

Using Electric Vehicle Onboard Data for Pavement Quality Assessment and Management

Mihai Marasteanu, Principal Investigator

Civil, Environmental, and Geo- Engineering

University of Minnesota

January 2026

Research Report

Final Report 2026-05

To get this document in an alternative format or language, please call 651-366-4720 (711 or 1-800-627-3529 for MN Relay). You can also email your request to ADArequest.dot@state.mn.us. Please make your request at least two weeks before you need the document.

Technical Report Documentation Page

1. Report No. MN 2026-05	2.	3. Recipients Accession No.	
4. Title and Subtitle Using Electric Vehicle Onboard Data for Pavement Quality Assessment and Management		5. Report Date January 2026	
		6.	
7. Author(s) Raphael Stern, Shaghayegh Nouhi, Mugurel Turoș, Qizhi He, Zihan Lin, and Mihai Marasteanu		8. Performing Organization Report No.	
9. Performing Organization Name and Address Department of Civil, Environmental, and Geo Engineering University of Minnesota 500 Pillsbury Dr SE, Minneapolis, MN 55455		10. Project/Task/Work Unit No. #2024005	
		11. Contract (C) or Grant (G) No. (c) 1036342 (wo) 93	
12. Sponsoring Organization Name and Address Minnesota Department of Transportation Office of Research & Innovation 395 John Ireland Boulevard, MS 330 St. Paul, Minnesota 55155-1899		13. Type of Report and Period Covered Final Report	
		14. Sponsoring Agency Code	
15. Supplementary Notes http://mdl.mndot.gov/			
16. Abstract (Limit: 250 words) This project investigated the feasibility of using electric vehicle onboard diagnostics (OBD-II) data for pavement quality assessment in Minnesota, inspired by Denmark's Live Road Assessment (LiRA) project. The research addressed the critical need for cost-effective, continuous pavement monitoring to supplement expensive instrumented vehicle surveys traditionally used by the Minnesota Department of Transportation (MnDOT). The methodology involved a comprehensive OBD-II scanner evaluation, selecting the OBDLink LX device for its superior compatibility with electric vehicles and ability to collect 694 Parameter IDs (PIDs). Data collection occurred across three test routes: MN-36 loop, Metro to Northfield loop, and MnROAD facility, generating around 10,000 data points synchronized with MnDOT's high-precision Pathways van measurements. The LiRA software framework was adapted through systematic data preprocessing, feature engineering using Sequential Feature Selection, and machine learning model development. Results demonstrated that hybrid Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) models achieved the best performance when using speed combined with vehicle orientation and acceleration features, yielding root mean square error (RMSE) values of 0.172 compared to MnDOT reference data. The methodology showed correlation with both MnDOT measurements and NIRA Dynamics network-level data across diverse road conditions, though performance varied based on environmental factors and data collection conditions. Implementation guidelines were developed for large-scale deployment, including equipment standardization protocols, data-processing pipelines, and risk-mitigation strategies. The research indicates that vehicle sensor data can provide useful pavement condition information for continuous infrastructure monitoring with reduced incremental cost. This approach shows potential for supplementing traditional pavement management systems, though successful implementation requires careful consideration of multiple operational factors and validation protocols to ensure reliable performance across varying conditions.			
17. Document Analysis/Descriptors Pavement condition, Quality assurance, Electric vehicles, Automatic data collection systems, Machine learning, Roughness, Monitoring		18. Availability Statement No restrictions. Document available from: National Technical Information Services, Alexandria, Virginia 22312	
19. Security Class (this report) Unclassified	20. Security Class (this page) Unclassified	21. No. of Pages 128	22. Price

Using Electric Vehicle Onboard Data for Pavement Quality Assessment and Management.

Final Report

Prepared by:

Raphael Stern
Shaghayegh Nouhi
Mugurel Turos
Qizhi He
Zihan Lin
Mihai Marasteanu
Department of Civil Engineering
University of Minnesota

January 2026

Published by:

Minnesota Department of Transportation
Office of Research & Innovation
395 John Ireland Boulevard, MS 330
St. Paul, Minnesota 55155-1899

This report represents the results of research conducted by the authors and does not necessarily represent the views or policies of the Minnesota Department of Transportation or the University of Minnesota. This report does not contain a standard or specified technique.

The authors, the Minnesota Department of Transportation, and the University of Minnesota do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to this report.

Acknowledgements

The research team extends sincere gratitude to the Minnesota Department of Transportation (MnDOT) for funding and supporting this research initiative under Contract 1036342, Work Order 93. Special appreciation is given to MnDOT personnel who facilitated data collection efforts and provided access to the Pathways van reference measurements.

We acknowledge the contributions of the MnROAD facility staff who coordinated controlled testing environments during field work phases.

Thank you to the members of this project's Technical Advisory Panel: Curt Turgeon, Barbara Fraley, Jed Falgren, Bernard Izevbekhai, Stephani Clark, Je Falgren, Marcos Sanchez-Pliego, Thomas Nordstrom, Adam Wellner, and Trisha Stefanski.

The authors appreciate the foundational work of the Danish Live Road Assessment (LiRA) project team, whose open-source methodology provided essential guidance for this adaptation to Minnesota conditions.

Contents

Chapter 1: Introduction	1
1.1 Project Overview	1
1.2 Implementation of Benefits.....	1
Chapter 2: Review of LiRA Project.....	2
2.1 LiRA Project Overview	2
2.2 Data Collection	3
2.3 LiRA Data Processing and Structuring.....	3
2.4 LiRA Data Analysis and Interpretation.....	4
2.5 Modeling.....	5
2.5.1 Prediction of International Roughness Index (IRI)	5
2.5.2 Estimating Unevenness	6
2.5.3 Estimating Friction.....	6
2.5.4 Classification and Regression Results.....	9
2.6 LiRA Software Architecture	10
2.7 LiRA Project's Goals	10
2.8 Challenges and Limitations of LiRA Project	11
2.9 Related Research Effort	11
Chapter 3: Adapt OBD-II scanners for the new car(s)	16
3.1 Finding Compatible OBD-II Scanners.....	16
3.2 Data Collection	19
3.2.1 Conventional Volkswagen vehicle.....	19
3.2.2 MnROAD Data Collection Efforts	19
3.2.3 Toyota Data Collection	20
3.2.4 Highway 55 Data Collection	20

3.2.5 Highway 36 Data Collection	20
3.3 Data Collection Results	20
3.3.1 Volkswagen Data Collection Descriptive Results	21
3.3.2 MnROAD Data Collection Descriptive Results.....	22
3.3.3 Toyota Data Collection Descriptive Results	25
3.3.4 Highway 55 Data Collection Descriptive Results.....	27
3.3.5 Highway 36 Data Collection Descriptive Results.....	30
3.4 MnDOT Reference Data Points and Next Steps.....	34
Chapter 4: Adapting LiRA Software	36
4.1 Data Pre-processing.....	36
4.2 Feature Selection	40
4.3 Prediction Modeling	40
4.3.1 Speed – 3D Acceleration	40
4.3.2 Speed-Vertical Acceleration	45
4.3.3 Speed-Other Features	50
Chapter 5: Collect Data	58
5.1 Data Collections	58
5.1.1 MN-36 Loop.....	59
5.1.2 Metro to Northfield Loop.....	59
5.1.3 MnROAD Loop.....	61
Chapter 6: Data Analysis and Recalibration of the Danish Models for Current Project Conditions	64
6.1 An Overview of the Data Cleaning Process	64
6.1.1 Initial Data Assessment and Quality Control.....	64
6.1.2 GPS Coordinate Validation and Spatial Consistency	65
6.2 Location Synchronization and Alignment	66

6.2.1 LiRA's Advanced Signal-Based Data Processing and Alignment Methodology:	67
6.2.2 Comparison with Our Geographic Information System-Based Processing Approach:	67
6.3 Feature Selection Process.....	68
6.3.1 Expansion from Chapter 4, SFS Implementation	68
6.3.2 Statistical Relevance and Physical Interpretability Analysis	69
6.3.3 Computational Efficiency and Implementation Constraints	69
6.4 An Overview of the Modeling Process	69
6.4.1 Data Preprocessing and Sequence Preparation.....	70
6.4.2 Model Architecture and Training Strategy.....	70
6.4.3 Hyperparameter Optimization and Calibration	71
6.4.4 Model Performance Evaluation Framework	71
6.5 Model Quality and Performance Analysis	71
6.5.1 Statistical Accuracy Assessment.....	71
6.5.2 Performance Evaluation Across Speed Conditions	77
6.6 Prediction Results Comparison with Reference Data.....	77
6.6.1 Model Performance Against MnDOT Ground Truth	77
Chapter 7: Guideline for Large-scale Implementation.....	84
7.1 NIRA Data as Alternative Reference Standard	84
7.1.1 Comparative Analysis: MnDOT vs. NIRA Reference Data	84
7.1.2 Model Prediction Accuracy Across Different Reference Standards.....	93
7.1.3 NIRA IRI Estimation Comparison: Daily and Seasonal Variations.....	94
7.2 Generalization of Scanners to Other Vehicles.....	98
7.2.1 Data collection with the MnDOT Van.....	98
7.2.2 IRI Prediction	100
7.3 Implementation Strategy and Framework for Large Scale.....	103

7.3.1 Phased Implementation Approach	103
7.3.2 Geographic Coverage and Prioritization	105
7.4 Operational Guidelines and Best Practices.....	105
7.4.1 Optimal Collection Conditions	105
7.4.2 Data Risk Assessment and Mitigation Strategies	105
7.5 Immediate and Medium-Scale Implementation Guidelines	106
7.5.1 Data Collection and Minimum Data Requirements	106
7.5.2 Recommended Procedure for Immediate Implementation	107
7.5.3 Limitations and Considerations.....	107
7.5.4 Recommended Validation and Calibration Steps.....	107
Chapter 8: Conclusions and Future Work	109
References.....	110

List of Figures

Figure 1. Envisioned components for live road condition assessment with internal vehicle sensors (Levenberg, et al., 2021).	2
Figure 2. Machine learning flow chart (Levenberg, et al., 2021).	5
Figure 3. Schematic of the multi-feature fusion network for TRFC estimation (Du, et al., 2023).	8
Figure 4. Ordered OBD-II scanners.	16
Figure 5. Research team conducting data collection at MnROAD.....	20
Figure 6. Speed (mph) over time (sec) for Volkswagen data collection.	21
Figure 7. Acceleration (ft/s ²) over time (sec) for Volkswagen data collection.....	22
Figure 8. Speed (mph) over time (sec) for MnROAD data collection.	22
Figure 9. Acceleration (ft/s ²) over time (sec) for MnROAD data collection.....	23
Figure 10. Speed color-coded GPS trace for MnROAD data collection drive.	23
Figure 11. Z component of Acceleration color-coded GPS trace for MnROAD data collection drive.	24

Figure 12. Color-coded map of total acceleration for MnROAD data collection.....	24
Figure 13. Speed (mph) over time (sec) for Toyota data collection.	25
Figure 14. Acceleration (ft/s ²) over time (sec) for Toyota data collection.....	25
Figure 15. RR and RL Wheels speed over time (sec) for Toyota data collection.	26
Figure 16. Speed color-coded GPS trace for Toyota data collection drive.	27
Figure 17. Z component of acceleration color-coded GPS trace for Toyota data collection drive.....	27
Figure 18. Speed (mph) over time (sec) for Highway 55 data collection.....	28
Figure 19. Acceleration (ft/s ²) over time (sec) for Highway 55 data collection.	28
Figure 20. Speed color-coded GPS trace for Highway 55 data collection drive.	29
Figure 21. Z component of acceleration color-coded GPS trace for Highway 55 data collection drive.	30
Figure 22. Speed (mph) over time (sec) for Highway 36 data collection.....	30
Figure 23. Acceleration (ft/s ²) over time (sec) for Highway 36 data collection.	31
Figure 24. Speed color-coded GPS trace of the first loop for Highway 36 data collection drive.....	32
Figure 25. Z component of acceleration color-coded GPS trace of the first loop for Highway 36 data collection drive.....	32
Figure 26. Speed color-coded GPS trace of the second loop for Highway 36 data collection drive.	33
Figure 27. Z component of acceleration color-coded GPS trace of the first loop for Highway 36 data collection drive.....	33
Figure 28. Framework of the proposed IRI prediction model.	36
Figure 29. ArcGIS map of the final aligned data points from both OBD and MnDOT datasets for the MN-36 loop.	37
Figure 30. ArcGIS map of the final aligned data points from both OBD and MnDOT datasets for Metro to Northfield loop.....	38
Figure 31. ArcGIS map of the final aligned data points from both OBD and MnDOT datasets for MnROAD inside loop.....	39
Figure 32. ArcGIS map of the final aligned data points from both OBD and MnDOT datasets for MnROAD outside loop.	39
Figure 33. Predicted vs. actual IRI for CNN-LSTM model testing set.	41

Figure 34. Predicted vs. actual IRI for RF model testing set.	42
Figure 35. Predicted vs. actual IRI for RF model testing set.	43
Figure 36. Predicted vs. actual IRI for SVR model testing set.	44
Figure 37. Predicted vs. actual IRI for CNN-LSTM model testing set.	46
Figure 38. Predicted vs. actual IRI for RF model testing set.	47
Figure 39. Predicted vs. actual IRI for KNN model testing set.	48
Figure 40. Predicted vs. actual IRI for SVR model testing set.	49
Figure 41. Predicted vs. actual IRI for CNN-LSTM model training set.....	51
Figure 42. Predicted vs. actual IRI for CNN-LSTM model testing set.	51
Figure 43. Training and validation loss vs epochs for showing the CNN-LSTM model performance.	52
Figure 44. Predicted vs. actual IRI for ANN model testing set.	53
Figure 45. Predicted vs. actual IRI for SVR model testing set.	54
Figure 46. Predicted vs. actual IRI for RF model testing set.	55
Figure 47. Predicted vs. actual IRI for KNN model testing set.	56
Figure 48. ArcGIS map of OBD-II data collection of MN-36 loop.....	59
Figure 49. Color-coded map of speed for Metro to Northfield loop data collection.	60
Figure 50. (a) Embedding the scanner in the vehicle for data collection. (b) MnDOT Pathway van leading our data collection vehicle on MnROAD.....	61
Figure 51. Map of OBD-II data collection on the MnROAD loop. Green points represent the inside loop, while red points indicate the outside loop.	62
Figure 52. Map of MnDOT data collection on the MnROAD loop. Green points represent the inside loop, while purple points indicate the outside loop.	62
Figure 53. The training and testing performance of the CNN-LSTM model on all datasets, with speed and other features as input features.	75
Figure 54. The training and validation loss vs epochs for the CNN-LSTM model on all datasets, with speed and other features as input features.	76
Figure 55. Actual and predicted IRI from the CNN-LSTM model output on all datasets, with speed and other features as input features.	76

Figure 56. True IRI color-coded map of MN-36 loop.	78
Figure 57. True IRI color-coded map of Northfield loop.....	78
Figure 58. Color-coded visualization of true IRI measurements along the inside loop.	79
Figure 59. Predicted IRI color-coded map of Northfield loop.....	80
Figure 60. Percentage difference of our IRI predictions with MnDOT’s ground truth measurements for the Metro to Northfield loop.	81
Figure 61. Percentage difference of our predictions with MnDOT’s ground truth measurements for the MN-36 loop.	82
Figure 62. Percentage difference of our predictions with MnDOT’s ground truth measurements for the inside loop of MnROAD, with a constant speed of 40 mph.....	82
Figure 63. IRI color-coded map of Metro to Northfield loop, MnDOT ground truth measurements.	85
Figure 64. IRI color-coded map of Metro to Northfield loop, NIRA predictions.....	86
Figure 65. IRI color-coded map of MN-36 loop, MnDOT ground truth measurements.	87
Figure 66. IRI color-coded map of MN-36 loop, NIRA predictions.	87
Figure 67. Percentage difference between our model's predictions and NiRA predictions across the Metro to Northfield loop.	89
Figure 68. Percentage difference between the NIRA model's predictions and MnDOT ground truth measurements across the Metro to Northfield loop.....	90
Figure 69. Percentage difference between the NIRA model's predictions and MnDOT ground truth measurements across the MN-36 loop.....	91
Figure 70. Percentage difference between our model's predictions and NiRA predictions across the MN-36 loop.	91
Figure 71. IRI percentage comparison between NIRA ground truth measurements and NIRA predictions for the MN-36 loop.	92
Figure 72. IRI percentage comparison between NIRA ground truth measurements and NIRA predictions for the Metro to Northfield loop.	93
Figure 73. IRI color-coded map of a portion of MN-36 - March 2nd.	94
Figure 74. IRI color-coded map of a portion of MN-36 - March 3rd.....	95
Figure 75. IRI color-coded map of a portion of MN-36 – November 2024.....	96

Figure 76. Figure 14. IRI color-coded map of a portion of MN-36 – January 2025.	96
Figure 77. IRI color-coded map of a portion of MN-36 - March 2025.	97
Figure 78. IRI color-coded map of a portion of MN-36 – July 2025.....	97
Figure 79. Scatter plots of key vehicle dynamics features versus time collected using OBD scanner with MnDOT van on the MnROAD loop.....	100
Figure 80. IRI percentage comparison between MnDOT ground truth measurements and Van collected data predictions for the MnROAD loop.	101
Figure 81. Settlement section of HMA next to PCC in MnROAD.	102
Figure 82. Settlement section of HMA next to PCC in MnROAD on the map.....	102
Figure 83. Transverse Ruts in PCC.....	102
Figure 84. PCC Drain Between HMA Cells.....	103
Figure 85. Three-phase implementation guidelines for large-scale OBD-II-based pavement quality assessment across Minnesota's road network.	104

List of Tables

Table 1. Comparison of Different sensor inputs for the ML models, where the best is used for motorways.	6
Table 2. Comparison of Different sensor inputs for the ML models, where the best is used for the combined motorway and municipality.	6
Table 3. Vehicle dynamics features are used for training the proposed network (Du, et al., 2023).....	9
Table 4. Threshold value of TRFC levels for the classification task (Du, et al., 2023).....	9
Table 5. Summary of OBD-II scanners.....	17
Table 6. Description of CAN signals utilized in the LiRA project.	18
Table 7. Custom PIDs for wheels speed in the Torque app.	19
Table 8. Optimized feature subset using speed and 3D acceleration.	40
Table 9. Performance metrics of CNN-LSTM method with speed and 3D acceleration input features.	41
Table 10. Performance metrics of RF method with speed and 3D acceleration input features.	42

Table 11. Performance metrics of KNN method with speed and 3D acceleration input features.	43
Table 12. Performance metrics of SVR method with speed and 3D acceleration input features.	44
Table 13. Optimized feature subset using speed and vertical acceleration.	45
Table 14. Performance metrics of CNN-LSTM method with speed and vertical acceleration input features.	46
Table 15. Performance metrics of RF method with speed and vertical acceleration input features.	47
Table 16. Performance metrics of KNN method with speed and vertical acceleration input features.	48
Table 17. Performance metrics of SVR method with speed and vertical acceleration input features.	49
Table 18. Optimized feature subset using speed and other features.	50
Table 19. Performance metrics of CNN-LSTM method with speed and other influential input features. .	52
Table 20. Performance metrics of ANN method with speed and other influential input features.	53
Table 21. Performance metrics of SVR method with speed and other influential input features.	54
Table 22. Performance metrics of RF method with speed and other influential input features.	55
Table 23. Performance metrics of KNN method with speed and other influential input features.	56
Table 24. Summary of data collection efforts.....	58
Table 25. Performance comparison of our proposed models with speed and vertical acceleration as input features.	72
Table 26. Performance comparison of LiRA proposed models with speed and vertical acceleration as input features. (LiRA, 2023)	72
Table 27. Performance comparison of our proposed models with speed and 3D acceleration as input features.	72
Table 28. Performance comparison of LiRA proposed models with speed and 3D acceleration as input features. (LiRA, 2023).....	73
Table 29. Performance comparison of our proposed models with speed and other features as input features.	73
Table 30. Performance comparison of LiRA proposed models with speed and other features as input features. (LiRA, 2023).....	74

Executive Summary

This project investigated the feasibility of using electric vehicle onboard diagnostics (OBD-II) data for pavement quality assessment in Minnesota, adapting methodologies from Denmark's Live Road Assessment (LiRA) project. The research addressed the critical need for cost-effective, continuous pavement monitoring to supplement expensive instrumented vehicle surveys traditionally used by the Minnesota Department of Transportation (MnDOT). The study involved comprehensive OBD-II scanner evaluation, systematic data collection across multiple test routes, adaptation of machine learning algorithms, and development of implementation guidelines for large-scale deployment.

Key Findings

The research successfully demonstrated that vehicle sensor data can provide useful pavement condition information, though with important limitations. Through testing six different OBD-II devices, the OBDLink LX scanner emerged as the optimal choice, successfully collecting 694 Parameter IDs from electric vehicles. Data collection across three test routes, MN-36 loop, Metro to Northfield loop, and MnROAD facility, generated more than 9,000 data points synchronized with MnDOT's high-precision Pathways van measurements.

Machine learning model development to predict pavement roughness (measured by international roughness index (IRI)) revealed that hybrid CNN-LSTM architectures achieved the best performance when using speed combined with vehicle orientation and acceleration features, yielding root mean square error (RMSE) values of 0.172 units IRI compared to MnDOT reference data. However, model performance varied significantly based on environmental factors, data collection conditions, and road types. The controlled MnROAD environment showed different optimal model architectures, with Artificial Neural Networks outperforming CNN-LSTM models due to reduced temporal dependencies in structured testing conditions.

Validation against both MnDOT measurements and NIRA Dynamics network-level data demonstrated correlation across diverse road conditions, though accuracy decreased in areas with rapid pavement condition changes or localized surface defects. The methodology showed stronger performance on highway segments compared to mixed municipal roads, reflecting the challenges of variable urban driving environments.

Implementation Considerations

Large-scale implementation requires careful attention to multiple operational factors. The research identified critical success factors including standardized equipment procurement (OBDLink LX scanners with GPS-enabled tablets), consistent data-collection protocols, and robust data-processing pipelines. Environmental conditions, vehicle maintenance status, and driver behavior significantly impact data quality, necessitating comprehensive training and quality-control procedures.

A three-phase implementation framework was developed, encompassing equipment standardization, data collection infrastructure establishment, and model deployment with validation networks. The framework addresses equipment reliability risks, data processing challenges, and environmental factors that affect sensor accuracy. Personnel requirements include data analysts, field support technicians, and regional coordinators, with scaling patterns that follow predictable geographic and vehicle count parameters.

Limitations and Constraints

The research identified several important limitations that affect practical implementation. Model accuracy depends heavily on having sufficient training data from similar road conditions, vehicle types, and environmental factors. Performance degradation occurs when applying models to completely new road segments without prior calibration data. Weather conditions, vehicle compatibility variations, and GPS accuracy limitations introduce systematic uncertainties that require ongoing monitoring and correction.

The methodology showed reduced effectiveness in detecting localized pavement defects and performed variably across different pavement types. Seasonal variations in pavement response and measurement equipment calibration introduce additional complexity requiring validation protocols and correction factors.

Recommendations

For immediate implementation, the research recommends starting with pilot deployments on major interstate highways where traffic volumes are high and existing MnDOT validation data is available. Continuous model monitoring, automated performance alerts, and periodic retraining protocols are essential for maintaining prediction reliability.

The research establishes that vehicle sensor data offers potential for supplementing traditional pavement management systems, enabling more frequent condition assessments with reduced incremental costs. However, successful implementation requires careful consideration of operational constraints, systematic validation procedures, and realistic expectations regarding accuracy limitations. This approach shows promise for transforming pavement monitoring through distributed data collection, although it should be viewed as complementary to rather than a replacement for existing high-precision measurement systems.

Future research should focus on expanding datasets to include additional road types, seasonal variations, and vehicle configurations while developing more generalized models that can be applied across all road conditions without requiring extensive site-specific calibration. Research priorities should include exploring real-time deployment strategies, integration with connected vehicle technologies, and developing robust algorithms that maintain accuracy across diverse pavement materials, construction standards, and geographic regions. The developed framework provides a foundation for cost-effective infrastructure monitoring that can support proactive maintenance strategies across Minnesota's

transportation network, with the ultimate goal of creating universally applicable models that perform reliably on any road segment regardless of local conditions or historical data availability.

Chapter 1: Introduction

1.1 Project Overview

Today's vehicles are constantly collecting onboard diagnostics data to assess driving conditions and internal component states. This is usually used to monitor and optimize vehicle performance but may soon be useful to monitor pavement quality using novel machine learning techniques. The benefit is that these data are already being collected as vehicle fleets traverse the roadway network as part of their typical daily operations. While this approach is not meant to immediately replace the current measurements, it could provide continual monitoring of the pavement network with very little cost.

Inspired by the recent Live Road Assessment based on modern car sensors (LiRA) project in Denmark funded by the Danish Road Ministry (LiRA, 2023) (Levenberg, et al., 2021), we plan to implement its open-source methods on a MnDOT vehicle (or small fleet of vehicles) to perform a feasibility analysis on conducting pavement monitoring using onboard vehicle data from fleets of vehicles in Minnesota. This will include identifying possible vehicles for use, decoding onboard data, recalibrating the machine learning models for the selected vehicle, and comparing obtained pavement quality estimates to established values from regular MnDOT pavement assessment measurements.

This memorandum outlines the benefits of the project and how they align with MnDOT's strategic priorities and how they will benefit people in Minnesota. This report also outlines how the deliverables that result from this project can be implemented.

1.2 Implementation of Benefits

Findings from this project and this final report can guide MnDOT in understanding the potential and limitations of using onboard vehicle data to estimate pavement quality.

The key deliverables this project will provide are listed below:

1. Literature review on the LiRA project conducted in Denmark, including findings, as well as the findings of other similar research efforts.
2. Software to collect onboard diagnostics (OBD-II) data off US vehicles in the MnDOT fleet and an estimate of pavement quality using the LiRA software.
3. Guidelines for a large-scale implementation of vehicle-based data for pavement monitoring.
4. Final report outlining the findings of this study and how MnDOT can implement these findings.

These deliverables can be used to understand what has been done in other places to leverage large volumes of onboard vehicle data to estimate pavement quality. This will allow MnDOT and local engineers to implement the project benefits by enabling data-driven pavement quality assessment.

Chapter 2: Review of LiRA Project

The foundation for implementing vehicle-based pavement quality assessment in Minnesota rests on understanding existing methodologies and their demonstrated capabilities in similar applications. This chapter is a literature review of the LiRA project, which represents the most comprehensive effort to date in utilizing onboard vehicle sensors for continuous road condition monitoring. The Danish initiative provides critical insights into both the technical feasibility and practical challenges of extracting meaningful pavement quality metrics from standard vehicle sensor data. Additionally, this review surveys related research efforts that have explored various approaches to vehicle-based infrastructure monitoring, establishing the broader context for sensor-driven pavement assessment technologies. The examination of these existing studies reveals the evolution of methodologies from simple accelerometer-based detection systems to sophisticated machine learning frameworks capable of predicting multiple pavement characteristics. Understanding the successes, limitations, and lessons learned from these previous implementations is essential for adapting and optimizing such approaches for Minnesota's specific road network conditions, climate factors, and operational requirements.

2.1 LiRA Project Overview

The LiRA project primarily focused on enhancing the safety, comfort, durability, and energy consumption related to road conditions. The project was initiated in early 2019, with financial support from the Innovation Fund Denmark, and four partners:

- (i) GreenMobility (GM), a car-sharing company operating Zoe car (100% electric) within the greater Copenhagen area
- (ii) Danish Road Directorate (DRD), a public authority responsible for operating, maintaining, and further developing the Danish highway network
- (iii) Sweco Denmark, a consulting company that owns, develops, and operates the current pavement management system for Copenhagen municipality, and
- (iv) the Technical University of Denmark (DTU).

The LiRA project results provide a pathway towards a transformational change in road infrastructure management. It introduced the use of sensors integrated into passenger vehicles to monitor road conditions in real time. The main components of this innovative approach are shown in Figure 1.

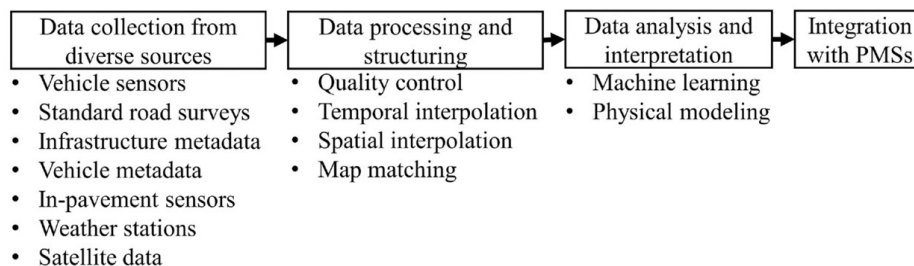


Figure 1. Envisioned components for live road condition assessment with internal vehicle sensors (Levenberg, et al., 2021).

2.2 Data Collection

The LiRA project's core strategy is centered on controlled data acquisition efforts using standard survey methodologies and GreenMobility vehicles that traveled specific road sections and lanes. These sections covered approximately 350 kilometers of roadways, both motorways and municipal roads. The selection of these road segments was deliberate, aiming to capture a wide array of pavement distress and defects. This choice considered the variation in physical conditions, as well as in travel patterns.

This project utilized data from two main categories of data sources. The first category is data from an Internet of Things (IoT) device called AutoPi, which is a significantly enhanced OBD-II device. OBD-II stands for On-Board Diagnostics version II, which is a standard implemented in most modern vehicles for monitoring sensor data. LiRA project has relied upon GreenMobility cars as its primary data collection platform. These vehicles are equipped with the AutoPi Dongle, which not only captures accelerometer data, but also collects various other vehicle parameters.

The second main category of data used in this project is from conventional pavement management instrumented vehicles. Specifically, data was gathered as described below:

- Profilometer (P79): This piece of equipment, equipped with 25-point lasers, measures road profiles in both longitudinal and transverse directions. It provides comprehensive data on cross-fall slope, vertical curvature, International Roughness Index (IRI), rut depth, and Mean Profile Depth (MPD).
- ViaFrik: Measures the friction coefficient of pavements on wet surfaces.
- CPX: Measures tire/road noise, using a pair of microphones positioned 200mm from each test tire.
- ARAN9000: Serves as a multi-functional road-testing equipment, and collects data on roughness, texture, Right of Way (RoW) images, and performs laser crack measurements.

The collected data within the LiRA project is categorized into three distinct areas:

- road functional properties,
- road distress or damage, and
- road condition indices.

By leveraging both IoT-connected vehicle data as well as data from conventional pavement management equipment, researchers were able to aggregate a robust transportation dataset covering numerous aspects of vehicle performance and location. The dual data sources allowed for a more comprehensive set of driving and infrastructure analytics. The vehicle's computers carefully collected data points and synchronized them with the onboard computers, which connected directly to the vehicle's Controller Area Network or CAN bus. This real-time data was then transmitted to a cloud-based system, where it was carefully stored in a structured database for further analysis.

2.3 LiRA Data Processing and Structuring

Data processing and structuring in LiRA included quality control, spatio-temporal interpolation, map matching, and data structuring. The LiRA project didn't simply collect and analyze data; it also sought to

structure and model the data effectively. This is achieved through the implementation of a three-tier data model: conceptual, logical, and physical. These stages progressively refined and structured the data, moving from a raw conceptual stage to a logically structured model and, finally, a physically structured data storage. The architecture incorporated two main categories of data sources: data from IoT devices such as AutoPi and OBD-II, and data from conventional pavement condition measurement vehicles, including P79, ARAN9000, CPX, ViaFrik, road progressive distance, and GPS coordinates. The architectural framework was designed with modularity in mind, incorporating microservices to ensure scalability and adaptability.

Processing the vast and diverse data entailed numerous challenges. The data had to be aligned, cleaned, and subjected to rigorous analysis. This is critical in the context of Machine Learning (ML), for which structured data becomes essential for training and deploying models.

The challenges in processing this vast and diverse dataset included:

- (i) Data Variety: Managing diverse data sources required thoughtful structuring to ensure compatibility and coherence, addressing challenges related to different data formats and structures,
- (ii) Data Volume: The substantial volume of data collected necessitated efficient storage solutions and retrieval mechanisms,
- (iii) Data Velocity: The LiRA project deals with real-time data streams, and ML models can benefit from a continuous influx of data. Handling the velocity of data ensures that ML models are trained on the latest information, enhancing their predictive capabilities.
- (iv) Relationship Establishment: Establishing meaningful relationships between new data and existing datasets is essential for both ML and maintaining a coherent dataset across distributed data collection. This involves accurate mapping, alignment, and cleaning of data to derive valuable insights.

By addressing these challenges, LiRA ensured that the structured data is not only suitable for ML applications but also supports the distributed nature of data collection, allowing seamless integration of information from various GreenMobility vehicles.

2.4 LiRA Data Analysis and Interpretation

Data analysis within the LiRA project encompassed a multi-pronged approach. It begins with data cleaning and alignment, ensuring that the collected data is accurately structured and synchronized. The project employed ML techniques, specifically linear models, such as Ridge, Lasso, and elastic net for regression, as well as logistic regression for classification modeling. Furthermore, physical models are developed to simulate the response of in-vehicle sensors to various road conditions. These models delve into longitudinal dynamics (including braking and acceleration), transverse dynamics (involving lateral vehicle movements), and vertical dynamics, such as the suspension's response to road surface irregularities. The heart of the project lay in its innovative use of ML, particularly hybrid models that synergize mechanics-based physical models with data-driven models. Various models, with a focus on supervised learning, were used to detect distress types, such as potholes, alligator cracking, roughness,

and skid resistance. The ability to detect these distress types depends on the availability of in-vehicle sensor data.

The project's framework comprised data preprocessing, dataset preparation, and deep learning, with convolutional neural networks (CNN) being a primary model for pattern recognition. CNN models were applied to identify distress types and severities in 10m road segments. The efficacy of distress-type detection was contingent on the availability of in-vehicle sensor data, highlighting the importance of a robust dataset. The ML flow chart is shown in Figure 2.

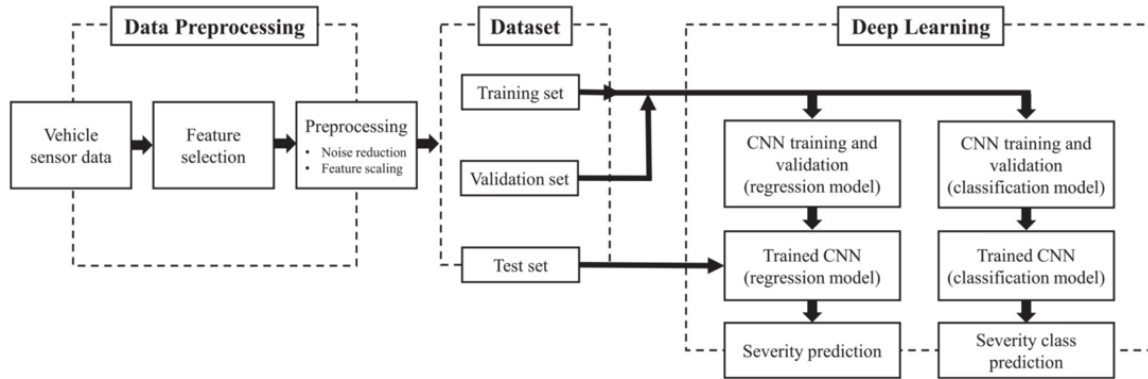


Figure 2. Machine learning flow chart (Levenberg, et al., 2021).

2.5 Modeling

2.5.1 Prediction of International Roughness Index (IRI)

The investigation into modeling the IRI included preprocessing, feature engineering, and machine-learning modeling. Preprocessing involved map-matching of reference measurements with in-vehicle sensor recordings. The feature engineering employed analyzing a set of data that was valid for training using Sequential Feature Selection (SFS), while the test dataset was refined to include only the selected features. This phase was important in laying the groundwork for subsequent machine-learning applications. The models were then applied to predict road conditions, with separate outcomes for motorways, municipalities, and a combined analysis.

For motorways, the evaluation encompassed the assessment of vertical acceleration and speed data using various ML models. These models, ranging from Baseline to Support Vector Regression (SVR) and Artificial Neural Networks (ANN), presented their unique strengths and limitations. The application of 3D acceleration and speed sensors demonstrated a slight improvement in the model performance, while the inclusion of additional sensors resulted in a deterioration of model performance, as shown in Table 1, emphasizing the importance of feature selection.

Table 1. Comparison of Different sensor inputs for the ML models, where the best is used for motorways.

Input sensor	MAE	RMSE	MRE	R2	Best model
Vertical Acceleration – Speed	0.18	0.26	0.14	0.59	Elastic Net (SFS)
3D Acceleration – Speed	0.17	0.25	0.12	0.62	SVR (PCA)
3D Acceleration – Speed - Additional	0.18	0.23	0.14	0.54	Random Forest (PCA)

The comparison of the best models' performance for different input sensor combinations using the combined motorway and municipality road data collection is summarized in Table 2. The best performance in terms of the R^2 was obtained using the vertical acceleration and speed sensors as inputs, while the 3D acceleration led to a better performance in terms of the mean relative error. The R^2 was more impacted by large outliers hence the model with the lower R^2 but also a low mean relative error showed a better average performance but more outliers.

Table 2. Comparison of Different sensor inputs for the ML models, where the best is used for the combined motorway and municipality.

Input sensors	MAE	RMSE	MRE	R^2	Best Model
Vertical Acceleration - Speed	0.57	0.76	0.23	0.75	Random Forest (SFS)
3D Acceleration – Speed	0.15	0.2	0.15	0.69	SVR (SFS)

2.5.2 Estimating Unevenness

The proposed methodology integrates measured vertical acceleration with a mechanical quarter-car model, facilitated by a proportional-integral-derivative (PID) controller. The matching process showcases a near-perfect correlation between calculated and measured acceleration traces (Skar and Levenberg, 2024).

2.5.3 Estimating Friction

To estimate the tire-road friction coefficient (TRFC) three distinct methods were used:

- A data-driven approach based on in-vehicle vibration sensors.
- A data-driven approach based on in-vehicle wheel rotation measurements.
- A machine-learning approach based on car sensors, images and a 2D vehicle dynamic model.

First, an exploratory test campaign was first conducted with the ViaFrik device, which measured skid resistance at a constant speed of 60 km/h with a relative sliding of 20%. A GreenMobility Renault Zoe followed the ViaFrik truck and in-vehicle accelerations, GPS data and some images were collected.

Using the first method, an estimation of skid resistance from vehicle vibrations was performed. When a car's tires, with a non-zero toe angle, are side-slid on the pavement, they produce horizontal vibrations. These vibrations can be theoretically linked to road skid resistance. In-depth analysis of ViaFrik data involved meticulous screening to retain readings within a narrow slip ratio range of 19.9% to 20.1%. Employing a first-order smoothing algorithm, with a bandwidth of 0.05, enhanced data reliability by eliminating variability in reported skid resistance values. Exploration into in-vehicle acceleration dynamics included preprocessing raw readings from a GreenMobility car. Sampling at 236 Hz, these readings underwent resampling, ensuring uniformity at 250 Hz, and corrections for yaw, pitch, and roll rotations to achieve standardized orientations. The model calibration phase applied a discrete Fourier transform to 256 readings, providing a spectral analysis in the 0 to 125 Hz frequency range. Logarithmic conversion to dB (after normalization the average) enabled comparison with ViaFrik skid resistance, revealing initial correlations between spectral magnitudes and road conditions. Preliminary results affirmed the potential of horizontal accelerations, showing meaningful correlations with skid resistance at different frequencies. This suggested further exploration in understanding road surface conditions through in-vehicle vibrations.

Skid resistance estimation based on in-vehicle wheel rotation measurements was conducted next. The friction coefficient (μ) is linked to $\alpha(V)$, a function incorporating speed, as well as the front-left and rear-left wheel radius (R_{FL} and R_{RL} , respectively). A constant β_{FL} interacts with RPM, and $f(\mu)$ represents a positive, unitless function of the tire-friction coefficient.

The analytical process revealed the resampling of car RPM data to a standardized 50 Hz frequency, alongside the resampling of left friction values from the ViaFrik system. Temporal alignment of car and ViaFrik data was based on speed profiles or GPS information and was followed by a smoothing procedure applied to both RPM and ViaFrik datasets. Assuming specific values for wheel radius ($R_{RL} = 0.31$ and $R_{FL} = 0.31$), the pavement friction coefficient $f(\mu)$ was calculated using the provided formula. The resulting friction coefficient was further transformed using an exponential function and logarithmic conversion, introducing the parameter β_0 in the process (Skar and Levenberg, 2024).

$$f(\mu) = \frac{\beta_{FL} R_{FL}}{R_{FL} RPM_{FL} - \alpha(V)}$$

$$f(\mu) = \text{Exp}(\mu^{\beta_0}) - 1$$

$$\mu = (\ln(f(\mu) + 1))^{\frac{1}{\beta_0}}$$

The calculated R^2 value of 0.0535 indicates the degree of correlation between the estimated friction (μ) derived from RPM analysis and the actual values obtained from ViaFrik. The optimal parameters, β_{FL} and β_0 , emerge as 4 and 1.2, respectively, showcasing the efficacy of this method in determining friction coefficients from wheel rotation data.

The last method consists of a novel approach that integrates car signals and images, harnessing the power of fused vehicle dynamics and machine vision information. Employing machine learning techniques, this Skid Resistance Estimation Method stands as a robust tool for measuring road conditions. A comprehensive field test campaign was executed, gathering standard TRFC values, vehicle dynamics, and road surface images to construct a meticulously labeled dataset.

Experiments were conducted to analyze the contribution of input data to overall performance and revealed that fused features were capable of adapting to a wide variety of driving conditions. The overarching methodology, illustrated in Figure 3, is divided into two branches for vehicle dynamics filtering and feature pre-extraction, with inputs from GoPro cameras and Autopi devices. Model-based vehicle state estimation becomes pivotal, incorporating signals such as longitudinal, lateral, and vertical accelerations, vehicle velocity, and engine torque.

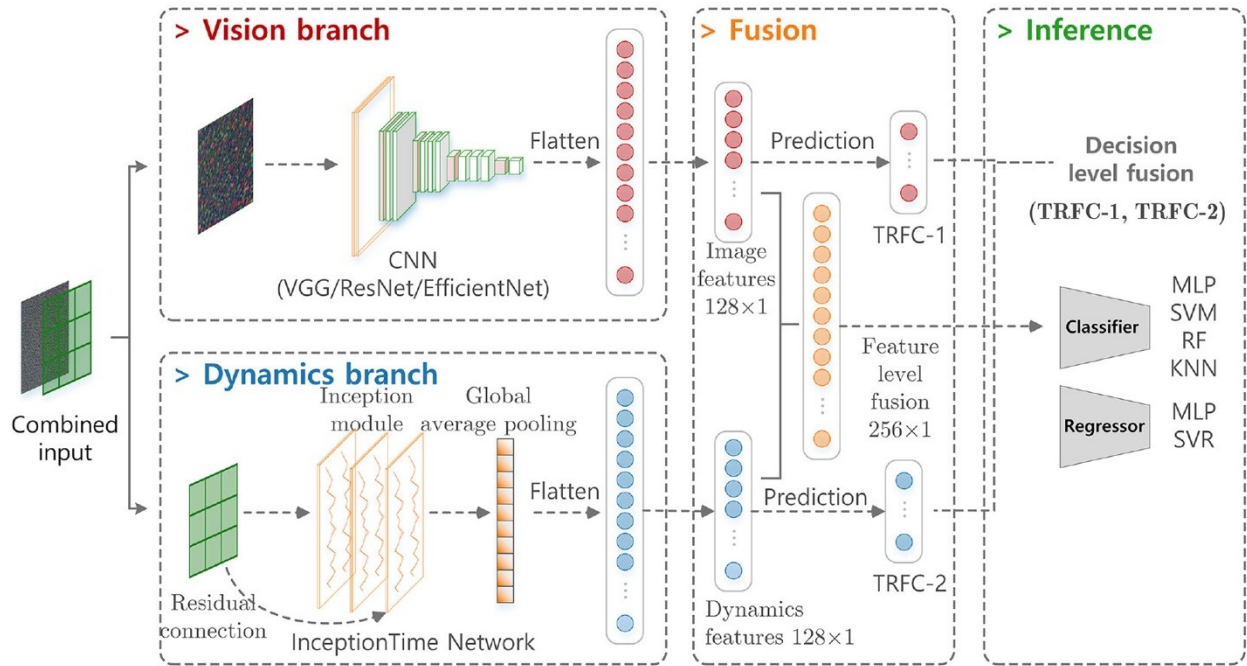


Figure 3. Schematic of the multi-feature fusion network for TRFC estimation (Du, et al., 2023).

Critical to the process was the calculation of the slip ratio, derived from wheel angle velocity, effective wheel radius, and longitudinal velocity, with slip dependency deemed crucial for maximum TRFC estimation. The integration of Vehicle Dynamics and Normalization employed a vertical load transfer model to compute normal forces on the front axle, ensuring input data is normalized to a uniform scale, avoiding local optima.

Table 3 catalogs the vehicle dynamics data used for training, providing a comprehensive overview of the corresponding ranges within the dataset. The road surface image processing strategy involved extracting video frames at regular intervals, segmenting them into regions of interest (ROIs), and strategically selecting ROIs to eliminate the influence of water film. The resulting grayscale ROI images were stacked

into layers, forming a singular input for training CNN models. This combination of methodologies established a comprehensive framework for accurate and scalable TRFC estimation.

Table 3. Vehicle dynamics features are used for training the proposed network (Du, et al., 2023).

Features	Unit	Range
Transversal Acceleration a_y	m/s ²	[-1, 1]
Vertical acceleration a_z	m/s ²	[-2, 3]
Yaw rate $\dot{\Phi}$	rad/s	[-0.3, 0.3]
Normal force f_{zf}	N	[7000, 8000]
Speed v	m/s	[15, 18]
Engine torque T	N.m	[180, 230]
Slip ratio λ	-	[-3, 3]
TRFC μ_{used}	-	[-0.1, 0.1]

2.5.4 Classification and Regression Results

2.5.4.1 Classification results

At the end of the significant effort in skid resistance estimation, the comprehensive Classification Results and Regression findings were discussed. The TRFC values were classified into three distinct levels, as outlined in Table 4.

Table 4. Threshold value of TRFC levels for the classification task (Du, et al., 2023)

TRFC level	Lower limit	Upper limit	Section number	Sample number
L1	-	0.6	980	8909
L2	0.6	0.7	998	9697
L3	0.7	-	120	1067

This approach systematically organized the dataset, which comprised a total of 2098 sections. Each section was segmented and labeled into three classes. and the data was partitioned based on a five-fold cross-validation allocation strategy.

Recognizing the challenges posed by imbalanced datasets, a class weight was calculated based on sample numbers, ensuring a judicious approach to model training. the implementation of the model was done within the Python-based deep-learning frameworks. The Stochastic Gradient Descent (SGD) optimizer, with an initial learning rate of 0.001 and a momentum of 0.98, was used in the optimization process.

2.5.4.2 Regression Results

The subsequent exploration, based on regression methods, used a transformation from a 3-unit layer to a single unit layer, to provide insights through metrics such as Mean Absolute Error (MAE), Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE), and R-square. The method had the highest R-square value of 0.62.

The results for the proposed network on TRFC level classification showed a notable F1 score increase ranging from 0.09 to 0.18. The proposed model not only achieved the highest R-square at 0.71 but also exhibited the lowest relative error within 20%, establishing its capabilities in distinguishing maximum TRFC levels effectively. This robust synthesis of findings underscored the efficacy of the proposed model in advancing the understanding and precision of tire-road friction coefficient estimation.

2.6 LiRA Software Architecture

An overview of the LiRA software architecture is presented here. From a high-level perspective, this architecture was developed considering three functionality groups:

- (i) LiRA Data Warehouse, covering activities to support data collection from diverse sources, data processing and structuring, and data preparation for subsequent analysis and interpretation, for example, initial training and periodic re-training of machine learning algorithms. Here users are also able to inspect the data and gain an overview of what is available; they are also provided with means of exporting any data needed for them to be used in external systems. The data format was also used for storing the inferred road conditions derived from the car sensors.
- (ii) LiRA ML, covering the application of machine learning to vehicle data and generation of road conditions. The derived live road conditions were added to the LiRA Data Warehouse.
- (iii) LiRA Map, covering the visualization of inferred road conditions as part of the integration effort with PMSs.

2.7 LiRA Project's Goals

The project goal was to establish an operative road assessment system that relies on sensors within a homogeneous car fleet. Additionally, a LiRA map demo, designed to provide real-time data, was another key objective.

The project involved the development of reliable algorithms and models to calculate road metrics based on data collected from car sensors. Notably, the goal was to ensure that at least six out of the ten identified road metrics achieved an accuracy rate exceeding 80%.

The researchers also dedicated significant time and effort to creating hardware configurations and set-up procedures that could be applied to other vehicles, all supported by a validated calibration process. Additionally, comprehensive guidelines were produced to aid others in the development of their own Live Road Assessment systems.

2.8 Challenges and Limitations of LiRA Project

The LiRA project focused on city roads, which present a unique set of difficulties, primarily due to their narrowness and the towering facades that line them. Achieving accurate GPS positioning was a challenging task, which was crucial for the project's success.

Moreover, both friction and noise measurements are normally performed at constant speeds, given that the measurements are speed-sensitive. In a city environment, it was practically impossible to operate these devices at a constant speed. Therefore, the actual operating speed, which is variable, was also saved in the data warehouse. These challenges represented critical aspects that the LiRA project had to navigate and overcome.

In addition, some important metadata was missing or inaccessible, for example, exact models of sensors installed in a car, their installed locations within the car, and mechanical parameters of the suspension system. This missing information further added to the challenge of developing and applying mechanics-based models for sensor data interpretation.

Currently, there is no standard protocol for translating CAN data into physical measures for all types of cars. Collaboration with the manufacturer or extensive reverse engineering was required to address this issue. Optimization of IoT devices and dual-core processing were needed to log data at the same frequency and resolution as the CAN bus offers.

2.9 Related Research Effort

EVs have created a paradigm shift in both passenger and freight transport. While their growing adoption offers many benefits, there has been a general lack of planning and preparedness regarding the impact of this new fleet on pavement assets. Mattinzioli (Mattinzioli, et al., 2023) proposed five key aspects regarding the effect of EVs and upcoming mobility trends on pavement degradation: a) increase in EV traffic, b) change in vehicle forces on pavement, c) influence of charging systems on pavements, d) change in load concentration due to autonomous transit, and e) funding implications from EV adoption. IoT technology is leveraged to connect electric vehicles, platforms, and individuals. This connectivity allows for real-time monitoring through PCs or smartphones, offering enhanced convenience and reliability. The integrated terminal modules and advanced GPS technology contribute to small size, low power consumption, high efficiency, and the accuracy of monitoring road assessment, enhancing the system's overall effectiveness (Liu, et al., 2014).

Purdue University's NIRA project, "Extraction of Vehicle CAN Bus Data for Roadway Condition Monitoring," (Li, et al., 2020) represents a significant research effort in the pursuit of obtaining timely information for monitoring road conditions and traffic operations throughout the state roadway

network. This ambitious initiative addressed a critical challenge in winter roadway maintenance—determining emerging or deteriorating conditions before they escalate into major accidents. The project capitalized on the advanced capabilities of modern vehicles, equipped with accelerometers, anti-lock brake (ABS), and traction control systems, to extract valuable data from the CAN. This data, when harmonized with GPS coordinates and cellular connectivity, provides a real-time, on-the-ground sampling of vehicle dynamics during adverse weather conditions. As the transportation sector transitions into an era where vehicle data can reshape agency management systems, the conventional reliance on fixed infrastructure sensors and telephone reports is being replaced. This transformative potential extends to the optimization of roadway weather information systems (RWIS) based on vehicle data.

The methodology employed in the project involved collaboration with industry partners, rigorous testing of vehicles from three manufacturers, and equipment testing. A comprehensive system for receiving, transmitting, and processing CAN data was piloted using a vehicle interface equipped with an embedded computer, cellular networks, and a back-office data system. The research findings underscore the successful piloting of an integrated CAN Bus road monitoring system, facilitating the real-time identification of roadway hazards during winter events. The team integrated CAN Bus data elements from nine categories into existing dashboard interfaces, effectively identifying winter roads and hazardous conditions. Notably, braking and wheel counter data emerged as particularly useful for providing early indications of winter conditions. This research lays a solid foundation for the future integration of vehicle-generated data into proactive roadway management strategies, fostering safer and more efficient transportation systems (Li, et al., 2020).

Several studies have made significant contributions to the exploration of utilizing sensors for real-time road assessment and monitoring. Ghose et al. (Ghose, et al., 2012) presented a pioneering demonstration of a smartphone-based road conditioning monitoring and alert system, marking a pivotal step in the integration of smartphones as part of road infrastructure assessment. Focusing on the detection of road surface anomalies, Seraj, F. (Seraj, et al., 2015) proposed a system equipped with inertial sensors, including accelerometers and gyroscopes. This system employed sophisticated techniques, such as stationary wavelet transform analysis, and an envelope technique to mitigate the effects of speed, slope, and drift from sensor signals.

Recent research, exemplified by Lima et al. (Lima, et al., 2016), marks a shift in road infrastructure assessment using an application. Crafted for real-time data recording and crowdsourced evaluations, it aligns with user-driven data collection trends in road monitoring. Souza et al. (Souza, et al., 2018) contributed with an app, using machine learning to assess road conditions, emphasizing the role of smart algorithms. Similarly, Xue et al. (Xue, et al., 2016) introduced an application for pothole profiling, demonstrating practical viability with low error rates. Collectively, these studies integrate tech-driven solutions, enhancing the accuracy of pavement assessments.

Furthermore, the landscape of road assessment methodologies has expanded significantly with the incorporation of sensors, as demonstrated by studies conducted by Agostinacchio et al. (Agostinacchio, et al., 2014) and Zhao et al. (Zhao, et al., 2015). These investigations represent an important shift in the

arsenal of techniques employed for road condition evaluation, particularly through the utilization of smartphone sensors to capture vehicle-induced vibrations and linear accelerations. These sensor-based approaches add a new dimension to road assessment, offering a diverse set of tools to analyze and understand the intricacies of road conditions.

Seraj et al. (Seraj, et al., 2017) explored crowd-sensing for predictive maintenance in the realm of ground transport infrastructures. Their study demonstrated the efficacy of utilizing sensor measurements, even with lower quality, coupled with smart sensor analytics to achieve results comparable to high-end measurement devices and vehicles. The introduction of RoVi, a framework for real-time monitoring of roads, railroads, and bike paths, stands as a pivotal aspect of their work.

In parallel, Stanie's work in 2021 introduces The Road Condition Tool (RCT), presenting an innovative and non-intrusive approach to road pavement condition assessment. RCT leverages real-time data from vehicles to offer insights into road conditions, traffic volume, and geolocation information. The functional validation of RCT emphasizes its effectiveness in evaluating road pavement conditions, showcasing its potential to revolutionize road infrastructure management. In the dynamic landscape of transportation infrastructure, the RCT serves as a compelling example, illustrating how technology can enhance the efficiency and accuracy of road pavement assessments while minimizing disruptions to traffic (Staniek, 2021).

Cafiso et al. (Cafiso, et al., 2022), in their study conducted in 2022, made significant contributions to the assessment of pavement conditions through innovative methods. They focused on examining the utility of specific Key Performance Indicators (KPIs) derived from data collected by devices mounted on bikes and e-scooters serving as probe vehicles. Specifically, they investigated the root mean square (RMS) of vertical acceleration and a weighted RMS (RMSw) based on ISO 2631–1 as potential KPIs. These metrics were explored for their effectiveness in evaluating pavement conditions, providing insights into distress features and severity rankings. The consistent and reliable results obtained from the application of these KPIs contributed valuable information to the expanding knowledge base concerning smartphone-based road assessment methods.

The integration of machine learning in road condition assessment has been a focal point of research, as evidenced by the work of Souza et al. (Souza, et al., 2018) and Allouch et al. (Allouch, et al., 2017). Souza et al. introduced an application, which harnesses the power of machine learning algorithms to evaluate road conditions and identify pavement defects. This innovative approach underscores the potential of smart algorithms in enhancing pavement assessment methodologies. In a parallel endeavor, Allouch et al. delved into an in-depth analysis of various classification algorithms. The comparative assessment contributed valuable insights into the selection of algorithms for efficient and accurate detection of road anomalies using smartphone sensor data. Together, these studies exemplify the ongoing efforts to leverage machine learning and classification algorithms to advance the field of road condition assessment, emphasizing the significance of smart technologies in paving the way for more effective and precise evaluation methodologies.

In a study conducted by Sholevar (Sholevar, et al., 2022), a thorough examination of machine learning techniques was performed, focusing on their potential role in pavement management systems (PMS). The review sheds light on various ML applications, including image classification, object detection, and segmentation, emphasizing their superiority in extracting rules and features from extensive datasets. In contrast to traditional methods, ML exhibits enhanced performance and generalization capabilities. The comprehensive analysis presented in the review encompasses state-of-practice data collection methods and pavement condition indices. It synthesizes insights from over 80 relevant studies on ML applications in pavement monitoring and distress detection. Notably, the paper addresses the limitations associated with the detection of complex distress patterns and the indication of severity, underscoring ML's transformative potential in optimizing road asset management. Employing a meticulous keyword combination search in major scholarly databases, the review effectively fills a critical gap in the literature by providing an in-depth exploration of ML's role in pavement condition evaluation and automated pavement monitoring. Overall, this research contributes valuable insights that can significantly influence the advancement of ML applications for enhanced pavement management strategies.

Other research efforts have been conducted on ML applications for pavement assessment. Bashar et al. (Bashar and Torres-Machi, 2021), examined the performance of ML algorithms in predicting pavement performance, focusing on the International Roughness Index. Analyzing 20 relevant studies, the findings reveal that ML algorithms, particularly random forests (RF), outperform traditional approaches by capturing 15.6% more variability in IRI. RF proves to be the most accurate ML technique with a performance score of 0.995, followed by ANN and support vector machines (SVM). The study recommends ANN for its consistent accuracy across various sample sizes and RF to prevent overfitting or achieve high accuracy. The study suggests a minimum sample size of 50 times the number of input variables for ML-based pavement performance modeling. Another paper by Abdelaziz (Abdelaziz, et al., 2020) explores the development of a predictive model for the IRI in flexible pavements, utilizing both Multiple Linear Regression and ANNs based on the Long-Term Pavement Performance database. Key findings include the effectiveness of an Asphalt Concrete overlay in significantly reducing IRI, while treatments like crack sealing and various patching methods show no immediate impact. The relationship between pavement age and IRI is generally linear, with some sections exhibiting an exponential trend. Despite the absence of initial IRI measurements for GPS sections in the Long-Term Pavement Performance database, the study's methodology for estimating these values aligns well with those measured for other sections, emphasizing the importance of rational assumptions and engineering judgment in handling such data. Influential parameters on flexible pavement roughness include pavement age, initial IRI, standard deviation of rut depth, and the presence of transverse and alligator fatigue cracks. Additionally, an ANN model, with one input layer, three hidden layers, and one output layer, achieves superior performance with an overall R^2 value of 0.75. This underscores the enhanced predictive capabilities of the neural network approach in flexible pavement IRI modeling.

Other authors investigated predicting other pavement features, such as friction. Three estimators were proposed by Feng et al. (Feng, et al., 2020), to enhance the accuracy of road surface friction coefficient estimation for four-wheel drive electric vehicles. The first method indirectly estimates the friction

coefficient using the utilization friction coefficient, showing an error of about 0.1 with significant fluctuations. The second method transforms the equation of the High-Speed Runway Friction Tester (HSRI) and achieves an error within 0.02 for most instances. The third estimator combines the strengths of the first two methods through optimal design, providing precise friction coefficient estimation that meets control requirements. Another study (Du, et al., 2023), introduced a data-driven approach to estimate the maximum TRFC based on vehicle operation signals and road surface images. A multi-feature fusion neural network was proposed, demonstrating superior performance in TRFC inference compared to models based solely on dynamics or vision data.

In summary, the existing research demonstrates feasible approaches to leverage sensors and data native to electric vehicles to evaluate and monitor road conditions. Methods range from simple uses of acceleration to integrated machine-learning predictions.

This literature review presented compelling evidence in support of assessing road conditions using in-vehicle sensor data, and introduced the LiRA project as a proof-of-concept implementation effort. The findings and challenges from the LiRA project serve as valuable insights for future endeavors in utilizing in-vehicle sensors for road condition assessment. This approach has the potential to revolutionize pavement management systems, providing real-time data and enabling more informed decisions in road maintenance and repair efforts. As technology and data analysis methods continue to advance, the future of road condition assessment holds promises for more efficient and effective infrastructure maintenance.

Chapter 3: Adapt OBD-II scanners for the new car(s)

Successful implementation of vehicle-based pavement quality assessment requires careful selection and validation of data collection hardware capable of reliably capturing the necessary sensor parameters from modern electric vehicles. This chapter documents the comprehensive evaluation process undertaken to identify optimal OBD-II scanning equipment for Minnesota's specific operational requirements. The assessment involved testing six different commercial OBD-II devices across multiple vehicle platforms and road conditions to determine their compatibility with electric vehicles, data collection capabilities, and practical usability constraints. The evaluation criteria emphasized the ability to capture the 694 Parameter IDs identified as essential for pavement quality prediction models, with particular attention to acceleration measurements across three axes, vehicle speed data, and GPS positioning accuracy. This systematic testing approach was necessary to ensure that selected hardware would provide consistent, high-quality data collection across diverse Minnesota road networks while maintaining compatibility with the state's emerging electric vehicle fleet infrastructure.

3.1 Finding Compatible OBD-II Scanners

In line with the focus of our project, we identified the most effective OBD-II scanners specifically designed for compatibility with Electric Vehicles (EVs). We ordered six different OBD-II scanners, shown in Figure 4, selected based on extensive online research and positive user experiences. These scanners were specifically chosen for their compatibility with EVs and their supported features, addressing our critical need in the project. We also acquired a high-capacity tablet, essential for managing and storing the data generated during diagnostics and capable of collecting accurate GPS data. Each scanner arrived with distinct features and operational protocols, necessitating a detailed exploration of their individual capabilities and applications. After each scanner was received, we performed multiple rounds of testing across various cars and roads to assess performance comprehensively.



Figure 4. Ordered OBD-II scanners.

Among the scanners, the OBDLink LX stood out due to its superior performance. It successfully collected data from 695 different Parameter IDs (PIDs), confirming its compatibility with EVs. Below, a summary highlighting the key features, advantages, and disadvantages of each scanner based on our evaluations has been provided in Table 5.

Table 5. Summary of OBD-II scanners.

OBD scanner	Important Features for project	Advantage	Disadvantage
Panlong (For both android and iOS)	-	-	Did not collect data. User-unfriendly. Requires a pro account for the Torque app.
OBDLink LX	694 PIDs, including features such as acceleration in 3 axes (x, y, z), speed, longitude, latitude, and 688 more.	Collects almost all needed features, user-friendly, compatible with all cars, easy to extract and store data. Great application (OBDLink).	Does not collect wheel speed data.
VEEPEAK OBDCheck	Average speed, distance travelled, OBD module voltage, vehicle acceleration, vehicle speed.	Compatible and easy to use.	Limited to five features; insufficient for our goals. Collected data was inaccurate and messy.
Vgate iCar Pro	-	-	Did not collect data.
AutoPi Telematics Unit	-	-	Difficult to operate; non-responsive support. Practical functionality was lacking.
KONNWEI KW902	Average speed, distance travelled, OBD module voltage, vehicle acceleration, vehicle speed.	Compatible and easy to use.	Only collects five features; insufficient for our goals. Data was inaccurate and messy.

Following the initial data collection phase, we decided to proceed with the OBDLink LX scanner as our primary device. With the data collected from this scanner, all the required features necessary for estimating the International Roughness Index (IRI) were available. Despite its numerous benefits, the OBDLink LX fell short in one critical area—it was unable to capture wheel RPM, a key metric essential for

estimating friction as specified by the LIRA project. However, direct measurement of CAN data may be able to collect the necessary RPM data.

Table 6. Description of CAN signals utilized in the LiRA project.

S_AutoPi	Use
ACC_LONG	Estimating inertial energy and friction
ACC_TRANS	Estimating friction
ACC_YAW	Aligning data across devices
BRK_TRQ_ELEC	Calculating braking energy
F_DIST ODO	Validating distance measurements
RPM_(FR, FL, RR, RL)	Estimating friction
SB_REM	Estimating number of passengers
SPD_VEH	All prediction models
TRAC_CONS	Calculating energy consumption
TRIP_CONS	Validating 'TRAC_CONS' and 'WHL_TRQ_EST'
WHL_PRS_(FR, FL, RR, RL)	Improving energy consumption
WHL_TRQ_EST	Calculating energy consumption
WW_F	Calc Indicating if surface is wet or dry

In our efforts to address the gaps identified compared to the LIRA project, specifically the lack of wheel RPM data, we initially chose to use the AutoPi Telematics Unit to capture more accurate data, including wheel speed. Despite extensive research, multiple testing sessions, and direct engagement with the

manufacturer, we were unable to effectively utilize this scanner due to inadequate support. After integrating the specific PIDs for wheel speeds into the Torque Pro app, we conducted tests across various vehicles and discovered that these custom PIDs functioned effectively only with Toyota vehicles.

Subsequently, the collection of wheel speed data was explored using alternative methods. Our investigations led us to the Torque app, where the Pro version offered the capability to add custom PIDs. The custom PIDs have been shown in Table 7.

Table 7. Custom PIDs for wheels speed in the Torque app.

Name	Shrot Name	Mode and PID	Equation	Min Value	Max Value	Units	Header
FR Wheel Speed	FR WS	2103	$A * 32 / 25 * 15625 / 25146$	0	120	mph	7B0
FL Wheel Speed	FL WS	2103	$B * 32 / 25 * 15625 / 25146$	0	120	mph	7B0
RR Wheel Speed	RR WS	2103	$C * 32 / 25 * 15625 / 25146$	0	120	mph	7B0
RL Wheel Speed	RL WS	2103	$D * 32 / 25 * 15625 / 25146$	0	120	mph	7B0

3.2 Data Collection

In addition to our initial trial to identify the optimal scanner, five major data collections were conducted for testing scanners, each detailed as follows:

3.2.1 Conventional Volkswagen vehicle

Before implementing our scanner on Electric Vehicles, we carried out a preliminary test to verify its operational functionality and feature collection capabilities. This initial trial entailed a brief trip around the UMN campus using a conventional Volkswagen vehicle. However, the data collection was performed using a tablet lacking GPS functionality, resulting in incomplete data capture.

3.2.2 MnROAD Data Collection Efforts

On December 4th, 2023 we conducted two rounds of data collection at Minnesota Department of Transportation roads (MnROAD) using a 2022 Chevy Bolt electric vehicle to ensure precision and enable comparative analysis. We tested three different devices during this process. The OBDLink LX proved effective, offering a range of important features. However, we faced significant challenges in exporting data from the AutoPi scanner. Additionally, the Veepeak OBD-II scanner yielded limited results,

providing only two features—Acceleration and Speed—which did not fully meet our project's requirements.



Figure 5. Research team conducting data collection at MnROAD.

3.2.3 Toyota Data Collection

Our second major data collection focused on testing custom PIDs for wheel speed, specifically with Toyota vehicles. On April 1st, 2024, we conducted these tests using custom PIDs on a Toyota, successfully capturing data on wheel speeds.

3.2.4 Highway 55 Data Collection

Utilizing the dataset provided by MnDOT from Hwy 55, we conducted drives in both directions along the highway to collect comprehensive data throughout the data collection. Subsequently, we extracted this data and made efforts to align our collected information with the existing MnDOT dataset.

3.2.5 Highway 36 Data Collection

The section of Highway 55 that was repaved last summer is currently in excellent condition, which diminishes its relevance for investigating IRI. Consequently, we opted to collect data from another section that also has corresponding MnDOT data available for comparison. This time, our data collection focused on a loop that includes Highway 36. To ensure accuracy and enable effective comparison, we drove this loop twice.

3.3 Data Collection Results

The following sections present key plots from each data collection phase to evaluate scanner performance.

3.3.1 Volkswagen Data Collection Descriptive Results

As previously mentioned, in this data collection, the tablet didn't have GPS functionality, so the most important extracted plots from data collection were the speed during the data collection and also acceleration across three axes of X, Y, and Z. This information is crucial for all models predicting IRI and friction as outlined in Table 6. Analyzing the plots, the speed fluctuates frequently in Figure 6; these changes indicate frequent accelerations and decelerations, likely reflecting urban driving patterns. In Figure 7, all three acceleration components exhibit significant fluctuations, with particularly sharp spikes in the X and Z components. These could indicate abrupt maneuvers or responses to uneven road surfaces. The sharp peak in the X-axis around the 0-second mark might suggest a sudden acceleration event or an abrupt start. The variability in acceleration reduces slightly across all axes, but still, persistent fluctuations suggest ongoing dynamic changes in movement.

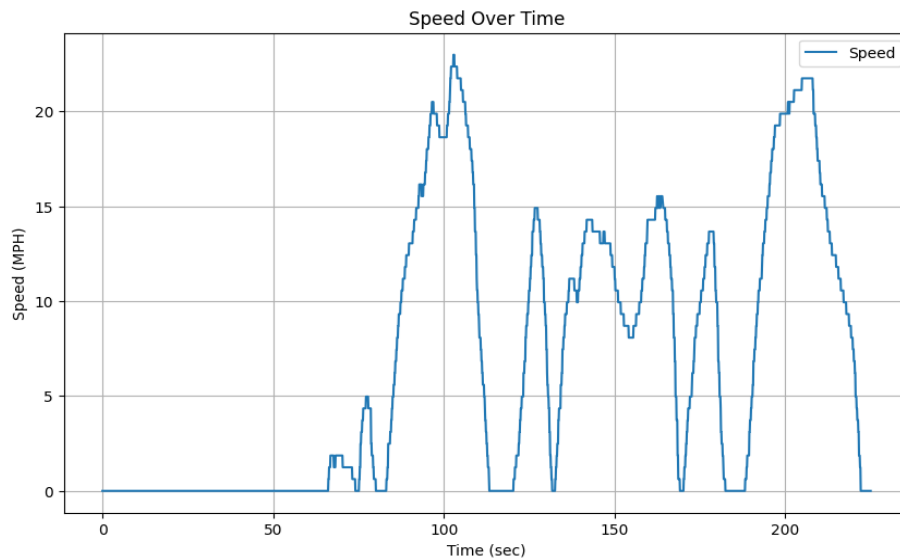


Figure 6. Speed (mph) over time (sec) for Volkswagen data collection.

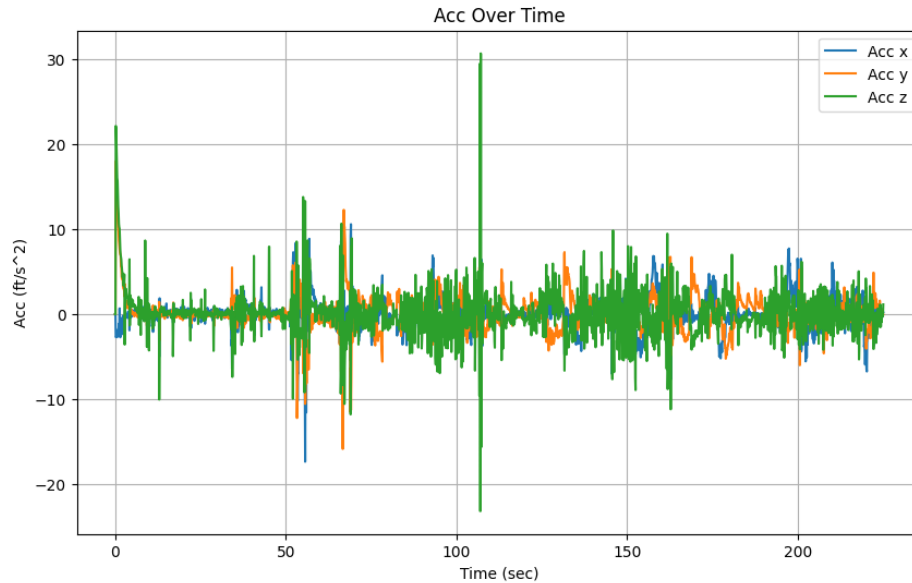


Figure 7. Acceleration (ft/s²) over time (sec) for Volkswagen data collection.

3.3.2 MnROAD Data Collection Descriptive Results

In Figure 8, the blue line shows consistent activity around zero with frequent but moderate fluctuations, indicating continuous adjustments in speed. Spikes seen in Figure 9 suggest moments of significant acceleration or deceleration. The Green line is much less variable than the X and Z axes but does show some variability. This indicates some lateral movement, possibly due to lane changes or navigating curves. The Red line exhibits the most variability and the highest spikes, confirming significant interaction with vertical forces. This is typical for vertical dynamics as they respond to road imperfections and indicate that the vehicle experienced various bumps or dips in the road.

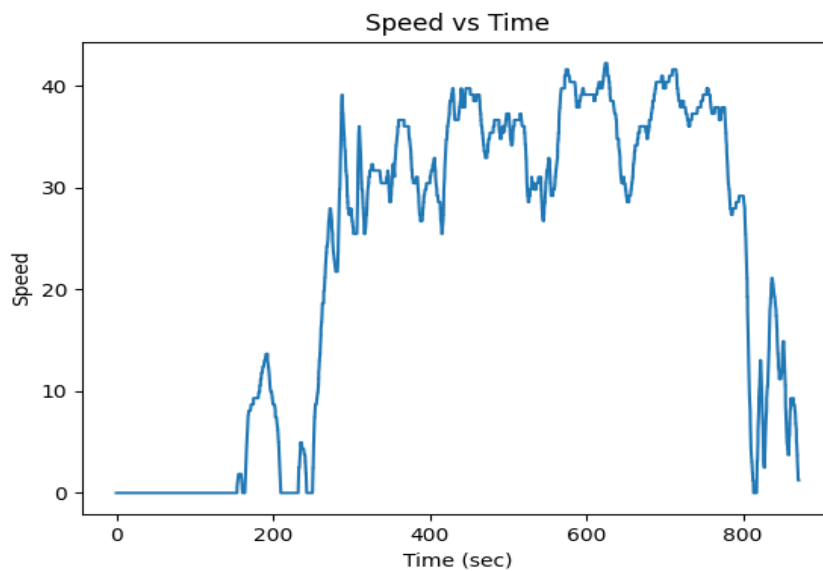


Figure 8. Speed (mph) over time (sec) for MnROAD data collection.

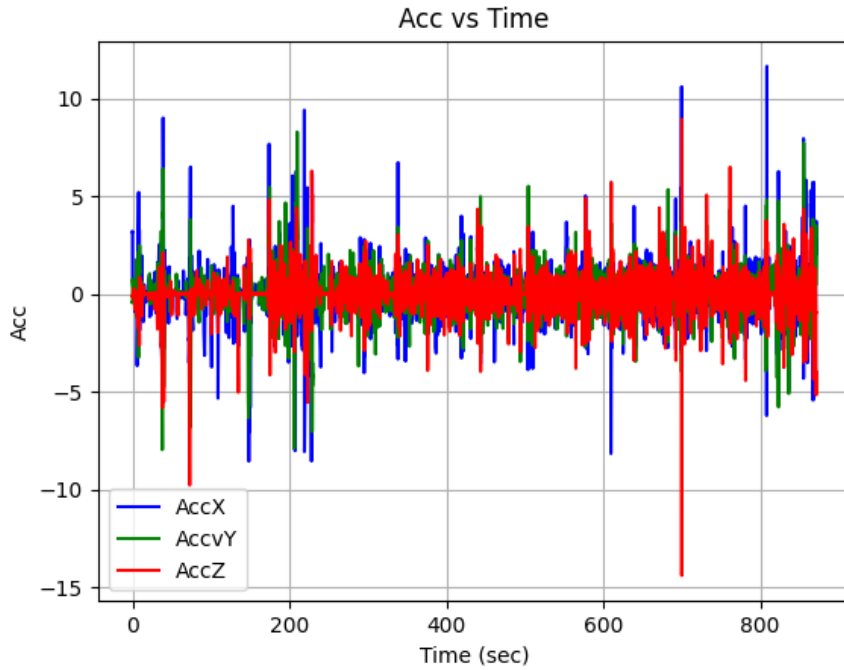


Figure 9. Acceleration (ft/s²) over time (sec) for MnROAD data collection.

In MnROAD data collection, longitude and latitude data were successfully gathered at each timestamp, enabling us to generate a detailed GPS trace of the road. Additionally, the road map color-coded by speed and acceleration in the Z-axis has been visualized in Figures 10, 11, and 12, respectively. The Z component of acceleration is particularly vital as it reveals vertical movements, such as bumps or dips, which reflect changes in IRI. This data is essential for tracking and visualizing these features and estimating IRI and road friction.

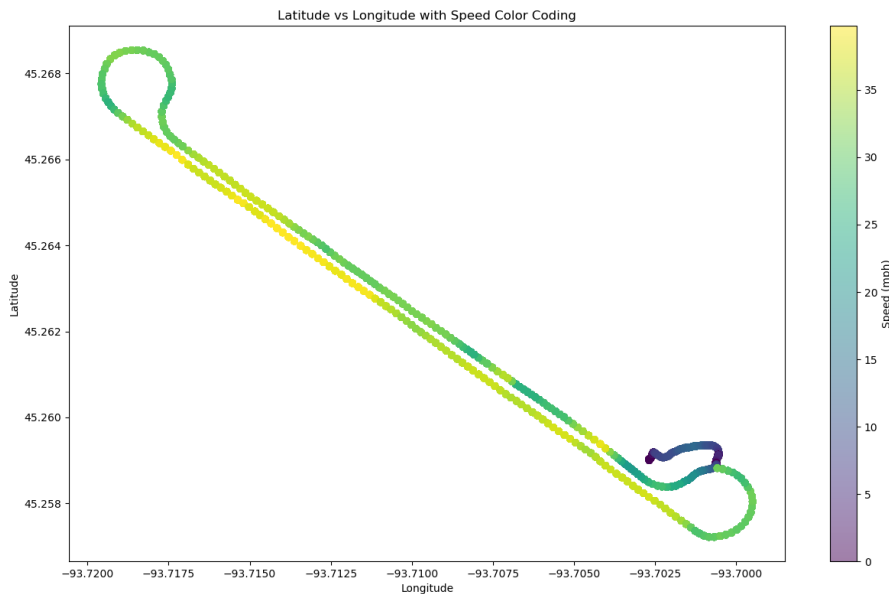


Figure 10. Speed color-coded GPS trace for MnROAD data collection drive.

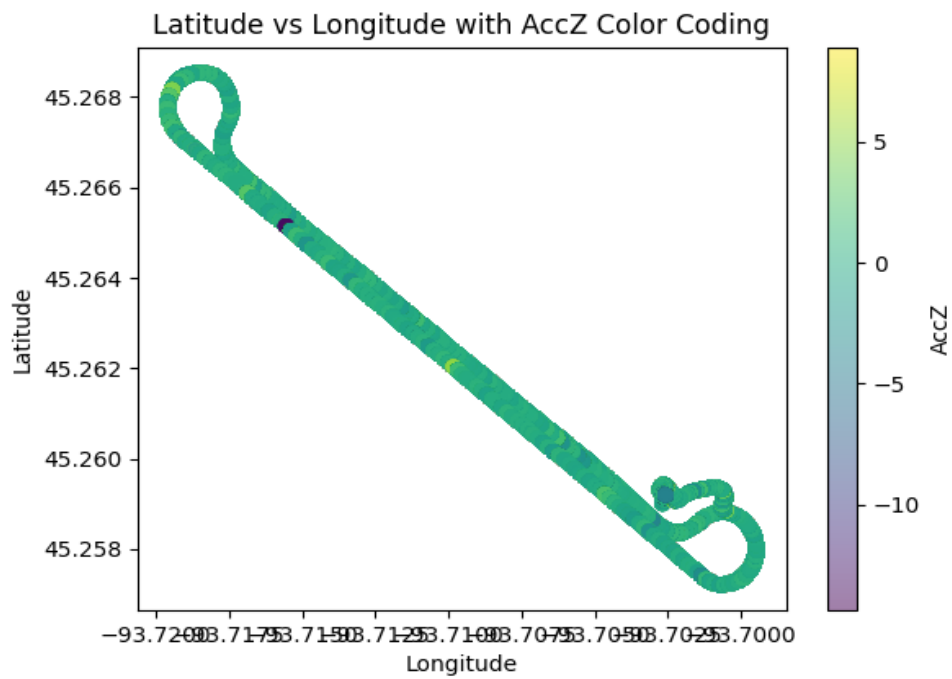


Figure 11. Z component of Acceleration color-coded GPS trace for MnROAD data collection drive.

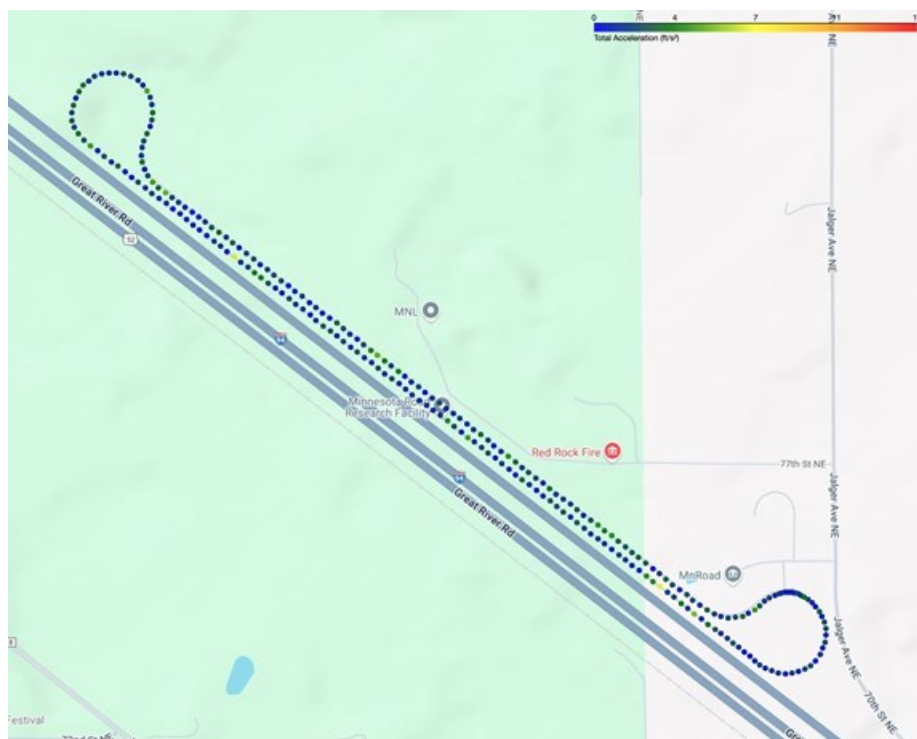


Figure 12. Color-coded map of total acceleration for MnROAD data collection.

3.3.3 Toyota Data Collection Descriptive Results

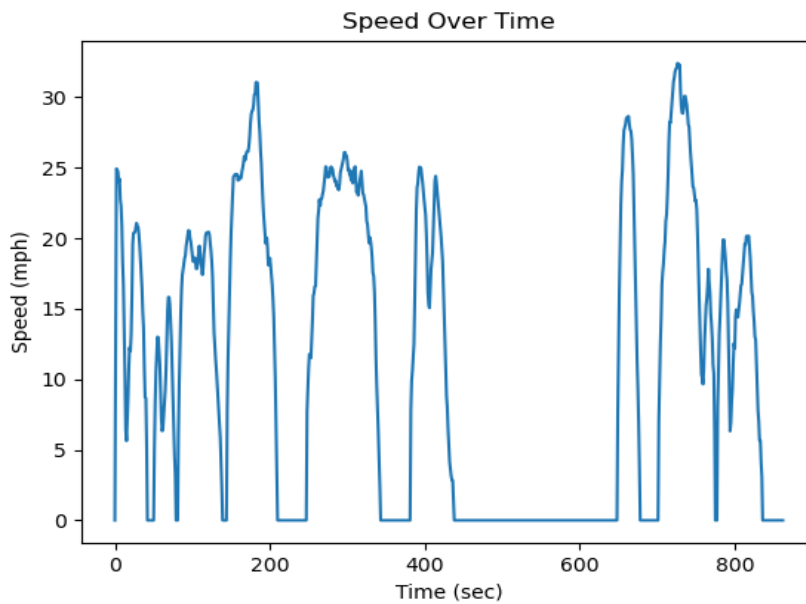


Figure 13. Speed (mph) over time (sec) for Toyota data collection.

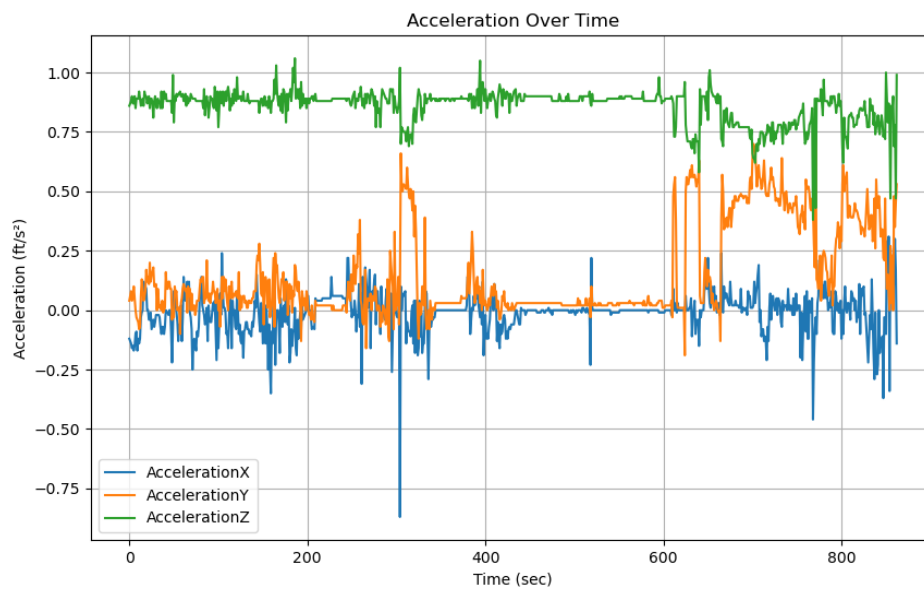


Figure 14. Acceleration (ft/s²) over time (sec) for Toyota data collection.

Based on Figure 13, speed varies significantly throughout the Toyota data collection drive, with several peaks above 20 mph and sharp declines where the speed drops to near zero, indicating potential stops or significant slowdowns, which is reasonable since this data collection was done around the UMN campus, which was crowded. In Figure 14, the data are relatively stable with minor fluctuations around

zero, which could reflect acceleration and braking dynamics. For the Y component, the fluctuations observed here are indicative of movements from side to side. For the Z Component, this line indicates vertical acceleration, which is relatively flat. This suggests that the driving surface was mostly smooth, with occasional bumps or dips as indicated by the minor peaks.

Data for the vehicle wheel speeds were successfully collected. The plots in Figure 15 illustrate a highly synchronized pattern between the rear right and left wheels, reflecting nearly identical speeds. Typically, during a turn, the outer wheels must travel a greater distance than the inner wheels, necessitating a faster rotation rate to cover the larger radius. However, the observed difference in wheel speeds during such maneuvers was minor—approximately 0.79 mph— which is too slight to be distinctly noted on the plots. The collected wheel speed data lacks the desired level of accuracy. We are currently investigating alternative techniques to enhance the accuracy of our data collection for future studies.

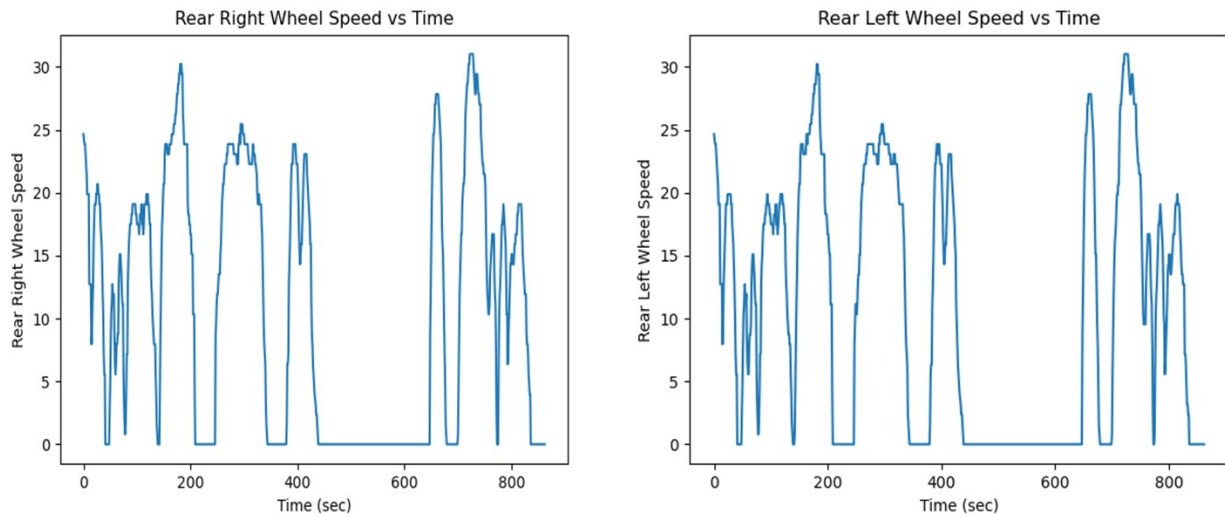


Figure 15. RR and RL Wheels speed over time (sec) for Toyota data collection.

The color-coded maps from our data collection, conducted on the crowded roads surrounding the campus, clearly illustrate the variability in speed and Z component of acceleration in Figures 16 and 17. This variation is particularly noticeable due to the diverse driving conditions encountered in these areas.

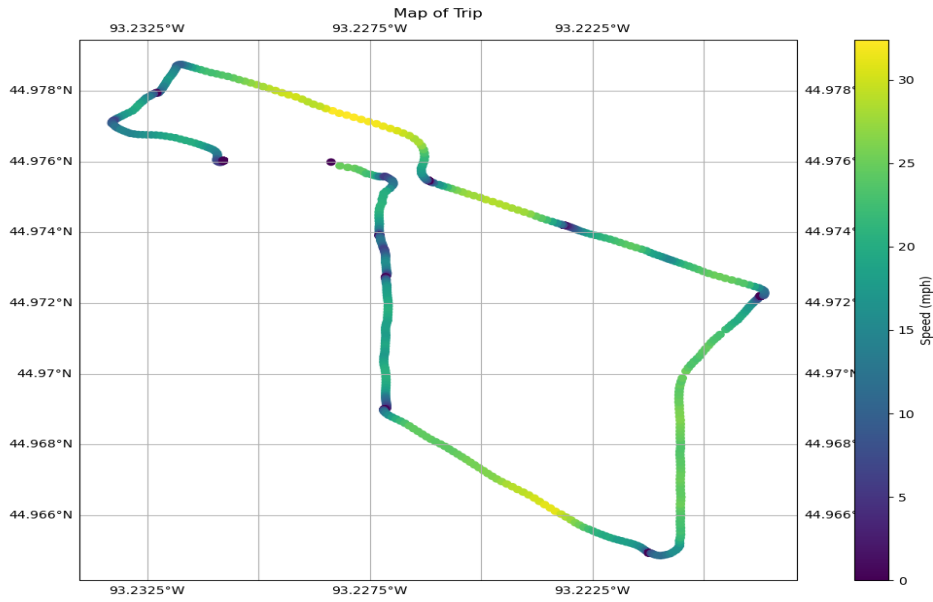


Figure 16. Speed color-coded GPS trace for Toyota data collection drive.

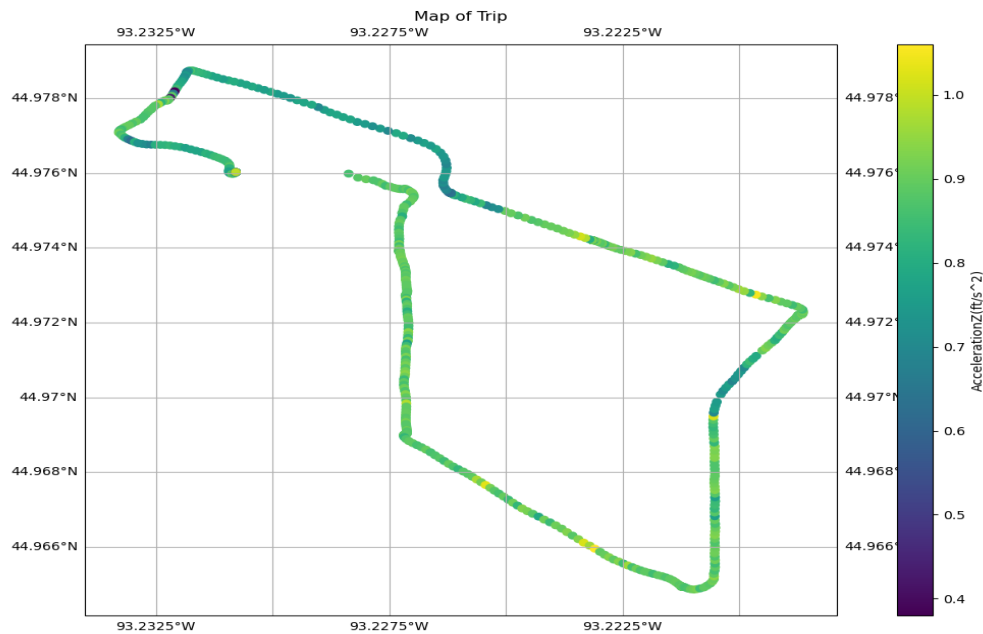


Figure 17. Z component of acceleration color-coded GPS trace for Toyota data collection drive.

3.3.4 Highway 55 Data Collection Descriptive Results

Speeds show substantial variability, with multiple sharp declines, suggesting frequent changes in velocity in Figure 18. These changes could be due to traffic conditions, stops, or varying speed limits. The peak speed reaches around 50 mph, indicating highway driving, while the lower speeds suggest residential areas. Acceleration vs Time in Figure 19 shows considerable variability; the X component, in particular,

shows the highest spikes, which are particularly notable around the 500 and 2500 time marks, suggesting dramatic stops, or speed changes. Despite the variability, the Z component maintains a tighter band of fluctuations compared to the X and Y components. This suggests that while the road may have had irregularities, the vertical impacts were not as severe as the lateral or longitudinal movements. There are points in the graph where all three components show simultaneous spikes (e.g., around time 2500). These moments may correspond to significant driving events such as taking a sharp curve on an uneven road surface.

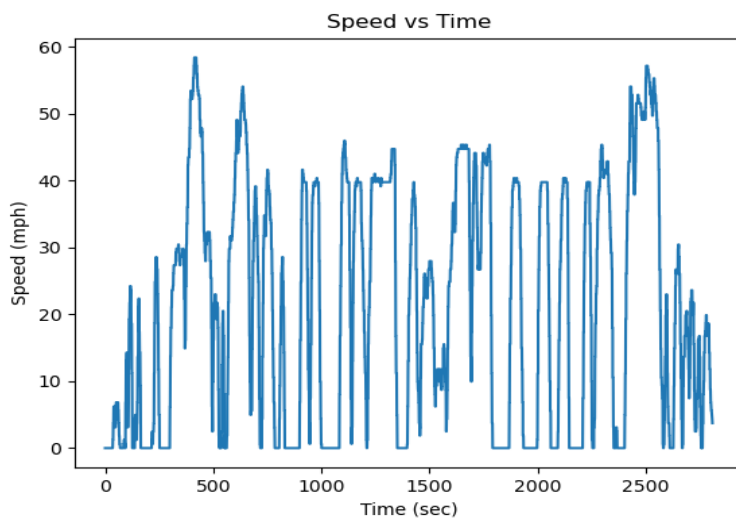


Figure 18. Speed (mph) over time (sec) for Highway 55 data collection.

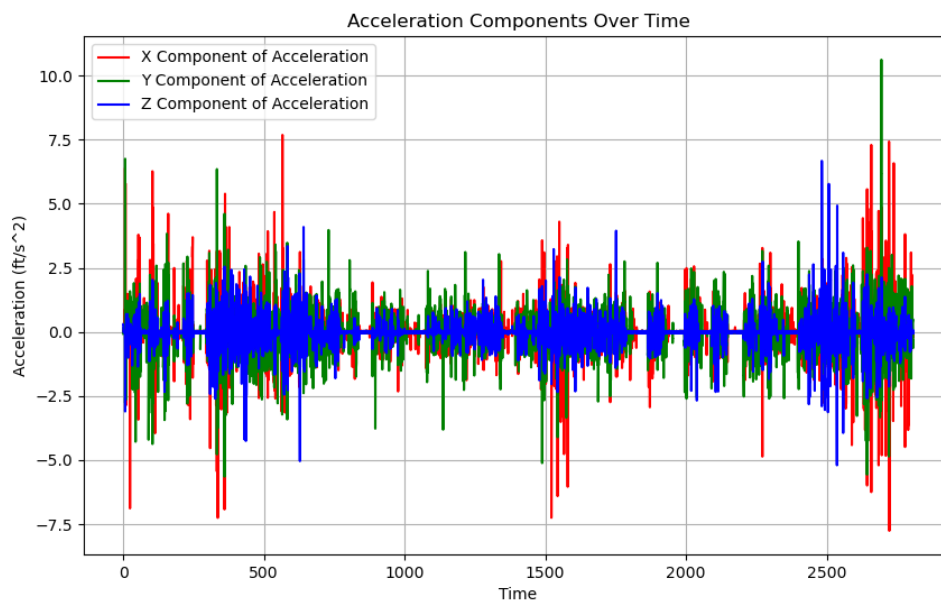


Figure 19. Acceleration (ft/s²) over time (sec) for Highway 55 data collection.

In the vertical acceleration color-coded plot, the first section shows varied colors from purple to green, indicating a change in vertical accelerations. This might suggest a series of bumps or a particularly uneven section of the road where the vehicle experienced both upward and downward forces. The color in the middle section is consistently closer to the middle of the scale, primarily displaying a light blue or teal color. This suggests relatively smoother conditions with minor undulations compared to the starting section. Similar to the starting section, the end shows more variability in colors with patches of purple and green. This area likely had more complex road features, such as potholes, speed bumps, or uneven terrain, causing more noticeable vertical movements in the vehicle.

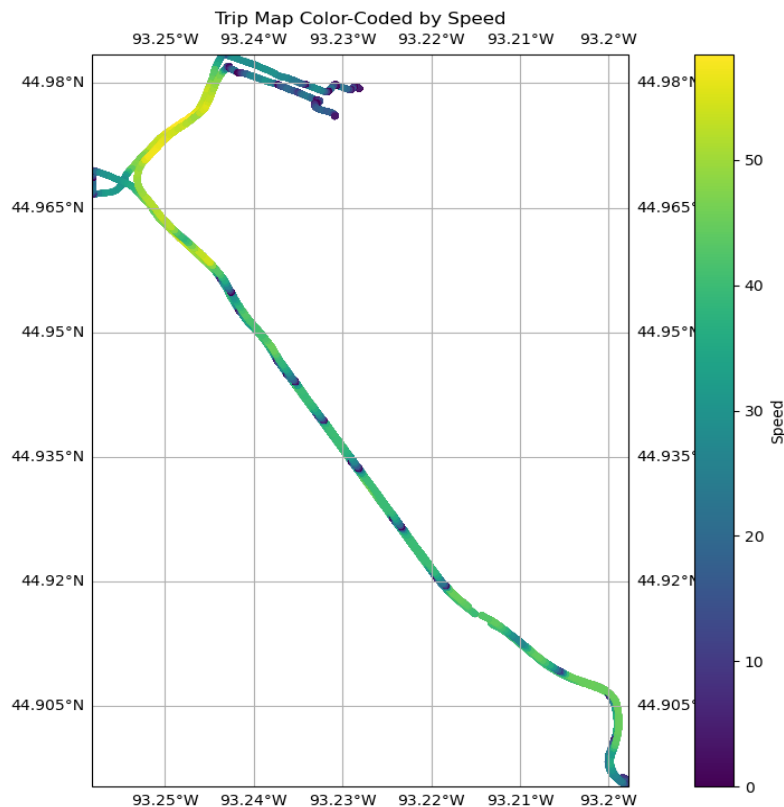


Figure 20. Speed color-coded GPS trace for Highway 55 data collection drive.

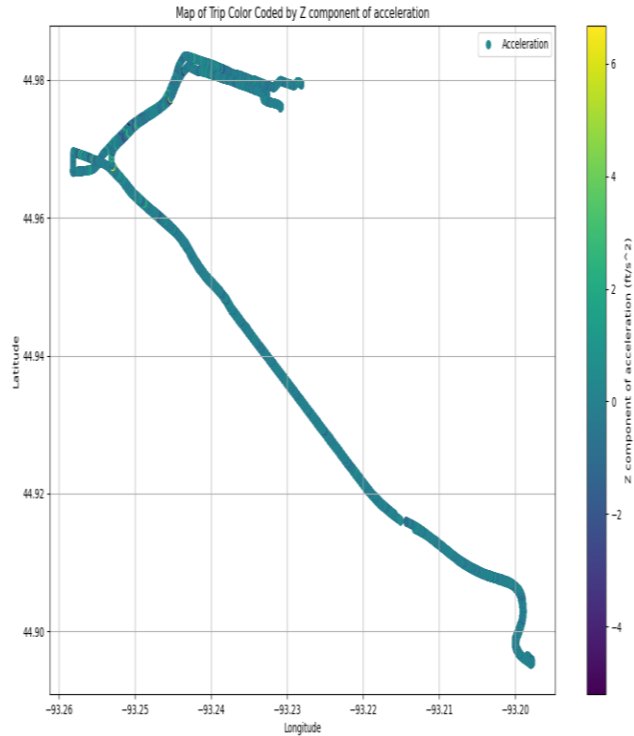


Figure 21. Z component of acceleration color-coded GPS trace for Highway 55 data collection drive.

3.3.5 Highway 36 Data Collection Descriptive Results

The speed plot shows frequent fluctuations in speed, with several peaks reaching up to 60 mph, suggesting periods of rapid acceleration followed by abrupt decelerations. The pattern indicates multiple stops and starts, or potentially varying traffic conditions. The deepest dips to zero suggest complete stops.

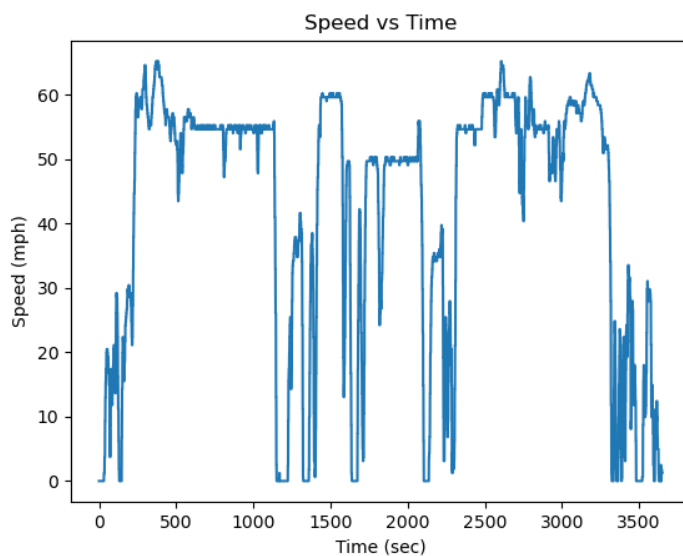


Figure 22. Speed (mph) over time (sec) for Highway 36 data collection.

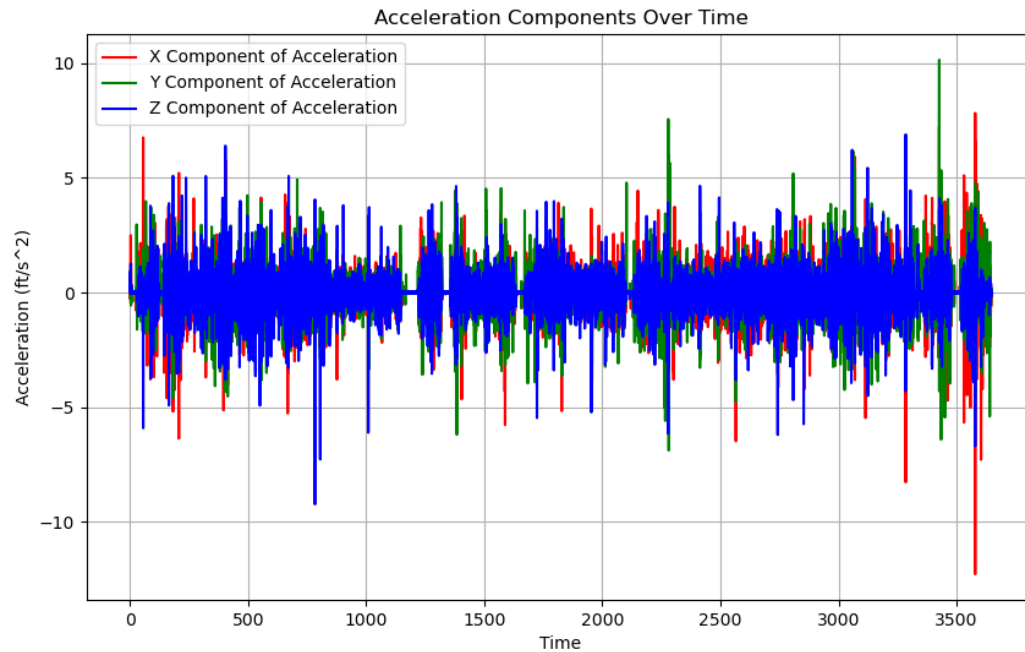


Figure 23. Acceleration (ft/s²) over time (sec) for Highway 36 data collection.

The X component of acceleration shows significant fluctuations and peaks throughout the time series. The spikes in the red graph may indicate sudden changes in speed or direction along the X-axis, which could be due to actions like turns, abrupt starts, or stops. The Y component also displays variability, although the peaks are not as extreme as the X component. This suggests moderate changes in the acceleration along the Y-axis, like lane-changing of the vehicle. The Z component is the most stable among the three, with fewer peaks and a more consistent range. This is typical since the Z-axis acceleration would mainly reflect vertical movements, such as bumps or dips on the road surface.

The considered loop, which contains MN-36 and I-694 was traversed twice, each time at varying speeds, to facilitate a comparative analysis and ensure the collection of more accurate data. Below, the findings from these loops are specifically detailed.

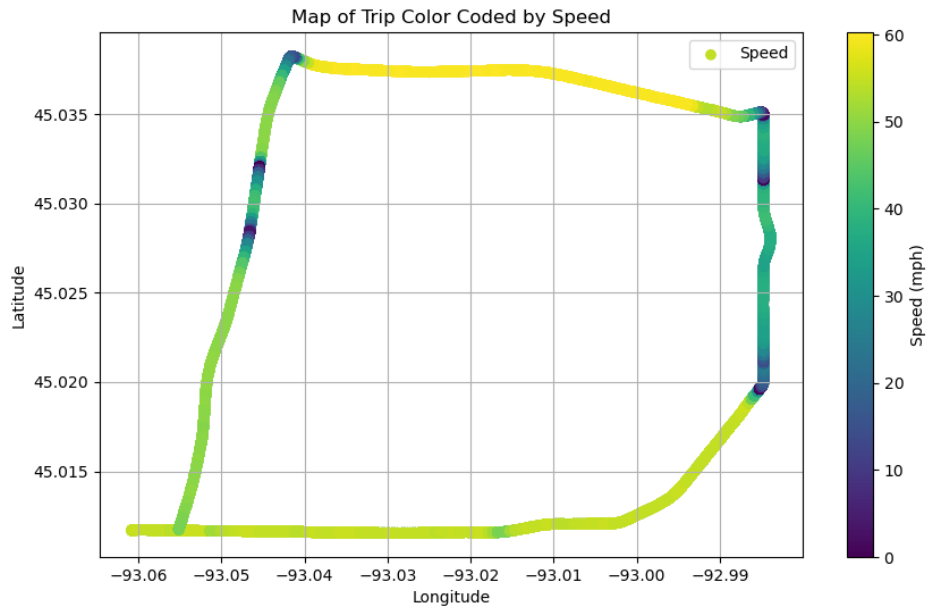


Figure 24. Speed color-coded GPS trace of the first loop for Highway 36 data collection drive.

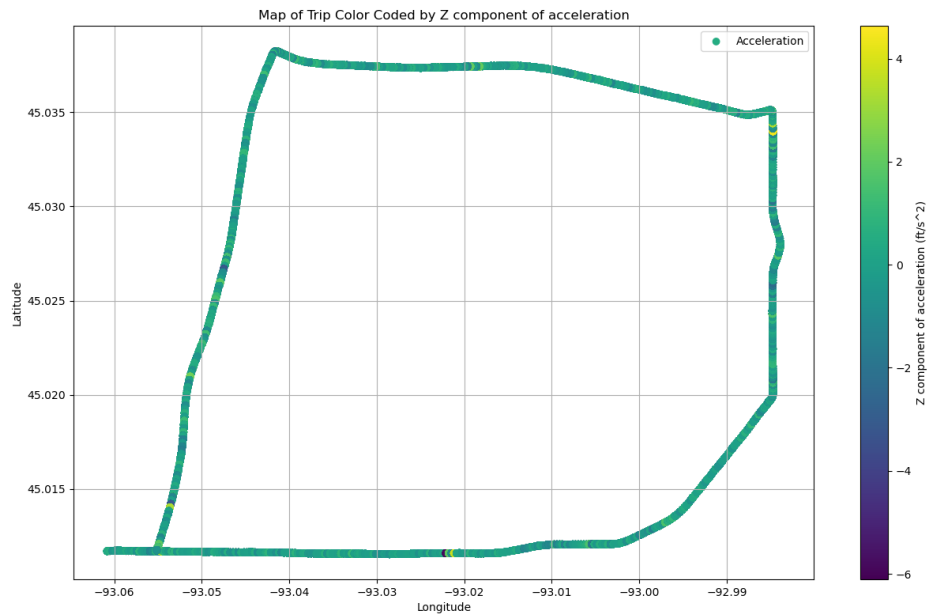


Figure 25. Z component of acceleration color-coded GPS trace of the first loop for Highway 36 data collection drive.

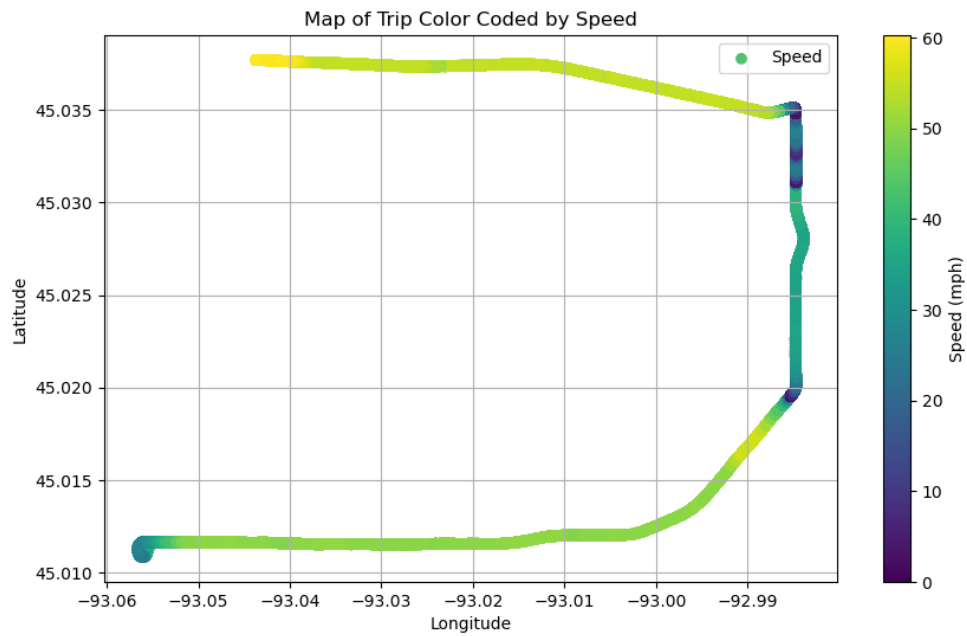


Figure 26. Speed color-coded GPS trace of the second loop for Highway 36 data collection drive.

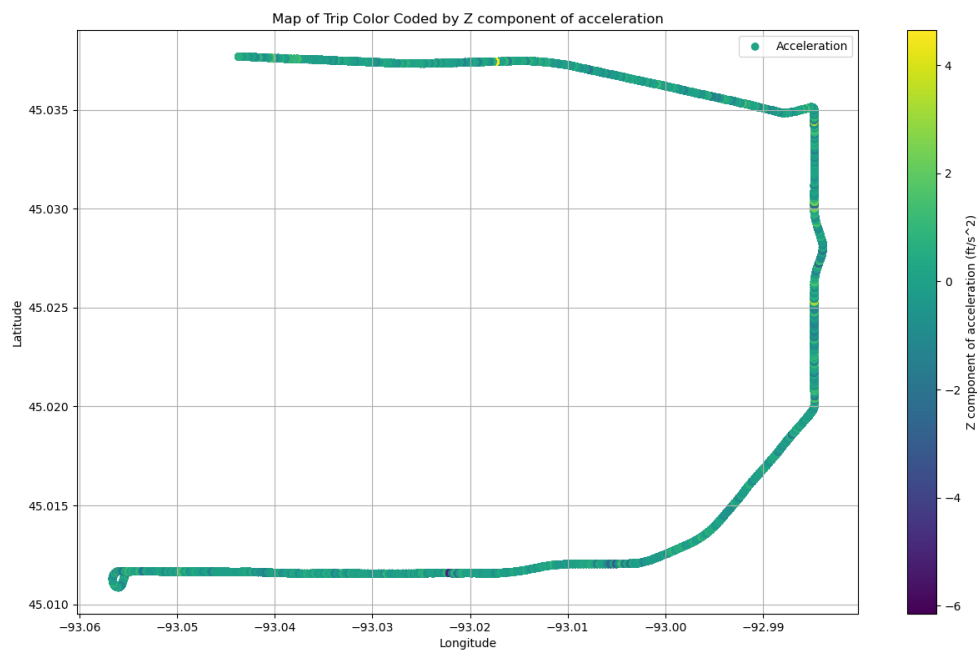


Figure 27. Z component of acceleration color-coded GPS trace of the first loop for Highway 36 data collection drive.

By comparing Figures 25 and 27, which utilize a color gradient from purple to yellow to depict speed variations, notable differences and similarities in traffic patterns emerge. Both figures track the same route, as evidenced by the identical paths shown, ensuring consistency in data collection across different sessions.

Figure 24 predominantly features segments of higher speeds, particularly visible in bright yellow across the northern and western sections of the loop, suggesting smoother traffic flow or less congestion during this particular drive. In contrast, Figure 26 shows these same areas with a mix of green to yellow hues, indicating a speed reduction which could be attributed to varied traffic conditions, time of day, and driving behavior. Notably, the top segment of the route in Figure 24 is almost entirely in bright yellow, signifying high speed, whereas Figure 24 shows a transition through green and yellow, highlighting a significant variation in speed. However, the bottom segment of the route maintains consistency between the two figures, predominantly exhibiting lower speeds. These differences underscore how external factors can influence driving speeds and behaviors on the same roadways.

Figures 25 and 27 represent the Z component of acceleration, ranging from purple (lower values) to yellow (higher values). The northern Section of Figure 25 Shows a relatively calm acceleration pattern, mainly in the teal to green range, indicating moderate vertical movements. As the route progresses southward, the color deepens to a darker green, particularly on the southward stretch, suggesting increased vertical accelerations, possibly due to rougher road surfaces or more pronounced road features like dips or bumps. In Figure 25 Similar to Figure 27, the north contains patches of darker green and yellow hues, indicating areas of greater vertical movement. The wider loop may cover more varied terrain or different road conditions. The intensity of the color changes is less pronounced in Figure 25 compared to the 27, especially along the eastern stretch. This could suggest smoother road conditions during this recording or differences in vehicle speed affecting the vertical acceleration measurements. Figure 27 shows a greater range of color changes and intensities along the route, suggesting more variability in the road surface conditions or the vehicle's response to these conditions at that time. Figure 25, while still displaying changes, shows a generally more consistent pattern of lighter greens along much of the route, which might indicate less severe vertical disturbances or a different driving style, possibly slower, which can reduce the perception of bumps.

The observation that both figures predominantly show similar overall patterns in the Z component of acceleration—mostly within a moderate range of green hues—suggests that the data collection process is consistent and reliable. This consistency across different datasets indicates that the acceleration measurements are stable and repeatable, supporting the accuracy and variability of the data.

3.4 MnDOT Reference Data Points and Next Steps

This chapter demonstrated the trial and testing of the scanner across various road segments, evaluating its capability to collect accurate and comprehensive data required for predictive modeling.

As previously mentioned, these highways were selected based on the availability of the MnDOT database, which provides an excellent reference dataset for this project's objectives. The database

contains precise features including latitude and longitude coordinates for segment start and end points, IRI measurements for both left and right wheel paths, and segment lengths. The data points from the MnDOT database are illustrated in the map below.

The MnDOT database points will be spatially linked to the nearest corresponding points from our OBD data collection, integrating OBD-derived features such as speed and acceleration. This integration will create a comprehensive dataset for each road segment, facilitating advanced regression and machine learning analyses detailed in subsequent chapters.

Chapter 4: Adapting LiRA Software

The LiRA software is specifically designed for predicting pavement characteristics, particularly the IRI and friction. To integrate their methodology into our datasets, we have adopted their multi-phase pipeline, which consists of preprocessing, feature engineering, and machine learning modeling.

The preprocessing phase involves cleaning the dataset to ensure readiness for map matching. Reference data points, defined for each 82-ft segment, are aligned with collected onboard sensor data based on GPS coordinates. This step ensures a seamless integration of reference and observed data, forming a unified and accurate dataset. Next, in the feature engineering phase, an extensive set of features is derived and extracted from each aligned segment. To refine the input variables, an optimal subset of relevant features is selected using the Sequential Feature Selection (SFS) algorithm. This selection process is performed on the train/validation dataset, leaving the test dataset to contain only the selected features to maintain consistency.

Finally, the machine learning phase involves training and fine-tuning predictive models on the train/validation dataset. These trained models are then applied to the test dataset to evaluate their generalization performance in estimating road conditions. This structured approach, illustrated in Figure 28, enables the effective prediction of pavement characteristics while ensuring the robustness and accuracy of the results.

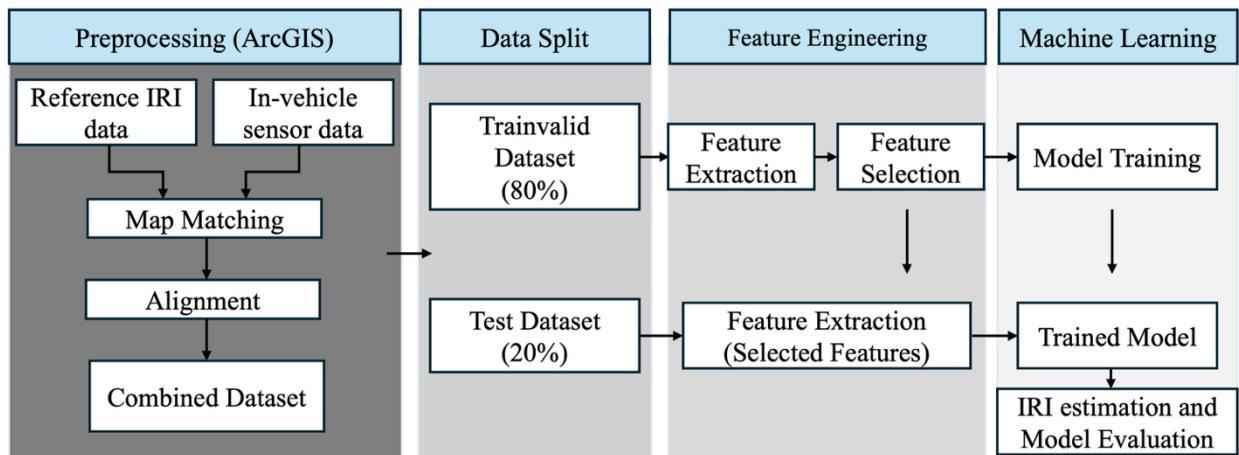


Figure 28. Framework of the proposed IRI prediction model.

4.1 Data Pre-processing

We have received MnDOT reference data points for IRI measurements from the same loops where we collected our data, with only minor differences in coverage. The next step involves merging the OBD-collected features with MnDOT ground truth measurements based on their locations and IRI values.

For the map-matching process, we utilized ArcGIS to integrate the datasets. Both the OBD and MnDOT datasets were imported into the ArcGIS environment and visualized as x, y coordinate points on the map. The integration of the OBD and MnDOT data was achieved by identifying the nearest OBD data points within a 25-meter radius for each MnDOT data point to account for GPS dither, which caused some OBD measurements to be recorded to the side of the road. Subsequently, the averaged values of the nearest OBD data parameters (PIDs) were calculated and joined to the corresponding MnDOT data points. This process resulted in a unified dataset that combines MnDOT data points, including their location information, the IRI, and 694 OBD-collected PIDs. This comprehensive dataset facilitates an in-depth analysis of the combined information. Figures illustrating the final merged dataset from both sources are provided below.

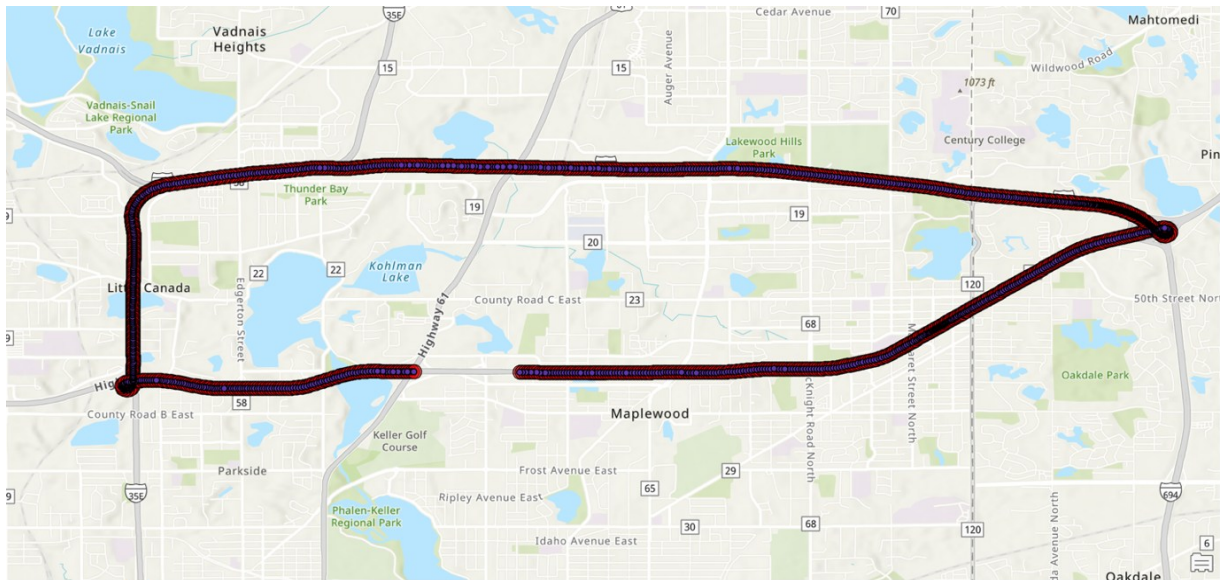


Figure 29. ArcGIS map of the final aligned data points from both OBD and MnDOT datasets for the MN-36 loop.

We traveled this loop twice, covering a 15-mile trip each time, to collect additional data points and compare the results for more reliable outputs. The combined data from both trips resulted in a total of 2,632 data points.



Figure 30. ArcGIS map of the final aligned data points from both OBD and MnDOT datasets for Metro to Northfield loop.

The Metro to Northfield loop encompasses a 95-mile trip, resulting in the collection of 5,494 data points. Combined with the MN-36 loop data, we have a total of 8,126 data points, including both MnDOT and OBD features, ready for modeling and prediction.

The third phase of data collection was conducted at MnROAD. This loop spans 2.6 miles; however, data was collected from both the inner and outer loops, each traversed four times at four different constant speeds. After alignment and map-matching, a total of 1,005 data points (528 data points for the inside loop and 477 data points for the outside loop) were obtained from this loop. The analysis of this data collection was conducted separately from the previous ones. Therefore, these data points were not combined with those from the two earlier loops.

Note that for the first and second data collection effort, the MnDOT data collection occurred before the OBD data collection. This suggests that some further pavement deterioration may have occurred between data collection efforts. However, the MnROAD data collection effort had both MnDOT data and OBD data collected at the same time.

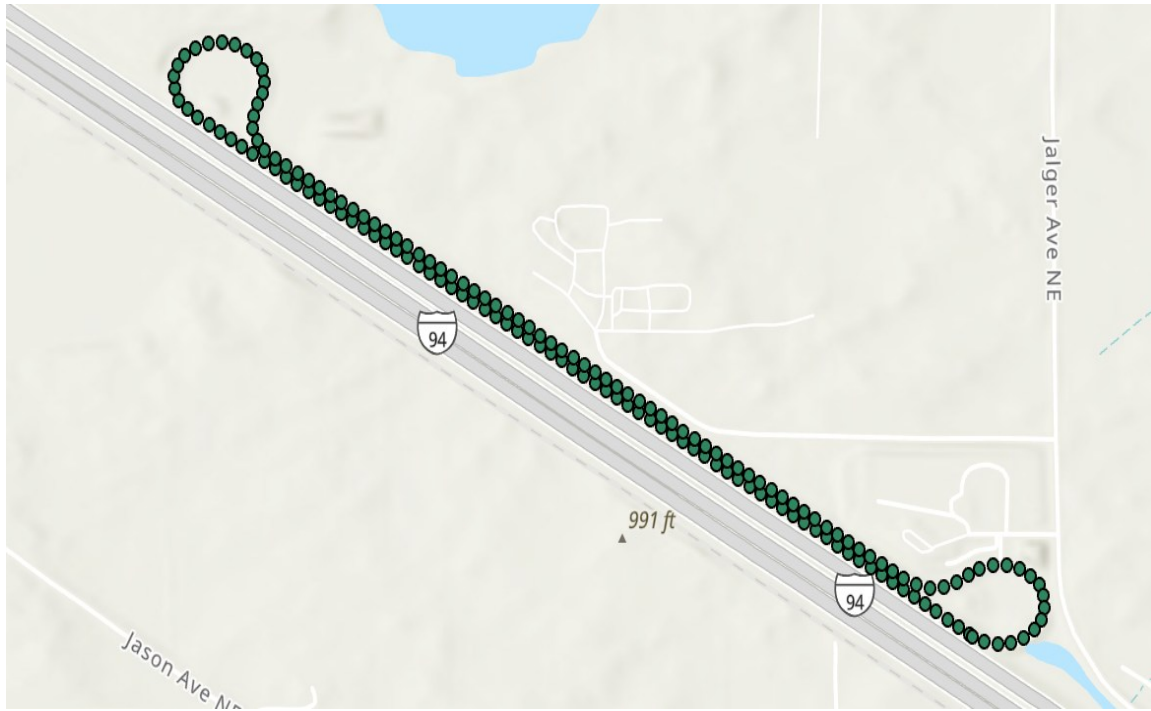


Figure 31. ArcGIS map of the final aligned data points from both OBD and MnDOT datasets for MnROAD inside loop.

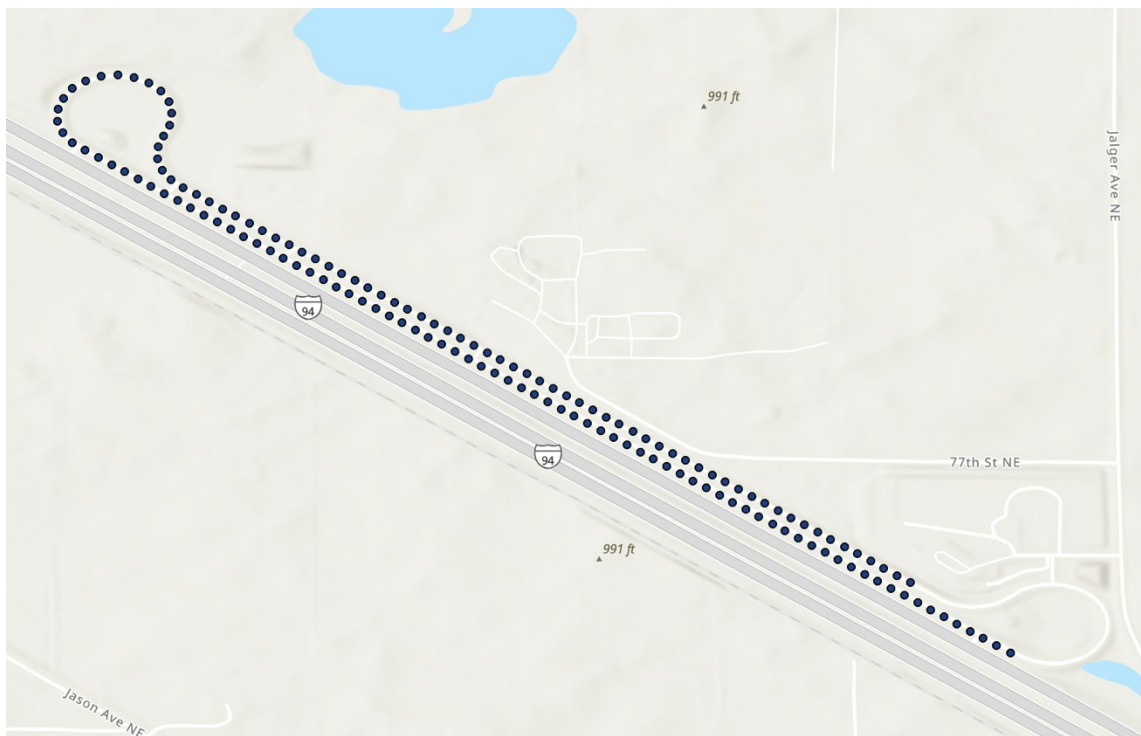


Figure 32. ArcGIS map of the final aligned data points from both OBD and MnDOT datasets for MnROAD outside loop.

4.2 Feature Selection

The data processing workflow began with segmentation, where the dataset was divided into overlapping segments using a window size of 100 and a step size of 50. Statistical and temporal features were extracted from each segment using the tsfel library. Sequential Feature Selection (SFS) was performed using a Random Forest model, selecting up to 30 features through 5-fold cross-validation. The selected features were then used to train a Random Forest model, which was evaluated using MSE. Finally, the importance of each feature was ranked to identify the most influential predictors. The feature selection process was implemented three times to ensure robustness and consistency in identifying the optimal feature set.

Since this chapter utilizes the LiRA software, it is essential to replicate their methodology accurately by first adopting their input feature categories. LiRA categorizes input features into three distinct groups: Speed-3D acceleration, Speed-vertical acceleration, and Speed-other features. The following sections present the results of these trials.

4.3 Prediction Modeling

4.3.1 Speed – 3D Acceleration

We selected speed and total acceleration in three dimensions as the primary features and extracted various combinations of these features to evaluate their influence on predicting IRI. Using feature selection methods, we identified the most impactful combinations, and the results are summarized in the following table.

Table 8. Optimized feature subset using speed and 3D acceleration.

Feature	Description
Vehicle_speed__MPH_Centroid	Centroid of vehicle speed distribution (MPH)
Vehicle_speed__MPH_Min	Minimum vehicle speed (MPH)
Vehicle_speed__MPH_Median	Median vehicle speed (MPH)
Vehicle_speed__MPH_Max	Maximum vehicle speed (MPH)
Vehicle_speed__MPH_RMS	Root mean square of vehicle speed (MPH)
Total_acceleration_Min	Minimum total acceleration (ft/s ²)
Total_acceleration_Median	Median total acceleration (ft/s ²)
Total_acceleration_Centroid	Centroid of total acceleration (ft/s ²)
Total_acceleration_Max	Maximum total acceleration (ft/s ²)
Total_acceleration_RMS	Root mean square of total acceleration (ft/s ²)
Total_acceleration_Entropy	Entropy of total acceleration (ft/s ²)

In this table, centroid represents the weighted average of a distribution: e.g., centroid of speed = $\frac{\sum(v_i w_i)}{\sum w_i}$, and entropy is a measure of uncertainty or randomness in a distribution: $H = -\sum p_i \log p_i$, where, w_i is the corresponding weights and p_i is the probability of each value.

4.3.1.1 Results for Speed and 3D Acceleration Input Features

We tried the Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) deep learning model, and a couple of classical machine learning models. Model performance was evaluated and compared using metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Relative Error (MRE). Scatter plots of predicted vs actual IRI for each method with the corresponding tables of results of the testing set are presented in the following.

CNN-LSTM METHOD

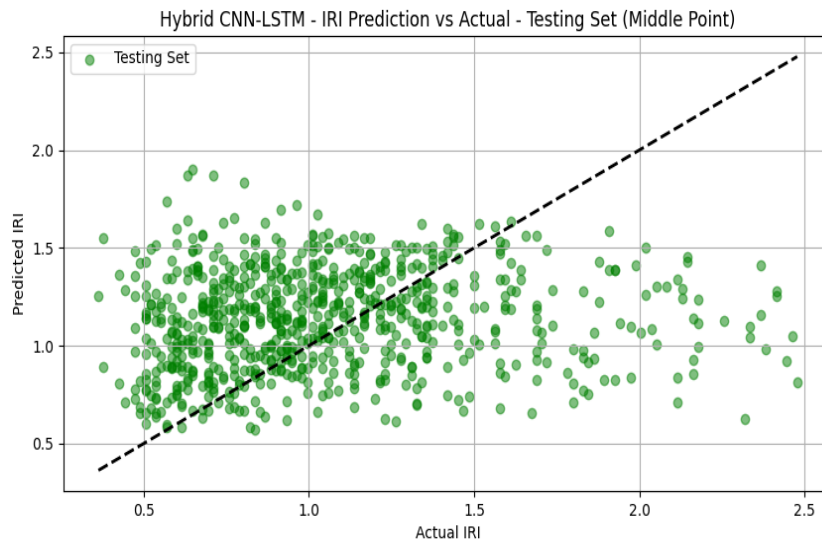


Figure 33. Predicted vs. actual IRI for CNN-LSTM model testing set.

Table 9. Performance metrics of CNN-LSTM method with speed and 3D acceleration input features.

Metric	Testing set
RMSE	0.842
MAE	0.531
MRE	0.312

RF METHOD

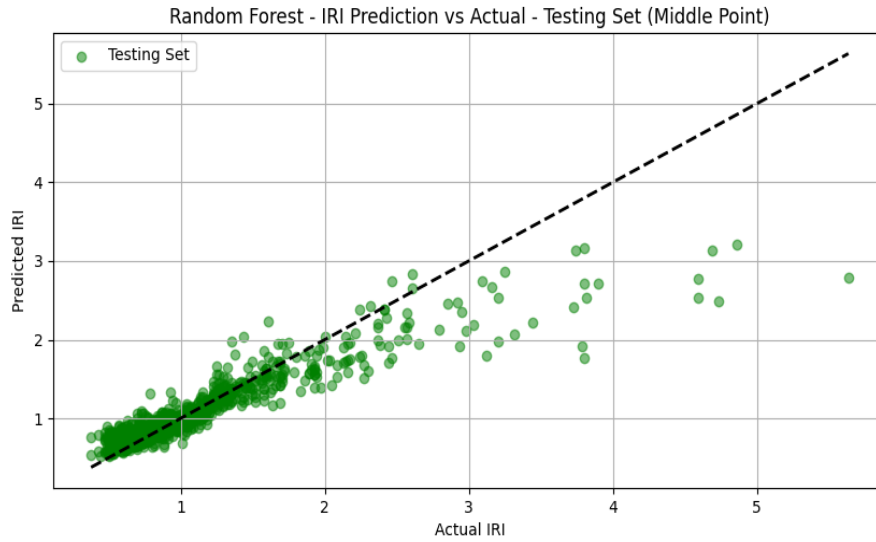


Figure 34. Predicted vs. actual IRI for RF model testing set.

Table 10. Performance metrics of RF method with speed and 3D acceleration input features.

Metric	Testing set
RMSE	0.288
MAE	0.199
MRE	0.209

KNN METHOD

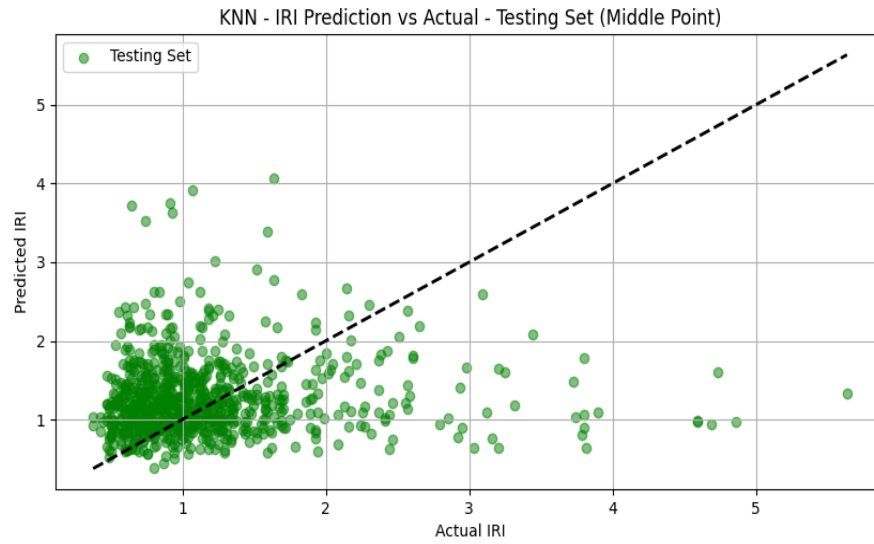


Figure 35. Predicted vs. actual IRI for RF model testing set.

Table 11. Performance metrics of KNN method with speed and 3D acceleration input features.

Metric	Testing set
RMSE	0.705
MAE	0.454
MRE	0.470

SVR METHOD



Figure 36. Predicted vs. actual IRI for SVR model testing set.

Table 12. Performance metrics of SVR method with speed and 3Dacceleration input features.

Metric	Testing set
RMSE	0.524
MAE	0.416
MRE	0.481

4.3.2 Speed-Vertical Acceleration

Speed and vertical acceleration (Z direction) were chosen as the primary features, and various combinations of these features were analyzed to determine their impact on predicting IRI. The most significant combinations were identified, with the results outlined in the following table.

Table 13. Optimized feature subset using speed and vertical acceleration.

Feature	Description
Vehicle_speed__MPH_RMS	Root mean square of vehicle speed (MPH)
Vehicle_speed__MPH_Min	Minimum vehicle speed (MPH)
Vehicle_speed__MPH_Median	Median vehicle speed (MPH)
Vehicle_speed__MPH_Max	Maximum vehicle speed (MPH)
Vehicle_speed__MPH_Centroid	Centroid of vehicle speed distribution (MPH)
Accel_Z__ft_s__Centroid	Centroid of vertical acceleration (ft/s ²)
Accel_Z__ft_s__Min	Minimum vertical acceleration (ft/s ²)
Accel_Z__ft_s__Max	Maximum vertical acceleration (ft/s ²)
Vehicle_speed_MPH_Entropy	Entropy of vehicle speed (MPH)
Accel_Z__ft_s__RMS	Root mean square of vertical acceleration (ft/s ²)
Accel_Z__ft_s__Entropy	Entropy of vertical acceleration (ft/s ²)

4.3.2.1 Results for speed and vertical acceleration input features

CNN-LSTM METHOD

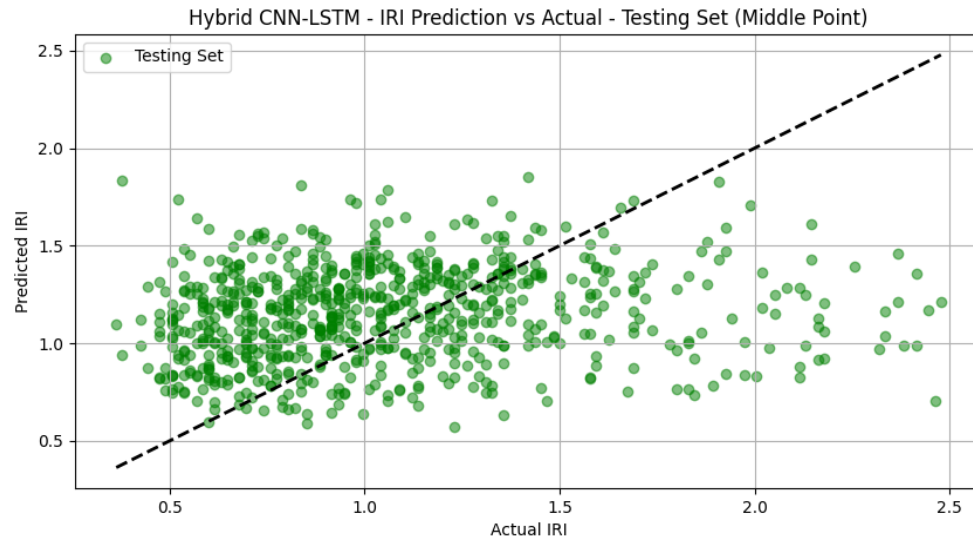


Figure 37. Predicted vs. actual IRI for CNN-LSTM model testing set.

Table 14. Performance metrics of CNN-LSTM method with speed and vertical acceleration input features.

Metric	Testing set
RMSE	0.852
MAE	0.649
MRE	0.837

RANDOM FOREST METHOD

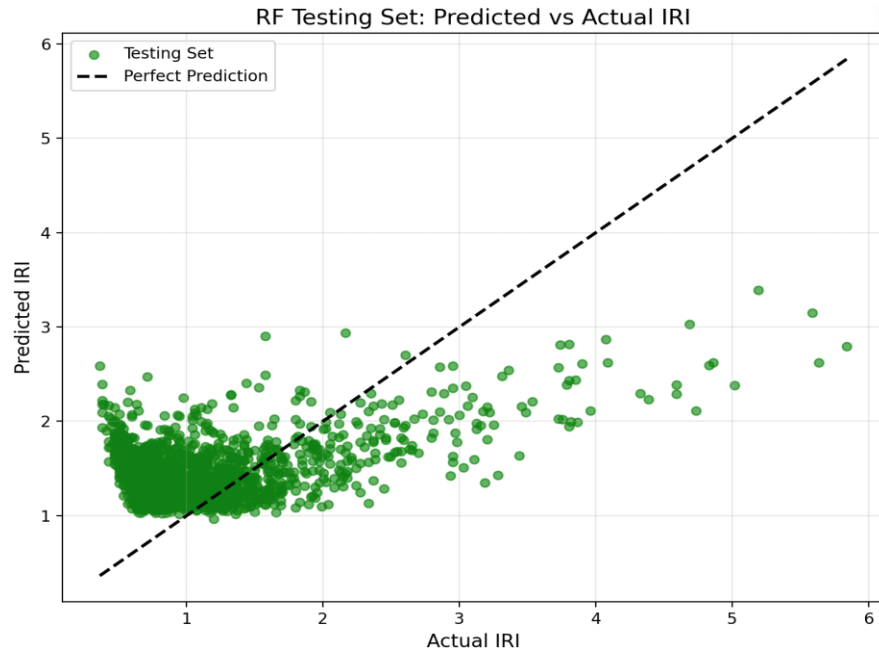


Figure 38. Predicted vs. actual IRI for RF model testing set.

Table 15. Performance metrics of RF method with speed and vertical acceleration input features.

Metric	Testing set
RMSE	0.67
MAE	0.533
MRE	0.637

K-NEAREST NEIGHBORS (KNN) METHOD

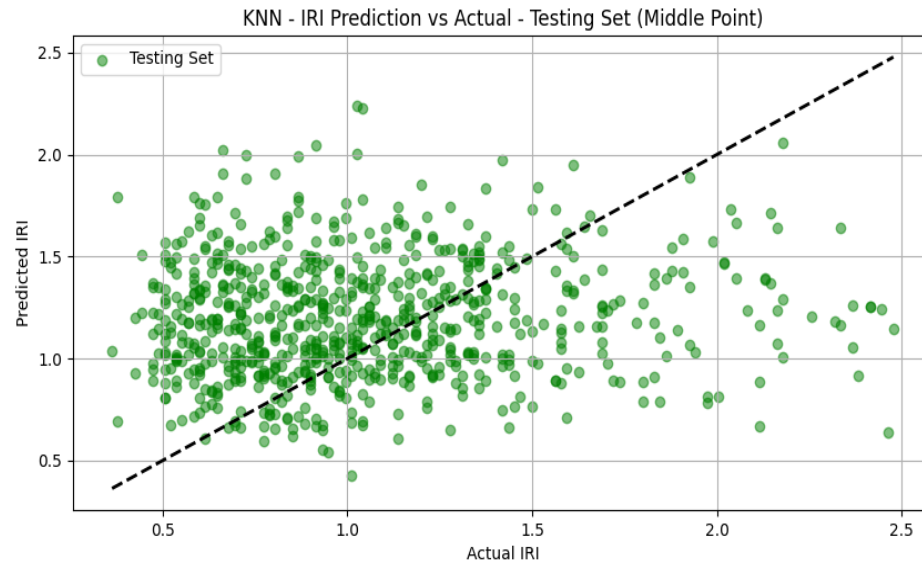


Figure 39. Predicted vs. actual IRI for KNN model testing set.

Table 16. Performance metrics of KNN method with speed and vertical acceleration input features.

Metric	Testing set
RMSE	0.918
MAE	0.672
MRE	0.758

SVR METHOD

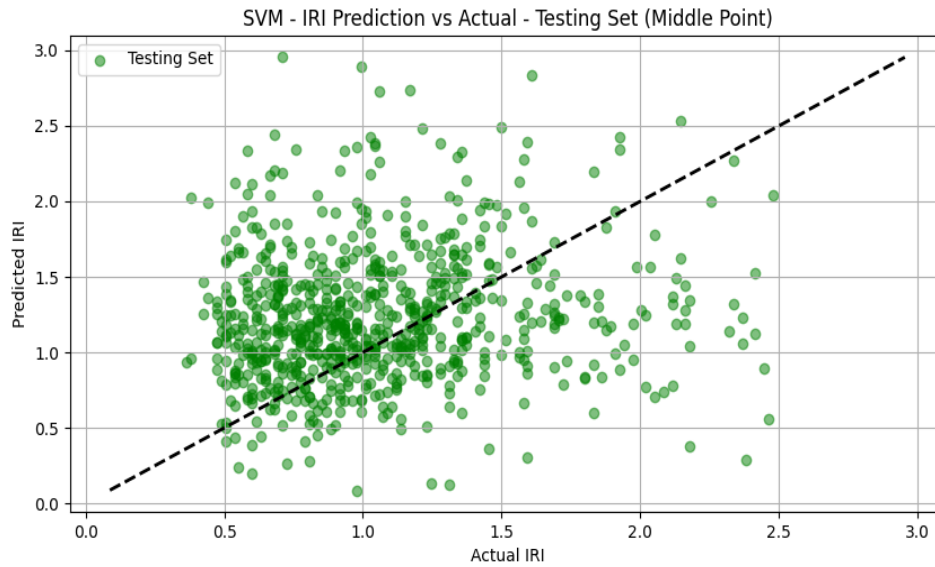


Figure 40. Predicted vs. actual IRI for SVR model testing set.

Table 17. Performance metrics of SVR method with speed and vertical acceleration input features.

Metric	Testing set
RMSE	0.818
MAE	0.664
MRE	0.827

4.3.3 Speed-Other Features

A feature selection process was then conducted across all collected PIDs to identify the most influential features. The results of this analysis are displayed in the following table.

Table 18. Optimized feature subset using speed and other features.

Feature	Description
Vehicle_speed__MPH	Vehicle speed in miles per hour
Roll__deg__	Vehicle tilt around longitudinal axis (degrees)
Pitch__deg__	Vehicle tilt around the lateral axis (degrees)
Accel_Y__ft_s__	Lateral acceleration (ft/s ²)
Accel_X__ft_s__	Longitudinal acceleration (ft/s ²)
Accel_Z__ft_s__	Vertical acceleration (ft/s ²)
Total_acceleration__ft_s__	Combined acceleration magnitude (ft/s ²)
Accel__Grav__X__ft_s__	Gravitational acceleration along X-axis (ft/s ²)
Accel__Grav__Z__ft_s__	Gravitational acceleration along Z-axis (ft/s ²)
+ Combinations of speed and total acceleration	

4.3.3.1 Results for Speed and Other Features as Input

CNN-LSTM METHOD

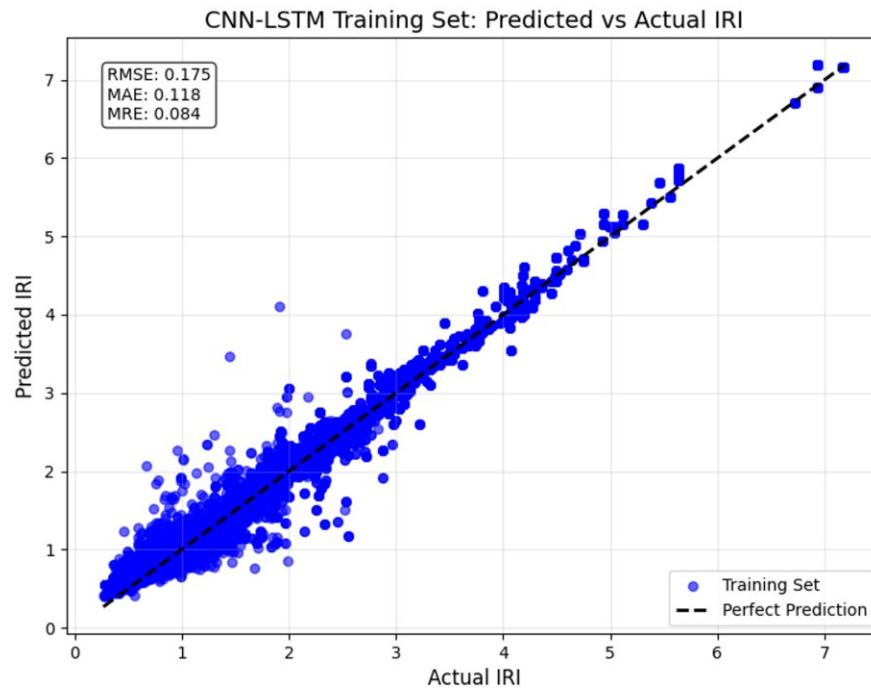


Figure 41. Predicted vs. actual IRI for CNN-LSTM model training set.

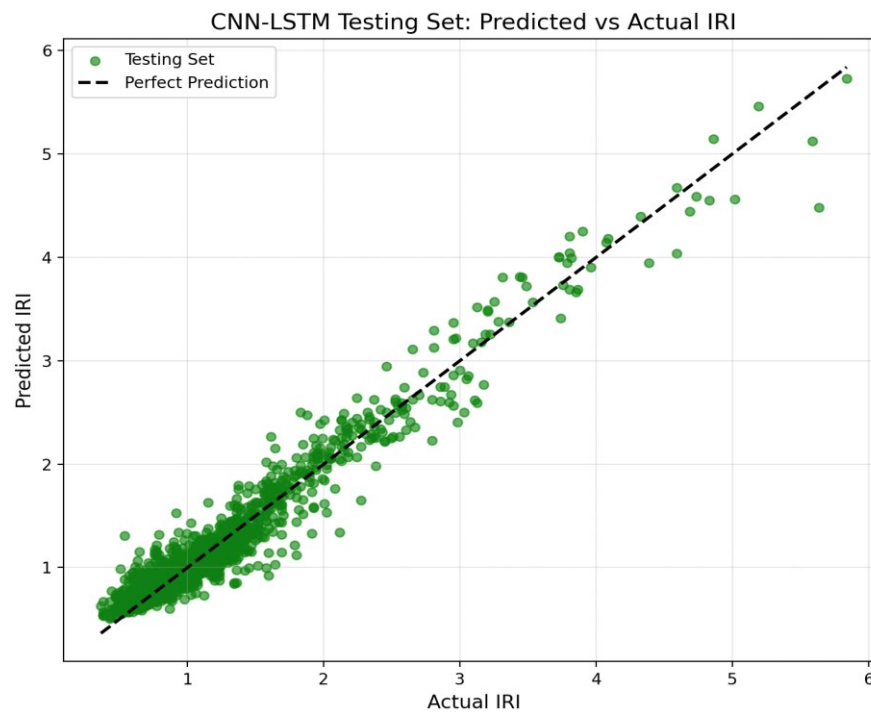


Figure 42. Predicted vs. actual IRI for CNN-LSTM model testing set.

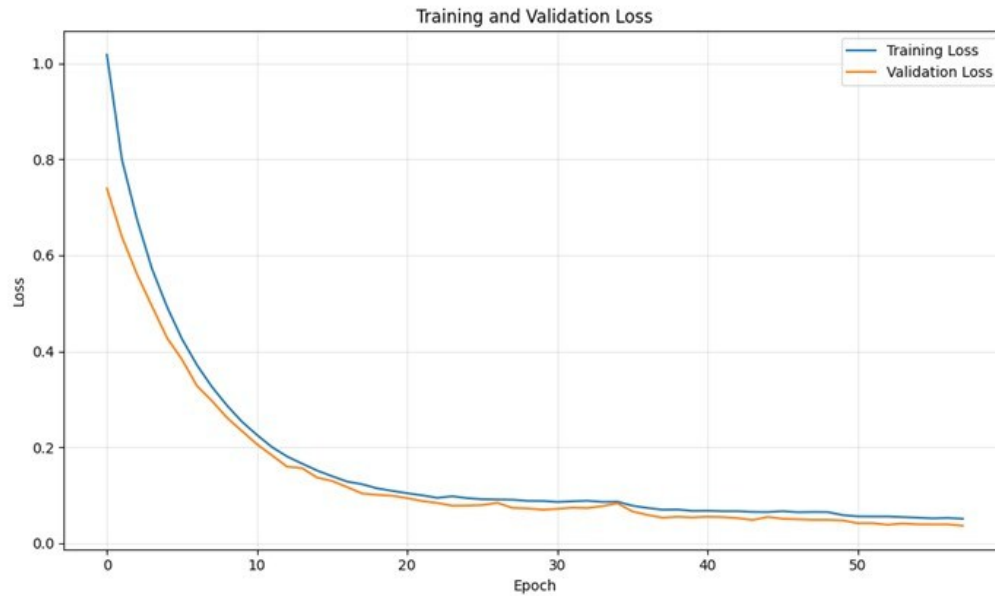


Figure 43. Training and validation loss vs epochs for showing the CNN-LSTM model performance.

Table 19. Performance metrics of CNN-LSTM method with speed and other influential input features.

Metric	Training set	Testing set
RMSE	0.175	0.172
MAE	0.118	0.128
MRE	0.084	0.131

Figure 43 illustrates the loss convergence of the Hybrid CNN-LSTM model over 50 training epochs. Note that an epoch refers to a complete pass through the training data. The training loss and validation loss both decrease steadily, indicating effective learning and generalization. Initially, both losses start at relatively high values (~1 for training and ~0.7 for validation) but rapidly decline within the first 10 epochs, demonstrating that the model quickly learns meaningful patterns from the data. Beyond 20 epochs, both losses stabilize around 0.1, showing that the model reaches a point of diminishing returns in learning improvements. The close alignment between training and validation loss suggests minimal overfitting, confirming that the model generalizes well to unseen data. This behavior indicates a well-trained model capable of accurately predicting IRI.

ANN METHOD

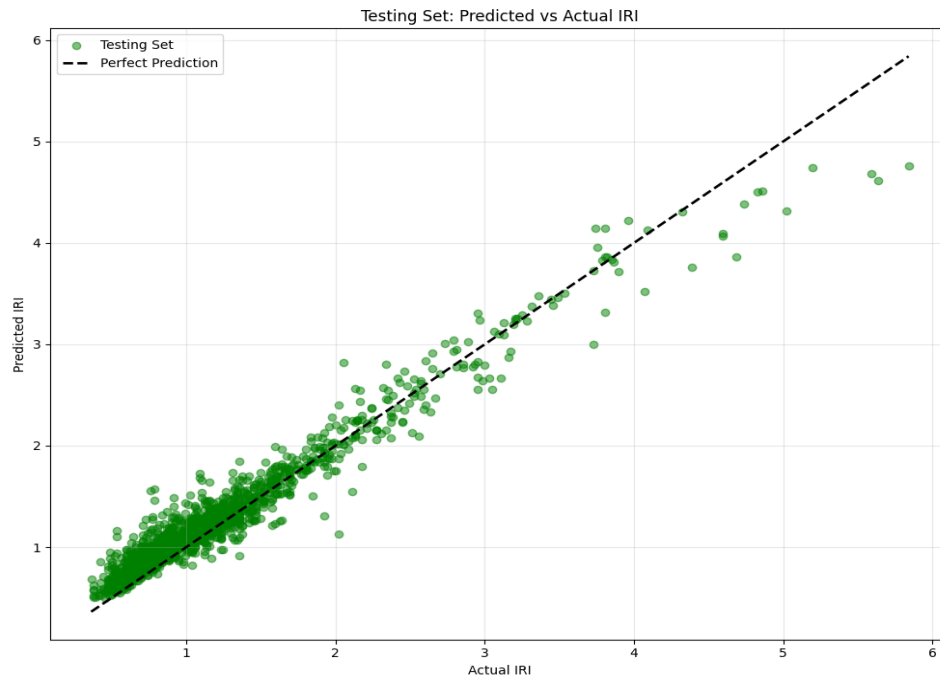


Figure 44. Predicted vs. actual IRI for ANN model testing set.

Table 20. Performance metrics of ANN method with speed and other influential input features.

Metric	Testing set
RMSE	0.182
MAE	0.136
MRE	0.146

SVR METHOD

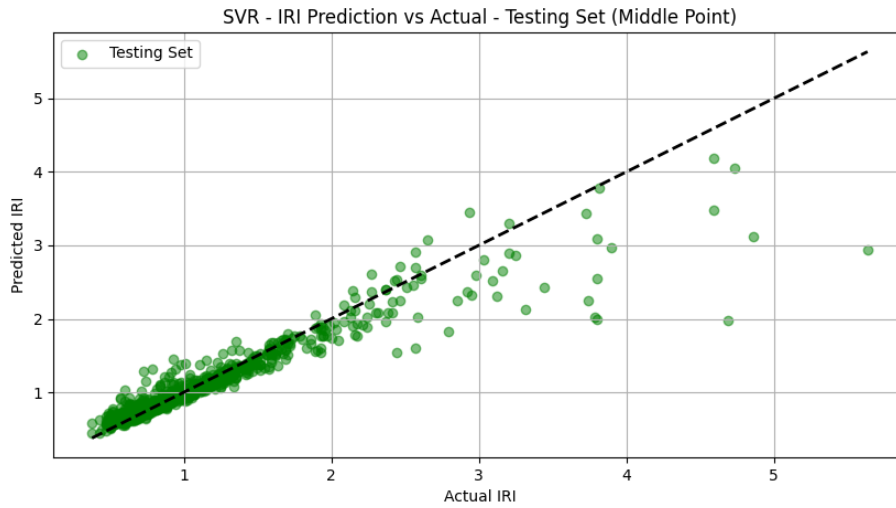


Figure 45. Predicted vs. actual IRI for SVR model testing set.

Table 21. Performance metrics of SVR method with speed and other influential input features.

Metric	Testing set
RMSE	0.226
MAE	0.177
MRE	0.186

RF METHOD

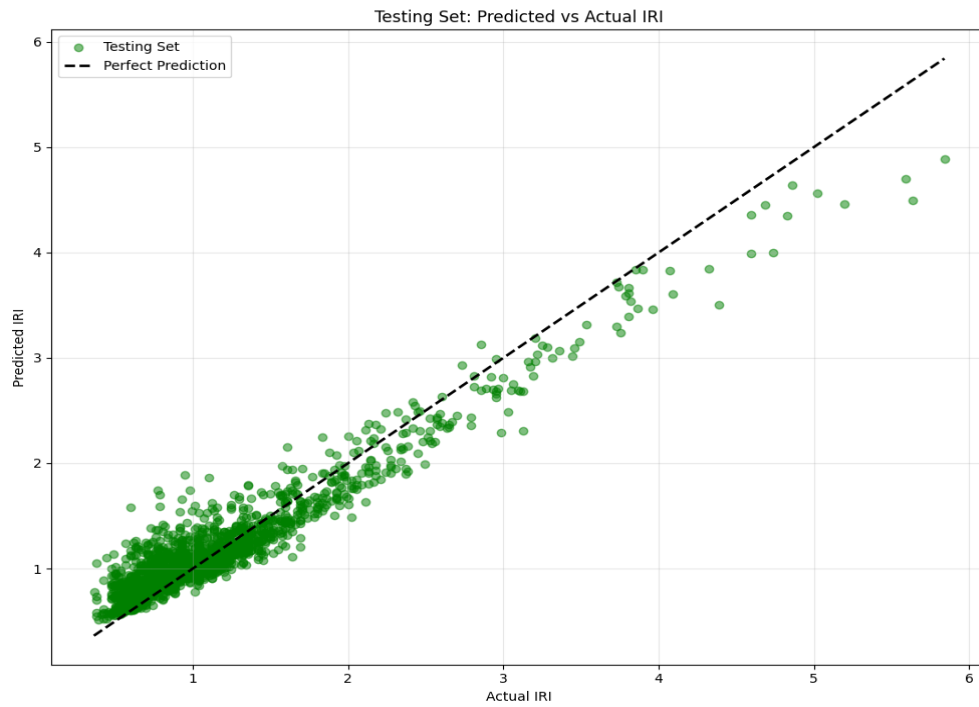


Figure 46. Predicted vs. actual IRI for RF model testing set.

Table 22. Performance metrics of RF method with speed and other influential input features.

Metric	Testing set
RMSE	0.224
MAE	0.169
MRE	0.180

KNN METHOD

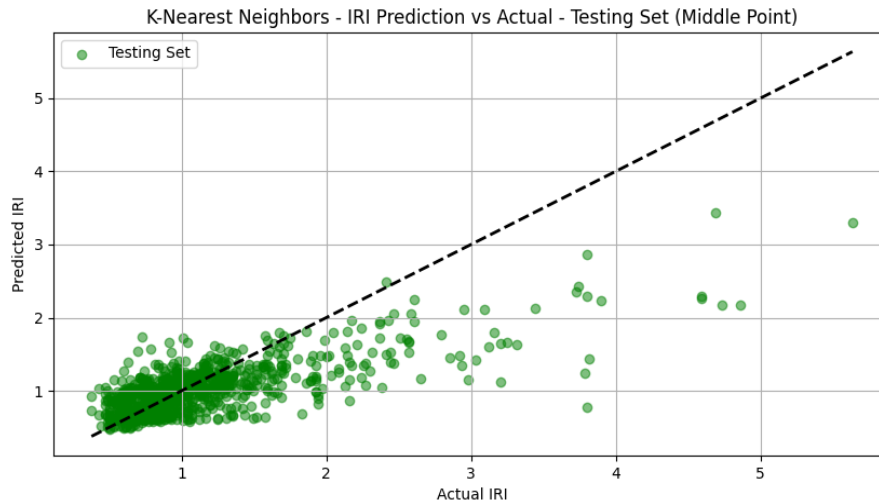


Figure 47. Predicted vs. actual IRI for KNN model testing set.

Table 23. Performance metrics of KNN method with speed and other influential input features.

Metric	Testing set
RMSE	0.544
MAE	0.395
MRE	0.418

The results presented in this chapter demonstrate that the Hybrid CNN-LSTM model achieves superior performance across all three input feature categories, with optimal accuracy when utilizing speed and related features as input variables. The model effectively predicts IRI values for both training and testing datasets, as evidenced by the strong correlation between predicted and actual IRI measurements. Although minor deviations occur at higher IRI values, particularly within the testing dataset, the overall performance remains robust, confirming the model's strong generalization capabilities.

These findings validate the effectiveness of the CNN-LSTM architecture in capturing both temporal and spatial dependencies inherent in the data, establishing it as a reliable methodology for road roughness estimation. Given its superior predictive performance, this model demonstrates significant potential for large-scale IRI prediction applications, supporting proactive road maintenance strategies and enhanced infrastructure management practices.

Further analysis will be conducted in subsequent chapters, including individual testing of models on separate datasets and comprehensive comparative evaluations to provide deeper insights into model performance and applicability.

Chapter 5: Collect Data

This chapter involves data collection to effectively implement the LiRA software components. A sufficient amount of independently collected data, along with reliable reference data from MnDOT, was required. To meet this need, three loops—MN-36, Metro to Northfield, and MnROAD—were selected based on the availability of MnDOT’s database, which serves as an excellent reference for this project's objectives. Data collection was conducted using the chosen onboard vehicle scanner, OBDLink LX, which was selected for its broad compatibility, extensive data retrieval capabilities, ease of use, and support for EVs. A map illustrating the data collection locations for these two loops is attached below.

5.1 Data Collections

To develop a robust and reliable pavement quality assessment model, multiple phases of data collection were conducted across different road networks. These efforts aimed to capture a diverse range of pavement conditions under real-world driving scenarios while ensuring synchronization with high-precision reference data. The collected datasets serve as the foundation for training and validating machine learning models for IRI prediction.

The data collection process was structured around three key test routes: the MN-36 loop, the Metro to Northfield loop, and the MnROAD loop. Each route was carefully selected based on its pavement characteristics, the availability of MnDOT reference data, and its relevance to infrastructure monitoring and assessment.

The following subsections present a comprehensive overview of each data collection effort, detailing the test routes, vehicle setup, and the spatial distribution of the collected data. Table 24 summarizes all data collection activities, including the dates and the number of data points obtained.

Table 24. Summary of data collection efforts.

Data collection location	Collection Date(s)	Trip Length (mile)	Number of rounds	Number of data points
MN-36 loop	May and September 2024	15	2	2,632
Metro to Northfield loop	November 2024	95	1	5494
MnROAD loop	January 2025	2.6	8	1005

5.1.1 MN-36 Loop

The MN-36 data collection route follows a shorter, enclosed loop in the northeastern Minneapolis-St. Paul metropolitan area as shown in Figure 48, covering urban and suburban road segments with consistent traffic flow. It is particularly useful for analyzing the model's performance under controlled, high-traffic conditions. Two separate data collections were conducted on this route—one in September and another in May 2024, specifically focusing on roads with old pavement to align with prepared MnDOT reference data points.

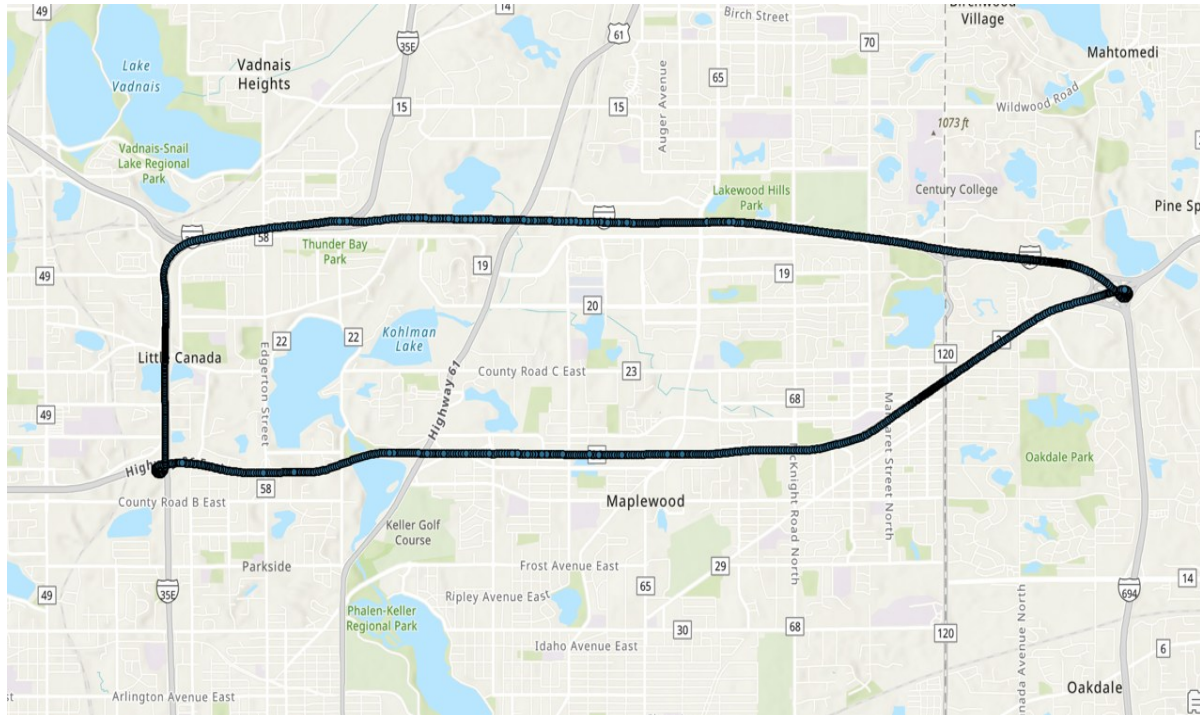


Figure 48. ArcGIS map of OBD-II data collection of MN-36 loop.

5.1.2 Metro to Northfield Loop

This data collection, conducted on November 8, extends from Minneapolis southward into Dakota County, reaching as far as Northfield before looping back, as illustrated in Figure 49. The route covers a mix of urban, suburban, and rural roadways, including major highways and arterial roads. This diverse coverage ensures a comprehensive dataset that captures variations in road roughness, traffic density, and speed conditions, making it valuable for evaluating model performance across different driving environments.



Figure 49. Color-coded map of speed for Metro to Northfield loop data collection.

The final output files are CSV files containing latitude and longitude coordinates along with 694 other PIDs. These PIDs include acceleration in three axes, speed, roll, pitch, and various other features, providing a comprehensive dataset for analyzing vehicle dynamics and road conditions.

5.1.3 MnROAD Loop

After collecting a sufficient amount of data and successfully implementing predictive models with acceptable performance in estimating IRI, we initiated a new phase of data collection at MnROAD. The objective of this phase was to drive the MnROAD loop, which consists of diverse pavement segments with varying surface qualities, at controlled speeds of 20 mph, 40 mph, 60 mph, and finally 80 mph. This dataset is particularly valuable as it encompasses a wide range of pavement conditions, making it a robust addition to the IRI prediction model's database. The primary focus of this data collection effort was to examine the influence and significance of vehicle speed in IRI prediction.

The data collection was conducted on January 30th. To ensure accuracy, a MnDOT Pathway van was positioned ahead of our vehicle, collecting reference IRI data while we closely followed it, maintaining temporal alignment. We collected data from both the inside and outside loops of MnROAD at the four designated speeds. As a result, this phase of data collection produced eight distinct datasets, each corresponding to a specific loop and speed combination.

Figure 50, which documents the data collection process, is attached below. It illustrates the setup, the MnDOT Pathway van collecting reference IRI data, and our vehicle following it to maintain synchronization.



Figure 50. (a) Embedding the scanner in the vehicle for data collection. (b) MnDOT Pathway van leading our data collection vehicle on MnROAD.

Figures 51 and 52, respectively, show the OBD-collected data and the reference data points shared by MnDOT on maps.

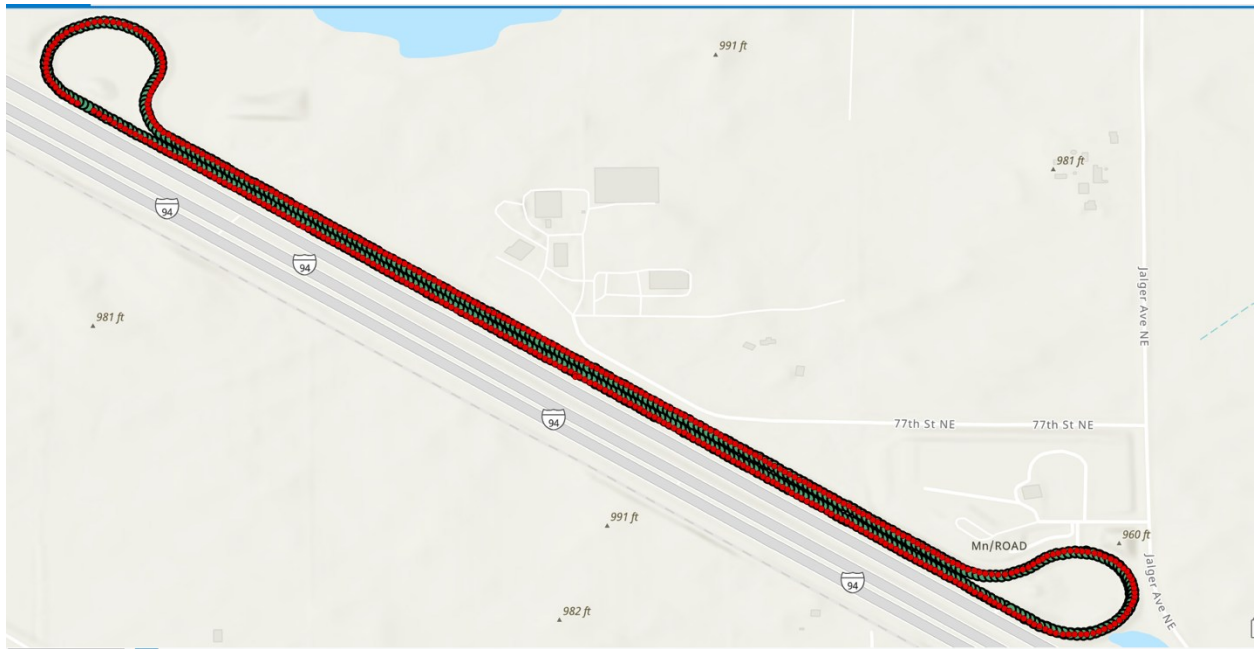


Figure 51. Map of OBD-II data collection on the MnROAD loop. Green points represent the inside loop, while red points indicate the outside loop.

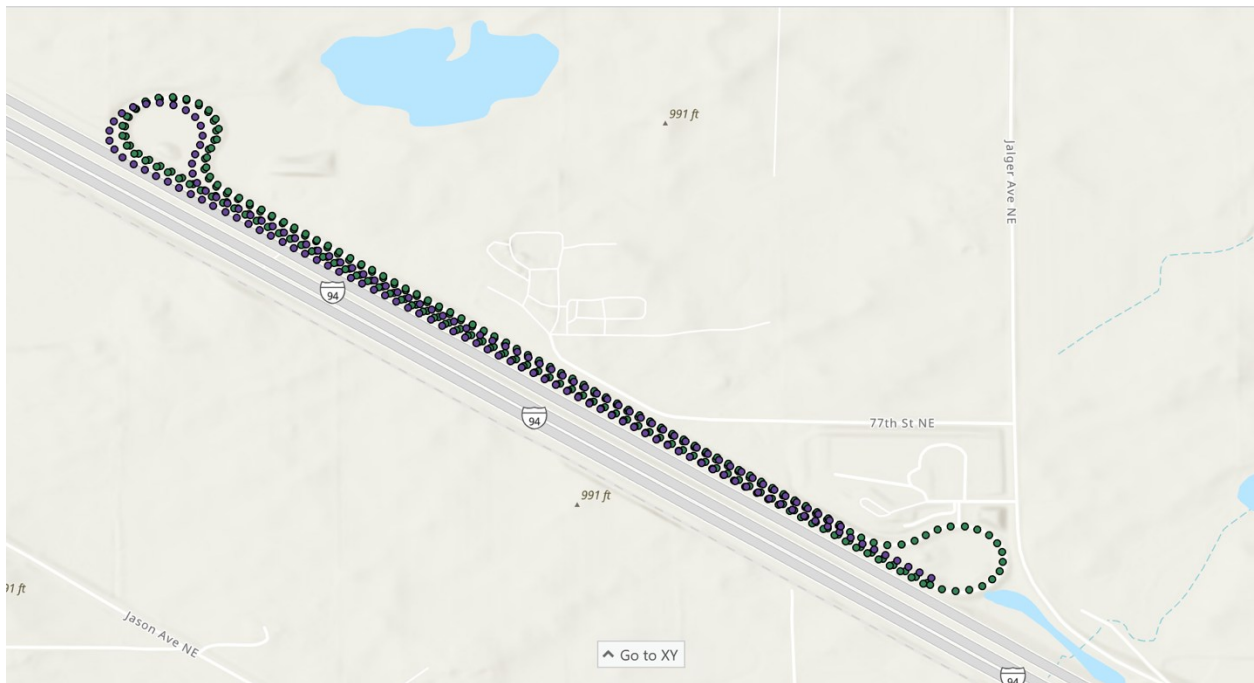


Figure 52. Map of MnDOT data collection on the MnROAD loop. Green points represent the inside loop, while purple points indicate the outside loop.

As illustrated in Figure 52, the MnDOT data collected from MnROAD does not include the east loop. Consequently, the final combined dataset for the outside loop will also lack the east loop, as our reference data points are derived from MnDOT's dataset.

This study aimed to explore the potential of using onboard vehicle data for pavement quality assessment by leveraging machine learning techniques and high-fidelity reference data from MnDOT. Through a structured data collection process across three key test routes—MN-36 loop, Metro to Northfield loop, and MnROAD loop—we gathered diverse datasets representing a range of pavement conditions, traffic environments, and speed variations. Each phase of data collection was carefully designed to ensure synchronization with reference measurements from MnDOT.

The incorporation of controlled-speed tests at MnROAD further enhances the dataset, enabling a detailed analysis of speed's impact on IRI prediction. These efforts contribute to a comprehensive database that advances pavement monitoring and supports data-driven maintenance strategies.

Moving forward, the collected datasets will be utilized to train and validate predictive models capable of estimating pavement conditions using OBD-II data. The insights gained from this study have the potential to advance infrastructure monitoring by reducing dependency on expensive instrumented vehicles, enabling a more scalable and cost-effective approach to pavement quality assessment. Further research will focus on refining model accuracy, integrating additional vehicle-based parameters, and evaluating the feasibility of large-scale implementation.

Chapter 6: Data Analysis and Recalibration of the Danish Models for Current Project Conditions

Building upon the successful data collection efforts described in Chapter 5 and the initial LiRA software adaptation outlined in Chapter 4, this chapter focuses on the comprehensive analysis and recalibration of the Danish LiRA models to optimize their performance for Minnesota's specific road conditions and vehicle characteristics. The original LiRA methodology, developed for European road networks and vehicle fleets, requires systematic adaptation to account for differences in pavement materials, construction standards, climate conditions, and vehicle dynamics prevalent in Minnesota. This recalibration process involves extensive data preprocessing, feature engineering optimization, model parameter tuning, and performance validation using the multi-loop datasets collected across MN-36, Metro to Northfield, and MnROAD facilities. The ultimate objective is to develop a robust, locally-calibrated predictive framework that can accurately estimate IRI values using onboard vehicle data while maintaining the scalability and cost-effectiveness advantages of the original LiRA approach.

6.1 An Overview of the Data Cleaning Process

The data cleaning process represents a critical foundation for ensuring model reliability and accuracy, particularly given the heterogeneous nature of the collected OBD-II datasets and the need for precise alignment with MnDOT reference measurements. This comprehensive preprocessing workflow begins with the identification and removal of unnecessary data points, including GPS coordinate anomalies, sensor reading outliers, and temporal inconsistencies that may have occurred during data collection across the three test routes. The cleaning process also addresses synchronization challenges between OBD-II measurements and MnDOT reference data, ensuring spatial alignment within acceptable tolerance limits to maintain the validity of the training dataset for subsequent machine learning applications.

6.1.1 Initial Data Assessment and Quality Control

Initial data assessment involves examining the raw datasets collected from the three test routes, providing statistical summaries of data completeness, sensor availability, and initial quality indicators. The assessment includes evaluation of data collection frequency, identification of systematic sensor failures, and analysis of GPS signal quality across different route segments. Documentation of the original dataset characteristics, including the distribution of the 694 PIDs across different vehicle systems (engine, transmission, acceleration sensors, etc.), establishes the baseline for subsequent cleaning operations.

A comprehensive visualization of the collected features across each route using multiple analytical approaches has been done. Color-coded mapping techniques were employed to display important features such as vehicle speed, acceleration profiles, and sensor readings directly on the trip maps, providing immediate visual verification of data reasonableness and spatial consistency. Additionally,

distribution analysis was conducted to examine feature variations over time, ensuring that collected parameters such as speed ranges and acceleration values fell within realistic and expected bounds for typical driving conditions. These visualizations served as both quality control measures and preliminary exploratory analysis tools, offering insights into the nature and characteristics of the collected data.

Following the initial visualization assessments presented in both Chapters 3 and 5, a systematic review of the 694 PIDs revealed significant data quality issues requiring immediate attention. A substantial portion of the collected PIDs contained only zero values, indicating either sensor unavailability or systematic collection failures that rendered these parameters unusable for analysis, like wheel speeds. Furthermore, many features were identified as irrelevant to the primary objective of IRI prediction, including parameters such as fuel rate, instant CO₂ rate, and CO₂ flow measurements. These environmentally-focused and fuel consumption-related PIDs, while potentially valuable for other applications, were removed from the dataset to focus on vehicle dynamics parameters most relevant to pavement quality assessment.

A critical challenge emerged from the data collection frequency settings, where OBD-II measurements were configured to sample at 0.1-second intervals while GPS coordinates remained static over longer periods. This temporal mismatch created entries where multiple sensor readings corresponded to identical geographic locations, necessitating a location-based aggregation approach. To address this issue, all features were averaged across each unique geographic location point, recognizing that while some parameters, such as vehicle speed and acceleration, varied within these time windows, others remained constant. This aggregation process resulted in a simplified, cleaner, and unified location-based dataset optimally structured for input into ArcGIS alignment procedures, ensuring that each geographic coordinate corresponded to a single, representative set of vehicle dynamics measurements. This cleaning process significantly reduced data complexity while preserving essential information for IRI prediction.

6.1.2 GPS Coordinate Validation and Spatial Consistency

The GPS validation process focuses on identifying and correcting coordinate anomalies that could compromise the spatial alignment between OBD-II data and MnDOT ground truth measurements. To validate the correctness of the collected coordinate data, maps of all datasets were plotted to ensure accurate positional data collection along the desired road segments. This visual verification process confirmed that the GPS coordinates correctly traced the intended routes across MN-36, Metro to Northfield, and MnROAD loops, identifying any systematic deviations or data collection errors that could affect subsequent spatial alignment procedures. Figures containing reference and OBD data points on maps for all data collections are presented in Chapters 4 and 5 previously.

Since the actual trips were not exactly the desired loops (for example, starting from the Civil Engineering department on campus and driving to MN-36), many data points were removed based on the plotted maps to retain only the desired loop segments that corresponded to available MnDOT reference data. This spatial filtering ensured that the final dataset contained only GPS coordinates within the target road segments where reference IRI measurements were available for model training and validation. These

small portions of the loops were removed from the input data to ensure that all remaining data points had associated ground truth IRI values for model training and validation purposes.

6.2 Location Synchronization and Alignment

This critical section addresses the challenge of synchronizing OBD-II measurements with MnDOT reference data to ensure accurate training pairs for machine learning models. The process involves location alignment, accounting for the potential distance between datapoints. The subsection describes the development of location matching algorithms and tolerance threshold establishment to maintain data quality while maximizing the number of valid training samples.

For this objective, the cleaned OBD-II collected data was first added to ArcGIS as the primary layer, followed by the addition of MnDOT's ground truth measurements as the second layer. This approach allowed visualization of each data point on the map using two different colors, revealing that the locations were not exactly synchronized. Typically, for each MnDOT data point, two or more OBD data points could be identified in close proximity, as the MnDOT location data represented an 82-foot road segment rather than a precise point location.

To create a unified dataset with one set of OBD data features corresponding to each MnDOT ground truth measurement, the MnDOT layer was maintained as the reference framework. A systematic ArcGIS workflow was implemented using the following sequence of commands: first, the "Near" tool to identify proximate OBD points for each MnDOT reference location; second, the "Statistics" function to calculate averaged values among the nearby OBD points; and finally, the "Join" operation to link each corresponding averaged OBD point set to its respective MnDOT point using the Near FID identifier. This process resulted in a unified layer containing MnDOT ground truth measurement locations, one column of reference IRI values, and multiple additional columns of averaged OBD features corresponding to each IRI measurement point.

By comparison, the LiRA project, which employed a dual-source data collection strategy using both IoT-equipped vehicles and conventional pavement measurement equipment. The primary platform consisted of GreenMobility's electric Renault Zoe vehicles equipped with AutoPi dongles (enhanced OBD-II devices) that captured accelerometer data and various vehicle parameters through the Controller Area Network (CAN bus). These vehicles traveled approximately 350 kilometers of deliberately selected road sections covering both motorways and municipal roads to capture diverse pavement conditions. Simultaneously, conventional pavement management instrumented vehicles collected reference data using specialized equipment including the P79 Profilometer (with 25-point lasers measuring road profiles, IRI, rut depth, and MPD), ViaFrik (friction coefficient measurement), CPX (tire/road noise measurement), and ARAN9000 (multi-functional road testing for roughness, texture, and crack measurements). All data was synchronized with onboard computers and transmitted to a cloud-based system for structured storage.

6.2.1 LiRA's Advanced Signal-Based Data Processing and Alignment Methodology:

The LiRA project implemented a sophisticated physics-informed approach to data alignment that fundamentally differed from traditional geographic matching methods. Their methodology centered on generating synthetic acceleration signals from P79 profilometer road profile measurements using quarter-car vehicle dynamics models, which were then cross-correlated with actual measured acceleration data from Green Mobility vehicles to identify optimal temporal alignments. LiRA addressed the challenge of variable sampling frequencies (ranging from 25-250 Hz) and GPS inaccuracies through a complex misalignment detection algorithm that utilized sliding window correlation analysis across different time shifts. Their approach involved segmenting reference IRI data into fixed-length windows (100m segments with 10m steps) and employing KDTree algorithms with Haversine distance calculations for spatial matching, followed by iterative correlation analysis to extract misalignment transformation maps. This physics-based methodology enabled them to account for the inherent temporal delays and spatial offsets that occur when different vehicles traverse the same road segments at varying speeds, ultimately creating a more robust alignment between high-fidelity reference measurements and consumer-grade vehicle sensor data.

6.2.2 Comparison with Our Geographic Information System-Based Processing Approach:

Our approach utilized a fundamentally different paradigm based on spatial proximity analysis and statistical aggregation through ArcGIS workflows, emphasizing geographic coordinate matching rather than signal correlation. While LiRA invested significant computational resources in synthetic signal generation and cross-correlation analysis to achieve temporal alignment, our methodology focused on location-based data consolidation by averaging OBD-II measurements across unique GPS coordinates to address the temporal mismatch between 0.1-second sensor sampling intervals and static GPS positions. This approach prioritized computational efficiency and practical implementation over the complex physics-based modeling employed by LiRA, recognizing that while LiRA's signal correlation method may provide superior theoretical accuracy in capturing vehicle-road dynamics, our GIS-based spatial statistical approach offers a more accessible and scalable solution for pavement condition assessment applications with acceptable accuracy for practical engineering purposes.

As a final step, we ensure consistency and comparability of cleaned datasets across the three test routes. This includes analysis of sensor collection consistency, speed profile comparisons, and statistical distribution matching to ensure the merged dataset provides a representative foundation for model training and validation while preserving the distinct characteristics of each individual route for subsequent analysis.

6.3 Feature Selection Process

The LiRA project implemented a comprehensive feature engineering approach that utilized the Python package `tsfresh` to extract an extensive set of statistical and temporal features from aligned sensor data segments. Their methodology began with computing statistical features including mean, variance, quartile ranges, and ECDF percentiles, alongside temporal characteristics such as autocorrelation, total energy, entropy, and various signal peak features for each input sensor type. To optimize model performance, LiRA employed a Sequential Feature Selection (SFS) algorithm - a greedy iterative process that systematically evaluated individual features and selected those resulting in the best model performance as measured by chosen metrics. The SFS algorithm started with an empty feature set and progressively added features one by one through forward search mode, continuing until performance plateaued or a predetermined limit was reached. Their feature engineering process was conducted separately for different sensor combinations (vertical acceleration with speed sensors, 3D acceleration with speed sensors) using Random Forest regressors, with careful attention to overfitting prevention through parameter tuning, including minimum impurity decrease thresholds ($1.5e-2$), maximum tree depths (limited to 5), and minimum samples per leaf constraints. Additionally, Principal Component Analysis (PCA) was applied as an alternative dimensionality reduction technique to explain 99% of the variance, allowing for comparison between SFS-selected features and PCA-transformed feature sets across both training and testing datasets to ensure robust model generalization.

6.3.1 Expansion from Chapter 4, SFS Implementation

In our calibration, the Sequential Feature Selection process was applied to the comprehensive set of engineered features derived from the ten core sensor measurements: vehicle speed, roll angle, pitch angle, Y-axis acceleration, X-axis acceleration, Z-axis acceleration, total acceleration, gravitational X-axis acceleration, gravitational Y-axis acceleration, and gravitational Z-axis acceleration. These fundamental parameters possess strong physical relevance for IRI prediction, as they directly capture vehicle dynamics and road surface interactions that manifest as roughness-induced vibrations and motion patterns. Following the feature engineering process that generated rolling statistical measures, interaction terms, entropy calculations, and temporal derivatives across multiple time windows, the SFS algorithm systematically evaluated 59 potential features to identify optimal combinations for model performance. The iterative selection process, employing cross-validated RandomForest evaluation with overfitting prevention constraints, converged on 17 features that demonstrated superior predictive capability. The results revealed a predominance of speed, pitch-related features, alongside vertical acceleration. The selected features demonstrate clear physical interpretability, pitch angle variations directly reflecting road surface irregularities, vertical acceleration capturing road elevation complexity, and multi-scale temporal features providing both immediate surface responses and longer-term road condition patterns. The following tables present the selected features for the three categories of LiRA input features: speed and vertical acceleration, speed and 3D acceleration, and speed and other features. Details of these three categories and their various combinations have been presented in Chapter 4.

6.3.2 Statistical Relevance and Physical Interpretability Analysis

The selected features demonstrate strong physical relevance for IRI prediction as they directly capture the vehicle-road interaction dynamics that characterize surface roughness. Pitch angle variations and their temporal changes reflect how the vehicle responds to road undulations and surface irregularities, with smoother roads producing minimal pitch fluctuations while rough surfaces cause pronounced angular deviations. Acceleration-related features quantify the intensity and complexity of vibrations transmitted from the road surface through the vehicle's suspension system. Vehicle speed features capture the dependency of road roughness perception on travel velocity, as the same surface defects produce different dynamic responses at varying speeds due to suspension dynamics and contact mechanics. The temporal windows enable the model to distinguish between localized surface defects that cause immediate responses and distributed roughness patterns that generate sustained vehicle motion, collectively providing a comprehensive representation of the road-vehicle system's dynamic behavior that directly correlates with IRI measurements.

6.3.3 Computational Efficiency and Implementation Constraints

The practical considerations of feature selection include computational efficiency requirements and implementation constraints for real-world deployment. At the scale of our dataset of around 8,000 data points, all features were implementable and realistic to use. The analysis covers feature reduction impacts on processing speed, memory requirements, and the trade-offs between model accuracy and practical usability for large-scale pavement monitoring applications.

While LiRA utilized the automated `tsfresh` package to systematically extract hundreds of predefined statistical and temporal features from sensor data, our approach manually engineered a focused set of 59 features based on rolling window statistics, entropy measures, differential calculations, and interaction terms across multiple time scales. Our targeted methodology proved more effective for our specific dataset characteristics, as it allowed precise control over temporal window sizes that matched the dynamic response patterns in our vehicle-road system, avoided computational overhead associated with irrelevant feature combinations, and enabled creation of physically meaningful interaction terms that captured the coupled behavior between vehicle orientation and acceleration responses. Additionally, our approach reduced the risk of including spurious features that might be statistically significant but physically irrelevant to road roughness detection, resulting in a more interpretable and computationally efficient feature set that achieved comparable performance with significantly fewer features.

6.4 An Overview of the Modeling Process

The LiRA modeling framework consists of several sections, including IRI prediction, friction prediction, and energy consumption analysis. However, we only focused on the IRI prediction section, as our equipment was not capable of collecting the required data for friction predictions. Additionally, they

provide separate IRI estimations for municipal areas and motorways, which fall outside the scope of our current analysis.

For IRI prediction, LiRA has employed several regression methodologies including Multiple Linear Regression, Lasso Regression, Ridge Regression, Elastic Net, KNN, RF, SVR, and ANN. To comprehensively evaluate predictive performance on our dataset, we implemented and tested all of these established methods as baseline approaches. Additionally, we extended the analysis by incorporating more sophisticated deep learning architectures to leverage the complex temporal and spatial patterns inherent in vehicle sensor data. These advanced methods included CNN for capturing local feature patterns, LSTM networks for modeling sequential dependencies, and hybrid CNN-LSTM architectures for combining spatial feature extraction with temporal sequence modeling. This comprehensive modeling approach allowed us to systematically compare traditional machine learning methods against state-of-the-art deep learning techniques, providing insights into the optimal balance between model complexity and predictive accuracy for road quality assessment applications.

6.4.1 Data Preprocessing and Sequence Preparation

The preprocessed data underwent comprehensive preparation to make it suitable for neural network training. Following the feature engineering process, which has been explained before, a balanced sequence creation approach was implemented to address the inherent class imbalance in road condition data. This method uses weighted sampling based on IRI value distribution, applying inverse frequency weighting with additional emphasis on high IRI values (above the 75th percentile) to ensure adequate representation of poor road conditions during training. The time series data was then transformed into sequential input-output pairs using a sliding window approach, where each input sequence contains multiple consecutive time steps of sensor data to predict the subsequent IRI value. Data normalization was achieved through RobustScaler to handle outliers effectively.

6.4.2 Model Architecture and Training Strategy

The neural network architecture consists of a hybrid CNN-LSTM model designed specifically for IRI prediction from sequential sensor data. The model begins with convolutional layers for local feature extraction, featuring two Conv1D layers with 64 filters each, a kernel size of 3, followed by batch normalization and dropout operations. The extracted features are then processed through three LSTM layers with 32 units each to capture temporal dependencies and sequential patterns in the sensor data. The final prediction is made through dense layers with 64 and 32 neurons, respectively, each incorporating ReLU activation functions, L2 regularization, batch normalization, and dropout layers to prevent overfitting and ensure stable training. Training optimization utilizes the Adam optimizer with gradient clipping and incorporates several callback mechanisms, including early stopping, learning rate reduction on plateau, and model checkpointing to ensure robust convergence. The model training process includes validation splitting and employs multiple regularization techniques to achieve optimal generalization performance on unseen road condition data.

Our focus is on using neural network layers instead of traditional machine learning models. For example, the hybrid CNN-LSTM or ANN model that showed superior performance in Chapter 4, along with exploration of alternative architectures such as ensemble methods, traditional machine learning approaches (Random Forest, SVR, KNN), and deep learning variants.

6.4.3 Hyperparameter Optimization and Calibration

This subsection explains the implementation of transfer learning techniques to leverage the adaptation of pre-trained Danish LiRA models to Minnesota roadway conditions requires systematic transfer learning protocols. The process includes parameter extraction from the Danish framework, weight transfer strategies, and fine-tuning approaches to adapt the trained models to our data structure and vehicle conditions while maintaining beneficial learned features. Hyperparameter optimization was conducted through systematic grid search and Bayesian optimization techniques, specifically tuning learning rates from 0.0001 to 0.01, regularization parameters (L1: 0.001-0.1, L2: 0.001-0.1), batch sizes (16, 32, 64), dropout rates (0.2-0.5), and network architecture parameters including hidden layer dimensions (64, 128, 256) and sequence lengths (5, 10, 15). The optimization process employed with early stopping (patience=10) and learning rate reduction callbacks. The training and validation approach utilized stratified data splitting across the three test loops (MN-36, Metro to Northfield, and MnROAD), ensuring generalization capability and techniques to handle diverse road conditions encountered across different test environments, thereby maintaining model robustness across varied Minnesota roadway conditions.

6.4.4 Model Performance Evaluation Framework

This subsection outlines the comprehensive evaluation framework used to assess model performance and guide the recalibration process. The framework includes multiple evaluation metrics (RMSE, MAE, MRE), statistical significance testing, and validation approaches specifically designed to ensure the recalibrated models meet the accuracy requirements for Minnesota's pavement monitoring applications. The performance comparison of all evaluated regression methods is presented in the Tables of Section 5.6.

6.5 Model Quality and Performance Analysis

6.5.1 Statistical Accuracy Assessment

This subsection presents a comprehensive statistical evaluation of the recalibrated models using traditional error metrics, including RMSE, MAE, and MRE. The analysis includes a comparison of training versus testing performance, examination of prediction residuals, and statistical significance testing to validate model accuracy. Performance distributions and confidence intervals provide insights into prediction reliability across the dataset for each input feature category. First, our predicted performance is presented, followed by the LiRA performance results.

6.5.1.1 Speed-Vertical acceleration

Table 25. Performance comparison of our proposed models with speed and vertical acceleration as input features.

Method	MRE	MAE	RMSE
Lasso	0.773	0.633	0.776
Ridge	0.783	0.650	0.819
Elastic Net	0.762	0.626	0.770
KNN	0.758	0.672	0.918
RF	0.637	0.533	0.67
SVR	0.827	0.664	0.818
ANN	0.837	0.686	0.859
CNN-LSTM	0.837	0.649	0.852

Table 26. Performance comparison of LiRA proposed models with speed and vertical acceleration as input features. (LiRA, 2023)

Method	MRE	MAE	RMSE
Lasso	0.20	0.62	0.91
Ridge	0.20	0.60	0.86
Elastic Net	0.20	0.58	0.82
KNN	0.19	0.53	0.77
RF	0.23	0.57	0.76
SVR	0.26	0.79	1.06
ANN	0.28	0.83	1.10

6.5.1.2 Speed-3D acceleration

Table 27. Performance comparison of our proposed models with speed and 3D acceleration as input features.

Method	MRE	MAE	RMSE
Lasso	0.302	0.278	0.385
Ridge	0.287	0.268	0.373
Elastic Net	0.299	0.275	0.382
KNN	0.470	0.454	0.705

Method	MRE	MAE	RMSE
RF	0.209	0.199	0.288
SVR	0.481	0.416	0.524
ANN	0.266	0.232	0.315
CNN-LSTM	0.312	0.531	0.842

Table 28. Performance comparison of LiRA proposed models with speed and 3D acceleration as input features. (LiRA, 2023)

Method	MRE	MAE	RMSE
Lasso	0.23	0.24	0.30
Ridge	0.21	0.23	0.30
Elastic Net	0.21	0.23	0.30
KNN	0.16	0.17	0.22
RF	0.16	0.16	0.21
SVR	0.15	0.15	0.20
ANN	0.15	0.16	0.21

6.5.1.3 Speed-acceleration and other features

Table 29. Performance comparison of our proposed models with speed and other features as input features.

Method	MRE	MAE	RMSE
Lasso	0.29	0.268	0.343
Ridge	0.277	0.279	0.394
Elastic Net	0.287	0.271	0.351
KNN	0.418	0.395	0.544
RF	0.18	0.169	0.224
SVR	0.186	0.177	0.226
ANN	0.146	0.136	0.182
CNN-LSTM	0.131	0.128	0.172

Table 30. Performance comparison of LiRA proposed models with speed and other features as input features. (LiRA, 2023)

Method	MRE	MAE	RMSE
Lasso	0.20	0.28	0.33
Ridge	0.18	0.26	0.31
Elastic Net	0.20	0.29	0.33
KNN	0.17	0.26	0.34
RF	0.20	0.27	0.31
SVR	0.18	0.26	0.31
ANN	0.33	0.44	0.48

Based on the comprehensive analysis of different input feature combinations and machine learning models, several key insights emerge regarding model performance and feature importance. With speed and 3D acceleration as input features, Random Forest achieves the best performance (MRE: 0.209, MAE: 0.199, RMSE: 0.288), followed closely by ANN (MRE: 0.266, MAE: 0.232, RMSE: 0.315). When using speed and vertical acceleration, Random Forest again demonstrates superior performance (MRE: 0.637, MAE: 0.533, RMSE: 0.67), while other models show significantly degraded performance with MRE values exceeding 0.75. For speed combined with other features, CNN-LSTM emerges as the top performer (MRE: 0.131, MAE: 0.128, RMSE: 0.172), followed by ANN (MRE: 0.146, MAE: 0.136, RMSE: 0.182). The superior performance of speed with other features as input, compared to other categories, demonstrates the importance of two additional features besides speed and vertical acceleration: roll and pitch, which represent the rotational movements of the vehicle around its longitudinal and lateral axes, respectively. These orientation parameters capture critical vehicle dynamics that reflect road surface irregularities and contribute significantly to accurate IRI prediction. Overall, the speed-other features combination yields the best results, suggesting that engineered features derived from speed and total acceleration, combined with vehicle orientation data, provide more valuable information than raw acceleration components alone. The best performing model across all scenarios is CNN-LSTM with speed and other features as input, achieving the lowest error metrics.

Comparing our results with the LiRA proposed models for both motorways and municipalities, our performance demonstrates competitive accuracy, particularly with speed-other features and speed-3D acceleration combinations. For speed-3D acceleration, our Random Forest model (MRE: 0.209) is close to LiRA's best model (SVR with MRE: 0.15). With speed-vertical acceleration, our Random Forest (MRE: 0.637) has the best performance, while for speed-other features, our CNN-LSTM (MRE: 0.131) and ANN (MRE: 0.146) models achieve comparable or better performance than LiRA's top results. However, there's an interesting divergence in optimal model selection: while LiRA identifies RF and SVR as their best performer across most scenarios, our analysis reveals that CNN-LSTM, RF and ANN specially with speed-other features provides the optimal combination, suggesting that deep learning approaches can

effectively leverage engineered features when properly configured, potentially capturing more complex temporal patterns in the IRI prediction task. We compare the recalibrated framework's performance against baseline models, original Danish LiRA performance, and alternative machine learning approaches tested during development. The comparative analysis of quantitative performance metrics provides context for the achieved improvements and validation of the recalibration approach for Minnesota conditions.

In the following figures, we can see the performance of our best model from Chapter 4, CNN-LSTM with speed and other features as input features, on testing and training datasets separately, and also over epochs. Figure 55 shows the predicted versus actual IRI values over test examples. As can be seen, the predicted line closely follows the actual values in most cases, demonstrating strong model performance and reliable IRI prediction capability.

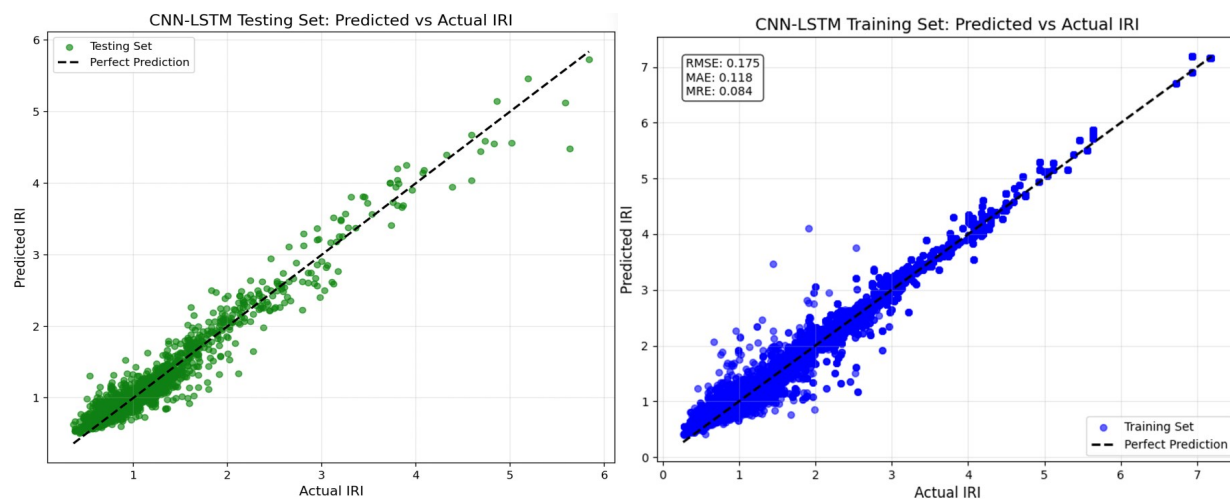


Figure 53. The training and testing performance of the CNN-LSTM model on all datasets, with speed and other features as input features.

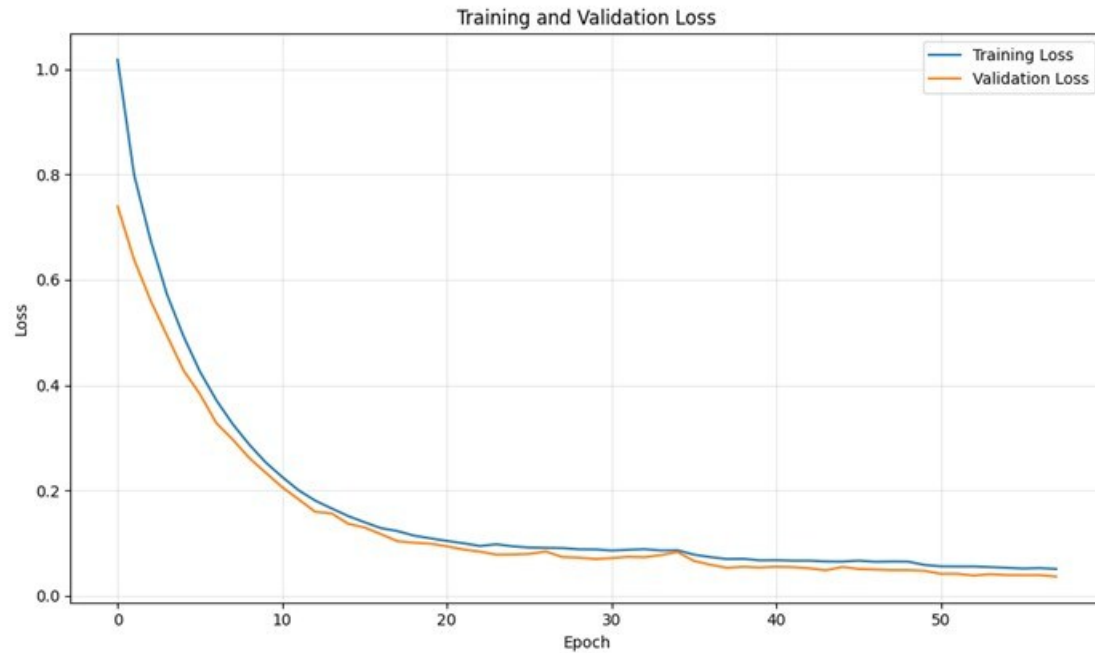


Figure 54. The training and validation loss vs epochs for the CNN-LSTM model on all datasets, with speed and other features as input features.

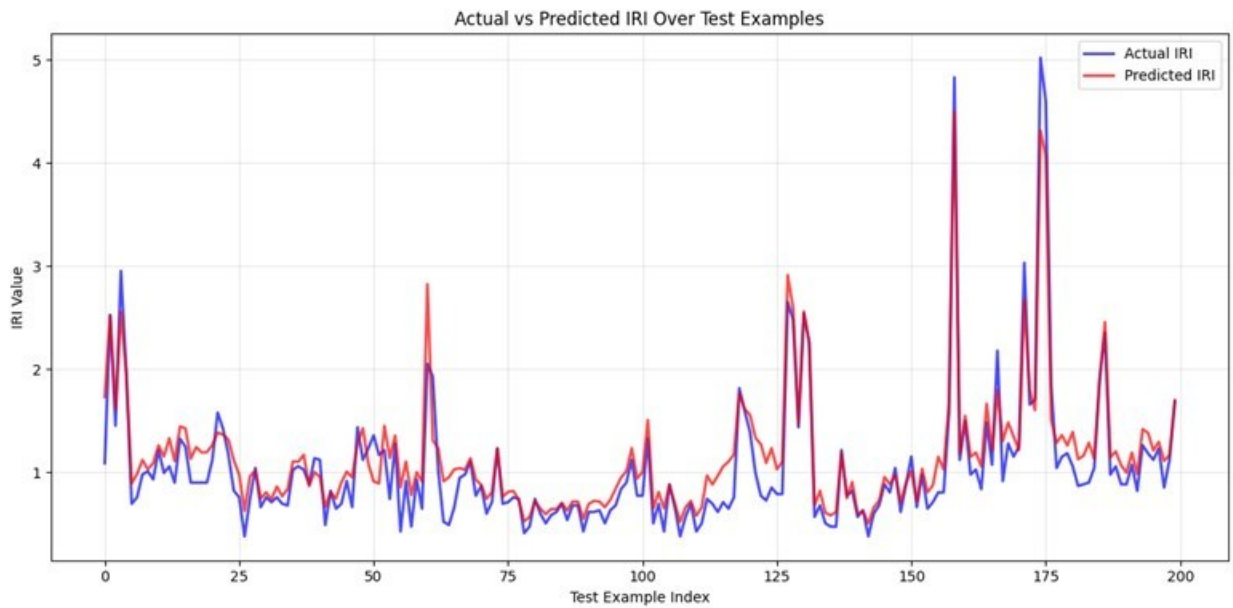


Figure 55. Actual and predicted IRI from the CNN-LSTM model output on all datasets, with speed and other features as input features.

6.5.2 Performance Evaluation Across Speed Conditions

Model performance was evaluated under controlled speed conditions at MnROAD (20, 40, 60, and 80 mph). This analysis provides insights into how the framework responds to different driving patterns and identifies optimal speed ranges for prediction accuracy. Since MnROAD testing involved constant speeds rather than variable driving conditions, the speed feature becomes less informative as an input parameter, resulting in slightly reduced model performance compared to real-world scenarios with dynamic speed variations. The evaluation then extends to examining model performance across three distinct test environments: the controlled MnROAD facility, the state highway conditions of MN-36, and the municipal road network from Metro to Northfield. This comprehensive analysis assesses whether specific pavement characteristics systematically enhance or diminish prediction accuracy and evaluates the model's ability to generalize across Minnesota's diverse pavement infrastructure and varying traffic conditions.

6.6 Prediction Results Comparison with Reference Data

6.6.1 Model Performance Against MnDOT Ground Truth

This section presents a comparative analysis between model predictions and MnDOT's high-precision Pathways van measurements, serving as the primary ground truth validation. The analysis includes overall performance across different IRI ranges, and color-coded maps of the three loops showing the percentage difference between our predicted IRI values and MnDOT's true data. Figures 56, 57 and 58 illustrate the true reference IRI values from MnDOT for the MN-36 and Metro to Northfield, and MnROAD loops, respectively. The color-coding on the map shows low to high values from blue to red shades. Figure 59 is the same color-coded map of Metro to Northfield, but this time with our predicted IRI values.

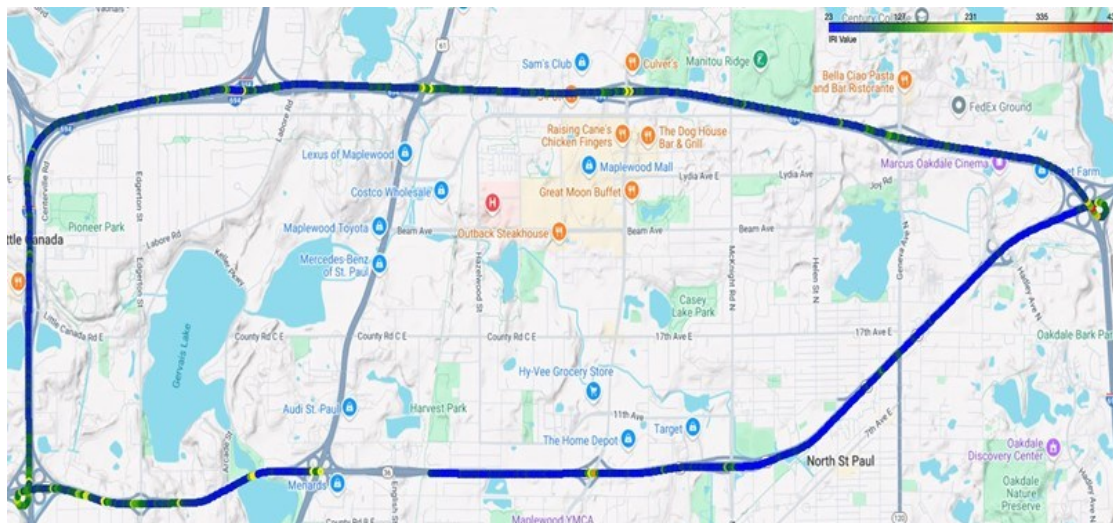


Figure 56. True IRI color-coded map of MN-36 loop.

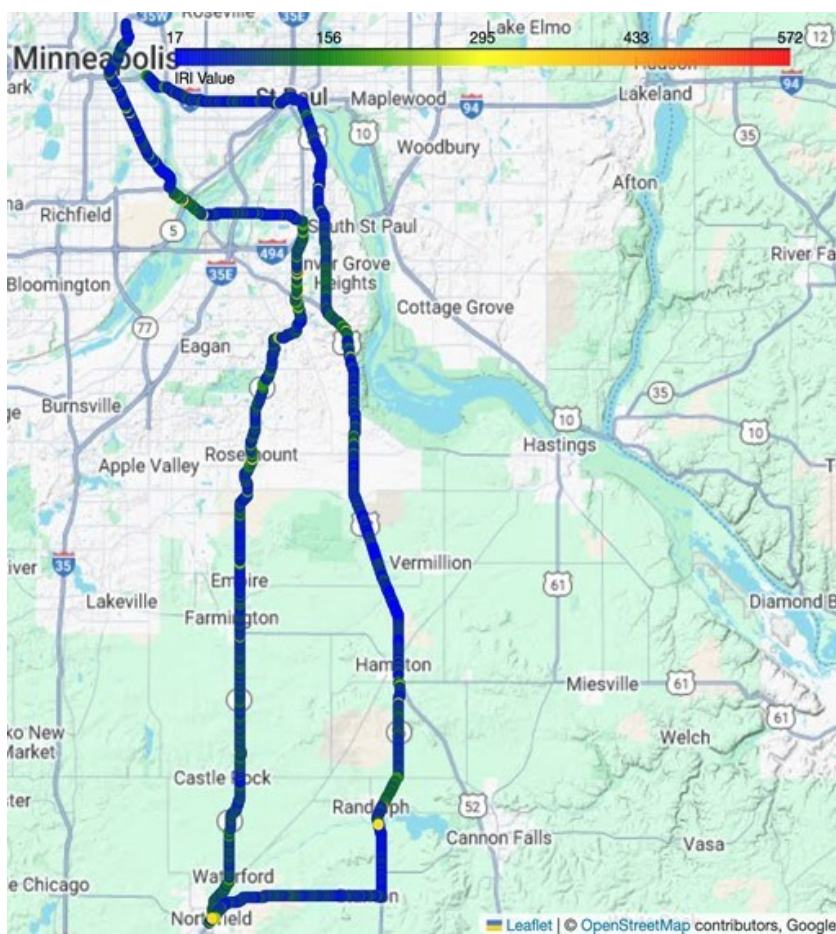


Figure 57. True IRI color-coded map of Northfield loop.



Figure 58. Color-coded visualization of true IRI measurements along the inside loop.

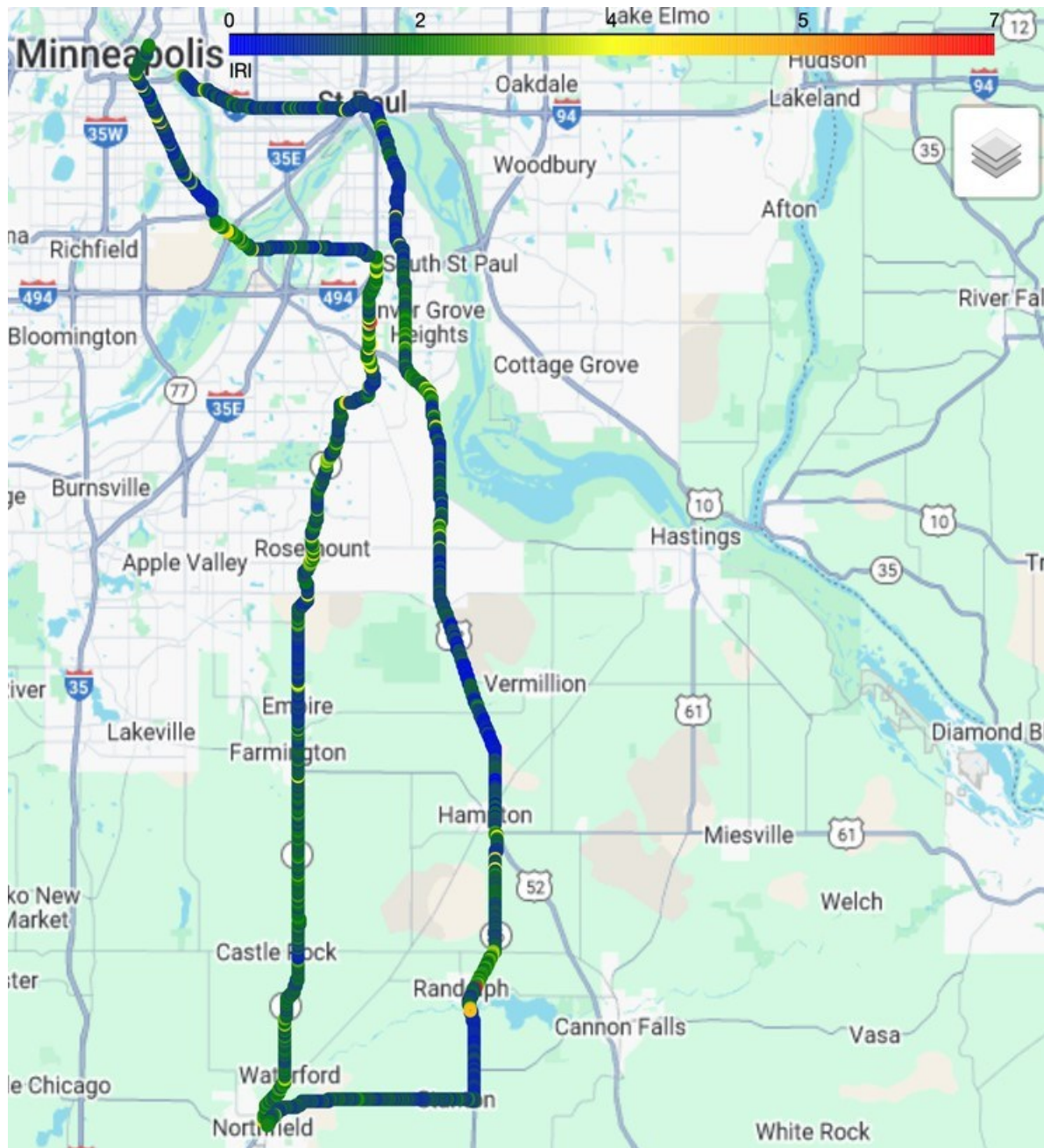


Figure 59. Predicted IRI color-coded map of Northfield loop.

In the following figures, instead of IRI values, we have illustrated the color-coded maps of the percentage difference between our predictions and MnDOT reference values in Figures 60, 61, and 62 for Metro to Northfield, MN-36, and MnROAD loops, respectively.

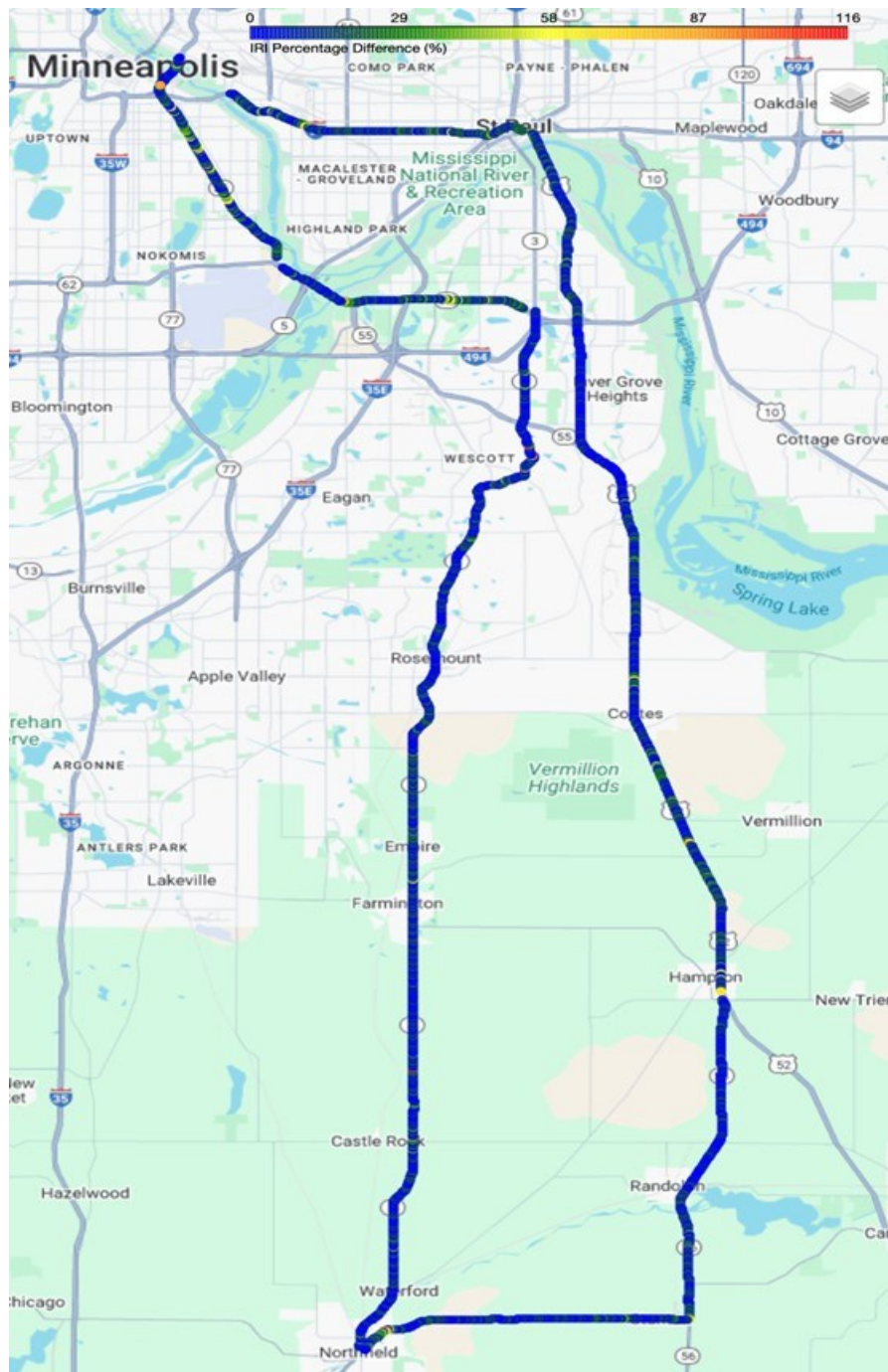


Figure 60. Percentage difference of our IRI predictions with MnDOT's ground truth measurements for the Metro to Northfield loop.

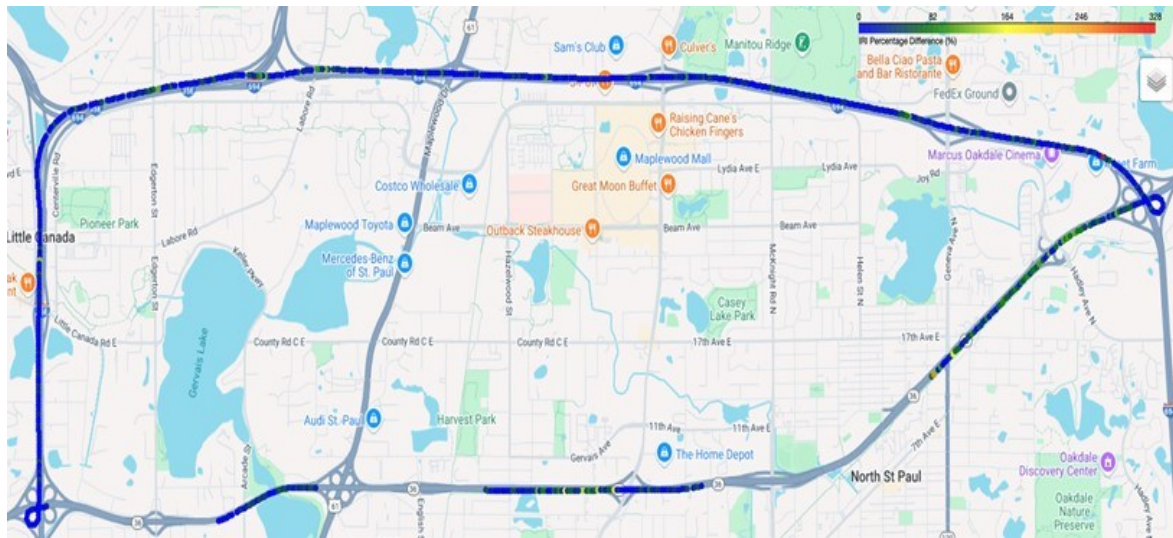


Figure 61. Percentage difference of our predictions with MnDOT’s ground truth measurements for the MN-36 loop.

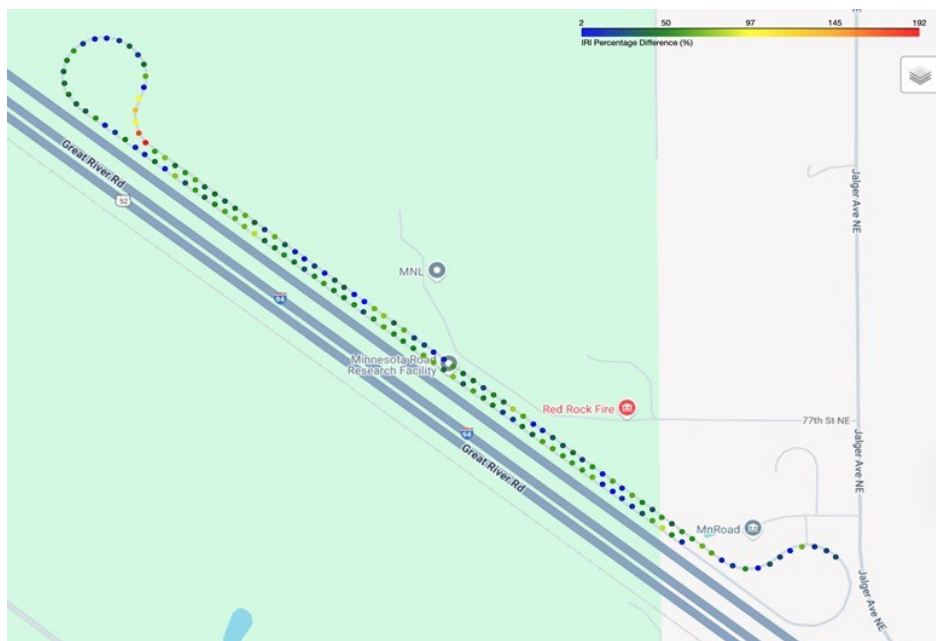


Figure 62. Percentage difference of our predictions with MnDOT’s ground truth measurements for the inside loop of MnROAD, with a constant speed of 40 mph.

As illustrated from the comparison results of IRI prediction, we can observe that the proposed model is able to predict IRI with high accuracy. However, like any other existing model, we are not without errors. As shown in the results, we have errors in points with colors other than blue, ranging from green and yellow to red, but the number of these points with high differences between predicted and true values is relatively small.

We can specify the performance across three different datasets. As observed, the highest error percentages are predominantly seen in the MnROAD loop compared to the two other datasets. This can be attributed to several factors: first, the MnROAD dataset did not have the influential speed factor that was present in other datasets. Additionally, some points (shown in red) in the MnROAD dataset are experiencing abrupt changes in z-acceleration and other sensor features, although the corresponding IRI values are not changing proportionally. Consequently, the model is predicting fluctuating high IRI values for these points. This discrepancy can be attributed to the fact that the model was trained on datasets where sudden changes in acceleration features were typically associated with higher IRI values. When the model encounters similar acceleration patterns in the MnROAD data, it predicts high IRI values even when the actual road conditions (and thus IRI values) remain relatively stable. This suggests that the relationship between sensor features and IRI may vary across different road types, data collection conditions, or geographic locations, highlighting the importance of dataset diversity in training robust IRI prediction models.

Based on the comprehensive analysis and implementation of the Danish LiRA methodology for Minnesota's pavement monitoring conditions, this study successfully demonstrates the feasibility and effectiveness of using onboard vehicle data for IRI prediction across diverse road environments. The systematic adaptation process, encompassing data cleaning, feature engineering, and model recalibration, resulted in a robust framework capable of achieving high prediction accuracy with CNN-LSTM and ANN architectures. The optimal feature combination of speed with other engineered features (including roll, pitch, and various acceleration parameters) yielded the best performance, with the CNN-LSTM model achieving remarkably low error metrics (MRE: 0.131, MAE: 0.128, RMSE: 0.172). This performance is competitive with the original Danish LiRA results, validating the successful transfer and adaptation of the methodology to Minnesota's road conditions, climate, and vehicle characteristics.

The successful integration of GIS-based spatial alignment methods with advanced deep learning architectures provides a practical alternative to more computationally intensive signal correlation approaches, making the methodology accessible for broader implementation across Minnesota's transportation network. Future research directions should focus on expanding the dataset to include additional road types, seasonal variations, and vehicle configurations to further enhance model robustness, while exploring real-time deployment strategies that could enable continuous, fleet-based pavement monitoring to support proactive infrastructure management and maintenance optimization.

Chapter 7: Guideline for Large-scale Implementation

This chapter establishes a framework for scaling the validated IRI prediction methodology from controlled test environments to statewide implementation across Minnesota's diverse road network. The primary objective involves developing robust guidelines that address the practical challenges of large-scale implementation, including the integration of alternative reference standards, validation protocols for different data availability scenarios, and cost-effective strategies for continuous pavement monitoring. The chapter encompasses comparative analysis between high-precision MnDOT measurements and network-level NIRA data to establish multiple validation pathways, ensuring the framework's adaptability to varying data availability conditions across different regions and road classifications within Minnesota's transportation infrastructure.

7.1 NIRA Data as Alternative Reference Standard

The validation of our OBD-II-based IRI prediction methodology extends beyond comparison with MnDOT's high-precision Pathways van measurements, to include NIRA data as an alternative reference standard. This dual validation approach is crucial for large-scale implementation, as high-precision van measurements are not available for every road segment across Minnesota's extensive transportation network.

7.1.1 Comparative Analysis: MnDOT vs. NIRA Reference Data

Figures 63-66, which have been obtained from the NIRA application, show the MnDOT ground truth measurements and NIRA predicted IRI points color-coded on the map for both the Metro to Northfield and MN-36 loops. We can see that NIRA predictions demonstrate excellent agreement with the reference data, with most points showing strong correlation despite some localized differences and occasional gaps in coverage.

The spatial patterns and color distributions show excellent alignment between the predicted and reference values across most of the loop. The general trends in pavement condition - from the smoother sections (lighter colors) to the rougher areas (darker/red sections) - are consistently captured by the NIRA model.

While there are some localized differences where individual segments show slight variations in IRI values between the two datasets, and some segments appear to have gaps in the NIRA predictions, these discrepancies are relatively minor compared to the overall strong correlation. The NIRA predictions successfully identify the same problem areas that show up in the MnDOT reference data, and the magnitude of roughness values generally corresponds well across both visualizations. Their model appears to be capturing the fundamental relationships between road conditions and IRI values

effectively, making it a valuable tool for pavement management despite the minor gaps and variations present in some locations.

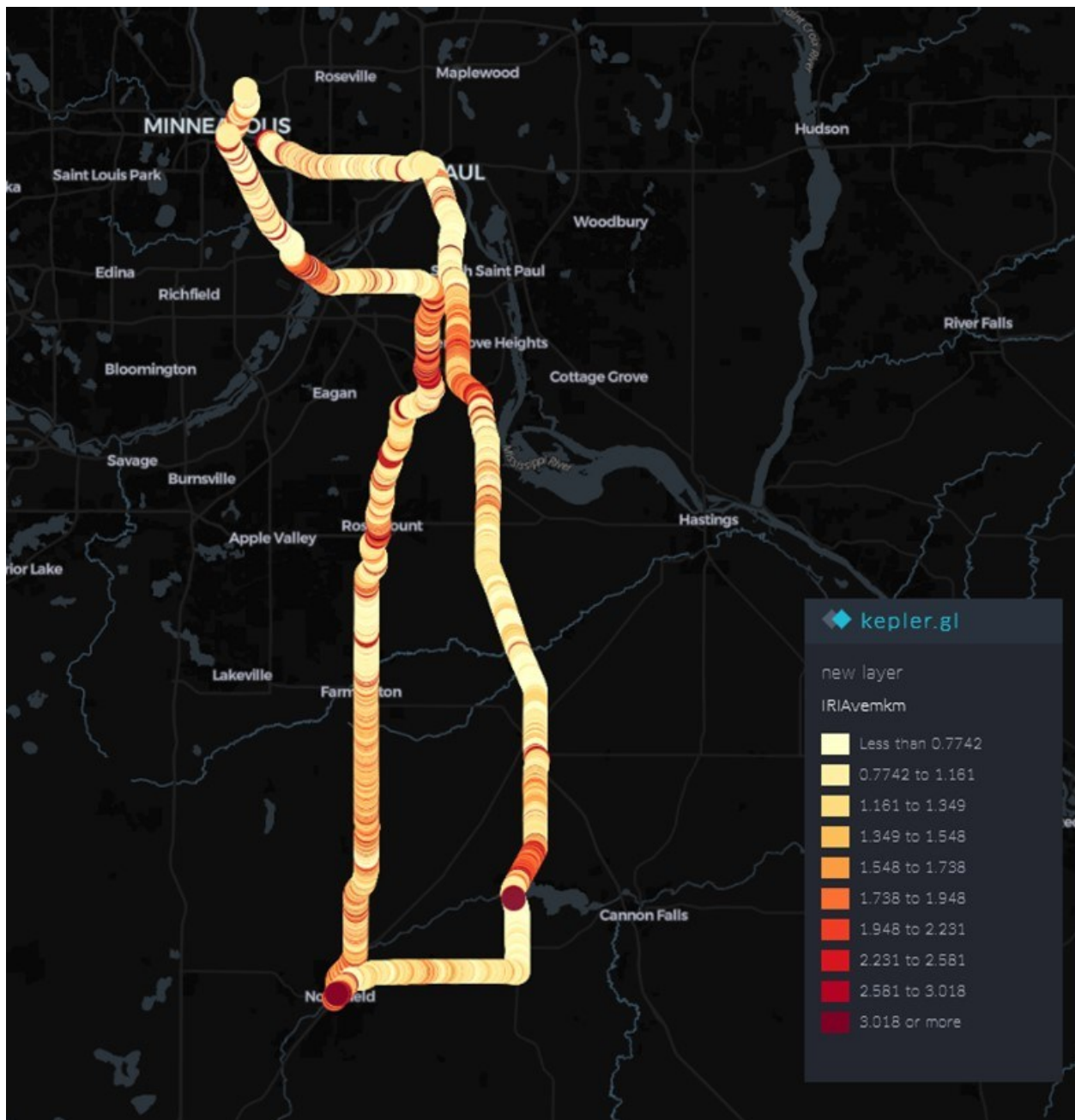


Figure 63. IRI color-coded map of Metro to Northfield loop, MnDOT ground truth measurements.

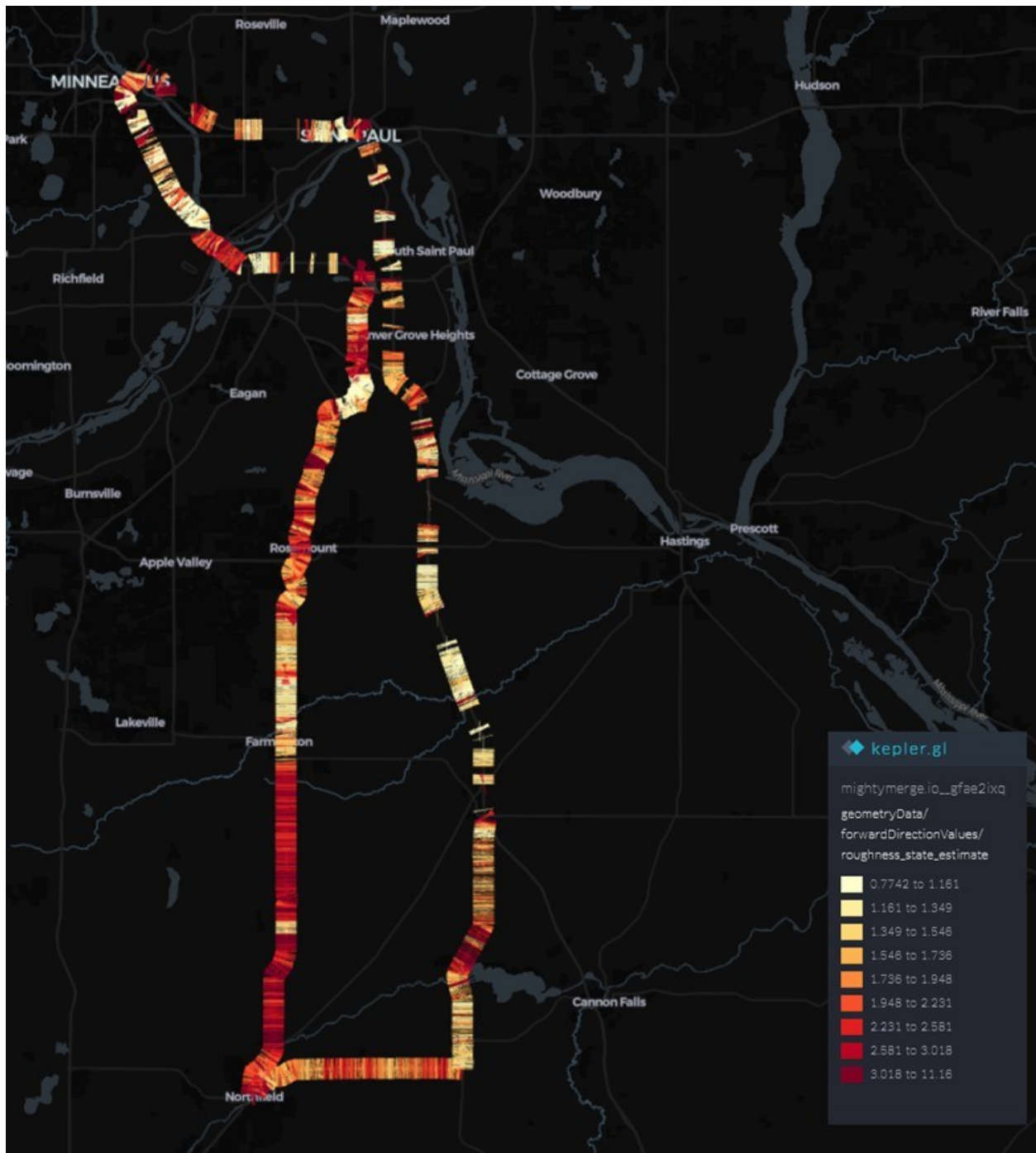


Figure 64. IRI color-coded map of Metro to Northfield loop, NIRA predictions.

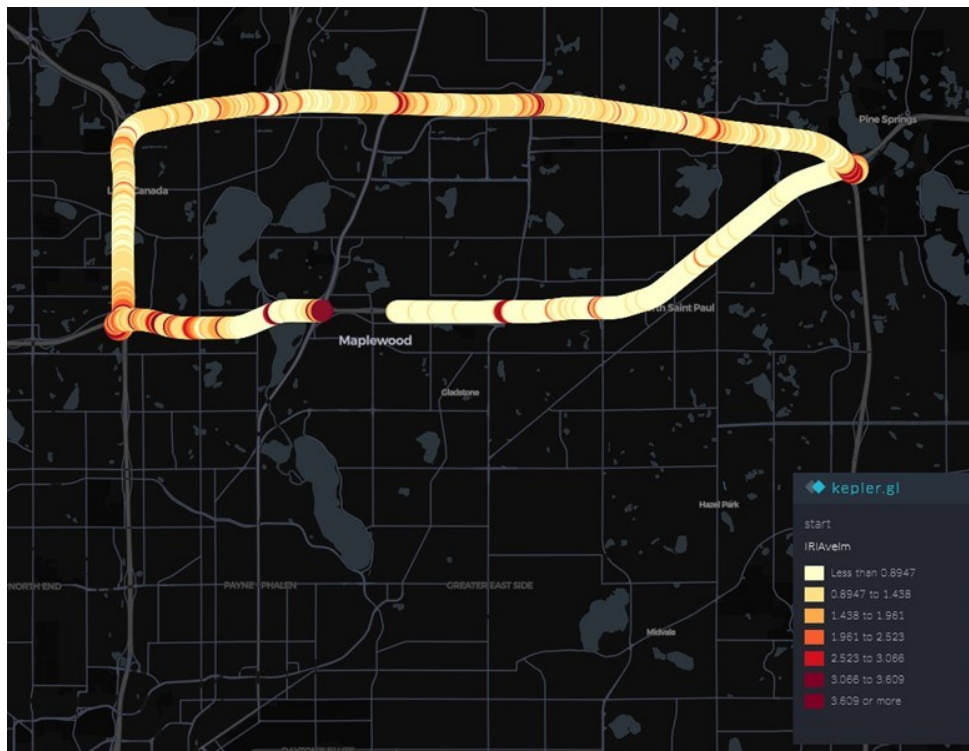


Figure 65. IRI color-coded map of MN-36 loop, MnDOT ground truth measurements.



Figure 66. IRI color-coded map of MN-36 loop, NIRA predictions.

Since we aimed to compare NIRA predictions with our predictions and MnDOT ground truth measurements ourselves, we obtained the NIRA estimations from them for the desired loops. Since NIRA sampling was denser than that of the MnDOT van, and the segmentation locations were not identical, we performed map matching in ArcGIS to find the closest NIRA point to each MnDOT point and averaged them to have one NIRA prediction for each MnDOT (and our) prediction data point. This map matching and location point alignment process slightly reduced the accuracy of our comparison compared to their own application comparison, which is illustrated in Figures 63-66. The systematic comparison between MnDOT Pathway van measurements and NIRA data across our test loops reveals key considerations for reference-data selection. Although MnDOT measurements offer superior precision, NIRA provides broader spatial coverage and more frequent updates, making it ideal for continuous monitoring.

In Figures 67 and 68, you can find the percentage difference between our model's predictions and NIRA predictions and the percentage difference between the NIRA model's predictions and MnDOT ground truth measurements across the Metro to Northfield loop, and in Figures 69 and 70, the same results for the MN-36 loop.

Our comparative analysis reveals that NIRA predictions show strong alignment with our OBD-II-based predictions, particularly in the MN-36, where about 55% of our predictions lie within $\pm 30\%$ of the NIRA values, and nearly 81% fall within $\pm 50\%$, with the largest discrepancies occurring in segments where pavement conditions change rapidly or localized surface defects introduce temporary variations.

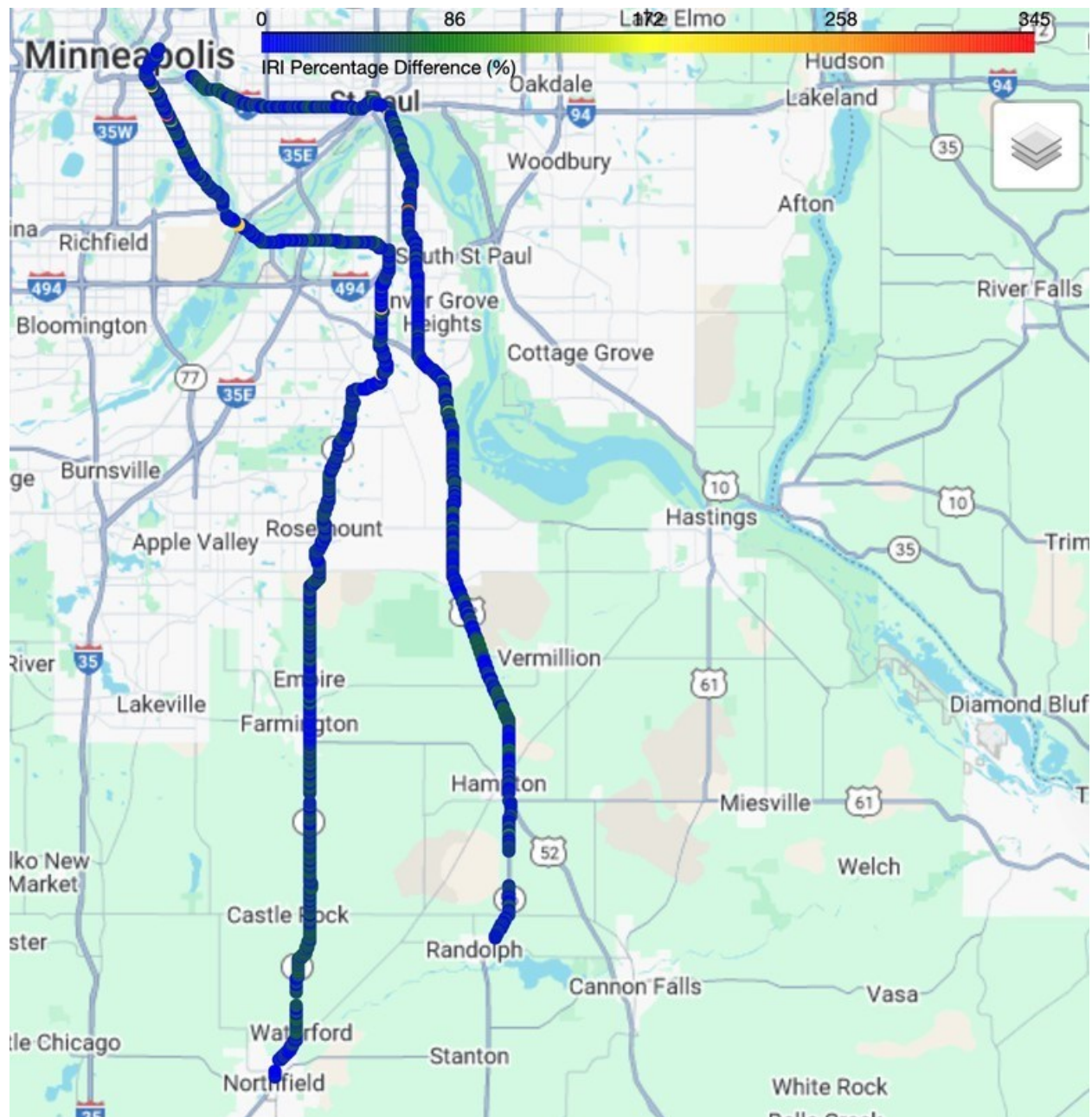


Figure 67. Percentage difference between our model's predictions and NiRA predictions across the Metro to Northfield loop.

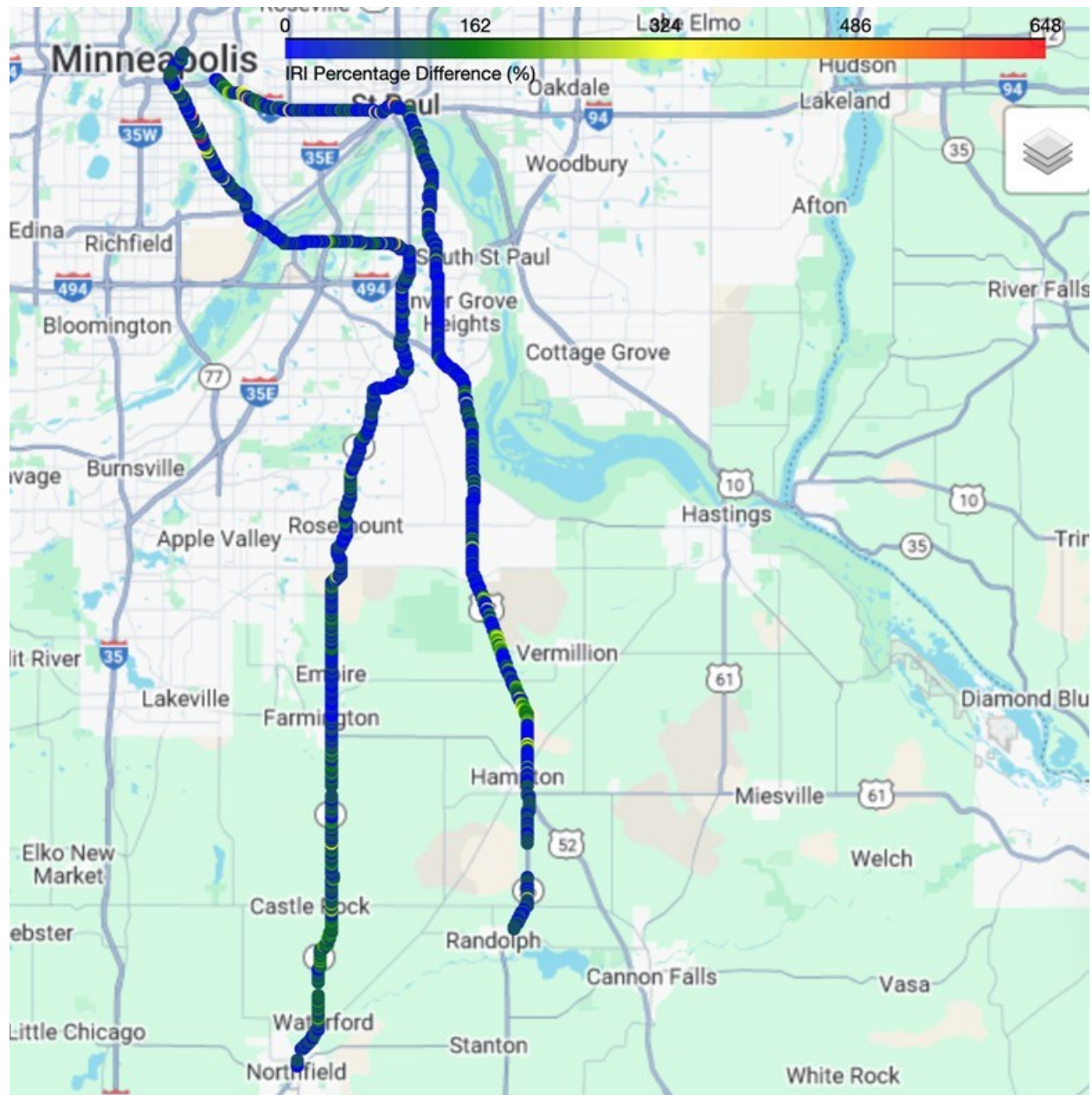


Figure 68. Percentage difference between the NIRA model's predictions and MnDOT ground truth measurements across the Metro to Northfield loop.

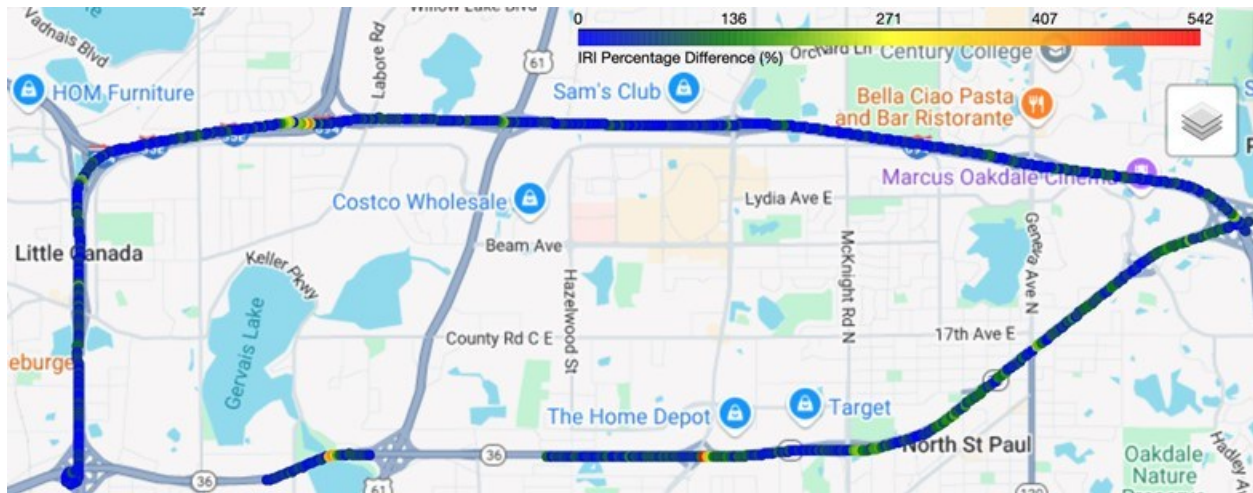


Figure 69. Percentage difference between the NIRA model's predictions and MnDOT ground truth measurements across the MN-36 loop.

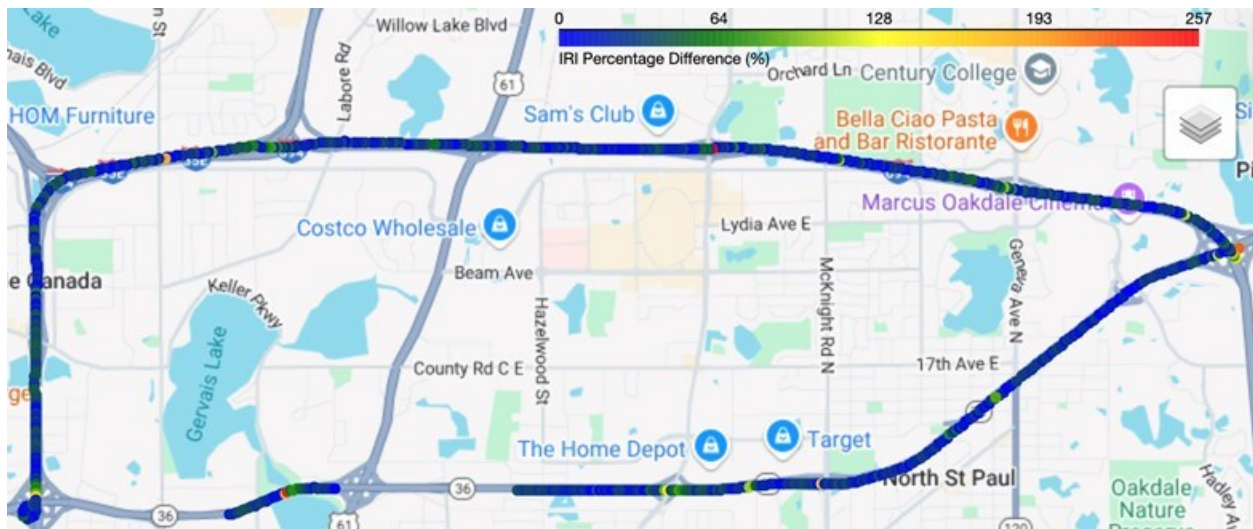


Figure 70. Percentage difference between our model's predictions and NIRA predictions across the MN-36 loop.

Based on the figures, we can see a percentage difference in all cases. One reason for this is that the positional data of NIRA and the data collection frequency were not identical with MnDOT or with our predictions. What we have done is average between the closest points to each MnDOT location to have NIRA estimations at the nearest point to ours and MnDOT's for comparison. Part of the errors can be attributed to this methodology. Another factor is that the timing, weather conditions, and many other variables have not been the same between us, MnDOT, and NIRA, so these percentage differences are completely expected and make sense; some errors are unavoidable. However, most points show very low differences between our predictions and NIRA predictions (blue points), which provides additional validation of our results.

Additionally, one source of NIRA prediction errors comes from the fact that they used their own ground truth measurements for training and prediction instead of MnDOT ground truth measurements, which

can obviously differ from MnDOT due to many factors, like timing and weather conditions. Therefore, for a better illustration of NIRA's ability to predict IRI, it may be better to compare their estimations with their own ground truth measurements, which we illustrate in Figures 71 and 72.

It should be noted that standardized driving procedures, such as those followed by MnDOT reference data collection, can also influence measurement consistency. According to our discussions, MnDOT operators are instructed to maintain the posted speed limit while driving exclusively in the first lane on multi-lane roads, consistently positioning the vehicle at the center of the lane. However, the extent to which these guidelines are strictly adhered to remains uncertain, which may introduce variability in reference measurements.

As the comparison of NIRA estimations with their own ground truth measurements shows, the errors are much less, demonstrating their strong performance, with just some errors in specific sections where yellow and green segments indicate moderate percentage differences.

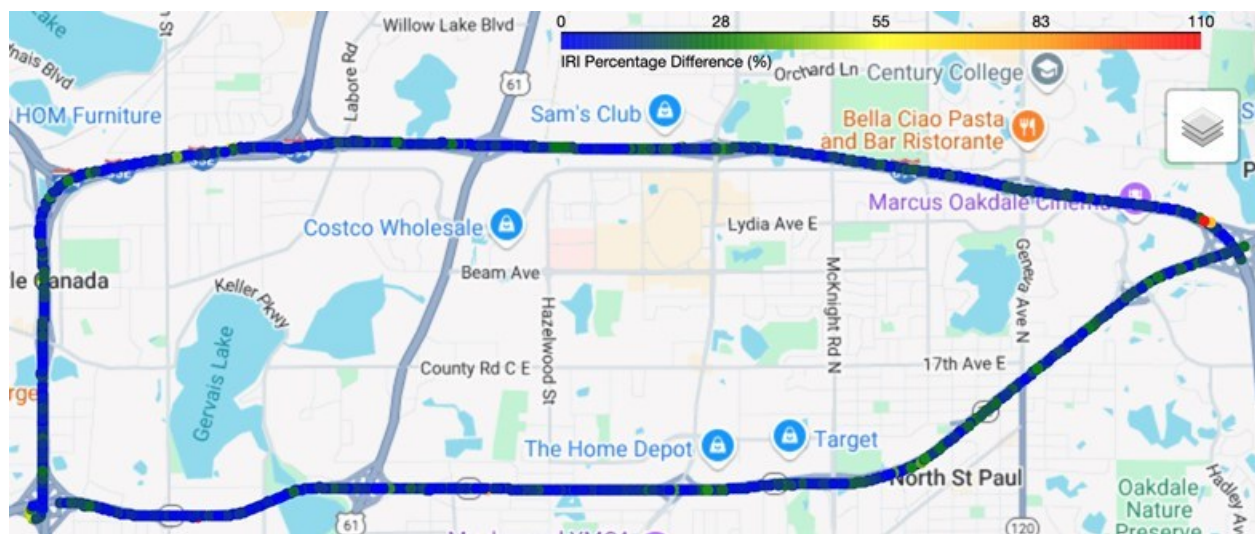


Figure 71. IRI percentage comparison between NIRA ground truth measurements and NIRA predictions for the MN-36 loop.

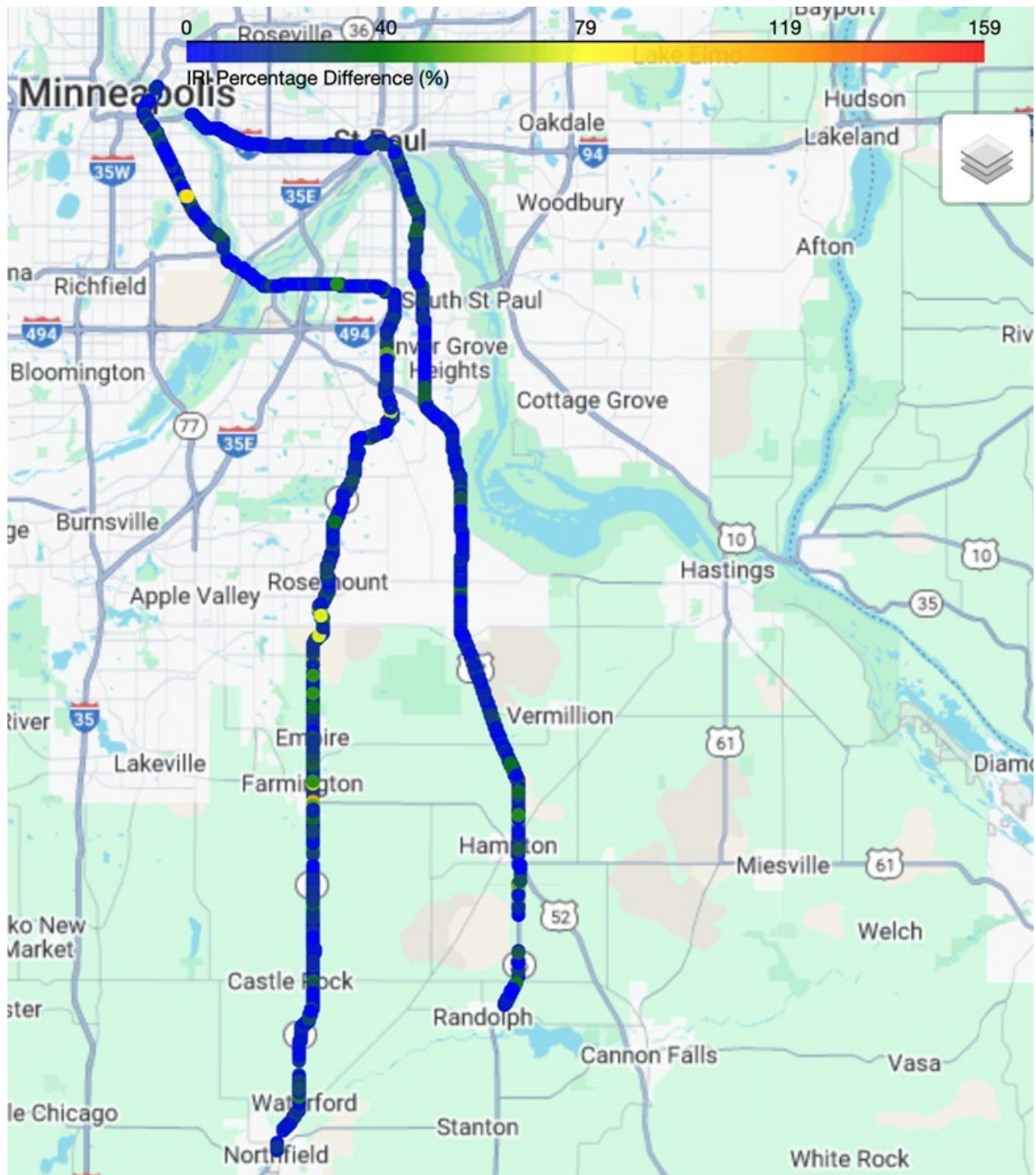


Figure 72. IRI percentage comparison between NIRA ground truth measurements and NIRA predictions for the Metro to Northfield loop.

7.1.2 Model Prediction Accuracy Across Different Reference Standards

Geographic analysis reveals that our predictions show strong agreement with NIRA data on highway segments MN-36 loop and slightly lower agreement on mixed municipal roads (Metro to Northfield).

This variation pattern aligns with the expected precision differences between controlled highway conditions and more variable urban driving environments. The dual validation approach establishes confidence in deploying our methodology across Minnesota's diverse road network, even in areas where high-precision reference measurements are unavailable or infrequent. The strong correlation with NIRA data provides a pathway for continuous validation and model refinement during large-scale implementation.

7.1.3 NIRA IRI Estimation Comparison: Daily and Seasonal Variations

Since NIRA relies on in-vehicle data collected continuously, it is possible that seasonable variability in pavement roughness (or even day-to-day fluctuations) could be observed in the NIRA pavement roughness estimates. Thus, we compare the estimated IRI from between different days and seasons. We consider a section of the aforementioned MN-36 loop (MN-36 highway from Maplewood Drive to the I-694 exit) for examination.

7.1.3.1 Analysis of Daily IRI Variation

For daily comparison of how NIRA estimations change from one day to the next, we selected two consecutive days in March (March 2nd and 3rd) and extracted IRI estimations for this section of MN-36 from the NIRA website. For better visualization of IRI estimations, color-coded maps of IRI estimations for these two days have been created and are illustrated in Figures 73 and 74.

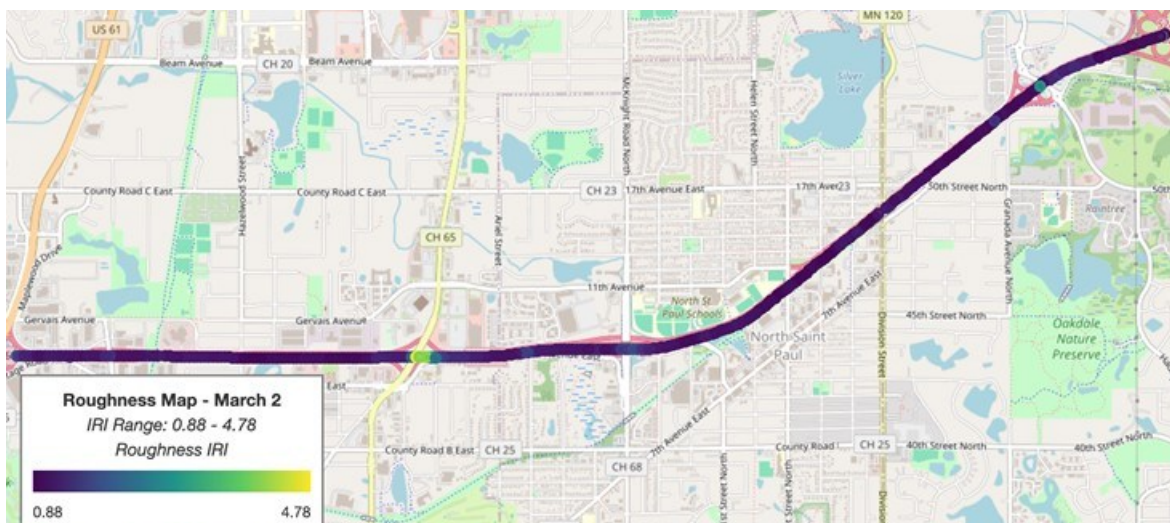


Figure 73. IRI color-coded map of a portion of MN-36 - March 2nd.

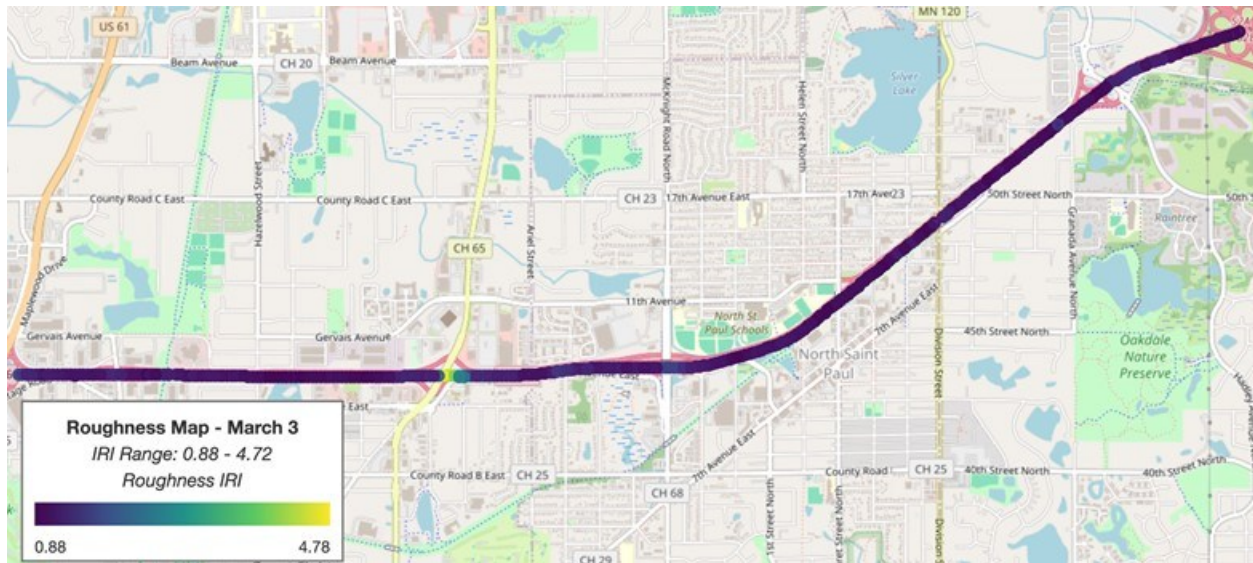


Figure 74. IRI color-coded map of a portion of MN-36 - March 3rd.

Based on these two consecutive IRI maps from March 2nd and 3rd, the road surface condition measurements show remarkable consistency between the two days, with both maps displaying nearly identical color patterns and spatial distribution of roughness values. While the maximum IRI range decreased slightly from 4.78 on March 2nd to 4.72 on March 3rd, this minimal difference of 0.06 units represents less than 1.3% variation, which falls well within the expected measurement precision and uncertainty of IRI collection systems. The virtually identical color coding and spatial patterns across both maps suggest that the road surface conditions remained stable over this 24-hour period, which is expected given that pavement roughness typically changes gradually over months or years rather than daily, demonstrating that IRI estimations collected on consecutive days should be nearly identical under normal conditions. These minor variations could be attributed to several factors including slight differences in vehicle speed during data collection, environmental conditions affecting the measurement equipment, GPS positioning accuracy variations, or normal sensor noise inherent in the IRI measurement process, rather than actual changes in pavement condition.

7.1.3.2 Analysis of Seasonal IRI Variation

To examine seasonal variations in pavement roughness, data from the 2nd day of four representative months were analyzed: November 2nd, 2024 (fall), January 2nd, 2025 (winter), March 2nd, 2025 (spring), and July 2nd, 2025 (summer). The estimated IRI values for the same section of MN-36 have been extracted for all four dates, and color-coded IRI maps are presented in Figures 75-78.

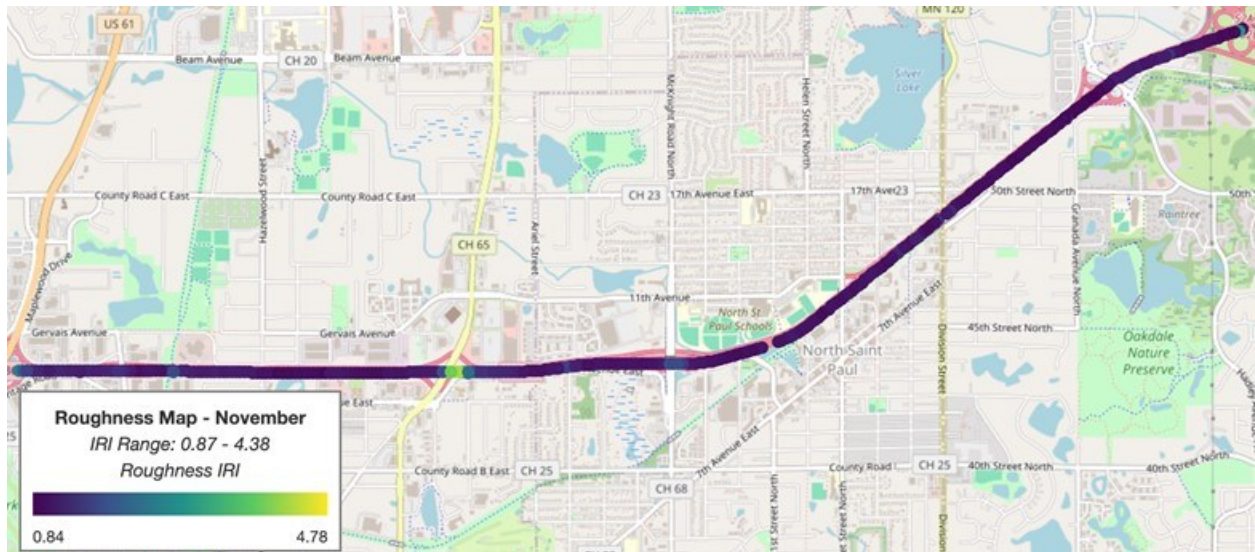


Figure 75. IRI color-coded map of a portion of MN-36 – November 2024.

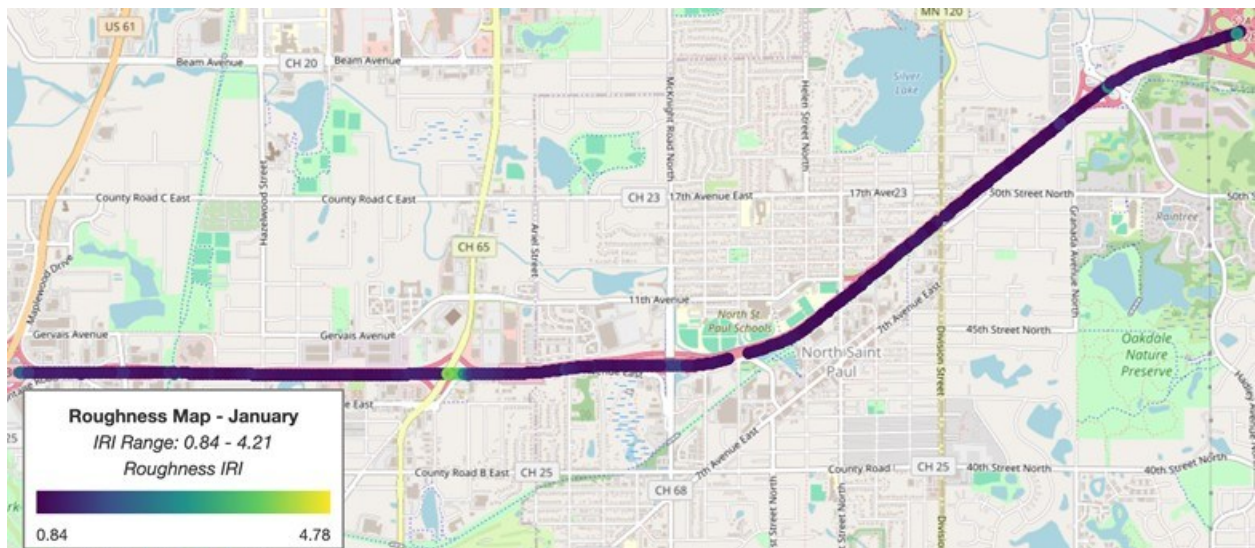


Figure 76. Figure 14. IRI color-coded map of a portion of MN-36 – January 2025.

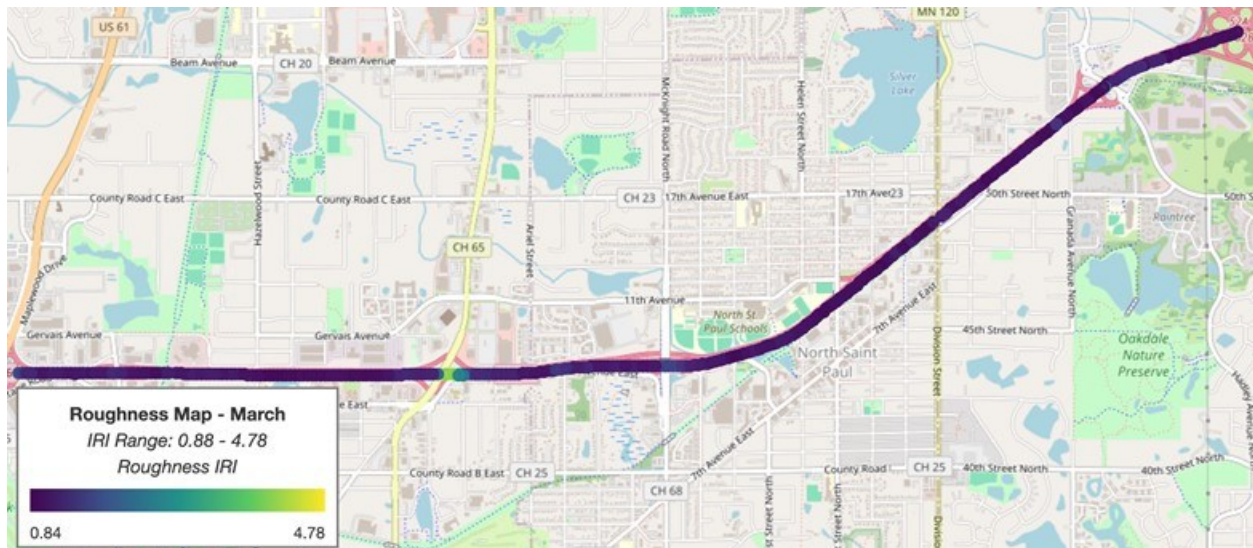


Figure 77. IRI color-coded map of a portion of MN-36 - March 2025.

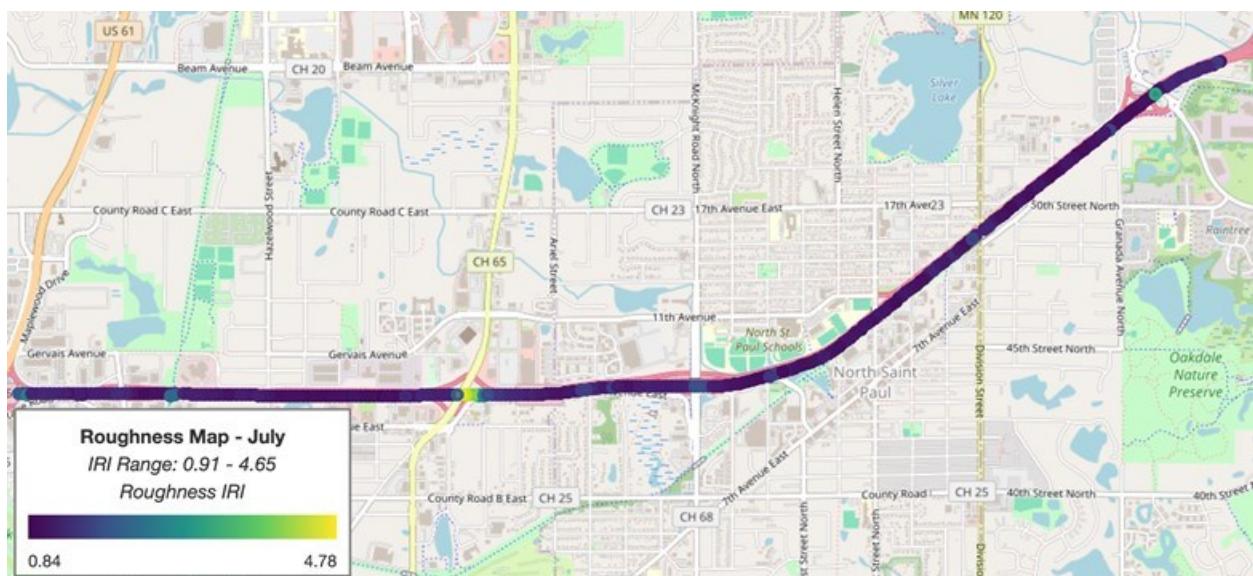


Figure 78. IRI color-coded map of a portion of MN-36 – July 2025.

Based on these seasonal IRI range comparisons, the data reveals that March and July exhibit the highest maximum IRI values (4.78 and 4.65 respectively), while November and January show slightly lower maximum values (4.38 and 4.21 respectively). However, these differences are quite minimal, with the largest variation being only 0.57 units between January's lowest maximum (4.21) and March's highest maximum (4.78), representing approximately a 13% difference. The minimum IRI values remain remarkably consistent across all seasons, ranging only from 0.84 to 0.91.

The differences are more apparent on road segments with an already high IRI indicating that sections with higher roughness are more affected by seasonal changes. This pattern suggests that already rough pavement sections experience greater seasonal variations, making them primary contributors to the

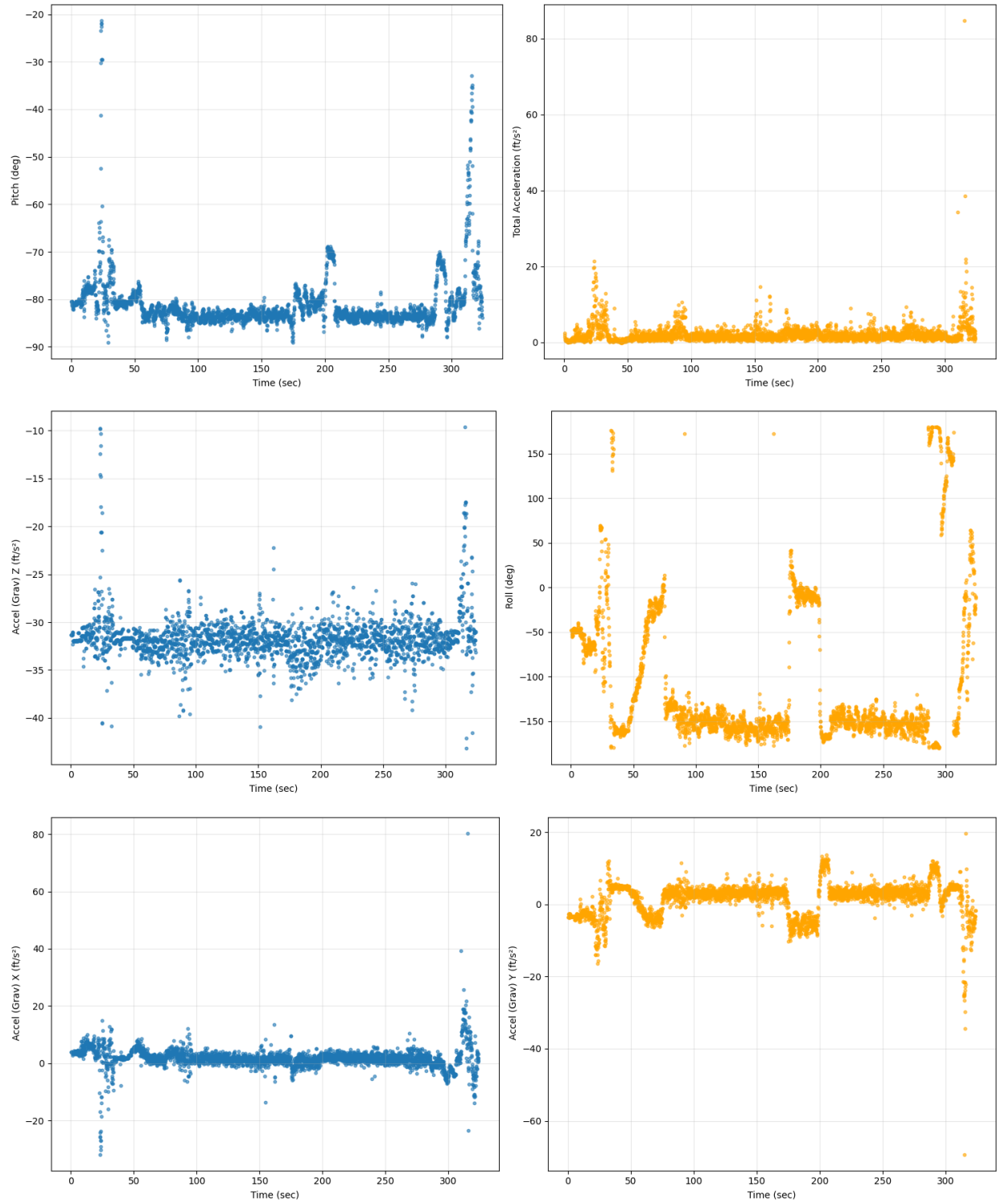
observed differences between seasons, while smoother pavement sections maintain relatively stable IRI values throughout the year.

These results suggest that while seasonal variations are larger than daily comparisons, the overall pavement roughness estimates remain relatively close throughout the year. However, small variations in IRI due to seasonal variability appear to be detectable using the NIRA pavement roughness estimates. However, these values cannot be validated, since no ground-truth data is available at such a fine time resolution. Moreover, the minor differences observed could be attributed to several factors, including seasonal temperature changes affecting measurement equipment calibration, varying vehicle dynamics due to temperature-related tire pressure changes, moisture content in pavement affecting surface characteristics, seasonal traffic loading patterns, or accumulated effects of freeze-thaw cycles during winter and spring months. Additionally, measurement uncertainties, GPS accuracy variations under different atmospheric conditions, and seasonal changes in pavement support conditions may contribute to these modest but observable seasonal trends in IRI values.

7.2 Generalization of Scanners to Other Vehicles

7.2.1 Data collection with the MnDOT Van

For this section, we want to investigate the compatibility of the MnDOT van in collecting OBD data with our scanners. On the day we collected data from MnROAD, the MnDOT van was equipped with our OBD scanner and driven once around the MnROAD loop to determine whether the scanner could collect data with this van or not. In the following figures, some scatter plots of the collected features with the MnDOT van have been illustrated in Figure 79 subplots.



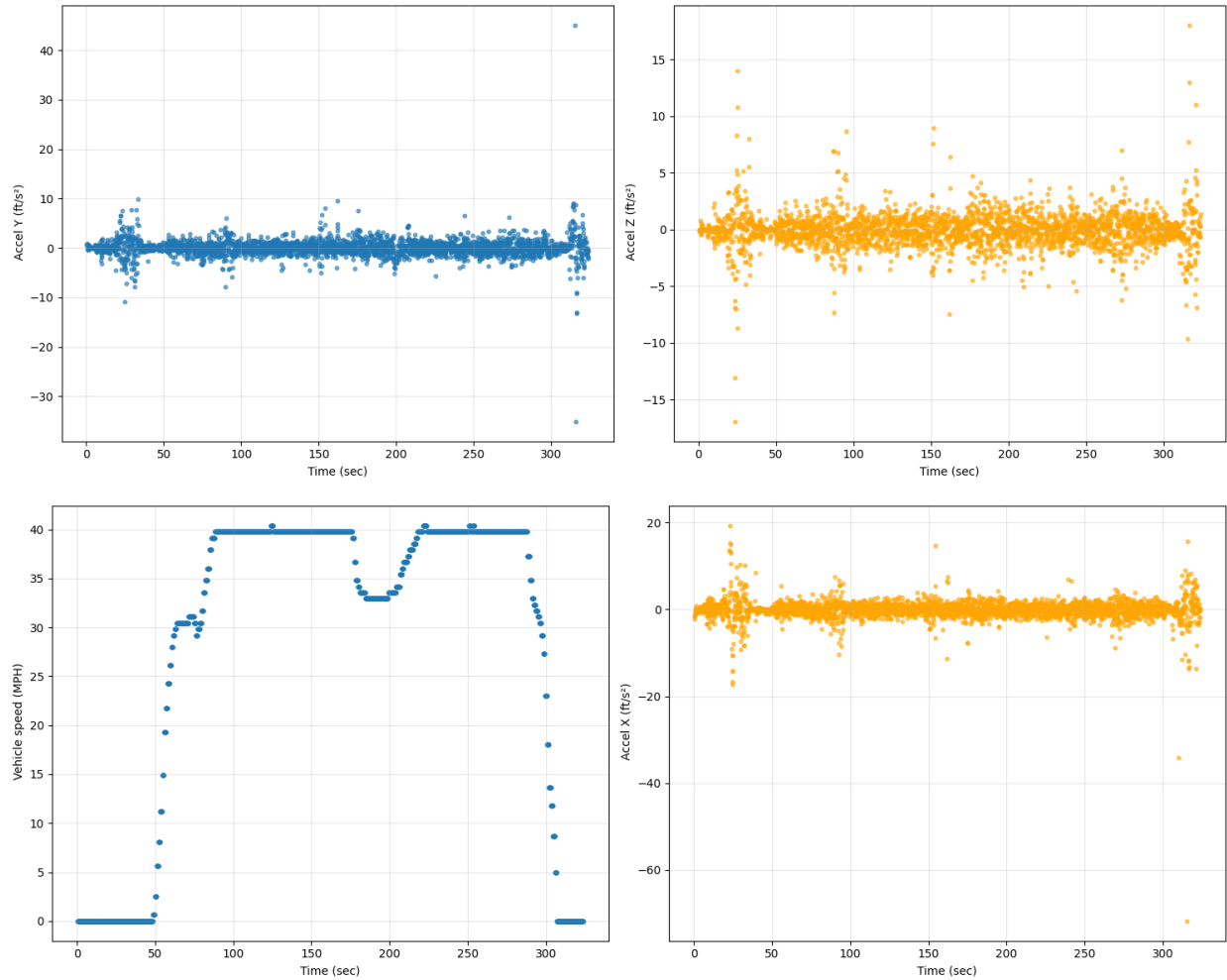


Figure 79. Scatter plots of key vehicle dynamics features versus time collected using OBD scanner with MnDOT van on the MnROAD loop.

The results demonstrate that OBD scanners are compatible with MnDOT's vans, and the distribution of the collected data makes sense with reality. The next step is to add this dataset to our previous datasets again to see how this data can predict the IRI with the proposed model.

7.2.2 IRI Prediction

Following the previous chapter guidelines, we conducted the map-matching process and alignment for the new van dataset and added it to all previous datasets to extract the predictions. In Figure 80, you can see the percentage difference between our predictions using the van dataset and the reference IRI data. As observed, at some points we have larger errors, and the total errors are slightly more frequent than our collected data comparison to ground truth measurements (referring to Chapter 6 results). However, the van data can still predict IRI at most points with high accuracy, with blue and green points showing acceptable compatibility of scanners with van data.



Figure 80. IRI percentage comparison between MnDOT ground truth measurements and Van collected data predictions for the MnROAD loop.

Van data may have less accuracy due to different vehicle characteristics (weight, suspension, tire properties) and varying OBD scanner mounting configurations compared to our original data collection setup. These systematic differences between the van and our training vehicle can introduce sensor response variations that affect prediction performance.

When we checked our prediction results with MnDOT, we received valuable feedback regarding the accuracy issues. The section of road in the left-side loop of MnROAD that showed low accuracy in IRI prediction has known problems that MnDOT had previously struggled with in their LVR Pathway data collection, and they continue to face challenges with the same areas on the LVR road. These are three areas of localized roughness, within otherwise fairly smooth sections. The localized roughness occurs at the transitions between test sections and, depending on the DMI settings, may appear in either or both of the adjacent test sections. Even with the Pathway Van, it is difficult to obtain consistent data or establish a good trendline for the test sections in these areas, as illustrated in the following figures.



Figure 81. Settlement section of HMA next to PCC in MnROAD.



Figure 82. Settlement section of HMA next to PCC in MnROAD on the map.



Figure 83. Transverse Ruts in PCC.



Figure 84. PCC Drain Between HMA Cells.

7.3 Implementation Strategy and Framework for Large Scale

7.3.1 Phased Implementation Approach

Phase 1: Equipment Standardization

Based on our extensive scanner evaluation in Chapter 3, large-scale implementation must standardize on the OBDLink LX scanner. While the AutoPi Telematics Unit was initially considered for its potential to capture wheel speed data, we encountered significant operational difficulties and inadequate manufacturer support that prevented its effective utilization. Among the feasible devices tested, the OBDLink LX demonstrated superior performance with 694 PID collection capability and reliable electric vehicle compatibility. The initial phase requires procuring OBDLink LX units along with high-capacity tablets featuring GPS functionality, having sufficient storage, and GPS accuracy within ± 3 meters. Vehicle selection criteria include compatibility with 2017+ model year vehicles for optimal OBD-II functionality, regular route coverage of target road segments, and driver availability for basic training procedures.

Personnel requirements include data analysts for processing pipeline management, field support technicians for equipment maintenance, and regional coordinators for program administration. The operational model scales efficiently with additional vehicles, requiring minimal incremental infrastructure investment. Personnel scaling follows predictable patterns, with additional analysts needed per geographic region or vehicle count.

Phase 2: Data Collection Infrastructure and Processing Pipeline

Drawing from our data collection experiences across MN-36, Metro to Northfield, and MnROAD, the processing pipeline must handle at least 8,000 data points based on our combined dataset experience. The data preprocessing and cleaning protocols based on our Chapter 6 methodology must be implemented to handle the volume and complexity of the collected data. ArcGIS map matching is essential for implementing our prediction model that successfully aligned OBD-II data with MnDOT ground truth measurements. Feature engineering processes convert raw sensor data into predictive

variables, while prediction model specialization ensures optimal performance for IRI estimation. Hyperparameter tuning across different geographic regions and road conditions maintains model accuracy and consistency throughout the network. The processing framework utilizes a Python-based and the machine learning model employs TensorFlow Serving. The visualization platform integrates with ArcGIS for mapping and spatial analysis capabilities. For future real-time implementation, platforms like Apache Kafka would be required for streaming data to handle 694 PIDs collected at 0.1-second intervals.

Phase 3: Model Deployment and Validation Network

The validation framework encompasses comprehensive performance monitoring and error tracking to ensure consistent model accuracy across Minnesota's diverse road network. The system includes checking IRI values with reference data points from MnDOT's high-precision measurements to maintain calibration standards. Additionally, checking IRI values with NIRA estimation provides continuous validation against network-level pavement data, ensuring the model performs reliably across different road segments and conditions. This dual validation approach establishes confidence in the methodology's performance while providing pathways for ongoing model refinement and quality assurance throughout the operational deployment.

The guidelines of the three phases can be seen in Figure 85.

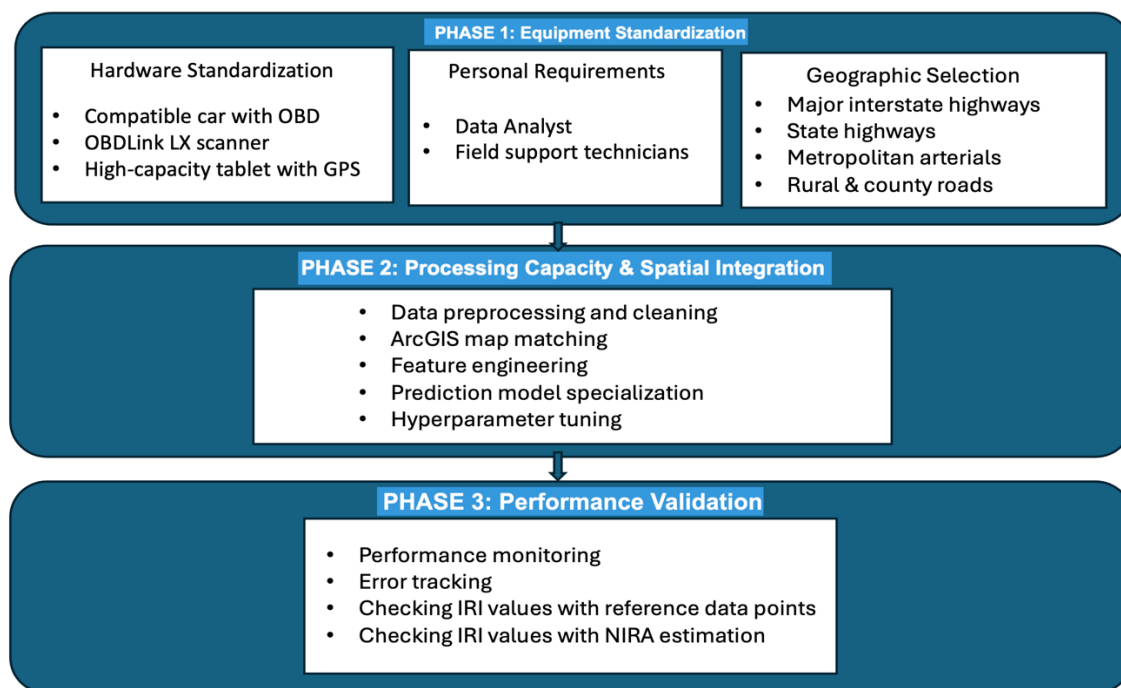


Figure 85. Three-phase implementation guidelines for large-scale OBD-II-based pavement quality assessment across Minnesota's road network.

7.3.2 Geographic Coverage and Prioritization

Initial focus areas are major interstate highways due to high traffic volumes and existing MnDOT validation data. Subsequent prioritization includes state highways with known pavement issues and metropolitan arterials, especially within the Twin Cities. Rural state highway networks and county roads follow as secondary priorities. Prioritization factors include traffic volume (AADT), pavement condition variability, economic impact, existing MnDOT reference data, and safety considerations.

7.4 Operational Guidelines and Best Practices

7.4.1 Optimal Collection Conditions

Based on our extensive testing across different conditions, specific environmental and operational parameters significantly affect data quality and prediction accuracy. Even the time and the car, and the way of getting the tablet, are so influential on the collected data quality and their sensitivity in predicting IRI. So, these all should be done in the least possible risky situation. Even the correct version of the sync application with the scanner is so effective, this is what we got from many times of trials and errors. Weather conditions play a critical role in data collection effectiveness. Daylight conditions are preferred for safety and optimal GPS signal quality. During operation, drivers should maintain steady speeds when possible while avoiding excessive acceleration and deceleration that could introduce artifacts in the sensor data. Route consistency requires following the same path for repeat measurements to ensure comparable data collection.

Vehicle requirements significantly impact data collection success, with model year providing the required OBD-II compatibility and comprehensive PID collection needed for prediction. Regular maintenance status is essential as engine diagnostic codes can interfere with OBD-II communication. Tire pressure within the manufacturer's specifications affects acceleration measurements and should be verified before data collection sessions.

7.4.2 Data Risk Assessment and Mitigation Strategies

7.4.2.1 Equipment Reliability Risks

GPS accuracy degradation represents a high probability, medium impact risk affecting spatial alignment accuracy through coordinate drift, poor satellite reception, and urban canyon effects. Mitigation involves high-sensitivity GPS tablets, signal quality monitoring, and alternative positioning methods, with contingency plans for post-processing correction using road network topology and manual verification for critical segments.

Data processing failures are low probability, high impact risks involving machine-learning model performance degradation due to data quality issues, indicated by high RMSE value and systematic bias

in predictions. Mitigation includes continuous model monitoring, automated rollback to previous model versions, and real-time performance alerts, with contingency plans for manual model retraining and temporary reversion to the Random Forest backup model.

7.4.2.2 Environmental and Operational Risks

Weather-related data quality issues present high probability, medium impact risks where precipitation, extreme temperatures, and road conditions affect sensor accuracy, indicated by abnormal acceleration patterns, GPS signal loss, and equipment malfunction. Mitigation involves weather-based collection protocols, equipment temperature ratings, and seasonal calibration adjustments, with contingency plans for increased validation frequency during adverse conditions and weather-specific correction factors.

Vehicle compatibility issues are medium probability, medium impact risks where OBD-II compatibility varies across vehicle makes, models, and years, indicated by reduced PID availability, communication errors, and inconsistent data collection. Mitigation includes vehicle compatibility testing, approved vehicle lists, and standardized equipment requirements, with contingency plans for alternative mounting solutions and vehicle-specific calibration procedures.

7.5 Immediate and Medium-Scale Implementation Guidelines

While large-scale implementation of the OBD-II-based IRI estimation framework has been extensively discussed, there is practical interest in immediate, smaller-scale applications. Districts often require rapid assessment of pavement conditions for individual road segments in their jurisdiction. The following guidelines outline how to effectively implement the OBD-II-based IRI prediction model in scenarios of limited scale and data availability.

7.5.1 Data Collection and Minimum Data Requirements

The accuracy of the proposed IRI prediction model is highly dependent on the volume and diversity of collected data points used both to calibrate the model for each vehicle type, as well as to evaluate the roughness of a particular stretch of pavement. Currently, our validated model performance (RMSE values of 0.172 and MRE of 0.131) was achieved using approximately 8,000 data points, distributed across various road conditions and segments within the Minnesota network. For smaller-scale implementations, such as assessing a new road, it is recommended to collect data at a higher sampling density.

7.5.2 Recommended Procedure for Immediate Implementation

To rapidly estimate the IRI for a new road:

1. Equip compatible vehicles (preferably those similar in type, weight, and suspension characteristics to previously trained models) with the standardized OBDLink LX scanners and GPS-enabled tablets.
2. Conduct multiple passes along the target roadway to enhance data reliability, capturing data during stable driving conditions.
3. Employ the established data collection procedures, ensuring consistent scanner mounting and correct application synchronization.
4. Process the collected data following the same feature extraction and preprocessing pipeline outlined in Chapter 6 guidelines. This process could be automated in future work for rapid data processing and model deployment if desired.

Note that the above procedure may require model calibration for the vehicle being used for data collection with ground truth data. However, this could be developed into a standalone piece of software that would allow easy calibration if ground truth (e.g., Pathways van) data is available.

7.5.3 Limitations and Considerations

Immediate, small-scale implementations without pre-existing training data from the same road segment will likely experience reduced accuracy compared to regions included in the original dataset. Currently, the model relies on data from road segments that have distinct training and testing partitions within the same general area. A completely new road, with no previous inclusion in training, may exhibit higher prediction variance. Additionally, our model's achieved accuracy is contingent upon having highly accurate ground truth measurements available for initial training. Without these high-accuracy ground truth measurements, prediction reliability may diminish significantly.

In cases where immediate high precision is required, a limited calibration effort using either MnDOT's high-precision van or NIRA reference data on a representative section of the new roadway is highly recommended. A brief calibration dataset, even if limited in spatial coverage, can substantially reduce uncertainty and improve model accuracy for immediate assessments.

7.5.4 Recommended Validation and Calibration Steps

To enhance prediction reliability:

- Conduct initial OBD-II data collection runs in parallel with at least a limited number of reference measurements (e.g., MnDOT van or NIRA).
- Utilize these ground truth measurements to perform localized calibration of the prediction model, adjusting for site-specific characteristics such as pavement type, vehicle speed distribution, and environmental conditions.

- Evaluate the calibrated model's performance through cross-validation techniques on the newly collected reference data.
- Since the existing model was developed and trained using roads for which we had ground truth measurements, and has been optimized for these specific conditions, applying it to completely new roads without ground truth measurements or with very limited reference data will likely require model optimization. The model should undergo hyperparameter tuning and customization to adjust to the new conditions and achieve the highest possible accuracy.

These pragmatic steps enable districts to effectively implement OBD-II-based pavement quality assessments at small or medium scales, providing immediate insights into pavement conditions with reasonable accuracy and manageable resource investments.

Chapter 8: Conclusions and Future Work

The comprehensive analysis and validation conducted during this project indicate that the OBD-II-based pavement quality assessment methodology is well-prepared for large-scale implementation across Minnesota's transportation network. The proposed method achieves strong alignment with established reference standards, demonstrating best-case RMSE values of 0.172 and MRE of 0.131 relative to MnDOT's high-precision measurements as well as comparable performance when validated against NIRA data.

A key advantage of our methodology is its ability to provide extensive coverage of Minnesota's entire state highway system through routine vehicle operations. This represents a significant improvement over traditional methods, which offer limited coverage due to operational constraints. By leveraging existing vehicle fleets, our OBD-II-based approach allows continuous pavement monitoring without the need for specialized instrumentation. Current MnDOT pavement assessments heavily depend on costly Pathways van operations that require extensive coordination, specialized personnel, and dedicated equipment deployments. The presented methodology can convert everyday vehicle operations into opportunities for ongoing pavement condition assessments, substantially altering the economics and effectiveness of infrastructure monitoring through distributed data collection.

The framework developed through this study establishes a foundation for the evolution of long-term pavement monitoring strategies. Its demonstrated compatibility with both high-precision MnDOT data and broader network-level NIRA data ensures adaptability across different data scenarios and regions. Future enhancements should explore integration with emerging connected vehicle technologies, advanced sensor systems, and artificial intelligence developments while continuing to leverage vehicle data for cost-effective infrastructure monitoring. For instance, integrating devices such as AutoPi or advanced sensors like Panda AI could enhance assessments beyond International Roughness Index (IRI) estimates, potentially measuring pavement friction and energy consumption.

The successful implementation of this methodology in Minnesota provides a valuable model. Future research directions include leveraging connected vehicle technologies for integration with Vehicle-to-Infrastructure (V2I) communications, real-time data streaming for immediate condition assessments, and enhanced sensor fusion techniques compatible with all vehicle platforms. Further development in advanced analytics and AI could encompass evolution beyond CNN-LSTM models, deep reinforcement learning for dynamic route optimization, computer vision applications for visual pavement assessments, and predictive analytics for maintenance scheduling.

Model improvements should also target better seasonal variation handling, vehicle-specific calibration, correlating pavement conditions with energy consumption, and enhancing the interpretability of machine-learning models. Extending application capabilities might include developing algorithms for estimating pavement friction, detecting and classifying potholes, assessing construction quality, and integrating environmental impact assessments.

References

- Abdelaziz, N., El-Hakim, R. A., El-Badawy, S. M., & Afify, H. A. (2020). International Roughness Index prediction model for flexible pavements. *International Journal of Pavement Engineering*, 21(1), 88-99.
- Agostinacchio, M., Ciampa, D., & Olita, S. (2014). The vibrations induced by surface irregularities in road pavements—a Matlab® approach. *European Transport Research Review*, 6(3), 267-275.
- Allouch, A., Koubâa, A., Abbes, T., & Ammar, A. (2017). Roadsense: Smartphone application to estimate road conditions using accelerometer and gyroscope. *IEEE Sensors Journal*, 17(13), 4231-4238.
- Bashar, M. Z., & Torres-Machi, C. (2021). Performance of machine learning algorithms in predicting the pavement international roughness index. *Transportation Research Record*, 2675(5), 226-237.
- Cafiso, S., Di Graziano, A., Marchetta, V., & Pappalardo, G. (2022). Urban road pavements monitoring and assessment using bike and e-scooter as probe vehicles. *Case Studies in Construction Materials*, 16, e00889.
- Du, Z., Skar, A., Pettinari, M., & Zhu, X. (2023). Pavement friction evaluation based on vehicle dynamics and vision data using a multi-feature fusion network. *Transportation Research Record*, 2677(11), 219-236.
- Feng, Y., Chen, H., Zhao, H., & Zhou, H. (2020). Road tire friction coefficient estimation for four wheel drive electric vehicle based on moving optimal estimation strategy. *Mechanical Systems and Signal Processing*, 139, 106416.
- Ghose, A., Biswas, P., Bhaumik, C., Sharma, M., Pal, A., & Jha, A. (2012). Road condition monitoring and alert application: Using in-vehicle smartphone as internet-connected sensor. In *2012 IEEE International Conference on Pervasive Computing and Communications Workshops*, pp. 489-491.
- Levenberg, E., Skar, A., Pour, S. M., Kindler, E., Pettinari, M., Bajic, M., Alstrøm, T. S., & Schlotz, U. (2021). Live road condition assessment with internal vehicle sensors. *Transportation Research Record*, 2675(10), 1442-1452. <https://doi.org/10.1177/03611981211016852>.
- Li, H., Saldivar-Carranza, E., Mathew, J. K., Kim, W., Desai, J., Wells, T., & Bullock, D. M. (2020). Extraction of Vehicle CAN Bus Data for Roadway Condition Monitoring. *Join Transportation Research Program*.
- Lima, L. C., Peixoto Amorim, V. J., Muzetti Pereira, I., Nunes Ribeiro, F. & Rabelo Oliveira, R. A. (2016). Using crowdsourcing techniques and mobile devices for asphaltic pavement quality recognition. In *2016 VI Brazilian symposium on computing systems engineering (SBESC)*, pp. 144-149.
- Live Road Assessment project (2023). Retrieved from <https://lira-project.dk/>

Liu, Z., Tao, W., Jiang, L., & Zhu, C. (2014, June). Design and application on electric vehicle real-time condition monitoring system by Internet of things technology. In 2014 IEEE 5th International Conference on Software Engineering and Service Science (pp. 744-747). IEEE.

Mattinzioli, T., Butt, A. A., & Harvey, J. (2023). Literature review on pavements and electric vehicle interaction: A research roadmap. *Transportation Research Part D: Transport and Environment*, 122, 103886.

Seraj, F., Meratnia, N., & Havinga, P. J. (2017, March). RoVi: Continuous transport infrastructure monitoring framework for preventive maintenance. In 2017 IEEE International Conference on Pervasive Computing and Communications (PerCom) (pp. 217-226).

Seraj, F., van der Zwaag, B. J., Dilo, A., Luarasi, T., & Havinga, P. (2015). RoADS: A road pavement monitoring system for anomaly detection using smart phones. In M. Atz-mueller, A. Chin, F. Janssen, I. Schweizer, & C. Trattner (Eds.) *Big data analytics in the social and ubiquitous context* (pp. 128–146). Springer.

Sholevar, N., Golroo, A., & Esfahani, S. R. (2022). Machine learning techniques for pavement condition evaluation. *Automation in Construction*, 136, 1 04190.

Skar, A., & Levenberg, E. (2024) Road profile inversion from in-vehicle accelerometers. *Journal of Transportation Engineering, Part B: Pavements*, 150(1), 04023040.

Souza, V. M. A., Giusti, R., & Batista, A. J. L. (2018). Asfalt: A low-cost system to evaluate pavement conditions in real-time using smartphones and machine learning. *Pervasive and Mobile Computing*, 51, 121-137.

Staniek, M. (2021). Road pavement condition diagnostics using smartphone-based data crowdsourcing in smart cities. *Journal of Traffic and Transportation Engineering*, 8(4), 554-567.

Xue, G., Zhu, H., Hu, Z., Yu, J., Zhu, Y., & Luo, Y. (2016). Pothole in the dark: Perceiving pothole profiles with participatory urban vehicles. *IEEE Transactions on Mobile Computing*, 16(5), 1408-1419.

Zhao, B., Takada, S., & Nagayama, T. (2015). Development of response-based road profile estimation using multiple observables. Paper presented at the JSCE Annual Meeting at Okayama.