

GEORGIA DOT RESEARCH PROJECT 19-13
FINAL REPORT

**COORDINATED ANTI-CONGESTION CONTROL
ALGORITHMS FOR DIVERGING DIAMOND
INTERCHANGES**



**OFFICE OF PERFORMANCE-BASED
MANAGEMENT AND RESEARCH**

**600 WEST PEACHTREE STREET NW
ATLANTA, GA 30308**

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. FHWA-GA-21-1913	2. Government Accession No. N/A	3. Recipient's Catalog No. N/A	
4. Title and Subtitle Coordinated Anti-Congestion Control Algorithms for Diverging Diamond Interchanges		5. Report Date April 2021	
		6. Performing Organization Code N/A	
7. Author(s) Sam Coogan, Ph.D. https://orcid.org/0000-0003-0495-1535 Makhin Thitsa, Ph.D. https://orcid.org/0000-0002-5448-1183		8. Performing Organization Report No. N/A	
9. Performing Organization Name and Address Georgia Institute of Technology, 225 North Avenue, Atlanta GA 30332 Mercer University, 1501 Mercer University Dr., Macon, GA 31207		10. Work Unit No. N/A	
		11. Contract or Grant No. PI# 0016831	
12. Sponsoring Agency Name and Address Georgia Department of Transportation Office of Performance-based Management and Research 600 West Peachtree Street NW Atlanta, GA 30308		13. Type of Report and Period Covered Final Report Final (July 2019—April 2021)	
		14. Sponsoring Agency Code N/A	
15. Supplementary Notes Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.			
16. Abstract The main objective of this research project is to develop data-driven traffic prediction and control algorithms for congestion management of diverging diamond interchanges (DDIs) and their surroundings. The project proposes an algorithm that leverages recent advances in data-driven modeling, namely, partial least squares multivariate regression, in order to address two fundamental issues: (1) how can future traffic flow through a network be predicted from current flow, and (2) how can these predictions be made robust to sensor fault and errors. Measured traffic flow data exhibit frequent missing values, necessitating a robust regression approach that is minimally affected by such outliers. The project results indicate strong predictive correlation between off-nominal traffic conditions in a single day separated by several hours. For example, off-nominal traffic in the morning—e.g., abnormally high flow in one direction—might correlate with off-nominal traffic in the evening caused by abnormally high flow in the opposite direction as drivers reverse their commute. The statistical methodology proposed in this project is able to determine and exploit this correlation. A proposed use case is employing such predictions over several hours in order to preemptively adjust traffic controllers in anticipation of off-nominal conditions. As case studies, the project considers data from the I-285 and Ashford Dunwoody Road DDI and the I-85 and SR 140/Jimmy Carter Boulevard DDI.			
17. Key Words Mathematical analysis, Statistical Analysis, Diverging Diamond Interchanges		18. Distribution Statement No restrictions. This document is available through the National Technical Information Service, Springfield, VA 22161. Enter any other agency mandated distribution statements. Remove NTIS statement if it does not apply.	
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 121	22. Price Free

GDOT Research Project 19-13

Final Report

COORDINATED ANTI-CONGESTION CONTROL ALGORITHMS FOR DIVERGING
DIAMOND INTERCHANGES

By

Sam Coogan, Ph.D.

Assistant Professor of Electrical and Computer Engineering and Civil and Environmental
Engineering
Georgia Institute of Technology

And

Makhin Thitsa, Ph.D.

Associate Professor of Electrical and Computer Engineering
Mercer University

Georgia Tech Research Corporation

Contract with
Georgia Department of Transportation

In cooperation with
U.S. Department of Transportation
Federal Highway Administration

April 2021

The contents of this report reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Georgia Department of Transportation or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

SI* (MODERN METRIC) CONVERSION FACTORS				
APPROXIMATE CONVERSIONS TO SI UNITS				
Symbol	When You Know	Multiply By	To Find	Symbol
LENGTH				
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
AREA				
in ²	square inches	645.2	square millimeters	mm ²
ft ²	square feet	0.093	square meters	m ²
yd ²	square yard	0.836	square meters	m ²
ac	acres	0.405	hectares	ha
mi ²	square miles	2.59	square kilometers	km ²
VOLUME				
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft ³	cubic feet	0.028	cubic meters	m ³
yd ³	cubic yards	0.765	cubic meters	m ³
NOTE: volumes greater than 1000 L shall be shown in m ³				
MASS				
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
TEMPERATURE (exact degrees)				
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
ILLUMINATION				
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m ²	cd/m ²
FORCE and PRESSURE or STRESS				
lbf	poundforce	4.45	newtons	N
lbf/in ²	poundforce per square inch	6.89	kilopascals	kPa
APPROXIMATE CONVERSIONS FROM SI UNITS				
Symbol	When You Know	Multiply By	To Find	Symbol
LENGTH				
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
AREA				
mm ²	square millimeters	0.0016	square inches	in ²
m ²	square meters	10.764	square feet	ft ²
m ²	square meters	1.195	square yards	yd ²
ha	hectares	2.47	acres	ac
km ²	square kilometers	0.386	square miles	mi ²
VOLUME				
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m ³	cubic meters	35.314	cubic feet	ft ³
m ³	cubic meters	1.307	cubic yards	yd ³
MASS				
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
TEMPERATURE (exact degrees)				
°C	Celsius	1.8C+32	Fahrenheit	°F
ILLUMINATION				
lx	lux	0.0929	foot-candles	fc
cd/m ²	candela/m ²	0.2919	foot-Lamberts	fl
FORCE and PRESSURE or STRESS				
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in ²

TABLE OF CONTENTS

EXECUTIVE SUMMARY	1
Task I: Literature Survey.....	1
Task II: Modeling and Data Analysis.....	3
Task III: Algorithm Development and Algorithm in Pseudocode	4
Task IV: Simulated Testbed for Algorithm Training and Evaluation.....	5
Task V: Interactive Web Engine Decision Support Tool.....	6
INTRODUCTION	8
LITERATURE SURVEY.....	10
DDI State of the Art	11
Experience of the First DDI in the United States	12
Missouri's Experience	14
Performance Evaluation Studies	15
Signal Timing.....	18
Safety.....	20
Miscellaneous.....	23
Conclusions	24
Active DDIs in the United States	24
MODELING AND DATA ANALYSIS.....	28
Detecting Anomalies from Traffic Data.....	29
Methodology	30
Results	32
Traffic Prediction	43
Methodology	43
Results	45
Optimal Time Segmentation for Time-of-day Plans	48
Methodology	49
Results	54
Conclusions	56
ALGORITHM DEVELOPMENT AND ALGORITHM IN PSEUDOCODE.....	57
Remarks on Corrupt Training Data	60
Robust Principal Component Analysis.....	62
Robust SIMPLS Methods and Results	63

Predictive Accuracy	65
Comparison to Various Data-driven Methods.....	69
Implications of Robust SIMPLS	70
ALGORITHM TRAINING AND TESTING IN SYNCHRO/SUMO ENVIRONMENT.....	73
Overview of Simulation Environment	73
Conversion from Synchro to SUMO.....	75
Geometry Conversion.....	76
Signal Timing and Phasing Conversion	77
Preparation of Traffic Data for Simulation	79
Testing and Training of Algorithm Using SUMO	80
INTERACTIVE WEB ENGINE FOR PREDICTING TRAFFIC FLOW AND LEVEL OF SERVICE.....	102
Overview of the Tool	102
ACKNOWLEDGMENTS	104
REFERENCES	105

LIST OF FIGURES

Figure 1. Scatter plot. Mondays northbound through right at signal 1662.....	34
Figure 2. Scatter plot. Mondays eastbound left at signal 1662.....	34
Figure 3. Scatter plot. Mondays southbound through right at signal 1662.....	35
Figure 4. Scatter plot. Mondays westbound right at signal 1662.	35
Figure 5. Scatter plot. Mondays westbound left at signal 1663.....	36
Figure 6. Scatter plot. Mondays westbound right at signal 1663.....	37
Figure 7. Scatter plot. Mondays westbound total at signal 1663.....	37
Figure 8. Scatter plot. Mondays northbound through at signal 1663.....	38
Figure 9. Scatter plot. Mondays northbound total at signal 1663.....	38
Figure 10. Scatter plot. Mondays eastbound left at signal 1664.....	39
Figure 11. Scatter plot. Mondays eastbound right at signal 1664.....	40
Figure 12. Scatter plot. Mondays eastbound total at signal 1665.....	41
Figure 13. Scatter plot. Mondays northbound total at signal 1665.....	41
Figure 14. Scatter plot. Mondays westbound total at signal 1665.....	42
Figure 15. Scatter plot. Mondays southbound total at signal 1665.....	42
Figure 16. Scatter plot. Mondays total at signal 1665.	43
Figure 17. Diagram. Partitioning of the data matrix for SIMPLS.....	45
Figure 18. Line graph. PLS prediction for the southbound through right movement at intersection 1662.....	46
Figure 19. Line graph. PLS prediction for the northbound through movement at intersection 1663.....	47
Figure 20. Line graph. PLS prediction for the eastbound left movement at intersection 1664. ...	47
Figure 21. Line graph. PLS prediction for the northbound through movement at intersection 1665.....	48
Figure 22. Algorithm 1. Optimal TOD segmentation via dynamic programming.	52

Figure 23. Line graphs. Optimal TOD segments using the original fit equation.	55
Figure 24. Line graphs. Optimal TOD segments using the modified fit.	55
Figure 25. Algorithm 2. Robust SIMPLS.	64
Figure 26. Photo. Map of a DDI and its adjacent intersections at I-85 and Jimmy Carter Blvd. in Norcross, GA.	65
Figure 27. Line graph. Northbound through right movements at intersection 1662 predicted between 10:00 a.m. and 12:00 p.m.	67
Figure 28. Line graph. Westbound total movements at intersection 1663 predicted between 10:00 a.m. and 12:00 p.m.	67
Figure 29. Line graph. Eastbound left movements at intersection 1664 predicted between 10:00 a.m. and 12:00 p.m.	68
Figure 30. Line graph. Northbound through movements at intersection 1665 predicted between 10:00 a.m. and 12:00 p.m.	68
Figure 31. Line graph. Predictions produced by Robust SIMPLS, an NN, and the average traffic volume for October 3, 2019.	71
Figure 32. Bar graph. Comparison of MAE produced by Robust SIMPLS, SIMPLS, an NN, and average predictions for October 3, 2019.	71
Figure 33. Line graph. Predictions produced by Robust SIMPLS, an NN, and the average traffic volume for October 1, 2019.	72
Figure 34. Bar graph. Comparison of MAE produced by Robust SIMPLS, SIMPLS, an NN, and average predictions for October 1, 2019.	72
Figure 35. Photo. Aerial view of the DDI at I-285 and Ashford Dunwoody Rd.	74
Figure 36. Map. DDI at I-285 and Ashford Dunwoody Rd. with Intersection Names.	74
Figure 37. Diagram. Overview of the conversion process from Synchro to SUMO.	75
Figure 38. Diagram. DDI at I-285 and Ashford Dunwoody Rd. in Synchro.	77
Figure 39. Diagram. DDI at I-285 and Ashford Dunwoody Rd. in SUMO.	77
Figure 40. Diagram. Ring-barrier diagram—the basis of Synchro’s phasing data.	78
Figure 41. Screen capture. Phasing scheme as portrayed in SUMO.	79
Figure 42. Scatter plot. Volume by Movement over day for Case Study 1.	81

Figure 43. Bar graph. Average Volume Difference by Movement for Case Study 1.....	82
Figure 44. Scatter plot. Volume by Movement over day for Case Study 2.....	83
Figure 45. Bar graph. Average Volume Difference by Movement for Case Study 2.....	83
Figure 46. Scatter plot. Volume by Movement over day for Case Study 3.....	84
Figure 47. Bar graph. Average Volume Difference by Movement for Case Study 3.....	85
Figure 48. Scatter plot. Volume by Movement over day for Case Study 4.....	86
Figure 49. Bar graph. Average Volume Difference by Movement for Case Study 4.....	86
Figure 50. Scatter plot. Volume over day by Case Study for 8173 Eastbound Left Movement.....	87
Figure 51. Scatter plot. Volume over day by Case Study for 8173 Eastbound Through-Left Movement.....	87
Figure 52. Scatter plot. Volume over day by Case Study for 8173 Eastbound Right Movement.....	88
Figure 53. Scatter plot. Volume over day by Case Study for 8173 Westbound Left Movement.....	88
Figure 54. Scatter plot. Volume over day by Case Study for 8173 Northbound Left Movement.....	89
Figure 55. Scatter plot. Volume over day by Case Study for 8173 Northbound Through Movement.....	89
Figure 56. Scatter plot. Volume over day by Case Study for 8173 Southbound Left Movement.....	90
Figure 57. Scatter plot. Volume over day by Case Study for 8173 Southbound Through Movement.....	90
Figure 58. Scatter plot. Volume over day by Case Study for 8172 Westbound Left Movement.....	91
Figure 59. Scatter plot. Volume over day by Case Study for 8172 Westbound Right Movement.....	91
Figure 60. Scatter plot. Volume over day by Case Study for 8172 Northbound Through Movement.....	92
Figure 61. Scatter plot. Volume over day by Case Study for 8172 Southbound Through Movement.....	92

Figure 62. Scatter plot. Volume over day by Case Study for 8172 Southbound Right Movement	93
Figure 63. Scatter plot. Volume over day by Case Study for 8171 Northbound Through Movement	93
Figure 64. Scatter plot. Volume over day by Case Study for 8171 Southbound Through Movement	94
Figure 65. Scatter plot. Volume over day by Case Study for 8122 Northbound Through Movement	94
Figure 66. Scatter plot. Volume over day by Case Study for 8122 Southbound Through Movement	95
Figure 67. Bar graph. Average Flows by Movement for Prediction Period for Case Study 1	96
Figure 68. Bar Graph. Percentage Error between Actual and Predicted Flows for Case Study 1.....	96
Figure 69. Bar Graph. Average Flows by Movement for Prediction Period for Case Study 2 ...	97
Figure 70. Bar Graph. Percentage Error between Actual and Predicted Flows for Case Study 2.....	97
Figure 71. Bar Graph. Average Flows by Movement for Prediction Period for Case Study 3 ...	98
Figure 72. Bar Graph. Percentage Error between Actual and Predicted Flows for Case Study 3.....	98
Figure 73. Bar Graph. Average Flows by Movement for Prediction Period for Case Study 4 ...	99
Figure 74. Bar Graph. Percentage Error between Actual and Predicted Flows for Case Study 4.....	99
Figure 75. Screen capture. Sample from the web engine.	103

LIST OF TABLES

Table 1. Crash reduction percent at the project level.....	21
Table 2. Crash reduction percent in the site-specific analysis.	21
Table 3. All DDIs active in the United States by the year 2012.....	25
Table 4. All DDIs active in the United States by the year 2015.....	26
Table 5. Number of DDIs active in 2015 by location.....	27
Table 6. Outliers for Mondays at signal 1662: SR 140 and Crescent Dr./Goshen Springs Rd. ..	33
Table 7. Outliers for Mondays at signal 1663: SR 140 and I-85 southbound.	36
Table 8. Outliers for Mondays at signal 1664: SR 140 and I-85 northbound.....	39
Table 9. Outliers for Mondays at signal 1665: Jimmy Carter Blvd. and Dawson Blvd./Live Oak Pkwy.....	40
Table 10. Summary of intersections examined in this study.	65
Table 11. Performance measures for Robust SIMPLS in various prediction scenarios.	69
Table 12. Total Delay, Average Delay, and Travel Time Savings for each Case Study.	101

EXECUTIVE SUMMARY

The main objective of this research project is to develop data-driven traffic prediction and control algorithms for congestion management of diverging diamond interchanges (DDIs) and their surroundings. The secondary goal is to report the research findings of the project to the Georgia Department of Transportation (GDOT) and disseminate them to the transportation research community, as well as the general public, through various publications and communication channels. Research objectives were achieved by undertaking the following tasks:

- I. Literature survey.
- II. Modeling and data analysis.
- III. Algorithm development.
- IV. Simulated testbed for algorithm training and evaluation development.
- V. Interactive web engine decision support tool construction.

TASK I: LITERATURE SURVEY

A thorough literature review was conducted to survey the state of the art of the current research on DDIs and traffic control schemes with a special focus on DDIs. The purpose of Task I was to help shape the scope and direction of the research in this project so that it would add the most value to the existing knowledge. The research team especially focused on two areas during the literature review:

- Best practice reports developed by state DOTs, which help home in areas that other DOTs have found challenging or difficult.

- Theoretical findings that have been published recently at venues such as Transportation Research Board (TRB) annual meetings.

Such a focus on the most recent research helped this project pursue truly novel directions and not cover existing results, while allowing the researchers to thoroughly examine and adopt any that were promising. According to this literature survey, the most common challenges reported by other DOTs and the current literature are the following:

- **Challenges posed by proximity to adjacent intersections** – Near-by intersections have been found to hinder the ability of the DDI to process traffic as efficiently as intended. Moreover, the mismatch in traffic processing efficiency between the DDI and adjacent intersections can result in queue spillbacks and demand starvations, and, therefore, can present increased operational challenges at the adjacent intersections upstream and downstream.
- **Unique features of individual DDIs** – DDIs have unique features and characteristics, including multimodal considerations, safety performance, operations, geometric design, spatial requirements, constructability, and maintenance. Therefore, a one-size-fits-all timing plan may be impossible.
- **Adjusting signal timings** – Currently, field engineers must manually adjust the signal timings since software tools, such as Synchro®(manufactured by CUBIC Trafficware), prove to be inefficient for this purpose. Rigorous studies on signal phasing schemes are needed in order to provide practitioners in the field with guidance to adjust signal timings. DDI studies in transportation literature are scarce; most focus on performance

evaluation of newly constructed DDIs, while the studies on signal phasing schemes are especially sparse.

TASK II: MODELING AND DATA ANALYSIS

Eight months of data were collected for two signals within the Norcross, Georgia, DDI at Jimmy Carter Blvd. (signal 1663, signal 1664) and two adjacent signals, one upstream and one downstream (signal 1662, signal 1665). Then, a thorough data analysis was conducted to extract traffic trends, as well as to construct data-driven models that can make predictions on future traffic flows. The purpose of Task II was to build the foundation for the development of a traffic predictive control algorithm. The information extracted and the models developed from data during this task were in Task III (algorithm development) for synthesizing algorithms that coordinate mainline traffic with downstream intersections and freeway exit ramp traffic. The following goals were accomplished during this task:

- Principal component analysis (PCA) was performed to detect anomalies in traffic trends. PCA is a powerful data visualization technique that helps extract the outliers in a large data set. This analysis tool was used to separate the days for which traffic trends are drastically different from the usual patterns and exclude them from the historical data later used to construct data-driven models. Also, analyzing the unusual flows and their characteristics helps determine how to adjust the signal timings when such trends are detected again in the future. The results demonstrated that the method was able to extract important features and characteristics of the traffic flow and detect anomalous trends effectively.

- Models were constructed from historical data using SIMPLS regression, which is an advanced regression method, and the technical methodology of which is detailed in this report. Utilizing the SIMPLS models, the research team was able to predict the afternoon and evening traffic from the morning data. However, the model parameters can be easily adjusted depending on the prediction time selected by the practitioner; thus, making predictions on new intersections and over different time periods is easily implementable by these methods. Being able to predict the future traffic flow helps traffic engineers to adjust the time-of-day (TOD) plans accordingly, as well as prepare for necessary arrangements if the prediction indicates that congestion is imminent. The algorithm used was able to make excellent predictions that match extremely well with the actual traffic.
- Optimal time segmentation of the daily traffic data was performed by dynamic programming. This process helped the researchers identify the optimal switching times between the TOD plans. This method allows the practitioner to select whether to prioritize the arterial traffic or the off-ramp traffic, depending on their respective volumes, and time segmentation is performed accordingly. Moreover, predicted future traffic flows can be incorporated into time segmentation to inform the practitioner when to continue with the current timing plan and when to adjust the switching time.

TASK III: ALGORITHM DEVELOPMENT AND ALGORITHM IN PSEUDOCODE

The information extracted and the models developed from Task II were used for synthesizing algorithms that coordinate mainline traffic with downstream intersections and freeway exit ramp traffic. During Task II (modeling and data analysis), the research team identified the issue that the turning movement count data obtained from GDOT's Automated Traffic Signal Performance Measures (ATSPM) website were highly corrupted. Therefore, the SIMPLS model constructed

during that task needed to be robustified against corruption in data. During Task III, the team developed a new data-driven traffic prediction algorithm—dubbed Robust SIMPLS—which performs well even with corrupted training data. The algorithm was first developed in pseudocode and subsequently coded in MATLAB®.

TASK IV: SIMULATED TESTBED FOR ALGORITHM TRAINING AND EVALUATION

To evaluate the Robust SIMPLS algorithm, a simulation testbed was developed using Synchro and SUMO (Simulation of Urban Mobility) software. The purpose of the simulation environment was twofold; it was intended to train the algorithm using different traffic demand profiles and to test the effectiveness of the recommendations made by the algorithm. Historical data from GDOT’s ATSPM website were used to train the prediction algorithm. In simulated scenarios where intervention was recommended by the prediction algorithm, SUMO was used to demonstrate how altering the signal timings and ramp metering would increase efficiency at the interchange and neighboring intersections. To evaluate the algorithm, the team decided to focus on the DDI located at the I-285 interchange with Ashford Dunwoody Rd. and the neighboring signals. GDOT provided the research team with Synchro files for several DDIs in the state of Georgia, including the one at I-285 and Ashford Dunwoody Rd. These Synchro files contained important data regarding the geometry, phasing, and signal plans of intersections within and surrounding the interchange. Based on the software’s simulation capabilities, the research team decided to use SUMO, an open-source, microscopic traffic simulator, rather than Synchro to test the algorithm. The team discovered, however, that there is not a straightforward way to import data from Synchro to SUMO. Thus, conversion files and scripts were created to convert the Synchro geometry and phasing data from Synchro to SUMO. After incorporating geometry,

phasing, volume, and detector data, some issues remained with the simulation, and the research team had to manually tweak some factors, such as road priorities and allowable U-turns, to accurately simulate real-world conditions at the DDI. Ultimately, the research team was able to successfully create a simulation environment for the I-285 and Ashford Dunwoody Rd. DDI in SUMO by converting the geometry and phasing files from Synchro to SUMO. The team was also able to convert available turning count movement data into vehicle routes that were required by SUMO. At this point, the research team was prepared to test and train the prediction algorithm using different traffic demand profiles.

TASK V: INTERACTIVE WEB ENGINE DECISION SUPPORT TOOL

This project involved the optimization of traffic flow at diverging diamond interchanges with ramp meters and neighboring intersections. The goal was to demonstrate the feasibility of a decision support tool that can notice unusual trends in the flow and notify GDOT that signal timing adjustments may be needed now or in the future; for example, an unusual morning peak period traffic profile may indicate to expect an unusual evening peak period profile. The utility of human–algorithm collaboration in traffic engineering is paramount. In an effort to close the gap between humans and computers in this collaboration, the research team devised a prototype web engine that integrates the productivity of algorithmic predictions with human oversight and decision-making. The functionality of the web engine is as follows:

1. The engine reads incoming traffic data from a network of intelligent infrastructure.
2. Then, data-driven algorithms provide predictions of future traffic flows/level of service across the network.

3. Lastly, the results of these predictions are displayed in a graphical user interface for use by traffic engineers. Suggestions for updated timing schemes and alerts are provided in this interface, if necessary.

INTRODUCTION

Diverging Diamond Interchanges (DDI) are a relatively new interchange design that moves traffic to the opposite side of the road at a freeway underpass or overpass, allowing vehicles to enter and exit the freeway via unimpeded left-hand turns. This arrangement improves safety, as vehicles do not need to turn left against oncoming traffic to enter the freeway, and has the potential to improve mobility, as this configuration requires only two two-phase traffic intersections. Since its introduction, both safety and mobility ramifications have been studied in simulation and, to a lesser extent, in theory. Since the concept of DDI itself is relatively new, the research devoted to the operational strategies of DDI's is few and far between. The main objective of this project is the optimization of traffic flow at diverging diamond interchanges with ramp meters and neighboring intersections. The primary goal is to develop data-driven traffic prediction and control algorithms for congestion management of diverging diamond interchanges (DDIs) and their surroundings. The secondary goal is to report the research findings of the project to the Georgia Department of Transportation (GDOT) and disseminate them to the transportation research community, as well as the general public, through various publications and communication channels. Research objectives were achieved by undertaking the literature survey, modeling and analysis, algorithm development and simulated testbed development for algorithm training and evaluation. An interactive web engine is also constructed to serve the practitioners as a decision support tool. This report is organized as follows. The literature survey is presented in Chapter 2. In Chapter 3, the modeling and data analysis, such as the methodology on the anomaly detection and traffic prediction, are discussed. In Chapter 4, the traffic prediction algorithm is robustified against high energy one-off errors and noise in the data, and the final traffic prediction algorithm is presented in pseudo code. Chapter 5 discusses the development of

a simulated testbed in Synchro/SUMO simulation for the algorithm. Development of the web engine decision support tool is presented in Chapter 6.

LITERATURE SURVEY

During Task I, the first 4 months of the project, a thorough literature review was conducted to survey the state of the art of the current research on diverging diamond interchanges (DDI) and traffic control schemes—with a special focus on DDIs. The purpose of this task was to help shape the scope and direction of the research in this project so it will add the most value to the existing knowledge. The researchers particularly focused on two areas during this literature review. The first area was best practice reports developed by state Departments of Transportation. This review helped home in on areas that other DOTs have found challenging or difficult. The second area was theoretical findings that have been published recently at venues such as Transportation Research Board (TRB) annual meetings. This focus on current research helped ensure that this project pursued truly novel directions and did not cover existing results, and it allowed the researchers to thoroughly examine and adopt existing results that are promising. According to the literature survey, the most common challenges reported by other DOTs and the current literature are the following:

- Proximity to adjacent intersections – Near-by intersections have been found to hinder the ability of the DDI to process traffic as efficiently as intended. Moreover, the mismatch in traffic processing efficiency between the DDI and adjacent intersections can result in queue spillbacks and demand starvations and, therefore, can present increased operational challenges at the adjacent intersections upstream and downstream.
- Unique features of individual DDIs – DDIs have unique features and characteristics, including multimodal considerations, safety performance, operations, geometric design,

spatial requirements, constructability, and maintenance. Therefore, a one-size-fits-all timing plan may be impossible.

- Signal timing adjustments – Currently, field engineers need to manually adjust the signal timings since software tools, such as Synchro®, prove to be inefficient for this purpose. Rigorous studies on signal phasing schemes are needed in order to provide practitioners in the field with guidance to adjust signal timings. DDI studies in transportation literature are scarce; most studies on DDI focus on the performance evaluation of newly constructed DDIs, while the studies on signal phasing schemes are especially sparse. The literature survey has been divided into five sections: (1) DDI state of the art, where the experiences of other DOTs with DDI are presented with a focus on some of the nation's first DDIs; (2) operational evaluation studies, where a summary of studies on DDI performance evaluated by state agencies, as well as reported in academic literature, is presented; (3) signal timings, where a summary of proposed signal timings for DDIs reported in the literature are surveyed; (4) safety, where a summary of safety-related issues regarding DDIs that are reported in the literature, as well as the literature on safety evaluation of DDIs, are presented; (5) miscellaneous, where a summary of notable studies that do not fit into any of the first four categories is presented.

DDI STATE OF THE ART

In 2009, the first DDI in the United States was built at the intersection of Interstate 44 and Missouri Route 13 (I-44 and MO-13) in the state of Missouri. In 2010, two more DDIs were built in Missouri, while Tennessee and Utah each built one DDI in their own state.^[1] By 2014, Missouri led in the number of DDIs with a total of 13 built or planned, while Utah followed, having a total of 6 DDIs within the state. By 2014, 16 states had built or planned at least 1 DDI.

The state of Georgia's first DDI was built in 2012 in Atlanta at I-285 and Ashford Dunwoody Rd. In 2013, another was built at I-85 and Pleasant Hill Rd. as the second DDI in Georgia.

Currently, there are six active DDIs in Georgia:

- I-285 and Ashford Dunwoody Rd.
- I-85 and Pleasant Hill Rd.
- I-75 and Wade Green Rd.
- I-75 and Windy Hill Rd.
- I-95 and SR 21.
- I-85 and SR 140/Jimmy Carter Blvd.

Experience of the First DDI in the United States

The DDI at I-44 and MO-13 in Springfield, Missouri, was the first intersection of its type constructed in the United States. In February 2011, a post-construction performance evaluation report was published on the nation's first DDI.^[2] In that report, the DDI's performance was evaluated by assessing three criteria: (1) traffic operations (2) safety, and (3) public perceptions.

Traffic Operations

The main takeaway from the 2011 evaluation report is that the DDI exhibits the potential for better management of increased traffic volumes compared to a conventional diamond interchange. This prediction is based on the observations during peak travel periods or when an over-dimension load negotiated the DDI. The recovery from traffic backups created by these conditions were normally eliminated within one to two signal cycle lengths. Over-dimension loads up to 18 ft wide and 200 ft long successfully moved through the DDI. It was also observed that left-turn movements within the DDI experienced a noticeable decrease in traffic delay and

traffic queuing. Through movements within the DDI experienced a slight increase in travel time in part due to the slow speed through the crossover areas during off-peak periods. However, the overall traffic flow through the DDI was concluded to be better.

Safety

Safety conclusions included the following. Total crashes were down by 46 percent in the first year of operation. Left-turn crashes were eliminated, and left-turn right-angle crashes were down 72 percent due to the way the left turns are handled within the DDI by free-flow movements or yield control. Rear-end crashes were down slightly as a result of how left turns are handled. The DDI's post-construction crash types were similar to any other signalized intersection, and no definite crash pattern was noticed. Therefore, the first DDI's operation was concluded to be safe, with a decrease in total crash number and no increase in any specific crash pattern attributed to the DDI.

Public Perception

The stakeholders considered in evaluating public perception included the general public, pedestrians, and operators of bikes and larger vehicles. A very high percentage (80 percent plus) of the general public expressed that traffic flow had improved and traffic delay had decreased. A very high percentage (87 percent) expressed that crashes were more likely to occur within a standard diamond when compared to a DDI. A very high percentage (around 80 percent) expressed that larger vehicles and pedestrian/bike movements through the DDI were better or similar to a standard diamond interchange. A very high percentage (91 percent) expressed good understanding on how the interchange operated with the current design of islands, signing, signals, and pavement markings.

Missouri's Experience

In May 2010, the Missouri DOT (MoDOT) released a document entitled *Missouri's Experience with a Diverging Diamond Interchange: Lessons Learned*.^[3] In this document, MoDOT reported that even though a DDI is not specifically designed to accommodate oversized loads, it operates well and is flexible for the navigation of oversized vehicles. It was observed that if the load can maneuver through a standard diamond interchange, it is able to maneuver through a DDI.

Oversized loads can be guided through by special accommodations, such as stopping other traffic to allow the load to move through or allowing it to use multiple lanes to maneuver. A super load is defined as any vehicle that exceeds 16 ft in height and width and 150 ft in overall length and 160,000 gross pounds. In Missouri, usually the arms of the super load are temporarily removed and replaced after the load has moved through the intersection with coordination by a certified signal contractor. MoDOT recommends structural analysis and a feasibility study for moving through super loads above 160,000 lb. In Missouri, signal heads are designed to be 16 ft above the pavement, and, therefore, signalized intersections can accommodate all other vehicles except super loads. MoDOT cautions that when the cross route passes over, DDIs can present complications for routing over-height loads past bridges with vertical height restrictions that span a limited access highway because the configuration of the ramp movements does not allow for reentry to the limited-access highway at the ramp termini.

MoDOT observed that when signaling the off-ramp lefts, the clearance intervals for the crossover throughs can become long, especially if the distance between the crossover intersection and “distance to clear” of the off-ramp left turns is significant. This increases the yellow and all-red intervals. Moreover, MoDOT reported that placing the pedestrian accommodations in the middle of the cross route and running pedestrian phases concurrently

with nonconflicting vehicle phases can help with achieving shorter clearance intervals.

According to MoDOT, the use of two controllers is recommendable in the case of an expansion of the interchange, for example, the addition of an off-ramp turn lane, since this can require that more phases be added. Moreover, two controllers may be the only option if there is a voltage drop due to the distance from the farthest signal head or the detecting device to the signal cabinet. However, using only one controller also offers advantages, such as elimination of communication issues between the two controllers, ensuring the two through movements approaching the DDI on the cross route are never on at the same time. Therefore, MoDOT recommends a careful consideration of deploying one controller versus two controllers depending on what movements are being signalized. Using software such as Synchro to optimize the signals for the DDI has proven inefficient for MoDOT; therefore, currently the signal timing is optimized manually.

PERFORMANCE EVALUATION STUDIES

Since DDIs are relatively new structures, studies on evaluation of the operational performance of the newly constructed DDIs have emerged. In most studies, the performance is evaluated based on the number of stops, total green time, approach capacity, and critical lane volume. Hunter et al. recently conducted one of these studies on the operational performance of DDIs in the state of Georgia.^[4] Their study included a sensitivity analysis of conventional diamond interchange (CDI) and DDI operational performance under various interchange lane configurations, including the selected study area of the I-85 and Jimmy Carter Blvd. interchange in Norcross, Georgia, under varying traffic demands and turn-movement ratios. The study sought to determine conditions in which one interchange configuration provides superior performance over the other. In that study, the sensitivity analysis was structured into a two-step process with a

critical lane volume (CLV) analysis as the first step, followed by a VISSIM microscopic simulation study as the second step. Hunter et al. concluded that a CDI is likely to be the preferred option at locations with traffic volumes well below capacity and cross-street left-turn traffic proportions below 30 percent of the total cross-street demand, and a DDI is likely to be preferred at locations with traffic volumes near capacity and cross-street left-turn proportions exceeding 50 percent of the total cross-street demand.

The unique geometry of DDI design allows for extra flexibility in signal phasing schemes—particularly with the two-phase operation, using overlaps best leverages the full power of the DDI's merit. In most two-phase operations, internal queue is allowed. Therefore, Pang et al. argued that the conventional capacity calculation methods may not be appropriate for the DDI if the effects of the internal queue need to be considered.^[5] Pang et al. investigated the approach capacity of DDIs, including the effect of internal queue. They developed a method to calculate the approach capacity of the DDI by using an analytical model that would consider the impacts of the internal queue on the traffic progression of either arterial roads or freeway ramps. They demonstrated that the model can be used to reliably calculate lost green time of traffic movements of the DDI. A similar method documented in chapter 22 of the *Highway Capacity Manual 2010* (HCM 2010) was also designed to calculate the lost green time of signalized interchanges due to the impact of an internal queue.^[6] Pang et. al, concluded that the HCM 2010 method is not appropriate for DDIs because the movements feeding into the internal space of a DDI are significantly different from the movements of a conventional signalized interchange.

Other simulation studies have aimed to determine if a DDI is suitable to replace an existing CDI in the same location. A case study was conducted by Khan et al. for an interchange in Athens, Alabama.^[7] The existing interchange was a conventional diamond interchange that was

compared with a DDI built at the same location by using Synchro/SimTraffic® as a simulation tool. The analysis focused on determining the level of service and vehicle delay for the two interchange types. Based on simulations in their study, DDI performed better than CDI in only four special cases and, therefore, it was concluded that a DDI is not an appropriate replacement for the current CDI network for the study area.

Still, other simulation studies have evaluated the performance of signal phasing schemes of DDIs. For example, Warchol et al. investigated the influence of crossover spacing and increased volumes on the performance of DDI phasing schemes using PTV Vistro® software and the dynamic bandwidth assessment tool to optimize the split, cycle length, and offset.^[8] Mean interchange delay and mean stops per vehicle were selected as measures of effectiveness. It was concluded that a two- or three-critical-movement phasing scheme usually resulted in the lowest mean interchange delay and the fewest stops. Overall, the results provide an initial signal timing scheme for practitioners.

Finally, some studies point out inherent inefficiency that resulted from two-phase configuration reflecting the two competing movements at the crossover points at each intersection of the DDI. Hainen et al., for example, identified the issues as: (1) the potential for internal queuing, and (2) the challenges faced when the inflow demand exceeded the outflow capacity of the interchange.^[9] They used high-resolution event data to develop performance measures to evaluate operations at a DDI in Salt Lake City, Utah. Their study investigated alternatives to the existing signal timing with the two-phase configuration, and then developed and deployed a new three-phase configuration. This three-phase configuration was found to address the internal queuing that occurs with two-phase timing. Under this new configuration, the flows from one DDI intersection to the other were balanced, and progression within the DDI was improved. The

study found that with the three-phase configuration, the percent of vehicles arriving on green at the heaviest internal movement within the DDI increased from 53 to 92 percent.

Several additional studies on performance evaluation of DDIs that the current research team surveyed include references 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, and 25.

SIGNAL TIMING

As the popularity of DDIs as the interchange of choice grows, the ability to provide practitioners in the field with guidance on signal timing adjustments becomes increasingly important. As several methods of timing schemes are being proposed in the literature, it becomes necessary to document the state of the practice in DDI signal phasing with principles established in the *Signal Timing Manual*.^[26] Cunningham et al. documented three fundamental phasing schemes where each scheme was described with a consistent naming convention.^[27] It used consistent phase numbering in a logical format to describe each of the phasing concepts being considered and discussed. The study also addressed the use of actuation and barriers to improve coordination, methods for reducing lost time, and issues relating to preemption and pedestrians to provide practitioners with basic guidance on when one phasing scheme may be more appropriate than another.

One of the key factors in enhancing the operating efficiency of DDI is being able to successfully coordinate the DDI along with neighboring intersections in an arterial. One of the studies that focused on this issue was conducted by Day and Bullock.^[28] Their study examined three different cycle length strategies for DDI coordination: (1) using the full cycle length of the corridor; (2) using a half-cycle; and (3) using a previously described three-phase scheme intended to manage queues within the DDI. Through a microsimulation study, six different

origin–destination (O–D) scenarios were tested and the results were presented in terms of the number of stops, movement delays at the DDI, queue lengths, and delay by O–D path. The results of the study indicate that the half-cycle and full-cycle strategies have both advantages and disadvantages. The half-cycle strategy yielded lower total and average delays but resulted in more stops along the arterial and higher arterial O–D path delays. On the other hand, the full-cycle option achieved fewer stops and lower delays for arterial routes, while it increased total and average delays for other movements. The study found that the three-phase strategy often reduced delays and queue lengths for the arterial movements exiting the DDI but increased those of other movements.

Other studies extended the optimization of the offset of signalized arterial using the Link Pivot algorithm to arterials with single-controller DDI. Day et al., for example, sought to optimize a DDI within an existing system to ensure smooth corridor operation, and their study presented a methodology for optimizing offsets on a corridor, including a single-controller DDI.^[29] This methodology uses high-resolution controller data and an enhancement to the Link Pivot algorithm that deconstructs the single-controller parameters into equivalent offset adjustments. Day et al. demonstrated their methodology by applying it to a five-intersection arterial route that included a DDI and assessing the outcomes by measurement of travel times by Bluetooth vehicle re-identification. A user benefit methodology was applied to the travel time data that considered the reliability of the travel times in addition to the central tendency, as well as to O–D paths that travel to and from the freeway in addition to routes along the arterial. The study reported a total annualized user benefit of approximately \$564,000.

Some studies suggested that the relationship between the crossover spacing and the signal offset need to be better understood. Cheng et al. pointed out that in the design of DDI, the two critical

and interdependent design components—crossover spacing and signal offsets—have not been addressed in most design guidelines.^[30] The estimated travel time between a DDI's two sub-intersections for all movement paths is essential for the design of signal offsets. Also, the crossover spacing needs to be designed to accommodate queues comprising mostly those vehicles not moving within the signal progression band, which is often designed with a given crossover spacing. Therefore, those authors proposed concurrent optimization of these two vital DDI design elements at the planning level. They conducted a case study at a DDI site with the proposed model to justify the necessity to perform the concurrent optimization under different operational conditions. They concluded that the design with the optimized crossover spacing and offset can yield the shortest total delay and the least number of stops for vehicles over the entire network, especially under near-saturated conditions. The study also found that the optimized crossover spacing can also prevent the formation of queue spillover across the crossovers in a DDI.

Other studies on signal timing optimization that the current research team surveyed include references 31, 32, 33, 34.

SAFETY

In addition to operational efficiency and signal timing optimization studies, research efforts have been devoted to safety evaluation studies. Claros et al. conducted a formal safety evaluation at the project level (i.e., interchange) and the site-specific level (i.e., ramp terminals) of DDIs.^[35] The investigators used three types of before–after evaluation methods: naïve, empirical Bayes (EB), and comparison group (CG). These three evaluation methods involved different trade-offs, such as data requirements, complexity, and regression to the mean. The safety evaluation at the

project level accounted for the influence of the DDI treatment in the entire footprint of the interchange; the site-specific approach focused on the influence at the ramp terminals only. All three methods showed that a DDI replacing a CDI decreased crash frequency for all severities. The percents of reduction in crash data presented in their report are summarized in table 1 and table 2.

Table 1. Crash reduction percent at the project level.

	Naive	Empirical Bayes (EB)	Comparison Group (CG)
Fatal or Injury Crashes	63.2%	62.6%	60.6%
Property Damage Only Crashes	33.9%	35.1%	49.0%
Total Crash Frequency	41.7%	40.8%	52.9%

Table 2. Crash reduction percent in the site-specific analysis.

	Naive	Empirical Bayes (EB)	Comparison Group (CG)
Fatal or Injury Crashes	64.3%	67.8%	67.7%
Property Damage Only Crashes	35.6%	53.4%	47.0%
Total Crash Frequency	43.2%	56.6%	53.3%

The concern for safety is typically higher in interchange traffic due to complex traffic patterns resulting from diverging, merging, and frequent lane changing. Nye et al. conducted a national-level safety evaluation of DDIs in the U.S.^[36] Their study updated the previous evaluations and expanded the treatment group size of previous studies to provide a more robust and reliable safety assessment of DDI deployments. They used a crash modification factor (CMF) to estimate the ability of the countermeasure to reduce the number of crashes and the crash severity. Nye et al. reported a total CMF of 0.633 based on the comparison group analysis of 26 DDIs in

11 states. The CG method was also applied to a variety of crash variables for their study. Angle, rear-end, and sideswipe crashes were found to have CMFs of 0.441, 0.549, and 1.139, respectively. Fatal-and-injury crashes provided a CMF of 0.461. Daytime and nighttime crashes provided CMFs of 0.648 and 0.638, respectively. Their study determined that, of the observational before–after evaluation methodologies, the comparison group approach yielded the best evaluation results.

One of the safety evaluation studies on interchange traffic management conducted outside of the U.S. was a study of driver behavior in a freeway interchange.^[37] Although the study did not specifically focus on DDIs, some of the results and conclusions can benefit future research on DDI safety considerations. In that study, Wang et al. analyzed driver behavior characteristics in low-volume freeway interchanges by conducting a field experiment on the Qingyin Expressway in China. Four freeway interchanges with relatively low volume were selected, and 12 qualified drivers (i.e., 6 car test drivers and 6 truck test drivers) drove vehicles according to the driving program. GPS and an eye-tracking instrument were employed to record parameters such as running speed in real time, running track, and fixation point. Wang et al. examined the drivers' fixation on exit guide signs, and their behavior in diverging and merging areas. Their results indicated that: (1) drivers recognize the exit direction signs at 170–180 m advance distance; (2) the diverging influence area is 1000 m upstream of the diverge point, while the merging influence area is 350 m downstream of the merge point; and (3) a NO OVERTAKING sign is recommended to be placed 350 m upstream of the diverging point.

One of the most important public safety concerns at DDIs is the risk of vehicles traveling down the wrong direction. Vaughan et al. investigated the long-term monitoring of wrong-way maneuvers at DDIs.^[38] Five DDIs were continuously monitored for 6 months to observe rates

and causes of wrong-way maneuvers by installing video cameras at each DDI and postprocessing the video by using video detection software. Their analysis showed that wrong-way maneuvers tended to occur more often when vehicles were first entering the DDI rather than after correctly moving through the first crossover and then going the wrong way at the second crossover that they approached (traveling outbound). Also, wrong-way maneuvers were found to occur more frequently at night than during the day. Their findings show that DDIs do experience a certain level of wrong-way maneuvers; however, no crashes could be identified from safety records that were associated with these events. Therefore, the study concluded that DDIs have generally proven to be safe and efficient movers of traffic when designed appropriately.

Other studies surveyed that address the safety of DDIs include references 39, 40, 41, 42.

MISCELLANEOUS

In this section, a number of notable works that do not fit into any of the above categories are presented. For example, Maji et al. developed a planning tool for an efficient way to evaluate DDIs by using the CLV as a criterion.^[43] Their study is different from the studies presented previously because its focus was not about the evaluation of a particular DDI but rather development of an easy and inexpensive way to evaluate DDIs in general. In a similar study on developing evaluation tools, Anderson et al. proposed discrete event simulation as a tool to evaluate DDIs.^[44] In a study by Duan and Abbas, the investigators conducted an analytic calculation of delay in DDIs in order to optimize the timing plans more efficiently.^[45] In another study, Yeom et al. pointed out that most of the lane use analysis models currently being used for lane configuration were originally developed for conventional interchanges.^[46]; thus, their study proposed unique multi-regime lane utilization models.

CONCLUSIONS

The research and reports discussed in this literature survey highlight the current progress and challenges of diverging diamond interchanges. In surveying the literature, the research on DDIs in the transportation field is still in its infancy, with research devoted to any aspect of DDIs being very limited. Coordinated control strategies for DDIs and their adjacent intersections have not been sufficiently explored in the transportation literature. No studies of utilizing closed-loop feedback control strategies have been reported. Therefore, from this literature survey, we concluded that the results produced from this project can be of great consequence to the field of transportation research.

Active DDIs in the United States

In this section, table 3 and table 4 summarize the active DDIs in the U.S. as of 2012 and 2015, respectively, with table 5 providing the location of the active DDIs in 2015.

Table 3. All DDIs active in the United States by the year 2012.

DDIs built in 2009: 1	
Interchange	Location
I-44 & MO-13	Springfield, MO
DDIs built in 2010: 4	
Interchange	Location
US-60 & National Ave.	Springfield, MO
I-270 & Dorsett Rd.	St. Louis, MO
US-129 & Bessemer St.	Alcoa, TN
I-15 & Pioneer Crossing/Main St.	American Fork, UT
DDIs built in 2011 : 5	
Interchange	Location
Harrodsburg Rd. & New Circle Rd. Lexington, KY	
US-65 & MP-248 Branson, MO	
I-15 & S 500 E St. American Fork, UT	
I-15 & SR-92 (Timpanogos Hwy.) American Fork, UT	
SR-201 & Bangerter Hwy. West Valley City, UT	
DDIs built in 2012 : 8	
Interchange	Location
I-285 & Ashford Dunwoody Rd.	Atlanta, GA
MD 295 & Arundel Mills Blvd.	Hanover, MD
US-67 & Columbia St.	Farmington, MO
I-435 & Front St.	Kansas City, MO
MO-150 & Botts Rd.	Kansas City, MO
US-65 & East Chestnut Expy.	Springfield, MO
I-580/US-395 & Moana Ln.	Reno, NV
I-590 & Winton Rd.	Rochester, NY

Table 4. All DDIs active in the United States by the year 2015.

DDIs built in 2013: 15	
Interchange	Location
I-85 & Pleasant Hill Rd.	Duluth, GA
U.S. 91/Yellowstone Hwy. & I-86	Chubbs, ID
I-35 & Homestead Ln.	Edgerton, KS
I-494 & 34th St.	Bloomington, MN
Highway 15 & CR. 120	St. Cloud, MN
US 52 & New Olmstead CR. 12	Oronoco, MN
I-70 & Stadium Blvd.	Columbia, MO
I-44 & Range Line Rd.	Joplin, MO
US 60 & MO 13	Springfield, MO
I-70 & Mid River's Mall Dr.	St Peter's, MO
I-270 & Roberts Rd.	Columbus, OH
I-95 & Thurbers Ave.	Providence, RI
I-15/84 & US-91	Salt Lake City, UT
I-15 & St. George Blvd.	St. George, UT
I-25 & College Dr.	Cheyenne, WY
DDIs built in 2014: 3	
Interchange	Location
I-70 & US 6/US 50	Grand Junction, CO
I-70 & Woods Chapel Rd.	Blue Springs, MO
I-64 & US 15	Zion Crossroads, VA

Table 5. Number of DDIs active in 2015 by location.

By Location	Number
CO, Colorado	1
GA, Georgia	2
ID, Idaho	1
KS, Kansas	1
KY, Kentucky	1
MD, Maryland	1
MN, Minnesota	3
MO, Missouri	13
NV, Nevada	1
NY, New York	1
OH, Ohio	1
RI, Rhode Island	1
TN, Tennessee	1
UT, Utah	6
VA, Virginia	1
WY, Wyoming	1
Total	36

MODELING AND DATA ANALYSIS

During Task I (months 1 to 4) of the project, a thorough literature review was conducted to survey the state of the art of the current research on DDIs and traffic control schemes. This initial task served to shape the scope and direction of this project's research so that the work of subsequent tasks could add the most value to the existing knowledge. During Task II, 8 months of data were collected for two signals within the DDI at Jimmy Carter Blvd. (signal 1663, signal 1664) and two adjacent signals, one upstream and one downstream (signal 1662, signal 1665). Then, a thorough data analysis was conducted to extract traffic trends, as well as to construct data-driven models that can make predictions on future traffic flows. The purpose of Task II was to build the foundation for the development of a traffic predictive control algorithm. The information extracted and the models developed from data during Task II were used in Task III (algorithm development) for synthesizing algorithms that coordinate mainline traffic with downstream intersections and freeway exit ramp traffic. The following items were accomplished during Task II:

- Principal component analysis (PCA) was performed to detect anomalies in traffic trends. PCA is a powerful data visualization technique that helps extract the outliers in a large data set. This analysis tool was used to separate the days for which traffic trends were drastically different from the usual patterns and exclude them from the historical data later used to construct data-driven models. Also, analyzing the unusual flows and their characteristics helps determine how to adjust the signal timings when such trends are detected again in the future. Our results demonstrated that the method was able to extract

important features and characteristics of the traffic flow and detect anomalous trends effectively.

- Models were constructed from historical data using SIMPLS regression, an advanced regression method—the technical methodology of which is detailed in this chapter. Utilizing the SIMPLS models, the researchers were able to predict the afternoon and evening traffic from the morning data. However, the model parameters can be easily adjusted depending on the prediction time selected by the practitioner—thus, making predictions on new intersections and over different time periods easily implementable by these methods. Being able to predict the future traffic flow helps the traffic engineers to adjust the time-of-day (TOD) plans accordingly, as well as prepare for necessary arrangements if the prediction indicates that congestion is imminent. The algorithm was able to make excellent predictions that match extremely well with the actual traffic.
- Optimal time segmentation of the daily traffic data was performed by dynamic programming. This helps identify the optimal switching times between the TOD plans. The method used allows the practitioner to select whether to prioritize the arterial traffic or the off-ramp traffic, depending on their respective volumes, and time segmentation is performed accordingly. Moreover, predicted future traffic flows can be incorporated into time segmentation to inform the practitioner when to continue with the current timing plan and when to adjust the switching time.

DETECTING ANOMALIES FROM TRAFFIC DATA

While constructing models from data, it is necessary to exclude outliers from the data set since their characteristics are drastically different from typical data. However, for large data sets, such as 8 months of traffic turning-movement counts for several different movements, it is usually not

possible to identify the outliers from time series plots alone. Principal component analysis is a powerful data visualization tool that allows efficient identification of anomalies in the data.^[47,48]

Methodology

Data on Mondays are used to demonstrate the method. Let D be the number of days for which the historical data were collected, let M be the number of movements, and let T be the number of times the turning movement count was recorded in a day. Then the data are represented as a matrix $A \in C^{D \times TM}$, as follows. Each row of A represents each Monday in the data set. The columns of A are partitioned into M blocks—one block is dedicated to each movement. Within each block, there are T columns and each column contains the traffic count at that particular measurement. It is well known that any matrix A can be decomposed into a product of three matrices, as in equation 1:

$$A = U\Sigma V^*, \quad (1)$$

where U and V are unitary matrices, and Σ can be partitioned into two submatrices: one with all zeroes and the other a diagonal matrix whose diagonal entries carry the singular values of A arranged in the descending order. Therefore, to approximate matrix A with a rank r matrix for any $r > 0$, the best rank r approximation of A denoted by \tilde{A}_r can be computed as in equation 2:

$$\tilde{A}_r = \tilde{U}_r \tilde{\Sigma}_r \tilde{V}_r^*, \quad (2)$$

where $\tilde{U}_r \in C^{D \times r}$, $\tilde{V}_r \in C^{TM \times r}$ are the first r columns of U and V , respectively, and $\Sigma_r \in C^{r \times r}$ is a diagonal matrix with r largest singular values arranged in descending order. The columns of U are the eigenvectors of AA^* and are called the left singular vectors of A . The columns of V are the eigenvectors of A^*A and are called the right singular vectors of A . The columns of U and V are

arranged in descending order of the magnitude of their corresponding singular values. The approximation \tilde{A}_r can be loosely interpreted as the matrix that contains r most important features of A . Therefore, PCA is a powerful method to extract low-rank structures from large-scale data.

Now let a_i be the i^{th} row of A and v_i be the i^{th} column of V and observe, as in equation 3:

$$AV = \begin{bmatrix} < a_1, v_1 > & < a_1, v_2 > & \cdots & < a_1, v_{TM} > \\ < a_2, v_1 > & < a_2, v_2 > & \cdots & < a_2, v_{TM} > \\ \cdots & \cdots & \cdots & \cdots \\ < a_D, v_1 > & < a_D, v_2 > & \cdots & < a_D, v_{TM} > \end{bmatrix}, \quad (3)$$

where $< , >$ denotes an inner product on C^n such that for any $x, y \in C^n$, $< x, y > = x^T y$, which is also known as the *dot product*. Since v_i are unit vectors, $< a_j; v_i >$ represents the magnitude of a_j projected in the v_i direction. In other words, it is the coordinate of the j^{th} row of data in the v_i direction. Now, consider taking the first three columns of AV ; doing so is equivalent to rank r truncation with $r = 3$. Therefore, these three columns contain the coordinates of the data along the three most important directions. Now, if v_1, v_2, v_3 are used as an orthogonal axes of a three-dimensional coordinate system, the data can be visualized effectively in a 3-D scatter plot. In that case, the outliers whose characteristics are drastically different from the rest can be expected to have significantly different coordinates in the three primary directions.

In this case study, data between 06/03/2019 and 02/03/2020 were collected for the two signals within the DDI at Jimmy Carter Blvd. (signal 1663, signal 1664) and two adjacent signals: one upstream and one downstream (signal 1662, signal 1665). Since the traffic trends are highly dependent on the day of the week, the data are separated into seven different matrices, one for each day of the week. Monday data are used to demonstrate the method without losing generality. There are 36 Mondays between these dates and, therefore, the Mondays matrix contains 36 rows. Even though PCA can be performed on the matrix with all blocks of

movements, the analysis reveals that such approach fails to capture the nuances of the individual movements. Therefore, PCA is performed on each movement separately. The researchers obtained the data by download from GDOT's Automated Traffic Signal Performance Measures (ATSPM) website at <https://traffic.dot.ga.gov/ATSPM/>. One issue encountered was that the database of the website erroneously registered zeroes for many days for several movements. The issue was communicated to GDOT, and GDOT informed the research team that the issue would be rectified soon. Meanwhile, the days with missing data are excluded from the analysis.

Results

Once the outliers have been identified by PCA, that data for those particular days can be more closely inspected to determine what is unique about those days. For example, figure 1 depicts the scatter plot of PCA results of Mondays data set for signal 1662 northbound through movements. It can be seen that days 7, 11, and 21 are clear outliers and unusual traffic trends should be expected in northbound through movements on these days. Indeed, on close inspection, it was observed that day 7 exhibited elevated northbound through and right movements between 6:30 a.m. and 1:45 p.m., and day 11 saw large spikes in the morning peak hours followed by a dip at 11:45 a.m. in the same movement. Day 21 turned out to be New Year's Eve, which had drastically different trends from a typical Monday. A thorough analysis of the behavior on each of the unusual days is reported in table 6. The summaries of analyses for signals 1662, 1663, 1664, and 1665 are summarized in table 6, table 7, table 8, and table 9, and their respective scatter plots of PCA analysis are shown in figure 1 through figure 16.

**Table 6. Outliers for Mondays at signal 1662:
SR 140 and Crescent Dr./Goshen Springs Rd.**

Day	Date	Findings:
2	July 24, 2019	The intersection observed abnormally large spikes in westbound right traffic at 12:00 p.m., 3:30 p.m., and 6:00 p.m. and abnormal dips in westbound right traffic at 10:00 a.m., 3:00 p.m., and 5:45 p.m.
5	July 22, 2019	Eastbound left traffic flow is discontinuous throughout the day—producing volumes that are both above and below the average flow.
7	August 5, 2019	Elevated northbound through right traffic volume between 6:30 a.m. and 1:45 p.m. Eastbound left flow saw an abnormally large spike in traffic between 9:30 a.m. and 12:30 p.m. Large spikes in westbound right traffic throughout the day. Large dips in southbound through right movements between 4:45 p.m. and 6:00 p.m.
10	September 9, 2019	Westbound right movements demonstrated a large dip at 11:45 a.m. followed by a sharp rise at 12:00 p.m. Additionally, these movements were above average in the morning.
11	September 16, 2019	On this day, the intersection sees large spikes in northbound through right traffic at 5:45 a.m. and 8:15 a.m. followed by a dip in traffic at 11:15 a.m. Eastbound left traffic was discontinuous throughout the day.
13	October 21, 2019	Increased westbound right movements in the morning and evening times. Increased southbound through right movements between 7:15 a.m. and 9:00 a.m.
14	October 28, 2019	Large spikes in westbound right movements at 3:30 p.m., 4:00 p.m., and 6:30 p.m. Elevated southbound through right movements between 6:30 a.m. and 9:15 a.m.
19	December 16, 2019	This intersection sees multiple 15-minute intervals in the morning and late evening during which no cars performed an eastbound left movement.
21	December 30, 2019	Holiday: New Year's Eve. Unusual northbound through right traffic trend: above average volume between 10:15 a.m. and 1:45 p.m., below average volume between 2:00 p.m. and 9:15 p.m.

1662, Mondays, Northbound Through Right

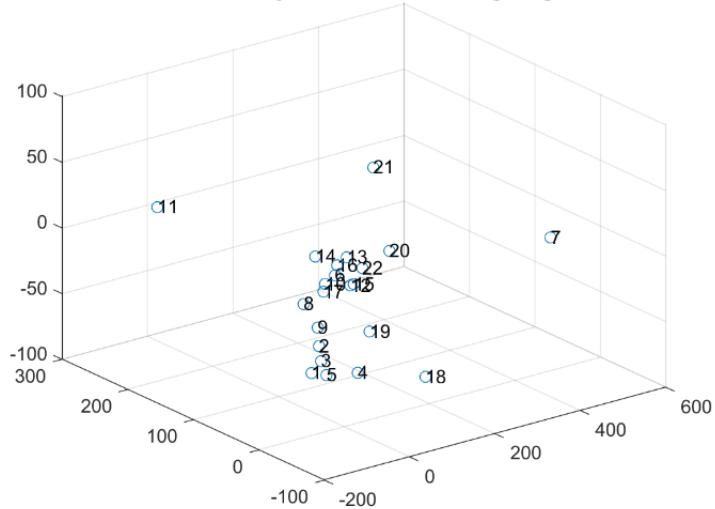


Figure 1. Scatter plot. Mondays northbound through right at signal 1662.

1662, Mondays, Eastbound Left

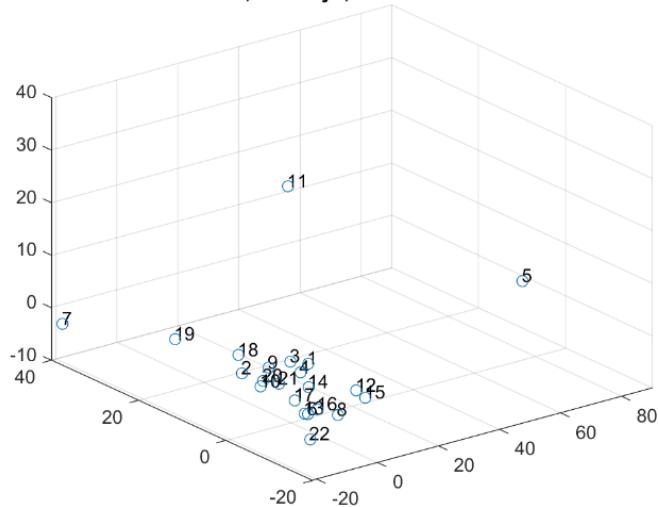


Figure 2. Scatter plot. Mondays eastbound left at signal 1662.

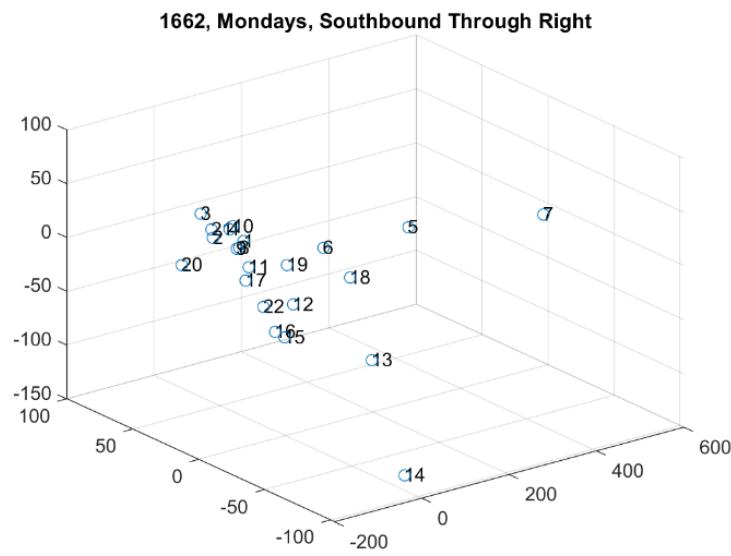


Figure 3. Scatter plot. Mondays southbound through right at signal 1662.

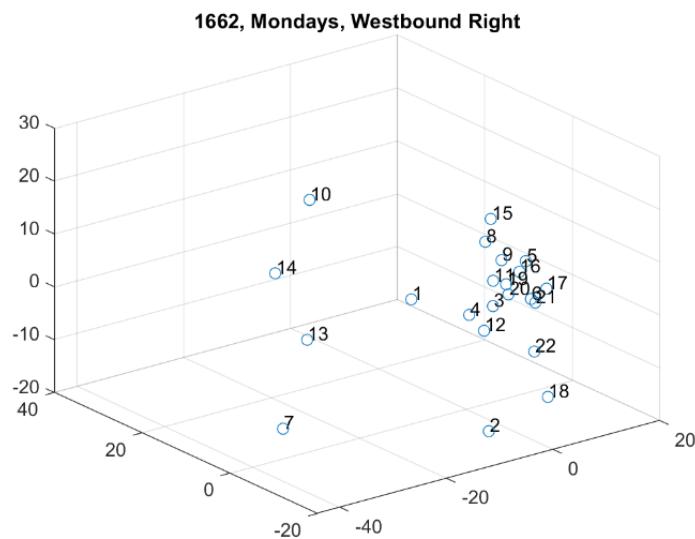


Figure 4. Scatter plot. Mondays westbound right at signal 1662.

**Table 7. Outliers for Mondays at signal 1663:
SR 140 and I-85 southbound.**

Day	Date	Findings:
2	June 10, 2019	Abnormally large spike in volume at 4:00 p.m. and an extremely low volume from 11:15 a.m. to 12:45PM in westbound left. Traffic strays from the usual trends at westbound total with drastically low volume at 9:00 a.m., 11:15 a.m., 1:00 p.m., and during the interval 2:45 p.m. to 3:30 p.m.
13	November 11, 2019	Abnormally large spike at 8:45 a.m. There is an extremely low volume of cars during the interval of 9:15 a.m. to 2:00 p.m. at westbound right.
15	December 2, 2019	At 12:00 a.m., there is an extremely high volume of cars, and this deviation from the trend does not end until 5:00 a.m.
18	December 23, 2019	Holiday weekend. Low volume in late afternoon and evening between the hours of 2:00 p.m. to 3:00 p.m. and 6:00 p.m. to 7:00 p.m. Slightly higher vehicle volume starting at 8:00 p.m.
19	December 30, 2019	Holiday Weekend. The volume of cars deviated from the trend multiple times throughout the day at northbound. Large spikes were observed at 4:00 a.m., 8:30 a.m. until 1:00 p.m., 3:15 p.m., and 4:45 p.m. There was also an abnormally low volume of cars around 4:00 p.m. At westbound total, there was also an abnormal spike between the hours of 6:15 a.m. to 9:30 p.m.

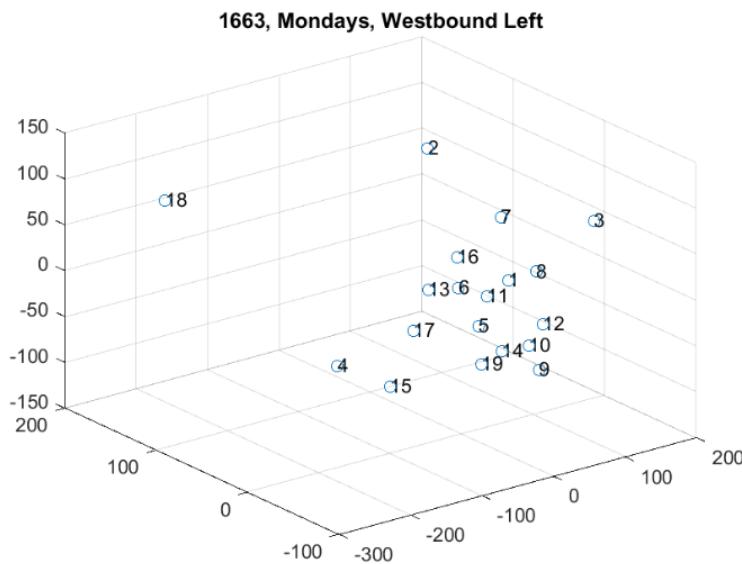


Figure 5. Scatter plot. Mondays westbound left at signal 1663.

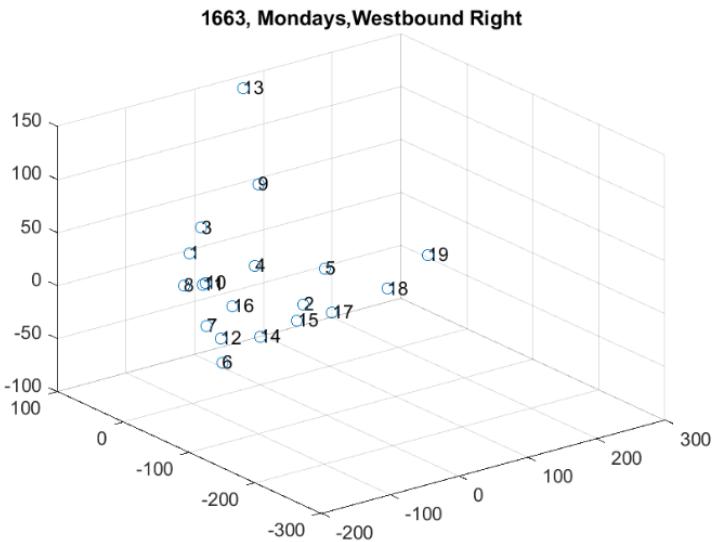


Figure 6. Scatter plot. Mondays westbound right at signal 1663.

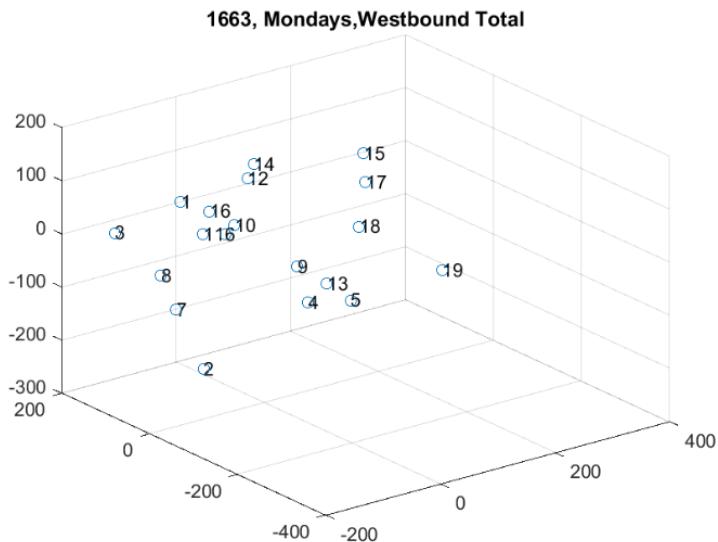


Figure 7. Scatter plot. Mondays westbound total at signal 1663.

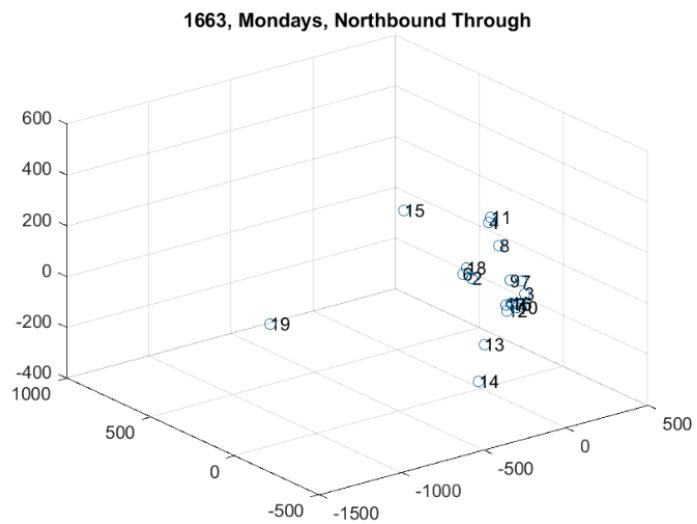


Figure 8. Scatter plot. Mondays northbound through at signal 1663.

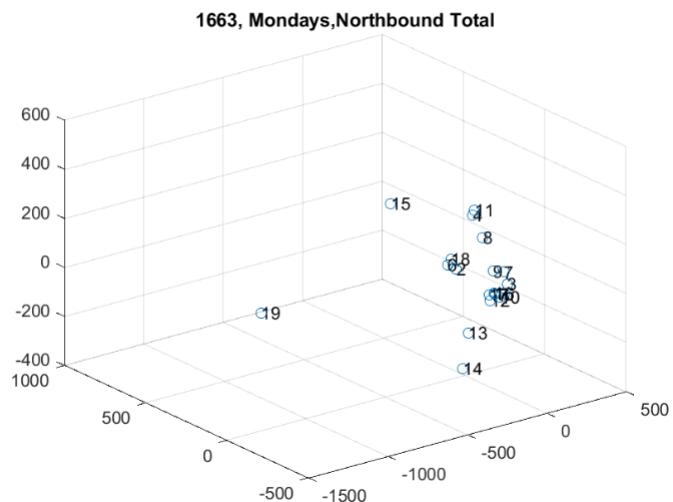


Figure 9. Scatter plot. Mondays northbound total at signal 1663.

**Table 8. Outliers for Mondays at signal 1664:
SR 140 and I-85 northbound.**

Day	Date	Findings:
4	February 10, 2020	After 5:00 p.m., traffic did not increase for rush hours in eastbound right.
7	December 16, 2019	Between 4:30 p.m. and 5:15 p.m., there was a sudden decrease of traffic in eastbound right; during a time in which hundreds of cars should be passing, only a couple dozen were detected.
9	September 2, 2019	Below-average traffic between 4:15 a.m. and 3:15 p.m. for eastbound left; however, a sudden large spike occurred between 5:00 p.m. and 5:15 p.m.
19	December 2, 2019	Large spike to 230 vehicles between 11:15 a.m. and 11:30 a.m. at eastbound left. After 6:30 a.m., the volume of traffic decreased to the point of zero vehicles detected between 7:30 a.m. and 7:45 a.m.
22	December 23, 2019	After an above-average spike at 7:45 a.m., there was a long period of below average traffic flow from 8:00 a.m. to 10:15 a.m. in eastbound left.

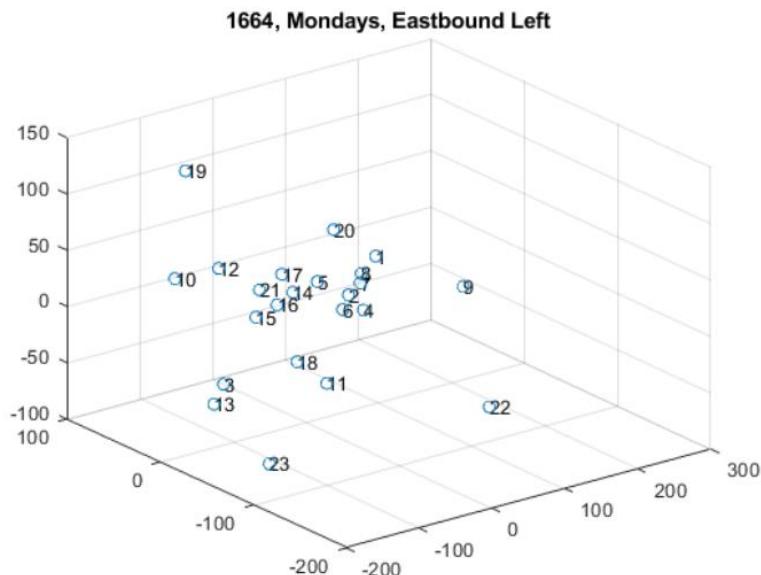


Figure 10. Scatter plot. Mondays eastbound left at signal 1664.

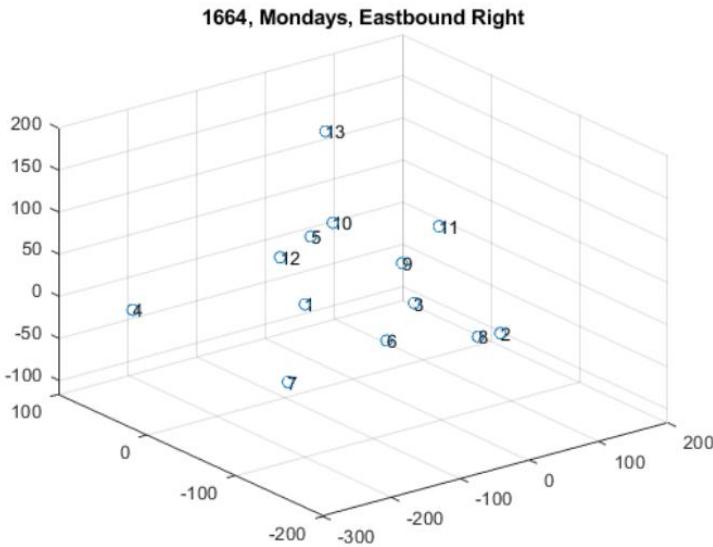


Figure 11. Scatter plot. Mondays eastbound right at signal 1664.

**Table 9. Outliers for Mondays at signal 1665:
Jimmy Carter Blvd. and Dawson Blvd./Live Oak Pkwy.**

Day	Date:	Findings:
1	January 13, 2020	Southbound traffic pattern began below average, increased in the middle of the day, and finished below.
3	February 3, 2020	Below-average northbound traffic.
6	June 24, 2019	Large dip in eastbound volume at 3:15 p.m.
7	July 1, 2019	Large spike in eastbound volume at 6:15 p.m.
9	July 15, 2019	Below-average northbound movements between 12:15 p.m. and 8:45 p.m. Elevated northbound movements into the late evening hours.
13	July 15, 2019	Low northbound traffic earlier in the day led into higher northbound traffic than usual.
16	September 2, 2019	Abnormally low eastbound movements throughout the day.
18	September 16, 2019	Westbound traffic elevated between 8:45 a.m. and 7:30 p.m.
19	September 23, 2019	Unusual westbound traffic volume. Large spikes at various points in the day, notably: 12:15 a.m., 1:30 a.m., 5 a.m., and 7 a.m.
30	December 23, 2019	Holiday season: Unusual traffic volumes across all movements. Traffic began below average and quickly soared above normal for the majority of the day.

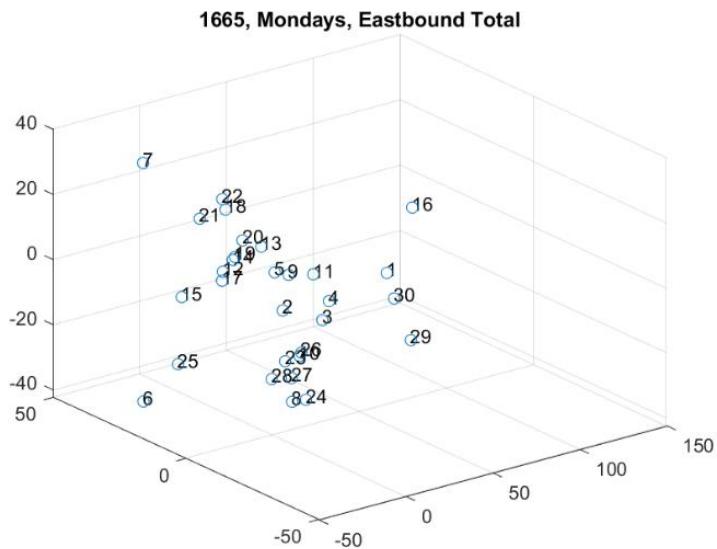


Figure 12. Scatter plot. Mondays eastbound total at signal 1665.

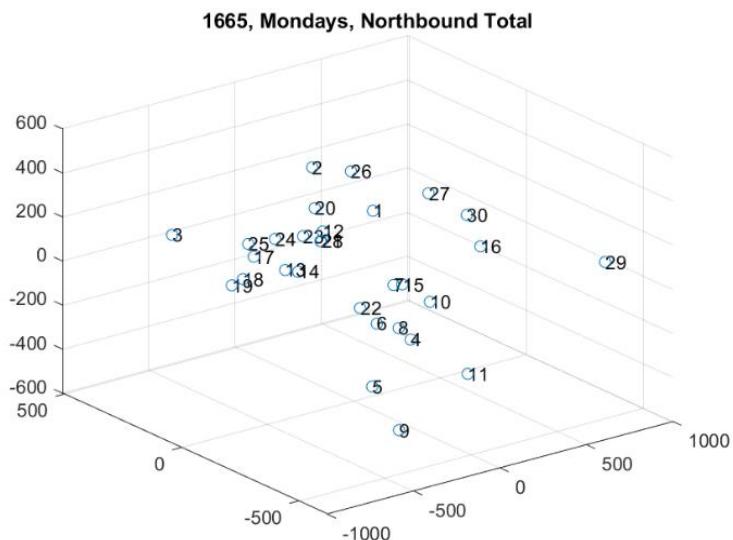


Figure 13. Scatter plot. Mondays northbound total at signal 1665.

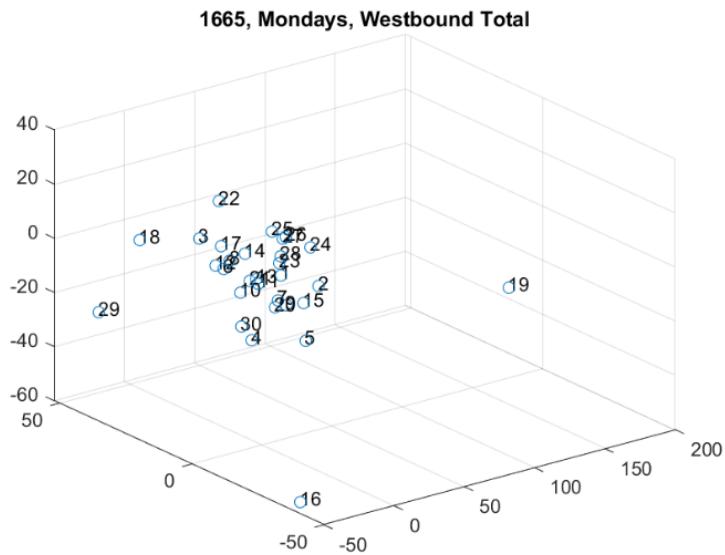


Figure 14. Scatter plot. Mondays westbound total at signal 1665.

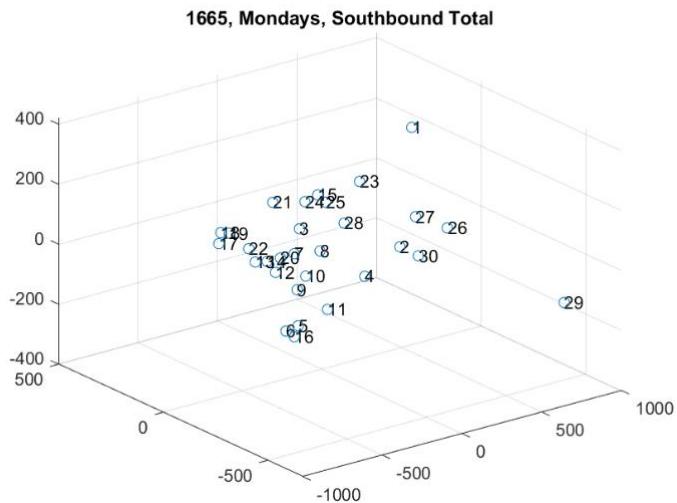


Figure 15. Scatter plot. Mondays southbound total at signal 1665.

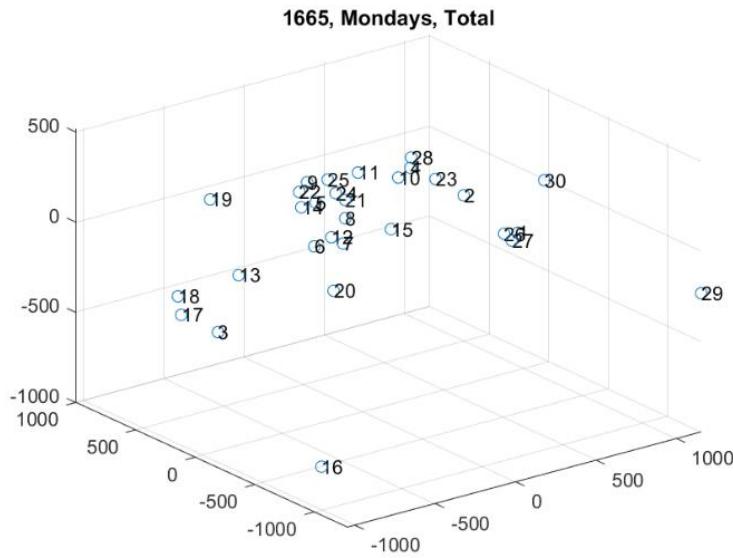


Figure 16. Scatter plot. Mondays total at signal 1665.

TRAFFIC PREDICTION

Methodology

Historically, partial least squares (PLS) regression has been known for its robust multivariate prediction capabilities.^[49,50] It has been shown to produce results comparable to powerful machine learning algorithms when predicting traffic flow.^[51] In this current study, the most recently developed form of PLS regression, known as SIMPLS regression, was used to construct models of turning movement behavior and to make predictions of future behavior. The collected historical data are partitioned into two blocks (see figure 17): Z (predictors) and Y (response). The model constructed from Z and Y will be used along with Z_s (new measurements) to predict Y_s (future response). The historical data are first preprocessed, i.e., they are mean subtracted and normalized by the standard deviation. Preprocessed Z and Y are denoted by \tilde{Z} and \tilde{Y} , respectively. From these historical data, three matrices T , P , and C are constructed. The matrix T contains orthogonal column vectors usually known as the *scores*. The P and C matrices are

known as the *latent variable* matrices. The matrices T , P , and C together describe the relationship between the predictor and the response, and they are designed to maximize the correlation between the prediction block and response block of the historical data. The matrices T , P , and C are constructed by an iterative method where one column is produced in each iteration. For example, the i^{th} iteration produces t_i , p_i , and c_i , the i^{th} column of matrix T , P , and C , respectively. To maximize the correlation between the predictor and response, first, begin by determining the unit vectors u^* and v^* such as in equation 4:

$$(u^*, v^*) = \arg \min_{\|u\|=1, \|v\|=1} (\tilde{Z}u)^T (\tilde{Y}v). \quad (4)$$

It can be verified that u^* and v^* are the first left and right singular vectors of $\tilde{Z}^T \tilde{Y}$. Now, we will use u^* to construct the first score vector and the latent vectors, as in equations 5, 6, and 7:

$$t_1 = \frac{\tilde{Z}u^*}{\|\tilde{Z}u^*\|}. \quad (5)$$

$$p_1 = Z^T t_1. \quad (6)$$

$$c_1 = Y^T t_1. \quad (7)$$

Notice that t_1 is a unit vector and we would like to have a set of such t_i vectors that are orthogonal to each other and a set of p_i and c_i vectors constructed from the corresponding t_i vectors. To ensure that the score vector t_2 produced in the second iteration is orthogonal to t_1 , the components of the data in the direction of t_1 must be subtracted out from the data matrix. Therefore, at the end of each iteration, the data matrices \tilde{Z} and \tilde{Y} are updated to prepare for the next iteration by subtracting their projection on the current score vector, as in equations 8 and 9:

$$\tilde{Z}^{++} = \tilde{Z} - t_1 t_1^T \tilde{Z}. \quad (8)$$

$$\tilde{Y}^{++} = \tilde{Y} - t_1 t_1^T \tilde{Y}. \quad (9)$$

This process is also known as *defeating the data matrices*. These updated matrices \tilde{Z}^{++} and \tilde{Y}^{++} will be used in the place of \tilde{Z} and \tilde{Y} in the next iteration. After N iterations, matrices T , P , and C each with N columns are constructed. It is known that the new measurements and future predictions are related through these matrices, as in equation 10:

$$Y_s - \bar{y} = (Z_s - \bar{z}) (P^T)^+ C^T, \quad (10)$$

where \bar{y} and \bar{z} are matrices carrying the column means of Y and Z_s , respectively.

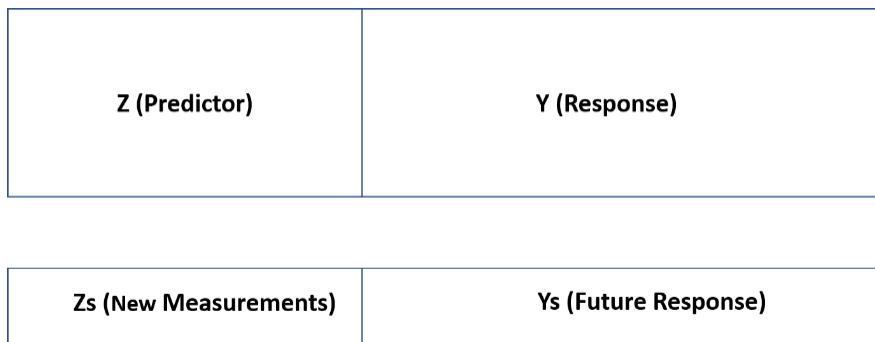


Figure 17. Diagram. Partitioning of the data matrix for SIMPLS.

Results

Data were gathered for signals 1662, 1663, 1664, and 1665 between the months of June 2019 and February 2020. Due to inherent sensor noise, the data were filtered prior to modeling as described in Xing et al.^[5] Then, a SIMPLS model was built for the turning movement behavior for each movement/intersection pair. Figure 18, figure 19, figure 20, and figure 21 show traffic predictions made for four different turning movement/intersection pairs on the day of Monday, February 3, 2020. Each model was trained over data from 17 prior Mondays. On the prediction

day, new measurements were made until 10:00 a.m. and predictions were made for the rest of the day. It was observed that the predictions matched the actual traffic very well.

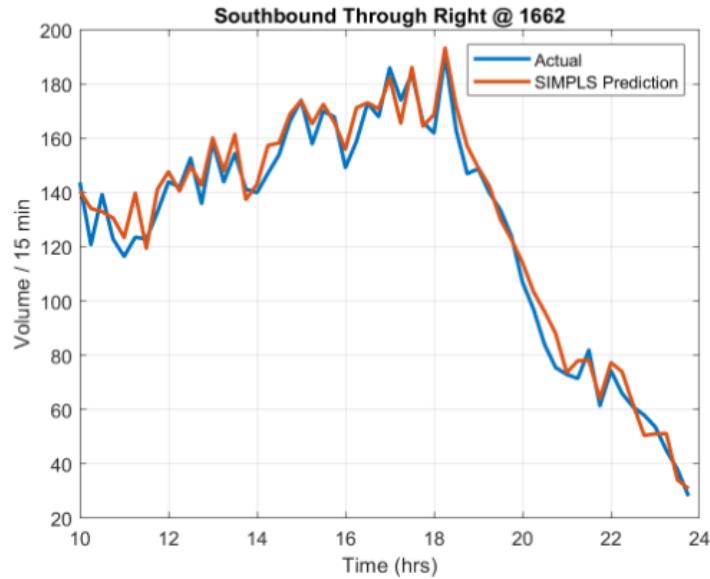


Figure 18. Line graph. PLS prediction for the southbound through right movement at intersection 1662.

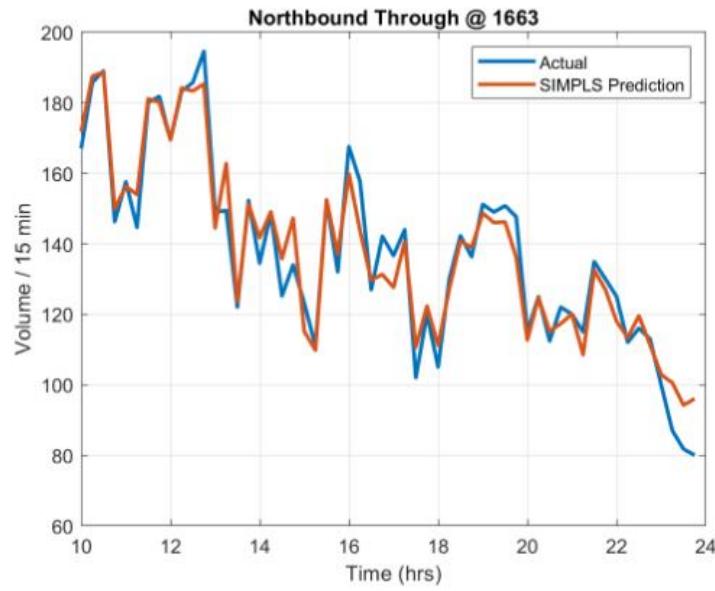


Figure 19. Line graph. PLS prediction for the northbound through movement at intersection 1663.

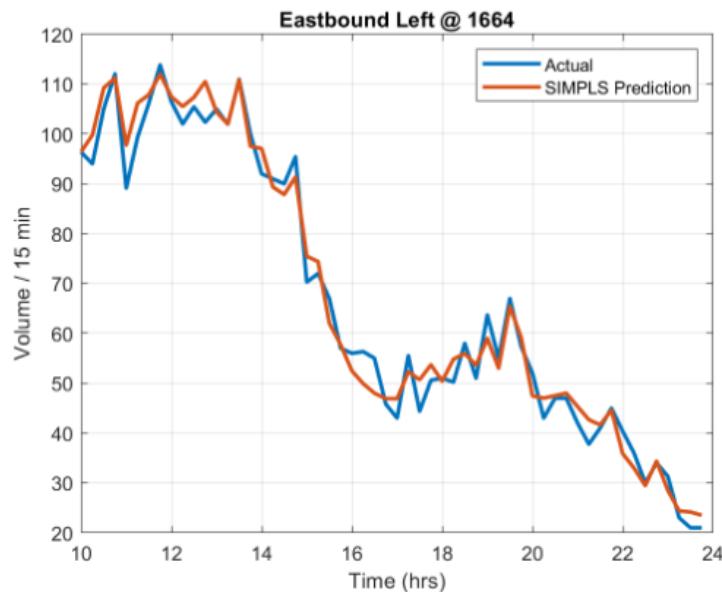


Figure 20. Line graph. PLS prediction for the eastbound left movement at intersection 1664.

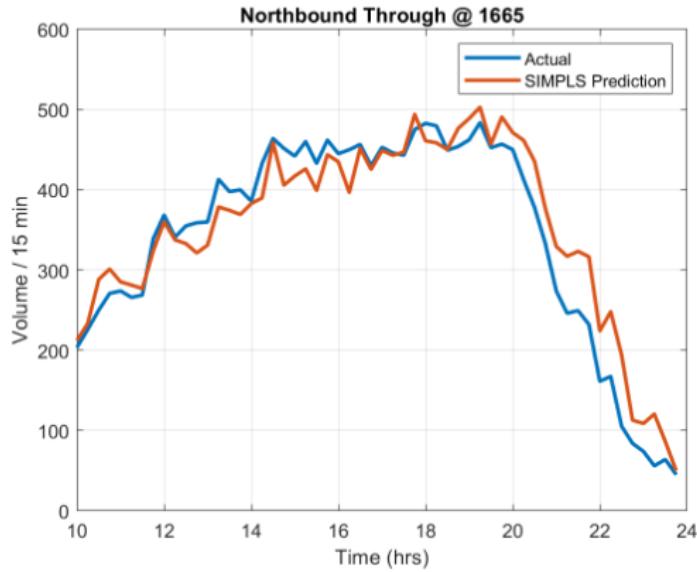


Figure 21. Line graph. PLS prediction for the northbound through movement at intersection 1665.

OPTIMAL TIME SEGMENTATION FOR TIME-OF-DAY PLANS

Time-of-day (TOD) scheduling is central to most traffic control schemes. By analyzing traffic trends at different periods throughout the day, traffic control engineers can design specific timing schemes for each period. However, segmentation of a day into such periods is often done heuristically or with the aid of simulation software. This current work uses the formal mathematical notation of *optimality* to derive *optimal* TOD segments from historical data such that the traffic is close to uniform over each segment. Previous work has highlighted the use of optimization algorithms for segmenting multivariate time series; however, much of this work uses *k-means* clustering, which requires heavy computational time.^[53,54] In this study, the researchers leveraged dynamic programming (DP)^[55] to perform optimal TOD segmentation as seen in references 56, 57, and 58. However, unlike previous approaches, the research team proposed a useful weighting scheme that allows the practitioner to weight the particular

movements in the optimization problem. This is useful when service of certain movements, pertaining to arterial or off-ramp traffic, is prioritized. In addition to providing optimal TOD segmentation times, the proposed algorithm provides the practitioner a model of traffic flow during each TOD interval. This model can be used to index readily available timing guides to arrive at optimal timing strategies for the DDI and its adjacent intersections.

Methodology

The problem of deriving optimal segments from a multivariate time series can be recast as determining the optimal piecewise constant function approximating said time series, since the goal is to find segments where traffic is most uniform. Given M movements over T time measurements, let $X \in R^{M \times T}$ represent a standardized, i.e., mean-centered and standard deviation of 1, data matrix with the average volume of traffic for movement i at time j stored at location $X(i,j)$. Since this problem is for TOD segmentation, the columns of this matrix represent different time stamps over a 24-hour day. Let $S \in \mathbb{Z}^+$ be the number of segments to divide the time series into. In traffic engineering, it is common for as many as 4 to 7 TOD plans to be used.^[58] At the end of the calculations, the goal is to derive a vector of $S + 1$ breakpoints that represent the boundaries of each TOD segment: $\tau = [\tau_0, \tau_1, \dots, \tau_S]$. Naturally, $\tau_0 = 0$, $\tau_S = T$, and $\tau_i < \tau_{i+1}$.

Over each segment, the goal is to find the optimal *neighborhood parameters* $\{\mu_1, \dots, \mu_S\}$, where $\mu_i \in R^M$ models the traffic flow in cars per hour over the i^{th} segmentation period. Formally, the goal is to minimize the error between the approximation μ_i and the actual traffic over the corresponding TOD period, as in equation 11:

$$\boldsymbol{\mu}_i = \arg \min_{\boldsymbol{\mu}_i} \sum_{t=\tau_{i-1}+1}^{\tau_i} (\mathbf{x}_i(t) - \boldsymbol{\mu}_i)^T (\mathbf{x}_i(t) - \boldsymbol{\mu}_i), \quad (11)$$

where \mathbf{T} is the transpose operator and $x_i(t)$ is the t^{th} column of X over the TOD segment i . This error is called the “fit” of the parameter set to the data, as in equation 12:

$$F(\tau_{i-1}, \tau_i, \mathbf{X}, \boldsymbol{\mu}_i) = \min_{\boldsymbol{\mu}_i} \sum_{t=\tau_{i-1}+1}^{\tau_i} (\mathbf{x}_i(t) - \boldsymbol{\mu}_i)^T (\mathbf{x}_i(t) - \boldsymbol{\mu}_i). \quad (12)$$

Note that the fit presented here is equivalent to summing the mean-square error of each of the movements across this TOD segment. Also, the fit is a positive definite function convex in its second argument. This unique property ensures that local minima are indeed global minima, thus making the optimization problem feasible.

Now that the error associated with approximating a given segment of a multivariate time series by a constant vector is known, the next step is to determine the optimal segmentation of the series. The goal is to minimize the error, or cost, of approximating the entire multivariate time-series by a piecewise constant vector-valued function. Precisely, the hope is to minimize the cost function, as in equation 13:

$$J(\mathbf{X}, \boldsymbol{\mu}, \boldsymbol{\tau}, S) = \min_{\{\boldsymbol{\mu}_i\}, \boldsymbol{\tau}} \sum_{i=1}^S \sum_{t=\tau_{i-1}+1}^{\tau_i} (\mathbf{x}_i(t) - \boldsymbol{\mu}_i)^T (\mathbf{x}_i(t) - \boldsymbol{\mu}_i). \quad (13)$$

The global minimum of this function would be computationally expensive to find by exhaustive enumeration. The following section proposes a dynamic programming approach to solving this optimization problem.

Dynamic Programming

Generally, dynamic programming is the act of performing computation recursively and storing computations that are performed in order to save computation time. This problem is rather involved and demands three answers:

1. What is the global minimum of the cost function outlined in equation 13?
2. What are the TOD boundaries, τ , that minimize this cost function?
3. Over each TOD interval, what are the optimal neighborhood parameters, μ , that best represent traffic flow?

The research team directly computed the answer to item 1, through backtracking the answers to items 2 and 3.

The Algorithm

Let $C_{i,j}^s$ be the optimal partitioning of the sequence $\{x_i, \dots, x_j\}$ into s segments. Let $Z_s(j)$ be the s^{th} segmentation point that minimizes $C_{i,j}^s$. Finally, let the vector $\mu_{i,j}$ represent the constant model of traffic flow across the interval of time i to j . The appropriate DP algorithm to find the global minimum of the cost and the optimal TOD segments is summarized as algorithm 1 in figure 22.

Step 1 preprocesses the data so that the segmentation algorithm has access to the average traffic volume for each movement at each time index throughout the day. Step 2 computes the measure of fit for all possible segments and memorizes these values for later use. Step 3 computes the optimal partitions iteratively, beginning with 2 segments and working its way up to S segments. Step 4 uses backtracking to find the optimal TOD breakpoints from the previous step's computations. When finished, the model parameters μ for each TOD segment can be indexed from Step 1. Note that the backtracking step computes the optimal partitions $s = 1, \dots, S$ segments. Thus, the algorithm does not need to be executed again to derive optimal lower-order segmentations.

As an elaboration on Step 4, consider segmenting a time series $\{x_1, \dots, x_T\}$ into S segments. Naturally, the final segment will be some $[s + 1, T]$. Subsequently, the optimization problem is

to segment the series $\{x_1, \dots, x_s\}$ into $S - 1$ segments. Step 4 uses this idea to backtrack over the results from Step 3 and extract the optimal TOD breakpoints.

Algorithm 1 Optimal TOD Segmentation via Dynamic Programming

```

1: procedure Pre-process Data (Step 1)
2:   Initialize data matrix  $\mathbf{A}$ , which contains historical time-series for movements 1 through  $M$ .
3:   Let the rows of matrix  $\mathbf{X}$  contain the average time-series for each movement from  $\mathbf{A}$ .
4:   Standardize the rows of matrix  $\mathbf{X}$  (mean zero/ standard deviation 1).
5: procedure Compute measure of fit for all possible segments (Step 2)
6:   for  $(i, j), i \leq j$ , and  $i, j \in [1, n]$  do
7:      $C_{i,j}^s := J(x_i \dots x_j, \mu', \tau', 1)$ 
8:      $\mu_{i,j} := \underset{\mu}{\operatorname{argmin}} J(x_i \dots x_j, \mu, \tau', 1)$ 
9:      $Z_1(j) := 0$ 
10: procedure Compute optimal partitions  $s = 2, 3, \dots, S$  segments (Step 3)
11:   for  $s = 2$  to  $S$  do
12:     for  $j = 1$  to  $T$  do
13:        $C_{1,j}^s := \underset{v \in [1, j]}{\min} (C_{1,v}^{s-1} + C_{v+1,j}^1)$ 
14:        $Z_s(j) := \underset{v \in [1, j]}{\operatorname{argmin}} (C_{1,v}^{s-1} + C_{v+1,j}^1)$ 
15: procedure Recover  $\tau$  via backtracking (Step 4)
16:   for  $s = 1$  to  $S$  do
17:      $\tau_s(S) := T$ 
18:     for  $i = s, s - 1, \dots, 1$  do
19:        $\tau_{s-1}(i) := Z_s(\tau_s(i))$ 
```

Figure 22. Algorithm 1. Optimal TOD segmentation via dynamic programming.

Convexity and Improved Fit Functions

This section is dedicated to improved fit functions, i.e., functions that make the segmentations and neighborhood parameters in this study more useful to traffic engineers. It is more important for a particular intersection to serve certain movements over the others. Therefore, the researchers present an improved fit function that allows traffic engineers to weight these movements over those less important.

As stated previously, convex optimization problems are much simpler to solve, in general, than nonconvex optimization problems. This is because local minima are, in fact, global minima in

convex optimization problems. Each fit function presented in this paper is convex, ensuring that the parameters μ are easily computed. However, the larger DP algorithm that computes the optimal segmentation times τ from the cost function $J(\cdot)$ does not use or rely on convexity.

Recall equation 12, which represents the optimal fit for an arbitrary interval $\tau_{i-1} + 1, \dots, \tau_i$. For notation's sake, rewrite it in the form given in equation 14:

$$F(\tau_{i-1}, \tau_i, \mathbf{X}, \boldsymbol{\mu}_i) = \min_{\boldsymbol{\mu}_i} \sum_{t=\tau_{i-1}+1}^{\tau_i} \Phi(\mathbf{x}_i(t), \boldsymbol{\mu}_i). \quad (14)$$

Allow, as in equation 15:

$$\Phi(\mathbf{x}_i(t), \boldsymbol{\mu}_i) := (\mathbf{x}_i(t) - \boldsymbol{\mu}_i)^T \mathbf{A} (\mathbf{x}_i(t) - \boldsymbol{\mu}_i), \quad (15)$$

where $\mathbf{A} \in \mathbb{R}^{M \times M}$ is a positive semi-definite (PSD) matrix. Due to this property and equation 15 being in quadratic form, the optimization problem remains convex.

The only remaining problem to address is the choice of a matrix A that appropriately weights the movements in this optimization problem. In general, it is simplest to choose a diagonal matrix with nonnegative entries. This ensures that the matrix is PSD and that its eigenvalues represent the weight of each movement. Consider a simple case in which there are two movements with a goal to weight the second movement twice as much as the first. It would result in the matrix shown in equation 16:

$$A := \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}. \quad (16)$$

In general, a nondiagonal PSD matrix could be chosen where the entries off of the diagonal represent the correlations between movements. However, the authors believe that this scenario may be inappropriate for practical implementation.

Results

The results presented here show the optimal TOD segmentation times derived by algorithm 1 for signal 1665 on Sundays. Training data from June 2019 to February 2020 were used. The movements considered were eastbound, southbound, northbound, and westbound totals. Results using the fit function reported in equation 12 are shown in figure 23.

At this intersection, northbound and southbound movements represent the majority of traffic flow. In fact, these movements feed into and receive traffic from an adjacent DDI. Therefore, the goal was to weight these movements heavily. The research team used the fit function proposed in equations 14 and 15. When these movements were weighted double their secondary counterparts, the algorithm produced the segmentation depicted in figure 24.

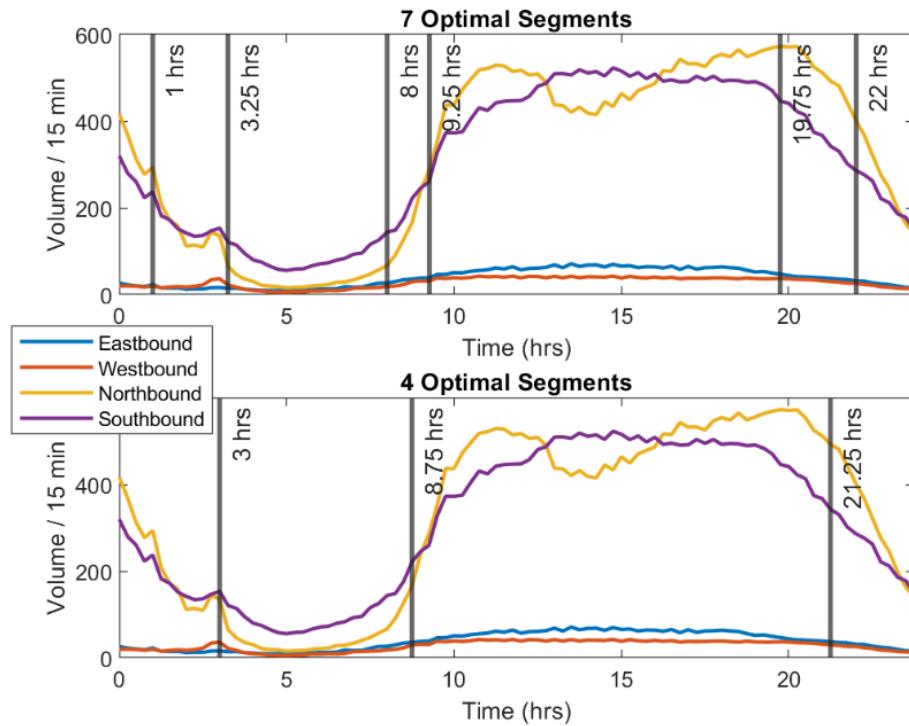


Figure 23. Line graphs. Optimal TOD segments using the original fit equation.

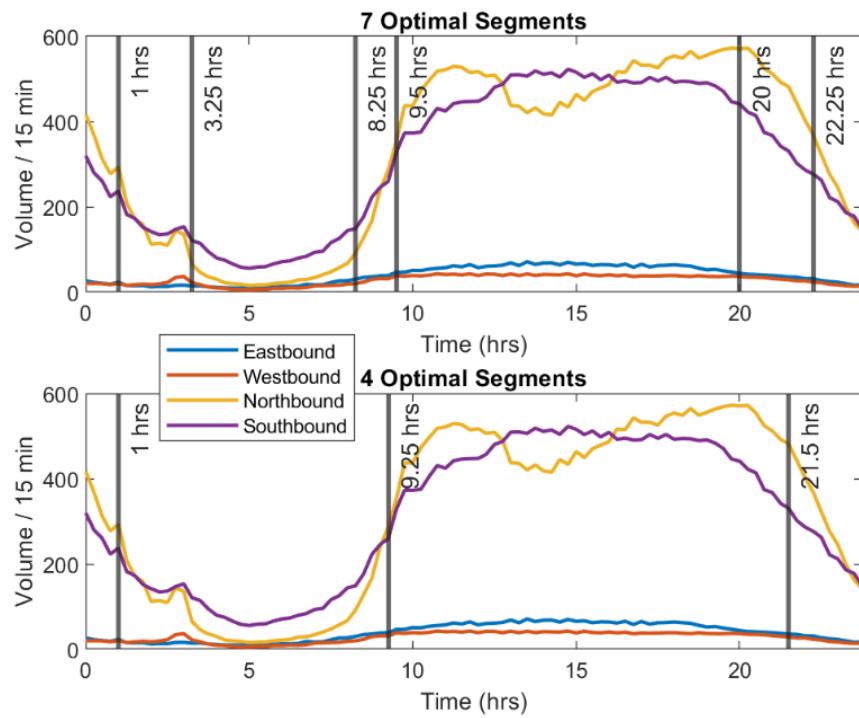


Figure 24. Line graphs. Optimal TOD segments using the modified fit.

CONCLUSIONS

During Task II, all necessary building blocks were built for the foundation of the traffic predictive control algorithm that would be developed in Task III. The important features and characteristics of the traffic trends were separated out from the noise in the data set through PCA and the results demonstrated that the method was able to detect anomalous traffic trends effectively. The low rank structures in the data obtained through PCA and singular value decomposition (SVD) were used to construct data-driven models by applying SIMPLS regression. It was observed that these models can make excellent predictions, which match the actual traffic extremely well. The researchers also developed an algorithm that performs optimal time segmentation for switching times between TOD plans using dynamic programming. This algorithm performs time segmentation depending on whether the practitioner selects to prioritize the arterial traffic or the off-ramp traffic. All of the building blocks and tools developed and the information gathered during Task II put the team on a solid foundation for the next step of developing the traffic predictive control algorithm in Task III.

ALGORITHM DEVELOPMENT AND ALGORITHM IN PSEUDOCODE

The main objective of this research project was to develop actuated traffic control algorithms for coordinated anticongestion control of diverging diamond interchanges and their surroundings based on recently developed control theory to produce closed-loop, data-driven, and model-free control strategies that provide guarantees on performance. The primary goal was to formulate coordinated signal control regimes between DDIs and adjacent intersections. The secondary goal was to report the project's research findings to GDOT and disseminate said findings to the transportation research community, as well as the general public, through various publications and communication channels. During Task I (literature review) of the project, a thorough literature review was conducted to survey the state of the art of the current research on DDIs and traffic control schemes. This task served to shape the scope and direction of this project's research so that the work of future tasks would add the most value to the existing knowledge. During Task II (modeling and analysis), 8 months of data were collected for two signals within the DDI at Jimmy Carter Blvd. (signal 1663, signal 1664) and two adjacent signals, one upstream and one downstream (signal 1662, signal 1665). Then, a thorough data analysis was conducted to extract traffic trends, as well as to construct data-driven models that can make predictions on future traffic flows. The purpose of this task was to build the foundation for the development of a traffic predictive control algorithm. During Task III (algorithm development), the information extracted and the models developed from data during Task II were used for synthesizing algorithms that coordinate mainline traffic with downstream intersections and freeway exit ramp traffic. The following were accomplished during Task III:

- The Robust SIMPLS traffic prediction and control algorithm was developed. The algorithm in pseudocode and the detailed technical aspects of the algorithm are presented in this report. During Task II, the team identified the issue that the turning moving count data obtained from the ATSPM website were highly corrupted. Therefore, the SIMPLS model constructed during Task II needed to be robustified against corruption in data. The team developed a new data-driven traffic prediction algorithm, dubbed Robust SIMPLS, which performs well even with corrupted training data. Results of this newly developed algorithm have been submitted for presentation at the TRB annual meeting and publication in the *Transportation Research Record* (TRR).
- The team developed a prototype of a decision support web engine to aid traffic engineers. The web engine receives traffic data on the DDIs through the ATSPM website. Based on this real-time incoming data, the predictions of the Robust SIMPLS algorithm will be displayed in a graphical user interface for use by traffic engineers. The web engine will also provide timing scheme adjustment recommendations or send out alerts to field engineers when the traffic trends indicate that intervention is necessary.
- In order to demonstrate the feasibility of the decision support tool web engine, the team created a Synchro/ SUMO simulated testbed. GDOT provided the research team with Synchro files for several DDIs in the state of Georgia. Data from these Synchro files were imported into SUMO, an open-source, microscopic traffic simulator. With SUMO, the traffic prediction algorithm was trained and tested with different traffic demand profiles. In cases where intervention is recommended by the prediction algorithm, SUMO can be used to demonstrate how altering the signal timings and ramp metering would increase efficiency at the interchange and neighboring intersections.

Traffic flow predictions play a central role in modern traffic predictive control algorithms. (See references 58, 59, 60, 61.) Recently, the importance of data-driven predictions—as opposed to analytic, model-based predictions—has been realized.^[62,63,64] Data-driven predictions are easily updated and are often computationally inexpensive to derive, which makes them attractive for practical applications. However, little work has been done to ensure that data-driven models of traffic flow are robust to corruptions in training data and real-time traffic measurements. Such corruptions may come from one-off traffic anomalies or faulty traffic flow sensors.

With the insurgence of deep learning in the transportation community, data-driven system identification has largely relied on heuristics—namely, various neural network (NN) architectures—rather than rigorous mathematical analysis. Ke et al. utilized convolutional NNs to predict multi-lane traffic speeds.^[65] Zhang et al. demonstrated the use of long short-term memory NNs to predict pedestrian crossing behavior.^[66] Perhaps the most notable NN-based approach to traffic flow prediction belongs to Li et al.^[67]; however, similar to Lu et al.,^[68] this method uses the concept of “network diffusion”—*i.e.* it relies on real-time network-wide measurements and communication to produce such predictions, which is not practical in many traffic control scenarios. All of these techniques replace traditional time series methods, such as seasonal mean, autoregressive moving average (ARMA) filters, and support vector regression (SVR), for traffic flow forecasting.^[69]

Despite the myriad of data-driven techniques for predicting traffic flow, most are impaired by corrupt data. The research team proposed combining Robust Principal Component Analysis (RPCA)^[70] and SIMPLS regression, whose name is best explained by Jong,^[71] to overcome this issue. The prediction scenario explored in this work is as follows: given traffic flow measurements from an intersection in the early morning, can traffic for the remainder of the day

be predicted? If the predictions differ substantially from historical averages, they can be used to automatically tune a future time-of-day plan or to alert a traffic engineer that manual adjustments may be required. This prediction scenario can, of course, be trivially generalized to different circumstances, prediction horizons, etc.

This section is organized as follows. Mathematical preliminaries are presented and then followed by a brief explanation of SIMPLS and RPCA. The proposed method, Robust SIMPLS, is then formalized. Then, the technique is used to predict real traffic flow through the DDI at Jimmy Carter Blvd. and its adjacent intersections. The results are compared to those produced by other data-driven approaches, which show that Robust SIMPLS is an improvement to the current state of the art.

REMARKS ON CORRUPT TRAINING DATA

In most data-intensive applications, the data are corrupt with noise. This noise could come in two forms: (1) low-energy, mean-centered Gaussian noise, (2) one-off, high-energy traffic anomalies—e.g., an accident that causes major delays for a brief period of time. This section address each of these sources of error below.

Low-energy, mean-centered Gaussian noise is likely to come from sensor error. This is equivalent to a sensor counting a few extra, or a few less, cars during each recording interval. Since these are low-energy, they are described by the final few rank-1 matrices of the data's SVD—i.e., they are described by the smallest singular values of the data. Thus, a simple rank-r truncation of the training data used in building SIMPLS models will keep SIMPLS from trying to predict this random process. However, the maximum correlation between predictors and responses is used in building SIMPLS models. Therefore, it is probable that there is no need to

address these low-energy corruptions in the data at all. Unfortunately, one-off, high-energy anomalies in training data are likely to severely corrupt SIMPLS models. Consider the following: a *single* entry in the training data, X , is highly corrupted. The corrupted data matrix can be described as in equation 17:

$$X + \theta e_i e_j^T, \quad (17)$$

where e_i is the i^{th} standard basis vector—i.e., a vector with zero entries everywhere except the i^{th} entry, which contains a 1—and θ is the size of the corruption. Note that $\theta e_i e_j^T$ is a rank-1 matrix and θ is its singular value. By the Weyl inequalities,^[72] it is known as in equation 18:

$$\sigma_{\max}(X + \theta e_i e_j^T) \leq \sigma_{\max}(X) + \theta. \quad (18)$$

Since the upper bound in equation 18 is achievable, it is possible for this single corruption to alter the largest singular value of the matrix—the singular value that is associated with representing the most variance in the data! This is dangerous for data scientists, and it prevents the practitioner from using the standard approach of using a truncated SVD in order to rid the data of disturbances. This easily extends to scenarios where the data are corrupt by multiple rank-1 matrices whose singular vectors are coherent with the standard basis vectors.

Imagine a data matrix representing turning movement counts. These data are corrupt with both low- and high-energy noise. It is also possible that the data have gaps expressed as zeros due to lapses in recording by a traffic controller. If these corruptions are even moderately present, the SVD of this data matrix may “waste” singular values/vectors on expressing such data corruptions. Quickly, the PCA (and SVD) of these data become tainted.

ROBUST PRINCIPAL COMPONENT ANALYSIS

RPCA was proposed as a matrix separation algorithm in order to denoise data.^[70] Specifically, it attempts to remove *sparse* corruptions in data. Such corruptions come in the form of $\theta e_i e_j^T$. However, (1) the singular vectors associated with these corruptions need only be *coherent* with the standard basis,¹ and (2) there may be more than one corruption in the data—i.e., the corruption matrix is expressed as in equation 19:

$$S = \sum_{i=1}^n \theta_i u_{s_i} v_{s_i}^T. \quad (19)$$

Consider an incoming data matrix as a superimposition of a low-rank matrix (L) and a sparse noise matrix (S) as in equation 20:

$$X = L + S. \quad (20)$$

The low-rank matrix is not corrupted by noise—this matrix can be used to inform SIMPLS models. L should be as low rank as possible and S should be as sparse as possible. In other words, these matrices are the solution to the optimization problem, as in equation 21:

$$\begin{array}{ll} \min & \text{rank}(L) + \|S\|_{\ell_0} \\ \text{subject to} & X = L + S \end{array}. \quad (21)$$

Unfortunately, neither the rank or the ℓ^0 norm of a matrix are convex; thus, this optimization problem is particularly difficult. However, Candès et al.^[70] have shown that the solution to the following convex optimization problem produces optimal solutions to equation 21, as in equation 22:

¹i.e., the singular vectors need only resemble the standard basis vectors. See reference 73 for an exact definition of coherence.

$$\begin{array}{ll} \min & \|L\|_* + \lambda \|S\|_{\ell_1} \\ \text{subject to} & X = L + S \end{array} . \quad (22)$$

Here, $\|\cdot\|_*$ denotes the nuclear norm (a proxy for rank), $\|\cdot\|_1$ denotes the ℓ^1 norm (a proxy for sparsity), and λ is a weighting parameter.

Equation 22 is known as *principal component pursuit* (PCP) and can be easily solved using augmented Lagrange multipliers (ALM).^[74] The principal components of L are robust to noise in the original data—hence, the name RPCA. If λ is chosen as in equation 23:

$$\lambda = \frac{1}{\sqrt{\max(m, n)}}, \quad (23)$$

PCP will recover L and S exactly, provided that the singular vectors of L are incoherent with the standard basis vectors and the rank of L is not too large. In general, this algorithm produces a low-rank matrix L that is significantly different than the low-rank approximation of the original data, A_r , obtained via the SVD/PCA.

ROBUST SIMPLS METHODS AND RESULTS

Given the likelihood of corrupt training data, it is only natural to extend the results of SIMPLS to a setting that is robust to such corruptions. This is easily achieved by first performing RPCA on the training data and then using the low-rank matrix L to construct SIMPLS models. A sample algorithm that may run throughout the day to either (1) deliver traffic predictions to control engineers, or (2) inform a traffic-predictive controller (as seen in reference 58), is outlined in algorithm 2 in figure 25.

To demonstrate the predictive accuracy of Robust SIMPLS, the research team evaluated its performance on highly corrupted data collected from a DDI in Norcross, Georgia, and its

adjacent intersections. A map of the intersections for which the results are reported is depicted in figure 26. Details about each of these intersections are summarized in table 10. Data between June 2019 and February 2020 were provided by GDOT for this study.

The predictions are constructed as such: given measurements of vehicle turning-movement counts between 12:00 a.m. and 10:00 a.m., predict turning-movement behavior for the rest of the day. This is a more difficult problem than the traffic flow predictions based on network diffusion mentioned in the introduction because: (1) traffic flow measurements from other movements/intersections are not available, and (2) traffic for the entire day is forecasted at once. Extensions and modifications to this prediction horizon are trivial and easily obtained. This sort of problem is realized in scenarios where predictions play a role in planning—such as TOD plan scheduling, which must be done prior to the switching of TOD plans.

Algorithm 2 Robust SIMPLS

```

1: loop
2:   if ← New Measurements Available then
3:      $X \leftarrow$  Record new traffic flow measurements
4:      $Z_s \leftarrow$  Extract new predictors from  $X$ 
5:      $Z \leftarrow$  Append  $Z_s$  to  $Z$ 
6:      $\hat{Y}_s \leftarrow$  Predict future behavior using SIMPLS
7:   if Update Model then
8:      $L, S \leftarrow$  Perform RPCA on  $X$ 
9:      $T, C, P \leftarrow$  Build new SIMPLS model using  $L$ 
```

Figure 25. Algorithm 2. Robust SIMPLS.



Figure 26. Photo. Map of a DDI and its adjacent intersections at I-85 and Jimmy Carter Blvd. in Norcross, GA.^[75]

© 2020 Google

Table 10. Summary of intersections examined in this study.

Int:	Primary Road:	Secondary Road:
1662	SR 140	Crescent Dr/ Goshen Springs Rd
1663	SR 140	I-85 Southbound
1664	SR 140	I-85 Northbound
1665	Jimmy Carter Blvd	Dawson Blvd/ Live Oak Pkwy

Predictive Accuracy

For an initial experiment, movements at the four intersections on Monday, February 3, 2020, were predicted. The results are shown in figure 27, figure 28, figure 29, and figure 30. In each of these figures, the measured turning movements are shown in a dotted blue line. These measurements are assumed to be corrupt by sensor noise and one-off traffic anomalies, as mentioned previously. The Robust SIMPLS algorithm was trained on data from just eight prior weekdays—defined as Monday through Thursday—for intersections 1662–1664. The model

derived for 1665 used 20 prior weekdays due to complex flows at this intersection. The predictor variables were not filtered by RPCA prior to making the predictions reported here. The results, however, show that the model built using Robust SIMPLS is robust to the noise in the predictor measurements. For comparison, RPCA was run over the data once the responses were measured (at the end of the day, after predictions had been made). The corresponding denoised response variables are shown in orange. Clearly, Robust SIMPLS closely predicted the denoised response variables despite receiving corrupt predictor variables as input. Since RPCA exactly recovers the denoised data under the conditions mentioned previously, Robust SIMPLS predictions that are close to those produced via RPCA ensure that the predictions are highly accurate.

11 summarizes the predictive accuracy of Robust SIMPLS for various movements and days. For each movement/day predicted, its data were removed from the data set used to inform the Robust SIMPLS model. Clearly, on some days Robust SIMPLS's accuracy was comparative to that of the average traffic trend. However, when traffic was anomalous, Robust SIMPLS provided a much better prediction of traffic flow. Thus, Robust SIMPLS was advantageous over the average for most traffic prediction scenarios.

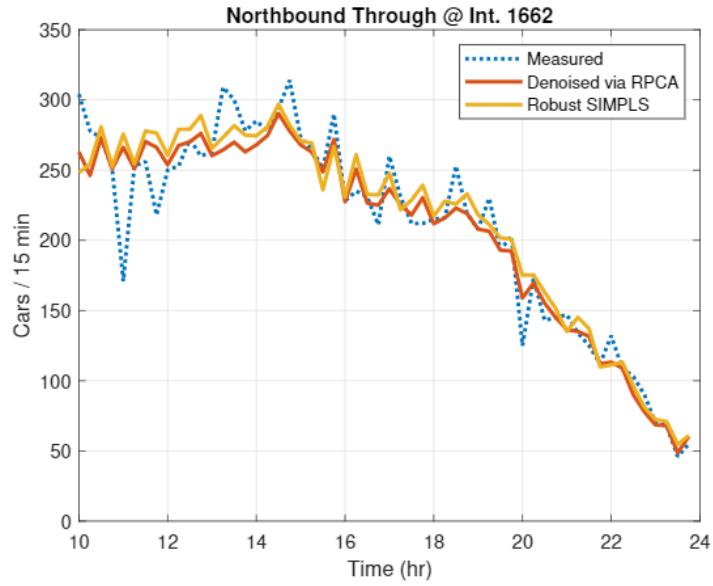


Figure 27. Line graph. Northbound through right movements at intersection 1662 predicted between 10:00 a.m. and 12:00 p.m.

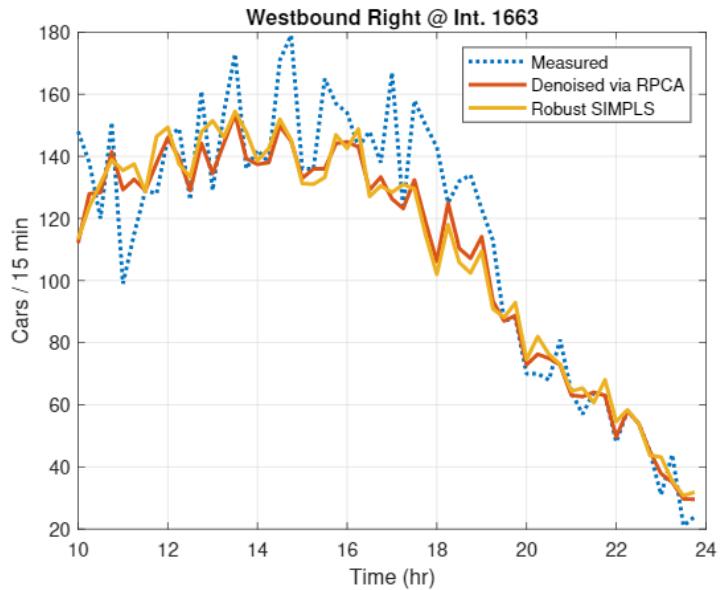


Figure 28. Line graph. Westbound total movements at intersection 1663 predicted between 10:00 a.m. and 12:00 p.m.

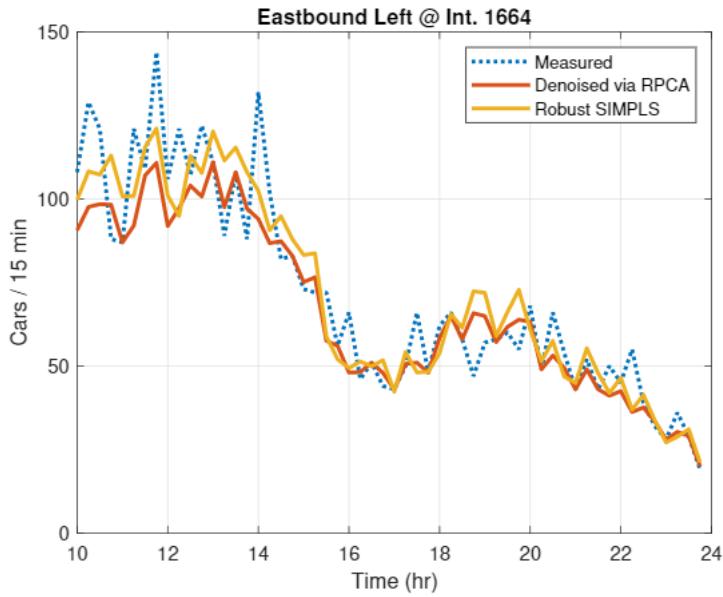


Figure 29. Line graph. Eastbound left movements at intersection 1664 predicted between 10:00 a.m. and 12:00 p.m.

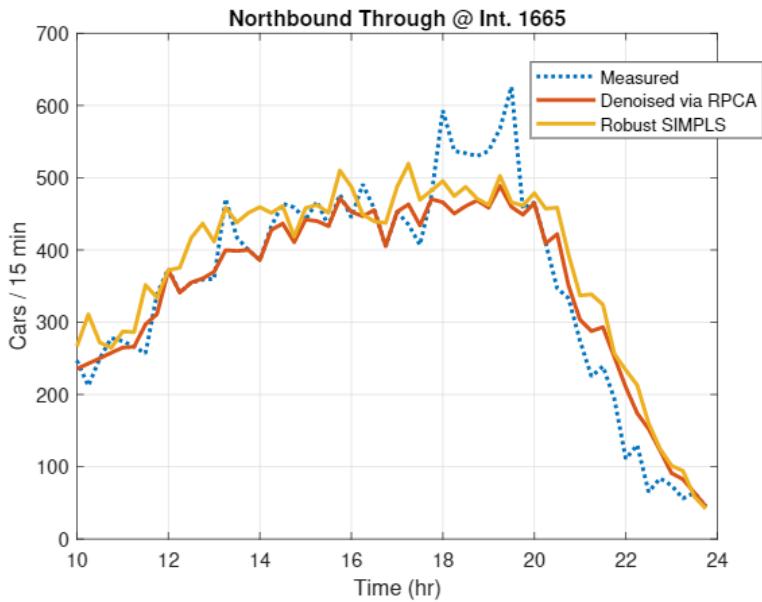


Figure 30. Line graph. Northbound through movements at intersection 1665 predicted between 10:00 a.m. and 12:00 p.m.

Table 11. Performance measures for Robust SIMPLS in various prediction scenarios.

Int:	Day:	Movement:	Robust SIMPLS MAE:	Simple Average MAE:	Improvement:
1662	2/6/2020	Northbound Through	17.95	33.01	15.06
1662	2/5/2020	Northbound Through	4.91	11.57	6.66
1662	2/4/2020	Northbound Through	3.56	11.80	8.24
1662	2/3/2020	Northbound Through	7.19	13.49	6.30
1663	2/6/2020	Westbound Right	13.93	20.96	7.03
1663	2/5/2020	Westbound Right	2.51	9.62	7.11
1663	2/4/2020	Westbound Right	3.70	7.60	3.90
1663	2/3/2020	Westbound Right	3.44	10.26	6.82
1664	2/6/2020	Eastbound Left	11.27	14.21	2.94
1664	2/5/2020	Eastbound Left	1.57	7.35	5.78
1664	2/4/2020	Eastbound Left	3.67	5.43	1.76
1664	2/3/2020	Eastbound Left	5.24	4.74	-0.50
1665	2/6/2020	Northbound Through	56.86	62.20	5.34
1665	2/5/2020	Northbound Through	32.49	41.12	8.63
1665	2/4/2020	Northbound Through	46.59	40.33	-6.26
1665	2/3/2020	Northbound Through	27.55	109.43	81.88

COMPARISON TO VARIOUS DATA-DRIVEN METHODS

Now, the predictive accuracy of Robust SIMPLS is compared to that of SIMPLS, nonlinear regression via a neural network, and the simple average. For this comparison, the total northbound movements at intersection 1665 on October 3, 2019, were studied. The neural network used in this study consisted of two hidden layers of 100 nodes, each followed by rectified linear unit (ReLU) activation functions and was trained for 8000 iterations with a batch size of 14 using the Adam optimizer^[76] and a learning rate of 0.001. Network weights were initialized using Glorot normal initialization.^[77] The network was trained using the same training data set as Robust SIMPLS. Care was taken that the network was trained sufficiently without overfitting, as undertraining or overtraining would decrease its efficacy.

The results of this comparison study are summarized in figure 31 and figure 32. Figure 31 shows the forecasted traffic flow via the average, the neural network, and Robust SIMPLS. Clearly, the

average traffic flow failed to capture the dynamics of traffic on this day. Both the neural network and Robust SIMPLS seem to capture the trend. However, as seen in Figure 32, Robust SIMPLS has a mean absolute error (MAE) that is lower than the neural network's prediction by over 33 cars per 15 minutes. For additional comparison, the MAE produced via traditional SIMPLS is also presented, which is over 27 cars higher per 15 minutes than that of its robust counterpart. Similar results are seen when comparing these methods on October 1, 2019—shown in figure 33 and figure 34.

IMPLICATIONS OF ROBUST SIMPLS

The results presented in this section have been submitted for presentation at the TRB annual meeting and for publication in the TRR. This work has identified issues associated with classic PCA-based regression models in corrupt environments, and it proposes Robust SIMPLS as a solution to predict traffic flow in such environments. The results presented show that Robust SIMPLS is a significant improvement to traffic flow prediction methods currently employed. Moreover, this method does not require the tuning of high-level hyperparameters, such as those encountered when designing neural networks—i.e., human-in-the-loop learning is not required. Thus, it is a highly effective method for traffic control engineers that may not be experts in deep learning or data science. This prediction technique can be used to inform traffic-predictive controllers that autonomously switch between TOD plans based on forecasted traffic flow or to alert traffic engineers of unusual traffic trends.

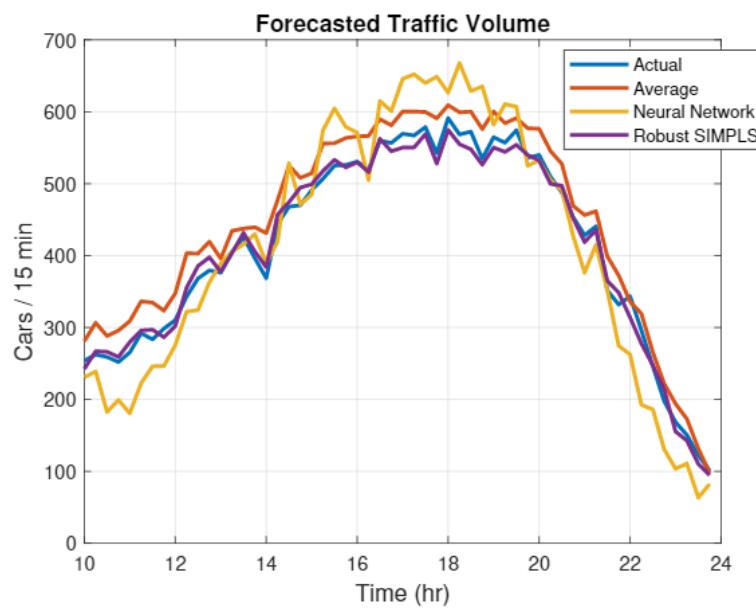


Figure 31. Line graph. Predictions produced by Robust SIMPLS, an NN, and the average traffic volume for October 3, 2019.

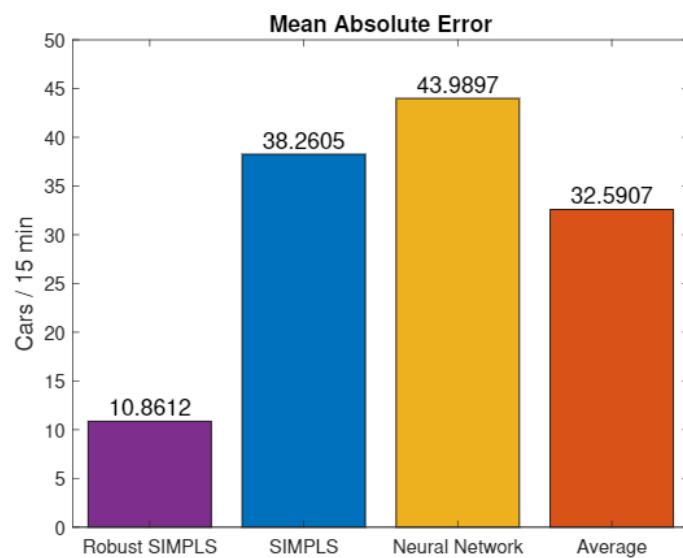


Figure 32. Bar graph. Comparison of MAE produced by Robust SIMPLS, SIMPLS, an NN, and average predictions for October 3, 2019.

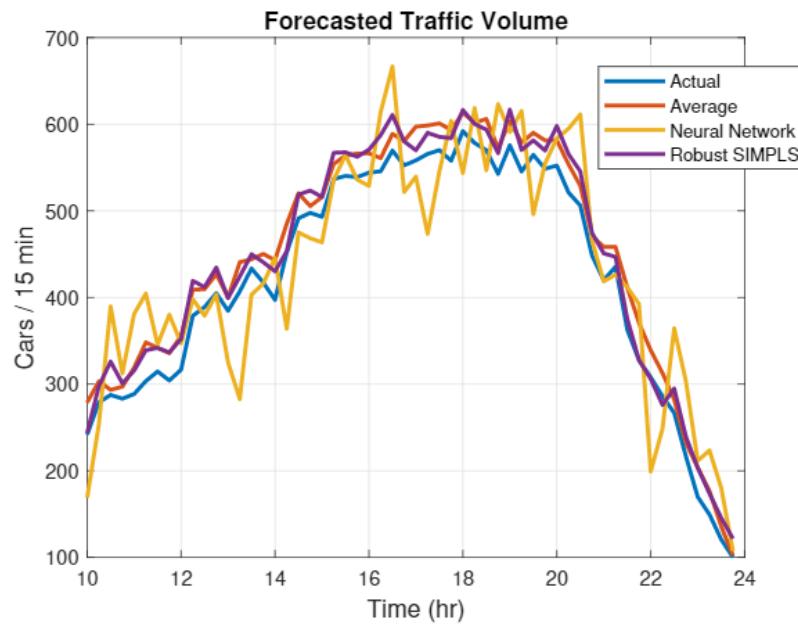


Figure 33. Line graph. Predictions produced by Robust SIMPLS, an NN, and the average traffic volume for October 1, 2019.

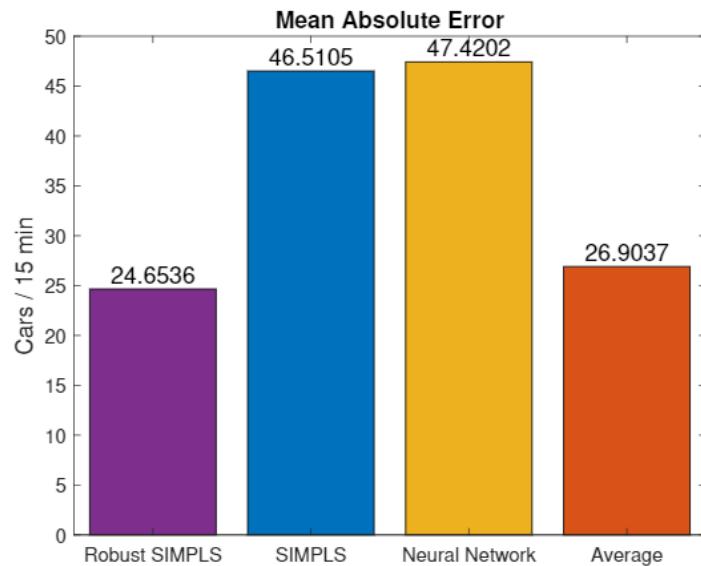


Figure 34. Bar graph. Comparison of MAE produced by Robust SIMPLS, SIMPLS, an NN, and average predictions for October 1, 2019.

ALGORITHM TRAINING AND TESTING IN SYNCHRO/SUMO ENVIRONMENT

OVERVIEW OF SIMULATION ENVIRONMENT

In order to evaluate the Robust SIMPLS algorithm, a simulation testbed was developed in Task IV using Synchro and SUMO software. The purpose of the simulation environment was twofold; it was intended to train the algorithm using different traffic demand profiles and to test the effectiveness of the recommendations made by the algorithm. Historical data from GDOT's ATSPM website were used to train the prediction algorithm. In simulated scenarios where intervention was recommended by the prediction algorithm, SUMO was used to demonstrate how altering the signal timings and ramp metering would increase efficiency at the interchange and neighboring intersections. To evaluate the algorithm, the team decided to focus on the DDI located at the I-285 interchange with Ashford Dunwoody Rd. and the neighboring signals; see figure 35 and 36. The Georgia Department of Transportation provided the research team with Synchro files for several DDIs in the state of Georgia, including the one at I-285 and Ashford Dunwoody Rd. These Synchro files contained important data regarding the geometry, phasing, and signal plans of intersections within and surrounding the interchange. Based on the software's simulation capabilities, the research team decided to use SUMO, an open-source, microscopic traffic simulator, rather than Synchro to test the algorithm. The team discovered, however, that there is not a straightforward way to import data from Synchro to SUMO. Conversion files and scripts were created to convert the Synchro geometry and phasing data from Synchro to SUMO.



Figure 35. Photo. Aerial view of the DDI at I-285 and Ashford Dunwoody Rd.
(Source: Google Earth)

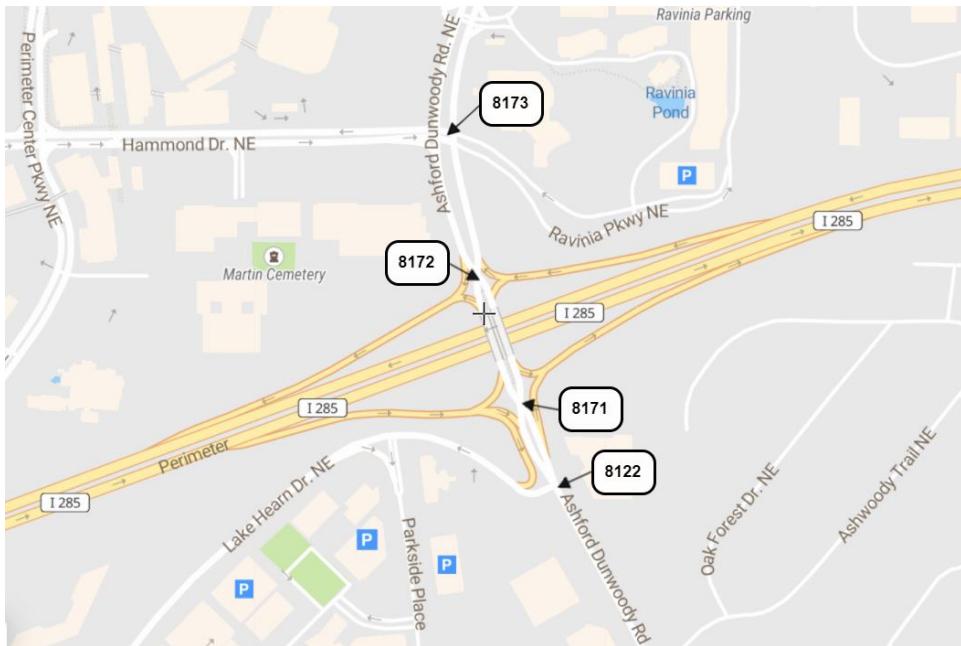


Figure 36. Map. DDI at I-285 and Ashford Dunwoody Rd. with Intersection Names.
(Source: Scribble Maps)

CONVERSION FROM SYNCHRO TO SUMO

Converting the DDI data from the provided Synchro files to the desired SUMO format was a complex task. Synchro is a data-heavy software that reads and outputs comma-separated value (CSV) files. SUMO, on the other hand, reads and outputs extensible markup language (XML) files. Synchro is able to read and interpret XML files that are output from SUMO, but the inverse is not true. SUMO is unable to independently interpret CSV files from Synchro. Because of this incompatibility, a need existed to create conversion files and scripts that could translate data from Synchro's CSV files into an XML format that SUMO could understand. For the conversion, the team used Python with the Pandas library. Figure 37 shows an overview of the conversion process used to move data from Synchro to SUMO. As shown in figure 37, Synchro and SUMO process and store road network data differently. To convert between the two software, both the geometry and phasing data needed to be translated.

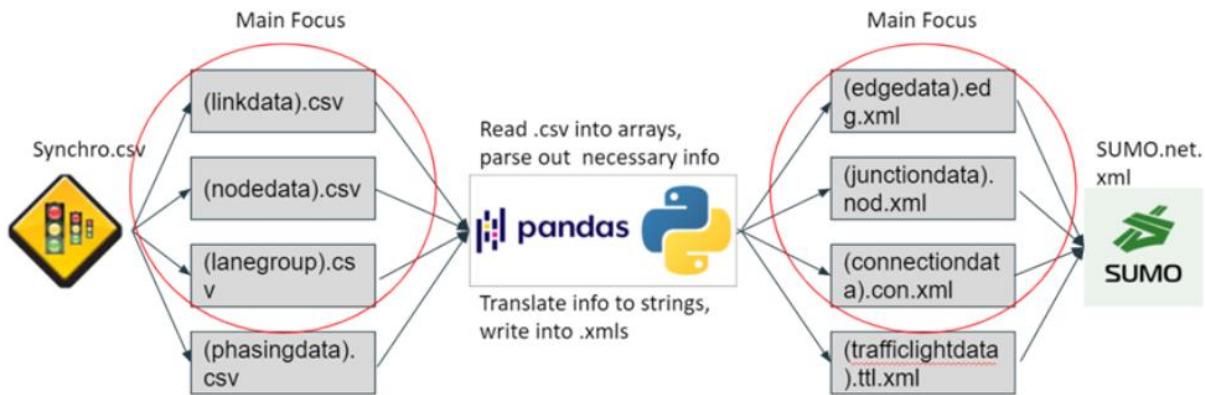


Figure 37. Diagram. Overview of the conversion process from Synchro to SUMO.

Geometry Conversion

In both Synchro and SUMO, there are geometry files that contain information regarding the layout of the road network. In Synchro, data regarding the road network are stored in three CSV files which, as shown in figure 37, are called link data, node data, and lane group. Road networks in Synchro are made up of links, which are segments of the roadway, and nodes, which are intersections that connect various links. Nodes can be either signalized or unsignalized.

The link data and lane group CSV files contain information regarding the links in the network, such as the number of lanes, the speed limit, and the volume flow rate. The node data CSV file contains information regarding the nodes in the network, including the type of control, phasing, and signal timing. SUMO has a somewhat different way of handling roadway geometry data. In SUMO, road networks are built using three XML files, which are referred to as edge data, junction data, and connection data. Thus, in SUMO, road networks are made up of edges and junctions. Edges in SUMO are the equivalent of links in Synchro. Junctions in SUMO are, in essence, the same as Synchro's nodes. The edge data XML file contains information regarding lanes, and the connection data XML file identifies which edges and nodes are connected in the network. The node data XML file contains data regarding where edges connect. Thus, the conversion of geometry files from Synchro to SUMO is not a one-to-one conversion. Each SUMO geometry file is made up of data that are contained in more than one Synchro geometry file. In spite of this, the research team was able to create scripts to convert the geometry files for the DDI at I-285 and Ashford Dunwoody Rd. from Synchro to SUMO. Converting the geometry files was an important first step in developing the testing environment in SUMO. Figure 38 shows the DDI at I-285 and Ashford Dunwoody Rd. in Synchro, and figure 39 shows the same DDI once converted to SUMO.



Figure 38. Diagram. DDI at I-285 and Ashford Dunwoody Rd. in Synchro.

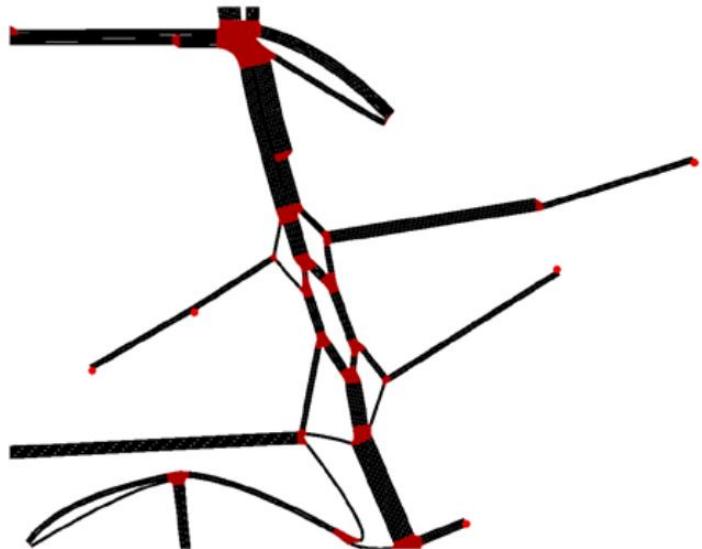


Figure 39. Diagram. DDI at I-285 and Ashford Dunwoody Rd. in SUMO.

Signal Timing and Phasing Conversion

For a road network in Synchro or in SUMO, there are phasing files that contain information regarding each signalized intersection's timing plans. Again, SUMO and Synchro handle these

phasing files differently. In Synchro, the phasing data are stored in a CSV file called *phasing data*. The Synchro phasing information is presented in terms of a classic ring-barrier diagram. As shown in figure 40, a ring-barrier diagram in Synchro is a linear timeline that displays the relationship and timing of an intersection's various phases. SUMO, as previously mentioned, handles phasing data in a different manner. In SUMO, the phasing data are stored in an XML file called *traffic light data*. Unlike Synchro, SUMO does not operate in terms of phases; rather, SUMO stores phasing data in terms of lanes. Figure 41 shows the timing scheme for an intersection in SUMO. Each of the positions in the “state” column represents a traffic signal for a single lane. The letter in that position describes the color of the light during that period (i.e., G means green, Y means yellow, etc.). Each row in the table represents a different period of time for the intersection throughout the cycle. The duration (“dur” column) of each row describes the amount of time that time period lasts. For example, say the last letter of the string in the “state” column represents a northbound left-turn lane at some intersection. During the highlighted period, which lasts 4.70 seconds, the northbound left-turn lane light will be yellow. Synchro describes phasing and timing like a ring-barrier diagram, whereas SUMO describes it as the status of all signals around the entire intersection on a step-by-step basis. This poses an issue when converting phasing files from Synchro to SUMO. Despite these challenges, the research team was able to develop a script to successfully translate the Synchro phasing files to SUMO.



Figure 40. Diagram. Ring-barrier diagram—the basis of Synchro’s phasing data.

Phases					
dur	min	max	state		nxt
9.30	8.00	9.30	srrrsrrrrrrrrrrrrGGGGGGG		
4.70			srrrsrrrrrrrrrrrGGGGGyy		
3.00			srrrsrrrrrrrrrrrGGGGGrr		
42.80	25.00	42.80	srrrGGGGGrrrrrrrrGGGGGrr		
4.50			srrrGGGGGrrrrrrrsyyyyrr		
2.70			srrrGGGGGrrrrrrrsrrrrrr		
31.70	8.00	31.70	srrrGGGGGGGGrrrsrrrrrr		
0.60			srrrGGGGGyyyyrrrsrrrrrr		
4.70			srrrsyyyyyrrrrrsrrrrrr		
3.00			srrrsrrrrrrrrrrrsrrrrrr		
27.20	26.00	27.20	GGGGsrrrrrrrrrrsrrrrrr		

Figure 41. Screen capture. Phasing scheme as portrayed in SUMO.

PREPARATION OF TRAFFIC DATA FOR SIMULATION

To test and train the prediction algorithm, the ability to input different traffic demand profiles into the SUMO simulation was essential. The historical data used for these purposes were pulled from GDOT’s ATSPM website. These traffic volume data are comprised of turning count movements at each intersection. The research team utilized one of SUMO’s built-in algorithms called “flowrouter.py.” This algorithm is able to calculate routes for all the vehicles within the simulation based on the turning count movements at detectors in the simulated network. To complete the simulation, the research team developed an XML file that describes where detectors are located within the system. The addition of these detectors to the simulation was imperative as they allowed traffic volumes to be input to the simulation using the “flowrouter.py” algorithm. After incorporating geometry, phasing, volume, and detector data, some issues still remained with the simulation. The research team had to manually tweak some factors, such as road priorities and allowable U-turns, to accurately simulate real-world conditions at the DDI.

TESTING AND TRAINING OF ALGORITHM USING SUMO

The research team was able to successfully create a simulation environment in SUMO for the I-285 and Ashford Dunwoody DDI by converting the geometry and phasing files from Synchro to SUMO. The team was also able to convert available turning-movement data into vehicle routes that were required by SUMO. At this point, the research team was prepared to test and train the prediction algorithm using different traffic demand profiles.

To test the algorithm using the SUMO simulation, four Case Studies with a day's worth of atypical traffic flows were created as follows. Data between November 2020 and January 2021 was collected for intersections 8171, 8172, 8173, and 8122. This data is represented as a large matrix containing turning movement counts across the four intersections over the three-month time period. First, a singular value decomposition of the measured traffic data was computed. Then, the singular values in this decomposition were randomly perturbed between plus or minus 15 percent of their nominal value to create four different synthetic profiles. The synthetic traffic profiles were selected from this newly generated synthetic data matrix. As a result, all synthetic profiles are within +/- 15% of the total volume on an average day. Each of the Case Studies are described below:

Case Study 1

The volume across the entire network for Case Study 1 is higher than average. The most substantial difference in volume between Case Study 1 and the average case occurs at the northernmost arterial (8173), with average volume differences ranging from 100 to 175 vehicles. Increased flow volume is given to this intersection's northbound-through, southbound-through, and eastbound-right movements, all of which contribute to high through-traffic going south

through the DDI. Though less substantial, increased flow volume is also given to the northbound-through movements in the southernmost arterial (8122) and the southbound-through movement in the southern DDI intersection (8171), with average volume differences ranging from 60 to 80 vehicles. This Case Study has the highest overall volume of traffic as well as some of the most volatile traffic patterns, with volume measurements changing rapidly every 15 minutes. Figure 42 shows a scatterplot of the traffic volumes for each movement over the course of the day for Case Study 1. Figure 43 shows a bar graph showing the difference in volume between the average day and the Case Study 1 day for each movement from 2:00-9:00PM.

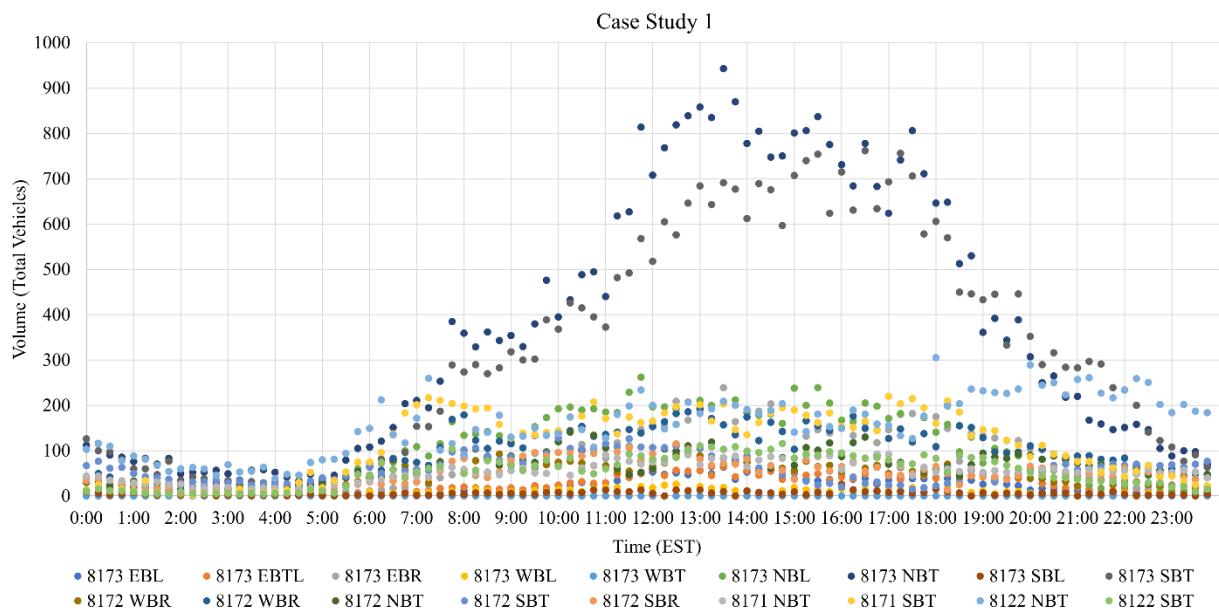


Figure 42. Scatter plot. Volume by Movement over day for Case Study 1.

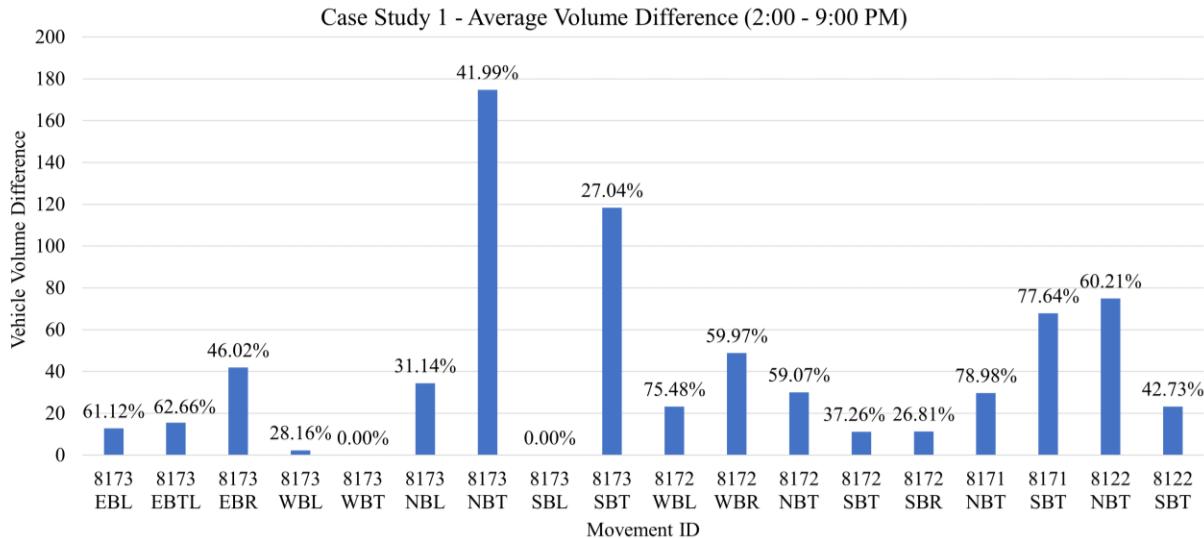


Figure 43. Bar graph. Average Volume Difference by Movement for Case Study 1.

Case Study 2

Like in Case Study 1, the volume across the entire network for Case Study 2 is higher than average. However, the trends in the flows over time are more steady and less volatile than Case Study 1. Overall, compared with Case Study 1, the volume differences for each movement are more uniform and less drastic between each intersection. For example, in the northernmost arterial (8173), the northbound-through movement has the highest average difference in traffic but this difference is consistently less than 100 vehicles. Decreased flow volume is given to 8173's southbound movements, whose difference in volumes is similar to that of the north DDI intersection (8172) westbound-right movement, the south DDI intersection (8171) southbound-through movement, and the southernmost arterial (8122) northbound-through movement, ranging from 35 to 45 cars. In Case Study 2, decreased flow volume is given to peripheral movements which contribute little to no vehicles to the main sources of traffic flow. Figure 44 shows a

scatterplot of the traffic volumes for each movement over the course of the day for Case Study 2.

Figure 45 shows a bar graph showing the difference in volume between the average day and the Case Study 2 day for each movement from 2:00-9:00PM.

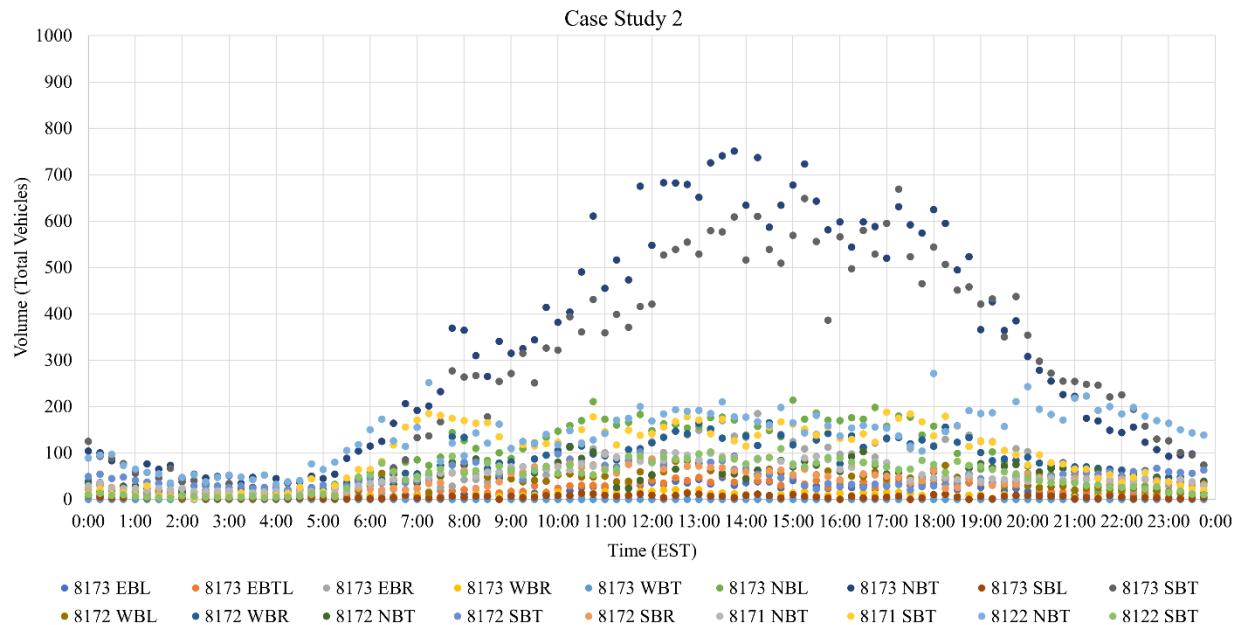


Figure 44. Scatter plot. Volume by Movement over day for Case Study 2.

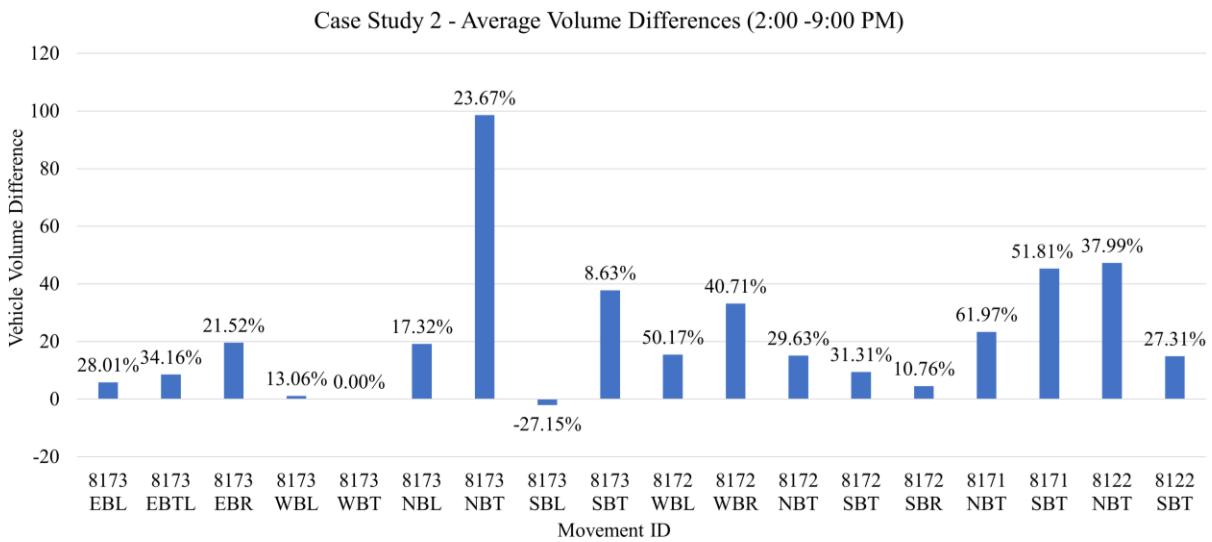


Figure 45. Bar graph. Average Volume Difference by Movement for Case Study 2.

Case Study 3

Like Case Studies 1 and 2, Case Study 3 generally has higher volumes throughout the day than the average day. The key difference between Case Study 3 and the prior two cases is that this case has a high volume volatility for each 15-minute interval. The volume changes are more drastic and less uniform than even Case Study 1. However, greatly increased flow volumes are given to the southern half of the network (8122 and 8171). This means that, while traffic patterns are very volatile, the average difference in volume between the northern arterial and the southern intersections are much closer together than they were in Case Studies 1 and 2. Overall, periphery movements behave similarly to the previous two cases. Figure 46 shows a scatterplot of the traffic volumes for each movement over the course of the day for Case Study 3. Figure 47 shows a bar graph showing the difference in volume between the average day and the Case Study 3 day for each movement from 2:00-9:00PM.

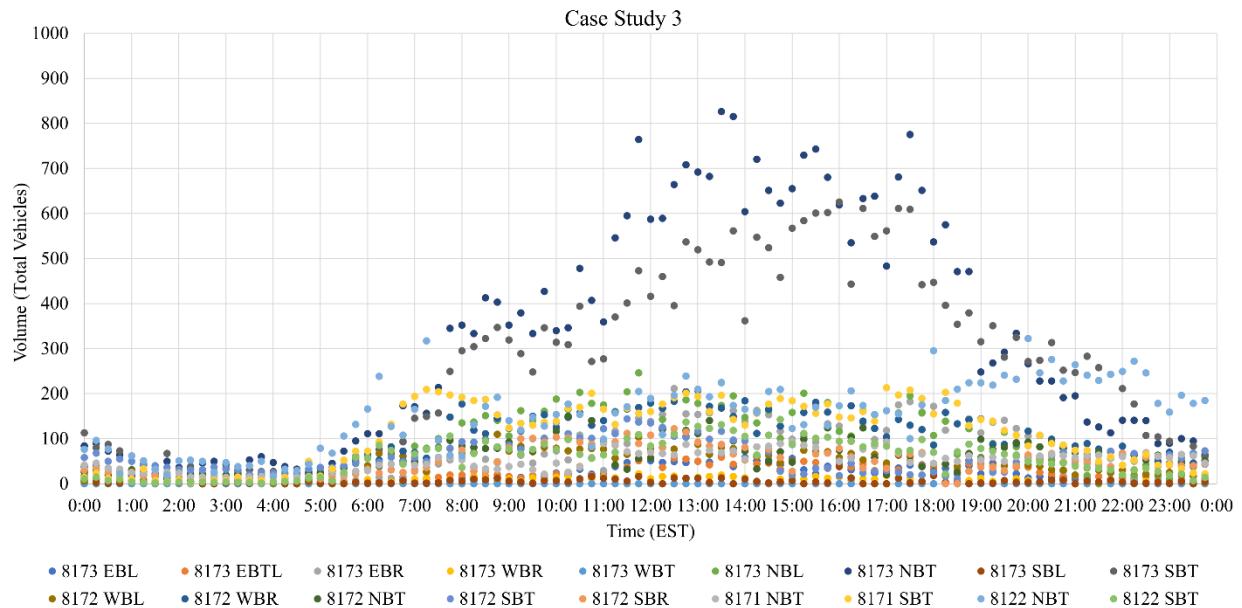


Figure 46. Scatter plot. Volume by Movement over day for Case Study 3.

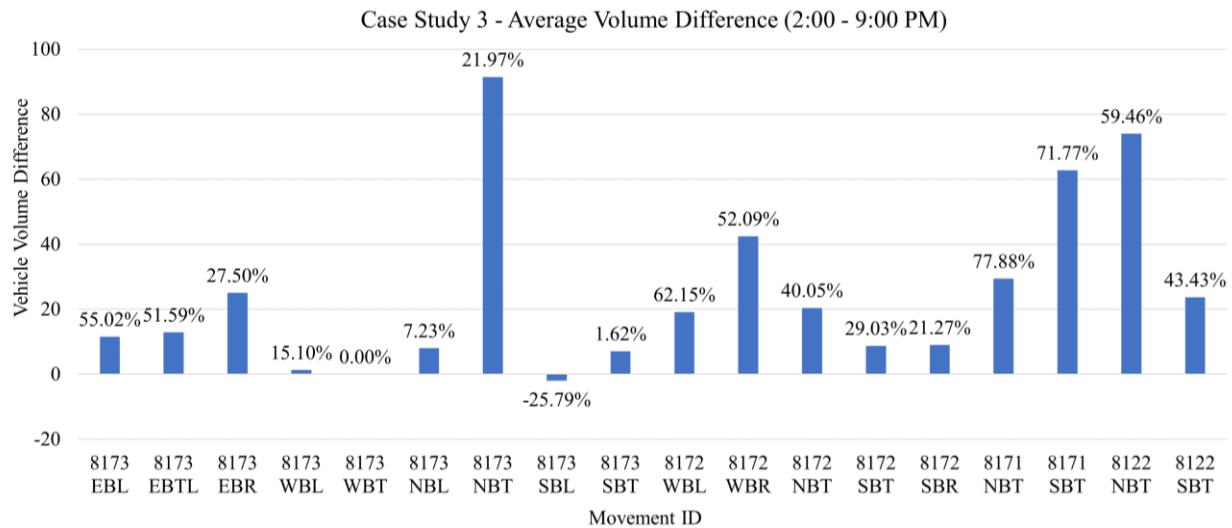


Figure 47. Bar graph. Average Volume Difference by Movement for Case Study 3.

Case Study 4

Case Study 4 is significantly different from the previous cases. Case Study 4 is a scenario where heavy traffic flows occur at the network's periphery movements and slightly increased flow volumes are given to northbound- and southbound-through movements. On average, flows for all through-movements in the network are lower than they are for the average day. All of the westbound- and eastbound-movements contribute higher volumes than they would on an average day. The greatest increases in flow volumes are given to the northern DDI intersection's (8172) westbound movement, representing traffic exiting I-285 westbound. The second greatest increases in flow volumes are given to the northernmost arterial (8173) eastbound-movements, with the eastbound right-movement contributing the most traffic from this direction. Flows for intersections in the southern half of the network are only slightly lower than average, while northbound-through and southbound-through traffic for 8173 ranges well below average. Figure 48 shows a scatterplot of the traffic volumes for each movement over the course of the day for

Case Study 4. Figure 49 shows a bar graph showing the difference in volume between the average day and the Case Study 4 day for each movement from 2:00-9:00PM.

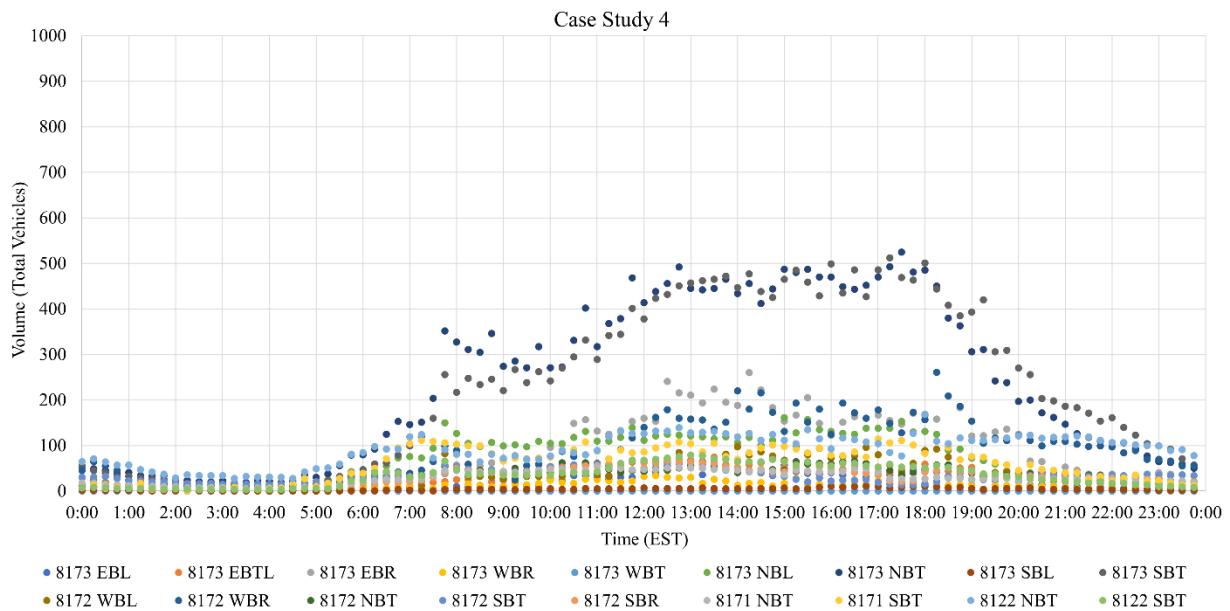


Figure 48. Scatter plot. Volume by Movement over day for Case Study 4.

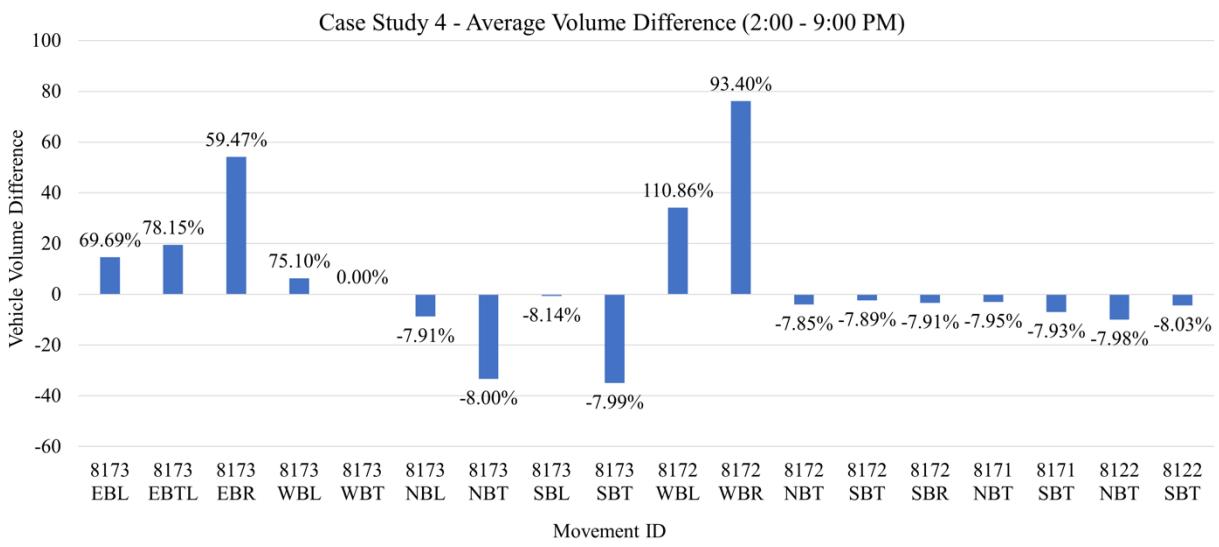


Figure 49. Bar graph. Average Volume Difference by Movement for Case Study 4.

Figures 50 through 66 show plots of the traffic flow volumes over the day for each movement in the network for all Case Studies.

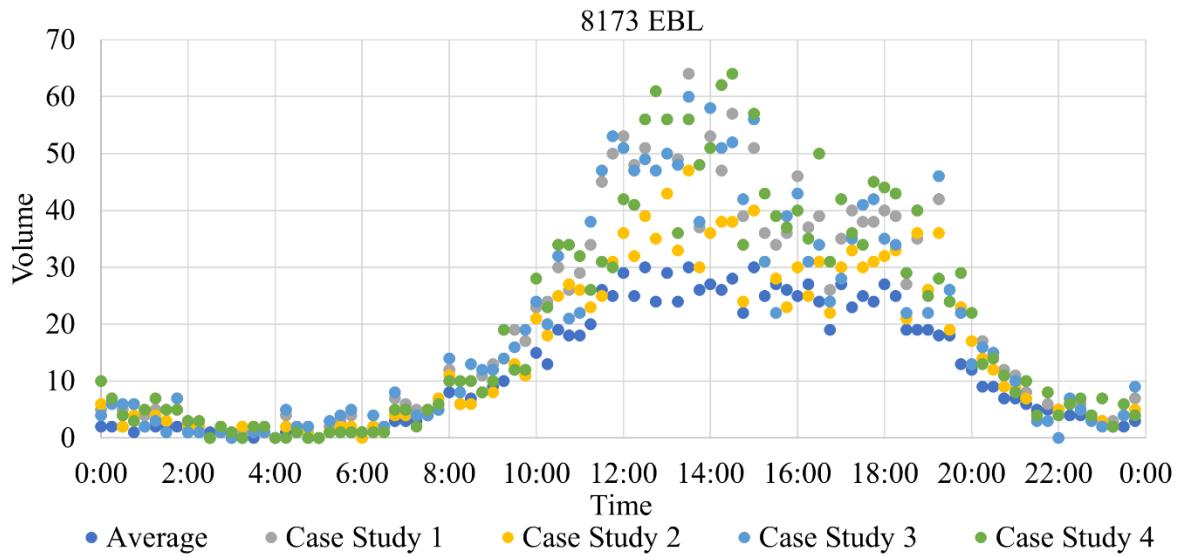


Figure 50. Scatter plot. Volume over day by Case Study for 8173 Eastbound Left Movement.

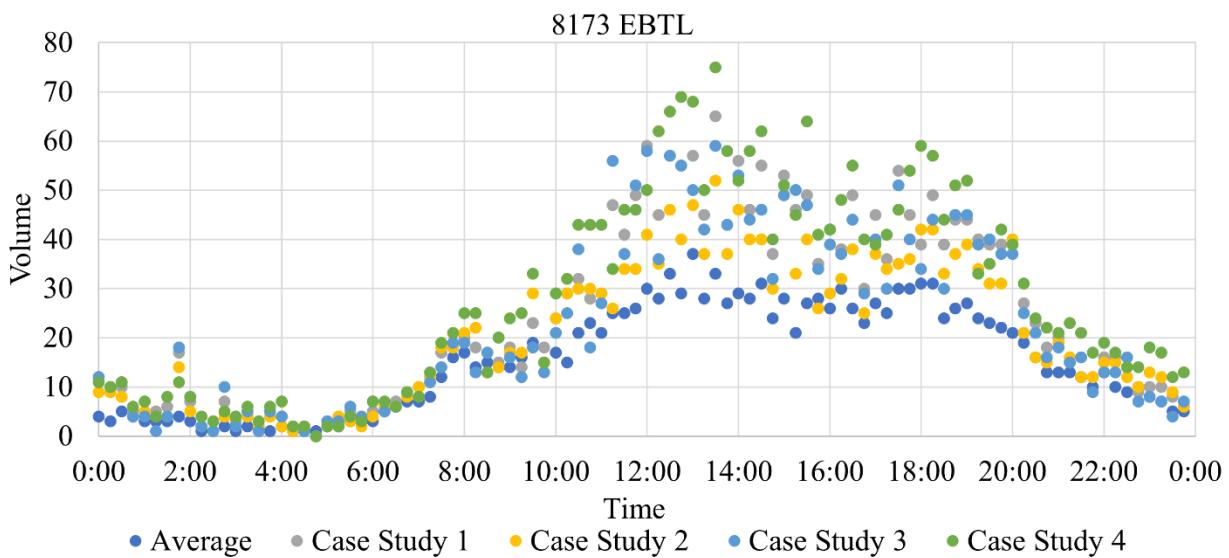


Figure 51. Scatter plot. Volume over day by Case Study for 8173 Eastbound Through-Left Movement.

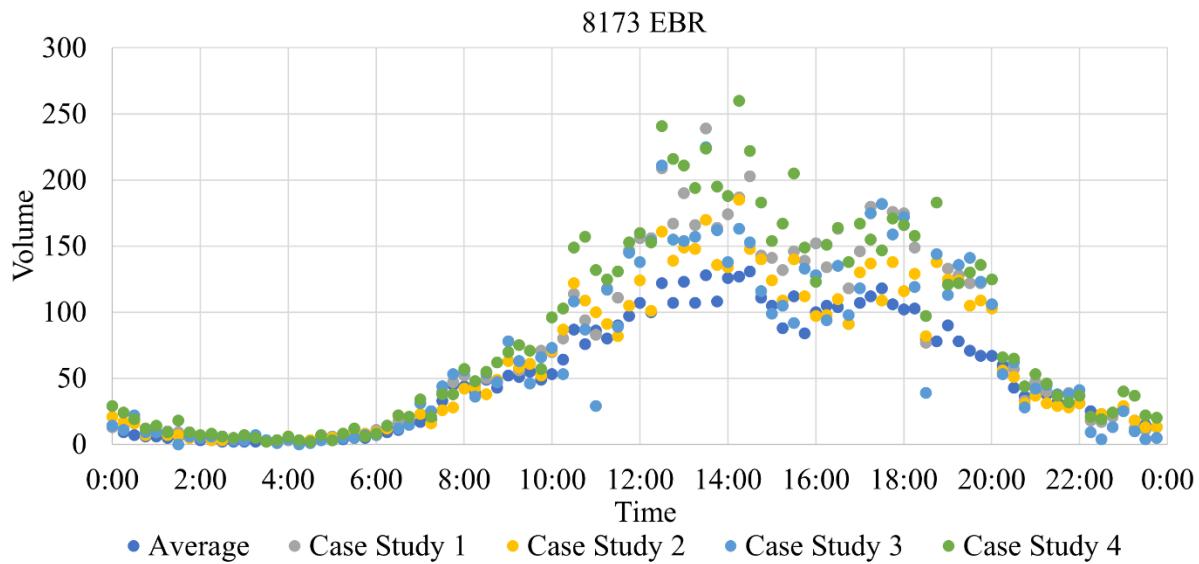


Figure 52. Scatter plot. Volume over day by Case Study for 8173 Eastbound Right Movement.

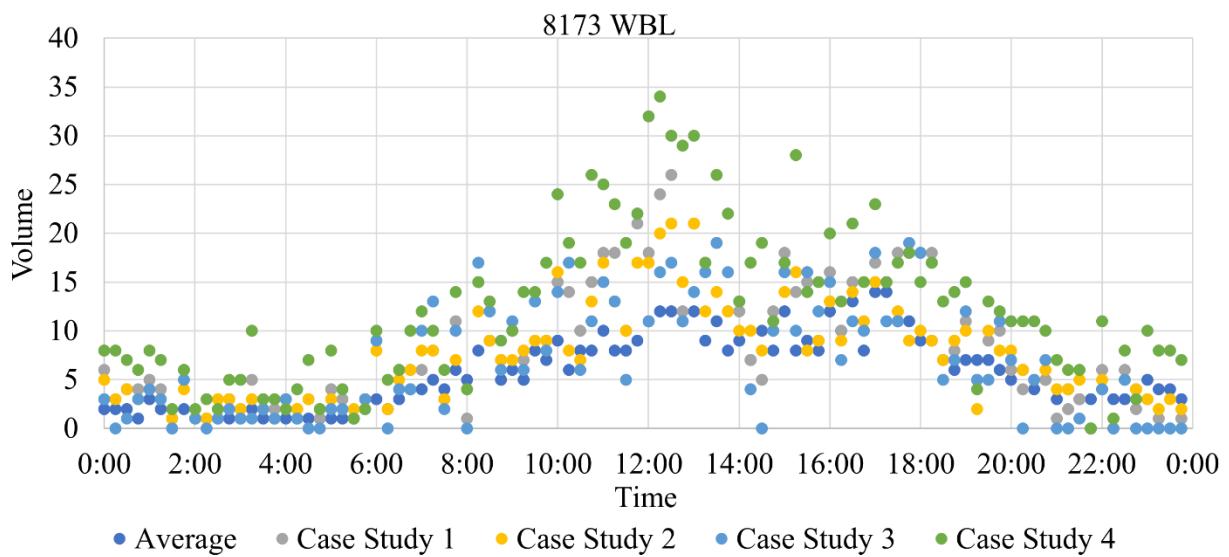


Figure 53. Scatter plot. Volume over day by Case Study for 8173 Westbound Left Movement.

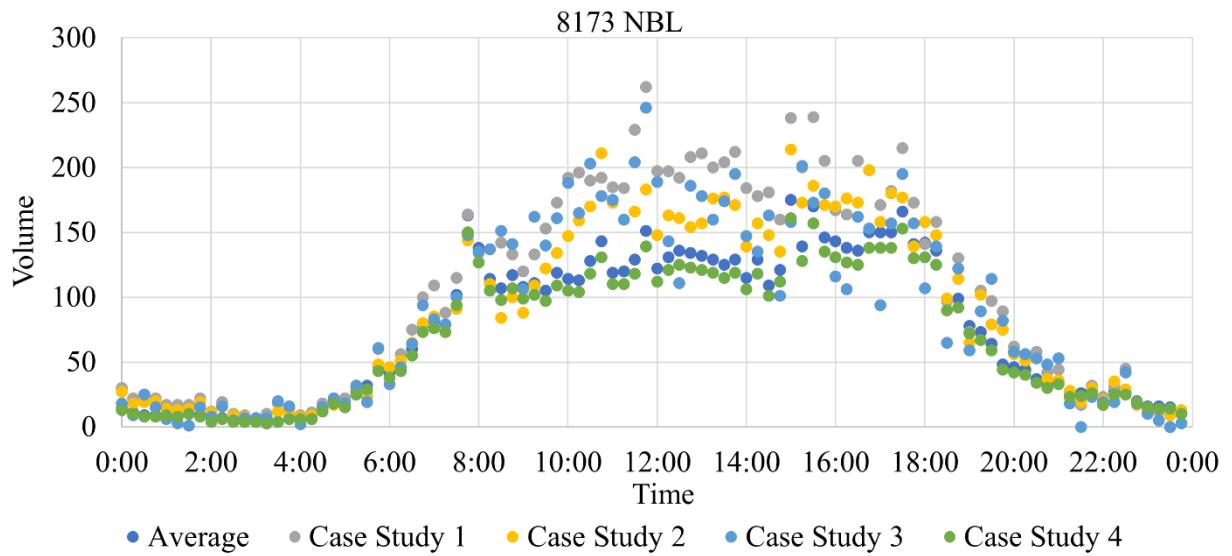


Figure 54. Scatter plot. Volume over day by Case Study for 8173 Northbound Left Movement.

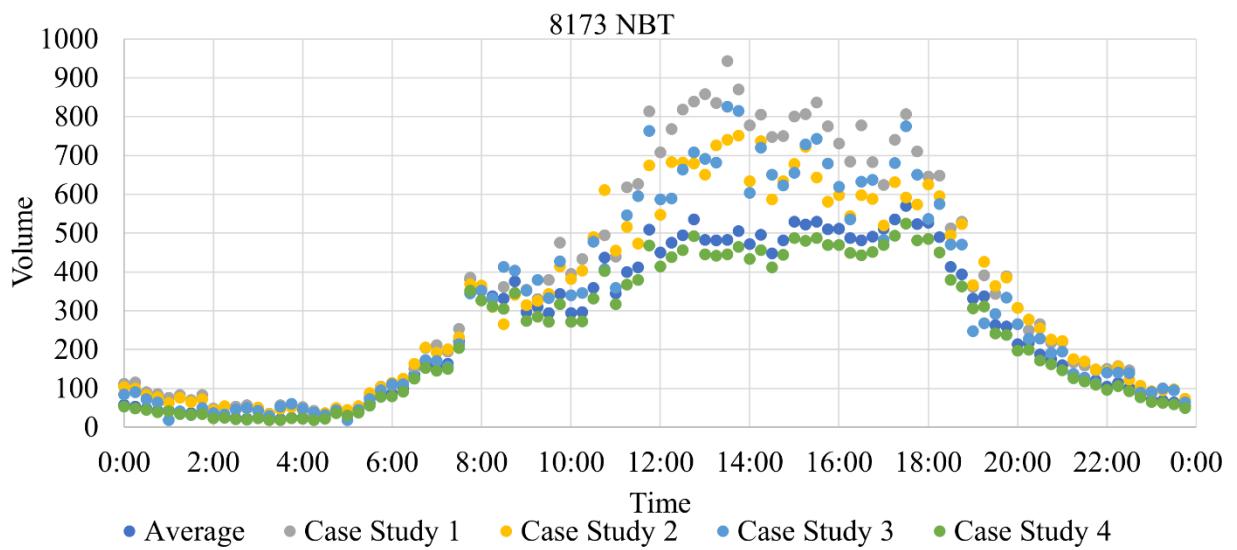


Figure 55. Scatter plot. Volume over day by Case Study for 8173 Northbound Through Movement.

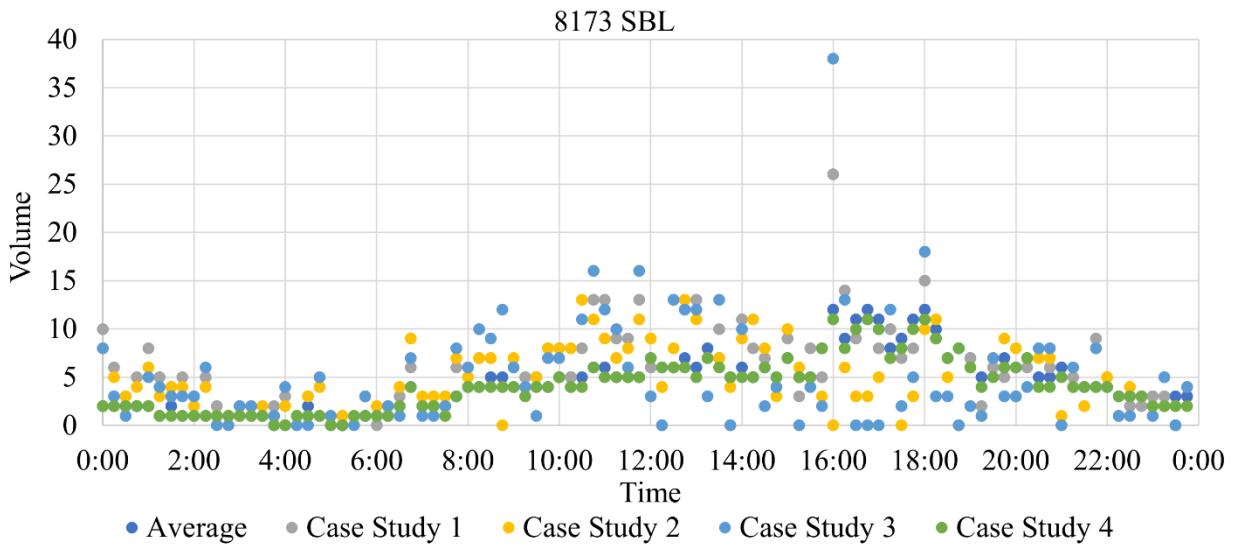


Figure 56. Scatter plot. Volume over day by Case Study for 8173 Southbound Left Movement.

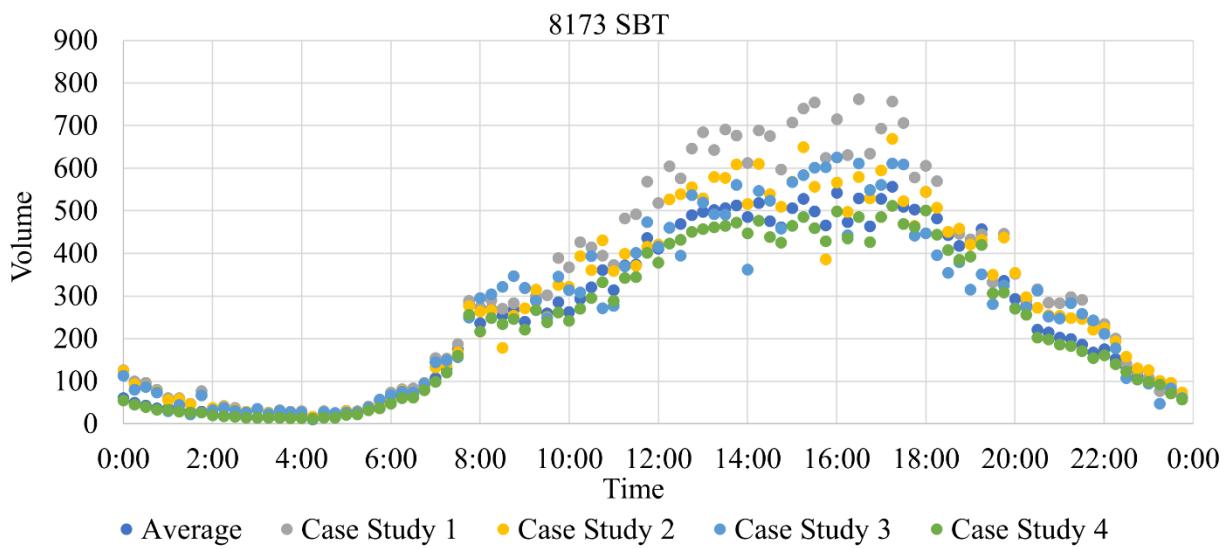


Figure 57. Scatter plot. Volume over day by Case Study for 8173 Southbound Through Movement.

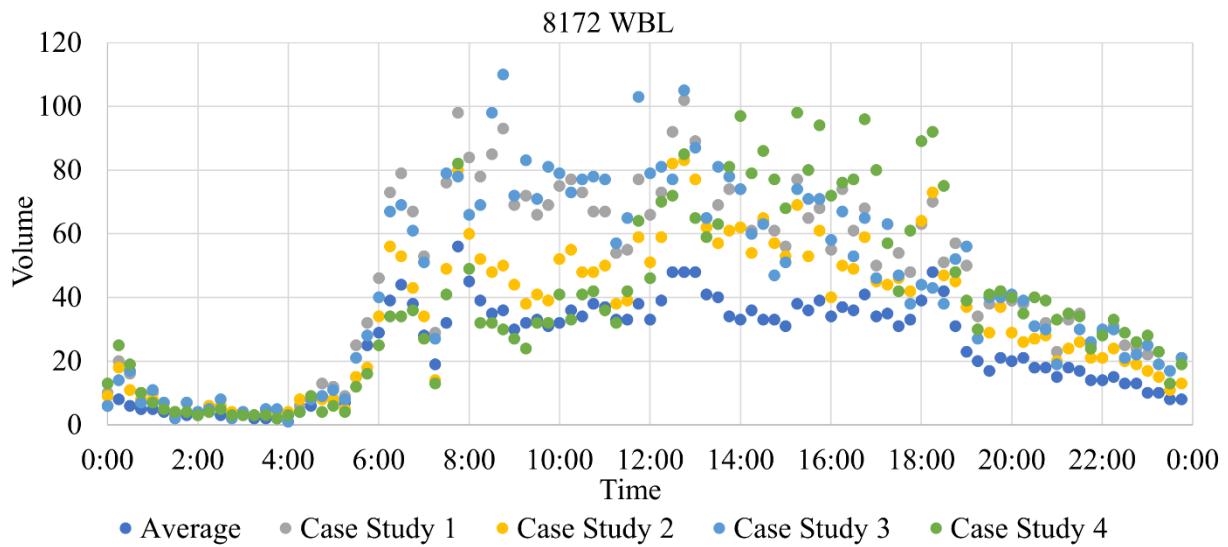


Figure 58. Scatter plot. Volume over day by Case Study for 8172 Westbound Left Movement.

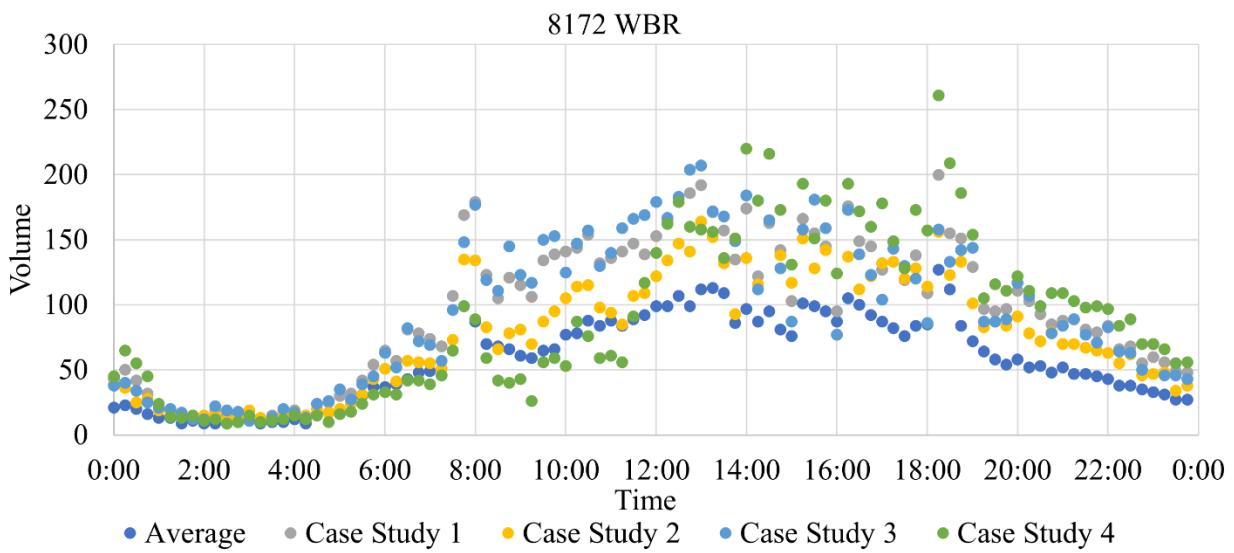


Figure 59. Scatter plot. Volume over day by Case Study for 8172 Westbound Right Movement.

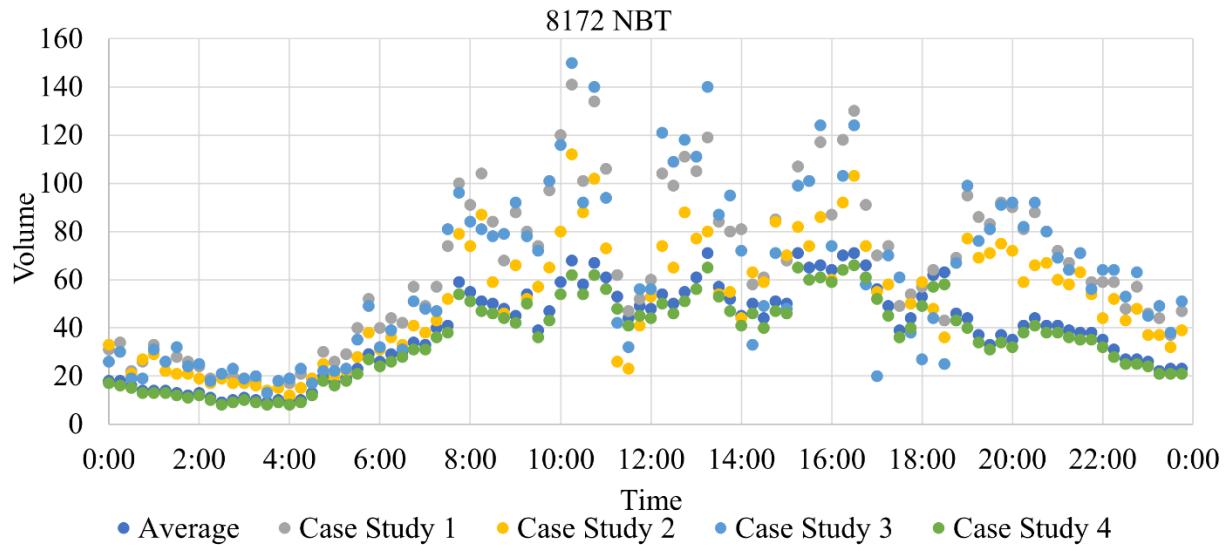


Figure 60. Scatter plot. Volume over day by Case Study for 8172 Northbound Through Movement.

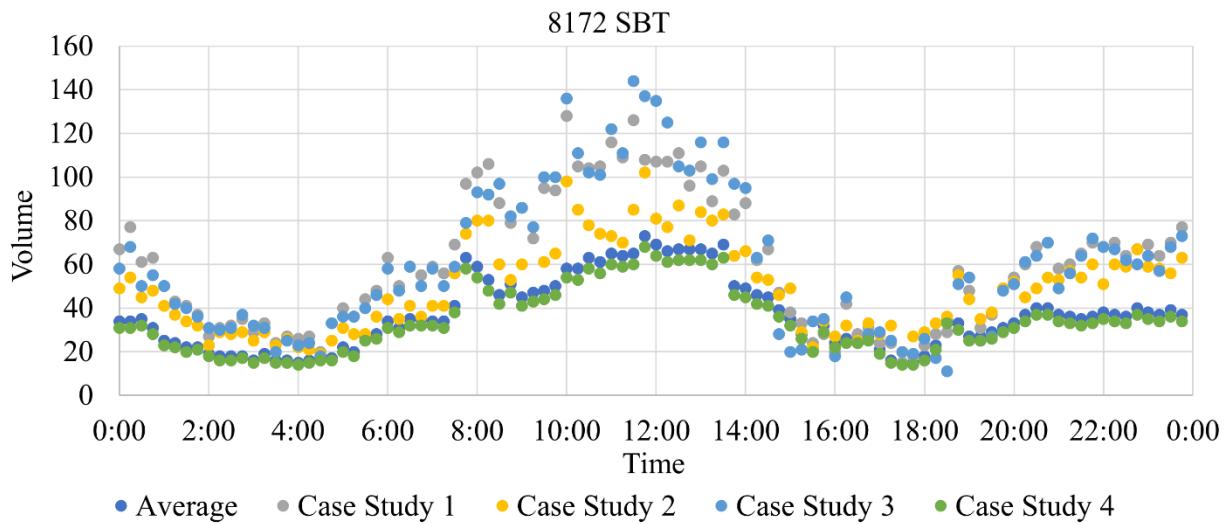


Figure 61. Scatter plot. Volume over day by Case Study for 8172 Southbound Through Movement.

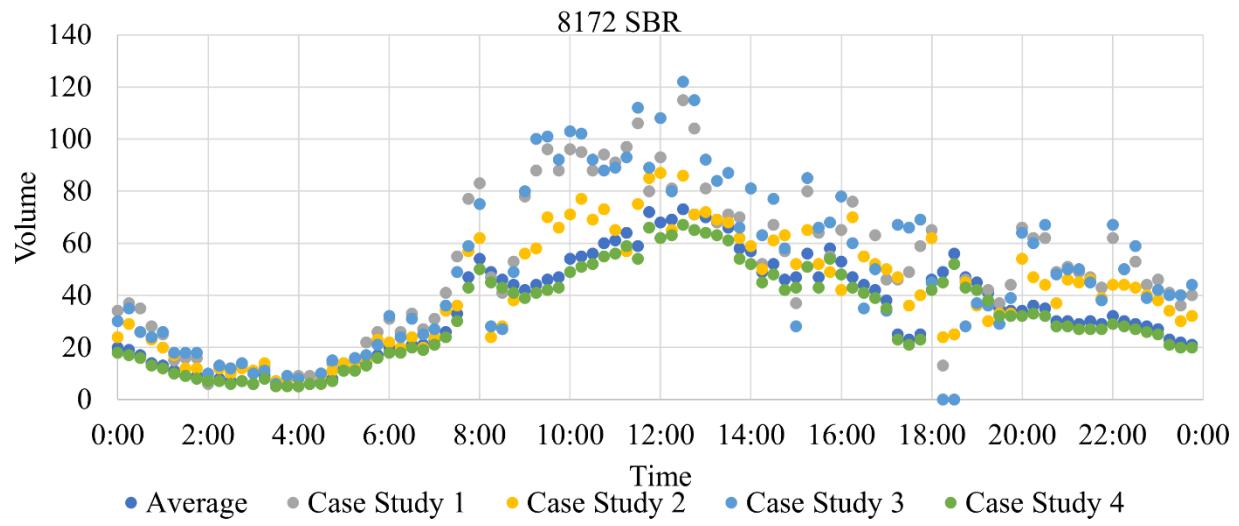


Figure 62. Scatter plot. Volume over day by Case Study for 8172 Southbound Right Movement.

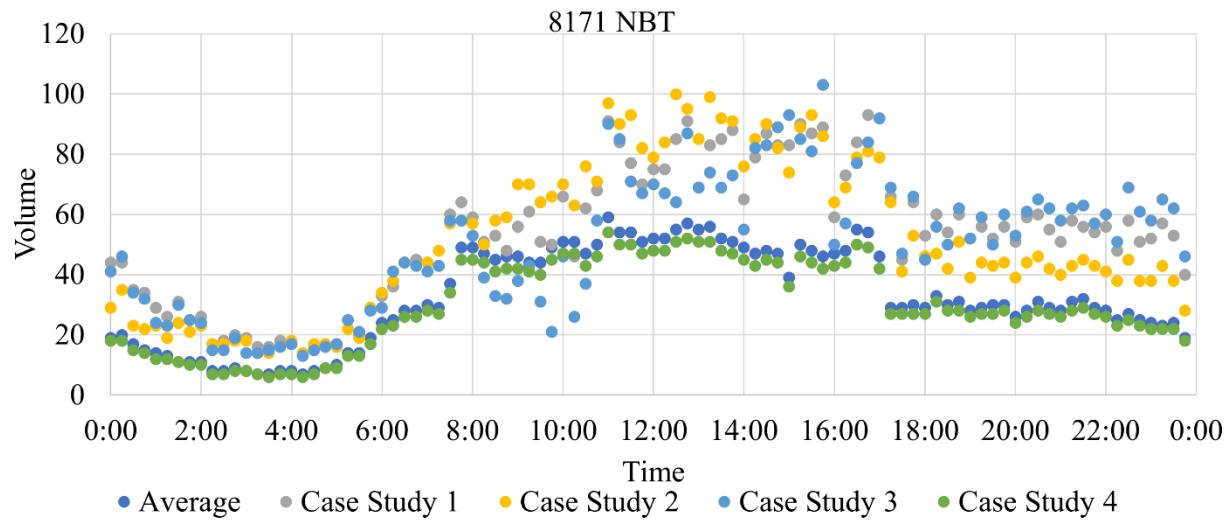


Figure 63. Scatter plot. Volume over day by Case Study for 8171 Northbound Through Movement.

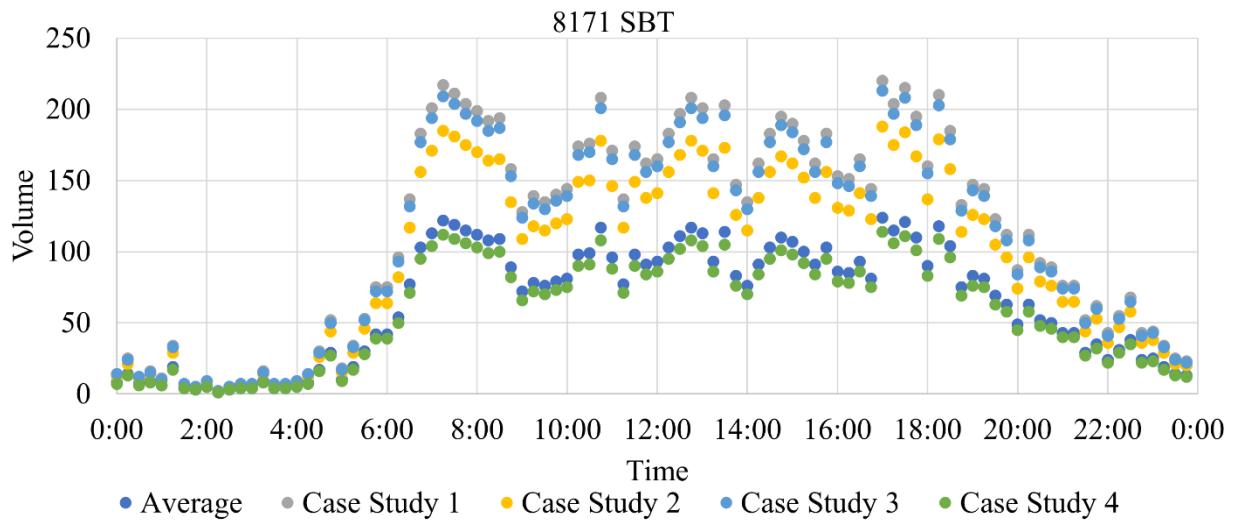


Figure 64. Scatter plot. Volume over day by Case Study for 8171 Southbound Through Movement.

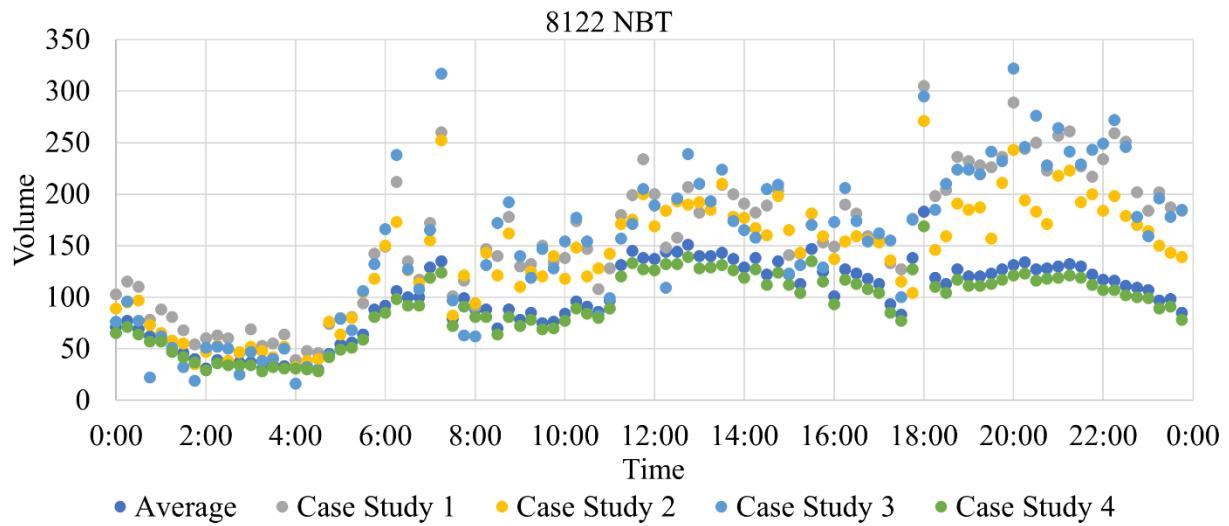


Figure 65. Scatter plot. Volume over day by Case Study for 8122 Northbound Through Movement.

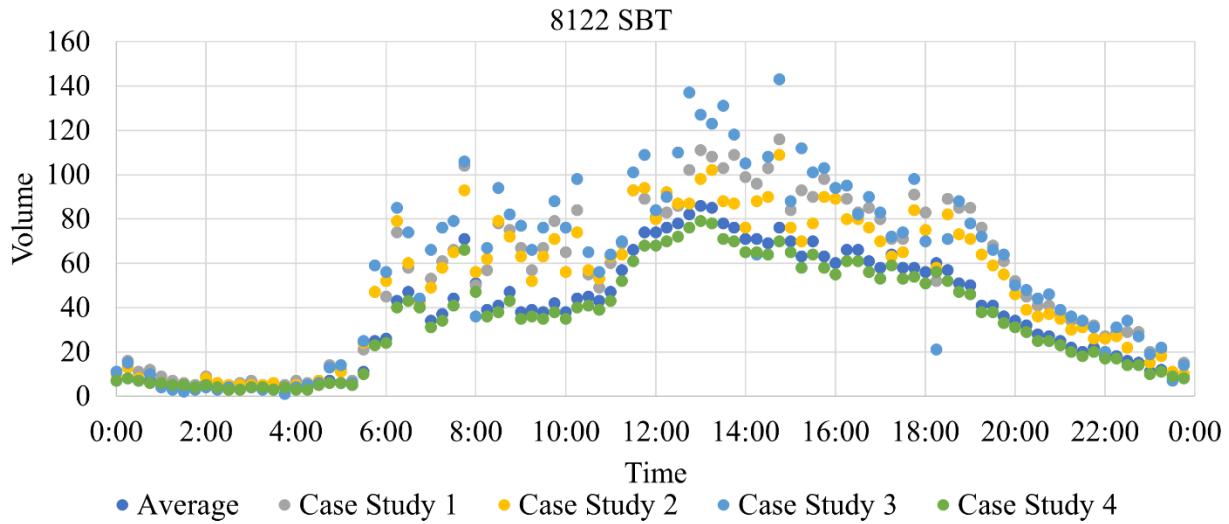


Figure 66. Scatter plot. Volume over day by Case Study for 8122 Southbound Through Movement.

Each of the Case Studies was used to assess the predictive accuracy of the algorithm. The traffic flow volumes from 12:00AM to 2:00PM were input into the algorithm. The algorithm was then used to predict the traffic flows from 2:00PM to 9:00PM. This procedure was used for each of the atypical volume Case Studies. Figure 67 shows the average predicted flows, the average actual flows, and the average day flows by movement for the period between 2:00PM and 9:00PM for Case Study 1. Figure 68 shows the percentage error between the average predicted flow and the average actual flows for Case Study 1.

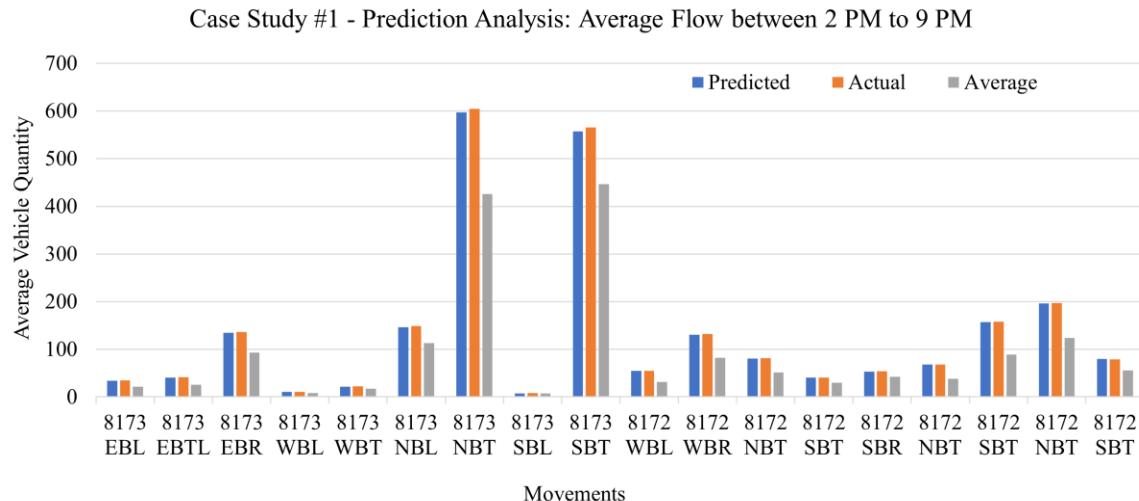


Figure 67. Bar graph. Average Flows by Movement for Prediction Period for Case Study 1.

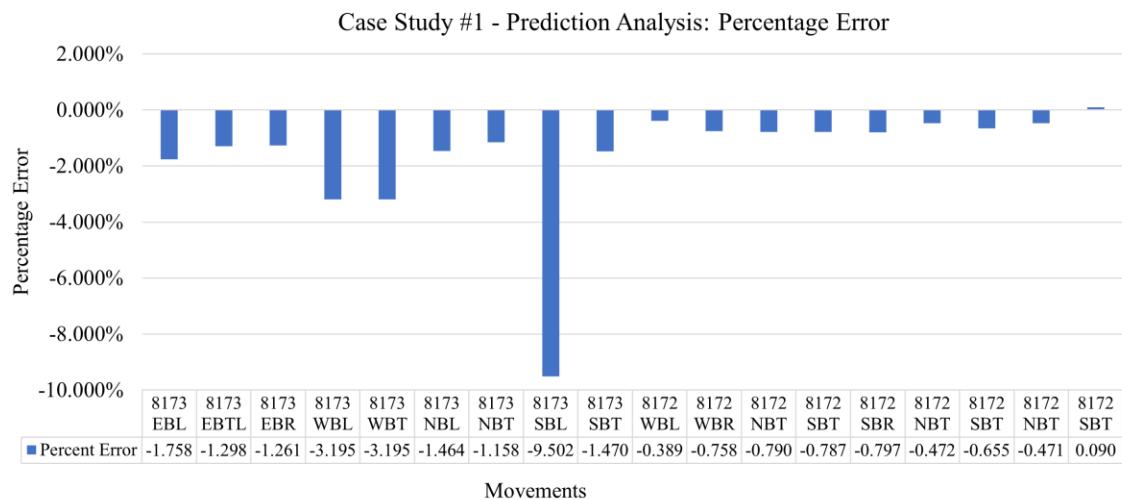


Figure 68. Bar Graph. Percentage Error between Actual and Predicted Flows for Case Study 1.

Figure 69 shows the average predicted, average actual, and average day flows by movement for the period between 2:00PM and 9:00PM for Case Study 2. Figure 70 shows the percentage error between the average predicted and average actual flows for Case Study 2.

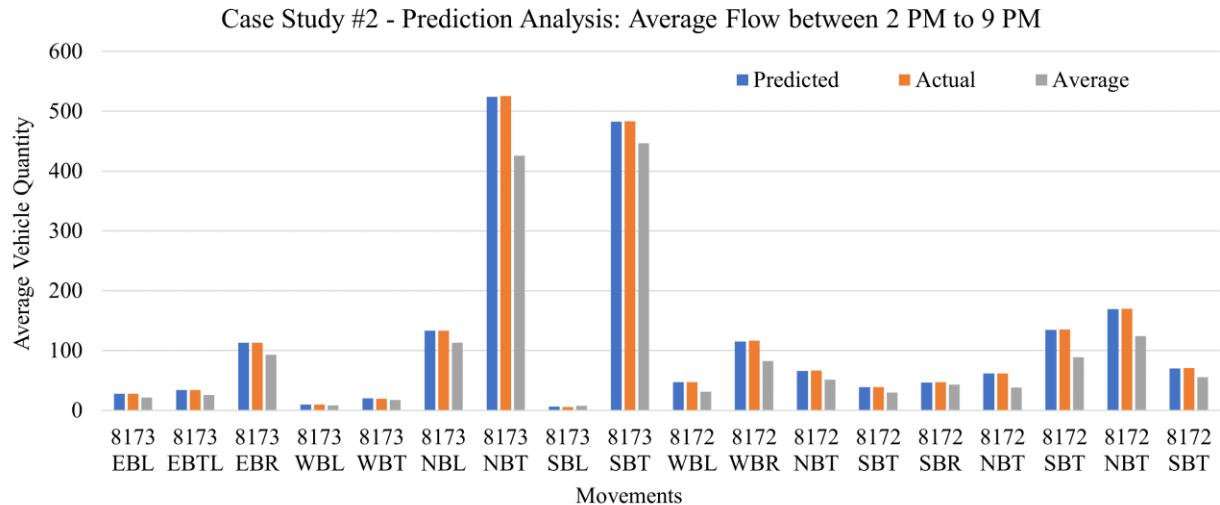


Figure 69. Bar Graph. Average Flows by Movement for Prediction Period for Case Study 2.

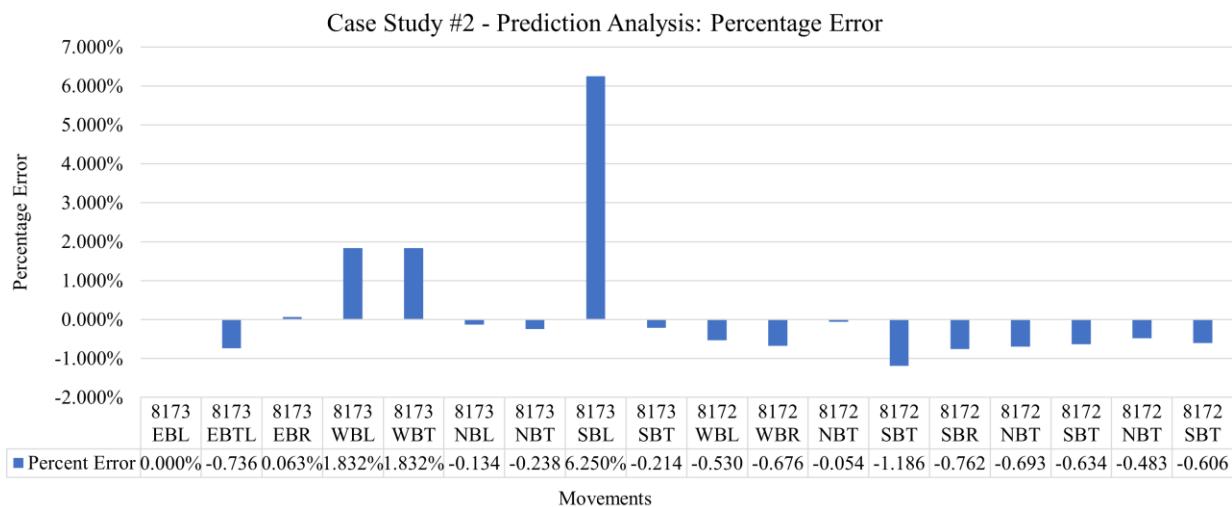


Figure 70. Bar Graph. Percentage Error between Actual and Predicted Flows for Case Study 2.

Figure 71 shows the average predicted, average actual, and average day flows by movement for the period between 2:00PM and 9:00PM for Case Study 3. Figure 72 shows the percentage error between the average predicted and average actual flows for Case Study 3.

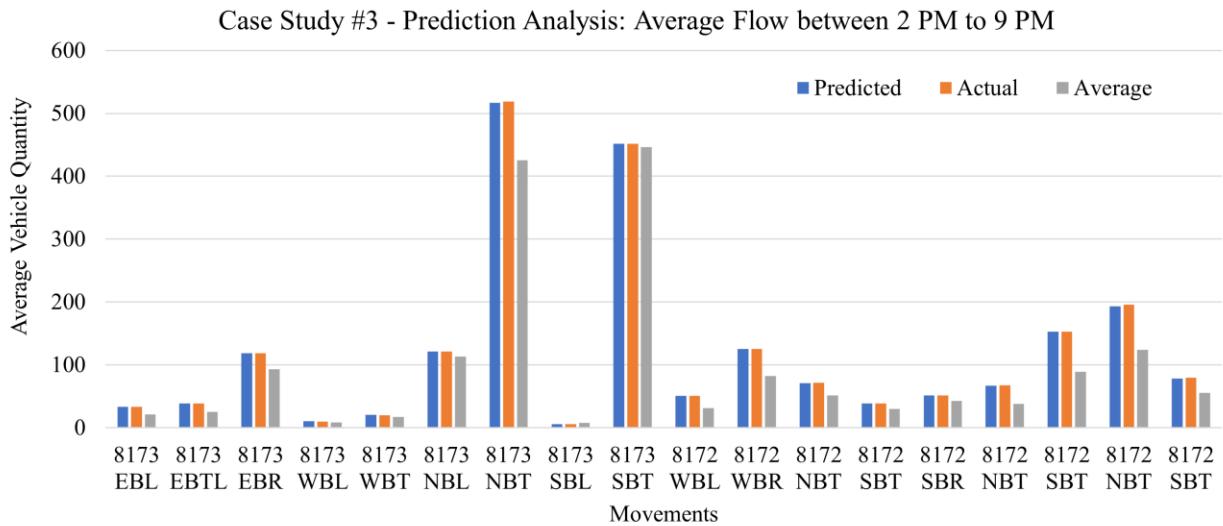


Figure 71. Bar Graph. Average Flows by Movement for Prediction Period for Case Study 3.

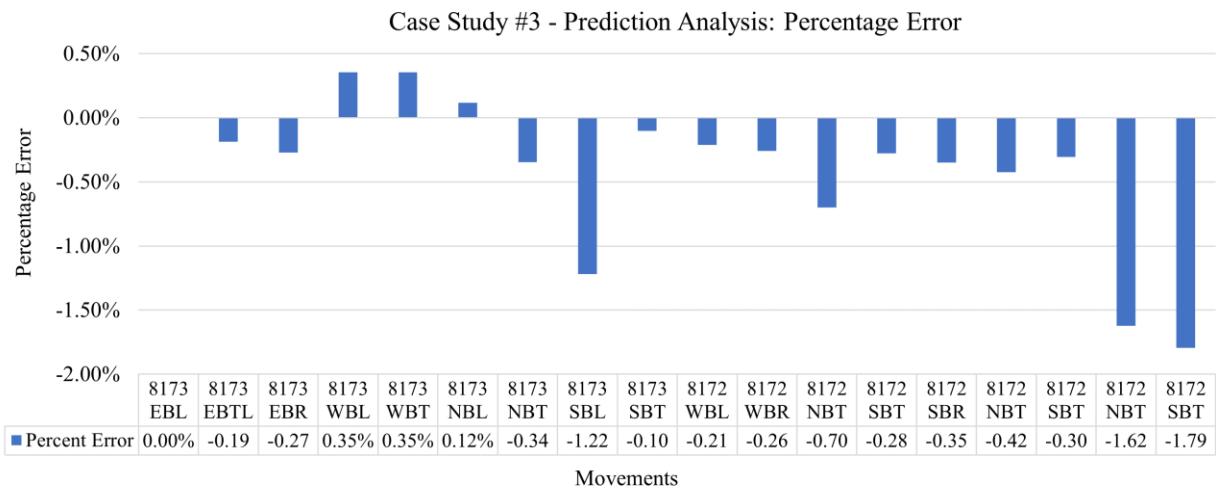


Figure 72. Bar Graph. Percentage Error between Actual and Predicted Flows for Case Study 3.

Figure 73 shows the average predicted, average actual, and average day flows by movement for the period between 2:00PM and 9:00PM for Case Study 4. Figure 74 shows the percentage error between the average predicted and average actual flows for Case Study 4.

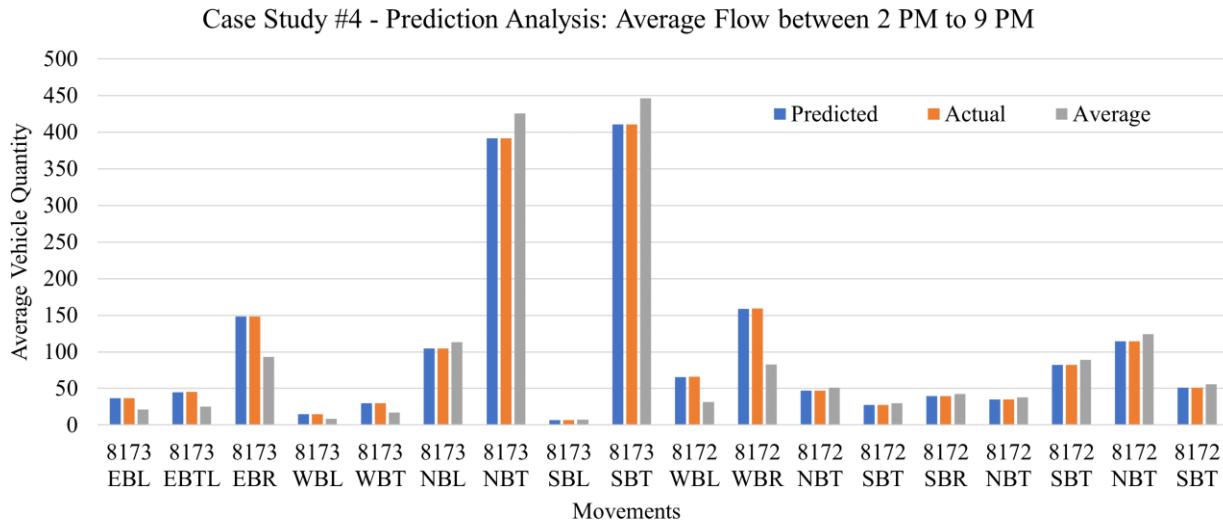


Figure 73. Bar Graph. Average Flows by Movement for Prediction Period for Case Study 4.

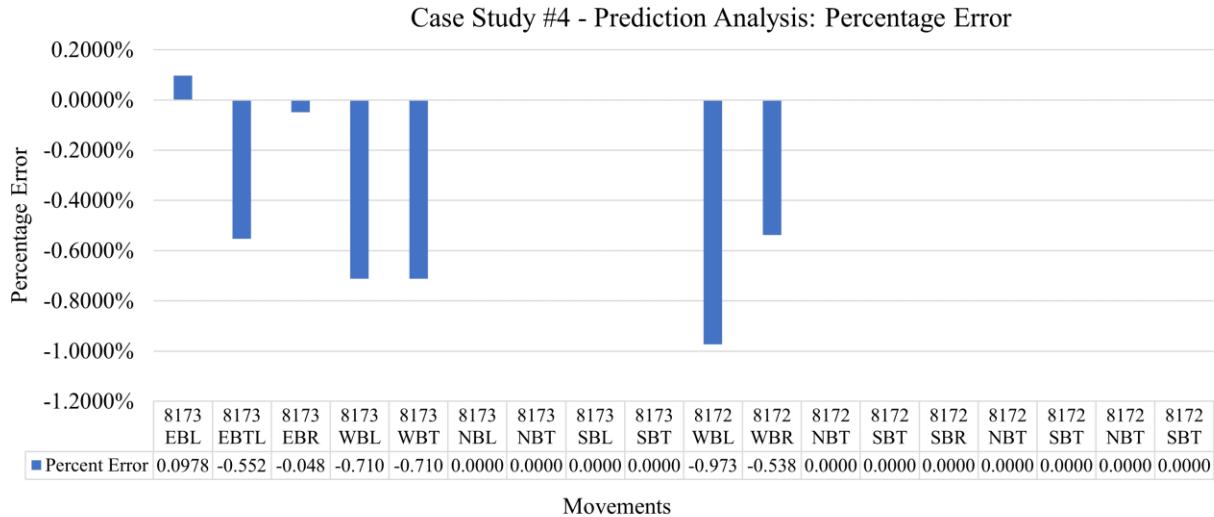


Figure 74. Bar Graph. Percentage Error between Actual and Predicted Flows for Case Study 4.

Several conclusions can be made by analyzing Figures 67-74. By observing the charts, it is clear that the Case Study traffic flows are significantly different than the average day flows. It is also

evident that the algorithm is quite good at predicting future volumes (in this case, from 2:00PM to 9:00PM) using the volumes from earlier in the day.

The simulation was used to determine the potential impact of the algorithm on the traffic flow within the network. The simulation was first run with the actual flows for each Case Study from 2:00PM to 9:00 PM using the existing signal timings, as found in MaxTime. This simulation represents the scenario where atypical flows occur within the network, and the algorithm is not used. In this scenario, no signal timing changes are made. The simulation was also run with the actual flows for each Case Study from 2:00PM to 9:00 PM using updated, optimized signal timings, as found using Synchro. This simulation represents the scenario where the algorithm notices atypical flows within the network and predicts future traffic volumes. In this scenario, signal timings are changed appropriately by traffic engineers so that they are optimal for the predicted flows. For both of these simulations, SUMO's edgeData function was used to measure the delay within the network. Table 12 contains the total and average delay for the unoptimized (control) and optimized signal timings, as well as the difference between the two. Use of the algorithm can save approximately 70 to 350 hours hours of delay, amounting to approximately \$1,000 to \$5,000 per day in travel time savings at the four intersections in the network when flow is atypical.

Table 12. Total Delay, Average Delay, and Travel Time Savings for each Case Study.

Case Study	Control		Optimized		Optimized - Control		
	Total Delay (hr)	Average Delay (s)	Total Delay (hr)	Average Delay (s)	Δ Total Delay (hr)	Δ Average Delay (s)	Travel Time Savings (per Day)*
1	4266	215.66	3983	201.37	283	14.29	\$3,985.81
2	3022	179.51	2677	159.02	345	20.49	\$4,863.36
3	3668	204.18	3594	200.04	75	4.15	\$1,050.72
4	4335	241.74	4258	237.44	77	4.29	\$1,085.59

*Hourly Travel Time Saving Rate \$14.10 per the 2016 U.S. Department of Transportation Revised Value of Travel Time Guidance

INTERACTIVE WEB ENGINE FOR PREDICTING TRAFFIC FLOW AND LEVEL OF SERVICE

OVERVIEW OF THE TOOL

This project involved the optimization of traffic flow at diverging diamond interchanges with ramp meters and neighboring intersections. The goal was to demonstrate the feasibility of a decision support tool that could notice unusual trends in the flow and notify GDOT that signal timing adjustments may be needed now or in the future; for example, an unusual AM peak period traffic profile may indicate to expect an unusual PM peak period profile. The utility of human–algorithm collaboration in traffic engineering is paramount. In an effort to close the gap between humans and computers in this collaboration, in Task V, the research team devised a prototype web engine that integrates the productivity of algorithmic predictions with human oversight and decision-making. The functionality of the web engine is as follows:

1. The engine reads incoming traffic data from a network of intelligent infrastructure.
2. Then, data-driven algorithms provide predictions of future traffic flows/level of service across the network.
3. Lastly, the results of these predictions are displayed in a graphical user interface for use by traffic engineers. Suggestions for updated timing schemes and alerts are provided in this interface if necessary. Figure shows a screen capture of the web engine that has been developed for GDOT. The user interface presents predictions for each movement at each intersection and the corresponding level of service over the prediction horizon. As a prototype, this web engine receives traffic data regarding the DDIs at Ashford Dunwoody Rd. and Jimmy Carter Blvd. from the ATSPM website. This, of course, could be expanded in the future if real-time communication were viable between traffic

infrastructure and the web engine itself. An example of how the portal will appear to users is depicted in figure. Measured traffic flow over the course of the day is reported with solid lines for various turning movements. Dotted lines represent this algorithms' prediction for traffic flow for the remaining portion of the day. The engine itself will be hosted in the cloud using computing resources allocated by the Georgia Institute of Technology and will be accessible via a web browser to GDOT. Additionally, the engine is being constructed in a modular framework in order to optimize the replacement of current data-driven algorithms with future generations. The research team believes that this prototype will set the standard for future human–algorithm collaboration in traffic engineering.

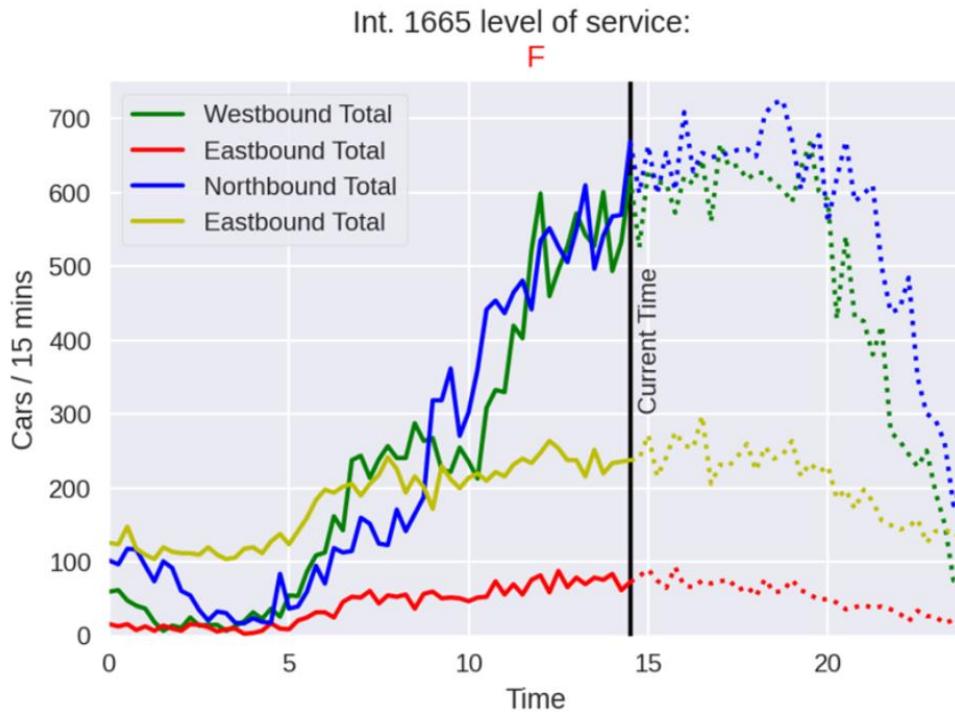


Figure 75. Screen capture. Sample from the web engine.

ACKNOWLEDGMENTS

The authors of this report wish to thank the Georgia Department of Transportation for its support and assistance throughout this effort, in particular the efforts of Benjamin Lempke and Sunil Thapa.

REFERENCES

- [1] Schroeder, B., Cunningham, C., Ray, B., Daleiden, A., et al. (2014). *Diverging Diamond Interchange Informational Guide*. Research Report 959, National Cooperative Highway Research Program, Transportation Research Board, Washington, DC.
- [2] Chilukuri, V., Siromaskul, S., Trueblood, M., Ryan, T., et al. (2011). Diverging Diamond Interchange Performance Evaluation (I-44 and Route 13). Research Report OR11.012, Organizational Results, Missouri Department of Transportation, Jefferson City, MO.
- [3] Missouri Department of Transportation. (2010). *Missouri's Experience with a Diverging Diamond Interchange: Lessons Learned*. Research Report OR10-021, Missouri Department of Transportation, Jefferson City, MO.
- [4] Hunter, M., Guin, A., Anderson, J., and Park, S.J. (2019). “Operating Performance of Diverging Diamond Interchanges.” *Transportation Research Record, Journal of the Transportation Research Board*, 0361198119855341.
- [5] Pang, B., Liu, H., and Xu, H. (2016). “Calculating Approach Capacity of Diverging Diamond Interchanges with Consideration of Internal Queue Effects.” *Transportation Research Record, Journal of the Transportation Research Board*, 2553, pp. 63–71.
- [6] National Research Council. (2010). *Highway Capacity Manual*. 5th edition, Transportation Research Board, Washington, DC.
- [7] Khan, T. and Anderson, M. (2016). “Evaluating the Application of Diverging Diamond Interchange in Athens, Alabama.” *International Journal of Traffic and Transportation Engineering*, 6, pp. 38–50.
- [8] Warchol, S., Chase, T., and Cunningham, C. (2017). “Use of Microsimulation to Evaluate Signal-phasing Schemes at Diverging Diamond Interchanges.” *Transportation Research Record, Journal of the Transportation Research Board*, 2620, pp. 10–19.
- [9] Hainen, A.M., Stevens, A.L., Day, C.M., Li, H., Mackey, J., Luker, M., Taylor, M., Sturdevant, J.R., and Bullock, D.M. (2015). “High-resolution Controller Data Performance Measures for Optimizing Diverging Diamond Interchanges and Outcome Assessment with Drone Video.” *Transportation Research Record, Journal of the Transportation Research Board*, 2487, pp. 31–43.
- [10] Leong, L.V., Mahdi, M.B., and Chin, K.K. (2015). “Microscopic Simulation on the Design and Operational Performance of Diverging Diamond Interchange.” *Transportation Research Procedia*, 6, pp. 198–212.

- [11] Bared, J.G., Edara, P.K., and Jagannathan, R. (2005). “Design and Operational Performance of Double Crossover Intersection and Diverging Diamond Interchange.” *Transportation Research Record, Journal of the Transportation Research Board*, 1912(1), pp. 31–38. Available online: <https://www.researchgate.net/deref/http%3A%2F%2Fdx.doi.org%2F10.1177%2F0361198105191200104>.
- [12] Chlewicki, G. (2011). “Should the Diverging Diamond Interchange Always be Considered a Diamond Interchange Form?” *Transportation Research Record, Journal of the Transportation Research Board*, 2223, pp. 88–95. Available online: <https://doi.org/10.3141%2F2223-11>.
- [13] Edara, P.K., Bared, J.G., and Jagannathan, R. (2005). “Diverging Diamond Interchange and Double Crossover Intersection—Vehicle and Pedestrian Performance.” Presented at 3rd International Symposium on Highway Geometric Design, Chicago, IL.
- [14] Galletebeitia, B. (2011). *Comparative Analysis Between the Diverging Diamond Interchange and Partial Cloverleaf Interchange Using Microsimulation Modeling*. Master’s thesis, Florida Atlantic University, Boca Raton, FL.
- [15] Sharma, S. and Chatterjee, I. (2007). “Performance Evaluation of the Diverging Diamond Interchange in Comparison with the Conventional Diamond Interchange.” Presented at Transportation Scholars Conference, Iowa State University, Ames, IA.
- [16] Siromaskul, S. and Speth, S.B. (2008). “A Comparative Analysis of Diverging Diamond Interchange Operations.” Presented at ITE 2008 Annual Meeting and Exhibit, Institute of Transportation Engineers (ITE), August 17–20, Anaheim, CA.
- [17] Lin, P.W. and Cogan, C. (2008). “The Design and Operations of Diverging Diamond Interchanges—A Case Study in Kansas City, MO.” Presented at 15th World Congress on Intelligent Transport Systems and ITS America’s 2008 Annual Meeting, ITS America, ERTICOITS JapanTransCore.
- [18] Ressel, N.R. (2012). *Insights into the First Three Diverging Diamond Interchanges in Missouri*. Ph.D. thesis, University of Missouri–Columbia.
- [19] Schroeder, B.J., Salamati, K., and Hummer, J. (2014). “Calibration and Field Validation of Four Double-crossover Diamond Interchanges in VISSIM Microsimulation.” *Transportation Research Record, Journal of the Transportation Research Board*, 2404, pp. 49–58.
- [20] Hughes, W., Jagannathan, R., Sengupta, D., Hummer, J., et al. (2010). *Alternative Intersections/Interchanges: Informational Report (AIIR)*. Technical Report FHWA-HRT-09-060, Federal Highway Administration, Office of Safety Research and Development.
- [21] Stanek, D. (2007). “Innovative Diamond Interchange Designs: How to Increase Capacity and Minimize Cost.” Presented at ITE 2007 Annual Meeting and Exhibit Institute of Transportation Engineers (ITE), August 5–8, Pittsburgh, PA.

- [22] Almoshaogeh, M. (2017). *Developing Warrants for Designing Continuous Flow Intersection and Diverging Diamond Interchange*. Ph.D. dissertation, University of Central Florida, Orlando, FL.
- [23] Autey, J., Sayed, T., and El Esawey, M. (2013). “Operational Performance Comparison of Four Unconventional Intersection Designs Using Micro-simulation.” *Journal of Advanced Transportation*, 47(5), pp. 536–552. Available online: <https://doi.org/10.1155/2017/7396250>.
- [24] Jagannathan, R. and Bared, J.G. (2004). “Design and Operational Performance of Crossover Displaced Left-turn Intersections.” *Transportation Research Record, Journal of the Transportation Research Board*, 1881, pp. 1–10.
- [25] Sayed, T., Storer, P., and Wong, G. (2006). “Upstream Signalized Crossover Intersection: Optimization and Performance Issues.” *Transportation Research Record, Journal of the Transportation Research Board*, 1961, pp. 44–54.
- [26] National Academies of Sciences, Engineering, and Medicine. (2015). *Signal Timing Manual*. 2nd edition, The National Academies Press, Washington, DC. Available online: <https://doi.org/10.17226/22097>.
- [27] Cunningham, C., Schroeder, B.J., Phillips, S., Urbanik, T., Warchol, S., and Tanaka, A. (2016). “Signal Timing for Diverging Diamond Interchanges: Fundamentals, Concepts, and Recommended Applications.” *Transportation Research Record, Journal of the Transportation Research Board*, 2557, pp. 1–10. Available online: <https://doi.org/10.3141%2F2557-01>.
- [28] Day, C.M. and Bullock, D.M. (2016). “Cycle-length Strategies for a Diverging Diamond Interchange in a Coordinated Arterial.” *Journal of Transportation Engineering*, 142, 04016067.
- [29] Day, C.M., Lavrenz, S.M., Stevens, A.L., Miller, R.E., and Bullock, D.M. (2016). “Extending Link Pivot Offset Optimization to Arterials with Single Controller Diverging Diamond Interchange.” In *TRB 95th Annual Meeting Compendium of Papers*, Paper No. 16-0111, Transportation Research Board, Washington, DC.
- [30] Cheng, Y., Chang, G.L., and Rahwanji, S. (2018). “Concurrent Optimization of Signal Progression and Crossover Spacing for Diverging Diamond Interchanges.” *Journal of Transportation Engineering, Part A: Systems*, 144(3), 04018001.
- [31] Yang, X., Chang, G.L., and Rahwanji, S. (2014). “Development of a Signal Optimization Model for Diverging Diamond Interchange.” *Journal of Transportation Engineering*, 140(5), 04014010. Available online: <https://ascelibrary.org/doi/10.1061/%28ASCE%29TE.1943-5436.0000657>.
- [32] Hu, P., Tian, Z.Z., Wu, X., and Xu, H. (2014). “A Proposed Signal Operation and the Effect of ITS Cycle Length on Diverging Diamond Interchanges.” *ITE Journal*, 84(11), pp. 28–35.

- [33] Xu, H., Liu, H., Tian, Z., and Zhang, W. (2011). “Control Delay Calculation at Diverging Diamond Interchanges.” *Transportation Research Record, Journal of the Transportation Research Board*, 2257, pp. 121–130.
- [34] Hu, P., Tian, Z.Z., Xu, H., and Andalibian, R. (2014). “An Advanced Signal Phasing Scheme for Diverging Diamond Interchanges.” *TRB 93rd Annual Meeting Compendium of Papers*, Transportation Research Board, pp. 16–32.
- [35] Claros, B.R., Edara, P., Sun, C., and Brown, H. (2015). “Safety Evaluation of Diverging Diamond Interchanges in Missouri.” *Transportation Research Record, Journal of the Transportation Research Board*, 2486, pp. 1–10. Available online: <https://doi.org/10.3141%2F2486-01>.
- [36] Nye, T.S., Cunningham, C.M., and Byrom, E. (2019). “National-level Safety Evaluation of Diverging Diamond Interchanges.” *Transportation Research Record, Journal of the Transportation Research Board*, 0361198119849589.
- [37] Wang, R., Hu, J., and Zhang, X. (2016). “Analysis of the Driver’s Behavior Characteristics in Low Volume Freeway Interchange.” *Mathematical Problems in Engineering*, 2016(1208), pp. 1–9. Available online: <https://doi.org/10.1155/2016/2679516>.
- [38] Vaughan, C., Jagadish, C., Bharadwaj, S., Cunningham, C.M., Schroeder, B.J., Hummer, J.E., Findley, D., and Rouphail, N.M. (2015). “Long-term Monitoring of Wrong-way Maneuvers at Diverging Diamond Interchanges.” *Transportation Research Record, Journal of the Transportation Research Board*, 2484, pp. 129–139.
- [39] Claros, B., Edara, P., and Sun, C. (2017). “When Driving on the Left Side is Safe: Safety of the Diverging Diamond Interchange Ramp Terminals.” *Accident Analysis and Prevention*, 100, pp. 133–142. Available online: <https://doi.org/10.1016/j.aap.2017.01.014>.
- [40] Burbidge, S.K. and Planning, A. (2012). *Identifying Characteristics of High Risk Intersections for Pedestrians and Cyclists: A Case Study from Salt Lake County*. Technical Report UT-11.305, 12-0947, Utah Department of Transportation, Research Division, Salt Lake City, UT.
- [41] El Esawey, M. and Sayed, T. “Analysis of Unconventional Arterial Intersection Designs (UAIDS): State-of-the-Art Methodologies and Future Research Directions.” *Transportmetrica A: Transport Science*, 9(10), pp. 860–895.
- [42] Tarko, A.P., Inerowicz, M., and Lang, B. (2008). “Safety and Operational Impacts of Alternative Intersections.”
- [43] Maji, A., Mishra, S., and Jha, M.K. (2013). “Diverging Diamond Interchange Analysis: Planning Tool.” *Journal of Transportation Engineering*, 139, pp. 1201–1210.
- [44] Anderson, M., Schroer, B., and Moeller, D. (2012). “Analyzing the Diverging Diamond Interchange Using Discrete Event Simulation.” *Modelling and Simulation in Engineering*, 2012(47), Article ID 639865. Available online: <https://doi.org/10.1155/2012/639865>.

- [45] Duan, X. and Abbas, M. (2016). “Macroscopic Calculation of Delay for Diverging Diamond Interchanges.” *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 589–594.
- [46] Yeom, C., Schroeder, B.J., Cunningham, C., Salamati, K., and Roushail, N.M. (2017). “Lane Utilization Model Development for Diverging Diamond Interchanges.” *Transportation Research Record, Journal of the Transportation Research Board*, 2618, pp. 27–37.
- [47] Nagendra, S.S. and Khare, M. (2003). “Principal Component Analysis of Urban Traffic Characteristics and Meteorological Data.” *Transportation Research Part D: Transport and Environment*, 8, pp. 285–297.
- [48] Wold, S., Esbensen, K., and Geladi, P. (1987). “Principal Component Analysis.” *Chemometrics Intelligent Laboratory Systems*, 2, pp. 37–52.
- [49] Abdi, H. (2010). “Partial Least Squares Regression and Projection on Latent Structure Regression (PLS Regression).” *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(1), pp. 97–106. Available online: <https://doi.org/10.1002/wics.51>.
- [50] Dauwels, J., Aslam, A., Asif, M.T., Zhao, X., Vie, N.M., Cichocki, A., and Jaillet, P. (2014). “Predicting Traffic Speed in Urban Transportation Subnetworks for Multiple Horizons.” *13th International Conference on Control Automation Robotics & Vision (ICARCV)*, pp. 547–552. Available online: <https://doi.org/10.1109/ICARCV.2014.7064363>.
- [51] Wang, W., Chen, S., and Qu, G. (2008). “Incident Detection Algorithm Based on Partial Least Squares Regression.” *Transportation Research Part C: Emerging Technologies*, 16, pp. 54–70.
- [52] Xing, X., Zhou, X., Hong, H., Huang, W., Bian, K., and Xie, K. (2015). “Traffic Flow Decomposition and Prediction Based on Robust Principal Component Analysis.” *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 2015, pp. 2219–2224.
- [53] Guo, R. and Zhang, Y. (2014). “Identifying Time-of-day Breakpoints Based on Nonintrusive Data Collection Platforms.” *Journal of Intelligent Transportation Systems*, 18(2), pp. 164–174.
- [54] Chen, P., Zheng, N., Sun, W., and Wang, Y. (2019). “Fine-tuning Time-of-day Partitions for Signal Timing Plan Development: Revisiting Clustering Approaches.” *Transportmetrica A: Transport Science*, 15(2), pp. 1195–1213.
- [55] Bellman, R. (1952). “On the Theory of Dynamic Programming.” *Proceedings of the National Academy of Sciences of the United States America*, 38(8), pp. 716–719.

- [56] Auger, I.E. and Lawrence, C.E. (1989). “Algorithms for the Optimal Identification of Segment Neighborhoods.” *Bulletin of Mathematical Biology*, 51, pp. 39–54. Available online: <https://doi.org/10.1007/bf02458835>.
- [57] Kehagias, A., Nidelkou, E., and Petridis, V. (2006). “A Dynamic Programming Segmentation Procedure for Hydrological and Environmental Time Series.” *Stochastic Environmental Research and Risk Assessment*, 20, pp. 77–94.
- [58] Coogan, S., Flores, C., and Varaiya, P. (2017). “Traffic Predictive Control from Low-rank Structure.” *Transportation Research Part B: Methodological*, 97, pp. 1–22.
- [59] Ling, E., Zheng, L., Ratliff, L.J., and Coogan, S. (2019). “Koopman Operator Applications in Signalized Traffic Systems.” arXiv preprint arXiv:1902.11289.
- [60] De Oliveira, L.B. and Camponogara, E. (2010). “Multi-agent Model Predictive Control of Signaling Split in Urban Traffic Networks.” *Transportation Research Part C: Emerging Technologies*, 18(1), pp. 120–139.
- [61] Lin, S., De Schutter, B., Xi, Y., and Hellendoorn, H. (2012). “Efficient Network-Wide Model-based Predictive Control for Urban Traffic Networks.” *Transportation Research Part C: Emerging Technologies*, 24, pp. 122–140.
- [62] Avila, A. and Mezic, I. (2020). “Data-driven Analysis and Forecasting of Highway Traffic Dynamics.” *Nature Communications*, 11, pp. 1–16.
- [63] Brunton, S. and Kutz, J. (2019). “Singular Value Decomposition (SVD).” Chapter 1 in *Data-driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*, Cambridge University Press, New York, pp. 3–46.
- [64] Strang, G. (2019). “Highlights of Linear Algebra.” Chapter in *Linear Algebra and Learning from Data*, Cambridge Press, Wellesley, MA, pp. 62–63.
- [65] Ke, R., Li, W., Cui, Z., and Wang, Y. (2020). “Two-stream Multi-channel Convolutional Neural Network for Multi-lane Traffic Speed Prediction Considering Traffic Volume Impact.” *Transportation Research Record, Journal of the Transportation Research Board*, 2674, pp. 459–470.
- [66] Zhang, S., Abdel-Aty, M., Yuan, J., and Li, P. (2020). “Prediction of Pedestrian Crossing Intentions at Intersections Based on Long Short-term Memory Recurrent Neural Network.” *Transportation Research Record, Journal of the Transportation Research Board*, 0361198120912422.
- [67] Li, Y., Yu, R., Shahabi, C., and Liu, Y. (2017). “Diffusion Convolutional Recurrent Neural Network: Data-driven Traffic Forecasting.” arXiv preprint arXiv:1707.01926.
- [68] Lu, H.P., Sun, Z.Y., and Qu, W.C. (2015). “Big Data-driven Based Real-time Traffic Flowstate Identification and Prediction.” *Discrete Dynamics in Nature and Society*, 2015, Article ID 284906, pp. 1–11.

- [69] Lippi, M., Bertini, M., and Frasconi, P. (2013). “Short-term Traffic Flow Forecasting: An Experimental Comparison of Time-series Analysis and Supervised Learning.” *IEEE Transactions on Intelligent Transportation Systems*, 14, pp. 871–882.
- [70] Candès, E.J., Li, X., Ma, Y., and Wright, J. (2011). “Robust Principal Component Analysis?” *Journal of the ACM* (JACM), 58(3), pp. 1–37. Available online: <https://doi.org/10.1145/1970392.1970395>.
- [71] Jong, S. (1993). “SIMPLS: An Alternative Approach to Partial Least Squares Regression.” *Chemometrics and Intelligent Laboratory Systems*, 18, pp. 251–263. Available online: [https://doi.org/10.1016/0169-7439\(93\)85002-X](https://doi.org/10.1016/0169-7439(93)85002-X).
- [72] Horn, R.A. and Johnson, C.R. (1994). “The Polar and Singular Value Decompositions,” In *Topics in Matrix Analysis*, Cambridge University Press, New York, pp. 448–452.
- [73] Candès, E.J. and Recht, B. (2009). “Exact Matrix Completion via Convex Optimization.” *Foundations of Computational Mathematics*, 9(6), pp. 717–772.
- [74] Lin, Z., Chen, M., and Ma, Y. (2010), “The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices.” arXiv:1009.5055.
- [75] Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. (2017). “Google Earth Engine: Planetary-scale Geospatial Analysis for Everyone.” *Remote Sensing Environment*, 202, pp. 18–27. Available online: <https://www.researchgate.net/deref/http%3A%2F%2Fd.x.doi.org%2F10.1016%2Fj.rse.2017.06.031>.
- [76] Kingma, D.P. and Ba, J. (2014). ”ADAM: A Method for Stochastic Optimization.” arXiv preprint arXiv:1412.6980 (2014).
- [77] Glorot, X. and Bengio, Y. (2010). “Understanding the Difficulty of Training Deep Feedforward Neural Networks.” *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics*, JMLR Workshop and Conference Proceedings, 9, pp. 249–256.