

Regionwide Traffic Performance Evaluation and Performance Measure Development Using Multi-Source Data

Final Report

Prepared by:



THE UNIVERSITY OF ARIZONA
COLLEGE OF ENGINEERING

Center for Applied
Transportation Sciences

Prepared for:



Pima Association of Governments

June, 2023

ACKNOWLEDGMENTS

This research project is sponsored by the Pima Association of Governments (PAG), funded by federal transportation sources such as STBGP, PL and 5305, and led by the University of Arizona team. The University of Arizona team would like to acknowledge PAG staff and engineers for their invaluable support and collaboration and for providing traffic data.

This report was prepared by:

Department of Civil and Architectural Engineering and Mechanics

The University of Arizona

Yao-Jan Wu, Ph.D., P.E., Associate Professor

Xiaofeng Li, Ph.D., Research Assistant Professor

Peipei Xu, Research Specialist



With the assistance of PAG staff:

Pima Association of Governments

Ryan James Hatch, Senior Transportation Data Scientist

James Tokishi, Transportation Data Science Coordinator

Xiaobo Ma, Ph.D., Transportation Data Scientist

Hyunsoo Noh, Ph.D., Data Science Administrator

Dave Adler, P.E., Deputy Executive Director



DISCLAIMER

The content of this report reflects the views of the authors, who are responsible for the facts and the accuracy of the information presented herein and do not necessarily reflect the official view or policies of the sponsoring organizations. These contents do not constitute a standard, specification, or regulation. The research work of this study was completed before Sept. 2023. Any new products/modules and the information released after the completion date of this study are not included in this study. This document is disseminated under the sponsorship of the Pima Association of Governments in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Authors

The University of Arizona – Consultant Team

Yao-Jan Wu, Ph.D., P.E., Associate Professor

Xiaofeng Li, Ph.D., Research Assistant Professor

Peipei Xu, Research Specialist

Pima Association of Governments

Ryan James Hatch, Senior Transportation Data Scientist

James Tokishi, Transportation Data Science Coordinator

Xiaobo Ma, Transportation Data Scientist

Hyunsoo Noh, Ph.D., Data Science Administrator

Dave Atler, P.E., Deputy Executive Director

TABLE OF CONTENTS

ACKNOWLEDGMENTS	1
DISCLAIMER	2
LIST OF FIGURES	8
LIST OF TABLES	16
ACRONYMS, ABBREVIATIONS, AND TERMS	18
EXECUTIVE SUMMARY	20
CHAPTER 1: INTRODUCTION	24
1.1 Background.....	24
1.2 Organization of report.....	24
CHAPTER 2: LITERATURE REVIEW	26
2.1 Traffic Performance Measure Estimation.....	26
2.2 GPS Data Application.....	30
2.3 Applications of Transportation Systems Management and Operations.....	31
CHAPTER 3: THIRD-PARTY PROBE VEHICLE DATA EXPLORATION	37
3.1 Data Cleaning and Processing	37
3.2 Traffic Delay.....	38
3.2.1 Calculation	38
3.2.2 Results and Analysis.....	39
3.3 Level-of-Service	44
3.3.1 Definition	44
3.3.2 Results and Analysis.....	44
3.4 Arrival-on-Green/Arrival-on-Red.....	45
3.4.1 Calculation	45

3.4.2 Results and Analysis	46
3.5 Split Failure.....	50
3.5.1 Calculation	50
3.5.2 Results and Analysis	50
3.6 Red-light Running.....	52
3.6.1 Calculation	52
3.6.2 Results and Analysis	52
CHAPTER 4: MOBILITY/RELIABILITY PERFORMANCE DATA COLLECTION AND ANALYSIS	55
4.1 Sensor-based performance Data collection and analysis	55
4.1.1 Simple Delay.....	56
4.1.2 Level-of-Service	67
4.1.3 Arrival-on-Green.....	70
4.1.4 Split Failure.....	75
4.2 Mobility/Reliability Performance Comparison.....	81
4.2.1 Delay Comparison	81
4.2.2 Level-of-Service Comparison.....	86
4.2.3 Arrival-on-Green Comparison	88
4.2.4 Split Failure Comparison	91
CHAPTER 5: REPRESENTATIVENESS EVALUATION OF PERFORMANCE MEASURES	94
5.1 Study Locations	94
5.2 Representativeness Evaluation of Control Delay	95
5.2.1 Through Movement	95

5.2.2	Left-turn Movement.....	104
5.3	Representativeness Evaluation of Arrival on Green	111
5.4	Representativeness Evaluation of Split Failure	114
5.5	Representativeness Evaluation of Reliability Performance	117
5.5.1	Delay Reliability	117
5.5.2	AoR Reliability	120
5.5.3	Split Failure Reliability.....	122
5.6	Data Quality and Availability Analysis	123
CHAPTER 6: MOBILITY/RELIABILITY PERFORMANCE ESTIMATION USING MAXVIEW AND MIOVISION SYSTEM.....		130
6.1	Data Description and study locations	130
6.1.1	Event-based Data Collected by MaxView System	130
6.1.2	Event-based Data Collected by Miovision System.....	131
6.2	Methodology	133
6.2.1	Data Processing for Input.....	133
6.2.2	Model-Agnostic Meta-Learning (MAML)	136
6.3	Implementation and ResULTs	138
6.3.1	Control Delay Estimation	139
6.3.2	Arrival-on-Green Estimation	145
CHAPTER 7: ACCEPTANCE CRITERIA DEVELOPMENT.....		152
7.1	Data Collection and Processing	152
7.2	Acceptance Criteria Development	152
7.2.1	Literature Review Summary	153

7.2.2	File Data Assessment.....	155
7.2.3	Acceptance Criteria.....	160
7.2.4	Validation Result Summary.....	161
CHAPTER 8: REGION-WIDE TRAFFIC MOBILITY/RELIABILITY PERFORMANCE ESTIMATION.....		163
8.1	Data Collection	163
8.2	Region-wide Control Delay Estimation and Analysis	164
8.2.1	Hourly Control Delay	164
8.2.2	Monthly Average Hourly Control Delay	169
8.3	Region-wide AOG Ratio Estimation and Analysis	172
8.3.1	Hourly AoG Ratio.....	172
8.3.2	Monthly Average Hourly AOG Ratio.....	177
APPENDIX A: DATA COLLECTION DESCRIPTION		181
APPENDIX B: MAML MODEL DESCRIPTION FOR ESTIMATING AOG AND DELAY		182
REFERENCES.....		183

LIST OF FIGURES

Figure 3-1. (a) Raw GPS points at one sample intersection; (b) on-road GPS data filter; (c) data processing procedure	38
Figure 3-2. Hourly delay and number of sample probe vehicles at Ina Rd. & La Cañada Dr.....	40
Figure 3-3. Average hourly control delay by movement at Ina Rd. & La Cañada Dr. during January 2021.....	41
Figure 3-4. Spatial-temporal distribution of delay.....	43
Figure 3-5. Intersection LOS distribution by hour on March 17, 2021	45
Figure 3-6. Heatmap of average AoG ratio by intersection and hour.....	47
Figure 3-7. Spatial-temporal distribution of AoG and AoR	49
Figure 3-8. Split failure ratio by intersection and hour.....	51
Figure 3-9. Average red-light running frequency of through vehicles	52
Figure 3-10. Red-light running trajectories at Ajo Way & Palo Verde Rd.	54
Figure 4-1. Miovision Sensor Locations.....	55
Figure 4-2. Miovision signal performance measures and network metrics	56
Figure 4-3. Temporal trend of simple delay at La Cholla Blvd. & River Rd., Tucson	57
Figure 4-4. 95 th percentile simple delay at La Cholla Blvd. & River Rd., Tucson during Jan-Mar 2021.....	59
Figure 4-5. Buffer delay at La Cholla Blvd. & River Rd., Tucson during Jan.-March 2021.....	60
Figure 4-6. Buffer index at La Cholla Blvd. & River Rd., Tucson during Jan.-March 2021.....	60

Figure 4-7. Spatial distribution of average through and left-turn simple delay at different hours	64
Figure 4-8. Spatial distribution of buffer delay index.....	66
Figure 4-9. Simple delay distribution by hour of the day	67
Figure 4-10. Intersection LOS distribution by hour on March 17, 2021	68
Figure 4-11. Temporal trend of AoG at La Cholla Blvd. & River Rd., Tucson.....	71
Figure 4-12. Spatial distribution of AoG at different hours.....	74
Figure 4-13. Buffer AoR at La Cholla Blvd. & River Rd., Tucson during Jan.-March 2021.	75
Figure 4-14. Temporal trend of average split failure percentage at La Cholla Blvd. / River Rd.	76
Figure 4-15. Spatial distribution of split failure percentage at different hours.....	79
Figure 4-16. 95 th percentile split failure percentage at La Cholla Blvd. & River Rd. during Jan.- March 2021	80
Figure 4-17. Buffer split failure percentage at La Cholla Blvd. & River Rd. during Jan.-March 2021	81
Figure 4-18. Comparison between Miovison delay and Wejo-based delay at La Cholla Blvd. / River Rd.....	82
Figure 4-19. Delay cumulative distribution function comparison at La Cholla Blvd. / River Rd.	83
Figure 4-20. Comparison between Miovison and Wejo-based 95 th percentile delay at La Cholla Blvd. / River Rd.....	84
Figure 4-21. Comparison between Miovison and Wejo-based buffer delay at La Cholla Blvd. / River Rd.....	85

Figure 4-26. Comparison between Miovision and Wejo-based buffer index at La Cholla Blvd. / River Rd.....	85
Figure 4-23. Reliability indices cumulative distribution function comparison at La Cholla Blvd. / River Rd.....	86
Figure 4-24. LOS comparison at La Cholla Blvd. / River Rd., Tucson.....	87
Figure 4-29. Comparison between Miovision AoG and Wejo AoG at La Cholla Blvd. / River Rd.	88
Figure 4-26. AOG cumulative distribution function comparison at La Cholla Blvd. / River Rd.	89
Figure 4-27. 95 th percentile AoR comparison at La Cholla Blvd. / River Rd.	90
Figure 4-32. Buffer AoR comparison at La Cholla Blvd. / River Rd.....	90
Figure 4-29. Comparison between Miovision Split Failure and Wejo Split Failure at La Cholla Blvd. / River Rd.....	91
Figure 4-30. Split failure cumulative distribution function comparison at La Cholla Blvd. / River Rd.	92
Figure 4-31. 95 th percentile split failure comparison at La Cholla Blvd. / River Rd.....	93
Figure 4-32. Buffer split failure comparison at La Cholla Blvd. / River Rd.	93
Figure 5-1. 62 study intersections in the PAG region.....	94
Figure 5-2. Comparison between Wejo- and Miovision-based delay of through movement	96
Figure 5-3. Distribution of the difference between two delays by sample size.....	96
Figure 5-4. Comparison between different distributions with various sample sizes	97

Figure 5-5. Comparison between different distributions with various sample sizes by hour of the day	98
Figure 5-6. Sensitivity analysis by changing the distance of Wejo delay estimation.....	99
Figure 5-7. Correlation between Wejo- and Miovision-based delay for through movement.....	101
Figure 5-8. Box plot of correlation coefficients for all directions at all intersections	101
Figure 5-9. Correlation coefficients under different sample sizes.....	102
Figure 5-10. Correlation coefficients for various traffic volume levels	103
Figure 5-11. Comparison between Wejo- and Miovision-based delay of left-turn movement. .	104
Figure 5-12. Plots of distributions of the error with various sample sizes (Left-turn traffic).....	105
Figure 5-13. Plots of distributions of the error with various sample sizes after removing outlier intersections (Left-turn traffic)	106
Figure 5-14. Sensitivity analysis by changing the distance. (Left-turn traffic)	107
Figure 5-15. Correlation between Wejo- and Miovision-based delay for left-turn movement. .	108
Figure 5-16. Box plot of correlation coefficients for all directions at all intersections (left-turn traffic).....	109
Figure 5-17. Correlation coefficients under different sample sizes.....	109
Figure 5-18. Correlation coefficients for various traffic volume levels	110
Figure 5-19. Comparison between the Miovision- and Wejo-based AoG at Ina Rd. & La Cañada Dr., NB.	111
Figure 5-20. Correlation between Wejo- and Miovision-based AOG for through movement...	112

Figure 5-21. Box plot of correlation coefficients for AoG for all intersections and road directions	112
Figure 5-22. Correlation coefficients of AOG under different sample sizes	113
Figure 5-23. Correlation coefficients of AOG for various traffic volume levels	114
Figure 5-24. Comparison between Wejo- and Miovision-based split failure of through movement at Ina Rd. & La Cholla Blvd.....	115
Figure 5-25. Correlation between Wejo-based and Miovision-based split failure for through movement	116
Figure 5-26. Box plot of correlation coefficients for split failure for all intersections and road directions	116
Figure 5-27. Correlation coefficients for split failure under different sample sizes	117
Figure 5-28. Box plot of correlation coefficients of 95 th percentile delay for all intersections and road directions	118
Figure 5-29. Correlation coefficients of 95 th percentile delay under different sample sizes	119
Figure 5-30. Correlation coefficients of 95 th percentile delay under different penetration rates	119
Figure 5-31. Box plot of correlation coefficients of 95 th percentile AoR for all intersections and road directions	120
Figure 5-32. Correlation coefficients of 95 th percentile AoR under different sample sizes	121
Figure 5-33. Correlation coefficients of 95 th percentile AoR under different penetration rates.	121
Figure 5-34. Box plot of correlation coefficients of 95 th percentile split failure for all intersections and road directions	122

Figure 5-35. Correlation coefficients of 95 th percentile split failure under different sample sizes	123
Figure 5-36. Correlation coefficients of 95 th percentile split failure under different penetration rates	123
Figure 5-37. Study locations	125
Figure 5-38 Different threshold distributions of Wejo data for through movement by the hour of the day	126
Figure 5-39. Different threshold distributions of Wejo data for left-turn movement by the hour of the day	127
Figure 5-40. Spatial distribution of Wejo data sample sizes for through movement (7 a.m.-8 p.m.)	128
Figure 5-41. Spatial distribution of Wejo data sample sizes for left-turn movement (7 a.m.-8 p.m.)	129
Figure 6-1. Typical detector configuration at signalized intersections in MaxView system.....	131
Figure 6-2. Typical detector configuration at signalized intersections in Miovision system	132
Figure 6-3. Event-based data collection process in MaxView and Miovision systems	132
Figure 6-4. Vehicle trajectory and signal timing	134
Figure 6-5. Signal status categorization based on detection events.....	135
Figure 6-6. The training process of the MAML algorithm.....	137
Figure 6-7. Comparison between estimated delay and ground-truth delay using the MaxView system at all study intersections	140

Figure 6-8. Ground-truth and estimated control delay comparison at Speedway Blvd. & Euclid Ave.	142
Figure 6-9. Method performance for delay estimation using the MaxView system by detector location and traffic movement.....	143
Figure 6-10. Comparison between estimated delay using Miovision system and ground-truth delay at all study locations	144
Figure 6-11. Method performance for delay estimation using the Miovision system by detector location and traffic movement.....	145
Figure 6-12. Comparison between estimated AOG and ground-truth AOG using the MaxView system at all study locations	147
Figure 6-13. Ground-truth and estimated AOG comparison at Speedway Blvd. & Euclid Ave.	149
Figure 6-14. Model performance for AOG estimation using the MaxView system by detector location and traffic movement.....	149
Figure 6-15. Comparison between estimated AOG using Miovision system and ground-truth AOG at all study locations	150
Figure 6-16. Model performance for AOG estimation using Miovision system by detector location and traffic movement.....	151
Figure 7-1. Comparison between estimated and ground-truth monthly average hourly control delay using the MaxView system.	157
Figure 7-2. Comparison between estimated and ground-truth monthly average hourly control delay using the Miovision system	157
Figure 7-3. Comparison between estimated and ground-truth monthly average hourly AOG using the MaxView system	158

Figure 7-4. Comparison between estimated and ground-truth monthly average hourly AOG using the Miovision system.....	159
Figure 8-1. The number of intersections with available event-based data	163
Figure 8-2. Estimated hourly control delay at Speedway Blvd. & Euclid Ave.	164
Figure 8-3. Estimated control delay distribution by hour of the day	165
Figure 8-4. Spatial distribution of estimated control delay on Sep 14, 2021.....	168
Figure 8-5. Monthly average hourly control delay at Speedway Blvd. & Euclid Ave. in Sept. 2021	169
Figure 8-6. Heatmap of monthly average estimated control delay at Speedway Blvd. & Euclid Ave.	171
Figure 8-7. Estimated hourly AoG ratio at Speedway Blvd. & Euclid Ave.	172
Figure 8-8. Estimated AoG ratio distribution by hour of the day	173
Figure 8-9. Spatial distribution of AoG ratio on Sept. 14, 2021.....	177
Figure 8-10. Monthly average estimated hourly AoG ratio at Speedway Blvd. & Euclid Ave. in Sept. 2021	178
Figure 8-11. Heatmap of monthly average estimated AoG ratio at Speedway Blvd. & Euclid Ave.	180

LIST OF TABLES

Table 1-1. Project tasks and schedule	25
Table 2-1. Performance Measures Summary from Miovision TrafficLink, CATT Lab RITIS, and INRIX Signal Analytics	27
Table 2-2. Applications and Major Data Sources for TSMO Programs.....	31
Table 2-3. Example Applications of AI Technologies for TSMO Applications	33
Table 2-4. Example Research of AI Technologies for TSMO Applications.....	33
Table 2-5. Case Studies for Advancing TSMO	35
Table 3-1. Signalized intersection LOS Criteria.....	44
Table 4-1. Intersection Ranking based on the LOS at 8 a.m.	69
Table 4-2. Intersection Ranking based on the LOS at 5 p.m.	69
Table 4-3. Intersection Ranking based on the LOS at 1 a.m.	70
Table 4-4. Confusion matrix of LOS comparison	87
Table 5-1. Threshold for different performance measures.	124
Table 7-1. Literature Review Summary of Control Delay Estimation	153
Table 7-2. Literature Summary of AOG Estimation	155
Table 7-3. Method performance metrics for hourly control delay estimation.....	155
Table 7-4. Method performance metrics for hourly AOG ratio estimation.....	156
Table 7-5. Method performance metrics for monthly average hourly control delay estimation	158
Table 7-6. Method performance metrics for monthly average hourly AOG estimation	160

Table 7-7. Preliminary Acceptance Criteria for Estimated Hourly Control Delay..... 160

Table 7-8. Preliminary Acceptance Criteria for Estimated Hourly AOG Ratio 160

Table 7-9. Preliminary Acceptance Criteria for Estimated Monthly Average Hourly Control Delay
..... 161

Table 7-10. Preliminary Acceptance Criteria for Estimated Monthly Average AOG Ratio 161

ACRONYMS, ABBREVIATIONS, AND TERMS

PAG	Pima Association of Governments
TMSO	Transportation Systems Management and Operations
UArizona	University of Arizona
STL	Smart Transportation Lab (at the University of Arizona)
STL-SQL	Smart Transportation Lab - Structured Query Language
COT	The City of Tucson
PCDOT	Pima County Department of Transportation
ADOT	Arizona Department of Transportation
ATMS	Advanced Traffic Management System
GPS	Global Positioning System
LOS	Level of service
AoG	Arrival on green
AoR	Arrival on red
GOR	Green occupancy ratio
ROR	Red occupancy ratio
ROR5	The red occupancy ratio during the first five seconds of red
TMC	Turning Movement Counts/Volumes
AADT	Annual Average Daily Traffic
CDF	Cumulative distribution function
ATMS	Advanced traffic management system
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
MAPE	Mean Absolute Percent Error
TOD	Time of Day

QA/QC	Quality assurance and quality control
DOT	Department of Transportation
RADS	Regional Archived Data System
CAV	Connected autonomous vehicles
TIM	Traffic Incident Management
FHWA	Federal Highway Administration
NPMRDS	National Performance Monitoring Research Dataset
RITIS	The Regional Integrated Transportation Information System
CATT Lab	The Center for Advanced Transportation Technology Laboratory at the University of Maryland
MMITSS	Multi-Modal Intelligent Traffic Signal System
SCATS	Sydney Coordinated Adaptive Traffic System
ATSMPS	Automated Traffic Signal Performance Measures
ASCT	Adaptive Signal Control Technology
RSUs	Roadside Units
OBUs	Onboard Units
RRFBs	Rectangular Rapid Flashing Beacons
LPI	Leading Pedestrian Interval
PPP	public-private partnerships

EXECUTIVE SUMMARY

Reducing traffic congestion and providing better transportation service in the region is one of the crucial roles for the Pima Association of Governments. Data-driven transportation planning and traffic operation management is not an imaginary goal anymore but a realistic and urgent target. For example, the federal FAST Act requires that state DOTs and MPOs develop data-driven performance measures using regional big data, and transportation agencies in the U.S. have been developing strategies for TSMO to improve the performance of the transportation system. Therefore, the goal of this study is to investigate the sources of traffic data and develop a method to support PAG's regional TSMO-related traffic data collection/maintenance and advanced modeling. This project will focus on identifying the traffic mobility/reliability performance measures from video- and event-based traffic data and crowdsourced data.

Most transportation agencies that have implemented TSMO systems relied on fixed traffic sensors for collecting traffic performance measures such as delay, arrival-on-green (AoG), and split failure, which has a strict requirement on the sensor configuration. However, most traffic detection systems originally configured for actuation signal control do not meet the requirement for collecting performance measures, because the only function of these traffic detection systems is to detect traffic arrival rather than collect traffic data. Even though most existing traffic detection configurations do not meet the requirement for collecting performance measures, they are still able to collect high-resolution event-based data, which is less informative compared to the standard detection configurations. However, these less informative events collected by existing traffic detection still have high correlation with traffic arrivals according to our previous project.

In addition to high-resolution event-based data, third-party connected-vehicle data (Wejo) has become available and has been applied to calculate the traffic performance measures at signalized intersections. Wejo data is collected and processed to explore its pertinence in extracting the performance measures. Several methods are proposed to calculate the control delay, AoG, and split failure without using any additional event-based data.

The Miovision sensors configured by PCDOT can collect performance measures from around 100 signalized intersections. These sensors can provide simple delay, AoG, arrival-on-red (AoR), and split failure through the TrafficLink portal. These Miovision-based traffic performance measures are collected from 2021 for analysis and can indicate the traffic condition. Also, these data serve as the ground-truth data for comparing with the Wejo-based traffic performance measures.

Comparing the Wejo-based performance measures and Miovision-based performance measures shows that the sample size of Wejo data significantly impacts the reliability and accuracy of Wejo-based performance measures. To control the quality of Wejo-based performance measures, QA/QC is conducted to determine the sample size threshold for ensuring a sufficient number of vehicle trajectories.

The traffic sensors at most intersections in the PAG region cannot provide performance measures due to the lack of a traffic data collection module. Event-based data, which is an existing data source generated by the default traffic detection module, is explored and utilized to estimate performance measures at these intersections without the data collection module. The proposed Model-Agnostic Meta-Learning (MAML) method was calibrated and evaluated using ground-truth data calculated using sufficient Wejo data for estimating delay and AoG. In order to ensure the quality of the estimated delay and AoG data, a QA/QC procedure was developed and acceptance criteria were decided based on literature review and sample data evaluation.

Data resource and limitation:

- Around 100 signalized intersections with Miovision sensors can provide traffic performance measures through TrafficLink portal. These data have been downloaded through TrafficLink portal, and no API is available currently.
- Due to a lack of advance detectors in Miovision sensors for left-turn movement, there is no AoG data available for the left-turn movement.
- In the greater Tucson area, there are around 700 signalized intersections managed by the MaxView system. The traffic sensors configured at these intersections are currently not capable of providing traffic delay and AoG data due to the detector configuration. However, these existing sensors are able to provide high-resolution event-based data, which is the major data source used for delay and AoG estimation in this project.

The following conclusions can be drawn from this project:

- The Wejo data can be used to provide control delay, AoG, and split failure without using other data sources. In addition, by combining Wejo data and signal event-based data, the red-light running information can also be extracted.
- According to the correlation analysis, the number of sample vehicles better indicates the necessary sample size of Wejo data than does the sample penetration rate. Comparing the Wejo-based performance measures and traditional sensor-based performance measures shows that the Wejo-based delay and AoG on through movements require at least 16 vehicles/hour to reach the maximum correlation coefficient.
- The Wejo-based delay on left-turn movements needs at least 6 vehicles/hour/lane to reach the maximum correlation coefficient. The two types of split failure data show a weak correlation because of excess zero split failure, which is challenging for determining the sample size threshold.
- By calculating the correlation between Wejo- and Miovision-based reliabilities, the delay reliability on through movements calculated using at least 500 vehicles can reach the maximum correlation coefficient. The delay reliability on left-turn movements calculated using at least 600 vehicles can reach the maximum correlation coefficient. The AoG reliability on through movements calculated using at least 1,000 vehicles can reach the maximum correlation coefficient.

-
- 153 signalized intersections in the greater Tucson metropolitan area are selected as study locations for calibrating and testing the proposed model-agnostic meta-learning method, which is an emerging AI method to address the impact of the intersection heterogeneity in terms of the traffic pattern and intersection layout on estimation accuracy and reliability. The evaluation results show that the through movements with advance and presence detectors in the MaxView system have a mean absolute percent error (MAPE) of 12%-15% and 15%-22% for control delay estimation, respectively, and the MAPE for left-turn movements is 22%-27%. For AoG estimation, MAPE for through movement with advance detectors is 13%-30%, and locations with presence detectors have a relatively higher error. The locations using the Miovision detection system also show a similar performance. The evaluation results indicate the proposed method can provide accurate and reliable network-level control delay and AoG regardless of the traffic detection system.
 - After aggregating into monthly hourly estimations, the estimation becomes more accurate and reliable. The MAPE of delay estimation is 9%-13% for through movements and 14%-18% for left-turn movements. The MAPE of AoG estimation is 7%-24% for through movements. The average monthly hourly AoG estimation for left-turn movement still has a relatively high error and lower reliability.
 - Since the through movements have accurate performance measure (PM) estimation but not left-turn and right-turn movements, it is challenging to obtain the intersection-level PM. Therefore, it is not recommended to estimate intersection-level using event-based data.
 - To understand the regional transportation system reliability, this study considered the buffer index of traffic delay, AoR, and split failure by comparing 95th percentile measure with the average measure. Buffer index shows a reliable system as close to 0 and the system would not be reliable as larger than 0.4. The buffer index shows worse performance at night and early morning, but the buffer delay index is under 0.2 for the daytime including morning and afternoon peaks and buffer AoR index shows similar patterns with worse performance at night and early morning but good performance under 0.2 in daytime. Buffer split-failure index was also evaluated but this study observed the scarcity of split failures in the region, and it is not recommended to measure split failure to represent system reliability.

Based on the conclusions above, we recommend the following to PAG regarding the use of existing data sources and future work:

- According to the representativeness evaluation results in Chapter 5, Wejo data with enough sample size can be used to derive accurate performance measures of control delay and AoG. Due to the limitation of sample size in using Wejo data, it is suggested to use Wejo data for monthly performance measures calculation in order to maximize the data usage and availability.

-
- The Miovision sensors in the greater Tucson area are able to provide valuable and reliable delay, AoG, and split failure data. So, it is recommended to consider and backup the provided Miovision performance measures.
 - Due to the heterogeneity of intersections, the delay and AoG estimation from the proposed meta-learning method using the event-based data is recommended to apply a network-level control delay and AoG. However, it is useful to review individual intersections with developed performance measures.
 - The developed acceptance criteria are based on the available data and references in this study. So, if there is any new crowd-sourced or traffic sensor data available, it would be suggested to evaluate the data and adjust the acceptance criteria accordingly.
 - Wejo data has been considered for calculating signalized intersection performance measures. However, beyond this project, it would be useful to understand other travel behaviors, such as travel purpose identification, parking behavior analysis, and parking time calculation.
 - An online data analytics platform would be a useful tool for visualizing and managing the estimated network- and intersection-level AoG and delay using event-based data.
 - This project and previous project focused on analyzing and estimating the performance measures for motorized traffic. It is also recommended to explore other existing data related to other traffic modes in Tucson, such as bike sharing, e-scooter, and transit, to obtain the performance measures for multi-modal transportation systems.

During the period of data collection and analysis for this report, Wejo was still considered a viable data source. However, in May 2023 following data collection, Wejo declared bankruptcy and ceased operations. Nevertheless, the analytical approach taken herein is considered sound in that it compares reliable crowd-sourced data to in-place, fixed systems. Results presented include an estimation of crowd-sourced (by percentage) data necessary to provide the region with a viable option to stationary (e.g. Miovision and Maxview) data collection systems.

For this study, Wejo data was collected for 8 months (January to April and September to December) in 2021, and Miovision and MaxView system data was compared with the collected Wejo data for the numerous analyses and calibration. We note that the travel behavior patterns in this period were influenced by COVID-19, and specifically that these travel behavior patterns differ significantly from patterns both before and after the pandemic period. However, as stated above, the purpose of this study was focused on comparing fixed data collection systems to connected vehicle data and travel behavior did not influence the results or conclusions of the analyses contained herein.

CHAPTER 1: INTRODUCTION

1.1 BACKGROUND

The Federal Highway Administration (FHWA), and Federal Transit Administration (FTA) encourage states, MPOs, and local governments to focus on transportation system management and operation, or TSMO, as a cost-effective set of strategies to address transportation challenges. As the MPO for the Eastern Pima County region of Arizona, the Pima Association of Governments (PAG) has developed TSMO strategies, which aim to improve the traffic operations of and optimize the safety, efficiency, and reliability of the existing transportation system. Performance measures enable PAG to track progress and quantify goals for TSMO strategies.

Getting the region-wide TSMO-related performance measures using multiple sources of data is challenging for most agencies, including PAG. Member jurisdictions are currently operating and maintaining various traffic sensors to collect data in the PAG region, including Max View and Miovision systems. Based on the previous project “*Comparative Analysis and Integration of Region-Wide Traffic Data*” exploring traffic sensor-based data, using these data to estimate TSMO-related performance measures at signalized intersections should be feasible and cost-effective. In addition to the traffic sensor data sources, third-party probe vehicle data can provide vehicle trajectory data using connected vehicle technologies and navigation software. The trajectory data has proven capable to calculate multiple traffic performance measures including intersection delay, level of service (LOS), arrival on green (AoG), and split failure. The sensor-based traffic data sources and third-party data sources are available within PAG and can be used to obtain TSMO-related performance measures.

The goal of this project is to investigate two sources of traffic data and develop a method to support PAG’s regional TSMO-related traffic data collection and advanced modeling. The performance measures will also be used in long-term planning, and the developed methodology guidelines will be efficiently applied to transportation planning and modeling and other regional traffic studies.

1.2 ORGANIZATION OF REPORT

The organization of the report is as follows. In **Chapter 2**, we summarize the TSMO-related traffic performance measure estimation methods, the application of the GPS data, and TSMO based on the literature review findings. In **Chapter 3**, we explore the third-party probe vehicle data in terms of data coverage and reliability, provide raw data cleaning and processing guidance, and develop traffic performance measure estimation methods using the processed trajectory data. **Chapter 4** calculates and compares analyzed mobility and reliability performance measures extracted from the Wejo data and the sensor-based data. The accuracy of the mobility performance measures extracted from the third-party probe vehicle data may vary with the penetration rate. **Chapter 5** evaluates the accuracy of the mobility performance measures under different penetration rates using other trajectory data collected by traffic sensors or ground-truth mobility data collected

manually. **Chapter 6** develops a method to estimate traffic performance measures using the sensor-based data from Max View and Miovision systems. To control the output data quality from the proposed estimation methods, acceptance criteria are developed with multiple sources of sample data based on the findings in **Chapter 7**. Based on the methods developed in previous tasks, we estimate the region-wide traffic mobility and reliability performance measures and summarize all results and findings in **Chapter 8**.

Table 1-1. Project tasks and schedule

	Task 9: Documentation	Task 8: Region-wide Traffic Mobility/Reliability Performance Estimation	Task 7: Acceptance Criteria Development	Task 6: Mobility/Reliability Performance Estimation using Max View and Miovision System	Task 5: Representativeness Evaluation of Performance Measures	Task 4: Mobility/Reliability Performance Data Collection and Analysis	Task 3: Third-Party Probe Vehicle Data Exploration	Task 2: Literature Review	Task 1: Kick-Off Meeting		
Feb-22	Dec-21										
Mar-22	Jan-22										
Apr-22	Feb-22										
May-22	Mar-22										
Jun-22	Apr-22										
Jul-22	May-22										
Aug-22	Jun-22										
Sep-22	Jul-22										
Oct-22	Aug-22										
Nov-22	Sep-22										
Dec-22	Oct-22										
Jan-23	Nov-22										
Feb-23	Dec-22										
Mar-23	Jan-23										
Apr-23	Feb-23										
May-23	Mar-23										
Jun-23	Apr-23										

CHAPTER 2: LITERATURE REVIEW

This project aims to investigate the sources of traffic data and develop a method and performance measures to support the PAG's regional TSMO-related traffic data collection/maintenance and advanced modeling. This project focuses on extracting the traffic performance measures from video-based and event-based traffic data and crowdsourced data. As a final output, the project develops cost-effective regional TSMO-related performance measures and enhances the QA/QC procedure and data integration based on the previous project "Comparative Analysis and Integration of Region-Wide Traffic Data," which explored the sources of regional traffic volume data and developed a model to estimate turning movement counts using event-based data.

This section summarizes the existing methods for traffic performance measure estimation, applications of GPS data in the transportation field, and major data sources and applications used in state TSMO programs.

2.1 TRAFFIC PERFORMANCE MEASURE ESTIMATION

Traffic performance measures such as traffic volumes and travel times are important indicators that quantify traffic conditions in order to monitor traffic and improve signal timing plans. However, these important performance measures are currently challenging to collect in real time for most transportation agencies. One of the major challenges is that the existing sensors currently used by most agencies are out of date and are unable to collect traffic performance measures. To obtain these performance measures without installing new sensors, some studies have been conducted to estimate performance measures from existing data sources, which can save a large amount of time and money.

Intersection volume data is a critical input for most traffic studies such as signal timing optimization. The traditional volume data collection methods using manual counts and intelligent sensors are time-consuming and costly (Li et al., 2019). Therefore, some studies have used existing data sources to estimate intersection volume. For example, GPS data has become more widely available and has been used to estimate volume. (Wang et al., 2019) proposed a framework that combines shockwave analysis and Bayesian networks to estimate intersection volume using trajectory data, with a mean absolute percentage error (MAPE) of 15%. In addition, (Zheng and Liu, 2017) and (Zhao et al., 2019) used GPS data collected from navigation apps to estimate through movement and left-turn volume at signalized intersections, and the results show that the model accuracy decreases as the penetration rate decreases. Also, it is common for only a limited number of sensors to be installed at sample locations rather than network-wide due to the high cost. Some studies (Sekula et al., 2018; Zhang et al., 2020) have used this limited volume data, collected from only a sample of all segments in the road network, in conjunction with crowdsourced data, that has wide coverage, to estimate network flow.

Delay and queue length are also important indicators of traffic conditions. (Bagdatli and Dokuz, 2021) used the data collected by camera sensors and controllers along with multiple machine learning methods to estimate intersection control delay. The evaluation results show that XGBoost

yields the highest accuracy, with a mean absolute error (MAE) of 2.8 seconds/vehicle. And machine learning-based methods are also commonly applied to estimate queue length using existing data sources such as GPS data and license plates (Liu et al., 2022; Zhan et al., 2015). Shockwave-based methods are also used for queue length estimation (An et al., 2017; Hao et al., 2014; (Jeff) Ban et al., 2011; Liu et al., 2009). In addition to intersection-level performance measures, corridor-level performance measures such as speeds and travel times are used for traffic studies, though they are challenging to collect. Therefore, various data sources such as loop detector data (Xiao et al., 2018), radar detector data (Ding et al., 2019), and GPS data (Bahuleyan and Vanajakshi, 2017; Wang et al., 2018) are leveraged to estimate speeds and travel times. In addition, some products on the market, such as Miovision TrafficLink (Miovision, 2022a), CATT Lab RITIS (CATT Lab, 2022), and INRIX Signal Analytics (INRIX, 2021), can provide various traffic performance measures, which are summarized in **Table 2-1**.

Table 2-1. Performance Measures Summary from Miovision TrafficLink, CATT Lab RITIS, and INRIX Signal Analytics

	Performance Measures	Interpretation	Data Sources	Resolution
Miovision TrafficLink	Simple Delay	The time between stop-bar detector actuation during the red phase and when the phase turns green	Video-based sensors	15 minutes
	Travel Time	Travel time between two intersections with Miovision sensors		Vehicle-based
	Arrivals on Red	The number of vehicles triggering advance detectors during red		5 minutes
	Green Allocation	Percentage of the green time allocated among the phases		15 minutes
	Occupancy Ratio	The percentage of time that the stop-bar detector is occupied during red, green, and the first 5 seconds of red		Cycle-based
	Pedestrian Compliance	Pedestrian red-light runner		Event-based

	Performance Measures	Interpretation	Data Sources	Resolution
	Pedestrian Delay	The time between pedestrian detector actuation during the Don't Walk phase and when the phase turns to Walk		Cycle-based, Binned into 15-min interval
	Red-Light Runner	Number of red-light runners using stop-bar detections		Event-based
	Split Failure	When both the Green Occupancy Ratio and the first 5 seconds of Red Occupancy Ratio are high (>80%)		Cycle-based
	Split Trend	Categories of split failure		15 minutes
CATT Lab RITIS	Speed (mph)	Average speed in the specified interval	HERE; INRIX; NPMRDS; TomTom	1 minute, 5 minutes, 10 minutes, 15 minutes, 1 hour, Day of week
	Congestion (%)	The percentage of vehicle miles traveled (VMT) under congested conditions in the specified interval		
	Buffer time (minutes)	The extra travel time that travelers must allow to arrive at a destination by the intended time on 95% of trips		
	Buffer index	The ratio of the buffer time to the average travel time		
	Planning time (minutes)	The total travel time that travelers must allow to arrive at a destination by the intended time on 95% of trips		

	Performance Measures	Interpretation	Data Sources	Resolution
	Planning time index	The ratio of the 95th percentile travel time to the uncongested (free flow) travel time		
	Travel time (minutes)	The time spent travelling		
	Travel time index	The ratio of the average travel time to the free-flow travel time		
INRIX Signal Analytics	Control delay	The difference between the travel time a vehicle reports to traverse a movement and the reference travel time	INRIX trip data; vehicle GPS waypoint data	Real-time (1-minute)
	Level of service	A qualitative measure describing operational conditions within a traffic stream, based on control delay		
	Arrival on green	The percentage of vehicles that traversed a movement without stopping		
	Approach speed	The highest speed reported by a vehicle within the 150-meter approach zone		
	Vehicle count	The total number of sampled vehicles in the database for a specified movement and time period		

	Performance Measures	Interpretation	Data Sources	Resolution
	Stopped vehicle count	The number of vehicles in a vehicle count that stopped at least once in the approach zone of an intersection (a vehicle has “stopped” if it reported a speed at or below 10 kph/6 mph)		
	Split failures	The number of occurrences when a green light fails to meet vehicle volume demand, resulting in a vehicle stopping more than once at a traffic light		

2.2 GPS DATA APPLICATION

According to the literature review of traffic performance measure estimation, most studies have used GPS data collected from connected vehicles and smart phone apps. This GPS data source has been attracting an increasing amount of attention and has been widely used in the transportation field. Traditionally, engineers or volunteers had to personally drive vehicles multiple times on routes through the study locations on the desired days at the desired times to collect sample GPS data. This sample GPS data was then used to evaluate traffic performance changes due to any new facilities and policies such as the installation of adaptive signal control (Hunter et al., 2012; Khattak et al., 2019; Tian et al., 2011). But this sample GPS data can only cover limited time periods and locations, and the collection process is also time-consuming.

Recently, some third-party companies have begun to provide 24/7 GPS data on a network level. For example, Wejo can provide connected vehicle trajectory data of individual vehicle waypoints with a 3-second interval in most cities. This third-party data has a penetration rate of around 5% on U.S. roads (Hunter et al., 2021). Due to the wide coverage and accessibility of this third-party data, most studies have used this data instead of traditionally collected sample GPS data to evaluate traffic performance. Wejo data has been used to calculate split failure, travel time, hard acceleration, control delay, speed, level of service, and stop delay (E. Saldivar-Carranza et al., 2021a) to (1) evaluate the impact of construction work zones (Desai et al., 2021); (2) evaluate changes to left-turn phasing (E. D. Saldivar-Carranza et al., 2021b); (3) evaluate the impact of speed signs (Mathew et al., 2021); (4) analyze roundabout performance (E. Saldivar-Carranza et al., 2021b); (5) assess diverging diamond interchanges (E. D. Saldivar-Carranza et al., 2021a). Because this trajectory data is only collected from a sample of all vehicles and not from all vehicles, some

studies used this sample data to estimate and predict traffic volume (Abdelraouf et al., 2022; Shoman et al., 2022), which enhances the functionality of Wejo data in deriving performance measures.

Even though third-party GPS data has been widely used to estimate performance measures, the accuracy of these derived performance measures has not been evaluated since the accuracy could vary with the penetration rate. Therefore, the relationship between the penetration rate and the accuracy of these derived performance measures needs to be investigated before using these measures for any traffic study.

2.3 APPLICATIONS OF TRANSPORTATION SYSTEMS MANAGEMENT AND OPERATIONS

Transportation system management and operations (TSMO) provides public agencies with a growing toolbox to address traffic congestion issues, traffic safety, system performance, and reliability and support long-term goals for the transportation system. TSMO is “an integrated set of strategies to optimize the performance of existing infrastructure through the implementation of multimodal and intermodal, cross-jurisdictional systems, services, and projects designed to preserve capacity and improve security, safety, and reliability of the transportation system” (FHWA, 2020). Some example TSMO strategies include (FHWA, 2020): Work Zone Management, Traffic Incident Management, Special Event Management, Road Weather Management, Transit Management, Freight Management, Traffic Signal Coordination, Traveler Information, Ramp Management, Congestion Pricing, Active Transportation and Demand Management, Integrated Corridor Management, Access Management, Improved Bicycle and Pedestrian Crossings, Connected and Automated Vehicle Deployment, Mobility on Demand. To make these strategies more efficient, effective, and robust, multiple data sources collected from a regional area are required to support the TSMO program. We summarized TSMO applications and the major data sources used for mobility and safety performance measures in different states in **Table 2-2**.

Table 2-2. Applications and Major Data Sources for TSMO Programs

State DOT	Major Data Sources
Arizona DOT (Kimley Horn, 2017)	Crash Reports; Loop Detectors; Regional Archive Data System (RADS); Highway Condition Reporting System; Weather
Florida DOT (FDOT, 2017)	Probe Vehicle Data; Traffic Detectors; Crash Reports; SunGuide Event Data; RITIS
Iowa DOT (Lakeside Engineers, 2016)	Wavetronix Detectors; INRIX; TransSuite; Waze; State Crash Records; Iowa DOT Motor Vehicle Division records; Weather
Minnesota DOT (MnDOT, 2019)	CAV data; Weather; Waze; INRIX; HERE; Google
Nevada DOT (NDOT, 2020)	State Crash Records; INRIX; Waze; Google; American Community Survey; VMT; Event-based Data; Waycare

State DOT	Major Data Sources
Oregon DOT (ODOT, 2021)	RITIS/INRIX; Event-based Data; Detector Data; Oregon State Police; Fire and Emergency Responders; 911 Emergency Dispatch; Tow Companies
Tennessee DOT (Cambridge Systematics, 2022)	INRIX; RITIS; Radar; Waze; GPS data; CAV TIM Reports; FHWA Performance Reports
Ohio DOT (Gannett Felming and BPS, 2017)	ATMS data; Crash Records; HERE

Artificial intelligence (AI) and machine learning processing methods have been widely used to advance TSMO because AI has the capacity to process multiple-sourced, large-scale, real-time data to model system behaviors, predict traffic conditions, and evaluate system performance. The use of AI technology for TSMO applications by several state DOTs has been reviewed and summarized in **Table 2-3**. The results of these applications show that AI methods can improve traffic safety and reduce the traffic staff workload. Some researchers also studied how to use AI technology to advance TSMO strategies, as summarized in **Table 2-4**. AI methods have been used for different TSMO applications, including management of different traffic modes such as connected vehicles, freight, transit, and pedestrians, and different traffic studies such as traffic safety and traffic signal coordination.

Table 2-3. Example Applications of AI Technologies for TSMO Applications

Agencies	TSMO Applications	AI method	Description	Benefits
Nevada DOT and Florida DOT (Gettman, 2019)	Traffic Incident Management	Neural Networks	Combine information from a variety of sources to detect and report suspected incidents.	Reduced the incidence detection time by 12 minutes and reduced crashes by 17%
Washington DOT (Meldrum and Taylor, 2000)	Ramp Management	Fuzzy Logic	Calculate a metering rate.	Improve the mainline efficiency
Delaware DOT (DelDOT, 2022)	Traffic Incident and Congestion Management	Neural Networks	Predict traffic, detect incidents, automatically disseminate traveler information.	Reduce the operator workload
Iowa DOT (Stolle, 2018)	Road Weather Management	Decision tree	Predict winter road conditions.	Field maintenance staff plan efficiently for winter road condition changes
Iowa DOT (Liu et al., 2021)	Work Zone Management	Artificial Neural Network	Predict work zones' traffic impacts.	Support work zone planning and management
Washington State DOT (Murthy and Yang, 2021)	Freight Management	Recurrent Neural Network	Predict parking occupancy.	Improve utilization of parking facilities

Table 2-4. Example Research of AI Technologies for TSMO Applications

TSMO Applications	AI Technologies	Research Area	Findings
Incident management (Chowdhury et al., 2006)	Support vector regression (SVR) and case-based reasoning (CBR)	Assess the impact of diversion strategies in response to incidents.	SVR performance is superior to CBR

TSMO Applications	AI Technologies	Research Area	Findings
Emergency management (Matveev et al., 2020)	Artificial neural network (ANN)	Assess the scale of an incident or emergency.	ANN identifies standard incidents similar to the current incident.
Freight management (Mahmud et al., 2020)	K-means clustering	Identify truck parking patterns.	Overnight and longer-duration parking was associated with facilities that had fewer amenities.
Connected and Autonomous vehicles (CAVs) (Ren et al., 2020)	Reinforcement learning (RL)	Adjust longitudinal position to find a safe gap to merge in the open lane at work zones.	The RL-based model outperforms conventional merge control strategies including late and early merges.
Freeway management (Sadek et al., 1998)	Stochastic search algorithm and CBR	Identify routing strategies to optimize highway network performance.	AI algorithms allow real-time information processing, experience learning, and deal with missing and incomplete data.
Transit management (Wai and Zhou, 2020)	eXtreme Gradient Boosting (XGBoost)	Predict bus departure times.	AI handles real transit agency challenges including missing models, timing points, and partially traveled segments.
Traffic signal coordination (Choy et al., 2003)	Reinforcement learning, fuzzy logic, and neural network	Recommend an appropriate signal policy at the end of each phase.	The AI-based system reduced delays and total vehicle stoppage time more than fixed-time traffic signal control.
Pedestrian and bicycle network (Lu et al., 2011)	Fuzzy logic	Determine dynamic pedestrian clearance interval while fulfilling multifaceted vehicle needs at intersections.	The fuzzy logic control system outperforms the NEMA control.

The integration of connected vehicle (CV) technologies and machine learning (ML) into transportation systems management and operations (TSMO) projects has been a significant development in recent years. As shown in **Table 2-5**, multiple transportation agencies have created a CV test-bed environment with pilot field locations to implement a real-world TSMO solution. The CV technology has been used for several TSMO strategies such as transit management, traffic safety management, and traffic signal performance. In addition to using different types of technologies to advance TSMO, other measures to improve TSMO include integrating ITS with TSMO and expanding TSMO capability for multimodal transportation. Four state DOTs with mature systems and technology programs within their TSMO programs are summarized in **Table 2-5**. These agencies provide valuable experience in using technology to enhance transportation network efficiencies and operations.

Table 2-5. Case Studies for Advancing TSMO

Projects/Cases	Practices	Implementation and Experience
Connected Streetcar Project in Portland (NITC, 2018)	Take CV technology for a Test Run on Portland Streetcar.	Deploy the MMITSS on the Portland Streetcar and at four intersections along the study corridor.
New Jersey Connected Technology Integration and Implementation (New Jersey Department of Transportation, 2022)	Deploy and integrate CV equipment into the existing NJDOT ITS architecture.	Deploy CV technology at five intersections with SCATS-enabled traffic signals.
Gainesville Signal Phase and Timing Trapezium Project (Banerjee et al., 2021)	Deploy Roadside and Onboard CV technologies on four roads to improve travel time reliability and safety.	<ul style="list-style-type: none"> • Deploy 27 RSUs and 71 OBUs on a variety of vehicles including emergency vehicles, transit buses, UF fleet, the City of Gainesville vehicles, and research vehicles. • Deploy a smartphone-based alert application to enhance bicyclist and pedestrian safety.

Projects/Cases	Practices	Implementation and Experience
Manatee County's Expansion of TSMO Capabilities to Enhance Multimodal (Manatee County, 2022)	Expand TSMO capabilities on multimodal transportation systems.	<ul style="list-style-type: none"> • Deploy ATSMPS at 75 signalized intersections. • Implement ASCT. • Disseminate traveler information using crowd-sourced platforms and websites. • Include RRFBs and LPI as part of the standard design increasing safety for pedestrians.
Four state DOTs with mature systems and technology programs within their TSMO programs (Atkins et al., 2019)	Make full use of the systems and technology component of TSMO.	<ul style="list-style-type: none"> • Oregon DOT: implement a statewide ITS architecture and embed the ITS architecture and planning efforts within broader regional planning efforts. • Georgia DOT: a Qualified Product List containing all the ITS equipment. • Ohio DOT: Move ITS network routing and switching from OIT to TSMO making coordination and communication easier. • Utah DOT: 1) Built PPP with private telecommunications companies to reduce the cost of fiber construction for UDOT; 2) Upgrade signals for ATMPS and include all traffic signals on the same communication network to enable UDOT to make decisions on a large scale.

CHAPTER 3: THIRD-PARTY PROBE VEHICLE DATA EXPLORATION

Third-party probe vehicle data (i.e., Wejo data) can provide vehicle trajectory data using connected vehicle technologies and navigation software. This chapter focuses on cleaning and processing the raw GPS data. Because the data can provide an individual vehicle's trajectory, individual driving behaviors can be mined and extracted from these trajectories. These individual driving behaviors can be used to indicate traffic conditions and performance. The feasibility of extracting potential traffic performance measures from the third-party probe vehicle data is investigated in this section.

3.1 DATA CLEANING AND PROCESSING

The raw Wejo GPS data contains all GPS points both on and off the road network, including the parameters of latitude, longitude, speed, moving direction (angle), timestamp, and trip ID. However, the raw data does not contain information about the intersections through which the vehicle passed nor does it contain the vehicle's direction (EB, WB, NB, or SB) or movement (left/right turn or through) at the intersection, which are necessary to calculate intersection-level performance measures. We first filter all on-road GPS points with a GIS map of most arterials in the Greater Tucson area to select only those points within a 50 m buffer of the arterials. As shown in **Figure 3-1(a)**, the adjacent intersections of the target intersection are used to filter vehicle trips that passed through or near the target intersection from all trips. These filtered GPS points are then further processed to add more information, and the processing procedure is shown in **Figure 3-1(b)**.

- **Generate new trips:** After filtering the GPS points using adjacent intersections, a few original trips such as back-and-forth trips with the same trip ID may be separated into multiple sub-trips in the study area. For example, if one trip is EB through movement, when this trip makes a U-turn at the next intersection, this trip then becomes WB through. This trip can be used to indicate EB and WB traffic conditions, so we split it into two trips. Specifically, when two consecutive GPS points in an original trip after trimming trajectories based on adjacent intersections have a time gap longer than 10 seconds, these two points are classified into two sub-trips with new trip IDs assigned.
- **Identify direction:** For most standard intersections like **Figure 1(a)**, the directions in which vehicles were driving can be identified based on the “direction” parameter in the data.
 - NB: $315^\circ \leq \text{direction} < 360^\circ$ or $0^\circ \leq \text{direction} < 45^\circ$
 - EB: $45^\circ \leq \text{direction} < 135^\circ$
 - SB: $135^\circ \leq \text{direction} < 225^\circ$
 - WB: $225^\circ \leq \text{direction} < 315^\circ$

There are some skewed intersections where the vehicle directions through the intersection cannot be identified using the method above. For these intersections, we first manually label the road angles using Google Maps and then identify the vehicle directions.

- **Identify turning movement:** The turning movements of vehicles at intersections can be determined from the trip directions identified in the previous step.

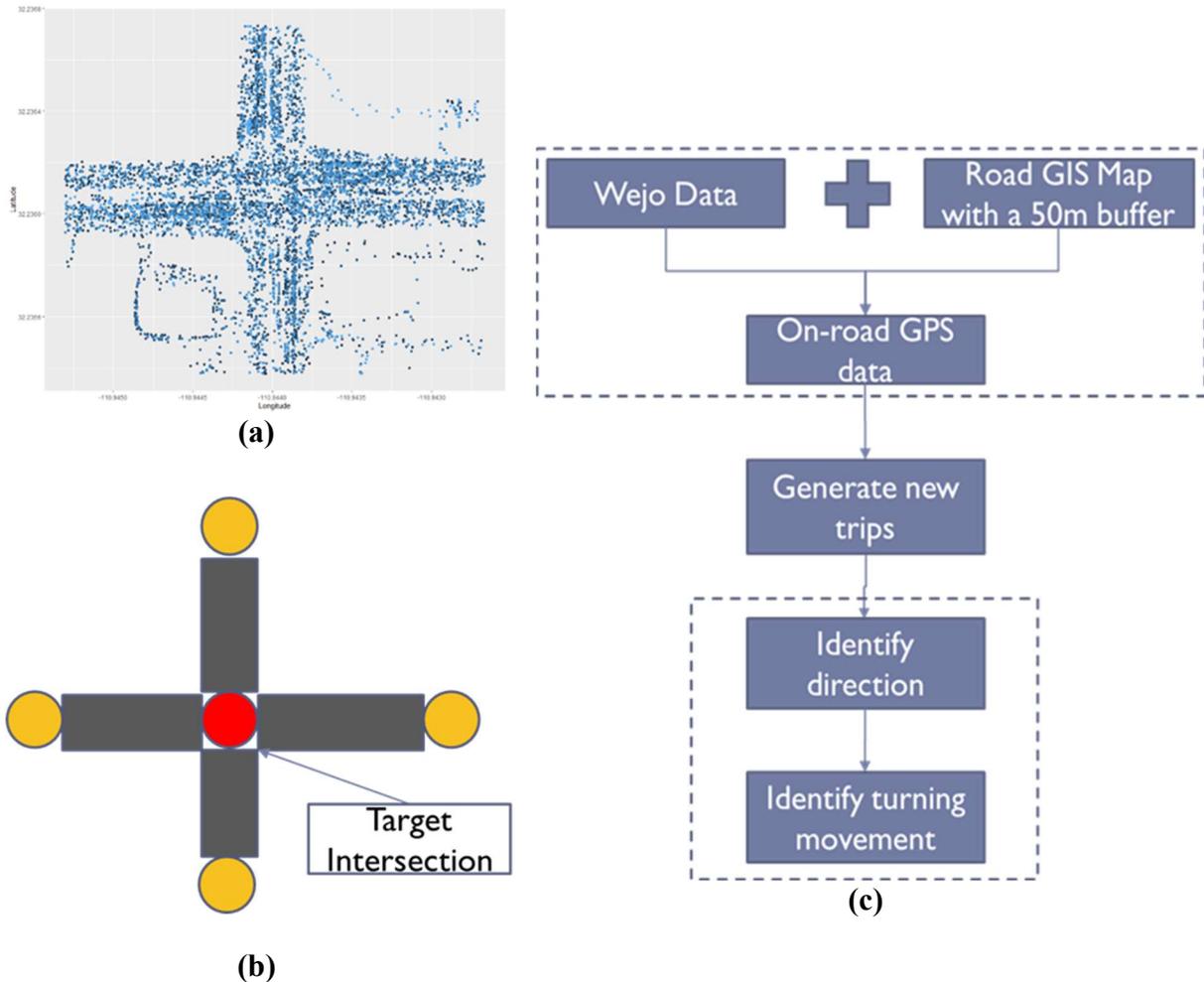


Figure 3-1. (a) Raw GPS points at one sample intersection; (b) on-road GPS data filter; (c) data processing procedure

3.2 TRAFFIC DELAY

3.2.1 Calculation

After cleaning and processing the raw probe vehicle data, it is used to calculate the traffic delay. Control delay consisting of deceleration delay, stop delay, and acceleration delay is commonly

used to indicate traffic performance at signalized intersections (Ko et al., 2008; E. Saldivar-Carranza et al., 2021a). Most studies applied probe vehicle data for delay calculation to compute the time difference between actual travel time and free-flow trajectory's travel time as the control delay, where the free-flow speed is assumed to be the posted speed limit. Even though the free-flow speed is correlated with the posted speed limit, these two speeds cannot be regarded as the same because drivers might drive slightly under or over the posted speed limit in the free-flow condition (Deardoff et al., 2011; Silvano et al., 2020). To compute the actual free-flow speed, the speeds of vehicles passing through an intersection without stops during nighttime (10 p.m. – 3 a.m., hours that typically have low traffic flow) are averaged. This average speed is then considered the free-flow speed for each movement at that intersection because these vehicles are not impacted by signal control and queues.

3.2.2 Results and Analysis

We calculate the hourly delay as the average of all vehicles' delay that pass through the target intersection within one hour, and the number of sample probe vehicles is also obtained. **Figure 3-2** shows hourly delay and number of sample probe vehicles for the through movement at Ina Rd. and La Cañada Dr. for three days in Jan. 2021. These three days of data consistently show that nighttime has very few and sometimes even zero probe vehicles, and so the delay cannot be calculated for many nighttime hours. During the daytime, more probe vehicles pass through the target intersection, where the number of sample probe vehicles is as high as 30 vehicles/hour for some directions during some peak hours. The hourly delay is slightly longer during peak hours than during other daytime hours, but overall the delay is similar among the various daytime hours, with most being shorter than 50 seconds.

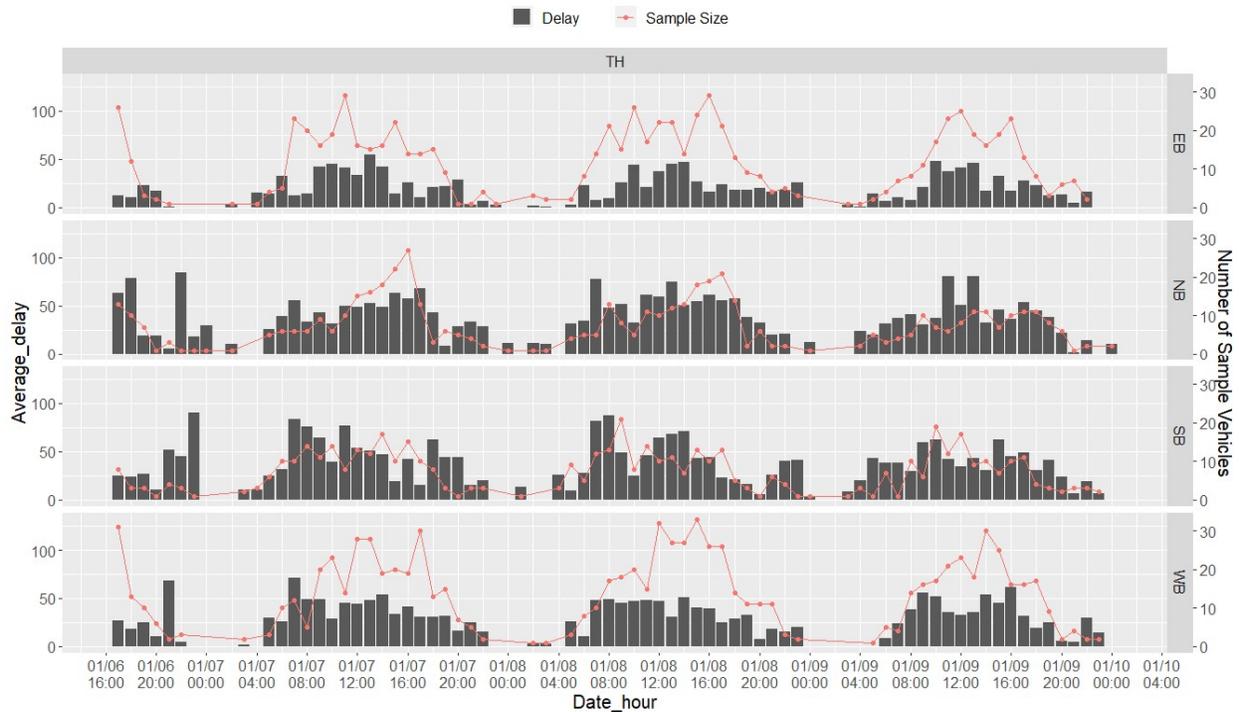


Figure 3-2. Hourly delay and number of sample probe vehicles at Ina Rd. & La Cañada Dr.

In addition to the through movement delay, the probe vehicle data can be used to calculate the left-turn and right-turn delay. **Figure 3-3** shows the average over the month of Jan. 2021 for the through, left-turn, and right-turn hourly delay at Ina Rd. and La Cañada Dr. These three movements have a similar hourly delay trend, with shorter delay during nighttime, longer delay during daytime, and peaks in the early morning and late afternoon. Also, left-turn vehicles typically experience the longest delay at the intersection and right-turn vehicles the shortest, which corresponds to the typical perceptions.

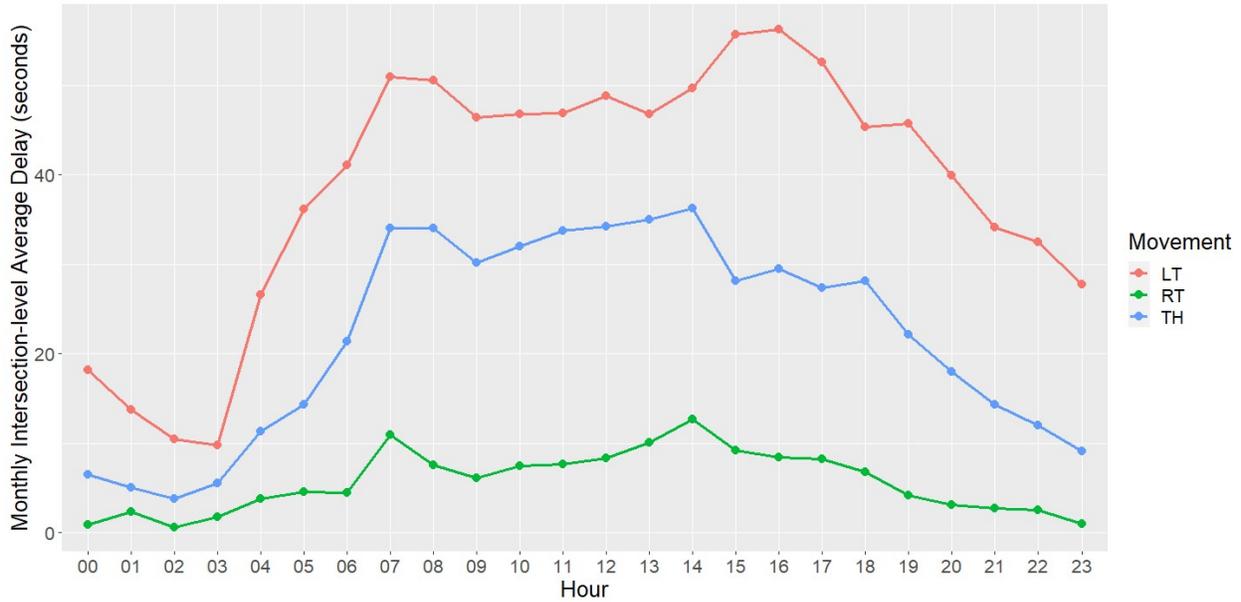
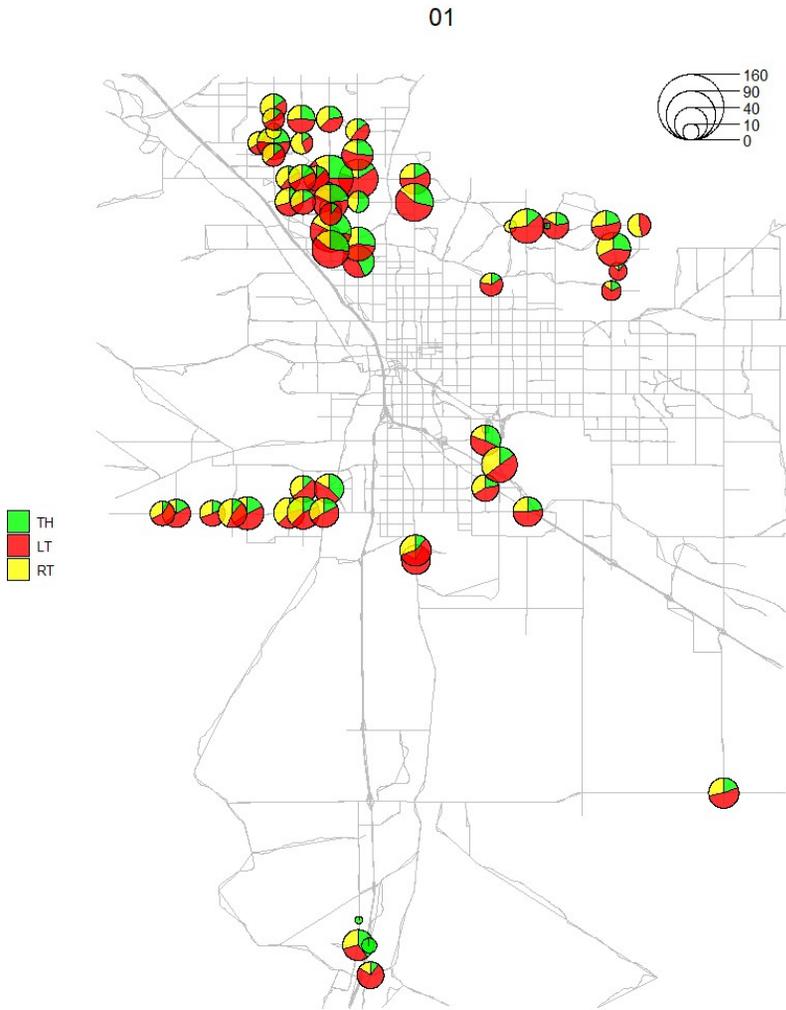
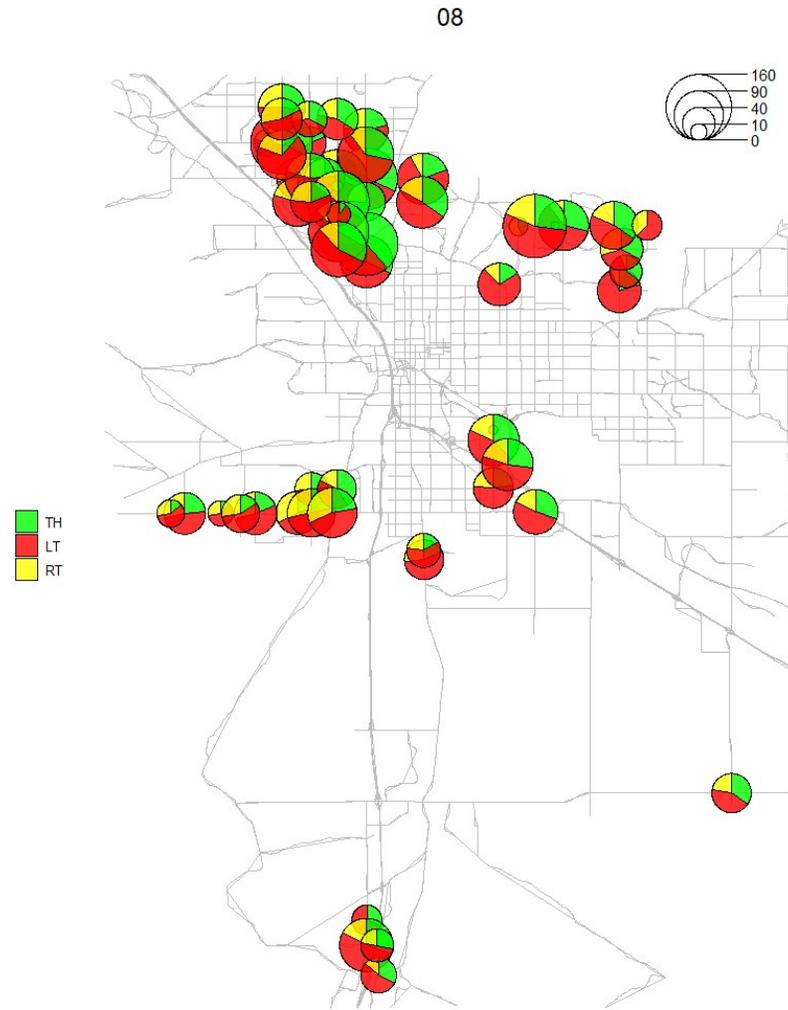


Figure 3-3. Average hourly control delay by movement at Ina Rd. & La Cañada Dr. during Jan. 2021.

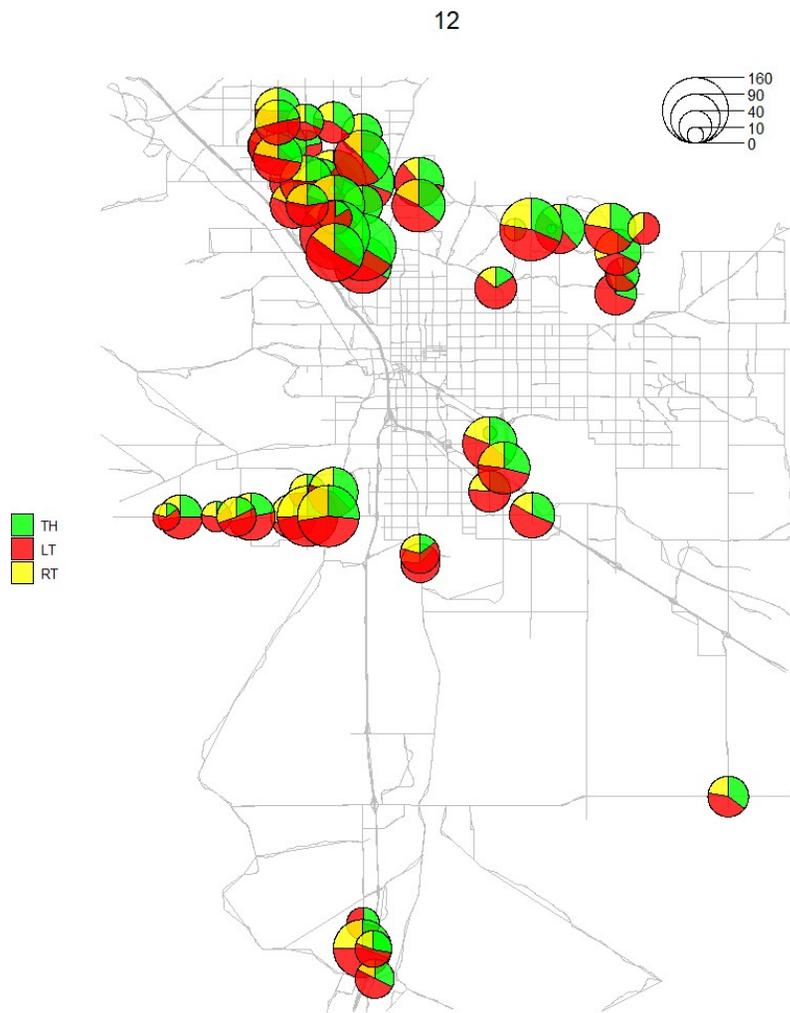
64 signalized intersections in the PAG region were selected, and the control delay at each intersection during January, February, and March of 2021 is calculated. **Figure 3-4** shows the spatial-temporal distribution of average delay for these months at these intersections, where the radius of the circle is the total of the average through, left-turn, and right-turn delay. The delay at most intersections at midnight is shorter than 40 sec/veh due to the light traffic. During the daytime, the total delay at most intersections becomes significant. In most situations, left-turn vehicles experience longer delay than through and right-turn vehicles, and often the left-turn delay is even longer than the sum of the through and right-turn delay.



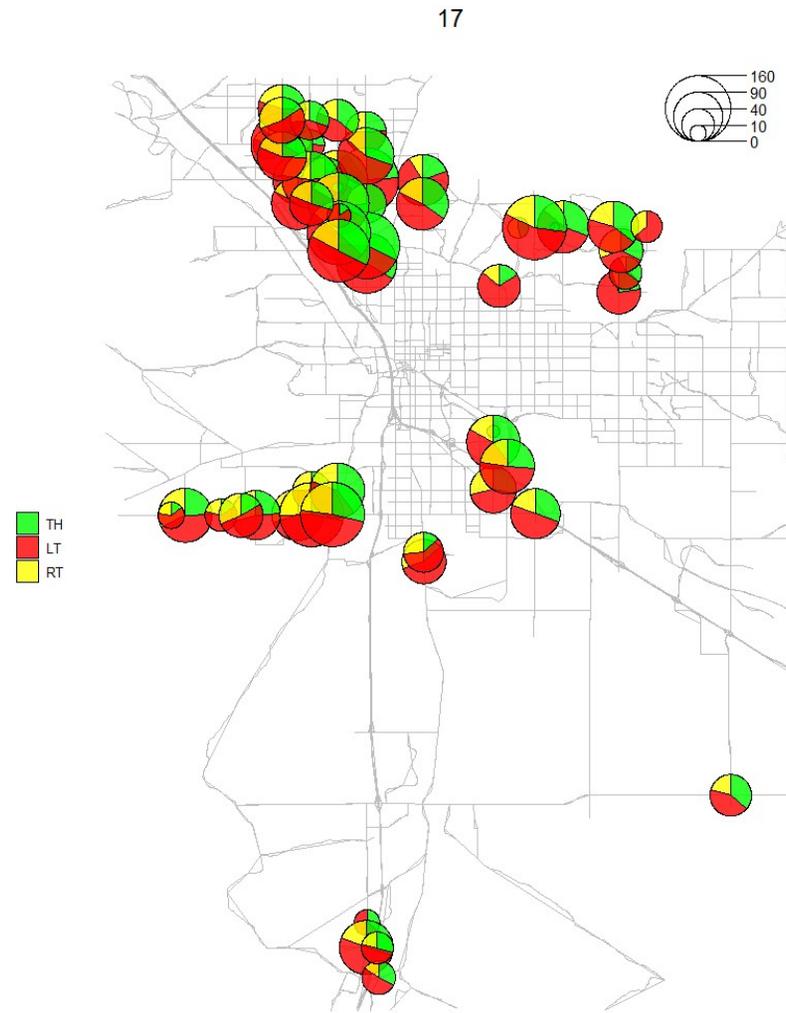
(a) delay distribution at 1 a.m.



(b) delay distribution at 8 a.m.



(c) delay distribution at 12 p.m.



(d) delay distribution at 5 p.m.

Figure 3-4. Spatial-temporal distribution of delay

3.3 LEVEL-OF-SERVICE

3.3.1 Definition

According to the 2016 Highway Capacity Manual (Transportation Research Board, 2016), the level of service (LOS) at a signalized intersection is defined in terms of the average vehicle delay of all movements (through, right-turn, and left-turn) that occur at the intersection. The definition of the LOS criteria is summarized in **Table 3-1**.

Table 3-1. Signalized intersection LOS Criteria

LOS	Average Delay (Sec/veh)	Description
A	≤10	Free flow
B	10 - 20	Stable flow (slight delays)
C	20 - 35	Stable flow (acceptable delays)
D	35 - 55	Approaching unstable flow (tolerable delay, occasionally wait through more than one signal cycle before proceeding)
E	55 - 80	Unstable flow (intolerable delay)
F	>80	Forced flow (jammed)

Source: (Transportation Research Board, 2016)

3.3.2 Results and Analysis

Figure 3-5 shows the intersection LOS distribution by hour, based on the delay data on March 17, 2021. The LOS at most intersections during most hours is A. At some intersections, the LOS decreases to B or C after 6 a.m. From 11 a.m. – 5 p.m., the number of intersections with LOS A, the number with LOS B, and the number with LOS C are similar. During every hour, there are few, if any, intersections with LOS of D or E.

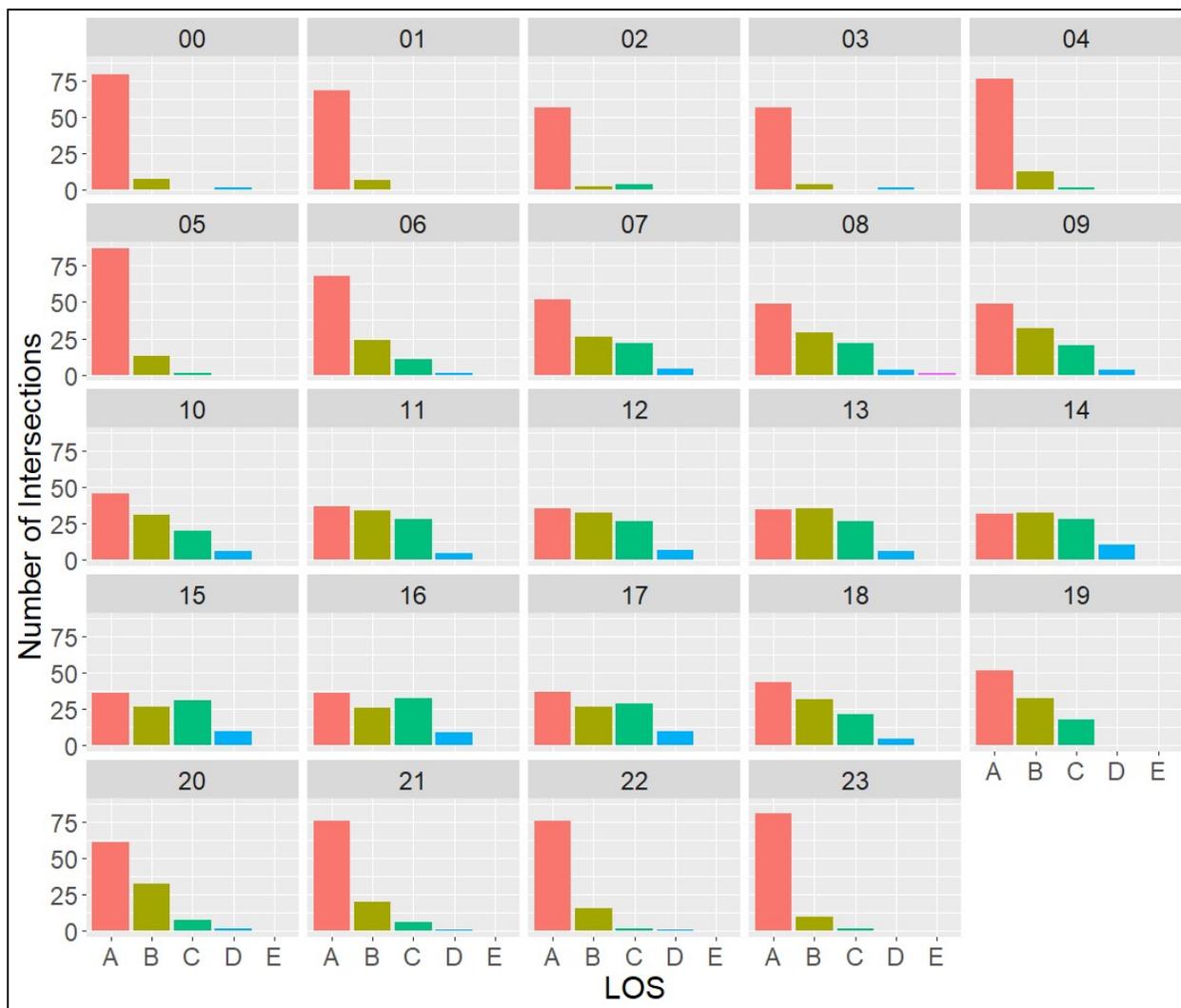


Figure 3-5. Intersection LOS distribution by hour on March 17, 2021

3.4 ARRIVAL-ON-GREEN/ARRIVAL-ON-RED

3.4.1 Calculation

Arrival on green (AoG) is the percentage of vehicles arriving at an intersection during the green phase and is used to indicate the signal coordination progression. Advance detectors are typically configured to count the number of vehicles triggering the detector during the green phase to calculate the AoG (Brennan et al., 2011; Day et al., 2014, 2008).

The distance of advance detectors from the intersection varies among intersections and approaches, and so the AoG percentages calculated from counts collected using advance detectors may lack the consistency needed to accurately compare the signal performance at different locations. In order to ensure consistency for performance evaluation, we applied the method proposed by (E.

Saldivar-Carranza et al., 2021a), which calculates AoG using probe vehicle data rather than detector data. The AoG is calculated as the ratio of the number of sample vehicles passing through an intersection without stopping to the total number of sample vehicles passing through the intersection during a defined time interval. The stopping status is identified as when a vehicle has a speed value lower than 1 mph before passing through an intersection.

3.4.2 Results and Analysis

The Arrival-on-Green (AoG) and Arrival-on-Red (AoR) for the 64 selected intersections during January, February, and March of 2021 are calculated and analyzed in this section. **Figure 3-6** shows a heatmap created from the average AoG ratio for each intersection and hour of the day. The AoG of 0 means all vehicles arrive during the red time, and the AoG of 100% means all vehicles arrive during the green time. Most intersections have a high AoG ratio (over 80%) during the nighttime, indicating free flow traffic during these hours, which is consistent with the delay results. Also consistent with the delay analysis, the AoG decreases after 5 a.m. and has a relatively low value during the daytime. In addition, major arterials such as La Cholla Blvd. and Ina Rd. typically have a lower AoG, indicating that these arterials are more likely to have congestion, which could be due to inefficient signal timing, heavy traffic, or both. **Figure 3-7** shows the average spatial-temporal distribution of AoG and AoR ratios. All intersections on the same arterial have similar AoG and AoR, which is likely because most or all signals on a major arterial have similar traffic conditions, and often the same signal timing plan is used to coordinate most or all signals on the corridor.

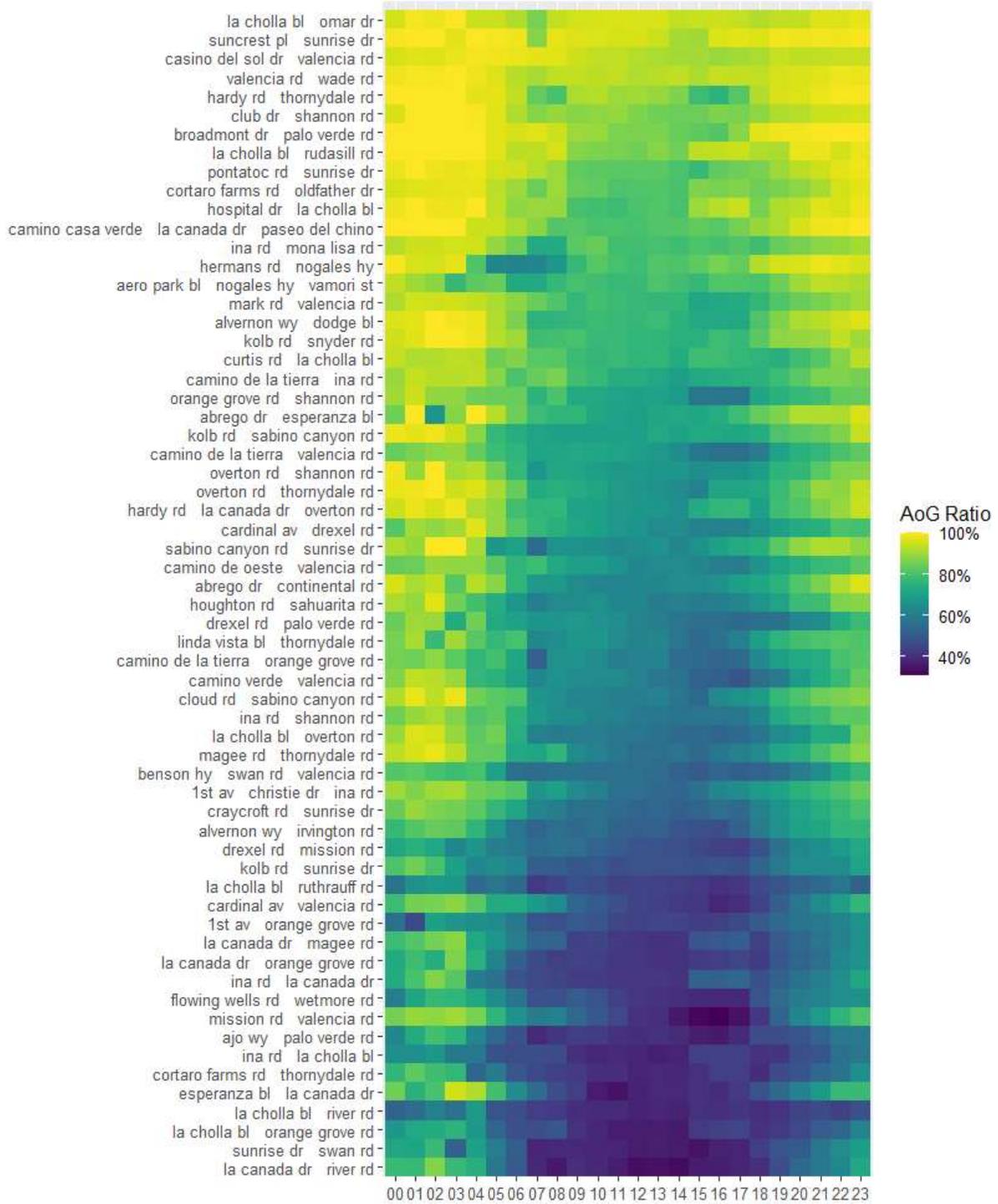
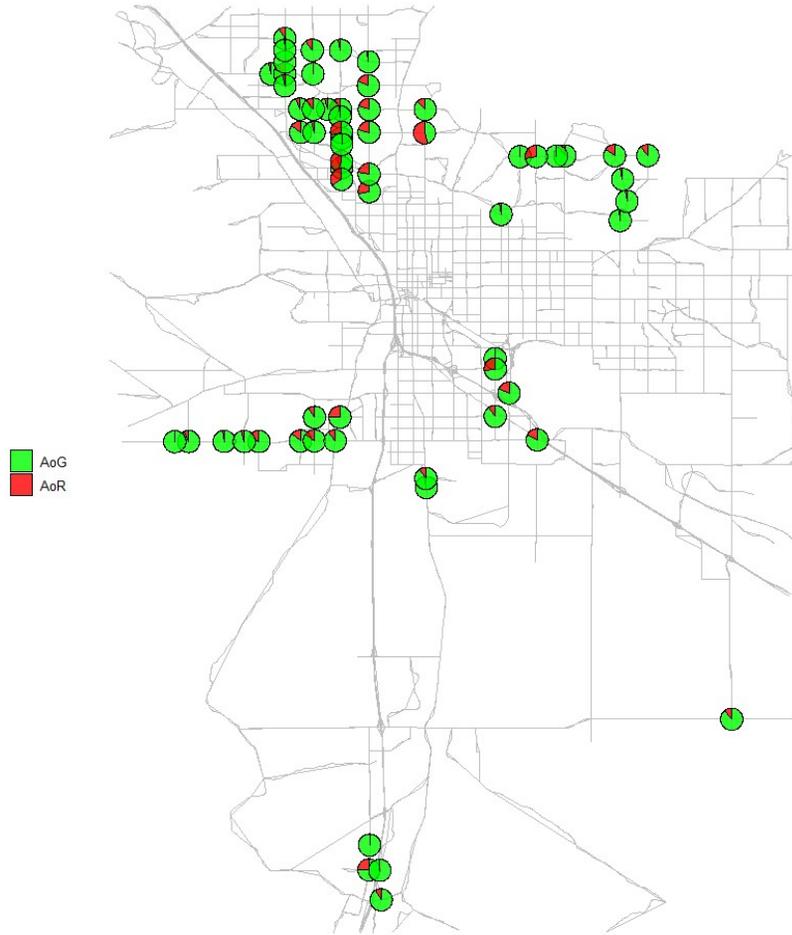


Figure 3-6. Heatmap of average AoG ratio by intersection and hour

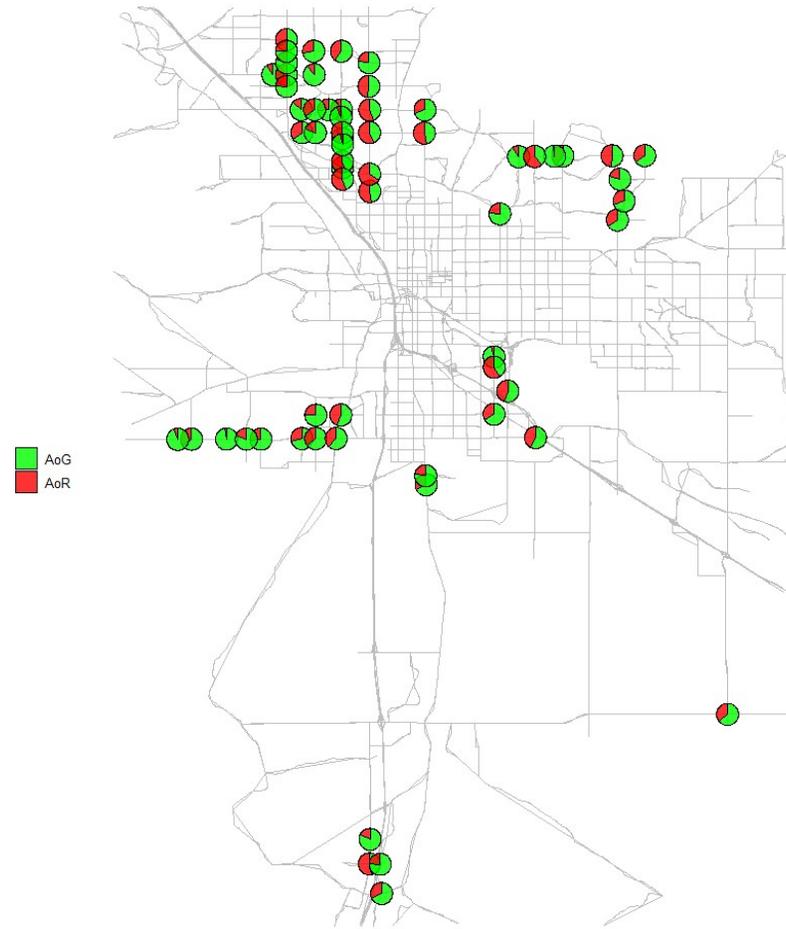


01



(a) AoG and AoR distribution at 1a.m.

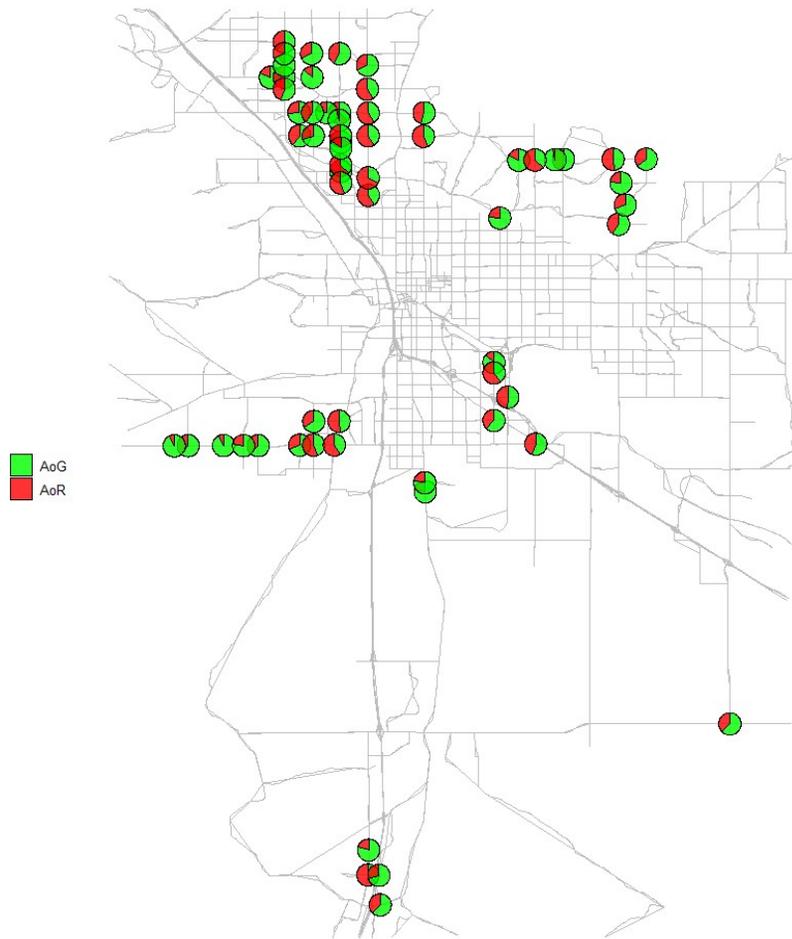
08



(b) AoG and AoR distribution at 8 a.m.

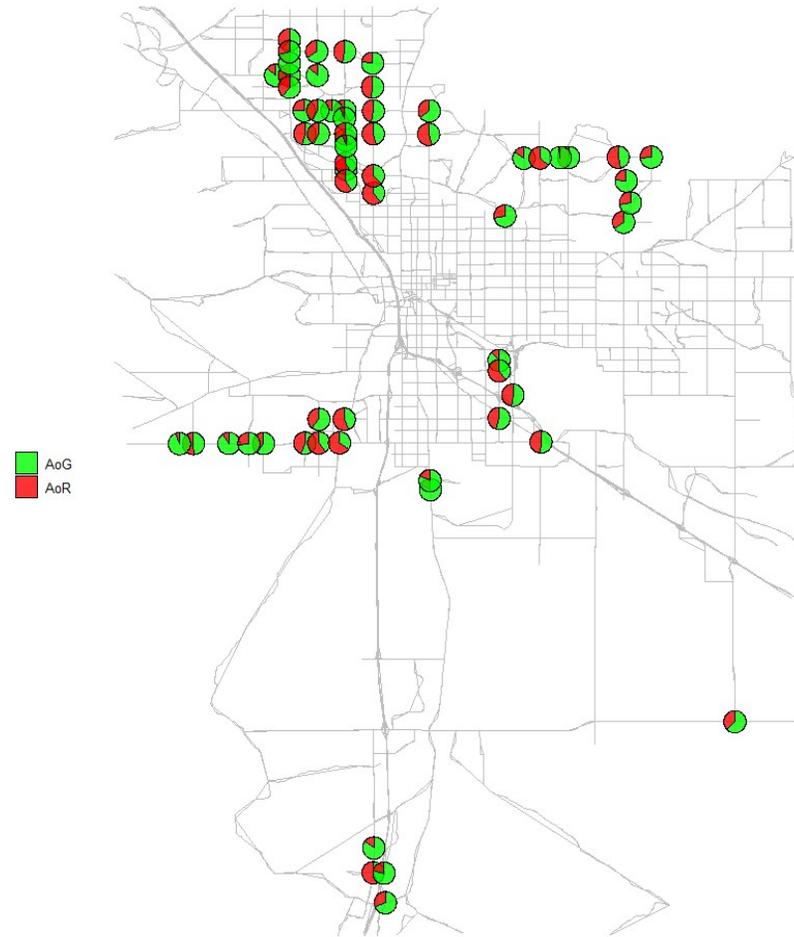


12



(c) AoG and AoR distribution at 12 p.m.

17



(d) AoG and AoR distribution at 5 p.m.

Figure 3-7. Spatial-temporal distribution of AoG and AoR

3.5 SPLIT FAILURE

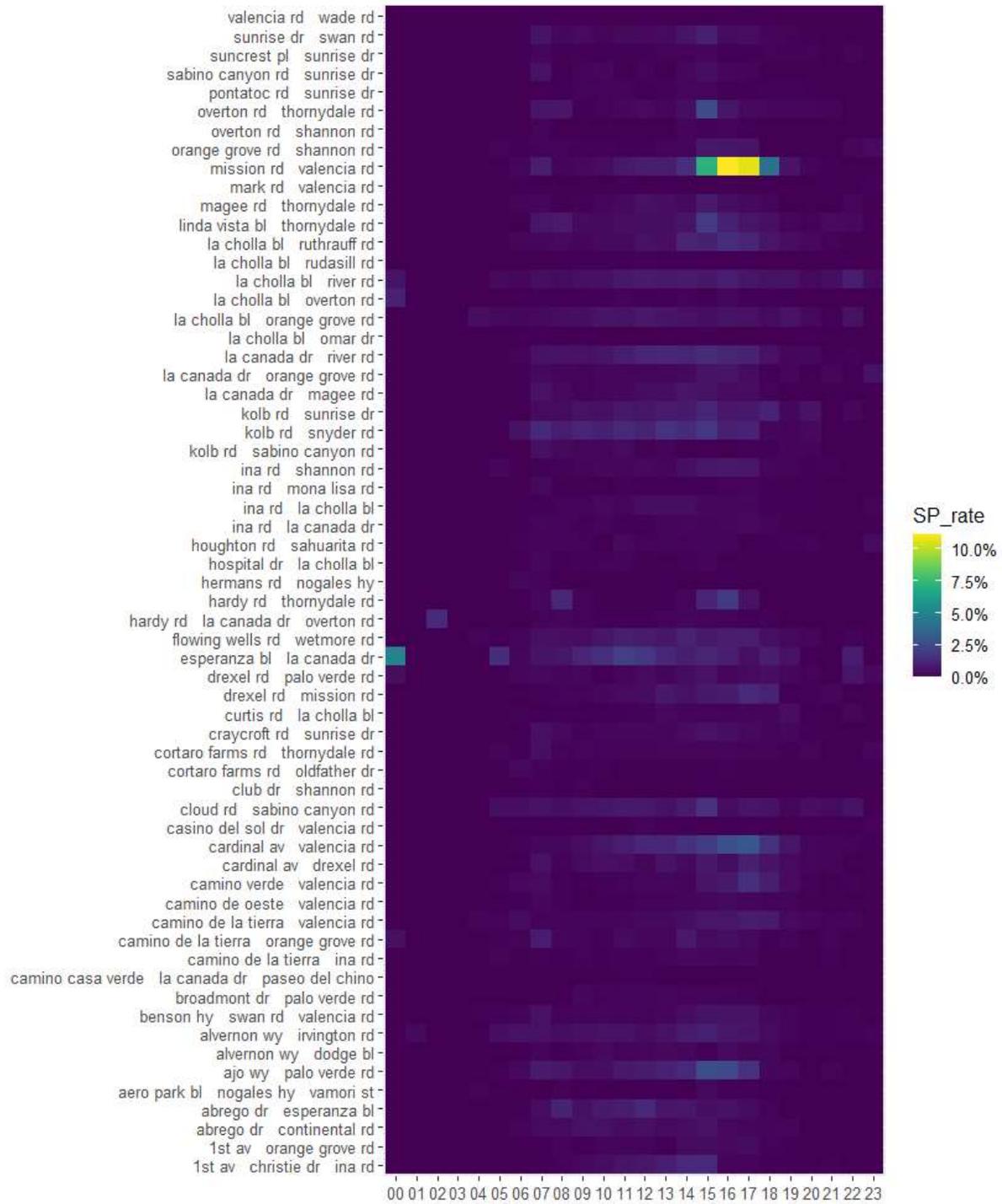
3.5.1 Calculation

Split failure is a performance measure to indicate when the traffic demand cannot be served within one cycle, and it is calculated using the green occupancy ratio (GOR) and red occupancy ratio (ROR) collected by presence detectors (Day et al., 2014, 2008).

We use the method proposed by (E. Saldivar-Carranza et al., 2021a). We first identify all sample vehicles that stop more than once before passing through an intersection during a defined time interval. These are the vehicles that are not served within one cycle. Then this number of sample vehicles that stop more than once is calculated as a percentage of the total number of sample vehicles that pass through that intersection during the time interval, and that percentage is the split failure ratio.

3.5.2 Results and Analysis

Figure 3-8 shows the average intersection-level signal split failure ratio during January, February, and March of 2021 for each intersection and hour of the day. Most intersections have a very low split failure ratio (under 2.5%), indicating that current signal timing plans at most intersections can serve most of the demand within one cycle with only a few vehicles that must wait for more than one cycle. However, the intersection at Mission Rd. and Valencia Rd. has a much higher split failure ratio (10%) from 4 – 5 p.m. This high ratio could be due to the inefficient signal plan, heavy traffic demand, or both.



3.6 RED-LIGHT RUNNING

3.6.1 Calculation

Using only probe vehicle data to identify red-light running is very challenging without signal events. We also use event-based signal data collected by the Miovision system to determine the signal status when a vehicle passes through an intersection to identify whether the vehicle ran a red light.

3.6.2 Results and Analysis

Figure 3-9 shows the average red-light running frequency of the through vehicles at all study locations. The frequency ranges from 0 to 6 vehicles per hour. During the nighttime, all intersections have very few, if any at all, red-light runners. During the daytime, an intersection typically has 1 to 3 red-light runners per hour. Friday has more red-light runners (5 to 6 per hour) than other days, especially during the p.m. peak hours.

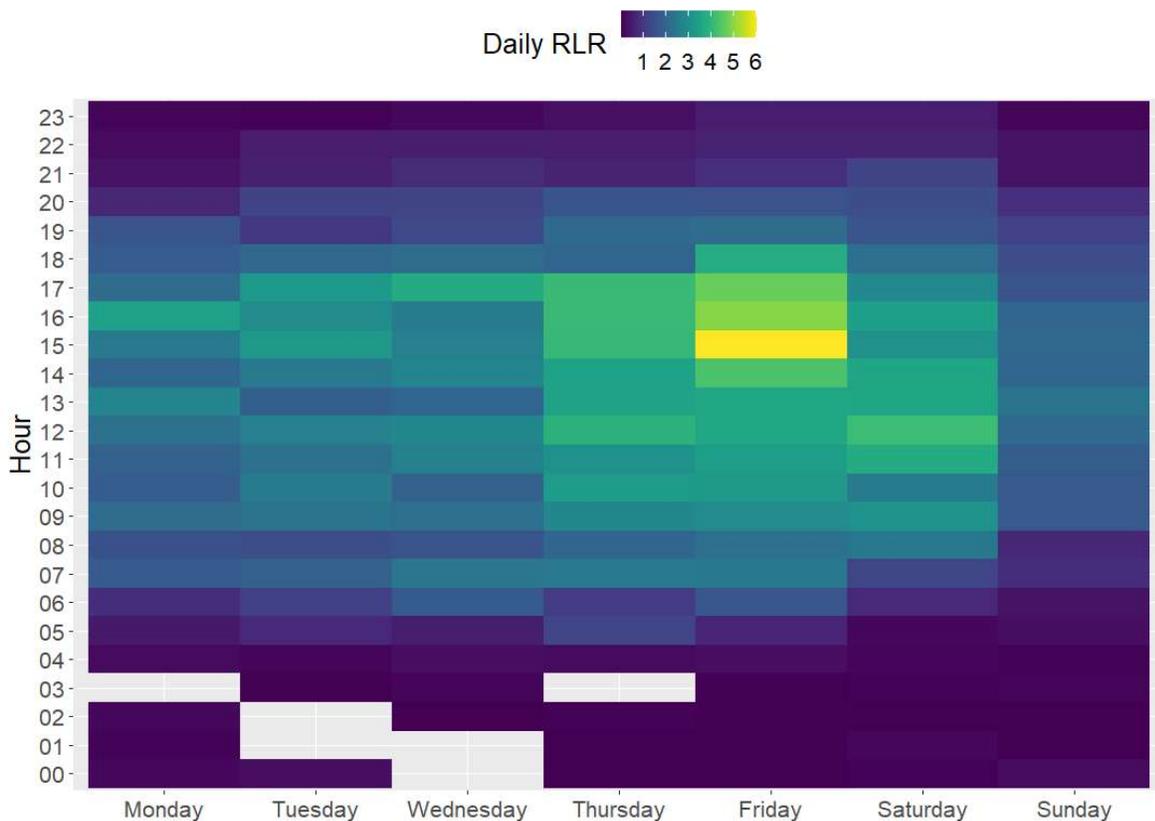


Figure 3-9. Average red-light running frequency of through vehicles

To further investigate the red-light running behaviors, probe vehicle trajectories at Ajo Way & Palo Verde Rd. and the associated red signal timing are visualized in **Figure 3-9**. The red bars in **Figure 3-10** represent the time for which the signal was red, and the black lines are vehicle trajectories. The results show that most red-light running occurred at the beginning of the red signal phase, which is likely during the clearance interval. Two trajectories, 679.11 and 501.18, apparently ran the red light after the red light starts for a moment, reflecting the aggressive driving. These two trajectories have a long stop at the intersection, so a long stop delay at the intersection could cause aggressive driving behaviors.

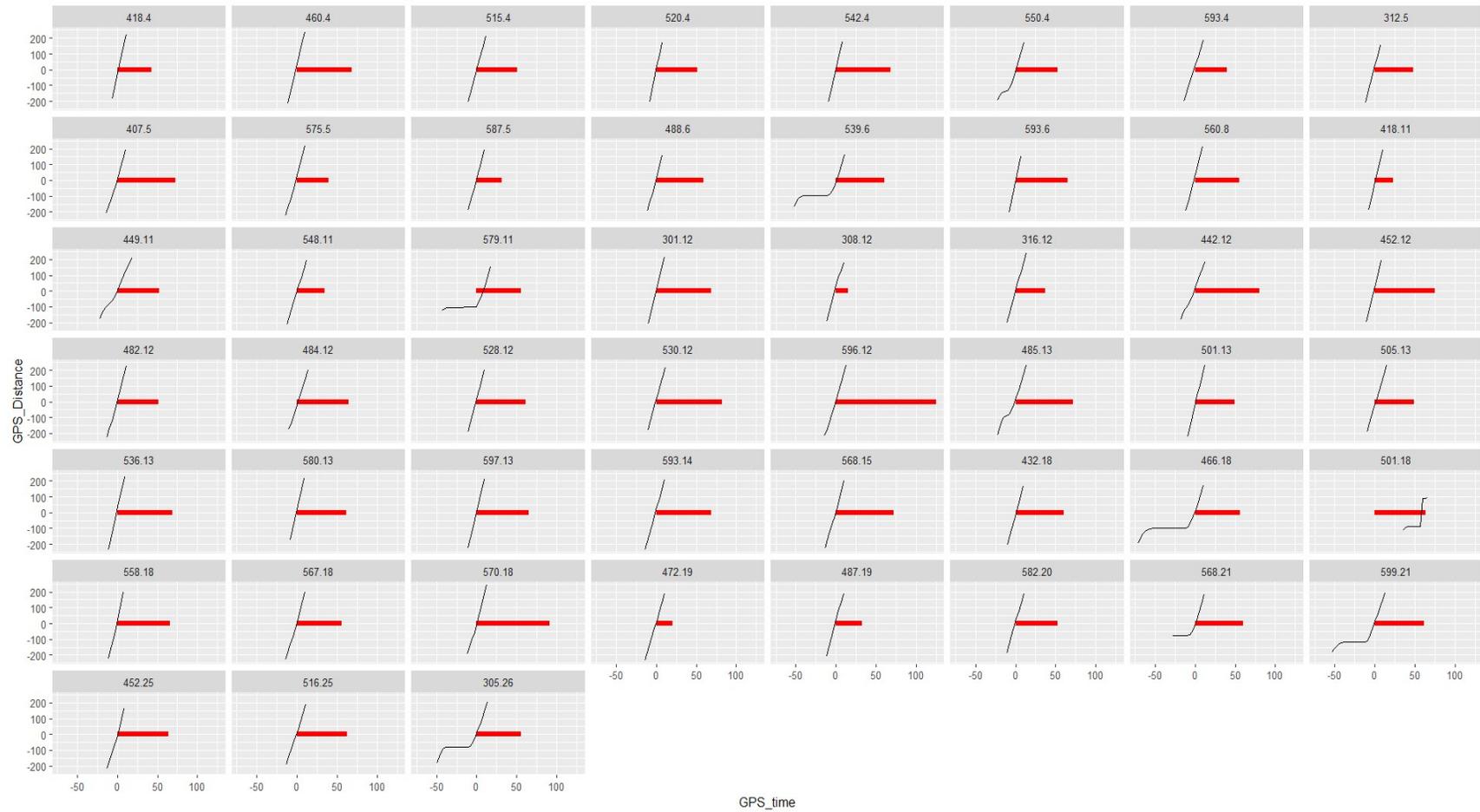


Figure 3-10. Red-light running trajectories at Ajo Way & Palo Verde Rd.

CHAPTER 4: MOBILITY/RELIABILITY PERFORMANCE DATA COLLECTION AND ANALYSIS

According to the previous chapter’s results, the third-party probe vehicle data provided by Wejo can be used to calculate multiple traffic performance measures including intersection delay, level of service (LOS), arrival on green (AoG), and split failure. In addition to mobility and reliability performance measures extracted from Wejo data, up-to-date traffic sensors installed at signalized intersections in the PAG region can also provide traffic performance measures. This chapter compares these sensor-based and Wejo-based performance measures to understand the capabilities, coverage, and differences.

4.1 SENSOR-BASED PERFORMANCE DATA COLLECTION AND ANALYSIS

In the PAG region, traffic sensors from various manufacturers, including Miovision, Autoscope, and Iteris, have been installed at many signalized intersections for traffic detection and signal control. However, most Autoscope and Iteris sensors cannot provide reliable and accurate mobility performance data due to issues such as detector layout configuration, cost, communication, and capability. As of 2022, Miovision sensors are the primary existing sensors that can collect 24/7 Automated Traffic Signal Performance Measures (ATSPM) data in the PAG region due to Miovision’s easy to use API, advanced AI detection technology, and cloud server. These Miovision sensors are managed by two agencies, the Town of Marana and Pima County Department of Transportation (PCDOT), and their locations are shown in **Figure 4-1**. Miovision provides a web interface called the TrafficLink portal (<https://trafficlink.miovision.com/>) for users to check and download multiple signal performance measures and network metrics, and

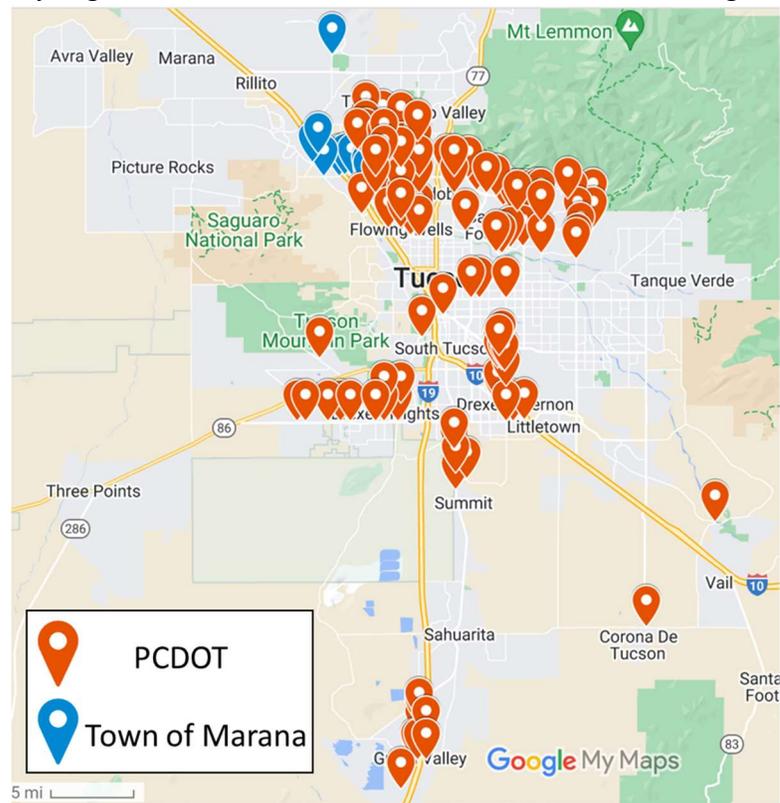


Figure 4-1. Miovision Sensor Locations

all the available measures are shown in **Figure 4-2**. All signal performance measures provided by the TrafficLink portal are defined and calculated according to event-based data. Therefore, to save time and effort in data collection, we first automatically collect event-based data through the API and then calculate arrival-on-green and split failure from this event data but download simple delay from the TrafficLink portal.

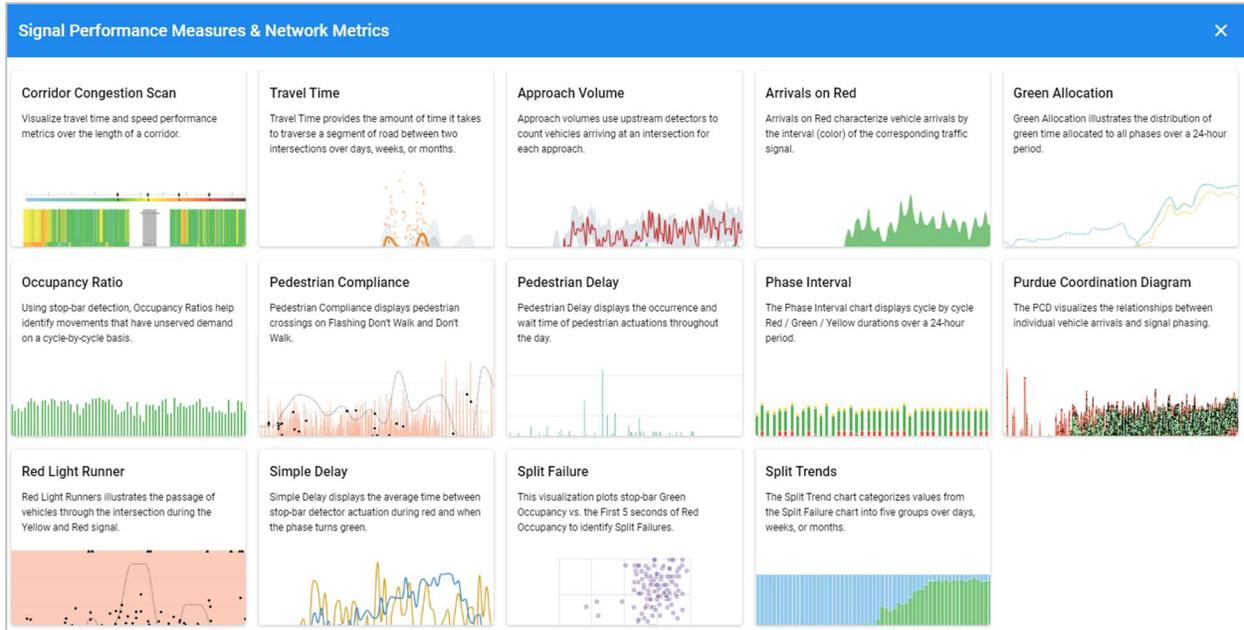


Figure 4-2. Miovision signal performance measures and network metrics

4.1.1 Simple Delay

The delay provided by Miovision is the simple stop delay rather than the control delay. The simple stop delay is defined as the time difference between stop-bar detector actuation during the red phase and the start of the next green phase (Miovision, 2022a). While the control delay includes the acceleration delay, deceleration delay, and the delay of vehicles in the queue outside the detection zone, the simple delay misses all these other components of the control delay.

The simple stop delay metric downloaded from the TrafficLink portal is aggregated into 15-minute intervals. In this section, the simple stop delay is analyzed by aggregation into 1-hour intervals. The intersection of La Cholla Blvd. and River Rd. is chosen as an example location, and three days of simple stop delay data in January 2021 is used to analyze the simple delay trends, as shown in **Figure 4-3**. **Figure 4-3** demonstrates that the left-turn delay and through delay in all four directions have similar temporal trends, with higher delay from 7 a.m. to 6 p.m. and lower delay during the night. The left-turn and through delay ranges from 0 to 130 seconds. In comparison with through delay, the left-turn delay has more fluctuation likely because of the stochastic arrival of left-turn vehicles and the permissive left-turn phase. In addition, the northbound and southbound

directions have greater delay for the left turn movement than the through movement during the daytime because of the high volume of left turns. Simple delay is calculated from the data generated by detectors at intersections, but most locations do not have detectors installed for right-turn lanes. Therefore, most Miovision sensors can only provide simple delay for the through and left-turn movements.



Figure 4-3. Temporal trend of simple delay at La Cholla Blvd. & River Rd, Tucson

In addition to hourly delay, the Miovision-based delay can be used to calculate the delay reliability performance, 95th percentile delay, buffer delay, and buffer index. The 95th percentile delay is the delay where 95% of data are under it at a specific location and hour, and other two reliability indicators are calculated using the following two equations.

$$\text{Buffer Delay} = 95^{\text{th}} \text{ Delay} - \text{Average delay} \quad (4-1)$$

$$\text{Buffer Index} = \frac{95^{\text{th}} \text{ Delay} - \text{Average delay}}{\text{Average delay}} \quad (4-2)$$

The three months of data collected from La Cholla Blvd. & River Rd. is used to calculate the four delay reliability indicators, as shown in **Figure 4-4 – Figure 4-6**. The 95th percentile delay show a smooth trend with lower value during night and higher value during daytime, indicating the delay during nighttime is more reliable than daytime. However, the buffer delay fluctuates more for 24 hours, and the trend depends on the direction and movement. The possible reason is that the average delay at different locations and times varies, leading to this fluctuating trend. In addition, the buffer index shows that the delay during daytime is more reliable than nighttime. When the buffer index is larger than 40%, the traffic delay at an intersection is deemed unreliable.

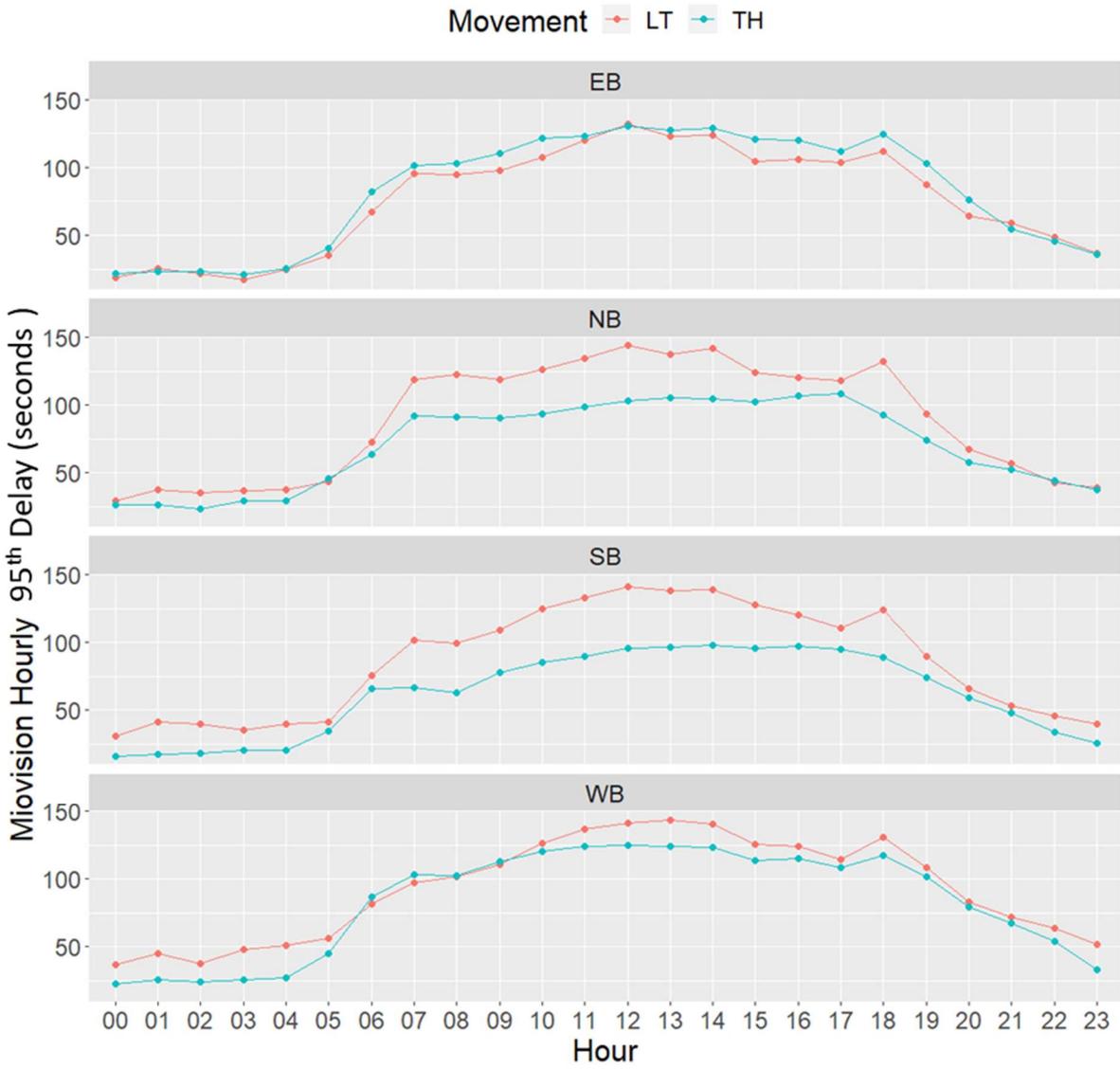


Figure 4-4. 95th percentile simple delay at La Cholla Blvd. & River Rd, Tucson during Jan.-March 2021.

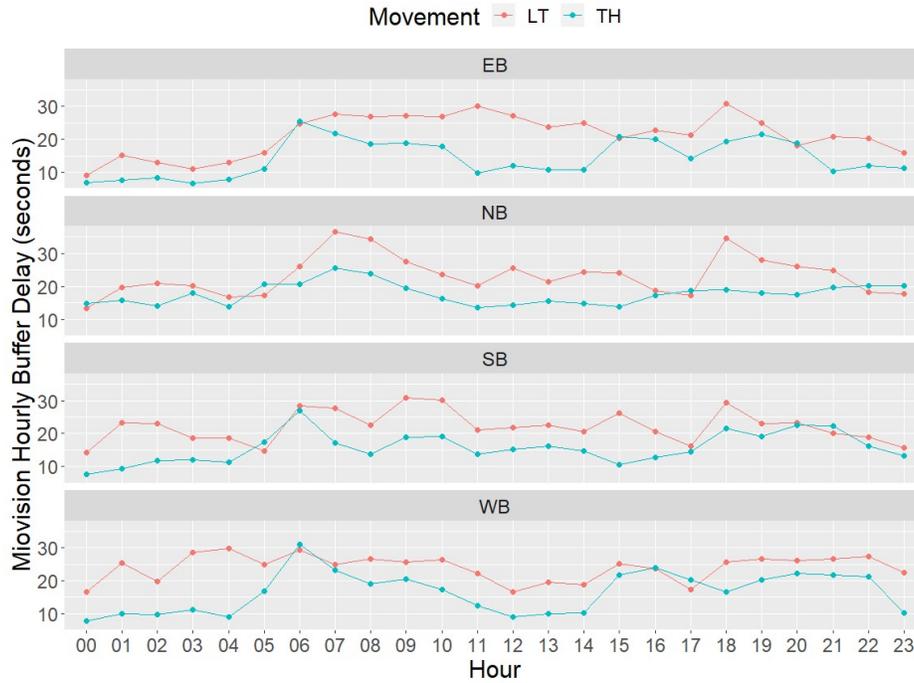


Figure 4-5. Buffer delay at La Cholla Blvd. & River Rd., Tucson during Jan.-March 2021.

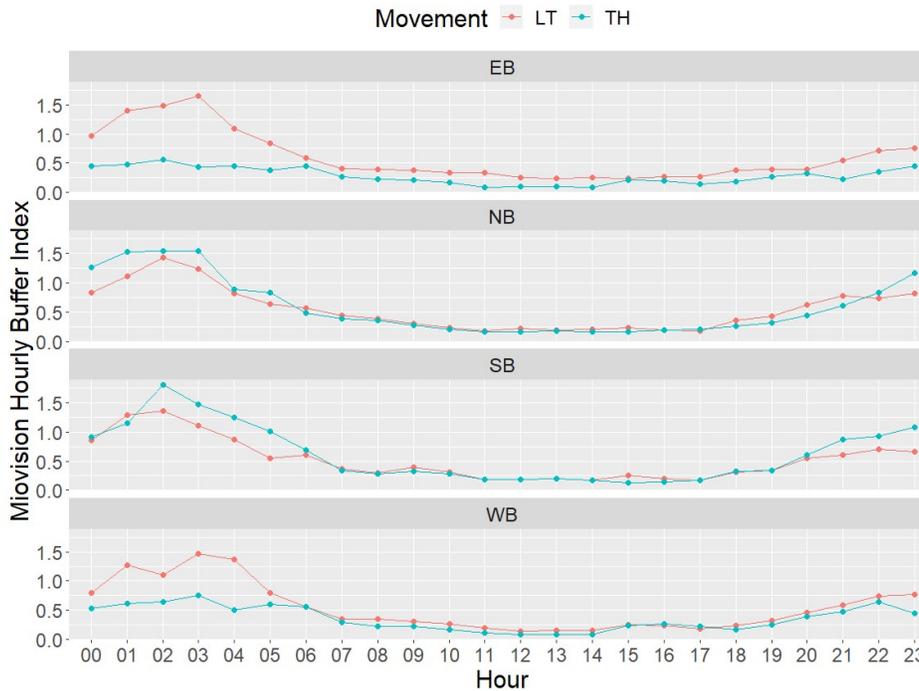


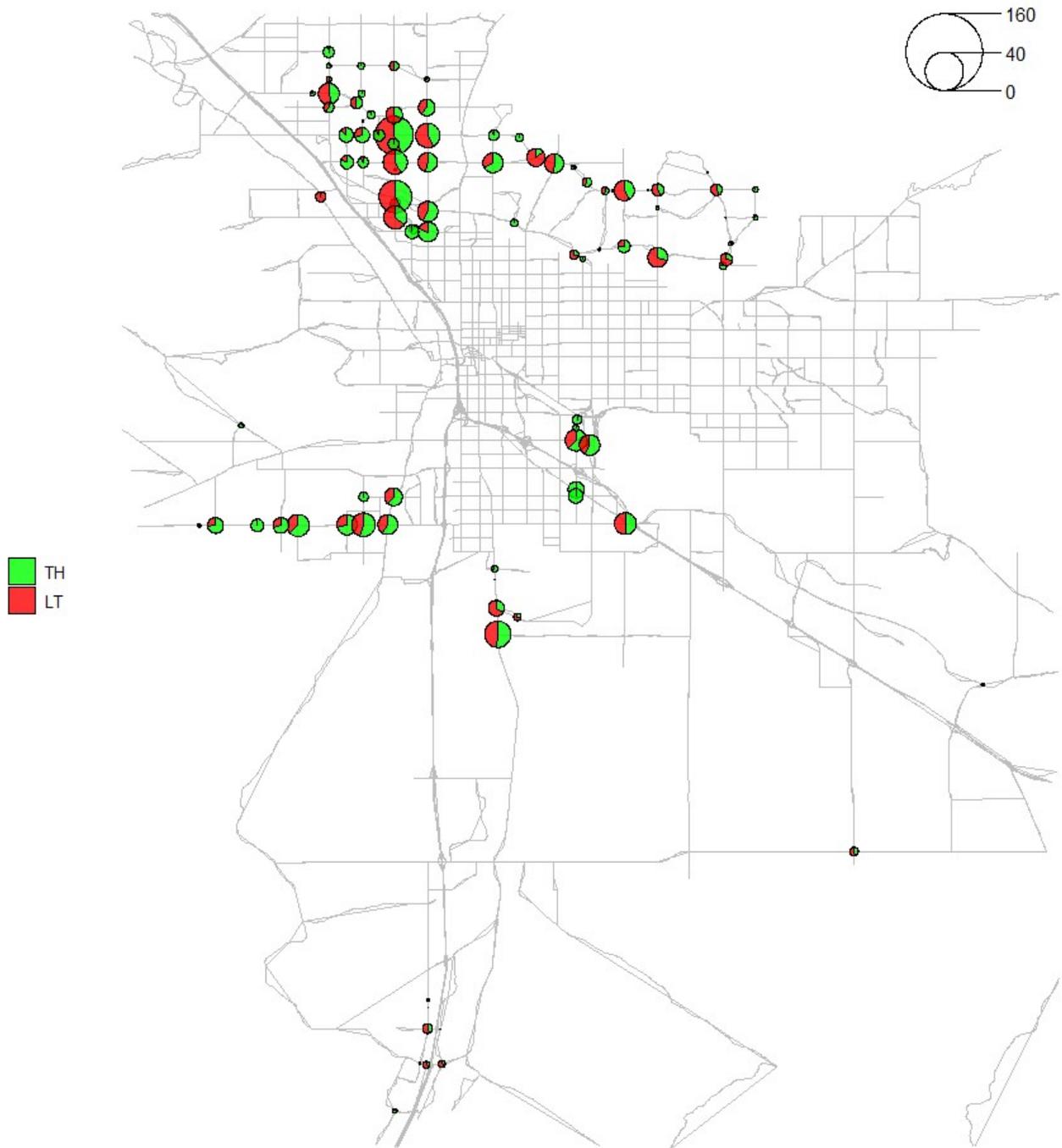
Figure 4-6. Buffer index at La Cholla Blvd. & River Rd., Tucson during Jan.-March 2021.

97 out of 108 signalized intersections with Miovision sensors have available simple delay data during Jan. – March 2021. The 1-hour interval simple delay is aggregated into intersection-level average hourly through/left-turn delay by averaging three months of data in different directions (EB, WB, NB, SB) at an intersection. The pie chart size in **Figure 4-7** is the sum of the average left-turn and through delay, and the areas of red and green slices indicate the proportions of left-turn and through delay, respectively.

Figure 4-7(a) shows the spatial distribution of simple delay at 1 a.m. At midnight, most intersections have a very low delay on left-turn and through movements, specifically lower than 10 seconds; some intersections even have zero delay due to a low volume. In addition, intersections on the La Cholla Blvd. corridor have a relatively higher delay but are still lower than 20 seconds. The proportions of the left-turn delay and the through delay are not consistent and vary with location. **Figures 4-7(b)** and **4-7(c)** show the simple delay spatial distribution at morning and afternoon peaks, respectively. Most intersections at peak hours have a higher left-turn and through delay, ranging from 60 to 80 seconds, which is relatively congested. At these intersections, left-turn delay is higher than through delay as indicated by the larger red slices and smaller green slices in the pie charts. However, some less congested intersections with a low delay have higher through delay than left-turn delay, which is likely because of the low left-turn volume. In addition, the proportions of left-turn delay and through delay at a location are similar at different hours during peak hours.

Additionally, **Figure 4-8** shows the spatial distribution of buffer index during off-peak and peak hours. Most intersections have more reliable delay during peak hours than off-peak hours, which is because the traffic pattern during peak hours is more consistent than during off-peak hours, when the traffic volumes vary more. The through movement has a more reliable delay than left-turn at most intersections.

01



(a) simple delay at 1 a.m.

08

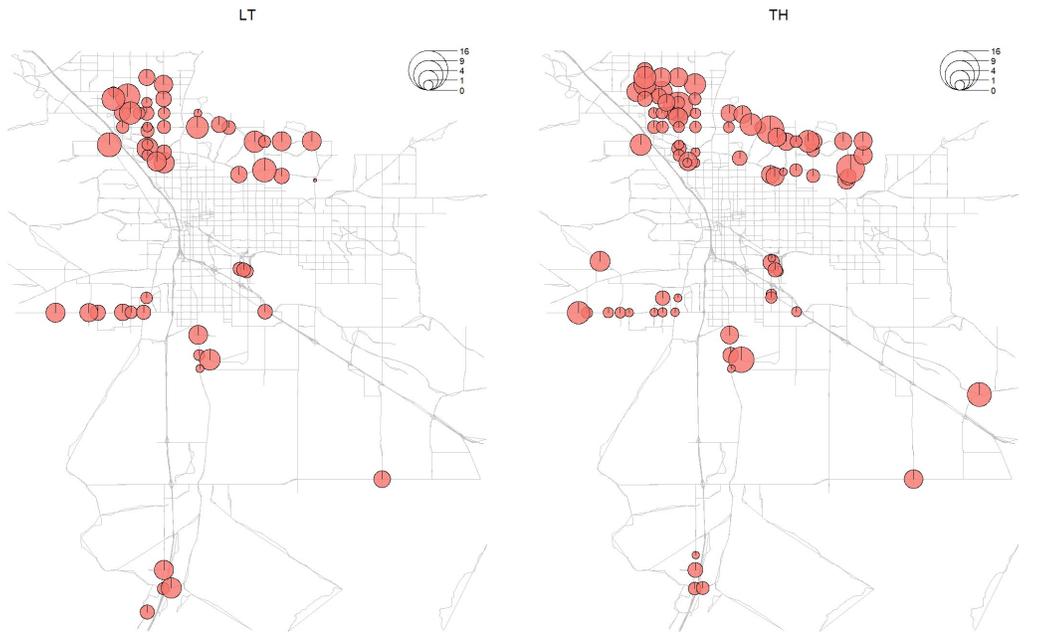


(b) simple delay at 8 a.m.



(c) simple delay at 5 p.m.

Figure 4-7. Spatial distribution of average through and left-turn simple delay at different hours



(a) 1 a.m.



(b) 8 a.m.

Figure 4-8. Spatial distribution of buffer delay index



(c) 5 p.m.

Figure 4-9. Spatial distribution of buffer delay index

Figure 4-9 shows the simple delay distribution of all intersections by the hour of the day. From midnight to 5 a.m., the distribution of through and left-turn delay approximates a log-normal distribution, with a mean value of lower than 5 seconds and a standard deviation of lower than 20 seconds, which indicates that most intersections have low delay during the night. Starting at 7 a.m., the delay distribution is more like a mixture of two distributions, and the distribution standard deviation becomes significantly larger. This mixed distribution is likely because some intersections far from the urban area are less congested all day, and so have a distribution with mean close to zero. After 7 p.m., both the mean and standard deviation of the distribution become smaller due to the reduction of traffic volume.

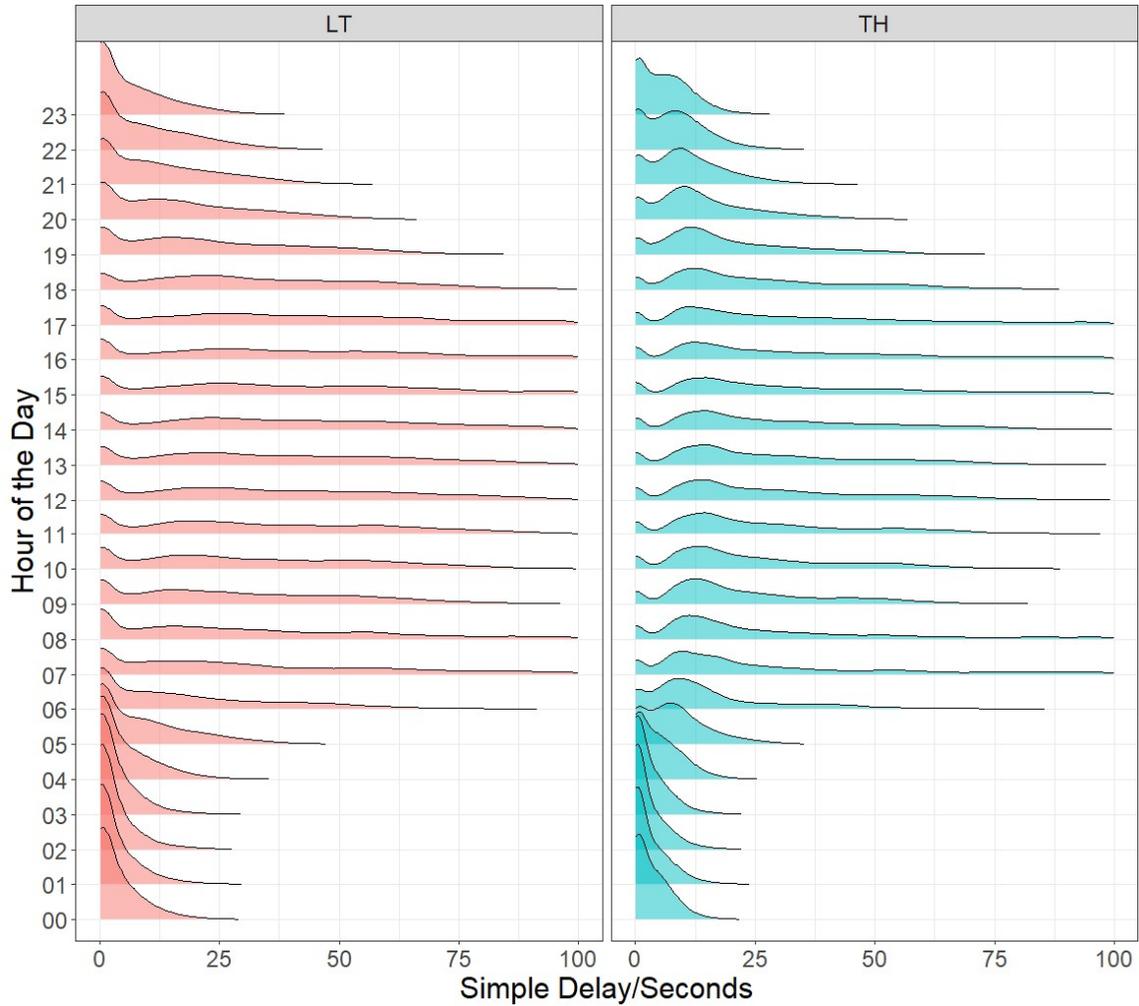


Figure 4-10. Simple delay distribution by hour of the day

4.1.2 Level-of-Service

The intersection level-of-service (LOS) is identified using simple delay according to the LOS definition in the Highway Capacity Manual (Transportation Research Board, 2016). **Figure 4-10** shows the intersection LOS distribution by hour on March 17, 2021. At midnight (12am-4am), around 80 of all 94 study intersections have the LOS of A. Starting at 6 a.m., the LOS at more intersections drops from A to B, C, and D, indicating these locations become more congested.

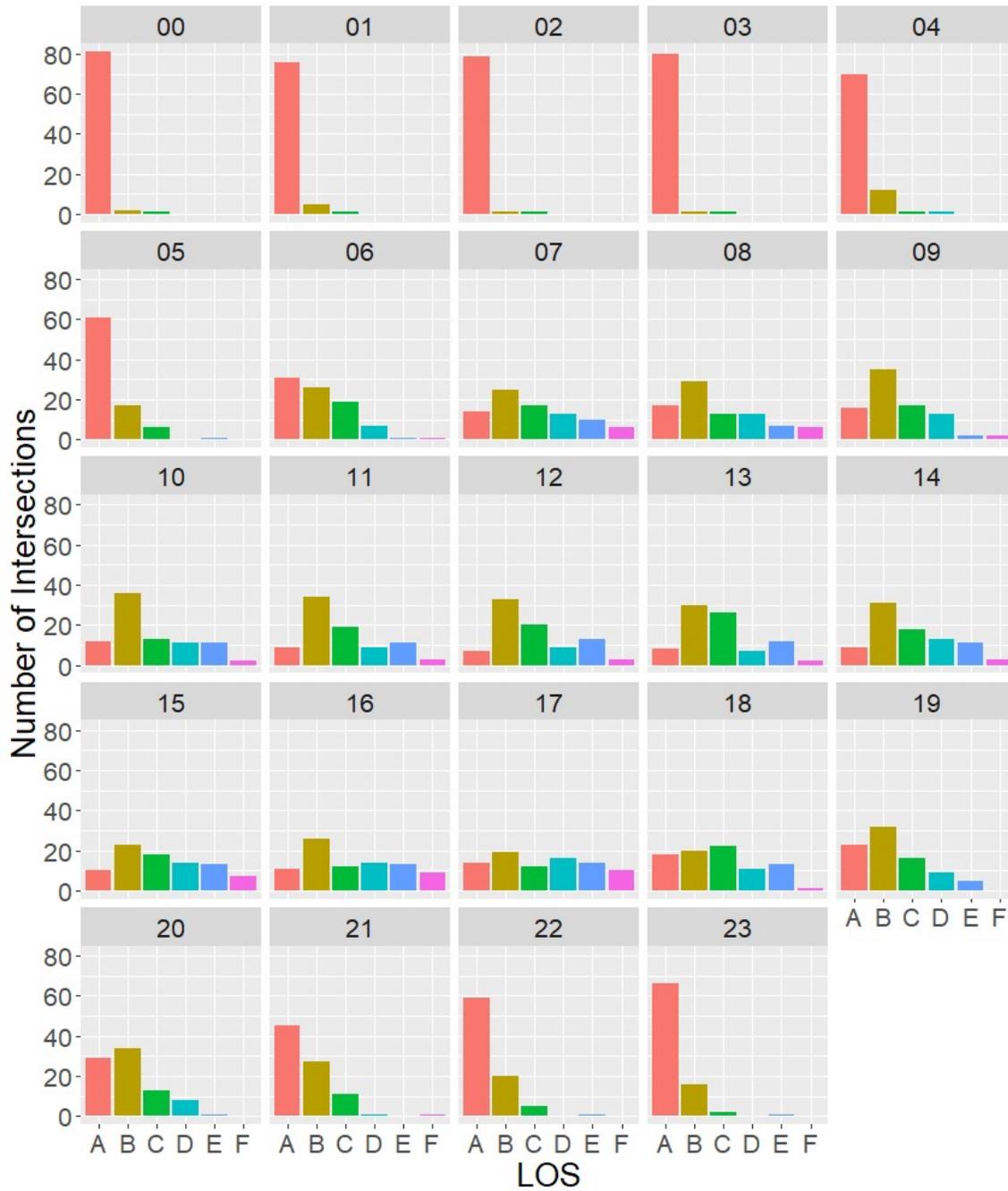


Figure 4-11. Intersection LOS distribution by hour on March 17, 2021

During the daytime (7 a.m. – 7 p.m.), only a limited number of intersections have LOS of F. During the morning peak, only 6 intersections have LOS of F, as shown in **Table 4-1**. Three of these intersections having LOS of F are located on the La Cholla Blvd. corridor, and one is located on the La Cañada Dr. corridor, and these two corridors are the major congested corridors in the area. In addition, the most congested intersection at morning peak is Pontatoc Rd. & Sunrise Dr.,

which is not located on the above-mentioned corridors. A possible reason is that there is a high school near the intersection, attracting more traffic during peak hours. Same as the morning peak, Pontatoc Rd. & Sunrise Dr. is also the most congested intersection with LOS of F in the afternoon peak, again possibly due to the nearby high school. As shown in **Table 4-2**, all top ten congested intersections have LOS of F, and most of these locations are located on one of the two corridors mentioned above. At midnight, only one intersection (Ina Rd. & La Cholla Blvd.) has a LOS worse than B, and the rest of the intersections have a LOS of A or B as shown in **Table 4-3**.

Table 4-1. Intersection Ranking based on the LOS at 8 a.m.

Rank	Hour	Intersection	Intersection Delay/s	LOS	Date
1	08	Pontatoc Rd. / Sunrise Dr.	97	F	2021-03-17
2	08	La Cholla Blvd. / Orange Grove Rd.	93	F	2021-03-17
3	08	Ina Rd. / La Cholla Blvd.	88	F	2021-03-17
4	08	Ina Rd. / La Cañada Dr.	86	F	2021-03-17
5	08	Cortaro Farms Rd. / Thornydale Rd.	86	F	2021-03-17
6	08	La Cholla Blvd. / River Rd.	83	F	2021-03-17
7	08	Sunrise Dr. / Swan Rd.	74	E	2021-03-17
8	08	La Canada Dr. / Magee Rd.	69	E	2021-03-17
9	08	Craycroft Rd. / Sunrise Dr.	69	E	2021-03-17
10	08	La Cholla Blvd. / Magee Rd.	67	E	2021-03-17

Table 4-2. Intersection Ranking based on the LOS at 5 p.m.

Rank	Hour	Intersection	Intersection Delay/s	LOS	Date
1	17	Pontatoc Rd. / Sunrise Dr.	105	F	2021-03-17
2	17	Ina Rd. / La Cholla Blvd.	96	F	2021-03-17
3	17	La Cholla Blvd. / River Rd.	95	F	2021-03-17
4	17	La Cholla Blvd. / Orange Grove Rd.	95	F	2021-03-17
5	17	Cortaro Farms Rd / Thornydale Rd.	94	F	2021-03-17
6	17	La Canada Dr. / Magee Rd.	87	F	2021-03-17
7	17	Flowing Wells Rd. / Wetmore Rd.	84	F	2021-03-17
8	17	La Cholla Blvd. / Ruthrauff Rd.	84	F	2021-03-17
9	17	La Canada Dr. / Orange Grove Rd.	82	F	2021-03-17
10	17	Ina Rd. / La Cañada Dr.	82	F	2021-03-17

Table 4-3. Intersection Ranking based on the LOS at 1 a.m.

Rank	Hour	Intersection	Intersection Delay/s	LOS	Date
1	01	Ina Rd. / La Cholla Blvd.	34	C	2021-03-17
2	01	Nogales Hwy / Old Nogales Hwy	14	B	2021-03-17
3	01	Camino De Oeste / Valencia Rd.	10	B	2021-03-17
4	01	La Cholla Blvd. / River Rd.	10	B	2021-03-17
5	01	La Cholla Blvd. / Ruthrauff Rd	10	B	2021-03-17
6	01	Cardinal Ave. / Valencia Rd.	10	B	2021-03-17
7	01	Flowing Wells Rd. / Wetmore Rd.	9	A	2021-03-17
8	01	Ajo Way / Palo Verde Rd.	8	A	2021-03-17
9	01	Benson Hwy / Swan Rd. / Valencia Rd.	8	A	2021-03-17
10	01	Mission Rd. / Valencia Rd.	7	A	2021-03-17

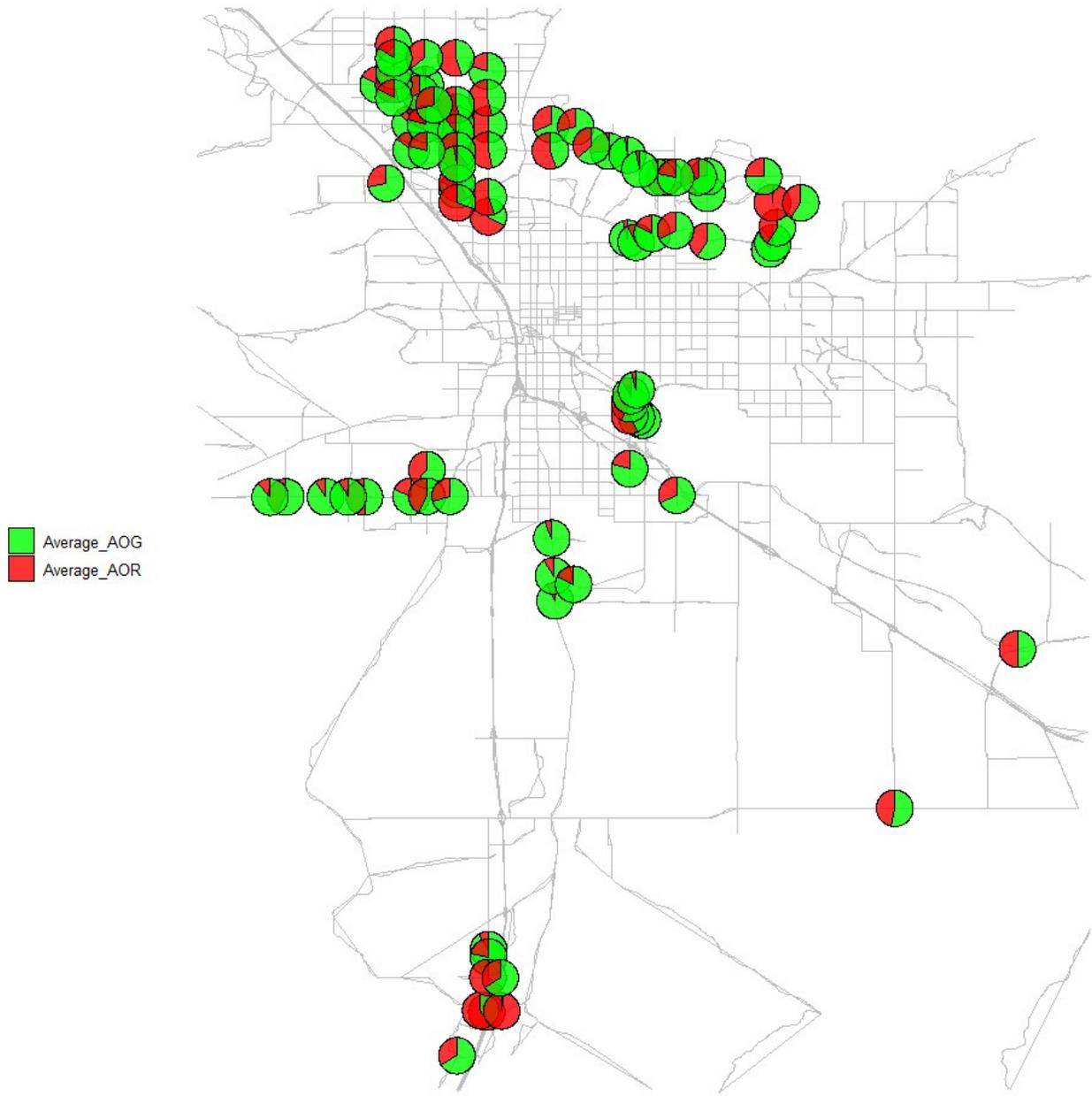
4.1.3 Arrival-on-Green

The arrival-on-green (AoG) is calculated based on the signal status when the advance detector is triggered by arriving vehicles. Because only the through movement is configured with advance detectors at study locations, we only calculated the AoG for the through movement. **Figure 4-11** shows the temporal trend of three days of AoG in January 2021 at La Cholla Blvd. & River Rd. The AoG at La Cholla Blvd. & River Rd. during daytime is consistent, ranging from 20% to 40%, except for WB. WB has a consistent peak during the afternoon from 3 p.m.-5 p.m. with the AoG around 60%, which is probably because the primary objective of the signal coordination design in the afternoon is to improve WB traffic. However, the nighttime does not have a clear trend of AoG as shown in **Figure 4-11**. Sometimes, EB and WB have a peak during midnight and NB and SB have a very low AoG such as on Jan. 6 and 8, 2021, but NB and SB have a peak and EB and WB have a very low AoG on Jan 7. The AoG trend during nighttime is inconsistent probably because of fewer vehicle arrivals and actuation signal timing control.



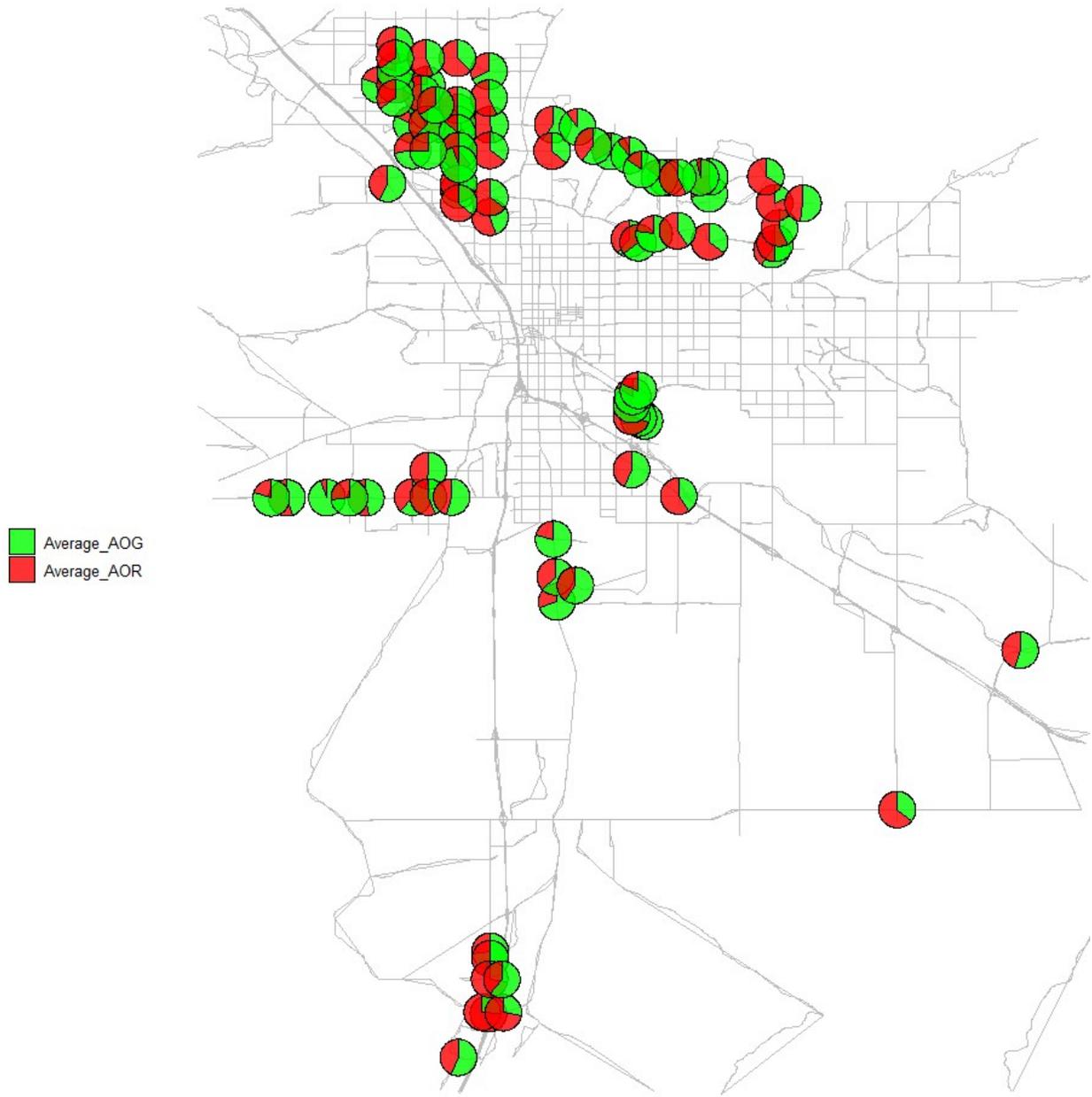
Figure 4-12. Temporal trend of AoG at La Cholla Blvd. & River Rd., Tucson

Figure 4-12 shows the spatial distribution of AoG and AoR at midnight, morning peak, and afternoon peak, respectively. The sum of AoG and AoR is 100% of all arriving vehicles, and the proportion of red and green pie slice indicates the average intersection-level AoR and AoG, respectively. As shown in **Figure 4-12(a)**, most intersections have an AoG larger than 50% at 1 am, but a few intersections still have an AoG smaller than 50%. This low AoG at 1am easily occurs at some large intersections where all directions have a similar volume but only one direction is coordinated, causing the other directions to have more vehicles arrive during the red time. During the morning peak and afternoon peak, the AoG is typically lower than 50% at most intersections as shown in **Figures 4-12(b)** and **4-12(c)**. Another finding is that AoG at an intersection during the morning peak is similar to the afternoon peak.



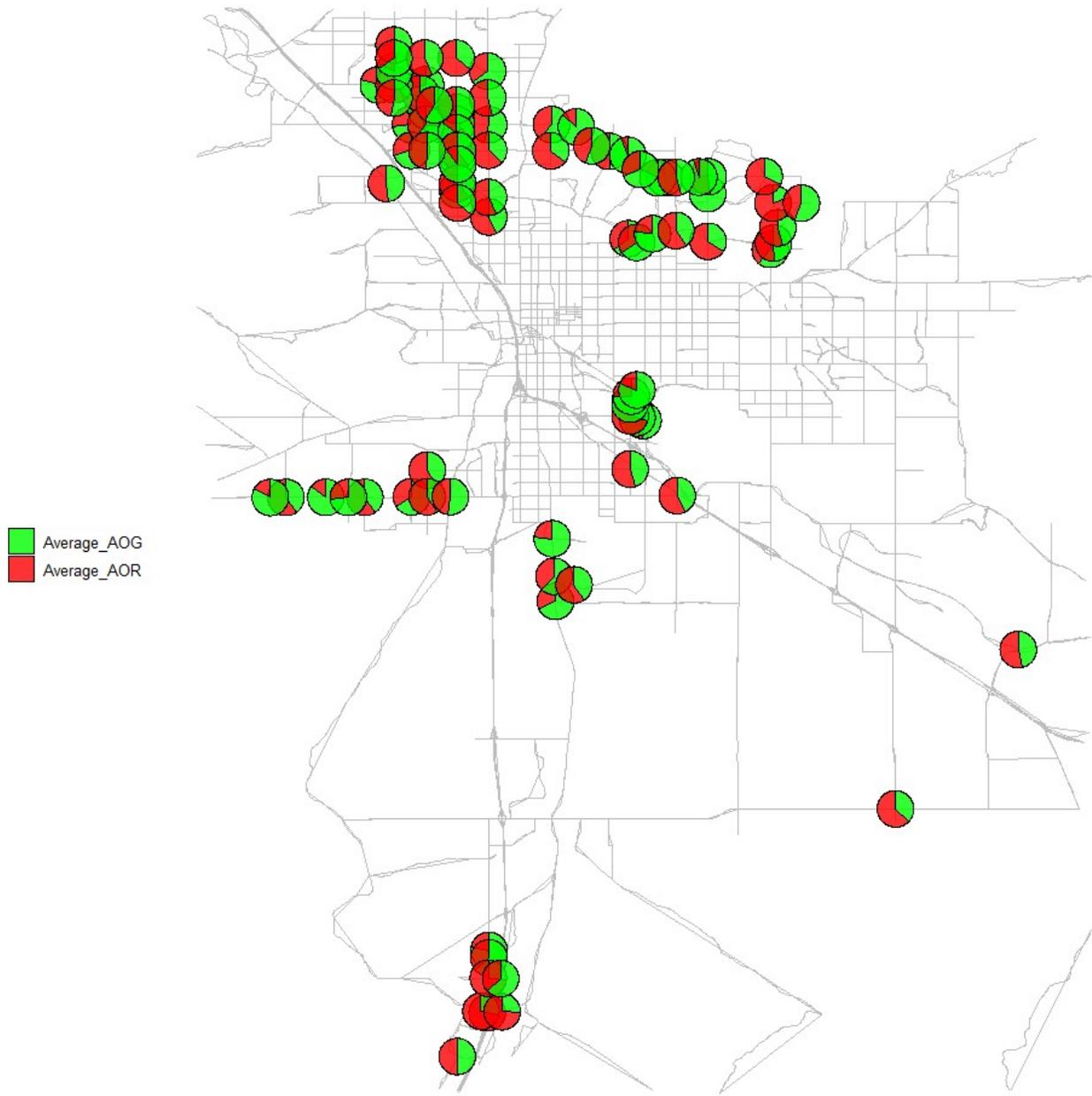
(a) AoG at 1 a.m.

Figure 4-13. Spatial distribution of AoG at different hours



(b) AoG at 8 a.m.

Figure 4-14. Spatial distribution of AoG at different hours



(c) AoG at 5 p.m.

Figure 4-15. Spatial distribution of AoG at different hours

The Miovision-based AoR can be used to calculate the AoR reliability performance, 95th percentile AoR, buffer AoR, and buffer AoR index. The 95th percentile AoR is the AoR where 95% of data

are under it at a specific location and hour, and other two reliability indicators are calculated using the following equation.

$$\text{Buffer AOR} = 95^{\text{th}} \text{ AOR} - \text{Average AOR} \quad (4-3)$$

Figures 4-13 shows the temporal trend of four AoR reliability indicators at La Cholla Blvd. & River Rd. The reason of using AoR rather than AoG is to ensure that a high reliability indicator value represents less reliability, and a low value represents more reliability, which is consistent with delay reliability indicators. These four figures show that the AoR reliability is similar for all 4 directions during most hours. However, the buffer AoR shows the minor directions (NB and SB) have a less reliable AoR during nighttime.

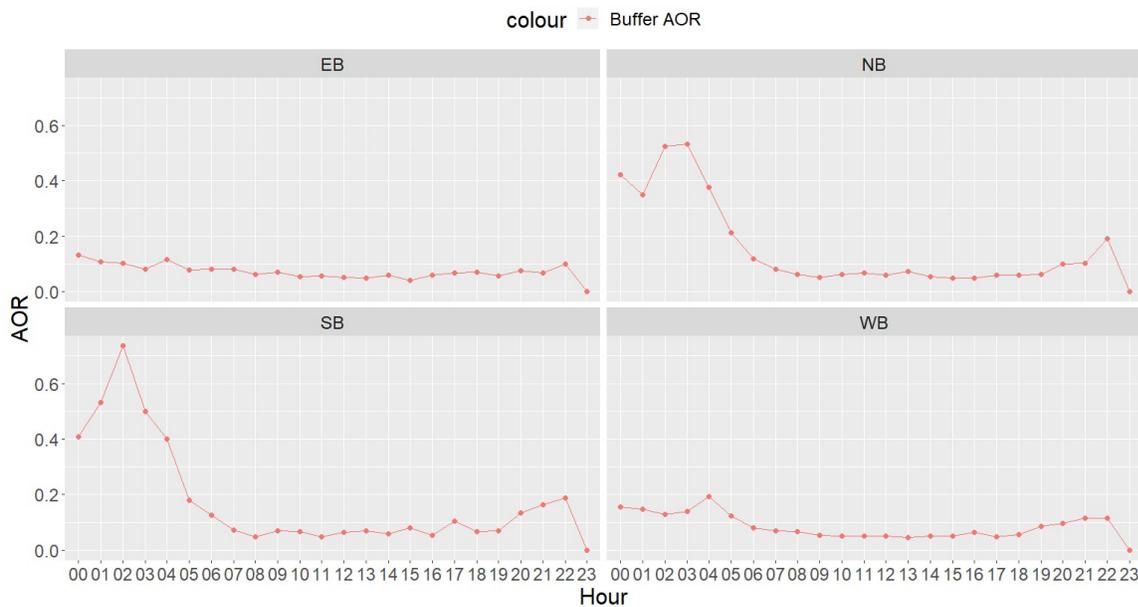


Figure 4-16. Buffer AoR at La Cholla Blvd. & River Rd., Tucson during Jan.-March 2021.

4.1.4 Split Failure

Split failure occurs when a signal phase cannot serve the demand within one cycle and is identified based on the green occupancy ratio (GOR) and the red occupancy ratio during the first five seconds of red (ROR5). The occupancy data is calculated using the event-based data collected by stop-bar detectors. When GOR and ROR5 are both higher than 80% during a phase, this indicates that a split failure occurred. We calculated the intersection-level average split failure percentage as the total number of split failures divided by the total number of cycles. **Figure 4-14** shows the temporal trend of the split failure percentage of left-turn and through movements at La Cholla Blvd. & River

Rd. The signal timing for the through movement has a low split failure likelihood, lower than 0.25%. The signal timing for the left-turn movement has a relatively higher split failure likelihood but still lower than 2%.

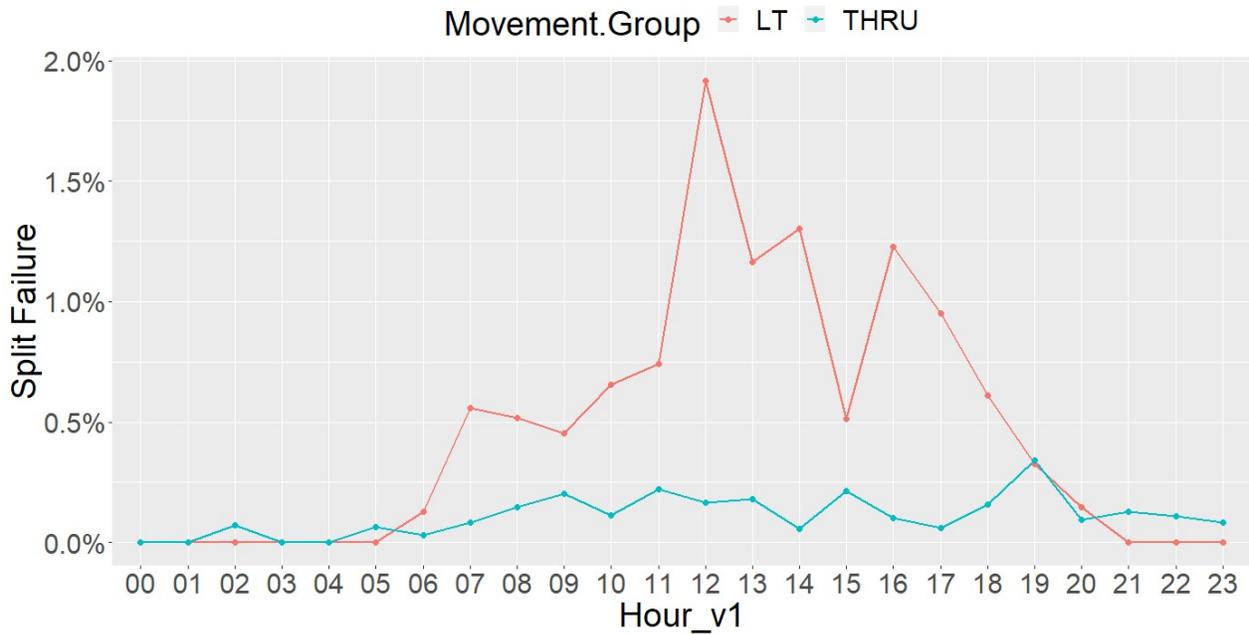
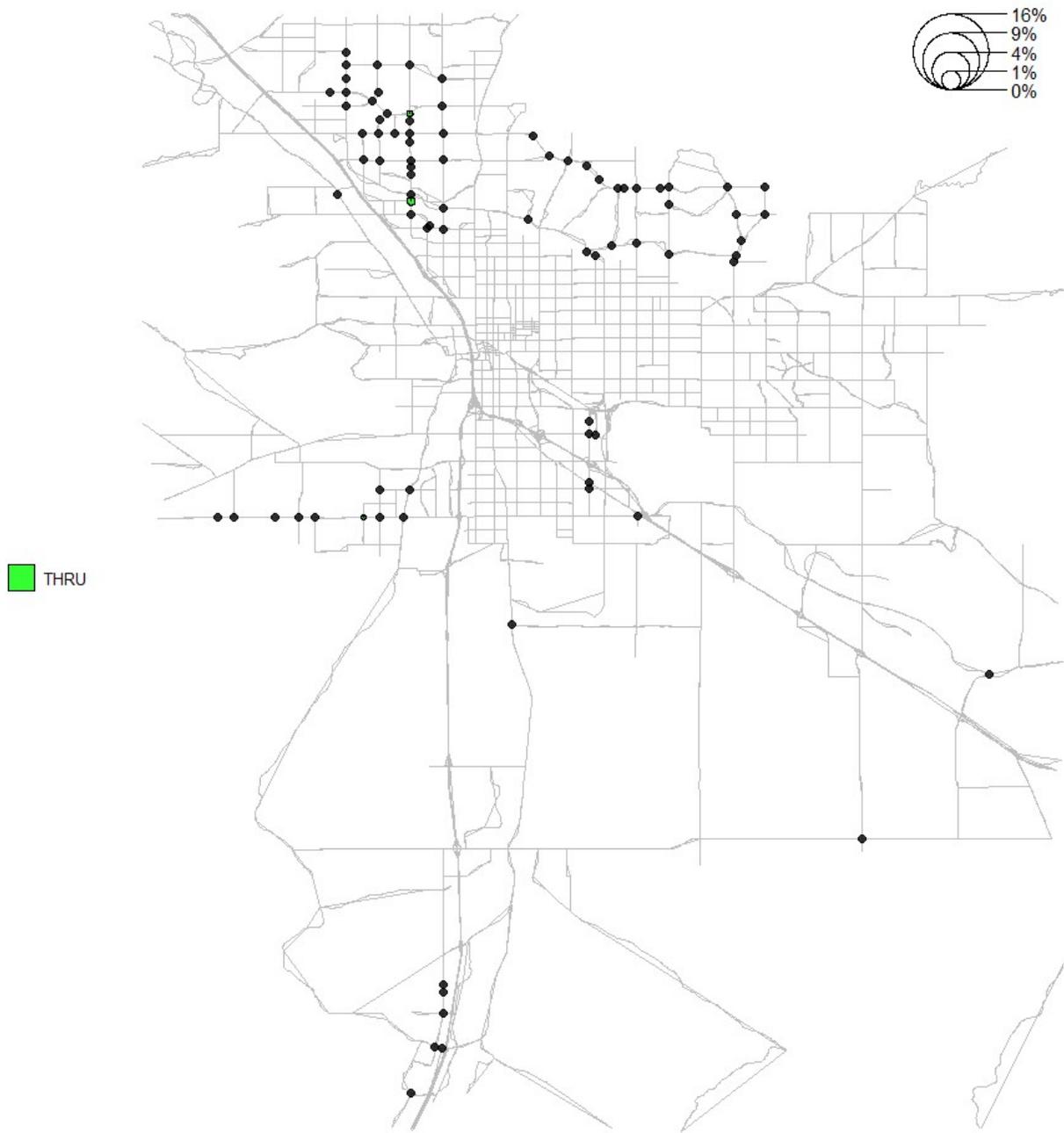


Figure 4-17. Temporal trend of average split failure percentage at La Cholla Blvd. / River Rd

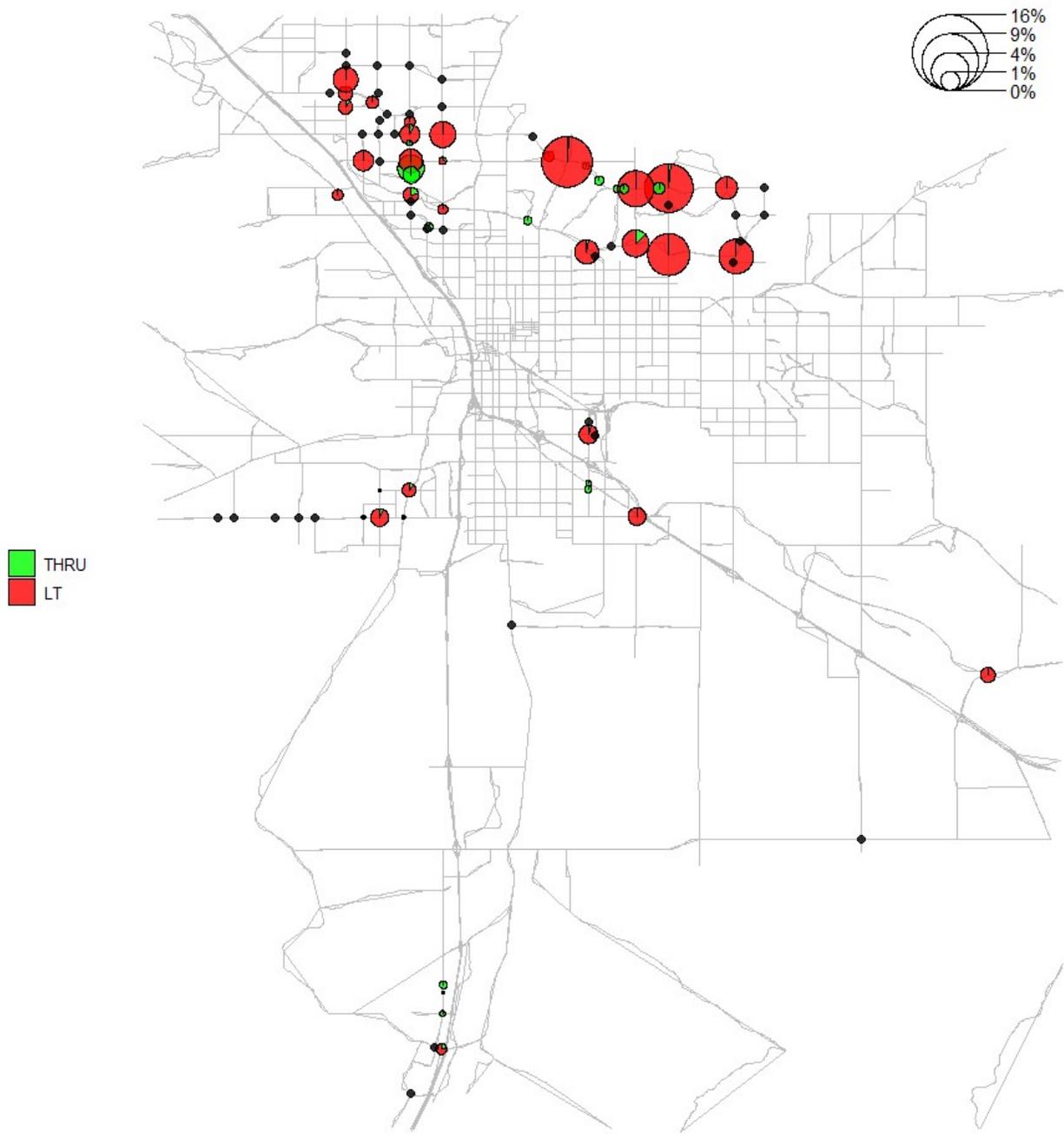
The signal timing has a very low likelihood of split failure during the nighttime because the traffic volume is very low, and all vehicles can be served within one cycle. In **Figure 4-15**, the size of the circle is the sum of the split failure percentages for the left-turn and through movements. The grey circles indicate the intersections do not have split failure. Curtis Rd. & La Cholla Blvd. and La Cholla Blvd. & Magee Rd. had split failure around 0.1% at 1 a.m., but no other intersections had split failure at 1 a.m. At peak hours, most intersections have split failure as shown in **Figures 4-15(b)** and **4-15(c)**, most intersections have split failure due to the higher traffic volume during peak hours. Most intersections have a split failure percentage for the left-turn movement that is significantly higher than the through movement.

01



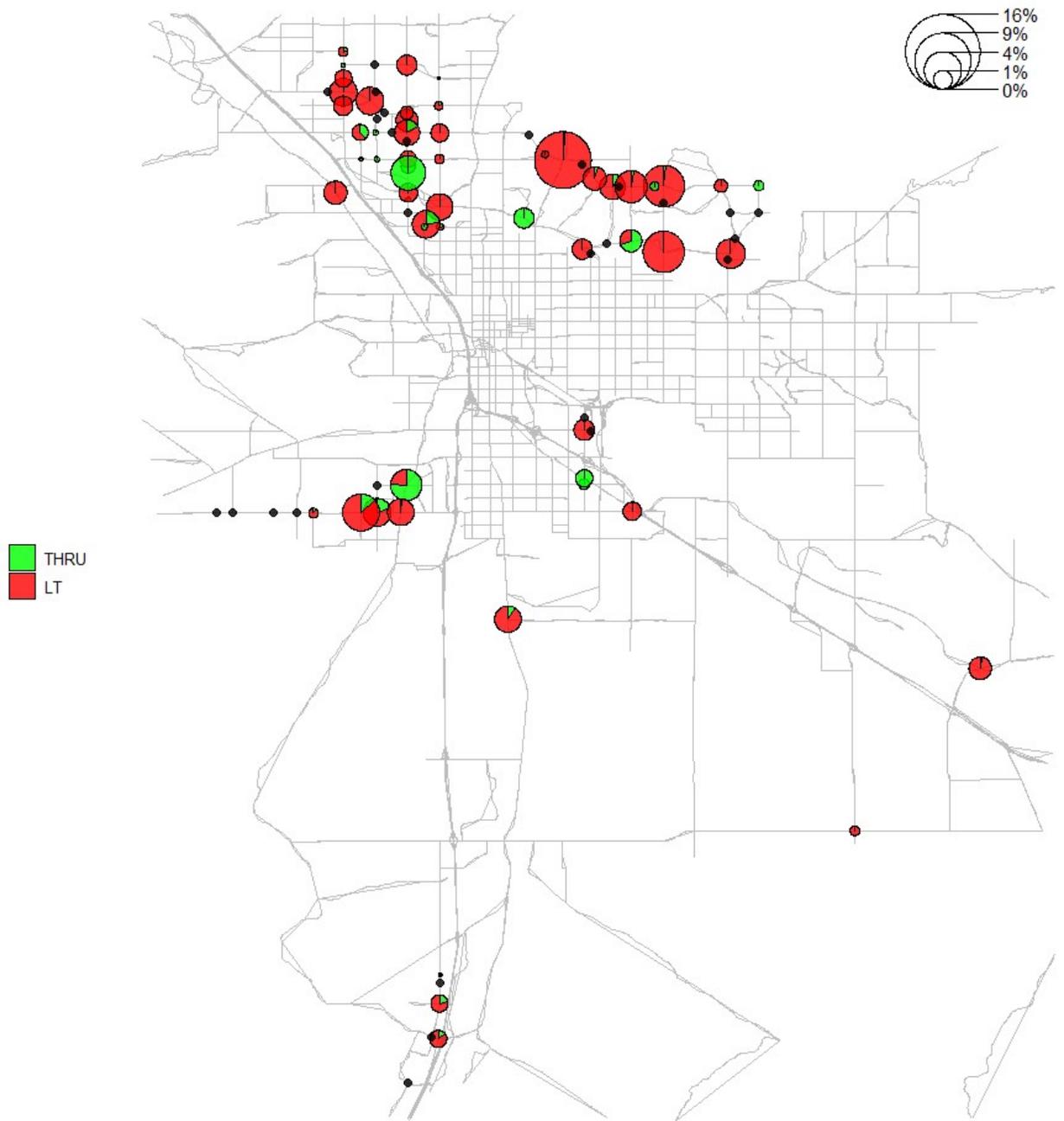
(a) Split failure percentage at 1 a.m.

Figure 4-18. Spatial distribution of split failure percentage at different hours



(b) Split failure percentage at 8 a.m.

Figure 4-19. Spatial distribution of split failure percentage at different hours



(c) Split failure percentage at 5 p.m.

Figure 4-20. Spatial distribution of split failure percentage at different hours

The Miovision-based split failure can be used to calculate the split failure reliability performance, 95th percentile split failure, buffer split failure, and buffer split failure index. The 95th percentile SP is the split failure where 95% of data are under it at a specific location and hour, and other two reliability indicators are calculated using the following equation.

$$\text{Buffer SF} = 95^{\text{th}} \text{ SF} - \text{Average SF} \quad (4-4)$$

Figures 4-16 and 4-17 show the temporal trend of 95th percentile and buffer SF percentages, respectively, at La Cholla Blvd. & River Rd., with 0 at most hours especially for the through movement. For the left-turn movement, nighttime tends to have more reliable split failure than daytime.

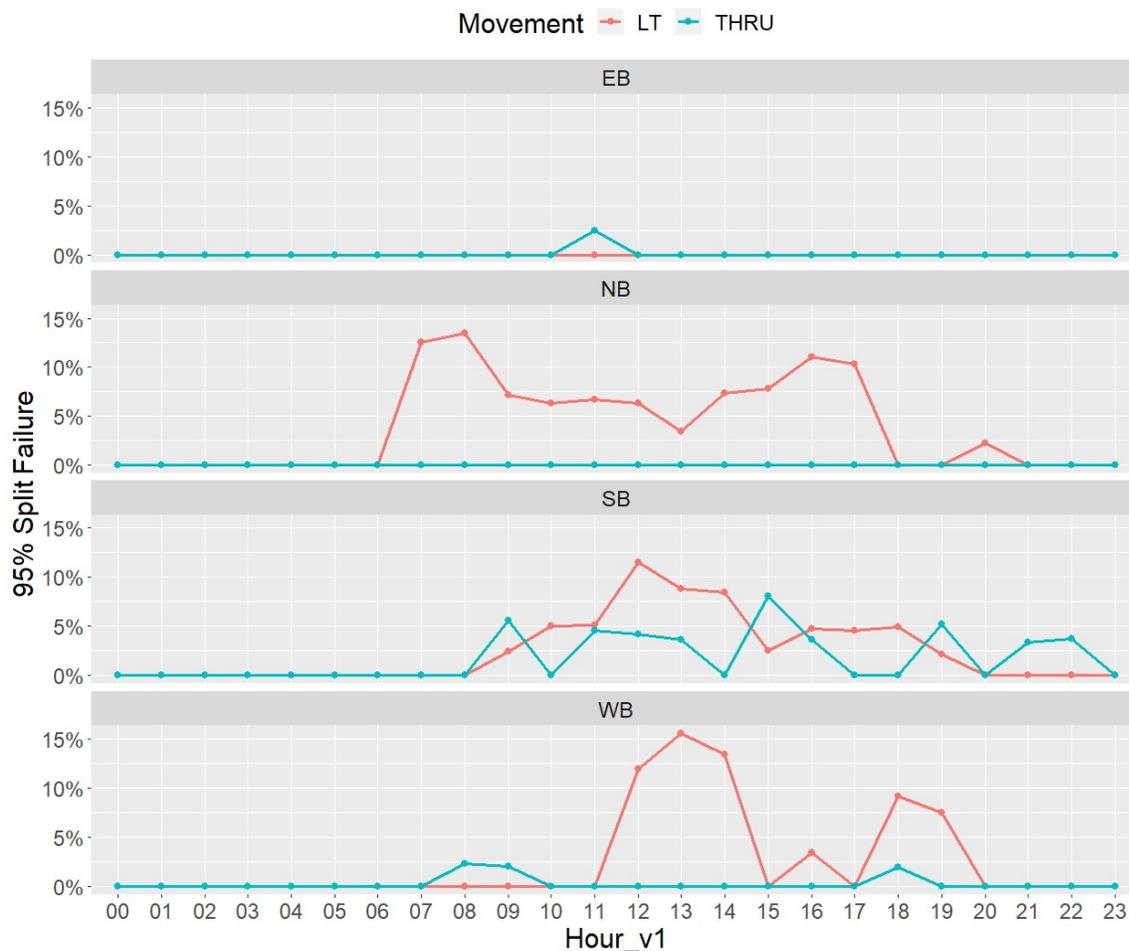


Figure 4-21. 95th percentile split failure percentage at La Cholla Blvd. & River Rd. during Jan.-March 2021

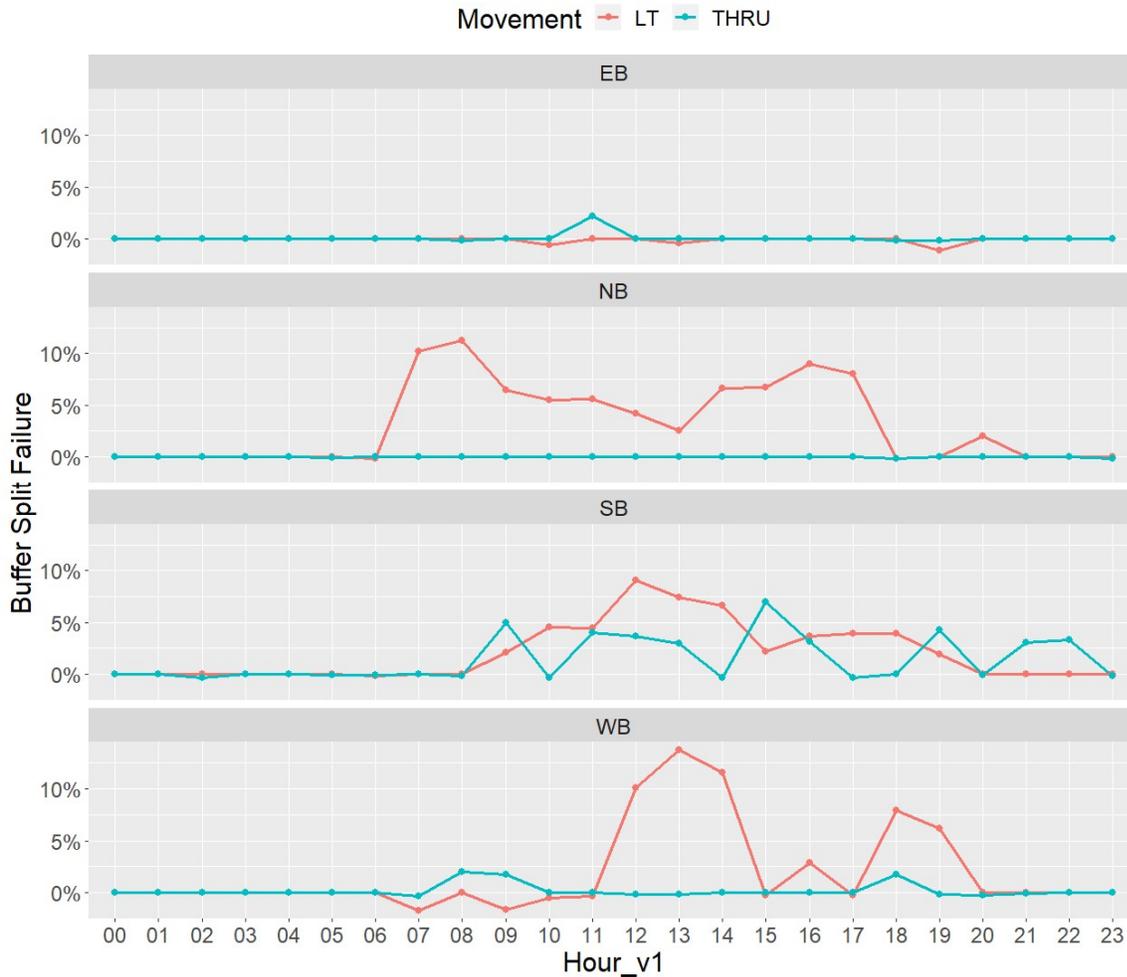


Figure 4-22. Buffer split failure percentage at La Cholla Blvd. & River Rd during Jan.-March 2021

4.2 MOBILITY/RELIABILITY PERFORMANCE COMPARISON

Because both Miovision data and Wejo data can provide mobility performance measures including delay, LOS, AoG, and split failure, we compare these two data sources and quantify the similarity in this section.

4.2.1 Delay Comparison

Figure 4-18 shows the comparison between Miovision delay and Wejo delay at La Cholla Blvd. & River Rd. This comparison uses two days of data and shows that both the left-turn and through delay calculated from Wejo data fluctuate more and have more outliers than Miovision delay,

especially during the nighttime due to the small sample of Wejo vehicles. In addition, the Wejo-based data is slightly lower than Miovision-based delay during the daytime, which could be because of the calculation method difference.

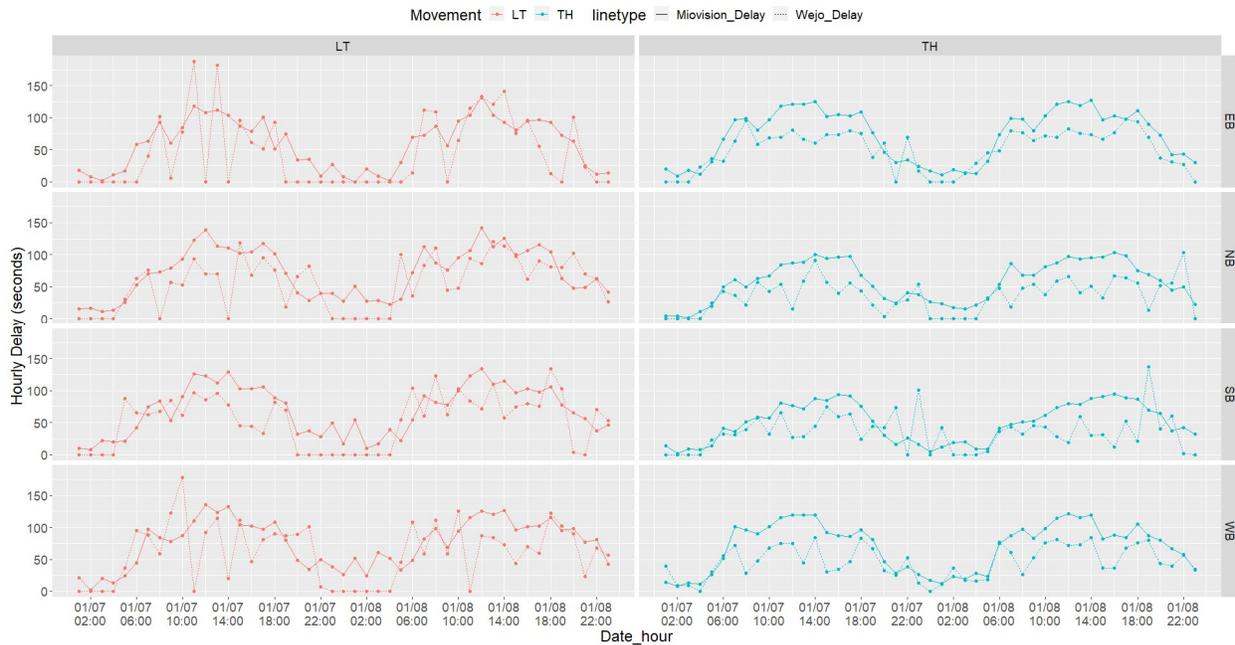


Figure 4-23. Comparison between Miovision delay and Wejo-based delay at La Cholla Blvd. / River Rd.

The Kolmogorov-Smirnov (K-S) test is used to decide if the delay data from these data sources are from the same distribution by statistically quantifying the distance between the two cumulative distributions. **Figure 4-19** shows the cumulative distribution function (CDF) comparison of two types of delay at La Cholla Blvd. / River Rd., where “D” represents the maximum vertical distance between the two CDFs and “P”, the p-value, is the probability that the two data sources are similar under the null hypothesis of the K-S test. All p-values of the K-S test for both the left-turn and through movements are zero, less than 0.05, indicating that Wejo delay is statistically different from Miovision delay. In addition, the probability of Wejo delay being zero is high because some time periods had no sample Wejo data, especially during the nighttime.

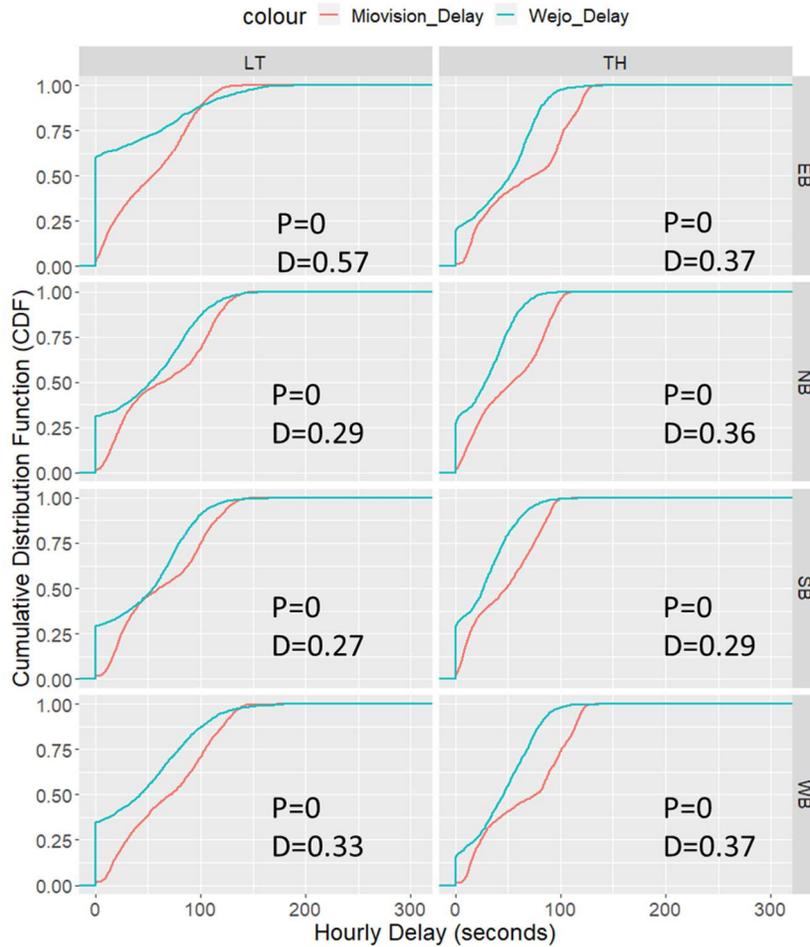


Figure 4-24. Delay cumulative distribution function comparison at La Cholla Blvd. / River Rd.

Additionally, the two types of delay reliability are compared using the four reliability indices. **Figures 4-20 to 4-22** show the temporal trends of Miovision-based and Wejo-based delay reliability at La Cholla Blvd. & River Rd. For the 95th percentile delay, the two types of data show a similar trend with higher reliability during nighttime than during daytime, but the measures derived from Wejo data fluctuate more due to the varying sample sizes. The temporal trend of the buffer index is different from the other two indices, but the two types of data show a similar trend, with the daytime reliability higher than nighttime.



Figure 4-25. Comparison between Miovison and Wejo-based 95th percentile delay at La Cholla Blvd. / River Rd.

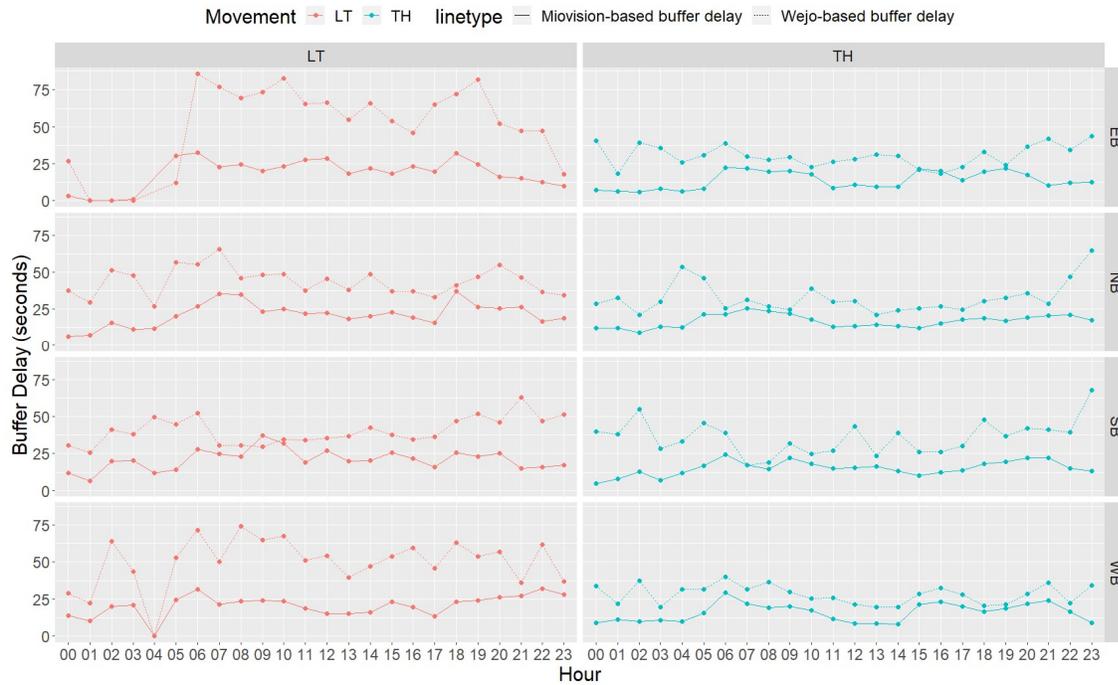


Figure 4-26. Comparison between Miovision and Wejo-based buffer delay at La Cholla Blvd. / River Rd.

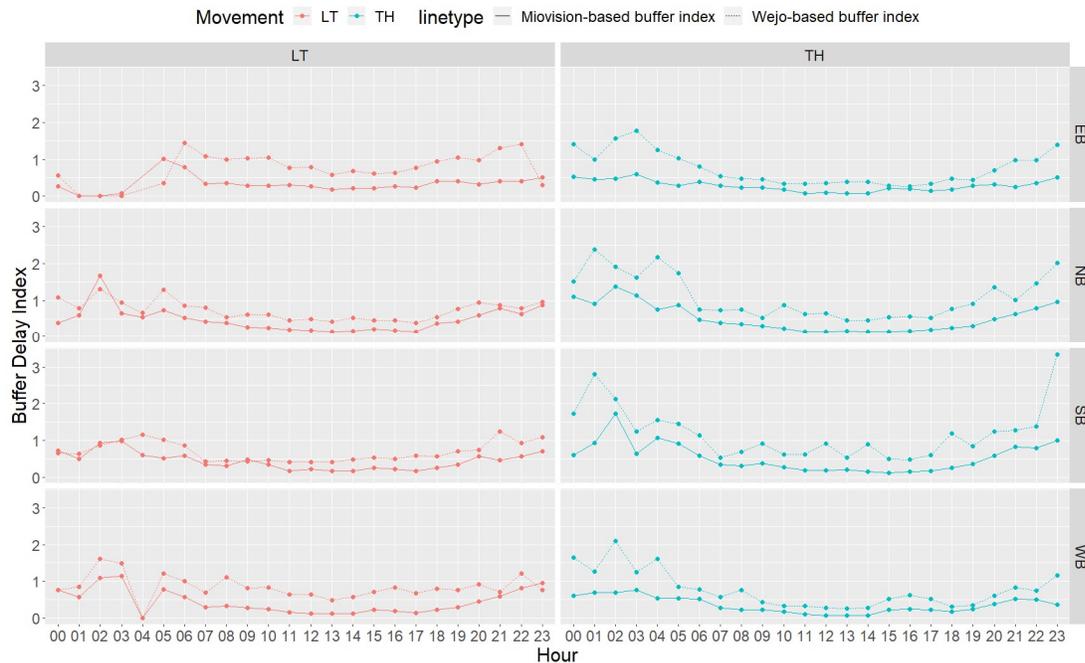


Figure 4-27. Comparison between Miovision and Wejo-based buffer index at La Cholla Blvd. / River Rd.

The K-S test is again used to quantify the similarity between the measures derived from Miovision data and the measures derived from Wejo data. As shown in **Figure 4-23**, the P-value for both left-turn and through movements for both the buffer index are zero, indicating that the distributions are significantly different for both indices. One possible reason for this difference is that the Wejo sample sizes vary too much. To provide more robust reliability performance measures, sufficient sample sizes are needed consistently.

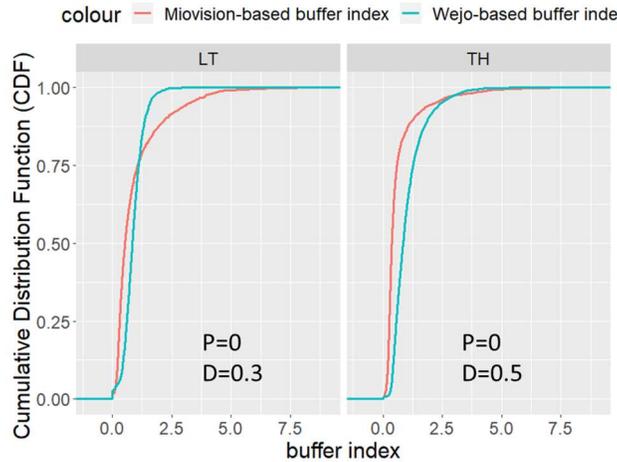


Figure 4-28. Reliability indices cumulative distribution function comparison at La Cholla Blvd. / River Rd.

4.2.2 Level-of-Service Comparison

The Wejo delay and Miovision delay are used to calculate the intersection LOS according to HCM. **Figure 4-24** shows the comparison between Wejo LOS (Wejo_LOS) and Miovision LOS (Mio_LOS) at La Cholla Blvd. / River Rd. Both types of LOS show the same temporal trend with a low LOS in the daytime and a high LOS at nighttime. The Wejo LOS tends to be one level higher than Miovision LOS, especially during the daytime, likely because the Wejo-based delay is lower than Miovision-based delay according to the results in Section 4.2.1.

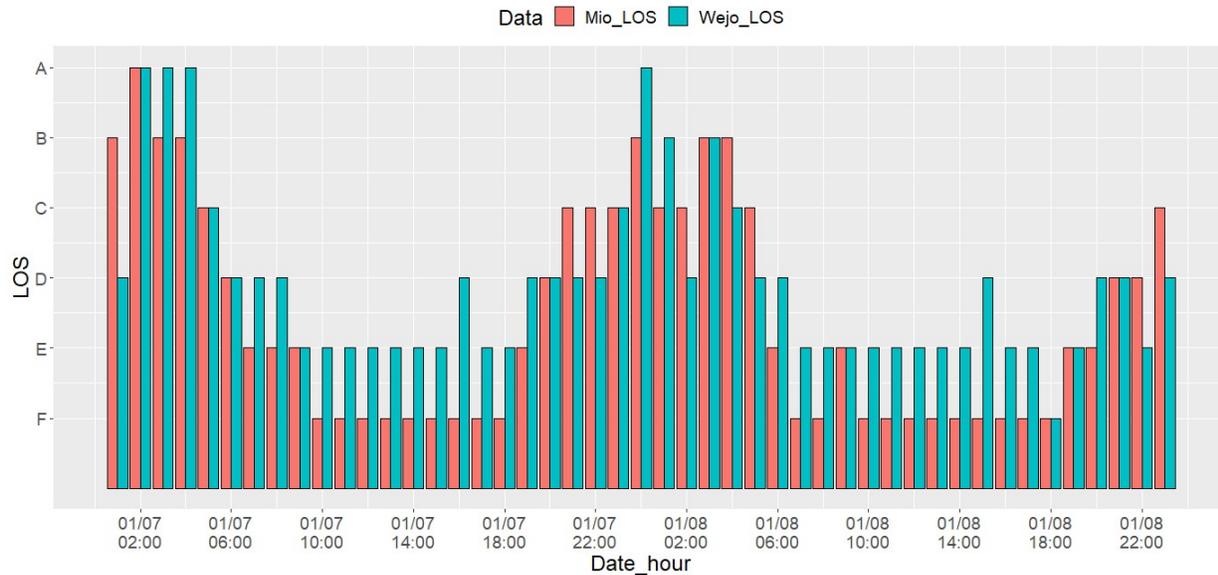


Figure 4-29. LOS comparison at La Cholla Blvd. / River Rd, Tucson

To quantify the similarity between Wejo-based and Miovision-based LOS, the data at all study intersections are used to calculate the confusion matrix to visualize the performance of two types of delay comparison, and the results are shown in **Table 4-4**. The total accuracy is 0.568.

Table 4-4. Confusion matrix of LOS comparison

		Miovision LOS					
		A	B	C	D	E	F
Wejo LOS	A	32,194	9,889	1,539	579	390	184
	B	5,766	10,855	3,454	373	83	17
	C	1,510	4,462	6,598	3,715	482	103
	D	239	304	1,335	4,466	4,643	850
	E	30	29	44	188	1,170	1,688
	F	8	9	9	18	8	81

4.2.3 Arrival-on-Green Comparison

Figure 4-25 shows the comparison between Miovison AoG and Wejo AoG at La Cholla Blvd. / River Rd. Wejo AoG fluctuates more than Miovison AoG, especially during nighttime when the number of Wejo vehicles is low. In addition, some periods had no Wejo sample vehicles to calculate the AoG, and so those values are missing. **Figure 4-26** shows the AoG CDF comparison at La Cholla Blvd. / River Rd., and the p-value of the K-S test for each of the four directions is 0, indicating that Wejo AoG is statistically different from Miovison AoG.

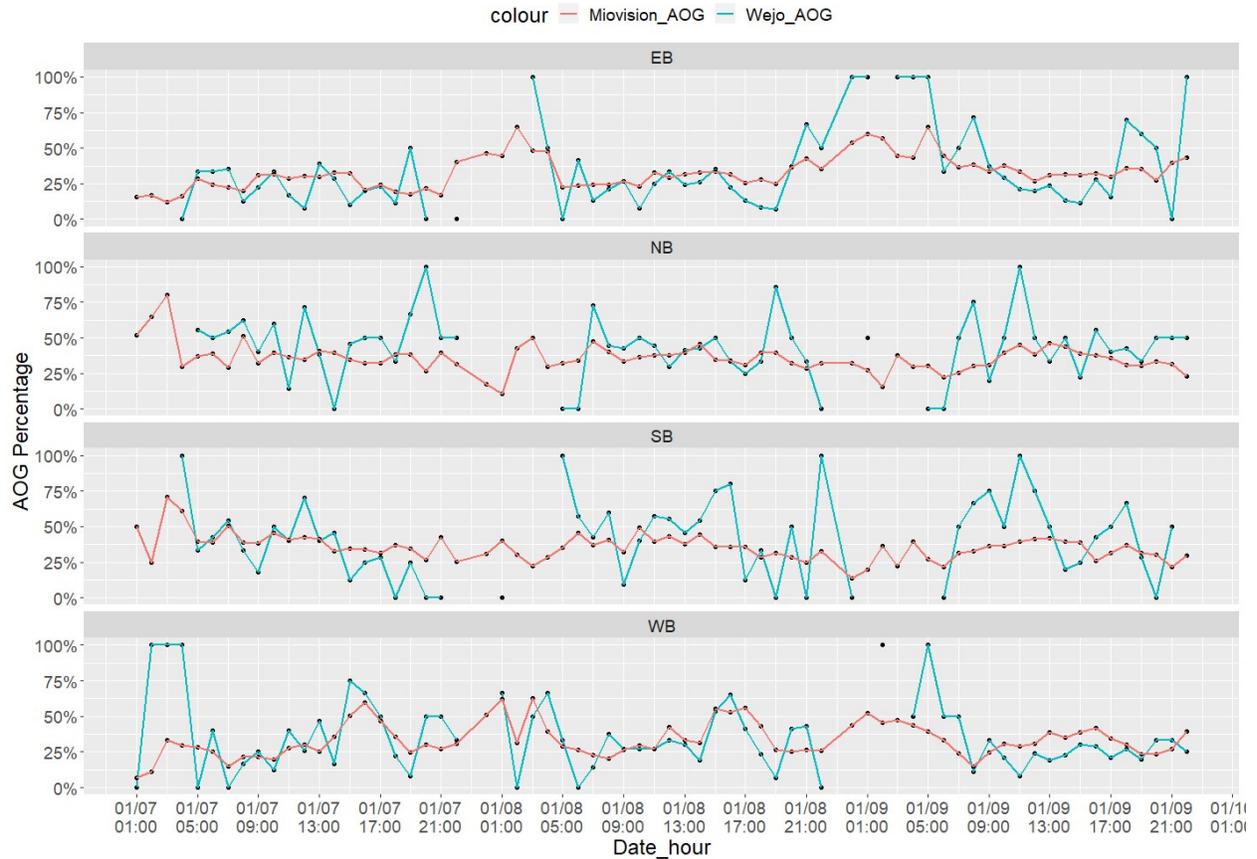


Figure 4-30. Comparison between Miovison AoG and Wejo AoG at La Cholla Blvd. / River Rd.

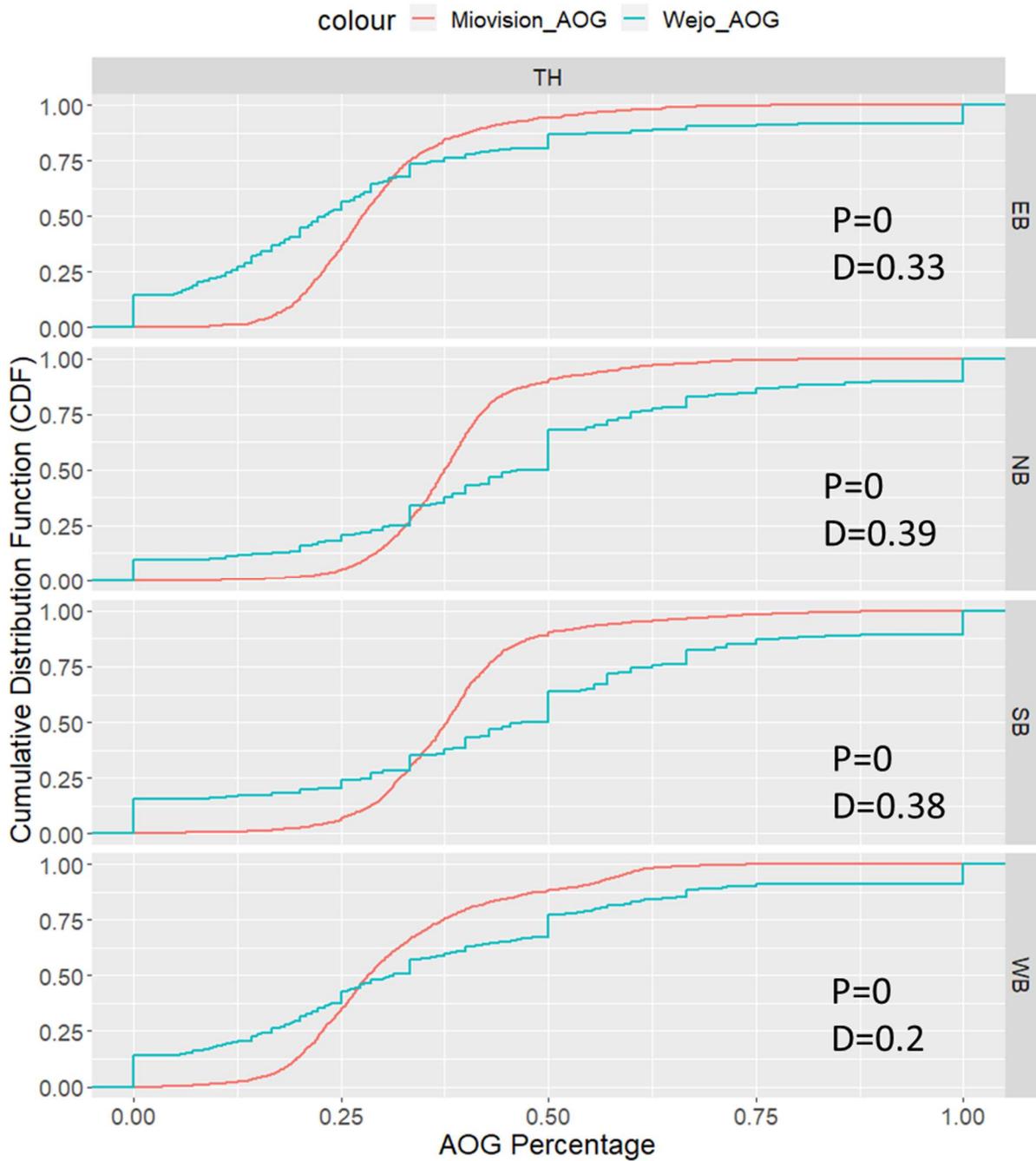


Figure 4-31. AOG cumulative distribution function comparison at La Cholla Blvd. / River Rd

Figures 4-27 to 4-28 compare the three reliability performance measures derived for AoR from Wejo data to those same measures derived from Miovision data at La Cholla Blvd. & River Rd.

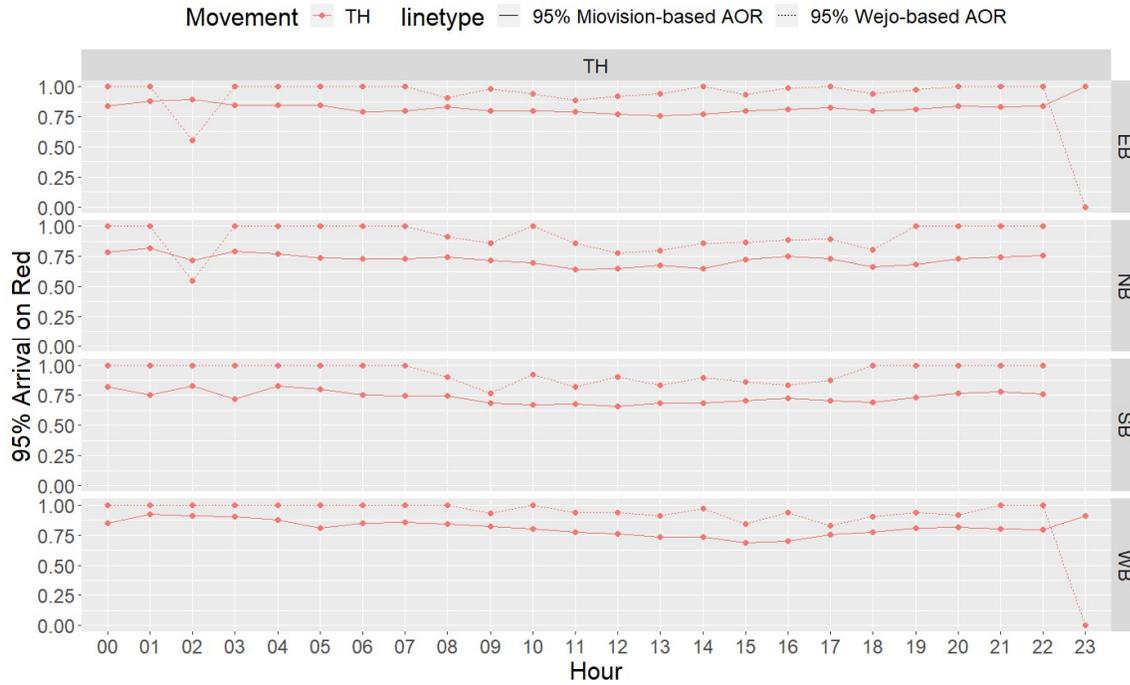


Figure 4-32. 95th percentile AoR comparison at La Cholla Blvd. / River Rd

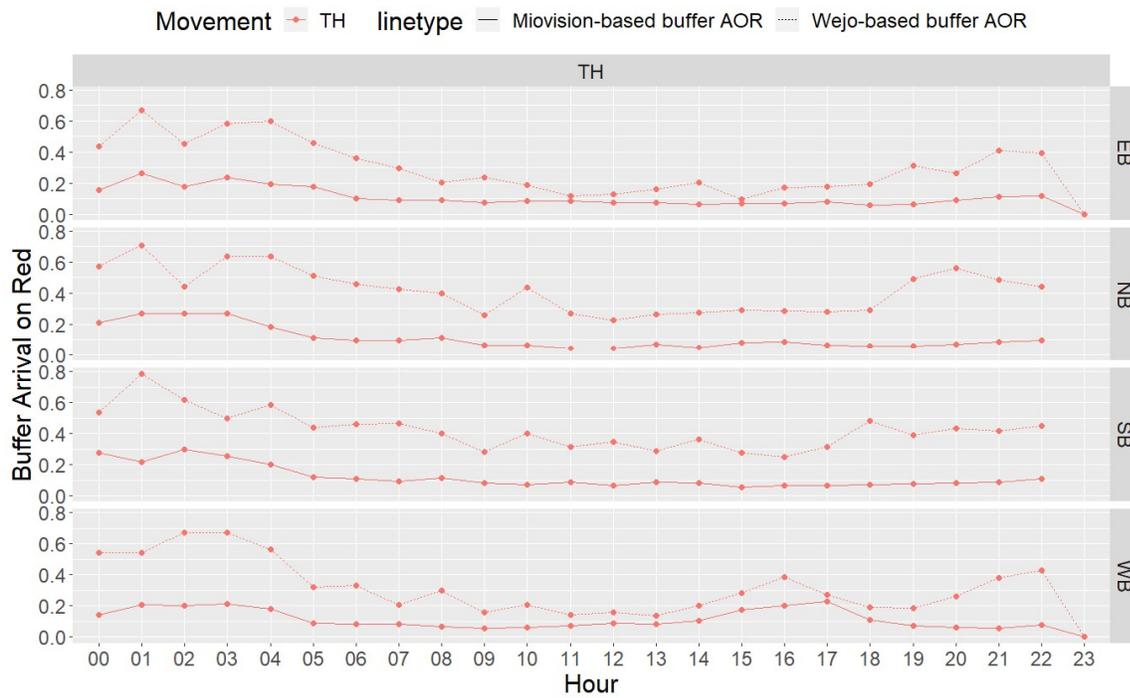


Figure 4-33. Buffer AoR comparison at La Cholla Blvd. / River Rd.

4.2.4 Split Failure Comparison

Figure 4-29 shows the comparison between Wejo split failure and Miovison split failure at La Cholla Blvd. / River Rd. Both data sources show that split failure rarely occurs at this signal. The left turn movement has a higher likelihood of split failure than the through movement. Wejo split failure data is missing because these times have zero Wejo vehicles and, therefore, no trips that could be used to calculate split failure.

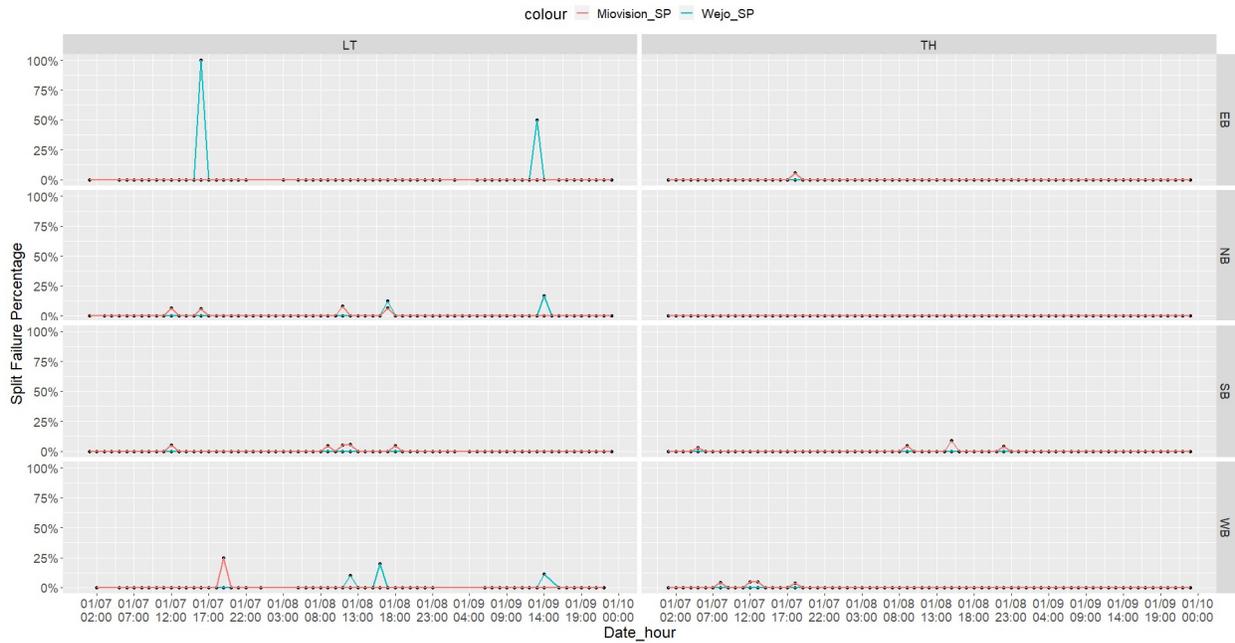


Figure 4-34. Comparison between Miovison Split Failure and Wejo Split Failure at La Cholla Blvd. / River Rd.

Figure 4-30 shows the split failure CDF comparison at La Cholla Blvd. / River Rd., and results show p-values for all directions except NB LT and SB LT larger than 0.1, indicating that Wejo split failure is statistically similar to the Miovison split failure. This similarity between the two data sources is likely because split failure is very rare at this signal. Specifically, about 99% of the calculated split failure percentages are zero.

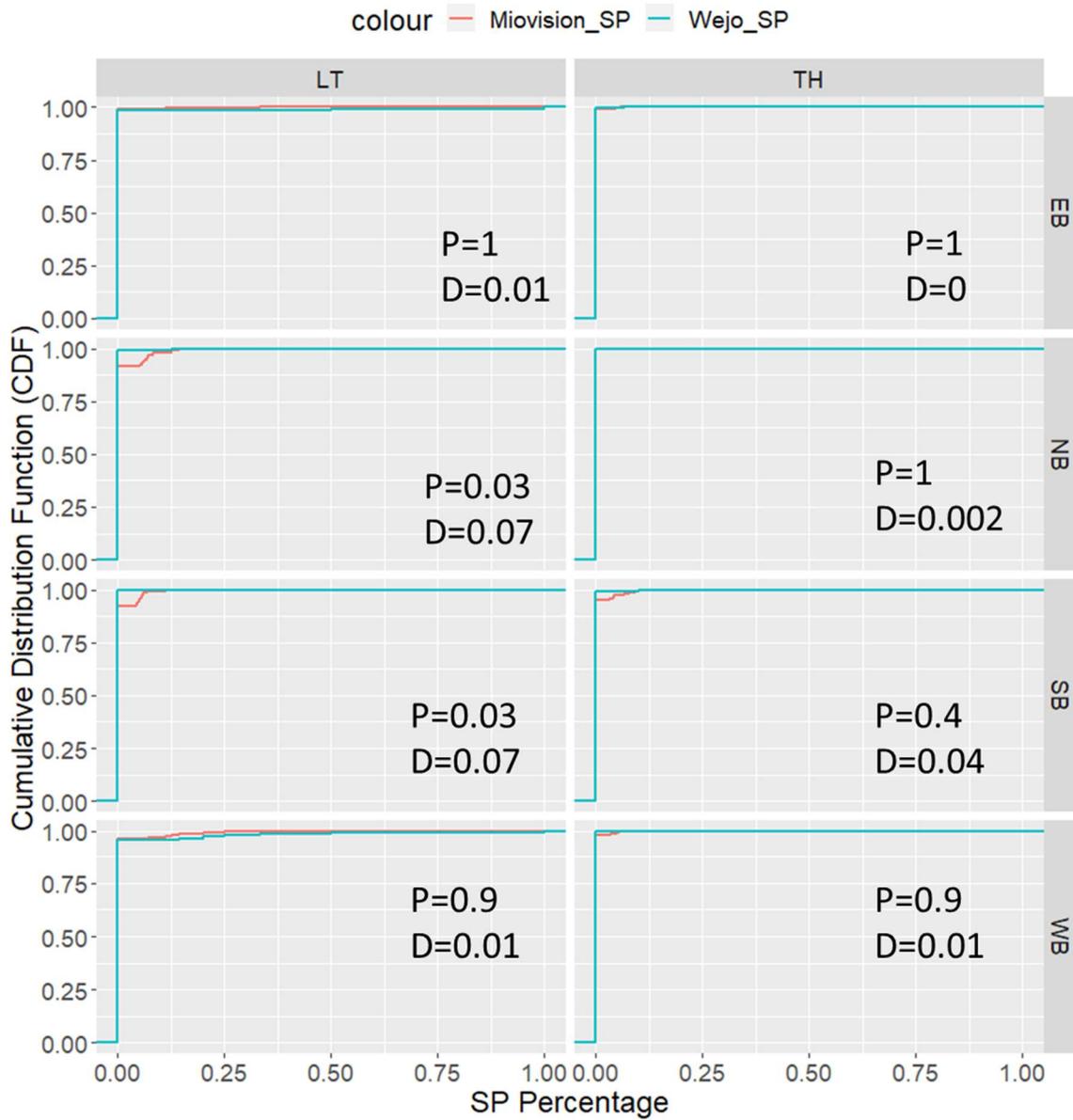


Figure 4-35. Split failure cumulative distribution function comparison at La Cholla Blvd. / River Rd.

Figures 4-31 to 4-32 compare two reliability performance measures for split failure derived from Wejo data to those same measures derived from Miovision data.

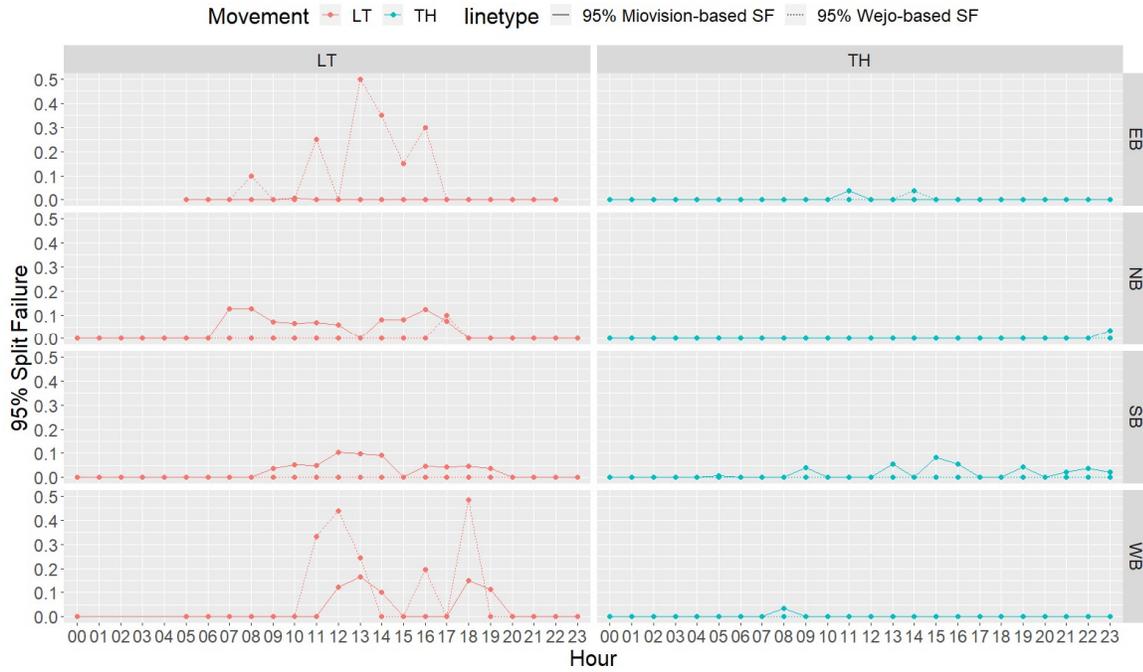


Figure 4-36. 95th percentile split failure comparison at La Cholla Blvd. / River Rd.

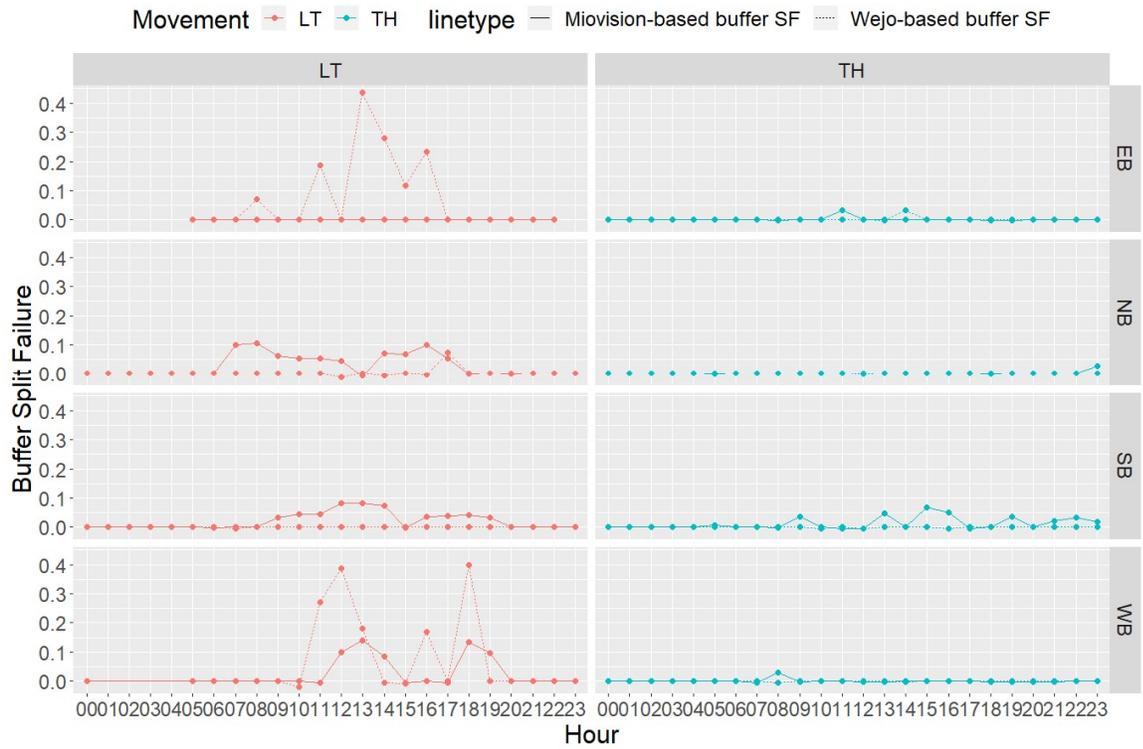


Figure 4-37. Buffer split failure comparison at La Cholla Blvd. / River Rd.

CHAPTER 5: REPRESENTATIVENESS EVALUATION OF PERFORMANCE MEASURES

Chapter 4 presents the traffic mobility performance results extracted from Wejo data and provided by the Miovision system, including delay and arrival-on-green (AOG). Through statistical analysis, differences were found between these two data sources. It is important to note that the difference between two data sources may vary based on the sample size of Wejo data at intersections. In this chapter, we will discuss the appropriate size of Wejo samples when using Wejo data as an independent measure or other model validation data. To quantify the impact of data sample size on performance accuracy, this chapter will assess the accuracy of various traffic mobility performance measures under different scenarios using traffic mobility performance data collected by the Miovision system.

5.1 STUDY LOCATIONS

To evaluate the accuracy of Wejo-based traffic performance measures at different penetration rates, this study utilizes Miovision traffic performance data. The selection of intersections for analysis is based on the availability of both data sources in the PAG region and intersection geometry. A total of 62 signalized intersections with unskewed geometry were selected for analysis, as shown in **Figure 5-1**.

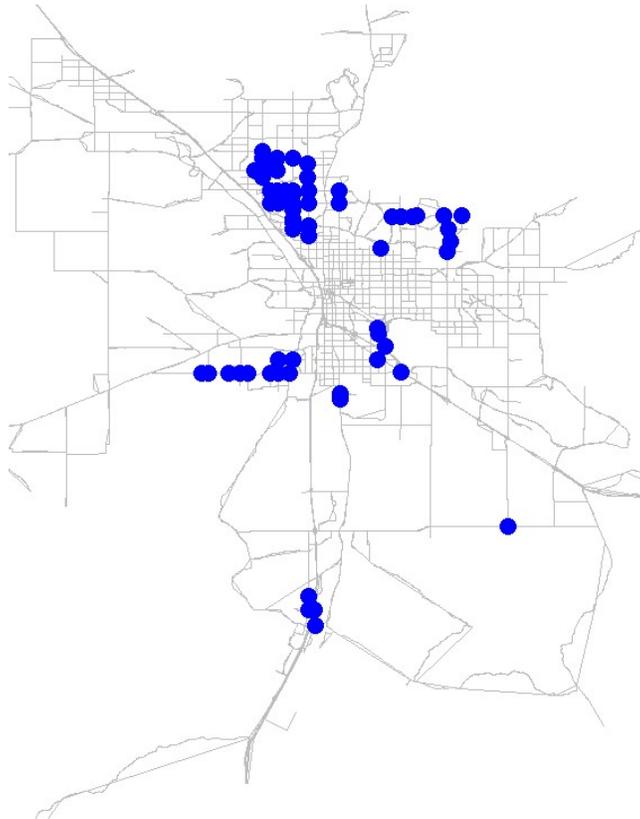


Figure 5-1. 62 study intersections in the PAG region

5.2 REPRESENTATIVENESS EVALUATION OF CONTROL DELAY

Control delay is the primary performance measure for signalized intersections. According to the proposed method in Chapter 3, the control delay is extracted from the Wejo data by calculating the travel time difference between actual speed and free-flow speed in a defined segment. The delay performance provided by the Miovision system is “simple stop delay” rather than control delay. These two types of delay are defined and measured differently and may not be identical even as the penetration rate of Wejo increases. In addition to comparing the similarity between the two data sources, the correlation between these two types of delay is investigated. Therefore, the Wejo-based delay, with sufficient sample size, should highly correlate with the Miovision-based delay. The left turn and through movements have different driver behaviors, causing inconsistent findings and results in representativeness evaluation, and so these two major movements are separated for analysis.

5.2.1 Through Movement

To compare the similarity between Miovision- and Wejo-based through movement delays, the intersection 1st Ave. & Orange Grove Rd. was selected as a sample location. We used four days of data from four directions at the study location to compare the two types of delay, as shown in **Figure 5-2**. The preliminary results indicate that the Wejo-based delay has a similar temporal trend to the Miovision-based delay, with a low delay during the night and a high delay during the daytime. However, the Wejo-based delay generally has a lower value than the Miovision-based delay. This issue is acceptable because the two delay types have different components. In addition, the Wejo-based delay fluctuates more than the Miovision-based delay, which has a smoother trend. This fluctuation of Wejo-based delay is probably caused by variations in the sample size of Wejo trajectories during different time intervals. For example, the Wejo-based delay may yield less accurate results when the sample size is too small. Therefore, the data sample size could be a significant factor to maintain the reliability and accuracy of the Wejo-based delay.

Based on this analysis we’d like to focus on the sample size analysis of Wejo data. All data collected from 62 intersections are categorized into different groups based on the sample size of Wejo data. As shown in **Figure 5-3**, all data is partitioned into 34 groups; each plot shows the results of each group; and each group contains two sample sizes. For example, the data with one and two Wejo trajectories per hour are assigned to the same group. Most difference distributions have a roughly normal distribution, but the distributions when the sample size is larger than 50 vehicles per hour become mixtures of distributions or show an unclear distribution. One explanation is the limited intersections and periods have a large sample Wejo data, displaying an unnoticeable pattern. According to the results with different sample sizes, the variance does not significantly decrease with the increase in sample size. **Figure 5-4** provides a clearer comparison of different distributions and shows that the absolute distribution mean increases while the variance does not change much with the increase in sample size. **Figure 5-5** shows the difference distributions by hours of the day.

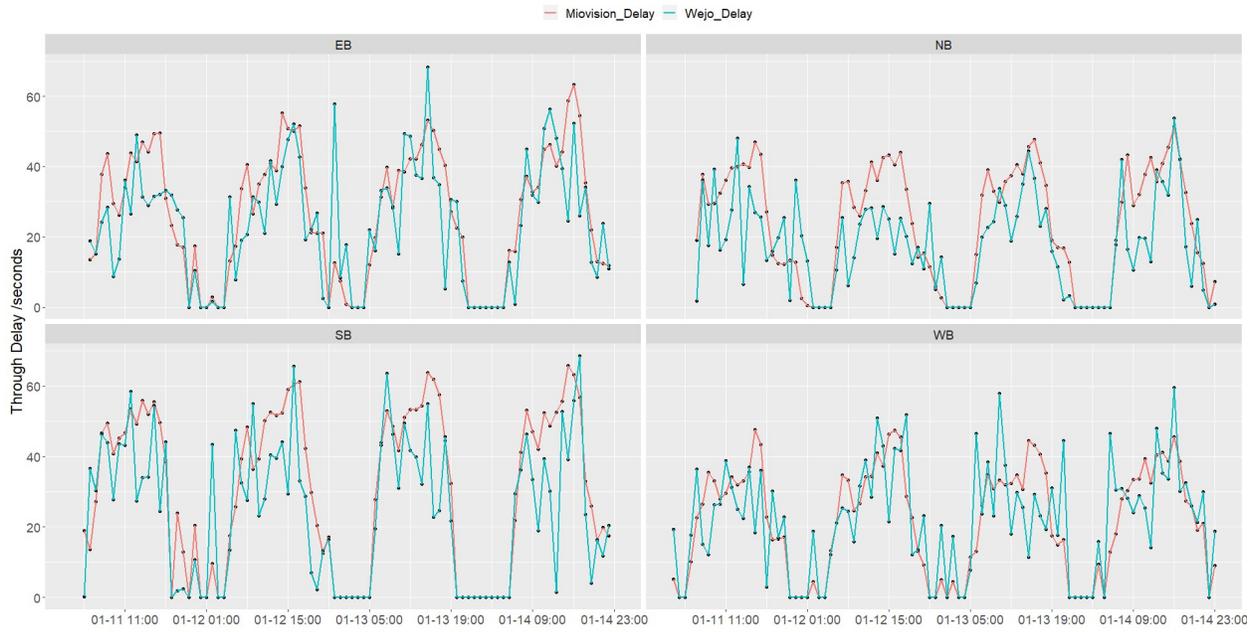


Figure 5-2. Comparison between Wejo- and Miovision-based delay of through movement

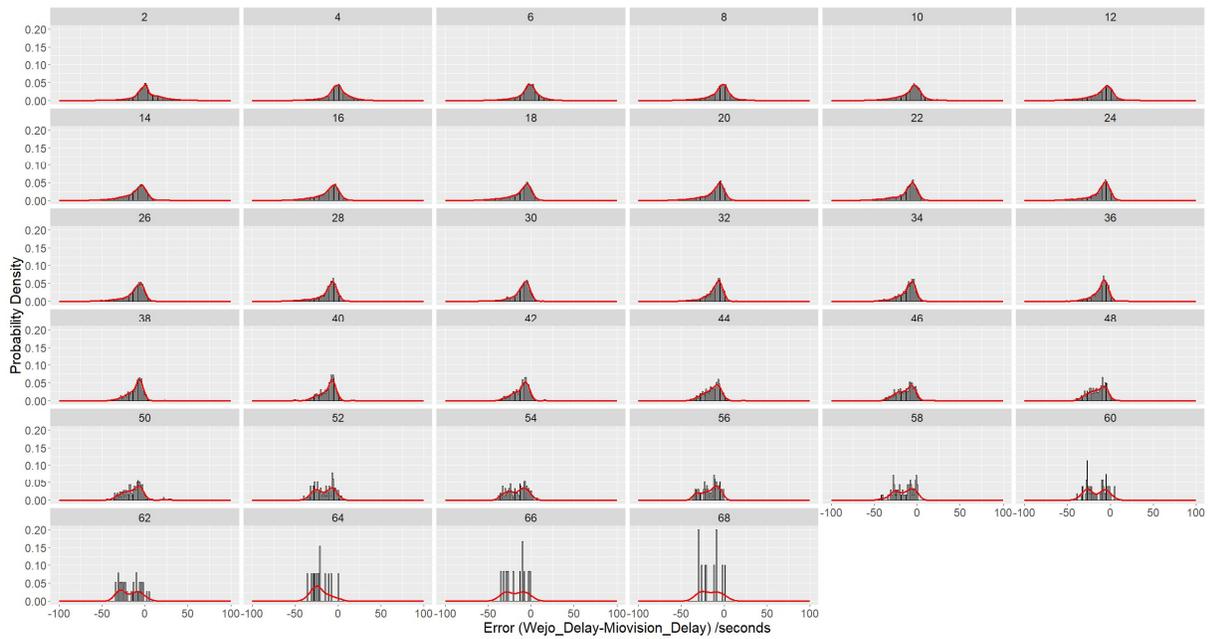


Figure 5-3. Distribution of the difference between two delays by sample size

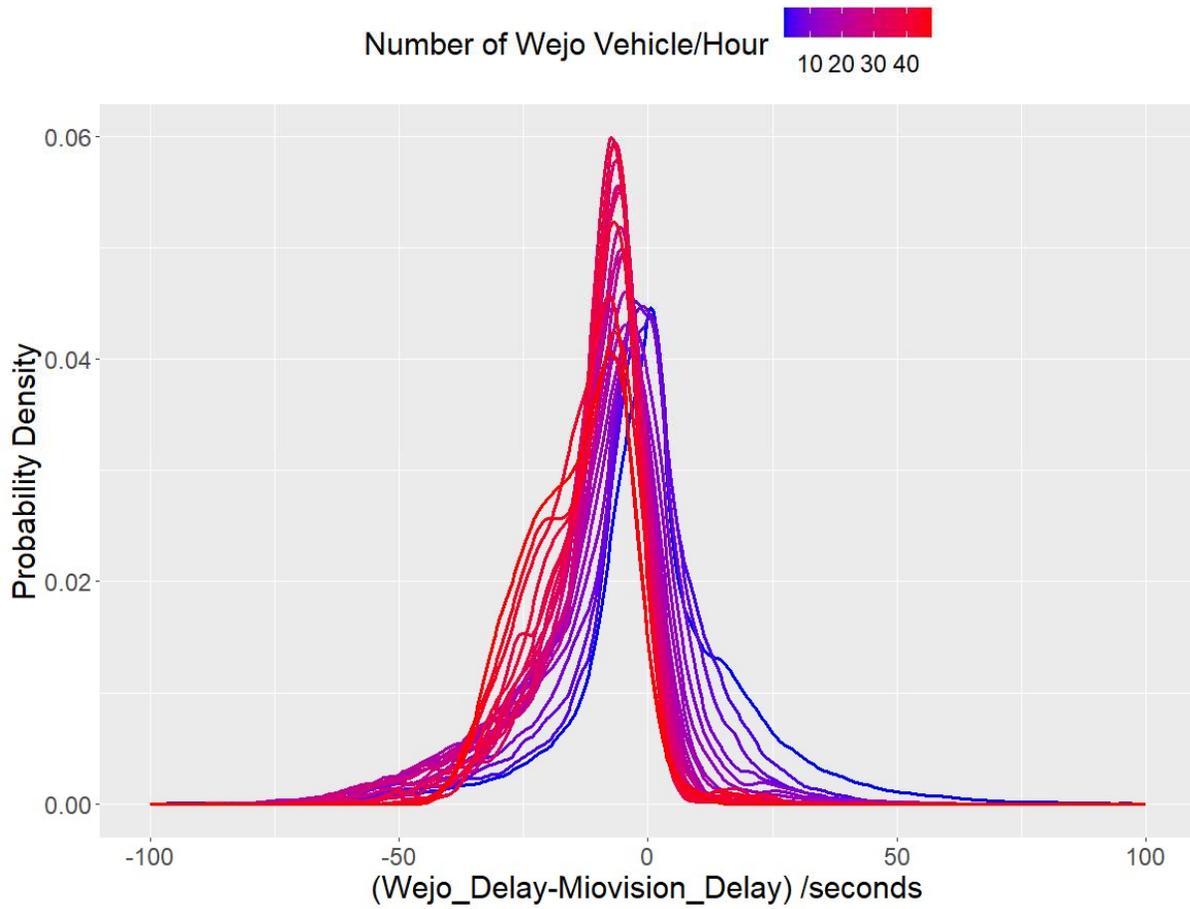


Figure 5-4. Comparison between different distributions with various sample sizes

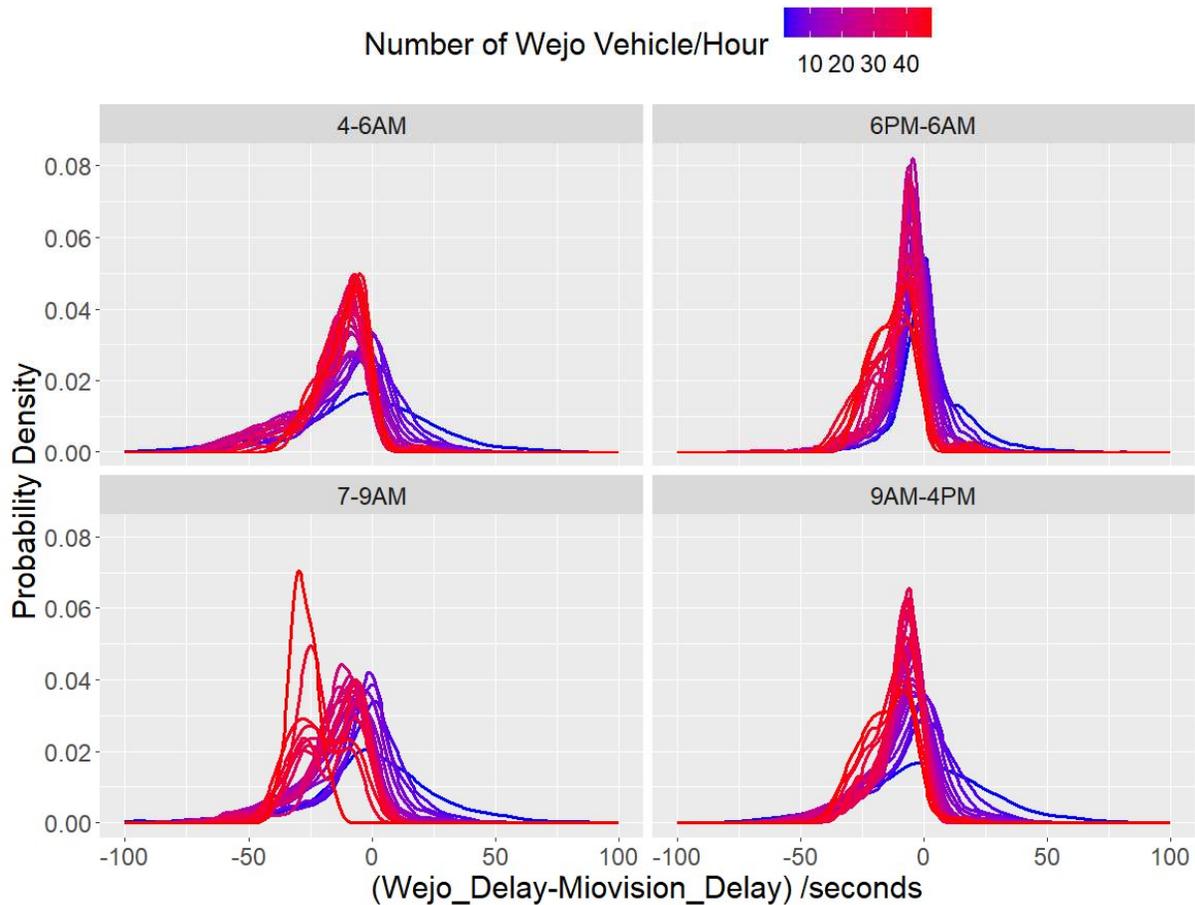


Figure 5-5. Comparison between different distributions with various sample sizes by hour of the day

Specifically, **Figure 5-4** shows that the distributions seem to shift towards the left. One possible reason for this shift to the left is the default distance selection for Wejo-based delay calculation. The original distance of 300 meters based on previous studies is used for delay calculation. In order to analyze and quantify the impact of the selected distance on the Wejo-based delay calculation, sensitivity analysis is conducted by changing the distance. The sensitivity analysis results are shown in **Figure 5-6**. When the selected distance is lower than 400 meters, the difference distribution mean decreases and becomes lower than zero with the number of Wejo trajectories increasing. When the selected distance is higher than 550 meters, the distribution mean increases with the number of Wejo trajectories increasing, and the difference between distribution variances becomes less significant. Choosing a distance between 500 and 550 meters for Wejo-based delay calculation yields a consistent distribution mean of zero, regardless of the sample size and a slight decrease in distribution variance as the sample size increases. This result of smaller variance indicates that the accuracy of Wejo-based can be enhanced by increasing the sample size, which is consistent with the prior research. However, when the sample size is larger than 15

trajectories per hour, increasing the sample size has no significant effect on the distribution variance, implying that increasing the sample size cannot improve the accuracy and reliability. One plausible explanation for this is the inherent difference between Miovision-based delay and Wejo-based delay, with Wejo-based delay potentially not matching the scale of Miovision-based delay despite having adequate sample data. Due to this intrinsic difference, using the error between the two types of delays as an indicator cannot reliably and reasonably evaluate the accuracy of the Wejo-based delay when Miovision-based delay is used as ground-truth data.

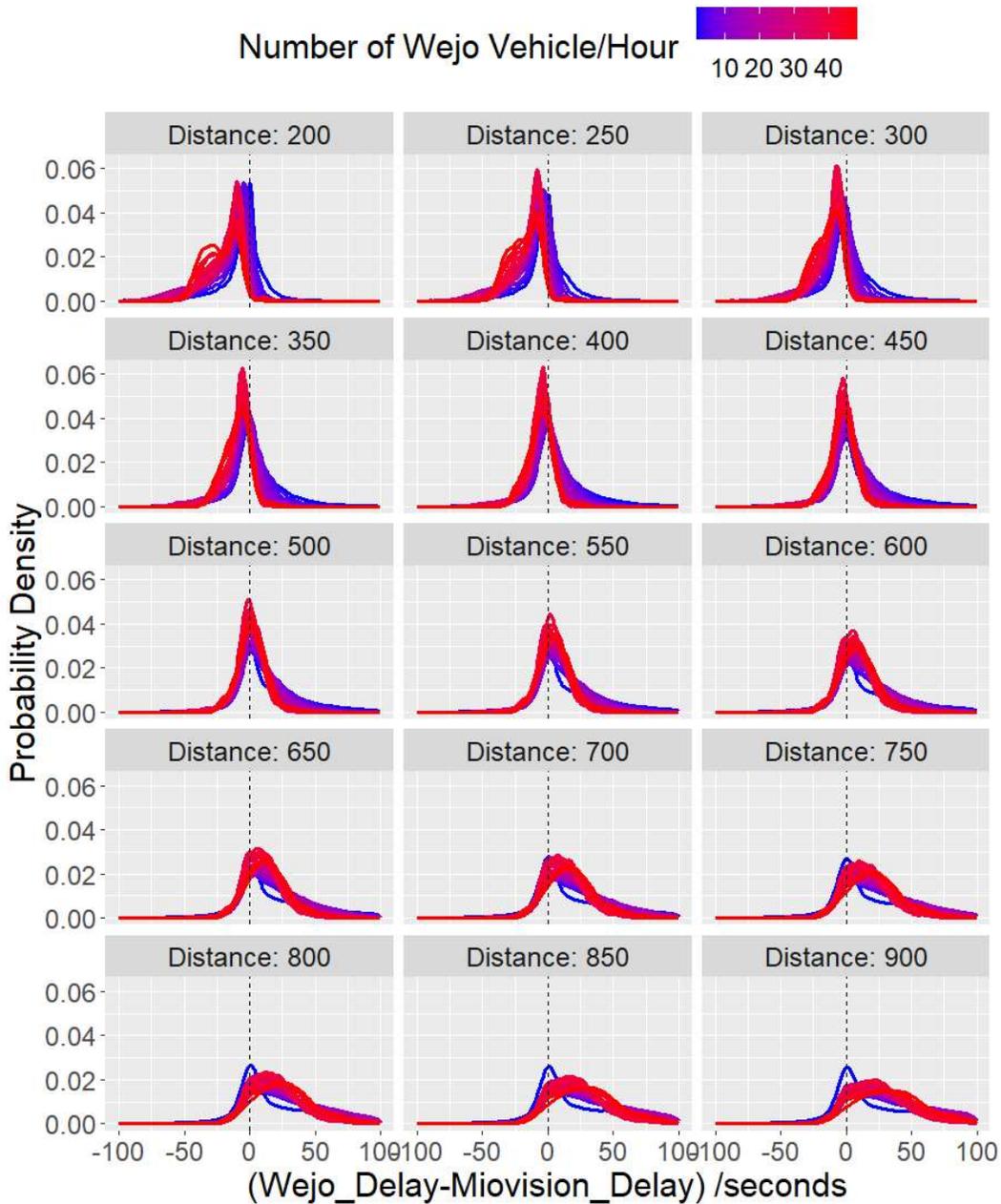


Figure 5-6. Sensitivity analysis by changing the distance of Wejo delay estimation.

Since the accurate Wejo-based delay with a sufficient sample size should qualitatively indicate the traffic conditions as the Miovision-based delay, the Pearson correlation coefficient is used to quantify the similarity between these two types of delay. The correlation coefficient r is calculated as **Eq.5-1**

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2(y_i - \bar{y})^2}} \quad \text{Eq. 5-1}$$

where x_i is the values of Wejo-based delay in a sample i ;

\bar{x} is the mean of Wejo-based delay;

y_i is the values of Miovision-based delay in a sample i ;

\bar{y} is the mean of Miovision-based delay.

The correlation between Wejo-based delay and Miovision-based delay was calculated using data collected from the through movement at 62 study intersections. **Figure 5-7** displays a high positive correlation coefficient of 0.75, indicating that the Wejo-based delay is highly correlated with Miovision-based delay. A box plot of the correlation coefficients for all directions at all intersections is presented in **Figure 5-8** and shows that, at most locations, the correlation between the two delay measures is between 0.45 and 0.75. The variation in correlation coefficients across locations is likely due to the sample sizes of Wejo trajectories, which vary with location and time. Most relevant studies use two common indicators to determine the sample size of probe vehicle data: penetration rate and the number of trajectories. To control the data quality of Wejo-based delay, it is important to determine 1) which indicator should be used to indicate the Wejo data sample size in our study, and 2) what the threshold of the sample size is for yielding accurate and reliable Wejo-based delay.

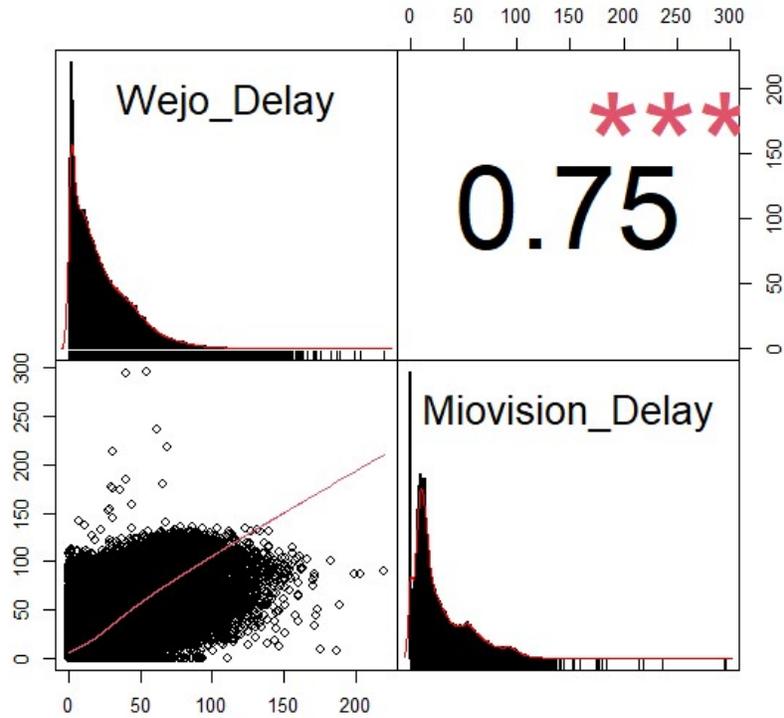


Figure 5-7. Correlation between Wejo- and Miovision-based delay for through movement

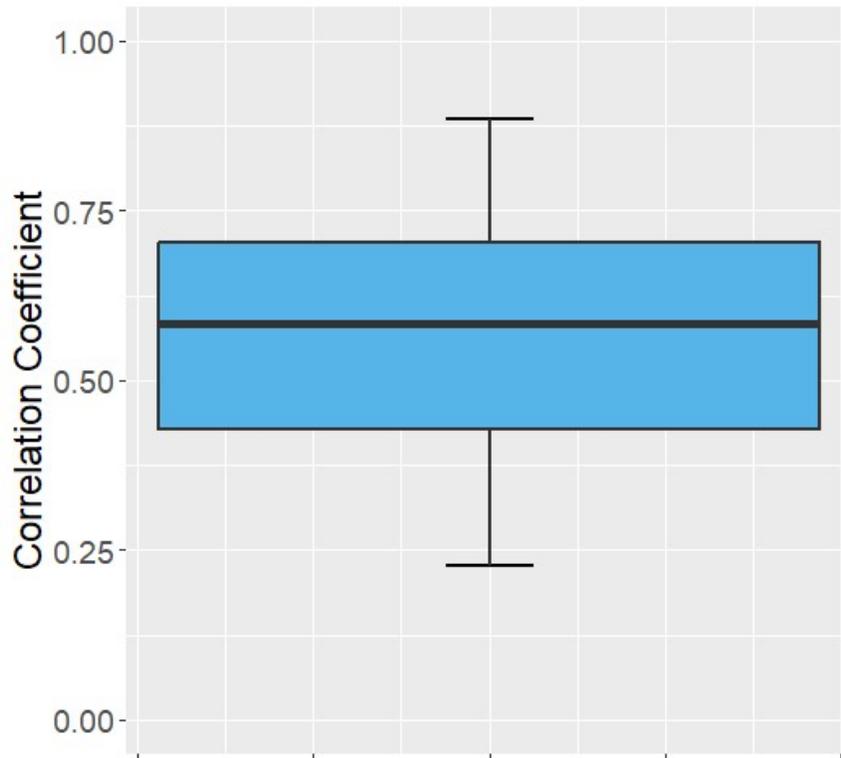


Figure 5-8. Box plot of correlation coefficients for all directions at all intersections

Figure 5-9 illustrates the correlation coefficients between Miovision-based delay and Wejo-based delay under different sample sizes. As seen in **Figure 5-9(a)** with the number of trajectories per hour as the sample size indicator, the correlation coefficient starts at 0.6 for the smallest sample size and increases as the sample size increases. Once the sample size exceeds 10 vehicles/hour, the coefficient is around 0.82, and only slightly increases even the sample size increases much more. With the penetration rate as the sample size indicator as shown in **Figure 5-9(b)**, the coefficient increases relatively smoothly as the penetration rate increases and remains below 5%. Once the penetration rate exceeds 5%, the coefficient has no clear trend and has many outliers. Comparing the results of using the number of trajectories with the use of the penetration rate, the coefficient sensitivity analyses for both indicators show a clear and smooth increase as the sample size increases until the maximum coefficient is reached. The maximum correlation coefficient is achieved when the Wejo-based delay has a sufficient sample size of 10 trajectories per hour, and the coefficient does not significantly increase regardless of the increase in sample size. However, the penetration rate has no clear sufficient minimum value because the correlation is lower with a higher penetration rate, which is unreasonable. Thus, these results suggest that the number of trajectories is a more reasonable and reliable indicator of sample size for Wejo data.

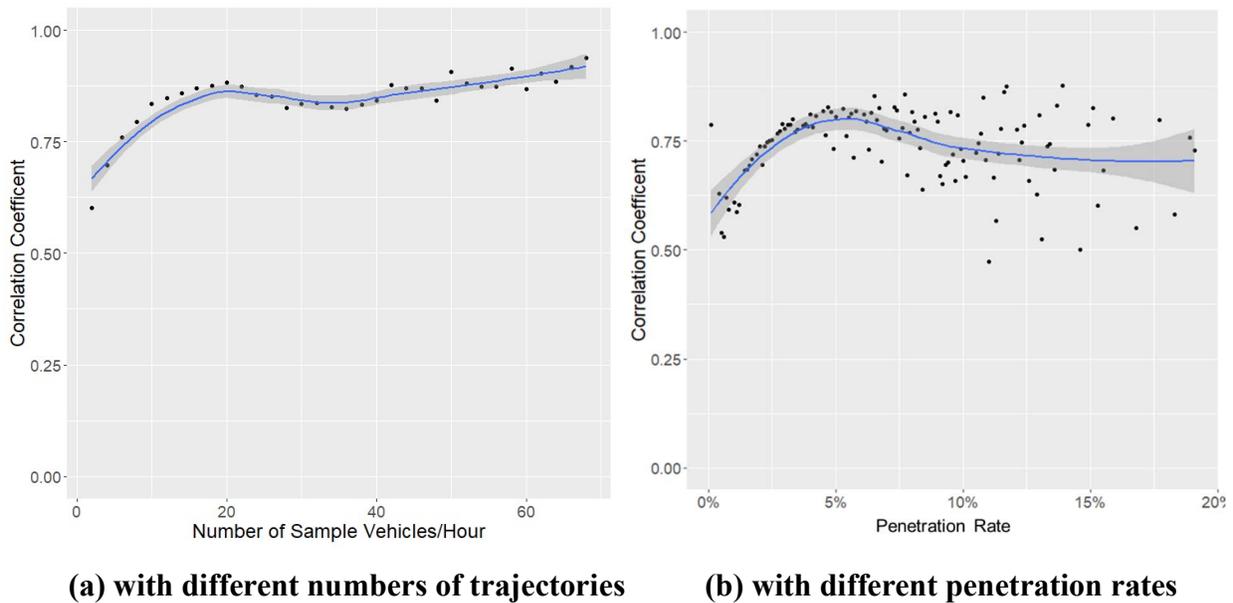


Figure 5-9. Correlation coefficients under different sample sizes

It is important to consider traffic volume when selecting indicators and threshold values for Wejo data sample size. The sensitivity analysis is conducted by calculating the correlation coefficient trend under different levels of traffic volume approaching an intersection. The traffic volume is categorized into four groups: Free-flow traffic (<100 vph), low volume (100-500 vph), moderate volume (500-1000 vph), and high volume (>1000 vph). **Figure 5-10(a)** shows the correlation

coefficient trend with the number of trajectories under different traffic volumes. The correlation coefficient has a consistent trend for different traffic volume levels, except for the high-volume scenario. When the number of trajectories is below 10, there is an increasing trend, but it becomes flat after that. In the scenario with over 1000 vph, the increasing trend is absent because no data is available with a sample size lower than 20, and the trend is flat when the sample size is over 20, which matches the trends for the other traffic volume levels. **Figure 5-10(b)** shows how the correlation coefficients vary with the penetration rate under different traffic volumes. The results show that the correlation coefficient has different trends for different traffic volume levels. In the free-flow traffic scenario, the correlation coefficient slightly increases as the penetration rate increases but remains below 0.75 even when the penetration rate exceeds 10%. Different traffic volume levels require different penetration rates to reach the highest correlation coefficient. For instance, when the traffic volume is 100-500 vph, a penetration rate of 5% can yield a coefficient of 0.82, while a traffic volume of 500-1000 vph only requires a penetration rate of 2.5% to achieve the same coefficient. Therefore, there is not a consistent penetration rate that allows samples with different traffic volume levels to produce a reliable and accurate delay (by achieving the highest correlation between Wejo-based and Miovision-based delay). Determining the threshold values for penetration rate for different scenarios required additional volume data, which is time-consuming and costly to collect. However, the number of trajectories as the sample size indicator shows consistent correlation coefficients regardless of traffic volume. Hence, the number of trajectories can consistently indicate data accuracy under different scenarios without the need for traffic volume information.

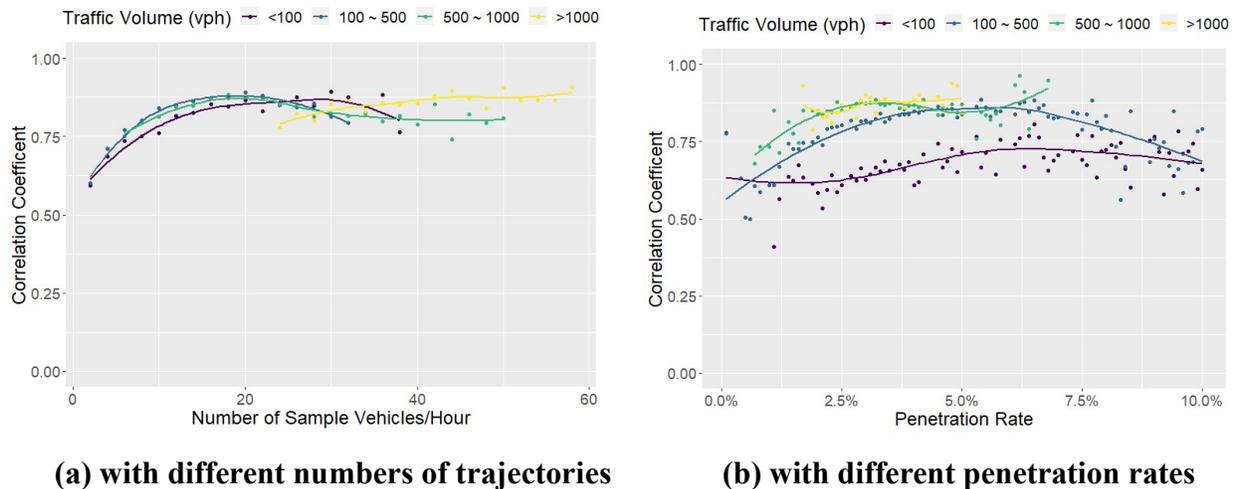


Figure 5-10. Correlation coefficients for various traffic volume levels

5.2.2 Left-turn Movement

Due to the travel behavior difference between through and left-turn traffic, the results regarding the indicator selection and associated threshold values from the above analyses may not apply to left-turn traffic. The same analysis method is applied to left-turn traffic for controlling the data accuracy of Wejo-based delay for the left-turn movement. Using 1st Ave. & Orange Grove Rd. as an example location, the Wejo-based delay is compared with the Miovision-based delay regardless of the sample size, as shown in **Figure 5-11**. These two types of delay consistently show that left-turn traffic has low delay or even no delay at night. During the daytime, both types of delay show that the left-turn traffic has a higher delay than nighttime. However, Wejo-based delay shows more fluctuation than Miovision-based delay, likely due to the low sample size of Wejo data, which may affect its reliability.

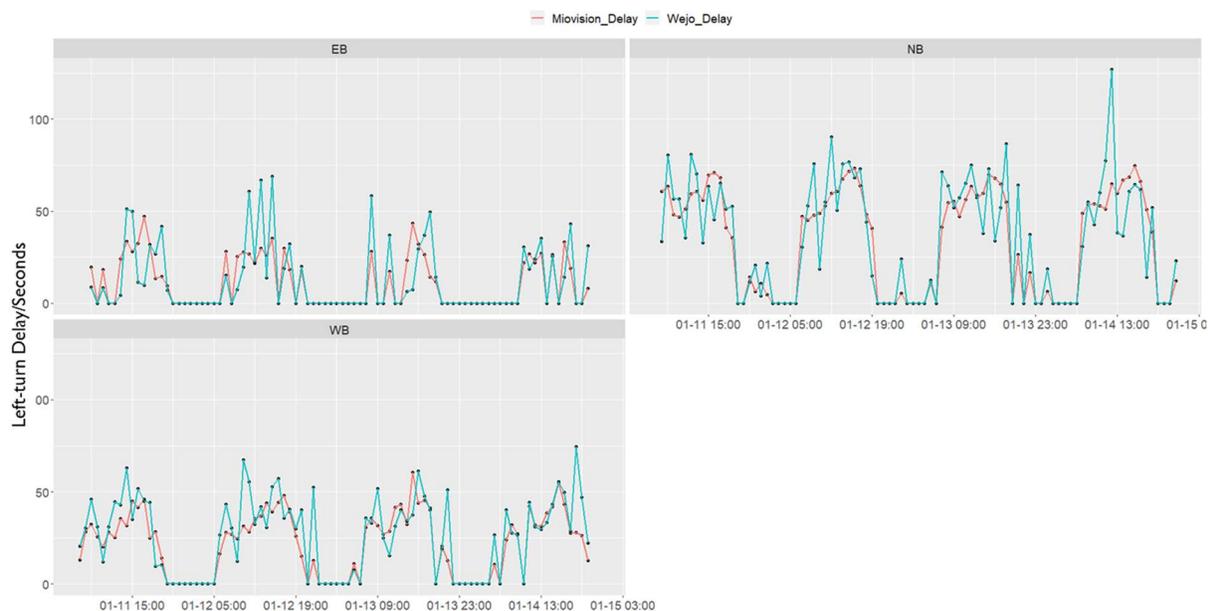


Figure 5-11. Comparison between Wejo- and Miovision-based delay of left-turn movement.

The difference between Miovision- and Wejo-based delay was calculated using left-turn data collected from all study intersections, and the difference distributions with various sample sizes are displayed in **Figure 5-12**. The 15 distributions show a leftward shift with a lower mean and similar standard deviation as the sample size increases. When the sample size is greater than 20 vehicles per hour, the distribution consists of two normal distributions. One of the normal distributions with a positive mean suggests that Wejo-based delay is higher than Miovision-based delay. Upon further examination of the data and traffic videos, it was discovered that the mixture distribution was caused by an outlier intersection, Hermans Rd. & Nogales Hwy, which had high left-turn volume but zero delay. This was likely due to communication loss resulting in missing Miovision-based delay data. Figure 12 shows the adjusted distributions after removing outliers,

with a higher error and a lower variance as sample size increases. When using the error as the data quality control indicator, the information as shown in **Figure 5-13** is unclear and difficult to interpret, which makes it challenging to draw any meaningful conclusions. The selected distance for Wejo-based delay calculation is one of the possible factors causing this interpretation difficulty. Therefore, a sensitivity analysis is conducted by changing the selected distance, and the results are shown in **Figure 5-14**. The results of the sensitivity analysis show the distribution trend as sample size varies becomes more interpretable when the distance is 400m or 450m, where distribution mean is zero regardless of the sample size, but the variance decreases as the sample size increases. These interpretable results indicate that larger sample size can improve the reliability of Wejo-based delay. However, as with the through traffic, due to this intrinsic difference between Miovision-based delay and Wejo-based delay, using the error between the two types of delays as an indicator cannot reliably and reasonably evaluate the accuracy of the Wejo-based delay for the left-turn movement when Miovision-based delay is used as ground-truth data.

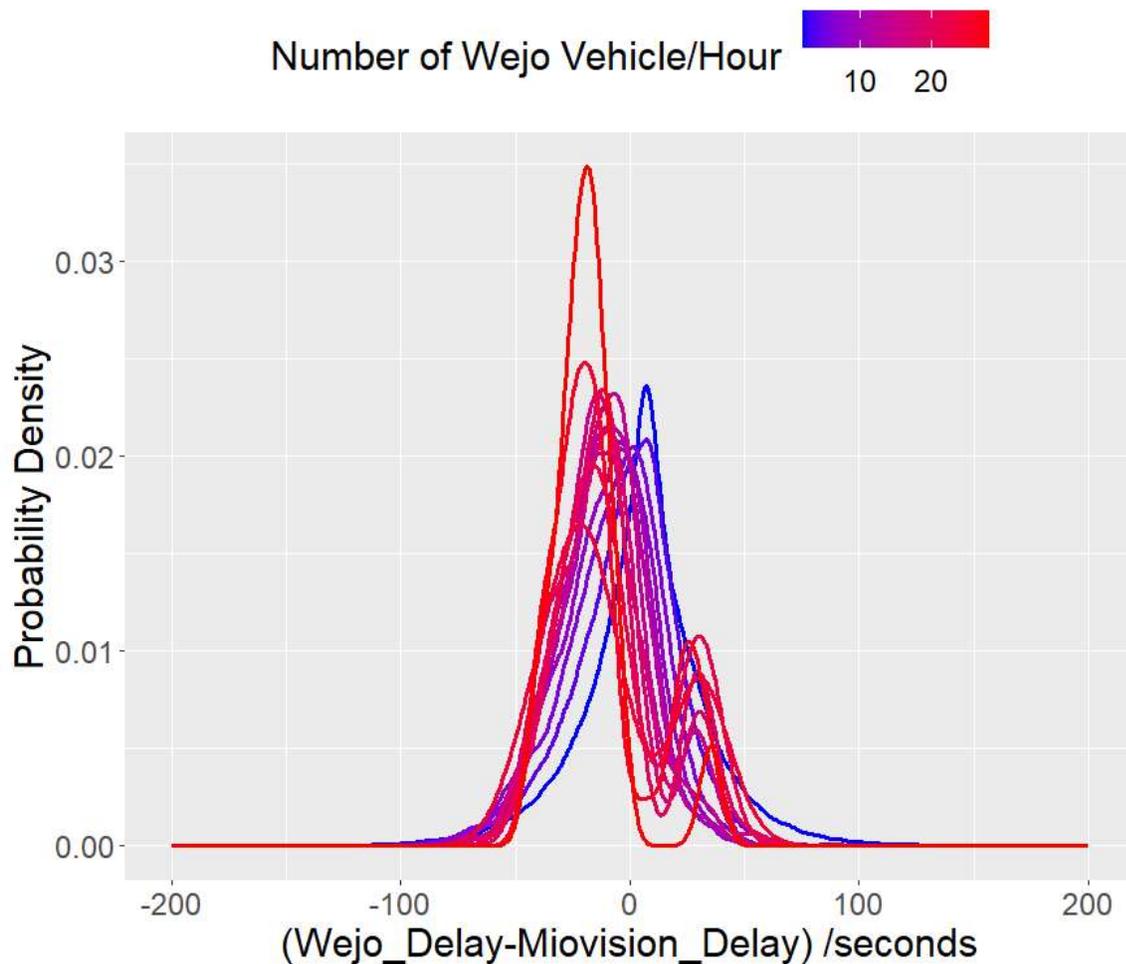


Figure 5-12. Plots of distributions of the error with various sample sizes (Left-turn traffic)

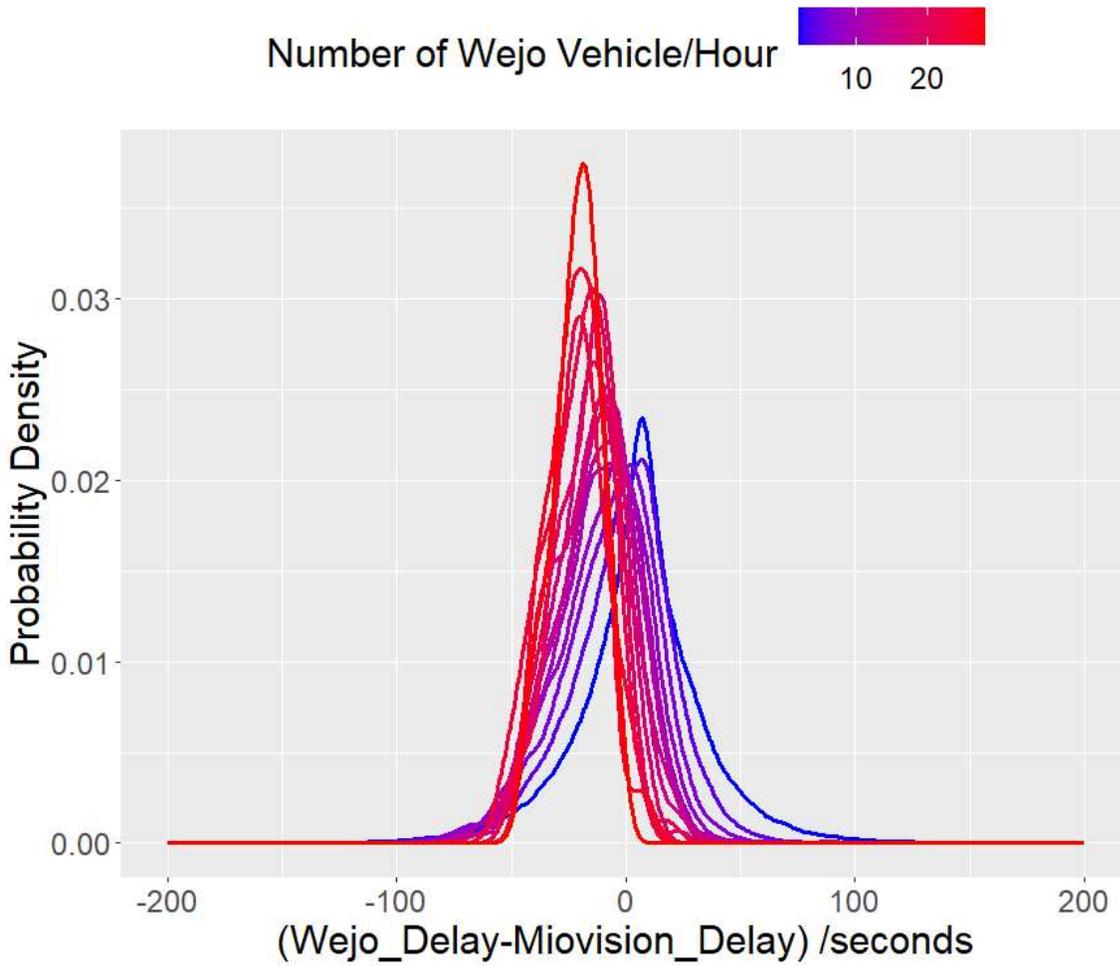


Figure 5-13. Plots of distributions of the error with various sample sizes after removing outlier intersections (Left-turn traffic)

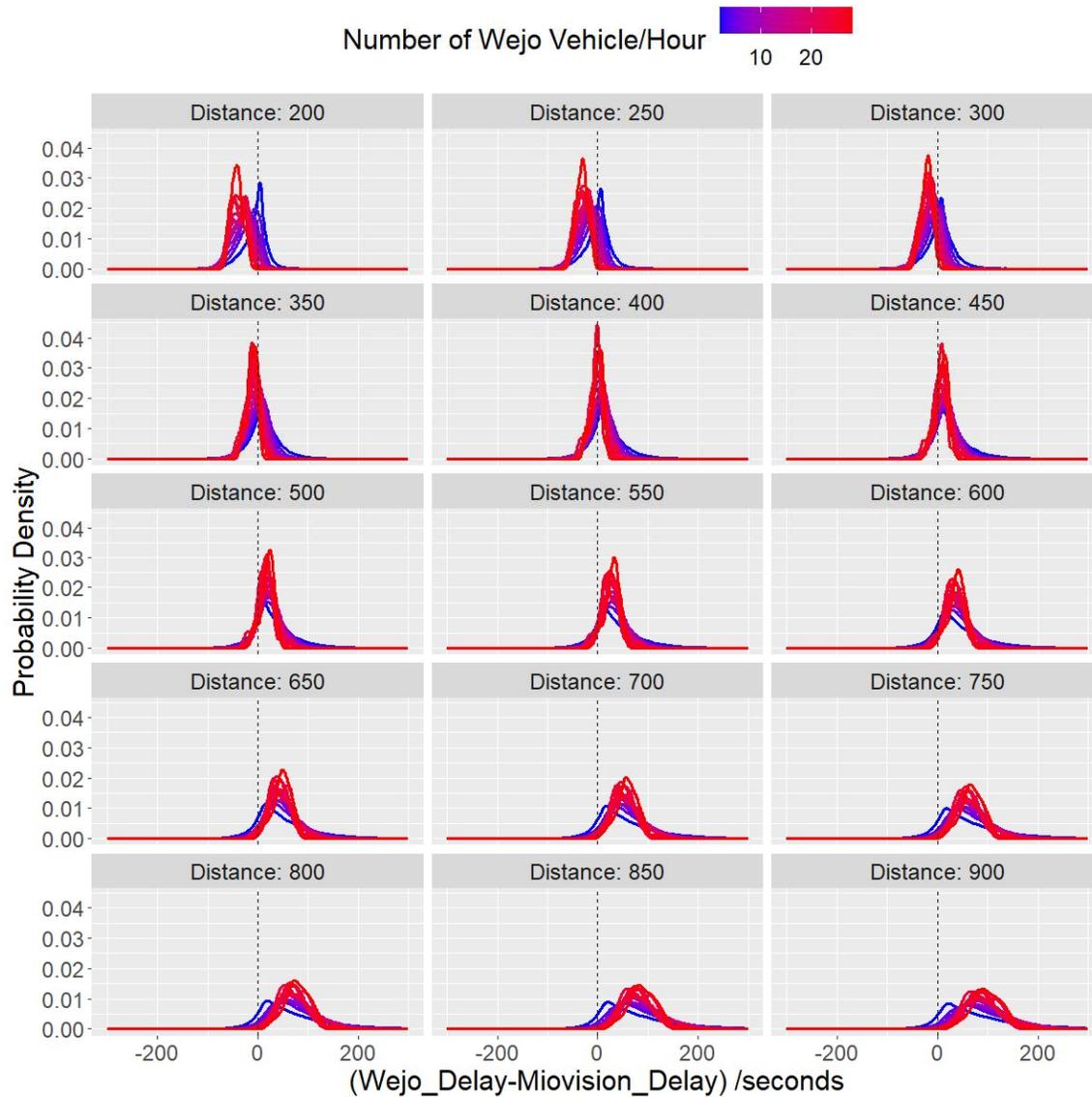


Figure 5-14. Sensitivity analysis by changing the distance. (Left-turn traffic)

Eq.1 and the left-turn data collected from all study intersections were used to calculate the correlation between the two types of delay. **Figure 5-15** shows the correlation coefficient is 0.63 between the two types of delay for left-turn traffic, indicating a strong positive correlation. Miovision-based delay has more data points with zero delay than Wejo-based delay because of missing Miovision data due to signal controller communication loss. **Figure 5-16** shows that the correlation coefficients between the two types of delay for all directions at all intersections range from 0.3 to 0.6. This inconsistent correlation could be due to the varying sample sizes of Wejo data at different locations. Two sample size indicators, the number of vehicles and the penetration rate, are applied to capture the impact of sample size on the correlation coefficient. **Figure 5-17(a)**

reveals that the correlation coefficient significantly increases as the number of vehicles increases. Once the number of vehicles exceeds 8 vph, the change of correlation coefficient with increasing sample size becomes insignificant. While using penetration rate, the change of correlation coefficient shows the same trend as using the number of vehicles, but with more outliers. Based on this comparison, the number of vehicles appears to be a more reasonable and reliable indicator for Wejo-based delay data quality control.

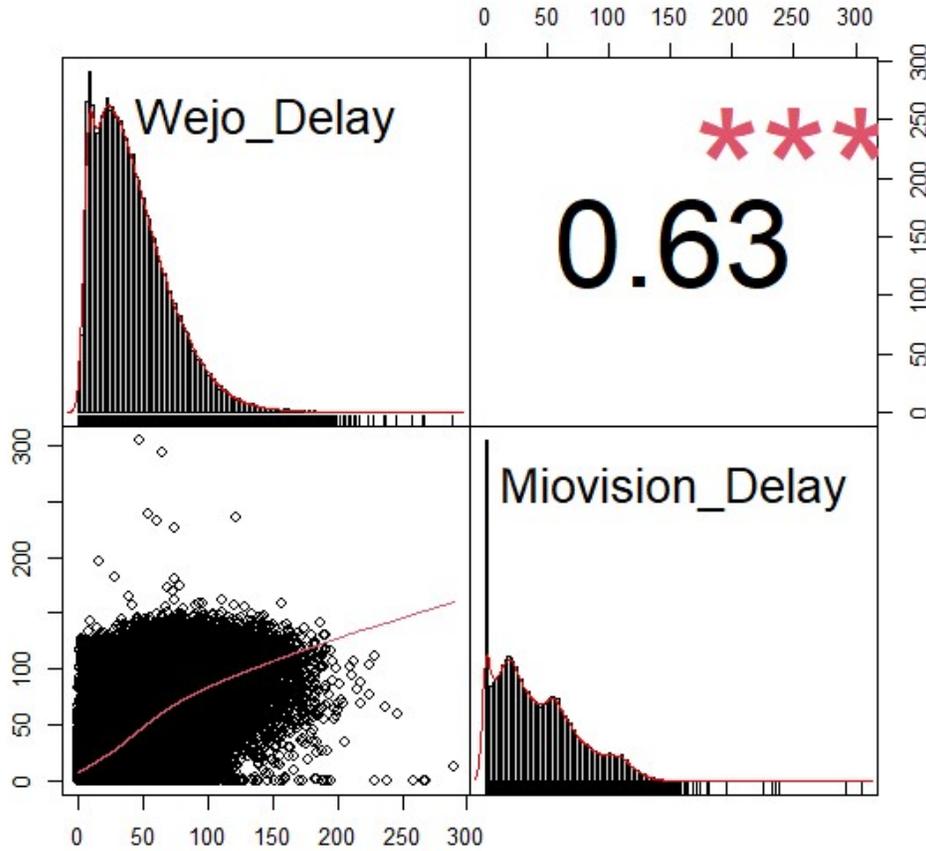


Figure 5-15. Correlation between Wejo- and Miovision-based delay for left-turn movement.

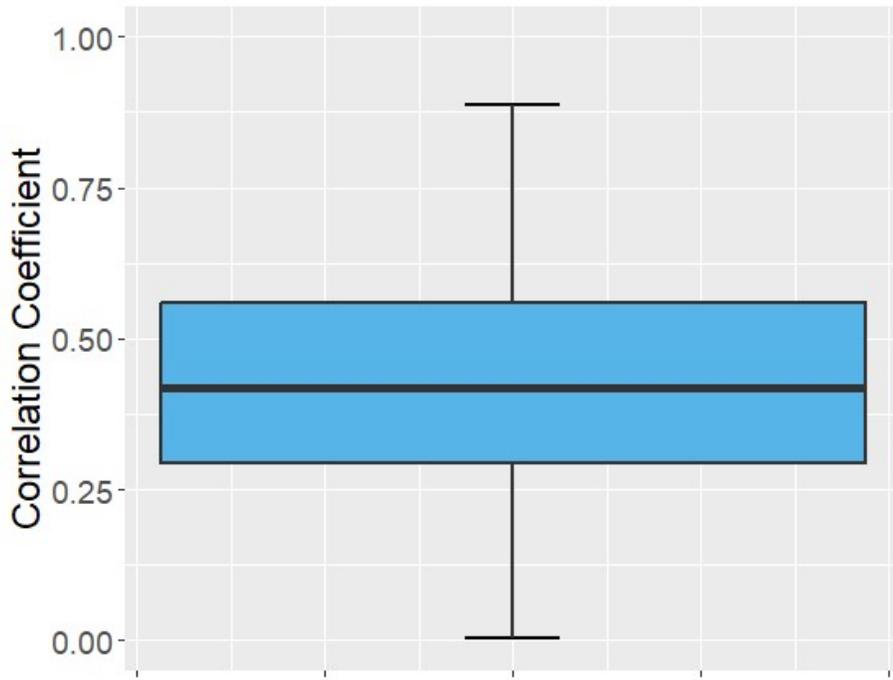


Figure 5-16. Box plot of correlation coefficients for all directions at all intersections (left-turn traffic)

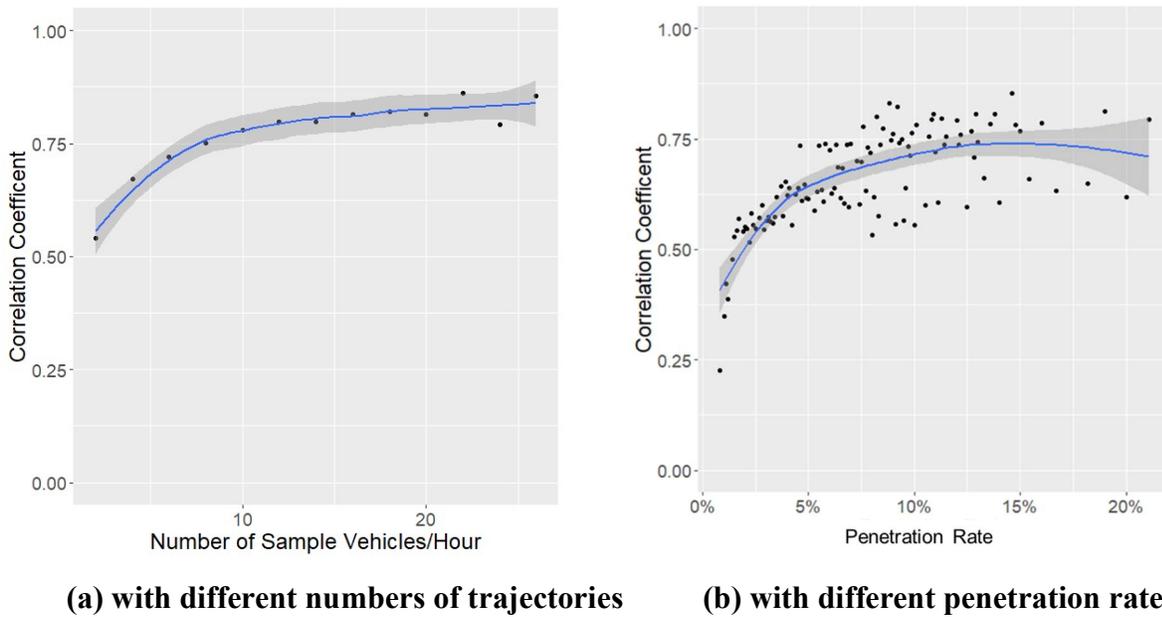


Figure 5-17. Correlation coefficients under different sample sizes

The relationship between the correlation coefficient and the two sample size indicators might be influenced by the traffic volume, which varies across different intersections and times. Therefore, the left-turn traffic data is categorized into four groups based on the traffic volume: less than 50 vehicles per hour (vph), 50-100 vph, 100-200 vph, and more than 200 vph, for correlation analysis. **Figure 5-18 (a)** shows the relationship between the correlation coefficient and penetration rate under various left-turn traffic volumes, and the results clearly demonstrate that the relationships vary with traffic volume. When the left-turn traffic volume is low, a higher penetration rate is required to achieve the same coefficient in comparison with the high-volume conditions. For example, when the left-turn volume is lower than 50 vph, the correlation coefficient is lower than 0.75 even though the penetration rate is 20%; however, left-turn traffic with a volume higher than 200 vph only requires a penetration rate of 5% to reach a coefficient of 0.75. **Figure 5-18 (b)** shows the relationship between the correlation coefficient and the number of sample vehicles under various left-turn traffic volumes, and the four relationships are similar. The correlation coefficient consistently increases as the number of sample vehicles per lane per hour increases and becomes stable when the number of sample vehicles per lane per hour is over 5 vehicles. Rather than using the total volume and number of all sample vehicles in the through moment analysis, the number of lanes has a significant impact on the correlation so the number of sample vehicles per lane is used to find the threshold. According to **Figure 5-18 (a)**, the left-turn traffic needs more than 5 sample vehicles per lane per hour to obtain the most accurate and reliable Wejo-based delay measurement.

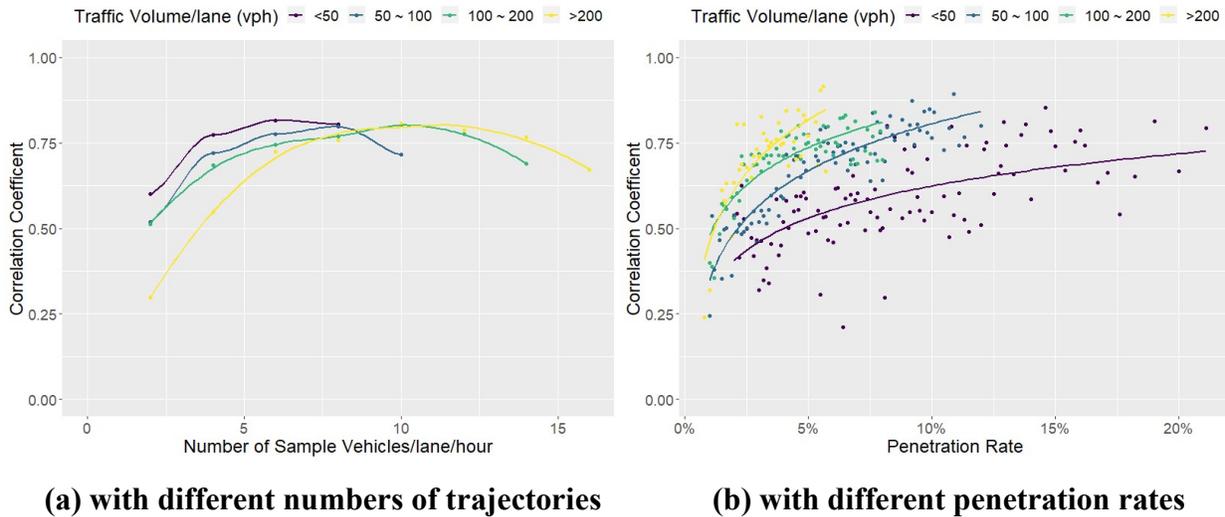


Figure 5-18. Correlation coefficients for various traffic volume levels

5.3 PREPRESENTATIVENESS EVALUATION OF ARRIVAL ON GREEN

In this section, the Wejo-based AoG is compared to the Miovision-based AoG to determine the indicator and threshold for ensuring the data quality of Wejo-based AOG. The Miovision-based AoG is the percentage of vehicles that arrive at the intersection using advance detectors during the green interval of a phase. The Wejo-based AoG ratio is calculated as the ratio between the number of vehicles passing an intersection without a stop and the total number of Wejo trajectories through that intersection during a defined interval.

We first use the data collected from one example intersection, Ina Rd. & La Cañada Dr., NB, to compare the temporal trend of these two types of AoG, as shown in **Figure 5-19**. The Miovision-based AoG can only be provided for through movements using advance detectors, which are not configured for left-turn lanes in the PAG region. **Figure 5-19** shows that the Wejo-based AoG has data gaps during nighttime due to the lack of Wejo trajectories. Moreover, Wejo-based AoG exhibits greater fluctuations than the Miovision-based AoG, possibly due to inconsistent sample sizes during different times. Despite these issues, both AoGs exhibit similar temporal trends during the daytime.

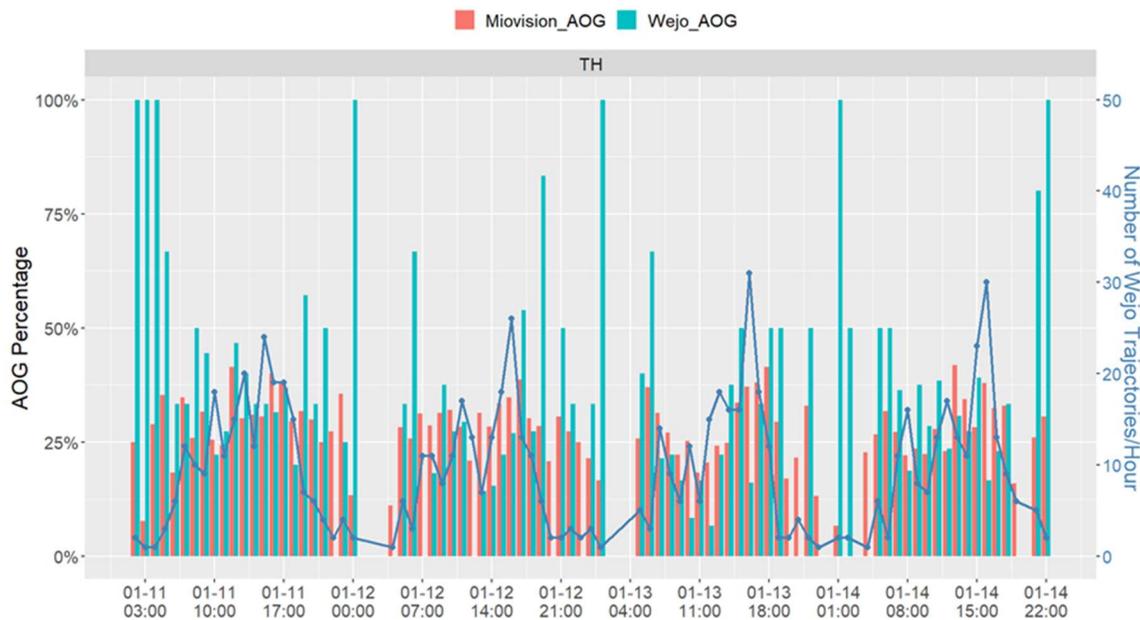


Figure 5-19. Comparison between the Miovision- and Wejo-based AoG at Ina Rd & La Canada Dr, NB.

Figure 5-20 displays the correlation between the Miovision and Wejo-based AOGs, with a correlation coefficient of 0.65 indicating a strong positive relationship. The regression fit line reveals that the Wejo-based AOG is slightly higher than the Miovision-based AOG, which could be due to the location of advance detectors. Inappropriately configured advance detectors can result in unreliable sensor-based AOGs, leading to overestimation issues. **Figure 5-21** presents a

box plot of the correlation coefficients from all study intersections, ranging from 0.25 to 0.55 and varying by location. The inconsistent correlation relationship is likely due to the different sample sizes of Wejo data at various locations.

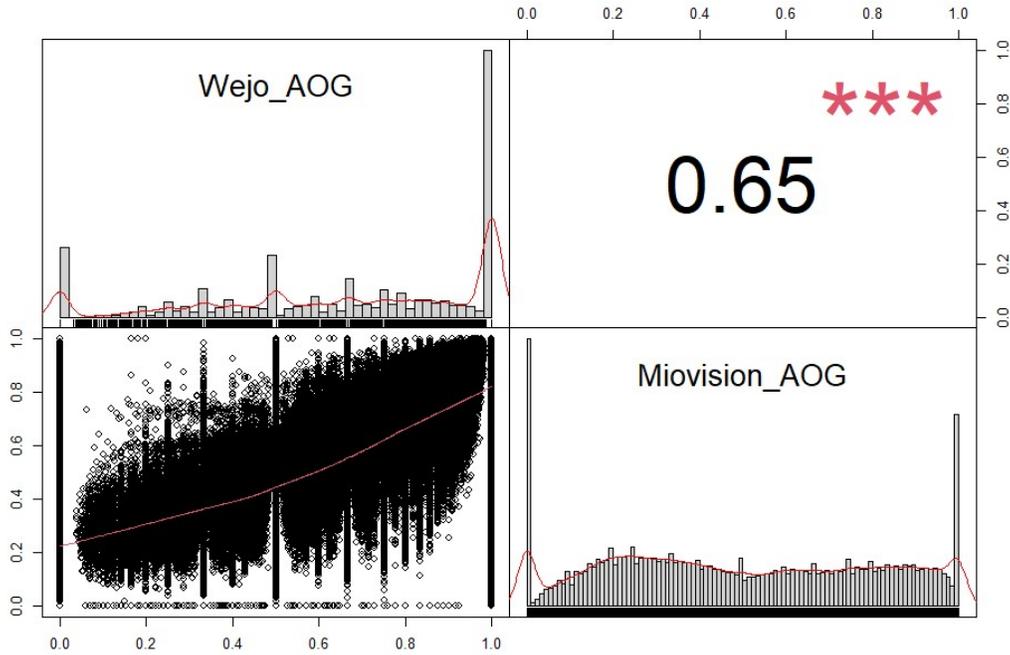


Figure 5-20. Correlation between Wejo- and Miovision-based AOG for through movement

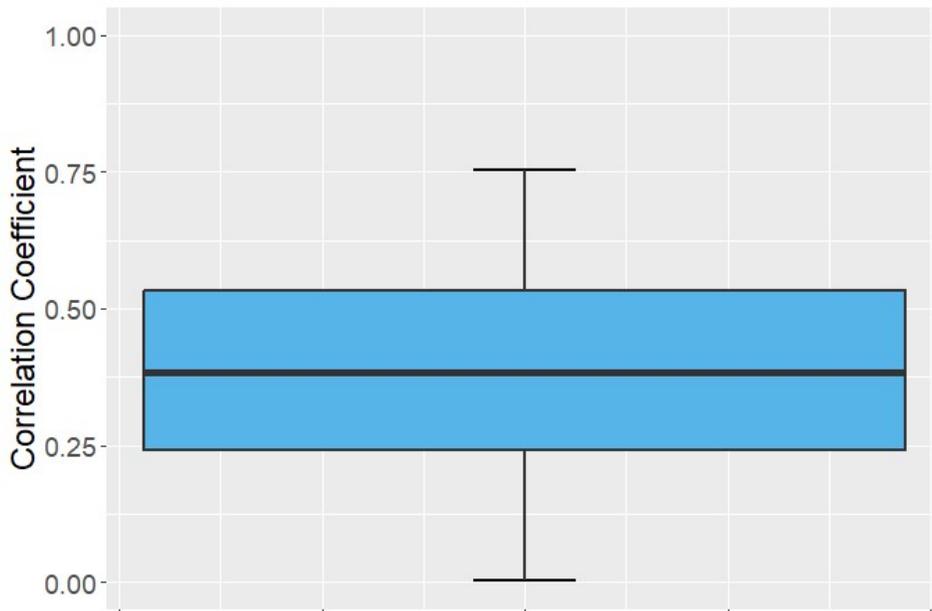


Figure 5-21. Box plot of correlation coefficients for AoG for all intersections and road directions

The correlation coefficient can be influenced by the sample size of Wejo data, which can be measured by two common indicators: the number of sample vehicles per hour and the penetration rate. **Figure 5-22(a)** illustrates the correlation coefficient under different numbers of sample vehicles. As the number of sample vehicles increases, the correlation coefficient also increases and becomes consistent regardless of the sample size once the number of vehicles surpasses 20 per hour. Therefore, the highest correlation coefficient, around 0.82, occurs when there are more than 20 sample vehicles per hour. **Figure 5-22(b)** displays the correlation coefficient under different penetration rates, calculated by dividing the number of sample vehicles by the total volume. The correlation coefficient initially increases with increasing penetration rate until reaching 5%. However, the correlation coefficient unexpectedly decreases as the penetration rate increases, which contradicts existing knowledge. One possible explanation is that low-volume locations can easily have a high penetration rate with only a few vehicles, but a few sample vehicles may not accurately reflect the traffic conditions, resulting in a low correlation coefficient. Moreover, more outliers are observed in **Figure 5-22(b)**, leading to unreliable results of the Wejo-based AOG, even with high penetration rates. These outliers and the decreasing trend with penetration rate could be caused by the traffic volume.

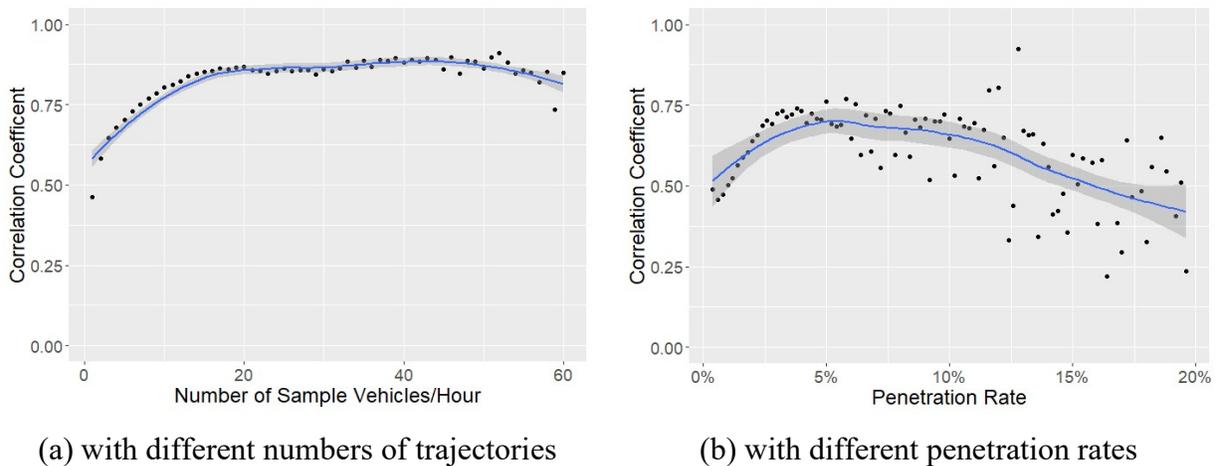


Figure 5-22. Correlation coefficients of AOG under different sample sizes

To analyze the correlation appropriately, all data is divided into four groups based on the total volume. **Figure 5-23** illustrates that all four groups display a similar trend for the correlation coefficient as the number of sample vehicles increases. The correlation coefficient of all scenarios increases with an increase in the number of sample vehicles and reaches the highest coefficient of around 0.82 when there are approximately 18-20 sample vehicles per hour. One exception is the scenario with the highest volume, >1000vph, and the correlation coefficient fit line does not include the increase trend because high-volume scenarios commonly have more sample vehicles, providing a high correlation coefficient. Thereafter, the correlation coefficient remains significantly unchanged regardless of the increase in the number of sample vehicles. Additionally,

the impact of changing the penetration rate on the correlation is analyzed and compared under different volumes, as shown in **Figure 5-23(b)**. According to **Figure 5-23(b)**, the relationship between correlation and penetration rate is not consistent, because different scenarios need different penetration rates to reach the same correlation coefficient. With an increase in traffic volume, a lower penetration rate is needed for the same correlation coefficient. When comparing the changes in the correlation coefficient with the number of sample vehicles and the penetration rate, using the number of sample vehicles is more reliable and convenient for controlling the data quality of Wejo-based AOG because it provides a consistent correlation coefficient regardless of location, time, and volume.

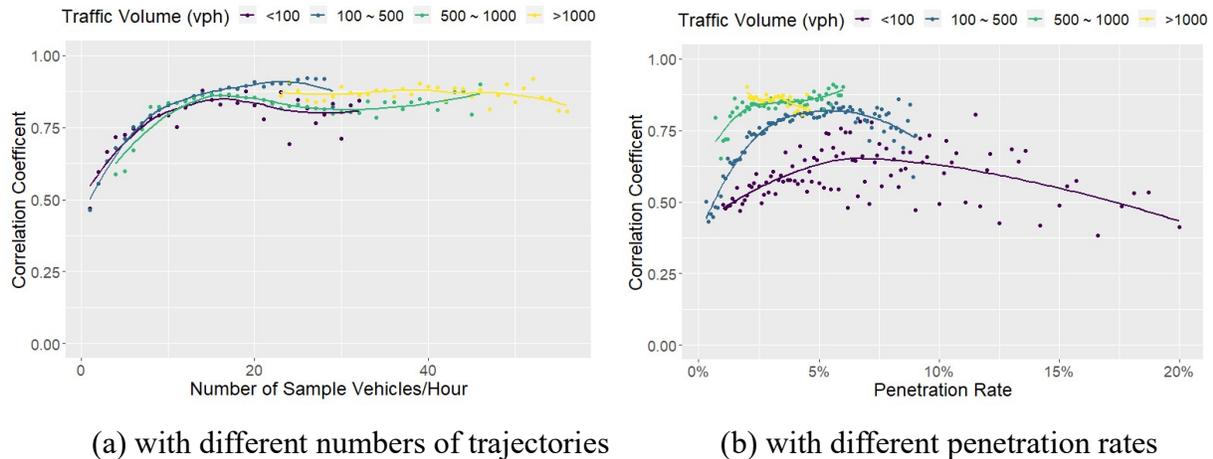


Figure 5-23. Correlation coefficients of AOG for various traffic volume levels

5.4 REPRESENTATIVENESS EVALUATION OF SPLIT FAILURE

In addition to control or stop delay and AOG, split failure is another important performance measure for quantifying traffic conditions, especially for identifying oversaturated conditions, because split failure occurs when a phase cannot serve all demand within one cycle.

Split failure is a performance measure to indicate when the traffic demand cannot be served within one cycle, and it is calculated using the green occupancy ratio (GOR) and red occupancy ratio (ROR) collected by presence detectors (Day et al., 2014, 2008). Movision sensors report split failure when both GOR and ROR5 are higher than 80% (Miovision, 2022b). E. Saldivar-Carranza et al. (E. Saldivar-Carranza et al., 2021a) has proposed a method to use GPS data only to calculate the split failure ratio. We first identify all sample vehicles that stop more than once before passing through an intersection during a defined time interval. These are the vehicles that are not served within one cycle. Then this number of sample vehicles that stop more than once is calculated as a percentage of the total number of sample vehicles that pass through that intersection during the time interval, and that percentage is the split failure ratio.

In this section, the Wejo-based split failure is compared with Miovision-based split failure to find the reliable indicator and threshold for controlling the data quality of Wejo-based split failure. **Figure 5-24** shows the comparison between Wejo- and Miovision-based split failure for through movements at Ina Rd. & Cholla Blvd. The data from both sources show that split failure occurs very rarely, with only one or two hours over three days showing split failure. In addition, as most of the data are zero, it is difficult to visually compare the two data sources or further investigate the relationship between Wejo- and Miovision-based split failure. **Figure 5-25** shows the correlation coefficient, which is almost zero, indicating a very weak or non-existent correlation. The correlation coefficient might vary with the location and be impacted by the Wejo data sample size. **Figure 5-26** displays the correlation coefficient by intersection and road direction, which ranges from 0 to 0.12. Furthermore, **Figure 5-27** shows that the correlation coefficient has an insignificant change, lower than 0.12, as the number of sample vehicles increases.

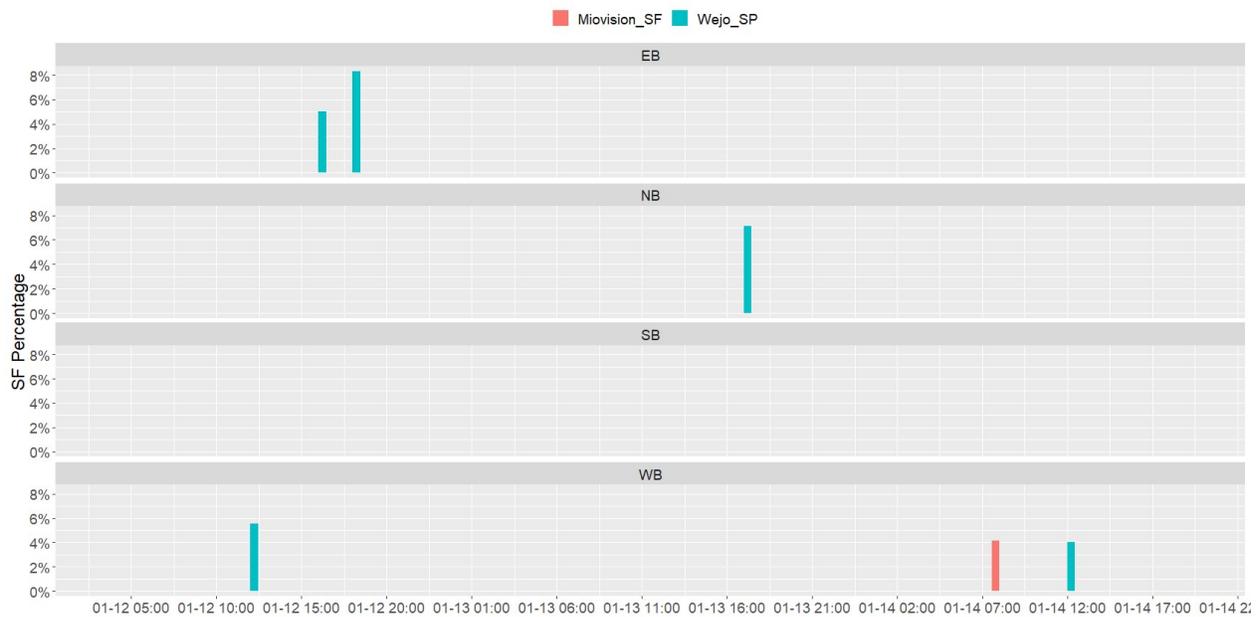


Figure 5-24. Comparison between Wejo- and Miovision-based split failure of through movement at Ina Rd. & La Cholla Blvd.

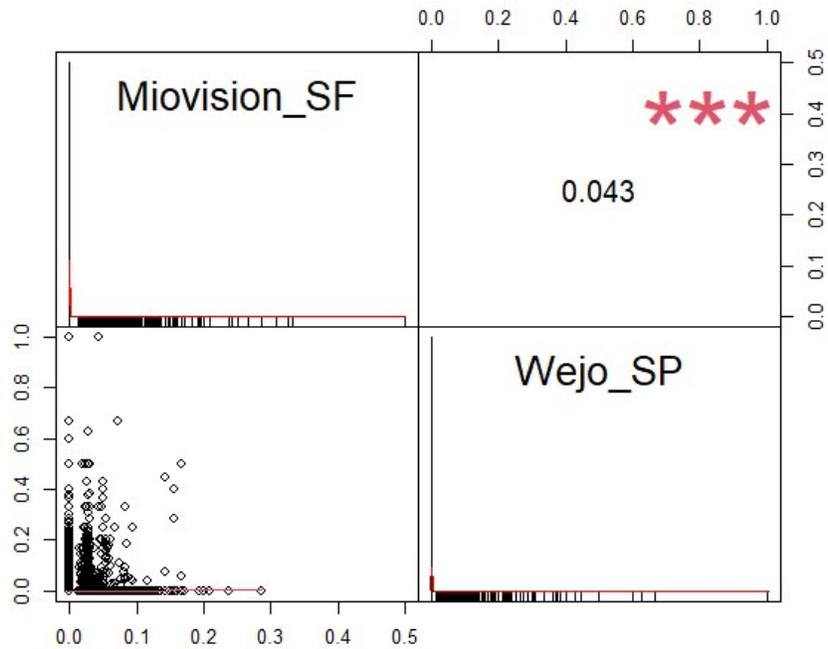


Figure 5-25. Correlation between Wejo-based and Miovision-based split failure for through movement

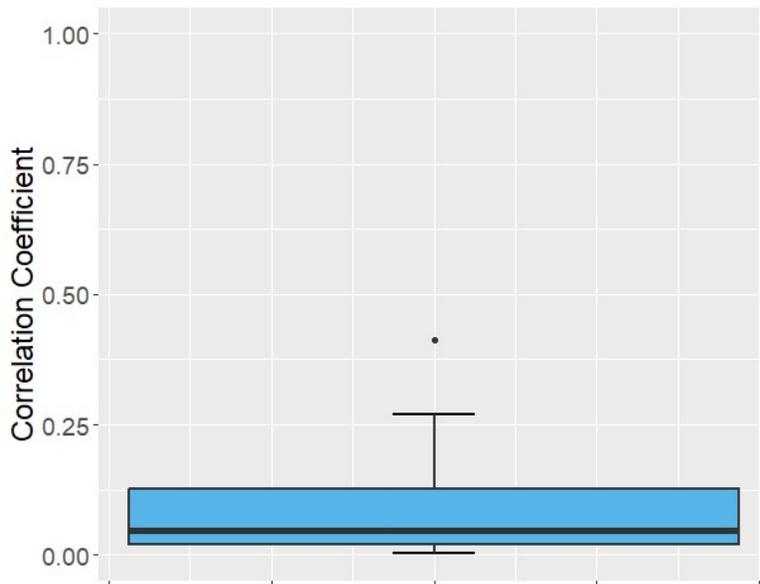


Figure 5-26. Box plot of correlation coefficients for split failure for all intersections and road directions

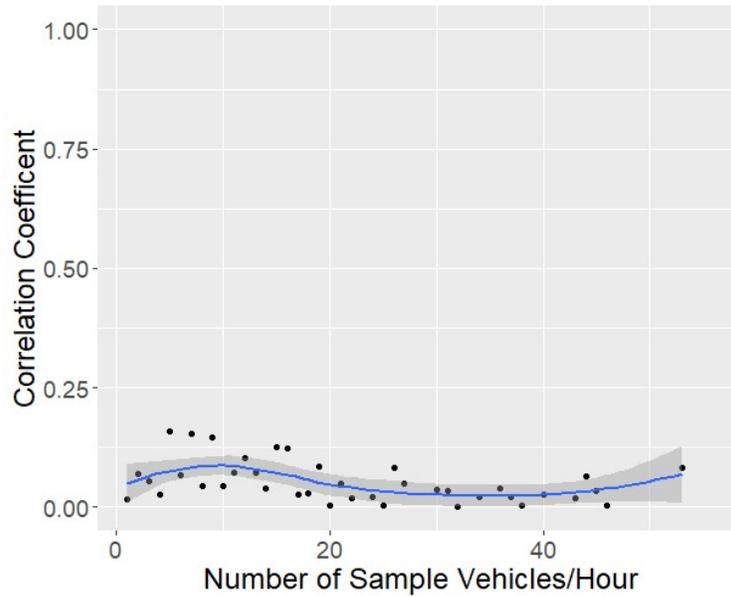


Figure 5-27. Correlation coefficients for split failure under different sample sizes

The low correlation coefficient between Wejo- and Miovision-based split failure is likely caused by so many zero values. According to the sensitivity analysis, increasing the sample size does not significantly improve the correlation coefficient because split failure rarely occurs. The Wejo-based split failure of zero could be because the sample size is not sufficient but split failure did occur or because there is no split failure in the actual situation. Due to so many zero values, accurately capturing the relationship between these two types of split failure is challenging, regardless of the sample size. Therefore, it is difficult to determine a threshold for ensuring the data quality of Wejo-based split failure.

5.5 REPRESENTATIVENESS EVALAUTON OF RELIABILITY PERFORMANCE

5.5.1 Delay Reliability

The delay reliability is calculated by location and movement using both Wejo and Miovision data, and the correlation between the delay reliability performance measures derived from both data sources is summarized in the box plots in **Figure 5-28**, which shows significantly higher correlation coefficients than those found in the delay comparison at the hourly level.

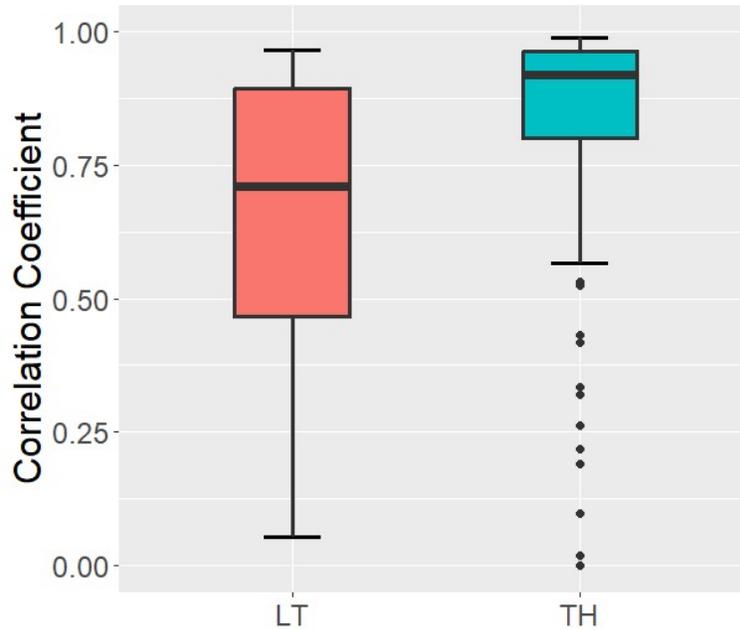
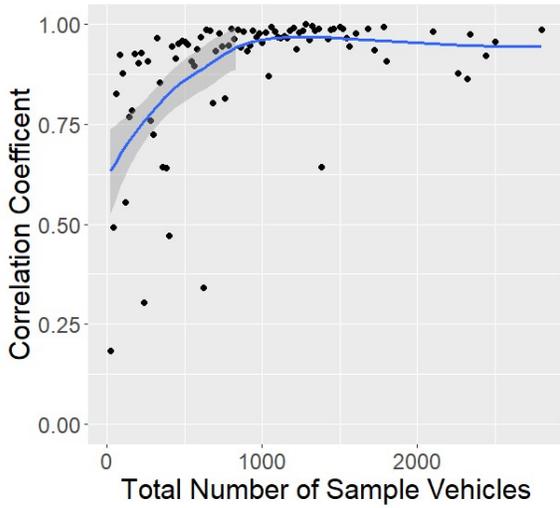


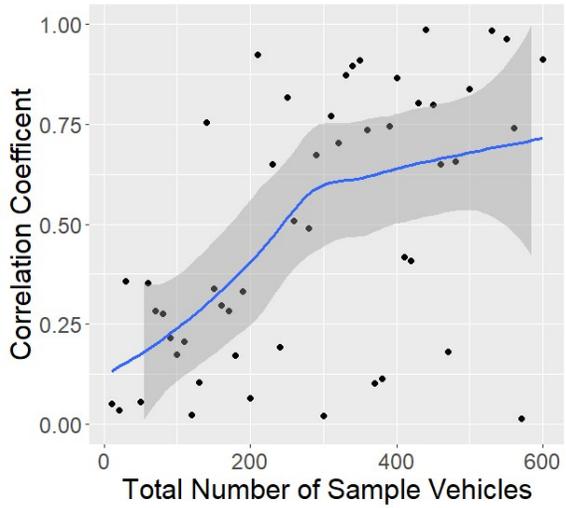
Figure 5-28. Box plot of correlation coefficients of 95th percentile delay for all intersections and road directions

Figure 5-29 shows the correlation coefficients as the total number of sample vehicles changes. As shown in **Figure 5-29(a)**, the correlation coefficient increases as the total number of sample vehicle increases until reaching a value of about 1 and then remains steady regardless of the increase in the total number of sample vehicles. When the total number of sample vehicles is higher than 500 vehicles per hour, the Wejo-based delay reliability performance is strongly correlated with Miovison-based reliability performance. And for the left-turn movement in **Figure 5-29(b)**, the correlation coefficient also increases as the total number of sample vehicles increases, but it does not show any clear relationship. The possible reason is the left-turn movement still needs 600 vehicles or more as the sufficient sample size.

Figure 5-30 shows the correlation coefficients as the penetration rate changes. As shown in **Figure 5-30(a)**, even though the correlation coefficient is around 1 when the penetration rate is higher than 1.5%, the relationship is not well defined, and the correlation coefficient even decreases slightly as the penetration rate continues to increase. For the left-turn movement, there is no clear relationship.

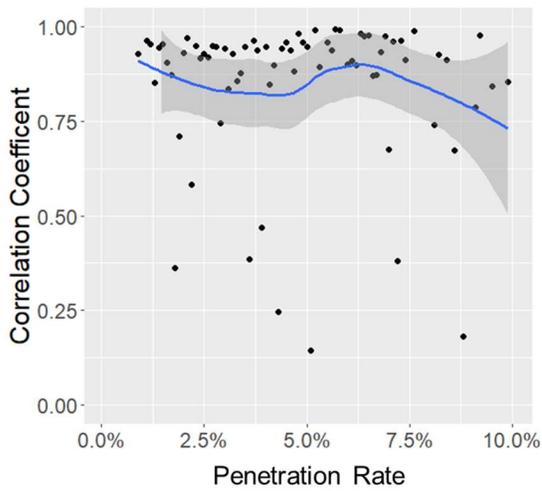


(a) Through movement

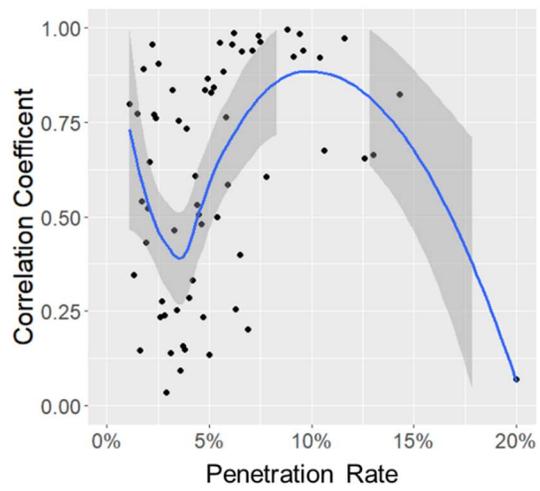


(b) left-turn movement

Figure 5-29. Correlation coefficients of 95th percentile delay under different sample sizes



(a) Through movement



(b) left-turn movement

Figure 5-30. Correlation coefficients of 95th percentile delay under different penetration rates

5.5.2 AoR Reliability

Figure 5-31 shows a box plot of the correlation coefficients between Miovision-based AoR and Wejo-based AoR, which are significantly higher than the AoG correlation coefficients calculated at the hourly level.

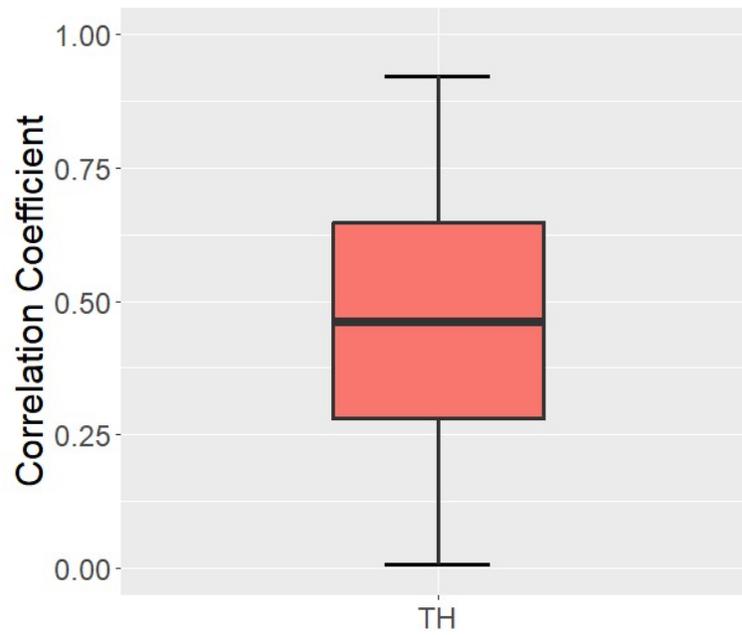


Figure 5-31. Box plot of correlation coefficients of 95th percentile AoR for all intersections and road directions

As shown in **Figure 5-32**, the correlation coefficient increases as the total number of sample vehicles increases until reaching about 1 at around 1000 vehicles per hour. Then, the coefficient remains steady regardless of the increase in the number of sample vehicles. As shown in **Figure 5-33**, the relationship between correlation coefficient and penetration rate is similar to the relationship with the total number of sample vehicles but with more outliers.

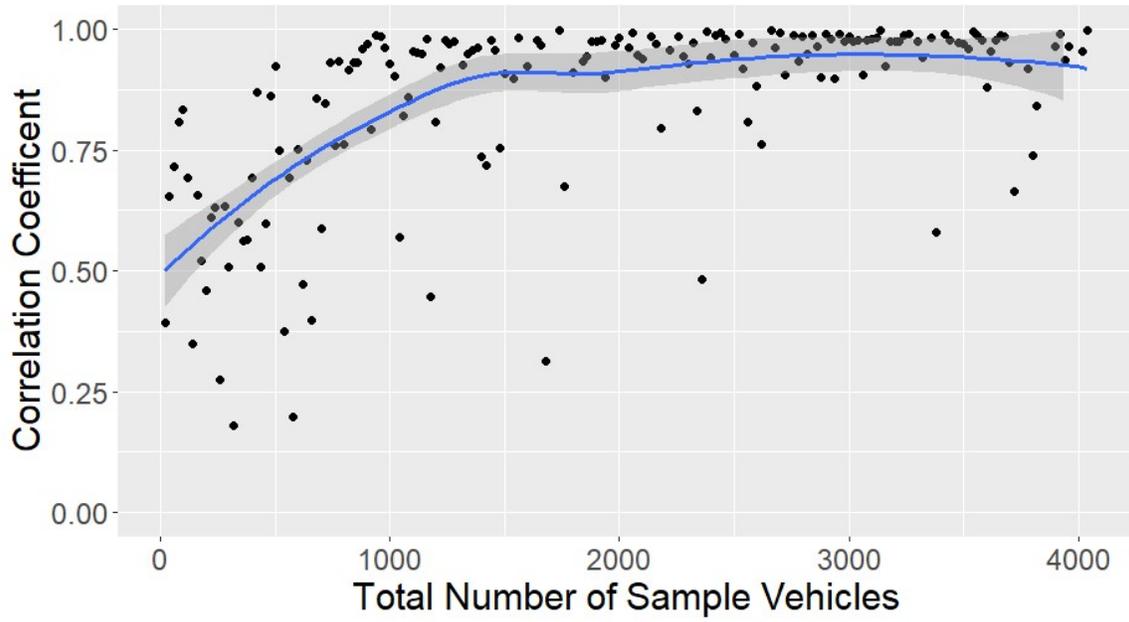


Figure 5-32. Correlation coefficients of 95th percentile AoR under different sample sizes

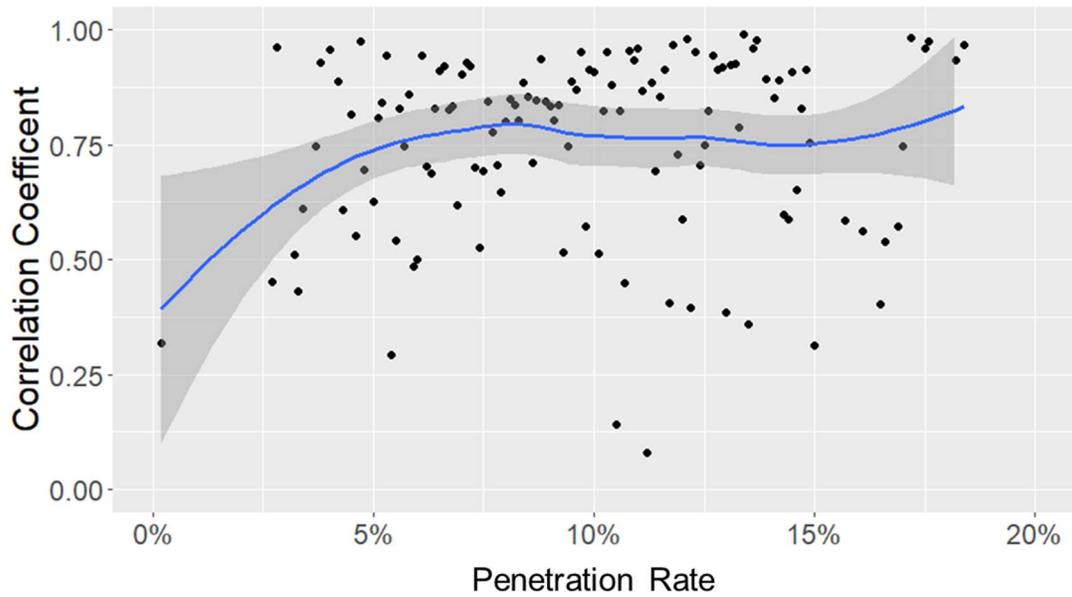


Figure 5-33. Correlation coefficients of 95th percentile AoR under different penetration rates

5.5.3 Split Failure Reliability

Figure 5-34 shows the correlation coefficients when comparing Wejo-based split failure and Miovision-based split failure, which are significantly higher than the correlation coefficients calculated at the hourly level for both the through and left-turn movements. Figures 5-35 and 5-36 show the correlation coefficient as the sample size changes and as the penetration rate changes, respectively, and no clear trend is observed.

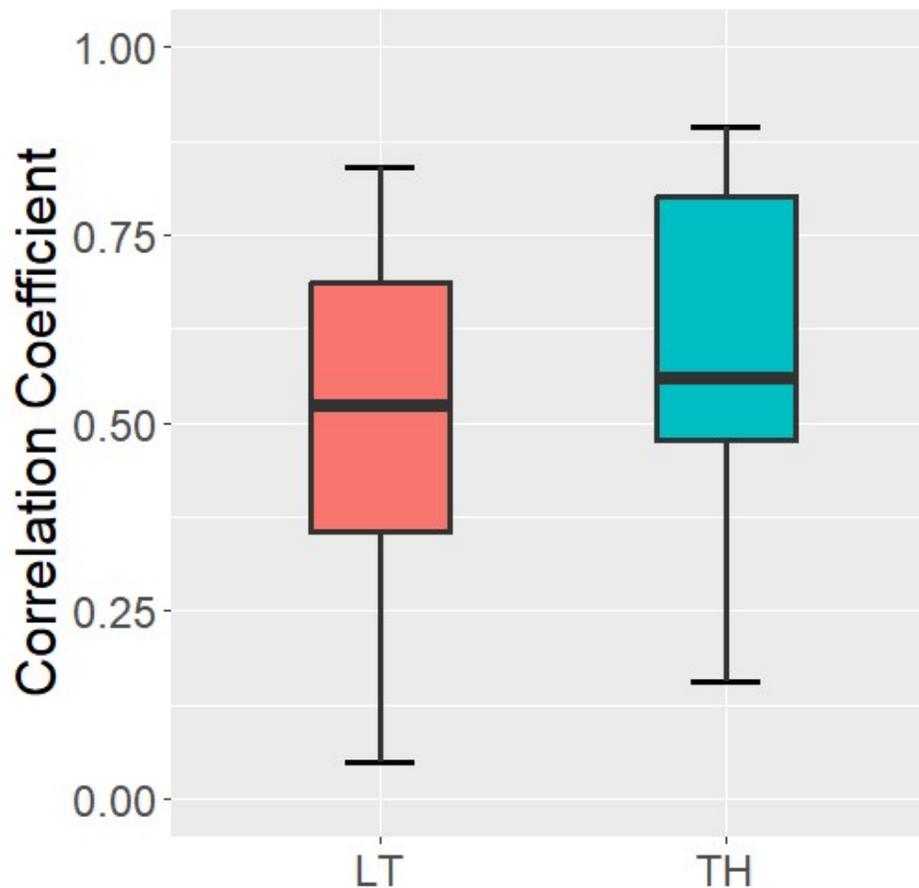


Figure 5-34. Box plot of correlation coefficients of 95th percentile split failure for all intersections and road directions

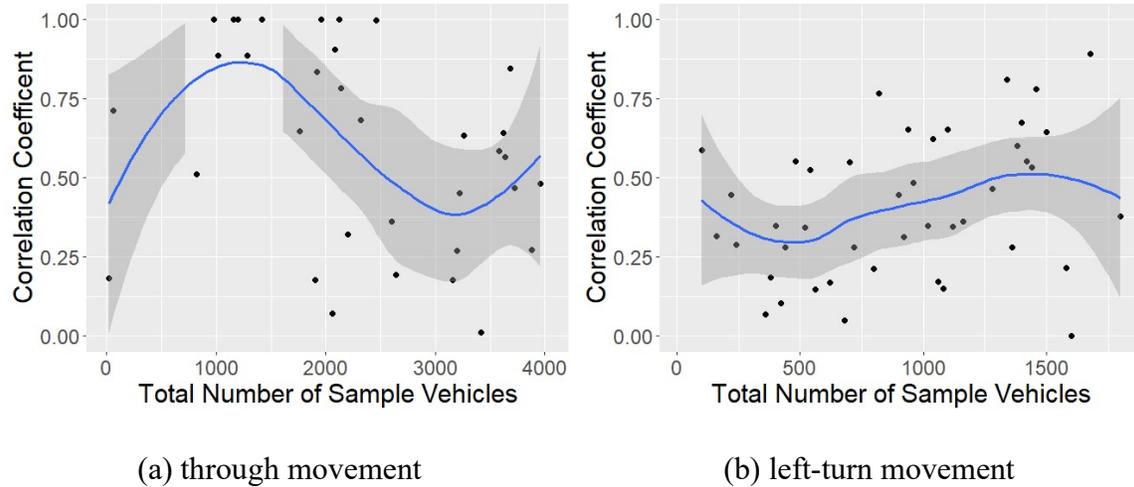


Figure 5-35. Correlation coefficients of 95th percentile split failure under different sample sizes

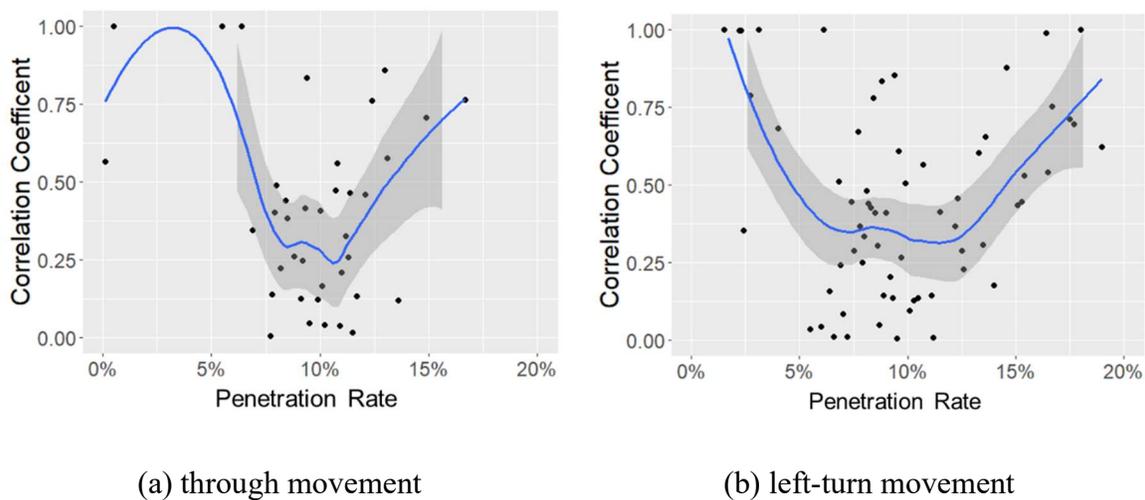


Figure 5-36. Correlation coefficients of 95th percentile split failure under different penetration rates

5.6 DATA QUALITY AND AVAILABILITY ANALYSIS

Based on the analyses above, it can be concluded that using the number of sample vehicles is a more reliable and convenient method for selecting the threshold to ensure the data quality of Wejo-based performance measures. **Table 5-1** summarizes the threshold values for the three types of performance measures. These values can be used to examine and summarize the data quality and availability of Wejo-based performance measures in the PAG region.

Table 5-1. Threshold for different performance measures.

	Movement	
	Through	Left Turn
Wejo-based Delay	16 vph	6 vph/lane
Wejo-based AOG	16 vph	-
Wejo-based Split Failure	-	-

The raw Wejo data collected from 560 major signalized intersections in the PAG region, as shown in **Figure 5-37**, are processed to calculate the number of sample vehicles. According to the threshold, the number of sample vehicles is categorized into three groups, no sample vehicles (0), insufficient sample vehicles (1-15), and sufficient sample vehicles (≥ 16). For the through movement, at least 16 vehicles per hour are appropriate for estimating Wejo-based performance measures, so fewer than 16 vehicles are considered insufficient. **Figure 5-38** shows the temporal distribution of sample size for the through movement. From midnight to 6 am, the percentage of samples greater than 16 vph is very low, lower than 2%, and most sample sizes are zero. After 7 a.m., around 20% of the samples meet the threshold, but around 60% are still lower than 16 vph. From noon to 6 p.m., more than 30% of the samples meet the threshold. After 8 p.m., the percentage of samples that meet the threshold becomes lower.

Although the temporal distribution shows that more samples meet the threshold during the daytime, it does not follow the traffic volume trend. Traffic volume in the PAG region usually has two peak periods: the AM and PM peak periods. However, the percentage of sufficient sample sizes gradually increases from the early morning until noon, with a peak period during the afternoon, as shown in **Figure 5-38**. In addition, **Figure 5-39** shows the temporal distribution of sample size for the left-turn movement. Similar to the through movement, the majority of the Wejo data collected lacks samples large enough to estimate performance measures of left-turn movement. Starting at 7 a.m., the percentage of sufficient sample sizes gradually increases and reaches approximately 5%. The afternoon shows the highest percentage of sufficient sample size, but it is still below 10%.

Based on the temporal analysis, the percentage of Wejo data with sufficient sample size for both through and left-turn movements varies with the hour of the day, with the highest percentage during the afternoon. However, the left-turn movement has a much lower percentage of data with sufficient sample size than does the through movement, indicating that accurate Wejo-based performance measures for the left-turn movement may be more challenging to obtain.

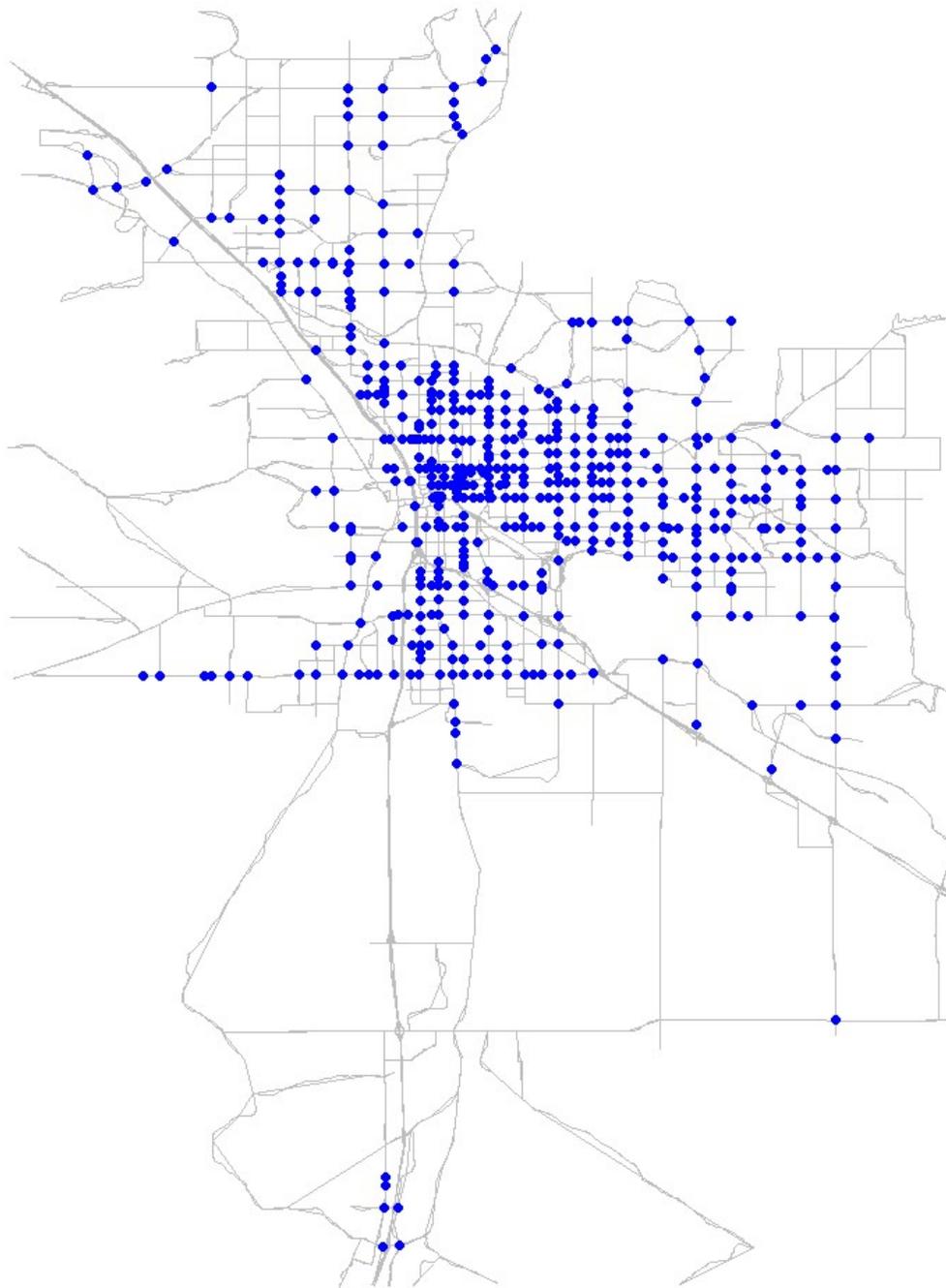


Figure 5-37. Study locations

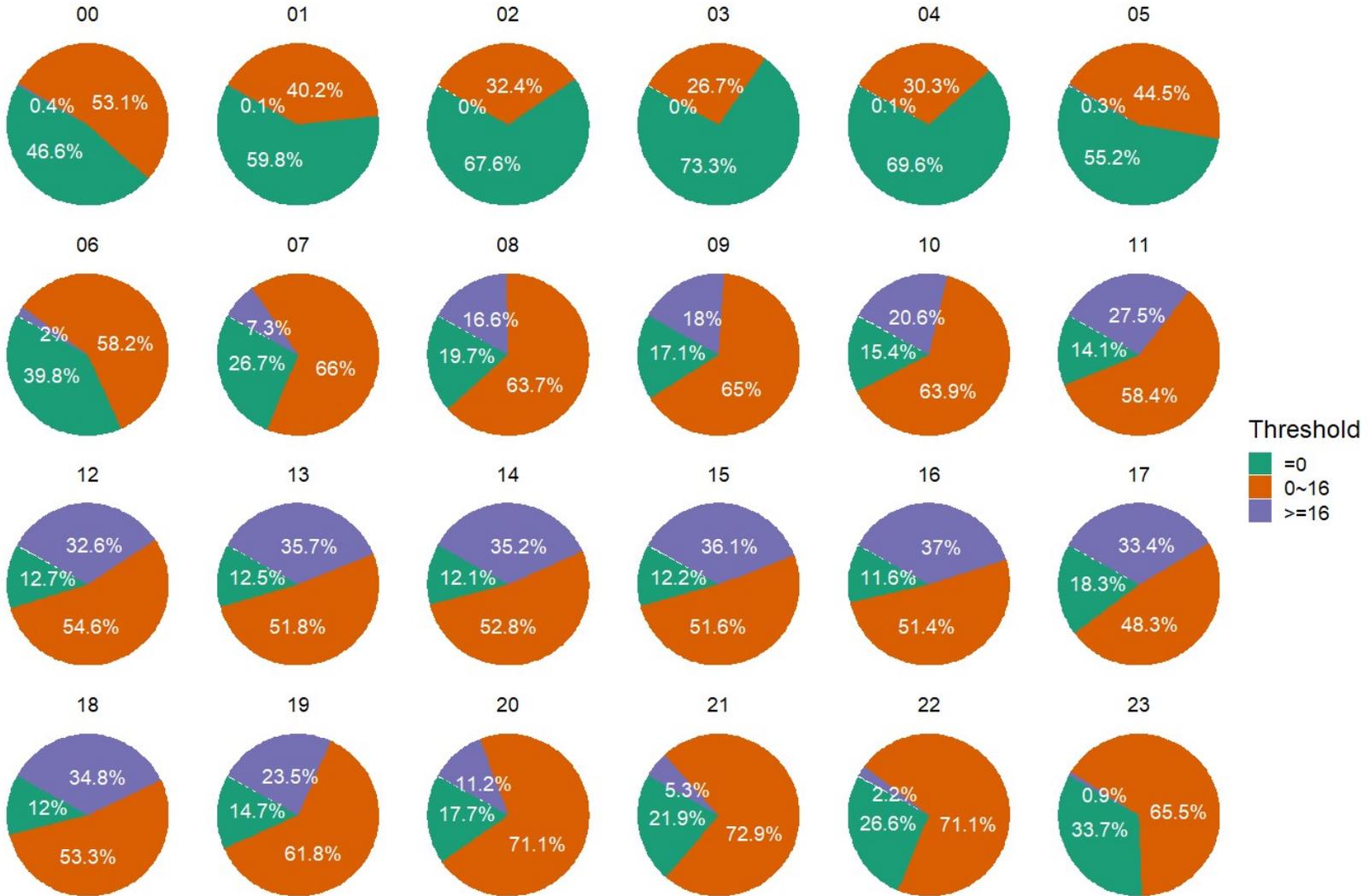


Figure 5-38 Different threshold distributions of Wejo data for through movement by the hour of the day

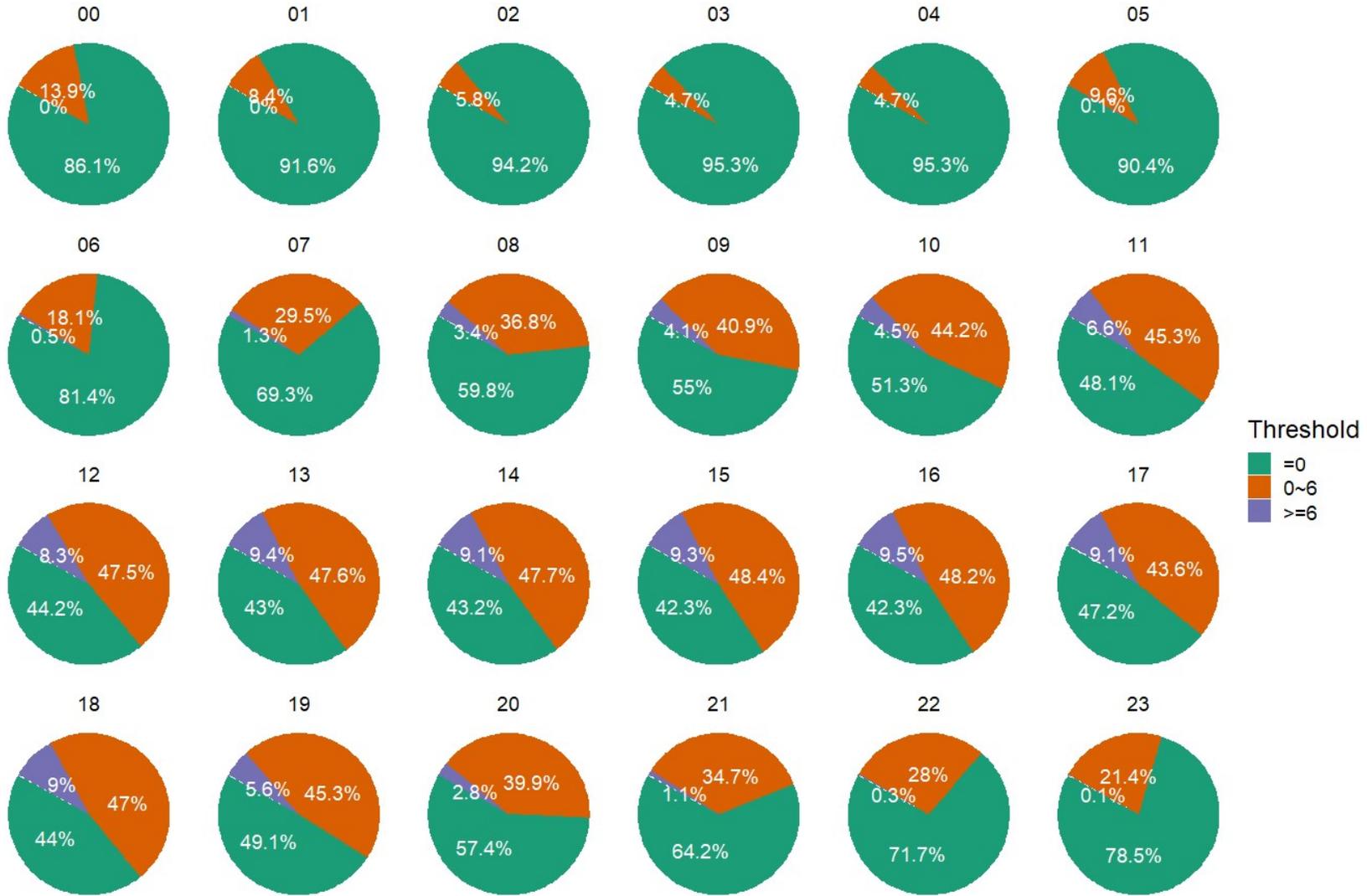


Figure 5-39. Different threshold distributions of Wejo data for left-turn movement by the hour of the day

In addition to analyzing the temporal trend of Wejo data sample size, a spatial distribution analysis is conducted to investigate the spatial variability and attempt to identify any pattern or trend. **Figure 5-40** displays the spatial distribution of different thresholds for the through movement during the day time. High-volume major corridors such as Valencia Rd., Kolb Rd., Speedway Blvd., and Wilmot Rd frequently have more than 50% of sample with sufficient size for the through movement. Although data for most intersections falls below the threshold, they still have Wejo data. The intersections with insufficient sample size are minor intersections located in suburban areas with relatively low volume throughout the day.

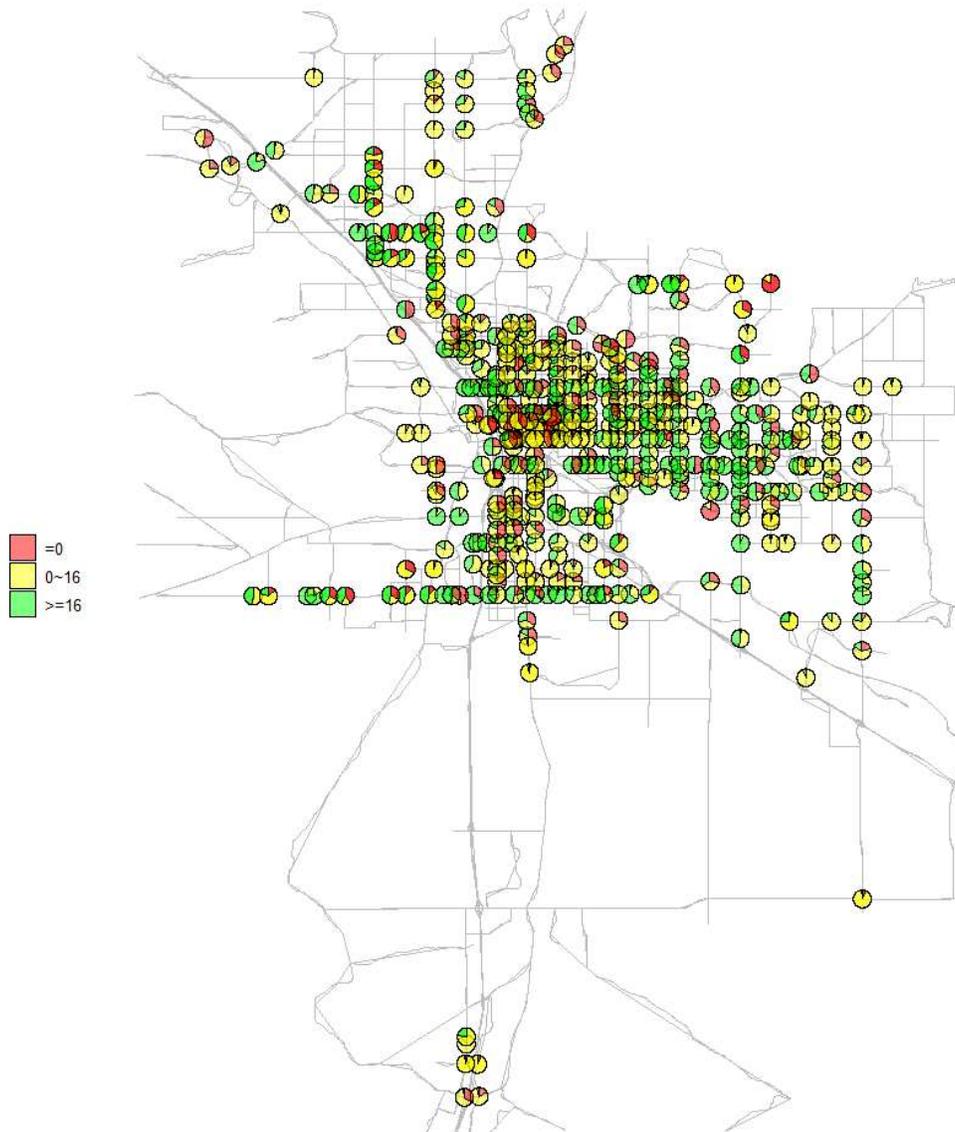


Figure 5-40. Spatial distribution of Wejo data sample sizes for through movement (7 a.m.-8 p.m.)

Figure 5-41 shows the spatial distribution of different sample sizes for the left-turn movement. Most intersections, even those on major corridors, have a very low percentage of Wejo data samples with sufficient size, and even percentages as low as zero, especially at some minor intersections in the suburban area. In addition, most intersections have 50% or more of the Wejo data missing for the left turn movement, which is a very significant amount of missing data for the left-turn movement.

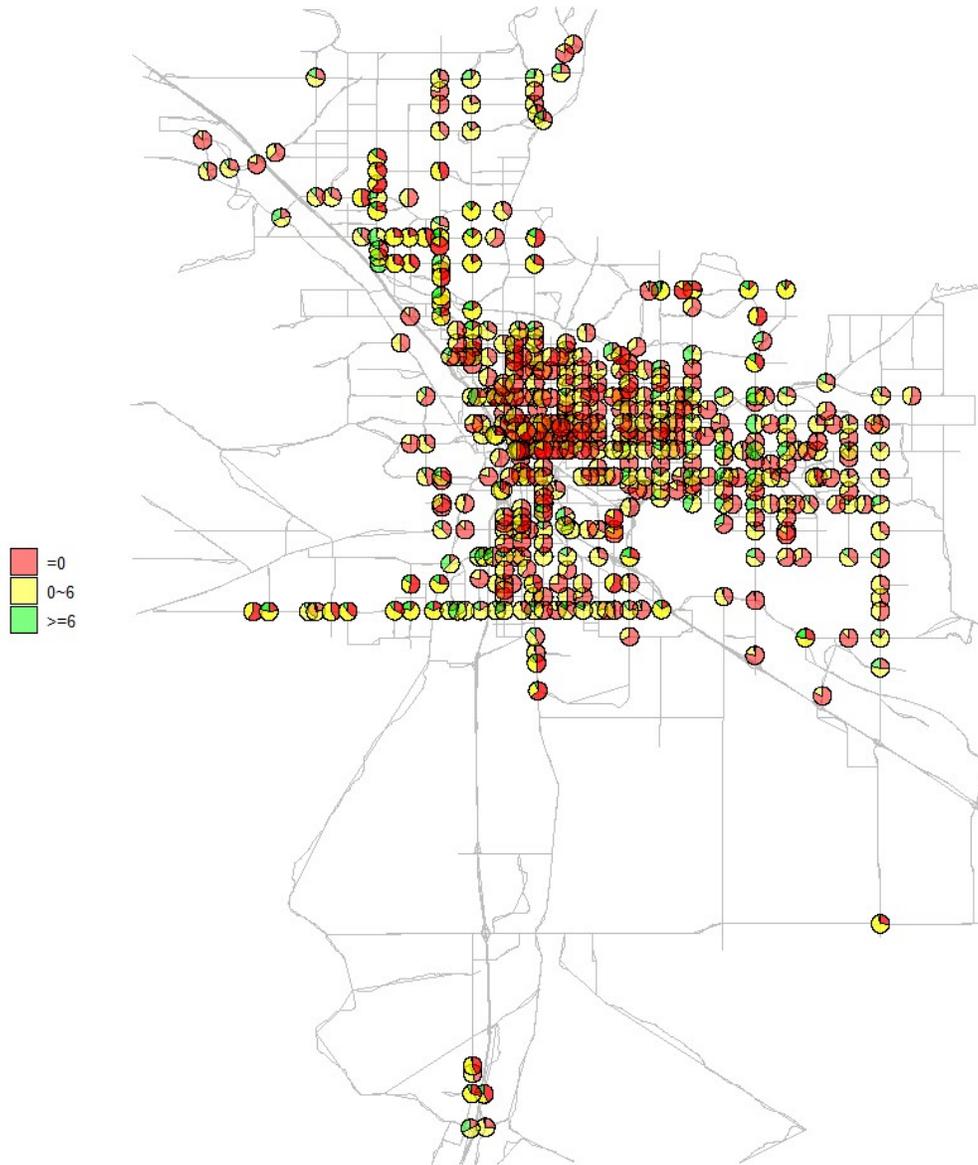


Figure 5-41. Spatial distribution of Wejo data sample sizes for left-turn movement (7 a.m.-8 p.m.)

CHAPTER 6: MOBILITY/RELIABILITY PERFORMANCE ESTIMATION USING MAXVIEW AND MIOVISION SYSTEM

The Wejo data with a sufficient sample size is able to provide reliable and accurate traffic performance measures, including control delay and arrival-on-green (AOG) for indicating traffic conditions at signalized intersections, as presented on Chapter 5. However, the data availability analysis conducted in Chapter 5 reveals that only a limited number of locations and time periods possess sufficient Wejo data. Consequently, the majority of locations lack Wejo-based performance measures, posing challenges for transportation agencies seeking consistent monitoring and management of traffic throughout the region when using crowd sourced data as a monitoring or management tool.

In the PAG region, an alternative data source known as signal event-based data exists, which has extensive coverage and is readily available. This event-based data is collected by the existing traffic sensors and has been archived into a database or online cloud without additional cost and time for sensor installation and data collection. Therefore, using this existing event-based data to estimate mobility is cost effective and saves time. Moreover, the estimated performance measures derived from this data can offer comprehensive insights into regional traffic conditions on a 24/7 basis, benefiting from the consistent availability and wide coverage of the event-based data. Consequently, the primary objective of this task is to develop a methodology for estimating traffic performance measures, including control delay and AOG, using the available event-based data.

6.1 DATA DESCRIPTION AND STUDY LOCATIONS

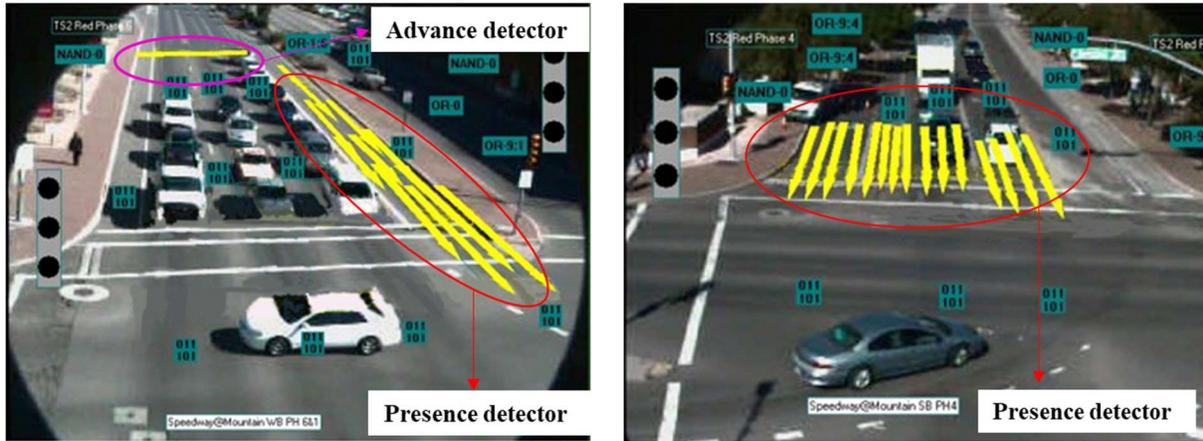
6.1.1 Event-based Data Collected by MaxView System

The MaxView system is an advanced traffic management system (ATMS) that utilizes various events generated by signal assets, such as detectors, signal heads, and pedestrian push buttons, to control traffic signals. Among these events, signal status and detection events are particularly crucial for assessing traffic performance. Figure 6-1 provides an illustration of the typical detector configuration within the MaxView system.

- The major roads feature an advance detector with a bar-shaped configuration specifically designed to cover all through lanes, as highlighted in **Figure 6-1(a)**.
- The major roads feature a presence detector with a long arrow-shaped configuration specifically designed to cover left-turn lanes, as highlighted in **Figure 6-1(a)**.
- The minor roads feature a presence detector with a long arrow-shaped configuration specifically designed to cover through lanes, as highlighted in **Figure 6-1(b)**.
- The minor roads feature a presence detector with a long arrow-shaped configuration specifically designed to cover left-turn lanes, as highlighted in **Figure 6-1(b)**.

The detection event consists of two types of events, the detector-on events when vehicles arrive at the detector and the detector-off events when vehicles exit the detector, and the associated

timestamps. The signal events consist of a series of events indicating the signal status changing such as green light on and red light on.



(a) Major road

(b) Minor road

Figure 6-1. Typical detector configuration at signalized intersections in MaxView system

6.1.2 Event-based Data Collected by Miovision System

Similarly, the Miovision system is also capable of collecting event-based data using detectors. Detector configurations in Miovision sensors are different from that managed by the MaxView system. **Figure 6-2** shows the typical detector configuration in Miovision sensors.

- The through movement in four directions has both presence and advance detectors configured to cover multiple through lanes. The presence detectors are long loops and advance detectors are short loops.
- The left-turn movement in four directions only has presence detectors configured, and one detector covers multiple left-turn lanes.

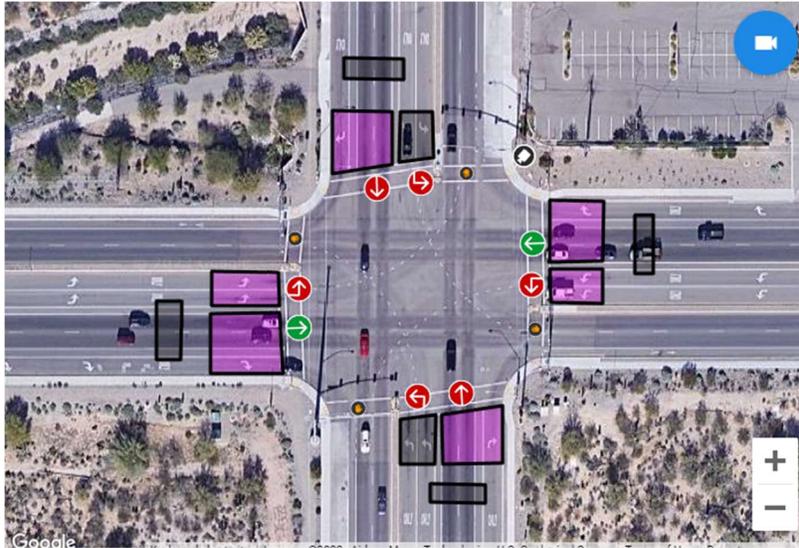


Figure 6-2. Typical detector configuration at signalized intersections in Miovision system

Compared to the MaxView system's detector layout, Miovision sensors have presence and advance detectors configured on through movements across all road directions. Another distinction lies in the storage of event-based data, where Miovision stores the data in an online cloud server, while the MaxView system stores the data in a local database server. **Figure 6-3** provides an illustration of the event-based data collection processes employed by these two systems.

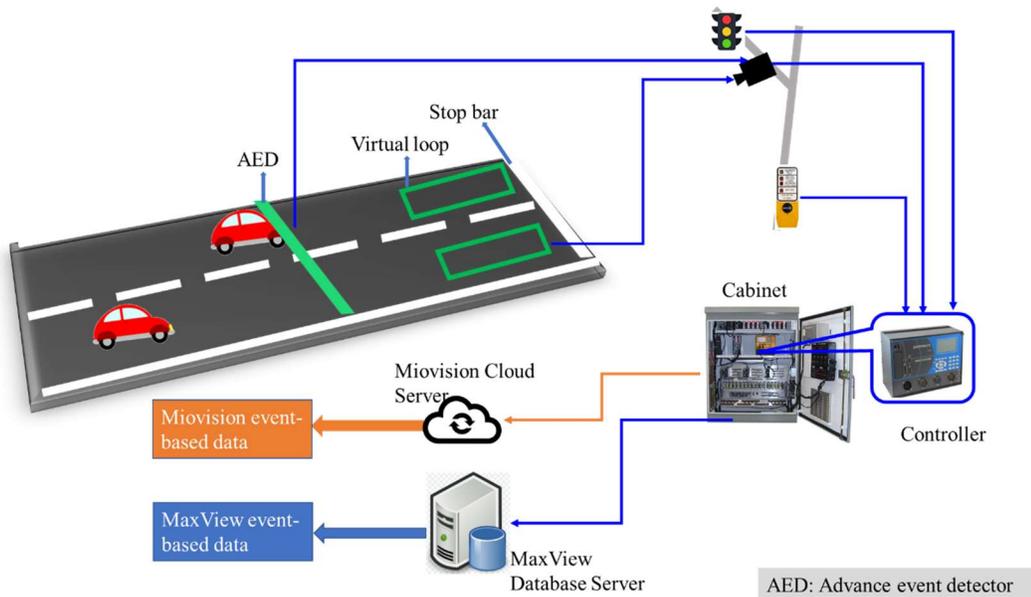


Figure 6-3. Event-based data collection process in MaxView and Miovision systems

6.2 METHODOLOGY

6.2.1 Data Processing for Input

The signal and detection events are highly related to performance measures, so the first step is to process event-based data to calculate associated input variables for the proposed estimation model. **Figure 6-4** shows an example of vehicle trajectory and signal status. Three variables are extracted from event-based data. Occupancy time is the time difference between a vehicle triggering and leaving a detector, which is calculated using **Eq. 6-1**.

$$Occ_i = t_i(d^f) - t_i(d^0) \quad \text{Eq. 6-1}$$

where Occ_i is occupancy time generated by the i^{th} detection event; $t_i(d^f)$ is the timestamp when the detector is off in the i^{th} detection event; and $t_i(d^0)$ is the timestamp when the detector is on in the i^{th} detection event.

The second related variable is the waiting time of the first vehicle arriving at the intersection during the red duration until the green light is on, which is calculated using **Eq. 6-2**.

$$W_j = G_{j+1} - t_j(d^f) \quad \text{Eq. 6-2}$$

where W_j is the simple waiting time for a vehicle during the j^{th} cycle; $t_j(d^f)$ is the timestamp of the first detector-on event during the red duration in the j^{th} cycle; G_j is the timestamp of the green duration starts of the j^{th} cycle. In addition, the number of detection events (C) is an important variable for indicating traffic conditions.

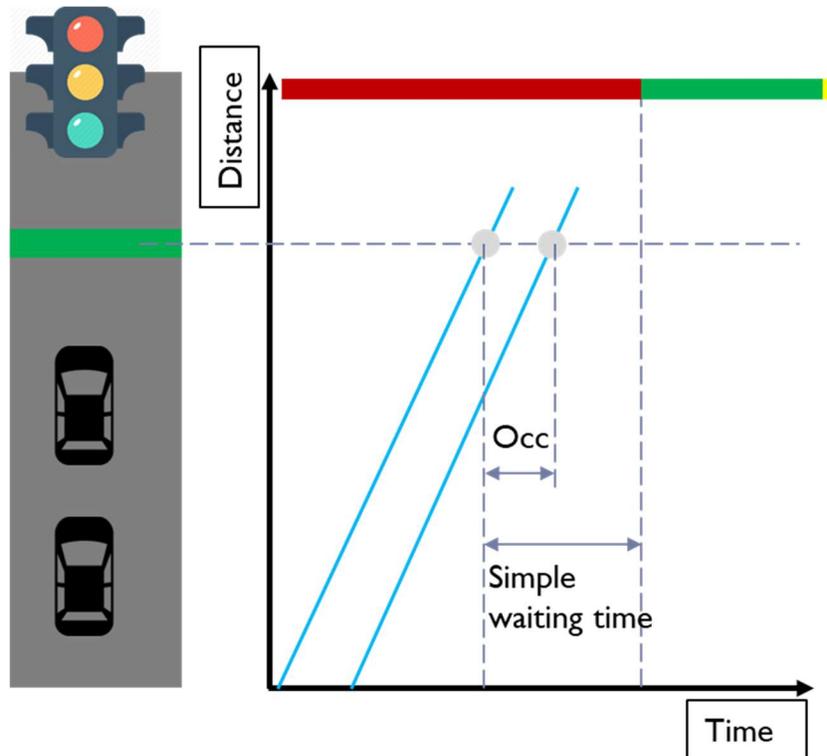


Figure 6-4. Vehicle trajectory and signal timing

Additionally, the signal status has a significant impact on the above-mentioned three extracted input variables. For example, given the same number of detection events, if these events occurred during the green duration, the variable indicates a light traffic condition in comparison with the same number of detection events that occurred during the red duration. Therefore, the signal status should also be considered when extracting the above three detection-related variables. All signal statuses are categorized into three groups based on the combination of the signal status when the detector is on and off.

- Red-to-Green: when vehicles trigger the detector during red duration and exit the detector during the green duration, as shown in **Figure 6-5(a)**.
- Red-to-Red: when vehicles trigger the detector during the red duration and exit the detector during the red duration as well, as shown in **Figure 6-5(b)**.
- Green-to-Green: when vehicles trigger the detector during the green duration and exit the detector during the green duration as well, as shown in **Figure 6-5(c)**.

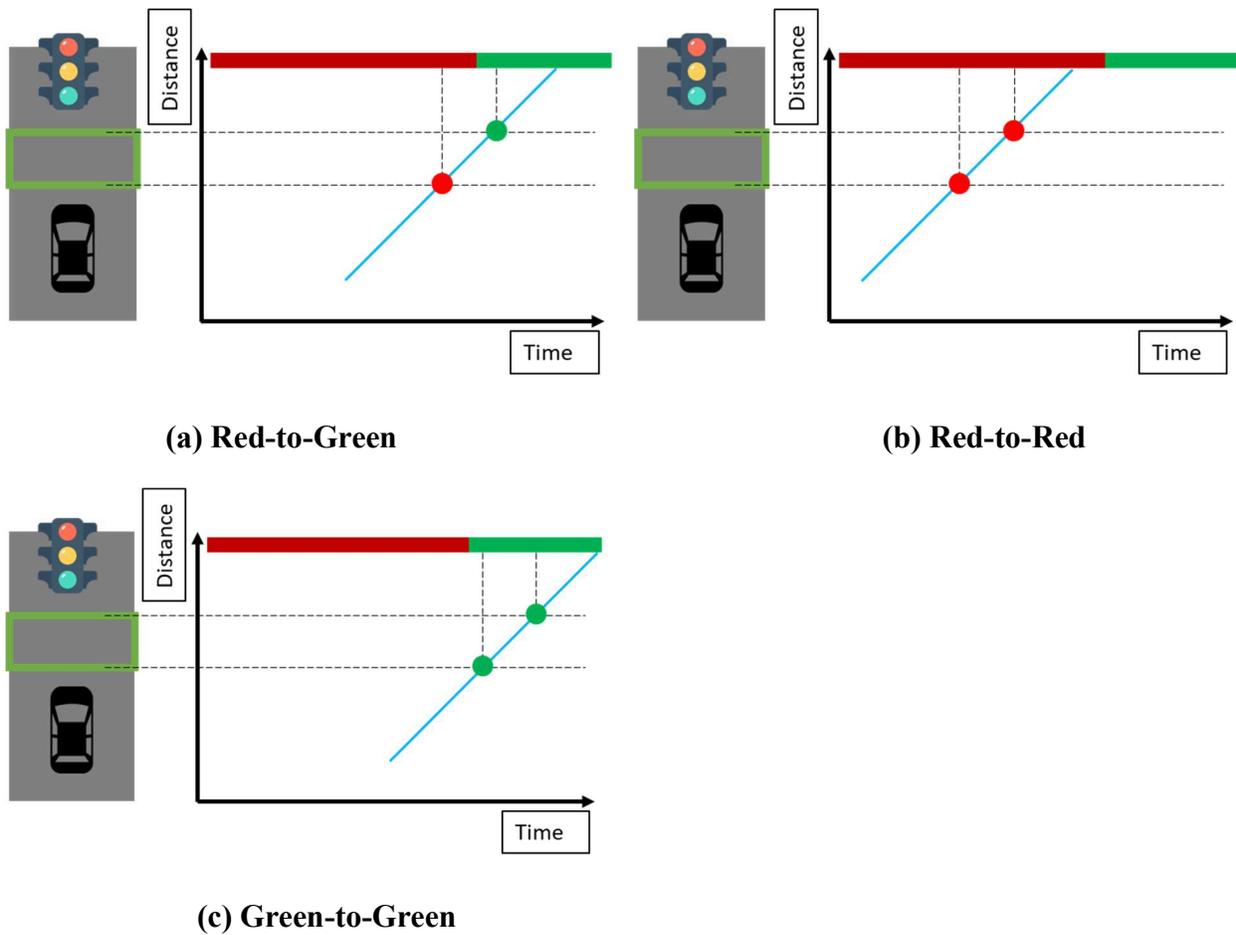


Figure 6-5. Signal status categorization based on detection events.

After combing the detection-related variables and signal status, the input variables can be formulated using **Eq. 6-3**.

$$X_{EB} = \begin{bmatrix} \overline{Occ(k)} \\ \overline{W(k)} \\ C(k) \end{bmatrix} \times \begin{bmatrix} RG \\ RR \\ GG \end{bmatrix} \quad \text{Eq. 6-3}$$

where X_{EB} is the input variables extracted from event-based data; $C(k)$ is the total number of detection events during k^{th} time period under three different signal statuses; $\overline{Occ(k)}$ and $\overline{W(k)}$ are average occupancy time and average simple waiting time during k^{th} period under three different signal statuses, which are calculated using **Eq.6-4** and **Eq.6-5**

$$\overline{Occ(k)} = \frac{\sum_{i=1}^{C(k)} Occ_i}{C(k)} \quad \text{Eq. 6-4}$$

$$\overline{W(k)} = \frac{\sum_{j=1}^{M(k)} W_j}{M(k)} \quad \text{Eq. 6-5}$$

where $M(k)$ is the number of cycles during k^{th} time period under three different signal statuses.

In addition to the variables extracted from event-based data, other relevant variables including speed limit, hour of the day, number of lanes, and shared (left or right turn) lane are included in input variables, so the input variables are formulated into **Eq. 6-6**.

$$X(k) = \{X_{EB}(k), S, H, L, A\} \quad \text{Eq. 6-6}$$

where S, H, L, A is the speed limit, hour of the day, number of lanes, and if there is a shared lane at the specific locations, and these four variables are converted into dummy variables.

6.2.2 Model-Agnostic Meta-Learning (MAML)

The raw data collected from various intersections are processed to calculate input $X(k)$, and estimating performance measures at each location can be considered as a learning task. However, applying traditional machine learning methods for estimating region-wide traffic performance measures is challenging because of the following two issues:

- Different intersections might have differences in terms of detector locations and lengths, which cause different relationships between the data collected by the detectors and performance measures. These various relationships are challenging to be accurately captured by traditional pre-trained machine learning models, especially when the pre-trained model is applied to a new intersection.
- The traditional machine learning models are challenging to handle the tasks with only limited data. The ground-truth traffic performance measures are extracted from probe vehicle trajectories data, but probe vehicle data currently has a relatively low sample size due to the high cost, especially at some small intersections.

In order to address the above two issues, the model-agnostic meta-learning (MAML) algorithm is applied to estimate region-wide traffic performance measures at various signalized intersections using event-based data. The MAML algorithm is one of the emerging meta-learning methods that was proposed (Finn et al., 2017) for general and fast learning of various problems with a small amount of data and is compatible with being adapted to new tasks after fine-tuning the model using a small amount of training data.

Figure 6-6 illustrates the training process of the MAML algorithm. The training dataset is categorized into a set of tasks $\mathcal{T} = \{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_i, \dots, \mathcal{T}_K\}$ according to the intersection and road

direction. The data in each task \mathcal{T}_i is then separated into support set $\mathcal{D}_{\mathcal{T}_i}^S$ and query set $\mathcal{D}_{\mathcal{T}_i}^Q$ likely a first and a second training data set respectively.

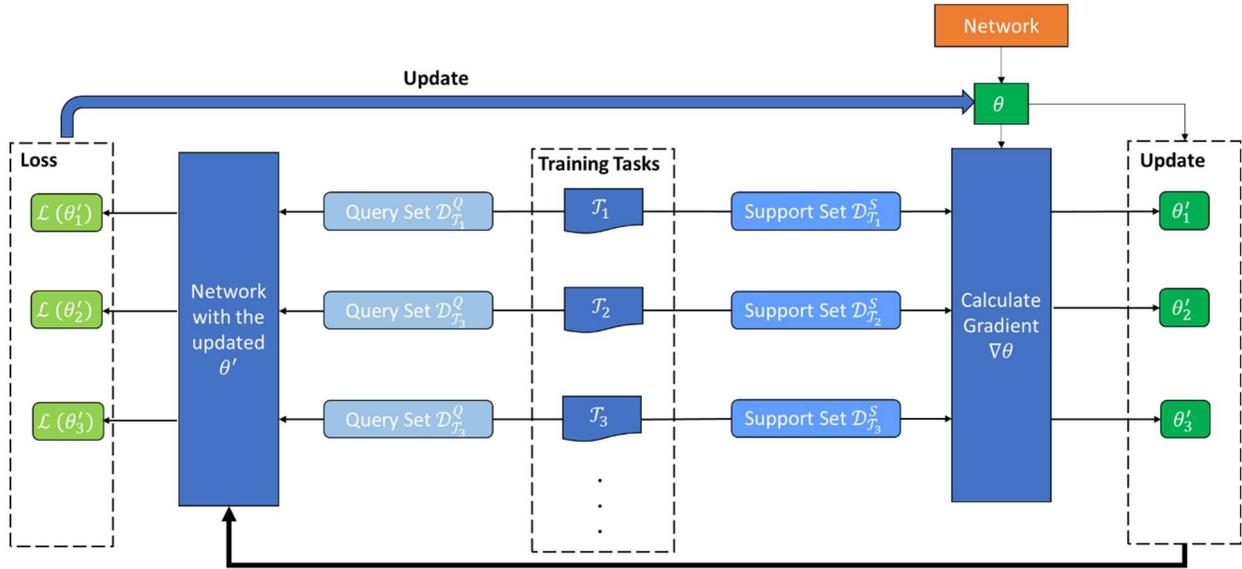


Figure 6-6. The training process of the MAML algorithm

The network is represented using f_θ with the parameters θ . The training process of the MAML algorithm is to optimize the parameters θ according to the loss function. First, the support set $\mathcal{D}_{\mathcal{T}_i}^S$ in each task is adapted into f_θ , and the initial parameters θ are updated as θ' using one gradient descent update, and the calculation is formulated as **Eq.6-7**. One gradient descent update is recommended for updating parameters in order to speed up the training procedure and avoid overfitting when using limited training data.

$$\theta' = \theta - \alpha \nabla_{\theta} \mathcal{L}(\theta | \mathcal{D}_{\mathcal{T}_i}^S) \quad \text{Eq. 6-7}$$

where α is the meta learning rate, and \mathcal{L} is the loss function.

After obtaining the updated parameter θ' for each task, the network with the updated parameters is applied to the query set $\mathcal{D}_{\mathcal{T}_i}^Q$ of each task for calculating the loss $\mathcal{L}(\theta' | \mathcal{D}_{\mathcal{T}_i}^Q)$. The objective of the model training is to find the optimal parameters θ by minimizing the total loss of the network using the updated parameters θ' across all tasks. Therefore, the meta-objective is formulated as **Eq.6-8**

$$\min_{\theta} \sum_{\mathcal{T}_i \sim p(\mathcal{T})} \mathcal{L}(\theta' | \mathcal{D}_{\mathcal{T}_i}^Q) = \min_{\theta} \sum_{\mathcal{T}_i \sim p(\mathcal{T})} \mathcal{L}(\theta - \alpha \nabla_{\theta} \mathcal{L}(\theta | \mathcal{D}_{\mathcal{T}_i}^S) | \mathcal{D}_{\mathcal{T}_i}^Q) \quad \text{Eq. 6-8}$$

The parameters θ are updated using stochastic gradient descent (SGD), which is formulated as **Eq.6-9**

$$\theta \leftarrow \theta - \beta \nabla_{\theta} \sum_{\mathcal{I}_i \sim p(\mathcal{I})} \mathcal{L}(\theta' | \mathcal{D}_{\mathcal{I}_i}^Q) \quad \text{Eq. 6-9}$$

where β is the meta step size.

When calculating the gradient decent for $\nabla_{\theta} \mathcal{L}(\theta' | \mathcal{D}_{\mathcal{I}_i}^Q)$ in Equation (9), $\frac{\partial \mathcal{L}(\theta')}{\theta(w)}$ is assumed to approximate $\frac{\partial \mathcal{L}(\theta')}{\theta'(w)}$. $\theta(w)$ and $\theta'(w)$ are w^{th} parameter in θ and θ' . Therefore, the gradient descent is formulated as Eq. 6-10.

$$\nabla_{\theta} \mathcal{L}(\theta' | \mathcal{D}_{\mathcal{I}_i}^Q) = \begin{bmatrix} \frac{\partial \mathcal{L}(\theta')}{\theta(1)} \\ \frac{\partial \mathcal{L}(\theta')}{\theta(2)} \\ \frac{\partial \mathcal{L}(\theta')}{\theta(3)} \\ \vdots \end{bmatrix} \approx \begin{bmatrix} \frac{\partial \mathcal{L}(\theta')}{\theta'(1)} \\ \frac{\partial \mathcal{L}(\theta')}{\theta'(2)} \\ \frac{\partial \mathcal{L}(\theta')}{\theta'(3)} \\ \vdots \end{bmatrix} \quad \text{Eq. 6-10}$$

Mean squared error (MSE) is used as the loss function, which is formulated as **Eq. 6-11**

$$\mathcal{L}(\theta' | \mathcal{D}_{\mathcal{I}_i}^Q) = \sum_{X(k), y(k) \sim \mathcal{I}_i} \|f_{\theta'}(X(k)) - y(k)\|_2^2 \quad \text{Eq. 6-11}$$

where $y(k)$ is the traffic performance measure during the k^{th} period at location i .

Once the parameters θ are optimized in the training process, the model f_{θ} can be applied to estimate the traffic performance measures at intersections that might or might not be included in the training dataset. The first step is to further tune up the model f_{θ} using a small amount of data in the new task to ensure the MAML can quickly learn how to estimate the traffic performance measure at this intersection. After tuning up the MAML, the model can be used to estimate traffic performance measures at the same intersection.

6.3 IMPLEMENTATION AND RESULTS

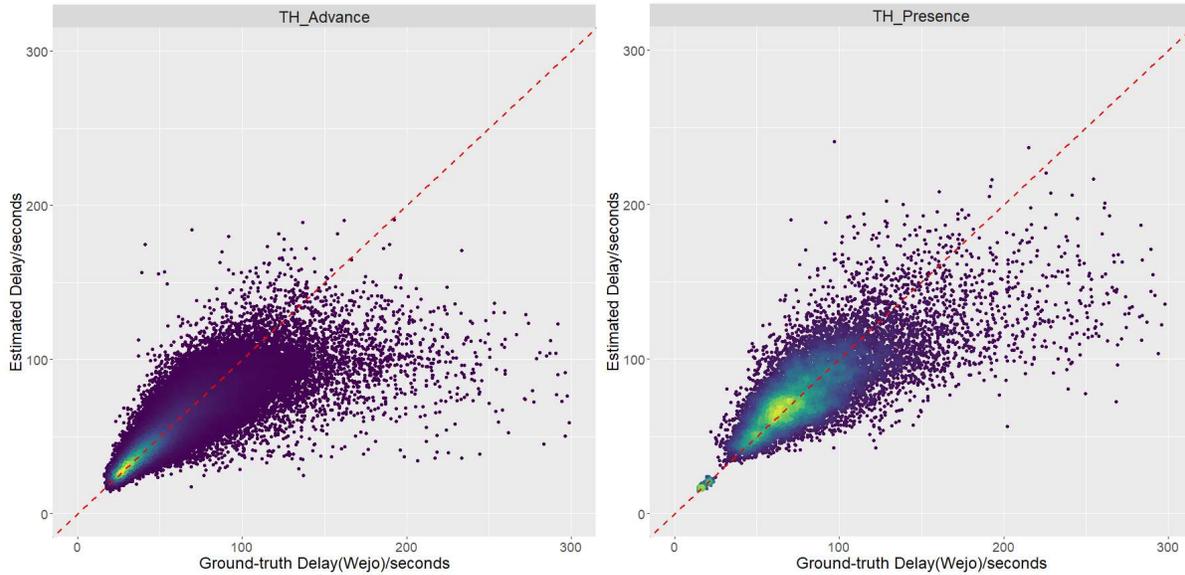
To evaluate the proposed method's performance, the data including event-based data and ground-truth control delay and arrival-on-green (AOG) ratio data is collected from 201 study intersections for model training and validation. Among these study intersections, 144 are managed by the

MaxView system and 57 are managed by the Miovision system. The data collected from Jan. 7 – March 30, 2021 is used as the training dataset, and 50% of the training data is support data, and the remaining training data is query data. The data from April 1-Nov. 30, 2021 is used as the validation dataset, with 20% of the validation dataset used for tuning up the model and the remaining 80% used to evaluate model performance.

6.3.1 Control Delay Estimation

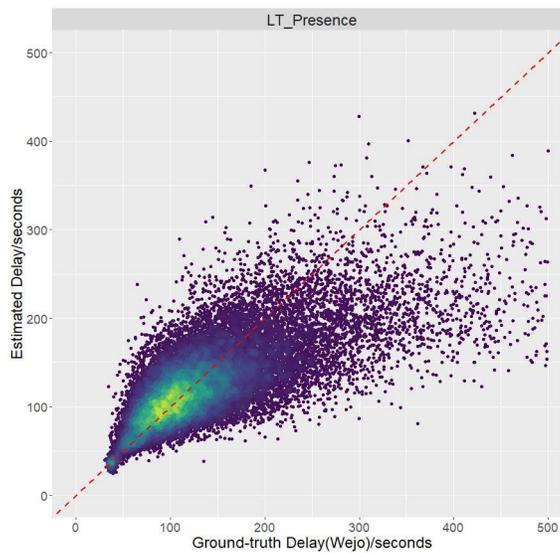
The ground-truth control delay is derived from the Wejo connected vehicle data that has a sufficient sample size for ensuring accuracy. Once the proposed model is trained and tuned up for a specific location, the model can be used to estimate the control delay using event-based data collected through the MaxView system. **Figure 6-7** illustrates the comparison between ground-truth and estimated control delay. Specifically, **Figure 6-7(a)** showcases the comparison for through movements where advance detectors are configured. In the plot, the red dashed line represents the ideal scenario where the estimated values align perfectly with the ground-truth values. Due to the large size of the validation dataset and the resulting overlap of data points, colors are utilized to indicate the density of overlapping points. Brighter colors indicate higher density. The comparison indicates that the points with the highest density align closely with the red dashed line, suggesting that the majority of estimated control delay values closely approximate the ground-truth delay. Some points with lower density are positioned on both sides of the red dashed line, indicating slight discrepancies between the estimated and ground-truth delay values but with a similar trend. However, when the control delay is larger than 150 seconds, the points appear below the red dashed line, indicating an underestimation issue with the estimated delay. Two possible reasons could cause this underestimation issue. According to observation, the data size for control delays exceeding 150 seconds is relatively limited compared to other data, which may affect the model's ability to accurately capture the relationship between input variables and control delay in that specific range. The other reason is there might be some outliers within the Wejo data, leading to incorrect ground-truth data that affects the accuracy of the estimation.

Figure 6-7(b) illustrates the comparison between ground-truth and estimated delay for through movements with presence detectors. Similar findings from **Figure 6-7(a)** are observed in **Figure 6-7(b)**. These points with high density align closely with the red dashed line, indicating the estimated delay closely approximates the ground-truth delay. On the other hand, points with low density are situated on both sides of the red dashed line, suggesting some discrepancies between the estimated and ground-truth delay values, although they exhibit a similar trend. Similarly, when the delay is higher than 150 seconds, the estimated control delay also has an underestimation issue, which might be caused by the above-mentioned reasons.



(a) through movement with advance detectors configured

(b) through movement with presence detectors configured



(c) left-turn movement with presence detectors configured

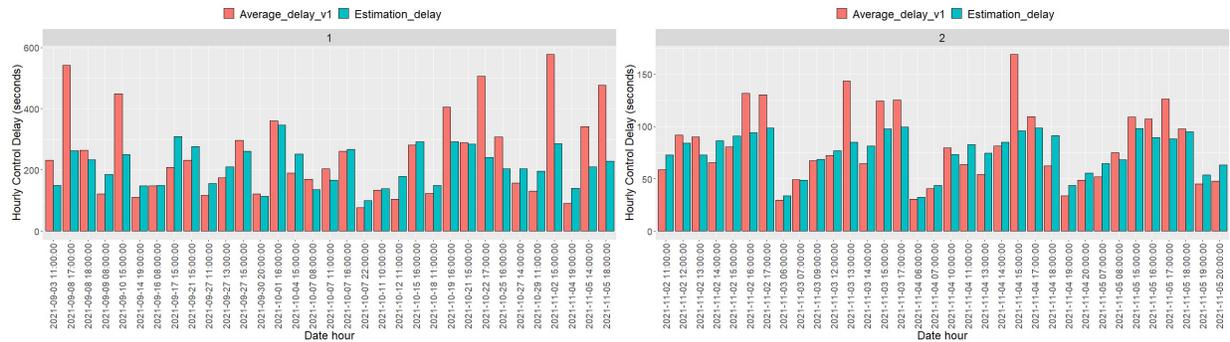
Figure 6-7. Comparison between estimated delay and ground-truth delay using the MaxView system at all study intersections

Figure 6-7(c) illustrates the comparison between ground-truth and estimated left-turn control delay, where all left-turn lanes are exclusively configured with presence detectors. Similar to the previous plots, points with high density align closely with the red dashed line, indicating that the estimated left-turn delay closely approximates the ground-truth delay. In comparison to the results for the

through movement, the left-turn data exhibits a higher number of outliers with significant errors. Moreover, when the left-turn delay exceeds 250 seconds, the estimated delay encounters an underestimation issue. In addition to the possible reasons mentioned earlier, the training data for the left turn has a much lower sample size than the through movement after processing and cleaning the Wejo data, likely further leading to biased output.

To provide a temporal comparison between the estimated and ground-truth delay for different movements and directions, the intersection of Speedway Blvd. & Euclid Ave. is chosen as an example. **Figure 6-8** showcases the comparison between the estimated and ground-truth delay for left-turn and through movements in four directions across various dates and hours.

In most instances, the estimated through and left-turn delay closely match the ground-truth data with only slight errors, particularly for EB and WB through movements where advance detectors are configured. One interesting finding is there are a few instances where there are significantly higher delays at certain hours compared to the regular delays, and the estimated delay consistently exhibits an underestimation issue during these periods. For example, the ground-truth delay at 3 p.m. on Nov. 4, 2021 in **Figure 6-8 (b)** and at 5 p.m. on Nov. 4, 2021 in **Figure 6-8(g)** has a delay of 160 and 700 seconds, respectively, which is much higher than the regular delays observed in the same location. During these two hours, the estimated delay significantly underestimates the ground-truth delay. One possible reason is that these two ground-truth values could be outliers caused by some outliers in Wejo data.



(a) WB left-turn movement

(b) EB through movement

Figure 6-8. Ground-truth and estimated control delay comparison at Speedway Blvd. & Euclid Ave.



Figure 6-9. Ground-truth and estimated control delay comparison at Speedway Blvd. & Euclid Ave.

To further quantify the performance of the proposed method on estimating control delay, the three metrics of mean absolute error (MAE), root mean square error (RMSE), RMSE%, and mean absolute percentage error (MAPE) are calculated for each location. All locations are categorized into three groups based on the movement and detector layout.

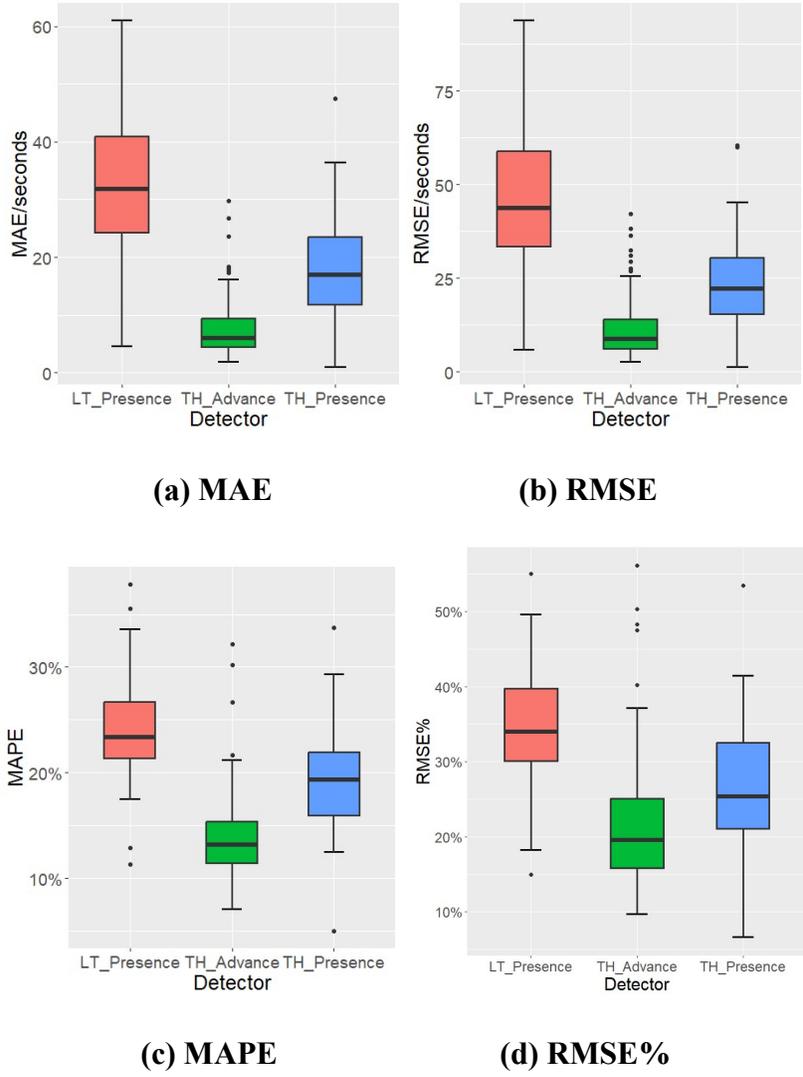


Figure 6-10. Method performance for delay estimation using the MaxView system by detector location and traffic movement.

Figure 6-9(a) shows the boxplots of MAE for three groups of locations. The through movement with advance detectors configured demonstrates the best performance, with the majority of locations having an MAE lower than 10 seconds. The through movement with presence detectors configured exhibits an MAE ranging from 11 to 22 seconds with an average value of 17 seconds, which is higher than the locations with advance detectors. Left-turn movements have the highest MAE, with a range of 25-40 seconds. Similar findings are observed in RMSE and MAPE. Through movements with advance detectors perform the best, with an RMSE of 5-12 seconds and MAPE of 12%-15%. Movements with presence detectors perform second best, with an RMSE of 12-30 seconds and MAPE of 15%-22%. Left-turn movements have the lowest performance, with an RMSE of 30-60 seconds and MAPE of 22%-27%. One possible reason for the lower performance of left-turn movements is the small size of the training data. Taking **Figure 6-8(g)** as an example,

the SB left-turn movement only has around 20 available data points over a span of five months. The limited amount of training data for left-turn movements may contribute to the lower accuracy in estimating the control delay.

The event-based data collected from the other system, the Miovision system, is also used to train and validate the model for control delay estimation. In contrast to the MaxView system, the Miovision system employs both advance and presence detectors on through lanes, and presence detectors only on left-turn lanes. Additionally, the virtual detectors in the Miovision sensors are rectangular in shape, as opposed to bar or arrow detectors. Therefore, a separate model is trained using the data collected from the Miovision system.

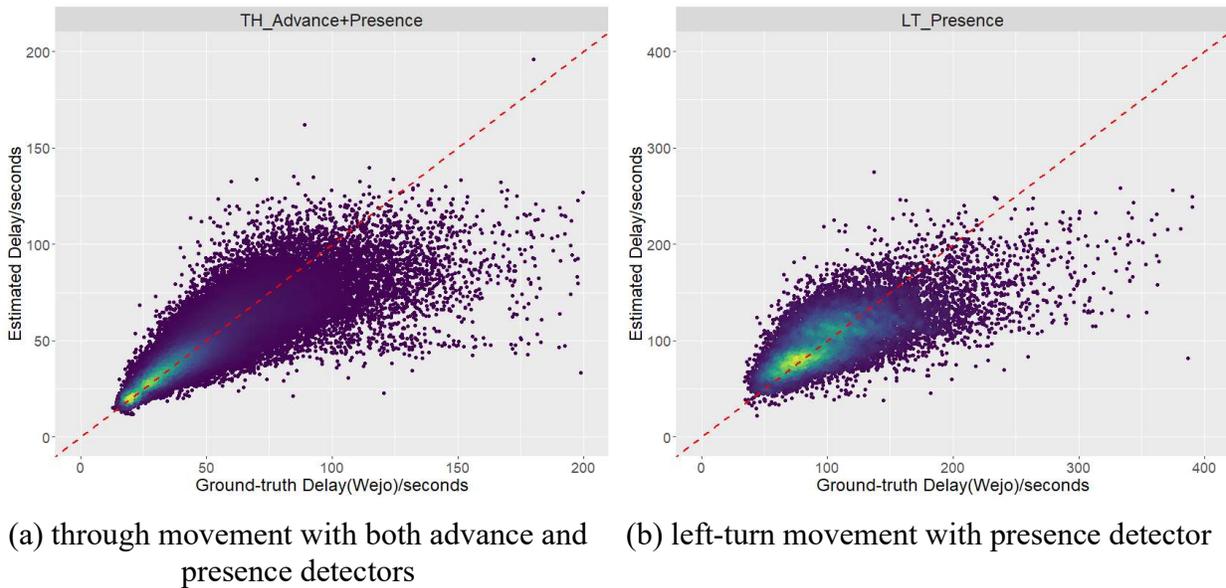


Figure 6-11. Comparison between estimated delay using Miovision system and ground-truth delay at all study locations

Figure 6-10 shows the comparison between the estimated delay and ground-truth delay. All study locations are categorized into two groups: through movement with both advance and presence detector configured and left-turn movement with presence detector configured. As shown in **Figure 6-10(a)**, most periods have a relatively low delay, and these data points closely align with the red dashed line, indicating a high level of accuracy in the estimated delay. However, a similar underestimation issue is also observed when the delay is higher than 100 seconds. In the case of left-turn movements, as depicted in **Figure 6-10(b)**, the majority of data points are situated in close proximity to the red dashed line. However, it is worth noting that the estimated delay tends to underestimate the ground-truth values when the delay exceeds 200 seconds.

To assess the accuracy and robustness of the model in estimating delay using data from the Miovision system, the three performance metrics MAE, RMSE, and MAPE are calculated. **Figure 6-11** presents the results of these metrics. It is observed that the through movement exhibits lower

error rates compared to the left-turn movement. For the through movement, the proposed method demonstrates a similar performance in estimating delay using data from both the Miovision and MaxView systems. However, when estimating the delay for left-turn movements, the model utilizing data from the Miovision system shows relatively higher performance compared to the MaxView system.

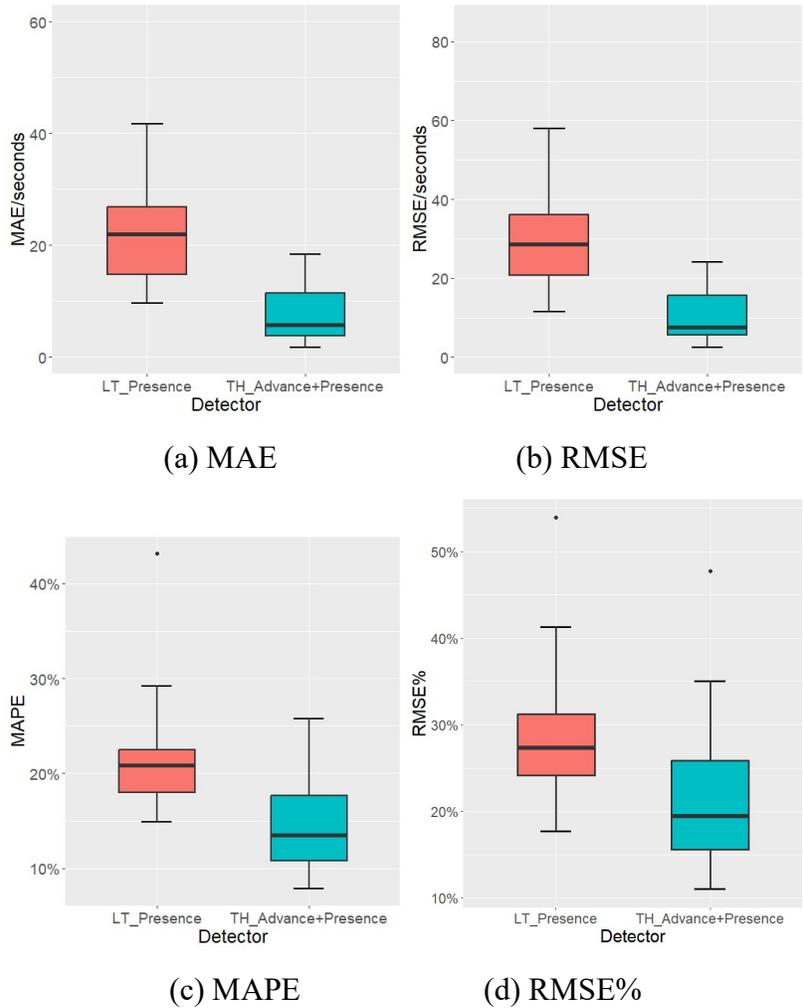


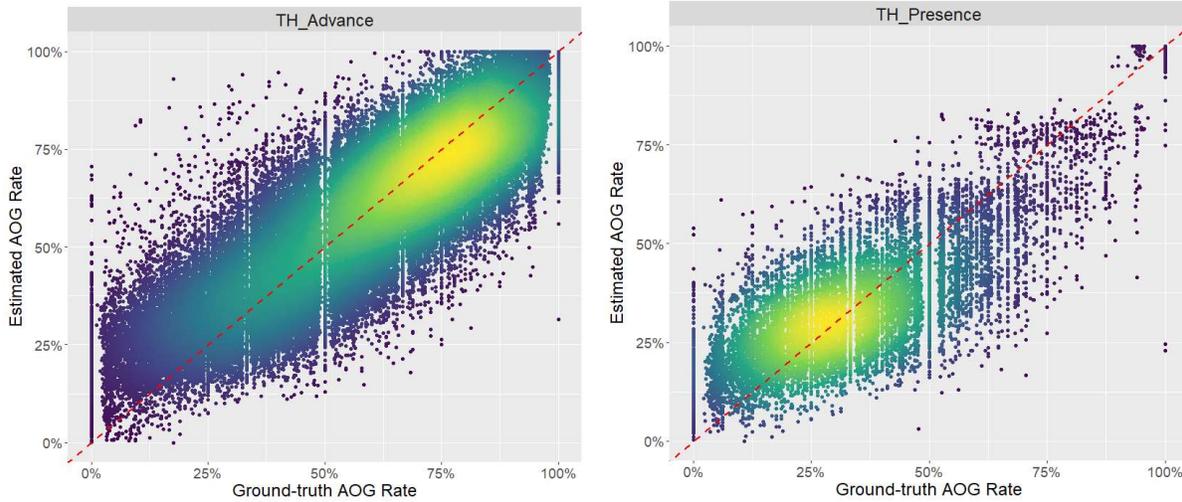
Figure 6-12. Method performance for delay estimation using the Miovision system by detector location and traffic movement

6.3.2 Arrival-on-Green Estimation

The data collected from the MaxView system is utilized to train and validate the model for estimating the AOG ratio. **Figure 6-12** provides a visual comparison between the estimated AOG and the ground-truth AOG at all study locations. The study locations are also divided into three groups based on the traffic movement and detector configurations: through movement with

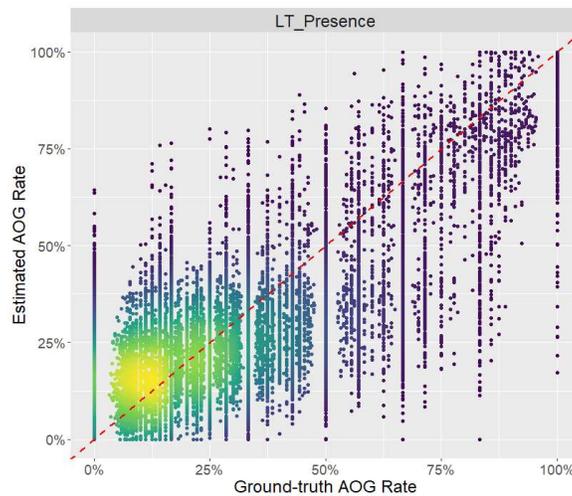
advance detector configured, through movement with presence detector configured, and left-turn movement with presence detector configured.

Figure 6-12(a) shows the comparison between estimated and ground-truth AOG ratio for the through movement at locations with advance detectors configured. Most data points align with or are close to the red dashed line, indicating that the estimated AOG values closely approximate the ground-truth values. In **Figure 6-12(b)**, the data comparison is shown for through AOG at locations with presence detectors configured. Most data points are close to the red dashed line, demonstrating that the proposed model provides reliable and accurate AOG data. However, an underestimation issue is observed when the AOG exceeds 60%. This can be attributed to the smaller amount of data in this higher AOG range compared to the lower AOG range, which affects the model's ability to accurately capture the relationship between input variables and AOG. **Figure 6-12(c)** shows the comparison between estimated and ground-truth left-turn AOG at locations with presence detectors. While most of the data points are close to the red dashed line, there are numerous outliers that deviate significantly from it, particularly when the AOG is higher than 30%. This could be due to the smaller sample size for left-turn movements or the lower reliability and accuracy of using Wejo data for left-turn AOG calculation, which requires further investigation.



(a) through movement with advance detectors configured

(b) through movement with presence detectors configured

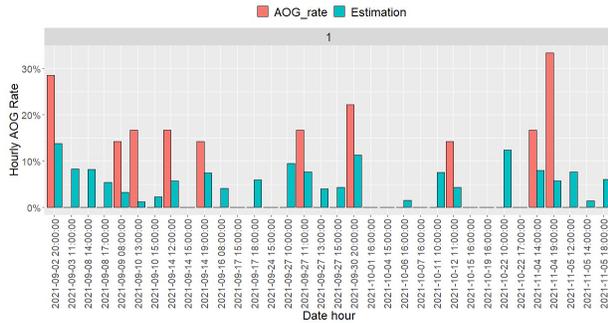


(c) left-turn movement with presence detectors configured

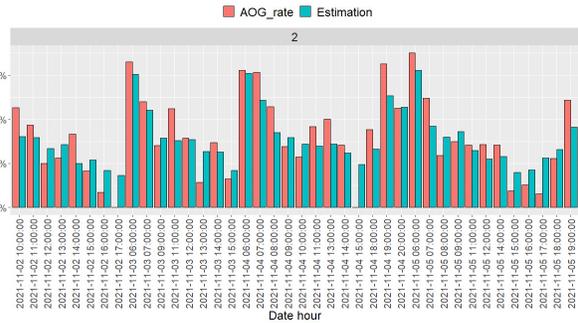
Figure 6-13. Comparison between estimated AOG and ground-truth AOG using the MaxView system at all study locations

Speedway Blvd & Euclid Ave is used as the example to temporally illustrate the comparison between estimated and ground-truth AOG for left-turn and through movements across four road directions, as shown in **Figure 6-13**. For through movements, the estimated AOG values closely match the ground-truth values during most time periods with only slight errors observed. Furthermore, the estimated AOG exhibits a similar temporal trend to the ground-truth AOG, indicating that it accurately captures the changes in traffic conditions. However, for left-turn

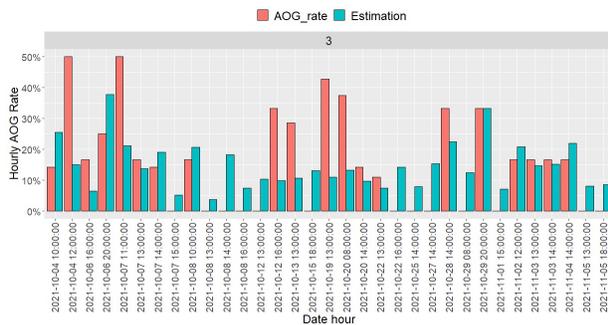
movements, the estimated AOG shows higher errors across most time periods. When the AOG is higher than the usual range, the estimated AOG values tend to underestimate the ground-truth values. Conversely, overestimation issues occur when the AOG is significantly lower, even reaching zero. These higher errors in estimated left-turn AOG can be attributed to two major factors: the smaller data size available for left-turn movements and the potential inaccuracies in the ground-truth data.



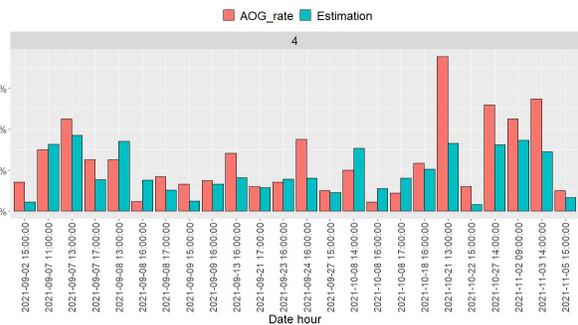
(a) WB left-turn movement



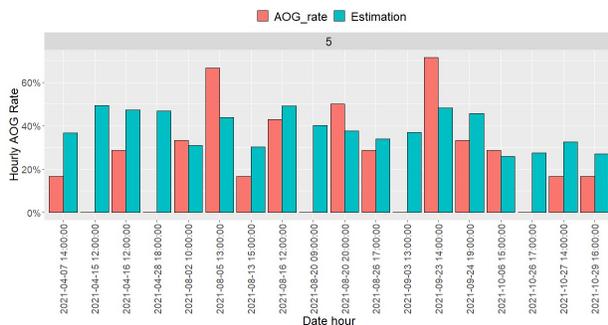
(b) EB through movement



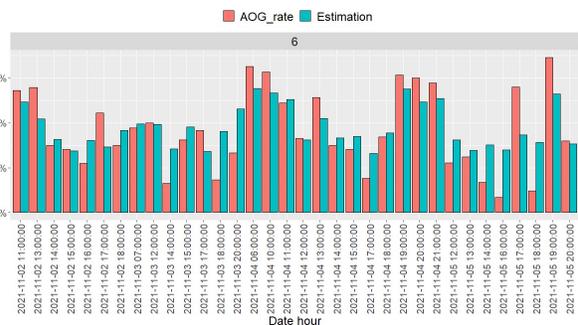
(c) NB left-turn movement



(d) SB through movement

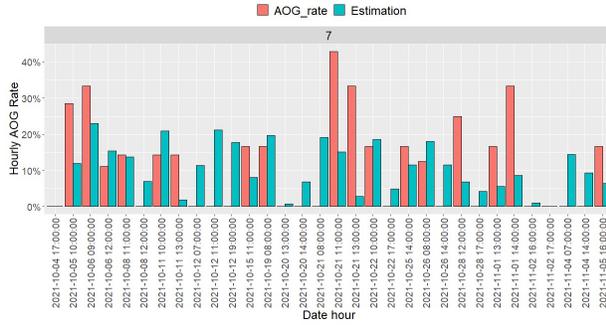


(e) EB left-turn movement

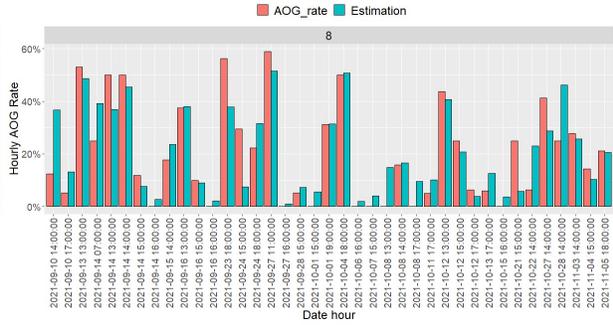


(f) WB through movement

Figure 6-14. Ground-truth and estimated AOG comparison at Speedway Blvd. & Euclid Ave.

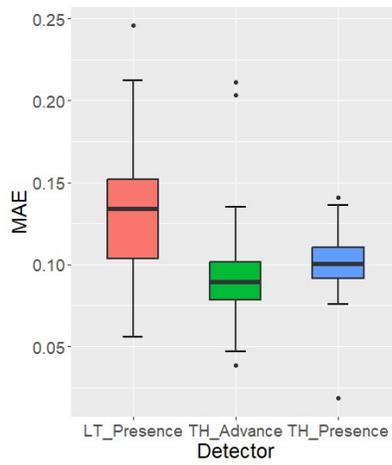


(g) SB left-turn movement

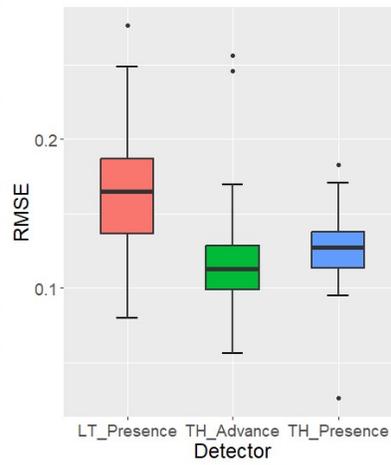


(h) NB through movement

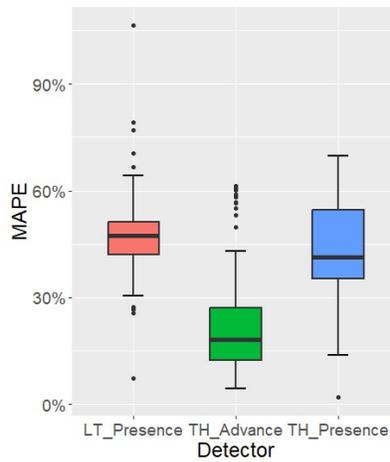
Figure 6-15. Ground-truth and estimated AOG comparison at Speedway Blvd. & Euclid Ave.



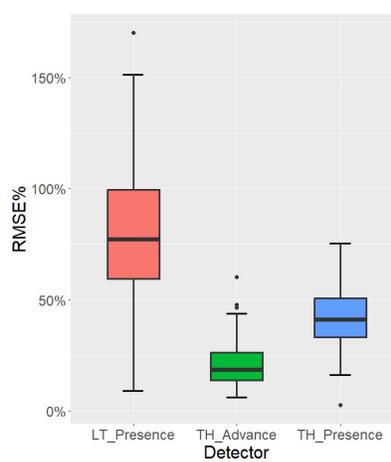
(a) MAE



(b) RMSE



(c) MAPE



(d) RMSE%

Figure 6-16. Model performance for AOG estimation using the MaxView system by detector location and traffic movement

To evaluate the performance of the model for AOG estimation, the three performance metrics MAE, RMSE, and MAPE, are calculated for different movements and detector configurations. The results, as presented in **Figure 6-14**, indicate varying levels of performance across different scenarios. For through movements with advance detectors configured, the model achieves the highest performance. The MAE ranges from 7% to 10%, the RMSE ranges from 10% to 13%, and the MAPE ranges from 13% to 30%. These results highlight the model's ability to accurately estimate AOG data using event-based data for through movements with advance detector configurations. Through movements with presence detectors show similar MAE and RMSE values to those with advance detectors. However, the MAPE value is higher for these locations, primarily because they tend to have lower AOG ratios. Nonetheless, the model still demonstrates a reasonably good performance for estimating AOG in these scenarios. The proposed method has the lowest performance for left-turn movement due to the above-mentioned two possible reasons.

The event-based data collected from the Miovision system is also used to train the proposed model for AOG estimation. **Figure 6-15(a)** depicts the comparison between the estimated and ground-truth AOG for the through movement. Most data points align closely with the red dashed line, indicating that the proposed method accurately estimates the AOG for the through movement. **Figure 6-15(b)** shows the AOG comparison for the left-turn movement. Obviously, the proposed model for left-turn AOG estimation exhibits a higher error, particularly when the AOG exceeds 50%. This finding is consistent with the left-turn AOG estimation using data collected from the MaxView system.

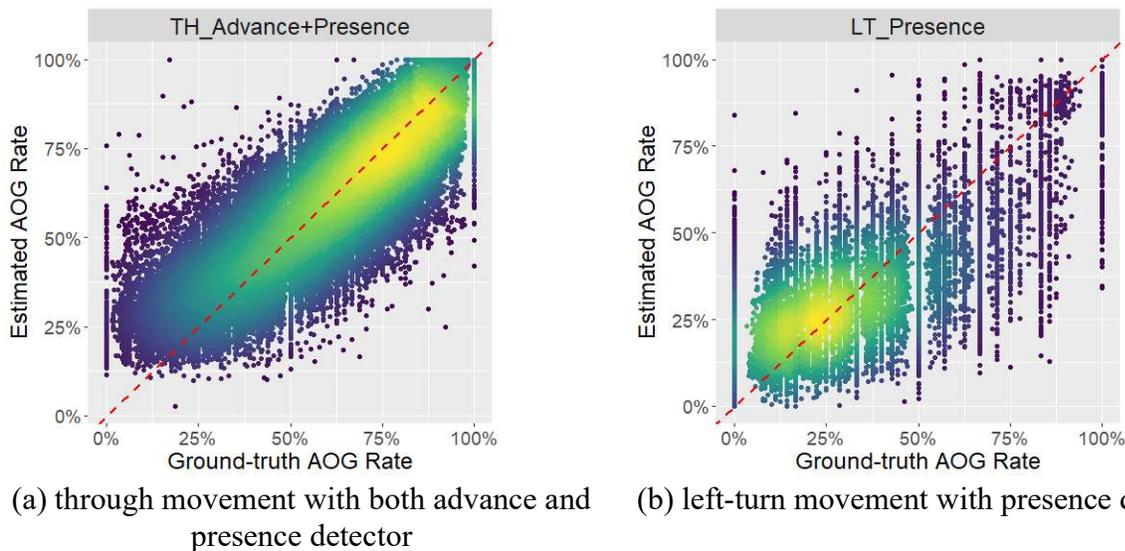


Figure 6-17. Comparison between estimated AOG using Miovision system and ground-truth AOG at all study locations

The model performance for left-turn and through AOG estimation is quantified using MAE, RMSE, and MAPE, as presented in **Figure 6-16**. For the through movement, the model achieves an MAE of 7%-10%, an RMSE of 10%-12%, and a MAPE of 10%-30%. These performance metrics

indicate that the proposed model performs well in estimating AOG for the through movement. The results are consistent with the performance observed when using the MaxView system, indicating that the model maintains a reliable and consistent performance regardless of the detector type. On the other hand, the model performance for left-turn traffic is relatively lower than that for the through movement. The left-turn AOG estimation exhibits higher errors compared to the through movement, as reflected in the higher MAE, RMSE, and MAPE values. This aligns with the findings observed in the MaxView system.

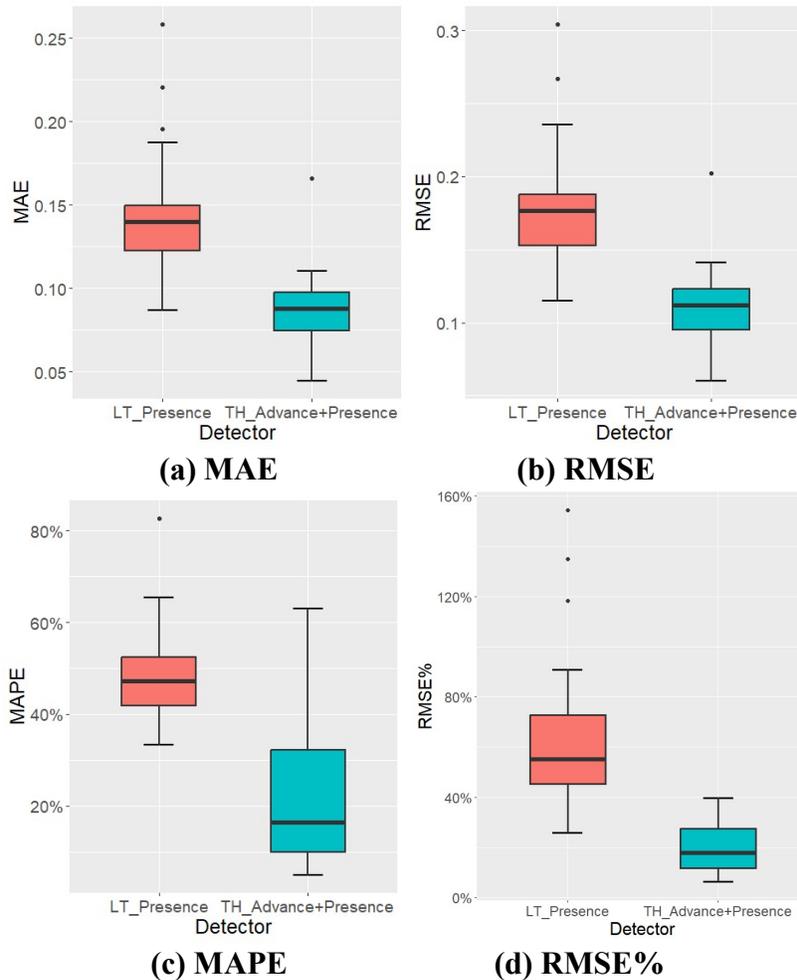


Figure 6-18. Model performance for AOG estimation using Miovision system by detector location and traffic movement

CHAPTER 7: ACCEPTANCE CRITERIA DEVELOPMENT

To control the output data quality from the proposed estimation methods, the project team will develop an integration method with multiple sources of sample data based on the findings from Tasks 3 through 6. A quality assurance and quality control (QA/QC) procedure is then developed for assessing data derived from Wejo connected vehicle data and traffic performance estimated by the proposed estimation model. Acceptance criteria, such as mean absolute error (MAE), mean absolute percentage error (MAPE), and root-mean-square error (RMSE), will be developed by the findings from sample traffic data assessment and literature review.

7.1 DATA COLLECTION AND PROCESSING

In this project, three major data sources are collected and analyzed: performance measures provided by the Miovision system, Wejo connected vehicle data, and high-resolution event-based data collected by the Miovision and MaxView systems. The Miovision system can provide simple (stop) delay, arrival-on-green, and split failure. These three performance measures can be directly downloaded through a specific platform developed by Miovision after undergoing data cleaning and processing, and therefore, further data processing is not needed. Simple delay and split failure can also be extracted from high-resolution event-based data according to their definition. This extraction process involves data cleaning and processing, particularly when dealing with any missing data within the event-based data.

Wejo data is used to extract accurate traffic performance measures as the ground-truth data, so necessary data processing is required to control the data quality. In addition to removing the trajectory outliers that fall outside the road network, the threshold for sample size has been determined to filter the data that can be used to calculate reliable and accurate performance measures. This step is important because the accuracy of Wejo data for performance measure calculation can vary based on the sample size. The details regarding the determination of this threshold are presented in Chapter 4.

High-resolution event-based data collected from the Miovision and MaxView systems are used to estimate traffic performance measures. The high-resolution data may have data missing due to communication loss, and so the data must be quality checked and cleaned before being used for performance measure estimation. After data cleaning, the data is input into the proposed method for estimating performance measures. To control and ensure the data quality of the estimated performance measures, the associated acceptance criteria are developed according to the literature review and data validation results.

7.2 ACCEPTANCE CRITERIA DEVELOPMENT

Developing acceptance criteria is one of the critical steps in QA/QC procedures, specifying the conditions that must be met for estimated performance measures, including control delay and AOG, to be considered valid, accurate, and acceptable.

The purpose of using the estimated performance measures as well as the data resolution should be determined and clarified first. In this project, the data is mainly used for two purposes, real-time traffic monitoring and long-term traffic planning. The hourly-based performance measures are estimated for the use of real-time traffic monitoring, and the hourly-based data is aggregated into monthly average hourly data for the use of long-term traffic planning.

7.2.1 Literature Review Summary

Multiple methods have been proposed for performance measure estimation, but the accuracy of these methods varies. To comprehensively understand the accuracy and reliability of the previously proposed estimation methods, a comprehensive literature review regarding control delay and AOG estimation is summarized in **Tables 7-1** and **7-2**. **Table 7-1** shows the literature review summary regarding control delay estimation in terms of methodology, data sources, ground-truth data collection, and accuracy. Both traditional methods such as the HCM method and emerging AI methods such as KNN have been applied to estimate the control delay at signalized intersections. Most studies leveraged traffic data such as volume and speed collected by traffic detectors for delay estimation, which is challenging to be collected through detectors directly. Their evaluation results show that most proposed methods have a relatively low error. One possible reason is some of these studies used simulation for the method validation, which cannot accurately reflect the model performance when using field data. The second reason is the number of study intersections is small, with only one or two intersections, which did not indicate the robustness and transferability of the proposed methods for network-level performance measure estimation.

A comparison study conducted by (Wang et al., 2016) evaluates the model performance of various delay estimation methods. The results show that the accuracy of the four methods is inconsistent, with a high standard deviation, and varies across the time periods and locations. Among these methods, the most accurate method has an absolute error ranging from 0.8 seconds to 46 seconds, and the Absolute Percentage Error (APE) ranging from 0.2% to 29%. In addition, the left-turn movement commonly has a higher error than the through movement.

Table 7-1. Literature Review Summary of Control Delay Estimation

Reference	Methods	Data Sources	Ground-truth Delay Collection	#Intersections	Data Resolution	Accuracy
(Bagdatli and Dokuz, 2021)	KNN, SVR, RF, XGBoost	Signal timing info, volume, queue.	Manually collecting via videos	12	Cycle-based	MAE: 0.8-2 sec. MSE: 1.4-7 sec.
(Anusha S. P. et al., 2016)	Kalman filter	Signal timing info, loop detector data	-	Two	5-minutes	RMSE: 0.2-0.8 sec.

(Wang et al., 2016)	Deterministic queuing model	Red/green duration, cycle length, saturation flow rate, demand flow.	Image processing	Three	15-minutes	Absolute error: 3-84 sec. APE: 4%-94%
	Webster model					Absolute error: 3-96 sec. APE: 5%-94%
	HCM 2000 model	Red/green duration, cycle length, saturation flow rate, demand flow, and various factors.				Absolute error: 0.8-46 sec. APE: 1%-30%
	Adjusted HCM 2000 model					Absolute error: 0.3-58 sec. APE: 2%-29%
(Dobrota et al., 2022)	Incremental queue accumulation (IQA) delay model	High-resolution event-based data	Simulation	Two	Cycle-based	MAPE: 4.3%-11.2%
(Saha Arpita et al., 2017)	Modified HCM method	Red/green duration, cycle length, saturation flow rate, demand flow, and various factors.	Manually collecting via videos	Seven	Cycle-based	MAPE: 3-5%
(Ban et al., 2009)	A least-squares-based algorithm	Sample vehicle trajectories	Simulation	One	Cycle-based	Error rate <15%
(Zheng et al., 2013)	Image processing algorithm	Traffic videos	Manually collecting via videos and simulation	One	Cycle-based	Error rate: 7%

Table 7-2 summarizes the literature regarding AOG estimation, where only one relevant study was found. The study conducted by (Gavric et al., 2023) applied a machine-learning method to estimate the number of arrival-on-green vehicles using the data collected by stop-bar detectors rather than

advance detectors. The data is collected from only one study intersection, and the simulation results show the model error is low with an MAE of 1 vehicle and RMSE of 1.2-1.5 vehicles. According to the average arrival-on-green vehicles per cycle, 5-10 vehicles, the rough error rate is around 10%-20%. In addition, this method is specifically proposed for the through movement only rather than the left-turn movement.

Table 7-2. Literature Summary of AOG Estimation

Reference	Methods	Data Sources	Ground-truth Delay Collection	#Intesections	Data Resolution	Accuracy
(Gavric et al., 2023)	Multigene genetic programming	Traffic data collected by stop-bar detectors	Simulation	One	Cycle-based	MAE: 0.9-1 vehicles; RMSE: 1.2-1.5 vehicles

7.2.2 File Data Assessment

To establish the acceptance criteria for estimated performance measures in the PAG region, the data collected from over 200 signalized intersections in the PAG region is used to estimate control delay and AOG. Three performance metrics are calculated based on the estimated and ground-truth data and summarized in **Tables 7-3** and **7-4**. **Table 7-3** summarizes the evaluation results for control delay estimation of through and left-turn movement using the data collected through the MaxView and Miovision systems. The median MAPE of through movements with advance detectors at all locations is around 13% with a standard deviation of 4%. Even though the MAPE varies by location, the MAPE remains below 20%. The MAPE of through movements with advance detectors and left-turn movements is relatively higher, but still lower than 30%. In comparison with previous studies, the error of the proposed method for control delay estimation is acceptable.

Table 7-3. Method performance metrics for hourly control delay estimation

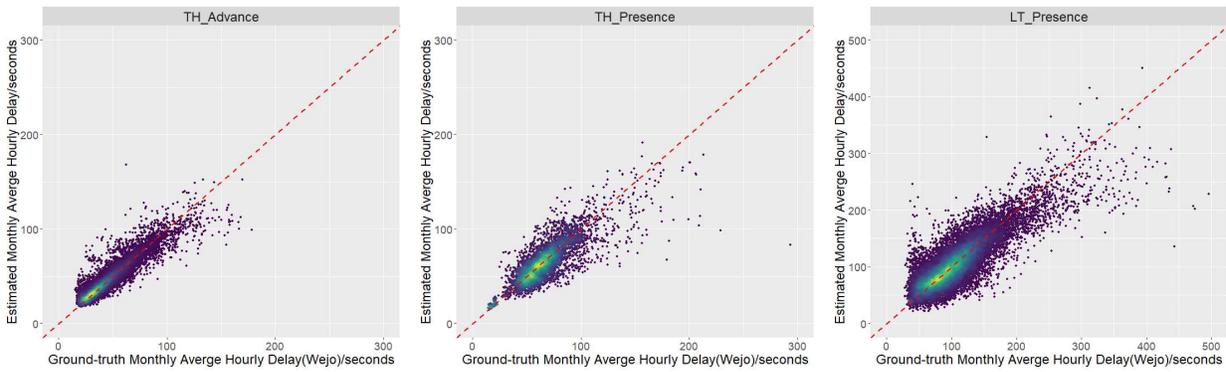
System	Movement and Detector	Performance Metrics		
		MAE	RMSE	MAPE
MaxView	Through (Advance Detectors)	6s (5s)	9s (7s)	13% (4%)
	Through (Presence Detectors)	17s (9s)	22s (13s)	19% (5%)
	Left-turn (Presence Detectors)	32s (12s)	44s (19s)	23% (4%)
Miovision	Through (Advance + Presence Detectors)	6s (5s)	7s (7s)	13% (4%)
	Left-turn (Presence Detectors)	21s (8s)	29s (11s)	20% (5%)

Table 7-4 summarizes the evaluation results for AOG ratio estimation of through and left-turn movements using the data collected through the MaxView and Miovision systems. The median MAPE for through movements with advance detectors is 18% and 16% for the MaxView and Miovision systems, respectively. Most locations have a MAPE of lower than 30% for through movements with advance detectors. However, left-turn movements and through movements with presence detectors have relatively higher errors, and most locations have a MAPE of higher than 40%.

Table 7-4. Method performance metrics for hourly AOG ratio estimation

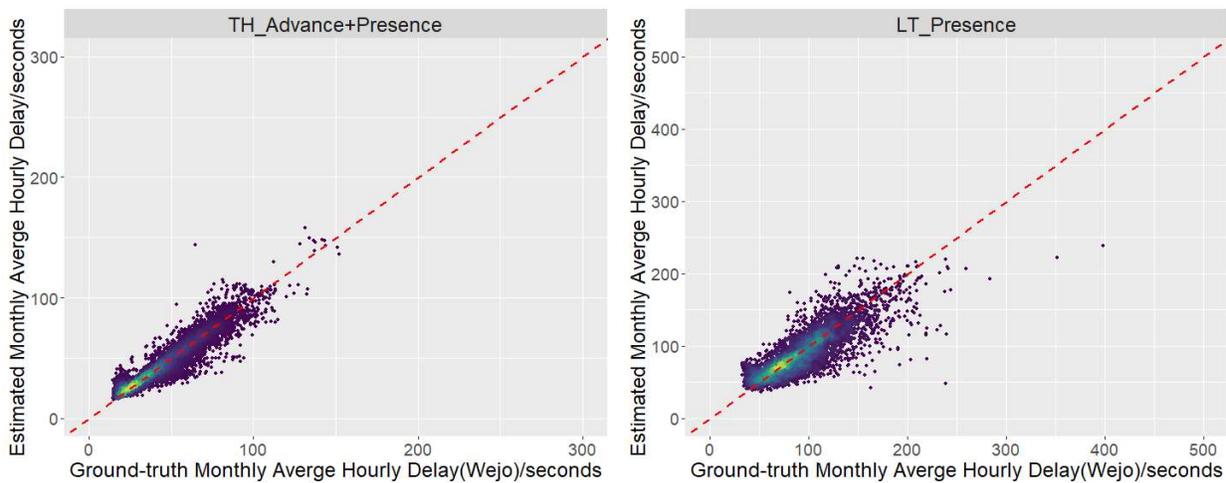
System	Movement and Detector	Performance Metrics		
		MAE	RMSE	MAPE
MaxView	Through (Advance Detectors)	9% (2%)	11% (3%)	18% (13%)
	Through (Presence Detectors)	11% (2%)	13% (3%)	41% (15%)
	Left-turn (Presence Detectors)	13% (4%)	16% (4%)	47% (12%)
Miovision	Through (Advance + Presence Detectors)	8% (2%)	11% (2%)	16% (14%)
	Left-turn (Presence Detectors)	14% (3%)	18% (4%)	47% (9%)

In addition to evaluating the method accuracy for hourly performance measures estimation, the hourly data is then aggregated into a month for calculating the monthly average hourly performance measures, and the method performance is accordingly evaluated. **Figure 7-1** shows the comparison between the estimated monthly average hourly control delay, using the data collected via the MaxView system, and the ground-truth control delay. All data is categorized into three groups according to the movement and detector layout for comparison, and the comparison results illustrate most of the data aligns closely with the dashed red line, indicating a close match between the estimated results and the ground-truth data. Among these three groups, the through movements with advance detectors have the highest accuracy, with only a few outliers. However, the other two groups with presence detectors have relatively lower accuracy, with more deviating significantly from the dashed red line. **Figure 7-2** shows the comparison between estimated values, using the data collected via the Miovision system, and ground-truth values. Similar to the MaxView system, most data points for both through and left-turn movements align closely with the dashed red line, indicating that the data collected by the Miovision system can also provide accurate and reliable monthly average hourly control delay. In addition, fewer outliers are observed for the left-turn movements in the Miovision system than the MaxView system. This difference can be attributed to the virtual presence detectors configured in the Miovision system, which have a shorter length and are more sensitive to left-turn vehicle arrivals, thereby providing more accurate information regarding vehicle arrivals.



(a) through movements with advance detectors (b) through movements with presence detectors (c) left-turn movements with presence detectors

Figure 7-1. Comparison between estimated and ground-truth monthly average hourly control delay using the MaxView system.



(a) through movements with advance and presence detectors (b) left-turn movements with presence detectors

Figure 7-2. Comparison between estimated and ground-truth monthly average hourly control delay using the Miovision system

Table 7-5 summarizes the method performance metrics for estimating monthly average hourly control delay using the data collected from the MaxView and Miovision systems. In comparison to the method performance for hourly control delay estimation, the method error for monthly average hourly control delay estimation is relatively lower for the two systems regardless of the metric employed. For example, the median MAPE for through movements with advance detectors,

through movements with presence detectors, and left-turn movements in the MaxView system are 9%, 13%, and 18%, respectively, and all are lower than 30%. In the Miovision system, the median MAPE for through movements and left-turn movements are 9% and 14%, respectively. Consistent with the observations in Figures 1(b) and 2(b), the data collected by the Miovision system demonstrates more accurate estimations of left-turn control delay compared to the MaxView system.

Table 7-5. Method performance metrics for monthly average hourly control delay estimation

System	Movement and Detector	Performance Metrics		
		MAE	RMSE	MAPE
MaxView	Through (Advance Detectors)	4s (3s)	5s (3s)	9% (4%)
	Through (Presence Detectors)	19s (5s)	12s (8s)	13% (5%)
	Left-turn (Presence Detectors)	18s (5s)	23s (6s)	18% (5%)
Miovision	Through (Advance + Presence Detectors)	3s (2s)	4s (3s)	9% (3%)
	Left-turn (Presence Detectors)	11s (6s)	15s (7s)	14% (5%)

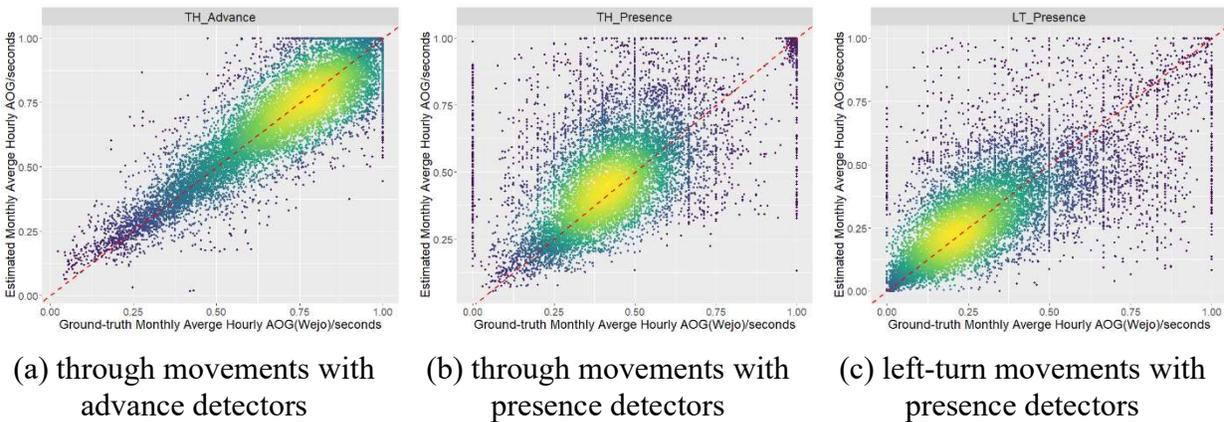
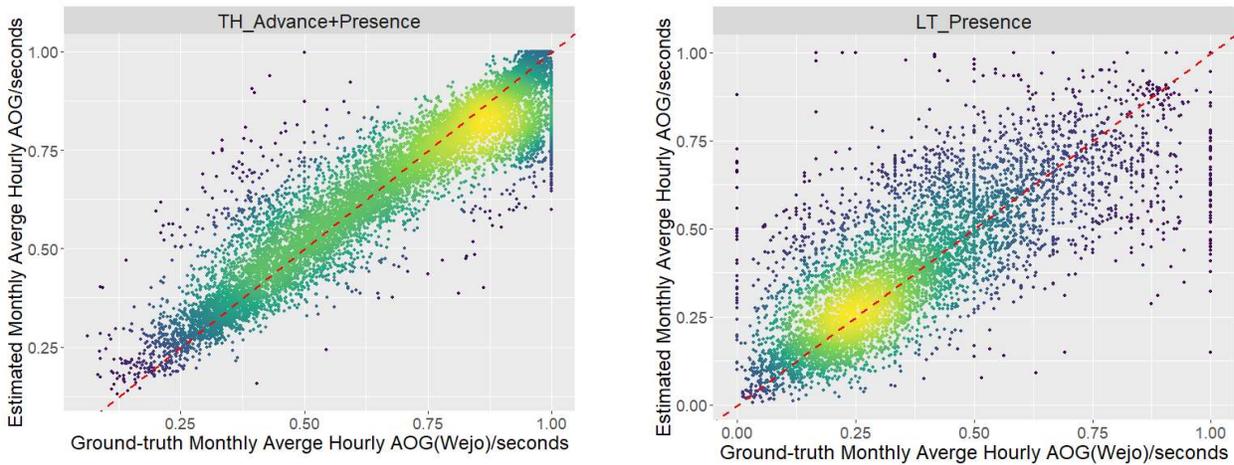


Figure 7-3. Comparison between estimated and ground-truth monthly average hourly AOG using the MaxView system

Figure 7-3 illustrates the comparison between the estimated monthly average hourly AOG, using the MaxView system, and ground-truth data. For through movements with advance detectors configured, most data points closely align with the dashed red line, indicating the estimated values match ground-truth values. For the locations with presence detectors, even though most data points are located near the dashed red line, more outliers deviate significantly from the dashed red line in comparison to the locations with advance detectors. **Figure 7-4** compares the estimated monthly average hourly AOG, using the data provided by the Miovision system, and the ground-truth values. Similar to the results when using the MaxView system, most data points for the through movement closely align with the dashed red line, indicating the data collected by the Miovision system can

provide accurate monthly average hourly AOG for through movements. For left-turn movements, while most data points are close to the dashed red line, numerous outliers deviate considerably from the line.



(a) through movements with advance and presence detectors (b) left-turn movements with presence detectors

Figure 7-4. Comparison between estimated and ground-truth monthly average hourly AOG using the Miovision system.

Table 7-6 summarizes the method performance metrics for monthly average hourly AOG ratio estimation using the data collected via the MaxView and Miovision systems. In comparison to hourly AOG estimation shown in **Table 7-4**, the monthly average AOG has a significantly lower error, especially for through movements. Specifically, for the through movements with advance and presence detectors in the MaxView system, the median MAPE drops from 18% and 41% (from **Table 7-4**) to 10% and 24%, respectively. The median MAPE for through movements in the Miovision system also decreases from 16% (from **Table 7-4**) to 7%. However, the method performance for the left-turn movements experiences insignificant improvement after aggregating the hourly AOG to a monthly level. In addition, the standard deviation of MAPE for the locations with presence detectors increases, indicating a higher variability in the estimation accuracy.

Table 7-6. Method performance metrics for monthly average hourly AOG estimation

System	Movement and Detector	Performance Metrics		
		MAE	RMSE	MAPE
MaxView	Through (Advance Detectors)	6% (4%)	8% (3%)	10% (5%)
	Through (Presence Detectors)	8% (5%)	11% (5%)	24% (27%)
	Left-turn (Presence Detectors)	10% (2%)	15% (3%)	41% (21%)
Miovision	Through (Advance + Presence Detectors)	5% (1%)	7% (2%)	7% (6%)
	Left-turn (Presence Detectors)	10% (3%)	14% (4%)	31% (25%)

7.2.3 Acceptance Criteria

According to the summary of the relevant literature and data assessment results, the preliminary acceptance criteria for different traffic performance measures, traffic movements, detector layouts, and systems are developed, as summarized in **Tables 7-7 to 7-10**. Considering that previous relevant studies used few study locations and simulation data for calculating the performance metrics, the criteria should be relatively less stringent than the assessment results using the real-world data collected from an entire region. Therefore, the acceptance criteria are set relatively higher than the performance metrics reported in previous studies to ensure the suitability and applicability of the criteria for real-world and region-wide applications.

Table 7-7. Preliminary Acceptance Criteria for Estimated Hourly Control Delay

System	Movement and Detector	Acceptance Criteria		
		MAE	RMSE	MAPE
MaxView	Through (Advance Detectors)	10s	15s	20%
	Through (Presence Detectors)	30s	30s	25%
	Left-turn (Presence Detectors)	30s	50s	30%
Miovision	Through (Advance + Presence Detectors)	10s	15s	20%
	Left-turn (Presence Detectors)	30s	50s	30%

Table 7-8. Preliminary Acceptance Criteria for Estimated Hourly AOG Ratio

System	Movement and Detector	Acceptance Criteria		
		MAE	RMSE	MAPE
MaxView	Through (Advance Detectors)	15%	15%	30%
	Through (Presence Detectors)	20%	20%	40%
	Left-turn (Presence Detectors)	20%	20%	40%
Miovision	Through (Advance + Presence Detectors)	15%	15%	30%
	Left-turn (Presence Detectors)	20%	20%	40%

Table 7-9. Preliminary Acceptance Criteria for Estimated Monthly Average Hourly Control Delay

System	Movement and Detector	Acceptance Criteria		
		MAE	RMSE	MAPE
MaxView	Through (Advance Detectors)	10s	10s	10%
	Through (Presence Detectors)	20s	20s	20%
	Left-turn (Presence Detectors)	25s	25s	25%
Miovision	Through (Advance + Presence Detectors)	10s	10s	10%
	Left-turn (Presence Detectors)	25s	25s	25%

Table 7-10. Preliminary Acceptance Criteria for Estimated Monthly Average AOG Ratio

System	Movement and Detector	Acceptance Criteria		
		MAE	RMSE	MAPE
MaxView	Through (Advance Detectors)	10%	10%	15%
	Through (Presence Detectors)	15%	15%	30%
	Left-turn (Presence Detectors)	15%	15%	30%
Miovision	Through (Advance + Presence Detectors)	10%	10%	15%
	Left-turn (Presence Detectors)	15%	15%	30%

7.2.4 Validation Result Summary

According to the developed acceptable criteria, the performance measures of the estimated hourly and monthly average hourly control delay for through and left-turn movements meet all the acceptance criteria, which would be used to control the quality of any other data. This indicates that the estimated control delay values are within the acceptable error range for accurate and reliable performance assessment and will be used for further application. For AOG ratio estimation, the through movements with the advance detector have a lower error in both hourly and monthly average hourly estimation than the three acceptance criteria. However, most left-turn locations and the through locations with presence detectors in both systems have a higher MAPE in hourly and monthly average hourly estimation than the acceptance criteria. In terms of MAE and RMSE, most locations with presence detectors can meet the acceptance criteria, because these locations usually have a lower priority in signal coordination, leading to a low AOG ratio. Therefore, the estimated control delay for both through and left-turn movements is considered acceptable and can be reliably used for real-time and long-term traffic studies. The estimated AOG ratio for through movements with advance detectors has an accepted error, but most locations with presence detectors have an error exceeding the acceptable criteria. When applying the estimated AOG ratio for traffic studies, the locations with presence detectors might provide less reliable results, which require further analysis and evaluation.

In addition to calculating control delay and AOG ratio, the Wejo data has been used to extract traffic signal split failure in previous studies. In this project, the split failure is not estimated using the event-based data due to the lack of reliable ground-truth data, so the associated acceptance criteria have not been developed. Even though the split failure has been extracted from the Wejo data using the method proposed by a previous study (E. Saldivar-Carranza et al., 2021a), the relationship between Wejo-based split failure and sensor-based split failure did not show a strong correlation regardless of the sample size of Wejo data due to a large amount of zero split failure. Consequently, the Wejo data cannot serve as a reliable source for determining ground-truth data for the split failure estimation method. Therefore, the estimation of split failure based on Wejo data is not feasible in this project, and alternative approaches should be explored for obtaining accurate and reliable ground-truth split failure data.

CHAPTER 8: REGION-WIDE TRAFFIC MOBILITY/RELIABILITY PERFORMANCE ESTIMATION

Based on the estimation method developed in previous tasks, this task will focus on estimating and analyzing traffic mobility/reliability performance measures including control delay and AoG at signalized intersections using event-based data from MaxView and Miovision within the PAG region, subject to QA/QC. The study intersections selected will depend on the communication over the regional communication network, which includes the MaxView and Miovision systems.

8.1 DATA COLLECTION

The estimation of performance measures heavily relies on event-based data collected through the Miovision and MaxView systems. The availability of event-based is dependent on the controller's communication quality. For example, poor communication quality can lead to missing data whereas traffic controllers with reliable communication can provide high-quality event-based data.

Figure 8-1 illustrates the number of intersections with event-based data available daily throughout 2021. On average, approximately 300 signalized intersections in the PAG region had event-based data available. The Miovision system demonstrates reliable and robust data quality, with 100 out of 110 signal controllers consistently providing event-based data on most days. Only a few days in June experienced some missing data. The MaxView system had event-based data available for around 200 intersections out of a total of 500 intersections, with a slight variation observed across different days of the year. Data missing during November was because the UA team did not back up the data for those specific days and was not due to communication issues.

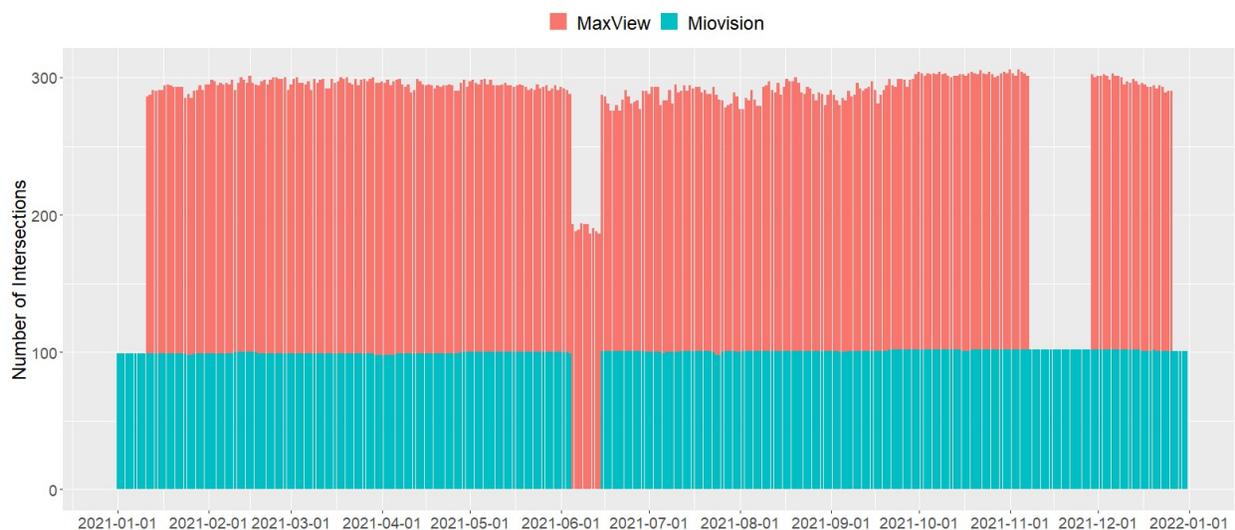


Figure 8-1. The number of intersections with available event-based data

8.2 REGION-WIDE CONTROL DELAY ESTIMATION AND ANALYSIS

8.2.1 Hourly Control Delay

The event-based data is utilized to estimate the left-turn and through control delay at all study intersections on a 24/7 basis. Taking Speedway Boulevard & Euclid Avenue as an example, **Figure 8-2** shows three days of estimated left-turn and through control delay for each of the four road directions. This estimated control delay shows a logical and clear traffic pattern, with a low delay during nighttime and a high delay during peak periods. In addition, four directions have a left-turn delay significantly higher than through delay, which is consistent with real-world observations.

These findings are consistent with existing knowledge in the field, indicating that the estimated control delay serves as a reliable and accurate measure for quantifying traffic patterns and conditions.

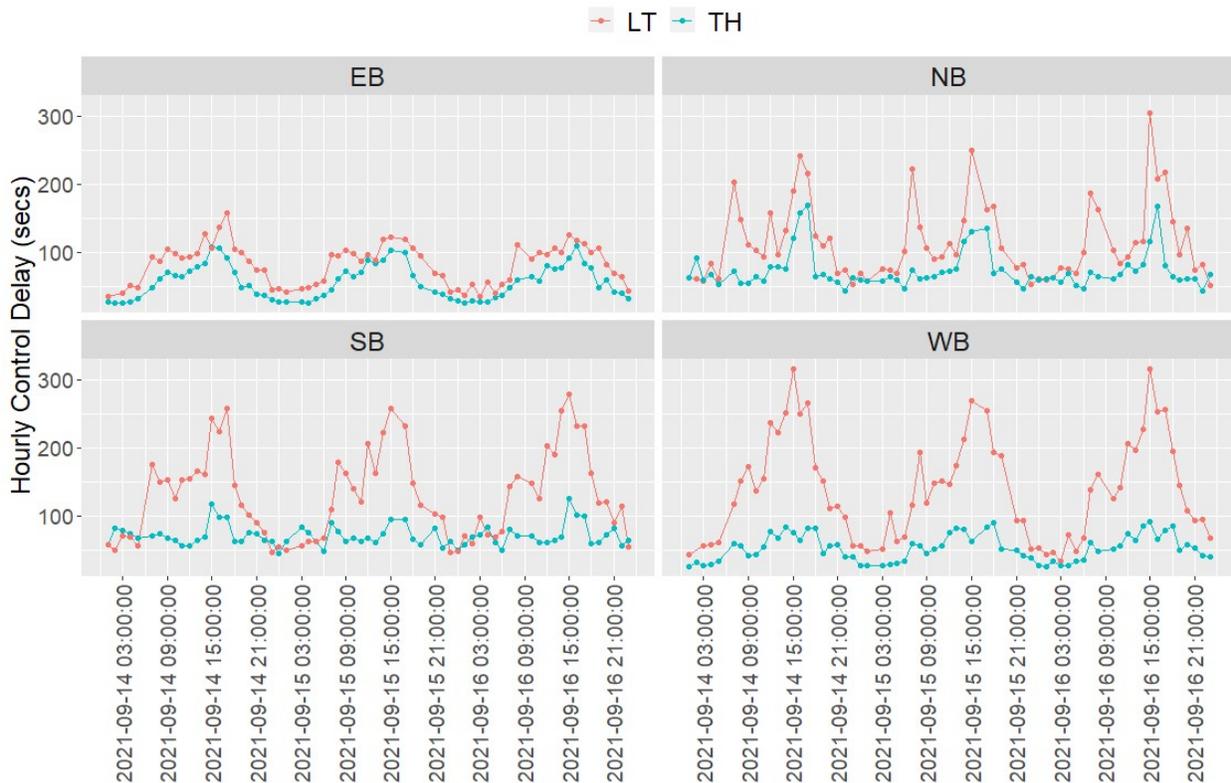


Figure 8-2. Estimated hourly control delay at Speedway Boulevard & Euclid Avenue

The estimated control delay obtained from all study intersections is used to analyze the distribution of delay by the hour of the day, as depicted in **Figure 8-3**. Both left-turn and through delay show a similar temporal trend. Before 5 a.m., the distribution displays relatively low median values and variance, which implies that the control delay is generally low. The distributions during the

daytime are similar, which is likely because traffic patterns in the region are relatively constant throughout the daytime hours. After 7 a.m., the median and variance of the distributions increase, indicating a rise in control delay that varies across different locations. Comparing left-turn delay and through delay, left-turn movements consistently experience higher delays than through movements throughout the entire 24-hour period. The control delay for most through movements remains below 50 seconds, while most left-turn movements exhibit delays exceeding 75 seconds during daytime hours.

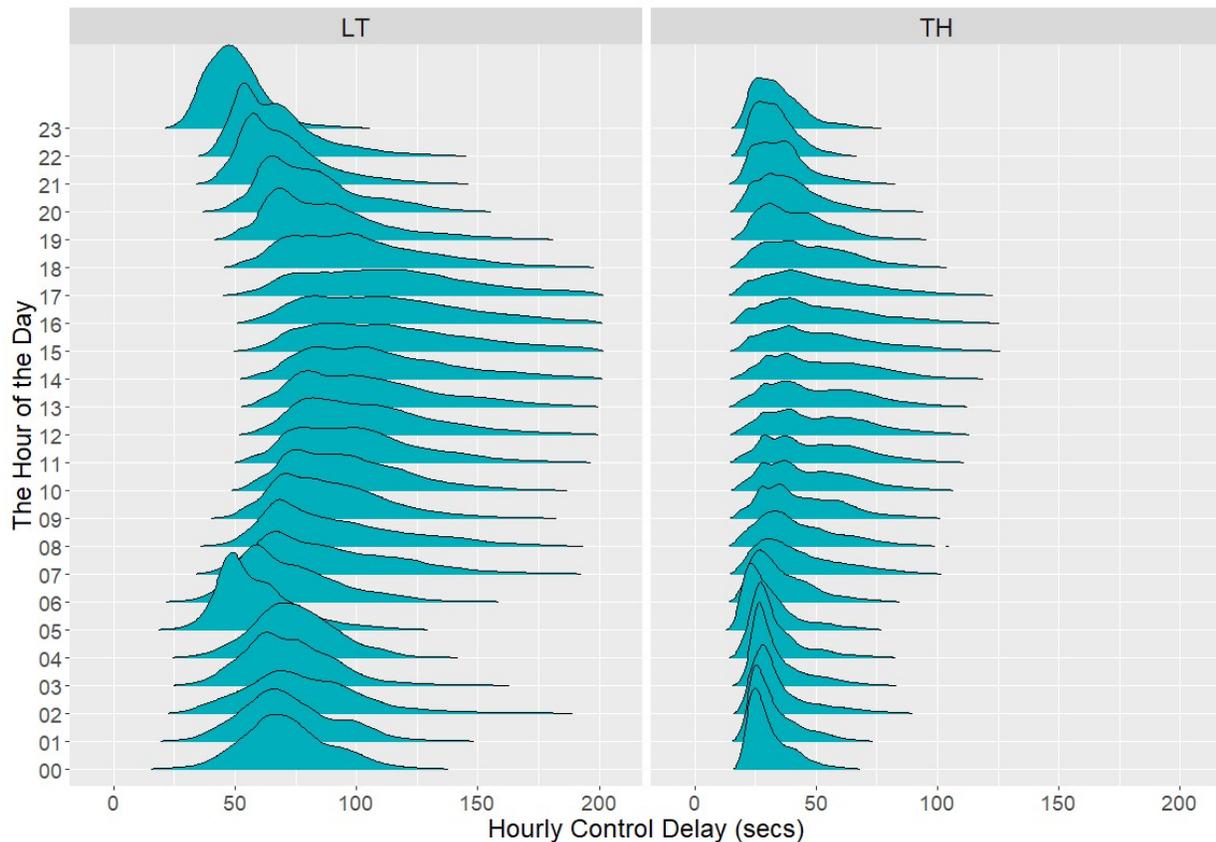
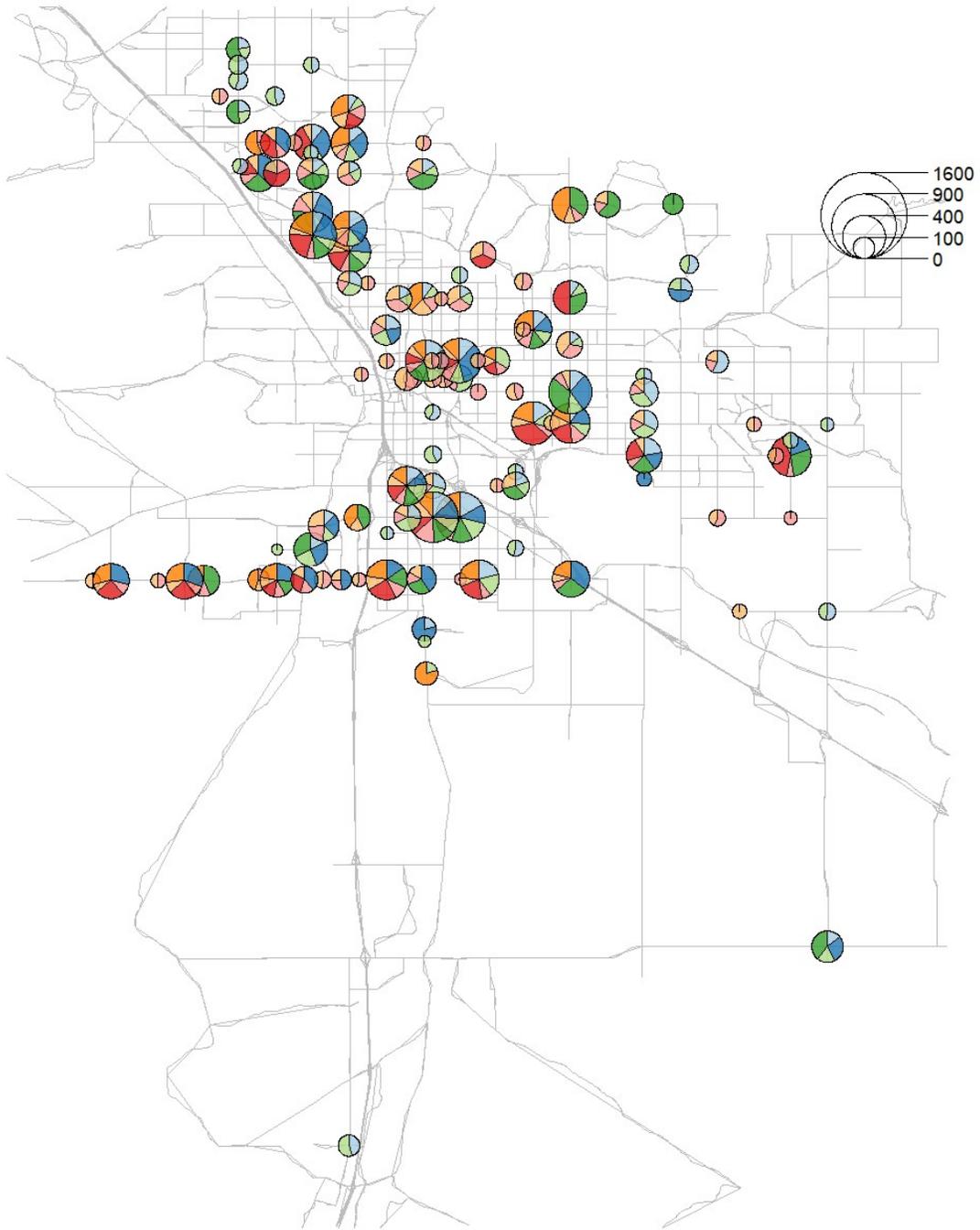
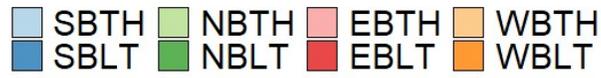


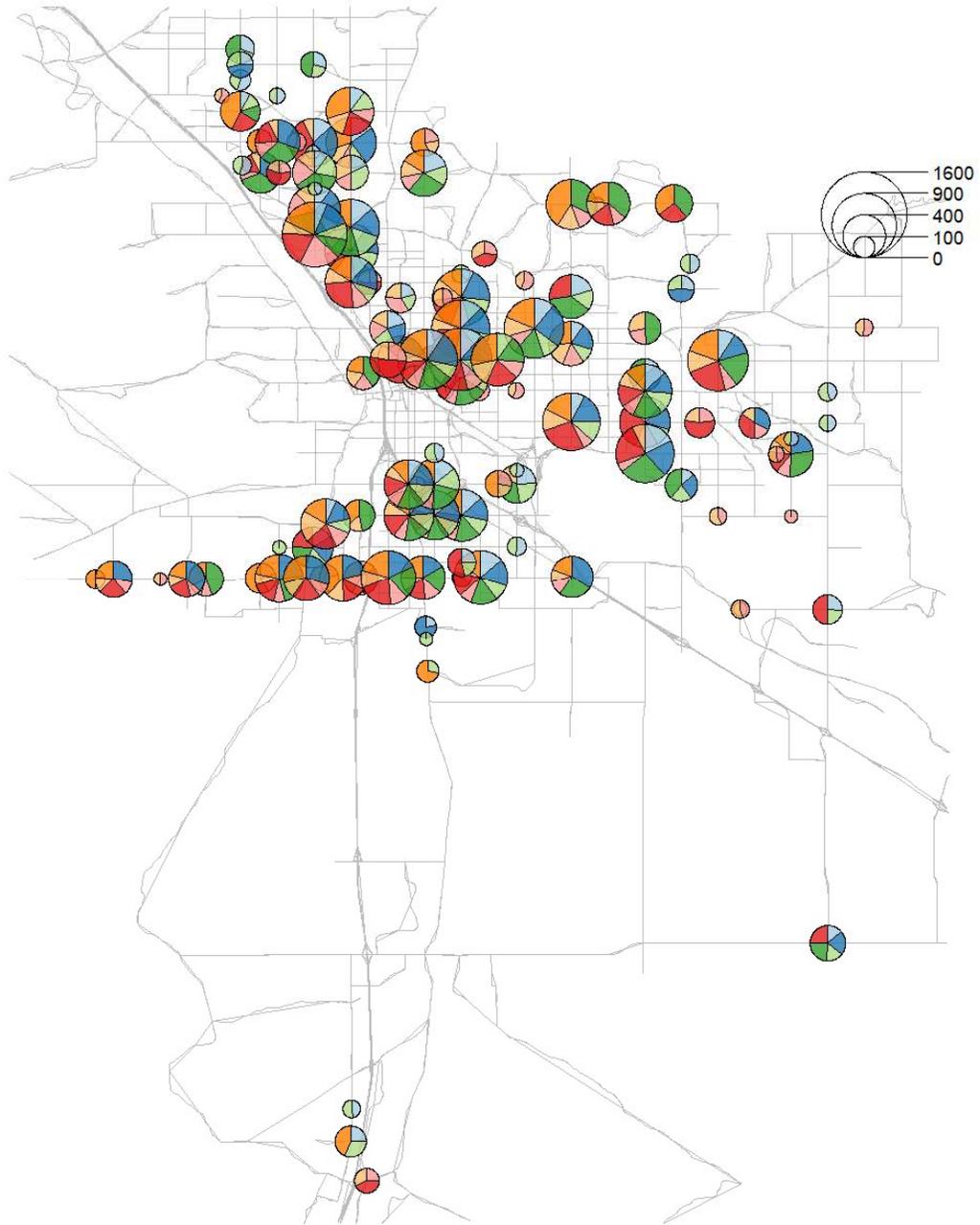
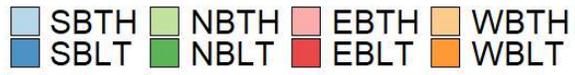
Figure 8-3. Estimated control delay distribution by hour of the day

In addition to temporally analyzing the estimated control delay, **Figure 8-4** illustrates the spatial distribution of control delay at different hours. Each color in the figure represents the control delay for a specific traffic movement in a particular road direction, while the radius of the circle indicates the total control delay for both through and left-turn movements at a given signalized intersection. **Figure 8-4 (a)** shows the spatial distribution at 1 a.m., during which most intersections exhibit low delay due to light traffic volume. However, there are still some intersections experiencing relatively high delay, primarily attributed to high left-turn delay. At 8 a.m., numerous locations, particularly along major corridors like Speedway Blvd. and Ajo Way, have a high delay, as shown in **Figure 8-4(b)**. **Figure 8-4(c)** demonstrates the spatial distribution at 5 p.m., which exhibits a similar pattern to 8 a.m.



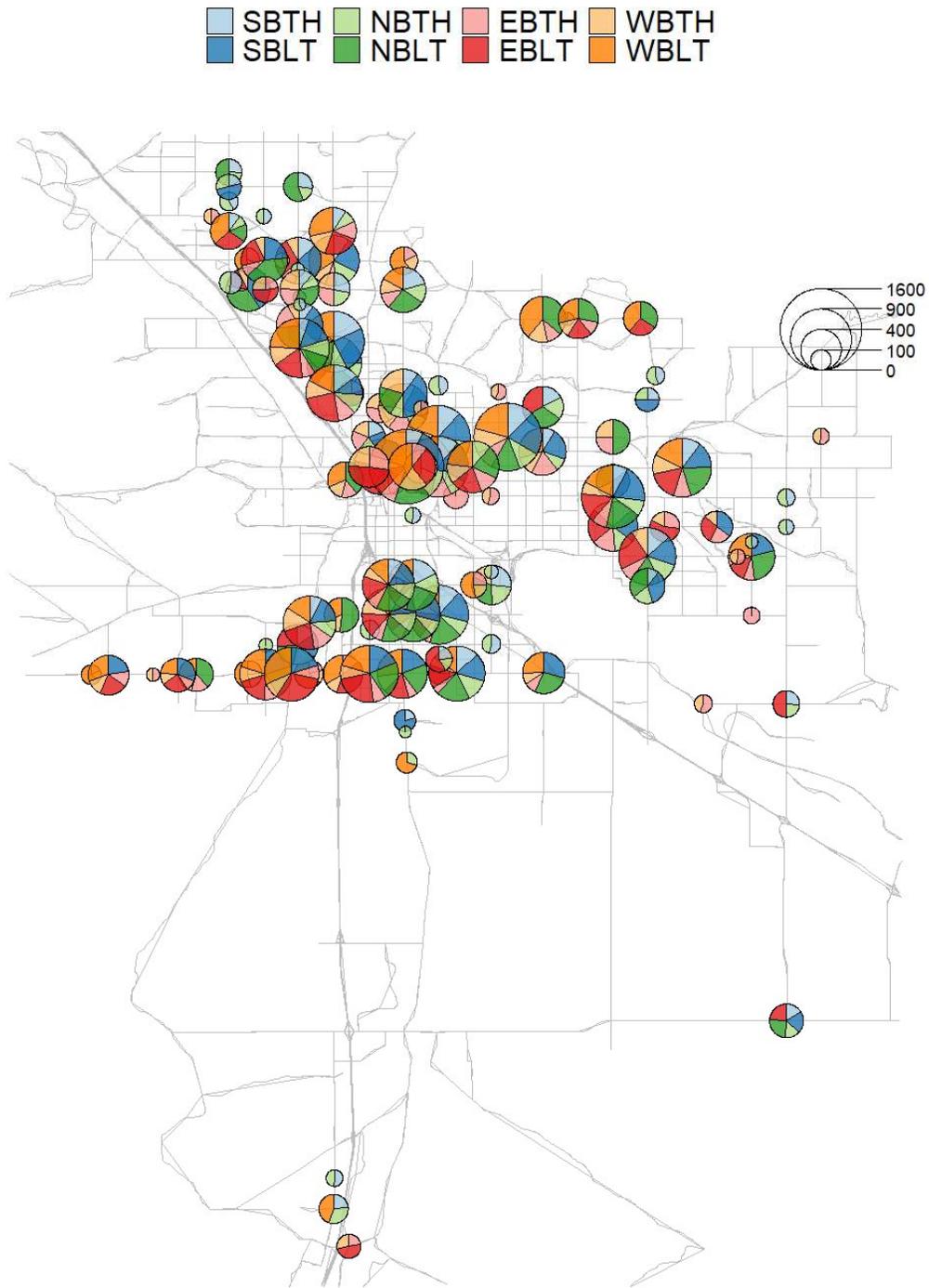
(a) Estimated control delay at 1a.m.

Figure 8-4. Spatial distribution of estimated control delay on Sep. 14, 2021



(c) estimated control delay at 8 a.m.

Figure 8-5. Spatial distribution of estimated control delay on Sep. 14, 2021



(c) estimated control delay at 5 p.m.

Figure 8-6. Spatial distribution of estimated control delay on Sep. 14, 2021

8.2.2 Monthly Average Hourly Control Delay

The estimated hourly control delay data is further aggregated to a monthly level to estimate the average hourly control delay for each month. **Figure 8-5** provides an example using data collected from Speedway Blvd. & Euclid Ave. in September 2021. The monthly average estimated hourly control delay exhibits clear traffic patterns for both left-turn and through traffic. Although the aggregated data follows a similar pattern as the hourly control delay, it presents a smoother trend due to the averaging process. According to the aggregated control delay, some directions have clear peak hours; for example, NB has its a.m. peak at 8 a.m. and PM peak from 4-6 p.m. In contrast, some directions exhibit similar values of hourly control delay throughout the daytime, such as the EB.

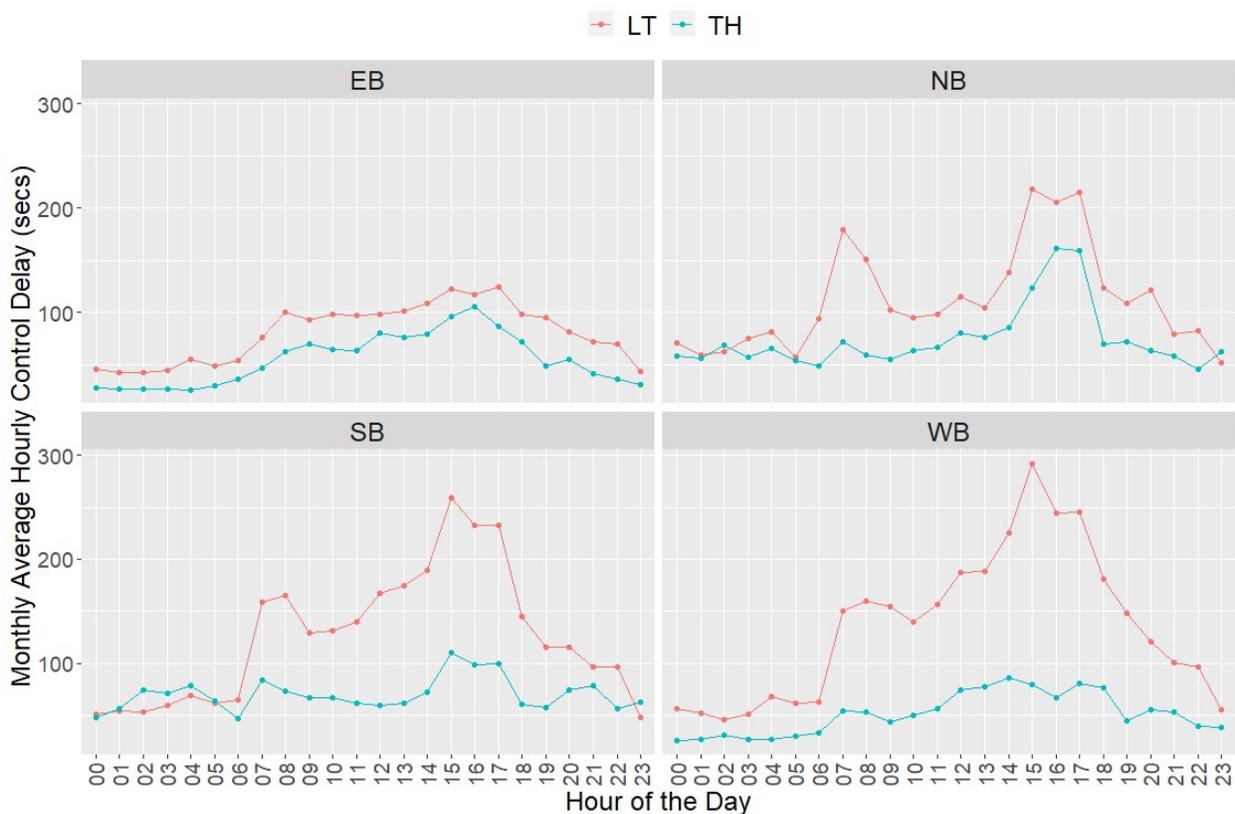
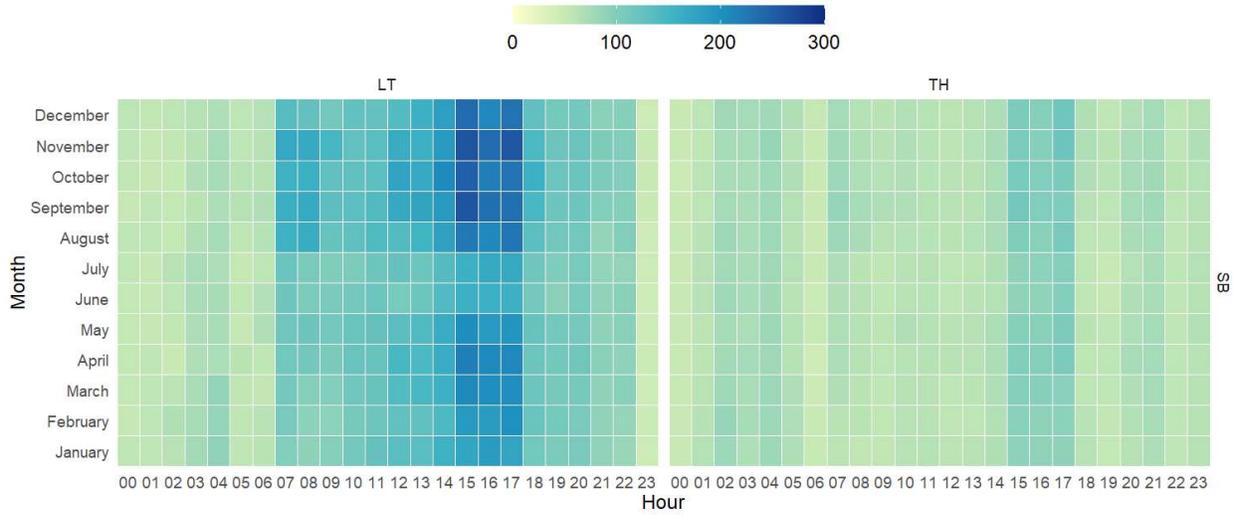


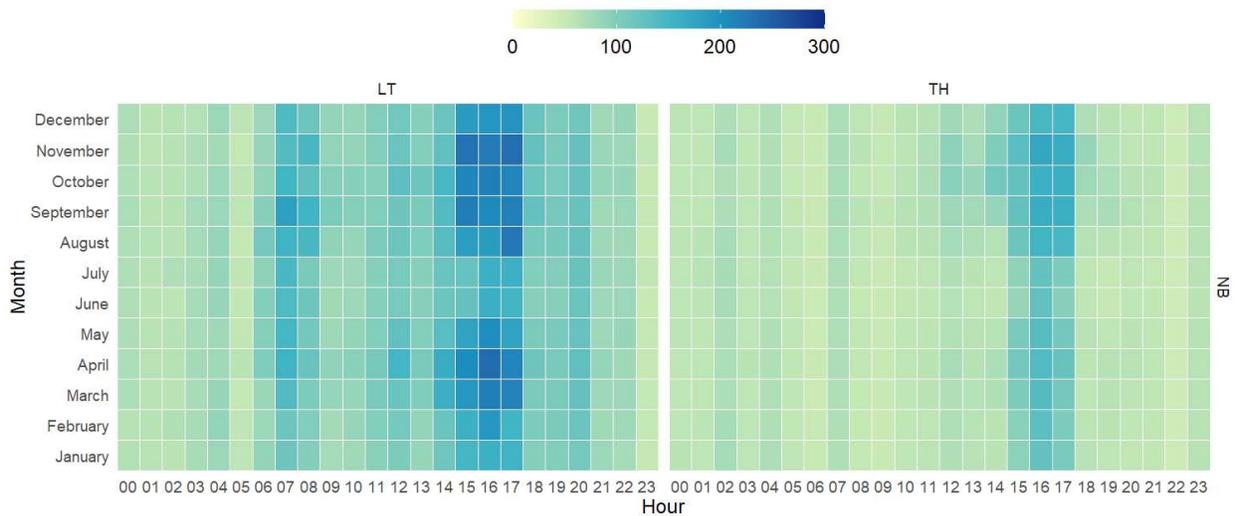
Figure 8-7. Monthly average hourly control delay at Speedway Blvd. & Euclid Ave. in September 2021

In addition to using September's data, the aggregated control delay from 12 months is utilized to depict the variations in traffic conditions across different months. As shown in **Figure 8-6**, the heatmaps illustrate the monthly average hourly control delay at 24 hours and 12 months for thorough and left-turn traffic in four road directions at Speedway Blvd. & Euclid Ave. In terms of the month dimension, the control delay during June and July is lower compared to other months,

particularly during daytime hours and peak periods. This decrease in delay can be attributed to lighter traffic volumes experienced during the summer break. The heatmaps provide valuable insights into the seasonal fluctuations in traffic conditions and highlight the impact of factors such as vacation periods on congestion levels.

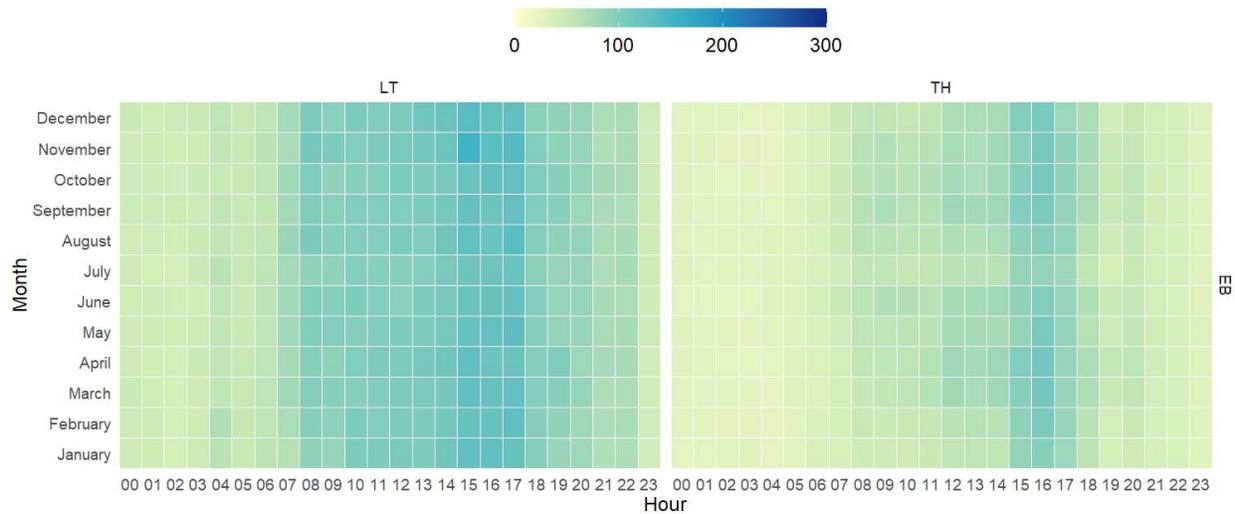


(a) Southbound

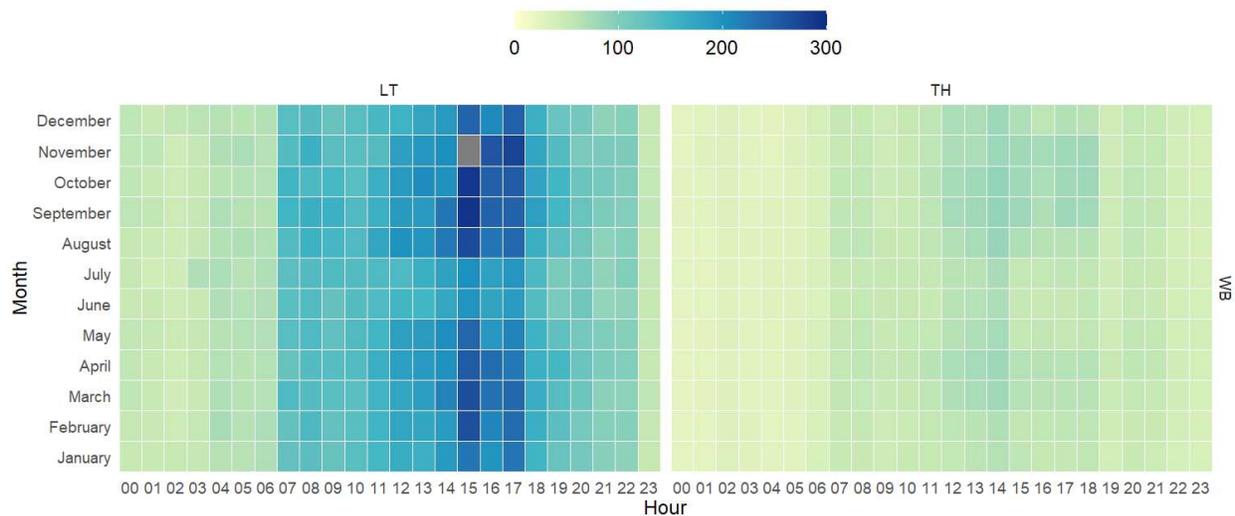


(b) Northbound

Figure 8-8. Heatmap of monthly average estimated control delay at Speedway Blvd. & Euclid Ave.



(c) Eastbound



(d) Westbound

Figure 8-9. Heatmap of monthly average estimated control delay at Speedway Blvd. & Euclid Ave.

8.3 REGION-WIDE AOG RATIO ESTIMATION AND ANALYSIS

8.3.1 Hourly AoG Ratio

The event-based data obtained from the MaxView and Miovision systems is utilized to estimate the 24/7 AoG ratio at all intersections within the PAG region. **Figure 8-7** takes Speedway Blvd & Euclid Ave. as one example to illustrate the temporal trend of AoG ratio specifically for through movements. The reason that the AoG ratio is estimated only for through movements is that some left-turn movements do not meet the acceptance criteria established in the previous section of the analysis. All directions except SB have a clear and similar pattern with low estimated AoG in peak hours and high estimated AoG at nighttime. The specific peak hours vary across directions. For instance, the EB direction peak hours span the entire daytime, while the NB direction peak hours occur from 3-6 p.m. In the case of the SB direction, there is no clear pattern due to the approach segment being a minor road with lower traffic volume and random vehicle arrivals.

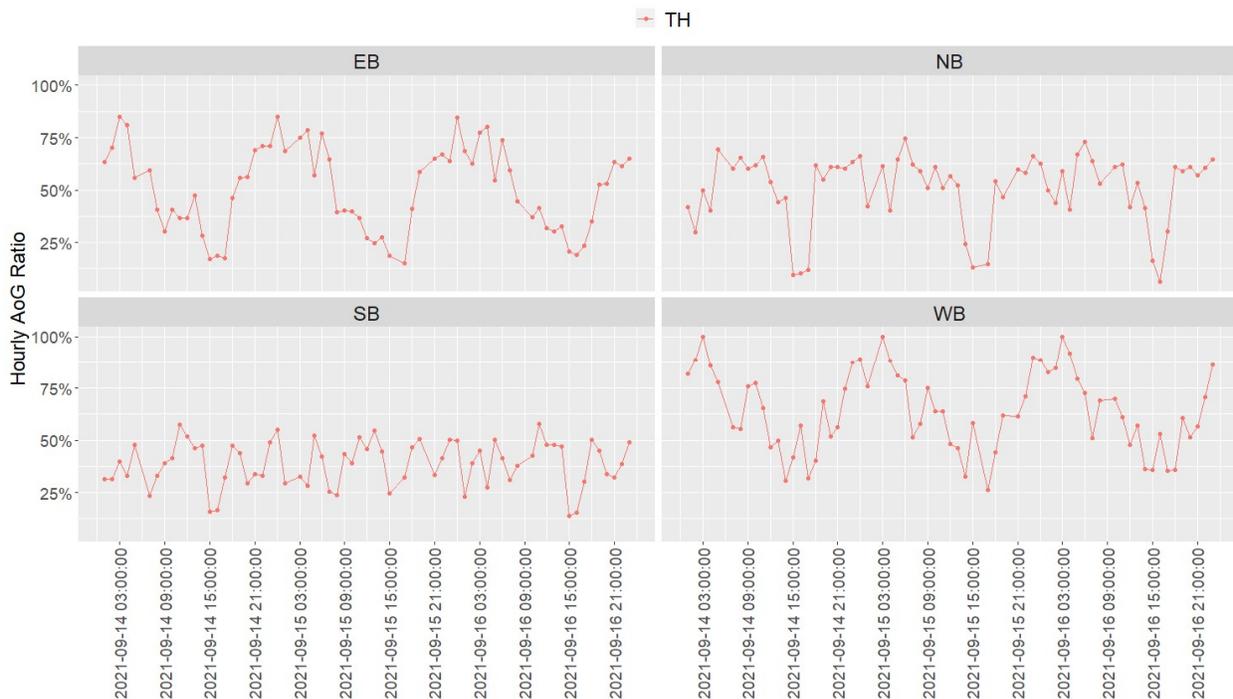


Figure 8-10. Estimated hourly AoG ratio at Speedway Blvd. & Euclid Ave.

Furthermore, the estimated AoG ratio obtained from all study intersections is utilized to generate the density distribution by the hour of the day, as shown in **Figure 8-8**. During nighttime hours, the estimated AoG ratio tends to be relatively high, with the median values of these distributions exceeding 75%. This indicates light traffic conditions during these periods, with vehicles experiencing less delay at signalized intersections. In addition, another distribution peak with a mean of 100% is observed during nighttime, indicating that some locations experience free-flow

traffic conditions during late night and early morning hours, where all vehicles can smoothly proceed through intersections without stops. However, during daytime hours, the median values of the distribution decrease, suggesting increased congestion and higher likelihood to stop at signalized intersections. The variance of the distribution is also higher during daytime, indicating greater variability in the AoG ratio at different intersections and during peak traffic periods.

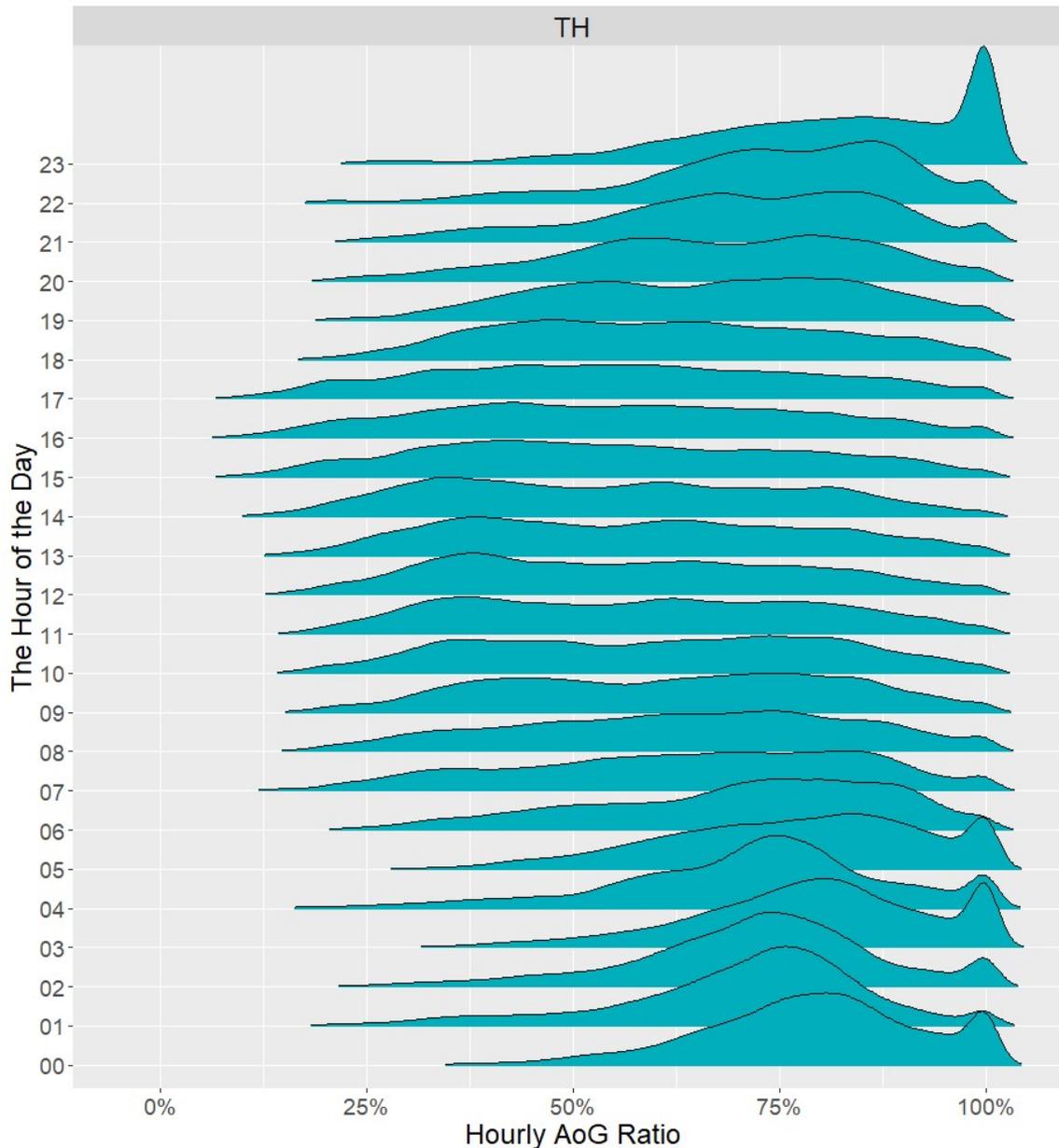
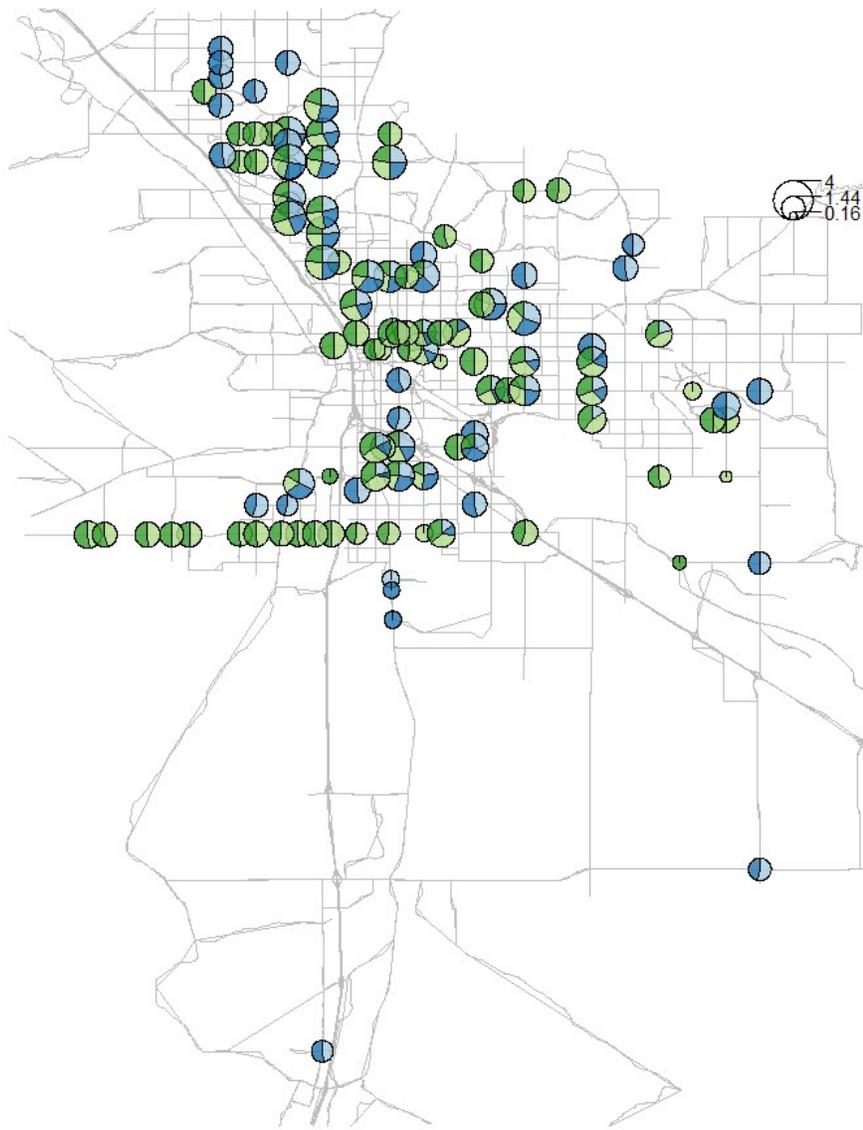


Figure 8-11. Estimated AoG ratio distribution by hour of the day

In addition to analyzing the estimated AoG ratio in a temporal dimension, **Figure 8-9** illustrates the spatial distribution at different hours on Sept. 14, 2021. At midnight, as shown in **Figure 8-9**

(a), most intersections have a relatively high estimated AoG ratio. Some intersections show a circle with a small radius, which does not mean these locations have a low AoG ratio, but rather that these locations such as intersections on Ajo Way only have one or two road directions with available event-based data or free-flow speed data caused by a very low volume. **Figures 8-9(b)** and 8-9(c) illustrate the spatial distribution of the AoG ratio during the morning and afternoon peak hours, respectively. These peak hours exhibit lower AoG ratios compared to the nighttime, indicating more traffic congestion and a higher likelihood that vehicles must stop at signalized intersections.

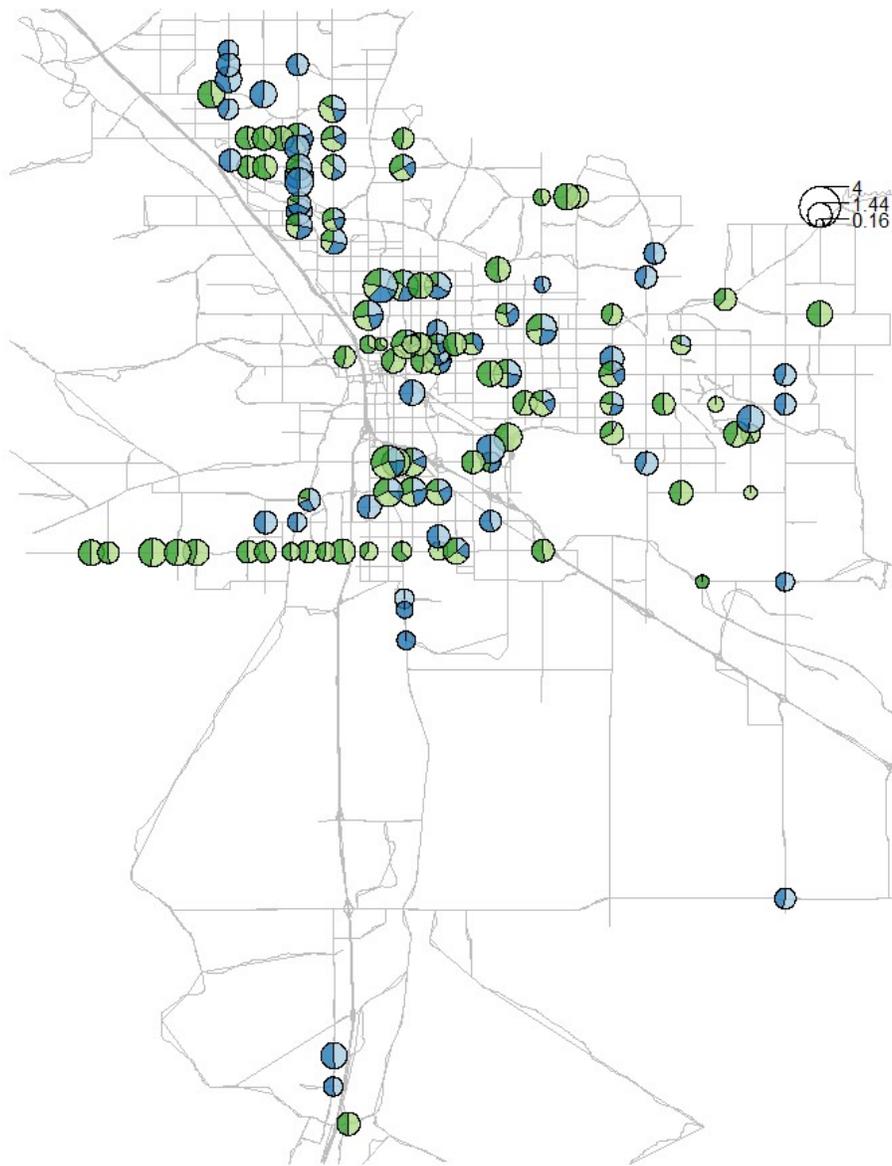
SBTH NBTH EBTH WBTH
01



(a) AoG ratio at 1am

Figure 8-12. Spatial distribution of AoG ratio on Sept. 14, 2021

SBTH NBTH EBTH WBTH
07

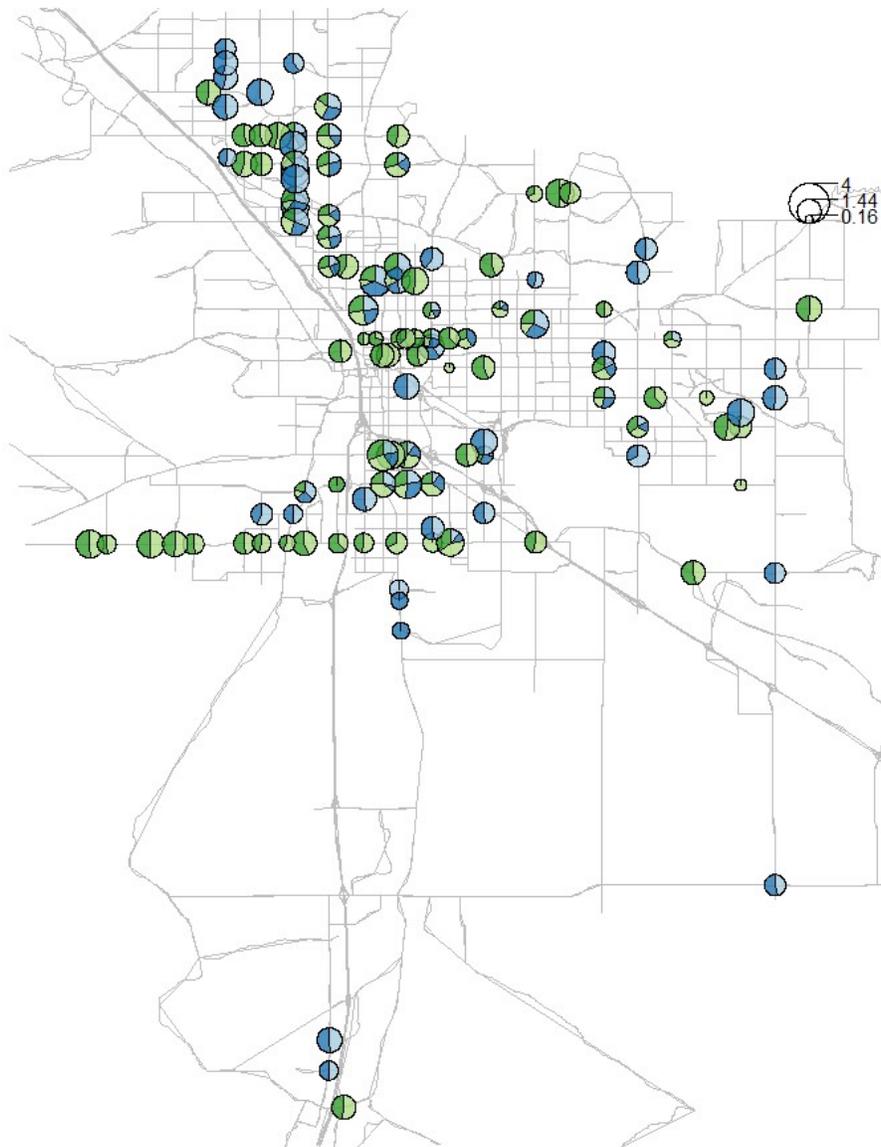


(b) AoG ratio at 7 a.m.

Figure 8-13. Spatial distribution of AoG ratio on Sept. 14, 2021

SBTH NBTH EBTH WBTH

17



(c) AoG ratio at 5 p.m.

Figure 8-14. Spatial distribution of AoG ratio on Sept. 14, 2021

8.3.2 Monthly Average Hourly AOG Ratio

The estimated hourly AoG ratio is aggregated for each month to estimate the monthly average hourly AoG ratio. Figure 8-10 illustrates the temporal trend of the monthly average estimated

hourly AoG ratio specifically at Speedway Blvd. & Euclid Ave. in Sept. 2021. Comparing the monthly average estimated hourly AoG ratio in **Figure 8-10** to the estimated hourly AoG ratio for the 3 days in **Figure 8-7**, it is evident that both display a similar pattern but with a smoother trend at the monthly level. The aggregation process provides a clearer and more consolidated view of the overall traffic conditions and the average level of congestion experienced at the intersection.

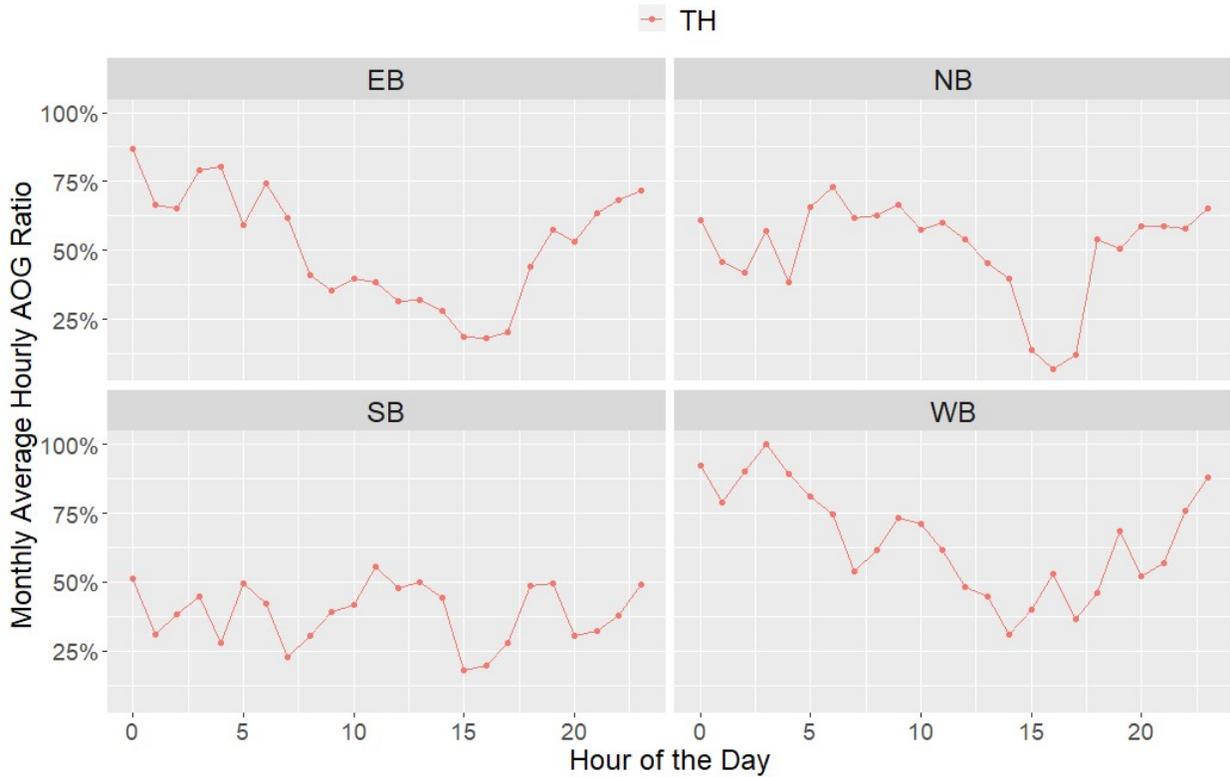
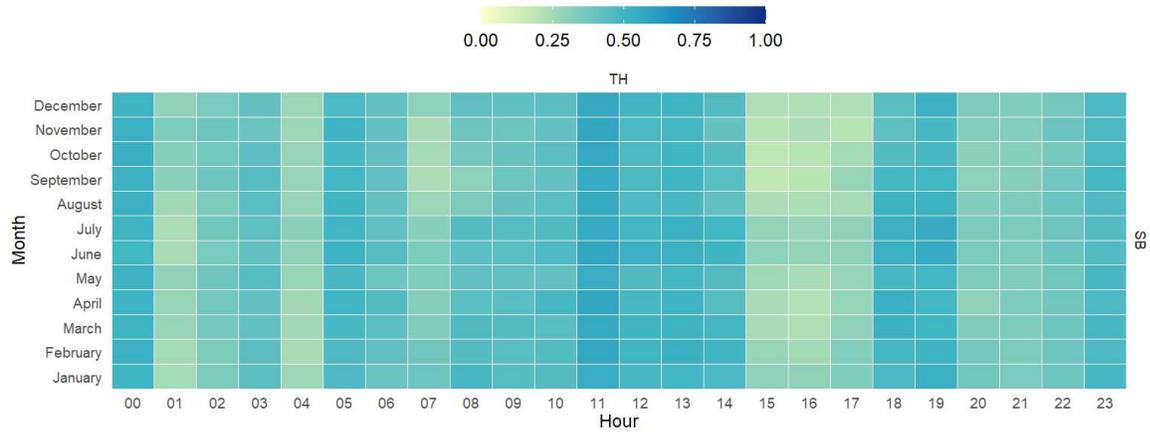
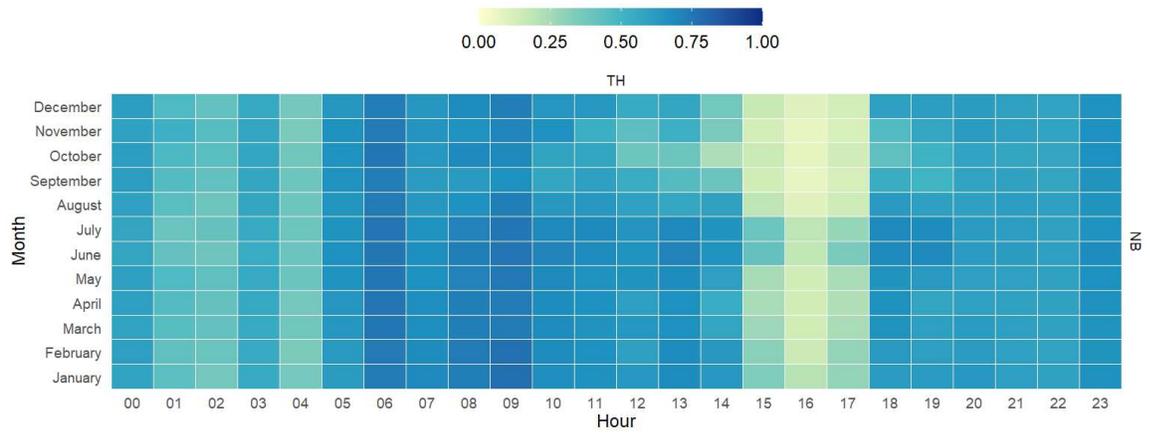


Figure 8-15. Monthly average estimated hourly AoG ratio at Speedway Blvd. & Euclid Ave. in Sept. 2021

Furthermore, the monthly average estimated AoG ratio at Speedway Blvd & Euclid Ave. is calculated for each of the 12 months to visualize the traffic conditions throughout the year. **Figure 8-11** shows the aggregated AoG ratios, revealing consistent patterns for many times of the day across different months. However, the AoG ratio from May to July is higher than in other months, especially during the afternoon peak, which can be attributed to the overall lower traffic volume during summer break.

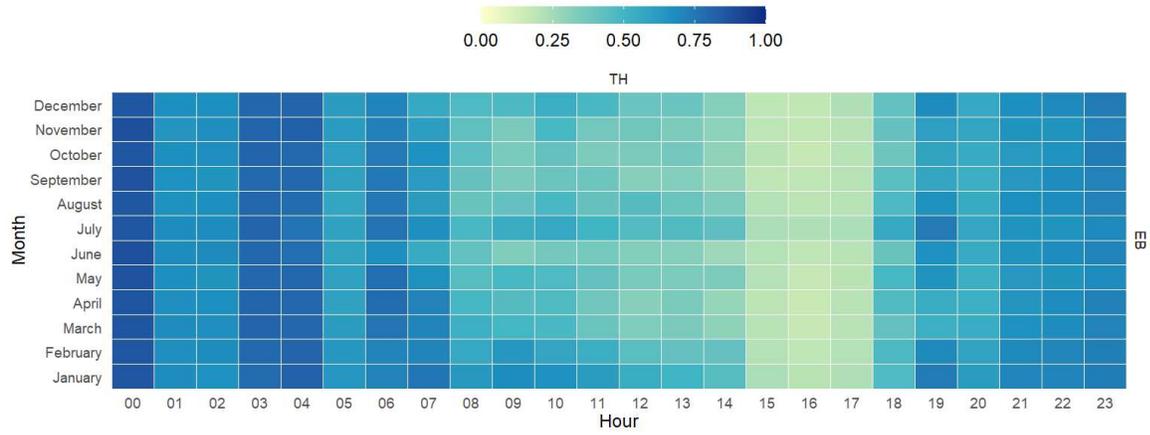


(a) Southbound

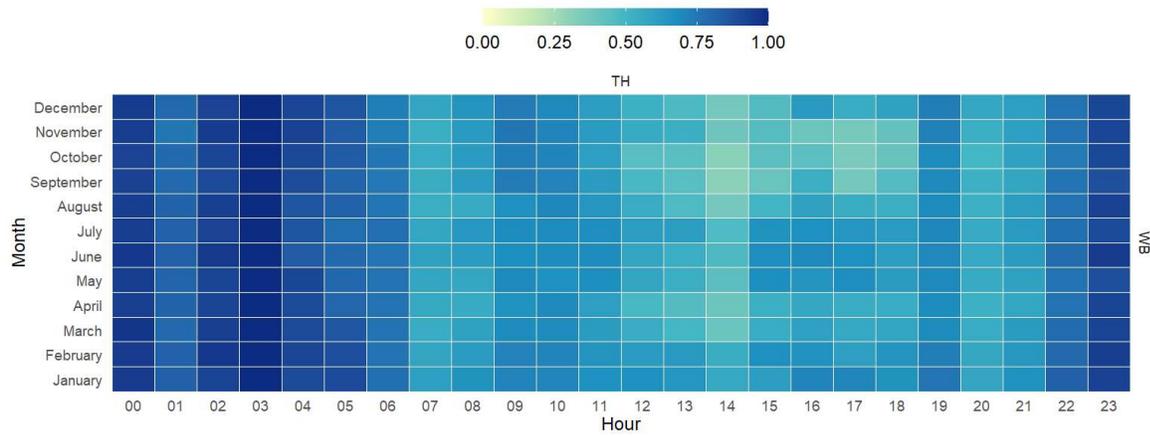


(b) Northbound

Figure 8-16. Heatmap of monthly average estimated AoG ratio at Speedway Blvd. & Euclid Ave.



(c) Eastbound



(d) Westbound

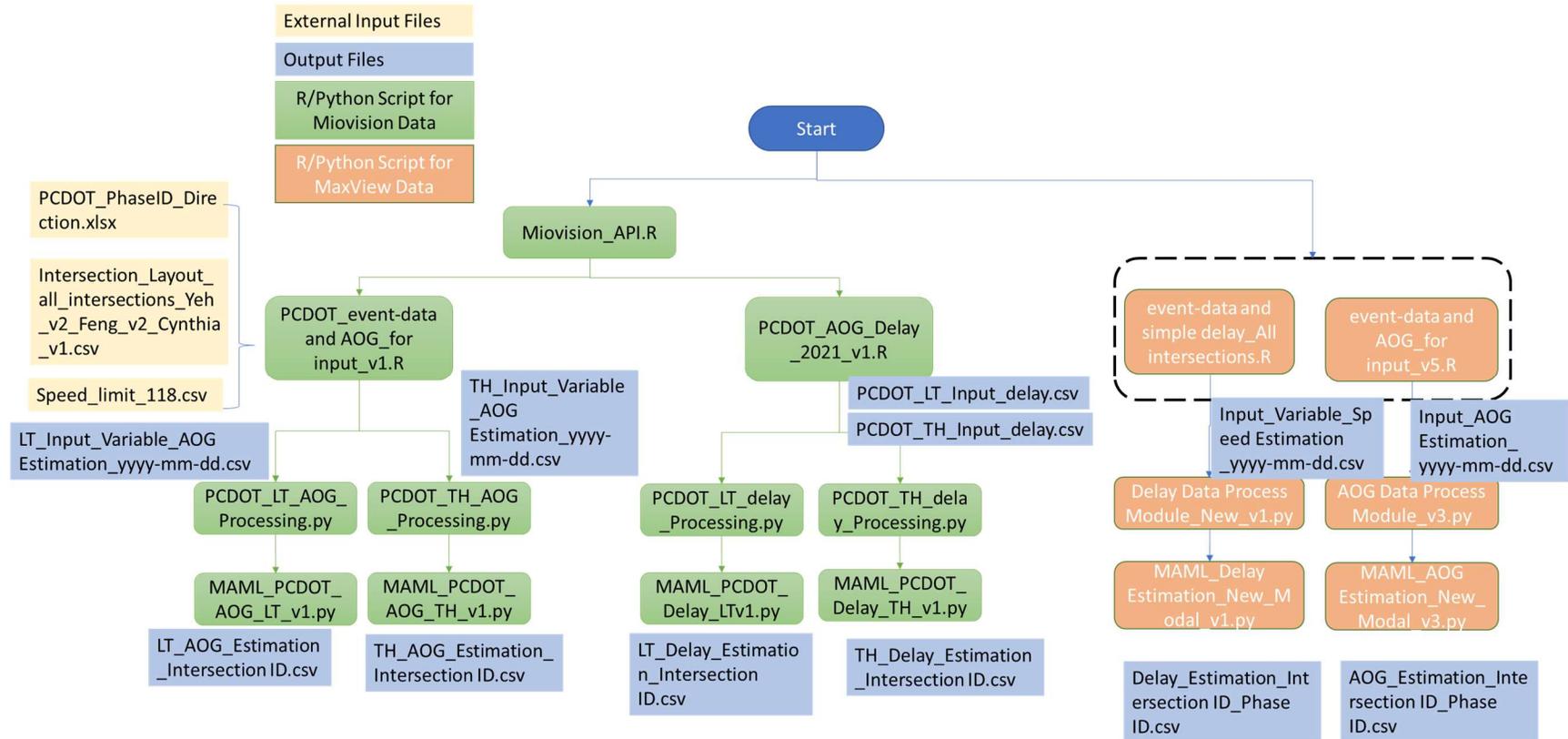
Figure 8-17. Heatmap of monthly average estimated AoG ratio at Speedway Blvd. & Euclid Ave.

APPENDIX A: DATA COLLECTION DESCRIPTION

Table A-1: The Summary of Data Collection Method

Data Source	Data Collection Method	Requirements	Collection Code File Name	Code Description
Event-based data	Query from MaxView database	VPN account & Database account	Read Event-based data. R	Read raw events of detection, signal, and communication
Event-based data	Using an API	API key	Miovision_API. R	Download the events from cloud server
Volume data	Using an API	API key	Miovision_API. R	Download the TMC from cloud server
AoG data	Request the data to Miovision or manually download from the TrafficLink	TrafficLink account	/	/
Simple delay data	Request the data to Miovision or manually download from the TrafficLink	TrafficLink account	/	/
Split failure	Request the data to Miovision or manually download from the TrafficLink	TrafficLink account	/	/
Wejo data	Provided by PAG	/	/	/
Intersection movement type	Manually collected from MaxView interface and Google Earth	VPN account & MaxView account	/	/
Intersection layout	Manually collected from Google Earth and MaxView interface	VPN account & MaxView account	/	/
Speed limit data	Manually collected from Google Earth	/	/	/

APPENDIX B: MAML MODEL DESCRIPTION FOR ESTIMATING AOG AND DELAY



REFERENCES

- Abdelraouf, A., Abdel-Aty, M., Mahmoud, N., 2022. Sequence-to-Sequence Recurrent Graph Convolutional Networks for Traffic Estimation and Prediction Using Connected Probe Vehicle Data. *IEEE Trans. Intell. Transp. Syst.* 1–11. <https://doi.org/10.1109/TITS.2022.3168865>
- An, C., Wu, Y.-J., Xia, J., Huang, W., 2017. Real-time queue length estimation using event-based advance detector data. *J. Intell. Transp. Syst.* 0, 1–14. <https://doi.org/10.1080/15472450.2017.1299011>
- Anusha S. P., Sharma A., Vanajakshi L., Subramanian S. C., Rilett L. R., 2016. Model-Based Approach for Queue and Delay Estimation at Signalized Intersections with Erroneous Automated Data. *J. Transp. Eng.* 142, 04016013. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000835](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000835)
- Atkins, L., Brey, O.R., Hoekstra, A., Lattimer, C.R., United States. Federal Highway Administration, 2019. Organizing for TSMO - Case Study 2: Systems and Technology - Utilizing ITS Architecture to Advance Transportation Systems Management and Operations (No. FHWA-HOP-19-064).
- Bagdatli, M.E.C., Dokuz, A.S., 2021. Vehicle Delay Estimation at Signalized Intersections Using Machine Learning Algorithms. *Transp. Res. Rec.* 2675, 110–126. <https://doi.org/10.1177/03611981211036874>
- Bahuleyan, H., Vanajakshi, L.D., 2017. Arterial Path-Level Travel-Time Estimation Using Machine-Learning Techniques. *J. Comput. Civ. Eng.* 31, 04016070. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000644](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000644)
- Ban, X., Herring, R., Hao, P., Bayen, A., 2009. Delay Pattern Estimation for Signalized Intersections Using Sampled Travel Times. *Transp. Res. Rec. J. Transp. Res. Board* 2130, 109–119. <https://doi.org/10.3141/2130-14>
- Banerjee, T., Manjunatha, P., Uddin, M.S., Kibet, L., Sengupta, R., Ranka, S., Elefteriadou, L., University of Florida. Transportation Institute, 2021. Data Analytics and Evaluation of the Gainesville Trapezium Connected Vehicle Signal Phasing and Timing (SPaT) Deployment Project.
- Brennan, T.M., Day, C.M., Sturdevant, J.R., Bullock, D.M., 2011. Visual Education Tools to Illustrate Coordinated System Operation. *Transp. Res. Rec.* 2259, 59–72. <https://doi.org/10.3141/2259-06>
- Cambridge Systematics, 2022. TDOT Transportation Systems Management and Operations (TSMO).

-
- CATT Lab, C., 2022. The Regional Integrated Transportation Information System (RITIS) [WWW Document]. CATT Lab. URL https://www.cattlab.umd.edu/?page_id=1395 (accessed 10.11.22).
- Chowdhury, M., Sadek, A., Ma, Y., Kanhere, N., Bhavsar, P., 2006. Applications of Artificial Intelligence Paradigms to Decision Support in Real-Time Traffic Management. *Transp. Res. Rec.* 1968, 92–98. <https://doi.org/10.1177/0361198106196800111>
- Choy, M.C., Cheu, R.L., Srinivasan, D., Logi, F., 2003. Real-Time Coordinated Signal Control Through Use of Agents with Online Reinforcement Learning. *Transp. Res. Rec.* 1836, 64–75. <https://doi.org/10.3141/1836-09>
- Day, C., Bullock, D., Li, H., Remias, S., Hainen, A., Freije, R., Stevens, A., Sturdevant, J., Brennan, T., 2014. Performance Measures for Traffic Signal Systems: An Outcome-Oriented Approach (Purdue University, West Lafayette). Indiana.
- Day, C., Smaglik, E., Bullock, D., Sturdevant, J., 2008. Real-Time Arterial Traffic Signal Performance Measures (No. FHWA/IN/JTRP-2008/09). Joint Transportation Research Program, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana.
- Deardoff, M.D., Wiesner, B.N., Fazio, J., 2011. Estimating Free-flow Speed from Posted Speed Limit Signs. *Procedia - Soc. Behav. Sci.*, 6th International Symposium on Highway Capacity and Quality of Service 16, 306–316. <https://doi.org/10.1016/j.sbspro.2011.04.452>
- DelDOT, 2022. Integrated Transportation Management Program - Delaware Department of Transportation [WWW Document]. URL <https://deldot.gov/Programs/itms/index.shtml?dc=projects> (accessed 12.13.22).
- Desai, J., Saldivar–Carranza, E., Mathew, J.K., Li, H., Platte, T., Bullock, D., 2021. Methodology for Applying Connected Vehicle Data to Evaluate Impact of Interstate Construction Work Zone Diversions, in: 2021 IEEE International Intelligent Transportation Systems Conference (ITSC). Presented at the 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), pp. 4035–4042. <https://doi.org/10.1109/ITSC48978.2021.9564873>
- Ding, F., Zhang, Z., Zhou, Y., Chen, X., Ran, B., 2019. Large-Scale Full-Coverage Traffic Speed Estimation under Extreme Traffic Conditions Using a Big Data and Deep Learning Approach: Case Study in China. *J. Transp. Eng. Part Syst.* 145, 05019001. <https://doi.org/10.1061/JTEPBS.0000230>
- Dobrota, N., Stevanovic, A., Mitrovic, N., 2022. A novel model to jointly estimate delay and arrival patterns by using high-resolution signal and detection data. *Transp. Transp. Sci.* 0, 1–33. <https://doi.org/10.1080/23249935.2022.2047126>
- FDOT, 2017. Transportation Systems Management & Operations - Strategic Plan.

-
- FHWA, 2020. What is TSMO? [WWW Document]. URL <https://ops.fhwa.dot.gov/tsmo/#q2> (accessed 10.7.22).
- Finn, C., Abbeel, P., Levine, S., 2017. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks.
- Gannett Felming, BPS, 2017. Performance Measures | Ohio Department of Transportation.
- Gavric, S., Dobrota, N., Erdagi, I.G., Stevanovic, A., Osman, O.A., 2023. Estimation of Arrivals on Green at Signalized Intersections Using Stop-Bar Video Detection. *Transp. Res. Rec.* 03611981221150394. <https://doi.org/10.1177/03611981221150394>
- Gettman, D., 2019. Raising Awareness of Artificial Intelligence for Transportation Systems Management and Operations - FHWA Office of Operations (No. FHWA-HOP-19-052). U.S. Department of Transportation Federal Highway Administration Office of Operations, Washington.
- Hao, P., Ban, X. (Jeff), Guo, D., Ji, Q., 2014. Cycle-by-cycle intersection queue length distribution estimation using sample travel times. *Transp. Res. Part B Methodol.* 68, 185–204. <https://doi.org/10.1016/j.trb.2014.06.004>
- Hunter, M., Mathew, J., Cox, E., Blackwell, M., Bullock, D., 2021. Estimation of Connected Vehicle Penetration Rate on Indiana Roadways. *JTRP Affil. Rep.* <https://doi.org/10.5703/1288284317343>
- Hunter, M.P., Wu, S.K., Kim, H.K., Suh, W., 2012. A Probe-Vehicle-Based Evaluation of Adaptive Traffic Signal Control. *IEEE Trans. Intell. Transp. Syst.* 13, 704–713. <https://doi.org/10.1109/TITS.2011.2178404>
- INRIX, 2021. IQ Signal Analytics Overview.
- (Jeff) Ban, X., Hao, P., Sun, Z., 2011. Real time queue length estimation for signalized intersections using travel times from mobile sensors. *Transp. Res. Part C Emerg. Technol.* 19, 1133–1156. <https://doi.org/10.1016/j.trc.2011.01.002>
- Khattak, Z.H., Magalotti, M.J., Fontaine, M.D., 2019. Operational performance evaluation of adaptive traffic control systems: A Bayesian modeling approach using real-world GPS and private sector PROBE data. *J. Intell. Transp. Syst.* 0, 1–15. <https://doi.org/10.1080/15472450.2019.1614445>
- Kimley Horn, 2017. Arizona Department of Transportation Transportation Systems Management and Operations (TSM&O) Plan.
- Ko, J., Hunter, M., Guensler, R., 2008. Measuring Control Delay Components Using Second-by-Second GPS Speed Data. *J. Transp. Eng.* 134, 338–346. [https://doi.org/10.1061/\(ASCE\)0733-947X\(2008\)134:8\(338\)](https://doi.org/10.1061/(ASCE)0733-947X(2008)134:8(338))

-
- Lakeside Engineers, 2016. Iowa Transportation Systems Management and Operations (TSMO) Program Plan.
- Li, X., Wu, Y.-J., Chiu, Y.-C., 2019. Volume Estimation using Traffic Signal Event-Based Data from Video-Based Sensors. *Transp. Res. Rec.* 2673, 22–32. <https://doi.org/10.1177/0361198119842120>
- Liu, D., An, C., Yasir, M., Lu, J., Xia, J., 2022. A Machine Learning Based Method for Real-Time Queue Length Estimation Using License Plate Recognition and GPS Trajectory Data. *KSCE J. Civ. Eng.* 26, 2408–2419. <https://doi.org/10.1007/s12205-022-0451-4>
- Liu, H., Ross, W., Nair, G., Ruiz-Juri, N., 2021. Using Smart Work Zone Trailer Data to Evaluate and Predict Lane Closure Impacts with a Consideration of Work Intensity Final Report (No. SWZDI TPF-5(438)). Federal Highway Administration U.S. Department of Transportation.
- Liu, H.X., Wu, X., Ma, W., Hu, H., 2009. Real-time queue length estimation for congested signalized intersections. *Transp. Res. Part C Emerg. Technol.* 17, 412–427. <https://doi.org/10.1016/j.trc.2009.02.003>
- Lu, G.X., Zhang, Y., Noyce, D.A., 2011. Intelligent Traffic Signal System for Isolated Intersections: Dynamic Pedestrian Accommodation. *Transp. Res. Rec.* 2259, 96–111. <https://doi.org/10.3141/2259-09>
- Mahmud, S., Akter, T., Hernandez, S., 2020. Truck Parking Usage Patterns by Facility Amenity Availability. *Transp. Res. Rec.* 2674, 749–763. <https://doi.org/10.1177/0361198120937305>
- Manatee County, 2022. Manatee County’s Expansion of TSMO Capabilities to Enhance Multimodal | National Operations Center of Excellence [WWW Document]. URL <https://transportationops.org/case-studies/manatee-countys-expansion-tsmo-capabilities-enhance-multimodal> (accessed 2.8.23).
- Mathew, J.K., Desai, J., Li, H., Bullock, D.M., 2021. Using Anonymous Connected Vehicle Data to Evaluate Impact of Speed Feedback Displays, Speed Limit Signs and Roadway Features on Interstate Work Zones Speeds. *J. Transp. Technol.* 11, 545–560. <https://doi.org/10.4236/jtts.2021.114034>
- Matveev, A., Maximov, A., Bogdanova, E., 2020. Intelligent decision support system for transportation emergency response. *Transp. Res. Procedia, XIV International Conference on Organization and Traffic Safety Management in Large Cities (OTS-2020)* 50, 444–450. <https://doi.org/10.1016/j.trpro.2020.10.058>
- Meldrum, D., Taylor, C., 2000. Algorithm design, user interface, and optimization procedure for a fuzzy logic ramp metering algorithm: a training manual for freeway operations engineers.

-
- Miovision, 2022a. Miovision TrafficLink Traffic Operations [WWW Document]. URL https://help.miovision.com/s/topic/0TO0d000000EoG9GAK/traffic-operations?language=en_US&tabset-a4a08=2 (accessed 10.11.22).
- Miovision, 2022b. Insights Split Failures in Miovision TrafficLink portal [WWW Document]. URL https://help.miovision.com/s/article/Insights-Split-Failures-in-Miovision-TrafficLink-portal?language=en_US (accessed 7.5.23).
- MnDOT, 2019. MnDOT Transportation Systems Management and Operations Implementation Plan.
- Murthy, K., Yang, H. (Frank), 2021. A COST-EFFECTIVE SOLUTION FOR TRUCK PARKING BASED ON ARTIFICIAL INTELLIGENCE.
- NDOT, 2020. NDOT Statewide TSMO Program Plan 2020.
- New Jersey Department of Transportation, 2022. New Jersey Connected Technology Integration and Implementation (NJCTII) | National Operations Center of Excellence [WWW Document]. URL <https://transportationops.org/case-studies/new-jersey-connected-technology-integration-and-implementation-njctii> (accessed 2.8.23).
- NITC, 2018. Smart City PDX Pilot: NITC Researchers Take Connected Vehicle Technology for a Test Run on Portland Streetcar | National Institute for Transportation and Communities [WWW Document]. URL <https://nitc.trec.pdx.edu/news/smart-city-pdx-pilot-nitc-researchers-take-connected-vehicle-technology-test-run-portland> (accessed 2.13.23).
- ODOT, 2021. ODOT-Operations Program Performance Management Plan.
- Ren, T., Xie, Y., Jiang, L., 2020. Cooperative Highway Work Zone Merge Control Based on Reinforcement Learning in a Connected and Automated Environment. *Transp. Res. Rec.* 2674, 363–374. <https://doi.org/10.1177/0361198120935873>
- Sadek, A.W., Smith, B.L., Demetsky, M.J., 1998. Artificial Intelligence-Based Architecture for Real-Time Traffic Flow Management. *Transp. Res. Rec.* 1651, 53–58. <https://doi.org/10.3141/1651-08>
- Saha Arpita, Chandra Satish, Ghosh Indrajit, 2017. Delay at Signalized Intersections under Mixed Traffic Conditions. *J. Transp. Eng. Part Syst.* 143, 04017041. <https://doi.org/10.1061/JTEPBS.0000070>
- Saldivar-Carranza, E., Li, H., Mathew, J., Hunter, M., Sturdevant, J., Bullock, D.M., 2021a. Deriving Operational Traffic Signal Performance Measures from Vehicle Trajectory Data. *Transp. Res. Rec.* 2675, 1250–1264. <https://doi.org/10.1177/03611981211006725>

-
- Saldivar-Carranza, E., Mathew, J.K., Li, H., Bullock, D.M., 2021b. Roundabout Performance Analysis Using Connected Vehicle Data. *J. Transp. Technol.* 12, 42–58. <https://doi.org/10.4236/jtts.2022.121003>
- Saldivar-Carranza, E.D., Li, H., Bullock, D.M., 2021a. Diverging Diamond Interchange Performance Measures Using Connected Vehicle Data. *J. Transp. Technol.* 11, 628–643. <https://doi.org/10.4236/jtts.2021.114039>
- Saldivar-Carranza, E.D., Mathew, J.K., Li, H., Hunter, M., Platte, T., Bullock, D.M., 2021b. Using Connected Vehicle Data to Evaluate Traffic Signal Performance and Driver Behavior after Changing Left-turns Phasing, in: 2021 IEEE International Intelligent Transportation Systems Conference (ITSC). Presented at the 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), pp. 4028–4034. <https://doi.org/10.1109/ITSC48978.2021.9564654>
- Sekuła, P., Marković, N., Vander Laan, Z., Sadabadi, K.F., 2018. Estimating historical hourly traffic volumes via machine learning and vehicle probe data: A Maryland case study. *Transp. Res. Part C Emerg. Technol.* 97, 147–158. <https://doi.org/10.1016/j.trc.2018.10.012>
- Shoman, M., Amo-Boateng, M., Adu-Gyamfi, Y., 2022. Multi-Purpose, Multi-Step Deep Learning Framework for Network-Level Traffic Flow Prediction. <https://doi.org/10.2139/ssrn.4063434>
- Silvano, A.P., Koutsopoulos, H.N., Farah, H., 2020. Free flow speed estimation: A probabilistic, latent approach. Impact of speed limit changes and road characteristics. *Transp. Res. Part Policy Pract.* 138, 283–298. <https://doi.org/10.1016/j.tra.2020.05.024>
- Stolle, S., 2018. Can A.I. Take Over Winter Road Condition Reporting?
- Tian, Z., Ohene, F., Hu, P., 2011. Arterial Performance Evaluation on an Adaptive Traffic Signal Control System. *Procedia - Soc. Behav. Sci.*, 6th International Symposium on Highway Capacity and Quality of Service 16, 230–239. <https://doi.org/10.1016/j.sbspro.2011.04.445>
- Transportation Research Board, 2016. Highway Capacity Manual 6th Edition: A Guide for Multimodal Mobility Analysis. The National Academies Press, Washington, DC. <https://doi.org/10.17226/24798>
- Wai, B., Zhou, W., 2020. Designing and Implementing Real-Time Bus Time Predictions using Artificial Intelligence. *Transp. Res. Rec.* 2674, 636–648. <https://doi.org/10.1177/0361198120947715>
- Wang, S., Huang, W., Lo, H.K., 2019. Traffic parameters estimation for signalized intersections based on combined shockwave analysis and Bayesian Network. *Transp. Res. Part C Emerg. Technol.* 104, 22–37. <https://doi.org/10.1016/j.trc.2019.04.023>

-
- Wang, Z., Chen, S., Yang, H., Wu, B., Wang, Y., 2016. Comparison of delay estimation models for signalised intersections using field observations in Shanghai. *IET Intell. Transp. Syst.* 10, 165–174. <https://doi.org/10.1049/iet-its.2014.0280>
- Wang, Z., Fu, K., Ye, J., 2018. Learning to Estimate the Travel Time, in: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Presented at the KDD '18: The 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, London United Kingdom, pp. 858–866. <https://doi.org/10.1145/3219819.3219900>
- Xiao, J., Wei, C., Liu, Y., 2018. Speed estimation of traffic flow using multiple kernel support vector regression. *Phys. Stat. Mech. Its Appl.* 509, 989–997. <https://doi.org/10.1016/j.physa.2018.06.082>
- Zhan, X., Li, R., Ukkusuri, S.V., 2015. Lane-based real-time queue length estimation using license plate recognition data. *Transp. Res. Part C Emerg. Technol.* 57, 85–102. <https://doi.org/10.1016/j.trc.2015.06.001>
- Zhang, Z., Li, M., Lin, X., Wang, Y., 2020. Network-wide traffic flow estimation with insufficient volume detection and crowdsourcing data. *Transp. Res. Part C Emerg. Technol.* 121, 102870. <https://doi.org/10.1016/j.trc.2020.102870>
- Zhao, Y., Zheng, J., Wong, W., Wang, X., Meng, Y., Liu, H.X., 2019. Various methods for queue length and traffic volume estimation using probe vehicle trajectories. *Transp. Res. Part C Emerg. Technol.* 107, 70–91. <https://doi.org/10.1016/j.trc.2019.07.008>
- Zheng, J., Liu, H.X., 2017. Estimating traffic volumes for signalized intersections using connected vehicle data. *Transp. Res. Part C Emerg. Technol.* 79, 347–362. <https://doi.org/10.1016/j.trc.2017.03.007>
- Zheng, J., Ma, X., Wu, Y.-J., Wang, Y., 2013. Measuring Signalized Intersection Performance in Real-Time With Traffic Sensors. *J. Intell. Transp. Syst.* 17, 304–316. <https://doi.org/10.1080/15472450.2013.771105>
- Abdelraouf, A., Abdel-Aty, M., Mahmoud, N., 2022. Sequence-to-Sequence Recurrent Graph Convolutional Networks for Traffic Estimation and Prediction Using Connected Probe Vehicle Data. *IEEE Trans. Intell. Transp. Syst.* 1–11. <https://doi.org/10.1109/TITS.2022.3168865>
- An, C., Wu, Y.-J., Xia, J., Huang, W., 2017. Real-time queue length estimation using event-based advance detector data. *J. Intell. Transp. Syst.* 0, 1–14. <https://doi.org/10.1080/15472450.2017.1299011>
- Anusha S. P., Sharma A., Vanajakshi L., Subramanian S. C., Rilett L. R., 2016. Model-Based Approach for Queue and Delay Estimation at Signalized Intersections with Erroneous

-
- Automated Data. *J. Transp. Eng.* 142, 04016013. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000835](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000835)
- Atkins, L., Brey, O.R., Hoekstra, A., Lattimer, C.R., United States. Federal Highway Administration, 2019. Organizing for TSMO - Case Study 2: Systems and Technology - Utilizing ITS Architecture to Advance Transportation Systems Management and Operations (No. FHWA-HOP-19-064).
- Bagdatli, M.E.C., Dokuz, A.S., 2021. Vehicle Delay Estimation at Signalized Intersections Using Machine Learning Algorithms. *Transp. Res. Rec.* 2675, 110–126. <https://doi.org/10.1177/03611981211036874>
- Bahuleyan, H., Vanajakshi, L.D., 2017. Arterial Path-Level Travel-Time Estimation Using Machine-Learning Techniques. *J. Comput. Civ. Eng.* 31, 04016070. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000644](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000644)
- Ban, X., Herring, R., Hao, P., Bayen, A., 2009. Delay Pattern Estimation for Signalized Intersections Using Sampled Travel Times. *Transp. Res. Rec. J. Transp. Res. Board* 2130, 109–119. <https://doi.org/10.3141/2130-14>
- Banerjee, T., Manjunatha, P., Uddin, M.S., Kibet, L., Sengupta, R., Ranka, S., Elefteriadou, L., University of Florida. Transportation Institute, 2021. Data Analytics and Evaluation of the Gainesville Trapezium Connected Vehicle Signal Phasing and Timing (SPaT) Deployment Project.
- Brennan, T.M., Day, C.M., Sturdevant, J.R., Bullock, D.M., 2011. Visual Education Tools to Illustrate Coordinated System Operation. *Transp. Res. Rec.* 2259, 59–72. <https://doi.org/10.3141/2259-06>
- Cambridge Systematics, 2022. TDOT Transportation Systems Management and Operations (TSMO).
- CATT Lab, C., 2022. The Regional Integrated Transportation Information System (RITIS) [WWW Document]. CATT Lab. URL https://www.cattlab.umd.edu/?page_id=1395 (accessed 10.11.22).
- Chowdhury, M., Sadek, A., Ma, Y., Kanhere, N., Bhavsar, P., 2006. Applications of Artificial Intelligence Paradigms to Decision Support in Real-Time Traffic Management. *Transp. Res. Rec.* 1968, 92–98. <https://doi.org/10.1177/0361198106196800111>
- Choy, M.C., Cheu, R.L., Srinivasan, D., Logi, F., 2003. Real-Time Coordinated Signal Control Through Use of Agents with Online Reinforcement Learning. *Transp. Res. Rec.* 1836, 64–75. <https://doi.org/10.3141/1836-09>

-
- Day, C., Bullock, D., Li, H., Remias, S., Hainen, A., Freije, R., Stevens, A., Sturdevant, J., Brennan, T., 2014. Performance Measures for Traffic Signal Systems: An Outcome-Oriented Approach (Purdue University, West Lafayette). Indiana.
- Day, C., Smaglik, E., Bullock, D., Sturdevant, J., 2008. Real-Time Arterial Traffic Signal Performance Measures (No. FHWA/IN/JTRP-2008/09). Joint Transportation Research Program, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana.
- Deardoff, M.D., Wiesner, B.N., Fazio, J., 2011. Estimating Free-flow Speed from Posted Speed Limit Signs. *Procedia - Soc. Behav. Sci.*, 6th International Symposium on Highway Capacity and Quality of Service 16, 306–316. <https://doi.org/10.1016/j.sbspro.2011.04.452>
- DelDOT, 2022. Integrated Transportation Management Program - Delaware Department of Transportation [WWW Document]. URL <https://deldot.gov/Programs/itms/index.shtml?dc=projects> (accessed 12.13.22).
- Desai, J., Saldivar–Carranza, E., Mathew, J.K., Li, H., Platte, T., Bullock, D., 2021. Methodology for Applying Connected Vehicle Data to Evaluate Impact of Interstate Construction Work Zone Diversions, in: 2021 IEEE International Intelligent Transportation Systems Conference (ITSC). Presented at the 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), pp. 4035–4042. <https://doi.org/10.1109/ITSC48978.2021.9564873>
- Ding, F., Zhang, Z., Zhou, Y., Chen, X., Ran, B., 2019. Large-Scale Full-Coverage Traffic Speed Estimation under Extreme Traffic Conditions Using a Big Data and Deep Learning Approach: Case Study in China. *J. Transp. Eng. Part Syst.* 145, 05019001. <https://doi.org/10.1061/JTEPBS.0000230>
- Dobrota, N., Stevanovic, A., Mitrovic, N., 2022. A novel model to jointly estimate delay and arrival patterns by using high-resolution signal and detection data. *Transp. Transp. Sci.* 0, 1–33. <https://doi.org/10.1080/23249935.2022.2047126>
- FDOT, 2017. Transportation Systems Management & Operations - Strategic Plan.
- FHWA, 2020. What is TSMO? [WWW Document]. URL <https://ops.fhwa.dot.gov/tsmo/#q2> (accessed 10.7.22).
- Finn, C., Abbeel, P., Levine, S., 2017. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks.
- Gannett Felming, BPS, 2017. Performance Measures | Ohio Department of Transportation.
- Gavric, S., Dobrota, N., Erdagi, I.G., Stevanovic, A., Osman, O.A., 2023. Estimation of Arrivals on Green at Signalized Intersections Using Stop-Bar Video Detection. *Transp. Res. Rec.* 03611981221150394. <https://doi.org/10.1177/03611981221150394>

-
- Gettman, D., 2019. Raising Awareness of Artificial Intelligence for Transportation Systems Management and Operations - FHWA Office of Operations (No. FHWA-HOP-19-052). U.S. Department of Transportation Federal Highway Administration Office of Operations, Washington.
- Hao, P., Ban, X. (Jeff), Guo, D., Ji, Q., 2014. Cycle-by-cycle intersection queue length distribution estimation using sample travel times. *Transp. Res. Part B Methodol.* 68, 185–204. <https://doi.org/10.1016/j.trb.2014.06.004>
- Hunter, M., Mathew, J., Cox, E., Blackwell, M., Bullock, D., 2021. Estimation of Connected Vehicle Penetration Rate on Indiana Roadways. *JTRP Affil. Rep.* <https://doi.org/10.5703/1288284317343>
- Hunter, M.P., Wu, S.K., Kim, H.K., Suh, W., 2012. A Probe-Vehicle-Based Evaluation of Adaptive Traffic Signal Control. *IEEE Trans. Intell. Transp. Syst.* 13, 704–713. <https://doi.org/10.1109/TITS.2011.2178404>
- INRIX, 2021. IQ Signal Analytics Overview.
- (Jeff) Ban, X., Hao, P., Sun, Z., 2011. Real time queue length estimation for signalized intersections using travel times from mobile sensors. *Transp. Res. Part C Emerg. Technol.* 19, 1133–1156. <https://doi.org/10.1016/j.trc.2011.01.002>
- Khattak, Z.H., Magalotti, M.J., Fontaine, M.D., 2019. Operational performance evaluation of adaptive traffic control systems: A Bayesian modeling approach using real-world GPS and private sector PROBE data. *J. Intell. Transp. Syst.* 0, 1–15. <https://doi.org/10.1080/15472450.2019.1614445>
- Kimley Horn, 2017. Arizona Department of Transportation Transportation Systems Management and Operations (TSM&O) Plan.
- Ko, J., Hunter, M., Guensler, R., 2008. Measuring Control Delay Components Using Second-by-Second GPS Speed Data. *J. Transp. Eng.* 134, 338–346. [https://doi.org/10.1061/\(ASCE\)0733-947X\(2008\)134:8\(338\)](https://doi.org/10.1061/(ASCE)0733-947X(2008)134:8(338))
- Lakeside Engineers, 2016. Iowa Transportation Systems Management and Operations (TSMO) Program Plan.
- Li, X., Wu, Y.-J., Chiu, Y.-C., 2019. Volume Estimation using Traffic Signal Event-Based Data from Video-Based Sensors. *Transp. Res. Rec.* 2673, 22–32. <https://doi.org/10.1177/0361198119842120>
- Liu, D., An, C., Yasir, M., Lu, J., Xia, J., 2022. A Machine Learning Based Method for Real-Time Queue Length Estimation Using License Plate Recognition and GPS Trajectory Data. *KSCE J. Civ. Eng.* 26, 2408–2419. <https://doi.org/10.1007/s12205-022-0451-4>

-
- Liu, H., Ross, W., Nair, G., Ruiz-Juri, N., 2021. Using Smart Work Zone Trailer Data to Evaluate and Predict Lane Closure Impacts with a Consideration of Work Intensity Final Report (No. SWZDI TPF-5(438)). Federal Highway Administration U.S. Department of Transportation.
- Liu, H.X., Wu, X., Ma, W., Hu, H., 2009. Real-time queue length estimation for congested signalized intersections. *Transp. Res. Part C Emerg. Technol.* 17, 412–427. <https://doi.org/10.1016/j.trc.2009.02.003>
- Lu, G.X., Zhang, Y., Noyce, D.A., 2011. Intelligent Traffic Signal System for Isolated Intersections: Dynamic Pedestrian Accommodation. *Transp. Res. Rec.* 2259, 96–111. <https://doi.org/10.3141/2259-09>
- Mahmud, S., Akter, T., Hernandez, S., 2020. Truck Parking Usage Patterns by Facility Amenity Availability. *Transp. Res. Rec.* 2674, 749–763. <https://doi.org/10.1177/0361198120937305>
- Manatee County, 2022. Manatee County’s Expansion of TSMO Capabilities to Enhance Multimodal | National Operations Center of Excellence [WWW Document]. URL <https://transportationops.org/case-studies/manatee-countys-expansion-tsmo-capabilities-enhance-multimodal> (accessed 2.8.23).
- Mathew, J.K., Desai, J., Li, H., Bullock, D.M., 2021. Using Anonymous Connected Vehicle Data to Evaluate Impact of Speed Feedback Displays, Speed Limit Signs and Roadway Features on Interstate Work Zones Speeds. *J. Transp. Technol.* 11, 545–560. <https://doi.org/10.4236/jtts.2021.114034>
- Matveev, A., Maximov, A., Bogdanova, E., 2020. Intelligent decision support system for transportation emergency response. *Transp. Res. Procedia, XIV International Conference on Organization and Traffic Safety Management in Large Cities (OTS-2020)* 50, 444–450. <https://doi.org/10.1016/j.trpro.2020.10.058>
- Meldrum, D., Taylor, C., 2000. Algorithm design, user interface, and optimization procedure for a fuzzy logic ramp metering algorithm: a training manual for freeway operations engineers.
- Miovision, 2022a. Miovision TrafficLink Traffic Operations [WWW Document]. URL https://help.miovision.com/s/topic/0TO0d000000EoG9GAK/traffic-operations?language=en_US&tabset-a4a08=2 (accessed 10.11.22).
- Miovision, 2022b. Insights Split Failures in Miovision TrafficLink portal [WWW Document]. URL https://help.miovision.com/s/article/Insights-Split-Failures-in-Miovision-TrafficLink-portal?language=en_US (accessed 7.5.23).
- MnDOT, 2019. MnDOT Transportation Systems Management and Operations Implementation Plan.

Murthy, K., Yang, H. (Frank), 2021. A COST-EFFECTIVE SOLUTION FOR TRUCK PARKING BASED ON ARTIFICIAL INTELLIGENCE.

NDOT, 2020. NDOT Statewide TSMO Program Plan 2020.

New Jersey Department of Transportation, 2022. New Jersey Connected Technology Integration and Implementation (NJCTII) | National Operations Center of Excellence [WWW Document]. URL <https://transportationops.org/case-studies/new-jersey-connected-technology-integration-and-implementation-njctii> (accessed 2.8.23).

NITC, 2018. Smart City PDX Pilot: NITC Researchers Take Connected Vehicle Technology for a Test Run on Portland Streetcar | National Institute for Transportation and Communities [WWW Document]. URL <https://nitc.trec.pdx.edu/news/smart-city-pdx-pilot-nitc-researchers-take-connected-vehicle-technology-test-run-portland> (accessed 2.13.23).

ODOT, 2021. ODOT-Operations Program Performance Management Plan.

Ren, T., Xie, Y., Jiang, L., 2020. Cooperative Highway Work Zone Merge Control Based on Reinforcement Learning in a Connected and Automated Environment. *Transp. Res. Rec.* 2674, 363–374. <https://doi.org/10.1177/0361198120935873>

Sadek, A.W., Smith, B.L., Demetsky, M.J., 1998. Artificial Intelligence-Based Architecture for Real-Time Traffic Flow Management. *Transp. Res. Rec.* 1651, 53–58. <https://doi.org/10.3141/1651-08>

Saha Arpita, Chandra Satish, Ghosh Indrajit, 2017. Delay at Signalized Intersections under Mixed Traffic Conditions. *J. Transp. Eng. Part Syst.* 143, 04017041. <https://doi.org/10.1061/JTEPBS.0000070>

Saldivar-Carranza, E., Li, H., Mathew, J., Hunter, M., Sturdevant, J., Bullock, D.M., 2021a. Deriving Operational Traffic Signal Performance Measures from Vehicle Trajectory Data. *Transp. Res. Rec.* 2675, 1250–1264. <https://doi.org/10.1177/03611981211006725>

Saldivar-Carranza, E., Mathew, J.K., Li, H., Bullock, D.M., 2021b. Roundabout Performance Analysis Using Connected Vehicle Data. *J. Transp. Technol.* 12, 42–58. <https://doi.org/10.4236/jtts.2022.121003>

Saldivar-Carranza, E.D., Li, H., Bullock, D.M., 2021a. Diverging Diamond Interchange Performance Measures Using Connected Vehicle Data. *J. Transp. Technol.* 11, 628–643. <https://doi.org/10.4236/jtts.2021.114039>

Saldivar-Carranza, E.D., Mathew, J.K., Li, H., Hunter, M., Platte, T., Bullock, D.M., 2021b. Using Connected Vehicle Data to Evaluate Traffic Signal Performance and Driver Behavior after Changing Left-turns Phasing, in: 2021 IEEE International Intelligent Transportation Systems Conference (ITSC). Presented at the 2021 IEEE International Intelligent

-
- Transportation Systems Conference (ITSC), pp. 4028–4034.
<https://doi.org/10.1109/ITSC48978.2021.9564654>
- Sekula, P., Marković, N., Vander Laan, Z., Sadabadi, K.F., 2018. Estimating historical hourly traffic volumes via machine learning and vehicle probe data: A Maryland case study. *Transp. Res. Part C Emerg. Technol.* 97, 147–158.
<https://doi.org/10.1016/j.trc.2018.10.012>
- Shoman, M., Amo-Boateng, M., Adu-Gyamfi, Y., 2022. Multi-Purpose, Multi-Step Deep Learning Framework for Network-Level Traffic Flow Prediction.
<https://doi.org/10.2139/ssrn.4063434>
- Silvano, A.P., Koutsopoulos, H.N., Farah, H., 2020. Free flow speed estimation: A probabilistic, latent approach. Impact of speed limit changes and road characteristics. *Transp. Res. Part Policy Pract.* 138, 283–298. <https://doi.org/10.1016/j.tra.2020.05.024>
- Stolle, S., 2018. Can A.I. Take Over Winter Road Condition Reporting?
- Tian, Z., Ohene, F., Hu, P., 2011. Arterial Performance Evaluation on an Adaptive Traffic Signal Control System. *Procedia - Soc. Behav. Sci.*, 6th International Symposium on Highway Capacity and Quality of Service 16, 230–239. <https://doi.org/10.1016/j.sbspro.2011.04.445>
- Transportation Research Board, 2016. Highway Capacity Manual 6th Edition: A Guide for Multimodal Mobility Analysis. The National Academies Press, Washington, DC.
<https://doi.org/10.17226/24798>
- Wai, B., Zhou, W., 2020. Designing and Implementing Real-Time Bus Time Predictions using Artificial Intelligence. *Transp. Res. Rec.* 2674, 636–648.
<https://doi.org/10.1177/0361198120947715>
- Wang, S., Huang, W., Lo, H.K., 2019. Traffic parameters estimation for signalized intersections based on combined shockwave analysis and Bayesian Network. *Transp. Res. Part C Emerg. Technol.* 104, 22–37. <https://doi.org/10.1016/j.trc.2019.04.023>
- Wang, Z., Chen, S., Yang, H., Wu, B., Wang, Y., 2016. Comparison of delay estimation models for signalised intersections using field observations in Shanghai. *IET Intell. Transp. Syst.* 10, 165–174. <https://doi.org/10.1049/iet-its.2014.0280>
- Wang, Z., Fu, K., Ye, J., 2018. Learning to Estimate the Travel Time, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Presented at the KDD '18: The 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, London United Kingdom, pp. 858–866.
<https://doi.org/10.1145/3219819.3219900>

-
- Xiao, J., Wei, C., Liu, Y., 2018. Speed estimation of traffic flow using multiple kernel support vector regression. *Phys. Stat. Mech. Its Appl.* 509, 989–997. <https://doi.org/10.1016/j.physa.2018.06.082>
- Zhan, X., Li, R., Ukkusuri, S.V., 2015. Lane-based real-time queue length estimation using license plate recognition data. *Transp. Res. Part C Emerg. Technol.* 57, 85–102. <https://doi.org/10.1016/j.trc.2015.06.001>
- Zhang, Z., Li, M., Lin, X., Wang, Y., 2020. Network-wide traffic flow estimation with insufficient volume detection and crowdsourcing data. *Transp. Res. Part C Emerg. Technol.* 121, 102870. <https://doi.org/10.1016/j.trc.2020.102870>
- Zhao, Y., Zheng, J., Wong, W., Wang, X., Meng, Y., Liu, H.X., 2019. Various methods for queue length and traffic volume estimation using probe vehicle trajectories. *Transp. Res. Part C Emerg. Technol.* 107, 70–91. <https://doi.org/10.1016/j.trc.2019.07.008>
- Zheng, J., Liu, H.X., 2017. Estimating traffic volumes for signalized intersections using connected vehicle data. *Transp. Res. Part C Emerg. Technol.* 79, 347–362. <https://doi.org/10.1016/j.trc.2017.03.007>
- Zheng, J., Ma, X., Wu, Y.-J., Wang, Y., 2013. Measuring Signalized Intersection Performance in Real-Time With Traffic Sensors. *J. Intell. Transp. Syst.* 17, 304–316. <https://doi.org/10.1080/15472450.2013.771105>
- Abdelraouf, A., Abdel-Aty, M., Mahmoud, N., 2022. Sequence-to-Sequence Recurrent Graph Convolutional Networks for Traffic Estimation and Prediction Using Connected Probe Vehicle Data. *IEEE Trans. Intell. Transp. Syst.* 1–11. <https://doi.org/10.1109/TITS.2022.3168865>
- An, C., Wu, Y.-J., Xia, J., Huang, W., 2017. Real-time queue length estimation using event-based advance detector data. *J. Intell. Transp. Syst.* 0, 1–14. <https://doi.org/10.1080/15472450.2017.1299011>
- Anusha S. P., Sharma A., Vanajakshi L., Subramanian S. C., Rilett L. R., 2016. Model-Based Approach for Queue and Delay Estimation at Signalized Intersections with Erroneous Automated Data. *J. Transp. Eng.* 142, 04016013. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000835](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000835)
- Atkins, L., Brey, O.R., Hoekstra, A., Lattimer, C.R., United States. Federal Highway Administration, 2019. Organizing for TSMO - Case Study 2: Systems and Technology - Utilizing ITS Architecture to Advance Transportation Systems Management and Operations (No. FHWA-HOP-19-064).
- Bagdatli, M.E.C., Dokuz, A.S., 2021. Vehicle Delay Estimation at Signalized Intersections Using Machine Learning Algorithms. *Transp. Res. Rec.* 2675, 110–126. <https://doi.org/10.1177/03611981211036874>

-
- Bahuleyan, H., Vanajakshi, L.D., 2017. Arterial Path-Level Travel-Time Estimation Using Machine-Learning Techniques. *J. Comput. Civ. Eng.* 31, 04016070. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000644](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000644)
- Ban, X., Herring, R., Hao, P., Bayen, A., 2009. Delay Pattern Estimation for Signalized Intersections Using Sampled Travel Times. *Transp. Res. Rec. J. Transp. Res. Board* 2130, 109–119. <https://doi.org/10.3141/2130-14>
- Banerjee, T., Manjunatha, P., Uddin, M.S., Kibet, L., Sengupta, R., Ranka, S., Elefteriadou, L., University of Florida. Transportation Institute, 2021. Data Analytics and Evaluation of the Gainesville Trapezium Connected Vehicle Signal Phasing and Timing (SPaT) Deployment Project.
- Brennan, T.M., Day, C.M., Sturdevant, J.R., Bullock, D.M., 2011. Visual Education Tools to Illustrate Coordinated System Operation. *Transp. Res. Rec.* 2259, 59–72. <https://doi.org/10.3141/2259-06>
- Cambridge Systematics, 2022. TDOT Transportation Systems Management and Operations (TSMO).
- CATT Lab, C., 2022. The Regional Integrated Transportation Information System (RITIS) [WWW Document]. CATT Lab. URL https://www.cattlab.umd.edu/?page_id=1395 (accessed 10.11.22).
- Chowdhury, M., Sadek, A., Ma, Y., Kanhere, N., Bhavsar, P., 2006. Applications of Artificial Intelligence Paradigms to Decision Support in Real-Time Traffic Management. *Transp. Res. Rec.* 1968, 92–98. <https://doi.org/10.1177/0361198106196800111>
- Choy, M.C., Cheu, R.L., Srinivasan, D., Logi, F., 2003. Real-Time Coordinated Signal Control Through Use of Agents with Online Reinforcement Learning. *Transp. Res. Rec.* 1836, 64–75. <https://doi.org/10.3141/1836-09>
- Day, C., Bullock, D., Li, H., Remias, S., Hainen, A., Freije, R., Stevens, A., Sturdevant, J., Brennan, T., 2014. Performance Measures for Traffic Signal Systems: An Outcome-Oriented Approach (Purdue University, West Lafayette). Indiana.
- Day, C., Smaglik, E., Bullock, D., Sturdevant, J., 2008. Real-Time Arterial Traffic Signal Performance Measures (No. FHWA/IN/JTRP-2008/09). Joint Transportation Research Program, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana.
- Deardoff, M.D., Wiesner, B.N., Fazio, J., 2011. Estimating Free-flow Speed from Posted Speed Limit Signs. *Procedia - Soc. Behav. Sci.*, 6th International Symposium on Highway Capacity and Quality of Service 16, 306–316. <https://doi.org/10.1016/j.sbspro.2011.04.452>

-
- DelDOT, 2022. Integrated Transportation Management Program - Delaware Department of Transportation [WWW Document]. URL <https://deldot.gov/Programs/itms/index.shtml?dc=projects> (accessed 12.13.22).
- Desai, J., Saldivar–Carranza, E., Mathew, J.K., Li, H., Platte, T., Bullock, D., 2021. Methodology for Applying Connected Vehicle Data to Evaluate Impact of Interstate Construction Work Zone Diversions, in: 2021 IEEE International Intelligent Transportation Systems Conference (ITSC). Presented at the 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), pp. 4035–4042. <https://doi.org/10.1109/ITSC48978.2021.9564873>
- Ding, F., Zhang, Z., Zhou, Y., Chen, X., Ran, B., 2019. Large-Scale Full-Coverage Traffic Speed Estimation under Extreme Traffic Conditions Using a Big Data and Deep Learning Approach: Case Study in China. *J. Transp. Eng. Part Syst.* 145, 05019001. <https://doi.org/10.1061/JTEPBS.0000230>
- Dobrota, N., Stevanovic, A., Mitrovic, N., 2022. A novel model to jointly estimate delay and arrival patterns by using high-resolution signal and detection data. *Transp. Transp. Sci.* 0, 1–33. <https://doi.org/10.1080/23249935.2022.2047126>
- FDOT, 2017. Transportation Systems Management & Operations - Strategic Plan.
- FHWA, 2020. What is TSMO? [WWW Document]. URL <https://ops.fhwa.dot.gov/tsmo/#q2> (accessed 10.7.22).
- Finn, C., Abbeel, P., Levine, S., 2017. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks.
- Gannett Felming, BPS, 2017. Performance Measures | Ohio Department of Transportation.
- Gavric, S., Dobrota, N., Erdagi, I.G., Stevanovic, A., Osman, O.A., 2023. Estimation of Arrivals on Green at Signalized Intersections Using Stop-Bar Video Detection. *Transp. Res. Rec.* 03611981221150394. <https://doi.org/10.1177/03611981221150394>
- Gettman, D., 2019. Raising Awareness of Artificial Intelligence for Transportation Systems Management and Operations - FHWA Office of Operations (No. FHWA-HOP-19-052). U.S. Department of Transportation Federal Highway Administration Office of Operations, Washington.
- Hao, P., Ban, X. (Jeff), Guo, D., Ji, Q., 2014. Cycle-by-cycle intersection queue length distribution estimation using sample travel times. *Transp. Res. Part B Methodol.* 68, 185–204. <https://doi.org/10.1016/j.trb.2014.06.004>
- Hunter, M., Mathew, J., Cox, E., Blackwell, M., Bullock, D., 2021. Estimation of Connected Vehicle Penetration Rate on Indiana Roadways. *JTRP Affil. Rep.* <https://doi.org/10.5703/1288284317343>

-
- Hunter, M.P., Wu, S.K., Kim, H.K., Suh, W., 2012. A Probe-Vehicle-Based Evaluation of Adaptive Traffic Signal Control. *IEEE Trans. Intell. Transp. Syst.* 13, 704–713. <https://doi.org/10.1109/TITS.2011.2178404>
- INRIX, 2021. IQ Signal Analytics Overview.
- (Jeff) Ban, X., Hao, P., Sun, Z., 2011. Real time queue length estimation for signalized intersections using travel times from mobile sensors. *Transp. Res. Part C Emerg. Technol.* 19, 1133–1156. <https://doi.org/10.1016/j.trc.2011.01.002>
- Khattak, Z.H., Magalotti, M.J., Fontaine, M.D., 2019. Operational performance evaluation of adaptive traffic control systems: A Bayesian modeling approach using real-world GPS and private sector PROBE data. *J. Intell. Transp. Syst.* 0, 1–15. <https://doi.org/10.1080/15472450.2019.1614445>
- Kimley Horn, 2017. Arizona Department of Transportation Transportation Systems Management and Operations (TSM&O) Plan.
- Ko, J., Hunter, M., Guensler, R., 2008. Measuring Control Delay Components Using Second-by-Second GPS Speed Data. *J. Transp. Eng.* 134, 338–346. [https://doi.org/10.1061/\(ASCE\)0733-947X\(2008\)134:8\(338\)](https://doi.org/10.1061/(ASCE)0733-947X(2008)134:8(338))
- Lakeside Engineers, 2016. Iowa Transportation Systems Management and Operations (TSMO) Program Plan.
- Li, X., Wu, Y.-J., Chiu, Y.-C., 2019. Volume Estimation using Traffic Signal Event-Based Data from Video-Based Sensors. *Transp. Res. Rec.* 2673, 22–32. <https://doi.org/10.1177/0361198119842120>
- Liu, D., An, C., Yasir, M., Lu, J., Xia, J., 2022. A Machine Learning Based Method for Real-Time Queue Length Estimation Using License Plate Recognition and GPS Trajectory Data. *KSCCE J. Civ. Eng.* 26, 2408–2419. <https://doi.org/10.1007/s12205-022-0451-4>
- Liu, H., Ross, W., Nair, G., Ruiz-Juri, N., 2021. Using Smart Work Zone Trailer Data to Evaluate and Predict Lane Closure Impacts with a Consideration of Work Intensity Final Report (No. SWZDI TPF-5(438)). Federal Highway Administration U.S. Department of Transportation.
- Liu, H.X., Wu, X., Ma, W., Hu, H., 2009. Real-time queue length estimation for congested signalized intersections. *Transp. Res. Part C Emerg. Technol.* 17, 412–427. <https://doi.org/10.1016/j.trc.2009.02.003>
- Lu, G.X., Zhang, Y., Noyce, D.A., 2011. Intelligent Traffic Signal System for Isolated Intersections: Dynamic Pedestrian Accommodation. *Transp. Res. Rec.* 2259, 96–111. <https://doi.org/10.3141/2259-09>

-
- Mahmud, S., Akter, T., Hernandez, S., 2020. Truck Parking Usage Patterns by Facility Amenity Availability. *Transp. Res. Rec.* 2674, 749–763. <https://doi.org/10.1177/0361198120937305>
- Manatee County, 2022. Manatee County’s Expansion of TSMO Capabilities to Enhance Multimodal | National Operations Center of Excellence [WWW Document]. URL <https://transportationops.org/case-studies/manatee-countys-expansion-tsmo-capabilities-enhance-multimodal> (accessed 2.8.23).
- Mathew, J.K., Desai, J., Li, H., Bullock, D.M., 2021. Using Anonymous Connected Vehicle Data to Evaluate Impact of Speed Feedback Displays, Speed Limit Signs and Roadway Features on Interstate Work Zones Speeds. *J. Transp. Technol.* 11, 545–560. <https://doi.org/10.4236/jtts.2021.114034>
- Matveev, A., Maximov, A., Bogdanova, E., 2020. Intelligent decision support system for transportation emergency response. *Transp. Res. Procedia, XIV International Conference on Organization and Traffic Safety Management in Large Cities (OTS-2020)* 50, 444–450. <https://doi.org/10.1016/j.trpro.2020.10.058>
- Meldrum, D., Taylor, C., 2000. Algorithm design, user interface, and optimization procedure for a fuzzy logic ramp metering algorithm: a training manual for freeway operations engineers.
- Miovision, 2022a. Miovision TrafficLink Traffic Operations [WWW Document]. URL https://help.miovision.com/s/topic/0TO0d000000EoG9GAK/traffic-operations?language=en_US&tabset-a4a08=2 (accessed 10.11.22).
- Miovision, 2022b. Insights Split Failures in Miovision TrafficLink portal [WWW Document]. URL https://help.miovision.com/s/article/Insights-Split-Failures-in-Miovision-TrafficLink-portal?language=en_US (accessed 7.5.23).
- MnDOT, 2019. MnDOT Transportation Systems Management and Operations Implementation Plan.
- Murthy, K., Yang, H. (Frank), 2021. A COST-EFFECTIVE SOLUTION FOR TRUCK PARKING BASED ON ARTIFICIAL INTELLIGENCE.
- NDOT, 2020. NDOT Statewide TSMO Program Plan 2020.
- New Jersey Department of Transportation, 2022. New Jersey Connected Technology Integration and Implementation (NJCTII) | National Operations Center of Excellence [WWW Document]. URL <https://transportationops.org/case-studies/new-jersey-connected-technology-integration-and-implementation-njctii> (accessed 2.8.23).
- NITC, 2018. Smart City PDX Pilot: NITC Researchers Take Connected Vehicle Technology for a Test Run on Portland Streetcar | National Institute for Transportation and Communities

-
- [WWW Document]. URL <https://nitc.trec.pdx.edu/news/smart-city-pdx-pilot-nitc-researchers-take-connected-vehicle-technology-test-run-portland> (accessed 2.13.23).
- ODOT, 2021. ODOT-Operations Program Performance Management Plan.
- Ren, T., Xie, Y., Jiang, L., 2020. Cooperative Highway Work Zone Merge Control Based on Reinforcement Learning in a Connected and Automated Environment. *Transp. Res. Rec.* 2674, 363–374. <https://doi.org/10.1177/0361198120935873>
- Sadek, A.W., Smith, B.L., Demetsky, M.J., 1998. Artificial Intelligence-Based Architecture for Real-Time Traffic Flow Management. *Transp. Res. Rec.* 1651, 53–58. <https://doi.org/10.3141/1651-08>
- Saha Arpita, Chandra Satish, Ghosh Indrajit, 2017. Delay at Signalized Intersections under Mixed Traffic Conditions. *J. Transp. Eng. Part Syst.* 143, 04017041. <https://doi.org/10.1061/JTEPBS.0000070>
- Saldivar-Carranza, E., Li, H., Mathew, J., Hunter, M., Sturdevant, J., Bullock, D.M., 2021a. Deriving Operational Traffic Signal Performance Measures from Vehicle Trajectory Data. *Transp. Res. Rec.* 2675, 1250–1264. <https://doi.org/10.1177/03611981211006725>
- Saldivar-Carranza, E., Mathew, J.K., Li, H., Bullock, D.M., 2021b. Roundabout Performance Analysis Using Connected Vehicle Data. *J. Transp. Technol.* 12, 42–58. <https://doi.org/10.4236/jtts.2022.121003>
- Saldivar-Carranza, E.D., Li, H., Bullock, D.M., 2021a. Diverging Diamond Interchange Performance Measures Using Connected Vehicle Data. *J. Transp. Technol.* 11, 628–643. <https://doi.org/10.4236/jtts.2021.114039>
- Saldivar-Carranza, E.D., Mathew, J.K., Li, H., Hunter, M., Platte, T., Bullock, D.M., 2021b. Using Connected Vehicle Data to Evaluate Traffic Signal Performance and Driver Behavior after Changing Left-turns Phasing, in: 2021 IEEE International Intelligent Transportation Systems Conference (ITSC). Presented at the 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), pp. 4028–4034. <https://doi.org/10.1109/ITSC48978.2021.9564654>
- Sekula, P., Marković, N., Vander Laan, Z., Sadabadi, K.F., 2018. Estimating historical hourly traffic volumes via machine learning and vehicle probe data: A Maryland case study. *Transp. Res. Part C Emerg. Technol.* 97, 147–158. <https://doi.org/10.1016/j.trc.2018.10.012>
- Shoman, M., Amo-Boateng, M., Adu-Gyamfi, Y., 2022. Multi-Purpose, Multi-Step Deep Learning Framework for Network-Level Traffic Flow Prediction. <https://doi.org/10.2139/ssrn.4063434>

-
- Silvano, A.P., Koutsopoulos, H.N., Farah, H., 2020. Free flow speed estimation: A probabilistic, latent approach. Impact of speed limit changes and road characteristics. *Transp. Res. Part Policy Pract.* 138, 283–298. <https://doi.org/10.1016/j.tra.2020.05.024>
- Stolle, S., 2018. Can A.I. Take Over Winter Road Condition Reporting?
- Tian, Z., Ohene, F., Hu, P., 2011. Arterial Performance Evaluation on an Adaptive Traffic Signal Control System. *Procedia - Soc. Behav. Sci.*, 6th International Symposium on Highway Capacity and Quality of Service 16, 230–239. <https://doi.org/10.1016/j.sbspro.2011.04.445>
- Transportation Research Board, 2016. Highway Capacity Manual 6th Edition: A Guide for Multimodal Mobility Analysis. The National Academies Press, Washington, DC. <https://doi.org/10.17226/24798>
- Wai, B., Zhou, W., 2020. Designing and Implementing Real-Time Bus Time Predictions using Artificial Intelligence. *Transp. Res. Rec.* 2674, 636–648. <https://doi.org/10.1177/0361198120947715>
- Wang, S., Huang, W., Lo, H.K., 2019. Traffic parameters estimation for signalized intersections based on combined shockwave analysis and Bayesian Network. *Transp. Res. Part C Emerg. Technol.* 104, 22–37. <https://doi.org/10.1016/j.trc.2019.04.023>
- Wang, Z., Chen, S., Yang, H., Wu, B., Wang, Y., 2016. Comparison of delay estimation models for signalised intersections using field observations in Shanghai. *IET Intell. Transp. Syst.* 10, 165–174. <https://doi.org/10.1049/iet-its.2014.0280>
- Wang, Z., Fu, K., Ye, J., 2018. Learning to Estimate the Travel Time, in: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Presented at the KDD '18: The 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, London United Kingdom, pp. 858–866. <https://doi.org/10.1145/3219819.3219900>
- Xiao, J., Wei, C., Liu, Y., 2018. Speed estimation of traffic flow using multiple kernel support vector regression. *Phys. Stat. Mech. Its Appl.* 509, 989–997. <https://doi.org/10.1016/j.physa.2018.06.082>
- Zhan, X., Li, R., Ukkusuri, S.V., 2015. Lane-based real-time queue length estimation using license plate recognition data. *Transp. Res. Part C Emerg. Technol.* 57, 85–102. <https://doi.org/10.1016/j.trc.2015.06.001>
- Zhang, Z., Li, M., Lin, X., Wang, Y., 2020. Network-wide traffic flow estimation with insufficient volume detection and crowdsourcing data. *Transp. Res. Part C Emerg. Technol.* 121, 102870. <https://doi.org/10.1016/j.trc.2020.102870>

-
- Zhao, Y., Zheng, J., Wong, W., Wang, X., Meng, Y., Liu, H.X., 2019. Various methods for queue length and traffic volume estimation using probe vehicle trajectories. *Transp. Res. Part C Emerg. Technol.* 107, 70–91. <https://doi.org/10.1016/j.trc.2019.07.008>
- Zheng, J., Liu, H.X., 2017. Estimating traffic volumes for signalized intersections using connected vehicle data. *Transp. Res. Part C Emerg. Technol.* 79, 347–362. <https://doi.org/10.1016/j.trc.2017.03.007>
- Zheng, J., Ma, X., Wu, Y.-J., Wang, Y., 2013. Measuring Signalized Intersection Performance in Real-Time With Traffic Sensors. *J. Intell. Transp. Syst.* 17, 304–316. <https://doi.org/10.1080/15472450.2013.771105>