Evaluation of Safety and Mobility Benefits of Connected and Automated Vehicles by Considering V2X Technologies

Md Hasibur Rahman
University of Central Florida
EVALUATION OF SAFETY AND MOBILITY BENEFITS OF CONNECTED AND AUTOMATED VEHICLES BY CONSIDERING V2X TECHNOLOGIES

by

MD HASIBUR RAHMAN
B.Sc., Bangladesh University of Engineering and Technology, Bangladesh, 2015
M.Sc. University of Central Florida, USA, 2019

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Civil, Environmental and Construction Engineering in the College of Engineering and Computer Science at the University of Central Florida Orlando, Florida

Summer Term 2020

Major Professor: Mohamed Abdel-Aty
ABSTRACT

The recent development in communication technologies facilitates the deployment of connected and automated vehicles (CAV) which are expected to change the future transportation system. CAV technologies enable vehicles to communicate with other vehicles through vehicle-to-vehicle (V2V) communications and the infrastructure through Vehicle-to-infrastructure (V2I) communications. Since the real-world CAV data is not currently available as of today, simulation is the most commonly used platform to evaluate the future V2X system. Although several studies evaluated the effectiveness of CAVs in a small roadway network, there is a lack of studies analyzing the impact of CAVs at the network level by considering both freeways and arterials. Also, none of the previous studies have attempted to differentiate the benefits of CAVs over only automated vehicles (AVs) by incorporating multiple preceding vehicles’ information (i.e., acceleration, position, etc.). On the other hand, most of the simulation-based studies assumed the uninterrupted communication between vehicles in the CAV environment which might not be feasible in reality. Hence, there is still a research gap that exists for which this study tried to fill this gap. Therefore, this study developed a calibrated and validated large-scale network for the deployment of CAV technologies by utilizing Dynamic Traffic Assignment (DTA) model in Orlando metropolitan area, Florida, using Multi-Resolution Modeling (MRM) technique. Also, the study proposed a signal control algorithm through V2I technology in order to elevate the performance of CAVs at intersections. Different car-following models were utilized to approximate different CAV technologies (CAV, AV, and CV (connected vehicle)) in the simulation environment. Hence, the study analyzed the benefits of CAV over AV with different market penetration rates (MPRs). Furthermore, the study considered the performance of different communication system along with the traffic condition by utilizing Dedicated Short-Range
Communications (DSRC or IEEE 802.11p) and wireless access (IEEE 1609 protocol) for the application of vehicle ad-hoc network (VANET). To this end, the study evaluated the safety effectiveness of different communication protocols under the CAV environment. Aimsun Next and SUMO & OMNET++ based Veins simulator were used as the simulation platform. Different car-following models, signal control algorithm, and communication systems were coded by using the application programming interface (API) and C++ language. For the traffic efficiency, the study utilized travel time and travel time rate (TTR) while for the safety evaluation, different surrogate safety measures; speed, and crash-risk models were used. Also, several statistical tests (e.g., t-test, ANOVA) and modeling techniques (e.g., generalized estimating equation, logistic regression, etc.) were developed to analyze both safety and mobility. The results of this study implied that CAV could improve both safety and efficiency at the network level with different MPRs. Also, CAV is more efficient compared to the only AV in terms of both traffic safety and mobility. Different communication protocols have a significant effect on traffic safety under the CAV environment. Finally, the results of this study provide insight to transportation planners and the decision makers about the benefits of CAV at the network level, different CAV technologies, and the performance of different communication systems under the CAV environment.
ACKNOWLEDGMENT

I like to convey my heartiest gratitude to my honorable supervisor Professor Dr. Mohamed Abdel-Aty, for his invaluable guidance, advice, support, motivation, enthusiasm, and encouragement toward successful completion of the doctorate degree. His guidance helped me in all the time of research and writing of this dissertation. I would also like to thank my respectable committee members, Dr. Mohamed Zaki Hussein, Dr. Yina Wu, and Dr. Xin Yan for their valuable suggestions for my Ph.D. work.

Meanwhile, I also want to acknowledge Dr. Jaeyoung Lee, and Dr. Qing Cai for providing valuable suggestions for my research work throughout this journey. Additionally, I am thankful to all my lab mates for helping me towards the successful completion of my Ph.D. degree.

I would also like to thank my beloved parents (Father: Md Shahidur Rahman, Mother: Afroza Akter), my Aunt (Naznin Akter), and my sister (Ruma Akter) who always encourage me a lot.

Last but not the least, many thanks to my beloved wife Hasina Akther Shima for always supporting me throughout my Ph.D. work. Without her support and encouragement, I cannot achieve anything.
# TABLE OF CONTENT

LIST OF FIGURES ........................................................................................................................................ x
LIST OF TABLES ......................................................................................................................................... xii

CHAPTER 1: INTRODUCTION ...................................................................................................................... 1
  1.1 Background .......................................................................................................................................... 1
  1.2 Objectives of the Research ............................................................................................................... 3
  1.3 Dissertation Structure ....................................................................................................................... 7

CHAPTER 2: LITERATURE REVIEW ........................................................................................................... 10

CHAPTER 3: CALIBRATION OF DYNAMIC TRAFFIC ASSIGNMENT MODEL FOR A LARGE-SCALE NETWORK BY USING MULTI-RESOLUTION MODELING TECHNIQUE FOR THE DEPLOYMENT OF CAV TECHNOLOGIES ......................................................... 23
  3.1 Introduction ......................................................................................................................................... 23
  3.2 Model Overview: Location and Data ................................................................................................. 25
    3.2.1 Study area .................................................................................................................................... 25
    3.2.2 Input data for model development ............................................................................................ 28
  3.3 Modeling Method ............................................................................................................................... 29
  3.4 DTA Model Development .................................................................................................................. 33
    3.4.1 Origin-destination matrix estimation (ODME) ......................................................................... 34
    3.4.2 Time dependent origin-destination demand ............................................................................. 40
    3.4.3 Calibration of mesoscopic simulation area ............................................................................... 42
    3.4.4 Validation of mesoscopic simulation area ................................................................................ 48
    3.4.5 Calibration and validation of microsimulation area ................................................................. 55
  3.5 Conclusions ........................................................................................................................................ 60
5.3.2 Modeling Automated vehicles (AVs) .......................................................... 103

5.3.3 Modeling Connected vehicles (CVs) and Connected & Automated vehicles (CAVs) .............................................................................................................. 105

5.4 Performance Evaluation Measures ................................................................. 110

5.4.1 Traffic Efficiency .......................................................................................... 110

5.4.2 Traffic Safety .................................................................................................. 112

5.5 Results and Discussion .................................................................................... 115

5.5.1 Traffic efficiency .......................................................................................... 118

5.5.2 Traffic Safety .................................................................................................. 125

5.6 Conclusions ...................................................................................................... 133

CHAPTER 6: EVALUATION OF TRAFFIC SAFETY BASED ON DSRC COMMUNICATION PARAMETERS UNDER CONNECTED AND AUTOMATED VEHICLES’ ENVIRONMENT ........................................................................................................ 137

6.1 Introduction ....................................................................................................... 137

6.2 Simulation setup and data ................................................................................. 139

6.2.1 Calibration and Validation ........................................................................... 140

6.3 Proposed methodology ..................................................................................... 144

6.3.1 Car-following model .................................................................................... 144

6.3.2 Communication protocol .............................................................................. 145

6.4 Evaluation measures ......................................................................................... 149

6.4.1 Time-to-collision (TTC) .............................................................................. 150
6.5 Results and discussion ................................................................. 151

6.5.1 Communication aspect......................................................... 152

6.5.2 Transportation aspect.......................................................... 156

6.5.3 Binary Logistic Regression Model ......................................... 162

6.6 Conclusions.............................................................................. 164

CHAPTER 7: CONCLUSIONS .............................................................. 168

7.1 Summary.................................................................................. 168

7.2 Implications............................................................................. 173

REFERENCES .................................................................................. 175
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Examples of MRM structure <em>(source: Cambridge Systematics, Inc. (FHWA, 2012))</em></td>
<td>11</td>
</tr>
<tr>
<td>3.1</td>
<td>Study area (a) full network OUATS regional model (b) mesoscopic area (Orange and Seminole County) (c) microscopic area</td>
<td>27</td>
</tr>
<tr>
<td>3.2</td>
<td>Modeling methods for the DTA based simulation model</td>
<td>31</td>
</tr>
<tr>
<td>3.3</td>
<td>Origin-destination matrix estimation results of OUATS (macroscopic) model</td>
<td>37</td>
</tr>
<tr>
<td>3.4</td>
<td>Origin-destination matrix estimation results of mesoscopic area model</td>
<td>39</td>
</tr>
<tr>
<td>3.5</td>
<td>Adjustment and observed link counts (every 15 minutes) for the Mesoscopic area</td>
<td>40</td>
</tr>
<tr>
<td>3.6</td>
<td>Profiled demand from OD adjustment procedure</td>
<td>41</td>
</tr>
<tr>
<td>3.7</td>
<td>Traffic count detectors locations in the mesoscopic area</td>
<td>43</td>
</tr>
<tr>
<td>3.8</td>
<td>Mesoscopic model calibration results</td>
<td>48</td>
</tr>
<tr>
<td>3.9</td>
<td>Travel time detectors locations in the mesoscopic area</td>
<td>50</td>
</tr>
<tr>
<td>3.10</td>
<td>Travel time detectors locations and validation results for the entire mesoscopic area.</td>
<td>53</td>
</tr>
<tr>
<td>3.11</td>
<td>Simulated and real-field travel time plot for Mesoscopic simulation</td>
<td>55</td>
</tr>
<tr>
<td>3.12</td>
<td>Volume and travel time detectors locations</td>
<td>57</td>
</tr>
<tr>
<td>4.1</td>
<td>Microsimulation study area</td>
<td>64</td>
</tr>
<tr>
<td>4.2</td>
<td>Illustration of vehicle-to-vehicle (a), vehicle-to-infrastructure (b), and signal control algorithm (c)</td>
<td>73</td>
</tr>
<tr>
<td>4.3</td>
<td>Travel time rate (TTR) for the (a) full network, (b) freeways, (c) arterials and (d) I-4 eastbound</td>
<td>82</td>
</tr>
<tr>
<td>4.4</td>
<td>Approach delay at intersections (base vs CAV100)</td>
<td>84</td>
</tr>
<tr>
<td>4.5</td>
<td>Standard deviation of speed on I-4 eastbound</td>
<td>86</td>
</tr>
</tbody>
</table>
Figure 4.6 Crash risk on freeway (I-4 westbound) ................................................................. 88
Figure 4.7 Standard deviation of speed on Arterials ............................................................... 89
Figure 4.8 Crash risk on arterials (full network) ................................................................. 90
Figure 5.1 Study Area .................................................................................................................. 98
Figure 5.2 Illustration of AV, CV, and CAV communication system ...................................... 109
Figure 5.3 Experimental design ................................................................................................. 116
Figure 5.4 Travel time of AV, CAV, and mixture of CV and CAV scenarios ......................... 120
Figure 5.5 Traffic conflicts based on different vehicle type ...................................................... 132
Figure 6.1 Study Area and detectors’ location ............................................................................. 140
Figure 6.2 Traffic Volume for Different Detectors ....................................................................... 141
Figure 6.3 Flow chart of the communication between OMNET++, SUMO, and Veins ............. 146
Figure 6.4 Effect of Queue size on packet loss ratio and traffic conflicts ................................. 153
Figure 6.5 Effect of Transmission power on packet loss and traffic conflicts ......................... 155
Figure 6.6 Effect of lane closure on traffic conflict and communication collisions .................. 157
Figure 6.7 Effect of traffic flow on traffic conflicts and communication collisions ............... 161
LIST OF TABLES

Table 2.1 Summary of previous Studies on connected and automated vehicles .......................... 17
Table 2.2 Detailed summary of recent studies ........................................................................... 21
Table 3.1 Network Attributes of OUATS model including mesoscopic and microscopic areas . 26
Table 3.2 Summary of Data Used for Simulation ....................................................................... 28
Table 3.3 Aimsun Next Calibration Parameters for Mesoscopic Simulation ............................. 52
Table 3.4 Aimsun Next Calibration Parameters for Microscopic Simulation Areas ..................... 58
Table 4.1 Aimsun Next Calibration Parameters for Microscopic Simulation Areas ..................... 67
Table 5.1 Calibration parameters in Aimsun Next ....................................................................... 101
Table 5.2 AVs model parameter settings .................................................................................... 105
Table 5.3 Comparison of CAV and AV for the same MPR .......................................................... 121
Table 5.4 GEE model estimation results ...................................................................................... 122
Table 5.5 Summary statistics of traffic conflicts (TTC) with different MPRs ............................... 126
Table 5.6 Comparison of traffic conflicts (TTC) for difference scenarios ................................. 127
Table 5.7 Sensitivity analysis of different scenarios .................................................................... 128
Table 5.8 Bayesian ZINB model estimation results ................................................................... 130
Table 6.1 Parameter settings in Sumo for calibration ................................................................. 143
Table 6.2 Communication parameters of IEEE 802.11p .......................................................... 147
Table 6.3 Message processing algorithm for CAV ...................................................................... 149
Table 6.4 Model estimation results for different scenarios .......................................................... 163
CHAPTER 1: INTRODUCTION

1.1 Background

With the advent of new technology, the road transportation system is changing rapidly. One of the major inventions in the transportation system is the connected and automated vehicles (CAV) which might have the potential for reducing traffic accidents and improving the traffic efficiency of the transportation system through real-time information about the surrounding traffic conditions e.g., position, speed, and acceleration, etc. Through these technologies, a vehicle can communicate with other vehicles through vehicle-to-vehicle (V2V) communications and infrastructures though Vehicle-to-infrastructure (V2I) communications, both referred to as V2X. It is known that traffic crashes are directly or indirectly related to human errors. According to the National Highway Traffic Safety Administration (NHTSA), most of the traffic crashes are caused by human driver error (Singh, 2015; Yue et al., 2018) which could be avoided with the application of CAV technologies (Fagnant and Kockelman, 2015). Therefore, it is worthwhile to evaluate the impact of CAV applications to enhance traffic safety and mobility.

Since the CAV technologies are still not fully developed and the real-world CAV data is not available as of today, simulation is one of the mostly used platforms to analyze the effectiveness of CAV. Also, to quantify the benefits of CAVs using real-world data would take a long time because millions of miles of real-world CAV operational data are needed. Hence, the microsimulation environment is the most used and feasible platform in this field. Also, several previous studies (e.g., Guériau et al., 2016; Talebpour and Mahmassani, 2016) evaluated the benefits of CAV either for a small freeway or arterial segment under limited traffic conditions. However, none of them analyzed the benefits of CAV at the network level considering both freeways and arterials. Hence, this study developed a dynamic traffic assignment (DTA) based
large scale microscopic model by considering multi-resolution modeling (MRM) technique which combines microscopic, mesoscopic, and macroscopic representations of traffic flow in the modeling effort of a single study. Additionally, the application of CAVs in the arterial segments is more challenging in order to maintain the connected and automated vehicles string especially at intersections. Hence, this study implemented a signal control algorithm and adjusted the signal phase and timing (SPaT) plan which could improve the overall performance of the intersection as well as the whole network.

Although recent studies (Milanés and Shladover, 2014; Shladover et al., 2012; Van Arem et al., 2006; Watzenig and Horn, 2017) evaluated the benefits of CAV under different traffic conditions, none of them have attempted to differentiate the benefits of CAVs over automated vehicles (AVs) by incorporating multiple preceding vehicle information (i.e., acceleration, position, etc.). Hence, this study incorporated different car-following models to approximate different CAV technologies (CAV, AV, etc.) in the simulation platform. Additionally, a different car-following model was used for the connected vehicles (CVs) without automation by addressing the human driver compliance factor. This study also utilizes mixed penetration of CAV and CV with no automation. Since 100% market penetration rate (MPRs) of CAV will not be feasible in the near future, there will be a long period of mixed traffic flow where human driven vehicles and CAVs move together. Thus, the dissertation also considered different MPRs to analyze the effectiveness of CAV technologies.

Since the acceleration/deceleration behavior of CAVs is largely depends on the information received from the nearby vehicles, an uninterrupted communication is required. Most of the studies in the microsimulation environment analyze the effectiveness of CAV by assuming undisturbed communication between vehicles which might not be feasible in the real world scenario. However,
none of the previous studies considered the safety effectiveness by using real communication system along with the different traffic condition for the application of CAV in the simulation environment. Therefore, this study utilized Dedicated Short-Range Communications (DSRC or IEEE 802.11p) and wireless access (IEEE 1609 protocol) for the application of vehicle ad-hoc network (VANET).

Most of the simulation was performed in Aimsun Next and SUMO & OMNET++ based Veins platform. Different car-following models, signal planning algorithm, and real communication systems were implemented by using application programming interface (API) and C++ language. Also, several statistical tests and modeling techniques were utilized for the evaluation of traffic safety and efficiency with different MPRs.

1.2 Objectives of the Research

The specific objectives for the dissertation are described here:

*Objective 1: Calibration of Dynamic Traffic Assignment Model for a Large-scale Network by using Multi-Resolution Modeling technique for the deployment of CAV technologies.*

   The objective of this task is to develop a data intensive framework for deployment, calibration, and validation of a Dynamic Traffic Assignment (DTA) model for Orlando metropolitan area, Florida, using Multi-Resolution Modeling (MRM) framework. The Regional Traffic Demand Model (RTDM) named Orlando Urban Area Transportation Study (OUATS) with a base year of 2009 was extracted from Cube Voyager modeling software and employed to develop DTA based mesoscopic and microscopic simulation model using the MRM framework. To develop the DTA model, we extracted Orange and Seminole counties from RTDM as the large subarea network as mesoscopic area, while relatively small subarea network including Downtown
Orlando and East Orlando were utilized as microsimulation area using Aimsun Next modeling software. This study extensively processed and summarized input data from multiple datasets. A detailed analysis of model output, calibration, and validation process is presented. In the mesoscopic area, a good calibration and validation results were found which were within the acceptable bound of error. In terms of the microscopic areas, this study also reached a calibration and validation criteria based on widely accepted guidelines. Finally, the calibrated and validated large-scale DTA model was utilized for the application of CAV technology.

**Objective 2: Application of Connected and Automated vehicles in a large-scale network by considering V2V and V2I technology**

The objective of this task is to analyze the effectiveness of CAVs at the network level by utilizing both vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication technologies. Also, the study proposed a new signal control algorithm through V2I technology in order to elevate the performance of CAVs at intersections. This study considered a car-following model (CACC) in order to approximate the driving behavior of CAVs in the Aimsun microsimulation environment. For the testbed, the research team selected Orlando CBD (central business district) area in Florida. To this end, the impacts of CAVs were evaluated based on traffic efficiency (e.g., travel time rate (TTR), and average approach delay, etc.) and safety (e.g., standard deviation of speed, real-time crash-risk models for freeways and arterials). The results showed that the application of CAVs reduced TTR significantly compared with the base condition even with the low market penetration level.
Objective 3: Assessing the Benefits of Connected and automated vehicles over only automated vehicles on a congested highway by considering multi-vehicle communication system

The applications of connected and automated vehicles (CAVs) are expected to increase traffic safety and mobility on freeway segments. Vehicle-to-vehicle (V2V) communication will help CAV to communicate with other vehicles and exchange information such as position, speed, and acceleration. With the help of this communication, CAVs will get the information from not only the immediate leader but also from the other preceding vehicles within the communication range. This advanced information will help CAVs to make decision about acceleration/deceleration choices for a safe car-following behavior. Most of the recent studies only assume the communication between adjacent two CV vehicles without taking into consideration about the acceleration behaviors of multiple leading vehicles. Therefore, the objective of this task is to evaluate the effectiveness of connected and automated vehicles (CAVs) over only automated vehicles (AVs) by incorporating multiple preceding vehicle information (i.e., acceleration, position, etc.). This study used a separate car-following model for both CAVs and AVs in order to approximate their driving behavior in the Aimsun Next simulation platform. Additionally, a different car-following model is used for the connected vehicles (CVs) without automation by addressing the human driver compliance factor. This study also utilizes mixed penetration of CAV and CV with no automation. A well calibrated and validated simulation testbed is developed for SR417 in Orlando, Florida which is the base scenario in this study. To this end, the impact of CAVs, AVs, and CVs are evaluated based on both traffic efficiency (i.e., travel time) and safety (i.e., traffic conflicts) under various market penetration rates (MPRs). Since 100% market penetration rate of CV will not be available in the near future, there will be a long period of mixed traffic flow where human driven vehicles and CAVs move together. Therefore, the study
also incorporated different market penetration rates (MPRs) to identify the possible benefits of CAVs.

**Objective 4: Evaluation of traffic Safety based on DSRC Communication Parameters under Connected and Automated Vehicles Environment**

This task investigated the effectiveness of real communication protocol in microsimulation under the connected and automated vehicle (CAV) environment. Since the benefits of CAV environment mostly depend on the successful communication system, this task evaluated how the communication protocol affects traffic safety under CAV environment. There have been several studies focusing on only communication system or traffic condition. However, the combination of communication and traffic condition under CAV environment is not yet analyzed. Also, most of the studies in the microsimulation environment analyze the effectiveness of CAV by assuming uninterrupted communication between vehicles which might not be feasible in the real-world scenario. Therefore, this study considered the real communication protocol as well as the car-following model in order to mimic CAV performance in the simulation environment by incorporating vehicle ad-hoc network (VANET) where CAV is used as on-board unit (OBU) which can communicate with other vehicles through V2V technology. This study used Dedicated Short-Range Communications (DSRC or IEEE 802.11p) and wireless access (IEEE 1609 protocol) for the application of vehicle ad-hoc network (VANET). Also, the intelligent driver model (IDM) was utilized to represent the longitudinal car-following behavior of CAV. The objective of this study is to analyze the performance of different communication parameters and traffic conditions in terms of traffic safety in the CAV environment. The simulation study was carried out on one of
the major expressways (SR408) in Orlando, Florida which was calibrated and validated based on real traffic data. The simulation was performed in SUMO and OMNET++ based veins platform.

1.3 Dissertation Structure

In Chapter 2, a detailed literature review is conducted on different approaches of connected vehicles and communication systems of connected vehicles. Recently, many researchers are going on to evaluate the effectiveness of connected and automated vehicles. Most of these studies analyze the safety and mobility benefits of CAV technology.

Chapter 3 presents the development of a data intensive framework for deployment, calibration, and validation of a Dynamic Traffic Assignment (DTA) model for Orlando metropolitan area, Florida, using Multi-Resolution Modeling (MRM) framework. In this research, the Regional Traffic Demand Model (RTDM) named Orlando Urban Area Transportation Study (OUATS) with a base year 2009 was extracted from Cube Voyager modeling software and employed to develop DTA based mesoscopic and microscopic simulation model using the MRM framework. Orange and Seminole counties from RTDM were analyzed as a mesoscopic area where Downtown Orlando and East Orlando were utilized as microscopic areas. The DTA based model was calibrated within the acceptable bound of error for both mesoscopic and microscopic level. The results of the calibration and validation process showed that the model can replicate the real-world traffic condition.

Chapter 4 illustrates the mobility and safety benefits of connected vehicle in a large-scale network. This study evaluates the effectiveness of CAV for both freeways and arterials simultaneously. The simulation experiments were set up by utilizing both V2V and V2I technologies. Also, the study proposed a new signal control algorithm through V2I technology in order to elevate the performance of CAVs at intersections. This study considered a car-following
model in order to approximate the driving behavior of CAVs. For the testbed, Orlando CBD (central business district) area in Florida was selected. The results showed that the application of CAVs reduced TTR significantly compared with the base condition even with the low market penetration level. Also, the proposed signal control algorithm reduced the approach delay at the intersection.

Chapter 5 shows the multi-vehicle communication system for the application of CAVs on freeway segment. This study analyzes the benefits of CAV over only automated vehicle in terms of both safety and mobility. Also, the study incorporates connected vehicle (CV) without automation by addressing driver’s compliance factor. The study considers three different car-following models in order to approximate the behavior of CAV, AV, and CV. Since the benefits of CV mostly depends on the driver compliance on the received information, the study utilized a mixed penetration of CAV and CV to mimic the real-world traffic scenario. A well calibrated and validated simulation testbed is developed for SR417 in Orlando, Florida which is the base scenario in this study. Finally, the results of this study indicate a significant improvement of both traffic efficiency and safety by implementing CAV and AV technologies on the freeway segments compared to the base scenario and CAV outperforms AV.

Chapter 6 depicts the application of CAV technology by utilizing real communication protocol in the simulation environment. This study uses Dedicated Short-Range Communications (DSRC or IEEE 802.11p) and wireless access (IEEE 1609 protocol) for the application of vehicle ad-hoc network (VANET). Also, the intelligent driver model (IDM) is utilized to represent the longitudinal car-following behavior of CAV. To this end, the study evaluates the impact of different communication on traffic safety under CAV environment. Also, the study considers different traffic conditions for the safety evaluation. The results of this study show that some
communication parameters have a significant effect on traffic safety. Finally, the results of this study provide an insight to the transportation planners and the decision makers about the safety effectiveness of CAV technology by using the real communication system.

Finally, Chapter 7 describes the conclusions and final outcome of the dissertation. Also, the implications of this study is discussed here.
CHAPTER 2: LITERATURE REVIEW

There are mainly three modeling approaches that have already been developed to analyze the transportation systems for both operational and planning purposes: macroscopic, mesoscopic, and microscopic. The Regional Travel Demand Models (RTDM) focuses on only major roadways and intersections referred to as links and nodes, respectively. The important parameters of the macroscopic model are link speed, capacity, and assigned traffic volume. However, vehicles are not individually modeled but are aggregated into the link demand as OD matrix (Zitzow et al., 2015). Nevertheless, macroscopic models have some key limitations in terms of incident management, infrastructure development, active traffic management, and decision support system.

Unlike macroscopic models, microscopic models have a high level of resolution including car following, lane changing, and other driving behavior models incorporating a fixed time-step framework. Detailed geometry, traffic, and signal timing information are needed to get the accurate representation of traffic such as queuing at intersections and congestion throughout the network. Calibration of the microscopic models is computationally intensive for a large area network. Due to the computational intensity of microscopic models and the limitations of macroscopic models, mesoscopic models bridge the gap between these two extremes (Mahut, 2001; Shafiei et al., 2018). Mesoscopic models use macroscopic rules to represent each individual vehicle in the network, which needs much less time than the microscopic model. Compared to macroscopic models, mesoscopic models are not limited in analyzing infrastructure development, active traffic management, and decision support system.

The MRM framework is an essential practice to simulate the DTA based simulation model. MRM refers to a modeling framework that combines microscopic, mesoscopic, and macroscopic representations of traffic flow in the modeling effort of a single project. The Federal Highway
Administration (FHWA, 2012) classified MRM into partial MRM and full MRM, as shown in Figure 2.1. The former one is to use demand forecasting models to provide initial demand estimates to mesoscopic or microscopic modeling tools. The latter one utilizes a mesoscopic simulation-based DTA model for a large sub-area using trip demands from the demand models and produces input data (Davis and Bigelow, 1998). Furthermore, the microscopic models can be used to provide the detailed analysis of selected sub-areas, corridors, or facilities. The full MRM approach is used in the network in which the initial demand matrices are estimated based on the approved RTDM and then wide-area diversion and bottleneck impacts are modeled using mesoscopic DTA models, followed by detailed analysis of traffic operations using microscopic models. Recently, researchers are implementing the MRM technique to develop a DTA based mesoscopic and/or microscopic simulation models utilizing large-scale networks (Hadi et al., 2015; Luo and Joshua, 2011; Massahi, 2017; Mirchandani et al., 2018; Qom, 2016; Tokishi and Chiu, 2013). Simulation-based dynamic traffic assignment (DTA) is an effective tool to analyze transportation systems for both operational and planning purposes (Ben-Akiva et al., 2012; Hadi et al., 2016a).

![Figure 2.1 Examples of MRM structure](source: Cambridge Systematics, Inc. (FHWA, 2012))

Very few attempts have been made at large-scale DTA models for major metropolitan
regions across the world. The first large-scale model was used to evaluate a variety of traffic system management strategies for a fifty-mile segment of US 101 in California (Pravinvongvuth and Loudon, 2011). The simulation model was successfully calibrated with the base-year according to the recommended calibration guidelines of the Federal Highway Administration (FHWA) (Wunderlich et al., 2019). However, the model was not validated to achieve high degree of assurance of good validation both the mesoscopic and microscopic level which is needed to develop a successful deployment of simulation model. Zitzow et al., (Zitzow et al., 2015) developed a large-scale model incorporating components of both microscopic and macroscopic modeling to achieve high resolution required along the light rail corridor in Minneapolis–Saint Paul, Minnesota. However, the study calibrated the DTA model, but they overlooked the validation process. Hence, none of the studies considered validation of the calibrated DTA model for every level of simulation: macroscopic, mesoscopic, and microscopic. In this study, the research team developed a calibrated and validated large-scale DTA based simulation model for all three levels. Thus, this study aims to develop a DTA based well calibrated and validated simulation model which will replicate the real-world traffic scenario.

Application of connected and automated vehicles (CAVs) have a great potential to improve crash risk which is caused by human drivers’ errors, and mobility during congestion in the transportation network. In most of the previous studies (Guériaud et al., 2016; Talebpour and Mahmassani, 2016), researchers evaluated the effectiveness of CAVs in a small network by incorporating different car-following models in the microsimulation environment and found the improvement of traffic safety. It is noticed that none of those studies considered the effectiveness of CAVs in a large-scale network (consisting of both freeways and arterials) with different level
of penetration rates. Hence, it is worthwhile to study the traffic efficiency and safety by utilizing varying penetration level of CAVs at a network level. In addition, the driving behavior of CAVs should be significantly different from conventional vehicles in terms of reaction time, acceleration profile, etc., which create a challenging task to model the driving behavior of CAVs in the simulation environment. Previously, very few studies (Guériau et al., 2016; Talebpour and Mahmassani, 2016) used car following model calibrated by connected vehicle real-field data while other studies (Mirheli et al., 2018; Tajalli and Hajbabaie, 2018) used default car-following model of the microsimulation environment in order to mimic the behavior of CAVs which would not replicate the real-world behavior of connected and automated vehicles. So, in this study, a widely accepted car-following model proposed by Van Arem et al., (Van Arem et al., 2006) was used to approximate the behavior of CAVs for both freeways and arterials.

Moreover, there are very few studies (Rahman et al., 2019, 2019) that examined the application of CAVs on arterial segments by considering V2I communication technology. Lee et al. (2013) proposed a connected vehicle intersection control algorithm under fully automated traffic condition only at four-leg intersections by using VISSIM default car-following model to simulate CAVs. Guler et al. (2014) analyzed the connected vehicle technology for two one-way roadways with a traffic optimization algorithm. There are other research (Fajardo et al., 2011; Zohdy and Rakha, 2016) which presented signal-free control strategies under fully connected vehicle environment but those studies were limited to 100% market penetration of CAVs which might not be feasible in the near future. Hence, it is more reasonable to study the CAVs application with low level of market penetration to maximize the benefits of safety and mobility in the whole network.
Furthermore, some researchers (Smith and Razo, 2016; Wang et al., 2019) used platoon control model in order to improve the intersection efficiency under CAVs condition. Sun et al. (2017) proposed dynamic lane assignment and green duration optimization based on automated vehicles. On the other hand, Zheng et al. (2017) developed a signal optimization algorithm by using GPS trajectory data from CVs. Hence, none of these studies consider the mixed traffic condition under low market penetration of CAVs. Previously, some studies (Feng et al., 2015; Zhuofei Li et al., 2014) adjust the SPaT plans automatically based on real-time information provided by the CV. All these studies examined the signal timing optimization based on the CV vehicle’s current speed and locations without considering the manually driven vehicles. Hence, the result of these studies showed the improvement of CV vehicles’ throughputs but could not improve the overall intersection efficiency. Also, most of the studies only optimized the signal control by considering each vehicle’s speed and location without considering of CV vehicle string which is a major disadvantage. This study proposes an algorithm in order to control CAVs string at the intersections by utilizing V2I communication technology and the signal control optimization based on both CAVs and manual driven vehicles. However, until now, no researcher has analyzed the application of CAV by utilizing V2V and V2I communication technology for a large-scale network.

In recent years, connected and automated vehicles (CAVs) have received tremendous attention from the government, automakers, and many researchers because of their potential to improve traffic safety and mobility in the transportation network. With the advent of various technologies, many car-manufacturers have introduced Advanced driving assistance system (ADAS) in the car such as adaptive cruise control (ACC), cooperative adaptive cruise control (CACC) systems, collision warning system (CWS), etc. This ADAS system helps the vehicles to
drive by itself under certain weather and traffic conditions. Since the adoption rates of CAVs is very low compared with the human driven vehicles and the availability of real field data on the CAV’s performance are limited, many researchers have used simulation-based studies to evaluate the impact of connected and automated vehicles in the transportation network.

There is a significant difference in driving behavior between human driven vehicles (HDVs) and CAVs. For example, CAVs can obtain information such as speed, acceleration, deceleration, position, etc., from the preceding vehicles within the communication range. Based on this information, the automated features help vehicles to drive by itself depending on different levels of autonomy stated by Society of Automotive Engineers (SAE) (Watzenig and Horn, 2017). Moreover, driving behavior parameters such as desired speed, headway, maximum acceleration, etc., of CAVs are different from the human driven vehicles. Therefore, modeling the driving behavior of CAV in the simulation environment is a challenging task.

In the literature, there are a considerable amount of studies (Kesting et al., 2008; Milanés and Shladover, 2014; M.S. Rahman et al., 2019; Rahman et al., 2019; Sun et al., 2018) which evaluated the effectiveness of connected and automated vehicles (CAVs) in the simulation environment by utilizing different car-following model to approximate the behavior of CAVs. For instance, Rahman et al., (2018) analyzed the longitudinal safety benefits of managed lane connected vehicles (CVs) platoons on an expressway by utilizing the intelligent driver model (IDM) in VISSIM simulation environment with different market penetration rates. To evaluate the safety effectiveness, the study estimated various surrogate safety measures. The results of this study showed that CVs could improve safety for each penetration level compared with the base condition. Li et al., (2017) evaluated the safety impact of adaptive cruise control (ACC) system on freeways by using IDM with different model parameters in Matlab based on time-to-collision
(TTC) surrogate measure. The study found that the performance of the ACC system for improving the traffic safety is significant compared with the field scenario. They have also tested the benefits of ACC with different penetration rates. Kesting et al., (2008) performed a study to evaluate the ACC system by implementing the IDM model in the simulation platform. The findings of this study implied that with a low penetration rate of ACC vehicles, traffic congestion was eliminated, and travel times were significantly reduced.

Along with the intelligent driver model (IDM), some previous researchers (Shladover et al., 2012; Talebpour and Mahmassani, 2016; Van Arem et al., 2006) used MIcrosopic model for Simulation of Intelligent Cruise control (MIXIC) model to implement the connected and automated vehicles (CAVs) which incorporates vehicle-to-vehicle (V2V) communication between the current and the preceding vehicle by sharing speed, acceleration, position, etc. Van Arem et al., (2006) evaluated the effectiveness of cooperative adaptive cruise control (CACC) by utilizing MIXIC model in the simulation environment. The results showed that the CACC model could improve the traffic flow stability, as well as traffic flow efficiency, compared with the non-equipped CACC vehicles. Shladover et al., (2012) analyzed the impact of CACC system on freeway traffic flow with varying market penetration rate in the Aimsun simulation environment. The results of this study illustrated that the CACC system could increase the capacity of the freeway segment significantly compared with the regular vehicles. Moreover, Liu et al., (2018) measured the benefits of CACC on a mixed traffic flow (combination of CACC vehicles and human driven vehicles) in a multi-lane freeway. The researchers found that CACC vehicles could increase the throughput of the freeway segments. Also, it could improve the throughput of the freeway on-ramp area.
It is worth mentioning that, some of the studies (Mirheli et al., 2018; Qian et al., 2014; Tajalli and Hajbabaie, 2018) used existing default car-following model present in the simulation environment in order to model the CAVs which would not replicate the real driving behavior of CAVs. Since many studies have evaluated the impact of CAVs either in the freeway segments or arterials, a detailed summary of the previous studies employing the simulation environment is presented in Table 2.1 which provides information about the car-following model, communication pattern, evaluation measures, and results.

<table>
<thead>
<tr>
<th>References</th>
<th>Car-following model (CAVs)</th>
<th>Communication pattern</th>
<th>Evaluation Measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Kesting et al., 2007)</td>
<td>IDM model</td>
<td>Following and the nearest leading vehicle</td>
<td>Throughput</td>
<td>Low market penetration (5%) of CAV can improve traffic flow.</td>
</tr>
<tr>
<td>(Zhou et al., 2016)</td>
<td>Cooperative Intelligent Driver Model (CIDM)</td>
<td>Following and the nearest leading vehicle</td>
<td>Traffic operation</td>
<td>AVs could reduce the total travel time and improve the traffic oscillations</td>
</tr>
<tr>
<td>(Rios-Torres and Malikopoulos, 2017)</td>
<td>Optimal control model for CAVs</td>
<td>Following and the nearest leading vehicle</td>
<td>Traffic flow and stability</td>
<td>CAVs can reduce the travel time and fuel consumption.</td>
</tr>
<tr>
<td>(Liu et al., 2018)</td>
<td>CACC model</td>
<td>Following and the nearest leading vehicle</td>
<td>Throughput analysis</td>
<td>Capacity of the freeway has been increased around 90% more compared with the manual driven vehicle</td>
</tr>
<tr>
<td>(Kesting et al., 2010a)</td>
<td>IDM with constant acceleration</td>
<td>Following and the nearest leading vehicle</td>
<td>Capacity of the Roadway</td>
<td>ACC system can improve the capacity of the roadway by about 0.3 percent for 1 percent increase in vehicles</td>
</tr>
<tr>
<td>(VanderWe rf et al., 2001)</td>
<td>ACC and CACC model</td>
<td>Following and the nearest leading vehicle</td>
<td>Capacity analysis</td>
<td>ACC vehicles could increase the capacity of the roadway</td>
</tr>
<tr>
<td>(Van Arem)</td>
<td>MIXIC model</td>
<td>Following and the nearest leading vehicle</td>
<td>Traffic</td>
<td>CACC could improve</td>
</tr>
<tr>
<td>References</td>
<td>Car-following model (CAVs)</td>
<td>Communication pattern</td>
<td>Evaluation Measures</td>
<td>Results</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>---------</td>
</tr>
<tr>
<td>et al., 2006</td>
<td>nearest leading vehicle</td>
<td>operation</td>
<td>the traffic-flow as well as traffic stability</td>
<td></td>
</tr>
<tr>
<td>(M.S. Rahman et al., 2019)</td>
<td>IDM</td>
<td>Following and the nearest leading vehicle</td>
<td>Traffic safety</td>
<td>CV could improve the traffic safety in the arterial segment</td>
</tr>
<tr>
<td>(Aria et al., 2016)</td>
<td>Default model (Wiedemann 99)</td>
<td>Following and the nearest leading vehicle</td>
<td>Density, travel time</td>
<td>Travel time of the automated vehicle was improved compared with the manual vehicle</td>
</tr>
<tr>
<td>(Li et al., 2013)</td>
<td>Default model in VISSIM</td>
<td>Following and the nearest leading vehicle</td>
<td>Intersection performance evaluation</td>
<td>Approach delay was reduced, and intersection performance was improved greatly</td>
</tr>
<tr>
<td>Talebpour and Mahmassani, 2016</td>
<td>IDM and CACC model</td>
<td>Following and the nearest leading vehicle</td>
<td>Stability analysis if CAVs</td>
<td>CAVs could improve the stability under mixed traffic condition</td>
</tr>
<tr>
<td>(Wu et al., 2015)</td>
<td>Default model in VISSIM</td>
<td>Following and the nearest leading vehicle</td>
<td>Traffic operation at intersection</td>
<td>CV could significantly decrease the delays and number of stops at intersection</td>
</tr>
</tbody>
</table>

From Table 2.1, it could be seen that different studies used different car-following behavior to model the CAVs in the simulation environment. Hence, the major limitation is that all the studies considered the speed/acceleration behavior of the nearest preceding vehicle only in order to approximate the longitudinal car-following model for CAVs which would not represent the real-world connected vehicle driving behavior. However, the main idea of vehicle connectivity is that a vehicle can communicate with all the preceding vehicles within the specified communication range through the V2V communication system. None of the previous studies considered multiple preceding vehicle communication systems for modeling connected and automated vehicles. Therefore, much more realistic driving behavior is required to simulation the CAVs in the simulation environment.
Since the 100% penetration of connected and automated vehicles (CAV) is not feasible in the near future and the real field CAV data are limited, many researchers are utilizing the simulation-based approach to evaluate the benefits of connected and automated vehicles. In the previous studies (Makridis et al., 2018; Milanés et al., 2013; Tajalli and Hajbabaie, 2018; Talebpour and Mahmassani, 2016; Wu et al., 2015), researchers examined the effectiveness of CAVs on freeways and arterials by implementing different car-following models in the simulation platform. The car-following behavior of CAVs is significantly different from conventional vehicles. The time headway between the following and the leading vehicle, reaction time, are expected to be less for CAVs compared with the non-CAVs. In terms of modeling the CAV behavior, two distinct modeling approaches are adopted in literature; models based on intelligent driver model (IDM) (to capture the longitudinal behavior of CAVs) and microscopic models for simulation of intelligent cruise control (MIXIC). The MIXIC models include adaptive cruise control (ACC) and Cooperative adaptive cruise control (CACC) system.

Recent studies (Mirheli et al., 2018; Tajalli and Hajbabaie, 2018) used the car-following model that is available in the microsimulation environment (i.e., VISSIM, Aimsun, SUMO, PARAMICS, NGSIM). Recently, some studies (M.S. Rahman et al., 2019; Wen-Xing and Li-Dong, 2018) used a car-following model (Bando’s model) for autonomous vehicles which considers mean expected velocity of the vehicle. None of the above studies considered an implementation of real communication system for the application of CAV technology. Also, all the previous studies (Makridis et al., 2018; Milanés et al., 2013; Tajalli and Hajbabaie, 2018; Talebpour and Mahmassani, 2016; Wu et al., 2015) only considered the following and the immediate leading vehicle speed to estimate the acceleration function of the car-following models. None of them considered the multiple vehicle information that is within the communication range.
For example, Guériau et al., (2016) used PARAMICS to model the connected vehicles car following by utilizing the IDM model without considering the real communication system between vehicles. Kesting et al., (2008) examined the benefits of connected vehicles (CV) on freeway segments by using IDM car-following model in the microsimulation environment. The results of this study showed a significant improvement in travel time compared to the non-CV condition. Paikari et al., (2014) used V2V and V2I technologies in the PARAMICS simulation by using the default car-following model. The findings of this study implied that a significant improvement of safety as well as mobility on the freeway segments for connected vehicles (CV) scenario. Olia et al., (2016) evaluated the CV technology in a city area by utilizing the default car-following model in PARAMICS and found that 45% improvement of safety under the CV environment. Also, the results of this study demonstrated that CV could improve mobility and reduce greenhouse gas emissions compared to the non-CV condition. Table 2.2 shows a detailed summary of car-following models, roadway type, and communication protocol of different studies on CAV.
Table 2.2 Detailed summary of recent studies

<table>
<thead>
<tr>
<th>References</th>
<th>Car-following model</th>
<th>Roadway Type</th>
<th>Evaluation Measures</th>
<th>Communication protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liu et al., (2018a)</td>
<td>CACC model</td>
<td>Freeways</td>
<td>Traffic operation (flow, capacity)</td>
<td>No real communication protocol</td>
</tr>
<tr>
<td>Talebpour and Mahmassani, (2016)</td>
<td>IDM and CACC</td>
<td>Freeways</td>
<td>Stability analysis</td>
<td>No real communication protocol</td>
</tr>
<tr>
<td>Kesting et al., (2007)</td>
<td>IDM model</td>
<td>Freeways</td>
<td>Traffic operation (travel time)</td>
<td>No real communication protocol</td>
</tr>
<tr>
<td>Wu et al., (2015)</td>
<td>Vissim default (Wiedemann)</td>
<td>Arterials</td>
<td>Traffic operation (delay)</td>
<td>No real communication protocol</td>
</tr>
<tr>
<td>Deluka Tibljaš et al., (2018)</td>
<td>Vissim default (Wiedemann)</td>
<td>Roundabout</td>
<td>Traffic safety</td>
<td>No real communication protocol</td>
</tr>
<tr>
<td>Ma et al., (2018)</td>
<td>CACC</td>
<td>Virtual network</td>
<td>String stability</td>
<td>No real communication protocol</td>
</tr>
<tr>
<td>Jin et al., (2013)</td>
<td>Sumo default (Krauss)</td>
<td>Arterials</td>
<td>Traffic operation (travel time)</td>
<td>No real communication protocol</td>
</tr>
<tr>
<td>Liu et al., (2018b)</td>
<td>CACC model</td>
<td>Freeways</td>
<td>Traffic operation (flow, capacity)</td>
<td>No real communication protocol</td>
</tr>
<tr>
<td>Aria et al., (2016)</td>
<td>Vissim default (Wiedemann)</td>
<td>Freeways</td>
<td>Traffic operation</td>
<td>No real communication protocol</td>
</tr>
<tr>
<td>M.S. Rahman et al., (2019)</td>
<td>IDM model</td>
<td>Arterials</td>
<td>Traffic safety</td>
<td>No real communication protocol</td>
</tr>
<tr>
<td>Olia et al., (2018)</td>
<td>Fritzsche model</td>
<td>Freeway merging</td>
<td>Traffic operation (capacity analysis)</td>
<td>No real communication protocol</td>
</tr>
</tbody>
</table>

From Table 2.2, it is observed that different studies used different modeling techniques to approximate the car-following behavior of connected vehicles by utilizing different simulation methods.
platforms. However, as stated earlier, none of them considered the connectivity technology from the communication perspective. Also, the studies assumed uninterrupted communication between CAVs, which might not be feasible in real-world CAV conditions. Additionally, none of them utilized multiple vehicle information in order to compute the acceleration function of the longitudinal car-following model.

Although very few previous studies (Pallewatta et al., 2017; Ramanathan, 2018) analyzed different communication parameters to evaluate different aspects of the connectivity technology, it is important to highlight that none of them analyzed the impacts of communication parameters on traffic safety under the CAV environment. This important research gap motivates the work in the current paper that seeks to analyze the performance of different communication parameters in terms of traffic safety under the CAV environment. As such, the contributions of this study are divided into three parts: (1) apply realistic car-following model by utilizing multiple vehicle information (2) apply the real communication protocols for modeling CAV, and (3) evaluate the safety benefits of CAV from both the communication and transportation perspectives.
CHAPTER 3: CALIBRATION OF DYNAMIC TRAFFIC ASSIGNMENT MODEL FOR A LARGE-SCALE NETWORK BY USING MULTI-RESOLUTION MODELING TECHNIQUE FOR THE DEPLOYMENT OF CAV TECHNOLOGIES

3.1 Introduction

Transportation planning organizations have been using Regional Travel Demand Models (RTDM) for demand forecasting, resource management, project scheduling, and traffic impact studies. In most cases, these tools are developed at a macroscopic level, including only the basic information about the road network, geometry, and traffic patterns. However, RTDM may not be used for evaluating important policy decisions, infrastructure development, traffic operations, incident management, and decision support system in the context of the network operation. To explore this issue, this paper develops and presents a Dynamic Traffic Assignment (DTA) based simulation model using Multi-resolution Modeling (MRM) framework which combines microscopic, mesoscopic, and macroscopic representations of traffic flow in the modeling effort of a single study.

Macroscopic regional models such as RTDM have some key limitations in terms of infrastructure development, active traffic management, and decision support system. Localized geometric effects are not captured by low-resolution (macroscopic) models, causing errors in speed and delay estimation (Zitzow et al., 2015). Moreover, macroscopic regional models fail to capture traffic control such as actuated signals, adaptive signals, and signal coordination. Thus, a research gap exists between current regional macroscopic models and their ability to evaluate active traffic management systems. To address this gap, the dynamic simulation was introduced with accurate
traffic, geometric, and signal information capturing the high-resolution modeling framework. There are two types of dynamic simulation (1) microscopic (2) mesoscopic. The former one incorporates sufficient details of high-resolution modeling throughout the corridor which becomes computationally prohibitive and difficult to construct for accurately describing the interaction between vehicles and the road network. The latter one is comparatively less detailed but can simulate vast areas efficiently compared to the microscopic model. Therefore, mesoscopic simulation can accommodate the tradeoff between the level of details and running efficiency.

Mesoscopic simulation is divided into two types: event-based simulation and time-stepped simulation. In the event-based simulation, the status of the roadway network and vehicle information are updated only when some events occur while in the time-stepped simulation, the information are updated at an approximately chosen time unit (Xu et al., 2014). In this study, the event-based method was used for the mesoscopic simulation in Aimsun Next since it has higher computation efficiency compared to the time-stepped simulation.

This study developed a calibrated and validated DTA model in Orlando, Florida by incorporating components of both microscopic and mesoscopic modeling using MRM approach. The application consists of a 2 hours morning peak network in Orange and Seminole counties, Florida, which is made up of 18350 links, 8942 nodes, 1416 Traffic Analysis Zones (TAZs), and 2417 signalized intersections. Many of the challenges involved to build the DTA model consists of processing data from multiple data sets, particularly the use of available RTDM, such as an origin-destination (OD) trip matrix to generate the data necessary for a dynamic model with time dependent vehicle demand. Finally, the team calibrated and validated the large-scale DTA model for both mesoscopic and microscopic area.
3.2 Model Overview: Location and Data

The overarching goal of this study was to develop a DTA based simulation model for a larger subarea network from a regional planning model using the MRM framework. According to full MRM, this platform utilizes three modeling levels: macroscopic model from regional planning model, mesoscopic model for a large subarea using trip demands from the RTDM model, and microscopic model to provide detailed analyses of selected corridors or facilities within the larger subarea already modeled using the DTA based mesoscopic simulation. Tasks are divided into five important steps: collecting the data, processing the data, implementing the data processes in the model, model calibration, and model validation.

3.2.1 Study area

The regional model along with the mesoscopic and microscopic study areas are presented in Figure 3.1. The regional model named OUATS with base year 2009 was extracted from Cube Voyager to develop a DTA based simulation model using the full MRM framework. It is worth mentioning that the OUATS covers Orange, Osceola, Seminole, and Lake Counties. In addition, the western portion of Volusia County and the northeastern part of Polk County are also included. The main reason for selecting the 2009 macroscopic model is the unavailability of a recent year travel demand model which is generally conducted every 10 or 15 years. The OUATS network covers over 6,732 miles section lengths and contains 2,438 Traffic Analysis Zones (TAZs), 26,094 links, 11,585 nodes, and 2,702 signalized intersections. In addition to the network data, the demand data was required as an important input for the modeling tools to run the traffic assignment procedure in the DTA model. The OD matrix from the Cube model were imported and used as a baseline initial matrix in the model development. To develop a mesoscopic DTA model, the team have selected the two largest counties in the region: Orange and Seminole Counties, as the subarea
network from the regional model. Two subareas/corridors (1) Downtown Orlando Area (including I-4, SR408, SR50, etc.) (2) East Orlando Area (including SR417, SR434, etc.), were utilized as the microsimulation areas to develop the microscopic DTA model. Figure 3.1(a) corresponds to the regional OUATS model while Figure 3.1(b) illustrates the subnetwork selected for the mesoscopic simulation. Among the road networks in the mesoscopic simulation, two critical areas of traffic generation and attraction were selected for the microscopic simulation (Figure 3.1(c)).

The DTA based mesoscopic and microscopic simulation model was developed for 2 hours in the morning peak (07:00 to 09:00 AM) using the real dataset of October 2017, with a total of eight time-intervals 15 minutes each. The network attributes of the full OUATS, mesoscopic, and microscopic areas are presented in Table 3.1.

**Table 3.1 Network Attributes of OUATS model including mesoscopic and microscopic areas**

<table>
<thead>
<tr>
<th>Network Level</th>
<th>Attributes Name</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full OUATS Model</td>
<td>Links</td>
<td>26,094</td>
</tr>
<tr>
<td></td>
<td>Nodes</td>
<td>11,585</td>
</tr>
<tr>
<td></td>
<td>Traffic analysis zone</td>
<td>2,438</td>
</tr>
<tr>
<td></td>
<td>Signalized Intersection</td>
<td>2,702</td>
</tr>
<tr>
<td>Mesoscopic Area</td>
<td>Links</td>
<td>18,350</td>
</tr>
<tr>
<td></td>
<td>Nodes</td>
<td>8,942</td>
</tr>
<tr>
<td></td>
<td>Traffic analysis zone</td>
<td>1,416</td>
</tr>
<tr>
<td></td>
<td>Signalized Intersection</td>
<td>2,417</td>
</tr>
<tr>
<td>Microscopic Area (Downtown Orlando)</td>
<td>Links</td>
<td>1,628</td>
</tr>
<tr>
<td></td>
<td>Nodes</td>
<td>707</td>
</tr>
<tr>
<td></td>
<td>Traffic analysis zone</td>
<td>179</td>
</tr>
<tr>
<td></td>
<td>Signalized Intersection</td>
<td>254</td>
</tr>
<tr>
<td></td>
<td>Section length (miles)</td>
<td>233</td>
</tr>
<tr>
<td>Microscopic Area (East Orlando)</td>
<td>Links</td>
<td>808</td>
</tr>
<tr>
<td></td>
<td>Nodes</td>
<td>428</td>
</tr>
<tr>
<td></td>
<td>Traffic analysis zone</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>Signalized Intersection</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>Section length (miles)</td>
<td>248</td>
</tr>
</tbody>
</table>
Figure 3.1 Study area (a) full network OUATS regional model (b) mesoscopic area (Orange and Seminole County) (c) microscopic area
3.2.2 Input data for model development

In this study, data from multiple sources and agencies were collected and processed for the demand estimation, model calibration, and model validation. The data contains records of traffic counts, speeds, and travel time which are aggregated at 15-minute intervals. The available data collection systems for freeways and arterials are different. For Freeways, data were collected from MVDS (Microwave Vehicle Detection System), AVI (Automatic Vehicle Identification), ATSPM (Automated Traffic Signal Performance Measures), Insync, NPMRDS (National Performance Measure Research Data Set), HERE, and Bluetooth. While arterials’ data were obtained from Bluetooth, HERE, SPM, and Insync. By conducting the comparative studies for the roadway network, the research team used the most reliable data for different segments on freeways (Chung et al., 2018), and arterials (Gong et al., 2018). For roads with different data availability, the data used for the simulation are summarized in Table 3.2. It is worth mentioning that data from various sources need to be combined to produce final traffic demand for representing the base year (2017).

<table>
<thead>
<tr>
<th>Roadway Type</th>
<th>County</th>
<th>Data Type</th>
<th>Simulation Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freeway</td>
<td>All</td>
<td>AVI, MVDS, HERE, NPMRDS</td>
<td>MVDS, AVI</td>
</tr>
<tr>
<td>Freeway</td>
<td>All</td>
<td>MVDS, HERE, Bluetooth, NPMRDS</td>
<td>MVDS, NPMRDS</td>
</tr>
<tr>
<td>Freeway</td>
<td>All</td>
<td>MVDS, HERE, NPMRDS</td>
<td>MVDS, NPMRDS</td>
</tr>
<tr>
<td>Arterial</td>
<td>Seminole</td>
<td>HERE, Bluetooth, SPM</td>
<td>SPM, Bluetooth</td>
</tr>
<tr>
<td>Arterial</td>
<td>Orange</td>
<td>HERE, Bluetooth, InSync</td>
<td>InSync, Bluetooth</td>
</tr>
<tr>
<td>Collector</td>
<td>Orange</td>
<td>HERE</td>
<td>----, HERE</td>
</tr>
<tr>
<td>Collector</td>
<td>Orange</td>
<td>Bluetooth, HERE</td>
<td>----, HERE</td>
</tr>
<tr>
<td>Collector</td>
<td>Orange</td>
<td>HERE</td>
<td>----, HERE</td>
</tr>
</tbody>
</table>
3.3 Modeling Method

The modeling method used to develop the large-scale DTA based simulation model using the MRM framework is summarized in Figure 3.2. The research team developed this workflow which was followed step by step to build such a large-scale simulation model. The first step of this framework was to update the available RTDM (OUATS) to the existing condition based on the roadway network and traffic volumes of 2017. Then, the two Counties (Orange and Seminole) were extracted from the regional model. This subarea network was used to develop the mesoscopic DTA model which was refined in terms of roadways, zones, and intersections to bring it as close as possible to the real-world. RTDMs were traditionally designed for the daily model that can only produce daily trip matrices. Hence, the OUATS model has only the daily trip matrices whereas the peak-hour traffic is simulated in this project. To achieve a reasonable trip matrix during the peak hour, two steps were conducted. First, the daily matrices were converted to 2-hour OD matrices multiplying by a peak factor. Then, an origin destination matrix estimation (ODME) process was performed to calibrate the 2 hours OD matrices at the macroscopic level using the available real-world traffic count dataset. ODME is a procedure for adjusting a prior OD matrix based on the input traffic counts. The solution algorithm is based on a bi-level model solved heuristically by a gradient algorithm using Aimsun Next which is discussed in the origin-destination matrix estimation section of this paper (Aimsun, 2018). Furthermore, the time dependent OD matrices were adjusted to set traffic counts slicing into shorter time intervals (15 minutes) in the mesoscopic area. Towards this end, static origin-destination departure adjustment step generated the whole 2-hour peak period demand into eight 15-minutes sliced demand. In this step, the team split the demand into different vehicle composition (i.e., cars and trucks). Since the data for vehicle classifications are not available for different types of detectors (NPMRDS, HERE, Bluetooth etc..),
we used MVDS detectors data which provide most accurate vehicle compositions based on the length of the vehicles. MVDS detectors contain seven important variables including volume, speed, and lane occupancy for each lane at 1 min interval, and also categorizes vehicles into four types according to their length; type 1: vehicles 0 to 10 ft. in length, type 2: vehicles 10–24 ft. in length, type 3: vehicles 24–54 ft. in length, type 4: vehicles over 54 ft. in length. Based on the vehicle length, we classified the vehicles into two categories: (1) Car and (2) Truck. A vehicle was considered as a passenger car (PC) if its length is equal to or less than 24 ft. (type 1 and type 2) (Shi, 2014; Wang, 2016) and truck if the length is larger than 24 ft. Based on the MVDS detectors in the study area, the overall percentages of cars and trucks are 90% and 10%, respectively (Shi, 2014). Hence, the percentages were employed in the simulation model. Hence, the percentages were employed in the simulation model.
Figure 3.2 Modeling methods for the DTA based simulation model.

To make the model dynamic, mesoscopic Dynamic User Equilibrium (DUE) step was performed for which the saved path file was utilized in the mesoscopic dynamic traffic assignment model. This step mainly focuses on the dynamic path assignment of all the vehicles throughout the entire mesoscopic area. In DTA, the path assignment of vehicles in the network is a very important and complex process based on the prevailing traffic condition for the large-scale simulation model. The probability of choosing a path $k$ in the DTA model can be obtained with any random utility model such as proportional, multinomial logit, and C-logit. The multinomial logit (MNL) model
is the preferred route choice model due to the closed form representation and easy estimation. However, the underlying modeling hypothesis of logit models is based on a critical assumption of Identical and Independent Distribution (IID) for the random terms in the utility functions. This assumption implies the independence of irrelevant alternatives (IIA) property which does not often hold in real-size networks because of the high correlation among various routes between any OD pair (Ben-Akiva et al., 2012). In other words, the multinomial logit model is unable to distinguish between two alternative routes when there is a high degree of overlapping resulting from the overestimation of choice probability. To overcome this issue, the C-logit model (Cascetta, 2009) is implemented to modify the MNL model by adding a term called ‘commonality factor’. This term is inversely proportional to path k’s degree of independence from other paths and is equal to zero if no other path shares links with path k (Barceló, 2010). The commonality factor is directly proportional to the degree of overlapping of path k with other alternative paths presented in the following equation.

\[ CF_k = \ln(1 + \sum_{l \neq k} \frac{L_{lk}}{\sqrt{L_l L_k}}) \]  

(1)

Where \( L_l \) and \( L_k \) are the cumulative values of the cost attribute over the links in path \( l \) and \( k \), respectively, and \( L_{lk} \) is the cumulative value of the cost attribute over the links that are shared by the two paths. Therefore, the C-logit model reduces the probability of choosing heavily overlapped paths and hence is a more realistic route choice model. Given \( CF_k \), the C-logit model is formulated as follows:

\[ p(k) = \frac{e(-g_k - \beta CF_k)}{\sum_{l \in K_{OD}} e(-g_l - \beta CF_l)} \]  

(2)

Where, \( p(k) \) is the probability of choosing path \( k \), \( g_k \) is the general cost of path \( k \), \( K_{OD} \) is the set of paths connecting OD pairs, and \( \beta \) is the proportion of the commonality factor. \( p(k) \) is calculated
for the maximum 3 shortest paths obtained from the well-known Dijkstra's label-setting algorithm (Cascetta, 2009) based on the current path costs. This specification of choosing a path selection procedure was applied in the DTA model using Aimsun Next. The above utility model was repeated in every assignment interval until all travel demand loaded into network.

After loading all the vehicles in the network, the next step was to calibrate and validate the entire mesoscopic area by using the real dataset of traffic counts and travel time, respectively. Furthermore, two congested microsimulation areas within the mesoscopic large area were further refined to represent the accurate traffic, geometric, and signal information. The microscopic areas of the DTA simulation network were also calibrated and validated based on the real dataset providing a final calibrated and validated large-scale DTA based mesoscopic and microscopic model which is the last step in the proposed workflow. Furthermore, this framework can be a good platform to test the active traffic management strategies in the microscopic model in selected areas.

3.4 DTA Model Development

This section discusses the development of the macroscopic model from the regional travel forecasting model along with the calibration. It is worth mentioning that the macroscopic model is the good starting point of any DTA based simulation model. In this study, The regional travel demand model named Orlando Urban Area Transportation Study (OUATS) provide valuable information such as the origin destination matrices, average trip length, network geometry, and attributes for each link such as length, number of lanes, free-flow speed, capacity, Bureau of Public Roads (BPR) coefficient, Bureau of Public Roads (BPR) exponent, and so on. Bureau of Public Roads (BPR) curves is a type of volume-delay function used to describe the speed-flow
relationships in a travel demand model network based on the available link capacity. The equation of BPR curve is formulated as follows:

\[ t_i = t_{o_i} \left[ 1 + \alpha \left( \frac{v_i}{c_i} \right)^\beta \right] \]

(3)

Where, \( t_i = \) Congestion flow travel time on link i;
\( t_{o_i} = \) Free-flow travel time on link i;
\( v_i = \) Volume of traffic on link i per unit of time;
\( c_i = \) Capacity of link i per unit of time;
\( \alpha = \) BPR coefficient;
\( \beta = \) BPR exponent;

The corresponding link and node attributes, such as the number of lanes, free-flow speed, link capacity, BPR coefficient, BPR exponent were configured in Aimsun Next. Afterwards, the macroscopic model was developed to adjust the demand matrix using the regional traffic demand model incorporating real datasets of traffic counts from multiple sets of detectors.

3.4.1 Origin-destination matrix estimation (ODME)

The estimation of time-variant trip matrices is an important step in DTA-based tools. DTA analysis requires dynamic or time-variant trip matrices specified for short time intervals (e.g., 15 minutes or 30 minutes) (Hadi et al., 2016b). However, regional demand models were traditionally designed for the daily model that can only produce daily trip matrices. More recently transportation modelers are developing “time-of-day” models that produce trip matrices representing the peak periods demand matrices. An ODME process is needed to fill this gap by estimating the trip tables for short intervals based on an initial matrix obtained from the regional demand models combined
with field data. In this study, the research team used the default ODME process in Aimsun since it gives better result compared with the previous study (Hadi et al., 2016a). The first set was produced by factorizing the demand matrix extracted from the regional demand model using static assignment. The second matrix is the calibrated demand matrix produced from static adjustment results. The result showed that the simulated link volumes cannot replicate the observed link volumes when factorization of the demand matrix was used since the corresponding R² is only 0.65. With calibrating the demands, the simulated link volumes get closer to the observed link volumes with a R² value of 0.95 (Figure 3.3). Furthermore, static adjustment results implied that the ODME procedure generates an approximately 61% improvement (RMSE from 1596.82 to 630.42) in terms of root mean square error (RMSE). The solution algorithm for the OD estimation is based on a bi-level optimization method (Codina and Barceló, 2004; Florian and Chen, 1995) solved heuristically by a gradient algorithm. The algorithm estimates the sequence of OD matrices which reduces the least squares error gradually between traffic volumes measured from real field and traffic flows obtained by a traffic assignment process. The OD matrix estimation process requires information about the routes used by trips contained in the OD matrix \( (d_{ij}) \). Also, the definition of route and the trip proportions relative to the total trips \( d_{ij} \) is used on each route originating at zone \( i \) and ending at zone \( j \). This kind of information is difficult to store in the database and can grow exponentially with the size of the network. Therefore, mathematical programming approach based on a traffic assignment algorithm is used, which is solved at each iteration without requiring the explicit route definition (Aimsun, 2018).

The Spiess (Spiess, 1990) bi-level optimization adjustment procedure solves the following bi-level non-linear optimization problem:
\[
\min F(v(g), \bar{v}) = \frac{1}{2} \left\{ \sum_{a \in A} [v(g)_a - \bar{v}_a]^2 \right\}
\]

\[
v(g) = \arg \min \sum_{a \in A} \int_0^{v_a} s_a(x) dx
\]

s.t. \[\sum_{k \in k_i} h_k = g_i, \forall i \in I\]

\[h_k \geq 0, \forall k \in k_i, \forall i \in I\]

\[v_a = \sum_{i \in I} \sum_{k \in k_i} \delta_{ak} h_k = \sum_{i \in I} g_i \sum_{k \in k_i} \delta_{ak} p_k, \left(p_k = \frac{h_k}{g_i}\right), \forall a \in A\]

where \(v(g)_a\) is the flow on link \(a\) estimated by the lower level traffic assignment problem with the adjusted trip matrix \(g\), \(h_k\) is the flow on the \(k\)-th path for the \(i\)-th O-D pair, and \(v_a\) is the measured flow on link \(a\). \(I\) is the set of all Origin-Destination pairs in the network, and \(k_i\) is the set of paths connecting the \(i\)-th O-D pair. \(s_a(v_a)\) is the volume-delay function for link \(a \in A\).
a. Static assignment of full regional network

b. Static adjustment of full regional network.

Figure 3.3 Origin-destination matrix estimation results of OUATS (macroscopic) model.
The previous assignment and adjustment procedures were performed for the full regional model. To develop a mesoscopic DTA model, the research team selected Orange and Seminole Counties as the subarea network from the regional model. This subarea large network is also well calibrated by static adjustment procedures using real datasets with $R^2$ and RMSE value of 0.96 and 590.03, respectively (Figure 3.4(a)) for the whole 2 hours which includes 655 links throughout the network. Furthermore, a static assignment is performed for this subarea network (Orange and Seminole Counties) in order to get the much accurate assigned volume based on adjusted volume, which would be used for signal timing of all the 2,417 signals in the mesoscopic area network. However, those signals would be utilized as dummy signal timing for the mesoscopic simulation area since the number of intersections is large (2,417 signals), while the accurate signal timing data were coded for the microscopic DTA simulation models within the mesoscopic area using full MRM techniques as previously described. This dummy signal timing was calculated based on the traffic demand in each approach during static assignment step in AIMSUN. The dummy signal timing is a sophisticated way to calibrate the mesoscopic area with very large number of signals. The static adjustment and the assignment results for mesoscopic level are presented in Figure 3.4(a) and Figure 3.4(b) with a comparison of simulated link volumes with real-world traffic counts for 655 links throughout the entire mesoscopic area.
a. Static adjustment of subarea network

Figure 3.4 Origin-destination matrix estimation results of mesoscopic area model
3.4.2 Time dependent origin-destination demand

The aforementioned ODME procedure was applied for the adjustment of initial demand matrix to peak period demand matrix with the 2-hour period. However, to prepare the demand for DTA based dynamic modeling (either mesoscopic or microscopic), it is important to adjust the set of traffic counts slicing them into 15-minute intervals. Hence, the time dependent OD matrices were adjusted to set traffic counts slicing into shorter time intervals (15 minutes) based on all the volume detectors (655 for the mesoscopic area) which have real traffic volume for every 15 minutes interval. Based on the real traffic volume and path assignment results of the static adjustment, the demand matrix of each 15-minutes interval of total 2 hours AM peak period was calculated. Figure 3.5 shows the 15 minutes sliced demand results for the whole two hours. From the figure, it could be seen that the $R^2$ and RMSE value is better for the sliced traffic demand.

![Figure 3.5](image)

\[ y = 0.9892x - 5.5586 \]

\[ R^2 = 0.949 \]

RMSE=78.27

**Figure 3.5 Adjustment and observed link counts (every 15 minutes) for the Mesoscopic area**
In a large-scale DTA model, the selection of the warm-up period is done by iterative process until the first 15 to 30 minutes of 2 hours dynamic simulation has a regression slope closer to 1. The optimal warm up period of this large network was found to be 45 minutes, so 2 hours 45 minutes OD matrices (from 6:15:00 to 9:00:00 A.M.) were obtained for each 15-minute time interval in the OD departure adjustment process. With optimal warm up period of 45 minutes, the simulated trip and link counts get closer to the observed trip and link counts with a R² value of 0.95 which confirmed reasonably good results. Moreover, the traffic demand for every 15 minutes interval was obtained from the results of OD departure adjustment procedure. The 2-hour 45 minutes profiled demand is presented in Figure 3.6 below in which first 45 minutes are considered for the warm up period in mesoscopic DTA simulation.

![Figure 3.6 Profiled demand from OD adjustment procedure](image-url)
However, the demand is only based on cars in the simulation. Based on the real-data, two types of vehicles including car and truck were considered to apply at each 15-minutes demand matrices with a proportion of 90% and 10%, respectively.

3.4.3 Calibration of mesoscopic simulation area

First of all, it is necessary to calibrate the entire model at the mesoscopic level to build a reliable framework of the microscopic model. Towards this end, vehicles are loaded into the network and then select routes through a C-logit route choice model. This process means that vehicles departing simultaneously will take routes that experience equal and minimal travel times for any given OD pair at any time during the simulation (Zitzow et al., 2015). To achieve this, a DTA approach is taken at each simulation step for which vehicles entering the network calculate the shortest routes for their OD pairs according to existing link speeds and distribute themselves onto the network. As they move through the road system, link speeds are updated for the next time interval. For the large-scale DTA model, the calibration process is a high-level iterative procedure that consisted of running the model, comparing the output with calibration and validation, changing the global parameters, identifying problem areas, comparing the model data with the real network, identifying appropriate changes, implementing the changes, and rerunning the model. The primary calibration metric used in this large-scale network was 15-minutes traffic counts throughout the whole mesoscopic area. The number of traffic count data was acquired that resulted in 655 calibration points from 655 locations (265 on freeways and 390 on arterials) throughout the entire mesoscopic network from 7:00 to 9:00 A.M. for every 15-minute interval. The count data covered 655 links in the network. Figure 3.7 shows the traffic count detectors locations within the mesoscopic area for freeways and arterials separately. The initial focus was on comparing model
output according to the 2-hours counts. With this approach, the model under predicted delay because of the cold start of the DTA model (i.e., the model began with an empty network). Thus, the optimal 45 minutes warm up period was utilized which is mentioned above.

Figure 3.7 Traffic count detectors locations in the mesoscopic area
Most of the DTA based simulation model used $R^2$ and Root Mean Squared Error (RMSE) value for the calibration criteria of large scale network (Hadi et al., 2016b; Zitzow et al., 2015). Therefore, the team analyzed the $R^2$ and RMSE for the mesoscopic area for every 15 minutes interval and found that the values ranged from 0.88 to 0.90 and 113.31 to 129.42 respectively which is much better than the previous studies. Figure 3.8 illustrates the results of $R^2$ and RMSE on the mesoscopic network for each 15 minute demand from 7:00:00 to 9:00:00 A.M for 655 links.

\[ y = 1.0245x - 27.533 \]
\[ R^2 = 0.901 \]

RMSE=124.64

(a) 7:00 – 7:15 AM
RMSE = 128.46

(b) 7:15 – 7:30 AM

RMSE = 127.43

(c) 7:30 – 7:45 AM
RMSE=129.42

(d) 7:45 – 8:00 AM

RMSE=123.92

(e) 8:00 – 8:15 AM
y = 1.032x - 40.133
$R^2 = 0.8796$

(f) 8:15 – 8:30 AM

y = 1.0382x - 42.06
$R^2 = 0.8961$

(g) 8:30 – 8:45 AM
3.4.4 Validation of mesoscopic simulation area

In practice, a simulation model should be well validated to confirm the credibility of the model that represents the true behavior of the real-world closely enough for decision-making purposes. To the best of our knowledge, none of the large-scale DTA models with the MRM framework (Hadi et al., 2016b; Shafiei et al., 2017; Zitzow et al., 2015) have considered the model validation because of the complexity of such kind of large networks except Shafiei et al. (Shafiei et al., 2018) which was a much smaller network in terms of mesoscopic area compared to this current study. However, previous large scale DTA models (Tokishi and Chiu, 2013; Zitzow et al.,
2015) performed only calibration as achievement of validation criteria is very time consuming. This research team validated the entire mesoscopic model with set of time-dependent path travel times obtained from Bluetooth, NPMRDS, HERE, and AVI detectors for both freeway and arterials. The travel time data that were collected resulted in 259 validation points with 38 on freeways and 221 on arterials throughout the network. The travel time detectors covered around 690 links in the mesoscopic area. Figure 3.9 shows segments (highlighted segments) having travel time data input in the whole mesoscopic area based on the availability of real data.

FDOT provides in general guidelines of travel time validation in two ways: (1) simulated travel time should be within ±1 minute for routes with observed travel times less than seven minutes (2) simulated travel time would be within ±15% for routes with observed travel times greater than seven minutes (FDOT Systems Planning Office, 2014).
There are two main parts to calibrate the parameters in Aimsun Next to achieve the good validation criteria: (1) dynamic traffic assignment or route choice and (2) mesoscopic parameters.
The network had much congestion in some corridors due to two possible reasons: (1) too many vehicles are assigned to that corridor which is governed by dynamic traffic assignment parameters; (2) the model is not properly calibrated in terms of mesoscopic parameters such as reaction time, look ahead distance, the speed acceptance for vehicle type etc. Sensitivity analysis is an acceptable practice to calibrate the simulation model based on adjusting the simulation software parameters (Rahman and Abdel-Aty, 2018b). Therefore, a sensitivity analysis was conducted on Aimsun Next traffic assignment parameters and mesoscopic parameters (behavioral calibration) based on engineering judgment with their allowable minimum and maximum values in the simulation model. For each parameter, a range of values between the minimum and maximum (include default value) were chosen to run the mesoscopic dynamic traffic assignment model and the corresponding percentages of travel time validation criteria were calculated. For each parameter, the maximum value of percentage locations achieving the aforementioned validation criteria is the corresponding calibrated value for that parameter. The calibrated values of both traffic assignment parameters and the mesoscopic parameters are presented with default values in Table 3.3.
### Table 3.3 Aimsun Next Calibration Parameters for Mesoscopic Simulation

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Unit</th>
<th>Default value</th>
<th>Calibrated value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Traffic assignment parameter</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model selection</td>
<td>N/A</td>
<td>uniform</td>
<td>C-logit</td>
</tr>
<tr>
<td>Attractiveness weight</td>
<td>N/A</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Maximum number of initial paths to consider</td>
<td>N/A</td>
<td>All</td>
<td>3</td>
</tr>
<tr>
<td>En-route</td>
<td>N/A</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Path Calculation</td>
<td>N/A</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>C-logit parameters (scale, beta, gamma)</td>
<td>N/A</td>
<td>(1,0.15,1)</td>
<td>(1,0.15,1)</td>
</tr>
<tr>
<td>Maximum Paths per interval</td>
<td>N/A</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td><strong>Mesoscopic parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reaction time</td>
<td>s</td>
<td>1.2</td>
<td>0.9</td>
</tr>
<tr>
<td>Reaction time at traffic light</td>
<td>s</td>
<td>1.6</td>
<td>1.2</td>
</tr>
<tr>
<td>Look ahead distance variability</td>
<td>%</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>Speed acceptance for Car</td>
<td>N/A</td>
<td>1.1</td>
<td>1.5</td>
</tr>
<tr>
<td>Speed acceptance for Truck</td>
<td>N/A</td>
<td>1.0</td>
<td>1.4</td>
</tr>
</tbody>
</table>

In terms of mesoscopic behavioral calibration, look ahead distance was iteratively adjusted in various problematic sections based on the results of unrealistic congestion, missed turns, etc.
The combination of those parameters and adjusting the behavioral calibration provided 85% locations that are within error bound ±1 minutes and ±15% for routes with observed travel times less than seven minutes and more than seven minutes, respectively.

Figure 3.10 Travel time detectors locations and validation results for the entire mesoscopic area.
Figure 3.10 illustrates the validation results of the mesoscopic network for eight 15-minute time intervals (2-hour simulation). The legends showed the number of validated time intervals from 1 to 8 by different colors (from red to green). For example, the red color (i.e., number of validated time interval is 1) means that the corresponding sub-path is satisfied the validation criteria for only one-time interval, while the green color (i.e., number of validate time interval is 8) represents all the 8 time intervals are satisfied with the validation criteria. From the figure, it can be depicted that most of the sub-paths are green, representing good validation. Also, we estimated the $R^2$ and RMSE value which is about 0.83 and 36.13 respectively based on simulated data and real data which is shown in Figure 3.11.

Therefore, the large mesoscopic network achieved reasonable validation within the acceptable bound of errors. Hence, within the DTA framework, the large-scale network is well calibrated and validated by the mesoscopic level of simulation.
3.4.5 Calibration and validation of microsimulation area

Based on the evaluation results, the research team has selected the critical corridors for testing integrated active traffic management (IATM) strategies in the microscopic level. Towards that end, two congested subarea networks (1) Downtown Orlando (including I-4, SR408, Colonial Drive, etc.) (2) East Orlando Area (including SR417, SR434, etc.) were selected as microsimulation areas to test the ATM strategies. The roadway along with the traffic control information of the two microscopic areas are presented in Table 3.1 above. In the microscopic areas, more accurate geometry and signal timing were provided as the original OUATS model was

RMSE=36.13

Figure 3.11 Simulated and real-field travel time plot for Mesoscopic simulation
a macroscopic model with less detailed geometric configuration and traffic. The signal timing data were collected from Orange County, Seminole County, and the City of Orlando. The research team coded the signal timing in the Downtown Orlando and East Orlando areas for 254 and 89 signals, respectively. Most of the intersections have actuated signal timing where very few of them have fixed signal timing.

As mentioned earlier, two microscopic areas, including Downtown Orlando and East Orlando area in Central Florida, were implemented at a microscopic resolution from previously calibrated mesoscopic level using full MRM techniques. To better assess traffic impact studies in an area of interest, it is necessary to calibrate and validate the selected microscopic area with the guidelines of small-scale simulation model used by previous study (Abdel-Aty et al., 2019; Chung et al., 2020; Khan and Rahman, 2016; Lee et al., 2018; M. H. Rahman et al., 2019; Rahman, 2019; Rahman et al., 2019, 2019). Traffic counts and the travel time with each 15-minutes interval were used for the calibration and validation purpose of the microscopic area. The number of volume detectors in Downtown Orlando and the East Orlando, areas are 93 and 84, respectively, while the travel time detectors are 28 and 41, respectively. Figure 3.12(a) and Figure 3.12(b) shows the volume and travel time detectors locations for the Downtown Orlando and East Orlando area, respectively.
Like mesoscopic calibration, there are two main parts to calibrate the parameters in order to achieve the good validation criteria: (1) dynamic traffic assignment or route choice and (2) microscopic parameters. Hence, a sensitivity analysis was also conducted to calibrate both traffic assignment and microscopic parameters to achieve the validation of the microscopic model. The calibrated values of both traffic assignment parameters and the microscopic parameters are presented in Table 3.4.
Table 3.4 Aimsun Next Calibration Parameters for Microscopic Simulation Areas

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Unit</th>
<th>Default value</th>
<th>Calibrated value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Microscopic Calibration Parameters (Downtown Orlando Area)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Traffic assignment parameter</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model selection</td>
<td>N/A</td>
<td>uniform</td>
<td>C-logit</td>
</tr>
<tr>
<td>Attractiveness weight</td>
<td>N/A</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Maximum number of initial paths to consider</td>
<td>N/A</td>
<td>All</td>
<td>3</td>
</tr>
<tr>
<td>Maximum Paths per interval</td>
<td>N/A</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>En-route</td>
<td>N/A</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Path Calculation</td>
<td>N/A</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>C-logit parameters (scale, beta, gamma)</td>
<td>N/A</td>
<td>(1,0.15,1)</td>
<td>(0.5,0.15,1)</td>
</tr>
<tr>
<td><strong>Microscopic parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reaction time</td>
<td>s</td>
<td>1.2</td>
<td>0.8</td>
</tr>
<tr>
<td>Reaction time at traffic light</td>
<td>s</td>
<td>1.6</td>
<td>1.2</td>
</tr>
<tr>
<td>Look ahead distance variability</td>
<td>%</td>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>Speed acceptance for Car</td>
<td>N/A</td>
<td>1.1</td>
<td>1.5</td>
</tr>
<tr>
<td>Speed acceptance for Truck</td>
<td>N/A</td>
<td>1.0</td>
<td>1.4</td>
</tr>
<tr>
<td><strong>Microscopic Calibration Parameters (East Orlando Area)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Traffic assignment parameter</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model selection</td>
<td>N/A</td>
<td>uniform</td>
<td>C-logit</td>
</tr>
<tr>
<td>Attractiveness weight</td>
<td>N/A</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Maximum number of initial paths to consider</td>
<td>N/A</td>
<td>All</td>
<td>3</td>
</tr>
<tr>
<td>Parameters</td>
<td>Unit</td>
<td>Default value</td>
<td>Calibrated value</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>------</td>
<td>---------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Maximum Paths per interval</td>
<td>N/A</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>En-route</td>
<td>N/A</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Path Calculation</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>C-logit parameters (scale, beta, gamma)</td>
<td>N/A</td>
<td>(1,0.15,1)</td>
<td>(1,0.15,1)</td>
</tr>
</tbody>
</table>

**Microscopic parameters**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Unit</th>
<th>Default value</th>
<th>Calibrated value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction time</td>
<td>s</td>
<td>1.1</td>
<td>0.90</td>
</tr>
<tr>
<td>Reaction time at traffic light</td>
<td>s</td>
<td>1.6</td>
<td>1.2</td>
</tr>
<tr>
<td>Look ahead distance variability</td>
<td>%</td>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>Speed acceptance for Car</td>
<td>N/A</td>
<td>1.1</td>
<td>1.5</td>
</tr>
<tr>
<td>Speed acceptance for Truck</td>
<td>N/A</td>
<td>1.0</td>
<td>1.4</td>
</tr>
</tbody>
</table>

To represent the driver’s behavior in the lane-changing decision process, distance zones, cooperation and aggressiveness parameters in the section were adjusted to match the calibration and validation criteria. After calibrating the parameters, the GEH values were calculated for both microsimulation areas. From the microscopic calibration results, both the Downtown Orlando area and East Orlando area achieved 86% of GEH less than 10, which represents good calibration for both microscopic areas based on simulation guidelines (Chung et al., 2020; Lee et al., 2018; Rahman, 2019; Rahman et al., 2019). Validation of the calibrated two microscopic DTA models have also been carried out. The combination of those calibrated parameters with behavioral modifications in Downtown Orlando and East Orlando areas provide 87% and 86% locations which are within ±1 minute and ±15% for routes with observed travel times less than seven minutes.
and more than seven minutes, respectively. Hence, within the DTA framework, the microscopic areas are well calibrated and validated based on accepted guidelines.

### 3.5 Conclusions

This study described the experience of building a large-scale DTA based simulation model with Multi-Resolution Modeling (MRM) framework in Orlando, Florida with the Cube based regional traffic demand model (RTDM). Such models were shown to allow investigation of active traffic management, infrastructure development, and decision support system in an integrated environment beyond the scope of existing regional traffic models. The Orlando Urban Area Transportation Study (OUATS) with base year 2009 was selected as the regional traffic demand model extracted from Cube Voyager. To develop a simulation model, Orange and Seminole counties’ data were extracted from RTDM indicating large subarea network as mesoscopic area, while small subarea network including downtown Orlando and east Orlando were utilized as microsimulation area. The project acquired and prepared numerous data sources, including the RTDM model from Cube, traffic counts, and travel times from multiple detector systems available at all corridors in the studied area. The deployment of data intensive calibration and validation techniques were presented for both mesoscopic and microscopic area. The results showed that both mesoscopic and microscopic areas were calibrated and validated within an acceptable bound of error. In terms of mesoscopic area calibration, $R^2$ was found to be 0.95 and the 85% of the 259 validation locations (travel time) with eight time-intervals spread across the network are within an acceptable bound of error. For the microscopic area, this study also achieved calibration and validation criteria based on widely accepted calibration and validation guidelines.

The calibration and the validation of any DTA based simulation model are highly required for the reliable investigation of traffic condition, incident management, etc., in the network. This
calibrated and validated DTA model has many opportunities for future extension, particularly for applications such as incident management, infrastructure development, and decision support system. Hence, this study utilized this calibrated network for the deployment of CAV technology.
CHAPTER 4: APPLICATION OF CONNECTED AND AUTOMATED VEHICLES IN A LARGE-SCALE NETWORK BY CONSIDERING V2V AND V2I TECHNOLOGY

(Rahman, M. H., M. Abdel-Aty, “Application of Connected and Automated Vehicles in a Large Scale Network by Considering V2V and V2I Technology” under review at Transportation Research Record, 2020.)

4.1 Introduction

Cities are growing rapidly which generate more traffic demand leading to deterioration in safety and efficiency of the transportation system (He et al., 2020). In addition, traffic accidents mainly occur due to drivers-related factors resulting from human errors (Petridou and Moustaki, 2000). During the past two decades, the advances of various technologies such as wireless communication, sensing, software, etc., have created revolutionized changes in the transportation system by incorporating Connected and Automated Vehicles (CAVs) which are capable to improve traffic efficiency and minimize drivers’ error. In particular, CAVs use different sensors and wireless communication systems which help to create an internet of vehicles where every vehicle can communicate with other vehicles through vehicle-to-vehicle (V2V) communications and infrastructure through vehicle-to-infrastructure (V2I) communications. Along with the communication systems, automated driving features enable CAVs to make safe and reliable decisions about acceleration/deceleration choices and lane changing maneuvers which can significantly improve traffic efficiency and safety.

The applications of CAVs are expected to reduce crash risk and improve mobility on the urban arterials and freeways by utilizing V2V and V2I communication systems. Since the market penetration rates (MPRs) of CAVs are increasing gradually in the roadway network, it is high time
to evaluate the overall effectiveness of CAVs in terms of safety and mobility in a large network (considering both freeways and arterials) instead of evaluating them in a small arterial or freeway segments. Therefore, this study aims to evaluate the effectiveness of CAVs in terms of safety and mobility at the network level with different level of market penetration through microsimulation environment in Aimsun Next. Hence, a car-following model, developed by Van Arem et al. (Van Arem et al., 2006), was used to represent the CAVs by using application programming interface (API) in Aimsun which can approximate the real driving behavior of connected and automated vehicles. Moreover, the application of CAVs in the arterials segments is more challenging in order to maintain the connected and automated vehicles string especially at intersections. Hence, this study considered the communication between vehicles and signals through V2I technology which will provide real-time traffic information such as speed of the vehicle, acceleration, location, signal timing, and current phase, etc. This information will CAVs to modify their speed so that they can safely pass through the intersection with reduced delay. Also, the signal controller may adjust the signal phase and timing plan (SPaT) which could improve the overall performance of the intersection as well as the whole network. Therefore, in this study, the research team developed a signal control algorithm based on connected and automated vehicles information in order to minimize the delay of the intersection.

### 4.2 Simulation Setup and Data

Since the objective of this study is to apply connected and automated vehicles at the network level, the research team developed a dynamic traffic assignment (DTA) based microsimulation network in the Orlando CBD (central business district) area in Florida (Figure 4.1) which is heavily congested especially in the morning and evening peak hours because of the presence of commercial offices, world largest theme parks, and other recreational facilities. The
DTA based model was built in Aimsun Next (version 8.3) microsimulation environment where the study period spans 2 hours of the morning peak from 7:00 to 9:00 A.M. on October 11, 2017. The length of the microsimulation network is 233 miles which consist of 1,628 links, 1,416 traffic analysis zones, and 254 signalized intersections. Also, Interstate 4 (I-4), State Road 408 (SR 408) are the major two freeways and several major arterials (SR50, SR423, etc.) are considered in the network.

Figure 4.1 Microsimulation study area
Traffic data such as volume, travel time, and signal timing were collected from multiple sources for the calibration and validation of the DTA based simulation model. For instance, freeways’ traffic information was collected from MVDS (Microwave Vehicle Detection System), AVI (Automatic Vehicle Identification), NPMRDS (National Performance Measure Research Data Set), and HERE, while arterials’ data were obtained from Bluetooth and InSync. Signal timing for 254 intersections was coded in the microsimulation environment from the real-field signal timing data collected from the Orange County and the city of Orlando. Also, different vehicle compositions were used based on MVDS detectors in the study area where car and truck percentages are 90% and 10%, respectively. For a large-scale network, warm-up period is important so that the full network is loaded with vehicles with the right proportion or proper distribution. The optimal warm-up period of this network was found to be 45 minutes which was chosen based on the iterative process. Moreover, the number of detectors used in the network for calibration and validation are 93 and 28 for the volume and travel time, respectively.

4.2.1 Model calibration and validation

In this study, the DTA based microsimulation model was calibrated and validated by using volume and travel time data from the real field which was aggregated into 15 minutes intervals. Hence, Geoffrey E. Heavers (GEH) was utilized to calibrate the model using traffic volume. The objective function for the calibration is to maximize the number of observations for which GEH value should be less than the standard value mentioned in the previous studies (M. H. Rahman et al., 2019).

As mentioned above, Aimsun Next uses the same car-following model (Gipps) for both freeways and arterials which captures the real field driving behavior. Therefore, different car-following parameters need to be adjusted to mimic the real-world condition. Hence, for achieving
a good calibration, the research team divided the parameters fine tuning process in two parts: (1) dynamic traffic assignment or route choice and (2) microscopic parameters. Also, a sensitivity analysis was conducted to calibrate both traffic assignment and microscopic parameters. After tweaking different parameters, the default values and the calibrated values of both traffic assignment and the microscopic parameters are shown in Table 4.1.

After adjusting all these parameters, it was found that in 86% of the cases, GEH value is less than 10 which represents good calibration for the microscopic network. Based on travel time validation guidelines by FDOT, the simulated and field travel time difference should be within ±1 minute for routes with observed travel times less than seven minutes or the difference should be within ±15% for routes with observed travel times greater than seven minutes (FDOT Systems Planning Office, 2014). Hence, the result of validation showed that about 87% of the cases, the validation criteria was satisfied. Finally, the DTA based microsimulation model is well calibrated and validated based on the accepted guidelines.
### Table 4.1 Aimsun Next Calibration Parameters for Microscopic Simulation Areas

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Unit</th>
<th>Default value</th>
<th>Calibrated value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Traffic assignment parameter</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model selection</td>
<td>N/A</td>
<td>uniform</td>
<td>C-logit</td>
</tr>
<tr>
<td>Attractiveness weight</td>
<td>N/A</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Maximum number of initial paths to consider</td>
<td>N/A</td>
<td>All</td>
<td>3</td>
</tr>
<tr>
<td>Maximum Paths per interval</td>
<td>N/A</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td><strong>Microscopic parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reaction time</td>
<td>s</td>
<td>1.2</td>
<td>0.8</td>
</tr>
<tr>
<td>Reaction time at traffic light</td>
<td>s</td>
<td>1.6</td>
<td>1.2</td>
</tr>
<tr>
<td>Look ahead distance variability</td>
<td>%</td>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>Speed acceptance for Car</td>
<td>N/A</td>
<td>1.1</td>
<td>1.5</td>
</tr>
<tr>
<td>Speed acceptance for Truck</td>
<td>N/A</td>
<td>1.0</td>
<td>1.4</td>
</tr>
</tbody>
</table>

### 4.3 Proposed Methodology

In this study, the research team analyzed the effectiveness of CAVs in mixed traffic condition by using vehicle to vehicle (V2V) and vehicle to infrastructure (V2I) communication in a large-scale network. The first part of this section describes the manual vehicle’s behavior. The car-following behavior of CAVs and the communication protocol of V2V and V2I technology are presented at the end of this section.

#### 4.3.1 Vehicle behavior modeling

Modeling driving behavior in the simulation environment for the manual vehicles and CAVs is a crucial component. CAVs are supposed to have different driving behavior compared with the conventional vehicles. Among them, the following distance or gap between the following
and the leading vehicle should be less for CAVs compared with the manual driven vehicles. In this research, the following distance of human-driven vehicles are calculated based on Gipps car following model which is calibrated by using real field data and for CAVs, the following distance estimation is described in the CAVs car-following model. Finally, the car following logic for manual vehicles and CAVs is completely different from each other where manual vehicles car-following models are mimicking human behavior and CAV’s car-following is based on a control logic model.

4.3.1.1 Car following model for Manual vehicles

In this study, the team assumed the behavior of human drivers follows Gipps model which is one of the most widely accepted car following models used by many researchers (Gipps, 1981; Yu et al., 2019). The Gipps model consists of two components (acceleration and deceleration) where the first component indicates the intention of a vehicle to reach a certain desired speed while the second part creates a limitation of the desired speed imposed by the preceding vehicle. The maximum speed to which a vehicle (n) can accelerate during a time period (t, t+T) is formulated in the first part as follows:

\[
V_a(n, t + T) = V(n, t) + 2.5 \alpha(n)T \left( 1 - \frac{V(n, T)}{V^*(n)} \right) \sqrt{0.025 + \frac{V(n, t)}{V^*(n)}}
\]  

where

- \(V(n, t)\) is the speed of the \(n^{th}\) vehicle at time \(t\),
- \(T\) is the reaction time,
- \(\alpha(n)\) denotes the maximum acceleration for vehicle \(n\),
- \(V^*(n)\) represents desired speed of \(n^{th}\) vehicle.
The second part of the model is formulated as follows:

\[
V_b(n, t + T) = d(n)T \\
+ \sqrt{d(n)^2T^2 - d(n)[2s(n, t) - V(n, t)T - \frac{V(n - 1, t)^2}{d(n - 1)}}
\]  \tag{6}

where

\[d(n)\] denotes the maximum deceleration of \(n^{th}\) vehicle,

\[s(n, t)\] is the gap of \(n^{th}\) vehicle.

The speed of the vehicle \(n\) at time \((t+T)\) is estimated as follows:

\[
V(n, t + T) = \min[V_a(n, t + T), V_b(n, t + T)]
\]  \tag{7}

So, the speed of the vehicle in the next step is estimated by using equation 7 which is the minimum of two values from equations 5 and 6. This means that if the gap > safety distance then the vehicle speed is calculated based on equation 5 and if the gap < safety distance, the value from equation 6 is considered as the vehicle speed. In addition, the reaction time and desired distance of manual driven vehicles were calibrated by using real-field traffic data during the aforementioned calibration and validation section.

4.3.1.2 Car following model for CAVs
Connected and automated vehicles have the ability to communicate and monitor the other vehicles constantly in their vicinity and also, they are certain about other drivers’ behavior. The reaction
time of CAVs is less than the manual vehicles which help them to react instantly with any kind of traffic changes. In this study, the reaction time for CAVs is set to be 0.3 seconds based on a previous study (Makridis et al., 2018). Hence, a deterministic car-following modeling behavior based on previous study (Van Arem et al., 2006) was used to model CAVs in this study. According to the car following model, the acceleration of the CAVs depends on the distance and speed difference between leader and follower which can be computed by using the following equation:

\[
a_{\text{ref}, d} = k_a * a_p + k_v * (v_p - v) + k_d * (r - r_{\text{ref}})
\] (8)

where

- \(a_{\text{ref}, d}\) denotes the acceleration of the following vehicle
- \(a_p\) is the acceleration of the leading vehicle
- \(v_p\) is the speed of the leading vehicle
- \(v\) is the current speed of the following vehicle
- \(k_a, k_v, k_d\) are constant factors which values are 1.0, 0.58, and 1 respectively.
- \(r\) represents the current clearance and \(r_{\text{ref}}\) is the reference clearance which is the maximum among safe following distance \(r_{\text{safe}}\), following distance based on reaction time \(r_{\text{sys}}\), and a minimum allowed distance \(r_{\text{min}}\) which is 2 m. The safe following distance is calculated as follows:

\[
r_{\text{safe}} = \frac{v^2}{2} * \left( \frac{1}{d_p} - \frac{1}{d} \right)
\] (9)
where \( d_p \) and \( d \) is the deceleration capabilities of following and leading vehicles, respectively.

Finally, the acceleration of CAVs are calculated using the following equation:

\[
a_{ref} = \min (a_{ref,d}, k \times (v_{int} - v))
\]  \hspace{1cm} (10)

where \( v_{int} \) is the intended speed of the following vehicle and \( k \) is a constant speed error factor which is equal to 1.

The resulting acceleration is between the maximum comfortable acceleration \( 2 \text{ m/s}^2 \) and maximum comfortable deceleration \(-3 \text{ m/s}^2\) based on the recommendation of the previous study (Van Arem et al., 2006).

The illustration of the CAVs car-following model implementation is presented in Figure 4.2(a) which shows the V2V communication between two CAVs when they are within the DSRC communication range (300 m). Hence, the following vehicle will adjust the speed by using the acceleration function expressed in equation 6.

4.3.1.3 Communication system for V2I technology

The application of CAVs is more complex on arterial segments compared with the freeways due to the interruption caused by the traffic signals. Hence, the string operation of CAV will be affected as it passes through the intersections which might result in less traffic efficiency in the arterials. Therefore, this study proposes a signal control algorithm by utilizing V2I communication technology which specifies the communication between signals of the intersections and CAVs. Figure 4.2(b) illustrates the V2I communication process of a typical intersection which shows that the signal will constantly communicate with all the vehicles within a certain communication range.
The V2I communication technology enables CAVs as well as signal controllers to get real-time traffic information e.g., the current phase of the signal, green time, yellow time, location, speed, acceleration, etc., which will help the CAVs and signals to perform better compared with the no communication system. Figure 4.2(c) describes the flowchart of the proposed V2I control model for a typical intersection. The whole process consists of two parts: (1) speed guidance and (2) signal timing optimization.
Figure 4.2 Illustration of vehicle-to-vehicle (a), vehicle-to-infrastructure (b), and signal control algorithm (c)

In order to reduce the approach, delay of an intersection, speed guidance model has been proposed which helps CAVs to adjust their speed based on current signal timing by using V2I communication technology. At each iteration step shown in Figure 4.2(c), the algorithm first collects the speed, current location for every CAV within the DSRC range, queue length for each approach and the speed limit of the roadway. All these parameters represent whether the CAV can
pass through the intersection within the current phase and signal timing. Based on these parameters, the required time to pass the intersection for a subject vehicle is as follows:

\[ t_{\text{req},n} = \frac{d_n - Q_{i,j}}{v_n} \]  \hspace{1cm} (11)

where

- \( d_n \) is the distance required to cross the intersection for the vehicle ‘n’ from its current position,
- \( Q_{i,j} \) is the queue length of \( i^{\text{th}} \) intersection for ‘j’ approach and
- \( v_n \) is the current speed of the vehicle ‘n’. If the time required to pass the intersection is less than the remaining green time of the current phase, then the vehicle ‘n’ can pass through the intersection. If the time required is higher than the remaining green time, then the speed of the vehicle is adjusted by using the following equation:

\[ v_{\text{adjusted},n} = \frac{d_n - Q_{i,j}}{t_{\text{req},n} - t_{\text{green}}} \]  \hspace{1cm} (12)

where \( t_{\text{green}} \) is the remaining green time of the current phase. The value of \( v_{\text{adjusted},n} \) should be within the posted speed limit of the roadway. Note that the subject vehicle will use the constant speed once the speed has been adjusted. Also, this study used simplified vehicle dynamics so that it can be executed in real-time traffic operation without any computational delay.

In addition, with the speed guidance model, the research team optimized the signal timing in order to minimize the approach delay and maximize the throughput for CAVs. In order to minimize the approach delay for intersections, a weight coefficient has been used which is calculated by using the headway of CAVs and manual vehicles. Thus, the maximum throughput optimization function in the algorithm is formulated as follows:
\[ T_c = \max \sum_{i=1}^n (w_i \cdot n_i(g_i,C)) \] (13)

where \( T_c \) is the maximum throughput vehicle number, \( n_i(g_i,C) \) is the actual throughput vehicles of the \( i^{th} \) approach, \( g_i \) is the green time, \( C \) is the cycle length, and \( w_i \) is the weighted coefficient which is calculated by using equation 14 (Liu et al., 2019). Based on this optimization function, the algorithm will assign more green time to an approach with more CAVs than the other approaches.

\[
w_i = \begin{cases} 
1 + \frac{h_{manual} - h_{CAVs}}{h_{manual}} & \text{for CAVs} \\
1 & \text{for Manual Cars}
\end{cases}
\] (14)

where \( h_{manual} \) and \( h_{CAVs} \) is the headway of manual cars and CAVs, respectively. The value of headway for CAVs is less than the manual driven vehicles (\( h_{CAVs} < h_{manual} \)). Thus, a higher weight is added to the CAVs since they would consume less green time compared with the manual vehicles. Moreover, the total cycle length is used as a constraint of the optimization function which is similar to the real-field. Also, the minimum green time from the real-field dataset has been used as another constraint to minimize the delay for the coordinated phase.

4.4 Performance Evaluation Measures

For the evaluation measures, the research team analyzed the effectiveness of CAVs at the network level as well as in the freeways and arterials separately, with different MPRs of CAVs. Also, the approach delay for each intersection in the network was analyzed to see the possible benefits of the proposed signal control algorithm.
Travel time rate, and approach delay at intersections were used for the evaluation of traffic efficiency. These measures were obtained by analyzing the data from the well calibrated and validated DTA based simulation model. For the safety assessment, standard deviation of speed was used as a surrogate measure. In addition, the real-time crash risk models based on previous studies (Shi and Abdel-Aty, 2015; Yuan et al., 2018a) were utilized to evaluate the safety performance of CAVs.

4.4.1 Traffic efficiency

Traffic data such as travel time was collected from the simulation for every 5 minutes interval in order to make the comparison between base condition and CAVs conditions. Hence, the collected travel time from the simulation was converted into travel time rate (TTR) by the length of each segment as follows:

\[
\text{Travel Time Rate (TTR; second/mile)} = \frac{\text{Travel Time (second)}}{\text{Length of each segments (mile)}} \tag{15}
\]

In addition, a statistical test was also conducted to verify whether traffic efficiency is significantly improved. For each scenario (base, CAV), traffic state generates twenty four data points for the whole 2 hours of simulation. Since the sample is relatively small (less than 30), a non-parametric test (Wilcoxon signed-rank) was conducted for the evaluation. Also, the average approach delay for each intersection was analyzed for the evaluation of the proposed signal control algorithm.
4.4.2 Traffic safety

For safety evaluation, different real-time crash-risk models were selected to analyze the safety performance for the freeways and arterials. In most of the previous real-time crash risk prediction models, the researchers used weather information and some other parameters which are not available in the simulation environment. Therefore, the research team used all the available parameters in the simulation related to actual prediction models to give an indication that the application of CAVs would not produce negative safety implications. In addition to the crash-risk models, standard deviation of speed was used as a surrogate measure of safety assessment.

4.4.2.1 Freeways crash-risk indicator

For freeway crash risk prediction, the random parameters logistic model for peak hours from the previous study (Shi and Abdel-Aty, 2015) was selected. From the study, the probability of total crash occurrence could be calculated as:

\[
\text{crash risk indicator}(p_i) = \frac{\exp\left(-3.315 + 0.823X_1 - 0.048X_2 + 6.190X_3\right)}{1 + \exp\left(-3.315 + 0.823X_1 - 0.048X_2 + 6.190X_3\right)}
\]

(16)

Where \(\text{crash probability}(p_i)\) is the predicted probability of total crash occurrence of the \(i^{th}\) observation; \(X_1\) is the logarithm of volume of the upstream section; \(X_2\) is the average speed of the upstream section; \(X_3\) denotes the downstream congestion index which is calculated as follows:

\[
\text{Congestion Index (CI)} = \begin{cases} 
\frac{\text{free flow speed} - \text{actual speed}}{\text{free flow speed}} & \text{when } CI > 0; \\
0 & \text{when } CI \leq 0
\end{cases}
\]

(17)
4.4.2.2 Arterials crash-risk indicator

For urban arterials and collectors, the random parameter conditional logistic model developed by (Yuan et al., 2018b) was chosen to evaluate the crash-risk of arterial segments. The odds ratio of crash occurrence of the i\textsuperscript{th} observation could be expressed as:

$$odds\ ratio_i = \exp[-0.027(X_1 - \bar{X}_1)]$$  \hspace{1cm} (18)

Where odds ratio\(_i\) is the predicted odds ratio of the i\textsuperscript{th} observation which is a crash relative to the other non-crash events; \(X_1\) is the average speed of the current segment; \(\bar{X}_1\) is the mean value of the average speed of the current segment during non-crash events. The predicted odds ratio may be larger than 1 and in order to be consistent with the probability of crash occurrence, all the odds ratios predicted were normalized by using min-max normalization.

4.5 Results and Discussion

The application of CAVs along with the signal timing optimization was written in python script which controls the simulation in real-time through the application programming interface (API) of Aimsun Next. Also, CAVs with varying market penetration rates e.g., 20\%, 40\%, 60\%, 80\%, and 100\% (presented in the study as CAV20, CAV40, CAV60, CAV80, and CA100 respectively) were compared with the base condition. In order to capture the randomness effect of the simulation, each scenario (e.g., base condition, CAV conditions with different market penetration) was run repeatedly 5 times for different random seeds and the average result (from 5 replications) was used for the final evaluation. Since the network is too big, the research team presented different results for different roadways although the improvement is uniform throughout the network.
4.5.1 Traffic efficiency

In order to analyze the traffic efficiency for the full network, the research team collected travel time data for each segment of a roadway which was further divided by the length of that segment to get travel time rate (second/mile). Afterwards average TTR was calculated for each time interval for the whole network which is shown in Figure 4.3(a). The horizontal axis shows the time interval from 7:00 to 9:00 A.M. while the vertical axis (left and right) indicates the value of TTR and percentage reduction, respectively. The line in Figure 4.3(a) represents the average TTR for a particular scenario for each time interval. For instance, the blue line indicates the TTR for the base scenario while the red line denotes TTR for CAV20 scenario. Therefore, from the figure, it could be seen that the value of travel time rate is lower for the CAV scenarios compared with the base scenario. Also, the value of TTR is decreasing with the increase of CAV market penetration rate and the lowest value was obtained for 100% market penetration of CAVs (Figure 4.3(a)). In addition, the reduction percentage of TTR was calculated by comparing the base condition with the CAV condition for different levels of market penetration which is shown in Figure 4.3(a) by using the column diagram. The value of negative percentage indicates that the application of CAVs could reduce TTR compared with the base condition and vice versa. From the figure, it is worth to mention that for a penetration level of 20% of CAVs, the TTR reduction ranged from 7% to 28%. Also, Figure 4.3(a) illustrates that the overall reduction percentage ranged from 7% to 45% for 5-minute aggregation intervals for the full network while the maximum reduction was obtained for the CAV100 scenario. Similarly, the TTR for freeways and arterials are analyzed separately for base condition and CAV conditions with varying market penetration levels which is shown in Figure 4.3(b) and Figure 4.3(c), respectively. The percentage reduction for freeways and arterials range from 20% to 63% and 5% to 45%, respectively. Hence, the results
indicate that the TTR reduction is much higher for the freeways compared with the arterials because of the uninterrupted flow of traffic. In order to test the statistical significance of TTR among all scenarios, Kruskal–Wallis test was conducted where the result showed that the value of TTR is significantly different among base condition and all CAV conditions for the full network as well as freeways and arterials.

For the detailed analysis, average travel time was collected for the I-4 eastbound direction for each segment for every 5 minutes interval which was further divided by the length of the segment to obtain TTR. Figure 4.3(d) illustrates the results of TTR for the base condition along with the CAV conditions which indicates the significant improvement of travel time rate compared for CAV conditions. In addition, the research team evaluated the percent reduction of TTR for the CAV conditions compared with the base condition which is shown in the vertical line of Figure 4.3(d). The negative value indicates TTR was reduced compared with the base condition and vice versa. From the figure, the maximum value of TTR for the base condition is about 300 seconds while for the CAV100 scenario, this value is about 60 seconds which means around 80% reduction of TTR. Moreover, for the low penetration level of CAVs (e.g., CAV20), the minimum reduction percentage of travel time is about 30%.
(a) Full network

(b) Freeways
(c) Arterials

(d) I-4 eastbound

Figure 4.3 Travel time rate (TTR) for the (a) full network, (b) freeways, (c) arterials and (d) I-4 eastbound
Along with the V2V communication, the research team also utilized V2I communication such as communication between vehicles and signals by using the aforementioned optimization algorithm in order to enhance the traffic operation by minimizing the approach delay of the intersections. Hence, the approach delays of each intersection (total of 254) for the base condition was compared with the CAV 100% market penetration condition to see the possible maximum benefit of V2I communication algorithm. A non-parametric (Wilcoxon signed-rank) paired test was conducted to see the significant improvement of the proposed algorithm. Figure 4.4 shows the map of the study area including all the roadways (black lines) and intersections (dots) where the green dot indicates the approach delay of that intersection has been reduced significantly compared with the base condition. On the other hand, red dot denotes approach delay has been increased but not significantly with respect to the base condition. From the analysis, it could be found that the approach delay has been reduced for around 94% of the total intersection under the full penetration level of CAVs. In addition, the average approach delays for all the scenarios (base conditions, CAV conditions) was calculated for the whole simulation period (2 hours). For base condition, the average approach delay was about 8400 seconds while 6400, 4800, 4700, 4400, and 4300 seconds were obtained for CAV20, CAV40, CAV60, CAV80, and CAV100 scenarios respectively. Thus, the results showed that the proposed algorithm could reduce approach delay significantly for all penetration rates of CAVs compared with the base condition.
Figure 4.4 Approach delay at intersections (base vs CAV100)
4.5.2 Traffic safety

Since travel time rate (TTR) under CAV conditions is less compared with the base scenario, average speed of the network (both freeways and arterials) might be higher for CAV scenarios. Therefore, the research team analyzed the standard deviation of speed as a surrogate measure of safety assessment since the higher value of standard deviation of speed might result in higher crash risk (Zheng et al., 2010). Hence, the standard deviation of speed for I-4 eastbound direction was analyzed for each segment separately for the whole 2 hours of simulation which is shown in Figure 4.5.
Figure 4.5 Standard deviation of speed on I-4 eastbound

In the figure, horizontal X-axis and Y-axis shows the time interval from 7:00 to 9:00 A.M. and segments of I-4 eastbound direction, respectively while the vertical axis shows the standard deviation of speed (km/h). The color map shows the value of standard deviation of speed (km/h).
with a range between 0 to 40 km/h. From the Figure 4.5, it could be observed that the standard deviation of speed for the base condition is higher (red peak) for most of the segments for each time interval compared with the all other CAV scenarios. Also, the variation of the value of standard deviation of speed between segments and time intervals is more uniform with the increase of MPRs of CAVs which means CAVs could reduce the crash risk.

In addition, aforementioned crash-risk model for the freeways was used to see the effect of CAVs compared with the base condition. The result of the crash-risk probability for all scenarios (base and CAV) for I-4 westbound direction is shown in Figure 4.6. In Figure 4.6, the horizontal and vertical axis indicates the time interval for 2 hours and segments of I-4 westbound direction, respectively. Also, the crash probability of base condition as well as CAV scenarios is presented in the vertical color map ranged from 0 to 1.

From the figure, it could be seen that the probability of crashes for most of the segments for each time step is very high for the base condition. On the other hand, the probability of crash-risk is less for all the CAV conditions. Also, the segments with high crash-risk in the base condition are turned into a low crash-risk with the increase of CAVs market penetration which means CAVs can improve traffic safety substantially. Also, the figure of CAV100 indicates that under full market penetration of connected and automated vehicles, the probability of crash-risk for all the segments is very small.
Figure 4.6 Crash risk on freeway (I-4 westbound)
For arterials safety evaluation, the study analyzed the standard deviation of speed which is shown in Figure 4.7. From the figure it could be seen that the standard deviation of speed is higher for the base condition compared to the CAV scenarios which implies that CAV could reduce the crash risk in the arterial segments. Also, minimum standard deviation of speed was observed for CAV100 scenario.

![Figure 4.7 Standard deviation of speed on Arterials](image)

Moreover, for analyzing the crash-risk in the arterial segments, aforementioned crash risk model for arterials was used to calculate the odds ratio for all the arterials individually and took the average crash risk for each time interval. Figure 4.8 shows the result of the odds ratio (normalized using min-max normalization) for the base scenario as well as all the CAV scenarios.
The result implies that for each time interval, the application of CAVs can improve traffic safety in the arterials for any percentage of connected and automated vehicle penetration level. For instance, 20% market penetration of CAVs can reduce average crash-risk by 14% compared with the base condition. Also, CAV with full market penetration level improves safety significantly more compared with the base condition.

Figure 4.8 Crash risk on arterials (full network)
4.6 Conclusions

It is expected that the introduction of CAVs will have a significant impact on the transportation system. In the foreseeable future, the market penetration rate of CAVs will increase which is expected to have a positive impact both on the traffic safety and mobility. Therefore, it is worthwhile to analyze the effect of CAVs in a large network under mixed traffic conditions (CAVs and Manual vehicles).

This study investigated the impact of CAVs by utilizing V2V and V2I technology at the network level which includes both freeways and arterials. Since the behaviors of manual vehicles and CAVs are different, two separate car-following models were used in order to approximate the manual vehicles and CAVs in the traffic simulation Aimsun. Also, a new signal control algorithm was proposed by using V2I communication technology to minimize the approach delay at intersections which gives priority to the CAVs vehicle strings. The algorithm can generate new signal timing by optimizing the CAV and manual vehicle trajectories within a certain range (DSRC) from the intersection. Moreover, the proposed algorithm provides the signal timing information e.g., current phase, green time, etc., to the CAVs through V2I technology which will help the vehicles to adjust their speed to pass through the intersection with reduced delay. For the study area, the research team selected the most important and busiest corridor (Orlando CBD area) in Orlando, Florida to evaluate how CAVs can perform in the complex traffic condition. The evaluation measures of the application of CAVs are divided into two parts: (1) traffic efficiency and (2) traffic safety. Travel time rate (TTR), and approach delay at the intersections were used for the performance measure of traffic efficiency. The results of the traffic efficiency showed that TTR for the full network was significantly reduced by a minimum of 7% when the market penetration rate of CAVs was 20%. Also, the minimum reduction percentage of TTR is 20% and
5% for the freeways and arterials, respectively for CAV20. Compared with the base condition, maximum reduction percentage of TTR was found for the full market penetration of CAVs for all the cases (full network, freeways, and arterials). Also, the result of the TTR for a specific roadway (I-4 eastbound) showed that the application of CAVs can reduce travel time substantially compared with the base condition where minimum reduction percentage was about 30% for the low level of penetration rate of CAV (CAV20). For the evaluation of the proposed signal control algorithm, the research team estimated the approach delay for all the intersections (total of 254) and found that approach delay was significantly reduced compared with the base condition for 94% of the total intersections under 100% market penetration of CAVs. In addition, the total average approach delay for the whole network is less for all the CAV scenarios (20%, 40%, 60%, 80%, and 100% penetration rate of CAVs) than the base condition.

For surrogate safety assessment, the research team analyzed the standard deviation of speed of I-4 eastbound direction where the result showed that standard deviation of speed is lower for all the CAV scenarios (e.g., CAV20, CAV40, etc.) compared with the based scenario. Also, the variation of the standard deviation of speed between segments (I-4 eastbound) and time intervals is more uniform when the penetration rates of CAVs is higher. Moreover, for the safety evaluation of the freeways and arterials, different real-time crash risk models were used. The result of the crash-risk model for the freeways showed that the probability of crash risk is higher for most of the segments in I-4 westbound direction under base condition. Also, the result implied that the application of CAVs can reduce the crash-risk for all the segments of I-4 westbound direction and this risk is close to zero for CAV100 scenarios. Moreover, the result of the crash-risk model for arterials implied that 20% market penetration of CAVs can reduce the average crash risk by 14% compared with the base condition for all the arterials in the network.
Although this study used different car-following models to depict the manual vehicles and CAVs, it is hard to mimic the actual field behaviors in the simulation environment. Also, this study assumed a constant uninterrupted communication between two CAVs within a DSRC range which might not be true in the real field condition. Thus, there is still a scope to improve the performance of CAVs by proposing new car-following model and signal control algorithm.
CHAPTER 5: ASSESSING THE BENEFITS OF CONNECTED AND AUTOMATED VEHICLES OVER ONLY AUTOMATED VEHICLES ON A CONGESTED HIGHWAY BY CONSIDERING MULTI-VEHICLE COMMUNICATION SYSTEM

(Rahman, M. H., M. Abdel-Aty, Yina Wu, “A multi-vehicle communication system to assess the safety and mobility of connected and automated vehicles” under review at Transportation Research Part C, 2020.)

5.1 Introduction

The recent development in communication technologies facilitates the deployment of connected and automated vehicle (CAV) which has been considered to significantly improve traffic safety and mobility in the transportation network. CAV technology has the potential to largely impact our economy by improving traffic safety as well as reducing traffic congestion (Fagnant and Kockelman, 2015) which has attracted government agencies, vehicle manufacturers, policymakers, and academia to focus more resources and investigation in this field. According to several previous studies (Singh, 2015; Yue et al., 2018), most of the traffic crashes occur due to the human drivers’ errors which could be avoided by 90% with the application of CAV at high market penetration rates (Fagnant and Kockelman, 2015). However, this estimation is not quantitively confirmed yet.

Since the real-world connected and automated vehicles data is not currently available as of today to analyze the effect of the safety and mobility benefits, there are several studies (Kesting et al., 2010b; Milanés and Shladover, 2014; Rahman et al., 2019; Talebpour and Mahmassani, 2016) used microsimulation-based study to measure this effectiveness. Moreover, a recent study (Kalra and Paddock, 2016) found that it would take several years to evaluate the benefits of CAVs using
real-world data, because millions of miles of real-world CAV operational data are needed. Therefore, in the meanwhile, microsimulation-based studies would provide a good representation of the benefits of CAV in the transportation network.

Since the driving behavior of connected and automated vehicles (CAVs) are significantly different from human driven vehicles, the successful representation of these in the microsimulation environment is a critical task. Although several previous studies (Milanés and Shladover, 2014; Rahman and Abdel-Aty, 2018a; Shladover et al., 2012; Van Arem et al., 2006; Watzenig and Horn, 2017) used different car-following models in order to approximate the behavior of connected and automated vehicles in the microsimulation environment, they only utilized car-following models which are limited to the communication between two vehicles (following vehicle and the immediate leading vehicle). However, in a real-world scenario, a connected and automated vehicle can communicate with all the preceding vehicles equipped with connectivity features within the specified communication range. None of the previous studies considered multiple vehicles communication in the car-following model to approximate the CAVs driving behavior. Hence, this study incorporated a longitudinal car-following model to mimic the real-world CAVs driving behavior which include multiple preceding vehicle information (such as speed, position, acceleration, etc.) within the DSRC (dedicated short-range communication) range. Moreover, we used a separate car-following model in order to model the behavior of automated vehicles (AV) and compared it with the CAV vehicles to evaluate the benefits of connectivity. Several recent studies (Eichelberger and McCartt, 2016; Reagan et al., 2018; Yue et al., 2019) found that around 40-60% people used their automated features while driving on the highway. Therefore, we added connectivity features for both who “turned on” and “turned off” the automated features and compared to the only automated vehicles (AVs) scenario.
Finally, this study incorporates three different car-following models in order to approximate the behavior of AVs, CVs, and CAVs in the Aimsun Next microsimulation environment by using application programming interface (API). For the deployment of connected and automated vehicles, we calibrated and validated one of the major freeways (SR417) in Orlando, Florida. Afterward, we analyzed the benefits of connected and automated vehicles over only automated vehicles in terms of both traffic efficiency and safety. For the traffic efficiency analysis, we used travel time while for the safety evaluation, traffic conflicts were calculated by using the Surrogate Safety Assessment Model (SSAM) (Gettman et al., 2008). Additionally, we separated the traffic conflicts based on different vehicles types to see the effect of driving behavior of CAVs in mixed traffic flow (CAVs and human driven vehicles (HDVs)). Moreover, we utilized a mixed penetration of CVs and CAVs to compare with the only automated vehicles (AVs) in order to estimate the benefits of connectivity features under the real-world traffic condition. To the best of our knowledge, this is the first study that evaluates the benefits of connected and automated vehicles on the freeway segments compared to the scenario of only automated vehicles. To be more specific, the research aims to contribute by identifying the benefits of connectivity by utilizing multiple preceding vehicles in connected and automated vehicles in terms of both traffic efficiency and safety. Since the full penetration of CAVs will not be available in the near future, this study incorporated different market penetration rates (MPRs) to identify the possible benefits of CAVs.

5.2 Simulation Setup and Data

In this study, the research team developed a dynamic traffic assignment (DTA) based simulation model in the microsimulation platform. The study area covered the major freeway SR417 in Florida which is known as Central Florida GreeneWay (Figure 5.1). The testbed is
around 14 miles long in each direction (eastbound (EB) and westbound (WB)) from Lake Underhill Road to SR434. Since the selected area is very close to the downtown area (Orlando Central Business District) and major freeway SR408, it is one of the most congested segments of SR417. There are around 6 on-ramps and off-ramps in each direction of the roadway. The research team simulated the model for both directions (EB, WB) of SR417. The DTA based simulation model was developed in the Aimsun Next microsimulation platform where the study period was in the morning peak from 7:00 to 9:00 A.M. on October 11, 2017. For the calibration and validation of the DTA based simulation model, the research team collected the traffic information such as volume, travel time from multiple sources such as MVDS (Microwave Vehicle Detection System), AVI (Automatic Vehicle Identification), and NPMRDS (National Performance Measure Research Data Set). In this study, two types of vehicle compositions (i.e., cars and trucks) were used which were based on the MVDS detectors data. MVDS detectors classify the vehicles into four types based on their length, type 1: vehicles 0 to 10 ft. in length, type 2: vehicles 10–24 ft. in length, type 3: vehicles 24–54 ft. in length, type 4: over 54 ft. in length. Based on the information, the cars are defined when the length is less than or equal to 24 ft. (type 1 and type 2) (Shi, 2014; Wang, 2016) and trucks are classified when the length is greater than 24 ft. (type 3 and type 4). Hence, the overall percentage of cars and trucks are 90% and 10%, respectively. For the calibration and validation, the research team used around 14 and 4 detectors for volume and travel time, respectively. Figure 5.1 shows the testbed (highlighted in blue) and the locations of volume detectors (green dots) as well as travel time detectors (highlighted in purple). Although the number of travel time detectors is few, they covered almost half of the total length of the testbed which was considered good enough for the validation. Since the testbed is reasonably large, we found
that the optimal warm-up time was around 30 minutes to load the network with enough vehicles. For the detailed analysis, no statistics were collected during the warm-up period.

Figure 5.1 Study Area
5.2.1 Model Calibration and Validation

The most important part of a simulation model is to calibrate and validate the model with the real-field data (i.e., traffic volume, travel time, speed, etc.) so that there is minimum difference between simulated and real-field conditions. In this paper, a DTA based simulation model was calibrated and validated by traffic volume and travel time based on real-field data aggregated into 15 minutes intervals. Hence, for the calibration, we used Geoffrey E. Heavers (GEH) statistics which is a modified Chi-square test that utilized both relative and absolute difference to compare between the real-world and simulated data. The objective function for the calibration is to maximize the number of observations for which the GEH value should be less than the specified criteria used in the previous studies (Chung et al., 2020; Li et al., 2020; M. H. Rahman et al., 2019; Zhang et al., n.d.).

The GEH is defined as follows:

\[
GEH = \sqrt{\frac{2(S - R)^2}{S + R}}
\]  

(19)

Where S is the hourly traffic volume from the simulation and R is the hourly traffic volume from the real field. According to the previous studies, GEH value should be less than or equal to 5 for at least 85% of the total observations.

For the travel time validation, Florida Department of Transportation (FDOT) provides a guideline which implies that the difference between simulated and field travel time should be within ±1 minute for routes with observed travel time less than seven minutes or the difference should be within ±15% for routes with observed travel time greater than seven minutes (FDOT Systems
Planning Office, 2014). Hence, the simulated travel time should be within the above criteria for at least 85% of the total observations.

Aimsun Next microsimulation environment uses Gipps car-following model to approximate the longitudinal behavior of vehicles which has different car-following parameters to change the driving behavior of vehicles. Hence, to get a good calibration and validation results, the research team tweaked the value of different car-following parameters within a certain range. The parameters are divided into two parts: (1) dynamic traffic assignment or route choice and (2) microscopic parameters. Also, a sensitivity analysis was utilized for both traffic assignment and microscopic parameters. After tweaking different parameters, the default values and the calibrated values for both traffic assignment and microscopic parameters are shown in Table 5.1.

From the calibration results, we have found that in around 86% of the cases, the GEH value is less than 5 which represents good calibration for the microscopic network. The validation results also implied that around 86% of the cases the travel time value satisfies the above-mentioned criteria. Hence, the model can represent the real-world traffic condition which is considered as base condition in this study.
### Table 5.1 Calibration parameters in Aimsun Next

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Unit</th>
<th>Default value</th>
<th>Calibrated value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Traffic assignment parameter</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model selection</td>
<td>N/A</td>
<td>uniform</td>
<td>C-logit</td>
</tr>
<tr>
<td>Attractiveness weight</td>
<td>N/A</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Maximum number of initial paths to consider</td>
<td>N/A</td>
<td>All</td>
<td>3</td>
</tr>
<tr>
<td>Maximum Paths per interval</td>
<td>N/A</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>En-route</td>
<td>N/A</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Path Calculation</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>C-logit parameters (scale, beta, gamma)</td>
<td>N/A</td>
<td>(1,0.15,1)</td>
<td>(1,0.15,1)</td>
</tr>
<tr>
<td><strong>Microscopic parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reaction time</td>
<td>s</td>
<td>1.1</td>
<td>0.90</td>
</tr>
<tr>
<td>Reaction time at traffic light</td>
<td>s</td>
<td>1.6</td>
<td>1.2</td>
</tr>
<tr>
<td>Look ahead distance variability</td>
<td>%</td>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>Speed acceptance for Car</td>
<td>N/A</td>
<td>1.1</td>
<td>1.5</td>
</tr>
<tr>
<td>Speed acceptance for Truck</td>
<td>N/A</td>
<td>1.0</td>
<td>1.4</td>
</tr>
</tbody>
</table>

### 5.3 Proposed Methodology

Modeling the car-following behavior in the simulation environment is one of the important components which requires tweaking of different car-following parameters to represent the exact behavior of vehicles. Therefore, the proposed methodologies of this study are mainly focused on modeling the longitudinal car-following behavior of automated vehicles (AVs), connected
vehicles (CVs), and connected and automated vehicles (CAVs) in the Aimsun simulation environment. It is obvious that the behavior of human driven vehicles (HDVs), AVs, CVs and CAVs should be significantly different from each other. At the beginning of this section, the behavior of human driven vehicles (HDVs) were discussed. Afterward, we explain the realistic car-following behavior of AVs, CVs, and CAVs in accordance with the recent literature that would approximate the exact representation of these technologies.

5.3.1 Modeling human driven vehicles (HDVs)

The behavior of manual vehicles is significantly different from connected and automated vehicles in many different ways such as reaction time, headway, acceleration, gap, etc. In this study, the Gipps car-following model was used to approximate the behavior of human driven vehicles (HDVs) based on many previous studies (Gipps, 1981; Yu et al., 2019). This model has two components (acceleration and deceleration) to control the longitudinal behavior where the acceleration component is used to reach a vehicle at a certain speed while the deceleration component provides restriction of the desired speed based on the preceding vehicle. The speed of the vehicle of the acceleration part at time \( (t + \Delta t) \) is expressed as follows:

\[
V_a(n, t + \Delta t) = V(n, t) + 2.5 \, a(n) \Delta t \left(1 - \frac{V(n, \Delta t)}{V^*(n)}\right) \sqrt{0.025 + \frac{V(n, t)}{V^*(n)}} \tag{20}
\]

where

- \( V(n, t) \) is the speed of the \( n^{th} \) vehicle at time \( t \),
- \( \Delta t \) is the reaction time,
- \( a(n) \) denotes the maximum acceleration for vehicle \( n \),
$V^*(n)$ represents desired speed of $n^{th}$ vehicle.

The deceleration part of the model is formulated as follows:

$$V_d(n, t + \Delta t) = d(n)\Delta t + \sqrt{d(n)^2\Delta t^2 - d(n)[2 * s(n, t) - V(n, t)\Delta t - \frac{V(n - 1, t)^2}{d(n - 1)}} \quad (21)$$

where

$d(n)$ denotes the maximum deceleration of $n^{th}$ vehicle,

$s(n, t)$ is the gap of $n^{th}$ vehicle.

The speed of the vehicle $n$ at time $(t + \Delta t)$ is estimated as follows:

$$V(n, t + \Delta t) = \min [V_a(n, t + \Delta t), V_d(n, t + \Delta t)] \quad (22)$$

Finally, the speed of the vehicle at time $(t + \Delta t)$ is determined through equation 22. Moreover, in order to approximate the real traffic behavior, we tweaked some car-following parameters in the simulation environment which are mentioned in the calibration and validation section of this paper.

5.3.2 Modeling Automated vehicles (AVs)

In order to model the car-following behavior of automated vehicles (AVs), we used the widely accepted intelligent driver model (IDM), introduced by Treiber et al., (2000). IDM is the simplest model among others which has the ability to model oscillations, stop and go traffic by providing realistic acceleration and deceleration profile (Kesting et al., 2008; Rahman and Abdel-Aty,
Also, a previous study (Pourabdollah et al., 2017) showed that the IDM model can provide better driving behavior compared to the other car-following models. The basic IDM is a non-linear car-following model where acceleration is calculated based on the speed difference $\Delta v$ and the dynamic desired gap distance $s^*$ between the following and the leading vehicle. The acceleration equation for IDM is expressed as:

$$a(t + \Delta t) = a_{\text{max}} \left[ 1 - \left( \frac{V_n}{V_0} \right)^\delta - \left( \frac{S^*(V_n, \Delta V)}{S_n} \right)^2 \right]$$

(23)

$$S^*(V_n, \Delta V) = S_0 + V_n T + \frac{V_n \Delta V}{2 \sqrt{a_{\text{max}} b}}$$

where $\Delta t$ = the perception-reaction time, $a_{\text{max}}$ = the maximum acceleration, $V_n$ = the speed of the following vehicle, $V_0$ = the desired speed, $\delta$ = the acceleration exponent, $S_n$ = the gap distance between two vehicles, $S_0$ = the minimum gap distance at standstill, $T$ = the safe time headway and $b$ = the desired deceleration.

Since the parameters of the IDM model for automated vehicles (AVs) should be significantly different from HDVs, many previous studies (Kesting et al., 2010b; Li et al., 2017a; Milanés and Shladover, 2014) used different values for the model parameters in order to approximate the behavior of AVs. The parameters of AVs behavior model are presented in the following Table 5.2.
<table>
<thead>
<tr>
<th>Model parameters</th>
<th>Automated vehicles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desired speed,</td>
<td>120 km/h</td>
</tr>
<tr>
<td>Acceleration exponent</td>
<td>4</td>
</tr>
<tr>
<td>Maximum acceleration</td>
<td>1.4 m/s²</td>
</tr>
<tr>
<td>Desired deceleration</td>
<td>2 m/s²</td>
</tr>
<tr>
<td>Minimum gap distance at standstill</td>
<td>2 m</td>
</tr>
<tr>
<td>Safe time headway</td>
<td>0.6 s</td>
</tr>
<tr>
<td>Maximum deceleration</td>
<td>2.8 m/s²</td>
</tr>
<tr>
<td>Time delay</td>
<td>1.5 s</td>
</tr>
</tbody>
</table>

5.3.3 Modeling Connected vehicles (CVs) and Connected & Automated vehicles (CAVs)

There is a significant difference between connectivity and automated technologies where connectivity system enables vehicles to get information such as speed, acceleration, position, traffic incident, etc., from the nearby vehicles or roadway infrastructure within the communication range while automated features help vehicles to drive by itself based on the information received. In this study, we modeled the connected and automated vehicles in two different ways by utilizing the automated features “turned off” and “turned on”. Also, the automated technology used in this study represents longitudinal control of the low level of automation.

In the previous studies (M.S. Rahman et al., 2019; Rahman and Abdel-Aty, 2018a; Talebpour and Mahmassani, 2016), the researchers used the IDM model by considering the speed and acceleration behavior of the immediate leading vehicle only to approximate the car-following model. However,
connectivity enables vehicles to communicate with all the preceding vehicles within the communication range. Therefore, the assumptions of the previous studies have some limitations. Also, the IDM model was originally modified for the adaptive cruise control (ACC) system (Kesting et al., 2008) in the previous research. In order to evaluate the effectiveness of CAVs and to avoid the confusion about the connectivity and automated technologies, this study introduces different approaches to model the connected and automated vehicles technologies in the simulation environment.

For modeling the behavior of CVs, we assumed that a vehicle can communicate with multiple preceding vehicles within the DSRC (dedicated short-range communication) range and get the traffic information such as speed, acceleration behavior, position, etc. In this case, we considered a vehicle that can communicate with maximum of five preceding vehicles which is based on the previous studies (Lioris et al., 2017; Zhao and Sun, 2013) on platoon size used for connected vehicles. Hence, a modified version of IDM model (Sun et al., 2018) was utilized which incorporates multiple preceding vehicles information. In terms of driving behavior, the main difference between CVs and CAVs is that CVs are driven by human drivers while CAVs are driven by equipped automated technologies. In order to get the maximum benefits of connected vehicles (CVs), the drivers should comply with the transmitted information. Therefore, for modeling the drivers’ compliance in CV environment, a compliance factor known as compliance utility (UT) was used in the modified IDM equation based on Sharma et al. (Sharma et al., 2017). Hence, we assumed that drivers will comply with all the transmitted information in a particular manner which heavily relies on the headway of the vehicle. The compliance factor is modeled based on the well-known decision-making theory named as prospect theory. The mathematical formulation of the UT is given below:
Usefulness value function

\[ V(\text{obs}) = \frac{1}{1 + e^{\lambda(\alpha h_{\text{obs}} - 1)}} \]

Weighting function

\[ W^{LC}(P_{LC}) = \frac{P_{LC}^{\gamma}}{(P_{LC}^{\gamma} + (1 - P_{LC})^{\gamma})^{1/\gamma}} \]

\[ P_{LC} = \min\left(\frac{h_{\text{obs}}}{h_{\text{max}}}, 1\right) \]

\[ W^{HC}(P_{HC}) = \frac{P_{HC}^{\gamma}}{(P_{HC}^{\gamma} + (1 - P_{HC})^{\gamma})^{1/\gamma}} \]

\[ P_{HC} = \min\left(\frac{h_{\text{min}}}{h_{\text{obs}}}, 1\right) \]

Compliance utility function

\[ UT = \max\left\{ UT^{LC} = V(\text{obs})W^{LC}(P_{LC}) \right\} \]

\[ UT^{HC} = V(\text{obs})W^{HC}(P_{HC}) \]

Where, \( \lambda \) and \( \alpha \) are the shape parameter of the usefulness value function, and \( h_{\text{obs}} \) is the observed headway. \( P_{LC}, P_{HC} \) are the low and high compliance range, respectively, and \( h_{\text{max}} \) is the maximum headway. Hence, in order to add the driver compliance in the modified IDM equation, the time gap parameter is multiplied with \((I + UT)\) in equation 25. Finally, the mathematical equation of the car-following model for CVs is shown in equation 25.

\[ a(t + \tau) = a_{\text{max}}[1 - \left(\frac{V_n}{V_0}\right)^\delta - \left(\frac{S^*(V_n, \Delta V)}{\sum_{j=1}^{m} y_m S_{n-m+1}}\right)^2] \]

\[ S^*(V_n, \Delta V) = S_0 + V_n T(1 + UT(h_{\text{obs}})) + \frac{V_n \sum_{j=1}^{m} y_m \Delta V_{n-m+1}}{2 \sqrt{a_{\text{max}} b}} \]

\[ \Delta V_{n-m+1} = V_n - V_{n-m+1} \]

\[ S_{n-m+1} = x_n - x_{n-m+1} - l_{\text{veh}} \]
\[
\sum_{j=1}^{m} \gamma_m \leq 1.
\]

\[
a(t + \tau) = a_{\text{max}}[1 - \left(\frac{V_n}{V_0}\right)^{\delta} - \left(\frac{S^*(V_n, \Delta V)}{\sum_{j=1}^{m} \gamma_m S_{n-m+1}}\right)^2]
\]

\[
S^*(V_n, \Delta V) = S_0 + V_n T + \frac{V_n \sum_{j=1}^{m} \gamma_m \Delta V_{n-m+1}}{2 \sqrt{a_{\text{max}} b}}
\]

\[
\Delta V_{n-m+1} = V_n - V_{n-m+1}
\]

\[
S_{n-m+1} = x_n - x_{n-m+1} - l_{\text{veh}}
\]

\[
\sum_{j=1}^{m} \gamma_m \leq 1.
\]

In the regular IDM model, \(\Delta V\) is the speed difference between following vehicle and the immediate leading vehicle where \(\Delta V_{n-m+1}\) indicates the relative speed based on all the leading vehicles within the communication range. \(S_{n-m+1}\) defines the bumper to bumper gap between following vehicle and all the leading vehicles. Also, we provided a weight factor \(\gamma\) for different leading vehicles based on the gap distance between the following and leading vehicle. For example, \(\gamma_1\) and \(\gamma_2\) are the weight factor for the first and second leading vehicle respectively where \(\gamma_1 > \gamma_2 > \ldots > \gamma_m\), and \(\sum_{j=1}^{m} \gamma_m \leq 1\).

Along with the utility function for driver compliance, we used the same values of reaction time, headway, maximum desired speed, etc., that are used for the HDVs for modeling the driving
behavior of CVs and all these parameters were calibrated and validated based on the real-world traffic conditions. On the other hand, CAVs driving behavior include both connectivity and automated features where connectivity technologies enable vehicle to communicate with multiple preceding vehicles within the communication range which is similar to the aforementioned model used for CVs without the driver’s compliance factor. Instead of the driver’s compliance factor, automated features were used for modeling CAVs which is based on the same model parameters used for AVs. Therefore, the automated modeling parameters for CAVs are same as mentioned in Table 5.2. The mathematical formulation of CAV car-following model is given in equation 26.

A graphical representation of driving behavior of AVs, CVs, and CAVs is illustrated in Figure 5.2.

Figure 5.2 Illustration of AV, CV, and CAV communication system
5.4 Performance Evaluation Measures

For the evaluation measures, the research team analyzed the benefits of AVs, CVs, and CAVs in terms of both traffic safety and efficiency. Travel data was collected for the evaluation of traffic efficiency while for safety assessment, Surrogate Safety Assessment Model (SSAM) was employed.

5.4.1 Traffic Efficiency

Traffic data such as travel time was collected from Aimsun Next directly for every 5 minutes interval for each roadway segment present in the simulation network for each scenario e.g., base, AVs, CAVs. At first, a statistical test was conducted to verify whether travel time is significantly different between each scenario (base, AVs, CAVs, etc.) with different market penetration rates (MPRs) by conducting one-way ANOVA.

Afterward, we developed a generalized estimating equation (GEE) model to analyze the impacts of AVs, CVs, and CAVs on travel time compared with the base scenario. GEEs provide an extension of generalized linear models (GLMs) which are usually used for repeated measurements studies. Also, GEEs are used for the analysis of clustered data to account for the spatial correlation of a given cluster. In this study, we considered each segment as a cluster where repeated measures are the travel time data for every 5 minutes interval for the whole two hours of simulation. Suppose, the travel time data frequency for a cluster $i$ at segment $j$ is $y_{ij}$ where $i=1,2,\ldots, K$ and $j=1,2,\ldots, n_i$ where $K$ is the total number of cluster and $n_i$ is the number of segment in cluster $i$. Therefore, total number of segments is denoted by $\sum_{i=1}^{k} n_i$. Let the vector of measurements on the $i$th subject be $Y_i = [y_{i1}, \ldots, y_{in_i}]'$ with corresponding vector of means, $\mu_i = [\mu_{i1}, \ldots, \mu_{in_i}]'$ and let $V_i$ be the covariance matrix of $Y_i$. Let the vector of independent, or explanatory, variables for the $j$th measurement on the $i$th subject be $x_{ij} = [x_{ij1}, \ldots, x_{ijp}]'$. The
generalized estimating equation for estimating \( p \times 1 \) vector of regression parameters \( \beta \) (Liang and Zeger, 1986) is given by

\[
S(\beta) = \sum_{i=1}^{k} f_i D_i' V_i^{-1} (Y_i - \mu_i(\beta)) = 0
\]

(27)

Where, \( D_i = \frac{\partial \mu_i}{\partial \beta} \). Since \( g(\mu_{ij}) = x_{ij}' \beta \) where \( g \) is the link function and \( p \times n_i \) matrix of partial derivative of the mean with respect to the regression parameters for the \( i \)th cluster which is given by

\[
D_i' = \frac{\partial \mu_i'}{\partial \beta} = \begin{bmatrix}
    x_{i11} & \cdots & x_{i1n_i} \\
    g'(\mu_{i1}) & \cdots & g'(\mu_{i1}) \\
    \vdots & \ddots & \vdots \\
    x_{i1p} & \cdots & x_{ipn_i} \\
    g'(\mu_{i1}) & \cdots & g'(\mu_{i1}) 
\end{bmatrix}
\]

(28)

The covariance matrix \( Y_i \) is modeled as \( V_i = \Phi A_l^{0.5} W_l^{-0.5} R(\alpha) W_l^{-0.5} A_l^{0.5} \) where \( A_l^{0.5} \) is an \( n_i \times n_i \) diagonal matrix with \( v(\mu_{ij}) \) as the \( j \)th diagonal element and \( W_l^{-0.5} \) is an \( n_i \times n_i \) diagonal matrix with \( w_{ij} \) as the \( j \)th diagonal where \( w_{ij} \) is a weight function. \( R(\alpha) \) is an \( n_i \times n_i \) working correlation matrix which is specified by the vector parameter \( \alpha \).

5.4.1.1 Independent \( R \)

The assumption behind the independence correlation matrix is that the travel time at different segments within a cluster are independent. The working correlation is not estimated in this case. Hence, the estimation result of GEE is similar to the GLM in this case, but the standard
errors are difference since GEE method accounts for the correlation by operating at the cluster level.

\[
corr(y_{ij}, y_{ik}) = \begin{cases} 
1 & j = k \\
0 & j \neq k
\end{cases} \quad \text{e.g., } R_{3\times3} = I_{3\times3} = \begin{bmatrix} 1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \end{bmatrix} \tag{29}
\]

5.4.1.2 Exchangeable \( R \)

In the exchangeable working correlation structure, a constant correlation is made between any two observations within a cluster.

\[
corr(y_{ij}, y_{ik}) = \begin{cases} 
1 & j = k \\
\alpha & j \neq k
\end{cases} \quad \text{e.g., } R_{3\times3} = I_{3\times3} = \begin{bmatrix} 1 & \alpha & \alpha \\
\alpha & 1 & \alpha \\
\alpha & \alpha & 1 \end{bmatrix}
\]

Where \( \hat{\alpha} = \frac{1}{(N' - p)\emptyset} \sum_{i=1}^{k} \sum_{j<k} e_{ij}e_{ik} \) and

\[
N' = 0.5 \sum_{i=1}^{k} n_i (n_i - 1) \tag{30}
\]

The dispersion parameter \( \emptyset \) is estimated by

\[
\emptyset = \frac{1}{N - p} \sum_{i=1}^{k} \sum_{j=1}^{n_i} e_{ij}^2
\]

5.4.2 Traffic Safety

For the safety evaluation, Surrogate Safety Assessment Model (SSAM) was utilized which calculates traffic conflicts from trajectory files generated in the simulation environment. In this
study, time-to-collision (TTC) was used for the traffic conflict analysis which is defined as the
time required to collide between two vehicles (leader and follower) if they travel on the same path
and same speed. If the value of TTC is less than the specified threshold (1.5s in this study), SSAM
considers it as a conflict. Hence, we aggregated the number of traffic conflicts for each segment
for each time interval (5 minutes) from 7:00 to 9:00 A.M.

In order to model the traffic conflicts based on the value of TTC, we used a Bayesian
hierarchical zero-inflated negative binomial model which takes into account different penetration
rates of AVs, CAVs, mixture of CV and CAVs, standard deviation of speed, and the logarithm of
flow. Data that contains higher number of zeros than expected for the underlying probability
distribution of counts can be modeled with a zero-inflated distribution. The probability distribution
of a zero-inflated negative binomial random variable Y is given by the following equation:

\[
Pr(Y = y) = \begin{cases} 
\omega + (1 - \omega)(1 + k\alpha)^{-1/k} & \text{for } y = 0 \\
\frac{1}{(1 - \omega) \Gamma(y + 1/k)} (k\alpha)^y (1 + k\alpha)^{-1} & \text{for } y > 0 
\end{cases}
\]

Where \(\omega\) is the zero-inflation probability, \(\alpha\) is the count mean, and \(k\) is the negative
binomial dispersion parameter. The variance of the negative binomial model is \(\alpha + k\alpha^2\).

The formulation of Bayesian zero-inflated negative binomial model is showed in the
equation below:

\[
\log(\mu_i) = (\beta_0 + \beta_1 x_i) + UH_i
\]
Where $UH_i$ is the unobserved heterogeneity, $\beta_o$ is the constant term and $\beta_i$ is the coefficient of the corresponding independent variable $X_i$ which follow non-informative normal distribution with zero mean. The Bayesian hierarchical model can be estimated by employing the Markov chain Monte Carlo method. The goodness of fit was evaluated by using deviance information criterion (DIC) which is expressed by the following equation:

$$DIC = D(\bar{\phi}) + 2p_D = \bar{D} + p_D$$

(33)

Where $D(\bar{\phi})$ is the deviance of the $\phi$ posterior means of the model, $p_D$ is the effective number of parameters, $\bar{D}$ is the posterior mean of the deviance $D(\bar{\phi})$. The smaller value of the DIC with least explanatory variables indicates the good fit of the model (Russell and Gray, 2013).

5.4.2.1 Time exposed time-to-collision (TET)

Since SSAM cannot provide the vehicle type information in the conflict analysis, this study used loop detectors on the roadway in order to calculate the traffic conflicts based on different vehicle type. In this case, traffic conflicts were estimated based on time exposed time-to-collision (TET) which is derived from the time-to-collision (TTC) notion. Minderhoud et al., (2001) developed TET based on $TTC_{brake}$ and TTC* which is expressed as:
\[
TTC_{\text{brake}}(t) = \frac{y_{n-1}(t) - y_n(t) - L_{n-1}}{v_n(t)}
\]

\[
TET(t) = \sum_{n=1}^{N} \delta_t \times \Delta t, \quad \delta_t = \begin{cases} 
1, & 0 < TTC_{\text{brake}}(t) \leq TTC^* \\
0, & \text{otherwise}
\end{cases}
\]

\[
TET = \sum_{t=1}^{\text{Time}} TET(t)
\]

where \( t = \) time ID, \( n = \) vehicle ID, \( N = \) total number of vehicles, \( \delta_t = \) switching variable, \( \Delta t = \) time step, which was 0.1 s in simulation, \( \text{Time} = \) simulation period, and \( TTC^* \) = the threshold of TTC is 1.5 s in this study.

5.5 Results and Discussion

In this paper, the benefits of AVs, CVs, and CAVs were evaluated in terms of both traffic operation and safety. Hence, the results and discussion section are divided into two parts i.e., traffic efficiency and traffic safety. For modeling different car-following behavior, the research team used application programming interface (API) available in the Aimsun Next simulation platform. Since the full market penetration rates (MPRs) of connected and automated vehicles will not be available in the near future, it is worthwhile to study the effect of AVs, CAVs with low penetration rate. Hence, this study utilized different market penetration rates of AVs, CVs, and CAVs. The experimental design of this study is shown in Figure 5.3.
In this study, we used five different market penetration rates for both AVs and CAVs which are denoted as AV20, AV40, AV60, AV80, AV100, and CAV20, CAV40, CAV60, CAV80, CAV100, respectively. AV20 means 20% of the total vehicles are equipped with automated
technologies where 80% of them are human driven vehicles (HDVs). Similarly, AV40 means 40% AVs, 60% HDVs and so on. Similarly, CAV20 indicates 20% of total vehicles are CAVs and 80% of them are HDVs and so on.

According to the previous studies (Eichelberger and McCartt, 2016; Reagan et al., 2018; Yue et al., 2019), around 40-60% people use their automated features while driving on highway/expressway which means the rest of the people feel comfortable to drive by themselves rather than of using AV technologies. Hence, this study assumes 40% will use automated features in order to emulate the current usage of automated technologies. With this regard, the study added the connectivity technologies for both conditions i.e., automated features “turned on” and “turned off”. Hence, the connectivity features might improve both automated and self-driving experience. Since most of the traffic crashes occur due to human driver errors (Wu et al., 2019; Yue et al., 2018), only the connectivity features might not provide better results compared with the automated features. Therefore, a fixed penetration rate of CVs and varying penetration rate of CAVs was incorporated based on the above-mentioned studies to investigate the possible maximum benefits under the current traffic condition. In Figure 5.3, market penetration rate CV60CAV20 means 60% of the total vehicles are equipped with connectivity features (CVs) while 20% have both connectivity and automated feature (CAVs) and the rest 20% are driven by human drivers (HDVs). For capturing the randomness effect of the simulation, the research team ran each scenario (e.g., base, AV, CAV, etc.) 5 times with different random seeds and reported the average results for the final evaluation.
5.5.1 Traffic efficiency

In order to evaluate the benefits of traffic efficiency, the research team collected the travel time data for each segment for every 5 minutes interval from 7:00 to 9:00 AM which was further aggregated to get the travel time for the eastbound direction (EB) of SR417. Figure 5.4(a) illustrates the travel time for different MPRs of CAVs while Figure 5.4(b) shows the travel time for AVs. On the other hand, Figure 5.4(c) depicts the travel time results for varying penetration rates of CVs and CAVs. From all the Figures, it could be clearly seen that CAVs and AVs with different penetration rates could improve the travel time of the roadway compared to the base condition and the maximum benefit was found for the full penetration rate of CAVs and AVs. In order to analyze the statistical significance of mean travel time of different scenarios of CAVs, AVs, and mixture of CVs and CAVs, one-way ANOVA (analysis of variance) test was conducted which captures the significant differences between means of several groups.

The results of the ANOVA test show that F-value (157.28) for all the scenarios in CAVs along with the base condition is considerably higher than the critical F-value (2.279) at 95% confidence interval. Similarly, F-values for AVs, and mixture of CVs and CAVs are 181.47 (F-critical =2.279) and 104.84 (F-critical =2.703), respectively. Therefore, the results of the ANOVA test imply that the base scenario along with all the scenarios in CAVs, AVs, and mixture of CAVs and AVs are significantly different from each other.
a. CAV scenarios

b. AV scenarios
c. CV and CAV scenarios

Figure 5.4 Travel time of AV, CAV, and mixture of CV and CAV scenarios

Moreover, to test the significant difference between the same penetration level of CAVs and AVs (e.g., CAV20 vs AV20, CAV40 vs AV40), we have utilized the non-parametric paired test Wilcoxon signed-rank test since the data points is relatively small (less than 30). The results of the non-parametric test are shown in Table 5.3.
Table 5.3 Comparison of CAV and AV for the same MPR

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAV20 vs AV20</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CAV40 vs AV40</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CAV60 vs AV60</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CAV80 vs AV80</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CAV100 vs AV100</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

From Table 5.3, it could be seen that with the same market penetration rate, all the CAV scenarios are significantly different from AV scenarios which means connectivity features has significant effect on the travel time of a roadway.

5.5.1.1 GEE model estimation

In order to evaluate the benefits of connectivity and automation technologies based on traffic efficiency; we utilized generalized estimating equation (GEE) model where travel time was used as a dependent variable. For independent variables, different penetration rates of AVs, CAVs, mixture of CVs and CAVs, flow, upstream flow, downstream flow, length of a segment, and number of lane (nblane) were utilized. Traffic data such as travel time, flow, upstream flow, downstream flow was collected for each segment for every 5 minutes interval from 7:00 to 9:00 A.M. Segment length and number of lanes are fixed variable for a specific segment.

Since the objective of this study is to evaluate the benefits of connected and automated vehicle over the existing base condition, the base scenario was used as a reference to analyze the GEE model. Also, we estimated the GEE model for both independent and exchangeable correlation
structures. The results of the GEE model with an identity link for two correlation structures (independent, exchangeable) are reported in Table 5.4.

Table 5.4 GEE model estimation results

| Parameter          | Estimate | Standard Error | Pr > |Z| | Estimate | Standard Error | Pr > |Z| |
|--------------------|---------|----------------|------|------|---------|----------------|------|------|
| Intercept          | 0.857   | 2.529          | 0.735|      | 2.905   | 3.976          | 0.465|
| CAV20              | -0.770**| 0.234          | 0.001|      | -0.766**| 0.236          | 0.001|
| CAV40              | -1.653**| 0.468          | <.0001|     | -1.650**| 0.468          | <.0001|
| CAV60              | -2.416**| 0.601          | <.0001|     | -2.409**| 0.600          | <.0001|
| CAV80              | -3.588**| 0.888          | <.0001|     | -3.574**| 0.888          | <.0001|
| CAV100             | -6.046**| 1.468          | <.0001|     | -5.985**| 1.471          | <.0001|
| AV20               | -0.043  | 0.191          | 0.824|      | -0.037  | 0.191          | 0.846|
| AV40               | -0.924**| 0.332          | 0.005|      | -0.922**| 0.331          | 0.005|
| AV60               | -1.633**| 0.471          | 0.001|      | -1.629**| 0.471          | 0.001|
| AV80               | -2.791**| 0.727          | <.0001|     | -2.784**| 0.727          | <.0001|
| AV100              | -5.303**| 1.468          | <.0001|     | -5.274**| 1.468          | <.0001|
| CAV0CV60           | -1.018**| 0.321          | 0.002|      | -1.016**| 0.323          | 0.002|
| CAV20CV60          | -2.953**| 0.794          | <.0001|     | -2.940**| 0.798          | <.0001|
| CAV40CV60          | -5.235**| 1.322          | <.0001|     | -5.164**| 1.392          | <.0001|
| Flow               | 0.002   | 0.002          | 0.354|      | 0.003**| 0.002          | 0.041|
| Upstream_flow      | -0.001  | 0.001          | 0.721|      | -0.002* | 0.001          | 0.099|
| Downstream_flow    | 0.0003  | 0.001          | 0.728|      | 0.001   | 0.001          | 0.444|
| Segment length     | 0.036** | 0.002          | <.0001|     | 0.035** | 0.002          | <.0001|
| nblane             | -0.773  | 0.792          | 0.329|      | -1.736* | 0.961          | 0.071|

*Significance at the 0.1 level. **Significance at the 0.05 level.

In order to compare the model fit between independent and exchangeable correlation structures, we have used QIC (Quasilikelihood under the Independence model Criterion) statistic proposed by Pan (2001) (Pan, 2001). The smallest value of QIC/QICu indicates the better model fit (Cui, 2007). From

122
Table 5.4, it could be seen that the value of QIC/QICu is smaller for the exchangeable correlation structure compared with the independent structure which means the model fits better in exchangeable structure. Also, the estimated coefficients and standard errors are different for independent and exchangeable correlation structure. For example, upstream flow is not significant (p-value=0.721) in the independent structure while it is significant for the exchangeable structure (p-value=0.099) at 90% confidence interval. The estimated working correlation value in the exchangeable structure is 0.746.

From the value of the estimated coefficients in Table 5.4, it could be seen that different market penetration rates of CAV and AV could decrease the travel time compared to the base condition. Also, the travel time improvement is higher with the increase of market penetration rate. For instance, the estimated coefficient for CAV40 (-1.650) is lower than the CAV20 (-0.766) which means 40 percent penetration rate of CAVs could improve the travel time more than the 20 percent penetration rate compared to the base condition. The results are also similar for different penetration rates of AV. The maximum benefit of travel time reduction was found for the full penetration rate of CAV (CAV100) as well as the full penetration rate of AV (AV100) compared to the base scenario. Moreover, from the p-value of the GEE model it could be found that all the different market penetration rates of CAV are statistically significant at 95% confidence interval (p-value<0.05). On the other hand, varying market penetration rates of AV are statistically significant except 20 percent penetration rate which means travel time improvement for 20 percent penetration rate of AV is not statistically significant compared to the base condition. Since CAV20 is statistically significant, it could be stated that at low penetration rate, connected and automated vehicles have higher benefits compared with the only automated vehicles in terms of traffic efficiency.
Moreover, from the value of estimated coefficients (Table 5.4), it was found that with the same penetration level, connected and automated vehicles (CAVs) could improve the travel time more than the only automated vehicles (AVs) compared to the base condition. For example, at 80 percent penetration rate, the value of the coefficient is -3.574 for CAVs and -2.784 for AVs which implies that at 80 percent market penetration level, the travel time reduction compared to the base condition is higher for CAV than AV scenario. Hence, it could be stated that connectivity along with the automated features improve travel time more compared with the only automated feature.

Furthermore,

Table 5.4 shows that the value of the coefficient for CAV0CV60 is -1.016 which is less than AV60 (-1.629) but greater than AV40 (-0.922) indicating that 60 percent penetration rate of CV is better than 40 percent penetration rate of AV but not better than 60 percent MPR of AV. Also, the value of the estimated coefficient of CAV20CV60 (-2.940) implies that when we added 20% connected and automated vehicles with the 60% connected vehicles, the travel time improvement is almost similar to the 80% penetration of AVs (-2.784). Similarly, travel time improvement for CAV40CV60 compared to the base scenario is almost identical to the AV100. Hence, we could say that connectivity technologies have a significant effect on travel time improvement.

In addition, flow was found to be a significant effect on the travel time of a segment. The higher traffic flow increases in the value of travel time of a segment. On the other hand, the higher the value of the upstream flow, the travel time value would be lower. Also, if the length of a segment is longer, then the value of travel time would be higher compared with the smaller segment. Moreover, we found that number of lanes have a significant effect on travel time. For
example, the travel time of two-lane highway with equal distance is higher compared with the three-lane highway for given traffic flow.

5.5.2 Traffic Safety

For the safety assessment, the research team utilized the Surrogate Safety Assessment Model (SSAM) which is a software tool to identify traffic conflicts automatically based on vehicle trajectory data in the microsimulation environment, developed and validated by the Federal Highway Administration (FHWA) (Gettman et al., 2008). SSAM uses several parameters such as Time-to-collision (TTC), Post encroachment time (PET), Maximum speed (MaxS), Speed difference (DeltaS), the second vehicle’s initial deceleration rate (DR), the second vehicle’s maximum deceleration (MaxD), and the maximum speed difference value among the two-conflicted vehicles (MaxDeltaV) (Gettman et al., 2008) in order to measure the traffic conflicts. Since this study only considered the longitudinal behavior of vehicles, we used Time-to-Collision (TTC) for the conflict analysis. SSAM reports a conflict when the calculated TTC value is less than the predetermined threshold which is 1.5s in this study. For the detailed analysis, default TTC value specified in the SSAM software were used and further sensitivity analysis was conducted for different values of TTC thresholds (ranging from 1 to 3s) mentioned in the previous studies (Charly and Mathew, 2017; Fan et al., 2013).

Aimsun Next provides the vehicle trajectory file at the end of each simulation run by selecting the corresponding option. SSAM tool provides several information along with the conflicts such as simulation time when the conflict occurred and the location (X and Y coordinates) of the conflict in the network. This study utilized this information in order to count the traffic conflicts for every 5 minutes interval from 7:00 to 9:00 A.M. and matched with the corresponding segment based on the location information.
The mean, minimum & maximum value, and standard deviation of traffic conflicts for different market penetration rates of AV, CAV, and mixture of CAV and CV along with the base scenario are reported in Table 5.5.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>MPR</th>
<th>Mean</th>
<th>Maximum</th>
<th>Minimum</th>
<th>Std deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>0%</td>
<td>4796</td>
<td>4997</td>
<td>4523</td>
<td>240.5057</td>
</tr>
<tr>
<td>AV</td>
<td>20%</td>
<td>4616</td>
<td>4686</td>
<td>4497</td>
<td>73.78143</td>
</tr>
<tr>
<td></td>
<td>40%</td>
<td>4252</td>
<td>4440</td>
<td>4128</td>
<td>117.3486</td>
</tr>
<tr>
<td></td>
<td>60%</td>
<td>3789</td>
<td>3911</td>
<td>3548</td>
<td>153.6376</td>
</tr>
<tr>
<td></td>
<td>80%</td>
<td>3182</td>
<td>3265</td>
<td>3097</td>
<td>80.31065</td>
</tr>
<tr>
<td></td>
<td>100%</td>
<td>2835</td>
<td>2937</td>
<td>2708</td>
<td>91.79706</td>
</tr>
<tr>
<td>CAV</td>
<td>20%</td>
<td>4277</td>
<td>4687</td>
<td>3613</td>
<td>421.3488</td>
</tr>
<tr>
<td></td>
<td>40%</td>
<td>4110</td>
<td>4234</td>
<td>3948</td>
<td>127.2918</td>
</tr>
<tr>
<td></td>
<td>60%</td>
<td>3297</td>
<td>3474</td>
<td>3030</td>
<td>181.1127</td>
</tr>
<tr>
<td></td>
<td>80%</td>
<td>2448</td>
<td>2512</td>
<td>2325</td>
<td>77.9596</td>
</tr>
<tr>
<td></td>
<td>100%</td>
<td>2038</td>
<td>2148</td>
<td>1816</td>
<td>141.8704</td>
</tr>
<tr>
<td>CV and CAV</td>
<td>CV60%CAV0%</td>
<td>4038</td>
<td>4205</td>
<td>3723</td>
<td>195.1533</td>
</tr>
<tr>
<td></td>
<td>CV60%CAV20%</td>
<td>3261</td>
<td>3302</td>
<td>3168</td>
<td>55.16067</td>
</tr>
<tr>
<td></td>
<td>CV60%CAV40%</td>
<td>2905</td>
<td>3188</td>
<td>2637</td>
<td>256.5925</td>
</tr>
</tbody>
</table>

Note: Above table shows the traffic conflicts for the 2 hours of simulation for both directions of SR417.

In the aforementioned table (Table 5.5), it could be seen that the number of traffic conflicts for different market penetration rates (MPRs) of AV and CAV decreased compared to the based scenario and the maximum reduction was found for the 100% penetration rate for both AV and CAV scenarios. Also, the mixture of CV and CAV scenarios could reduce the number of traffic conflicts compared to the base scenario. For each MPR, CAV could decrease the number of traffic conflicts more compared with the AV. For example, looking at the 60% MPR, the mean value of TTC for CAV scenario (3297) is lower compared to the AV scenario (3789). Therefore, the
scenario with CAV for each MPRs has the lowest crash risk compared to the AV scenario and the maximum crash risk was found for the base scenario.

<table>
<thead>
<tr>
<th>Table 5.6 Comparison of traffic conflicts (TTC) for difference scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scenarios</strong></td>
</tr>
<tr>
<td>AVs</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>CAVs</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>CVs and CAVs</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Moreover, this study used ANOVA to compare the mean and their associated variance of several groups. Table 5.6 shows the results of the ANOVA test of traffic conflicts between the base scenario and different MPRs of AV, CAV, and mixture of CV and CAV. The results of the ANOVA test illustrate that traffic conflicts for base scenario along with the different MPRs in AV, CAV, and mixture of CAV and CV are significantly different at 95% confidence level.

The aforementioned results were based on 1.5s of TTC threshold. Therefore, a sensitivity analysis was also conducted by using different value of TTC threshold from 1 to 3 seconds. For simplicity, Table 5.7 shows the sensitivity analysis for 40% and 80% penetration rate only although the results are similar for other penetration rates. The results (Table 5.7) of the sensitivity analysis showed that crash risks on the freeway segment are consistent with different values of
TTC threshold. For different values of TTC threshold, all the TTC values were decreased within 19-33% for AV and 40-52% for CAV when the MPR is 80%.

Table 5.7 Sensitivity analysis of different scenarios

<table>
<thead>
<tr>
<th>Scenario</th>
<th>MPR</th>
<th>1.5s</th>
<th>2s</th>
<th>3s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>0%</td>
<td>4796.0</td>
<td>12046.8</td>
<td>45588.2</td>
</tr>
<tr>
<td></td>
<td>% reduction</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>AV</td>
<td>40%</td>
<td>4252.0</td>
<td>11644.8</td>
<td>42801.0</td>
</tr>
<tr>
<td></td>
<td>% reduction</td>
<td>11.3</td>
<td>3.3</td>
<td>6.1</td>
</tr>
<tr>
<td></td>
<td>80%</td>
<td>3182.0</td>
<td>9700.4</td>
<td>31165.0</td>
</tr>
<tr>
<td></td>
<td>% reduction</td>
<td>33.7</td>
<td>19.5</td>
<td>31.6</td>
</tr>
<tr>
<td>CAV</td>
<td>40%</td>
<td>4110.0</td>
<td>11365.6</td>
<td>40813.0</td>
</tr>
<tr>
<td></td>
<td>% reduction</td>
<td>14.3</td>
<td>5.7</td>
<td>10.5</td>
</tr>
<tr>
<td></td>
<td>80%</td>
<td>2448.0</td>
<td>7219.8</td>
<td>21985.8</td>
</tr>
<tr>
<td></td>
<td>% reduction</td>
<td>49.0</td>
<td>40.1</td>
<td>51.8</td>
</tr>
<tr>
<td>CV and CAV</td>
<td>CV60%CAV20%</td>
<td>3261.0</td>
<td>9835.8</td>
<td>35234.6</td>
</tr>
<tr>
<td></td>
<td>% reduction</td>
<td>32.0</td>
<td>18.4</td>
<td>22.7</td>
</tr>
</tbody>
</table>

5.5.2.1 Bayesian Zero-inflated negative binomial model

To better understand the benefits of connectivity feature along with the automation (CAV), this study developed a Bayesian zero-inflated negative binomial (BZINB) hierarchical model by using the traffic conflicts as a dependent variable and different MPRs of AV, CAV, mixture of CAV and CV, standard deviation of speed, and logarithm of flow as independent variables. The reason for using zero-inflated model is that there are many segments where the number of conflicts is zero compared to the other segments. Also, the research team conducted a test proposed by Vuong for the appropriateness of using zero-inflated model rather than traditional model. The result of the test showed that z-value is significant which means zero-inflated model is better than the traditional model (Lee and Mannering, 2002) for our dataset.

Table 5.8 shows the results of the Bayesian zero-inflated negative binomial model where base scenario was used as a reference to compare all the scenarios in AV, CAV, and mixture of
CV and CAV. From the table, it could be seen that all the MPRs in both AV and CAV could decrease the crash risk compared to the base condition. Also, different MPRs in the mixture of CV and CAV scenarios are also effective in reducing the traffic conflicts compared to the based scenario. The maximum reduction of crash risk was found for the full penetration rate of CAV. For the same MPRs, the CAV scenario could reduce the crash risk more than AV scenario compared to the base condition. For instance, for 80% penetration rate, the value of the coefficient is less for the CAV scenario (-0.672) compared to the AV scenario (-0.368). Thereby, CAV outperforms AV in reducing crash risk. The coefficient for the CV60CAV0 (-0.126) is greater than the coefficient of (-0.172) AV60 which implies that automated vehicle is better than the only connected vehicle in reducing crash risk. On the other hand, 20% added penetration of CAV along with the 60% CV (CV60CAV20) in the mix penetration scenario shows a higher reduction of crash risk than the 80% penetration rate of automated vehicles (AV80) compared to the base condition. The result implies that connectivity along with the automated technology have significant positive effect on the crash risk reduction of a freeway segment. Also, from Table 5.8, standard deviation of speed is found to be a significant variable in traffic conflict analysis where higher standard deviation indicates higher crash risk and vice versa.

In summary, the deployment of CAV and AV significantly decreased the number of traffic conflicts on the freeway segments compared to the base scenario. However, it is clearly found that CAV could improve traffic safety more compared to the AV scenario without connectivity.
Table 5.8 Bayesian ZINB model estimation results

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>MCSE</th>
<th>Median</th>
<th>Equal-tailed [95% Cred. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Std. Dev.</td>
<td>MCSE</td>
<td>Median</td>
<td>95% Cred. Interval</td>
</tr>
<tr>
<td><strong>Negative binomial regression part</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAV20</td>
<td>-0.039</td>
<td>0.084</td>
<td>0.007</td>
<td>-0.038</td>
<td>-0.201</td>
</tr>
<tr>
<td>CAV40</td>
<td>-0.130</td>
<td>0.072</td>
<td>0.012</td>
<td>-0.130</td>
<td>-0.270</td>
</tr>
<tr>
<td>CAV60</td>
<td>-0.258</td>
<td>0.065</td>
<td>0.011</td>
<td>-0.256</td>
<td>-0.385</td>
</tr>
<tr>
<td>CAV80</td>
<td>-0.672</td>
<td>0.078</td>
<td>0.017</td>
<td>-0.667</td>
<td>-0.826</td>
</tr>
<tr>
<td>CAV100</td>
<td>-0.813</td>
<td>0.063</td>
<td>0.008</td>
<td>-0.815</td>
<td>-0.936</td>
</tr>
<tr>
<td>AV20</td>
<td>-0.012</td>
<td>0.065</td>
<td>0.008</td>
<td>0.010</td>
<td>-0.115</td>
</tr>
<tr>
<td>AV40</td>
<td>-0.032</td>
<td>0.084</td>
<td>0.011</td>
<td>-0.030</td>
<td>-0.197</td>
</tr>
<tr>
<td>AV60</td>
<td>-0.172</td>
<td>0.061</td>
<td>0.005</td>
<td>-0.171</td>
<td>-0.287</td>
</tr>
<tr>
<td>AV80</td>
<td>-0.368</td>
<td>0.075</td>
<td>0.006</td>
<td>-0.371</td>
<td>-0.514</td>
</tr>
<tr>
<td>AV100</td>
<td>-0.506</td>
<td>0.073</td>
<td>0.011</td>
<td>-0.506</td>
<td>-0.645</td>
</tr>
<tr>
<td>CV60CAV0</td>
<td>-0.126</td>
<td>0.085</td>
<td>0.011</td>
<td>-0.126</td>
<td>-0.296</td>
</tr>
<tr>
<td>CV60CAV20</td>
<td>-0.434</td>
<td>0.064</td>
<td>0.005</td>
<td>-0.433</td>
<td>-0.558</td>
</tr>
<tr>
<td>CV60CAV40</td>
<td>-0.598</td>
<td>0.076</td>
<td>0.008</td>
<td>-0.599</td>
<td>-0.751</td>
</tr>
<tr>
<td><strong>Logistic regression part</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed_D</td>
<td>0.007</td>
<td>0.008</td>
<td>0.001</td>
<td>0.007</td>
<td>-0.010</td>
</tr>
<tr>
<td>constant</td>
<td>1.547</td>
<td>0.044</td>
<td>0.008</td>
<td>1.546</td>
<td>1.467</td>
</tr>
<tr>
<td>lnalpha</td>
<td>2.018</td>
<td>0.033</td>
<td>0.003</td>
<td>2.018</td>
<td>1.957</td>
</tr>
</tbody>
</table>

5.5.2.2 *Traffic conflicts based on vehicle type*

Almost all the previous studies (Fan et al., 2013; M.S. Rahman et al., 2019) analyzed the benefits of traffic safety under connected and automated vehicle environment by using different surrogate safety measures (such as traffic conflicts) without considering the vehicle type. Hence, this study evaluated traffic conflicts based on different types of vehicles. Since SSAM doesn’t provide any vehicle type information from the trajectory file, this study incorporated loop detectors on the critical segments (such as near ramp, lane closure) of the roadway to obtain the vehicle type.
information such as speed, arrival time on the detector, vehicle type, etc. Based on this information, the number of conflicts was calculated by using time exposed time-to-collision (TET) which is derived from the time-to-collision (TTC) notion. In total, six detectors (three in each direction) were placed to evaluate traffic safety based on vehicle type. In this study, conflicts were divided into four categories for each scenario (e.g., CAV, AV) based on different vehicle types. For example, if the following vehicle is CAV and the leading vehicle is also CAV, we categorized it as CAV to CAV conflict (A-A). Similarly, if the following vehicle is CAV and the leading vehicle is human driven vehicle (HDV), it is called as CAV to HDV conflict (A-B). The other two types are HDV to CAV conflict (M-C) and HDV to HDV conflict (M-D). The above conflict types are based on CAV scenarios. For AV scenarios, the conflict types are AV to AV conflict (A-A), AV to HDV conflict (A-B), HDV to AV conflict (M-C), and HDV to HDV conflict (M-D). The results of the traffic conflicts (TET) for different types of vehicle for both scenarios (e.g., CAV, AV) are illustrated in Figure 5.5.
Figure 5.5 Traffic conflicts based on different vehicle type

From Figure 5.5(a), it could be seen that the number of A-A conflicts (TET) is significantly less for the connected and automated vehicle (CAV) compared with the only automated vehicle (AV) which also supports the previous results. From Figure 5.5(b), it is found that the value of A-B conflicts is significantly lower for the CAV scenario compared to the AV scenario which implied that CAVs produce less conflicts with the manual vehicle than AVs. Therefore, CAVs driving behavior is safer in the mixed traffic condition compared to the AVs. On the other hand, human
driven vehicles create more conflicts with the CAVs (M-C conflicts) compared to the AVs which implied that connectivity features provide more confident driving behavior to CAVs than AVs. Moreover, the conflicts between two human driven vehicles (M-D conflicts) are significantly less for CAV scenario compared to the AV scenario.

Since the full market penetration rate will not be available in the near future, therefore it is worthwhile to study the driving behavior of CAV under mixed traffic flow. From the above results, it could be seen that the driving behavior of CAV creates less traffic conflicts under mixed traffic flow compared to the AV.

5.6 Conclusions

This paper investigated the traffic efficiency and safety impact of connected and automated vehicles (CAVs), automated vehicles (AVs) and the mixture of CAVs and connected vehicles (CVs) on the freeway corridor in the Aimsun next microsimulation platform. In order to differentiate the driving behavior of AVs, CAVs, and CVs, the research team used different longitudinal car-following models through an application programming interface (API) in the Aimsun Next. For example, AVs can only communicate with the nearest preceding vehicle whereas CAVs have the ability to communicate with all the preceding vehicles within the DSRC range. Also, CVs could get information from all the preceding vehicles within the DSRC range, but they are driven by the human drivers based on the information provided.

For the deployment of connected and automated vehicles, this study developed a testbed along the major roadway (SR417) in Orlando, Florida which was well calibrated and validated based on real-world traffic data (i.e., volume and travel time). For the traffic efficiency analysis, travel time data was used for the different market penetration rates (MPRs) of AVs, CAVs and
mixture of CVs and CAVs. The results of the analysis showed that the travel time value was less
for all MPRs of AVs, CAVs and mixture of CVs and CAVs compared to the base scenario where
the maximum reduction of travel time was observed for the full penetration rate of both AVs and
CAVs. One-way ANOVA analysis was conducted to test the statistical difference of mean for
several groups which showed that base scenario and different MPRs of AVs, CAVs, and mixture
of CVs and CAVs were significantly different from each other. Moreover, a generalized estimating
equation (GEE) was developed by using the travel time data to compare the impacts of CAVs and
AVs separately for different MPRs. In this case, travel time data was aggregated into 5 minutes
intervals for each segment for the whole 2 hours of simulation. The results of the GEE model
implied that for any penetration rate of CAV, AV, and mixture of CV and CAV, it could decrease
the value of travel time compared to the base scenario. For the same value of penetration rate,
CAV scenario could reduce the travel time more than the AV scenario compared to the base
condition which means connectivity features have a significant effect on travel time improvement.
For example, the travel time improvement for AV20 was not significant where CAV20 could
reduce the travel time significantly compared to the base scenario. The maximum reduction of
travel time was found for the full penetration rate of CAVs. Also, the mix penetration rate of CVs
(60% MPR) and CAVs (20% MPR) showed that the travel time reduction is almost similar to the
80% MPR of AVs.

For the safety assessment of CAVs and AVs, the study used traffic conflicts based on time-
to-collision (TTC) which was estimated by using the Surrogate Safety Assessment Model (SSAM).
In general, both CAVs and AVs technologies reduce the crash risk by providing significant
reduction of TTC and the maximum reduction was found for the 100% penetration rate of both
CAVs and AVs. A Bayesian hierarchical zero-inflated negative binomial model was developed by
using traffic conflicts which were aggregated into 5 minutes interval for each segment. The results of the model showed that a significant reduction of crash risk for all the MPRs of CAVs, AVs, and mixture of CVs and CAVs. Also, the higher the MPR of any connected and automated vehicle scenario (CAVs, AVs, and mixture of CVs and CAVs), the higher safety benefits were achieved. For a specific MPR, CAV outperformed AV in terms of crash risk reduction compared to the base scenario. For example, at 80% penetration rate, CAV scenario could reduce the crash risk more than AV scenario compared to the base traffic conditions.

This study also analyzed the traffic conflicts based on different vehicle types which were calculated by utilizing loop detectors on the roadway. For the conflicts’ calculation, time exposed time-to-collision (TET) was used. Hence, the traffic conflicts were divided into four parts based on different vehicle types of the following and leading vehicle such as CAV/AV to CAV/AV conflicts (A-A), CAV/AV to HDV conflicts (A-B), HDV to CAV/AV conflicts (M-C), and HDV to HDV conflicts (M-D). The results of the analysis showed that the value of A-A and A-B conflicts were less for the CAV scenarios compared to the AV scenarios which indicates the driving behavior of CAV do not have any negative impact on the other vehicles.

This research investigated the safety and operational benefits of connected and automated vehicles and compared it to only automated vehicles in the simulation environment which might be a key to answer the question of how the connectivity features might impact in the real-field CAV environment. One of the limitations of this study was to assume some driving behavior of CAVs/AVs/CVs which are not calibrated using the real-world data due the fact that those technologies are still not used adequately in real traffic. Since the full market penetration of both technologies (CAV and AV) will not be available in the near future, there will be a long transition of mixed traffic flow of conventional vehicles and CAVs or AVs or CVs. With this regard, this
study modeled the interaction between conventional vehicles and CAVs/AVs/CVs by changing the default car-following model in Aimsun Next. Thus, there is clearly a scope to improve the car-following model based on real-world data.
CHAPTER 6: EVALUATION OF TRAFFIC SAFETY BASED ON DSRC COMMUNICATION PARAMETERS UNDER CONNECTED AND AUTOMATED VEHICLES’ ENVIRONMENT


6.1 Introduction

Urban areas are developing rapidly, which creates more traffic demand leading to deterioration in the safety and efficiency of the transportation network (He et al., 2020). As traffic crashes primarily stem from human errors (Petridou and Moustaki, 2000), transportation today is shifting rapidly toward automation and connectivity, leading to the development of promising Connected and Automated Vehicles (CAVs) technologies. As the transportation research in CAVs is emerging and real data is still limited, a majority of the studies in this field are simulation oriented.

This study aims to evaluate the traffic safety under CAV environment by implementing both the realistic car-following model and the communication system in a simulation environment. The objective is to address the predominant shortcomings associated with the current practice in simulating CAV. In previous studies (Shladover et al., 2012; Zhao and Sun, 2013), researchers tend to identify the car following new parameters that can closely relate to real-life automation and connectivity. Yet, it is not feasible to evaluate the connectivity between vehicles through only car-following implementation as many of the elements of connectivity cannot be expressed with a car-following model. Rather than emulating the connectivity between vehicles in car-following models as adopted in the previous studies, the paper analyzes the performance of different communication
parameters in terms of safety effectiveness under the CAV environment. Also, in a connected vehicle environment, a vehicle can communicate with all the vehicles within the communication range and not only the leading and following vehicles. This latter approach is dominant in the literature review (Ioannou and Stefanovic, 2005; M.S. Rahman et al., 2019; Shladover et al., 2012), where the researchers adopted car-following models with acceleration is a function of only the following and the next leading vehicle. The current paper improves on state of the art by also considering all vehicles within the communication range.

The current study integrates OMNET++ (OMNeT++, 2020) and SUMO (SUMO, 2020) by using Veins (Veins, 2020) to analyze the performance of different communication parameters and traffic conditions based on traffic safety. The study adopts a longitudinal car-following model to mimic the driving behavior of CAV by considering multiple vehicle speeds within the communication range. Also, the study incorporates the realistic communication system of Dedicated Short-Range Communications (DSRC) (IEEE802.11p) in the simulation platform. For the deployment of CAV application, the study develops a calibrated and validated (based on the real traffic conditions) microsimulation testbed in one of the most congested roadways (SR408) in Orlando, Florida. Since the communication system might have a significant effect on traffic safety under the CAV environment, the study evaluates the crash risk by changing different parameters in the communication protocol. Also, the study considers different traffic conditions for the evolution of crash risk under the CAV environment. For the safety evaluation, surrogate safety measure (time-to-collision (TTC)) is utilized. Also, the study develops a binary logistic regression model using traffic conflicts. For the evaluation of the communication system, packet loss, packet loss ratio, and communication collision are used. To the best of our knowledge, this is the first study where
the performance of different communication parameters and traffic conditions were analyzed based on traffic safety under the CAV environment.

6.2 Simulation setup and data

For the study area, we used one of the most congested expressways in Orlando, Florida named Holland East-West Expressway (SR408). The length of the roadway segment was 7-miles long which includes 7 on-ramps and 11 off-ramps from “Old Winter Garden Rd” to “Lake Underhill Rd”, Orlando. In the study area, traffic data was collected from MVDS (Microwave Vehicle Detection System) detectors. MVDS provides volume, speed, and lane occupancy for each lane at 1 min interval. Also, the collected dataset categorizes vehicles into four types according to their length where type 1: vehicles up to 3 m in length, type 2: vehicles 7.5 m in length, type 3: vehicles 7.5 to 16.5 m in length, type 4: vehicles over 16.5 m in length. In this study, vehicle composition was classified into two categories i.e., (1) passenger car (PC) and (2) heavy goods vehicle (HGV) based on the length of the vehicle. If the length of a vehicle is equal or less than 7.5m, it was classified as passenger car (PC) and the others are termed as heavy goods vehicles (HGV). The study area along with the MVDS detectors is shown Figure 6.1.
6.2.1 Calibration and Validation

Calibration and validation are important tasks for a microsimulation-based study in order to replicate the real-world traffic condition. In this study, simulations were performed in the SUMO platform. Existing traffic volume data were collected at the selected study area (SR408) during weekday morning period from 6 to 10 AM. After that, we analyzed the data for the selected detectors for the specified period to identify the peak traffic volume. Figure 6.2 shows the traffic volume data for different detectors for the different time intervals from 6:00 to 10:00 AM. From the figure, it could be seen that traffic volume is higher from 7:00 to 9:00 AM which was used for the simulation in this paper.
In order to calibrate and validate the simulation model, traffic data including traffic volume and speed were aggregated into 5 minutes intervals. The simulation time was set from 6:30 to 9:00 am, where first 30 minutes were selected for warm-up period. For calibration, traffic volume was used and for validation, speed data was utilized. To compare the simulated traffic volume with the real-world data, Geoffrey E. Heavers (GEH) and Correlation Coefficient (CC) statistics were analyzed (M. H. Rahman et al., 2019). In order to achieve a good representation of the simulation model, different parameter settings available in the Sumo simulation environment were changed to replicate the real field traffic conditions. Hence, the objective function to calibrate the model was to minimize the differences between simulation and real-field traffic volume so that the GEH value should be less than 5 and CC value should be close to 1.
GEH is a statistical parameter based on the Chi-square test which calculates the difference between simulated and field traffic volume. The definition of GEH is as follows:

\[
GEH = \sqrt{\frac{2 \times (V_{obs} - V_{sim})^2}{(V_{obs} + V_{sim})}}
\]  

(35)

where \(V_{obs}\) is the traffic volume (veh/hr) from field detectors and \(V_{sim}\) is the simulated traffic volume (veh/hr). In order to get a good calibrated network, GEH should be equal or less than 5 for at least 85% of the total observations (Abdel-Aty and Wang, 2017; Rahman et al., 2019; Rahman and Abdel-Aty, 2018c; Yu and Abdel-Aty, 2014). Additionally, the Correlation Coefficient (CC) was calculated, it measures the linear relationship between field and simulated traffic volume. The expression of CC as follows:

\[
CC = \frac{1}{n-1} \sum_{i=1}^{n} \frac{(y_{i,sim} - \bar{y}_{sim})(y_{i,obs} - \bar{y}_{obs})}{S_{sim}S_{obs}}
\]  

(36)

where \(n\) is the total number of observations in traffic measurement, \(\bar{y}_{sim}\) and \(\bar{y}_{obs}\) are average value of simulation and field volume, respectively. \(S_{sim}\) and \(S_{obs}\) are the standard deviations of simulation and field traffic volume, respectively. For a perfect correlation, the CC value should be within ±1 where a value of 0.85 is considered acceptable for the calibration (El Esawey and Sayed, 2011; FDOT Systems Planning Office, 2014; Hollander and Liu, 2008). For validation, this study utilized speed data from both simulation and real field aggregated into 5 minutes intervals. \(R^2\) and RMSE value were used to compare the field and simulated average
speeds (Ciuffo and Punzo, 2010). Sumo simulation platform uses krauss car-following model to control the longitudinal behavior of vehicles. In order to calibrate and validate the simulation model, different parameters available in the car-following model were tweaked within a certain range. For example, the default value for sigma and tau are 0.5 and 1, respectively where the calibrated value is 0.4 and 0.9, respectively. Table 6.1 shows the default value and calibrated value for different parameters which were used in this study.

Table 6.1 Parameter settings in Sumo for calibration

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Default value</th>
<th>Calibrated value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sigma</td>
<td>0.5</td>
<td>0.4</td>
</tr>
<tr>
<td>Tau</td>
<td>1</td>
<td>0.9</td>
</tr>
<tr>
<td>Max speed</td>
<td>55 m/s</td>
<td>40 m/s</td>
</tr>
<tr>
<td>Speed factor</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td>Speed Deviation</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Acceleration</td>
<td>2.6 m/s²</td>
<td>3.6 m/s²</td>
</tr>
<tr>
<td>Deceleration</td>
<td>4.5 m/s²</td>
<td>4.5 m/s²</td>
</tr>
</tbody>
</table>

The calibration results showed that around 94% of the cases, the GEH value was less than 5 which meets the above mention criteria. Also, the Correlation coefficient value was 0.96 which means almost perfect and direct correlation. From the validation results, it was found that $R^2$ value is around 0.93 and RMSE is 2.12 which indicates good validation based on the average speed.
Hence, the simulation model (base scenario) is now well calibrated and validated based on the real field data.

6.3 Proposed methodology

Since CAV technology depends on both communication system and car-following behavior, the research team incorporated both in this study.

6.3.1 Car-following model

Since the driving behavior of connected and automated vehicles (CAVs) is significantly different from non-CAV vehicles, it is one of the most challenging tasks to approximate their behavior in the simulation platform. Several previous studies used different car-following models to mimic the CAV behaviors in the simulation environment. Among them, intelligent driver model (IDM) is the most used car-following model since it has the ability to model different traffic conditions such as oscillations, stop and go traffic, etc. The IDM was first introduced by Treiber et al. (Treiber et al., 2000) which was further modified for the automated vehicle (adaptive cruise-control system) by adjusting different parameters in acceleration function in some previous studies (Kesting et al., 2010b) where the acceleration function is computed based on the speed difference between following and the immediate leading vehicle. In a connected vehicle environment, a vehicle can communicate with all the vehicles within the communication range. Therefore, this study considered the average speed of all the vehicles within the communication range based on the previous studies (Peng et al., 2020; Sun et al., 2015). Hence, the acceleration function of the modified IDM model for the connected and automated vehicle (CAV) is given below:
\[ a(t + \tau) = a_{max} \left[ 1 - \left( \frac{V_n}{V_0} \right)^\delta - \left( \frac{S^*(V_n, \Delta V)}{S_n} \right)^2 \right] \]

\[ S^*(V_n, \Delta V) = S_0 + V_n T + \frac{V_n \Delta V}{2\sqrt{a_{max} b}} \]

(37)

\[ \Delta V = \gamma_1 (V_{\text{following}} - V_{\text{leading}}) + \gamma_2 (V_{\text{following}} - \frac{1}{n} \sum_{i=1}^{n} \bar{V}) \]

where \( V_{\text{following}}, V_{\text{leading}} \) are the following and the immediate leading vehicle speed, respectively and \( \bar{V} \) is the average speed of all the vehicles within the communication range, \( n \) is the total number of vehicles, \( \gamma_1 \) and \( \gamma_2 \) are the weight factor. Also, \( \tau = \) the perception-reaction time, \( a_{max} = \) the maximum acceleration, \( V_n = \) the speed of the following vehicle, \( V_0 = \) the desired speed, \( \delta = \) the acceleration exponent, \( S_n = \) the gap distance between two vehicles, \( S_0 = \) the minimum gap distance at standstill, \( T = \) the safe time headway and \( b = \) the desired deceleration.

Based on the previous studies (Kesting et al., 2010b), the parameter settings in order to mimic the IDM model for the automated system is as follow: minimum gap (\( S_0 \)), safe time headway (\( T \)), maximum acceleration (\( a_{max} \)), desired speed (\( V_0 \)), and acceleration exponent (\( \delta \)) are 2 m, 0.6 s, 1.4 m/s\(^2\), 2.8 m/s\(^2\), and 4, respectively.

6.3.2 Communication protocol

In the connected vehicle environment, the following vehicle gets information such as acceleration, position, speed, etc., from the leading vehicles through the wireless communication system. Based on this information, the following vehicle adjusts their acceleration behavior to
avoid collision. Packet loss is one of the most important parameters that will affect the performance of this application. A vehicle receives the information about speed and acceleration which is transmitted from the nearest vehicles. Based on this information, the vehicle adjusts the speed or acceleration behavior by using the modified IDM model (mentioned above). The process of message broadcasting, receiving, and communication system of a vehicle is utilized by using Veins which integrates OMNET++ and SUMO bi-directionally (Sommer et al., 2010). Figure 6.3 shows the whole process of integrating SUMO, OMNET++, and Veins.

![Flow chart of the communication between OMNET++, SUMO, and Veins](image)

**Figure 6.3 Flow chart of the communication between OMNET++, SUMO, and Veins**

A packet loss may cause slow reaction of the following vehicle, and inaccurate acceleration prediction which might cause vehicle to vehicle conflicts. To implement the vehicle ad-hoc networks (VANETs), this study used Veins which bridges SUMO and OMNET++ by coupling them bi-directionally through TraciScenario Manager. By using the TCP socket and traffic control interface (Traci), Veins helps SUMO and OMNET++ to run in parallel (Sommer et al., 2010).
In this study, we used the dedicated short-range communication (DSRC) or IEEE 802.11p in order to replicate the communication protocol. Also, the message transmission module requires ConnectionManager, Decider80211p, PyhLayer80211p, Mac1609_4, and BaseWaveApplLayer. ConnectionManager to determine whether two vehicles are within the communication range or not based on the maximum interference distance. Decider80211p analyzes the packet received or dropped at the receiver end based on the parameters such as received power, signal to interference plus noise ratio (SINR), etc. PyhLayer80211p determines the code needed to be inserted in the analogue models. BaseWaveApplLayer works as an application layer which helps a vehicle to send each packet with different transmission rate and power. The Mac1609_4 module incorporates the MAC layers such as 1609.4 and 802.11p. The detailed communication parameters are shown in Table 6.2.

### Table 6.2 Communication parameters of IEEE 802.11p

<table>
<thead>
<tr>
<th>Parameters</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmission rate</td>
<td>6 Mbps</td>
</tr>
<tr>
<td>Transmission power</td>
<td>100 mW, 50mW</td>
</tr>
<tr>
<td>Maximum Interference Distance &amp; sensitivity</td>
<td>300 m, -89dBm</td>
</tr>
<tr>
<td>Carrier Frequency</td>
<td>5.890e9 Hz</td>
</tr>
<tr>
<td>Queue size</td>
<td>10, 20</td>
</tr>
<tr>
<td>Queue type</td>
<td>Drop tail</td>
</tr>
<tr>
<td>Physical layer</td>
<td>802.11p</td>
</tr>
<tr>
<td>MAC layer</td>
<td>1609.4</td>
</tr>
</tbody>
</table>
By using these communication parameters, vehicles transmit and receive their information and adjust the longitudinal behavior by utilizing the acceleration function of modified IDM model (mentioned above). At first, vehicle get the information such as speed, position from the immediate leading vehicle and the average speed of all vehicles within the DSRC range. Based on this information, the acceleration was estimated by using equation 3. Finally, the acceleration is used by the vehicle if it is less than the maximum acceleration otherwise vehicle use the maximum acceleration. Hence, the message processing algorithm for the estimation of acceleration profile of CAV is shown in Table 6.3 below:
6.4 Evaluation measures

Surrogate safety measures are a widely accepted technique in order to evaluate the crash risk in the microsimulation-based study. Almost all the car-following models in the simulation environment are crash free models. Hence, surrogate safety measures are the only technique to analyze the safety effectiveness, which is calculated by utilizing the trajectory files generated in
the simulation run. In the SUMO simulation platform, vehicles are equipped with SSM device which collects the trajectory information for each vehicle for every time interval and hence estimates the surrogate measures. Different studies used different surrogated measures such as time-to-collision (TTC), deceleration rate to avoid a crash (DRAC), and post encroachment time (PET), etc. Since the study modeled the longitudinal control of CAV vehicle, TTC was used for the safety evaluation.

6.4.1 Time-to-collision (TTC)

Since the study considered the longitudinal behavior of CAV, time-to-collision (TTC) was utilized for the surrogate safety measure which represents the rear-end conflicts. TTC is defined as the time that is required to have a collision between two vehicles (leading and the following vehicle) if the vehicles maintain the same travel path and the current speed. Hence, if the value of the TTC is less than the specified threshold (1s in this study), it is considered as a conflict. If the speed of the following and leading vehicle is $v_n$ and $v_{n-1}$, respectively then the TTC is estimated based on the following equations:

$$\text{TTC}_n(t) = \begin{cases} 
\frac{y_{n-1}(t) - y_n(t) - L_{n-1}}{v_n(t) - v_{n-1}(t)}, & \text{if } v_n(t) > v_{n-1}(t) \\
\infty, & \text{if } v_n(t) \leq v_{n-1}(t)
\end{cases} \quad (38)$$

where $\text{TTC}_n(t) =$ the TTC value of vehicle $n$ at time $t$, $y =$ the positions of vehicles, $v =$ velocities of the vehicles, $L_{n-1} =$ Length of leading vehicles.
Based on this information, the study developed a binary logistic regression model to evaluate the crash risk of different communication protocols and traffic conditions. The binary logistic regression model estimates the probability of happening a conflict based on the explanatory variables. If $Y$ denotes the dependent variable (where $Y=1$ means conflict and $Y=0$ means no-conflicts) and $X = (X_1, X_2, ..., X_n)$ be the set of explanatory variables then the probability that a conflict will occur is estimated based on the following equation:

\[
\pi_i = \Pr(Y = 1|X_i = x_i) = \frac{e^{f(x_i)}}{1 + e^{f(x_i)}}
\]

where, $\pi$ denotes the probability of conflicts which is estimated by using the total number of conflicts divided by the total number of observations (conflicts and no-conflicts). $\beta$ denotes the coefficients of each explanatory variables. The odds of an event (conflicts) are estimated based on the probability of conflicts divided by the probability of no-conflicts. Hence, the odds ratio is the ratio of two odds that is equal to the $e^\beta$ which implies the association between two scenarios (e.g., Q10 vs Q20).

### 6.5 Results and discussion

In order to evaluate the safety effectiveness of CAV technology by utilizing the real communication parameters, this study analyzed traffic conflicts based on the time-to-collision (TTC) measure. For the analysis of TTC the study analyzed the number of conflicts for every 5
minutes interval from 7:00 to 9:00 AM. No data were considered for the final evaluation from the warm-up period. Also, multiple runs were conducted to address the randomness effect of simulation. Since the simulation was implemented by using OMNET++, we recorded packet loss, packet loss ratio, and number of communication collisions, to compare different communication protocols. Finally, this study analyzed the safety effectiveness by utilizing both different communication parameters and traffic conditions. Hence, this section is divided into two parts: communication and transportation aspects.

6.5.1 Communication aspect

6.5.1.1 Queue Size

Queue size is the maximum number of packets that can be stored or received by a node (vehicle). If the size of the queue is greater than the maximum queue size, the packet is dropped (Marbach and Lu, 2006; Wu et al., 2008). Hence, the queue size is an important parameter in terms of CAV communication system. Vehicles send their information such as speed, acceleration, position, etc., to the nearby vehicles as a packet. If the number of packets is very high or the queue size of the received vehicle is low, the packet might be lost. Hence, the vehicle (receiver end) cannot get the actual traffic information from its leader, which might result in the inaccurate prediction of acceleration/speed. Based on the previous studies (Gong et al., 2006; Klingler et al., 2016; Saritha and Viswanatham, 2014; Zheng and Atiquzzaman, 2001), the study used two types of queue size: queue size 10 which is denoted as Q10 and queue size 20 which is termed as Q20. In this study, we have used drop tail type for the queueing analysis which uses the mechanism where each packet is considered as identical, and when the size of the queue is maximum, the newly incoming packets is dropped (Gong et al., 2006; Wu et al., 2008). Figure 6.4(a) illustrates
the number of packets loss ratio (number of packet loss/sent packet) for different values of queue size. Figure 6.4(b) shows the relation between queue size and traffic conflicts.

Figure 6.4 Effect of Queue size on packet loss ratio and traffic conflicts
From Figure 6.4(a), it could be seen that number of packets loss ratio for each node is higher for queue size 10 (Q10) compared to the queue size 20 (Q20). The total packet loss ratio for all nodes is 87199 and 83617 for Q10 and Q20, respectively. Since the packet loss ratio is higher for Q10 compared to Q20, the number of traffic conflicts is also higher for Q10 compared to the Q20, which is shown in Figure 6.4. The total number of traffic conflicts (based on TTC) is 4488 and 3808 for Q10 and Q20, respectively. The study also conducted a paired sample t-test, which determines the mean difference between two groups. The results of the analysis showed that traffic conflicts and packet loss ratio are significantly higher for Q10 compared to the Q20 at 95% significance level.

6.5.1.2 Transmission Power

Transmission power is one of the most important factors in the wireless communication system as well as in CAV technology. At high transmission power, there is a high chance of creating interference, which might result in packet drop (Panichpapiboon et al., 2006). Therefore, communication between CAVs might be interrupted. Hence, this study utilized two different values for the transmission power i.e., 50mW and 100mW to investigate the possible effect on CAVs communication in the transportation network. Figure 6.5(a) shows the packet loss value for each node in the simulation environment where TP-50 means transmission power 50mW and TP-100 indicates transmission power 100mW. For almost all the nodes, packet loss is higher for the TP-100 scenario compared to the TP-50 scenario.
a. Packet loss of different nodes (TP100 and TP50)

b. Traffic conflicts

Figure 6.5 Effect of Transmission power on packet loss and traffic conflicts
From Figure 6.5(b), it could be found that the number of traffic conflicts (between vehicles) is higher for the TP-100 scenario compared to the TP-50 scenario because of the high communication interference, which causes a higher packet drop. The total packet loss for all nodes is about 626165 and 598069 for TP-100 and TP-50 scenario, respectively. Also, the number of traffic conflicts is about 4488 and 3849 for TP-100 and TP-50 scenarios, respectively. The study also conducted a paired t-test by utilizing both packet loss and traffic conflict separately. The results of the analysis showed that TP-100 and TP-50 are significantly different from each other at 95% significance level.

6.5.2 Transportation aspect

6.5.2.1 Lane closure

The likelihood of crash risk near the bottleneck area is higher compared to the regular segment of a roadway (Li et al., 2014). This bottleneck could be created due to several reasons such as roadside construction, traffic incidents, etc. Connected and automated vehicles might have the potential to improve traffic safety near the bottleneck area by providing traffic information to the upstream vehicles through vehicle ad-hoc network (VANET) system (Li et al., 2017) where successful communication between vehicles is one of the major challenges. Due to the higher traffic volume near the upstream segment of the bottleneck area, the number of transmitted packets would be higher, which might increase the collision between packets. Thus, the communication disruption of the wireless system might occur due to the communication collisions which could cause higher crash risk especially in the bottleneck area. Hence, this study analyzed the safety effectiveness of CAV technology by utilizing with or without real communication protocol (DSRC) near the bottleneck area and compare between them. Also, the study compared the results
between CAV and the base scenario. To conduct the study, we developed a bottleneck scenario by closing a lane in the eastbound direction of SR408 roadway. Communication disruption and traffic safety evaluation results for the lane closure (LC) scenario are shown in Figure 6.6.

![Diagram showing traffic conflicts and communication collisions for different scenarios]

**Figure 6.6 Effect of lane closure on traffic conflict and communication collisions**
Figure 6.6 (a) illustrates the number of traffic conflicts for both lane closure (LC) and without lane closure (WLC) scenarios. The study used two ways for the CAV application: (1) CAV with DSRC and (2) CAV without DSRC, where it represents CAV scenario with no real communication system, which is similar to previous studies (Rahman et al., 2019; Wen-Xing and Li-Dong, 2018). In CAV without DSRC, researchers didn’t use any communication system and assumed uninterrupted sharing of information between vehicles which would not replicate the real CAV condition. On the other hand, CAV with real communication system is denoted as CAV with DSRC scenario which would approximate the real CAV scenario. In this case, the study utilized the actual communication protocol. It should be noted that CAV without DSRC is used to analyze the effect of the communication system. Also, the study analyzed the effectiveness of CAV and compared to the base condition. From Figure 6.6 (a), it could be depicted that the number of traffic conflicts for both scenarios (WLC, LC) is lower for CAV conditions (with and without DSRC) compared to the base condition, which means crash risk is improved with the application of CAV. Since the objective of this study is to capture the communication effect on traffic safety, we analyzed the performance of CAV considering with and without real communication system (DSRC) while CAV without DSRC doesn’t replicate the real traffic condition. The performance of CAV conditions (with and without DSRC) for the Lane closure scenario (Figure 6.6 (a)) showed that the number of traffic conflicts is higher for the CAV with DSRC compared to the CAV without DSRC which implied that the communication system has a significant effect on crash risk. Also, the study found that crash risk is about 28% lower for CAV without DSRC compared to the CAV with DSRC. These increased number of traffic conflicts possibly occurred due to the higher number of communication collisions between packets. Additionally, crash risk is 26% higher for LC scenario compared to the WLC scenario when considering CAV with DSRC. Thus, the results
also implied that higher communication collisions might increase the crash risk under CAV conditions. Hence, the study evaluated the number of communication collisions of each node for both LC and WLC scenarios. A communication collision occurs when two or more nodes transmit data/packets over the same node at the same time. Traffic density would be higher near the bottleneck area and vehicles would receive higher number of packets from the nearby vehicles at the same time, which might result in collision between transmitted packets as well as packet drops. Hence, data cannot be transferred to the receiver end, which can increase the likelihood of crash risk. Figure 6.6 (b) shows that the number of communication collisions for the lane closure (LC) scenario is significant compared to the without lane closure (WLC) scenario. Due to the communication disruption, crash risk is significantly higher for the LC scenario (traffic conflicts = 4669) compared to the WLC scenario (traffic conflicts = 3694). Thus, it could be implied that the communication system has significant effect on traffic safety when considering lane closure situation.

6.5.2.2 Vehicle Flow

In order to evaluate the safety effectiveness of CAV technology by utilizing real communication protocol, this study utilized four different percentages of total traffic flow levels based on the base (field) condition such as 25% (F25), 50% (F50), 75% (F75), and 100% (F100). For example, F25 means 25 percent of the total vehicle (base scenario) is utilized in this scenario and so on. Similar to the lane closure, the study considered two ways for the CAV application (with and without DSRC) and compared to the base scenario. Figure 6.7 shows the results of the communication collision and crash risk (traffic conflicts between vehicles) analysis for different percentage of traffic flow. Figure 6.7(a) illustrates that crash risk is lower for both CAV conditions
(with and without DSRC) compared to the base scenario. Also, the results of the analysis showed that CAV (with or without DSRC) could reduce the crash risk more for high traffic flow than low traffic flow compared to the base condition. For example, CAV with DSRC in F100 scenario can reduce the crash risk by 42% compared to the base condition, while for F25 scenario, the reduction percentage is about 21%. For CAV conditions (with or without DSRC), the results of the analysis for different percentages of traffic flow illustrated that CAV with DSRC has higher crash risk compared to the CAV without DSRC. Therefore, the communication system has a significant effect on traffic safety as well. For instance, crash risk is 25% higher for CAV with DSRC compared to the CAV without DSRC when considering F100 scenario. Additionally, CAV with DSRC for different percentages of traffic flow results implied that the number of traffic conflicts is higher for a higher percentage of traffic flow. Therefore, more communication collisions are occurring with the increase in traffic flow, which leads to the higher crash risk. Figure 6.7(b) illustrates the total number of communication collisions for the same number of nodes (vehicles) for each percentage of traffic flow under CAV with DSRC condition. The results of the analysis inferred that with the increase of traffic flow, the number of communication collision is higher, which indicates a higher number of packet drops. Thus, crash risk is higher for 100% traffic flow compared to the other scenarios (i.e., 25%, 50%, 75%).
a. Total communication collisions

b. Total communication collisions

Figure 6.7 Effect of traffic flow on traffic conflicts and communication collisions
Moreover, the study evaluated the analysis of variance (ANOVA) to test the difference of mean and their associated variance among several groups. The results of the ANOVA test for the traffic conflicts with different percentage of flow showed that the number of traffic conflicts is significantly different (F-value=105.12 > F_{critical}) from each other at 95% confidence interval level. Therefore, the number of packets loss is higher with the increase in traffic flow due to the increased number of communication collisions between packets. Hence, the crash risk (conflicts between vehicles) could be higher with the increase in traffic flow.

6.5.3 Binary Logistic Regression Model

The study developed a binary logistic regression model to evaluate the binary outcome (conflict vs no-conflict) by utilizing the traffic conflicts generated in SUMO simulation run. In this model, the dependent variable Y is termed as binary outcome where Y=1 represents conflict, and Y=0 indicates no-conflict. The model was developed for both communication aspect (Queue size, Transmission power) and transportation aspect (Lane closure, Vehicle flow). Hence, the odds ratio was estimated for each scenario by considering CAV with DSRC scenario. Also, we compared the results by utilizing the CAV without DSRC conditions. The results of the binary logistic regression model are shown in Table 6.4.
Table 6.4 Model estimation results for different scenarios

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Reference Category</th>
<th>Coefficient (p-value)</th>
<th>Odds ratio</th>
<th>95% Wald confidence limit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower limit</td>
</tr>
<tr>
<td>Q20</td>
<td>Q10</td>
<td>-0.0486 (0.0281)</td>
<td>0.953**</td>
<td>0.912</td>
</tr>
<tr>
<td>TP50</td>
<td>TP100</td>
<td>-0.0979 (&lt;0.001)</td>
<td>0.907**</td>
<td>0.868</td>
</tr>
<tr>
<td>Intercept</td>
<td>--</td>
<td>-4.7070 (&lt;0.001)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>LC</td>
<td>WLC</td>
<td>0.1038 (&lt;0.001)</td>
<td>1.109**</td>
<td>1.062</td>
</tr>
<tr>
<td>F25</td>
<td>F100</td>
<td>-0.1968 (&lt;0.001)</td>
<td>0.821**</td>
<td>0.768</td>
</tr>
<tr>
<td>F50</td>
<td>F100</td>
<td>-0.1156 (&lt;0.001)</td>
<td>0.891**</td>
<td>0.844</td>
</tr>
<tr>
<td>F75</td>
<td>F100</td>
<td>-0.953 (&lt;0.05)</td>
<td>0.953**</td>
<td>0.908</td>
</tr>
<tr>
<td>Intercept</td>
<td>--</td>
<td>-4.7913 (&lt;0.001)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>LC (CAV</td>
<td>LC (CAV without</td>
<td>0.0752 (&lt;0.001)</td>
<td>1.078**</td>
<td>1.031</td>
</tr>
<tr>
<td>with DSRC)</td>
<td>DSRC)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>--</td>
<td>-4.7627 (&lt;0.001)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>F100(CAV</td>
<td>F100 (CAV without</td>
<td>0.0478 (0.05)</td>
<td>1.049**</td>
<td>1.008</td>
</tr>
<tr>
<td>with DSRC)</td>
<td>DSRC)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>--</td>
<td>-4.8391 (&lt;0.001)</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

**Significance at 0.05 level.

The goodness of fit results showed that the developed logistic regression model was significant (chi-square = 126.13, df=4, p-value<0.001). From Table 6.4, it could be seen that the odds of having conflicts for Q20 is 4.7% lower than the odds of having conflicts for Q10 scenario. Looking at the transmission power scenarios (TP50 vs TP100), it was found that odds ratio is 0.907, which means the odds of having conflicts for TP50 is 9.3% less than the odds of experiencing conflicts in the TP100 scenario. On the other hand, the odds of having conflict in LC
(lane closure) scenario is 1.109 times to the chances of having conflicts in WLC (without lane closure) scenario. In addition, odds of having for CAV with DSRC is 7.8% higher compared to the CAV without DSRC when considering LC scenario. Moreover, the odds ratio of different percent of traffic flow indicates that the odds of having conflicts is 17.9%, 10.9%, and 4.7% lower for F25, F50, and F75 scenarios, respectively than the odds of having conflicts for the F100 scenario. CAV with DSRC showed that the odds of having conflicts is 1.049 times higher compared to the CAV without DSRC for F100 scenario which implies that communication system has significant effect on traffic safety.

6.6 Conclusions

This study analyzed the safety effectiveness of CAV technology by utilizing real communication protocol (IEEE 802.11p) in the simulation platform. To evaluate the CAV technology, the study developed a testbed in a SUMO simulation environment along the major roadway (SR408) in Orlando, Florida. To represent the real traffic condition of SR408, the research team calibrated and validated the simulation model based on volume and speed data, respectively. Calibration and validation results were found to be appropriate to represent the real traffic condition in the simulation model. As the driving behavior of CAV can be different from non-CAV, this research used the modified version of the intelligent driver model (IDM) to approximate the longitudinal car-following behavior of CAV. The application of CAV technology does not only depend on the car-following behavior but also the uninterrupted communication between vehicles. Therefore, this study incorporates a real communication system by using OMNET++ simulation environment. SUMO and OMNET++ simulation platforms that were coupled bi-directionally by using Veins. For the safety evaluation, the study estimated the traffic conflict by
using time-to-collision (TTC) surrogate measure, which is calculated from the trajectory file generated in SUMO. Also, the study developed a binary logistic regression model based on traffic conflicts.

To evaluate the safety effectiveness of CAV technology due to the communication disruption, the study considered information transmission through the implementation of the IEEE 802.11p protocol. The study used different values for the different communication parameters such as transmission power, queue size to investigate the impact of traffic safety. Two different values of queue size (Q10, Q20) were proposed to evaluate the safety of CAV technology. The results of the queue size analysis showed that with the smaller number of queue size, more packets are dropped, which causes loss of transmitted information. Hence, the number of traffic conflicts (TTC) is higher for smaller queue size (Q10) compared to the large queue size (Q20). Also, the logistic regression model showed that the odds of having traffic conflicts for Q20 scenario are 4.7% lower than the odds of experiencing traffic conflicts for the Q10 scenario. Transmission power is an important parameter in wireless communication where high transmission power indicates higher interference. The study adopted two values of transmission power (TP-50, TP-100) to evaluate the crash risk of CAV technology. From the analysis, it was found that crash risk (conflicts between vehicles) is significantly lower for TP-50 compared to the TP-100 scenario. Additionally, the odds ratio for TP50 vs TP100 scenarios is 0.907, which also indicates a significant reduction of traffic conflicts. From the transportation aspect, the study found that communication system has a significant effect on traffic safety. Hence, the study evaluated the crash risk based on lane closure and different percentages of traffic flow. The study utilized two ways for CAV application: (1) CAV with DSRC, (2) CAV without DSRC. Due to the higher
number communication collisions, the number of traffic conflicts is 28% lower for CAV without DSRC compared to the CAV with DSRC when considering lane closure scenario which implies a significant impact of communication system on traffic safety. Also, crash risk is 26% higher for lane closure (LC) scenario compared to without lane closure (WLC) when considering CAV with DSRC. Communication results inferred that LC scenario creates higher communication disruption compared to the WLC scenario. Different percentages of traffic flow, i.e., (25%, 50%, 75%, and 100%) showed that CAV with DSRC experience a higher number of traffic conflicts compared to the CAV without DSRC for every percentage of traffic flow. On the other hand, with the increase in vehicles flow, the communication collision between packets is higher. Thus, the crash risk (conflicts between vehicles) is higher with the increase in traffic flow when considering CAV with DSRC.

Finally, the results of the study implied that the communication systems have a significant effect on traffic safety in a connected and automated vehicle (CAV) environment. Also, different traffic conditions might deteriorate communication between CAV, which could increase the crash risk. It should be noted that both CAV conditions (with and without DSRC) are better compared to the base condition for every scenario. The finding of this study implies that not only the acceleration/deceleration behavior of CAV but also a successful communication between CAVs is necessary to elevate the performance CAVs. Therefore, the results of this study provide insights to transportation planners and decision makers about the performance of communication systems under the CAV environment. To enhance the performance of the VANET system, different control strategies could be used under CAV environment. For example, different types of queue management strategies could be utilized to improve the safety under CAV environment. Also, distributed fair power, dynamic priorities to messages, etc., could be used to improve the effect of
transmission power in the VANET system. It is important to account for communication in studying CAV as the simplistic representation in CAV is used in many previous research (and also illustrated in this paper in the CAV without DSRC scenario) could over estimate the benefits of CAV without accounting for the restrictions of communication.

There are some limitations to the current study and scope for further investigation of this research. First, the car-following model is not calibrated based on the field data given the lack of deployed CAV technologies. Future experiments can then be expanded by providing more realistic behavior of CAV. Also, further research about the improvement of the VANET system could be the future extension of this study.
CHAPTER 7: CONCLUSIONS

7.1 Summary

The dissertation investigated the benefits of connected and automated vehicle (CAV) in terms of both safety and mobility in the simulation platform. For the deployment of CAV technologies, the study developed a large-scale calibrated and validated network based on real traffic conditions by using dynamic traffic assignment (DTA) model. Different CAV technologies e.g., connected and automated vehicle (CAV), automated vehicle (AV), and connected vehicle (CV) were analyzed by utilizing different car-following models. For the evaluation of safety and mobility benefits, the study used freeway, arterial, and network level which considers both freeways and arterials. The driving behavior of different CAV technologies was coded by using application programming interface (API). Also, the study used a signal control algorithm to elevate the performance of CAV in the arterial segments. Furthermore, different market penetration rates (MPRs) were analyzed to investigate the benefits of CAV in mixed traffic flow. To implement the real communication system in the simulation environment, the study utilized Dedicated Short-Range Communications (DSRC or IEEE 802.11p) system. Also, several statistical tests (t-test, ANOVA, Wilcoxon signed-Rank test, etc.) were used for the analysis of safety and mobility benefits of CAV with different MPRs. Additionally, different modeling techniques e.g., generalized estimating equation (GEE), binary logistic regression, etc., were utilized for the evaluation of CAV technologies.

In Chapter 3, the study described the development of a large-scale DTA based simulation model for the deployment of CAV technologies using Multi-Resolution Modeling (MRM) framework in Orlando, Florida. The model was developed by using Cube based regional traffic demand model (RTDM) with a based year of 2009. Afterward, the model was calibrated and
validated with the real field data (volume and travel time) of 2017. For the calibration and validation process, the study collected data from multiple sources such as MVDS, AVI, HERE, etc. Also, the study collected real time signal information from different sources. The results showed that both mesoscopic and microscopic areas were calibrated and validated within an acceptable bound of error. In terms of mesoscopic area calibration, $R^2$ was found to be 0.95 and the 85% of the 259 validation locations (travel time) with eight time-intervals spread across the network are within an acceptable bound of error. For the microscopic area, this study also achieved calibration and validation criteria based on widely accepted calibration and validation guidelines. To this end, the study used this calibrated and validated network for the deployment of CAV technologies.

In Chapter 4, the dissertation discussed the application of CAV technologies at the network level which considered both freeways and arterial segments. Since the full penetration of CAV technologies will not be available in the foreseeable future, the study utilized different market penetration rates (MPRs) of CAV and analyzed the benefits in terms of safety and mobility. Also, the study used vehicle-to-infrastructure (V2I) technology which enables signals to communicate with CAV and share real time information such as vehicle speed, position, green time, etc. Based on this information, CAVs acceleration/deceleration profile and signal phase and timing (SPaT) were adjusted through the signal control algorithm. For the deployment of CAV technologies, the study used the most important and busiest corridor (Orlando CBD area) in Orlando, Florida. For the traffic efficiency analysis, travel time rate (TTR), approach delay at the intersection were used and for the safety evaluation, speed standard deviation, real-time crash risk model (freeways and arterials separately) were utilized. The traffic efficiency results showed that TTR was reduced by 7% when the MPR of CAV is 20% for the whole network while maximum benefits were found in
100% penetration rate. Also, the reduction percentage is around 20% and 5% for freeways and arterials, respectively when the penetration level is 20%. The proposed signal control algorithm showed that approach delay was significantly lower compared to the base condition for around 94% of the total intersections (total=254) when the penetration rate is 100%. The safety evaluation results showed that the standard deviation of speed is less for the CAV condition compared to the base condition. The real time crash-risk model for both freeways and arterials showed that CAV technologies can improve traffic safety at all levels of penetration rates and the maximum improvement was found for the full penetration level. Finally, the study provides an insight about the application of CAV technologies at the network level rather than in a small roadway and showed their benefits in terms of both safety and mobility.

In Chapter 5, the study investigated the impact of connected and automated vehicles (CAVs), automated vehicles (AVs) and the mixture of CAVs and connected vehicles (CVs) on the freeway corridor in the Aimsun next microsimulation platform. Therefore, different car-following models were utilized to approximate the behavior of CAV, AV, and CV. A CAV can get the vehicle information from multiple preceding vehicles within the communication range where AV has the ability to get information from the immediate-preceding vehicles. Additionally, the driving behavior of CV was modeled by utilizing the driver compliance factor. For the deployment of connected and automated vehicles, this study developed a testbed along the major roadway (SR417) in Orlando, Florida which was well calibrated and validated based on real-world traffic data (i.e., volume and travel time). For the traffic efficiency analysis, travel time data was used for the different market penetration rates (MPRs) of AVs, CAVs and mixture of CVs and CAVs. Also, the study developed ANOVA and generalized estimating equation (GEE) model for the traffic efficiency analysis. The results of the analysis showed that all the CAV technologies (CAV, AV,
and CV) can improve the travel compared to the field condition and maximum benefit was found for full penetration level. Additionally, for the same value of penetration rate, CAV scenario could reduce the travel time more than the AV scenario compared to the base condition which means connectivity features have a significant effect on travel time improvement. Also, the mix penetration level of CAV and CV showed that a similar improvement compared to the AV scenario. For the safety assessment, the study utilized Surrogate Safety Assessment Model (SSAM) to calculate the traffic conflicts. Based on this information, a Bayesian hierarchical zero-inflated negative binomial model was developed for the safety evaluation of CAV, AV, and CV. The results of the analysis showed that CAV, AV, and mixture of CAV and CV could improve traffic safety compared to the base condition and CAV outperforms AV for all penetration levels. Also, the study conducted a conflict analysis based on different vehicle types and found that CAV is safer compared to the AV scenario. Finally, the results of this study showed how the connectivity and automation improves both safety and efficiency compared to the only automation.

In Chapter 6, the dissertation conducted the simulation-based study by using real communication system (IEEE 802.11p) under CAV environment. The study developed a testbed in the SUMO simulation environment along the major roadway (SR408) in Orlando, Florida. To represent the real traffic condition of SR408, the research team calibrated and validated the simulation model based on volume and speed data, respectively. The driving behavior of CAV should be significantly different from non-CAV. Hence, the study used the intelligent driver model (IDM) in order to approximate the longitudinal car-following behavior of CAV. The application of CAV technology not only depends on the car-following behavior but also the uninterrupted communication between vehicles. Therefore, this study incorporates a real communication system by using OMNET++ simulation environment. SUMO and OMNET++ simulation platforms that
were coupled bi-directionally by using Veins. For the safety assessment, the study used time-to-collision (TTC) surrogate measure. A binary logistic regression model was developed by using the conflict data. Hence the study analyzed the safety effectiveness of CAV technology due to the communication disruption which is divided into two parts: (1) Communication aspect and (2) Traffic aspect. In the communication side, we used different values for the different communication parameters such as transmission power, queue size, etc., to investigate the impact of traffic safety. The results of the analysis showed that queue size has a significant effect on traffic safety and at the higher queue size value, the number of traffic conflicts was less under the CAV condition. Also, higher transmission power creates more interference which also creates more traffic conflicts. From the transportation aspect, the results showed that the number of communication collisions is higher for the lane closure scenario compared to the without lane closure scenario which causes more traffic conflicts. Moreover, different percentages of traffic flow i.e., (25%, 50%, 75%, and 100%) showed that with the increase in vehicle flow, the communication collision between packets is higher. Thus, the crash risk (conflicts between vehicles) could be higher with the increase in traffic flow. Finally, the results of this study provide an indication about the importance of communication system under CAV environments.

This research investigated the safety and mobility benefits of various CAV technologies and their applications in different ways by utilizing multiple simulation platforms. There is no study without limitations. Since the CAV technology is still not fully developed yet, the car-following models used in this study are not calibrated with the real CAV data. Also, the real-communication system could be improved through different control strategies. Hence, there is still scope for further investigation of this study.


7.2 Implications

This study developed a framework for the calibration and validation of a large-scale dynamic traffic assignment (DTA) model using Multi-Resolution Modeling (MRM) framework in Orlando, Florida. The model was developed by using the regional travel demand model (RTDM) in Orlando known as Orlando Urban Area Transportation Study (OUATS) with the base year of 2009. The study used origin-destination matrix estimation (ODME) process to project the base year data to the planning year by utilizing different detector systems. The model results showed a good calibrated and validated simulation model at each level (macroscopic, mesoscopic, microscopic). Such models were shown to allow investigation of active traffic management, infrastructure development, and decision support system in an integrated environment beyond the scope of existing regional traffic models. Also, the study provides a framework for the development of a large-scale model which could be helpful for transportation engineers.

The study showed also the application of connected and automated vehicles (CAVs) in a large-scale network by considering both freeways and arterials. The study utilized vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) technologies for the application of CAVs. Also, the study used a signal control algorithm to improve the performance of CAV at the intersection. The results of this study proved that CAV could improve both safety and mobility at the network level. Also, CAV could improve the freeway operation more compared to the arterials. Therefore, it is suggested that V2I technology could be useful to elevate the performance of CAV in arterial segments.
The analysis of different CAV technologies (CAV, AV, and CV) by multiple preceding vehicle information and driver compliance factor provided some important implications about CAV technologies. All the CAV technologies (CAV, AV, and CV) could reduce crash risk and improve traffic efficiency at the freeway segment. CAV outperforms AV, CV at all penetration levels and CAV behavior is safer under mixed traffic conditions compared to AV. Also, the high penetration rate of CV with low penetration rate of CAV showed similar improvement compared to the AV scenario. Finally, the finding of this study implies that CAVs with multiple preceding vehicle information provides more safe and efficient driving behavior compared to only automated vehicles (AV).

Finally, this research incorporated real communication system by using dedicated short-range communication (DSRC) for the application of connected and automated vehicles (CAV) environment. The results of this study proved that communication systems have a significant effect on traffic safety under the CAV environment. Also, different traffic conditions might deteriorate the communication between CAV which could increase the crash risk. Therefore, the finding of the study implies that not only the acceleration/deceleration behavior of CAV but also a successful communication between CAVs is necessary to improve the performance CAVs.
REFERENCES


Kesting, A., Treiber, M., Helbing, D., 2010b. Enhanced intelligent driver model to access the


Li, Y., Xu, C., Xing, L., Wang, W., 2017b. Integrated cooperative adaptive cruise and variable speed limit controls for reducing rear-end collision risks near freeway bottlenecks based on


Ma, J., Zhou, F., Huang, Z., James, R., 2018. Hardware-in-the-loop testing of connected and


Rahman, M.S., Abdel-Aty, M., 2018c. Longitudinal safety evaluation of connected vehicles’


https://doi.org/10.1016/j.trc.2019.01.029


Spiess, H., 1990. A gradient approach for the OD matrix adjustment problem. a\in \hat{A} 1, 2.


