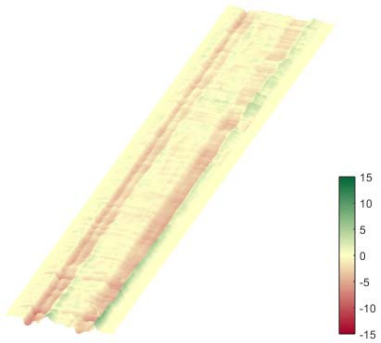
(a) 7/13/2012



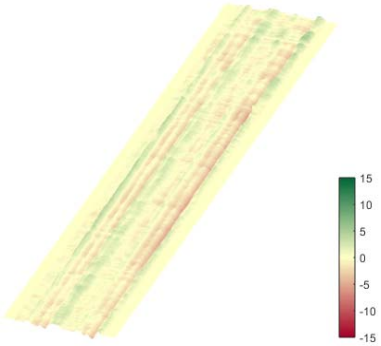
(b) 7/18/2014



(c) 6/15/2015

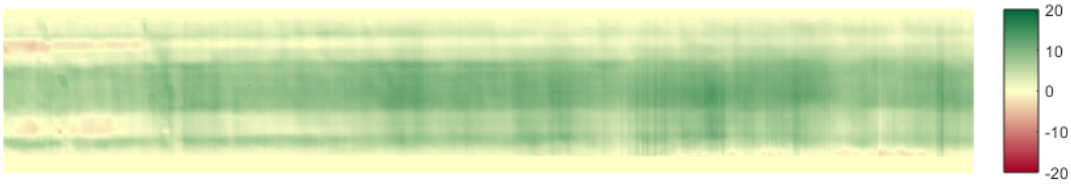


(d) 7/13/2012 to 7/18/2014

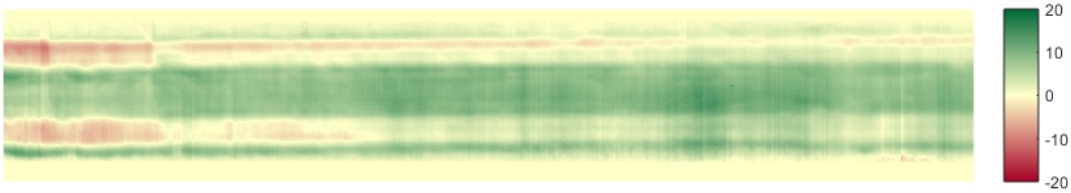


(e) 7/18/2014 to 6/15/2015

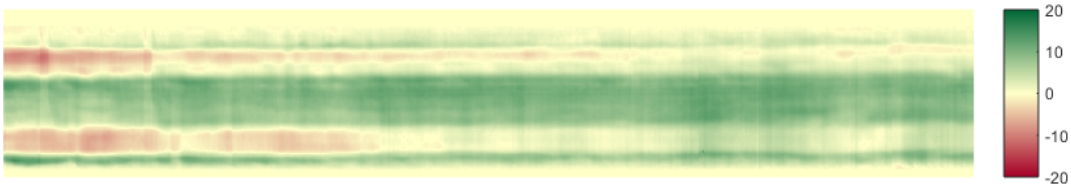
FIGURE 6.26: 3D visualization of rut deterioration on SR26 Westbound



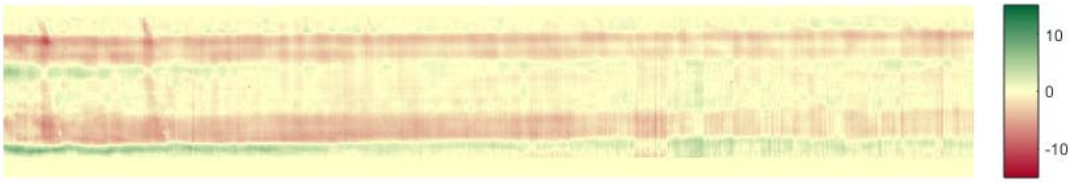
(a) 7/13/2012



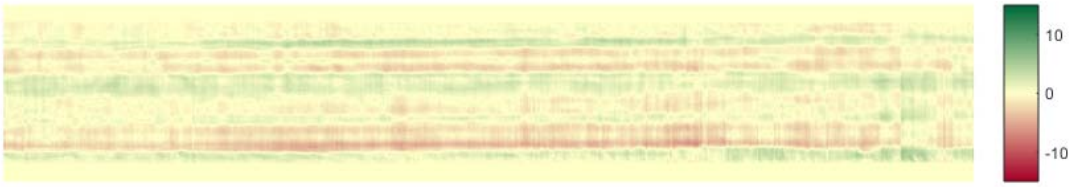
(b) 7/18/2014



(c) 6/15/2015



(d) 7/13/2012 to 7/18/2014



(e) 7/18/2014 to 6/15/2015

FIGURE 6.27: 2D visualization of rut deterioration on SR26 Westbound

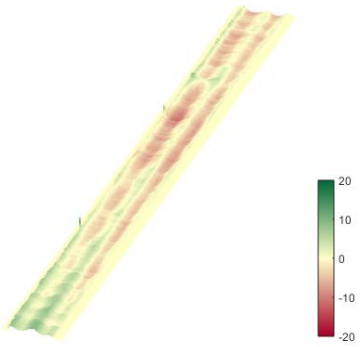
4.2 Rut Deterioration of a Selected Section on SR 275

A 45-meter section on SR 275 Northbound was selected to demonstrate how the proposed method can be used to visualize and quantify the deterioration of rutting. As shown in FIGURE 6.28 and FIGURE 6.29, rutting in both wheelpaths grew not only in depth, but also in length and volume. While the deterioration was not as obvious as the previous example on SR 26, the change in the longitudinal shape was appreciable. This finding further affirms the importance of the use of longitudinal parameters for rut characterization and deterioration analysis.

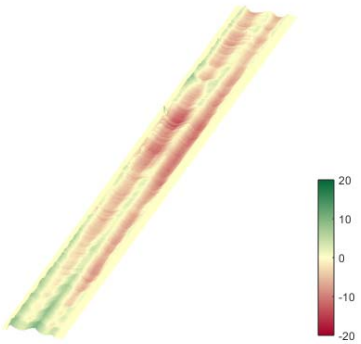
4.3 Discussion

The two examples shown in this section demonstrate how the proposed boundary-based data registration method can effectively register multi-timestamp 3D pavement data for rut characterization and deterioration analysis at large scales, such as project and segment-levels; it can also support deterioration analysis at the individual rut-level.

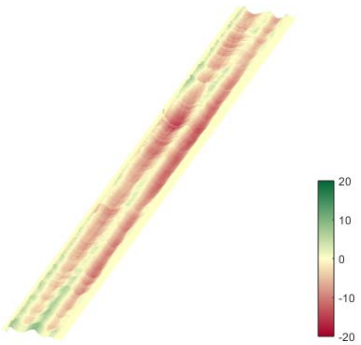
The examples above also demonstrate two types of data visualization for long-term 3D pavement data. Each of these two types of visualization has its advantages and disadvantages. For example, the 3D models can help engineers evaluate the pavement at a scale and detailed level that cannot be achieved even in the field. Many 3D modeling software programs can provide interactive functions that allow engineers to examine the pavement section at any desired scale and angle. These 3D graphics, however, are limited by the viewing angle if plotted on a 2D surface (i.e., paper). The 2D visualization, on the other hand, can effectively provide a holistic view of how the section deteriorates, using color, even though it does not visually show the 3D shape.



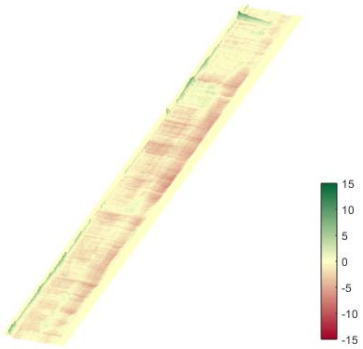
(a) 7/13/2012



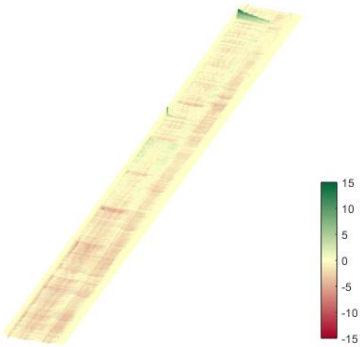
(b) 7/18/2014



(c) 6/15/2015

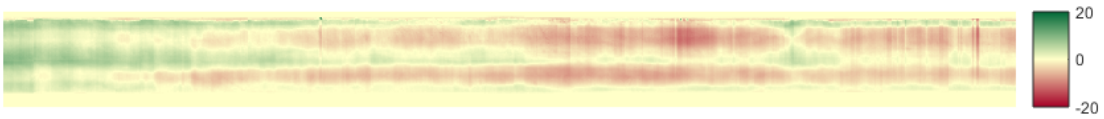


(d) 7/13/2012 to 7/18/2014

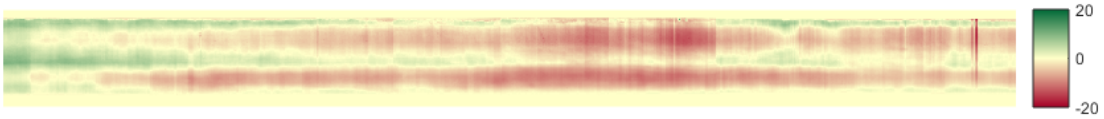


(e) 7/18/2014 to 6/15/2015

FIGURE 6.28: 3D Visualization of Rut Deterioration on SR 275 Northbound

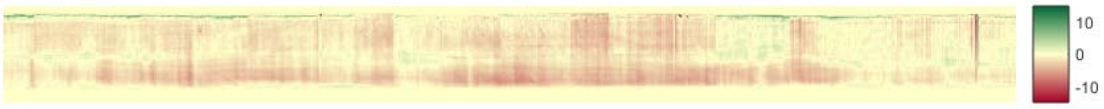


(a) 7/13/2012



(b) 7/18/2014

(c) 6/15/2015



(d) 7/13/2012 to 7/18/2014

(e) 7/18/2014 to 6/15/2015

FIGURE 6.29: 2D Visualization of rut deterioration on SR275 Northbound

5. Preliminary Assessment of Rut Parameters

A fair number of rut parameters have been defined in the literature to characterize the shape of 3D rutting. While some parameters can provide a fair representation of the severity of rutting, they may not be sufficient to describe the deterioration behavior of ruts. In this section, some preliminary assessment of rut parameters used in this study, in terms of their correlation and some observed issues, are summarized.

5.1 Correlation among Rut Parameters

Correlation among rut parameters has been studied in the literature for identifying the potential variation of different data collection techniques. For example, Simpson examined the correlation between rut depths derived using different methods (e.g., straightedge and stringline). Simpson (1999) also explored the correlation among other parameters, such as rut width, positive area, negative area, and fill area. Qiu (2013) also conducted correlation analysis to examine rut depth measurements calculated using the AASHTO PP69 provisional standard and the straightedge method.

In this section, the correlation among rut parameters used in this study is examined. FIGURE 6.30 to FIGURE 6.33 show the correlation matrices among rut parameters on the available routes. From these matrices, several findings in the multi-scale deterioration analysis can be confirmed. For example, it is noted that several rut parameters, including PD, LRD, RRD, LCA, RCA, and TAD, show high correlations. This indicates that these parameters generally reflect the severity of rutting; the higher these parameters are, the more severe the rutting is. Longitudinal parameters also show high correlations with the aforementioned parameters, especially on SR 26, indicating

that ruts grow not only in depth and area but also in length and volume. Temporal parameters share high correlation among themselves; however, they do not correlate well with other spatial parameters. This implies that temporal parameters represent certain rut features (e.g., deterioration) that other spatial parameters cannot. The correlation matrix of SR 275 rut parameters shows less correlation among them, which can be caused by the lack of diversity in rut conditions in this dataset.

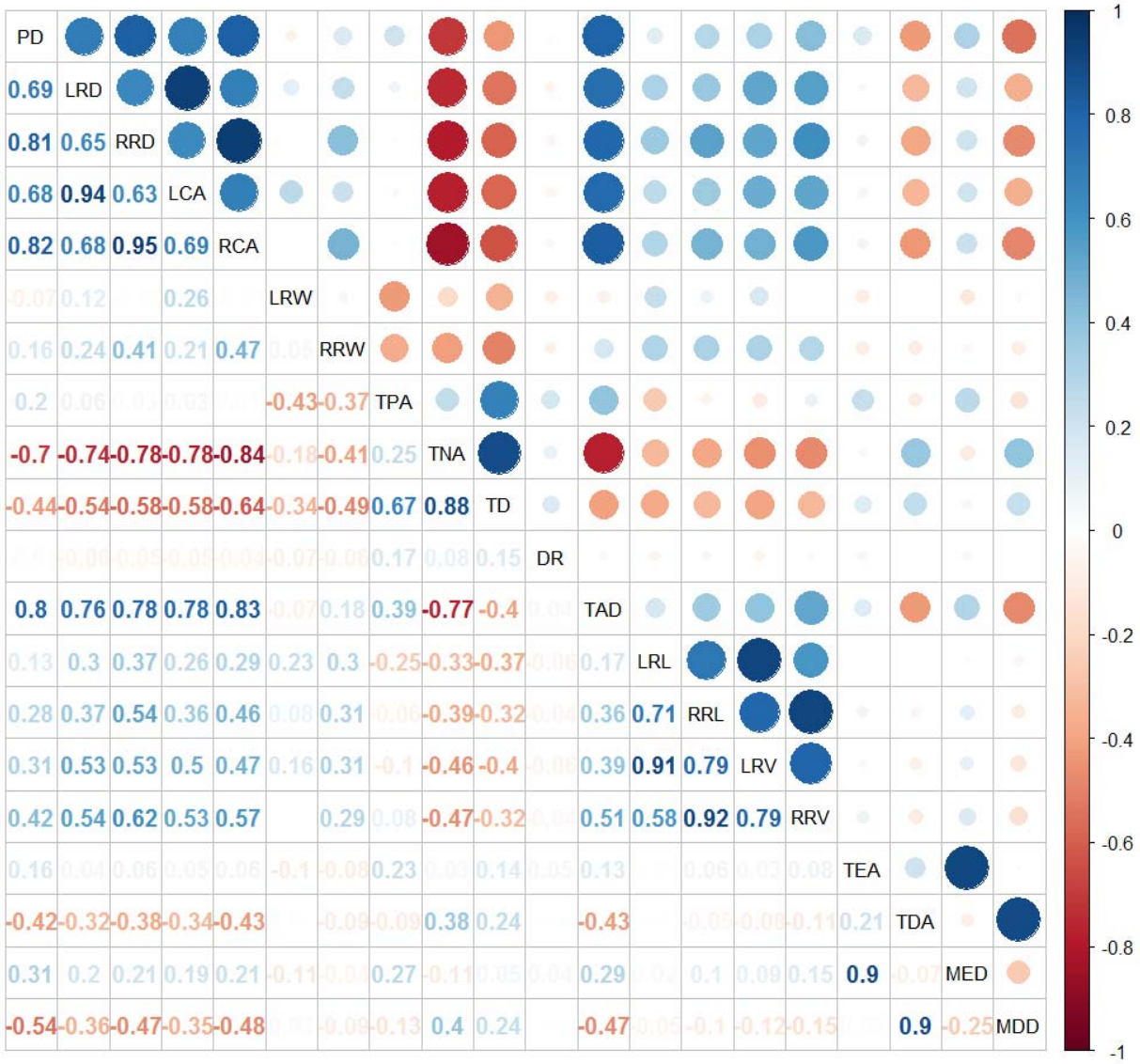


FIGURE 6.30: Correlation matrix of rut parameters on SR 26 Eastbound

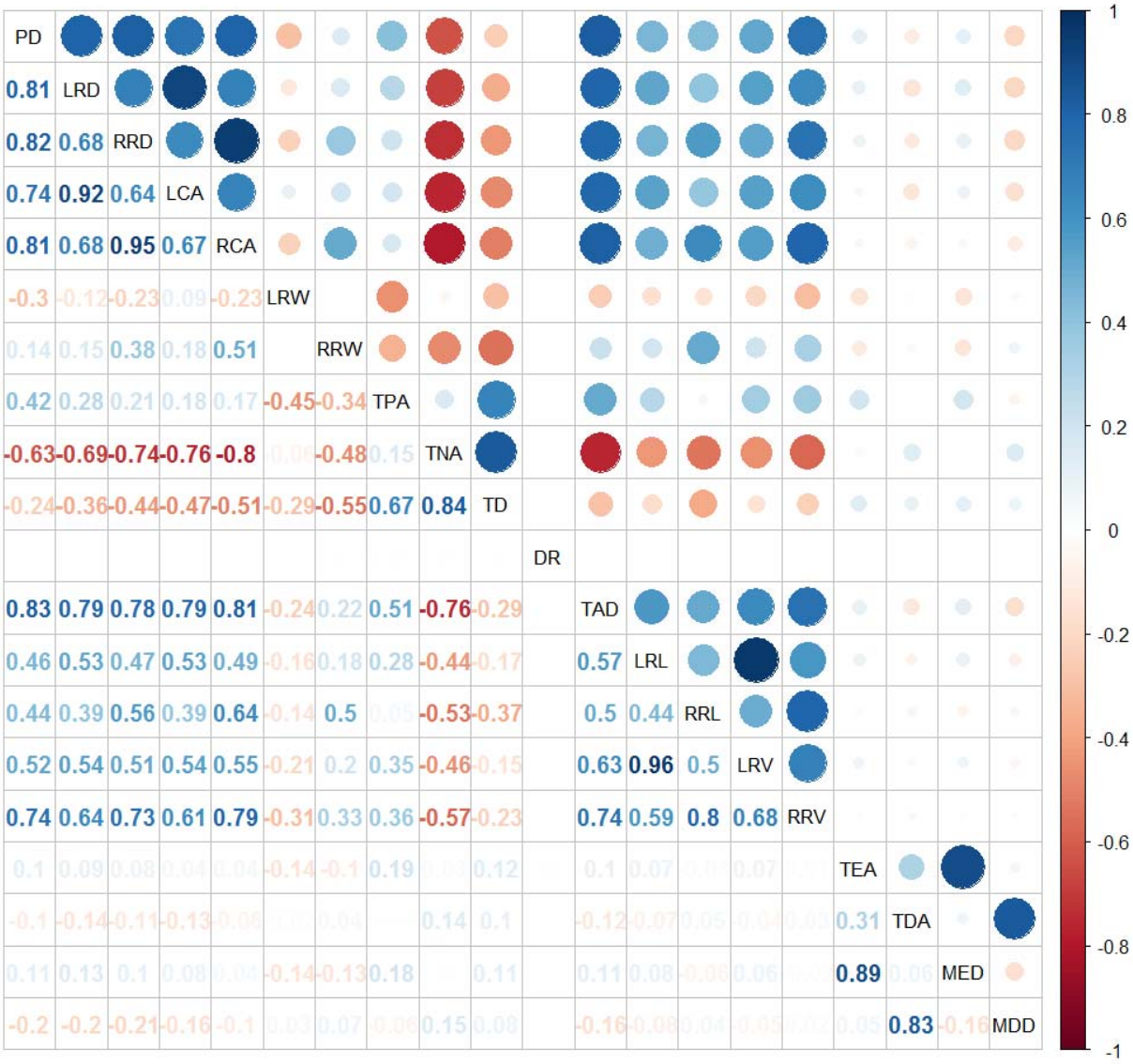


FIGURE 6.31: Correlation matrix of rut parameters on SR 26 Westbound

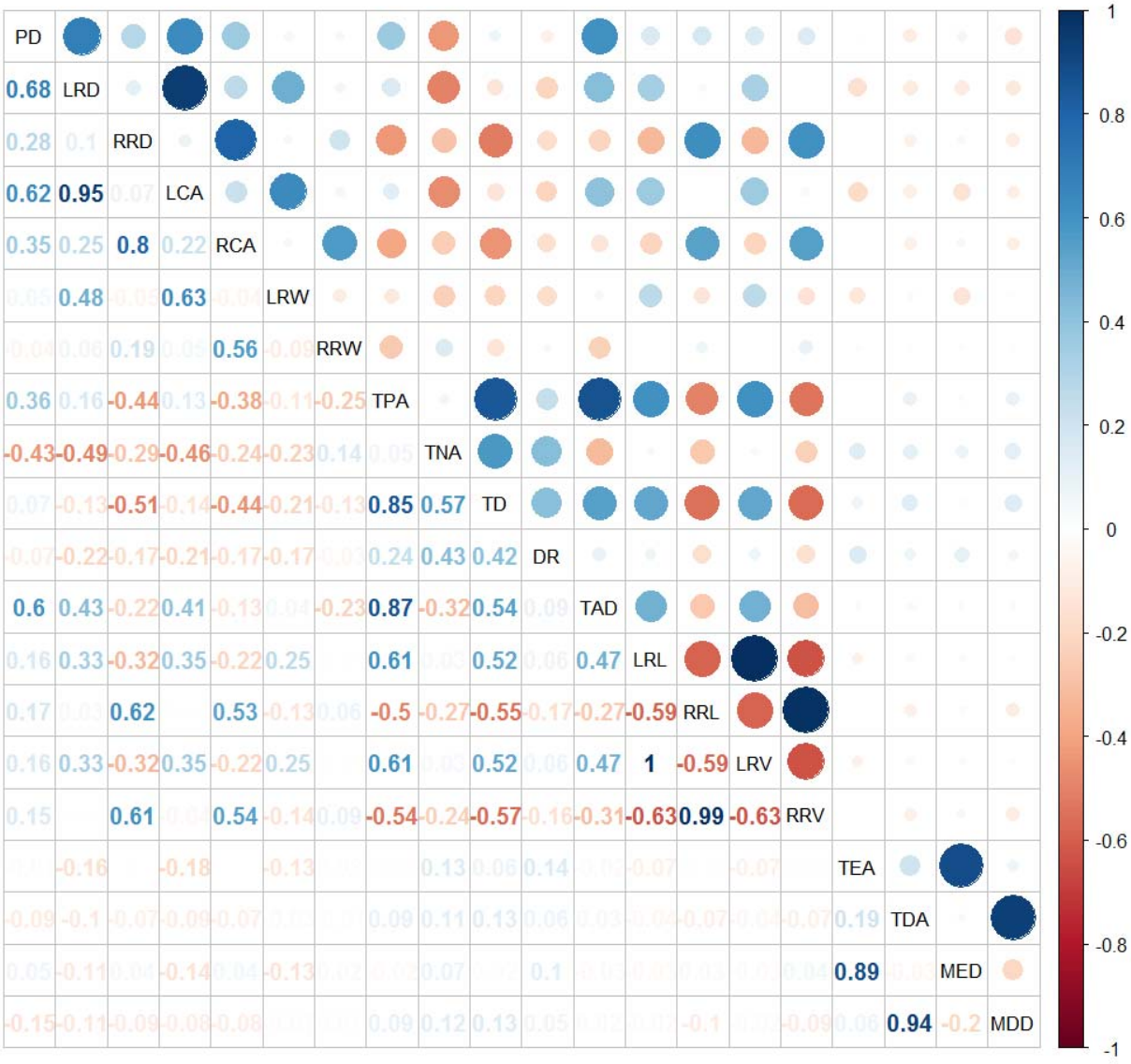


FIGURE 6.32: Correlation matrix of rut parameters on SR 275 Northbound

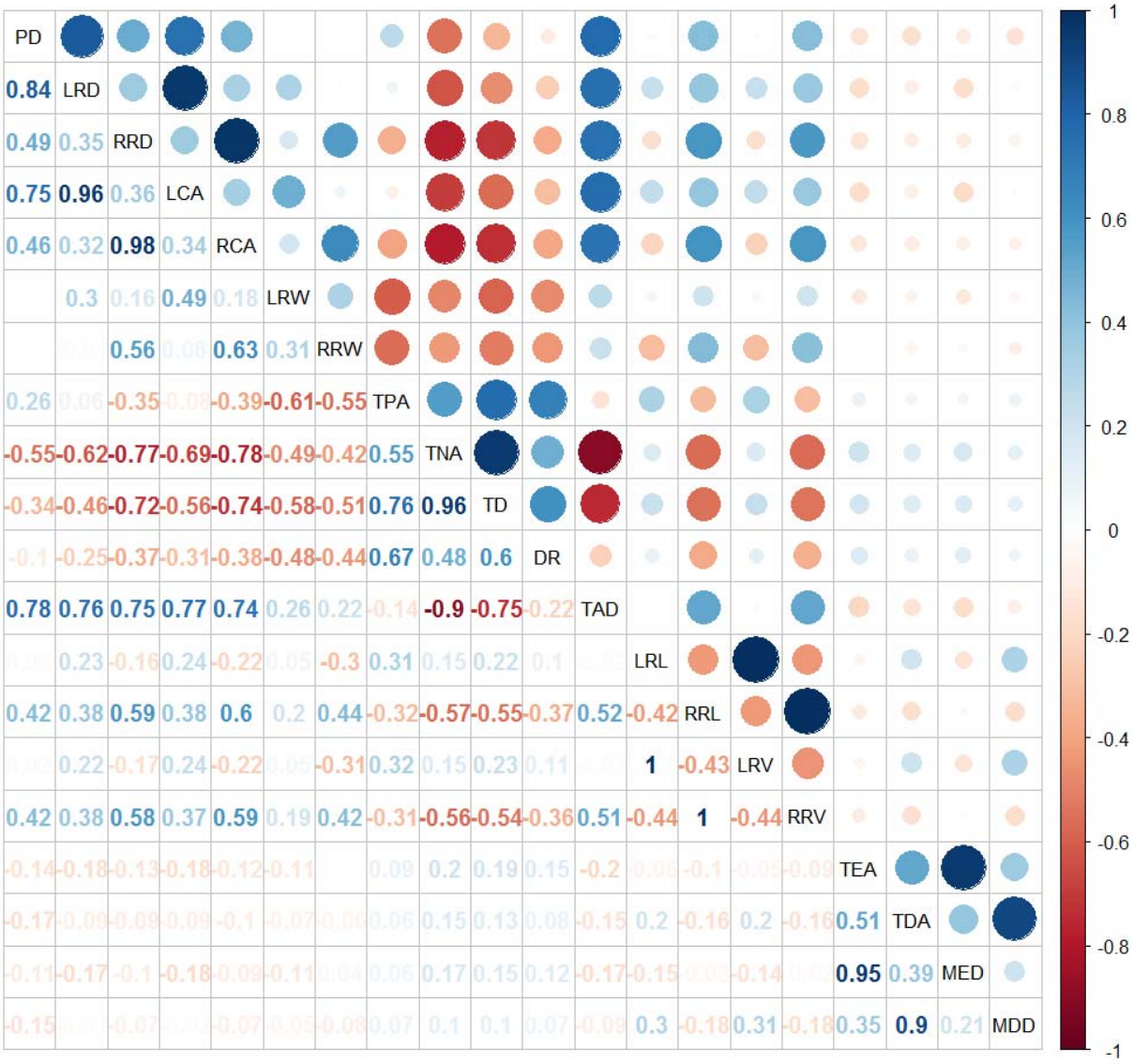


FIGURE 6.33: Correlation matrix of rut parameters on I-95 Southbound

5.2 Observed and Potential Issues of Parameters

Partial Lane Profile-based Parameters

Partial profile-based parameters in this study refer to rut parameters that are derived from part of the transverse lane profile. For example, for calculating right wheelpath rut parameters, as depicted in FIGURE 6.34, AASHTO PP69 rotates the transverse profile so that the elevations at the right edge (Spot 5) and the lane center (Spot 1) are zero. Rut parameters, such as right rut depth, width, and cross-sectional area, can be calculated using just the right half of the profile. Moreover, some studies and software for two sensor systems (e.g., LCMS) use profiles collected from the two sensors separately to calculate left and right rut parameters (Li 2012). For instance, profiles collected by the right sensor of the GTSV were used to calculate right rut depths in the study conducted by Li (Li 2012).

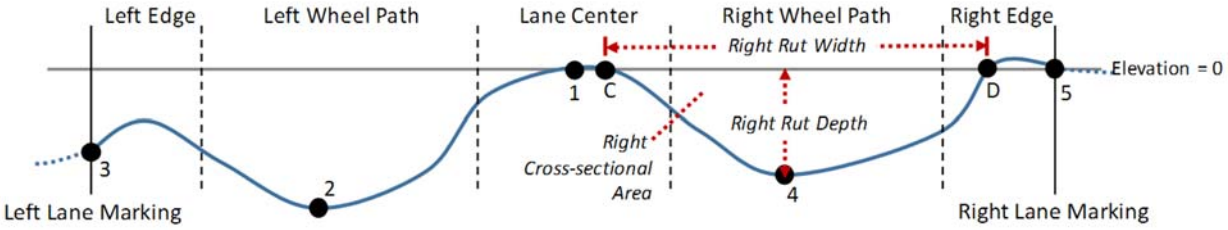


FIGURE 6.34: Deriving PP69 right rut parameters

A potential issue with calculating rut parameters using partial lane profiles is that these partial profiles may not reveal the entire shape if the rut is not entirely in the wheel path. For instance, FIGURE 6.35 shows an isolated rut on 11th Street in Atlanta. This rut is wide, and its basin exceeds the typical width of the wheelpath (e.g., 1,000 mm). A stitched transverse profile of this rut is shown in FIGURE 6.36. The straightedge rut

depth measured at this location was approximately 38.1 mm (1.5 in). However, if rut depth is calculated using the right half of the rotated profile (FIGURE 6.37), the calculated rut depth becomes 32.4 mm. Calculation of other right wheelpath rut parameters, such as rut width and cross-sectional area, would also result in underestimation. Moreover, similar results are expected if only the half profile collected by the right laser unit was used to calculate rut parameters.

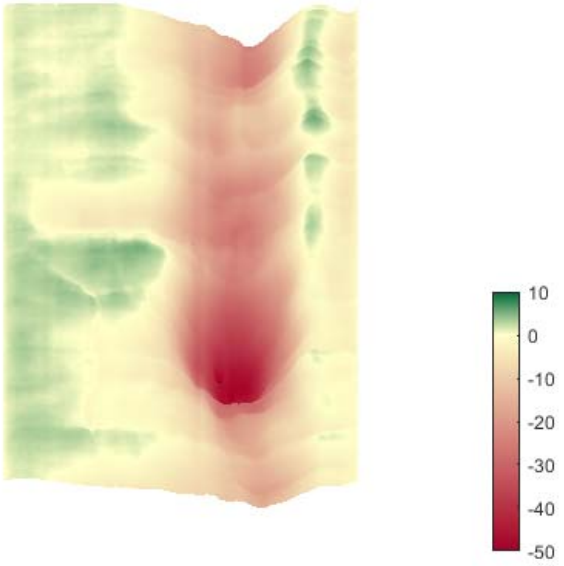


FIGURE 6.35: An Example of Wide Rutting on 11th Street in Atlanta

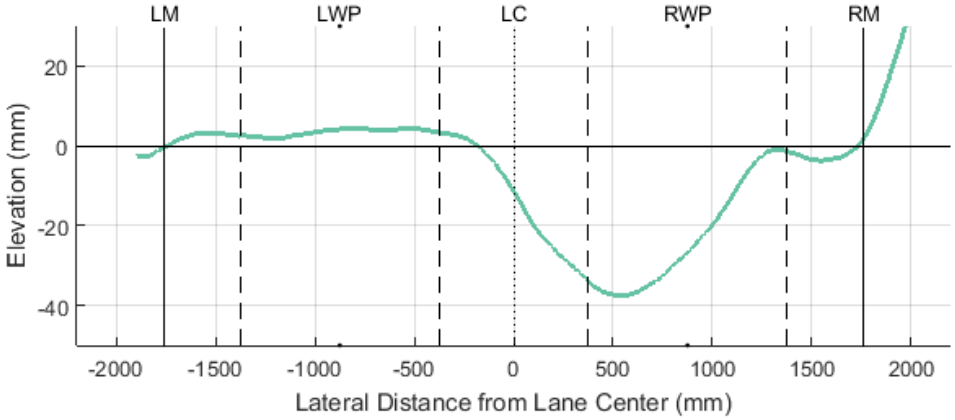


FIGURE 6.36: Transverse Profile of a Wide Rut in Right Wheelpath

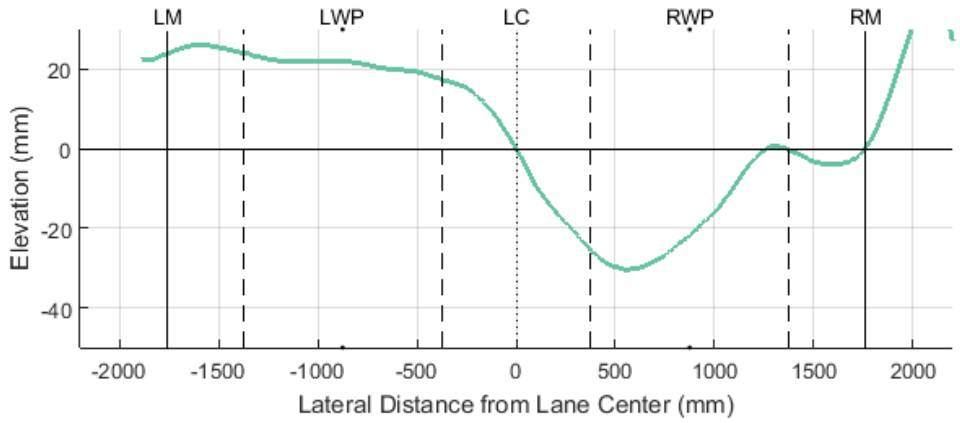
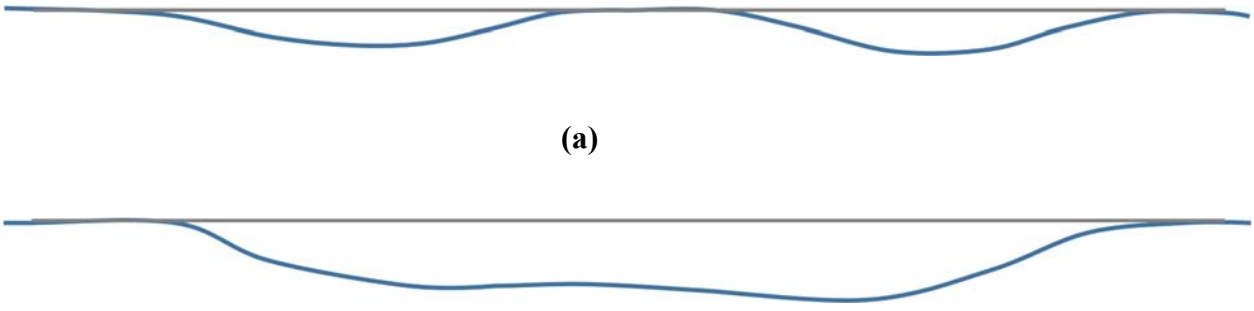


FIGURE 6.37: Rotated transverse profile for calculating right rut parameters

Another potential issue with calculating rut parameters using partial lane profiles is that the same measurement may be the outcome of different rut shapes. As depicted in FIGURE 6.38, the two profiles show two distinct shapes of rutting; however, their left rut parameters (e.g., depth, width, and cross-sectional area) may be the same based on PP69 definition.



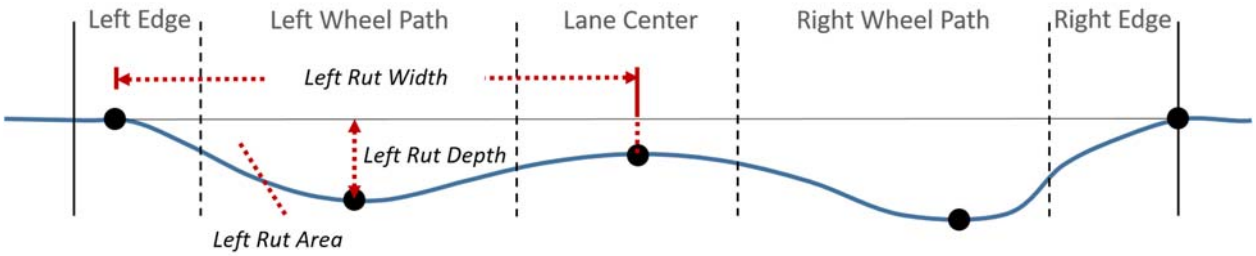
(a)

(b)

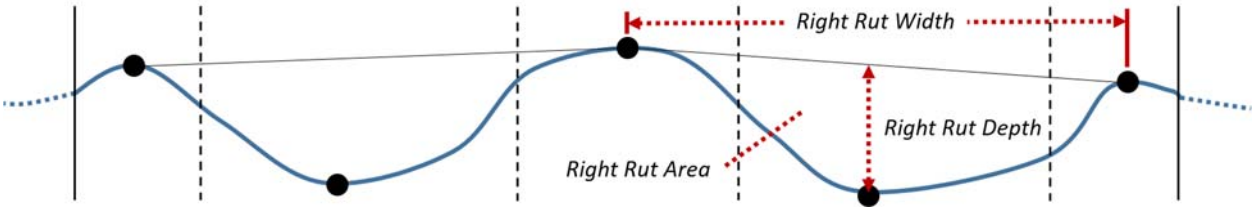
FIGURE 6.38: Illustration of different rut shapes with same left rut depth

A possible solution for the above issues with regard to partial lane profile-based parameters is that, instead of rotating the transverse profile, the normalized lane profile can be used to derive wheelpath-based parameters. The following discussion describes the solution in more details.

As depicted in FIGURE 6.39, the 5 zones defined in PP69 are adapted to determine the lowest point (valley) in each wheelpath zone and the highest point (peak) in each of the other three zones. Since left rut parameters and right rut parameters can be calculated in the same manner, the following figures describe how left profile-based parameters are defined and calculated.



(a) Lower Center Elevation



(b) Higher Center Elevation

FIGURE 6.39: Proposed of wheelpath-specific rut parameters

Reference Surface Plane Assumptions

Positive and negative areas defined in the literature were determined by connecting both edges of the lane with an imaginary straight line, as depicted in FIGURE 6.40. However, this reference line does not represent the actual horizon plane or the original pavement surface. Consequently, the derived positive and negative areas and the total distortion and the distortion ratio cannot represent the severity of rutting well, and direct comparison of these parameters may not be meaningful.

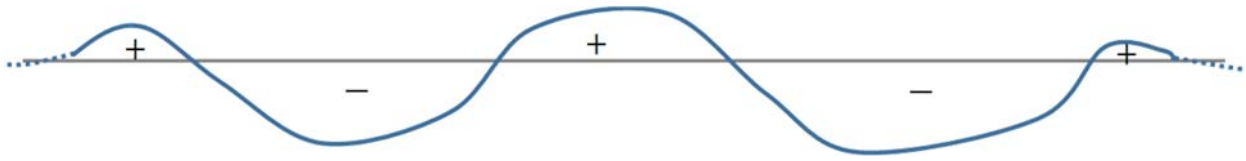


FIGURE 6.40: Illustration of positive and negative areas

6. Summary

In this chapter, 3D rut shape and its deterioration behaviors were characterized and analyzed at multiple scales using long-term 3D pavement data. Rut parameters, including transverse profile-based parameters, longitudinal parameters, and temporal parameters, were defined and proposed. Descriptive statistics and 2D and 3D visualizations were used to analyze the deterioration behaviors of rutting. Some key findings of this chapter are summarized as follows:

- Transverse profile-based rut parameters, including rut depth, rut cross-sectional area, percent deformation, and total absolute distortion show good correlation among each other and provide consistent information about the conditions of ruts.
- Longitudinal parameters also showed similar trends as the aforementioned parameters, indicating that ruts not only grow in depth and area but longitudinally in length and volume.
- Temporal parameters provide a direct means to quantify the deterioration of rutting. These parameters can be very useful, especially when the rate of deterioration is high.
- The multi-scale analysis results show that seasonal variation can have an appreciable effect on the deterioration of ruts. Temporal parameters derived from a shorter period (e.g., a few months), can be affected by the seasonal variation. This finding suggests that for different deterioration analysis applications, different analysis periods should be applied to obtain consistent results.
- The comparison among three different routes showed that traffic and roadway characteristics also play an important role in the deterioration of ruts. Other factors, such as the age and design of the pavement, can also contribute to the actual deterioration behavior of ruts.
- 2D and 3D visualization of ruts at the individual level shows the importance and benefits of having registered long-term 3D pavement data. Details of how ruts deteriorate, e.g., the development of dual-wheel shape ruts and the growth in rut length and volume, were able to be identified in the selected sections. This

information can further support the diagnosis of rutting and inform data-drive maintenance decisions.

- Some parameters, such as PP69's partial lane profile-based parameters, total positive area, total negative area, distortion ratio, and total distortion, can possibly introduce larger variation because they are calculated using a reference that tends to be variable. Future research can be conducted to further analyze the possible effect of these parameters.

References

Alard, C. (2000). "Image subtraction using a space-varying kernel." *Astronomy and Astrophysics Supplement Series*, EDP Sciences, 144(2), 363–370.

Alard, C., Lupton, R. H., Alard, C. . & G. J., Alcock, C. . et al., Aubourg, E. . et al., Crotts, A. . & T. A., Kochanski, J. . G. P. . T. A. . & F. P. F., Phillips, A. C. . & D. L. E., Renson, P., Roddier, F. . C. L. . G. J. E. . & S. A., Schechter, P. . & M. M., Shao, M. . & C. M. M., Tomaney, A. . & C. A., and Udalski, A. . et al. (1998). "A Method for Optimal Image Subtraction." *The Astrophysical Journal*, IOP Publishing, 503(1), 325–331.

Bertozzi, M., and Broggi, A. (1998). "GOLD: A parallel real-time stereo vision system for generic obstacle and lane detection." *IEEE transactions on image processing*, article, IEEE, 7(1), 62–81.

Brown, L. G. (1992). "A Survey of Image Registration Techniques." *ACM Comput. Surv.*, article, ACM, New York, NY, USA, 24(4), 325–376.

Chang, T.-H., Hsu, C.-S., Wang, C., and Yang, L.-K. (2008). "Onboard Measurement and Warning Module for Irregular Vehicle Behavior." *IEEE Transactions on Intelligent Transportation Systems*, 9.

Dawn, S., Saxena, V., and Sharma, B. (2010). "Remote Sensing Image Registration Techniques: A Survey." *Proceedings of the 4th International Conference on Image and Signal Processing*, ICISP'10, inproceedings, Springer-Verlag, Berlin, Heidelberg, 103–112.

Fonseca, L. M. G., and Manjunath, B. S. (1996). "Registration Techniques for Multisensor Remotely Sensed Imagery." *Journal of Photogrammetry Engineering and Remote Sensing*, article, 62(9), 1049–1056.

Freeborough, P. A., Woods, R. P., and Fox, N. C. (1996). "Accurate registration of serial 3D MR brain images and its application to visualizing change in neurodegenerative disorders." *Journal of computer assisted tomography*, article, LWW, 20(6), 1012–1022.

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J.-B., Heather, J. D., and Frackowiak, R. S. J. (1995). "Spatial registration and normalization of images." *Human Brain Mapping*, Wiley Subscription Services, Inc., A Wiley Company, 3(3), 165–189.

Kano, A., Doi, K., MacMahon, H., Hassell, D. D., and Giger, M. L. (1999). "Digital image subtraction of temporally sequential chest images for detection of interval change." *Medical Physics*, American Association of Physicists in Medicine, 21(3), 453.

Li, F. (2012). “A Methodology for Characterizing Pavement Rutting Condition using Emerging 3D Line Laser Imaging Technology.” Ph.D. Dissertation, Georgia Institute of Technology.

Lowe, D. G. (1999). “Object recognition from local scale-invariant features.” *Proceedings of the Seventh IEEE International Conference on Computer Vision*, IEEE, 1150–1157 vol.2.

Lucas, B. D., and Kanade, T. (1981). “An iterative image registration technique with an application to stereo vision.” *International Joint Conference on Artificial Intelligence, IJCAI*, inproceedings, 674–679.

McCall, J. C., and Trivedi, M. M. (2004). “An integrated, robust approach to lane marking detection and lane tracking.” *Intelligent Vehicles Symposium, 2004 IEEE*, inproceedings, 533–537.

McGill, R., Tukey, J. W., and Larsen, W. A. (1978). “Variations of Box Plots.” *The American Statistician*, article, [American Statistical Association, Taylor & Francis, Ltd.], 32(1), 12–16.

Otsu, N. (1979). “A Threshold Selection Method from Gray-Level Histograms.” *IEEE Transactions on Systems, Man, and Cybernetics*, article, 9(1), 62–66.

Qiu, S. (2013). “Measurement of Pavement Permanent Deformation based on 1mm 3D Pavement Surface Model.” Oklahoma State University.

Simpson, A. (1999). “Characterization of Transverse Profile.” *Transportation Research Record: Journal of the Transportation Research Board*, article, Transportation

Research Board of the National Academies, 1655, 185–191.

Transfund New Zealand. (1997). *RAMM - Road Condition Rating and Roughness Manual*. book, Transfund New Zealand, Wellington.

Tsai, Y. (James), Hu, Z., and Alberti, C. (2010). "Detection of Roadway Sign Condition Changes using Multi-Scale Sign Image Matching (M-SIM)." *Photogrammetric Engineering & Remote Sensing*, 76(4), 391–405.

White, T. D., Haddock, J. E., Hand, A. J. T., and Fang, H. (2002). "NCHRP Report 468: Contributions of Pavement Structural Layers to Rutting of Hot-Mix Asphalt Pavements." report, Washington D.C.

Wu, Y. J., Lian, F. L., and Chang, T. H. (2006). "Traffic Monitoring and Vehicle Tracking using Roadside Cameras." *2006 IEEE International Conference on Systems, Man and Cybernetics*, inproceedings, 4631–4636.

Chapter 7 Implementation of Research Outcomes

We give thanks to USDOT/OST-R and GDOT's sponsorship of RS-GAMS Phase 1 and Phase 2. These two research projects focused on the fundamental validation of CRS&SI technologies, including 3D line laser and mobile LiDAR, to enhance pavement, traffic sign, pavement marking, and roadway geometry data inventory, condition assessment, and management. The benefit of these two research projects is profound because the research results have been implemented in numerous completed and/or ongoing projects for GDOT. As part of our effort to quantify the benefits from this research project, as well as those from Phase 1 (Tsai & Wang, 2016a), this chapter presents other completed and ongoing projects that are based on the research outcomes from this and the Phase 1 project. In addition, some future projects are also proposed.

1. Research Focuses

As shown in FIGURE 1.1, RS-GAMS Phase 1 and Phase 2 covered a broad area of transportation assets, including traffic sign, asphalt pavement, concrete pavement, pavement marking, cross slopes, horizontal curvature, and pavement width. Table 7.1 lists the research focuses in each project. For convenience of reference, each research focus was assigned a research focus number.

RS-GAMS focused on the validation of using sensing data collected by the GTSV to scientifically enhance various transportation asset data inventory and condition assessment. The research approaches include 1) laboratory tests in a well-controlled environment; 2) field tests on carefully selected test road sections; 3) acquisition of data from other transportation agencies; 4) development of algorithms and applications; 5)

statistical analysis; and 6) extensive discussion with TAC members and other participating transportation agencies. The validated results can be implemented in transportation agencies' practice, and/or applied in other relative research projects.

Table 7.1: Research Focuses in RS-GAMS Phase 1 and Phase 2

Research Project	Research Focus ID	Research Focus
RS-GAMS Phase 1	RF(1,2)-1	Development of Georgia Tech Sensing Vehicle (GTSV) and streamline data collection procedures
	RF(1)-1	Network-level and isolated rutting measurement using 3D laser data
	RF(1)-2	Quantitative performance measure of automatic pavement crack detection
	RF(1)-3	Automatic asphalt pavement crack detection using 3D laser data
	RF(1)-4	An enhanced sign inventory procedure
	RF(1)-5	Sign retroreflectivity condition assessment using mobile LiDAR
	RF(1)-6	Prototype GIS-based sign management
	RF(1,2)-2	Long-term monitoring of pavement conditions
RS-GAMS Phase 2	RF(1,2)-1	Improvement of GTSV
	RF(2)-1	Automatic asphalt pavement crack classification
	RF(2)-2	Concrete pavement distress detection using 3D laser data
	RF(2)-3	Pavement marking retroreflectivity condition assessment using mobile LiDAR
	RF(2)-4	Extraction of roadway geometric characteristics (cross slopes, horizontal curvature, and pavement width)
	RF(2)-5	Prototype GIS-based asset management for concrete pavement and curvature
	RF(1,2)-2	Long-term monitoring of pavement conditions

2. Implementation and Applications

The Georgia Tech research team has actively worked with GDOT's Office of Research and other offices to further implement the research outcomes through numerous GDOT sponsored research projects. Thus, the fundamental research conducted in RS-GAMS can be actually implemented in GDOT's practices to enhance GDOT's transportation asset, e.g. pavements, signs, roadway safety, data inventory, and condition assessment. Table 7.2 lists some completed, ongoing, and proposed projects that have implemented and/or applied the research outcomes as shown in Table 7.1. The following will briefly present how the RS-GAMS research outcomes are implemented and/or applied in other

research projects. Please note that the project RP 13-18 is an extension to RS-GAMS because, in this project, raveling detection algorithms were developed and validated. Nevertheless, the GTSV, data collection, and data preprocessing procedures developed in RS-GAMS were applied in RP 13-18. In addition, the result outcomes from RP 13-18 were successfully implemented in conducting pavement condition evaluation for Georgia's interstate highways (RP 15-11).

Table 7.2: Completed, Ongoing, and Proposed Projects

RP #	Title	Status	Applied Research Focuses
13-18	Development of an Asphalt Pavement Raveling Detection Algorithm Using Emerging 3D Laser Technology and Macrotexture Analysis	Completed	RF(1,2)-1
13-19	Enhancing GDOT's Jointed Plain Concrete Pavement (JPCP) Rehabilitation Program Using Emerging 3D Sensing Technology and Historical Concrete Condition Survey Data	Completed	RF(1,2)-1, RF(2)-2
14-37	Next Generation Crack Sealing Planning Tool for Pavement Preservation	Completed	RF(1)-3, RF(2)-1
15-11	Implementation of Automatic Sign Inventory and Pavement Condition Evaluation on Georgia's Interstate Highways	Completed	RF(1,2)-1, RF(1)-1, RF(1)-3, RF(1)-4, RF(2)-1, RP 13-18
15-04	Developing Georgia's High Friction Surface Treatment (HFST) Program - HFST Site Characteristics (HFST-SC) Data Collection and Analysis	Ongoing	RF(1,2)-1, RF(2)-4
15-05	Curve Identification for High Friction Surface Treatment (HFST) Installation Recommendation	Ongoing	RF(1,2)-1, RF(2)-4
N/A	Validating Change of Sign and Pavement Conditions and Evaluating Sign Retroreflectivity Condition Assessment on Georgia's Interstate Highways using 3D Sensing Technology	Proposed	RF(1,2)-1, RF(1)-1, RF(1)-3, RF(1)-4, RF(1)-5, RF(2)-1, RP 13-18

- **RP 13-18: Development of an Asphalt Pavement Raveling Detection Algorithm Using Emerging 3D Laser Technology and Macrotexture Analysis (Tsai & Wang, 2015)**

  Raveling has become the most predominant and concerning distress on Georgia's interstate highways with open graded friction course (OGFC), which is a progressive disintegration of a pavement surface. Raveling shortens pavement life and results in

various safety concerns, such as flying stones that damage windshields and vehicle bodies, and rough and uneven pavement surface that increase road/tire noise and degrade riding safety. It is critical to detect raveling at its early stage in order to apply inexpensive preventive maintenance methods, such as fog sealing. However, the current visual method makes it difficult for an accurate raveling survey because the change of lighting conditions and vehicle speed significantly affects a survey's perception of the pavement surface texture. Thus, there is an urgent need for an accurate and objective raveling survey. This research project, sponsored by the National Cooperative Highway Research Program (NCHRP) Innovations Deserving Exploratory Analysis (IDEA) program and GDOT, comprehensively tested and validated the automatic raveling detection, classification, and measurement algorithms using 3D laser technology. The GTSV, data collection, and data preprocessing procedures developed in RS-GAMS were applied in this research project. The raveling condition survey protocol used in GDOT was adopted in the testing and validation, though it can be easily extended to other highway agencies' protocols. Though this research project is not a direct implementation of RS-GAMS, its research outcomes have been successfully implemented in RP 15-11 (see Table 7.2).

- **RP 13-19: Enhancing GDOT's Jointed Plain Concrete Pavement (JPCP) Rehabilitation Program Using Emerging 3D Sensing Technology and Historical Concrete Condition Survey Data (Tsai, 2016)**

  The JPCPs have carried more than 20% of Georgia's truck traffic. A majority of these JPCPs (e.g., I-16) have been in service for more than four decades with minor

maintenance and no or little rehabilitation. These aging pavements are in great need of maintenance, rehabilitation, and reconstruction (MR&R), including actions such as broken slab replacement, grinding, re-sealing, etc., or full lane replacement. Faced with limited funding and the increasing needs for aged JPCP MR&R, GDOT needs to enhance its data-driven JPCP maintenance and management program to ensure its highway maintenance program goals and objectives are met, including the concrete pavement condition evaluation system (CPACES), MR&R practices, and quantity estimation, all in support of JPCP MR&R planning and programming. For this purpose, this research project refined the JPCP distress protocol based on different levels of severity (e.g., shattered slab) to support slab replacement prioritization when funding is limited, accurately estimated the MR&R quantity (e.g., slab replacement) to prevent project overruns for better budget planning, and predicted the future JPCP condition and MR&R needs for better planning. An enhanced slab replacement quantity estimation method was developed and successfully demonstrated a promising capability to effectively identify distresses and accurately estimate slab replacement quantities using 3D laser data (see FIGURE 7.1). Results show a significant improvement (approximately 26%) in the accuracy of slab replacement quantity estimation as compared to the current windshield survey. This method is capable of simulating different slab replacement alternatives, e.g., replacing only the severe distresses (e.g., shattered slab) or all distressed slabs, and calculating corresponding costs. This allows OM to analyze MR&R alternatives based on different treatment criteria and estimate corresponding costs.

- **RP 14-37: Next Generation Crack Sealing Planning Tool for Pavement Preservation (Tsai & Wang, 2016b)**

  State DOTs like GDOT have insufficient work forces to conduct routine pavement preservation, e.g., crack sealing. Therefore, GDOT is considering outsourcing its crack sealing projects. However, it has become a challenge to select the most cost-effective projects and accurately estimate project costs, which have hindered the outsourcing effort. To address the above issues, this research project proposes a systematic framework to study the cost-effectiveness of crack sealing and crack filling (CS/CF) and incorporate CS/CF planning into a pavement management system (PMS). The proposed methodology has been evaluated using different case studies and has demonstrated promising results. In this research, the crack detection and classification methods were carried over from the RS-GAMS. In the next step, we need to work with Office of Maintenance to evaluate the implementation procedures and specifications.

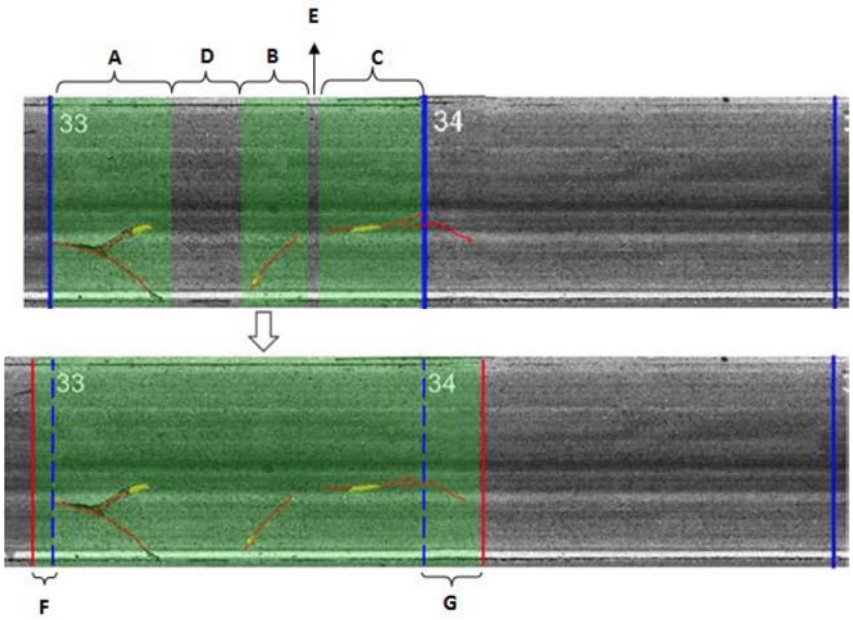


FIGURE 7.1: Example of processing slab replacement (Tsai, 2016)

- **RP 15-11: Implementation of Automatic Sign Inventory and Pavement Condition Evaluation on Georgia's Interstate Highways (Tsai et al., 2017)**

  To establish a data-driven sign asset management, it is required that there be a complete sign inventory, including locations, types, and conditions. The complete sign inventory is also very valuable to assist timely sign maintenance operations, e.g. determining the type of maintenance work based on sign conditions and planning field trip in advance based on sign locations and pictures of the signs. Since GDOT has not had a comprehensive sign inventory in the past, this research project has established a complete and comprehensive sign inventory for interstate highways, covering 22,344 signs, along with their MUTCD codes, locations, and visual conditions, in which the enhanced sign inventory procedures developed in RS-GAMS have been successfully implemented (see FIGURE 7.2). The Georgia Tech research team has also worked with a GDOT Area Office to conduct a pilot test of a timely sign maintenance program, which can be further implemented in GDOT's working districts.

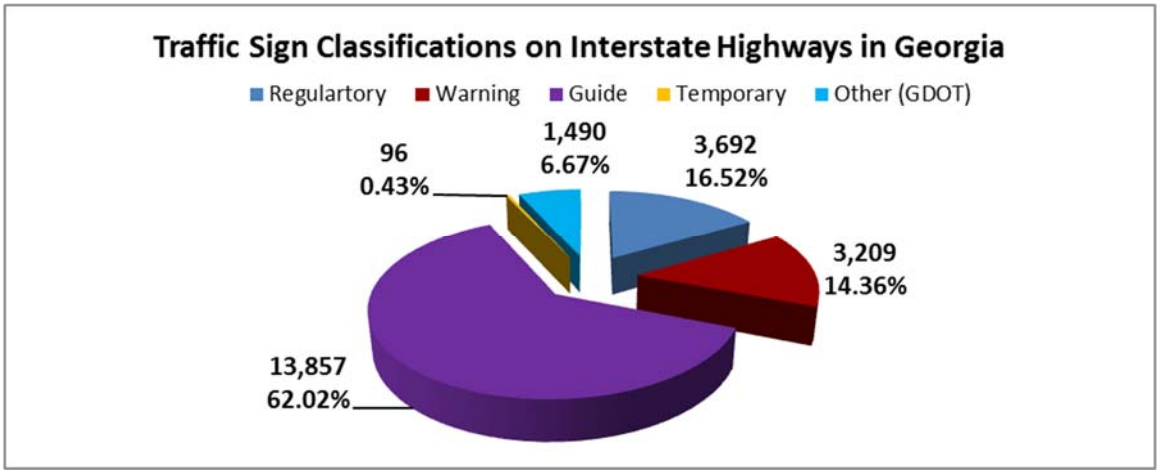


FIGURE 7.2: Distribution of traffic signs on interstate highway (Tsai et al., 2017)

GDOT has collected statewide pavement condition data since 1986 based on its distress protocol, PACES. The current manual method works well on state-maintained, non-interstate roadways, but it is difficult and, sometimes, prohibitive for surveying interstate highways due to the safety concerns caused by limited parking space for field engineers and high-speed, high-volume traffic. Therefore, GDOT needs a safer and more cost-effective method to complement its manual method of data collection on interstate highways. By implementing the research outcomes in RS-GAMS (e.g., rutting measurement, crack detection and classification, and raveling detection and measurement), a comprehensive pavement condition evaluation for all asphalt-surfaced interstate highways in Georgia was established; it covers 1,513 miles of asphalt pavements on interstate highways, and rated them using streamlined pavement condition evaluation procedure and the 3D laser data (see FIGURE 7.3).

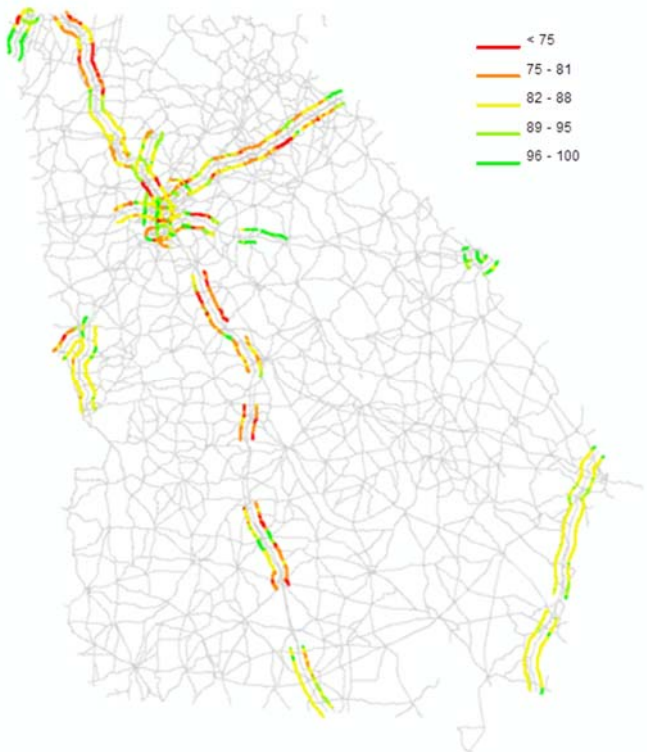


FIGURE 7.3: COPACES ratings on interstate highways (Tsai et al., 2017)

- **RP 15-04: Developing Georgia's High Friction Surface Treatment (HFST) Program - HFST Site Characteristics (HFST-SC) Data Collection and Analysis**

  HFST is an effective safety treatment method that enhances pavement friction to prevent run-off-road crashes. However, it is a relatively new and expensive countermeasure compared to the traditional ones, such as chevrons, rumble strips, etc., and the process for selecting the most effective HFST sites is still evolving. As a leading state in HFST implementation with millions of dollars invested (and more will be invested) for HFST, GDOT needs to determine the suitable sites for HFST to maximize the return on investment. This ongoing research project/study is to enhance GDOT's HFST program by developing a systematic site selection procedure and developing a procedure that establishes a location-referenced, detailed, HFST site characteristics information (e.g., curve radius, super-elevation, vertical grade, pavement texture, signage, etc.) method. Emerging sensing technologies, including 2D imaging, lasers, 3D LiDAR, IMU, and GPS/GIS technologies with machine learning and artificial intelligence will be applied and have already been comprehensively validated in RS-GAMS.

- **RP 15-05: Curve Identification for High Friction Surface Treatment (HFST) Installation Recommendation**

  According to the Every Day Counts (EDC) initiative, more than 25 percent of fatal crashes in the United States occur on horizontal curves. Horizontal curve information, including curve location, radius, the point of a curve (PC), and the point of a tangent (PT), is essential for curve safety analyses to identify sharp curves and determine appropriate countermeasures. However, transportation agencies lack an

effective method for obtaining such curve information. The objectives of this ongoing research project are to develop and deploy a curve identification tool based on the algorithms developed by the Georgia Tech research team to automatically extract detailed horizontal curve information, including the curve radius, PC, and PT, of every curve along a corridor using GPS data and GIS data that are widely available for transportation agencies. The research team further developed two applications to demonstrate the capacity of the developed tool. The extracted curve information, including PC and PT, is essential for recommending the starting and ending points for HFST installation. Also, the developed tool for curve information extraction will enable transportation agencies to establish a large-scale horizontal curve inventory by using widely available GPS data to support HFST site selection.

- **Proposed: Validating Change of Sign and Pavement Conditions and Evaluating Sign Retroreflectivity Condition Assessment on Georgia's Interstate Highways using 3D Sensing Technology**

  As a continuous effort to RP 15-11, this proposed project is to track the existence and condition change of signs and pavement surface distresses on interstate highways. Unlike a new inventory process, the second-round data collection should be faster and more accurate because the inventory of sign and pavement surface distresses has been previously done. In the second round, the previous data can be utilized to speed up the data collection in the second round and to control the second-round data quality. The major objective of this research project is to validate the above concept and, in the meantime, establish a tracked condition change of signs and pavement surface distresses.

Sign retroreflectivity conditions are critical for nighttime driving safety. In GDOT's current practice, nighttime visual inspection is applied to assess sign retroreflectivity conditions, but this is a subjective, inaccurate, and time consuming method. In RS-GAMS, the Georgia Tech research team has proposed and validated a systematic approach to assess sign retroreflectivity conditions using mobile LiDAR. In comparison to the nighttime visual inspection method, the proposed method is more objective, more accurate, and faster. In this proposed research project, the mobile-LiDAR-based sign retroreflectivity condition assessment will be implemented. After all the signs on interstate highways are inventoried, the mobile LiDAR data will be used to assess their retroreflectivity conditions.

3. Summary

The research conducted in the two phases of RS-GAMS has produced a solid foundation for us to promote the use of CRS&SI technologies in GDOT for enhancement of pavement, sign, and other safety-related asset data collection, condition assessment, and maintenance decision-making. The full implementation and application of the RS-GAMS research results has the potential to pave a new way for GDOT to take advantage of the developed and validated 3D sensing technologies for determining the right projects, right timing, and right treatments for pavement and sign maintenance and safety improvement. The enhanced pavement condition evaluation method will be very valuable and enable GDOT to meet the new pavement performance measures.

References

Tsai, Y. and Wang, Z. (2015) “Detecting Asphalt Pavement Raveling Using Emerging 3D Laser Technology and Macrotexture Analysis,” Final Report GA-14-1308.

Tsai, Y. (2016) “Enhancing GDOT’s Jointed Plain Concrete Pavement (JPCP) Rehabilitation Program Using Emerging 3D Sensing Technology and Historical Concrete Condition Survey Data.” Final Report FHWA-GA-16-1319.

Tsai, Y. C. and Wang, Z. (2016a) “A Remote Sensing and GIS-Enabled Asset Management System (RS-GAMS)” Final Report GA-15-1008.

Tsai, Y. and Wang, Z. (2016b) “Next Generation Crack Sealing Planning Tool for Pavement Preservation,” Final Report GA-15-1437.

Tsai, Y., Wang, Z., and Ai, C. (2017) “Implementation of Automatic Sign Inventory and Pavement Condition Evaluation on Georgia’s Interstate Highways,” Final Report FHWA-GA-16-1511.

Chapter 8 Conclusions and Recommendations

This research project scientifically validated the use of CRS&SI technologies, including 3D line laser and mobile LiDAR, to improve the asphalt pavement crack classification, concrete pavement distress detection, and pavement marking retroreflectivity condition assessment. Through the analysis of the long-term monitoring data, a comprehensive study has been performed on cracking and rutting deterioration. The following discussion summarizes the major research findings and recommends future study projects.

1. Conclusions

The following conclusions are categorized in terms of the research topics.

- **Automatic asphalt pavement crack classification**

  This research validated the performance of an automatic crack classification algorithm that uses 3D laser data and developed previously by the PI. The algorithm is based on a multi-scale CFE model and uses crack detection results, i.e. crack maps, as inputs, which have been validated in RS-GAMS Phase 1. The classification of two of the most commonly occurring cracks, load cracking and block cracking, defined in GDOT's pavement distress survey manual, PACES, was implemented, tested, and validated. The promising results demonstrate that the algorithm is capable of transforming raw sensing data and detected crack maps into useful decision-support information, including crack types, severity levels, and extents.

In a test on 1,069 pavement 3D laser images (each image covers a pavement section about 5-meters long and 4-meters wide), GDOT pavement engineers visually reviewed each image and established the ground truth. Based on the comparison between the ground truth and the automatically classified results, the algorithm showed an accuracy of 92.2% on classifying load cracking at four severity levels and 98.1% on classifying block cracking at three severity levels. Another test was conducted on ten 100-ft test sections that were selected on State Route (SR) 236, SR 275, and SR 67 in Georgia. In each test section, GDOT pavement engineers visually identified the crack types, severity levels, and extents in the field, which was used as ground truth. Among ten test sections, four were surveyed by accurate measurements using a measuring wheel, while the other six sections were surveyed by visual estimation following GDOT's current survey practices. Then, comparison was made on deducts derived from the automatic crack classification and the field visual survey. For the wheel-measured sections, the average absolute difference of total deducts was 3.25 out of 100 (a pavement rating is between 0 and 100), and for the visually-estimated sections, the average absolute difference was 5 out of 100. Both differences were within the error tolerance based on GDOT's current practice (5 out of 100).

The validation results show that the use of 3D laser data and the corresponding algorithm could improve the productivity and efficiency of collecting pavement distress information. Moreover, the fine-grained sensing data also opened the opportunity to improve existing pavement management by adding more detailed decision-support information that, previously, could not have been acquired.

- **Concrete pavement distress detection**

  This study validated the detection and measurements of various concrete pavement distresses, including cracking, faulting, spalling, and shoulder joint distress, using 3D laser pavement data. The validation results demonstrate the potential of using 3D laser data for automatically detecting distresses in concrete pavements. The test sites were selected on interstate highways I-16 and I-516.

  The validation of concrete pavement cracking detection using 3D laser data showed acceptable performance. The automatic crack detection results were compared to the manually digitized ground truth using a buffered Hausdorff scoring method that was developed in RS-GAMS Phase 1. The results showed that detection of cracks on I-516 (mainly longitudinal cracks) is accurate and robust; however, the detection of cracks on I-16 (mainly transverse cracks) is not as good as on I-516. The larger data acquisition interval along the driving direction, which was about 5 mm, might be the reason that some transverse cracks cannot be captured by 3D laser data. In comparison, the transverse resolution is about 1 mm, which can better capture the longitudinal cracks. Limited to the laser data resolution, hairline cracks (thinner than 2 mm) were still challenging for automatic detection.

  The validation of concrete joint faulting measurement showed that it is very feasible to use 3D laser data for collecting faulting data at highway speed. Using the regression-based method, the automatic faulting measurements were consistent with manually measured ground truth using the Georgia Faultmeter in both a well-controlled lab test and a field test.

The accuracy of automatic spalling detection varied for different sizes of spalling. Spalling with widths greater than 90 mm can be successfully detected; the detection accuracy was reduced, but still acceptable for widths between 50 and 90 mm wide, while it was hard to detect when widths are less than 50 mm. Though some small spallings were not successfully detected, they can be clearly observed on the laser data. Thus, the automatic detection algorithm could be further improved to handle such cases.

Since there is no dedicated application that is commercially available for shoulder joint distress detection, we explored the feasibility of using an automatic spalling detection algorithm to detect shoulder joint distress. The larger extent and depth of shoulder joint distress make them distinctive on laser range data and easier to detect. On the selected representative cases, the automatic detection results were visually consistent with field observations. However, it should be noted that due to the transverse coverage of the current pavement surface (about 4 meters), the shoulder area might be missed when the vehicle wanders. In addition, a specific shoulder joint distress detection algorithm is needed to further ensure an accurate and robust detection.

- **Pavement marking retroreflectivity condition assessment**

  This research was to establish the correlation between the retroreflectivity measured by handheld retroreflectometers and the retro-intensity acquired from a LiDAR point cloud. Establishing a reliable correlation is the key step for assessing pavement marking retroreflectivity conditions using a mobile LiDAR. In this preliminary

study, thermoplastic and waterborne paint, which are the most commonly used pavement marking materials, were tested. Test sites were selected on Ferst Drive, Hemphill Avenue, and 17th Street on/near the Georgia Tech campus.

It was discovered that the retro-intensity values acquired from mobile LiDAR are not sensitive to ambient temperatures, with an average standard deviation less than 0.0041. The retro-intensity acquired from mobile LiDAR held good repeatability on the tested thermoplastic and waterborne materials with an average standard deviation of 0.0044.

There was an exponential correlation between retroreflectivity and retro-intensity with an R-square of 0.9525 for thermoplastic and 0.9267 for waterborne paint. The correlation between retroreflectivity and retro-intensity might be sensitive to different bead formulas of the pavement marking material. Separate correlation curves might be needed not only for different pavement marking material category, e.g. thermoplastic, waterborne, etc., but also for different bead formulas in the same material category. Based on the correlation results, a preliminary retro-intensity threshold corresponding to the minimum retroreflectivity (100 mcd/m^2/lux) defined in the MUTCD could be defined as 0.4263, with a 95% confidence interval ranging from 0.4035 to 0.4505 for thermoplastic and 0.3521, with a 95% confidence interval ranging from 0.2973 to 0.4264. Using the established correlations, a mobile LiDAR-based pavement marking retroreflectivity condition assessment method can be further developed.

- **Long-term monitoring of crack deterioration**

  This research studied pavement deterioration behavior over time using long-term monitoring of 3D laser data. A detailed-level representation of crack information, including length, width, orientation/direction, position/location, intensity, pattern, etc. that are derived using 3D laser data, crack detection algorithms, and the multi-scale CFE model, was applied to study the temporal crack deterioration behavior. The results showed that it could potentially be used to support the following applications: 1) the fundamental study of pavement mechanistic and pavement deterioration behavior, 2) validation of current pavement design methods and development of new design concepts and methods, 3) determination of adequate treatment methods and strategies based on pavement distress characteristics and their deterioration behaviors, 4) development of accurate and reliable forecasting models, and 5) development of cost-effective pavement management operations/practices, like intelligent crack sealing planning. Compared to the frequently used lab simulation data or well-controlled experimental data (e.g. AASHO road test), the in-service pavement data used in this study represent true crack deterioration mechanisms and enables the ability to explicitly connect pavement condition deterioration with the real-world factors that cause it. While it is not feasible to comprehensively cover the entire field of research, this study is a concrete first step and is believed to be transformative in changing the way researchers have approached sensing data based infrastructure condition monitoring and risk assessment.

- **Long-term monitoring of rutting deterioration**

  This research characterized the 3D rut shape and analyzed its deterioration behaviors at multiple scales using long-term pavement 3D laser data. Rut parameters, including transverse profile-based parameters, longitudinal parameters, and temporal parameters were defined and proposed. Descriptive statistics and 2D and 3D visualizations were used to analyze the deterioration behaviors of rutting.

  Transverse profile-based rut parameters, including rut depth, rut cross-sectional area, percent deformation, and total absolute distortion, show good correlation among each other and provide consistent information about the conditions of ruts. Longitudinal parameters also showed similar trends as the aforementioned parameters, indicating that ruts not only grow in depth and area but also longitudinally in length and volume. Temporal parameters provide a direct means to quantify the deterioration of rutting. These parameters can be very useful, especially when the rate of deterioration is high. The multi-scale analysis results show that seasonal variations can have appreciable effect on the deterioration of ruts. Temporal parameters derived from a shorter period (e.g., a few months) can be affected by the seasonal variation. This finding suggests that, for different deterioration analysis applications, different analysis periods should be applied in order to obtain consistent results. The comparison among three different routes showed that traffic and roadway characteristics also play an important role in the deterioration of ruts. Other factors, such as the age and design of the pavement, can also contribute to the actual deterioration behavior of ruts. 2D and 3D visualization of ruts at the individual level shows the importance and benefits of having registered long-term 3D pavement data. Details of how ruts deteriorate, e.g.,

the development of dual-wheel shaped ruts and the growth in rut length and volume, were able to be identified in the selected sections. This information can further support the diagnosis of rutting and inform data-drive maintenance decisions.

2. Recommendations

The following suggest future research for implementation:

- With the promising results from the automatic asphalt pavement crack classification for load cracking and block cracking, it is recommended that automatic classification be extended to other types of distresses defined in GDOT's pavement distress manual. In addition, the algorithms can be easily extended to other crack survey protocols used by different state highway agencies because of the flexibility provided by the crack CFE model.
- The validation results for automatic concrete pavement faulting measurements showed very good consistency with manual measurements made with a Georgia Faultmeter. A large-scale pilot study with a state DOT, e.g. GDOT, to automate the network-level faulting measurements is suggested. This can significantly improve productivity, data accuracy, and data coverage.
- The concrete pavement crack detection shows promising results. However, it is difficult to detect hairline, transverse cracks due to the relatively coarser data resolution in the driving direction using the current 3D laser device. Thus, to capture hairline cracks, the data capture frequency and resolution of a 3D laser device needs to be further improved. In addition, automated crack evaluation for concrete

pavements, automatic crack classification algorithms need to be developed; it can be based on the work we have done for asphalt pavements.

- New algorithms need to be developed because the automatic detection for concrete spalling doesn't work well on spalling widths less than 50 mm. In addition, new algorithms are needed for automatic shoulder joint distress detection.
- The pavement marking validation results indicate that pavement marking retroreflectivity conditions can be measured and evaluated using mobile LiDAR that can be operated at highway speed. However, the testing samples in this research project are limited. It is suggested that large-scale testing on more marking materials be conducted before it is implemented.