



Pavement Evaluation 2019



September 17-20, 2019
Roanoke, Virginia

Development of a Machine Learning-Based Quality Control Approach for Automated Pavement Condition Data

Amir Arshadi, PhD, PE
Abbas Kachwalla, PE



September 19, 2019

Introduction

- *Extensive advancements in computer technology → Machine learning and artificial intelligence*
- *Using machine learning algorithms to identify road surface distresses from pavement downward images has been gaining momentum in the pavement management industry in recent years.*
- *There are numerous advantages in using automotive technologies to collect pavement condition data including safety, speed, and comprehensiveness*

Advantages of Automated Pavement Data Collection Technology

| | |
|--|--|
| Safety | Traffic control and lane closures are <u>not</u> required. |
| | Pavement inspectors are not exposed to traffic hazards. |
| | Pavement distress surveys are performed by inspectors in the office. |
| Speed | Highways and Roadways – Data collection rates of over 150 lane miles per day on highways and 30-40 miles per day on city streets. |
| | Highways and Roadways – Data collection is performed at posted traffic speeds, up to 60 mph. |
| | Airfields – 10,000 ft. x 150 ft. runway survey can be completed in approximately 5 hours. |
| Comprehensive, Geocentric Pavement Data | Airfield ... on airside operations. |
| | Sub-meter elementation for each pavement data |
| | Accurate |
| | High resolution provided forward-facing ROW images |
| | ASTM compliant measurements along with accurate, |
| | Pavement and longitudinal grade. |

**RELIABILITY
OF THE
DATA?!**

Challenges with Automated Pavement Condition Data

- Identification of *micro/macro texture-based distresses* such as weathering, raveling, and bleeding;
- Confusion with the existence of *stationary objects* on the pavement road and consequently in downward photos such as security cones, foliage and debris such as plastic bottles;
- Challenges with pavement condition rating of asphalt concrete roads with *concrete gutters*;
- Issues with rutting measurements when the data is collected on a road with a *steep slope, sharp turn, or sudden change* in direction or lane;
- Incorrectly including the road shoulder in rating of *narrow roads*;

Challenges with Automated Pavement Condition Data (continue)

- Identification of *paint cracking* as a pavement distress;
- Identification of *cobblestone cross-walks and speed bumps* as surface distresses;
- Incorrect detection of *manhole leads and catch basin covers* as pavement distresses;
- Difficulty in identifying *patches*;
- Difficulty in identifying and classifying *block cracking* separately from other cracking; and
- Challenges with machine learning algorithms to distinguish between *raveling and potholes*.

Problem Statement

- *There are still serious limitations with machine learning algorithms in identifying and classifying pavement distress data.*
- *Accuracy and consistency in machine learning algorithm outputs are not certain and should be constantly verified.*
- *Currently, the most common QC program for APCD includes random sample audits*

Goal

- *The main goal of this study is to develop a systematic approach for quality control of automated pavement condition data.*
- *This study includes a systematic approach to flag the pavement sections (or sample units) for further checks based on known challenges and deficiencies of APCD in addition to random quality checks.*

Data Patterns

- *It is unlikely that medium and high severity potholes occur in sections with no weathering distress;*
- *Usually potholes cannot be the only distress type present on the sample unit;*
- *Potholes are load related distress and generally exist with other load related distresses such as alligator cracking;*
- *In asphalt roads which are older than five years, weathering at either low, medium, or high severity is usually present;*

Data Patterns (continue)

- *Sections (or samples) exhibiting medium or high severity weathering generally have other distresses present;*
- *Alligator cracking evolves from longitudinal and transverse (L&T) cracking. Therefore, other quantities of L&T cracking will likely be present if alligator cracking is present;*
- *High severity distresses such as alligator cracks, potholes, and raveling are not generally observed on pavement sections with age less than 5 years;*

Data Patterns (continue)

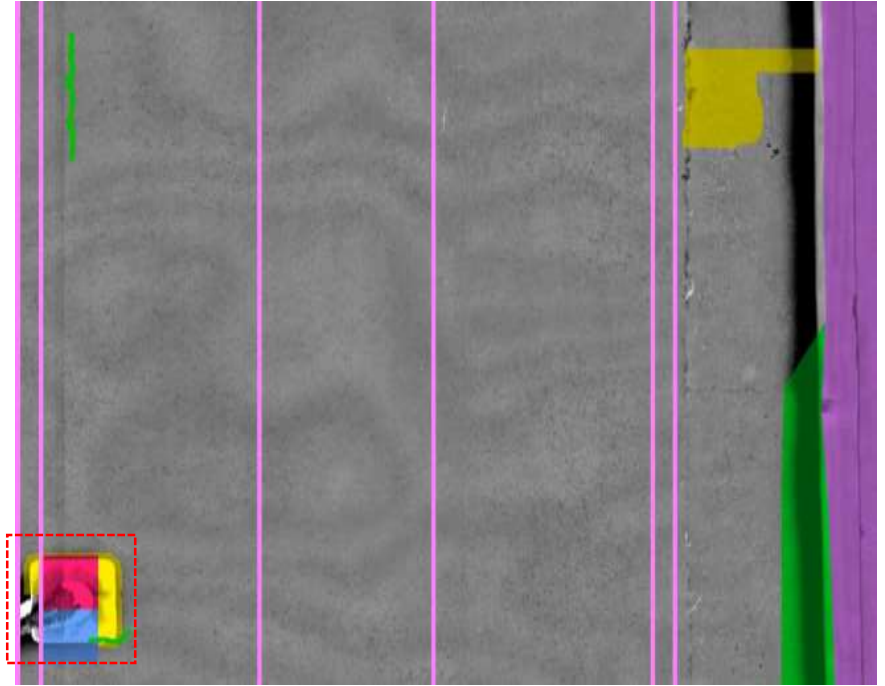
- *In a specific road section, it is uncommon to see significant instances of rutting distress in only one sample unit while not in other sample units (with the exception of intersections); and*
- *In a specific road section, it is uncommon to see significant quantities of alligator cracking distress in only one sample unit and not in other sample units.*

Pavement Data Quality Management (PDQM) Software

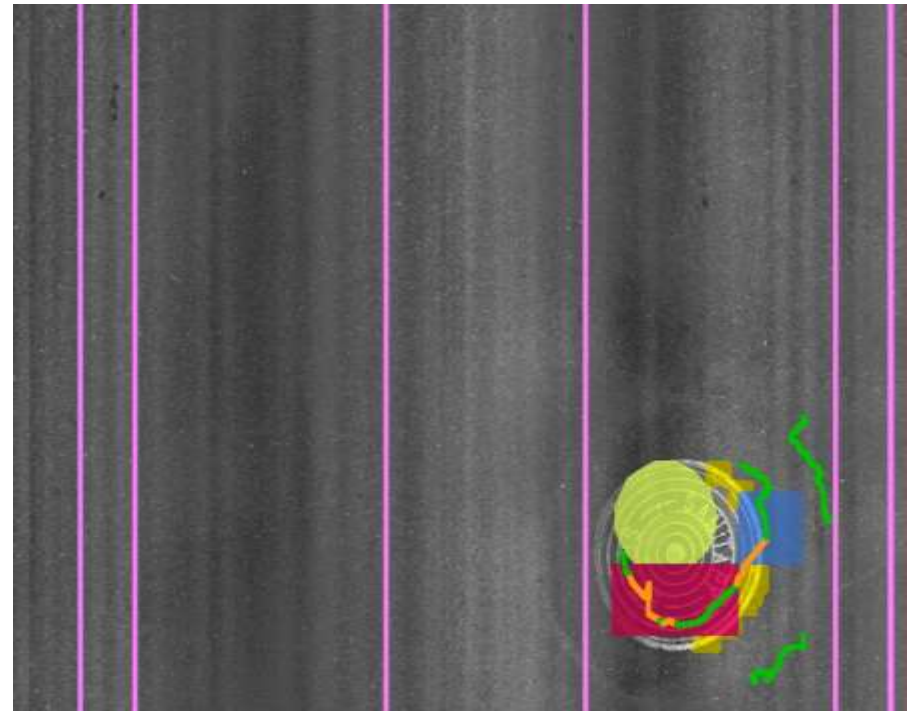
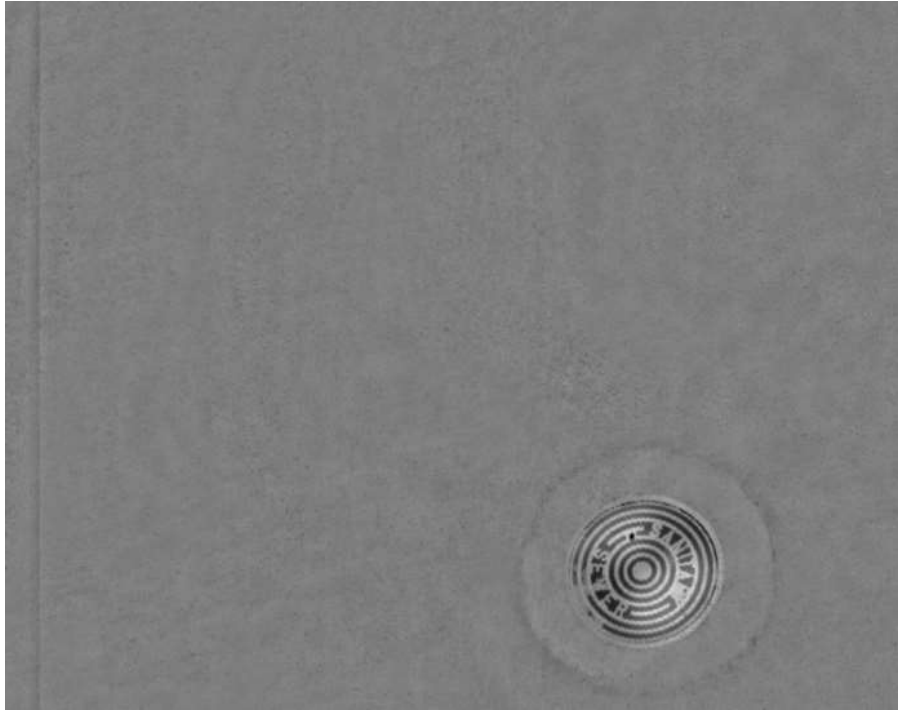
The screenshot displays the PDQM 1.3 software interface, which is used for managing pavement data quality. The interface is divided into several functional areas:

- Inventory:** Shows project details such as 'ASTM-Landside', 'ASTM-Airside', 'Pave-TSD', 'DeIDOT', 'HPMS', and 'PASER'. It includes 'Inventory Statistics' for 'Network 1 - Bmc 40 - Scn: 3858 - SU: 38590' with a total length of 239.61 miles. The X-axis lists 'Network', 'Branch', 'Section', 'SU', 'Surf Type', and 'Inspect Date'. The Y-axis lists 'Size', 'SU Size', 'PCI', 'Histogram', and 'IRI'. Buttons for 'plot' and 'Table' are available.
- Quality Control:** Features a 'QUALITY' logo and 'Database Related Checks' including 'Data File Structure', 'Data Completeness', 'Location Accuracy', 'Data Format', 'Duplicated Records', and 'Payment Type'. It also includes 'Statistical Analysis Checks' for 'Min and Max Tolerance Parameters', 'Reasonable Extent of Distress', and 'Extent of Area Based Distresses'.
- Input Files for PAVER 7.0.6:** Contains options to 'Generate Shape File', 'Save the Shapefile', and 'Generate the XML File'. It also has a 'Distress File' section with 'Generate the XML File' and 'Open the Distress File' buttons.
- Statistical Visualization:** Includes a 'Show the Road on Map' section with buttons for 'AC', 'PCC', 'AC-A', and 'PCC-A'. Below this, it shows the 'Network' (SRL - Lane 1), 'Branch' (001017010), 'Section' (1F19.13.36), and 'Sample Unit' (All). A 'Find the Road' button is also present.
- Maps:** Two maps are shown. The top map is a regional map with latitude (lat) from 29.2 to 29.6 and longitude (lon) from -76.0 to -75.2. The bottom map is a zoomed-in view of a road segment with latitude (lat) from 39.556 to 39.557 and longitude (lon) from -75.663 to -75.661, highlighting a specific road section in red.

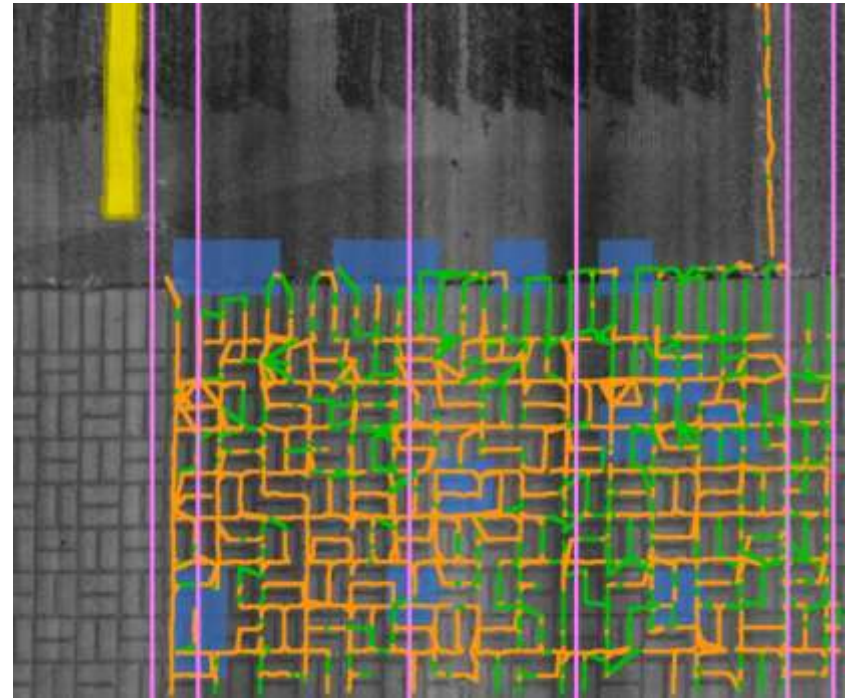
Incorrect Distress Identification



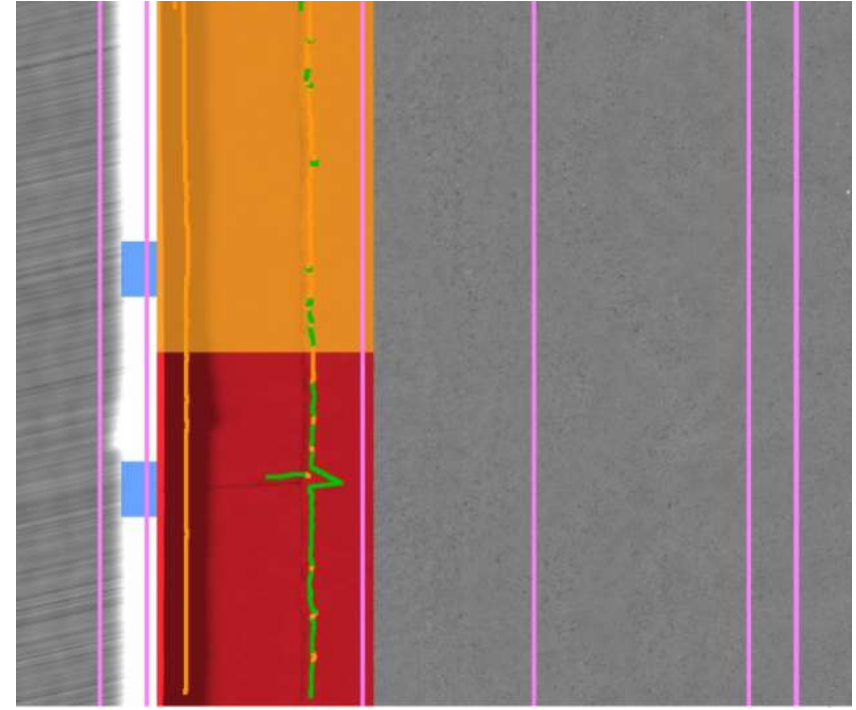
Incorrect Identification of Raveling within Manhole Cover



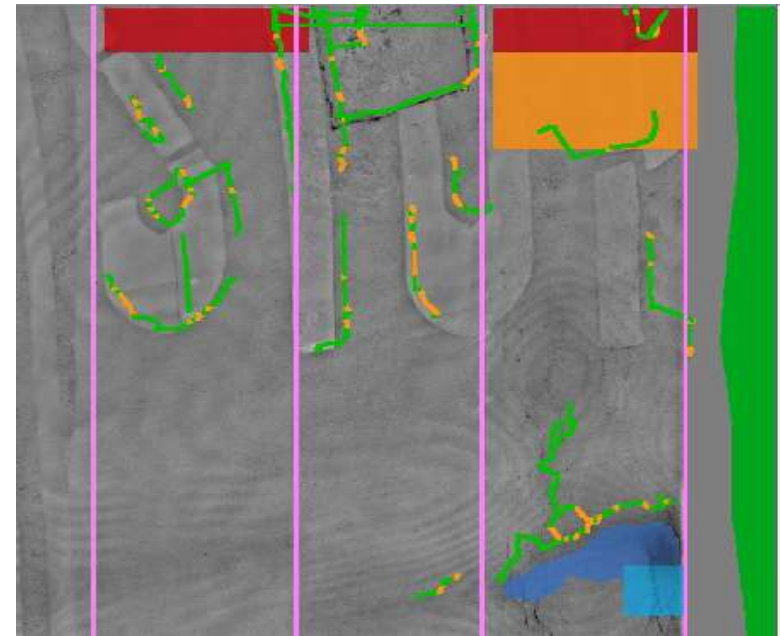
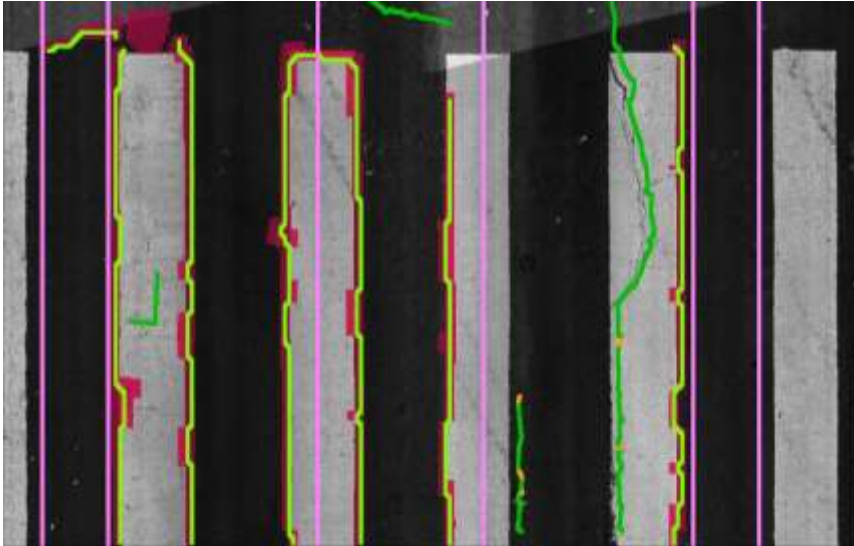
Identification of Cobblestone Crosswalk as Surface Distress



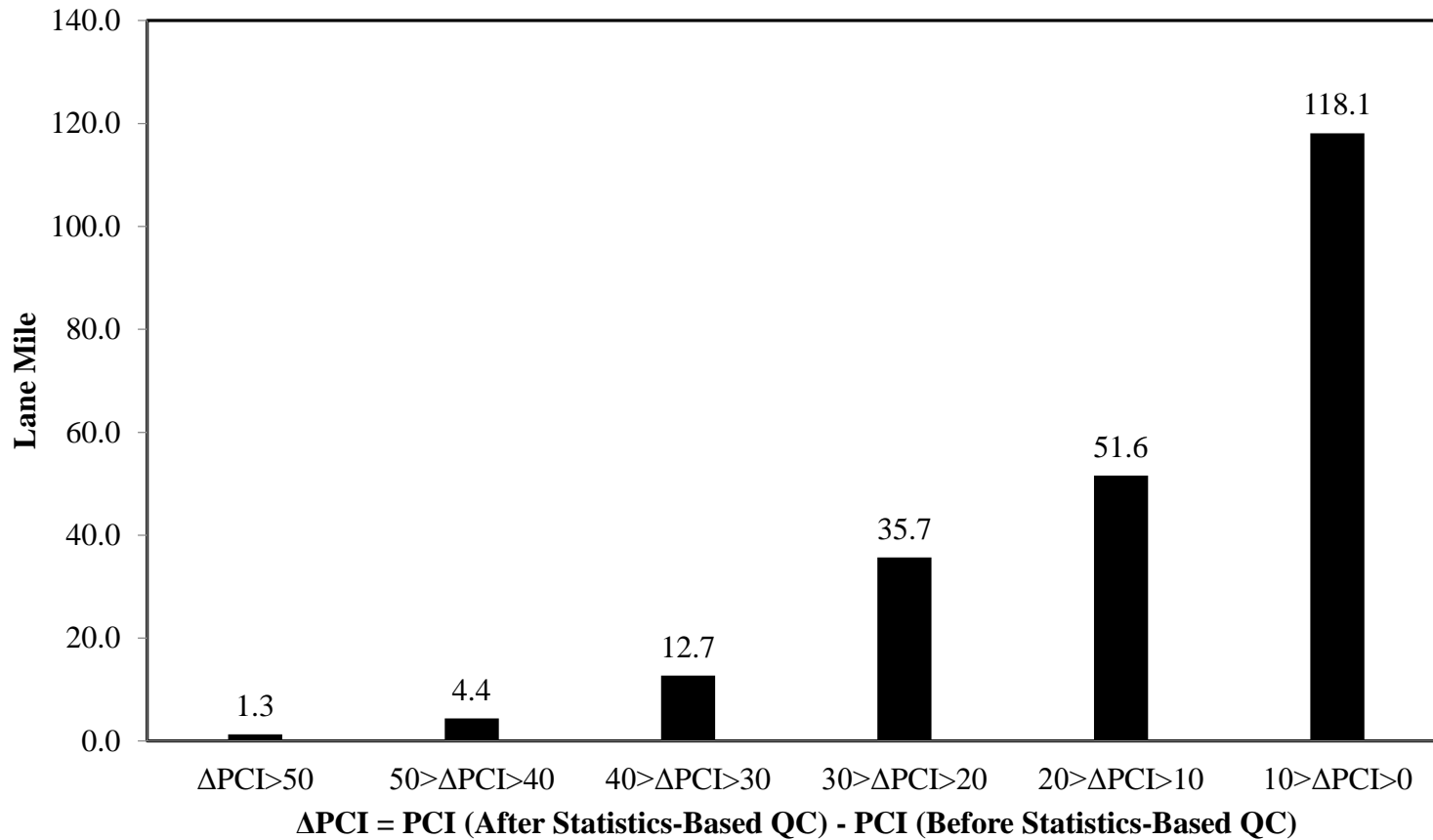
Incorrect Pavement Condition Rating in Narrow Streets



Identification of Road Markings as Pavement Distresses



Effect of Statistics-Based QC on PCI Values



Conclusions

- A statistics-based quality control approach was presented as an effective method to improve the automated pavement condition data.
- This approach was successfully tested on one thousand miles of automated pavement condition data collected on urban and suburban asphalt roads.
- The results indicate that typical quality control approaches that are based on random spot checks of the data are not sufficient due to the nature of the automated pavement condition data processing.

Thank You

Amir.Arshadi@aecom.com

September 19, 2019

AECOM