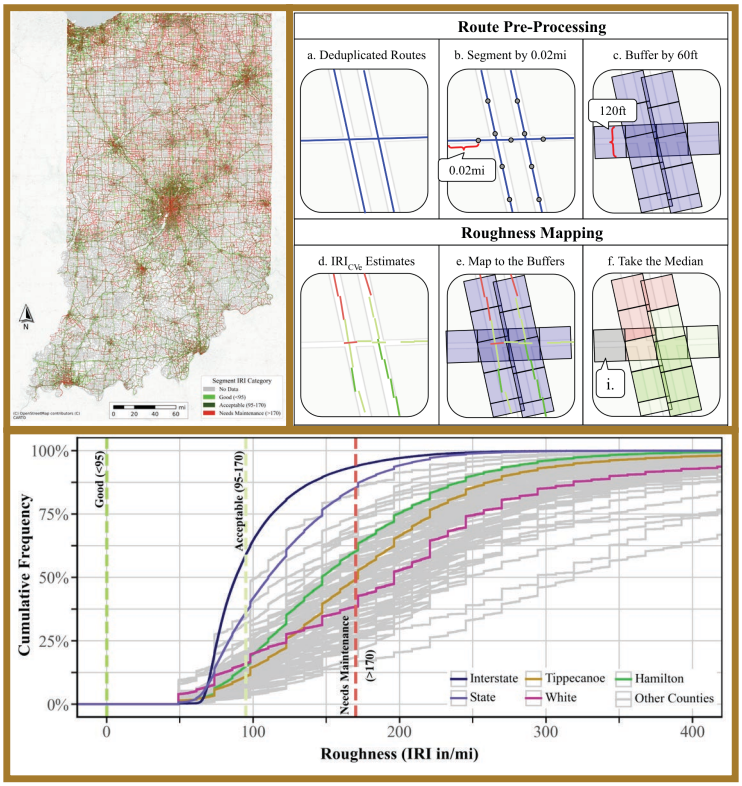
JOINT TRANSPORTATION
RESEARCH PROGRAM
INDIANA DEPARTMENT OF TRANSPORTATION
AND PURDUE UNIVERSITY

Research Support to INDOT on Systemwide Asset Condition Assessment Using Connected Vehicle Data



Andrew Thompson, Jairaj C. Desai, Rahul Suryakant Sakhare, and Darcy M. Bullock

SPR-4907 • Report Number: FHWA/IN/JTRP-2026/04 • DOI: 10.5703/1288284318612

RECOMMENDED CITATION

Thompson, A., Desai, J. C., Sakhare, R. S., & Bullock, D. M. (2026). *Research support to INDOT on systemwide asset condition assessment using connected vehicle data* (Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2026/04). West Lafayette, IN: Purdue University. https://doi.org/10.5703/1288284318612

AUTHORS

Andrew Thompson
Purdue University
(630) 219-9418
thomp907@purdue.edu
Corresponding Author

Jairaj C. Desai, PhD
JTRP Transportation Research Engineer
Lyles School of Civil and Construction Engineering
Purdue University

Rahul Suryakant Sakhare, PhD
JTRP Transportation Research Engineer
Lyles School of Civil and Construction Engineering
Purdue University

Darcy M. Bullock, PhD, PE
Lyles Family Professor of Civil Engineering
JTRP Director
Lyles School of Civil and Construction Engineering
Purdue University

ACKNOWLEDGMENTS

The connected vehicle roughness data used in this study was provided by NIRA Dynamics AB. Google Cloud Platform (GCP) and Google BigQuery were used in this study for cloud computing and storage. The Indianapolis PCI data was provided by Crossroad Engineers, and the Noblesville PCI10 and PASER datasets, including the PASER rating imagery, were provided by the City of Noblesville. The PASERCVi rating data was collected by Cyvl inc.

This report was adapted from the following papers:

- Mathew, J. K., Desai, J., Sakhare, R. S., Hunter, J., & Bullock, D. M. (2024). Spatiotemporal analysis of pavement roughness using connected vehicle data for asset management. *Journal of Transportation Technologies*, *15*(1), 1–16. https://doi.org/10.4236/jtts.2025.151001
- Thompson, A., Desai, J., & Bullock, D. M. (2025). Evaluation of Connected Vehicle Pavement Roughness Data for Statewide Needs Assessment. *Infrastructures*, *10*(9), 248. https://doi.org/10.3390/infrastructures10090248
- Thompson, A., Desai, J., & Bullock, D. M. (2026). A comparison of connected-vehicle roughness and traditional pavement condition index. *Future Transportation*, *6*(1). https://doi.org/10.3390/futuretransp6010047
- Thompson, A., Desai, J., & Bullock, D. M. (2026, June 22–25). *A comparison of connected vehicle roughness data with computer vision pavement condition ratings* [Manuscript submitted for publication and presentation]. IEEE Intelligent Vehicles Symposium (IV), Detroit, MI, United States.

JOINT TRANSPORTATION RESEARCH PROGRAM

The Joint Transportation Research Program serves as a vehicle for INDOT collaboration with higher education institutions and industry in Indiana to facilitate innovation that results in continuous improvement in the planning, design, construction, operation, management and economic efficiency of the Indiana transportation infrastructure. Learn more at engineering.purdue.edu/JTRP.

Published reports of the Joint Transportation Research Program are available at docs.lib.purdue.edu/jtrp/.

NOTICE

COVER IMAGE

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. FHWA/IN/JTRP-2026/04	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle Research support to INDOT on systemwide asset condition assessment using connected vehicle data		**5. Report Date** February 4, 2026 **6. Performing Organization Code**
7. Author(s) Andrew Thompson (https://orcid.org/0009-0004-7255-8289) Jairaj C. Desai, PhD (https://orcid.org/0000-0003-2885-203X) Rahul Suryakant Sakhare, PhD (https://orcid.org/0000-0001-7843-5707) Darcy M. Bullock, PhD, PE (https://orcid.org/0000-0002-7365-1918)		**8. Performing Organization Report No.** FHWA/IN/JTRP-2026/04
9. Performing Organization Name and Address Joint Transportation Research Program Hall for Discovery and Learning Research (DLR), Suite 204 207 S. Martin Jischke Drive West Lafayette, IN 47907		**10. Work Unit No.** **11. Contract or Grant No.** SPR-4907
12. Sponsoring Agency Name and Address Indiana Department of Transportation (SPR) State Office Building 100 North Senate Avenue Indianapolis, IN 46204		**13. Type of Report and Period Covered** Final Report **14. Sponsoring Agency Code**

15. Supplementary Notes
Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.

16. Abstract
Crowdsourced pavement roughness data using connected vehicles now provides cost-effective, high-resolution condition monitoring in near real-time. This study evaluates how connected vehicle data can support effective investment prioritization across Indiana's road network and improve upon traditional data sources using surveys that are often done annually or biennially. A key challenge is handling terabyte-scale geospatial roughness datasets and integrating them meaningfully into asset management workflows. A scalable data processing methodology is outlined to address this issue and is used for the subsequent analysis.

This report's analysis uses more than 3 billion daily CV-estimated roughness measurements (IRI_{CVe}) from 2022–2025, in addition to multiple visual condition sources: 2024 PCI data from Indianapolis and Noblesville and 2025 computer-vision PASER data from Noblesville. Statewide local paved roads show a spatial data coverage increase from 46.5% to 53.2% between 2023 and 2024. Case studies from I-65, Noblesville, and Indianapolis demonstrate the potential of these developing data sources even at the segment and route levels. Additionally, multiple metric comparisons confirm a weak roughness-surface condition correlation (with R^2 of 0.15 to 0.34) and were used for network-wide screening for outlier detection and quality control. These results support implementation recommendations for using IRI and condition data together for a more comprehensive estimation of pavement quality. Although the evaluated data sources show substantial potential, they are not yet ready to replace traditional data for all use cases.

17. Key Words asset management, IRI, connected vehicles, pavement condition monitoring	**18. Distribution Statement** No restrictions. This document is available through the National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 28, including appendices	**22. Price**

EXECUTIVE SUMMARY

Motivation

The recent widespread availability of crowdsourced pavement roughness measurements using connected vehicles (CV) has enabled cost-effective and high-resolution condition monitoring in near real-time. At the same time, computer vision–based data collection is being deployed to mitigate the inherent subjectivity of human surface condition ratings. The purpose of this study is to assess the feasibility of utilizing CV data sources for agile and effective prioritization of maintenance and investment projects across Indiana's road network.

Study

Traditional methods of monitoring pavement condition have involved manual human surveys often done annually or biennially. CV data can be used to obtain much higher resolution view of a network's condition, make consistent comparisons across local agency jurisdictions, and reduce network condition monitoring costs, among other benefits. Case studies from Interstate 65 (I-65), Noblesville, and Indianapolis are included in the report. Two challenges that must be addressed before effective agency integration is feasible are the efficient handling of terabyte-scale roughness data and the careful interpretation and integration of the data for use in asset management strategies. This study uses more than 3 billion daily CV-estimated International Roughness Index (IRI_{CVe}) measurements from across Indiana in the years 2022–2025, in addition to multiple visual surface condition data sources used for comparison. The surface condition data includes Pavement Condition Index (PCI) data from Indianapolis and Noblesville in 2024 and computer vision-derived Pavement Surface Evaluation and Rating ($PASER_{CVi}$) data from Noblesville in 2025. The data source comparisons reveal the similarities and differences between the automated and manual data sources and provide insight into where the automated data can be used most effectively and immediately by agencies.

Results

The following results were identified after the evaluation and comparison of the various pavement condition data sources:

- Cost-effective condition monitoring at scale is feasible using IRI_{CVe} data.
  - Statewide IRI_{CVe} coverage of paved local roads increased from 46.5% to 53.2% between 2023 and 2024.
  - County-level network condition and IRI_{CVe} coverage analysis revealed additional spatiotemporal trends and enabled large-scale comparisons.
  - A localized case study shows how IRI_{CVe} can be used to identify segment-level condition changes at network scale.
- The automated and manual data sources are not interchangeable, yet still useful.
  - IRI_{CVe} and $PASER_{CVi}$ data sources weakly correlated to the manual PCI datasets with R^2 of 0.15 to 0.34 when stratified by roadway class and surface type.
  - Metric agreement varied by roadway class and prominent distress types, but with trends that were supported by prior studies done with manual data.

Recommendations

Based on the results outlined above, the relevant recommendations are as follows: Roughness (IRI) and surface condition (PCI or PASER) should still be used together to obtain a clear picture of the roadway condition. IRI_{CVe} is best used as a network-wide screening tool for near-term use by agencies. This use case alone can allow agencies to quickly focus their efforts on portions of their network that are emerging with near-term maintenance needs. As the CV coverage and quality continues to grow and IRI estimation improves, it is probable that IRI_{CVe} will be able to replace profiler-based IRI in the long-term, after further validation studies. $PASER_{CVi}$ cannot be used interchangeably with manual PCI surface condition data; however, the correlation trends and agreement analysis showed that it has potential as an unbiased alternative to human surface condition rating. A comparison of automated and manual PASER would be necessary for more specific implementation recommendations regarding this data source.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. PROJECT OVERVIEW

The recent widespread availability of crowdsourced pavement roughness measurements using connected vehicles (CV) has enabled cost-effective and high-resolution condition monitoring in near real-time. At the same time, computer vision–based data collection is being deployed to mitigate the inherent subjectivity of human surface condition ratings. The purpose of this study is to assess the feasibility of utilizing these automated data sources for agile and effective prioritization of maintenance and investment projects across Indiana's road network. The study is structured as follows:

- Section 1. Project Overview: An overview of the project's scope and objectives.
- Section 2. Data Description: An explanation of the various data sources and locations used for the project.
- Section 3. Methodology: Discusses the foundational data processing methodology.
- Section 4. Results: Summary of the case-specific methods and analysis findings.
- Section 5. Conclusion: A review of the key insights and the implications of the study.
- Section 6. Future Research: Outlines promising future directions for this research.

1.1 Introduction

CV data have been explored in depth for use in traffic condition monitoring and intersection performance, however, a relatively new application of this data is for large-scale pavement condition monitoring (Thompson et al., 2025). Condition data are fundamental to effective roadway maintenance and investment prioritization and have traditionally been collected by manual surveys. These are either done visually or with a dedicated sensor such as an inertial profiler (Federal Highway Administration [FHWA], 2024), or more recently, Light Detection and Ranging (LiDAR) and computer vision (Abohamer et al., 2021; De Blasiis, et al., 2021).

Three widely used pavement condition metrics are Pavement Condition Index (PCI), Pavement Surface Evaluation and Rating (PASER), and the International Roughness Index (IRI). PCI and PASER are visual surface condition ratings, while IRI is a ratio of vertical displacement to longitudinal movement which is used to represent the roughness of a road segment. PCI ranges from 0–100 and PASER ranges on a scale from 1–10, and in both metrics a segment with a lower rating is in worse condition than a segment with a higher rating. Although PCI and PASER have been used for decades, their manual collection induces a tradeoff between the speed of collection and the subjectiveness of the rating data. For example, the ASTM PCI D66433 document states that inspection should be performed by manually walking the length of the road segment to take measurements (ASTM International, 2020), which, of course, takes a substantial amount of time and funding for any sizeable road network. This cost limits agencies to usually only survey their network annually or biennially and puts surveyors in dangerous conditions during traffic. Alternatively, the Michigan PASER manual recommends data collection to be performed by using a rating criteria cheat sheet while driving over the segment (Michigan Infrastructure Council, n.d.). This makes collection faster than PCI; however, the cheat sheet contains vague language such as "slight," "moderate," or "severe" distress severity, which introduces substantial inter-rater variability, despite efforts to consistently train surveyors through programs like the Local Technical Assistance Program (LTAP) PASER certification (Purdue University, n.d.). PCI also has this problem, as only a sample of segments are rated, and this sampling methodology may vary by agency, breaking down any comparisons made at the network level. IRI can be a more objective condition metric due to its quantitative definition, although it is still subject to many of the same inefficiency limitations when collected by traditional survey methods. Previous metric comparison studies have established that the IRI and surface condition metrics like PCI and PASER have weak but still significant correlations (Fakhri & Shahni Dezfoulian, 2019; Gerst, 2009; Hannedeh et al., 2022). This can be intuitively explained by the concept that they capture different aspects of roadway degradation. The IRI, being a measure of roughness, may not capture all types of visible distresses, and PCI and PASER, being solely based on visible surface condition, may not account for certain distresses or anomalies that impact ride quality without indicating poor condition, such as manholes or rumble strips. Therefore, it is evident that the IRI is most useful in combination with other condition metrics to obtain a more complete view of the roadway condition (Kırbaş, 2018).

In recent years, advances in sensing and computing technology have resulted in more advanced methods of pavement condition data collection that attempt to mitigate many of the limitations that have been mentioned. Two such automated methods that will be discussed in this report are CV-estimated IRI (IRI_{CVe}) and Computer Vision–derived PASER ($PASER_{CVi}$). IRI_{CVe} is collected by the Original Equipment Manufacturer (OEM) sensors from a fleet of connected vehicles that report roughness in near real-time. A sensor-fusion model was developed using a Kalman filter that combines sensor output into an estimated longitudinal profile of the roadway, which can then be used for IRI calculations (Agebjär et al., 2025). Multiple validation studies have found R^2 values ranging from 0.7–0.79 between IRI and IRI_{CVe}, indicating moderate to strong correlation (Llopis-Castelló et al., 2024; Mahlberg et al., 2022, 2024). The benefits of this IRI collection approach include its cost-effectiveness from crowdsourcing, robustness due to measurement aggregation, and higher spatial and temporal resolution. Another advanced method of interest is $PASER_{CVi}$. This approach uses dashcam imagery collected during a survey to extract PASER ratings from visible distresses at a high resolution. Although it still often uses an intentional survey to collect the imagery data, there have been studies exploring the use of crowdsourced dashcam imagery to extract these ratings at a large scale (Dadashova et al., 2021; Randhawa et al., 2024). Regardless, this approach addresses many of the objectivity limitations of traditional manual methods through the use of a consistent computer vision model. Therefore, the collection speed and

subjectivity tradeoff can be eliminated: this approach potentially enables more objective data than manual PCI and faster collection than traditional PASER. Even with an imperfect model, the resulting condition ratings are at minimum derived using a known consistent methodology across all networks for accurate comparisons, as opposed to using many different human raters under varying local agency standards.

1.2 Scope and Objectives

The objective of this study is three-fold: to propose a scalable methodology for the processing of billions of statewide IRI_{CVe} observations, to introduce the reader to what insights this data source can provide, and to compare the automatically collected IRI and PASER data to manually collected data to determine how it can and should be used in existing agency asset management practices. The findings and discussion in this study will include recommendations for both the implementation and interpretation of the automated pavement condition data sources.

1.3 Dissemination of Research Results

The following research studies were prepared in part during this project to facilitate an agile dissemination of results for public and private sector stakeholders:

- Mathew, J. K., Desai, J., Sakhare, R. S., Hunter, J., & Bullock, D. M. (2024). Spatiotemporal analysis of pavement roughness using connected vehicle data for asset management. *Journal of Transportation Technologies*, *15*(1), 1–16. https://doi.org/10.4236/jtts.2025.151001
- Thompson, A., Desai, J., & Bullock, D. M. (2025). Evaluation of Connected Vehicle Pavement Roughness Data for Statewide Needs Assessment. *Infrastructures*, *10*(9), 248. https://doi.org/10.3390/infrastructures10090248
- Thompson, A., Desai, J., & Bullock, D. M. (2026). A comparison of connected-vehicle roughness and traditional pavement condition index. *Future Transportation*, *6*(1). https://doi.org/10.3390/futuretransp6010047
- Thompson, A., Desai, J., & Bullock, D. M. (2026, June 22–25). *A comparison of connected vehicle roughness data with computer vision pavement condition ratings* [Manuscript submitted for publication and presentation]. IEEE Intelligent Vehicles Symposium (IV), Detroit, MI, United States.

The findings from this study were presented to stakeholders over the course of this project at various SAC meetings. The following sections of this report summarize the methodology, results, and implications of the three above-mentioned studies.

2. DATA DESCRIPTION

2.1 US Roads Dataset

The US Roads Dataset was used as an independent route source, which is available publicly free of charge on the BigQuery Marketplace (Google Cloud Console, n.d.). The data are derived from the US Census Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER) database and contain comprehensive route lines for all US states and territories. Notably, the data include functional classification classes that were used for determining local and state-managed routes.

2.2 Statewide Connected Vehicle Roughness Data

This project utilizes IRI_{CVe} data collected by onboard sensors from a fleet of connected vehicles, with anonymized data obtained through a commercial third-party vendor. A Kalman filter is used to fuse data from multiple sensors, such as the Global Positioning System (GPS) and Inertial Measurement Unit (IMU), which are used to measure vehicle speed and chassis vibration. This multisensor fusion approach estimates the road's longitudinal profile, which is then used to calculate its IRI value (Agebjär et al., 2025). The data are provided as short road segments ranging from 50 to 85 ft in length, updated on a daily basis. Each segment's IRI_{CVe} estimate represents a 60-day moving average over the CV measurements, and each measurement is converted from experienced roughness to the standard IRI scale with units of inches per mile (in./mi), representing the vertical displacement in inches over the longitudinal movement in miles along the roadway. To classify roughness, FHWA-defined (FHWA, 2014) thresholds were used: segments were labeled as:

- Good (< 95 in./mi),
- Acceptable (95–170 in./mi), and
- Needs Maintenance (> 170 in./mi).

Approximately 3.1 billion records were analyzed over 78,000 mi of roadway across Indiana. Figure 2.1 is included to convey the scale of the data by visualizing the roughness category of the Indiana road network in 2024. Subsets of this data were used in comparisons with other metrics where data was available.

2.3 Noblesville Condition Data

Multiple pavement condition datasets were provided by the City of Noblesville, Indiana, comprising of manual PCI surveys from July 2023 (with maintenance updates in November 2024) and computer vision–derived $PASER_{CVi}$ surveys from May 2025. There were 127 mi of roadway used in comparisons where IRI and both condition datasets were available. For comparison consistency, PCI scores were divided by 10 and rounded (producing PCI_{10}), while $PASER_{CVi}$ ratings, which are ~60 ft segments with annotated distress imagery, were classified as Good (8–10), Fair (5–7), or Poor (1–4) following TAMC guidelines (Michigan Transportation Asset Management Council, 2026), with half-point values assigned to the lower classification. Each condition dataset was spatially aligned with median IRI_{CVe} roughness data from April 2024 and April 2025, respectively, yielding two matched datasets: (1) 2024 IRI_{CVe} with PCI_{10} and (2) 2025 IRI_{CVe} with $PASER_{CVi}$. These datasets were used to separately model the condition-roughness relationship and to facilitate comparisons between the manual and automated data. Figure 2.2 shows the set of road segments, along with their 2025 $PASER_{CVi}$ categories, that were used for modelling and comparisons.

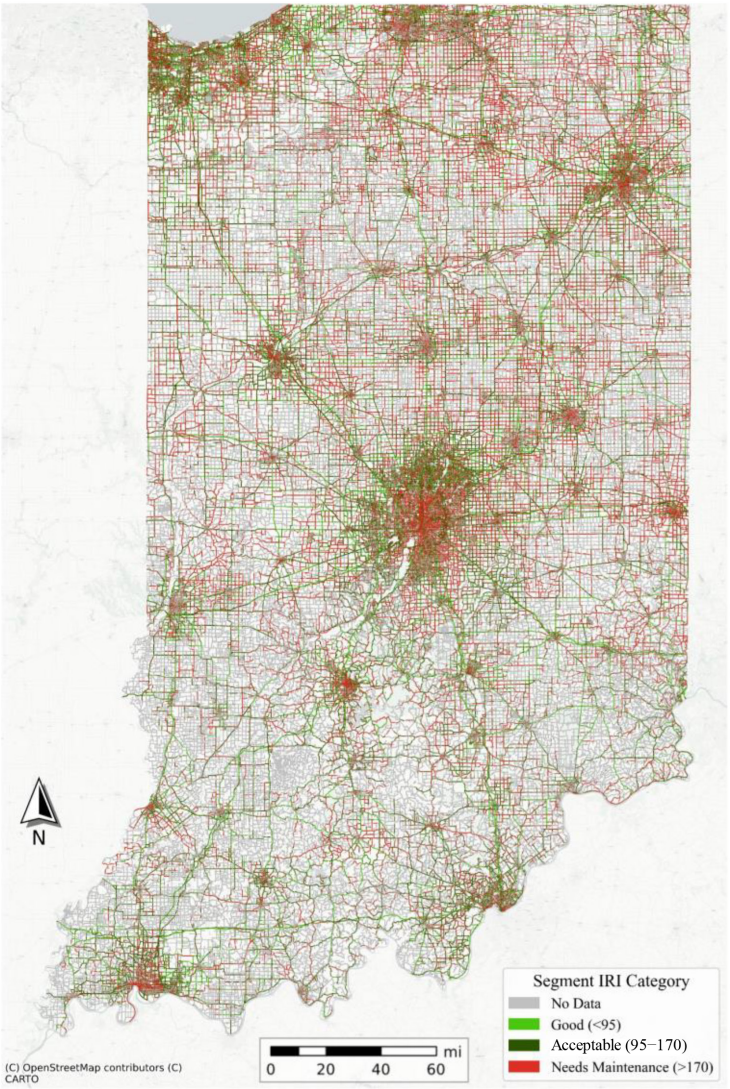


Figure 2.1 Indiana Road Network IRI Category Map (2024).

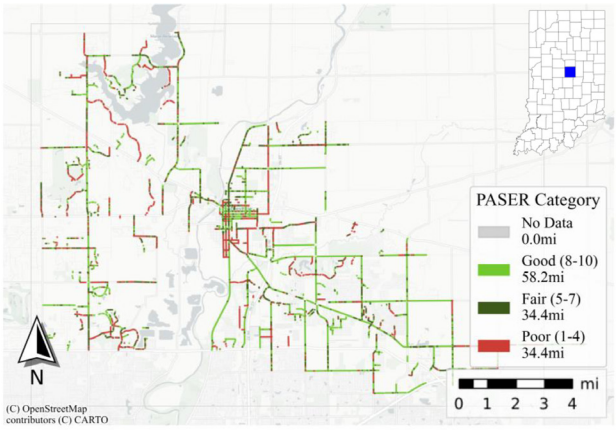


Figure 2.2 Noblesville PASER$_{CVi}$ Category Map May 2025 with Hamilton County Inset Map.

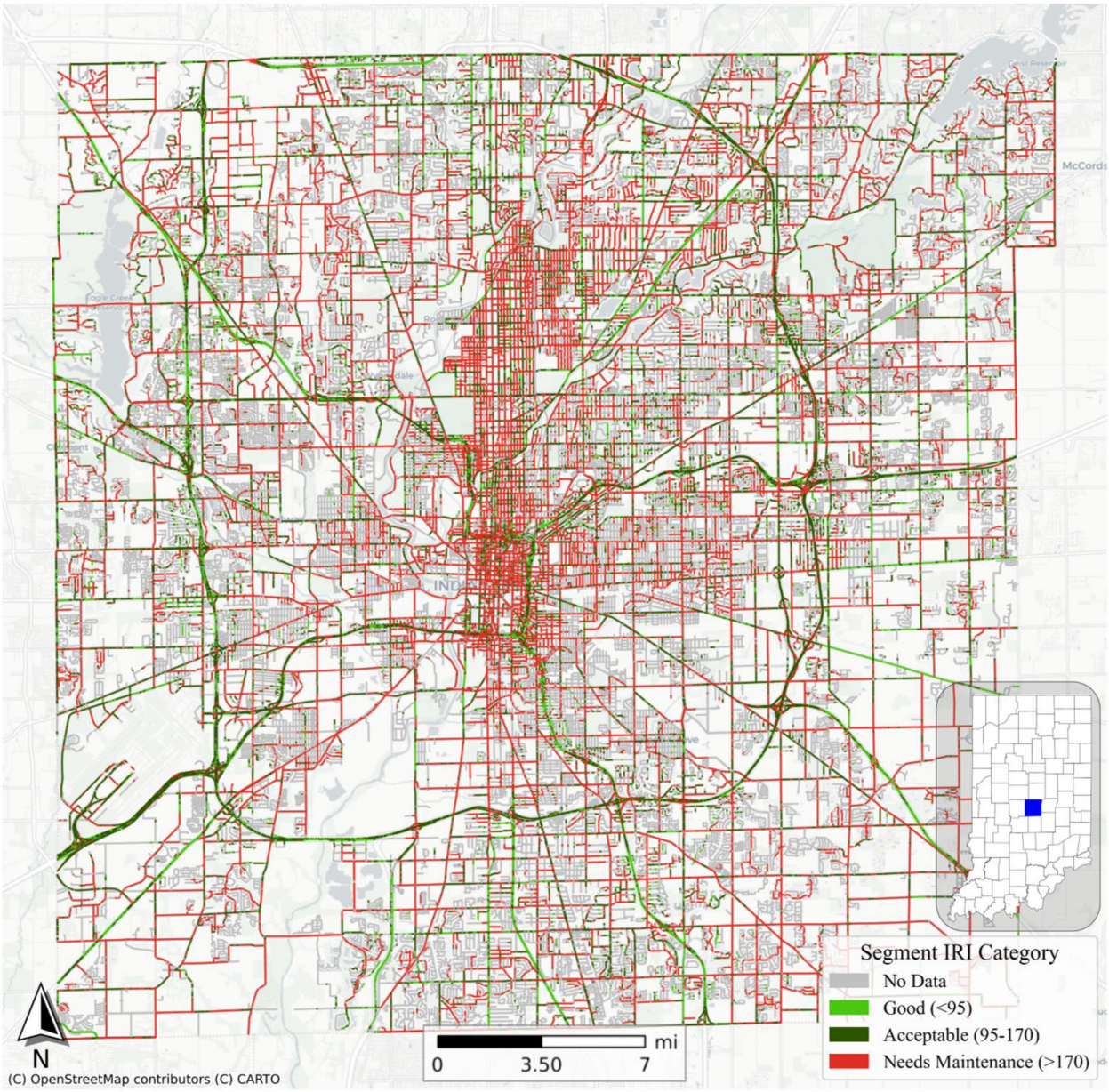


Figure 2.3 Marion County IRI Roughness Category Map Nov. 2024 With Inset Map.

2.4 Indianapolis Condition Data

Indianapolis also provided manual PCI data for metric comparisons. There were 1,963 mi (98,150 segments) used in comparisons where IRI data was also available. Approximately one-third of the PCI data from this source were manually collected each year between 2022 and 2024 and integrated into a common dataset for the Indianapolis network as of November 2024. The dataset represents the most accurate estimation of the Indianapolis road network condition in 2024 as reported by the City of Indianapolis, although not all measurements were collected in 2024. The routes were segmented by intersection, meaning every section of roadway between intersections was evaluated independently. No specific distress data used in the PCI calculation were recorded in the dataset, and only the PCI ratings were available. Each segment was assigned a PCI rating from 0–100, and these ratings were later categorized as Good (86–100), Satisfactory (71–85), Fair (56–70), Poor (41–55), Very Poor (26–40), Serious (11–25), and Failed (0–10), as per the ASTM standard outlined in their D6433 document (ASTM International, 2020). Additional attributes such as roadway class and surface type were made available by this same data source. The roadway classifications, shown in Figure 2c, correspond to the FHWA (2023) National Functional Classification system. Figure 2.3 shows the route network for this county and corresponding IRI category data in November 2024.

3. METHODOLOGY

3.1 Route Network Preparation

The US Roads dataset contains approximately 30% overlapping geometry, often from historical name or route changes. A simple deduplication algorithm was developed to address the issue as the first step in the data processing pipeline. The algorithm works by iteratively merging overlapping route geometries (combining metadata such as the route ID) until none remain. Additional metadata, such as county and district, were sourced from IndianaMap and spatially joined with the route segments by maximum overlap. As illustrated in Figure 3.1, Steps b and c, these routes were then segmented by 0.02 mi and buffered by 60 ft into road segments. A 0.02-mi segment length was chosen to capture the ~75 ft (0.014 mi) resolution of the IRI_{CVe} data. As the segments were created, the undirected heading was calculated by using the azimuth bearing between each segment's start and end points, converting it to degrees, and projecting it to a 0–180° range using modulus division. This transformation accounts for inconsistent line directionality in the base network (i.e., some segments may face opposite directions along the same corridor), while still allowing for directional filtering when joining with other data in subsequent steps. This filtering is particularly useful at intersections where crossing data need not be accounted for.

Once the base route segments were prepared, the IRI_{CVe} data was mapped and aggregated over them to obtain one representative roughness measurement per segment per month. Figure 3.1, Steps e and f illustrate this process. First, each IRI_{CVe} estimate was assigned to a single route segment by choosing the one with the lowest segment ID to prevent oversampling segments that span multiple segments. Assignments were also filtered on their headings with a tolerance of ±10° to ignore intersecting data that is very often from other routes. Daily per-segment medians were computed over all of these assigned segments, and then a second median was taken over each month. The PCI and PASER data were mapped and aggregated to these segments in the same way as the IRI data, but the temporal medians were not necessary. Although this study used monthly and yearly aggregation, this time period can be adjusted as needed. A longer time period gives greater data coverage at the cost of temporal precision. Segments with no IRI_{CVe} data in a given time period were assigned a roughness value of -1 to indicate this but were not removed from the dataset to facilitate data coverage analysis (Figure 3.1, Callout i).

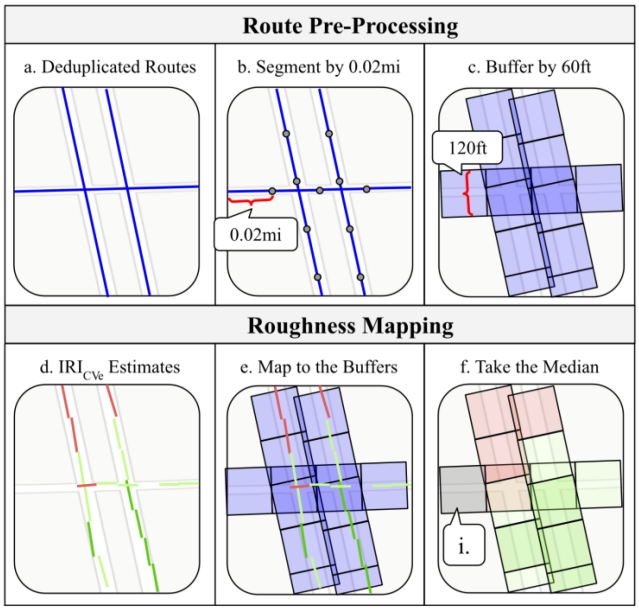


Figure 3.1 An Illustration of Data Processing Steps. Blue is Used to Represent Unmapped Geometry. Additional Colors in Steps d–f Represent Different IRI Categories.

3.2 Threshold Optimization and Metric Agreement

Threshold optimization was conducted on the PCI-IRI and PASER-IRI data to determine the thresholds that most closely correlate to segments categorized as "Needs Maintenance" by their IRI value. Maximizing Youden's statistic (True Positive Rate – False Positive Rate) yielded a PCI threshold of 50 and a PASER threshold of 5, meaning that PCI values less than 50 and PASER values less than 5 correlate best with IRI values greater than 170 in./mi in both datasets. Notably, these thresholds lie at the midpoints of their respective metric scales.

A measure of "metric agreement" was developed to assess the consistency between condition and roughness data. A segment is in metric agreement when both metrics convey similar information about pavement condition. Agreement was assessed using the derived thresholds of 5 and 50. Four possible outcomes were defined:

- True Positive (TP): both condition ≤ threshold and IRI > 170 classify the segment as needs maintenance.
- True Negative (TN): both condition > threshold and IRI ≤ 170 classify the segment as not needing maintenance.
- False Negative (FN): condition > threshold suggests the segment does not need maintenance, but IRI > 170 indicates it does.
- False Positive (FP): IRI ≤ 170 suggests the segment does not need maintenance, but condition < threshold indicates it does.

3.3 Statistical Modeling Assumptions

Because the PCI_{10} and $PASER_{CVi}$ data provided by Noblesville are ordinal, the relationship between visual condition and roughness was modeled using an ordered logit model, which assumes an unobserved latent variable

$$Y^* = \beta_0 + \beta_1 IRI + \varepsilon \qquad \text{(Equation 3.1)}$$

Representing true pavement quality. Thresholds dividing Y* into discrete categories are estimated by Maximum Likelihood. The logistic error assumption, rather than normal as in a probit model, provides robustness to outliers and interpretable coefficients in terms of odds ratios. Model calibration was evaluated using Randomized Quantile Residuals (RQRs), which convert predicted category probabilities into continuous residuals. The Q–Q plots (Figure 3.2) show RQRs closely following the

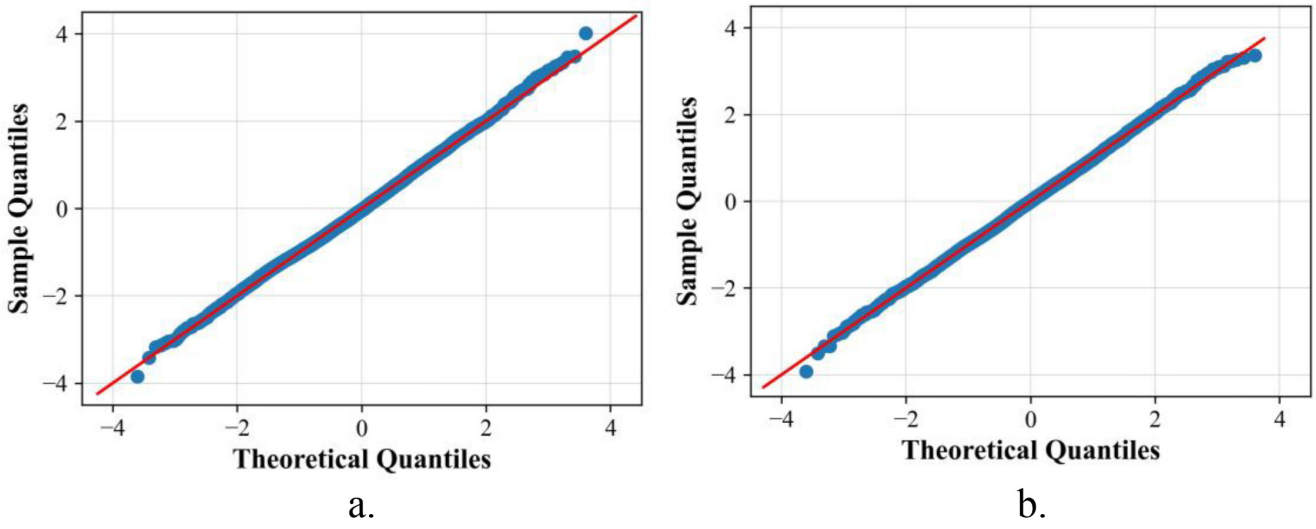


Figure 3.2 Ordinal Model Q–Q Plots a. PCI10, b. $PASER_{CVi}$.

normal distribution with only minor tail deviations, confirming that the logistic error assumption holds. Brown–Forsythe tests indicated strong heteroscedasticity in both PCI_{10} ($p < 10^{-14}$) and $PASER_{CVi}$ ($p < 10^{-5}$) data, validating the choice of an ordinal model as opposed to a linear model. Finally, Brant tests showed that the proportional-odds assumption held for PCI_{10} ($p = 0.16$) but not for $PASER_{CVi}$ ($p \approx 0$), suggesting that IRI_{Cve}'s effect varies across rating thresholds in the automated data. Despite this violation, the model remained well calibrated, so ordered logit was retained for comparability, with $PASER_{CVi}$ coefficients interpreted as averaged effects due to the variation by PASER level.

4. RESULTS

4.1 Statewide Evaluation of Connected Vehicle Roughness Data

Before further analysis was conducted, the availability of the IRI_{CVe} data was evaluated using the measure of spatial coverage, meaning the percentage of road segments in a given area or network that had at least one observation in the time period. Figure 4.1 shows the county-level IRI_{CVe} data coverage for 2023 and 2024 (Subfigures a and b, respectively). Callouts i, ii, and iii correspond to White, Tippecanoe, and Hamilton counties, which serve as reference locations for following analyses. Statewide coverage increased from 46.5% to 53.2% between 2023 and 2024. Many individual counties also saw substantial increases, and only two decreased. As expected, coverage tends to be highest in urban areas where connected vehicle density and traffic is greatest. Rural areas with low population density generally showed lower data availability. The increasing availability of this IRI_{CVe} data emphasizes the growing viability of its use for monitoring pavement conditions across local road networks.

Figure 4.2 provides a detailed comparison of paved local roadway conditions by county and the two blue lines represent statewide state and interstate networks. Each line is a Cumulative

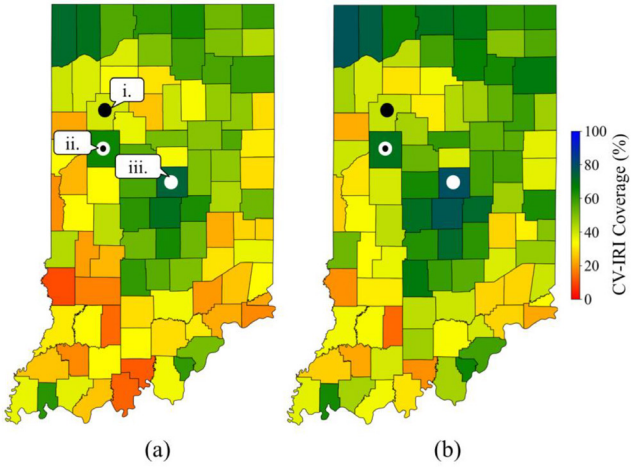


Figure 4.1 Indiana Paved Local Roads IRI Data Coverage: (a) 2023; (b) 2024.

Frequency Distribution (CFD), which visualizes the proportion of segments with a given IRI value or better at every IRI value along the x axis. Lines that are further to the left have more segments in better condition than those that are on the right. The green, light green, and red vertical lines represent the FHWA IRI thresholds for their respective categories. The three counties indicated in Figure 4.1 are colored for reference. The figure shows how state and interstate managed roads are smoother than nearly all local networks, as expected. Among the three highlighted counties, Hamilton's roads are the smoothest, followed by Tippecanoe and White. This type of visualization enables rapid comparison of pavement condition distributions across multiple road networks at a high level.

The Hamilton, Tippecanoe, and White County local road networks were analyzed in further detail in Figure 4.3. This figure illustrates the power of IRI_{CVe}'s high temporal resolution by

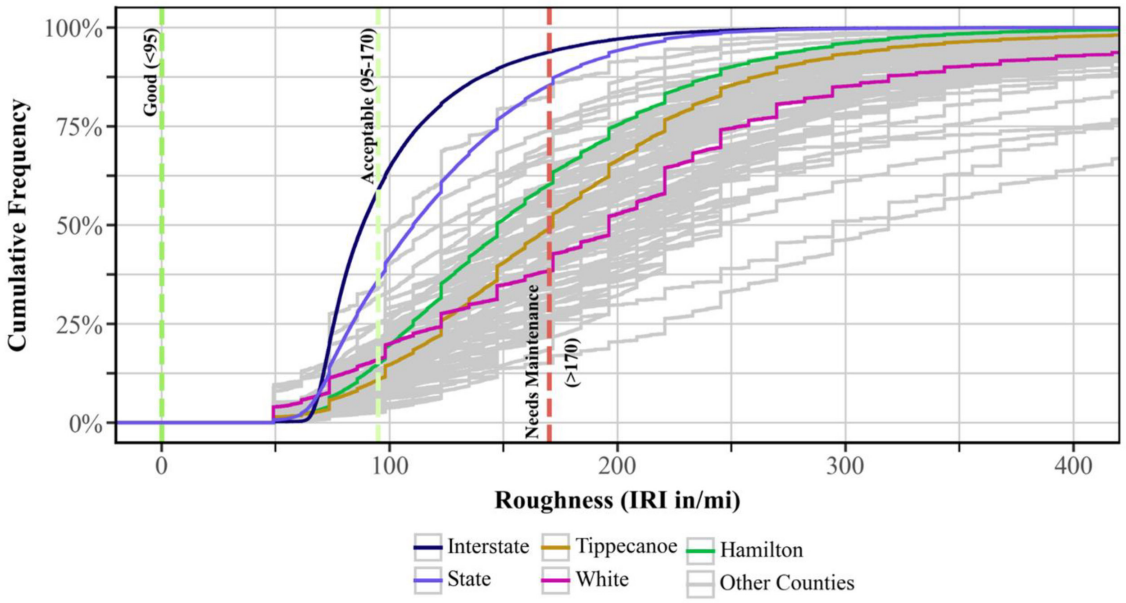


Figure 4.2 Segment IRI CFD by Interstate, State, and Local County Roads in 2024.

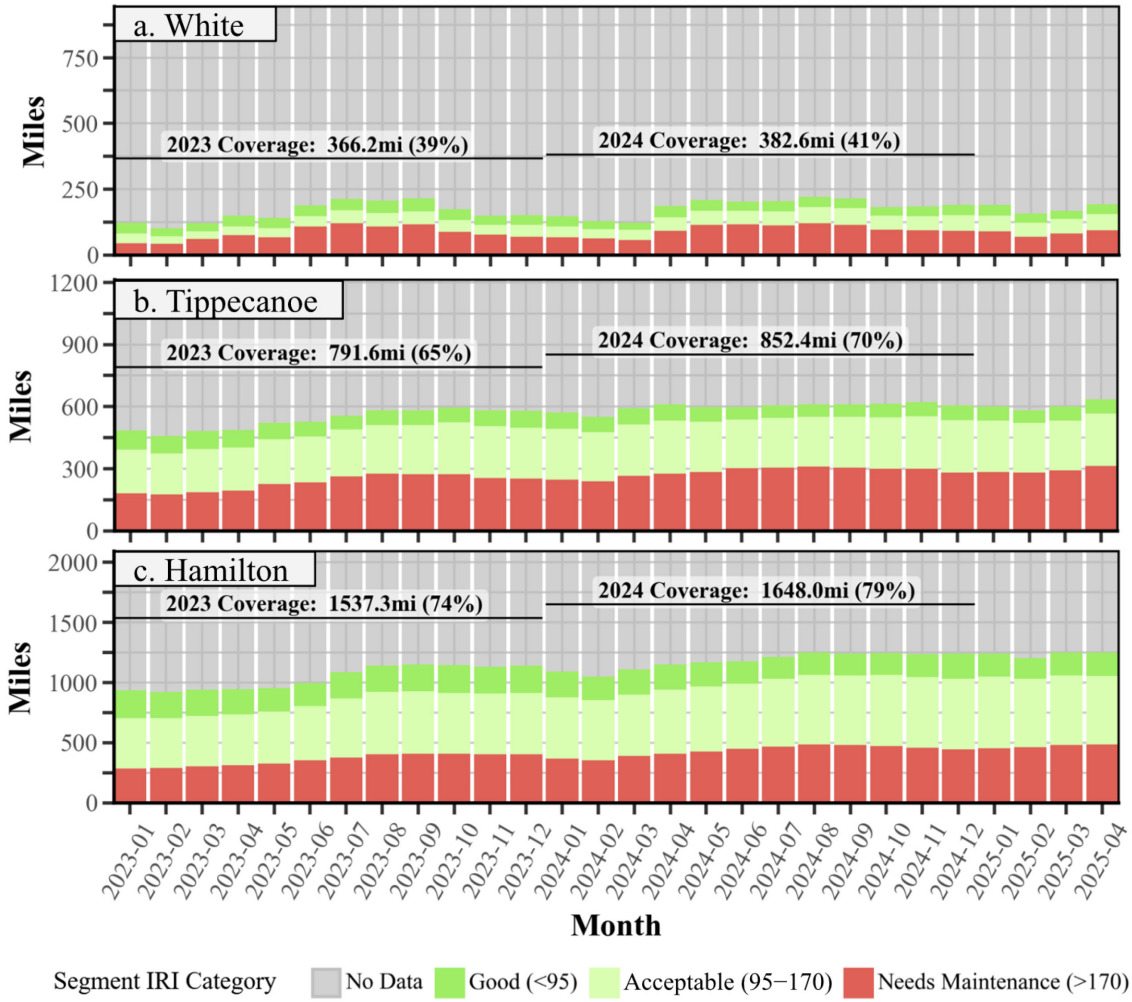


Figure 4.3 Paved Local Roadway IRI_{CVe} Coverage and IRI Categories by Month for Select Counties. (a) White County, (b) Tippecanoe County, (c) Hamilton County.

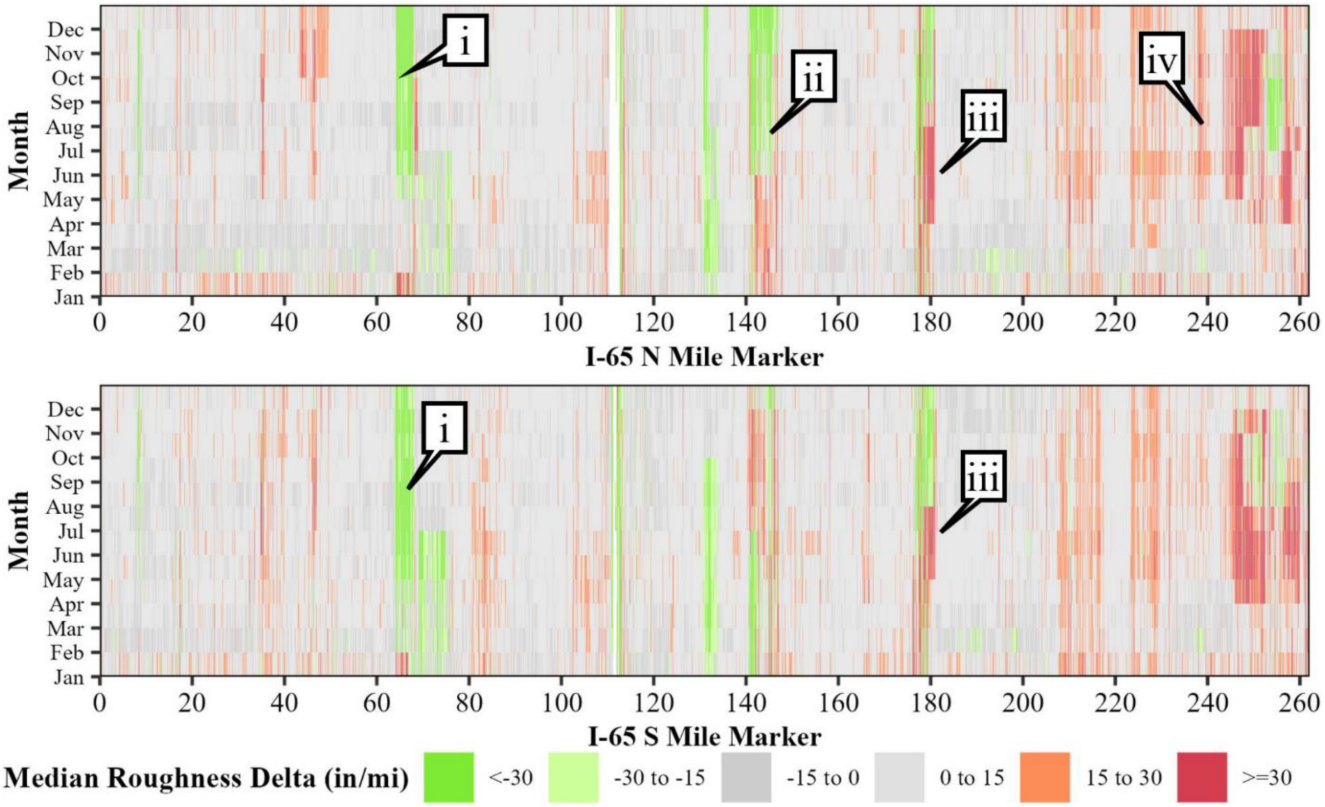


Figure 4.4 Delta IRI Heatmap Between 2022 and 2023 on I-65.

visualizing the change in IRI category distributions, including gray for no data to also convey the county's change in coverage over the 28-month time period from January 2023 to April 2025. This monthly resolution enables the inspection of seasonal changes, like those visible for White County in Subfigure a. This shows how the county has substantially higher data coverage in the summer, which may be explained by seasonal tourists. Additionally, the increase in data coverage for Hamilton and Tippecanoe visible in Figure 4.1 can be seen in greater detail here. The greatest increases were in 2023 but then began to plateau in 2024.

To illustrate the effect of temporal aggregation on data coverage, annual coverage lines were added to all three subfigures. These coverage lines show how an annual aggregation window size leads to roughly 15% higher coverage when compared to that of a single month. This increase is due to the accumulation of IRI_{CVe} observations over time, where longer periods increase the likelihood of receiving observations for low traffic segments that would otherwise have no data.

Individual routes can be analyzed in detail with this data, which can provide valuable insight into localized changes, either degradation, or improvements due to construction. Figure 4.4 is a heatmap showing the relative change in IRI along I-65 between 2022 and 2023. Areas in green and red saw considerable change, while areas in light and dark grey had little to no change in IRI. For example, conditions improved near mile marker (MM) 60 (Callout i) and MM 140 (Callout ii), and conditions worsened near MM 180 (Callout iii) and MM 240 (Callout iv). Route-level visualizations such as this will continue to be an important use case for IRI_{CVe} data due to their relevance for resource allocation. In particular, delta heatmaps will be an important tool for agencies to understand the time and location of changes in pavement condition to inform asset management planning. This case study is discussed in detail by Mathew et al. (2024).

A case study was conducted in Hamilton County to demonstrate how IRI_{CVe} data can also be used for localized condition monitoring at a high resolution, enabling agencies to detect pavement changes with minimal fieldwork. Figure 4.5 uses maps of the case study location in 2023 (a) and 2024 (b). Callouts i and ii indicate a target segment that experienced significant IRI_{CVe} improvement within one year, decreasing from 221 in./mi to 94 in./mi. The following two subfigures use Google Street View imagery to show the transition from a visibly cracked surface in July 2019 (c) to a newly resurfaced condition in September 2024 (d). These findings confirm that roadway maintenance was conducted to improve the surface condition and that the IRI_{CVe} data successfully captured the resulting improvement in ride quality.

Although this case study focused on a single segment, the same analysis can be applied at scale for proactive network monitoring. By continuously tracking segment-level IRI_{CVe} changes over time, agencies can implement automated flagging systems to identify statistically significant improvements or

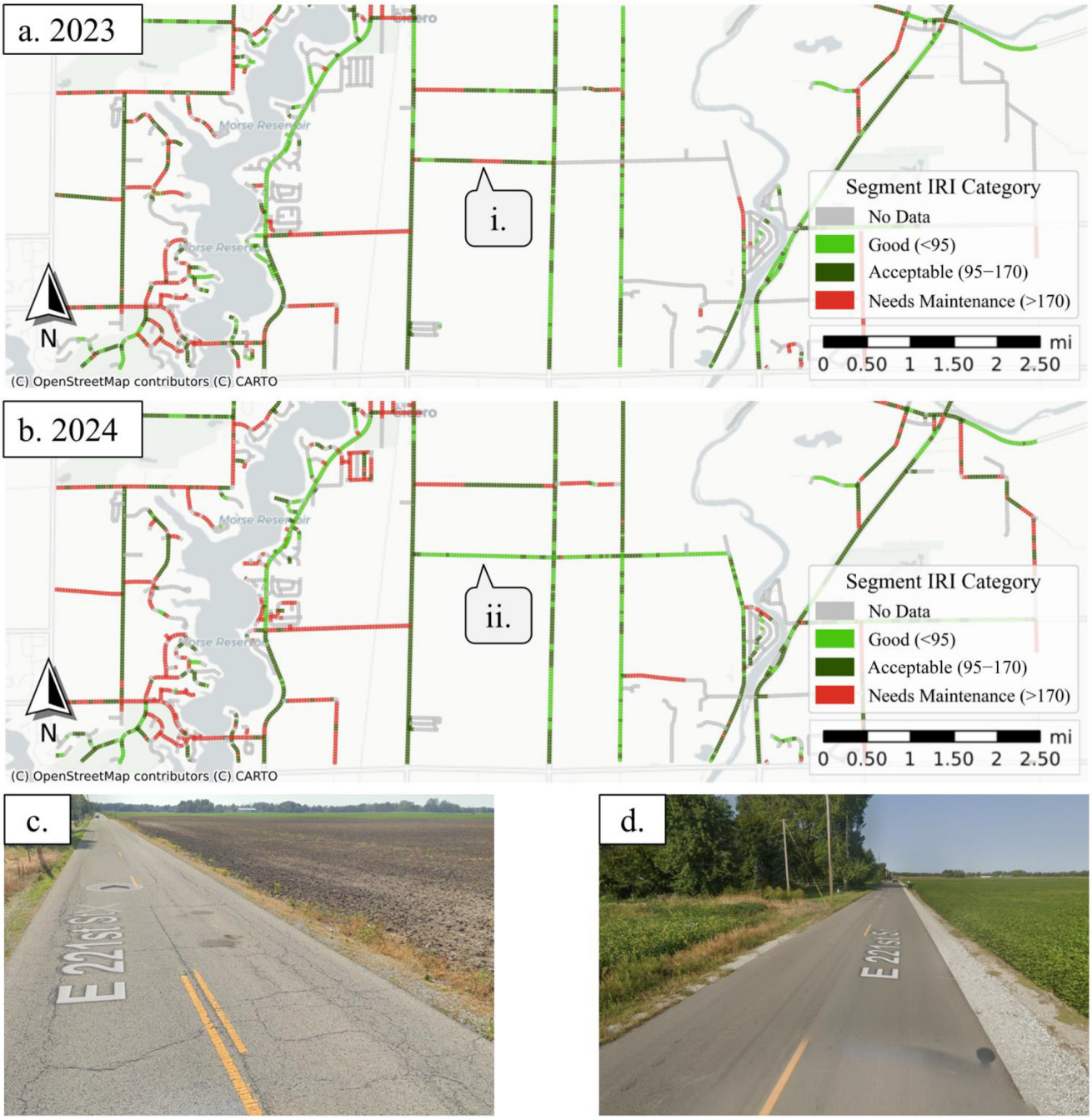


Figure 4.5 Hamilton County E 221st Street Case Study: (a) 2023 IRI Category Map, (b) 2024 IRI Category Map, (c) Google Street View for Location i in July 2019, and (d) Google Street View for Location ii in September 2024.

deteriorations. Such change detection algorithms can help prioritize field inspections, verify completed maintenance, and support data-driven decision making. Notably, roughness change alerts are already offered by some third-party data vendors, making integration into existing workflows feasible without developing additional data-processing infrastructure.

To illustrate the longitudinal variation in roughness over time, a 1-mi section of roadway was analyzed in Figure 4.6. This figure shows the IRI values when using daily, monthly, and yearly aggregation using the blue, pink, and black lines, respectively.

The black yearly aggregated lines indicate that the roughness consistently increased year-over-year between 2022 and 2024; however, the monthly and daily lines make it clear that there are significant interyear variations in roughness. In early 2022, the section deteriorated rapidly (Callout i), but conditions improved after subsequent maintenance that summer (Callout ii). The conditions worsened again in the late winter/early spring of the

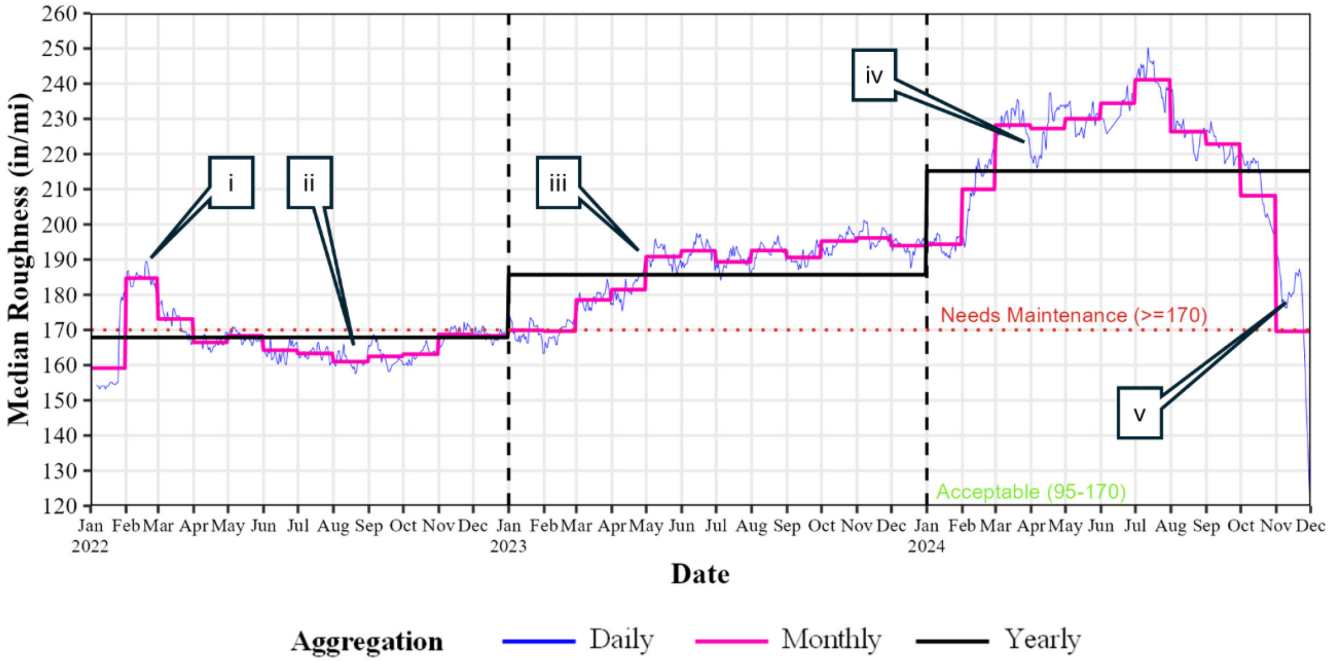


Figure 4.6 Daily, Monthly, and Yearly Roughness Variation on I-65 N MM 238-239 Section.

following two years (Callouts iii and iv). In the second half of 2024, substantial maintenance activities were performed, resulting in substantial reduction in the IRI (Callout v). The sharp reduction in IRI in the last 2 weeks of December 2024 should be disregarded, as it is an artifact of the data filtering.

This analysis also illustrates the impact of various aggregation levels. For example, daily or monthly aggregation would be most appropriate to understand the degradation performance at a project level, while yearly aggregation may be more appropriate for systemwide evaluation. Although Figure 4.3's results include the ~15% increase in coverage when aggregating by year, Figure 4.6 illustrates how temporal variations in the year can be missed. Notably that while the monthly estimates align closely with the daily estimates, the yearly aggregation lines can substantially deviate from them. Agencies should be aware of the tradeoff between the better coverage and easier storage from yearly aggregation and the improved resolution of monthly aggregation.

4.2 Condition-Roughness Relationship Analysis with Automated Data Sources

To assess where $PASER_{CVi}$ and IRI_{CVe} are best used in an existing agency's monitoring practices, they had to be compared to baseline data collected by traditional methods. This section will make three comparisons: Marion County IRI_{CVe} versus PCI, Noblesville PCI_{10} versus IRI_{CVe}, and Noblesville $PASER_{CVi}$ versus IRI_{CVe}. These comparisons will help ground the performance of the new collection methods to the methods that are commonly used in practice today.

We will start with the Marion County dataset that was used to analyze the IRI_{CVe} – PCI relationship in Figure 4.7 and Figure 4.8. Figure 4.7 is a scatterplot with PCI and IRI category lines for reference. The figure indicates a weak inverse relationship between PCI and IRI_{CVe}, with notable outliers in the upper-left and lower-right quadrants representing segments where the two metrics provide conflicting information. A large cluster of segments can be seen around PCI = 87. This may reflect category bias from the human surveyors, as this PCI value is the very minimum required to be considered in the "Good" category as defined by ASTM. If this is truly the result of category bias, it indicates how manually rated PCI is subject to systematic human error and inconsistencies. The scatterplot also shows how the relationship weakens as the pavement condition gets worse. As PCI decreases, the variation in the IRI values increases dramatically, meaning that a segment in good condition is also likely to have a good (low) IRI value, while segments in poor condition are far more unpredictable.

Figure 4.8 uses the metric agreement classifications defined in Section 3.2 stratified by roadway class to assess their relative performance. Primary collector performed the best with 70% agreement, and Local streets performed the worst with only 59% agreement. The most likely explanation for the relatively poor Local street agreement is the higher prevalence of intentional surface anomalies like manholes that impact IRI but do not affect the PCI rating.

To represent the IRI-PCI relationship numerically, the coefficient of determination (R^2) and Matthews Correlation Coefficient (MCC) values were computed across combinations of roadway class and surface type as shown in Table 4.1. The MCC ranges from -1 (complete disagreement) to 1 (perfect agreement), with 0 indicating no correlation. It is calculated with Equation 4.1, using the PCI threshold of 50 and IRI threshold of 170 like defined previously. In comparison to R^2, which

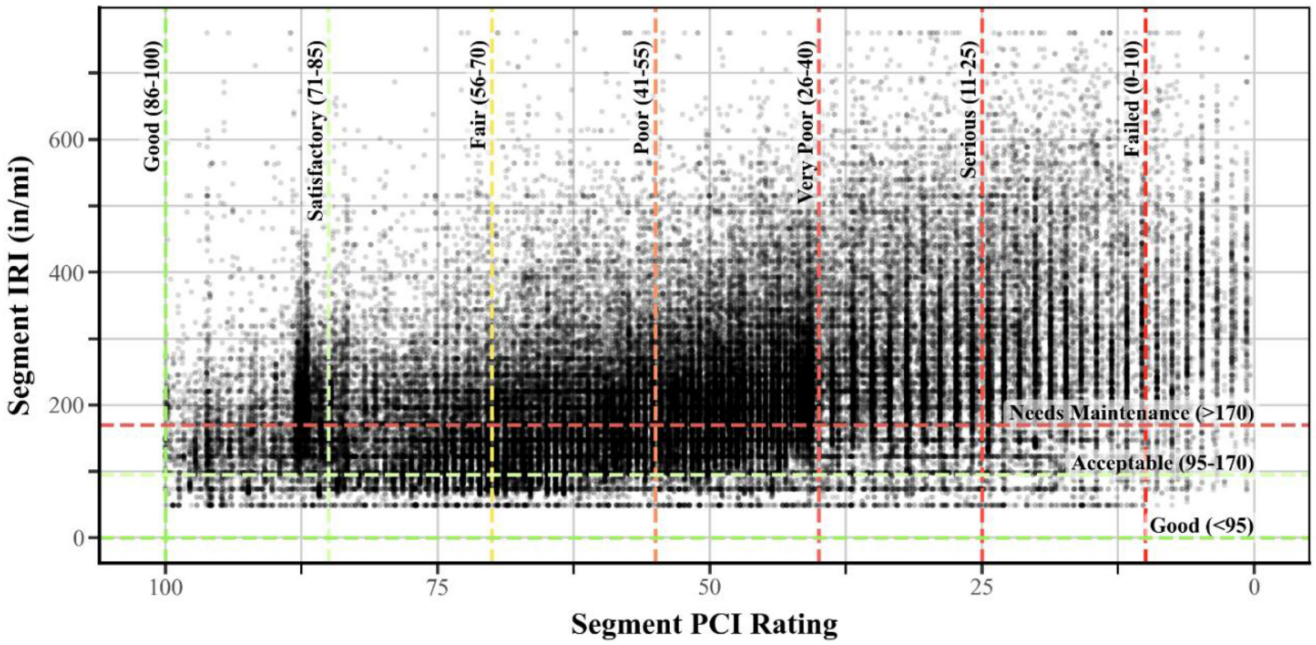


Figure 4.7 Scatterplot of PCI Value vs. IRI Values With PCI and IRI Category Lines in Marion County.

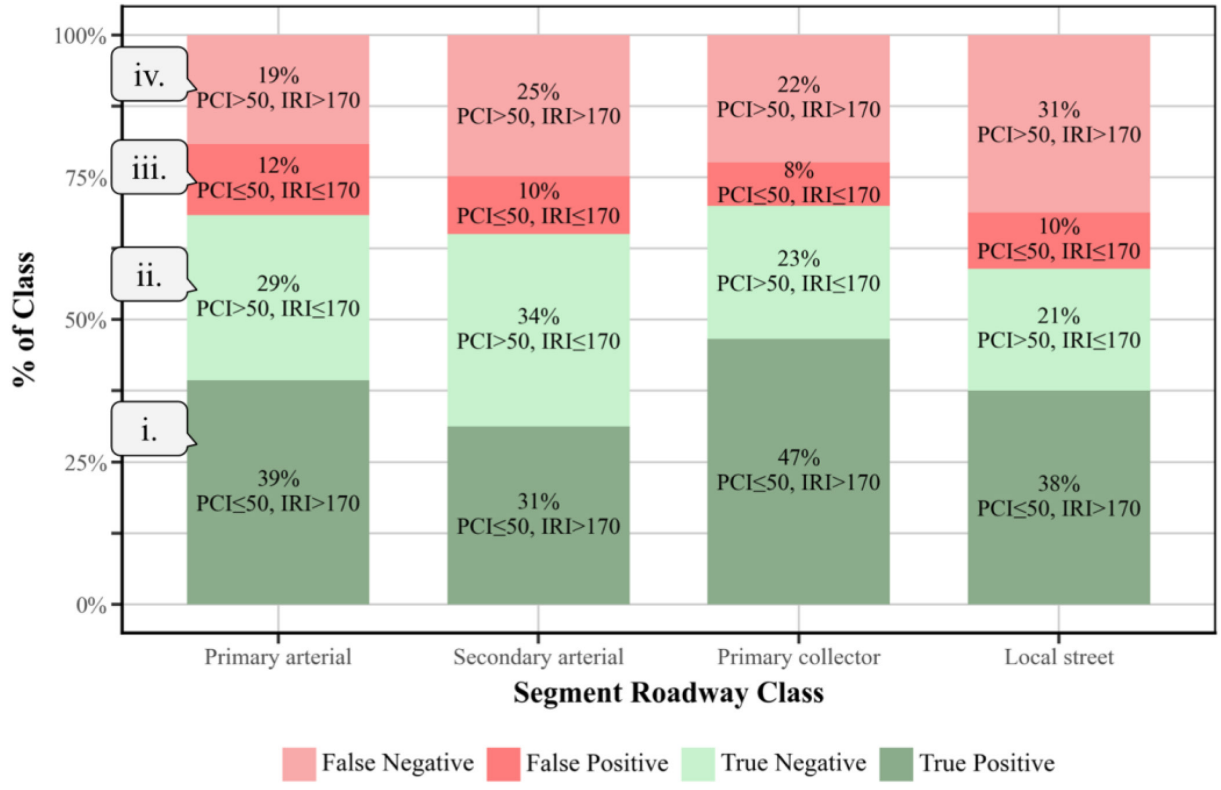


Figure 4.8 IRI and PCI Confusion Matrix Metric Agreement Outcomes by Roadway Class in Marion County.

TABLE 4.1
Marion County MCC and R^2 Correlation Strengths by Roadway Class and Surface Type.

Roadway Class	$MCC_{Asphalt}$	$MCC_{Concrete}$	MCC_{Total}	$R^2_{Asphalt}$	$R^2_{Concrete}$	R^2_{Total}
Primary Arterial	0.37	0.17	0.35	0.18	0.14	0.17
Secondary Arterial	0.3	0.00	0.33	0.21	0.34	0.21
Primary Collector	0.40	0.41	0.40	0.21	0.23	0.21
Local Street	0.21	0.00	0.21	0.15	0.03	0.15
Total	0.30	0.08	0.30	0.15	0.03	0.15

MCC derived from PCI ≤ 50 predicting IRI > 170

measures raw data correlation, the MCC specifically measures the classification correctness of the defined thresholds.

$$MCC = \frac{(TP \cdot TN) - (FP \cdot FN)}{\sqrt{(TP + FP) \cdot (TP + FN) \cdot (TN + FP) \cdot (TN + FN)}} \quad \text{(Equation 4.1)}$$

Results show that the primary collector roadway class agrees most closely with IRI, which is also supported by the findings from Figure 4.3. Asphalt segments tended to have higher MCC values than concrete, likely due to the far greater representation of asphalt in the dataset. Asphalt had 1,900 mi of data, compared to only 14.7 mi for concrete. This makes the PCI threshold more optimized particularly for asphalt segments. R^2 values similarly suggest a weak overall correlation between PCI and IRI, though correlations generally improved when filtered by roadway class and surface type. The highest observed R^2 was 0.34 for secondary arterial roads with a concrete surface and the overall R^2 was 0.15.

Beyond the Marion county dataset, the Noblesville data enable two more comparisons: IRI_{CVe} versus PCI_{10} and IRI_{CVe} versus $PASER_{CVi}$. The boxplots in Figure 4.9 indicate a consistent inverse relationship between roughness and condition in both the PCI_{10} (a) and $PASER_{CVi}$ (b) data. The relationship again shows heteroscedasticity (nonconstant variance), where IRI becomes noisier and the relationship with condition weakens as pavement condition gets worse. The comparison also highlights differences between the manual and automated condition assessment methods. $PASER_{CVi}$ ratings have a higher 0.5-point resolution and lack perfect 10 ratings, as the computer-vision model cannot account for pavement age, which is required for a "like-new" rating under the Michigan PASER manual (Michigan Transportation Asset Management Council, 2026).

Because condition ratings are ordinal, model fit was evaluated using McFadden's pseudo-R^2 rather than the traditional R^2. This metric compares the log-likelihood of the fitted, or full, model detailed in Section 3.3 with that of a null model containing only an intercept term.

$$R^2_{McFadden} = 1 - \frac{log\ L_{full}}{log\ L_{null}} \quad \text{(Equation 4.2)}$$

Pseudo-R^2 values of 0.10 (PCI_{10}) and 0.12 ($PASER_{CVi}$) match those reported in prior profiler-based IRI studies (Hanandeh et al., 2022). Likelihood ratio tests ($p < 10^{-4}$) confirmed that IRI_{CVe} significantly improves prediction of condition rating.

The IRI_{CVe} condition distribution remained stable over the time period across all roadway classes. In contrast, the condition distributions shifted significantly toward the extremes in 2025. The proportion of "Fair" ratings (rating 5–7) declined from 44 % to 24 % for local roads and from 53% to 29% for primary roads, while the shares of both "Good" and "Poor" categories increased greatly. This polarization is likely caused by methodological differences, since the PCI_{10} ratings were collected by manual PCI-based assessment and the $PASER_{CVi}$ ratings by a PASER-oriented computer-vision model. A chi-square test confirmed the year-to-year shift in condition distributions was highly significant ($\chi^2 = 713.8$, $p < 10^{-16}$), while roughness distributions showed no significant change ($\chi^2 = 4.4$, $p = 0.11$). These results suggest that the condition datasets are not drawn from the same distribution, whereas the IRI measurements, collected with a consistent method, remained comparable between years.

The ordered logit models were then used to estimate the probability of each condition level as a function of IRI (Figure 4.1). As expected, "Good" probability decreases, and "Poor" probability increases, when roughness increases. The "Fair" category peaks at intermediate roughness levels (around IRI ≈ 200 in./mi), representing the transition between good and poor pavements. These curves can be interpreted directly: for any given IRI value, the plot gives the likelihood of that segment receiving each condition category. The uncertainty reflected in the overlap among categories mirrors real-world subjectivity. For example, two surveyors might reasonably assign different condition ratings to a pavement with moderate roughness. In the PCI_{10} model, "Fair" has the highest probability in the midrange of IRI (100–300 in./mi), whereas $PASER_{CVi}$ shows stronger polarization with lower "Fair" probabilities and higher "Good"/"Poor" probabilities. This is consistent with earlier findings.

Model performance was calculated with the Mean Absolute Error (MAE), defined below:

$$MAE = \frac{1}{n}\sum\left|y_i - \hat{y}_i\right| \quad \text{(Equation 4.3)}$$

where y_i is the true condition rating and $\hat{y}_i$ is the predicted rating (the rating level with highest probability). The MAE for the 2024 and 2025 models was 1.87 and 1.73, respectively, indicating comparable and moderate performance. This modest

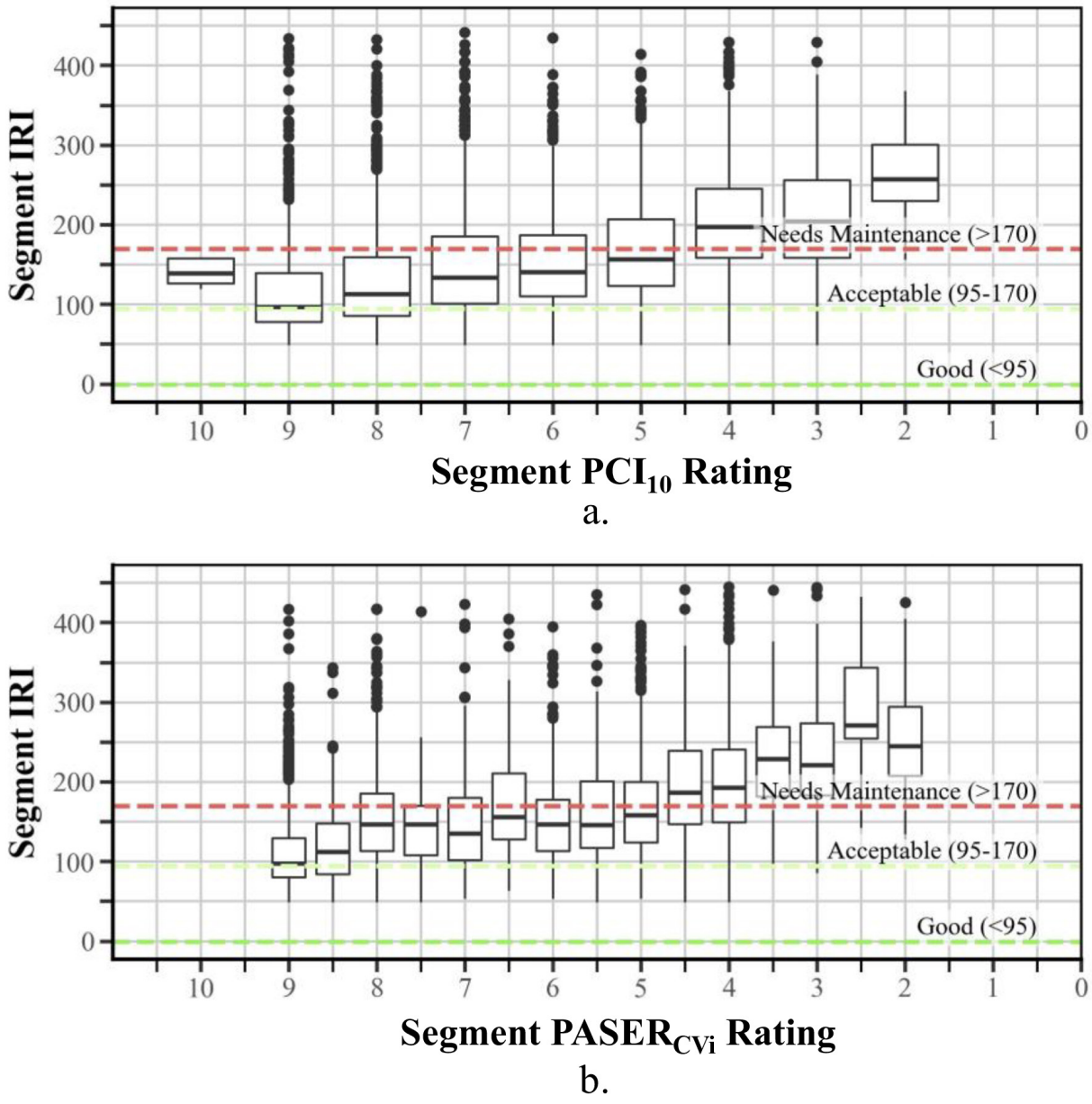


Figure 4.9 Noblesville Condition-Roughness Rating Boxplots, (a) PCI_{10}, (b) $PASER_{CVi}$.

performance is expected given that IRI alone cannot capture the full impact of surface distresses on road condition, as the roughness of a given distress will not always align with its influence on road condition.

Like with the Marion County dataset, the metric agreement distributions were analyzed for each comparison using the thresholds defined in Section 3.2. As shown in Figure 4.11, in both datasets, TP proportions increased and TN proportions decreased from primary to local roads, reflecting poorer surface quality on lower-class roads. For PCI_{10}, FN rates rose substantially for collector and local roads, indicating many rough segments (IRI > 170 in/mi) rated > 5 by surveyors. This is likely due to localized features such as manholes, patches, or speed bumps that inflate IRI without indicating poor surface condition and was also identified in Figure 4.8. $PASER_{CVi}$ showed more uniform disagreement across classes, suggesting reduced bias toward higher-function roads, though FN cases remained higher on locals, again likely reflecting roughness from these anomalies. Overall, disagreement rates (FN + FP) were non-negligible for both datasets, emphasizing that IRI and condition must be interpreted jointly to assess roadway quality.

Using the detailed distress data reported by the $PASER_{CVi}$ survey in 2025, additional insight can be drawn into how specific surface distresses influence the IRI–PASER relationship in particular. Figure 4.12 extends the metric agreement framework introduced in Figure 4.2 by stratifying agreement outcomes

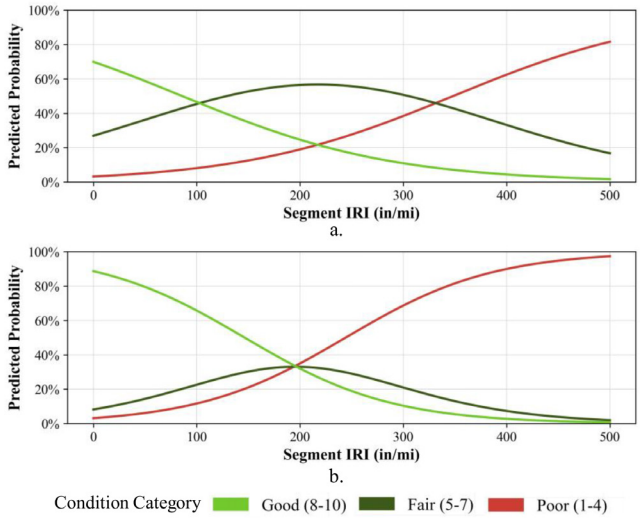


Figure 4.10 Ordered Model IRI_{CVe} Probability Curves by Category, (a) PCI_{10}, (b) $PASER_{CVi}$.

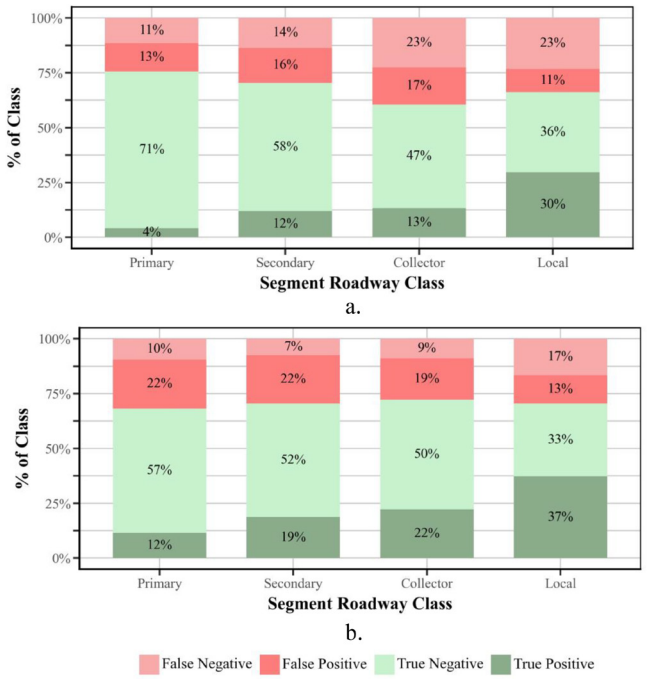


Figure 4.11 Noblesville Condition-Roughness Metric Agreement by Roadway Class, (a) PCI_{10}, (b) $PASER_{CVi}$.

across individual distress types. Because the distress data were reported as the percentage of surface coverage by severity level, a composite distress score was computed by weighting the coverage by severity (1× for light, 2× for moderate, and 3× for severe). The weighted scores were normalized over within-distress distributions using the range between the 5th and 95th percentiles to filter outliers and clipped to the interval (0, 1). Next, these values were averaged to yield percentage values by distress type.

The distresses were ordered by their overall proportion of metric agreement (TP + TN), revealing substantial variation among distress types. Alligator cracking exhibited the highest agreement (78%), while rutting showed the lowest (48%). This is physically intuitive: alligator cracking strongly affects both surface roughness and visual condition; its presence always leads to a PASER ≤ 3 (Michigan Transportation Asset Management Council, 2026; Wisconsin Transportation Information Center, 2013). Accordingly, only 23% of its distress score fell in FN or FP categories, indicating strong PASER–IRI consistency. In contrast, rutting often produces a smooth but deformed surface, yielding low IRI but poor PASER scores (a high FP rate). This discrepancy highlights how surface distresses not directly captured by vertical roughness can still drive PASER deterioration.

Other trends emerge as well. Sealant had the highest proportion of FN outcomes, as it has little influence on PASER ratings but often still increases IRI. Alligator cracking, block cracking, and weathering had the highest TP rates, indicating agreement on poor condition, while longitudinal cracking, distortion, and sealant showed the highest TN rates, which is consistent with these distresses being visually apparent yet having minimal effect on ride quality. Longitudinal cracking in particular can be explained by the use of IRI_{CVe} data, where drivers are more likely to avoid wheel paths with these cracks. When alligator cracking, block cracking, and patching co-occur, the TP rate is at 69%; however, when only patching is present, that rate drops to 14%, demonstrating that distress interaction drives much of the agreement for some distresses like patching. This concept will be explored in greater detail by Table 4.2.

Table 4.2 summarizes $PASER_{CVi}$ rating distributions against guidelines in the literature and shows how the PASER–IRI relationship varies by distress type, accounting for distress co-occurrence.

The empirical PASER values for each distress type were found by calculating the adjusted lift and choosing the top three ratings. The adjusted lift is a co-occurrence adjusted measure of how much a PASER level is over-represented for a given distress. Let $Y \in \{1..10\}$ be the floor of the PASER rating and $D = (D_1, \ldots D_{10})$ be vectors of normalized distress scores. The adjusted lift is calculated as follows:

$$\text{Lift}_{adj}\left(d,k;d^*\right)=\frac{E[\hat{P}(Y=k \mid D_d=d^*)]}{E[\hat{P}(Y=k \mid D_d)]} \qquad \text{(Equation 4.4)}$$

Where $\hat{P}$ is the predicted PASER outcome from an ordered logit model similar to Equation 3.1 but with 10 fitted β_d values, one for each distress type d. The numerator fixes distress d to $d^*=0.9$, while other distress scores remain observed; the denominator uses observed values for all. The top three lift scores identify the most prevalent empirical PASER levels for each distress, which are compared to reference PASER ranges from the Michigan PASER cheat sheet (Michigan Infrastructure Council, n.d.). In Table 4.2, the empirical PASER ratings appear in decreasing order of their lift scores, while the theoretical PASER ratings were simply sorted by their value.

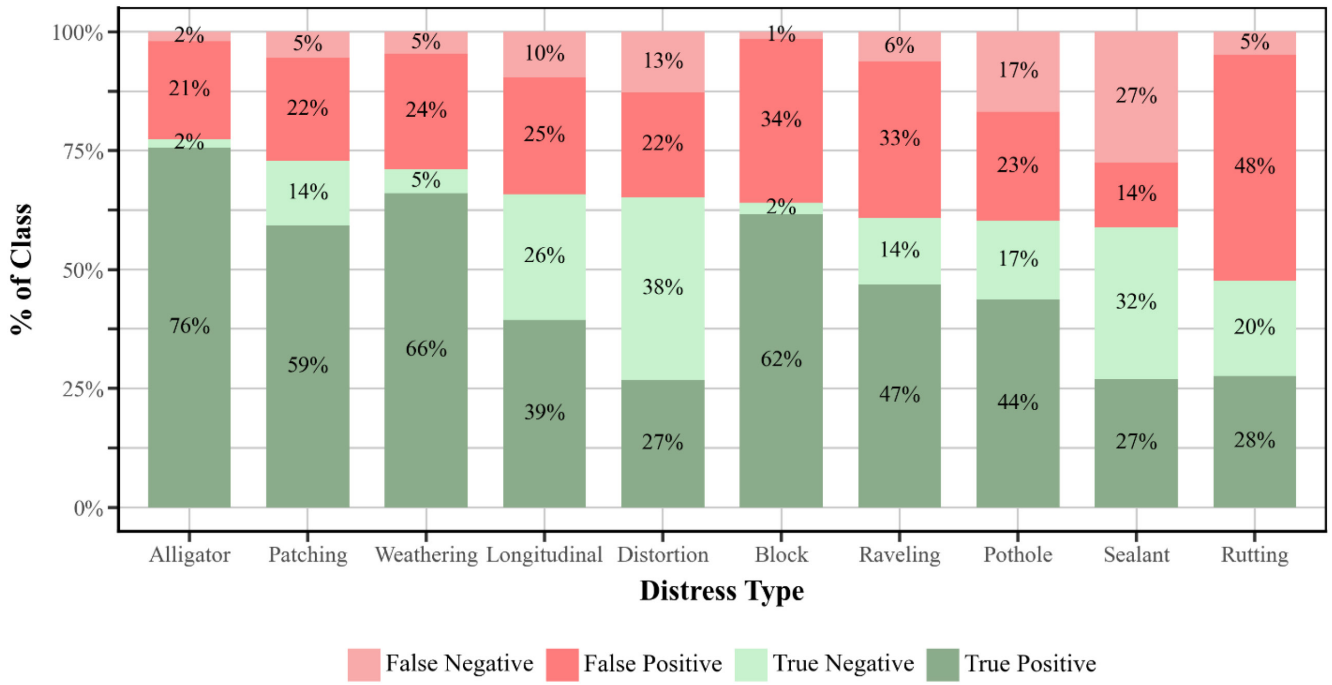


Figure 4.12 Noblesville IRI-PASER Metric Agreement by Distress Type, 2025.

TABLE 4.2
Prevalent PASER Levels and IRI-PASER Correlation by Distress Type.

Distress Type	PASER Empir.	PASER Theor.	τ Naïve	τ Adj.	|Δτ|	Num. Seg.
Alligator Cracking	3, 2, 4	2, 3	-0.27	-0.19	0.08	3,487
Longitudinal Cracking	5, 6, 4	3, 4, 5	-0.39	-0.19	0.20	7,445
Distortion	5, 6, 2	2	-0.36	-0.17	0.19	3,662
Raveling	5, 2, 6	4, 5, 6	-0.33	-0.12	0.21	1,938
Patching	5, 4, 2	2, 3, 4, 5	-0.35	-0.11	0.24	1,354
Block Cracking	4, 2, 3	3, 4, 5	-0.29	-0.10	0.19	1,968
Rutting	4, 2, 3	2, 3, 4	-0.29	-0.10	0.19	872
Pothole	9, 3, 8	2, 3	-0.35	-0.08	0.27	137
Weathering	4, 5, 3	4, 5, 6	-0.20	-0.08	0.12	2,319
Sealant	5, 2, 6	8	-0.23	-0.06	0.17	2,562

Subsequent columns provide insight into the distress-level IRI–PASER correlations, both when adjusting for co-occurrence and when not. The adjusted τ statistic is the residualized Kendall's τ-b partial rank correlation obtained by correlating residuals from two linear models: IRI ~ distress scores and ranked PASER ~ distress scores. This is done to correct for co-occurrence. The naïve τ simply calculates the correlation between IRI and PASER for segments where the given distress is present, and |Δτ| is the magnitude of change between the two, indicating how much the correlation depended on co-occurrence. All τ values are negative, consistent with the expected inverse IRI–PASER relationship.

The final column lists the number of segments exhibiting each distress (score > 0). Because many segments contain multiple distresses, totals exceed the 7,970 unique $PASER_{CVi}$ segments with valid IRI and PASER data, which is higher than the 6,350 segments used elsewhere since the 2024 overlap constraint was removed.

Empirical PASER levels match with reference ranges for most distresses, with two notable exceptions: potholes and sealant. Potholes are rare in this dataset (n = 137), which makes estimates unreliable. Sealant is far more common but typically is associated with other distresses, making it difficult for the ordered model to isolate its effect, which is the likely explanation for the discrepancy. In total, $PASER_{CVi}$ appears to be capturing the correct impact of a segment's distresses on its PASER rating.

The τ correlations quantify how each distress affects the strength of the IRI–PASER relationship. Naïve τ values were consistently higher because they reflected shared variation among co-occurring distresses. After correcting for co-occurrence, adjusted τ values decreased, and the isolated contribution of each distress is derived. Distresses such as alligator and longitudinal cracking retained the strongest negative correlations, while sealant and weathering were weakest.

The magnitude of τ change (|Δτ|) reveals how much apparent correlation arose from distress overlap. Patching, for example, showed a large |Δτ|, which is consistent with earlier results that its agreement depends heavily on co-occurring alligator and block cracking. Meanwhile, alligator cracking showed the smallest change, indicating its correlation with IRI is largely self-driven. This is consistent with existing results from Figure 4.12.

Overall, the distress analysis presented in Figure 4.12 and Table 4.2 highlights how the IRI–PASER relationship depends strongly on distress type. Distresses that directly impact both

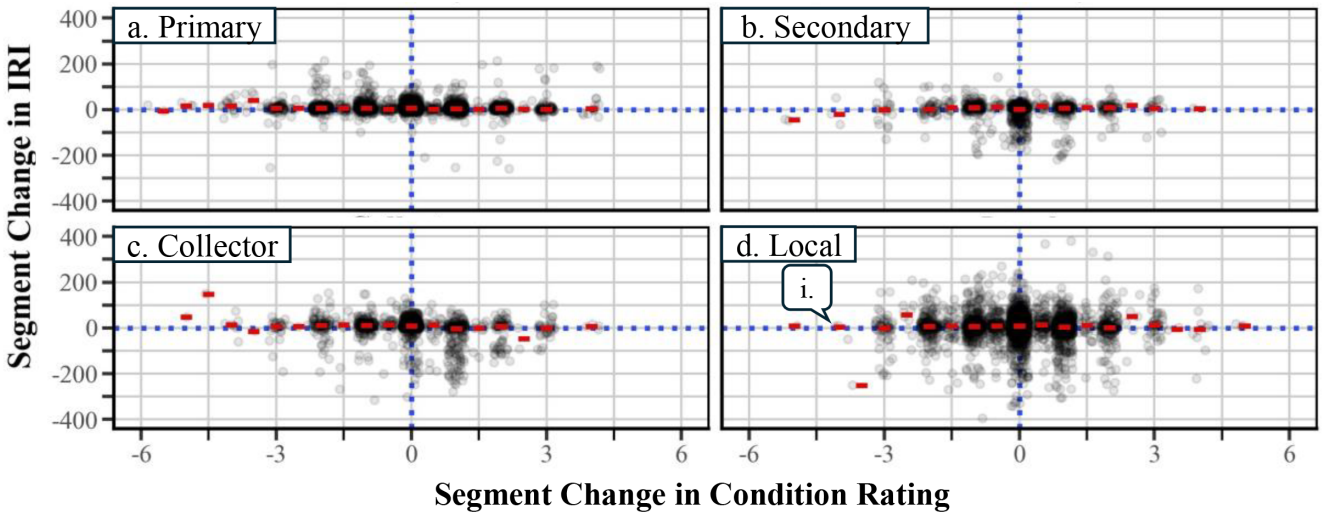


Figure 4.13 Noblesville Delta IRI - Delta Condition Rating Distributions by Roadway Class With Red Median Indicators, (a) Primary, (b) Secondary, (c) Collector, (d) Local.

IRI 2024	294
IRI 2025	300
Delta IRI	6
PCI_{10} 2023	7
$PASER_{CVe}$ 2025	3
Delta Condition	-4

Figure 4.14 Slateford Rd. Case Study, Google Street View Aug. 2024 With Segment IRI and Condition Values. Corresponds to Figure 4.13, Callout i.

visible surface condition and ride quality such as alligator cracking, patching, and distortion tend to correlate strongly, while those affecting surface condition without ride quality such as rutting do not. These results emphasize the complementary nature of IRI and PASER: while IRI captures functional ride quality, PASER reflects visual and surface deterioration, and combining both offers a more comprehensive view of pavement condition that can be very useful. One such use case is explained in detail in the following section.

4.3 Application for Long-Term Condition Monitoring and Quality Control

Figure 4.13 shows year-to-year changes in condition and roughness for matched segments across roadway classes. Each point represents one segment, with the red bars indicating medians and blue dashed lines marking zero change on each axis. Most points cluster tightly around the origin, indicating minimal change in either measure between 2024 and 2025. However, many segments have large changes in condition despite minimal change in roughness. This trend is visible in all roadway classes but is particularly clear in Primary and Secondary roads as they have larger horizontal variance than vertical variance. A localized case study (Figure 4.14) examines the Slateford Road outlier segment identified with Figure 4.13, Callout i. Between 2024 and 2025, its IRI rose only 6 in./mi (from 294 to 300), yet its condition rating fell sharply from 7 to 3. Google Street View imagery from August 2024 shows extensive block and early alligator cracking, inconsistent with the original PCI_{10} = 7, indicating the manual survey likely overestimated condition. Thus, the $PASER_{CVi}$ = 3 rating appears to more accurately reflect the true surface condition.

Such discrepancies illustrate how combined condition-roughness trends enable automated quality control. While an individual error may have a small impact, systematic over- or under-rating in manual datasets can lead to funding misallocation and less effective maintenance prioritization across entire networks. Outlier detection using combined change in roughness and condition can effectively flag inconsistent segments for review, reducing manual workload. Commercial systems already leverage continuous IRI monitoring for roughness alerts; integrating IRI trends with condition changes further improves classification reliability. In this case, the automated approach correctly identified a probable human rating error, demonstrating its value as a scalable validation tool for pavement management.

5. CONCLUSION

This research acts as a synthesis of many aspects of large-scale condition monitoring: current challenges, new approaches that can mitigate them, how these approaches can be implemented in practice, their use cases, and how they compare to traditional methods. Findings include coverage and network condition visualizations made possible by IRI_{CVe} data, as well as correlations and trends from multiple metric comparisons.

- An increase in statewide paved local road IRI_{CVe} coverage from 46.5% to 53.2% was observed between 2023 and 2024.

- A localized case study highlighted a segment showing an IRI improvement from 221 in./mi in 2023 to 94 in./mi in 2024, illustrating how IRI_{CVe} can be used for network-wide continuous condition change detection.
- Multiple condition-roughness comparisons in Marion County and Noblesville, Indiana, indicate a weak inverse correlation with R^2 around 0.15. The relationship is strongest among segments in good condition and weakens with degradation. Stratification by roadway class and surface type tended to improve the correlation to a max of 0.34 among secondary arterial concrete road segments.
- Metric agreement was evaluated for all three metric comparisons stratified by roadway class and found that local roads consistently were most weakly correlated. The $PASER_{CVi}$ data was used to extend this analysis to distress types, finding that presence of alligator cracking is the largest influence on metric agreement, among other insights.
- Outlier segments can be detected using a combination of roughness and condition data, which exemplifies a promising IRI_{CVe} use case for network-wide screening.

The results indicate that large-scale condition monitoring is feasible using a mix of connected vehicle and computer vision data sources, however, they must be implemented with careful consideration and interpretation. These findings suggest that the combination of connected vehicle and computer vision data is a significant advancement in pavement management. Together, they enable objective, high-resolution, and cost-effective condition assessment at scales that are not feasible with traditional manual surveys.

6. FUTURE RESEARCH

There are multiple promising future research directions that would expand and refine the results of this project. Although the study uses statewide IRI_{CVe} data, the relatively small sample of $PASER_{CVi}$ data (127 mi) limits the effect size of conclusions drawn from all comparisons. A future study utilizing $PASER_{CVi}$ derived from crowdsourced dashcam imagery would enable larger comparisons, as well as further insights into large-scale automated data processing and interpretation. On a similar note, the collection of more manually rated condition data reported by local agencies would enable network-level comparisons between agencies to potentially identify bias from varying interagency data collection practices.

Another possible direction for future research is to address the sparsity in IRI_{CVe} measurements for rural and low traffic routes where only a fraction of segments may have data in a given month or year. Advanced modelling approaches such as Spatiotemporal Graph Neural Networks (STGNNs), and particularly Graph Recurrent Imputation Networks (GRINs), can be explored for IRI_{CVe} data imputation, which have recently demonstrated promising results in prior studies involving CV trajectory prediction and traffic forecasting (Chen et al., 2025; Gao et al., 2025). The goal of this approach would be to derive statewide segment-month-level IRI_{CVe} confidence intervals for segments with missing data. The natural graph structure of the road network would be leveraged, along with a fusion of multiple other features that encode pavement degradation such as traffic volume, speed limit, and surface type. The impact of this research would be the ability to extend this project's IRI_{CVe} analysis to underserved rural networks that may otherwise only have <20% data coverage.

REFERENCES

Abohamer, H., Elseifi, M., Dhakal, N., Zhang, Z., & Fillastre, C. N. (2021). Development of a deep convolutional neural network for the prediction of pavement roughness from 3D images. *Journal of Transportation Engineering, Part B: Pavements*, *147*, 04021048. https://doi.org/10.1061/jpeodx.0000310

Agebjär, M., Zetterqvist, G., Gustafsson, F., Wahlström, J., & Hendeby, G. (2025). Road roughness estimation via fusion of standard onboard automotive sensors. In *2025 proceedings of the 28th International Conference on Information Fusion (FUSION)*. IEEE. https://doi.org/10.23919/FUSION65864.2025.11123970

ASTM International. (2020). *Standard practice for roads and parking lots pavement condition index surveys* (ASTM D6433–20). https://doi.org/10.1520/D6433-20

Chen, Y.-T., Liu, A., Li, C., Li, S., & Yang, X. (2025). Traffic flow prediction based on spatial-temporal multi factor fusion graph convolutional networks. *Scientific Reports*, *15*, 12612. https://doi.org/10.1038/s41598-025-96801-1

Dadashova, B., Dobrovolny, C. S., & Tabesh, M. (2021). *Detecting pavement distresses using crowdsourced dashcam camera images* Safe-D National UTC and Texas A&M Transportation Institute Report No. TTI-Student-07). https://vtechworks.lib.vt.edu/server/api/core/bitstreams/9b20748e-c8a5-44fb-83e8-ab211d22d9aa/content

De Blasiis, M. R., Di Benedetto, A., Fiani, M., & Garozzo, M. (2021). Assessing the road pavement roughness by means of LiDAR technology. *Coatings*, *11*(1), 17. https://doi.org/10.3390/coatings11010017

Fakhri, M., & Shahni Dezfoulian, R. (2019). Pavement structural evaluation based on roughness and surface distress survey using neural network model. *Construction and Building Materials*, *204*, 768–780. https://doi.org/10.1016/j.conbuildmat.2019.01.142

Federal Highway Administration. (2014). *Office of Highway Policy Information.* https://www.fhwa.dot.gov/policyinformation/pubs/hf/pl11028/chapter7.cfm

Federal Highway Administration. (2023). *Highway functional classification: Concepts, criteria and procedures* (2023 edition). Washington, D.C. https://rosap.ntl.bts.gov/view/dot/72430/dot_72430_DS1.pdf

Federal Highway Administration. (2024). *Pavements—Inertial profiler—Pavement (IP)*. https://infotechnology.fhwa.dot.gov/inertial-profiler-road-pavement/

Gao, Y., Yang, K., Yue, Y., & Wu, Y. (2025). A vehicle trajectory prediction model that integrates spatial interaction and multiscale temporal features. *Scientific Reports*, *15*, 8217. https://doi.org/10.1038/s41598-025-93071-9

Gerst, K. J. (2009). *Predicting pavement evaluation from the international roughness index* (Master's thesis, Purdue University). Purdue e-Pubs. https://docs.lib.purdue.edu/dissertations/AAI1469853/

Google Cloud Console. (n.d.). US Roads [Dataset]. https://console.cloud.google.com/marketplace/product/united-states-census-bureau/all-roads?hl=en&inv=1&invt=Ab27_Q&project=yt-dl-443015

Hanandeh, S., Hanandeh, A., Alhiary, M., & Al Twaiqat, M. (2022). Application of soft computing for estimation of pavement condition indicators and predictive modeling. *Frontiers in Built Environment*, *8*, 895210. https://doi.org/10.3389/fbuil.2022.895210

Kırbaş, U. (2018). IRI sensitivity to the influence of surface distress on flexible pavements. *Coatings*, *8*(8), 271. https://doi.org/10.3390/coatings8080271

Llopis-Castelló, D., Camacho-Torregrosa, F. J., Romeral-Pérez, F., & Tomás-Martínez, P. (2024). Estimation of pavement condition based on data from connected and autonomous vehicles. *Infrastructures*, *9*(10), 188. https://doi.org/10.3390/infrastructures9100188

Mahlberg, J. A., Li, H., Zachrisson, B., Leslie, D. K., & Bullock, D. M. (2022). Pavement quality evaluation using connected vehicle data. *Sensors*, *22*(23), 9109. https://doi.org/10.3390/s22239109

Mahlberg, J. A., Li, H., Zachrisson, B., Mathew, J. K., & Bullock, D. M. (2024). Applications of using connected vehicle data for pavement quality analysis. *Frontiers in Future Transportation*, *4*, 1239744. https://doi.org/10.3389/ffutr.2023.1239744

Mathew, J. K., Desai, J., Sakhare, R. S., Hunter, J., & Bullock, D. M. (2024). Spatiotemporal analysis of pavement roughness using connected vehicle data for asset management. *Journal of Transportation Technologies, 15*(1), 1–16. https://doi.org/10.4236/jtts.2025.151001

Michigan Infrastructure Council. (n.d.). *Pavement surface evaluation and rating (PASER).* https://www.michigan.gov/mic/tamc/training/paser

Michigan Transportation Asset Management Council. (2026). *TAMC data collection manual: PASER rating system for asphalt, concrete, and brick & block surfaces.* Michigan Department of Transportation, Center for Technology & Training, Michigan Technological University. https://www.ctt.mtu.edu/sites/default/files/resources/paser/tamc-data-collection-manual.pdf

Purdue University. (n.d.). *PASER–Local technical assistance program (LTAP) asset management.* Retrieved November 21, 2025, from https://www.purdue.edu/inltap/asset-management/PASER.php

Randhawa, S., Aygün, E., Randhawa, G., Herfort, B., Lautenbach, S., & Zipf, A. (2024). Paved or unpaved? A deep learning derived road surface global dataset from Mapillary street-view imagery. *arXiv*. Retrieved November 12, 2025, from https://arxiv.org/html/2410.19874v1

Thompson, A., Desai, J., & Bullock, D. M. (2025). Evaluation of connected vehicle pavement roughness data for statewide needs assessment. *Infrastructures*, *10*(9), 248. https://doi.org/10.3390/infrastructures10090248

Wisconsin Transportation Information Center. (2013). *PASER manual: Asphalt pavement surface evaluation and rating* (2nd ed.). University of Wisconsin–Madison. https://ltap.engr.wisc.edu/wp-content/uploads/2019/12/Asphalt-PASER_02_rev13.pdf

APPENDICES

Appendix A. List of Acronyms

Appendix A. List of Acronyms

ASTM	American Society for Testing and Materials
CFD	Cumulative Frequency Distribution
CV	Connected Vehicle
FHWA	Federal Highway Association
FN	False Negative
FP	False Positive
GPS	Global Positioning System
GRIN	Graph Recurrent Imputation Network
IMU	Inertial Measurement Unit
IRI	International Roughness Index
IRI_{CVe}	Connected Vehicle-estimated IRI
LiDAR	Light Detection and Ranging
LTAP	Local Technical Assistance Program
MAE	Mean Absolute Error
MCC	Matthews Correlation Coefficient
MM	Mile Marker
OEM	Original Equipment Manufacturer
PASER	Pavement Surface Evaluation and Rating
$PASER_{CVi}$	Computer Vision-derived PASER
PCI	Pavement Condition Index
PCI_{10}	PCI Divided by 10
RQR	Randomized Quantile Residual
STGNN	Spatiotemporal Graph Neural Network
TAMC	Transportation Asset Management Council
TIGER	Topologically Integrated Geographic Encoding and Referencing
TN	True Negative
TP	True Positive

About the Joint Transportation Research Program (JTRP)

On March 11, 1937, the Indiana Legislature passed an act which authorized the Indiana State Highway Commission to cooperate with and assist Purdue University in developing the best methods of improving and maintaining the highways of the state and the respective counties thereof. That collaborative effort was called the Joint Highway Research Project (JHRP). In 1997 the collaborative venture was renamed as the Joint Transportation Research Program (JTRP) to reflect the state and national efforts to integrate the management and operation of various transportation modes.

The first studies of JHRP were concerned with Test Road No. 1—evaluation of the weathering characteristics of stabilized materials. After World War II, the JHRP program grew substantially and was regularly producing technical reports. Over 1,600 technical reports are now available, published as part of the JHRP and subsequently JTRP collaborative venture between Purdue University and what is now the Indiana Department of Transportation.

Free online access to all reports is provided through a unique collaboration between JTRP and Purdue Libraries. These are available at docs.lib.purdue.edu/jtrp/.

Further information about JTRP and its current research program is available at engineering.purdue.edu/JTRP.

About This Report

An open access version of this publication is available online. See the URL in the citation below.

Thompson, A., Desai, J. C., Sakhare, R. S., & Bullock, D. M. (2026). *Research support to INDOT on systemwide asset condition assessment using connected vehicle data* (Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2026/04). West Lafayette, IN: Purdue University. https://doi.org/10.5703/1288284318612