Utilization of Connected Vehicle Data to Support Traffic Management Decisions

FDOT Project BDV29-977-21

Final Report

Prepared for

Florida Department of Transportation

By

Lehman Center of Transportation Research
Florida International University

and

Southwest Research Institute (SwRI)

September 2017

DISCLAIMER

The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the State of Florida Department of Transportation.

METRIC CONVERSION CHART

APPROXIMATE CONVERSIONS TO SI UNITS

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
LENGTH				
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
AREA				
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
VOLUME				
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
NOTE: volumes greater than 1000 L shall be shown in m^3				
MASS				
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
TEMPERATURE (exact degrees)				
^oF	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^oC
ILLUMINATION				
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
FORCE and PRESSURE or STRESS				
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Utilization of Connected Vehicle Data to Support Traffic Management Decisions		5. Report Date September 2017
		6. Performing Organization Code
7. Author(s) Mohammed Hadi, Ph.D., P.E.; Yan Xiao, Ph.D., P.E.; Md Shahadat Iqbal; Samaneh Khazraeian; Purser Sturgeon II		8. Performing Organization Report No.
9. Performing Organization Name and Address Lehman Center for Transportation Research Florida International University 10555 W. Flagler Street, EC 3680, Miami, FL 33174 Southwest Research Institute 6220 Culebra Rd, San Antonio, TX 78238		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. BDV29-977-21
12. Sponsoring Agency Name and Address Office of Research and Development State of Florida Department of Transportation 605 Suwannee Street, MS 30, Tallahassee, FL 32399-0450		13. Type of Report and Period Covered Final Report February 2015 – November 2017
		14. Sponsoring Agency Code

15. Supplementary Notes
Ms. Elizabeth Birriel at the Florida Department of Transportation served as the project manager for this project.

16. Abstract
Given the accelerated advancements in connected vehicles (CV), there is a need for the Florida Department of Transportation (FDOT) to start preparing for the next generation of advanced traffic management strategies that utilize connected-vehicle-to-infrastructure (V2I) technologies. This project investigated the utilization of data from connected vehicle technologies in support of SunGuide traffic management center operations. An assessment of the data that are currently available and are expected to be available from connected vehicles was first conducted. Nine SunGuide functions that make use of field data and can possibly be supported by connected vehicle data were reviewed, and discussions are provided on how connected vehicle data can be used to support these SunGuide modules. This project then demonstrated the used of connected vehicle data in supporting traffic management center processes by utilizing combinations of the collected data from onboard units and emulated CV data based on real-world vehicle trajectories, as well as microscopic simulation. The examined methods and applications included the prediction of future year link level CV proportions, the benefits of queue warning in a CV environment based on surrogate safety measures, and the time lines when the CV technology can provide sufficient data quality to replace existing data acquisition systems in terms of travel time estimation, incident detection, and traffic volume estimation. As part of the project, a field test was conducted by installing cell-based CV technologies in the service patrols of the Florida Department of Transportation (FDOT District 5). CV data were collected, transferred in real-time to a central server, processed, and used in an analysis of mobility measures of a roadway section along the I-4 in Central Florida.

17. Key Word	18. Distribution Statement		
Connected Vehicle Data, Traffic Management, Performance Measurement, Travel Time	Unrestricted		
19. Security Classif. (of this report)	**20. Security Classif. (of this page)**	**21. No. of Pages**	**22. Price**
Unclassified	Unclassified	199	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

EXECUTIVE SUMMARY

The accelerated advancements in connected vehicle (CV) technologies are expected to have significant influence on agency management and operations, starting in the next few years. There is a need for the Florida Department of Transportation (FDOT) to start preparing for the next generation of advanced traffic management strategies that utilize connected-vehicle-to-infrastructure (V2I) technologies. The goal of this project is to investigate the utilization of data from connected vehicle technologies in support of SunGuide traffic management center operations. A number of tasks were conducted in this project to achieve this goal; they are summarized below.

Connected Vehicle Data Assessment

An examination of the data that are currently available and are expected to be available from connected vehicles was performed through a detailed review of the literature. An analysis of the connected vehicle data collected using the SunGuide software from the Orlando deployment and archived in the Research Data Exchange (RDE) database and additional data collected from the Safety Pilot deployment in Ann Arbor, Michigan, was also conducted in this study. The analysis of the Orlando CV data reveals the importance of data preprocessing and cleaning due to the existence of erroneous CV data. The measures of travel time, speed, and acceleration distributions along the I-4 WB study corridors were calculated from the Orlando CV data. However, these results are limited by the small sample size of the data.

SunGuide Modules and Potential Use of Connected Vehicles

Nine SunGuide functions that make use of field data and can possibly be supported by connected vehicle data were reviewed in this study. Below is the list of SunGuide Modules that were identified for this purpose.

- Incident detection
- Traffic detection and performance measurements
- Ramp metering
- Managed lanes
- Road Ranger support
- Road weather information system
- Variable speed limit
- Wrong way driving
- Safety barrier device

The required input data for these modules and the sources for these data were summarized in this study. Based on literature review, recommendations were provided on how connected vehicle data can be used to support these SunGuide modules.

Orlando Connected Vehicle Deployment and Data Analysis

One important task of this project was to decide on what communication technology to use in the demonstration of this project for transmitting data between connected vehicles and the central server. Based on the review conducted in this study, it is concluded that both dedicated short range communications (DSRC) and cellular communications can be used to support most dynamic mobility applications. In this project, Onboard Units (OBUs), previously purchased as part of the World Congress deployment, were activated or re-installed on FDOT District 5 Road Ranger service patrol vehicles and cellular modems were also installed to allow the communications of CV data to a central server. The OBUs were also connected to the vehicle Onboard Diagnostic II (OBD-II) to collect additional CAN data. Tests were conducted to determine what additional CAN data elements can be retrieved from the OBD-II port. The collected data from this field deployment were processed and used in an analysis of the mobility measures of a roadway section along the I-4 eastbound (EB). The comparison of the speed measurements collected from CV data and those from detector data indicates that connected vehicle data not only can be used to supplement the measurements of point traffic detectors along the freeway mainline and ramps, but also can be applied to verify the accuracy of detector data.

Assessment of Link Level Variation of CV Market Penetration

Estimation of the market penetration of connected vehicles is important to identifying the impacts of these technologies. Past efforts have assumed the growth in the CV market penetration without considering the variations in the socioeconomic characteristics between regions and zones within a region. This study proposes a methodology to determine the variations of the CV market penetrations between regions, zones within a certain region, links within the region, and time-of-day. The methodology can be implemented with various CV implementation scenario assumptions and considers the variations in the socioeconomic characteristics of travelers of a region. Applying this methodology to a case study indicates that the distribution of the link-specific CV market penetration follows a lognormal distribution. The percentage variation in the market penetration between zones and links is shown to be the highest in the first year of CV implementation and decreases exponentially with the number of years passing since the implementation. The market penetration variations between links are the highest on collectors, followed by arterials, followed by freeways. The study also shows that the average percentage increase in the CV market penetration grows in the first several years then remains almost constant before dropping sharply.

Accuracy and Reliability of Estimated Travel Time Using Data Collected from Connected Vehicles

The expected implementations of connected vehicles in the next few years will provide a promising alternative to existing technologies in providing data for travel time estimation. However, it is important to assess the data quality provided by this new technology compared to existing technology. This study assessed the quality of travel time estimates based on CV data on freeway and urban street segments. The data quality was examined under different market penetration scenarios considering the randomness in CV presence on the links and the variation in the market penetration between links in the same region due to the variation in the socioeconomic characteristics of the zones in the region. Based on the results of the study, it can be stated that the CV market penetration will be sufficient for use in planning and real-time operations of the investigated freeway segment in the first year after the expected mandate for installing CV technology on all new vehicles will become effective. However, for the urban street, it will take one to three years for the data quality to be sufficient for use for planning purposes and three to six years for operation purposes depending on the market penetration (MP) of the CV, considering the variations in the socioeconomic characteristics in the region.

Assessment of the Benefits of Queue Warning in a CV Environment Based on Surrogate Safety Measures

Queue warning systems (QWS) have been implemented to increase traffic safety by informing drivers about the queued traffic ahead, so that they can react in a timely manner to the presence of queues. Existing QWS rely on fixed traffic sensors to detect the back of queue. It is expected that if the transmitted messages from the connected vehicles are utilized for this purpose, the detection can be faster and more accurate. In addition, with connected vehicles, the delivery of the messages can be done using onboard units instead of dynamic message signs (DMS), providing more flexibility on how far upstream of the queue the messages are delivered. This study investigated the accuracy and benefits of the QWS based on connected vehicle data. The study evaluated the safety benefits of the QWS under different market penetrations of CV in future years based on safety surrogate measures estimated using simulation modeling combined with the Surrogate Safety Assessment Model (SSAM) tool. The results from this study indicate that a relatively low market penetration, around 3% to 6% for the congested freeway examined in this study, is sufficient for accurate and reliable estimation of the queue length. Even at 3% market penetration, the CV-based estimation of back of queue identification is significantly more accurate than that based on detector measurements. It is also found that CV data allows faster detection of the bottleneck and queue formation. Further, it is concluded that the QWS improved the safety condition of the network by reducing the number of rear-end conflicts. The safety impacts become significant when the compliance percentage with the queue warning messages is more than 15%.

A Methodology to Assess the Quality of Travel Time and Incident Detection Based on CV Data under Different Demand Levels

This study investigated the use of CV data as an alternative to existing data acquisition techniques in providing two critical functions, travel time estimation and incident detection, in order to support Transportation Management Center (TMC) under different demand levels. In support of this investigation, the study developed regression models to estimate travel time measurement accuracy, travel time measurement reliability, and incident detection latency as functions of the traffic demand level and the CV proportion in the traffic stream. The developed regression models were used in conjunction with a prediction of CV proportions in future years to determine when the CV technology can provide sufficient data quality to replace existing data acquisition systems under different demand levels. The results can be used by Transportation System Management and Operations (TSM&O) programs and agencies to plan their investment in data acquisition alternatives in future years.

Identifying a Timeline for Future Utilization of Connected Vehicle Data to Support Traffic Volume Estimation on Urban Streets

Although it will be possible to estimate measures such as speed, travel time, delay, and number of stops with relatively low market penetrations of CV, a high CV proportion in the traffic stream will be needed to estimate parameters such as traffic volume and density at the required levels of accuracy. This study developed a method to determine the approximate time in the future when connected vehicle data can replace or complement existing detectors on urban streets in estimating segment traffic volumes on urban streets. The result from applying the methodology to a case study indicates that after four years of the mandate of CV of new vehicles, CV data can be used to improve the estimation of volumes on the street links with no detectors and without removing the existing detectors on the other links. Depending on the adopted volume accuracy thresholds utilized by agencies, it will be possible to start removing some of the detectors after 5 to 8 years. The agencies can remove all detectors for volume measurements after 10 to 15 years, depending on the accuracy threshold.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The Florida Department of Transportation (FDOT) SunGuide centers are increasingly investing in the implementation of advanced strategies and associated technologies, realizing the significant benefits and cost effectiveness of these implementations. The SunGuide Traffic Management Centers (TMCs) of the FDOT are considered to be the nerve centers for monitoring and managing traffic and incidents and disseminating information to the traveling public. There are currently 12 TMCs in Florida (FDOT, 2016).

The FDOT TMC functionalities are enabled by the SunGuide® software which supports the management of operating equipment, integration of data sources, and automating incident detection and response (FDOT, 2014). Emerging connected vehicle (CV) technologies promise to provide a significant increase in data quantity and quality to support real-time operations and off-line planning of transportation systems. The FDOT demonstrated a connected vehicle implementation at the 2011 Intelligent Transportation System (ITS) World Congress in Orlando, FL.

The demonstration included onboard units (OBU), dedicated short range communications (DSRC) roadside units (RSU) connected to a FDOT fiber network, backend servers at FDOT District 5 RTMC, and an enhancement made to the SunGuide software to allow it to capture and process the connected vehicle data. With the anticipated introduction of connected vehicles in the transportation system, these technologies are expected to have significant influence on agency management and operations, starting in the next few years. Thus, the FDOT needs to start preparing for the next generation of advanced traffic management strategies that utilize connected vehicle-to-infrastructure (V2I) technologies. The draft Federal Highway Administration (FHWA) V2I (vehicle-to-infrastructure) Deployment Guidance (FHWA, 2016) encourages V2I deployments, but it states that the United States Department of Transportation (USDOT) will not require public agencies to implement V2I technology or applications, and recommends that this implementation should be based on agency assessments. The National CV Field Infrastructure Footprint Analysis document produced by the American Association of State Highway and Transportation Officials (AASHTO) (Wright et al., 2014) stated that "Public agencies will assess and trade-off the opportunities to use connected vehicle probe data aggregation and processing versus the continued deployment, operations and maintenance of traditional ITS vehicle detection versus purchasing commercial traffic information services." Given the accelerated advancements in technologies, there will be increasing challenges for agencies to determine the trade-offs between existing and emerging technologies.

The goal of the "Utilization of Connected Vehicle Data to Support Traffic Management Decisions" project funded by the FDOT, was to investigate the utilization of data from connected vehicle technologies in support of SunGuide traffic management center operations. The specific objectives of the project are:

- Examination of the data that can be obtained from connected vehicles at the present time and in the future
- Identification of the SunGuide traffic management center processes and software modules that will benefit from the collected data and how the data can be used to support these processes and modules
- Examination of the current connected vehicle deployment in Orlando in relation to its ability to support traffic management operations and identifying potential updates to this deployment
- Comparing the use of DSRC and cellular technologies in communicating probe data from connected vehicles to traffic management centers
- Providing recommendations regarding future developments of the SunGuide system as related to connected vehicles

To achieve the above goal, a number of tasks were conducted in this project. The results of these tasks are documented in the remaining sections of this report, as shown below.

- Section 2 presents an assessment of the data that are currently available and are expected to be available from vehicles, an analysis of the connected vehicle data collected using the SunGuide software from the Orlando deployment and archived in the Research Data Exchange (RDE) database, and an analysis of additional data collected from newer deployments like the Safety Pilot deployment in Ann Arbor, Michigan.
- Section 3 lists an inventory of SunGuide modules that are supported by field data and have the potential to be supported by CV data.
- Section 4 discusses how CV data can be used to support the SunGuide modules.
- Section 5 reviews the current state of the connected vehicle technologies utilized in the Orlando implementation and describes the use of connected data collection processes and cellular communication technology to transmit probe vehicle data to central location. An analysis of the collected CV data is also presented in this section.
- Sections 6 to 10 demonstrates the effectiveness of connected vehicle data in supporting selected traffic management center processes by utilizing combinations of the collected data from the onboard units and data from existing traditional and ITS infrastructure as well as simulation. Specifically, Section 6 presents an assessment of link level variation of connected vehicle market penetration. Section 7 discusses the accuracy and reliability of estimated travel time using basic safety message data. Section 8 assesses the benefits of queue warning in a connected vehicle environment based on surrogate safety measures. A methodology is developed in Section 9 to determine when the CV technology can provide sufficient data quality to replace existing data acquisition systems in terms of travel time estimation and incident detection. Similar methodology is further applied in Section 10 to identify a timeline for future utilization of CV data to support traffic volume estimation along urban streets.

2. CONNECTED VEHICLE DATA ASSESSMENT

This section presents an assessment of the data that are currently available and are expected to be available from vehicles.

2.1. Connected Vehicle Data Elements

The connected vehicle (CV) message types and components are specified in the Society of Automotive Engineers (SAE) J2735 standards (SAE International, 2016). The latest standards are the fifth edition.

The J2735 standards specify a number of message types. The basic safety message (BSM) is one of these message types that will be used for vehicle-to-vehicle communication. The BSM contains vehicle safety-related information broadcasted to surrounding vehicles, but can also be sent and/or captured by the infrastructure. The BSM, as defined in the J2735 standards, consists of two parts. Part 1 is sent in every BSM message broadcasted 10 times per second and will be mandated to be broadcasted by the NHTSA ruling. It contains core data elements, including vehicle position, heading, speed, acceleration, steering wheel angle, and vehicle size. BSM Part 2 consists of a large set of optional elements such as precipitation, air temperature, wiper status, light status, road coefficient of friction, Antilock Brake System (ABS) activation, Traction Control System (TCS) activation, and vehicle type. BSM Part 2 elements are sent based on criteria that are not specified in the J2735 standards. However, not all of these parameters are currently available from vehicles, as described later in this section, and they will not be mandated by the USDOT.

Figure 2-1 shows the message format of the BSM (Hong at al., 2014). Table 2-1 shows the potentially useful data elements of the BSM Part 1 and BSM Part 2. A preliminary assessment of the use of BSM messages to support Dynamic Mobility Applications (DMA) found that BSM Part 1 is useful for a limited subset of mobility applications, but is not solely sufficient for most applications (McGurrin, 2012). However, Barbaresso and Johnson (2014) pointed out that simple probe data, such as vehicle location, speed and heading, can be useful for very important applications like traffic monitoring, Advanced Traveler Information Systems (ATIS), traffic signal timing analyses, and for planning purposes.

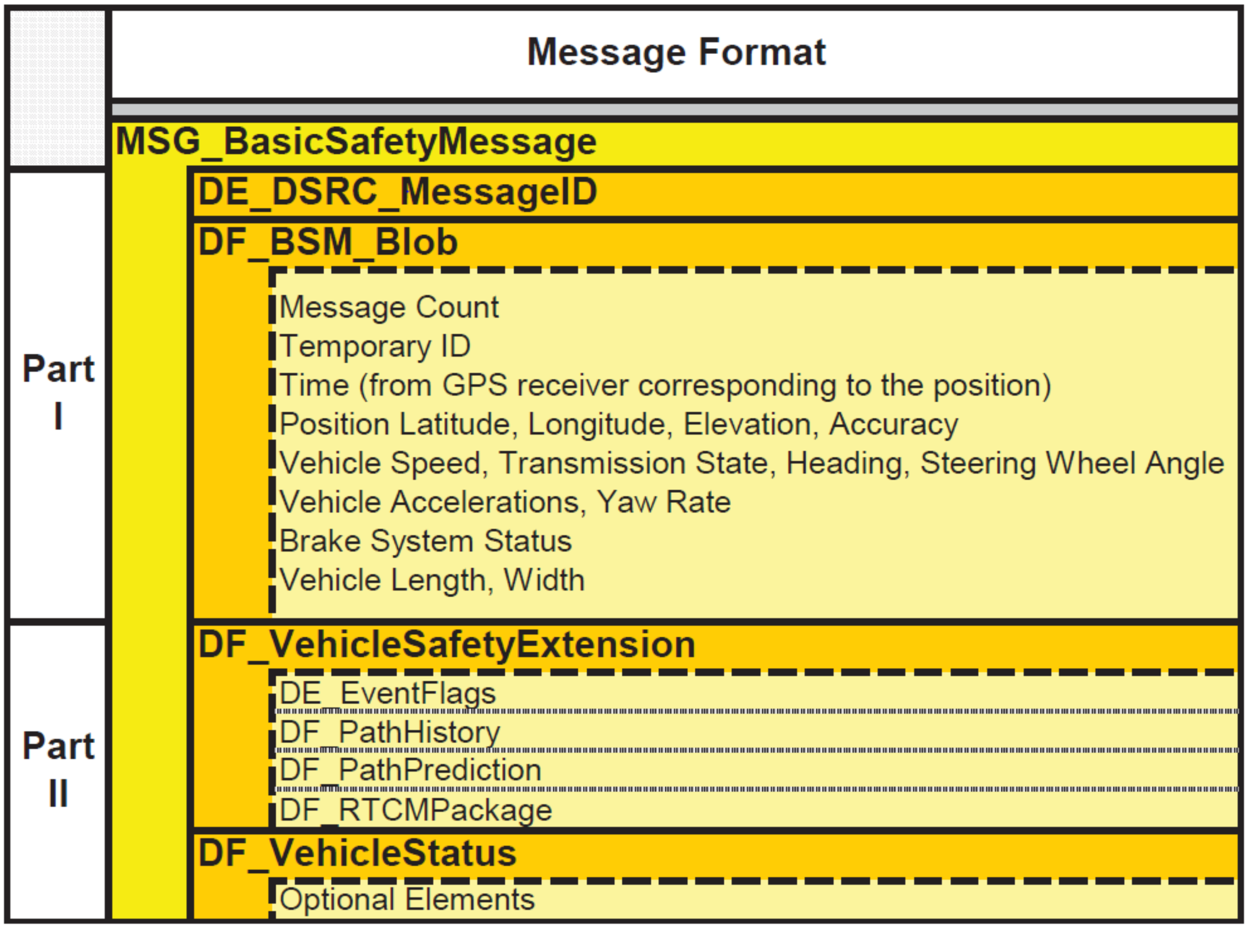


Figure 2-1 The Format of Basic Safety Messages (Hong et al. 2014)

Table 2-1 Potentially Useful BSM Parameters Specified in J2735 (McGurrin, 2012)

Data	# Applications	BSM Part 1	BSM Part 2	Exchange Type
Airbag deployment	2		✓	Event driven
Ambient air pressure	8		✓	Periodic
Ambient air temperature	9		✓	Periodic
Antilock brake system active over 100 msec	9	✓		Event driven
Brake system status	3	✓		Periodic
Cargo weight	2		✓	Periodic
Compliance with target speed	1			Periodic
Confidence-position	2		✓	Periodic
Confidence-speed/heading/throttle	2		✓	Periodic
Confidence-time	2		✓	Periodic
Current lane	1			Periodic
Descriptive vehicle identifier	1		✓	Periodic
Engine RPM	1			Periodic
Engine torque	1			Periodic
Exhaust diagnostics	1			Periodic
Exterior lights (status)	8		✓	Periodic
Fleet owner code	1		✓	Periodic
Hazard lights active	2		✓	Periodic
HAZMAT status	1		✓	Periodic
Impact sensor status	1			Event driven
Incident report from traveler	1			Event driven
Level of brake application	2		✓	Periodic
Lights changed	8		✓	Event driven
Motion	11	✓		Periodic
Pavement temperature	1			Periodic
Position (local 3D)	14	✓		Periodic
Rain sensor	9		✓	Periodic
Rate of charge of steering wheel angle	3		✓	Periodic
Recent or current hard braking	2		✓	Event driven
Relative humidity	1		✓	Periodic
Road coefficient of friction	10		✓	Periodic
Traction control system active over 100 msec	11		✓	Event driven
Trailer weight	3		✓	Periodic
Vehicle data	2		✓	Periodic

Table 2-1 Potentially Useful BSM Parameters Specified in J2735 (McGurrin, 2013) (Continued)

Data	# Applications	BSM Part 1	BSM Part 2	Exchange Type
Vehicle mass	4		✓	Periodic
Vehicle placarded as HAZMAT carrier	1		✓	Periodic
Vehicle size	6	✓		Periodic
Vehicle type (fleet vehicles)	6		✓	Periodic
Wiper status	9		✓	Periodic
Wiper charged	9		✓	Event driven

There are a number of CV message types specified in the J2735 standards, in addition to the BSM messages. The J2735 Standard Fact Sheet (USDOT, 2015a) provides an overview of these messages. However, these message types will not be mandated by the USDOT and are not addressed by the upcoming FHWA V2I Deployment Guidance (FHWA, 2015). Table 2-2 presents a list of the J2735 message types and their applicability to V2V, V2I, and I2V communication. As one of CV message types, the probe vehicle data message contains snapshots of vehicle information and sensor data collected from and sent to a vehicle's onboard unit. These snapshots can be generated in three ways: periodically, event triggered, and during starts and stops. This or a similar type of messages can be used if the information is cached on the vehicle and sent at less frequent interval to the infrastructure.

Table 2-2 SAE J2735 Messages and Communication Modes (AASHTO, 2013)

	V2V	V2I	I2V
Basic Safety Message Part 1	✓	✓	
Basic Safety Message Part 2	✓	✓	
Emergency Vehicle Alert	✓		
Common Safety Request	✓		
Probe Vehicle Data		✓	
Signal Request Message		✓	
Roadside Alert			✓
Traveler Information			✓
MAP Data			✓

It is possible that several data items that connected vehicle messages, as defined in SAE standards, are not able to provide will instead be provided by connected travelers that opt to do so. Due to privacy concerns, the J-2735 standards specify changing the vehicle identification at frequent intervals. Thus, information like vehicle routing and origin-destination cannot be obtained based on connected vehicle data collected according to these standards. Information service providers are already collecting data by using location and speed data from mobile devices belonging to captive vehicle fleets or travelers that opt in to provide data in exchange for information services (Wright et al., 2014). This type of data has been used to estimate travel time, and in some cases, origin-destination trips. New data collection and processing methods will provide estimation of additional information that is useful for transportation network management and next generation traveler information system. The EnableATIS Operation concepts describe an environment where disaggregate traveler decision data and vehicle/traveler itinerary data and preferences are collected from the travelers (Adler et al., 2014). The USDOT funded two research efforts: one conducted by the University of Minnesota and the other by Massachusetts Institute of Technology (MIT). These efforts identified methods and technologies to infer disaggregate traveler behavior from the collected data.

2.2. DSRC versus Cellular Based Technologies

The CV Infrastructure Footprint Analysis conducted by AASHTO estimated that it will take until about 2040 for the full deployment of DSRC and even then there will not be a full coverage of the network with DSRC (Wright et al. 2014). Around the time of the full deployment, it is expected that 80% (250,000) of the nation's traffic signal locations and up to 25,000 other roadside locations will be V2I-enabled. Non-signalized locations with significant vehicle-infrastructure interactions are also candidates for deployment. On freeways, the above mentioned study of Wright et al. (2014) assumed that 50% of the existing ITS sites on the freeways such as vehicle detection stations, DMS, CCTV, and environmental sensor stations will be equipped with DSRC units. Installing DSRC at these locations and at intersection locations provides an opportunity to leverage existing power sources, cabinet space, and backhaul communications, which will minimize deployment costs. The USDOT V2I Guidance document provides additional guidance on the deployment of the road side units (FHWA, 2015).

As stated above, as currently envisioned, there will be no full network coverage with DSRC units (Wright et al., 2014), at least for years to come. Thus, both DSRC and non-DSRC technologies (e.g., cellular, Wi-Fi) are expected to be used to support mobility applications. Many mobility applications do not require the information to be received by the infrastructure at 1/10 seconds. The information needs to be either continually sent to the infrastructure, at a lower update rate, if alternative mobile wireless technology such as cellular is used, or be cached onboard the vehicle and then sent in a message containing both current and stored snapshots (McGurrin, 2012). With regard to the use of cellular technologies, automakers are increasing building applications in their vehicles to provide navigation, notification of required maintenance, emergency services, and in-vehicle "infotainment" (Wright et al., 2014) utilizing fourth-generation (4G) mobile communications. 4G commercial networks are increasingly being implemented around the nation. These implementations will further support data collection and traveler information systems. It is interesting to note that Toyota announced in January 2016 that they are investing in satellite vehicle connectivity, considering that only 10% of earth is covered by 4G technology. Toyota estimates that the technology could be market ready between 2020 and 2025 at the same time as 5G LTE is expected to be available. There has been increasing discussion of the use of 5G cellular technologies that are expected to start to be available in the near future. Such technologies will play an important role in Smart Cities and the Internet of Things applications and will allow V2V and vehicle to everything (V2X) communications including vehicle to the infrastructure, pedestrians, cell tower, and the network. The 5G Automotive Association has been formed to support the use of 5G in CV applications and include OEMs, cellular industry companies, and third partner technology vendors. Public agencies should continue monitoring the trends in using DSRC versus 5G technologies, the proposed utilization, technology progress, the business models, and the roles that they can play with each technology deployment.

Public sector respondents to a Michigan Delphi study (MDOT, 2012) indicated that the respondents believed that DSRC will be used for urban intersections while cellular technology will be the more likely transmission mode for urban freeways. The respondents also indicated that 3G and 4G cellular technology will be the main communication medium for probe data collection, fleet management, commercial and private applications, and asset management while DSRC will be used for in-vehicle safety warnings. Barbaresso and Johnson (2014) pointed out that some interviewed public agency staff believed that DSRC is only needed for safety-critical, low-latency applications and that cellular technology may provide an alternative for many mobility and traveler information applications that do not require V2V communications.

Cronin (2012) suggested that the data should be transmitted via DSRC when available and cellular otherwise and presented the diagrams in Figures 2-2 and 2-3 to illustrate the combined transmission of data from private and public vehicles, respectively. The European Cooperative ITS (C-ITS) platform document (European Commission, 2016) pointed out that currently neither DSRC nor cellular systems can provide the full range of necessary services for C-ITS. The document concluded that a hybrid communication concept is needed to take advantage of these technologies with the use of the existing cellular communications infrastructure to increase the geographic coverage to locations without DSRC deployment.

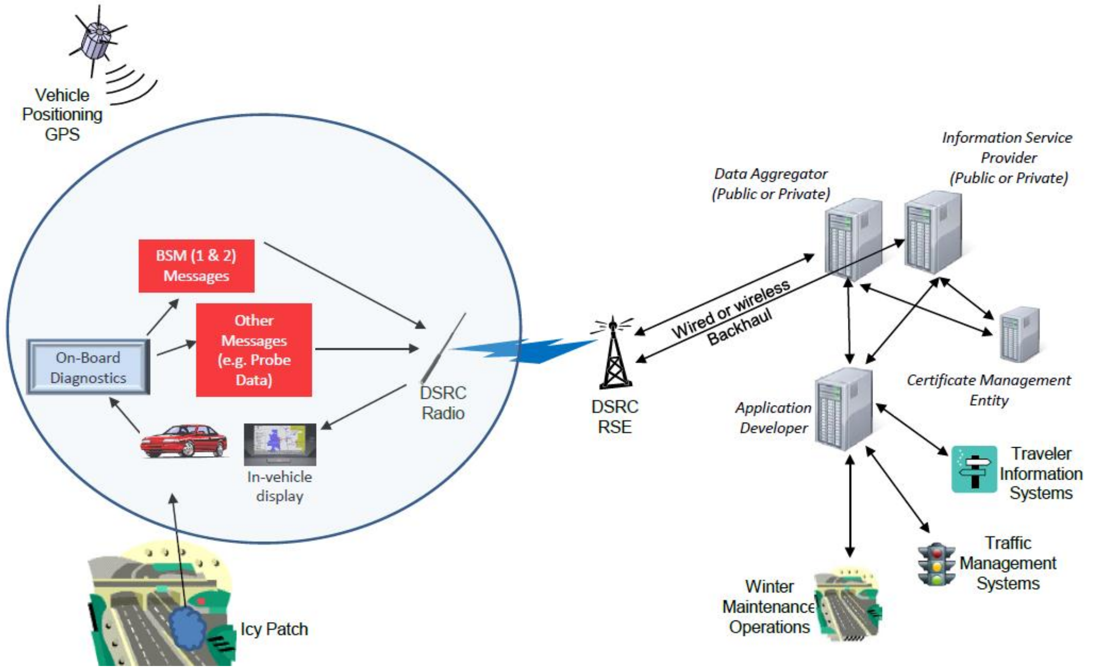


Figure 2-2 Combined DSRC and Cellular Transmission of Data from Private Vehicles (Cronin, 2012)

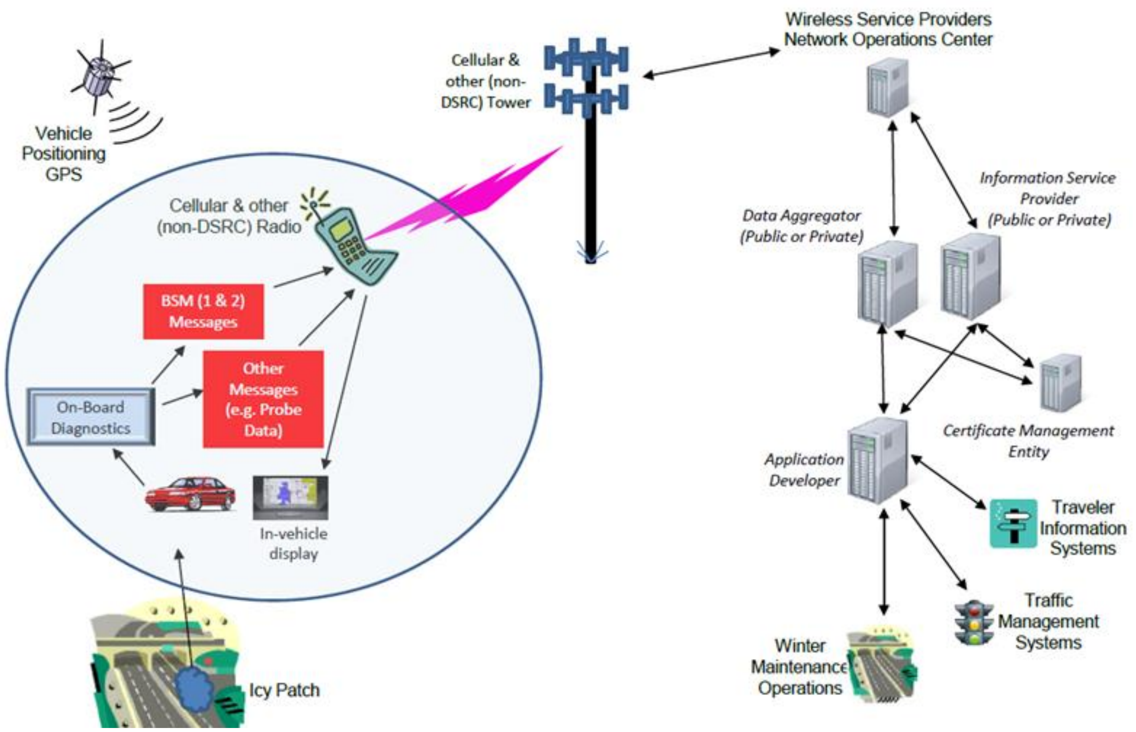


Figure 2-3 Combined DSRC and Cellular Transmission of Data from Public Sector Vehicles (Cronin, 2012)

As an example of utilizing the cellular network, instead of DSRC, the Colorado Department of Transportation (CDOT) and a private sector company announced in January 2016 the implementation of a cellular-based connected vehicle alert system on the I-70 Mountain Corridor as part of CDOT RoadX project, with 1,000 vehicles participating in the implementation. The private sector system collects, analyzes, and distributes data to travelers and CDOT traffic management systems. Data is collected from onboard sensors as well as infrastructure-based detectors. This implementation is similar to the Finnish Transport Agency's Nordic Way project that is also based on the same private sector technology. A similar project is planned in the Netherlands.

2.3. Data Accessibility

V2V BSM messages sent through DSRC can be captured by an RSU, where the connected vehicle is transmitting data in the vicinity of an RSU. This information can then be sent to traffic management centers for use in different traffic management applications over the backhaul communications link. These data may also be provided directly to the center element by a vehicle using the wide area link. The BSM Part 1 message is the only message which is widely recommended to be transimitted (AASHTO, 2013). It has been reported that data collected using BSM Part 1 via DSRC are not sufficient for all dynamic mobility applications. Many of the applications require data beyond the BSM Part 1. However, as stated earlier, BSM Part 1 data are useful for several applications**.** Although additional messages such as the BSM Part 2 and the probe vehicle data message have been proposed, the details on the availability and transmission of these messages are not defined yet.

Messages delivered through wide-area communications such as those in the J2735 standards, the previously proposed Basic Mobility Message (BMM), and opt-in data provided by travelers can provide significant additional support of mobility applications. As it stands now, the BSM messaging will be mandated and regulated for safety applications only, but other types of messages will likely be market driven. Thus, there is a need for a business model to accompany the wide area messaging mentioned earlier (Thompson, 2013). At least some information regarding various aspects of this issue is expected to be available in the near future, based on on-going efforts and coordination between OEMs, third party operators/service providers, and public agencies. An effort by the European Commission found that there are strong disagreements between vehicle manufacturers and the third party providers on issues such as data accessibility, onboard application platform, and the central data server platform. Also, there are various views regarding data accessibility implementation and the need and details of possible legislation (European Commission, 2016).

Most existing passenger vehicle models are equipped with a controller area network (CAN), which is a vehicle bus that allows communications and data sharing between vehicle components. Although, the bus contains a variety of data that can be used to populate V2V and V2I messages, the ability to access the CAN data varies by manufacturer and model. In addition, the definitions of some data elements are not standard. Furthermore, not all data goes through the CAN bus. The Onboard Diagnostic (OBD) connector has been used for open real-time data access for the automotive aftermarket community utilizing the OBD II standard. Currently, a standard 16 PIN onboard diagnostics provides access to in-vehicle data and used mainly for diagnostic and repair purposes and for the monitoring control of emission. There are commercial retro-fit devices for wireless transmission of in-vehicle data (plugged into the OBD port) to access the vehicle data. The information available through the OBD is different by different vehicle manufacturers and in some cases the access to the information is through proprietary connections to OEM servers (European Commission, 2016). There is a need for standards for upgraded OBD interface connectivity to allow collecting and processing the data onboard the vehicles. The market implementation has been estimated to be five years after the availability of the necessary standards (European Commission, 2016).

Depending on how the data accessibility issues will be resolved (e.g. setting new standards for data accessibility), it is possible that there will be some differences between the amount of data that can be accessed by different aftermarket vehicle devices, depending for example if the aftermarket device is OEM approved or it may be limited to data available through the vehicle's onboard diagnostics (OBD-II) connector. Systems without access to OEM data are likely to be limited in functionality due to the lack of access to sensors generally available within a vehicle (AASHTO, 2013).

Even if the data is available through the OBD-II, the lack of standard definitions and format make the access of the information very difficult. There are currently aftermarket companies that are specialized in reverse engineering data on the CAN bus to allow aftermarket devices to be integrated into the vehicle (Hong et al., 2014). Table 2-3 shows sample data elements available from a single CAN bus on each vehicle that were used to populate BSMs within the Southeast Michigan Test bed.

Table 2-3 USDOT Test Bed Vehicle BSM Availability (Hong et al., 2014)

Jeep	Ford
BSM Part I	**BSM Part I**
BrakeSystemsStatus\StabilityControlStatus	AccelerationSet4Way
SpeedAndTransmission	SteeringWheelAngle
SteeringWheelAngle	BrakeSystemsStatus\BrakeAppliedStatus
AccelerationSet4Way	SpeedAndTransmission
BrakeSystemsStatus\BrakeAppliedStatus	BrakeSystemsStatus\TractionControlState
AccelerationSet4Way	BrakeSystemsStatus\TractionControlState
	BrakeSystemsStatus\StabilityControlStatus
	BrakeSystemsStatus\AntilLockBrakeStatus
BSM Part II	**BSM Part II**
ExteriorLights	ExteriorLights
ThrottlePosition	ThrottlePosition
WiperStatusFront	AmbientAirTemperature
AmbientAirTemperature	AmbientAirPressure
AmbientAirPressure	

Another issue is the central server platform that will receive, store, and distribute the collected data from private vehicles through wide-area network. There is a strong feeling that this will be established based on public-private sector initiatives. However, the details about the transmission of the data to this platform, data provided from the vehicles to this platform, the management and control of the platform, vehicle sharing, and the associated business models have not been established.

2.3.1. Communication Technology Performance

An important component of V2I data communications with different technologies is to determine the data loss and data latency. This is an important consideration when selecting and implementing technologies and also when simulating the technologies, as will be done in this study. An analysis of a test performed in 2009 of the 5.9 GHz DSRC technology performance (SAE, 2009) shows that the average data loss for OBU to RSU transmission via DSRC is about 12% with typical values in the 10 to 20% range. Based on experience with current DSRC implementations, a good assumption for use in a simulation environment, the data packet loss is in the 5-10% range. This is largely dependent on the availability of line of sight and the distance between the vehicle and the RSU. Close range communications is expected to be at 5% or better when within 3,300 ft and the line of sight. Longer distances can increase up the data loss by 10% to 50%, depending on the line of sight issues.

The latency of the DSRC communication is affected by whether it is using the Wave Short Message (WSM) transmission or it is establishing an IP link over a WAVE Basic Service Set (WBSS). The WSM transmission is normally used for BSM communication and generally gives better performance, but is more size constrained (less than 1400 byte payloads). The IP over WBSS transmission is appropriate for longer probe data messages. An AASHTO report states that the latency of a DSRC system depends on the channel congestion and that the average latency is between 5 msec and about 100 msec, for a non-switched system, and 55 msec to 100 msec for a switched system (AASHTO, 2013). The same report states that the LTE latency is 35 msec to 60 msec. It appears based on the above discussion that although DSRC technology latency is widely recognized as better than the cellular option; the cellular option latency is reasonably close to the average DSRC latency. The cellular option is adequate for many V2I and I2V applications.

2.3.2. Prediction of Technology Market Penetration

DSRC technology adoption scenarios were developed with consideration of the timing of the NHTSA DSRC mandate decision. Some vehicles with embedded DSRC are expected in the 2017 models, and all new models could be equipped with the technology by 2020. Figure 2-4 illustrates the fraction of connected vehicles within the vehicle fleet under different assumptions, as presented by Wright et al. (2014). It is important to recognize that vehicle cellular connectivity is expected to increase ahead of the DSRC implementations. As stated earlier, wide area communication using cellular technology is expected to play an important role in providing V2I and I2V connectivity for many mobility applications. In addition, an important consideration that has not been considered when discussing market penetration is that although the market penetration in a region on average could be a certain percentage, the actual penetration on a given link can vary depending on the socioeconomic characteristics of the travelers on the link. The market penetration of DSRC and cellular connectivity will be discussed in a later section.

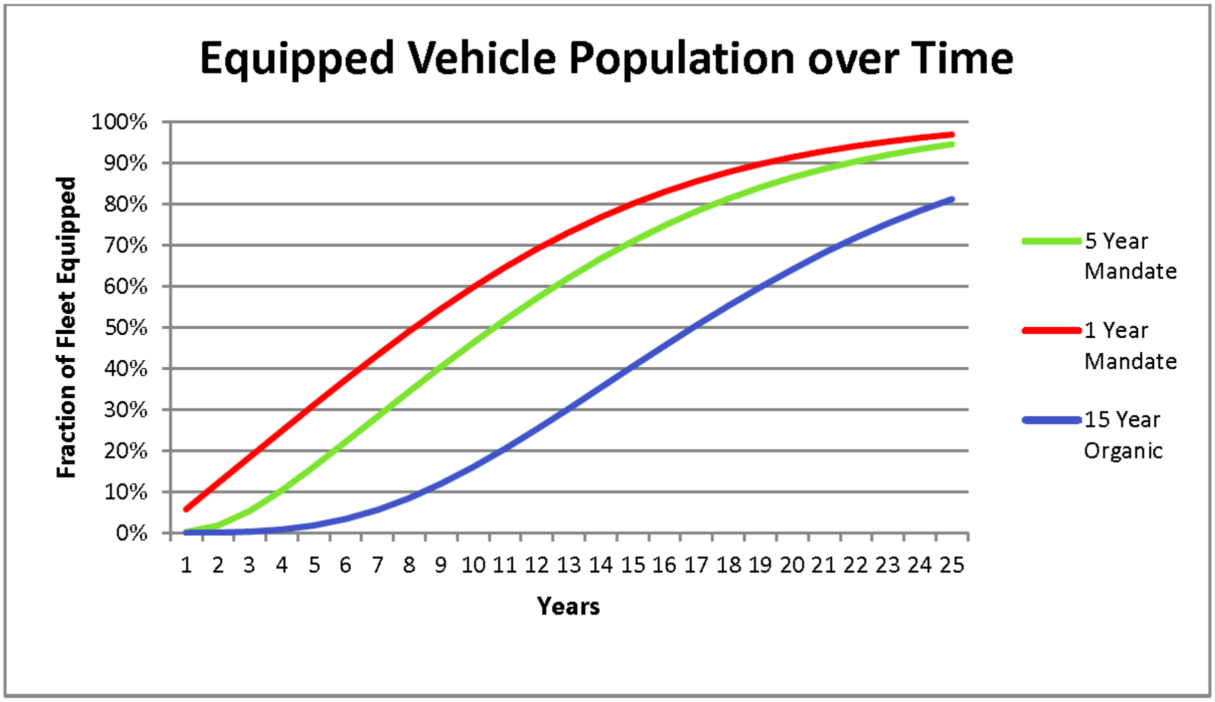


Figure 2-4 Equipped Vehicle Population over Time (Wright et al., 2014)

2.4. Analysis of Data Currently Available in the USDOT Research Data Environment (RDE)

The Research Data Exchange (RDE) is a transportation data sharing system that provides data from different sources to support ITS researchers, developers including connected vehicle research activities. Currently, the RDE website includes the following data sets:

- Safety Pilot Model Deployment Data
- ITS World Congress Connected Vehicle
- Road Weather Demonstration
- FDOT Orlando ITS World Congress
- NCAR 2010
- San Diego
- NCAR 2009
- Pasadena
- Portland
- Seattle
- Safety Pilot Model Deployment
- Vehicle Infrastructure Initiative Proof of Concept
- Leesburg VA Vehicle Awareness Device
- Minnesota DOT Mobile Observation Data

The amount of data in the above datasets varies. Table 2-4 provides a brief description of each data set.

Table 2-4 A Brief Description of Data Sets in RDE

Data Set Name	Location	Time Period	Sample Size	Purpose
Safety Pilot Model Deployment Data	Ann Arbor, MI	10/1/2012 – 4/30/2013 (The data in October 2012 and April 2013 are provided in RDE)	Over 2700 vehicles	• Explore the real-world effectiveness of connected vehicle safety applications in multi-modal driving conditions • Evaluate how drivers adapted to the use of this connected vehicle technology • Identify potential safety benefits as a result of this connected vehicle technology.
ITS World Congress Connected Vehicle	The City of Detroit	9/8/2014 – 9/10/2014	9 vehicles	• Support a queue estimation algorithm • Demonstrate a real-world implementation of a connected vehicle environment • Showcase the operation of the associated Data Warehouse and Data Clearinghouse with an intention to support connected vehicle research.
Road Weather Demonstration	A short loop on Belle Isle, Detroit, MI	9/5/2014-9/11/2014		• For the public demonstration in September 2014. • Showcase the ability of instrumented vehicles to collect vehicle sensor data under the simulated road weather conditions and trig advisories or warnings to travelers.
FDOT Orlando ITS World Congress	Orlando FL	9/1/2010-10/22/2010	A set of Lynx transit buses	• Test the capability of Vehicle Awareness Devices (VADs) to capture and store data in the format of the J2735 Basic Safety Message (BSM) at a large scale.

Table 2-4 A Brief Description of Data Sets in RDE (Continued)

Data Set Name	Location	Time Period	Sample Size	Purpose
NCAR 2010	Michigan Test Bed	1/28/2010 – 3/29/2010 (Only the best RSE and OBE data in six days are included in this dataset.)	A small set of vehicles	• Concentrate on comparing atmospheric data from vehicle-mounted sensors to data from a nearby fixed weather observing station
NCAR 2009	Michigan Test Bed	4/6/2009 – 4/22/2009 (Nine-day RSE data in April 2009 and six-day good RSE and OBE data are available in the data set.)	A small set of vehicles	• Concentrate on collecting data during periods of rainy or snowy weather.
San Diego	I-5 in San Diego	1/1/2010 – 12/31/2010	10,000 trips	• Provide multi-modal data and contextual information (weather and incidents) that can be used to research and develop applications for the USDOT Dynamic Mobility Applications (DMA) program • One of the four test data sets acquired by the USDOT Data Capture and Management program.
Pasadena	Diverse roadway network in and around the City of Pasadena, California	9/1/2011 – 10/31/2011		• Provide multi-modal data and contextual information (weather and incidents) that can be used to research and develop applications for the USDOT Dynamic Mobility Applications (DMA) program • One of the four test data sets acquired by the USDOT Data Capture and Management program.

Table 2-4 A Brief Description of Data Sets in RDE (Continued)

Data Set Name	Location	Time Period	Sample Size	Purpose
Portland	Portland (including freeways and arterials)	9/15/2011 – 11/15/2011		• Provide multi-modal data and contextual information (weather and incidents) that can be used to research and develop applications for the USDOT Dynamic Mobility Applications (DMA) program • One of the four test data sets acquired by the USDOT Data Capture and Management program.
Seattle	Seattle	5/1/2011 – 10/31/2011		• Provide multi-modal data and contextual information (weather and incidents) that can be used to research and develop applications for the USDOT Dynamic Mobility Applications (DMA) program • One of the four test data sets acquired by the USDOT Data Capture and Management program.
Safety Pilot Model Deployment – One Day Sample	Ann Arbor, MI	4/11/2013	Over 2700 vehicles	• Serve as the preview of the larger Safety Pilot Model Deployment data • Intend to help prepare researchers and application developers to ingest these data to inform their research and development activities.
Vehicle Infrastructure Initiative Proof of Concept	Michigan Test Bed	8/21/2008 – 8/29/2008 (RSE data for the public application tests were available for eight days in August 2008. The data in this data set consists of RSE and OBE data for the middle six of these days.)	52 RSEs within 45 square miles and 27 vehicles configured with OBEs	• Proof of Concept (POC) trials • Three major phases were included in this testing program: subsystem test, system integration and test, and public and private applications test.

Table 2-4 A Brief Description of Data Sets in RDE (Continued)

Data Set Name	Location	Time Period	Sample Size	Purpose
Leesburg VA Vehicle Awareness Device	Trips in and around Leesburg VA and one long road trip from Ann Arbor, MI to Leesburg, VA by way of eastern Indiana	10/18/2012 – 12/19/2012	One test vehicle	• The data set was produced to give researchers an early sample of the large data set being collected as part of the Safety Pilot Model Deployment.
Minnesota DOT Mobile Observation data	Minnesota	6/26/2013 – 12/21/2015	310 instrumented snowplows and 19 instrumented light duty trucks as of May 2013	• Provide sample weather and vehicle engine status data that is transmitted in near-real time from vehicles to MnDOT over cellular media. (Vehicle-to-Infrastructure)

As shown in Table 2-4, 10 out of 14 datasets are connected vehicle-related data. The Safety Pilot Model Deployment is one of the largest real-world applications of connected vehicles. The data from this deployment, as well as data from the Orlando deployment, were analyzed in this project.

The Safety Pilot Model was deployed in Ann Arbor, Michigan, between October 1, 2012, and April 30, 2013 (Booz Allen Hamilton, 2015). Approximately 3,000 vehicles participated in this deployment, including passenger cars, commercial trucks, and buses. Four types of devices were installed in these vehicles, including integrated safety device (ISD), aftermarket safety device (ASD), retrofit safety device (RSD), and vehicle awareness device (VAD). Twenty-six roadside units (RSUs) were installed at signalized intersections and strategic freeway locations.

The dataset of the Safety Pilot Model consists of eight components, including two driving datasets (DAS1 and DAS2), basic safety message (BSM), RSE, and three types of contextual data (weather, network, and schedule). DAS1 is the data collected by the Data Acquisition System developed by the University of Michigan Transportation Research Institute (UMTRI), while DAS2 represents the data collected by the system developed by Virginia Tech Transportation Institute (VTTI). BSM is the basic safety message transmitted to and from an equipped vehicle. RSE is the data received and transmitted by roadside equipment. The contextual data of weather, network configuration and performance, and transit and special event schedule show the conditions under which the data were collected. The detailed attributes of each component are explained in the Appendix.

This section presents an analysis of the connected vehicle data collected using the SunGuide software from the Orlando deployment. The data files analyzed as discussed in this section were downloaded from the RDE and include data collected between September 1, 2011 and October 22, 2011. The attributes included in each data file are listed in Table 2-5.

A total number of 143 CSV files were downloaded from the RDE website. Each file contains data collected from one OBU, also referred to as Vehicle Awareness Device (VAD) for one day. Figure 2-5 and Figure 2-6 present a visualization of speed and elevation data collected from a VAD with an ID of 11935 on September 27, 2011.

Table 2-5 Attributes of Orlando VDS Data

Attribute Label	**Attribute Definition**
Timestamp	Time at which the data was recorded
DSRCmsgID	Identifies the type of message being recorded, in this dataset, the value is always 38
msgCnt	Message count. This value increments by one each 0.1 second until it reaches 127, then resets to 0
TemporaryID	Vehicle temporary identification number. It remains the same throughout one data file.
Dsecond (msec)	Count of milliseconds. After reaching 60000 (60 seconds) it resets to zero.
Latitude (deg)	Vehicle latitude in degrees
Longitude (deg)	Vehicle longitude in degrees
Elevation (m)	Vehicle elevation in meters
Semi-Major Accuracy (m)	Always 0
Orientation of Semi-Major Axis (deg)	Always 0
Speed (mph)	Vehicle speed in miles per hour
Heading (deg)	Vehicle heading in degrees (0 degrees is North)

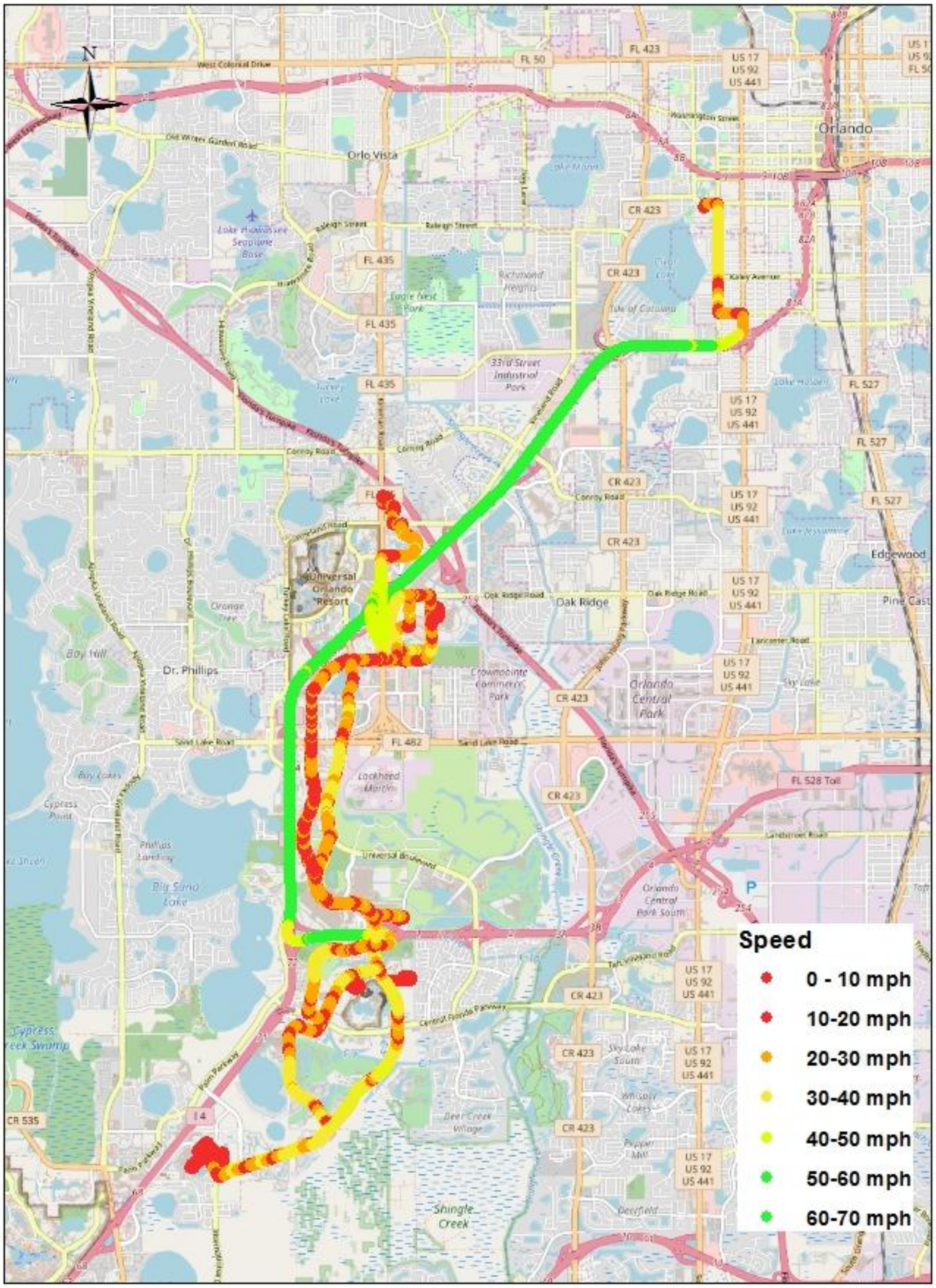


Figure 2-5 Visualization of Speed Measurement in Miles per Hour Based on Orlando VAD Data (9/27/2011)

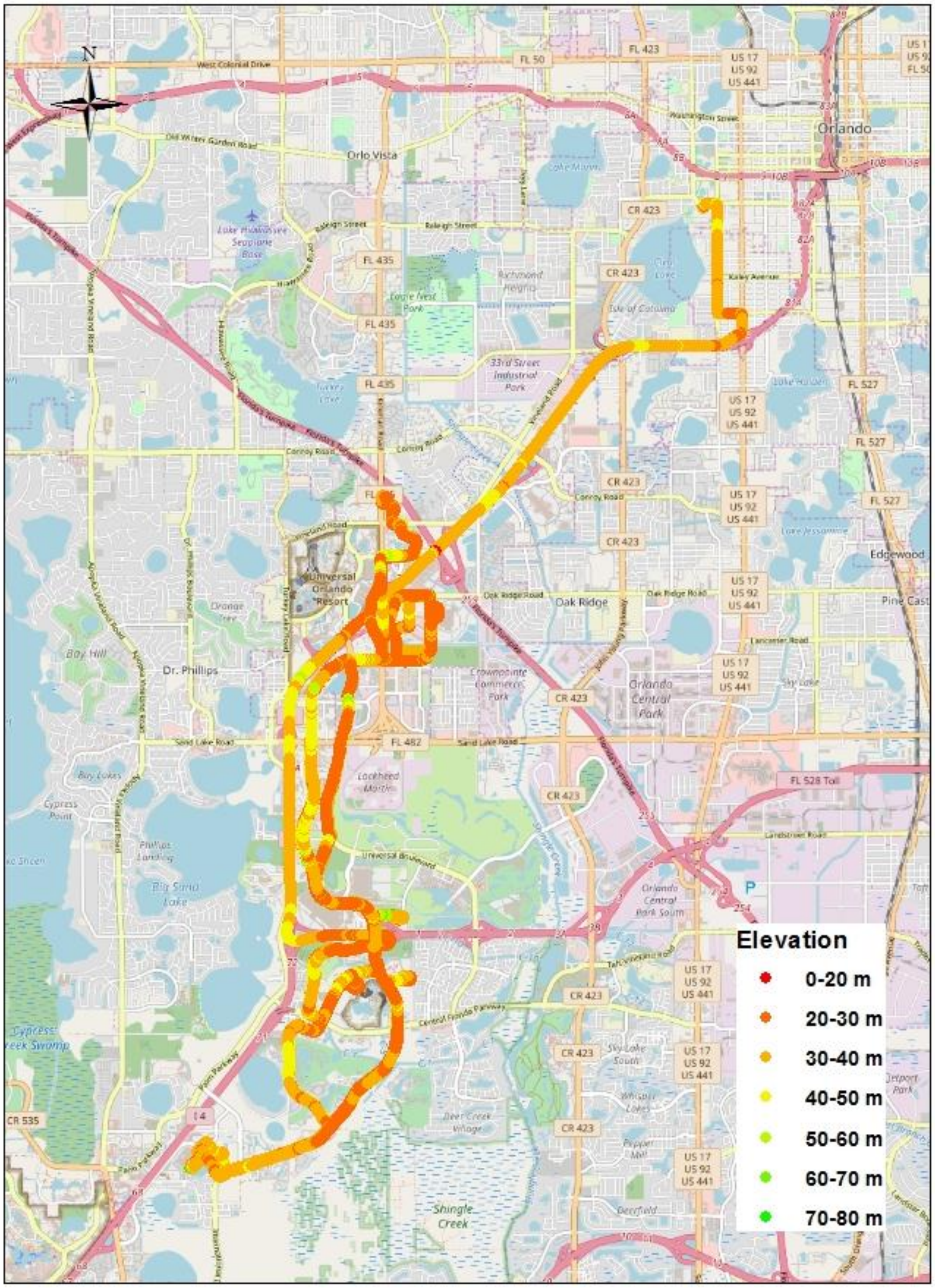


Figure 2-6 Visualization of Elevation Measurements Based on Orlando VDS Data (9/27/2011)

2.4.1. Data Preprocessing

As a case study of Orlando data utilization, the performance of a segment along the I-4 WB facility was examined in this study. As shown in Figure 2-7, this segment has a total length of 6.3 mile with a starting location at the south of L B McLeod Road and an ending location at the north of SR 528. As the raw VDS data file consists of continuous trajectories, the trips that travel along this study segment has to be identified and extracted first. A program was written to automatically fulfill such a function based on the distances between any point on the trajectory and the starting and ending locations. The red, yellow, blue, and green coloring of the links in Figure 2-7 correspond to their classifications according to the Highway Capacity Manual (HCM); including basic, merge, diverge, and weaving segments, as described later in this document.

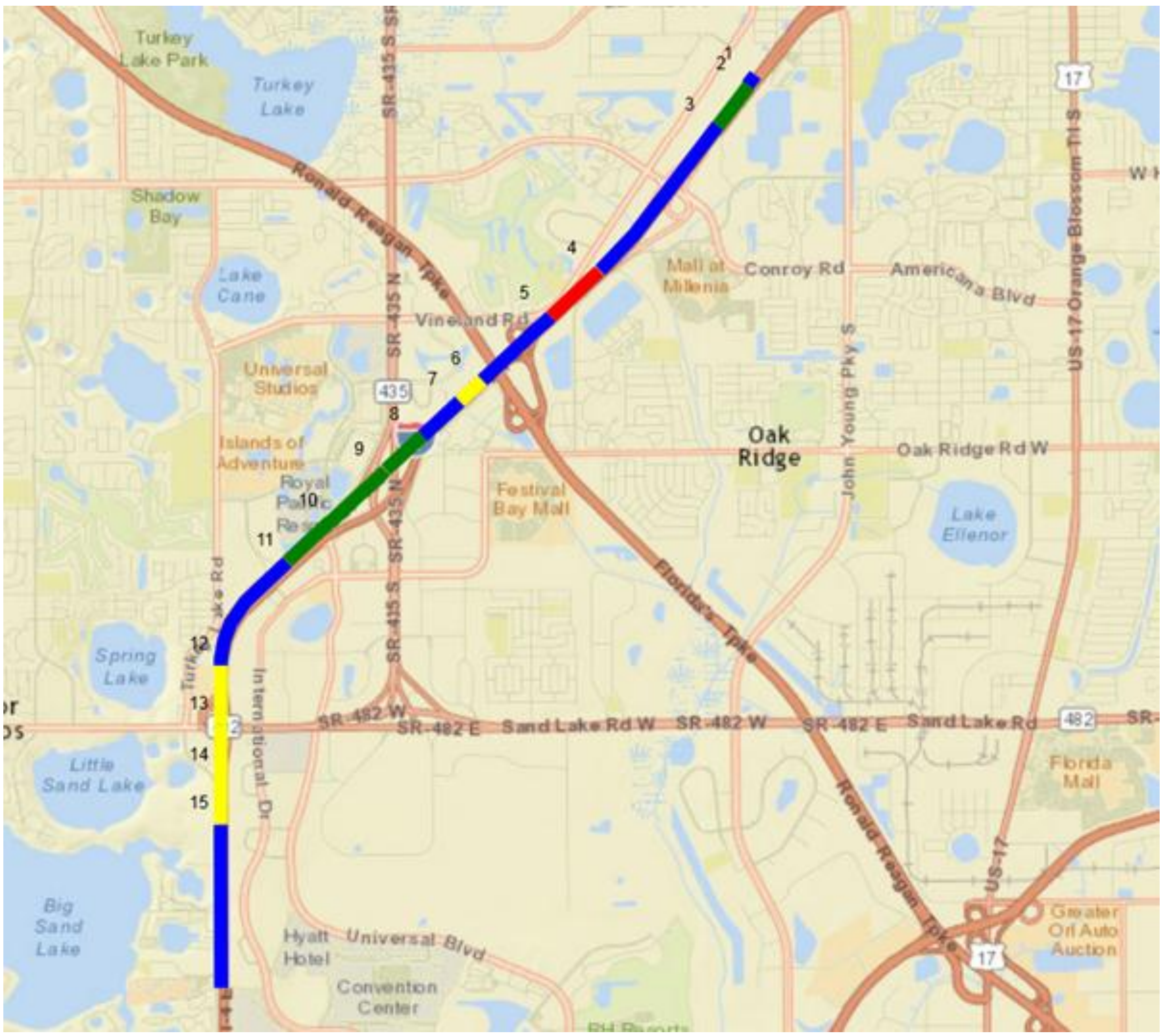


Figure 2-7 Location of Study Segment

During the data extraction process, a number of issues with the raw VDS data were found. The first issue is the data completeness. Even though there are 143 data files, most of the data files only have very limited records of trajectories. Some of data files have missing data with a value of “unavailable”. Records with zero latitude and/or longitude or wrong latitude and longitude were found in few data files. Extremely high speed values, for example, 366 mph, were also identified in few records.

The above discussed errors were removed during the data preprocessing. It is also noticed that the timestamps in the VDS data files do not differentiate between the AM and PM peaks. Therefore, it is assumed in this study that the timestamps after the occurrence of hour 12 are for afternoon measurements.

2.4.2. Result Analysis

Six trips that travelled between the origin and destination along the study corridors were extracted from all the available VDS data files. These were the only six trips that could be identified between the explored origin and destination pairs in the data set due the limited available valid data. Figure 2-8 shows the travel time for these trips. As shown in this figure, the travel time is roughly between 6.9 minutes and 8.7 minutes. Note that this corridor has a total length of 6.3 miles with a speed limit of 55 mph for the first 4.5-mile segment and a speed limit of 60 mph for the remaining segment. Thus, the travel time when traveling at the speed limit can be calculated as 6.7 minutes. It can be seen that five of the six of the extracted trips traveled at the speed limit with one vehicle experiencing congestion with a travel time of 8.7 minutes and thus a speed of 43 mph.

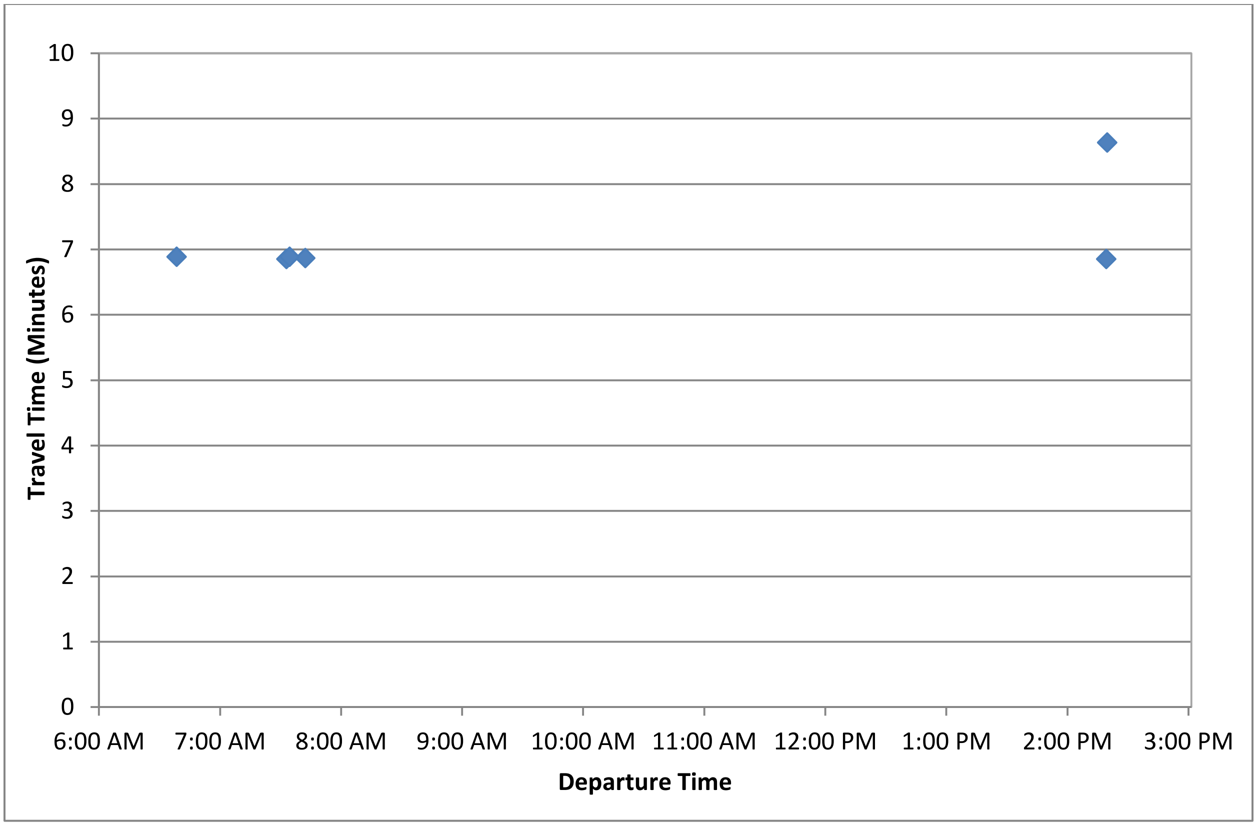


Figure 2-8 Travel Time along the Study Segment

As shown in Figure 2-7, the study corridor was further divided into 15 segments including basic, on-ramp merge, off-ramp diverge, and weaving segments, as defined in the Highway Capacity Manual. Figure 2-9 presents the average speed for each segment. Note that the segment number in this figure is ordered from the start to the end of the study corridor. The segment type is indicated by colors. As shown in this figure, most of the segments have an average speed of 53 to 54 mph except a weaving segment between Conroy Rd and Florida Turnpike with a much lower speed of 37 mph and the basic segment that follows it at an average speed of about 48 mph. There is no significant variation in the average speeds among the remaining basic segments, on-ramp segments, and off-ramp segments. Figure 2-10 shows the average speed along each segment for the trip with the longest travel time among the extracted trips, which has a total trip time of 8.7 minutes. It can be seen from this figure that the longer travel time of this trip is mainly caused by the congestion at the weaving segment and its following basic segment, where the average speed is about 14 mph and 30 mph, respectively. This again illustrates the power of connected vehicle data for identification of problem locations that is difficult to achieve with current surveillance technologies.

The boxplot in Figure 2-11 shows the variations of the speeds along each segment. Note that the top and bottom borders of the box correspond to the first and third quantile of the speed, respectively. It is seen from this figure that the weaving segment and the segment following it have a much larger variation in speeds than the other segments. It can also be noticed that the widths of the boxes for the on-ramp merge areas are relatively wider compared to the basic and off-ramp diverge segments, which indicates a larger variation in speed along the on-ramps.

Similarly, Figure 2-12 presents the variation of accelerations along the study segments. As shown in this figure, the median accelerations are close to zero. However, there is a wide distribution of accelerations and decelerations, particularly in the weaving and diverge areas. It may be possible to associate these acceleration and deceleration statistics with safety problems

Figure 2-13 summarizes the frequency of unchanged locations by comparing the current location of a vehicle in terms of its latitude and longitude with the previous record. This is another measure that can be used for quantifying the congestion level of the roadway segment. As the traffic becomes more congested, vehicles are more likely to stop, resulting in a higher frequency of unchanged locations. It is seen from Figure 2-13 that the basic segments No. 3 (that precedes the weaving segment) and No. 11 have a much higher frequency of unchanged locations, which indicates that these two locations have the potential to become congested segments. Figure 2-14 shows the relationship between 95^{th} percentile of acceleration and deceleration and the frequency of unchanged locations. As shown in this figure, there is no clear relationship between the frequency of unchanged locations and the absolute value of the 95^{th} percentile of acceleration or deceleration. These measures could be used as important new measures of system performance and will be explored in future.

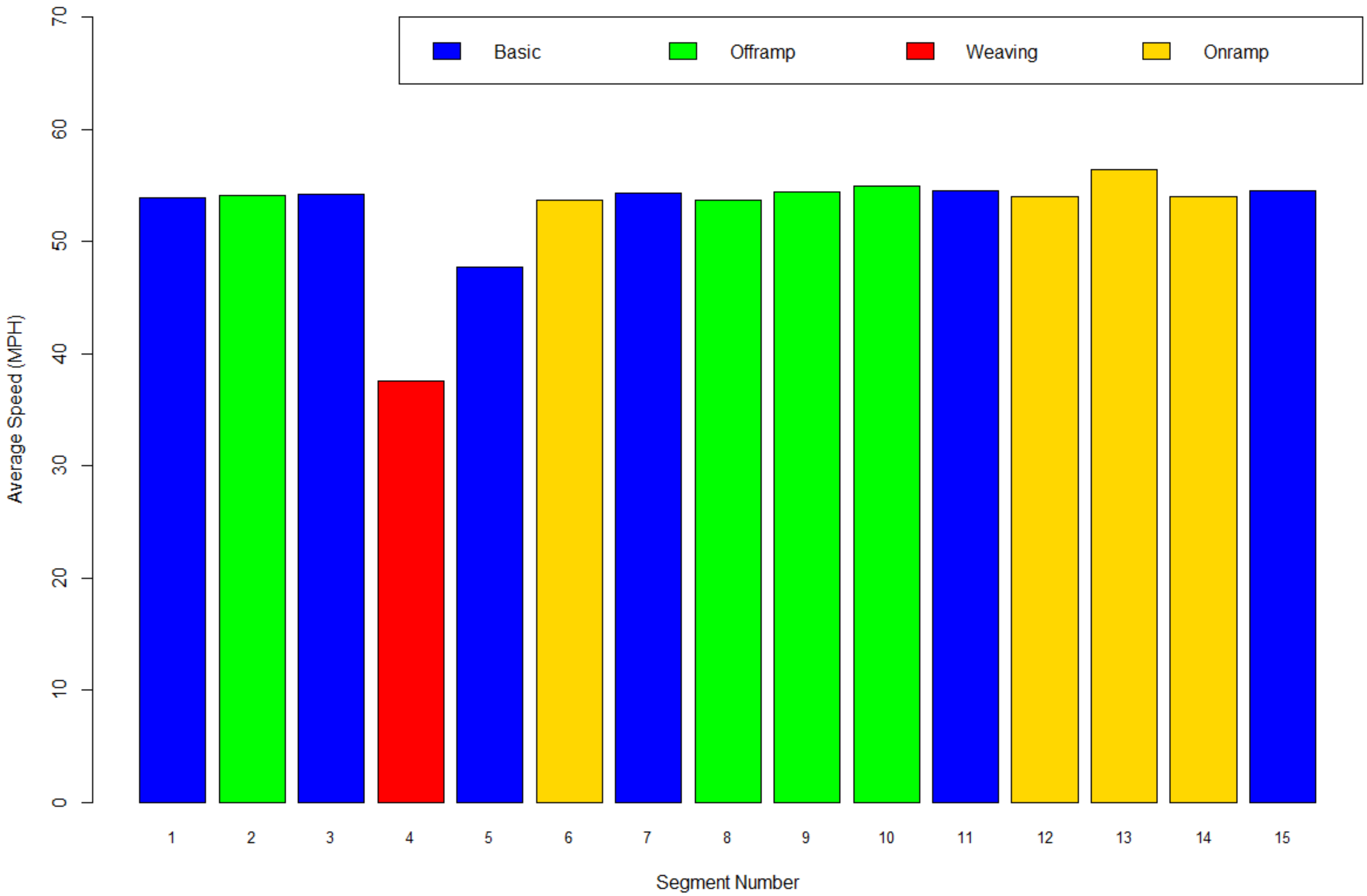


Figure 2-9 Average Speed along the Study Corridor

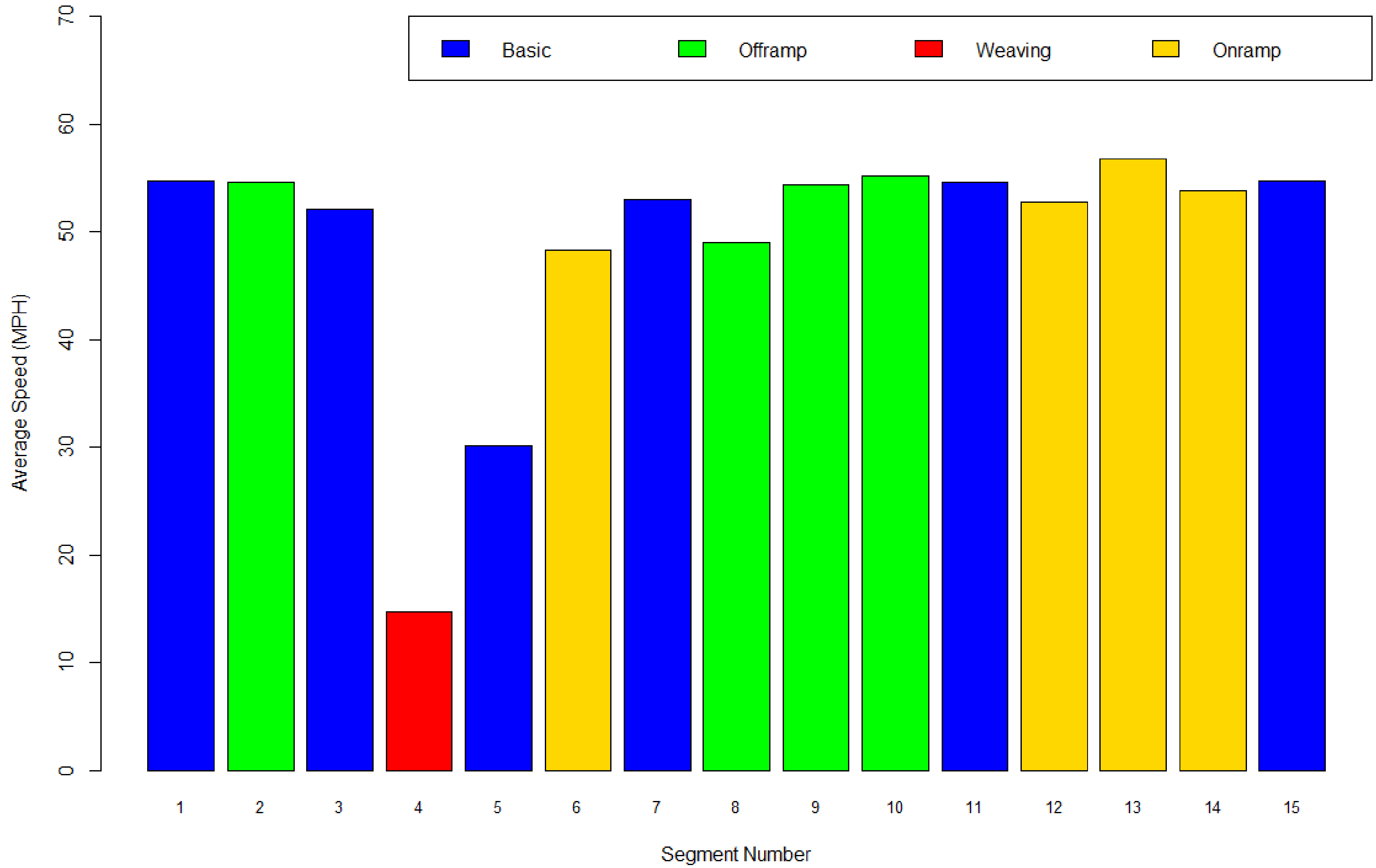


Figure 2-10 Average Speed along the Study Corridor for the Trip on October 16, 2011, Departing at 2:19:49 PM

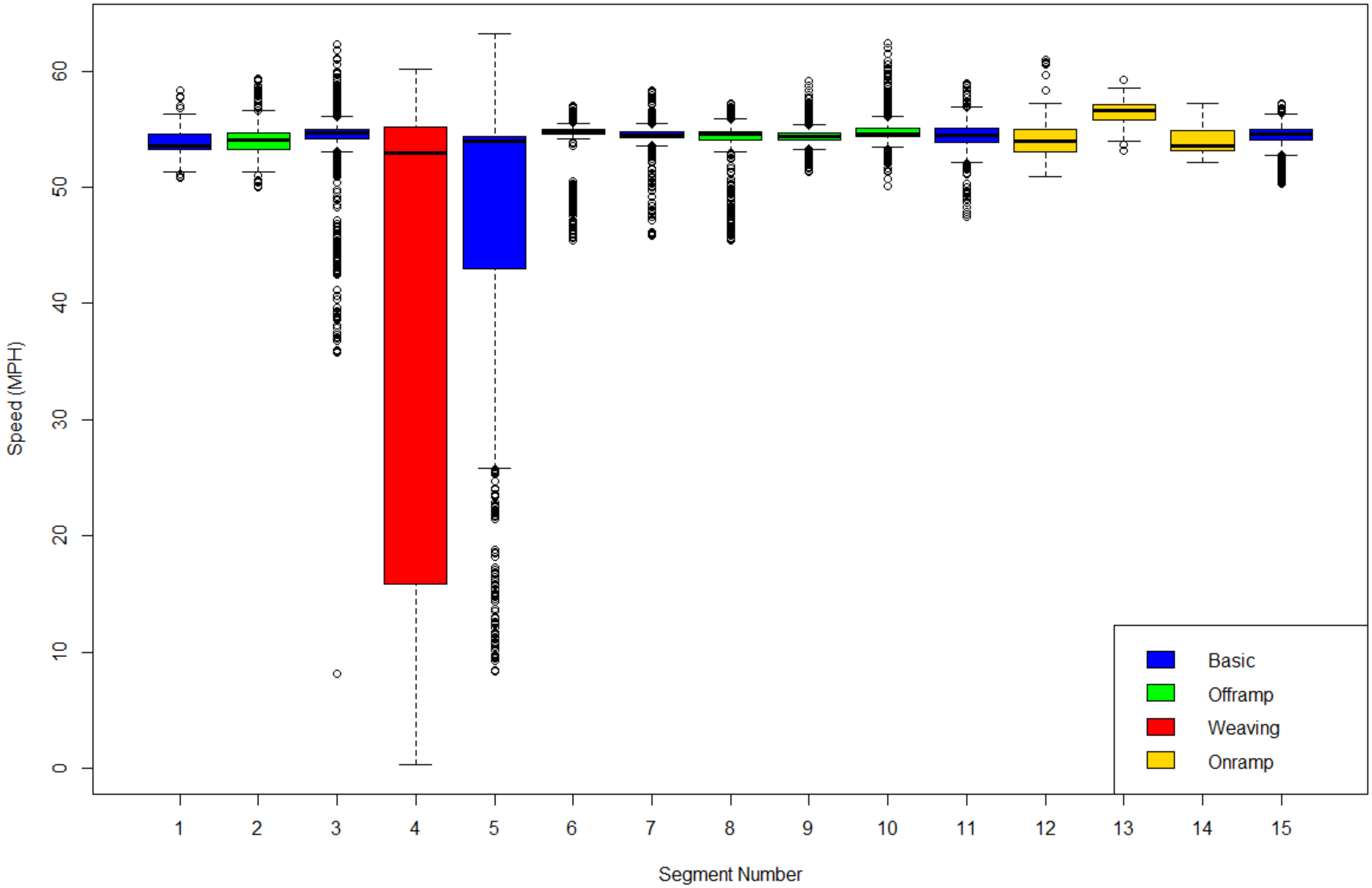


Figure 2-11 Variation of Speeds along the Study Corridor

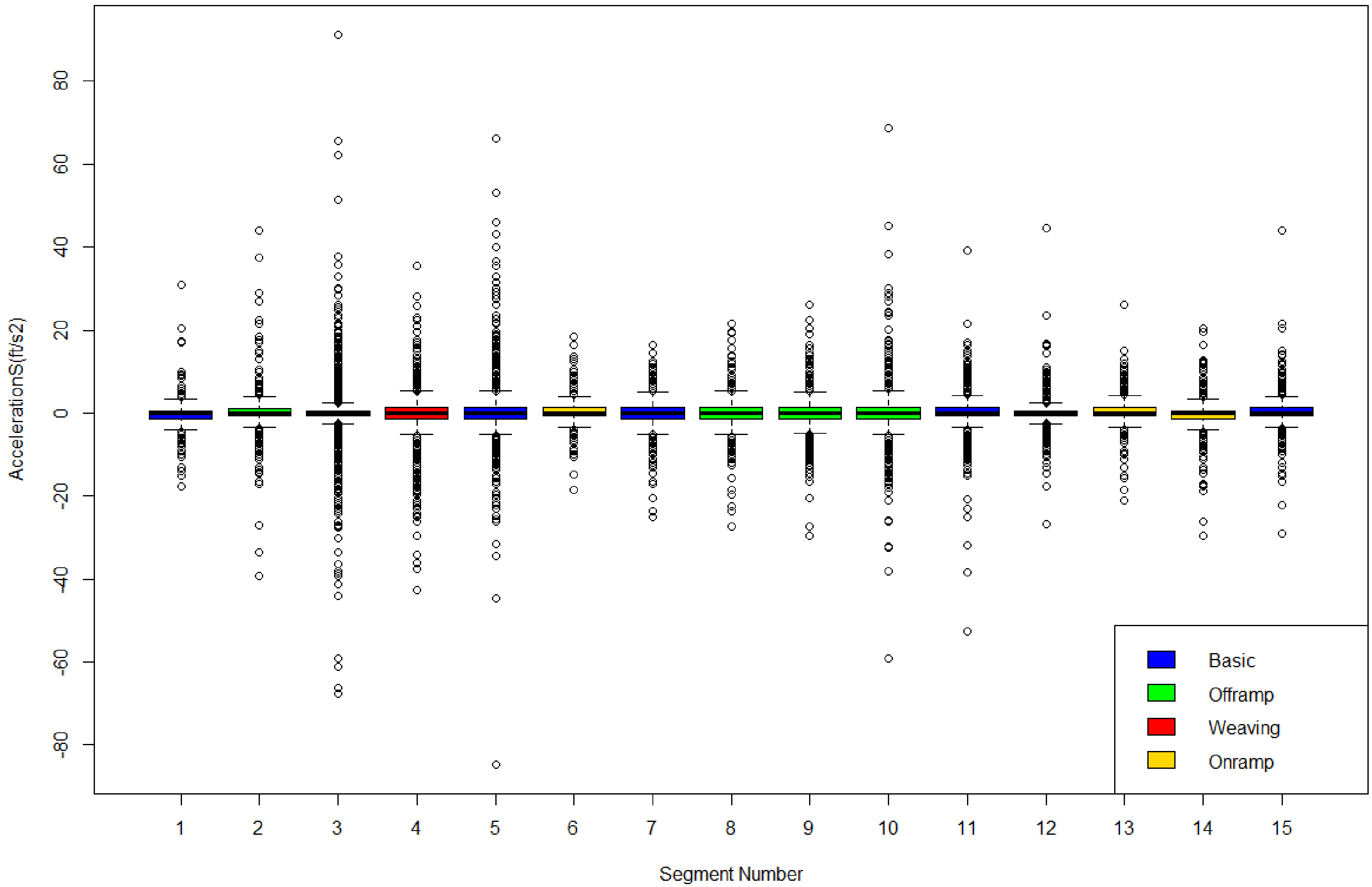


Figure 2-12 Variation of Accelerations along the Study Corridor

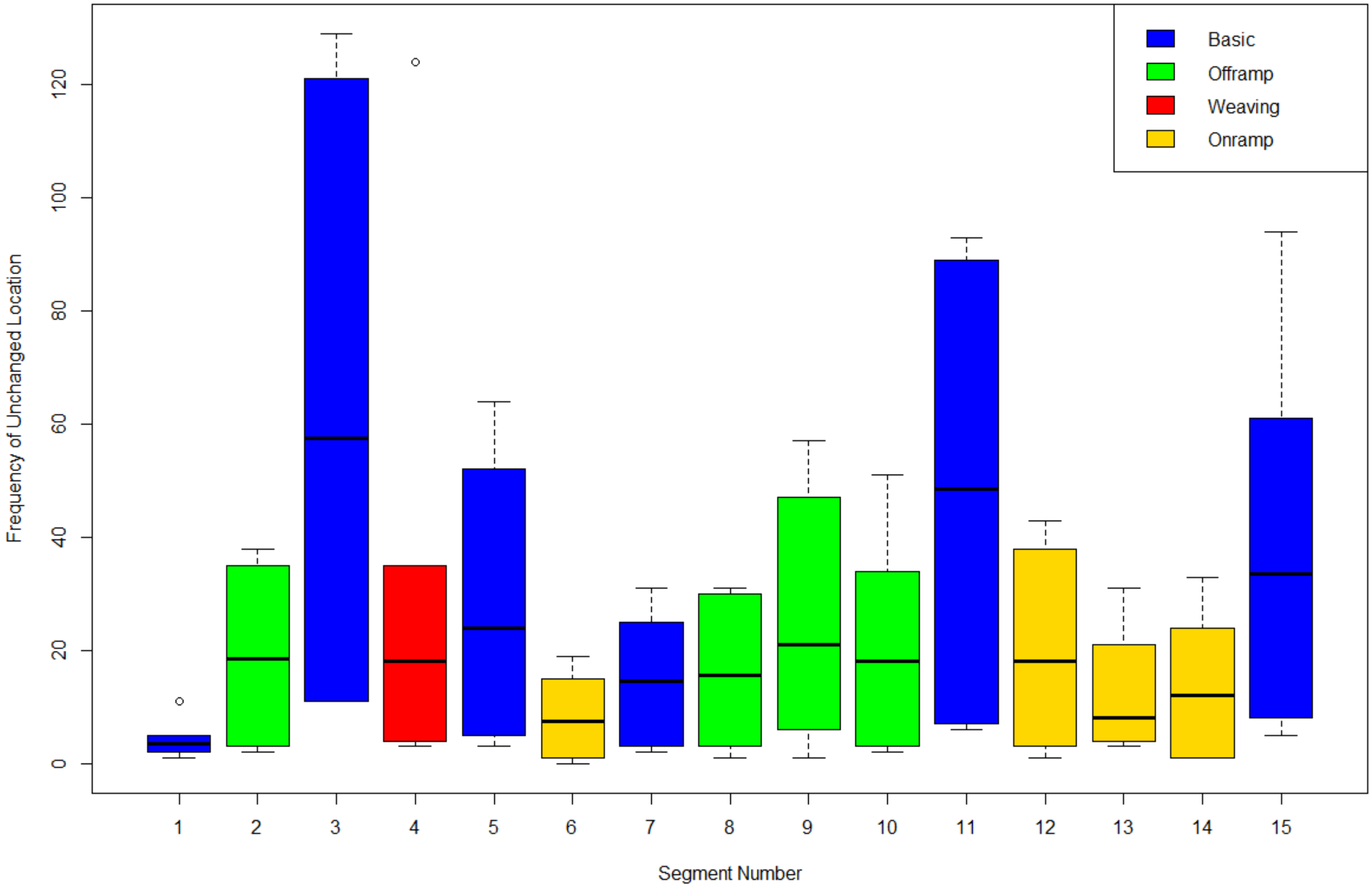


Figure 2-13 Frequency of Unchanged Locations along the Study Corridor

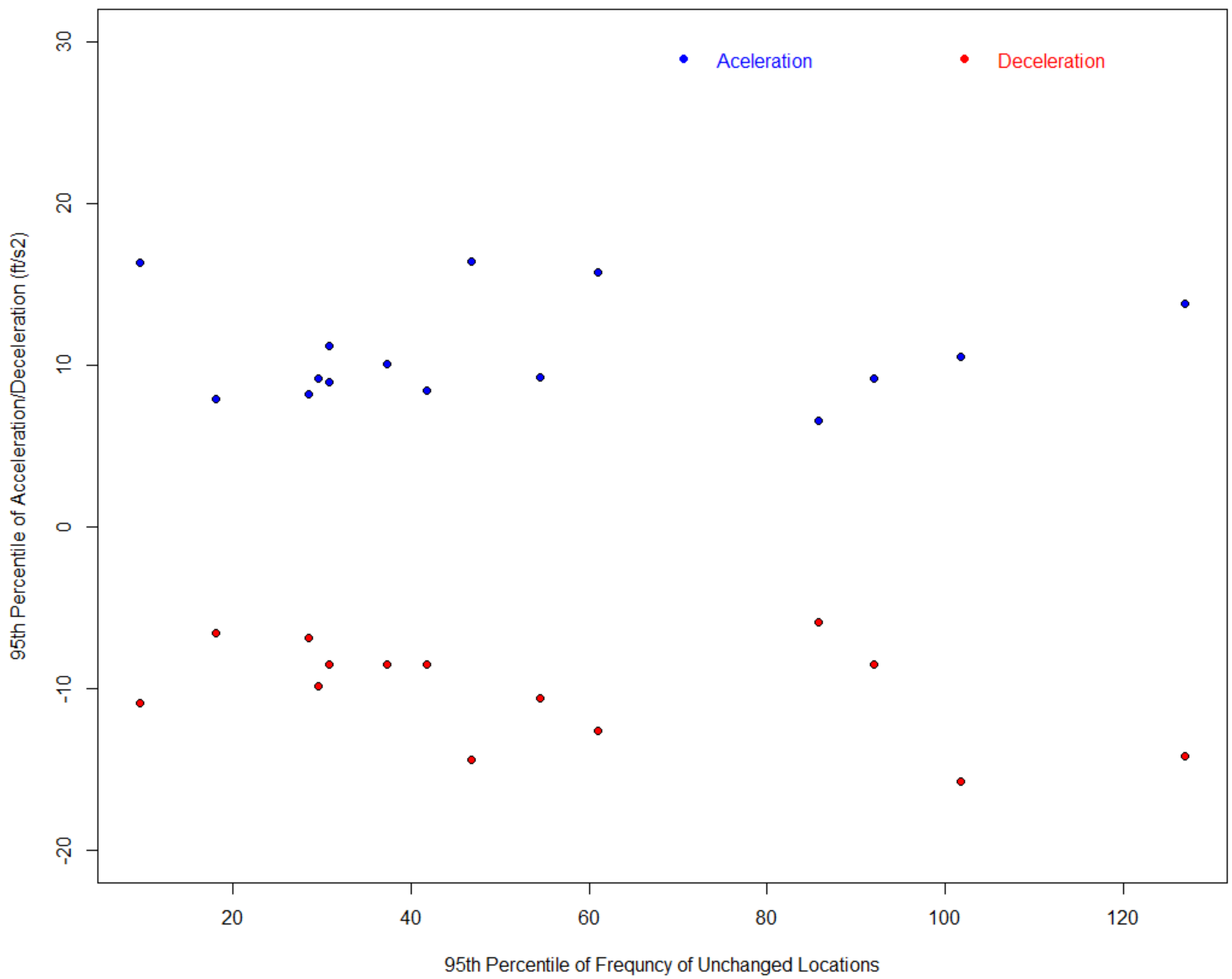


Figure 2-14 Relationship between 95th Percentile of Acceleration and Frequency of Unchanged Locations

3. SUNGUIDE MODULES

The central software used by the FDOT TMCs is the SunGuide software, The SunGuide software is an integration of a set of modules that allows the control of roadway devices as well as information exchange across a variety of transportation agencies and is deployed throughout the state of Florida. Figure 3-1 provides a graphical view of the SunGuide software architecture. The managed ITS devices by the SunGuide software includes traffic detection devices, cameras and associated encoders and decoders, video walls, dynamic message signs, highway advisory radios, road weather information systems, connected vehicle basic probe data, reversible lane systems, vehicle safety barriers, ramp metering, variable speed limits, wrong way driving, and express lanes. The software has automated incident and adverse weather conditions detection. In addition to collecting data from variety of point detection devices and automatic vehicle re-identification technologies, data are collected from third-party data feeds including HERE and WAZE. The traffic information is displayed on the operator map interface for use by the operator and used as an input to various SunGuide software modules such as automatic incident detection, travel time calculation, setting variable speed limits, and express lanes pricing. In addition, traffic and incident data are shared through center-to-center facilities with external systems including the Florida's advanced traveler information system, referred to as FL511. The central software also directly interacts with mobile devices utilized by the service patrols, referred to as Road Rangers in Florida. The software recommends response plans, stores event details, and produces customizable reports in graphical or tabular formats. The software also share data with the statewide data archived hosted in the Regional Integrated Transportation Information System (RITIS).

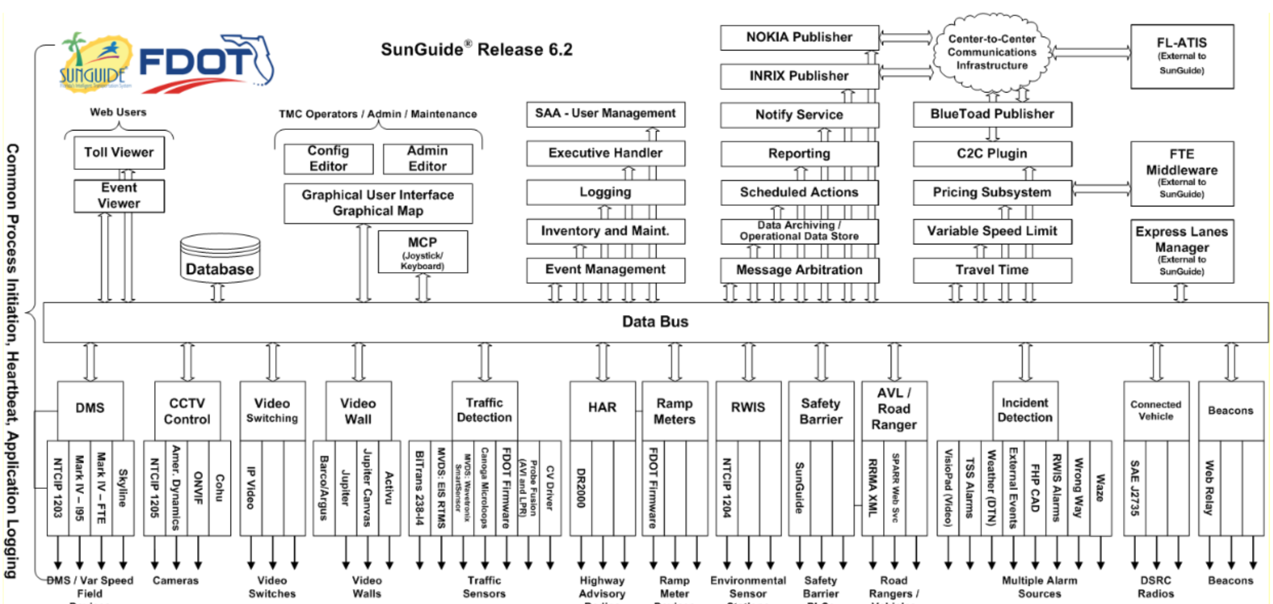


Figure 3-1 SunGuide Architecture

As shown in Figure 3-1, the SunGuide software consists of a number of modules including: Incident Detection, Traffic Detection, Travel Time Estimation, AVL / Road Ranger, Road Weather Information System (RWIS), Safety Barrier, CCTV, Center to Center, Connected Vehicle (travel time estimation), DMS and HAR, Ramp Meters, Express Lanes, Wrong Way Driving, Variable speed limits, Event Management, Inventory Maintenance, Message Arbitration, Video Switching, and Video Wall. Below is a brief description of the modules that are expected to be directly affected by CV technology.

3.1. Incident Detection

One of the most important functions of the TMCs is incident management. An important aspect of incident management is timely and reliable incident detection. Incidents are detected based on data from multiple sources including detecting traffic anomaly abnormality based on point detector measurements. The incident detection subsystem includes a number of other methods to detect incidents. These methods include, external notifications, manual communications with the Florida Highway Patrol (FHP)/police, center-to-center (C2C) connection to Florida Highway Patrol (FHP) Computer Aided Dispatch (CAD), third part feeds, through Road Ranger's (service patrol's) reporting, and in some district utilizing video analytic (SunGuide supports the VisiPad product) based on CCTV cameras. SunGuide has a module to receive crowdsourcing incident data obtained from a third party vendor (WAZE) and a module to interface to the Florida Highway Patrol Computer Aided Dispatch (FHP CAD) system.

The incident detection module based on point detection in the SunGuide software issues an alarm if it determines that traffic anomalies have occurred. The algorithm utilizes a rolling average of the current traffic measurements. The alarm is triggered when the speed on a freeway segment falls below a threshold associated with that segment. These thresholds are configurable by segment, time of day, and day of week. It should be stated here that some FDOT districts have used statistical limits on the mean at a significant level such as the 95th percentile to set the detection thresholds that produce the alarms.

After the TSS alerts functionality was added to SunGuide® software, operational experience has shown that SunGuide software often produced a large number of these alerts, resulting in operators ignoring most of them. A study was performed for the FDOT to address this issue of excessive alerts by better-calculating average speeds and applying an additional alert filter. This enhancement included providing alerts based on volume-weighted speeds and applying a volume filter that results in ignoring speeds based off of very low volumes (three vehicles or less). Using the volume-weighted averaging plus a volume filter eliminates a large number of TSS alerts. However, the SunGuide documentation states that, although significantly reducing the false alarms, these methods do not guarantee that the remaining TSS alerts reported indicate a traffic event. In some cases, TSS alerts are still triggered due to sensor failures or other data quality issues. However, the number of TSS alerts is significantly less using these methods, making operator validation of the alerts more operationally viable.

SunGuide also supports manual detection of incidents by allowing the operator to use manual entry screen to enter incident information from non-automated sources (e.g., service patrol, FHP, 911, and other sources) into the software. The SunGuide user interface allows the operator to enter additional details including verifying the location at which the incident was reported. The initial location is based on operator knowledge or location of the traffic detector which reported abnormal traffic flows. The operator may note any adverse weather conditions, pavement conditions, note the number and types of vehicles involved, specify any hazardous materials associated with the incident, and/or indicate which lanes or shoulders are blocked. The SunGuide system also provides the operator with the ability to associate alternate routes and events with the incident.

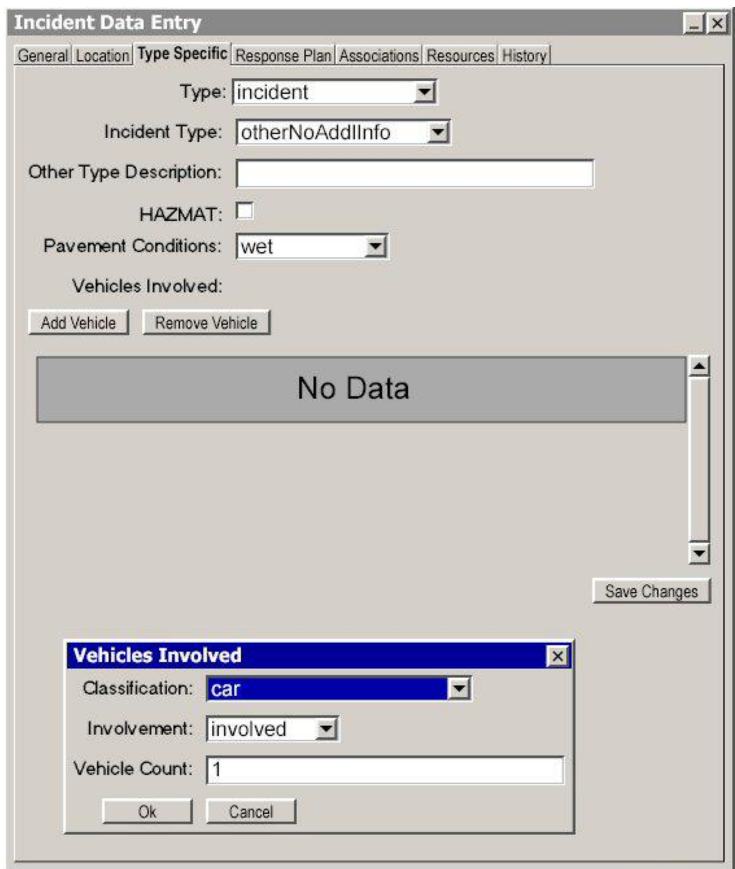


Figure 3-2 Incident Data Entry Form in SunGuide

3.2. Traffic Detection and Travel Time Estimation

One of the most important functions of traffic management centers is the real-time monitoring of the performance of the facilities managed by the centers. This monitoring is important to allow operators and central software to assess traffic conditions and to identify management actions to improve performance. The estimated travel times and potentially other measures are also disseminated to travelers using DMS and shared with the FL511 and other traveler information systems. In addition, the measured performance parameters are archived for use in planning and planning for operations.

The basic modules that support real-time traffic performance monitoring and estimation in the SunGuide software are the Traffic Detection (also known as the Transportation Sensor Subsystem (TSS)) module, the Travel Time module, and the data sharing with third party vendors (Inrix and HERE) modules.

The TSS subsystem in SunGuide acquire data (speed, volume, and occupancy) from point traffic detectors or speed/travel times from automatic vehicle re-identification (AVI) technologies to allow the assessment of traffic conditions. The TSS subsystem supports a number of point detectors communication protocols including Wavetronix microwave detectors, RTMS microwave detectors, Canoga magnetometers, and controller firmware. In addition, SunGuide supports collecting travel time/speed data based on automatic vehicle identification (AVI), as discussed later in this section. As mentioned in the introduction section, a “Connected Vehicle” module was introduced in SunGuide that allows the estimation of travel time based on connected vehicle data, in preparation for the 2011 ITS World Congress Connected Vehicle deployment in Orlando.

The SunGuide operator can visualize the current traffic conditions collected using point detectors for a specific link by double-clicking on that link in a map, and viewing the TSS Details panel that is opened. The panel displays speed, occupancy, and volume information by lane, for a selected link with point detection (see Figure 3-3, as an example). The top number (in a larger font) represents the “rolling average” and the bottom number (in a smaller font) represents the most recent value from the detectors. The rolling average timeframe and the polling cycle can be configured so the number of polls included in the rolling average can be computed by dividing the rolling average value by the polling cycle.

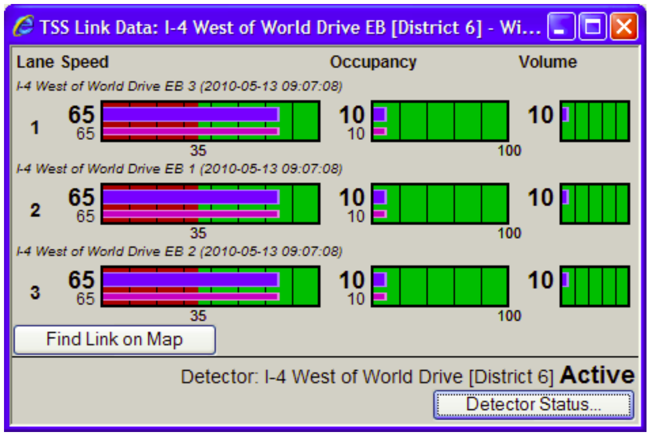


Figure 3-3 Point Detector Data Displayed on the TSS Details Panel

Point detectors do not measure travel time directly. Rather, the SunGuide Travel Time module calculates travel time based on smoothed rolling average of spot speeds measured at the point detector locations. A TSS link is defined by the user to be associated with one or more detector station. The speed of the TSS link is calculated based on the average of the lane speeds of the associated detector stations. What is referred to as Travel Time (TVT) links can also be defined, for use, for example, when disseminating travel time information on dynamic message signs (DMS) or for publication through the Florida ATIS system (FL-ATIS), as combinations of TSS links. The travel time of a TVT link is computed as the sum of the travel times of the associated TSS links. Travel time links are defined as one or more TSS links, as shown in Figure 3-4.

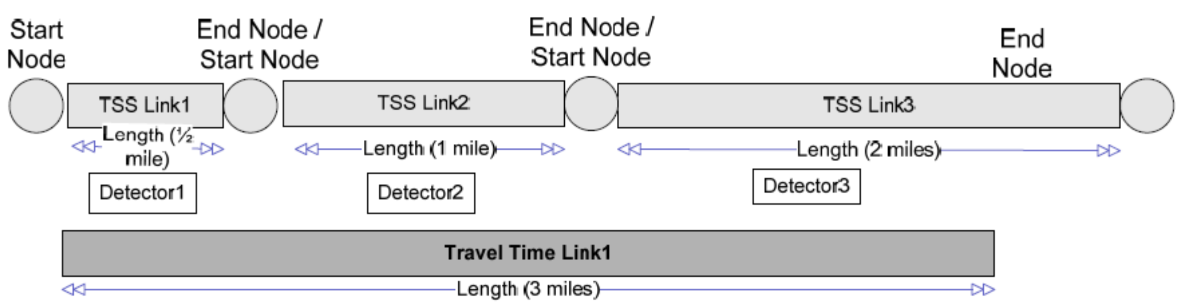


Figure 3-4 Travel Time Links Relationship to TSS Links

In addition to point detectors, SunGuide interfaces with AVI technologies. The supported AVI technologies includes electronic toll tag (two different vendors), license plate readers (two different vendors), and Bluetooth readers. As shown in Figure 3-5, when clicking on a link covered by AVI readers rather than showing speed, volume, and occupancy in the TSS panel; speed and current link traversal time are displayed, using the same rolling average and most recent display approach described for point detectors. Additionally, the counts of vehicle matches over the last detection period and over the last four hours are displayed, along with the current delay over free flow time and whether dynamic linking across the link is currently allowed. Dynamic linking, if enabled, will compensate for a failed probe detector in the middle of other operational probe detectors of the same type by creating a longer (dynamic) TSS link around the failed probe detector.

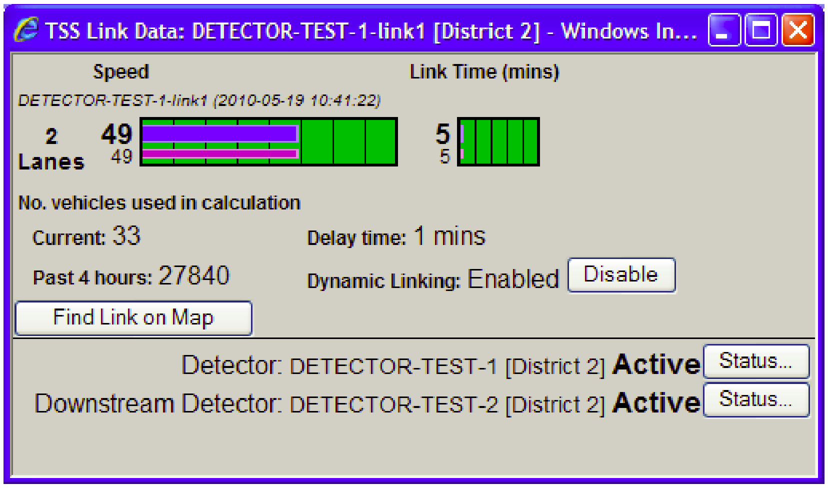


Figure 3-5 Probe (AVI) Data Displayed on the TSS Details Panel

Recent releases of SunGuide also allows receiving travel time information from private sector data providers (Inrix and HERE) through center-to-center (C2C) connections. The received data through these connections are made available for use by other system components. The private sector data providers use location and speed data from mobile devices to estimate travel time, and in some case origin-incident (O-D) information. SunGuide does not receive or use O-D information. Operators can use this travel time/speed data by viewing additional traffic speed links which are visible on their standard SunGuide Operator Map interface. Operators may also selectively enable or disable the display of this additional data. The data can also be used to produce travel time estimates of TVT links that can be used for posting travel time messages on DMSs and for publication through FL-ATIS.

The collected traffic and incident data are archived for use for planning and planning for operations. The data are archived utilizing the Regional Integrated Transportation Information System (RITIS) maintained by the University of Maryland.

3.3. Ramp Metering

The ramp metering (signaling) subsystem in SunGuide utilizes the Washington Department of Transportation for data acquisition and ramp signal control. The subsystem consists of vehicle detection devices (induction loops), field controllers, ramp meters, and a central system module. The field controller includes a Model 170 Controller firmware that processes volume and occupancy data based on single detectors as well as speed, length, and vehicle classification from dual-detector speed traps. Data are stored every 20 seconds and over-written every minute. This provides the firmware with a rolling one-minute data set. This data is reported every 20 seconds to the center in response to a poll request.

The SunGuide system makes use of the traffic information received from the detectors using the Fuzzy Logic Ramp Metering algorithm at Central. The data inputs to the fuzzy logic algorithm are (a) mainline occupancy just before the ramp outlet, (b) mainline speed just before the ramp outlet, (c) ramp queue occupancy, (d) advance queue detector occupancy, (e) high occupancy vehicle (HOV) bypass volume, (f) downstream speed from assigned one or more detector station(s), and (g) downstream occupancy from one or more assigned detector station(s). The outputs from the fuzzy logic algorithm is a vehicle metering rate that is updated every 20 seconds.

The firmware works in two control modes: central or local control. The central source was normally used for metering, except in special cases and after a communication failure. In special cases, the operator commands the controller to meter, using the local algorithms. The basic metering algorithms and metering rate adjustments are controlled by a set of parameters downloaded from the central computer. The central software allows the operator to view the status of a ramp meter and to change the current plan implemented at the ramp meter.

3.4. Express (Managed) Lanes

An important function of SunGuide TMCs is the operation and management of express lanes, or managed lanes (MLs). The first ML project in Florida is the 95 Express project in South Florida. The I-595 reversible lane in Broward County came next. There are several other ML in the implementation or planning stages. Early in the planning/design stages of the 95 Express project, the FDOT decided to add a module (pricing subsystem) to the existing SunGuide software for express lane operations. Due to a tight 95 Express project schedule, the initial version of the SunGuide software pricing subsystem module did not provide dynamic pricing functionality. The scope of the SunGuide software Pricing Subsystem module was limited to posting toll rates and lane status on DMS, scheduled time-of-day rate tables, manual entry of toll rates, alerts, limited reporting, and the transmission of toll rates to the Florida's Turnpike Enterprise (FTE) Toll Operations system. FDOT D6 TMC developed a separate software application to implement dynamic pricing based on real-time measurements. This application set the toll rates based on the level of service defined based on density in accordance with the Highway Capacity Manual (HCM). The application calculates density as a function of the speed and volume measured by point detectors. The initial deployment of dynamic pricing utilized the average density for the entire 95 Express facility, as the decision variable. Later, an option was added to utilize the density on selected detectors (such as at bottlenecks) for the toll calculation. An Operator Interface is included to alert the operator of toll rates changes, allowing modifications to these rates.

3.5. Road Ranger Support

The service patrol vehicles, referred to in Florida as the Road Rangers, communicate detailed information to the TMC that is critical for event management. The communications between SunGuide TMCs and road rangers have been done using tablets or laptops with attached or embedded GPS units communicating through a cell card. SunGuide's interface supports VANUS and IBI Road Ranger tablets using defined sets of communication schemas. More recent versions of SunGuide also support the use of smart phones, instead of the tablets, as a platform for this communication.

With Road Ranger (RR) applications, service patrol location updates need to be routinely sent to SunGuide so the operators can track the position of Road Rangers and dispatch them appropriately. Location information includes latitude, longitude, heading, and speed. The vehicle locations are displayed on the SunGuide Operator Map to show the real-time Road Ranger movement. Road Rangers can create events and enter the events into the SunGuide system. These events will contain basic information including the location and roadway direction of the event. When a Road Ranger arrives at an event they did not create, they will send an update to operators indicating they have arrived. While the Road Ranger is on scene, the Road Ranger is able to view and report the details, services rendered, and vehicles involved. Road Rangers will also be able to call the TMC from the application in order to report additional event details. When a Road Ranger has finished their work at an event, the application will allow them to depart and an event and continue patrolling for other events.

Additional sets of functionality was identified for potential future applications in SunGuide including:

- Additional event editing such as modifying lane blockage or adding comments
- Event photos allowing images taken by the RR driver to be stored as part of the event record, displayed to an operator, or possibly emailed to operation staff
- Event Routing by integrating the application with a navigation application on the device to provide turn-by-turn directions to the Road Ranger

3.6. Road Weather Information System (RWIS)

The SunGuide RWIS subsystem interfaces to field weather sensor stations that collect weather conditions information. The system is capable of warning motorists of fog, low visibility, wind and other events. In addition, better operation and management activities can be implemented based on alerts provided to the operator (see Figure 3-6 for an example of these alerts). Default thresholds are provided for activation that can be configured by the user. In addition, automated response plans including generated messages and alerts are produced by the system that can be fine-tuned or overwritten by the operator based on additional information and human understanding of the situation. When the RWIS device data generates an alert in the SunGuide software, the software automatically selects the appropriate beacons based on a configurable distance radius search from the device in tenths of miles. SunGuide software can be configured to either automatically activate the selected beacons immediately upon selection, to place them in a response plan for the operator to activate as a part of an event, or to take no action. Multiple beacons could be activated. The response can involve DMS. HAR, and/or FL511 messaging for low visibility events or activating beacons.

.

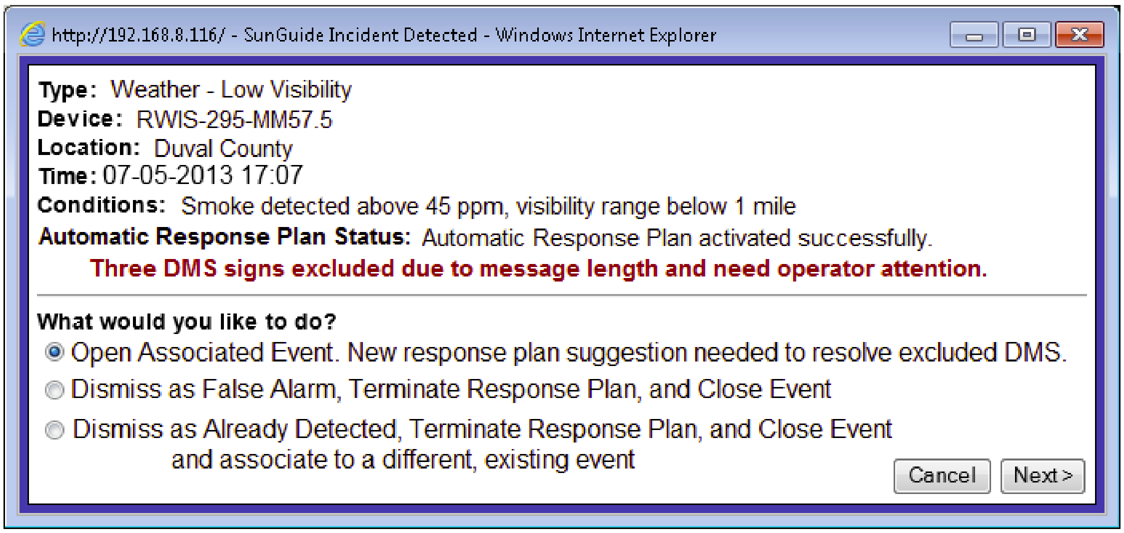


Figure 3-6 Weather Alert messages Based on RWIS

3.7. Variable Speed Limit (VSL)

The SunGuide System has a VSL module that allows operators to configure VSL messages and activation parameters. As implemented by FDOT District 5 on I-4 in Orlando, the VSL signs between Maitland Boulevard and Orange Blossom Trail adjust the speed limits leading into highly congested areas to reduce congestion and crashes during rush hours. Figure 3-7 shows a picture of these signs. The traffic management software analyzes sensor data and makes recommendations on adjusting speeds. The operators then review videos from CCTV cameras and reports from other sources to determine if the speed limits should be changed. As implemented by FDOT District 4 in Broward County, VSL is designed to improve safety along a busy section of SR25/US27 (Okeechobee Road) near West Broward High School. The purpose of the VSL signs is to inform motorists of adjustments to the existing SR 25/US27 speed limits on school days, during which time the school's flashing school beacons (FSB) will be in operation, and also for periods of school activities (e.g. sporting events) associated with West Broward High School. At other times, the VSL signs display the normal speed limit of 65 mph. On school days, the VSL signs gradually reduce the speed limit before the school zone flashing lights are turned on. The signs display transition from 65 mph to 55 mph for a short period. The signs then display transition down to 45 mph just before the lights are turned off. The transition time will allow motorists to adjust to the lower speed limits as they approach the school zone. This system includes closed-circuit television cameras to monitor the VSL system devices and traffic flow for the purposes of verifying proper VSL operation.

Figure 3-7 The I-4 VSL Sign Display

3.8. Wrong Way Driving (WWD)

The WWD system in SunGuide receives WWD detection incident data from field devices capable of detecting WWD. The WWD event in SunGuide can be created manually by operators or based on automatic detection resulting in preconfigured response plan activation. When a WWD incident is detected, up to five closed-circuit television (CCTV) and the associated presets can be automatically invoked. Specific DMS warnings and email and text recipients can be included in the response plans. The WWD detection information is logged in the database for future reporting and analysis. SunGuide immediately displays an alert handling and video viewing window on operator workstations without the operator having to click on the alert (see Figure 3-8). WWD events can be detected by Wavetronix microwave detection and the TAPCO system that combines camera and microwave detection.

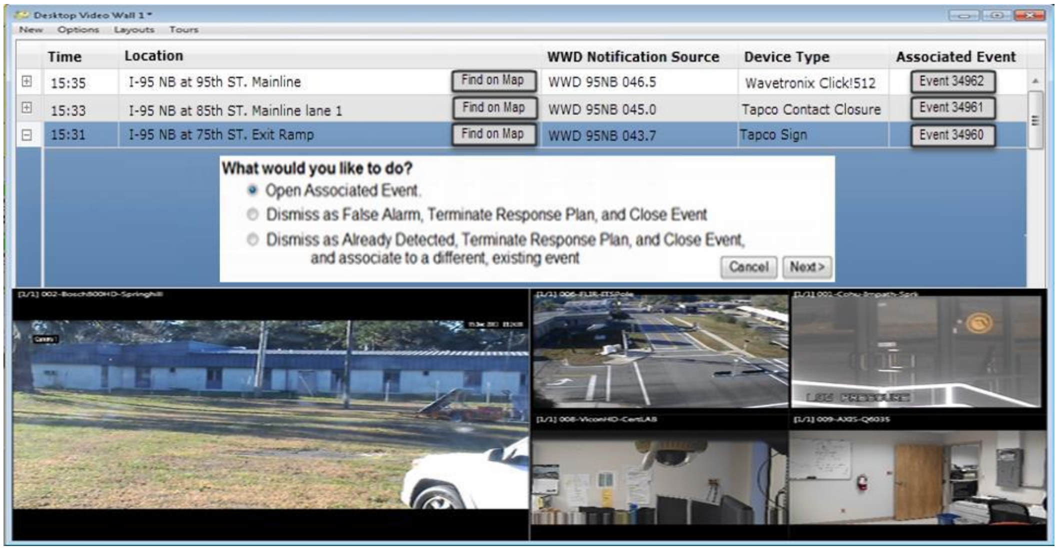


Figure 3-8 WWD Alert Display in SunGuide

The SunGuide documentations state that connected vehicle on board units could be used in a future enhancement of SunGuide to disseminate traveler advisory messages to motorists, warning them of the wrong way driver and to use extreme caution.

3.9. Safety Barrier Device

The Safety Barrier subsystem provides the capability to receive barrier events from Programmable Logic Controllers (PLCs) when cars breach the cable. As the crash breaks the cable contact, a nearby strobe light is activated and crash location information is transmitted to the Center. The SunGuide Safety Barrier software subsystem communicates with the field controllers using the TCP/IP protocol.

3.10. Summary

Table 3-1 lists the SunGuide functions that make use of field data that can possible supported by connected vehicle data, as reviewed in the previous subsections. The table includes the collected data for use as inputs to these functions and the sources for these data. Comments are also given regarding the used data.

Table 3-1 SunGuide Functions and Associated Data

SunGuide Function	Collected data	Current Data Sources	Comment
Incident Detection	Incident occurrence, incident attributes, vehicles involved, lanes blocked, incident cleared, Hazmat information	Alarms based on point detection, FHP/Police notifications, RR service patrol notifications and reporting, WAZE crowdsourcing data, Video Analytics	Many of the incident and involved vehicle attributes are obtained using manual methods through notifications, RR reporting, and CCTV camera monitoring
Traffic detection and travel	Volume, speed, occupancy, travel time	Point detectors. Travel times also collected from vehicle re-identification technologies (e.g., Wi-Fi and Bluetooth) and third party vendors	Accuracy of measures obtained using different technology varies as reviewed later in this document. There is potential for obtaining more detailed and advanced measures to support management
Ramp Metering	Volume and occupancy data based on single detectors as well as speed, length, and vehicle classification	Point detectors (inductive loops)	Utilizes aggregate measures in space and time. Potential for use high resolution measures to support metering and merging.
Managed Lanes	Volume and speed (density is calculated)	Point detectors	Density calculation based on point detection is less accurate. Vehicle classification may be useful for enforcement.
Service Patrol Support	Latitude, longitude, heading, and speed; in addition to event information entered by the driver	AVL technology (GPS). Event information entered manually by driver in mobile application.	Most of the event information has to be entered manually. Also, there is a potential to install equipment of the RR to support system performance measurements
RWIS	Rain intensity, visibility, fog	RWIS stations. Additional information obtained from national weather services.	RWIS station deployments have been limited

Table 3-1 SunGuide Functions and Associated Data (Continued)

SunGuide Function	Collected data	Current Data Sources	Comment
Variable Speed Limit	Traffic speed and occupancy	Point detectors.	
Wrong Way Driving	Wrong driving vehicles	Point detectors (e.g., microwave and camera based)	
Safety Barrier	Vehicles breaching barrier cables	Cable contact detection	

4. POTENTIAL USE OF CONNECTED VEHICLES

This section discusses the potential use of connected vehicle technologies to support existing SunGuide functions that utilize real world data. For each of the functions, it first discusses the national experience and performance of these functions when utilizing existing technologies. Then, a discussion is presented of the use of CV data to support these functions.

4.1. Incident Detection

As stated in the previous section, a critical function of SunGuide is the fast detection and validation of incident occurrence and attributes to reduce lane and shoulder blockage durations, as well as to allow fast notification of responding agencies. As stated earlier, SunGuide centers currently detect incidents utilizing a number of methods, including processing of data collected from point detectors (and in some cases AVI data), external notifications including notifications from the police, service patrols, WAZE data, and video analytics. SunGuide utilizes a relatively simple method for detecting incidents based on point detectors based on thresholds, although advanced statistical methods can be used to set the thresholds.

There are numerous automatic incident detection algorithms based on point detectors that have been reported and researched in the literature. These algorithms are of six types: comparative algorithms, statistical algorithms, time-series algorithms, filtering based algorithms, traffic modeling algorithms, and artificial intelligence algorithms. Incident detection methods have been assessed utilizing three main parameters: Detection Rate (DR) in percentage, Mean Time to Detect (MTTD) in seconds, and False Alarm Rate (FAR) in percentage. Previous experience with incident detection algorithms indicate a possible MTTD of one to four minutes when using various incident detection algorithms with point detector data and false alarms lower than 2%. The reported detection rate varies between 30% and 100%. Some existing incident detection algorithms, however, cannot detect the incident until the queue caused by the incident reaches the upstream detectors, which may take a long time or may never happen if the queues due to incidents are short or do not exist.

There has been interest in using AVI technologies for incident detection. Mouskos et al. (1998) reported on the experience with incident detection based on data collected by electronic toll readers placed at 0.5 to 2.1-mile intervals in New Jersey and New York. The reported findings show that the tested incident algorithm can produce results comparable to the common loop-based algorithms.

Connected vehicle data has the potential to be used in incident detection. Depending on the CV market penetration, the CV data can be used to detect critical parameters to support incident management including congestion/incident detection, queue length and back of queue location, shockwave speed, blocked lanes, types of involved vehicles, and incident clearance. The Connected Vehicle Reference Implementation Architecture (CVRIA) (USDOT, 2015b) describes how connected vehicle data can be used to support incident management, as part of the Vehicle Data for Traffic Operation (VDOT) application. This application, according to the description, uses probe data information obtained from vehicles in the network to support traffic operations, including incident detection and the implementation of localized operational strategies. The CVRIA specifies the use of connected vehicle data in incident detection such as changes in vehicle speeds indicating the disruption of traffic flow, activation or deployment of vehicle's safety systems, or sudden vehicle turns or deceleration at a specific location (indicating a potential obstacle in the roadway). Basic Safety Message (BSM) Part 1 parameters such as location, speed, vehicle size, and acceleration/deceleration rates can be used to detect incidents. Such parameters are expected to be mandated to be broadcasted by NHTSA and can be captured by DSRC roadside units as they broadcasted as part of vehicle-to-vehicle (V2V) communications. However, there are many other parameters that are included in BSM part 2 that will be useful for incident management, particularly with high market penetrations of CV vehicles. These include air bag deployment, antilock brake system active over 100 msec, brake system status, rate of charge of steering wheel angle recent or current hard braking, hazmat status, traction control system active, vehicle placarded as hazmat carrier, and vehicle type. Most likely these will have to be brought to a center using a wide-area communication technology like cellular communications.

The use of information included in BSM Part 1 can support incident detection. Crabtree and Stamatiadis (2007) concluded, based on simulation results, that for a market penetration of 30% of connected vehicles, the MTTD ranges from 2 to 4 minutes for a reader spacing of 2 miles, and 2.5 to 14 minutes for a reader spacing of 10 miles. Asakura et al. (2015) found that using an incident detection algorithm based on data from probe vehicles at a 1% market penetration can produce a detection rate, false alarm rate, and MTTD of 19.1%, 0.0021%, and 7.9 minutes, respectively. A simulation study by Chue et al. (2002) found that incident detection based on 5% to 50% of probe vehicles have an MTTD of 12 minutes to 4 minutes. Another study (Parkany and Bernstein, 1995) indicates that the MTTD is 0.8 minutes with 50% market penetration. Further analysis was conducted in this study to determine the impacts of market penetration on incident and back of queue detection, which is documented in the later section of this report.

4.2. Traffic Detection and Performance Measurements

Another important function of the SunGuide centers is system performance measurement function that includes obtaining data that can be used to estimate measures for use in various planning, operation, and management functions, as well as to support the dissemination of information to travelers and third parties. As stated earlier, the SunGuide TMCs currently collect, estimate, use, and disseminate few measures including speed/travel time, occupancy, and volume. There accuracy of the measures obtained using different technology varies as reviewed in this section. There is a potential for use CV data to replace or at least complement (at lower market penetrations) the existing data collected by existing technologies.

Connected vehicle promise to provide several parameters currently provided by other technologies, as well as parameters that cannot be collected by these other technologies. New measures collected or estimated based on connected vehicle data will provide additional opportunities to enhance current management strategies and algorithms. Such measures can include for example vehicle acceleration/deceleration, vehicle classification, number of stops and brakes, potential for crashes, and emission/fuel consumption. The CV market penetration required to collect accurate and reliable measurements vary depending on the specific measure, type of facility (freeway versus arterial), demand and congestion levels, and the density of intersections on the corridor. It is expected that among the above listed measures traffic volume, classification, and occupancy/density will require significantly higher market penetration, necessitate combining CV data with data from other sources to achieve acceptable accuracy levels. Further details are presented below and in later sections of this document.

4.2.1. Travel Time/Speed

It is expected that travel time and speed will be among the first measures that agencies will start estimating based on CV data since the required percentage of equipped vehicles to produce accurate results is relatively small, particularly for urban freeways during the peak periods. Most agencies specify an accuracy of 90% or 95% for travel time/speed estimation. Section 660-2.3 Section of the 2016 FDOT standard specifications for Road and Bridge Construction specifies traffic data detection system performance requirements as follows: "Provide a vehicle detection system capable of meeting the minimum total roadway segment accuracy levels of 95% for volume, 90% for occupancy, and 90% for speed for all lanes, up to the maximum number of lanes that the device can monitor as specified by the manufacturer."

Point detectors have been used to measure traffic speeds at a point on uninterrupted facilities like freeways, and these speeds are then used to estimate the travel times between segments. An assessment of speed measurements in Nebraska (Grone, 2012) estimated the speed measurement accuracy to be 95% when testing two widely used true presence microwave detectors, although there may be evidence that at least one of the technologies may overestimate speeds under congested conditions. A Minnesota study (Minge et al. 2010) found a speed inaccuracy of less than 1 mph of non-intrusive point detection technologies. A study of dual loop detectors in Arizona showed an average error in 5-minute speed measurements of 6-7%, with an error range of 3-12% (Samuelson, 2011). However, one should note that the segment travel times estimated based on these speeds will have higher errors than the speed measurement errors at the point detection locations. This depends on the error in the point detector speed measurements, but also depends on the distances between detectors, the method of travel time estimation from point detections, the variations in the congestion conditions between the detection locations, and the speed of the congestion shockwave. Previous research conducted by this project's researchers (Xiao, 2011) and based on simulation analysis found that the mean absolute percentage errors were 1.3% and 1.6% for uncongested conditions when the detector spacing was 0.3- and 0.6-mile, respectively, if the point detection had a random error of 90%. The corresponding values for congested conditions during a one-lane blockage incident with fast-growing queue were about 10% and 20%, respectively. That study also examined the reliability of the estimated travel time, which was defined as the percentage of vehicles with travel times that are within the range of the travel time posted on a traveler's information device. The results showed that 99.75% to 100% of travel time estimates were within the posted travel time range for uncongested conditions, with a detector spacing between 0.3 and 0.6 miles, while the reliability of travel time estimates dropped to 74.5% and 54.0% with these two detector spacing for congested conditions during the simulated one-lane blockage incident.

Increasingly, the FDOT districts are using Bluetooth and/or Wi-Fi data to measure travel time, particularly on urban arterials, in addition to the utilization of third party vendors for these purposes. It is expected that the accuracy of these measurements depends on the sample size, which is a function of the demand levels on the facility at different times of the day. A study conducted for Washington State DOT found that the mean absolute percentage error (MAPE) of the data collected from the Bluetooth devices and the Bluetooth/Wi-Fi combined device ranging between 13% and 20% for an arterial street. The accuracy of measurements of freeway travel time is expected to be better. An evaluation conducted for the I-95 coalition in 2013 found that third-party probe data can adequately detect congestion on arterial streets when the number of signalized intersections per mile is less than or equal to 1 on principal arterials with average annual daily traffic (AADT) of 40,000 vpd or more, but increasingly underestimates congestion, as the number of intersections increases due to the increase in the variations in travel times and the decrease in volume due to having smaller sample sizes than statistically required (Young et al., 2015). As private sector companies increase the sample size and have new sources of data, the provided travel time estimates will be able to better capture the congested conditions on arterials. Such improvements have been found in an ongoing investigation that will be published in an update to the I-95 Coalition report mentioned earlier (Young et al., 2015).

In general, the sample sizes of Bluetooth and Wi-Fi data and third-party vendor data are not sufficient for low traffic volume conditions. In addition, the Bluetooth and Wi-Fi readers cannot be placed short distances apart due to the inaccuracy of the identification of the position of the vehicles. Thus, they cannot be used to determine travel time on short urban street links between intersections.

There are few studies that appear in the literature regarding travel time estimation using connected vehicle data. Zou et al. (2010) estimated travel time based on CV and found an average error percentage of 27.6%, 12.5%, and 8.2% for 1%, 5%, and 10% market penetrations, respectively. These estimates were based on traffic simulations of a hypothetical network that simulates vehicles broadcasting PDM data according to J2735 standards. Vasudevan and O'Hara (2015), in a workshop presentation about an unpublished work, summarized a method to estimate the travel time and back of queue location using BSM and PDM data. Their methodology does not use the vehicle ID and the trajectories of the vehicles. The BSM includes temporary vehicle IDs but does not keep the same ID for a long period of time to protect privacy. Argote et al. (2012) estimated measures of effectiveness based on real-world vehicle trajectories. These measures of effectiveness include queue length, speed, number-of-stops, acceleration noise, and average delay per unit distance. They used the Next Generation Simulation (NGSIM) data for the testing purpose. A drawback of this study is that it uses the vehicle ID, but does not consider the change of vehicle ID during its course of travel, as specified in the J2735 standards. In the current study, the quality of travel time estimates based on CV data on freeway and urban street segments have been assessed. The results are discussed in this and later sections.

4.2.2. Occupancy and Density

Density is an important measure to assess freeway level of service and is used as the decision criteria in the 95 Express dynamic pricing algorithm. As stated in Section 3.2.1, a density accuracy of 90% is required by FDOT standard specifications of traffic detection. At the present time, the density of a segment is estimated based on point traffic detection using different methods. A study of a freeway segment by Hadi et al. (2014) showed that the density estimation MAPE values for a segment with a 0.5-mile detector spacing range between about 11% and 15%. Qiu et al. (2010) reported that in moderate traffic situations, when the actual traffic density range is between 60 and 100 vehicles per mile per lane (vpmpl), the error of estimating density based on loop detectors is about 17%. In congested traffic situations, when the actual traffic density range is between 80 and 140 vpmpl, the study reported that the corresponding relative errors are 30% to 40%. Gazis and Knapp (1971) estimated traffic densities from flow and speed data collected at the entrance and exit points of a section and showed that the error percentage is less than 10% for most of the investigated cases.

A study by Khan (2015) showed, based on simulation modeling, that the use of CV data as input into an advanced estimation algorithm can provide an accuracy of at least 85% when the connected vehicle penetration level was 50% or more with the estimation accuracy increasing with the increase in the market penetration, as shown in Figure 4-1. In this figure, "SVM" and "CBR" are two different methods that estimate density. The same study showed that density estimations that used an algorithm based on point detector data produced an accuracy rate between 42.5% and 62.2%. An incremental benefit-cost analysis indicated that the use of CV provides a higher return on investment, compared to the use of loop detectors. Unfortunately, the study did not assess the accuracy of CV data utilization for market penetrations below the 50% market penetration level.

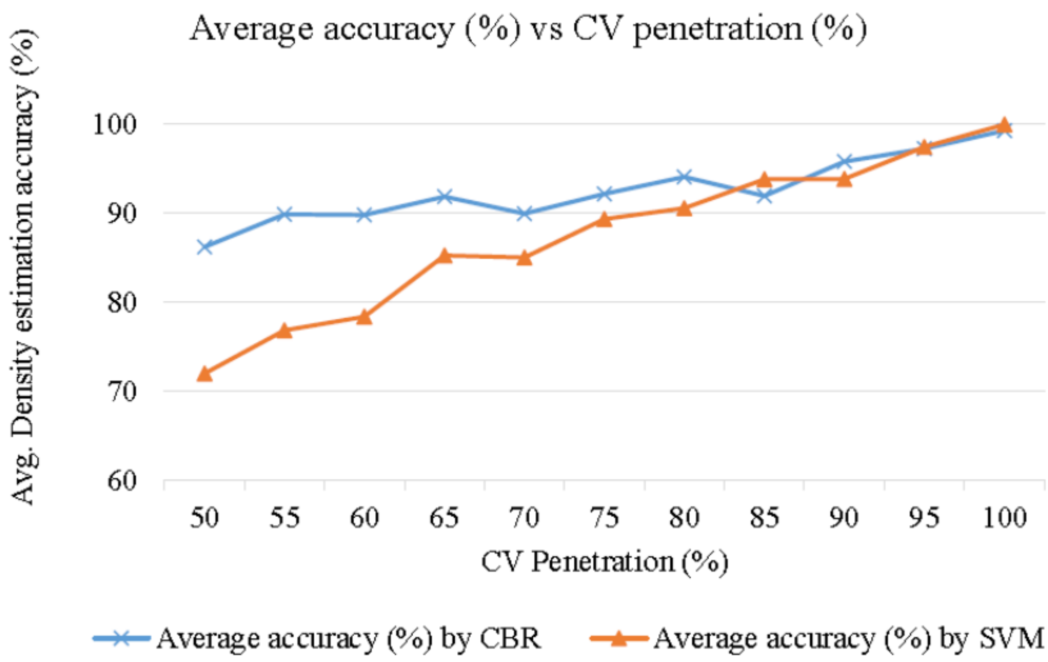


Figure 4-1 Average Density Estimation Accuracy for Different CV Penetration Levels (Khan, 2015)

Panichpapiboon and Pattara-Atikom (2011) evaluated a method to estimate density by an equipped vehicle based on the number of neighboring connected vehicles in its vicinity of the equipped vehicle. The study showed that the accuracy of density depends on the transmission range of communicating vehicles, number of vehicles with communication capabilities, and roadway traffic conditions. At lower market penetrations, the mean absolute error percentage was as large as 70%. It was found that the error percentage exponentially decreases with the increase of the market penetration reaching less than 10% when the market penetration is 50%. Garelli et al. (2011) used a similar approach to estimate traffic density. They performed the test for 30%, 50% and 100% CV market penetrations. They found that the methodology estimates the density accurately at low traffic volumes and with an error of 15% to 20% at high volumes. Barrachina et al. (2013) and Barrachina et al. (2015) showed that the use of V2I and V2V technology can produce an accurate estimation of density with an absolute mean error lower than 3% at very high market penetration.

A number of studies examined the potential of using a low sample size of probe vehicles in combination with point detector data to improve density estimation accuracy. Al-Sobky and Mousa (2016) conducted a study to determine traffic density using two smartphones inside two vehicles and an observer to obtain count data. The results show that the measured and ground truth densities had no significant difference at the 5% significance level. This indicates that at a relatively very low market penetration of connected vehicles, the CV data can be combined with counts obtained using point detection, and can produce good results. The error of the density estimated using this method ranges from 1.3% to 15%, with an average of 8%.

Qiu et al. (2010) combined detector data with probe data to estimate density and found that the relative error for the given periods can be improved from 30% based on point sensor data, to 4-6% based on point sensor data plus probe vehicle data. They used two loop detectors 1,000 feet apart, with two probe vehicles driven five round trips along the section. Once again, this indicates the potential of using CV data with point detection to estimate density at low market penetrations of CV.

4.2.3. Traffic Volume

A 95% to 98% accuracy rate for volume measurements has been specified as a requirement by transportation agencies. As stated in the previous section, a volume accuracy of 95% is required by FDOT standard specifications of traffic detection. The main permanent automated source of traffic volume measurements are point detectors, although video analytic products that utilize CCTV camera images have been recently proposed for this purpose. Point detectors like the inductive loop, video image detectors, magnetometers, and microwave detectors were found to produce acceptable volume count accuracy, although some of these technologies are subject to errors, particularly during congested conditions when the proximity of vehicles to each other can result in counting more than one vehicle as one vehicle. Video-based detection products are also affected by adverse weather and lighting conditions. A Minnesota DOT (Minge et al., 2010) study found that four tested non-intrusive detection technology products produced a volume accuracy comparable to loops (typically within 1.6 percent), during both free-flow and congested conditions. However, a per-vehicle analysis revealed some occlusion with the microwave detection technology when slow moving trucks in the lane nearest to the sensor blocked subsequent lanes, resulting in the undercounting of about 20% in the occluded lanes in periods of heavy congestion when considering short counting intervals. This is expected to be a function of the number of lanes and trucks on the freeway. A study in Nebraska (Grone, 2012) found an error in a one-minute traffic count ranging from 5.5% to 8.2% for four widely used non-intrusive point detectors. However, this error dropped at higher aggregation levels (5 or 15 minutes). Nihan et al. (2002) found an error of just 1-3% of volume measurements using loop detectors when aggregated at the 60-minute levels. However, when examined at the 20-second level, 22.1% of the intervals had incorrect values. A study of loop detectors in Arizona showed an average error in 5-minute counts of 3% to 6%, with an error range of 1-20% (Samuelson, 2011).

With regard to video analytics based on existing CCTV cameras, an ENTERPRISE pool funded study (Preisen and Deeter, 2014) found average traffic volume errors of 9% during daytime conditions and 17% average errors for nighttime conditions. The study found a 14% average error for the AM peak period, and 9% average error for the PM peak period.

In this study, an assessment was made of partially utilizing CV data to estimate traffic volumes to potentially allow the removal of some of the detectors used to collect traffic volumes. This will be discussed later in this document.

4.3. Ramp Metering and Merging Assistance

As stated in Section 3.3, the FDOT ramp metering subsystem utilizes the Fuzzy Logic algorithm developed in the Washington State. This is an advanced adaptive metering algorithms that utilizes loop detectors on the freeway mainline and on-ramps. Connected vehicle data have the potential for use to support ramp metering but requires newly developed or modified ramp metering algorithms. In fact, the fuzzy logic algorithm could be modified to utilize CV data possibly, at least initially at low market penetrations, combined with point detector data. This could be as simple as adding new fuzzy rules based on speed and ramp queues measured and estimated using CV data to the current set of fuzzy logic rule sets. The use of CV data has an advantage over using aggregated spot data collected from point sensors in that CV data allows the consideration of the actual vehicle arrivals and vehicle gaps in setting the on-ramp rates and possibly supporting the merging operations. As such, connected vehicle technology will be able to provide higher resolution in time and space that will allow new adaptive ramp metering algorithms or the modifications of existing algorithms to take full advantage of the new collected data. In addition, the resolution of the collected data combined with technologies such as cooperative adaptive vehicle control will allow a smooth merging of the vehicles by controlling the accepted gaps, cooperative behavior of mainline vehicles, and lane selection of main line traffic ahead of merge areas. However, some of these applications will require higher CV market penetrations.

Because traffic flow information is not a reliable indicator of congestion, Kattan and Saidi (2013) developed a probe-based adaptive ramp metering based on CV data and compared the results with a detector-based and pre-timed ramp metering approach using PARAMICS micro-simulation. The probe-based approach takes as its main input the space mean speed extracted from vehicle probes moving constantly on the entire freeway. The results indicated that the probe-based algorithm outperformed the two other algorithms. The sensitivity analysis showed that larger penetration rates would not significantly change the results. A 10% penetration rate is expected to be enough for a reliable probe-based ramp metering. The results showed that the probe-based ramp metering still performs better than other algorithms at low penetration rates like 3%. However, for very low penetration, such as 1%, the detector-based algorithm produces better results.

A research project conducted by the University of Virginia Center for Transportation Studies (Park, 2008; Park and Smith, 2012; Park et al., 2011) investigated the potential of using individual connected vehicle data to enhance metering strategies. Three connected vehicle-enabled local ramp metering algorithms (the variable speed limit, the lane changing advisory, and the GAP) were proposed. These three algorithms provide advisory to CV vehicles regarding speed, lane selection, and gaps between vehicles, respectively, to improve the merging of vehicles. The PARAMICS microscopic simulation was used to model a large congested network in Orange County, California to evaluate the three proposed algorithms. The results showed that the connected vehicle-enabled ramp metering algorithms improved the ramp metering performance by providing 4.3% more vehicle miles traveled while reducing vehicle hours traveled by 4.6%, which resulted in 9.3% higher average speeds. However, the proposed algorithm requires very high compliance of drivers and near full deployment of connected vehicles to achieve the highest possible benefits. Park and Smith (2012) examined the effect of gap size, compliance with the advisory message, and the level of service (LOS) at the left lane (the lane to which the lane change is going to happen) on the network performance. They found out that LOS B for the left lane, gap size of 50 feet, and compliance of 100% resulted in the best network performance.

Scarinci et al. (2013) presented a new ramp metering strategy that takes advantage of the presence of vehicles equipped with Cooperative Adaptive Cruise Control (CACC) technology. The control strategy, called Cooperative Ramp Metering (CoopRM), requested the cooperation of the main line vehicles for facilitating merging maneuvers of on-ramp vehicles released by ramp metering signals. A microscopic simulation of the CoopRM system showed a reduction in the occurrence of congestion, because of better merging maneuvers that was estimated to be between 50% and 70%, depending on the on-ramp flow. Also, the number of vehicles not able to find a suitable gap decreases between 60% and 80%. Davis (2004) analyzed mixed traffic flow consisting of vehicles equipped with adaptive cruise control (ACC) and manually driven vehicles using car-following simulations in merging area. The results showed a significant improvement in throughput (18%) for 50% ACC mixed flow relative to the flow with all manual vehicles. Pueboobpaphan et al. (2010) assessed an algorithm for on-ramp merging, intended to assist the merging process using microscopic simulation. The algorithm encourages smooth deceleration of the mainline vehicles upstream of the merging area in order to create gaps for ramp vehicles. They considered a merging assistant algorithm for situations that the mainline traffic is composed of manual and CACC vehicles and the ramp traffic is purely manual. The results showed that the effectiveness of the merging assistant algorithm can be different depending on demand and percentage of CACC vehicles on the mainline.

4.4. Managed Lanes

As stated in Section 3.4, managed/expressed lanes are being planned and implemented around Florida. The dynamic pricing model of the 95 Express in South Florida requires the density as an input to estimate the level of service on the managed lane. This density is currently being calculated as a function of point detection measurements of speed and volume. Utilizing speed measurements based on CV data are expected to produce better estimates of the density, along the highway segments, as mentioned when discussing the estimation of density as a performance measure earlier in this chapter.

As CACC technology becomes available and used, it is expected that ML strategies will benefit from giving preference to CVs to use ML. This is because these vehicles are able to run with less headway compared with manual vehicles. Thus, having higher percentages of them on the ML will improve the capacity and thus the system performance.

4.5. Variable Speed Limits

Variable Speed Limit (VSL) systems implement active traffic management strategies that dynamically adjust the speed limit based on the prevailing traffic condition, road surface condition, and weather conditions information. Such strategies are used to deal with congestion, incidents, weather and/or special events by reducing congestion impacts and crash risk. Infrastructure-based dynamic message signs are normally used to disseminate the VSL to drivers, although in-vehicle information devices can also be used. Two general applications have evolved in the use of speed limits. The first emphasizes the safety benefits of VSL, such as reducing the number of rear-end collisions and traffic homogenization (Harbord, 1995); whereas the second is more focused on avoiding or mitigating traffic flow breakdown by reducing the input flow at bottlenecks using VSL (Lenz et al., 1999). For this second type of application, the VSL signs are installed upstream of the bottlenecks, with recurring congestion as a way to reduce the speed of the congestion build-up shockwave produced once congestion starts. Lee et al. (2004) used a crash prediction model to assess the safety effects of VSL based on the simulation model. The results showed that the reduction in speed limits can decrease the average total crash potential, and the greatest reduction in crash potential is expected to occur at the locations with high traffic turbulence, such as at a bottleneck. However, the VSL also resulted in an increase in travel time. Elefteriadou et al. (2012) concluded using simulation that the VSL algorithms tested improved the mobility at bottlenecks and areas upstream of the bottleneck, and increased the throughput by a maximum of 120 to 360 veh/hr.

CV V2I applications can be used to more effective VSL implementations by collecting more detailed information about the bottleneck locations with a lower latency compare to information gathered using point detectors. In addition, the VSL recommendations can be delivered to the vehicles instead or in addition to using DMS messages. Piao and McDonald (2008) assessed the safety benefits of in-vehicle VSL instead of roadside VSL using the microscopic simulation model AIMSUN. VSLs were applied when the speed difference between a queuing section and the upstream section was greater than 12.5 mph. It was assumed that all vehicles were equipped with in-vehicle devices to communicate speeds and receive VSL. The posted speed limits ranged between 62 mph and 37 mph, with a 5 mph increment. The simulation results showed that the VSL reduced speed differences, small time headways, small time-to-collision events, and lane change frequency. This overall reduction creates homogenization and reduces crash potential. The authors also indicated that large speed variations could occur because some vehicles did not have the in-vehicle device.

Dowling et al. (2015) evaluated the impacts of the developed prototype of a variable speed limit also referred to as Speed harmonization (SPD-HARM) by Balke et al. (2014) (The SPD-HARM /Q-WARN prototype was written in the VISSIM com interface). The VISSIM micro-simulation was used to model an 8.5 miles of the US 101 freeway in San Mateo, CA. However, researchers pointed out that the Q-WARN application could not be assessed in the micro-simulation due to the lack of information on how drivers would react to the queue warning messages. Therefore, only the performance of the SPD-HARM was tested in the micro simulation. It was assumed that 100% of drivers comply with the recommended speed generated by the SPD-HARM algorithm. The study concluded that the prototype reduced the magnitude of shockwaves (speed drops between vehicles) at 10% market penetration level. It also showed the rapid increase in the benefits for the first 20% of the vehicles that are both connected and complying with the SPD-HARM recommendations. After this 20%, the rate of increase in the benefits is lower but still increasing.

4.6. Queue Warning

SunGuide does not include a module for queue warning. Nevertheless, it is included in this discussion since it has been identified as a high priority near-term CV V2I deployment by the V2I coalition. One-third of all collisions have been reported to be rear-end incidents (National Transportation Safety Board, 2001a). Recurrent congestion (bottlenecks), incidents, and work zones are three main causes of slow/stopped traffic and can lead to queued traffic conditions and consequently rear-end collisions. Queue warning systems are designed to inform drivers about the queued traffic ahead so that they can react in a timely manner. Based on the results of one study by Daimler-Benz ((National Transportation Safety Board, 2001b), 60% of the rear-end collision could be prevented if the drivers had an extra half a second. The study also indicated that 90% of the rear end collision could be prevented if an additional second would be given to the drivers.

Findings from the queue warning system (QWS) evaluation in Amsterdam showed that, the system reduced overall accident by 23% and secondary accidents by 46%. An evaluation of queue warning system and freeway lane control found a 20% reduction in accident rates (S.T.

Team, 1999). A queue warning system in Madison County, Illinois showed a 13.8% reduction in incidents (Enterprise, 2014). An end-of-Queue warning system implemented on I-35 along 96 miles in the central Texas reduced crashes by 45% and fewer rear-end collisions were observed (ARTBA Work Zone Safety Consortium, 2015). Different queue detection techniques have been tested so far. The smart work zone in Illinois (Nemsky, 2014), which is composed of Doppler speed detectors, Bluetooth readers and portable DMS were found to reduce the number of rear end crashes by 14%. This reduction occurred despite of increase in traffic volume and the higher number of temporarily closed lanes during the project. Ullman et al. (2016) investigated the safety effects of portable End-Of-Queue (EOQ) warning system and found out that it reduced the number of crashes by 44 percent and the crash cost by $1.36 million over the study period.

Most of the existing queue estimation methods are point detector-based and user either speed or cumulative volume to estimate the queue length. In the speed-based methods, downstream and upstream speed measurements are compared with a threshold. If both are less than a threshold, the queue length is assumed to be equal to the length the segment and if only downstream measurements are less than the threshold, the queue length is assumed to be half of the segment length. The cumulative volume-based methods use volume measurements instead of speed. The cumulative downstream detector arrival volume count is compared with the cumulative departure volume count and the difference between these two is estimated as the number of vehicle in queue (Nam and Drew, 1999; Zhang, 2006; Vanajakshi et al., 2009).

Petersen et al. (2013) investigated the accuracy and latency of the queue warning system for the Minnesota I-94 Intelligent Work zone (IWZ) project, which is based on point detector measurements. The results showed that the accuracy of the back of queue estimation were within one mile across different queue lengths (Note that detectors were spaced at one mile intervals). All the aforementioned QWS rely on fixed traffic sensor or cameras to detect the back of queue. So, the location of back of queue cannot be detected exactly. If the transmitted messages from the connected vehicles are utilized, the detection has the potential to be faster and more accurate. As stated earlier, queue warning is one of the four priority CV V2I applications identified by the V2I Coalition.

Balke et al. (2014) developed speed harmonization and queue warning algorithms to generate recommended speeds and queue warning information to the drivers, as a part of the USDOT Intelligent Network Flow Optimization (INFLO) prototype. The study also addressed how the prototype use the recommended speed and queue warnings generated by the algorithms to produce both infrastructure and vehicle-based warning messages. Three types of queue warning algorithms were included in the prototype: Traffic Management Entity (TME)-based, Cloud-based, and vehicle based. The TME queue warning algorithm fuses the data obtained from the traffic sensors and connected vehicles to detect the back of queue (BOQ) and generate queue warning messages through both infrastructure signs and connected vehicles.

A small-scale demonstration was conducted by TTI, Battelle and Washington State Department of transportation in order to equip 21 vehicles with connected vehicle systems traveling in a 23 mile corridor of I-5 from Tukwila to Edmonds through downtown Seattle during the week of January 12, 2015 (Stephens et al., 2015). The connected vehicle data was transmitted and gathered using bot cellular phone and DSRC. The purpose of the small-scale demonstration was to implement the INFLO prototype and test its functionality and performance in a real traffic environment. TME-based queue warning and TME-based speed harmonization with Weather Responsive Traffic Management (WRTM) were implemented. Speed data were collected from both the WSDOT infrastructure-based detectors and the connected vehicles. The collected data were analyzed in real time as the Q-WARN and SPD-HARM messages were delivered to drivers. The study concluded that no loss of BSM data was observed and there was no disruption in the algorithm due to any loss of BSM data. The data capture, processing and delivery of messages to the drivers took less than 10 second. This guaranteed that drivers receive the queue warning message 1 mile in advance of the back of queue. The Q-WARN was found to detect the back of queue 3 min sooner and could locate the back of queue more accurately (0.5 to 1.5 miles farther upstream) than the road loop detectors. Since the INFLO algorithms capture the speed data in each 0.1 mile interval they can provide a better estimation of vehicle speed in the queue than the infrastructure-based sensors that capture speed every 0.5 mile.

Li et al. (2013) developed an event-based method that uses probe data and signal timing to estimate the queue length on signalized urban arterials. The results showed when the penetration rate is 50%, the mean absolute percentage error (MAPE) is less than 18%; and, for low penetration such as 10%, MAPE is around 60%.

4.7. Weather Applications

Weather applications can provide significant mobility and safety benefits under adverse weather and mobility conditions. Weather and lighting conditions data with increased resolution, coverage and frequency can be obtained from connected vehicles and be used in combinations with data from other sources. This data will support central software decision support systems and operators to make traffic management decisions. The data can be also disseminated to travelers using connected vehicle and traveler technologies in addition to existing traveler information systems. CV data can be used to support weather applications including maintenance decision Support (particularly in areas with icy and snowy conditions), surface condition monitoring, traveler information, real-time situational awareness, weather-related crash monitoring and analysis. Weather-responsive traffic management applications such as ramp metering, VSL, and signal control can be activated based on the collected data. The Connected Vehicle Reference Implementation Architecture (CVRIA) developed by the U.S. DOT include a number of weather applications (Iteris, 2016). Table 4-1 lists these weather-related applications.

Table 4-1 Weather-Related Applications in CVRIA

Type	Group	Application Name
Environmental	Road Weather	Enhanced Maintenance Decision Support System
		Road Weather Information and Routing Support for Emergency Responders
		Road Weather Information for Freight Carriers
		Road Weather Information for Maintenance and Fleet Management Systems
		Road Weather Motorist Alert and Warning
		Variable Speed Limits for Weather-Responsive Traffic Management
Safety	V2I Safety	Spot Weather Impact Warning

The U.S. DOT has initiated a Road Weather Connected Vehicle Applications program that aims at using connected vehicle data to improve mobility, safety, productivity, and operations during adverse weather conditions (U.S. DOT, 2017). Since 2010, this program has instrumented over 600 vehicles owned by the Michigan, Minnesota, and Nevada DOTs and used these vehicles as mobile weather stations. The collected data have been used for maintenance decision support and traveler information. The program also developed a national and open data sharing system, Clarus, that can provide near real-time weather and pavement data based on all surface transportation weather observations in North America. Recently, this system was replaced by the Weather Data Environment (WxDE). A Maintenance Decision Support System (MDSS) was also developed through this program, which can provide maintenance recommendations based on route-specific prediction of weather and pavement conditions. The inclusion of connected vehicle data allows spot-specific weather forecasting and recommendations.

As a part of the FHWA Road Weather Management Program (RWMP), the Michigan Department of Transportation (MDOT) developed a Weather Responsive Traveler Information (Wx-TINFO) System that aims at assisting traveler travel decisions by providing near-time weather-related advisories and alerts (Toth et al., 2016). The Wx-TINFO system integrated data from multiple sources, including the mobile observations collected from 60 vehicles equipped with special sensors for atmospheric and pavement conditions as well as CAN bus data, weather information from environmental sensor stations, connected vehicle data use analysis and processing (DUAP) system, vehicle-based information and data acquisition system (VIDAS) deployed in a small number of MDOT fleet vehicles, and data from the National Weather Service (NWS). These data undergo a quality check and a logic/decision tree analysis to further improve the data confidence and create weather event messages. Such messages are disseminated to travelers either through DMS or the MDOT traveler information website, MiDrive. An evaluation of this system was conducted using the methods of before-after analysis and with-without analysis based on survey data and data for DMS messages, alert, events, traveler information website usage, and user delay costs. The evaluation results indicate that the Wx-TINFO system can improve the performance of real-time traffic management during adverse weather and reduce the user delay costs by 25% to 67% during NWS advisories and warnings. The system also obtained positive feedback from traffic operators. Table 4-2 lists the evaluation findings.

Table 4-2 Michigan Wx-TINFO System Evaluation Hypotheses and Findings (Toth et al., 2016)

Hypothesis	Summary of Results
1: The system will improve the real-time traffic management capabilities of MDOT TOC Operations staff during weather events.	• Six out of seven maintenance regions' operations during winter weather showed an increase in the average number and percentage of DMSs that display weather-related information during NWS Advisories and Warnings. • For DMSs that display weather-related messages during and up to three hours before NWS Advisories and Warnings, the advance notification time remains about the same if not slightly earlier for six out of seven maintenance regions. • In a survey of TOC Operators, about half had a favorable impression of the initial system. The provision of advance information was perceived as helpful, and some system changes were suggested. Adjustments were made to the system after the evaluation period to improve performance.

2: The system will improve the timeliness and content of road weather condition reporting updates to the traveling public.	• Mi Drive website visits increased by about 20 percent (associated with storms of one inch of snow or greater).
	• Mi Drive website DMS selections (mouse-clicks) was mixed—increased for three regions, remained constant for three regions and decreased for one region (associated with storms of one inch of snow or greater).
	• Travelers responding to the web-based survey indicated the following:
	  ○ They are familiar with weather-related DMS messages (76 percent)
	  ○ They responded to weather-related DMS messages by slowing down (65-87 percent) and/or changing trip plans (20-58 percent)
	  ○ They feel DMS improve safety (88 percent) and reduce delay (75 percent)
	• Weather-related messages during NWS alerts are more likely to contain specific alert terms and travel times, and fewer general announcements such as ICE AND SNOW TAKE IT SLOW and USE EXTREME CAUTION
	• Incident rate decreased for two regions, remained constant for two regions, and increased for two regions.
	• User delay costs during NWS Advisory and Warning alerts decreased statewide (25 to 67 percent).

3: The system will provide road weather condition information that MDOT Maintenance staff perceive as valuable for possible use in road weather maintenance operations.	• A limited survey of Maintenance staff indicates the value a Wx-TINFO system could provide during Maintenance operations. In general, roughly half of the respondents had a favorable opinion of the system potential, while others felt that the current system was fine. • Staff who supported the system citied the potential for improved accuracy and detail as the reason for increased value, while others perceived potential challenges with interpreting the large quantities of data, and potential distractions for maintenance vehicle drivers

The National Center for Atmospheric Research developed a Pikalert Vehicle Data Translator (VDT) that can match vehicle-based roadway and atmospheric measurement (for example, CAN messages) with weather data from traditional weather sources such as RWIS, NWS and so on, and produce current and forecasted road and weather conditions (FHWA, 2016). Currently, this tool has been running in real-time as part of Integrated Mobile Observation project in collaboration with Michigan, Minnesota, and Nevada Department of Transportation. The tool is available for downloading from the FHWA Open Source Application Development Portal (OSADP) website.

Hammit and Young (2015) tested the possibility of using connected vehicle weather data for the operation of rural variable speed limit. In this study, one research vehicle was equipped with connected vehicle technology and the collected vehicle data was then input to different algorithms including the Pikalert VDT system to determine the usefulness and accuracy of such data for setting up variable speed limit. The results of this study show that connected vehicle technology provides a new method for addressing adverse weather conditions. However, it also concluded that using the off-the-shelf CAN-bus technology is not adequate for collecting a complete vehicle data without standardized data from all vehicles.

4.8. Wrong Way Driving

According to the data from the National Highway Traffic Safety Administration (NHTSA) Fatality Analysis Reporting System, approximately 360 fatalities that occur along controlled-access highway every year are due to wrong-way driving (National Transportation Safety Board, 2012). Wrong-way driving is usually caused by distracted or confused driving, or driving with impaired driving such as driving under the influence (DUI) or driving while intoxicated (DWI). Currently, many transportation agencies rely on traditional ITS technologies to detect WWD (for example, radar detector, camera, thermal sensors, and inductive or magnetic in-pavement sensors), and in-pavement marking or signs to alert the wrong-way driver or neighboring drivers. However, there are some limitations associated with the existing detection technologies (Randolph, 2016). High false detection rate is reported for Doppler and microwave radar detectors. Camera-based video imaging processing products do not work well under the low light conditions. Detection from thermal sensors do not always occur, while in-pavement sensors such as inductive loop or magnetic have high installation and maintenance costs, and also require the vehicles to remain in the same lanes. The verification of WWD event and activation of warning message on signs need to be accomplished manually.

The advancement in connected vehicle technologies provide alternatives to detect WWD, alert vehicles, and communicate with traffic management centers. Connected vehicle technologies allow the collection and monitoring of vehicle location, speed, acceleration, and heading information. This allows a comparison between vehicle current path with underlying geographical map data providing another way to detect WWD. The processing of data for WWD can be conducted either within the roadside unit or in vehicles. An application example of this type of technologies is the driver assistance program developed by the BMW Group Research and Technology in Germany (BMW Group, 2007). It was reported that this program can reduce wrong-way drivers. The wrong-way driving can also be detected by the vehicles driving in the opposite direction that are equipped with onboard collision avoidance systems (Finley et al., 2016). Once a WWD event is detected, an alert can be immediately issued to the wrong-way driving vehicles through in-vehicle alert system, or to nearby connected vehicles using vehicle-to-vehicle (V2V) communications, or to vehicles not equipped with CV technologies or vehicles in the large area through roadside information devices. The traffic management centers can be notified of the WWD event and disseminate the WWD alert through the vehicle-to-infrastructure (V2I) technologies. Autonomous vehicle technologies can identify no-entry signs using camera installed inside the windscreen as a part of an in-vehicle alert system and provide both an audio and visual alert to the driver, as is the case with the system developed by Daimler AG and planned for potential implementation in the Mercedes-Benz S-Class and E-Class model vehicles (Szczesny, 2013).

Based on an extensive review of the state-of-the-art of connected vehicle technologies and also expert opinions, Zeng et al. (2012) produced two list of potential connected vehicle applications. After prioritizing these applications based on their deployment feasibility, the research team selected wrong-way driving, as one of five CV applications, for full concept of operation development. Figure 4-2 illustrates the concept design of the wrong-way warning system. As shown in this system, it includes four components, roadside units, onboard unit, a detection, and a GPS. A WWD event is detected through loop detectors and the detection signals are processed by the first roadside unit. The first roadside unit also processes the signals from onboard units. The second roadside unit is optional and its main function is disseminating WWD message. One disadvantage of this system is that it requires onboard units to be installed on each vehicle.

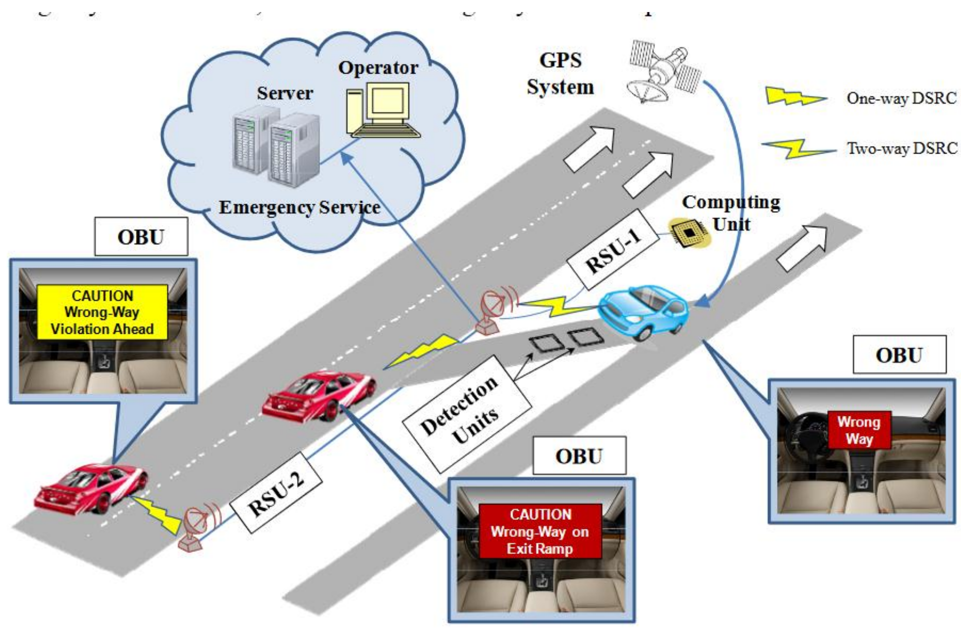


Figure 4-2 Concept of Wrong-Way Warning System (Zeng et al., 2012)

Finley et al. (2016) developed a concept design of high-level wrong-way detection and management system for the Texas Department of Transportation, which covers each stage of wrong-way driving, including WWD detection, verification, notification, alert, and clear. Figure 4-3 illustrates this concept of operation. As shown in this figure, this system takes both vehicles equipped with and without connected vehicle technologies into consideration. Traditional sensors are used to detect wrong-way driving vehicles without CV technologies while roadside unit detects the wrong-way driving connected vehicles.

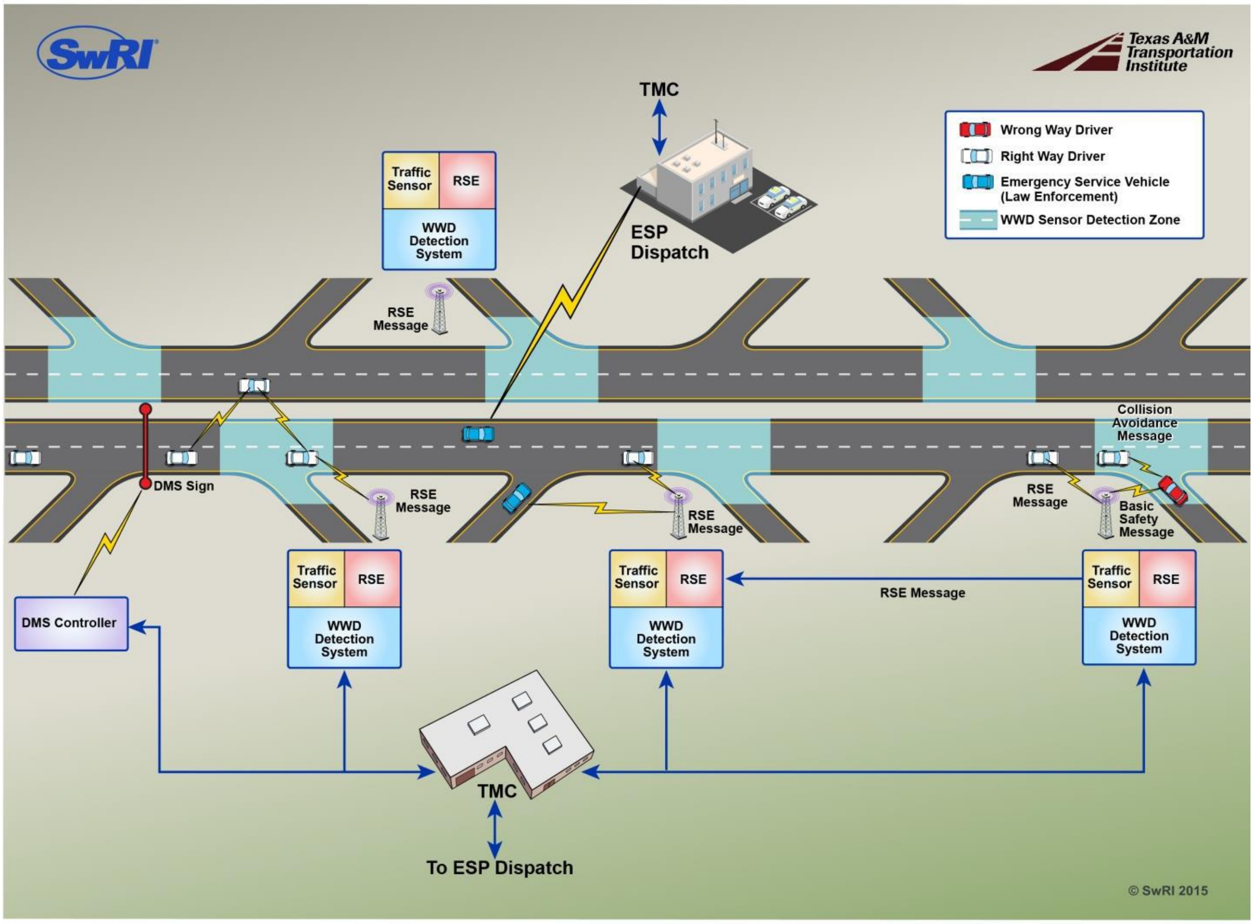


Figure 4-3 WWD Detection Zones and Integration with TMC and ESP (Finley et al., 2016)

4.9. Safety Barrier

The application of CV technologies for safety barrier are similar to those for wrong-way driving. No current or conceptualized applications of CV technologies to provide this functionality was found in the literature.

5. CONNECTED VEHICLE DEPLOYMENT DEMONSTRATION

This chapter reviews the current state of the connected vehicle technologies utilized in the Orlando implementation and describes the use of connected data collection processes and cellular communication technology to transmit probe vehicle data to central location. An analysis of the collected CV data is also presented in this section.

5.1. Needed Update to the Orlando Connected Vehicle Deployment

This section presents a discussion of the connected vehicle deployment of the ITS World Congress in Orlando, current status of the deployment, alternative deployment analysis for updating and making the deployment operational for the purpose of this project, and the needed upgrades to the existing OBU and RSU.

5.1.1. Alternative Analysis

The FDOT demonstrated a connected vehicle implementation at the 2011 Intelligent Transportation System (ITS) World Congress in Orlando, FL. The demonstration included onboard units (OBU) installed on the Road Ranger service patrol vehicles, Lynx buses, and I-Ride Trolleys, 5.9 GHz Dedicated Short Range Communications (DSRC) roadside units (RSU) connected to the center through the FDOT District 5 fiber network (a total of 29 units on three corridors), backend servers at FDOT District 5 Regional Transportation Management Centers (RTMC), and an enhancement made to the SunGuide software to allow it to capture and process the connected vehicle data. This section discusses the status of the original units and alternatives for establishing an operational implementation for the purpose of this study.

<u>DSRC Connectivity</u>

The original RSU deployment for the 2011 ITS World Congress are no longer operational. Many were removed as a result of the construction project on I-4 and the remaining were either removed for other reasons or do not have a serviceable backhaul communications connection. Thus, if a DSRC solution is to be adopted for this project, a new installation of RSU will be required. The following options have been considered for DSRC connectivity.

Road Ranger Connectivity: To support DSRC connectivity for the I-4 Road Rangers, any new DSRC RSU should be installed on I-4 corridor segments that are not within the activities of the current major construction on the corridor. Considering this constraint, it was determined that the best segment for this purpose is the segment from Exit 58 (the CR 532 - Poinciana / Kissimmee exit) to Exit 72 (the SR 528 exit). This segment corresponds to Segments No. 5 and 6 of the I-4 Road Ranger coverage. It should note that according to the FDOT District 5 Road Ranger schedule, different trucks are used for each segment for different shifts and different days of the week. This is an important consideration since the OBU are expected to be only available on a limited number of trucks. Thus, with this option, coordination with FDOT District 5 is necessary to ensure that all the road ranger vehicles selected for this segment are equipped with OBU.

I-Ride Trolley Connectivity: As stated earlier, the OBUs were also originally installed on the I-Ride trolleys, operating on International Drive; a major tourist, convention, and other visitor corridor in Orlando. Based on discussions with FDOT District 5, the installation of RSU on International Drive for the purpose of this project is not advisable because the communications between the RSU and the central location on that corridor is likely to be unstable. International Drive is constantly under construction for various developments, resulting in frequent and long interruption to communications through fiber-optic lines. In addition, the OBU installed on these vehicles during the 2011 deployment are out of service.

LYNX Bus Connectivity: If LYNX transit vehicles are to be included in this project, the best candidate locations to install the RSU based on discussion with FDOT District 5 are Orange Avenue, Robinson Avenue, South Street, or near the LYMMO lanes on city streets (LYMMO is a free circulator bus in downtown Orlando). However, since these streets are not state roads, installing equipment on these streets will create additional complexity regarding coordinating the installation and operation activities. In addition, LYNX fleet has changed significantly since the 2011 deployment and it is expected that only few devices on LYNX buses are operational, if any.

Based on the above, it was apparent that installing DSRC RSU units on the I-4 segment between Exit 58 and 72 is the most attractive option, if a DSRC option is utilized. This means that OBU should be available on the service patrol vehicles and preferably other public agency or their contractor vehicles that use this segment.

The make of the original OBU equipment installed on vehicles in the 2011 deployment was Arada on FDOT District 5 Road Ranger’s vehicles, Savaris on Lynx buses, ITRI on Mears buses, and Cohda. Currently there are ten FDOT District 5 service patrol trucks equipped with Arada devices (trucks 05-07, 05-08, 05-09, 05-10, 05-11, 05-12, 05-14, 05-16, 05-17, and 05-18). Two trucks (Trucks 05-19 and 05-20) are new trucks and there are plans to equip them with old devices from the trucks that went out of service.

It should be noted that the existing units are not connected to the controller area network (CAN) bus of the vehicles. Thus, limited types of data can be obtained from these units based on the measurements of the GPS module within the OBU. This data includes timestamp, location, speed, heading, and elevation. Vehicle acceleration can also be derived based on these measurements. Connecting the OBU to the CAN will allow the collection of additional data items specified in the SAE 2735 standards, which are not available in the current setup, as described later in this document.

Although the DSRC option was initially the preferred option, it was determined that it is not feasible for the purpose of this study, as described in Section 5.2. Other vehicle-to-infrastructure communication options were investigated for this project, as discussed next.

Cellular Connectivity

This option would use a cell modem connected to the OBU that provides a continuous stream of data. However, this option will require a cell data plan. The cellular based solution can provide the same connected vehicle message sets and/or data items that would be obtained with the DSRC solution, although it is more appropriate for PDM types of messages. A benefit of this solution compared to the DSRC solution is eliminating the geographic constraints because of the generally ubiquitous nature of cellular communications in the target deployment area. Minor software modifications are necessary in order to port the existing software of the ITS World Congress.

Connectivity from the Maintenance Yard

Retrieving data from vehicle OBU at the maintenance yard is an option even if one of the other two options mentioned above is implemented. This option would allow retrieving the data that are stored onboard the vehicles when the options of sending the data through DSRC or cellular communication is not possible. If the DSRC option on I-4 is implemented, this option can provide data for other segments of the highway with no RSU. The following options can be considered to collect data at the maintenance yard:

The WiFi Option: This solution needs to be configured so the OBU are connected to the FDOT District 5 network or an access point at the maintenance yard, where the Road Ranger vehicles park when off-duty to download the data. Similar to the cellular option, a WiFi access point would need to be installed with the existing OBU and configured such that it would connect to an access point installed at the maintenance yard to allow a process to automatically retrieve data logged locally on the OBUs. This is the preferred option, if feasible, since it is easier and less costly to implement.

The DSRC at the Maintenance Facility Option: If installing DSRC RSU on the roadside is determined to be not feasible, an option would be to install the RSU at the maintenance yard where the Road Ranger vehicles operate from and the data could be downloaded whenever the vehicles get back to the yard. This is a similar solution to the WiFi solution, but is implemented over DSRC instead. This option requires addressing network logistic issues and other considerations.

The Manual Retrieval Process: This option is the least preferred option, as it involves the tedious task of requiring District 5 to pull data at a frequent interval such as every week or month.

5.1.2. Required Updates to the OBU and RSU Devices

This section presents a discussion of how the technology deployed as part of the World Congress deployment is compared to the latest connected vehicle technology and the need for updating these devices.

The DSRC devices deployed as part of the 2011 World Congress were among the first at the transition from the Vehicle to Infrastructure Proof of Concept (VII-POC) era devices to devices that implement a new set of standards. Both the SAE (J2735) and suite of IEEE standards (1609.x) were updated, breaking backwards compatibility with the previous generation of devices. Since these new devices were early adopters of the new standards, they will need to be evaluated for accuracy and correctness of their implementation of the standards, as many manufacturers were continuously providing firmware updates for these devices as these issues were identified and addressed. It is unclear what version of the standard implementations is currently implemented by the devices.

In addition, none of the OBUs deployed as part of World Congress was integrated into the vehicle CAN bus to provide additional data elements or higher accuracy and detailed data. A goal of this project is to leverage additional vehicle data where possible, which will involve software updates to devices to facilitate the CAN bus integration.

Similar to the deployment in 2011, the Connected Vehicle community is again at a transition to a revised set of standards that will break backwards compatibility with the 2011-2015 era devices. While the revised standards are recently complete or nearly complete, discussions with various manufacturers indicate that they will not be implemented on devices for another 6-9 months, which will not allow the new standards to be used on this specific project. Fortunately, manufacturers indicate that updates from the current generation of device firmware to the version(s) that will implement the new standards will only require software updates, allowing current hardware to be upgraded without having to replace the hardware itself.

5.2. Communication Technology Considerations

The overview presented in the previous section indicates that an important task of this project was to decide on what communication technology to use for transmitting data between traffic management applications and CV. Based on the review conducted in this study, it can be concluded that both DSRC and cellular communications can be used to support most dynamic mobility applications including those that involve collecting probe vehicle data, which is the application selected for testing in this study. Although DSRC technology latency is widely recognized as better than the mobility applications. The expected introduction of the 5G technology will increase the applications that can be supported by cellular communications.

Initially, the DSRC option was investigated for use in this study, particularly considering the initial desire for the potential use of the DSRC-based RSU and OBU installed as part of the FDOT demonstration of connected vehicles at the 2011 Intelligent Transportation System (ITS) World Congress in Orlando, FL, as described above. However, it was determined that the RSU installed as part of that effort is not usable for the purpose of this study. Thus, if a DSRC solution was to be adopted for this project, a new installation of RSU would be required. This DSRC option was considered to be not practical for the purpose of this research considering the time required to get the permits for the installation of the devices, the cost involved, and the constraints of I-4 construction activities, which prohibit the installation of new devices while the construction is on-going. Thus, it was decided to utilize cellular communication to the vehicles, in this project.

It was decided to use OBU on the FDOT District 5 service patrol (Road Rangers) vehicles, for the purpose of this project, since the OBU on these vehicles can be updated and made operational for the purpose of this study. It should be mentioned that a developed cellular based CV application of the type developed in this study for FDOT road ranger vehicles can also be used by other districts. It should be also mentioned that capturing CV data from agency connected vehicles has the advantage of the ability to collect all the data available from the vehicle Controller Area Network (CAN) on the Onboard Diagnostic System (OBD-II) of the vehicle. Only a subset of this data is expected to be available as part of vehicle-to-vehicle (V2V) message exchanges utilizing the Basic Safety Message Part I (BSM I) standard. This will most likely restrict the type of information that can be captured by the RSU from general CV traffic.

The CV connectivity was achieved using a cellular modem connected to the OBU that provides a continuous stream of data to a central server. Software and hardware updates were made to the OBU installed during the ITS World Congress designed and supporting equipment to allow the OBU to communicate data using cellular in addition to DSRC, obtain CAN data through connection to the OBD-II, form messages based on the collected data and communicate the message to a central sever, and upgrade them to the new industry standards. The software and hardware set-ups are explained in this document.

One consideration is whether to use 3G or 4G cell modems, which makes a difference in the data latency and a small difference in the price of the modems. In the application of this project, the 3G option was used since the data latency of this technology is acceptable for probe data collection applications.

Another important consideration with the cellular solution is the amount of data that will be generated by each device since this affects the cost of the data plan. Limiting the size of broadcasted messages can save on the cost of the data plan. The research team produced rough estimates of the data usage based on the expected amount of data given the Road Ranger shifts. The estimates indicated that a 1GB plan for each device is sufficient for the applications of this project.

Overall, the cost for this solution is estimated to around $3,300 per vehicle (3,000 for the RSU plus $300 for the Ethernet-to-cell 3G modem, $11.60 for a Bluetooth adapter, and $7.00 for the DC-DC converter for the cell modem). Please note that in this project, the existing RSU from the Orlando ITS World Congress deployment was used, and thus no new RSU were purchased. A disadvantage of the cellular option is the requirement of a cellular data plan with monthly payments for each vehicle. However, this cost has dropped significantly recently, and the 1G data plan used in this project costs $21 per month per device.

5.3. CAN Data Collection from Road Ranger Vehicles

As noted in the previous section, the 2011 deployment does not include connections to the vehicles to obtain CAN bus data from the vehicles. Thus, limited types of data can be obtained from these units based on the measurements of the GPS module within the OBU. This data includes timestamp, location, speed, heading, and elevation. Vehicle acceleration can also be derived based on these measurements. Connecting the OBU to the vehicle OBD-II will allow the collection of additional CAN data items specified in the SAE standards, which are not available in the 2011 setup. One of the important tasks in this regard was to determine what basic CAN data elements could be obtained from the OBD-II port of each of Road Ranger vehicles. As stated above, the information available from each vehicle model is different. Car manufactures do not report what data elements are available on the OBD-II. Thus, the research team made a visit to Orlando in August 2016 and applied the process described in this section to determine these elements for each Road Ranger vehicles. Due to time and available driver constraints, only data elements that could be determined while the vehicles were stationary were identified. These included items such as brake and accelerator pedal position, exterior lights status, windshield wiper status, and transmission state.

A software module was used to determine what data items can be collected from the OBD-II port of each service patrol vehicle. The software installed on a laptop that is connected to the OBD-II to collect and monitor data from the vehicle OBD-II port. The process would receive data from the vehicle and track it in a hash table based on the CAN frame identifier for each received message. Data frames in the CAN messages are at most 8 bytes long and thus can be plotted into a 64-bit-wide image by the utilized software, as shown in Figure 5-1. The list on the left shows the CAN frame identifiers received for a specific vehicle. The black and white image on the right shows the visual representation of the CAN data received. Each row is correlated to an individual

CAN message by the identifier, in this case, 13 unique message types. Each row is represented as a sequence of black and white boxes, each representing a bit within the eight (8) byte CAN messages. Black squares indicate a 0 and white squares indicate a 1. As messages are received from the vehicle, the image is redrawn with the new data, visually showing changes of each individual bit in real-time. This identification of the changes to the bits allowed the researchers to determine what bit is changing with the changes in in vehicle inputs such applying breaks, wiper on/off, lights on and off, and so on. The following subsections include descriptions of the steps of this process.

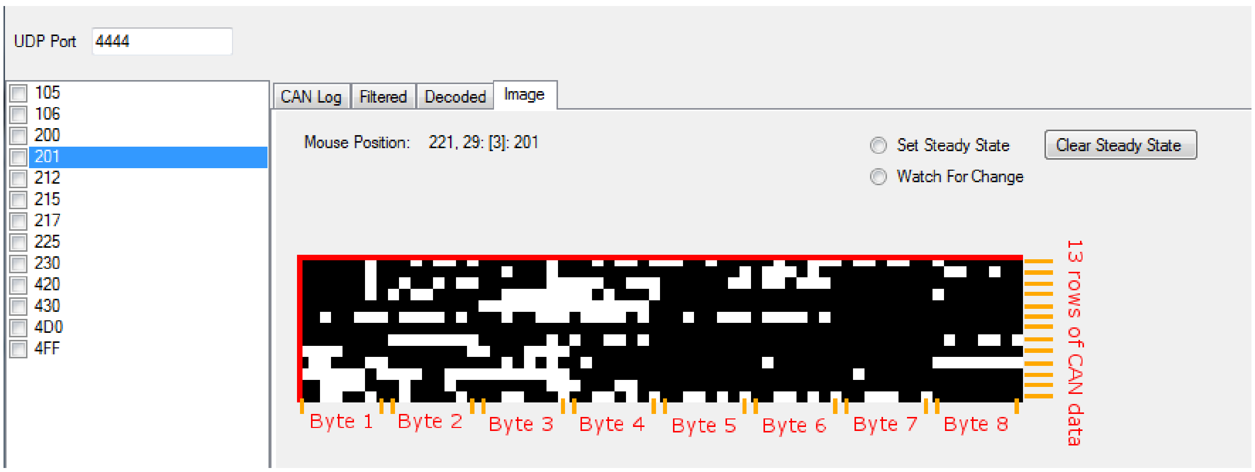


Figure 5-1 Raw CAN Data Visualization

5.3.1. Baseline State Identification

The utilized software module described above has a mechanism to indicate that the vehicle (or data of interest) is in a steady state, allowing changes during a period of time to be identified and filtered out. Figure 5-2 shows an example of this identification. The black and white image on the left shows the real time data feed and the image on the right shows the values that have changed while monitoring the vehicle in a steady state. White squares indicate that a value has changed and are masked out (setting bits to 1 in the baseline). Data fields such as engine RPM, mass air flow, temperature, and several other fields routinely change over very short periods of time, even when the vehicle is stationary and idling. These changes can make it difficult to isolate infrequent changes that result from user (or driver) input, even with the utilized software.

The longer the baseline state is monitored and the mask is built up, the more accurate and complete the mask will be. When a specific data item is being evaluated, other controlled changes not related to the data item were made during the baseline collection in order to increase the filter in a measured manner. For example, when the transmission state is being evaluated, other systems can be activated or changed during the baseline, such as pressing the brake and accelerator pedal, turning on and off each of the lights (including headlights – both high beam and low beam, turn signals, and interior lights), turning the steering wheel, opening and closing the doors, among other items. This will increase the coverage of the mask, reducing the number of unfiltered or unmasked bits that are included for processing in the next step.

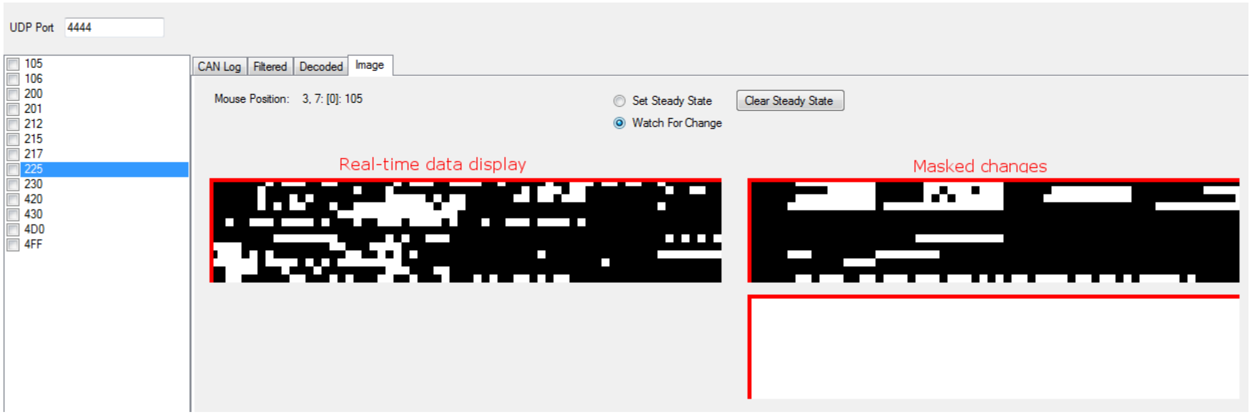


Figure 5-2 CAN Data Steady State Mask

5.3.2 Monitoring of Controlled Changes

Once the steady state mask is identified and verified to be stable, the process then focused to monitoring for changes in data compared to the steady state due to controlled changes in the vehicle operator's inputs. When in this mode, a controlled change can be made and the resulting bits that change are shown in the bottom right image, as illustrated in Figure 5-3. In this example, the exterior lights were turned on. The bits that changed as a result of turning the lights on are indicated by the green squares in the bottom right image. Hovering the cursor over those bits shows that they are in the 6th byte of CAN frame 430. Once this change is identified, the specific data for that message can be monitored on another interface in the application to identify the specific bytes/bits associated with the data items. A single message can be selected as a filter and the raw data for that message can be analyzed as the user/driver input is changed.

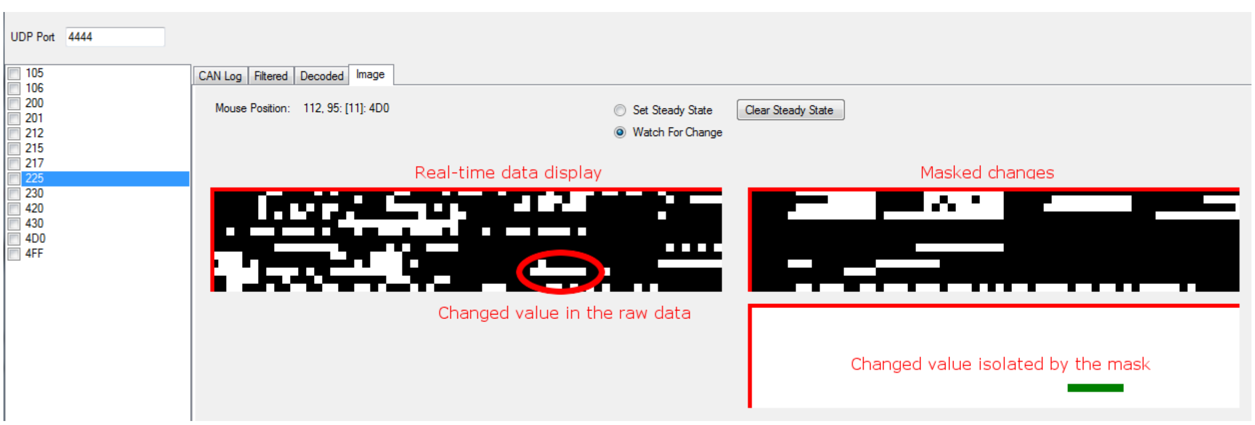


Figure 5-3 Detection of the CAN Data Change

The specific field for each data item from the OBD-II is different for different vehicle models. Thus, this process had to be repeated for each of the examined service patrol vehicles. Figure 5-4 shows one view of the data from another examined vehicle involving the identification of the data field indicating that the doors were open. In the view in Figure 5-4, data from CAN Frame ID 334 is isolated and shown streaming in real-time. The right hand panel shows complete messages over time as they are received with the newest messages at the bottom. In this particular instance, the 3rd byte changes from 08 to 00 to 10, which correspond to the rear driver's side door being open, all doors closed, and the rear passenger door being open, respectively.

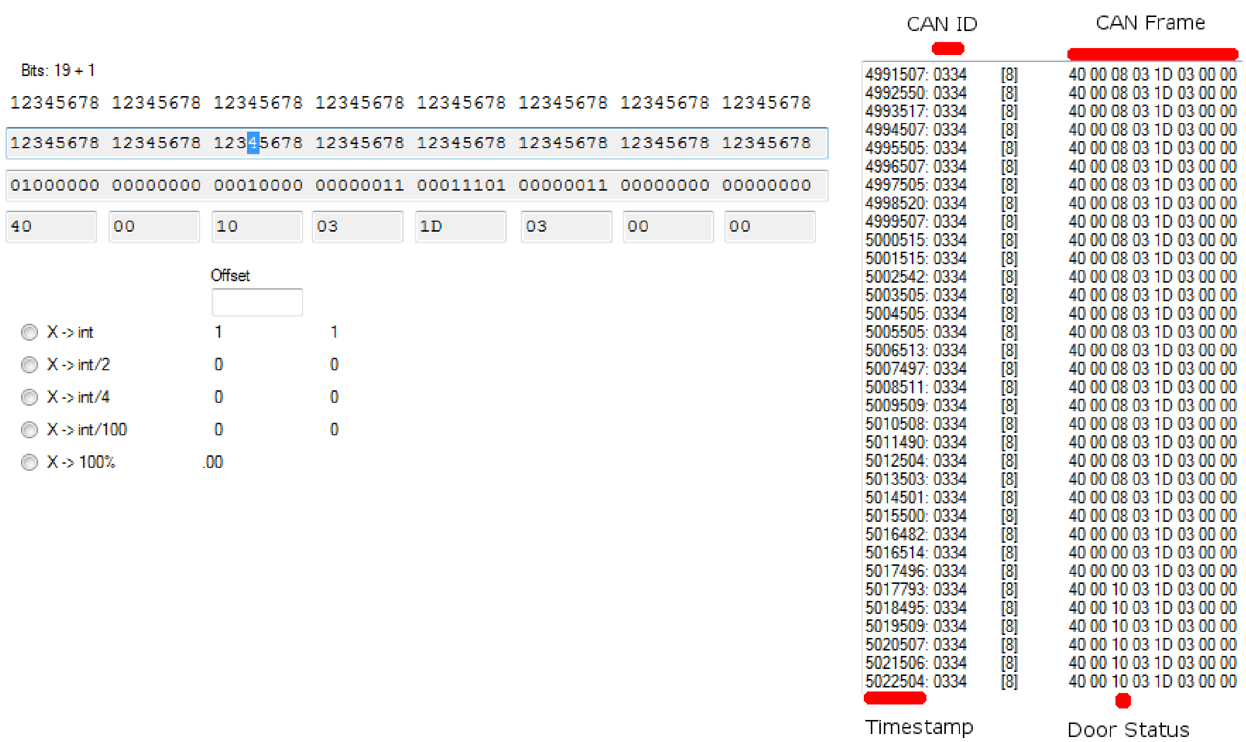


Figure 5-4 Raw CAN Data Stream Filtered That Identified the Door Open Data Item

For other data items that involve continuous rather than discrete changes in values, such as the revolution per minute (RPM) and steering wheel angle, the streaming data view also allows real-time decoding and scaling within the selected ranges of the received data using the common CAN scaling values, shown in Figure 5-5. The scaling is in accordance with the SAE J1979 standards to allow the transmission of data items that have values wit decimals. The user highlights a range for analysis, such as bytes 6 and 7 in the example of Figure 5-5. The value is decoded and scaled in real-time as data is received for that CAN message and printed in five common scaling: 1:1, 2:1, 4:1, 100:1, and as a percent of the maximum value of the range of the highlighted item. In this example, the 1:1 scaling matches the vehicle RPM, which can be watched in real-time as the accelerator pedal is pressed and released.

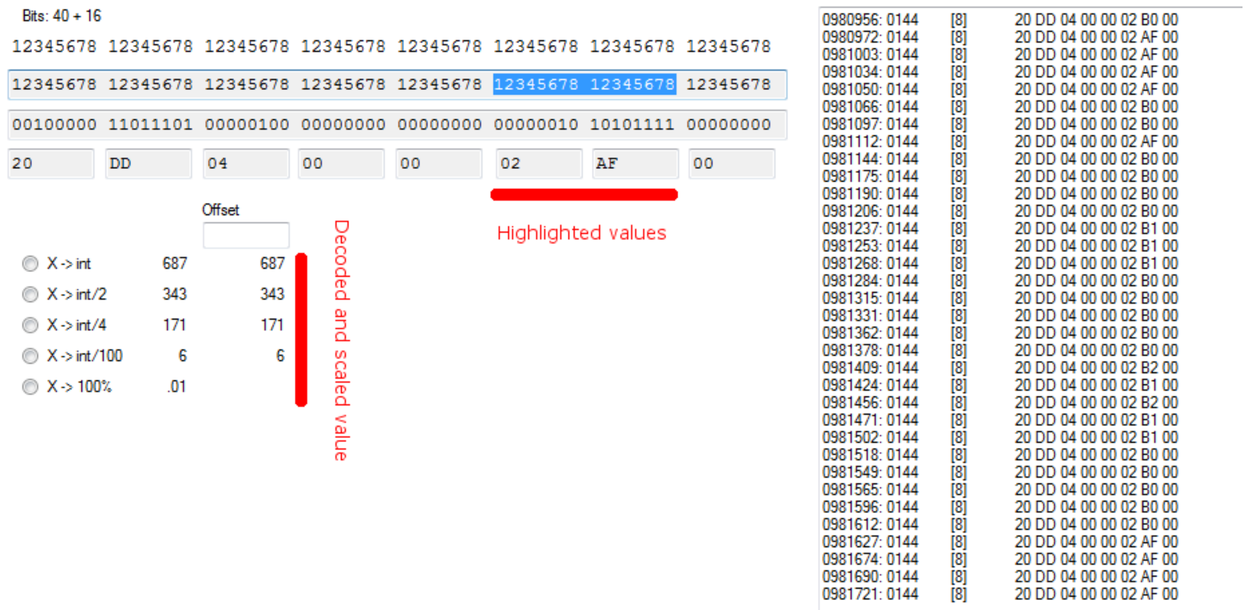


Figure 5-5 Streaming Values Decoded for Continuous Data Items

The available data was identified for ten FDOT District 5 Road Rangers trucks, as summarized in Figure 5-6. As can be seen from this figure, the amount of available data varies significantly between the trucks. The newer models and, in particular the 2015 Ram model (Truck 5-019 and 5-020), have significantly more data than other models. On the other extreme, only a single data item is available from truck 005-18, and that item was obtained by a request/response message and not broadcasted as a streaming message like the other trucks data. In addition to the items in Figure 5-5, it was found that most of the tested trucks support common request/response CAN messages, including vehicle speed, mass air flow, air intake temperature, ambient air pressure and few others, which are defined as Onboard Diagnostics Parameter IDs (PIDs) in J1979 standards.

		005-14				
	005-10	005-16		005-19	005-22	005-24
	005-11	005-17	005-18	005-20	005-23	005-25
	2005 Ford F350	2008 GMC 3500	2012 Ram 2500	2015 Ram 3500	2015 Chevy 3500	2016 Chevy 3500
Brakes On/Off	X	X		X	X	
Brake Position	X			X	X	
Brakes (Parking)	X					
Doors Open/Closed (each door)				X		
Doors Locked/Unlocked (each door)				X		
Lights On/Off	X	X		X	X	
Lights (High/Low/Running)	X	X		X	X	
RPM	X	X		X		X
Signal Lights				X	X	
Signal Lights Hazards				X		
Steering Angle				X		X
Steering Rate of Change				X		
Throttle Position		X	R	X		
Traction Control				X	X	X
Transmission State	X	X		X	X	X
Tow Haul Status	X					
Vin		X		X	X	
Wipers				X		

Figure 5-6 Identified CAN Data

The research team utilized the results from the process described in this section to customize PID filters and masks specific to each vehicle to allow isolating useful CAN messages. This is then used to identify the message and associated data frames, in which to include the available vehicle CAN data, when transmitting the data from the vehicle to the central server.

5.4. Identification of the OBU Status

The research team and District 5 staff and contractors worked together to determine the status of the original OBU on the Road Ranger trucks. It was determined that the original OBU equipment installed on the vehicles in the 2011 deployment were manufactured by Arada. The original devices were installed on ten FDOT District 5 service patrol trucks (trucks 05-07, 05-08, 05-09, 05-10, 05-11, 05-12, 05-14, 05-16, 05-17, and 05-18). The research team evaluated these OBU status and the software running on them. It was determined that several of the original units had been removed from the vehicles and most of the devices had various issues associated with the firmware, antenna, wiring, power, etc. FDOT District 5 staff checked, and fixed where necessary, cabling and connections for power to the OBUs in the vehicles, as well as the GPS antenna and DSRC antenna on each. As a result, it was possible to make the following five trucks operational for the purpose of this project: 05-25, 05-20, 05-19, 05-16, and 05-11. As stated previously, Trucks 05-19 and 05-20 has the most amount of data of the examined truck, as described in Section 3.

5.5. Device Installation and Update

Following the investigation of the availability of the CAN data on FDOT District 5 Road Ranger trucks, described in the Section 5.3, software modification were made to ensure that the messages generated from the vehicles include all the needed collected data items. The standard BSM Part I definition was extended to include a custom CAN data frame to allow inclusion of additional available CAN data from the vehicles into the generated and transmitted messages. The transmitted messages then are according to the J2735-2016-03 standard, with an extension of that standard in order to add the additional CAN data elements to the BSM. The messages are encoded using the Abstract Syntax Notation One (ASN.1) Unaligned Packed encoding rules (UPER) per the SAE standard.

It was found that the original Arada OBU of the FDOT District 5 Road Ranger vehicles does not include a hardware CAN interface as initially believed. It should be noted that the world congress deployment did not involve connections to collect CAN data. To address this, the team had to integrate and test a OBS-II Bluetooth adaptor into each of the vehicle interface for the Arada OBU. As part Bluetooth OBD-II adapters had to be purchased for this purpose. This required use of a Bluetooth adaptor rather than a cable to read the OBD-II data resulted in limitations on the data rates and buffer sizes within the communications stack. This restricted the end volume of the CAN data gathered from each vehicle, resulting in buffer errors when there is a significant amount of CAN traffic. To mitigate this issue, a custom data filters and masks were implemented for each vehicle depending on the specific useful data generated by the vehicle, in order to limit the total number of messages. This solved the above mentioned issue to a certain extent.

With the support of FDOT District 5 staff, the OBUs and the supporting equipment were installed on the five Road Ranger vehicles identified in Section 5.4 to support the data collection phase. District 5 staff installed a power switch in-line to the OBU allowing the device to be power cycled independently from the vehicle ignition. The power switch was uniquely installed in each vehicle depending on the available mounting points and where the power cable was run. See Figure 5-7 for an example of the power switch installed in a new vehicle. Older vehicles with an existing installation generally had the switch installed, as part of the center console structure.

The research team configured cellular modems that were installed in each truck allowing them to be a gateway to a backhaul server for data collection, eliminating the need for periodic on-site data collection from each unit. The cell modems selected for this project were configured as Ethernet-to-cellular routers, making the configuration transparent to the OBU. Lacking a hardware CAN interface card, as described earlier, the OBUs were configured to use a USB Bluetooth adaptor to connect to the vehicle CAN bus, shown in Figure 5-7 and Figure 5-8.

Figure 5-7 CAN Bluetooth Adapter

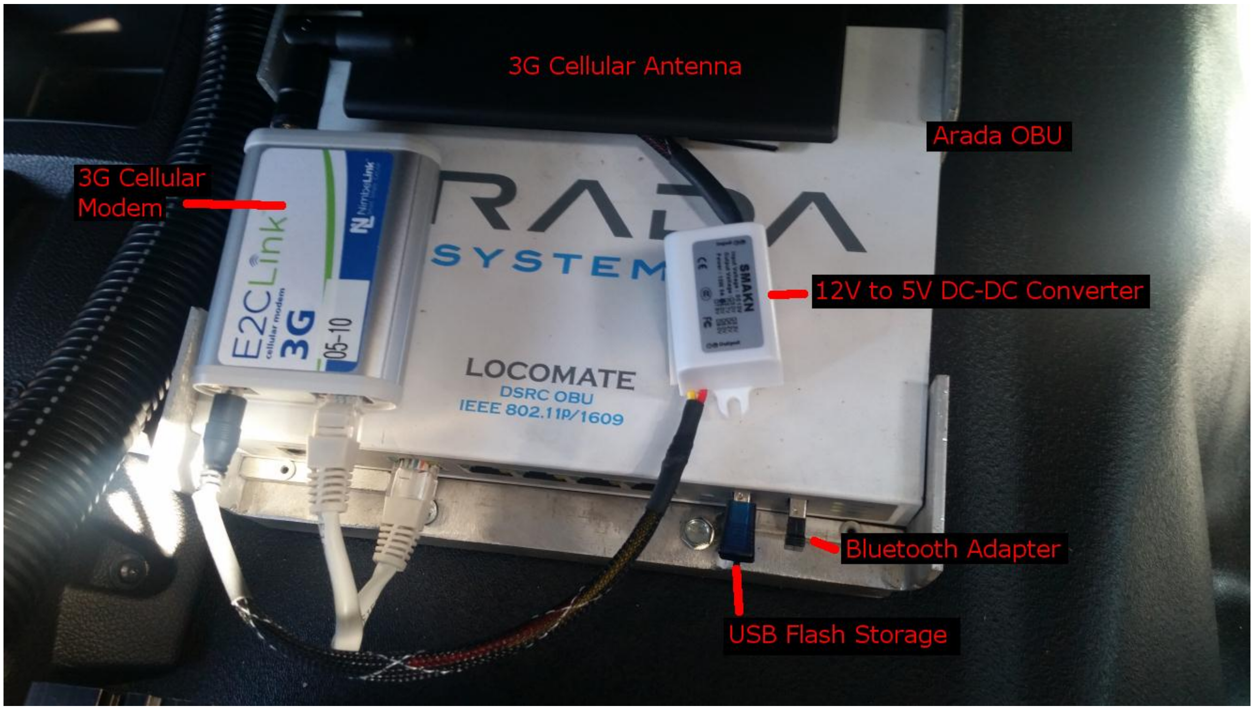


Figure 5-8 OBU Hardware and Cellular Modem

Once on site, the Arada OBUs were tested to identify the firmware version and the general state of the device. While some OBUs had been updated to the most recent firmware version, each had some level of flash storage corruption that needed to be resolved. This issue was most often with the removable USB storage on which the updates would be applied. The corruption with the unit was identified by attempting to print the configuration and firmware information for each of the devices. Corruptions were indicated by unexpected ASCII characters being printed instead of the expected actual values, as shown in Figure 5-9. Local flash corruptions within the devices were resolved by manually overwriting low level configuration values for the product name, serial number, MAC address, and hardware version. In other cases, the removable flash storage could be recovered via standard Linux utilities on a development laptop, however three USB storages were not recoverable. For these three, the USB storage was replaced with USB media taken from other Arada OBUs that were considered as spares, generally because they had other hardware or power issues that could not be addressed.

```
[root@Arada0124A3 /root]# cli
Arada0124A3#show configuration

ap information
        apname Arada0124A3
        ProductId ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
        serialNo ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ

ûT*µäÔ
        macaddress 00:26:AD:01:24:A4
        firmware-version WAVE_LOCOMATE-200_1.90.0.18
        country/region unitedstates-public-safety
        mac-address-randomization enable
        Deployment default
        LCM-status disable
        Wireless interface macaddress for 5GHz band      00:26:AD:01:20:CC
```

Figure 5-9 Flash Corruption on OBU

Once the firmware and configurations were corrected, an updated BSM generation process was installed on the OBU that would read the configured CAN data from the vehicle bus and populate the new frames/elements in the transmitted messages.   The message generation process was configured to route these messages over the cell modem interface to the backhaul collection server, in addition to transmitting them locally over DSRC.

As part of the hardware installation, a power harness was built that power the cell modem in-line with the same connector that the OBU was connected to. The cell modem required a 5V power supply, while the OBU had a 12V supply run to it, so a 12V-5V DC-DC stepdown converter was used, as shown in Figure 5-8. Figure 5-10 shows the final layout of the equipment in relation to each other for one of the vehicles. Other vehicles had similar layouts, however the exact location of the OBU varied in some of them. Aside from in front of the center console, the other typical installation location was under the front passenger seat.

Figure 5-10 Final Installation Layout

During the installation, the data collection and transmission process was monitored locally first to verify that each identified CAN data element was correctly parsed and populated in the generated messages. Once the data was verified, remote staff monitored the backhaul server to verify the data was being generated and transmitted to the end collection process and logged correctly.

5.6. Data Capture and Archiving

The data collected and communicated as described in the previous section are transmitted to a central server. The server implement a process to capture the streaming data packets in raw network packet capture files (PCAP files) with the extension name ".pcap". The packet capture process utilizes an application programming interface (API) for capturing network traffic. A captured PCAP file has a global header containing some global information followed by zero or more records for each captured packet, as shown below.

Global Header	Packet Header	Packet Data	Packet Header	Packet Data	…

Figure 5-11 PCAP File Format

There are tools that are available to read and display the data captured in PCAP files. Figure 5-12 shows a sample captured packet displayed in Wireshark, a free and open source packet analyzer tool, which was used in this study.

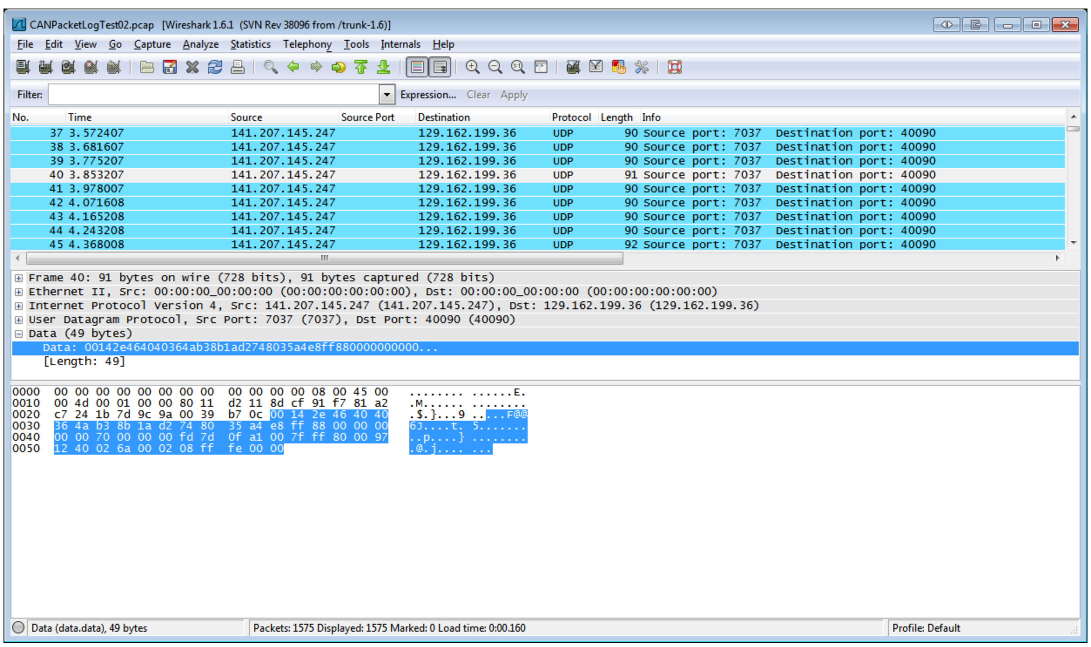


Figure 5-12 Wireshark Display of Sample Packet in the UPER Format

The constructed connected vehicle messages are encoded in ASN.1 UPER format, as described earlier. There are several options for encoding formats of the packets. The PCAP packet data portion shown in Figure 5-12 is 91 bytes, and the 48 bytes highlighted is the connected data payload. The UPER encoding format uses additional information, such as the lower and upper limits for numeric values, based on the ASN.1 specifications, to represent the data units using the minimum number of bits. Thus, it is very efficient for transferring the data. However, the information is difficult for humans to understand and requires more time to process. The highlighted portion in Figure 5-12 is the portion of the packet that contains the vehicle information to be decoded.

Another possible encoding format with the same ASN.1 data structure is the XML Encoding Rules (XER). The XER provides greater human readability compared to the UPER format. The research team developed a data conversion tool to convert the encoded payload information from the vehicles into the XER format. Figure 5-13 shows an example packet from the Road Ranger vehicles, decoded into the ASN XER. The conversion of the messages from the UPER format to the XER format increased the size of the captured data files significantly, requiring additional server space. The transmitted messages, as specified in the ASN definition file of the process described above, is an extension to the SAE J2735-201603 BSM Part 1 standard that includes additions to support the vehicle CAN data such as ambient air pressure, ambient air temperature, engine RPM, fuel rate, and so on. All other BSM (Basic Safety Message) Part 1 elements are defined and used per the standard.

The data from the five units has been monitored daily since the installation to ensure the devices are still transmitting as expected. In the approximately one-month period between the dates of the start of the installation (March 6, 2017) and April 4, 2017, approximately 1.91 Gigabytes of data have been recorded, which corresponds to 14,604,659 messages received from these five vehicles. One of the identified issues is that the equipment on two of the trucks still had the original wiring harness installation for the 2011 ITS World Congress demonstration. This setup has the unit wired to always be powered on, even when the ignition is off. It is possible that these two devices will generate more data than originally calculated and use their data plan allotment on the cell modems before each month billing cycle is over. The mitigation for this issue is to continue to monitor data usage per vehicle and try to predict if any of the vehicles will exceed their data plan. If it looks like they will, one option is to increase the data plan from 1GB to 2GB; a second option is to request those units be turned off for some or all of the month.

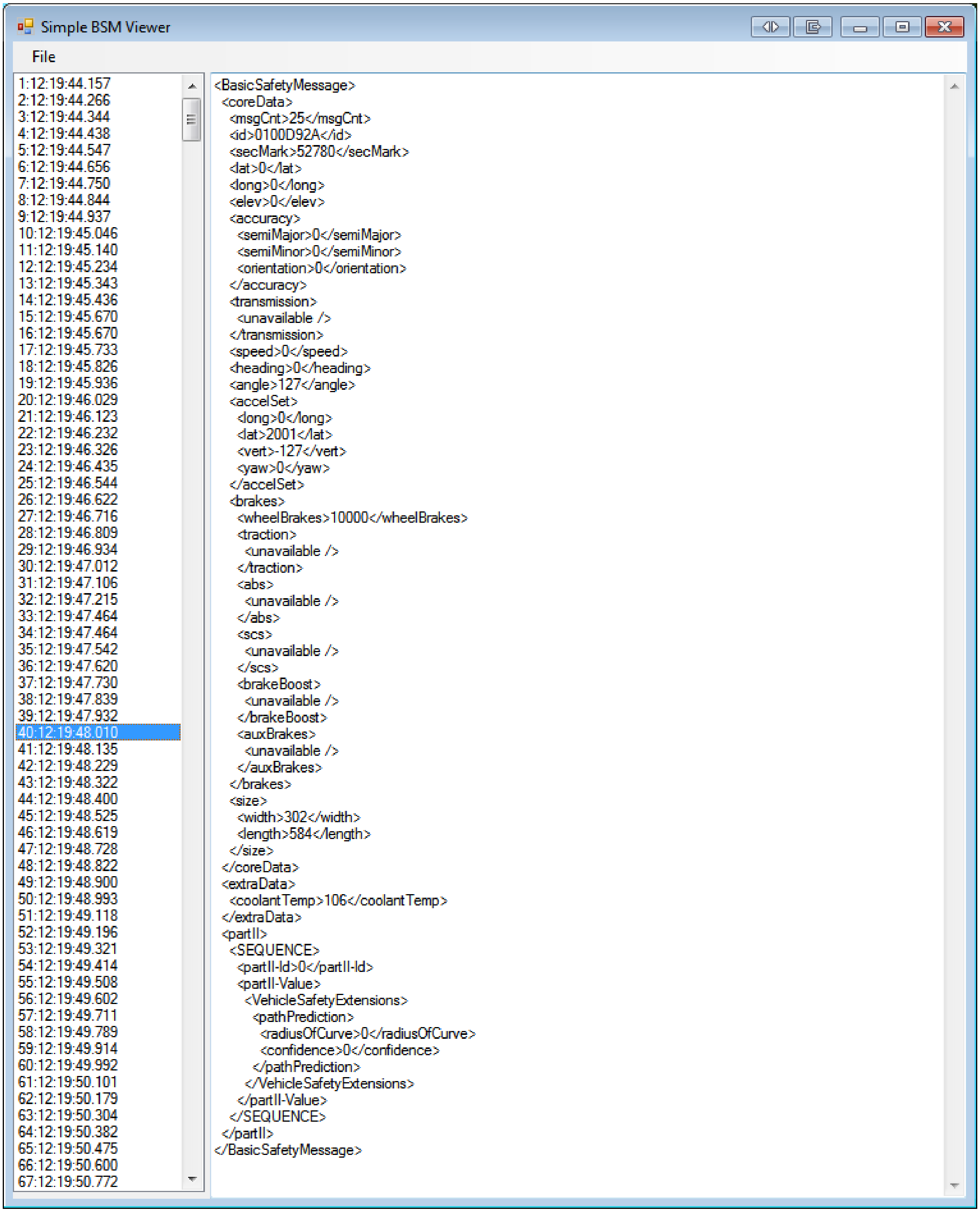


Figure 5-13 Data Decoded in the XER Format

5.7. Data Analysis

As described in the previous sections of this chapter, connected vehicle equipment were installed on road ranger vehicles in Orlando, FL. The decoded XER BSM data between March 6, 2017 and April 7, 2017 were imported into a SQL server database. A detailed analysis was conducted using the data and the corresponding results are presented in this section.

5.7.1. BSM Data Preprocessing

The collected BSM data consists of trajectories of Road Ranger vehicles. These data have to go through a set of data preprocessing procedures to be associated with specific roadway segments. Also, abnormal data need to be identified and filtered out. In this study, the section between Osceola Polk Line Rd and US 192, along the I-4 eastbound direction (EB) was selected as study segment, as shown in Figure 5-14.

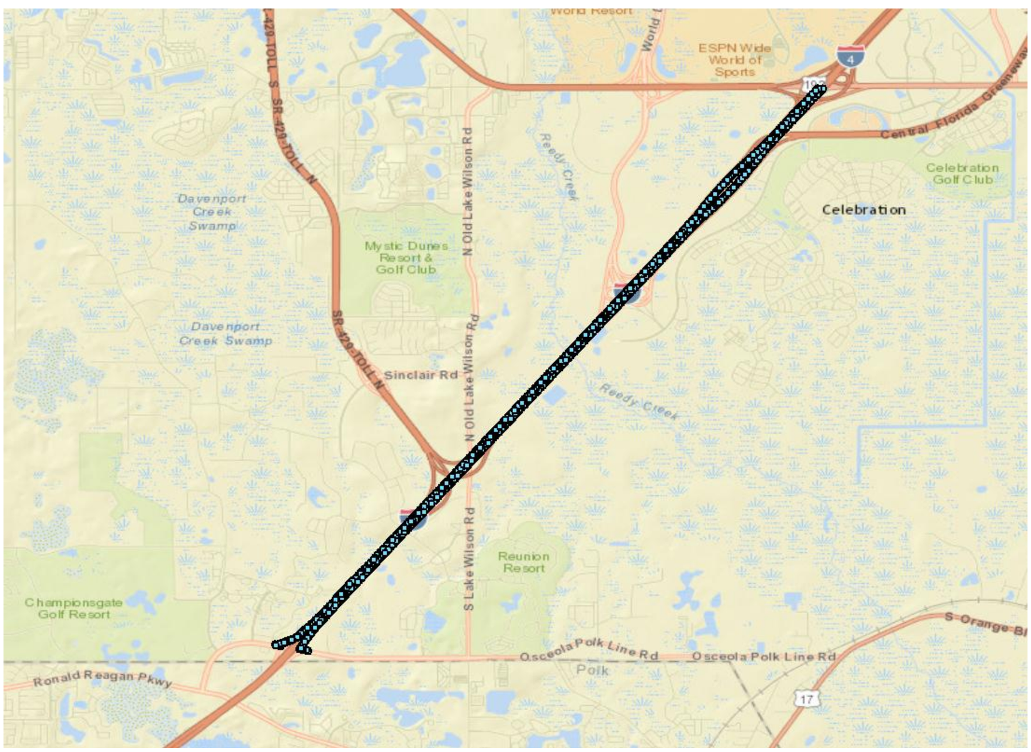


Figure 5-14 Example of BSM Trajectory Data and Location of Study Segment

The first step of the data preprocessing was to remove the BSM data with coordinates beyond the boundaries of Osceola Polk Line Rd and US 192. To reduce the large amount of data and facilitate the analysis, the BSM data were aggregated from 0.1-second into 1-second level and only the data in the AM peak period on weekdays were kept for further analysis.

The next step of data preprocessing is to filter out the BSM data along the I-4 WB and arterial streets as this analysis focuses on the I-4 EB segments. In this study, this step was accomplished by spatially joining the BSM data with the directional links included in the link shapefile of the Central Florida Regional Planning Model (CFRPM) in ArcGIS. Road Ranger vehicles may travel along the shoulder or stay off-roads to help clear incidents and the corresponding BSM data points do not represent the typical traffic conditions. Figure 5-15 shows an example of such trajectories. In this study, these data points were manually removed by overlaying the trajectories with the satellite map of the World Imagery in ArcGIS.

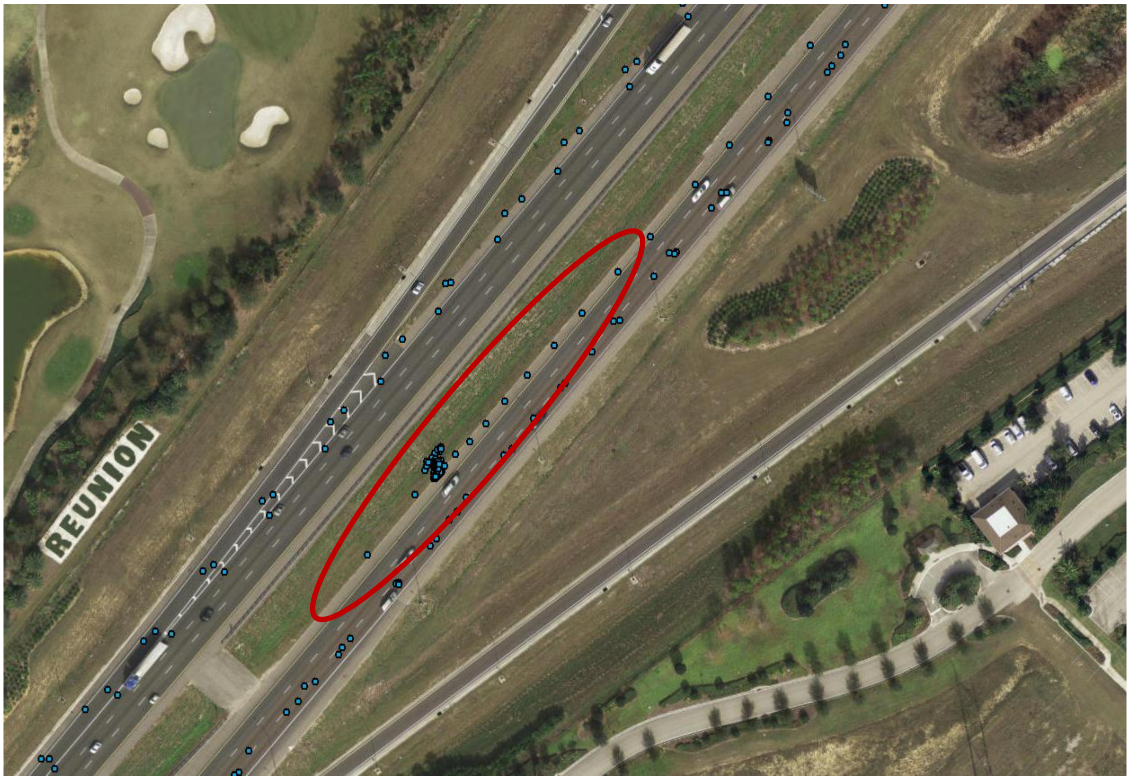


Figure 5-15 Example of BSM Data Points Located at Shoulders

The location information included in the BSM data is in terms of latitudes and longitudes. This location information was transformed to milepost-based location referencing to facilitate the calculation of performance measures for the analysis segments. For this purpose, the correspondence between the latitude/longitude and milepost location referencing was extracted from the FDOT state road shapefile using the ArcGIS split function and the geometry calculation function. The FDOT state road shapefile consists of state roads within Florida. Each state road has a roadway length associated with it. The ArcGIS split function can be used to split the roads into small segments, for example, 0.001 mile, and the ArcGIS geometry calculation function can then be applied to find latitude and longitude of each segment starting and ending points. The latitude and longitude information collected for each BSM data points was matched to the closest milepost with minimum distance from the location of the BSM data point.

During the analysis, it was also found that some data records report zero speed but with changes in latitude and longitude (that is, an indication of movement), as shown in Figure 5-16. These records with zero speeds were considered as abnormal data and removed from the analysis.

1	ts_time	ts_date	msgCnt	id	vehicleID	secMark	lat	long	elev	accuracyS	accuracy_	accuracyOr	transmissi	speed	headi
2	2017-03-08 07:47:49.000	3/8/2017	101	20-05-6F-60	20-05	56606	28.26622900000	-81.60770830000	358	0	0	0	unavailable	0	
3	2017-03-08 07:47:50.000	3/8/2017	112	20-05-6F-60	20-05	57694	28.26644420000	-81.60752500000	365	0	0	0	unavailable	0	
4	2017-03-08 07:47:51.000	3/8/2017	121	20-05-6F-60	20-05	58584	28.26657070000	-81.60740160000	361	0	0	0	unavailable	0	
5	2017-03-08 07:47:52.000	3/8/2017	3	20-05-6F-60	20-05	59576	28.26674070000	-81.60724500000	361	0	0	0	unavailable	0	
6	2017-03-08 07:47:53.000	3/8/2017	14	20-05-6F-60	20-05	661	28.26695090000	-81.60705830000	361	0	0	0	unavailable	0	
7	2017-03-08 07:47:54.000	3/8/2017	23	20-05-6F-60	20-05	1552	28.26707740000	-81.60693330000	359	0	0	0	unavailable	0	
8	2017-03-08 07:47:55.000	3/8/2017	34	20-05-6F-60	20-05	2639	28.26725070000	-81.60677330000	360	0	0	0	unavailable	0	
9	2017-03-08 07:47:56.000	3/8/2017	44	20-05-6F-60	20-05	3628	28.26742240000	-81.60661330000	362	0	0	0	unavailable	0	
10	2017-03-08 07:47:57.000	3/8/2017	54	20-05-6F-60	20-05	4617	28.26759400000	-81.60645330000	362	0	0	0	unavailable	0	
11	2017-03-08 07:48:02.000	3/8/2017	105	20-05-6F-60	20-05	9661	28.26849420000	-81.60562160000	369	0	0	0	unavailable	0	
12	2017-03-08 07:48:10.000	3/8/2017	57	20-05-6F-60	20-05	17573	28.26975240000	-81.60443160000	367	0	0	0	unavailable	0	
13	2017-03-08 07:48:11.000	3/8/2017	68	20-05-6F-60	20-05	18661	28.26994750000	-81.60424990000	365	0	0	0	unavailable	0	
14	2017-03-08 07:48:12.000	3/8/2017	77	20-05-6F-60	20-05	19551	28.27006570000	-81.60413160000	363	0	0	0	unavailable	0	
15	2017-03-08 07:48:13.000	3/8/2017	88	20-05-6F-60	20-05	20639	28.27022740000	-81.60398330000	361	0	0	0	unavailable	0	
16	2017-03-08 07:48:18.000	3/8/2017	10	20-05-6F-60	20-05	25584	28.27107570000	-81.60318660000	333	0	0	0	unavailable	0	
17	2017-03-08 07:48:19.000	3/8/2017	20	20-05-6F-60	20-05	26573	28.27125570000	-81.60301990000	327	0	0	0	unavailable	0	
18	2017-03-08 07:48:20.000	3/8/2017	30	20-05-6F-60	20-05	27563	28.27143570000	-81.60284830000	312	0	0	0	unavailable	0	
19	2017-03-08 07:48:21.000	3/8/2017	41	20-05-6F-60	20-05	28651	28.27161740000	-81.60267660000	309	0	0	0	unavailable	0	
20	2017-03-08 07:48:22.000	3/8/2017	51	20-05-6F-60	20-05	29640	28.27180240000	-81.60250490000	306	0	0	0	unavailable	0	
21	2017-03-08 07:48:23.000	3/8/2017	61	20-05-6F-60	20-05	30629	28.27198570000	-81.60233330000	306	0	0	0	unavailable	0	
22	2017-03-08 07:48:24.000	3/8/2017	71	20-05-6F-60	20-05	31618	28.27216900000	-81.60216000000	308	0	0	0	unavailable	0	
23	2017-03-08 07:48:25.000	3/8/2017	81	20-05-6F-60	20-05	32608	28.27235400000	-81.60198490000	308	0	0	0	unavailable	0	
24	2017-03-08 07:48:26.000	3/8/2017	91	20-05-6F-60	20-05	33598	28.27253900000	-81.60181160000	303	0	0	0	unavailable	0	
25	2017-03-08 07:48:27.000	3/8/2017	101	20-05-6F-60	20-05	34585	28.27272240000	-81.60163830000	295	0	0	0	unavailable	0	
26	2017-03-08 07:48:28.000	3/8/2017	111	20-05-6F-60	20-05	35577	28.27290400000	-81.60146330000	291	0	0	0	unavailable	0	

Figure 5-16 Example of BSM Data with Zero Speeds

5.7.2. Point Traffic Detector Data Collection and Processing

In order to compare the performance measures resulted from the BSM data with those from traffic detectors, the 5- minute detector data along the I-4 EB within the study segment for the same time periods were downloaded from the Regional Integrated Transportation Information System (RITIS) website. A total number of 10 detectors are located in the study segment as shown in Figure 5-17, however, detectors 6519, 6592, 6956, and 7059 always reported null data in this study periods and thus were not utilized in this study. This left the study with data from six detectors available for the analysis.

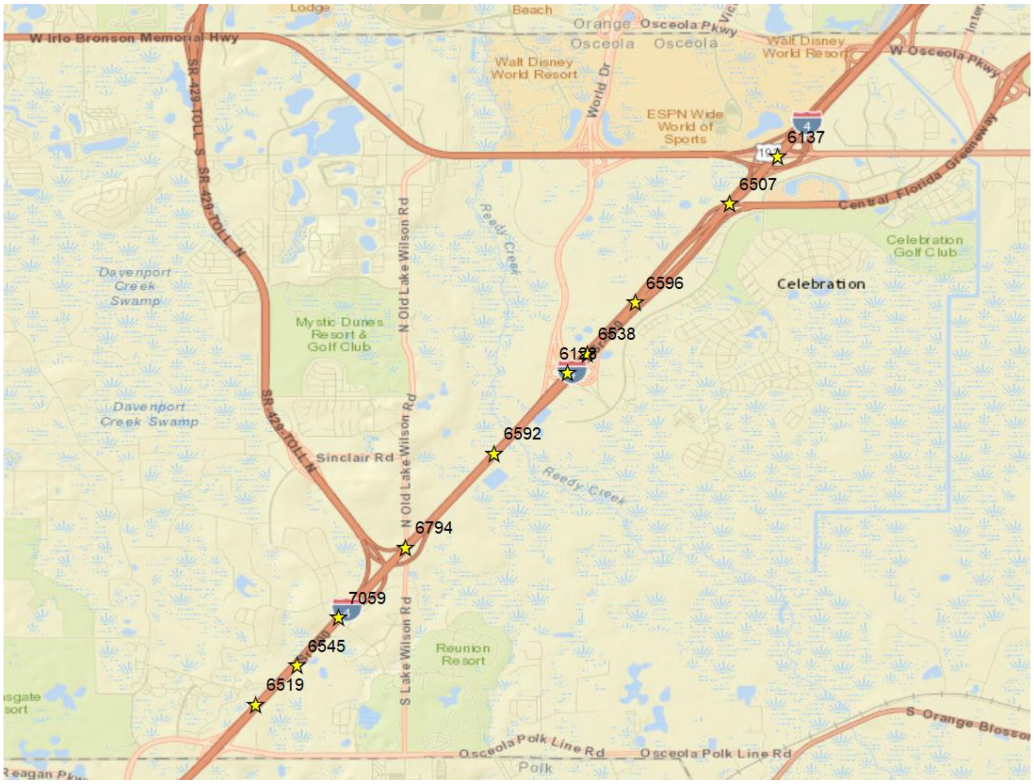


Figure 5-17 Locations of Point Traffic Detectors

5.7.3. Analysis Results

The cleaned BSM data consists of the movements along the I-4 EB on-ramp at Osceola Polk Line Rd as well as the movements along the I-4 EB mainline. The corresponding analysis results are presented separately in this section.

Freeway Analysis Results

Figure 5-18 presents the boxplots of speeds collected from the BSM data in each 300-ft segment along the I-4 EB mainline for the AM peak period. 106 mainline segments, each with a length of 300 ft were used in the analysis. The followings are the segments in the on-ramp merge or off-ramp diverge areas:

- Segment 1: the first segment after the gore area at the Osceola Polk Line Rd on-ramp.
- Segment 17: the last segment before the off-ramp to Western Beltway NB
- Segment 38: the first segment after the gore area at the Western Beltway SB on-ramp
- Segment 56: the last segment before the off-ramp to Word Drive SB
- Segment 84: the first segment after the gore area of the on-ramp at Word Drive
- Segment 95: the last segment before the off-ramp to US 192

As shown in Figure 5-18, the speed has large variations from Segment 1 to Segment 47, which is just downstream of the on-ramp from US 429 SB, but the variation becomes less at the downstream segments. It can also be seen that congestion occurs between these two on-ramp areas, especially in the area downstream of the first on-ramp from the Osceola Polk Line Rd during the AM peak period.

For comparison purpose, the point traffic detector data were also plotted, as shown in Figure 5-19, for the same period. The results in Figure 5-19 show a similar trend as that in Figure 5-18, in that the downstream locations have less variations in speed and congestion compared to the upstream locations. The speeds reported from the point traffic detectors are higher than those from BSM data, possibly due to aggregating fast and slow lane and measuring at points rather than along the segment. It may be also a function of the accuracy of speed measurement at low speeds of the utilized point detector technology. However, it is noticed that the detector data reports much lower speeds at Segment 64 (located at the World Dr. off-ramp area) compared to the BSM data. A further examination of the detector data reveals a malfunction of Lane 1 detector at this location, which only reports volume count as shown in Figure 5-20(a). It seems that this lane data was included in the calculation for the aggregated zone readings, resulting in very low zone speeds as shown in Figure 5-20(b). Also, due to the missing of data and large space between detectors, the detector data cannot capture the congestions at the first few segments. This case study shows that connected vehicle data not only can be used to supplement the measurements of point traffic detectors, but also can be utilized to verify the accuracy of detector data.

Figure 5-21 presents the distribution of the longitudinal acceleration in each 300-ft segment based on BSM data. It is seen from this figure that for most of the segments, the longitudinal acceleration is between -10 ft/s^2 and 10 ft/s^2. This was unexpected since it was expected that the acceleration variation is higher for the congested segments upstream of segment 47. The median longitudinal acceleration was about zero. It should be mentioned that the acceleration was calculated as a derived measure based on the speed measurements due to difficulty in obtaining the acceleration directly from the OBUs. This may have contributed to the unexpected results and should be checked in future studies.

The lateral acceleration, vertical accelerations, and yaw rate[1] were not available in the collected BSM data. The lateral acceleration and Yaw rate have the potential to be used to determine the number of lane changes in a section.

[1] Yaw is an indication of a vehicle's rotation from it's vertical axis indicating how far is the vehicle angled to the left or right from its center straight course.

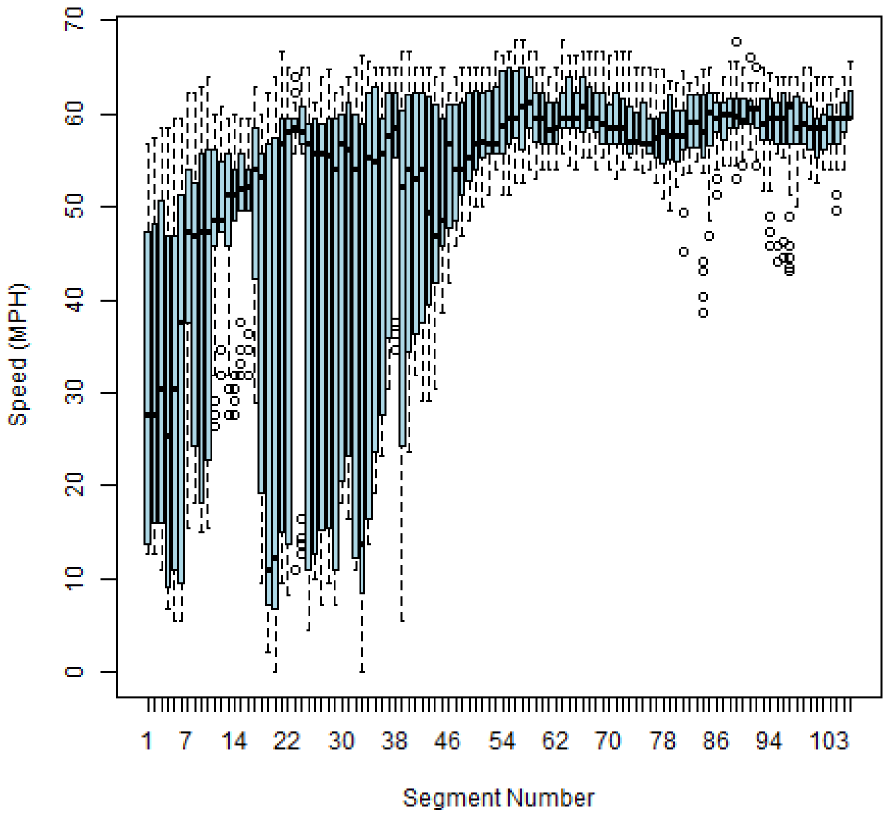


Figure 5-18 BSM Speed Distribution in Each 300-ft Segment along the I-4 EB Mainline

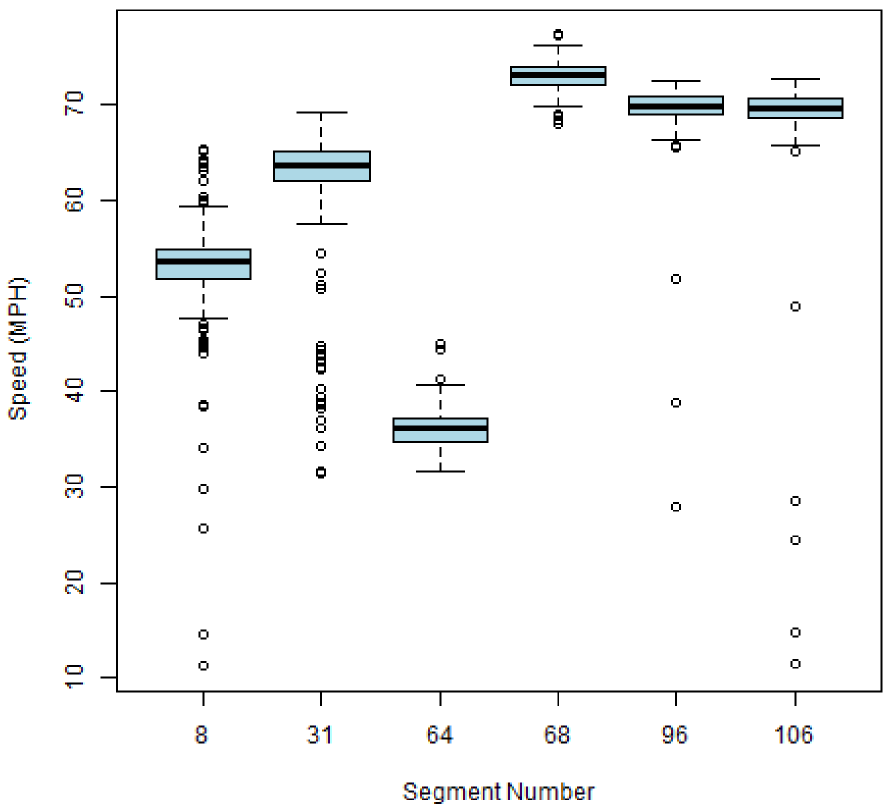
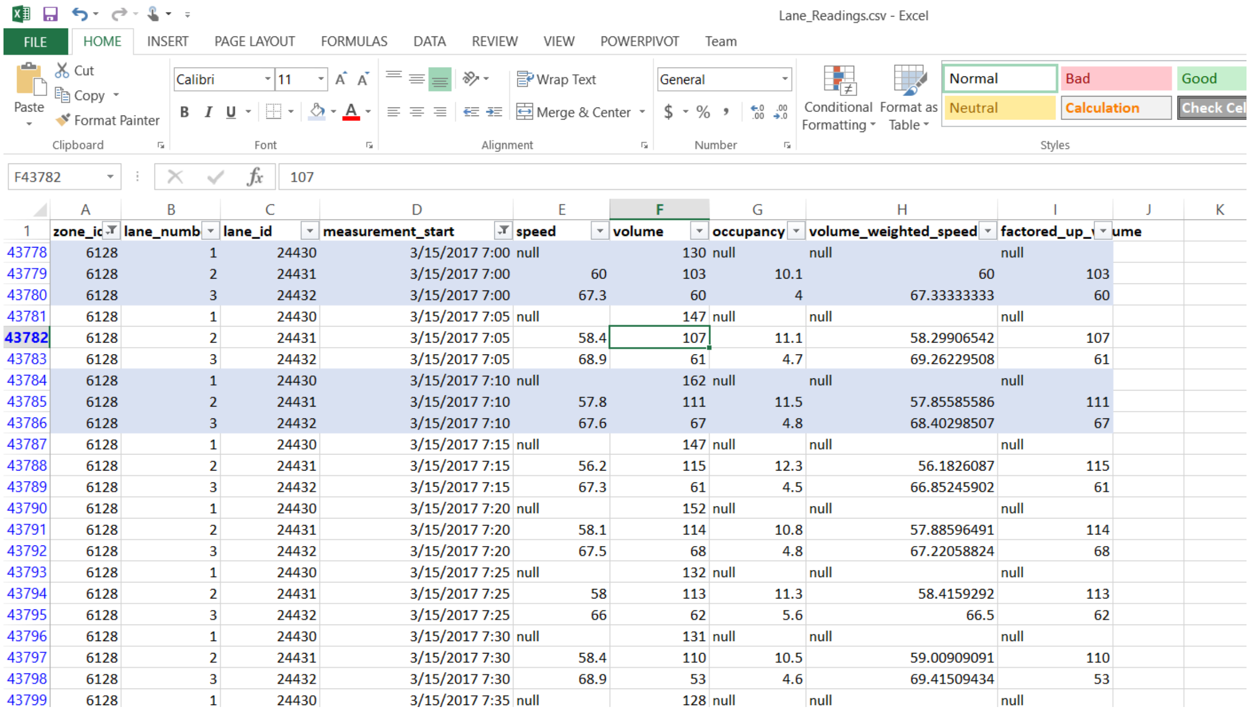
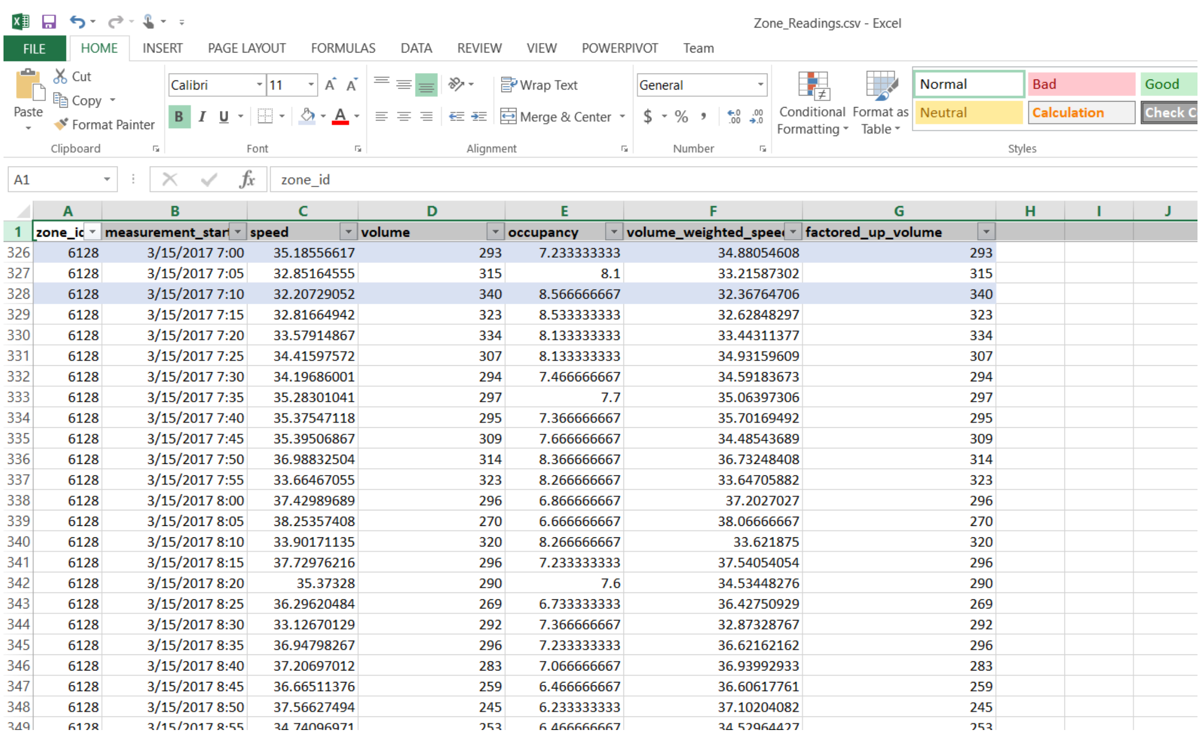


Figure 5-19 Detector Speed Distributions along the I-4 EB Mainline



	zone_ic	lane_numb	lane_id	measurement_start	speed	volume	occupancy	volume_weighted_speed	factored_up_ ume
43778	6128	1	24430	3/15/2017 7:00	null	130	null	null	null
43779	6128	2	24431	3/15/2017 7:00	60	103	10.1	60	103
43780	6128	3	24432	3/15/2017 7:00	67.3	60	4	67.33333333	60
43781	6128	1	24430	3/15/2017 7:05	null	147	null	null	null
43782	6128	2	24431	3/15/2017 7:05	58.4	107	11.1	58.29906542	107
43783	6128	3	24432	3/15/2017 7:05	68.9	61	4.7	69.26229508	61
43784	6128	1	24430	3/15/2017 7:10	null	162	null	null	null
43785	6128	2	24431	3/15/2017 7:10	57.8	111	11.5	57.85585586	111
43786	6128	3	24432	3/15/2017 7:10	67.6	67	4.8	68.40298507	67
43787	6128	1	24430	3/15/2017 7:15	null	147	null	null	null
43788	6128	2	24431	3/15/2017 7:15	56.2	115	12.3	56.1826087	115
43789	6128	3	24432	3/15/2017 7:15	67.3	61	4.5	66.85245902	61
43790	6128	1	24430	3/15/2017 7:20	null	152	null	null	null
43791	6128	2	24431	3/15/2017 7:20	58.1	114	10.8	57.88596491	114
43792	6128	3	24432	3/15/2017 7:20	67.5	68	4.8	67.22058824	68
43793	6128	1	24430	3/15/2017 7:25	null	132	null	null	null
43794	6128	2	24431	3/15/2017 7:25	58	113	11.3	58.4159292	113
43795	6128	3	24432	3/15/2017 7:25	66	62	5.6	66.5	62
43796	6128	1	24430	3/15/2017 7:30	null	131	null	null	null
43797	6128	2	24431	3/15/2017 7:30	58.4	110	10.5	59.00909091	110
43798	6128	3	24432	3/15/2017 7:30	68.9	53	4.6	69.41509434	53
43799	6128	1	24430	3/15/2017 7:35	null	128	null	null	null

(a) Lane Readings



	zone_ic	measurement_star	speed	volume	occupancy	volume_weighted_spee	factored_up_volume
326	6128	3/15/2017 7:00	35.18556617	293	7.233333333	34.88054608	293
327	6128	3/15/2017 7:05	32.85164555	315	8.1	33.21587302	315
328	6128	3/15/2017 7:10	32.20729052	340	8.566666667	32.36764706	340
329	6128	3/15/2017 7:15	32.81664942	323	8.533333333	32.62848297	323
330	6128	3/15/2017 7:20	33.57914867	334	8.133333333	33.44311377	334
331	6128	3/15/2017 7:25	34.41597572	307	8.133333333	34.93159609	307
332	6128	3/15/2017 7:30	34.19686001	294	7.466666667	34.59183673	294
333	6128	3/15/2017 7:35	35.28301041	297	7.7	35.06397306	297
334	6128	3/15/2017 7:40	35.37547118	295	7.366666667	35.70169492	295
335	6128	3/15/2017 7:45	35.39506867	309	7.666666667	34.48543689	309
336	6128	3/15/2017 7:50	36.98832504	314	8.366666667	36.73248408	314
337	6128	3/15/2017 7:55	33.66467055	323	8.266666667	33.64705882	323
338	6128	3/15/2017 8:00	37.42989689	296	6.866666667	37.2027027	296
339	6128	3/15/2017 8:05	38.25357408	270	6.666666667	38.06666667	270
340	6128	3/15/2017 8:10	33.90171135	320	8.266666667	33.621875	320
341	6128	3/15/2017 8:15	37.72976216	296	7.233333333	37.54054054	296
342	6128	3/15/2017 8:20	35.37328	290	7.6	34.53448276	290
343	6128	3/15/2017 8:25	36.29620484	269	6.733333333	36.42750929	269
344	6128	3/15/2017 8:30	33.12670129	292	7.366666667	32.87328767	292
345	6128	3/15/2017 8:35	36.94798267	296	7.233333333	36.62162162	296
346	6128	3/15/2017 8:40	37.20697012	283	7.066666667	36.93992933	283
347	6128	3/15/2017 8:45	36.66511376	259	6.466666667	36.60617761	259
348	6128	3/15/2017 8:50	37.56627494	245	6.233333333	37.10204082	245
349	6128	3/15/2017 8:55	34.74096971	253	6.466666667	34.52964427	253

(b) Zone Readings

Figure 5-20 Detector Data Reported from RITIS at Detector 6128

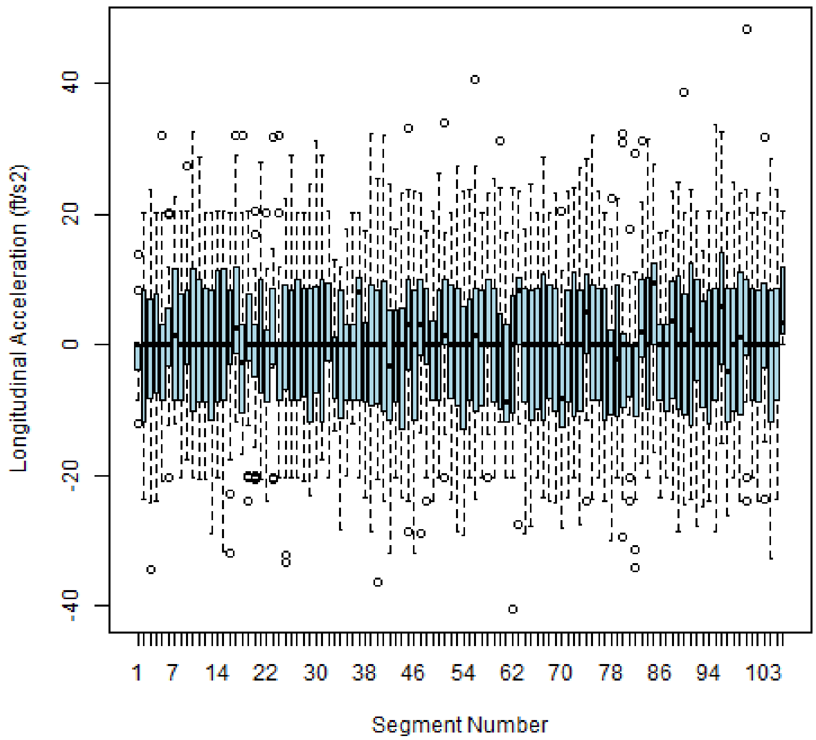


Figure 5-21 Distribution of BSM Longitudinal Acceleration in Each 300-ft Segment along the I-4 EB Mainline

In addition to speed and acceleration; a number of CAN data items are also available in the collected data including ambient air pressure, air temperature, engine revolutions per minute (RPM), brake applied status, coolant temperature, mass air flow rate, and fuel rate. However, it was found that a null value was reported for most of measurements and only a very limited number of non-zero values were archived. As an example of additional measures for which sufficient data are available, Figure 5-22 shows the distribution of engine RPM and Figure 5-23 presents the results for fuel rate. It is seen from Figure 5-22 that the engine RPM values are higher when Road Ranger vehicle travels at high speeds at the uncongested study segments downstream of Segment 47, as expected. There is a relationship between vehicle speed and the RPM for a specific vehicle. The speed can be calculated based on RPM if the Transmission Gear Ratio, Differential Gear Ratio, and Loaded Tire Radius are known. The RPM value can be used in combination with speed to estimate the number of stops. There is also a relationship between the Mass Air Flow and the Fuel Consumption and between the RPM and fuel consumption. There are relationships between these variables and emission rate. The boxplot in Figure 5-23 reveals that the fuel rate has a larger variation at the congested upstream segments.

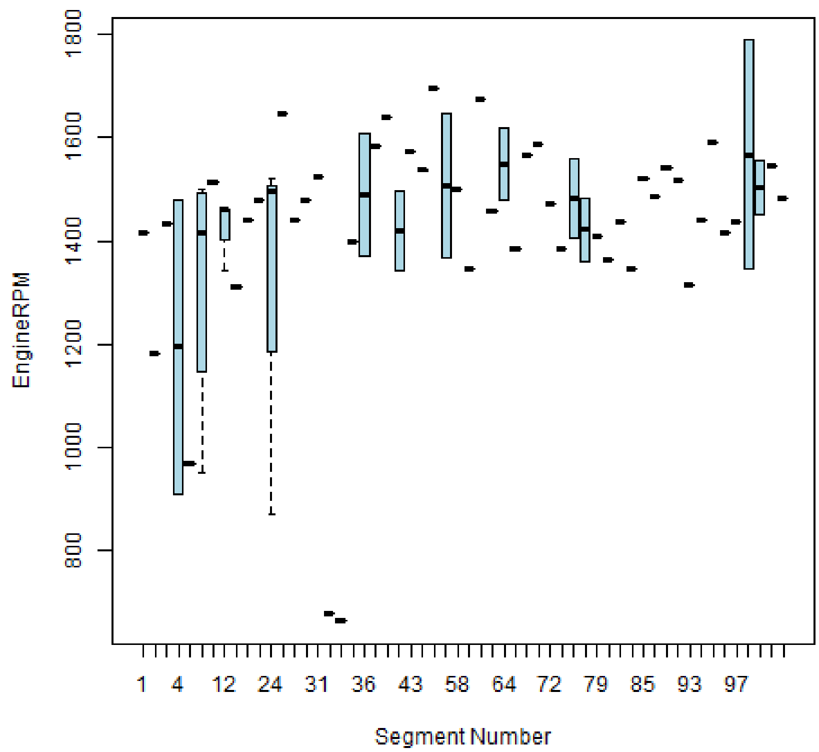


Figure 5-22 Distribution of Engine RPM in Each 300-ft Segment along the I-4 EB Mainline

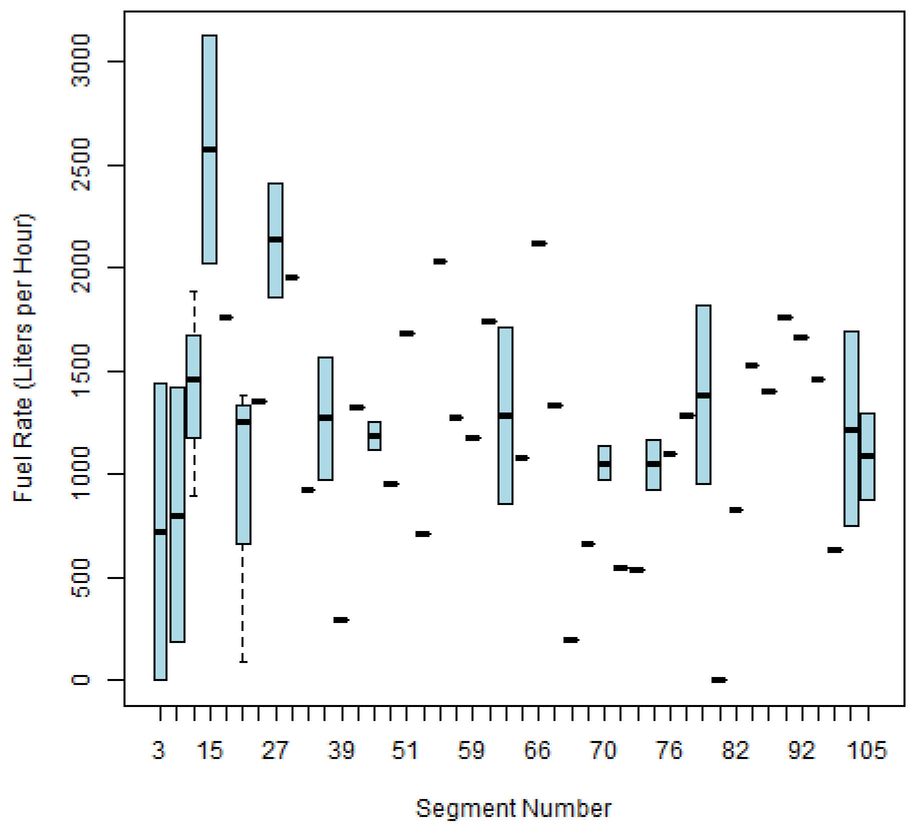
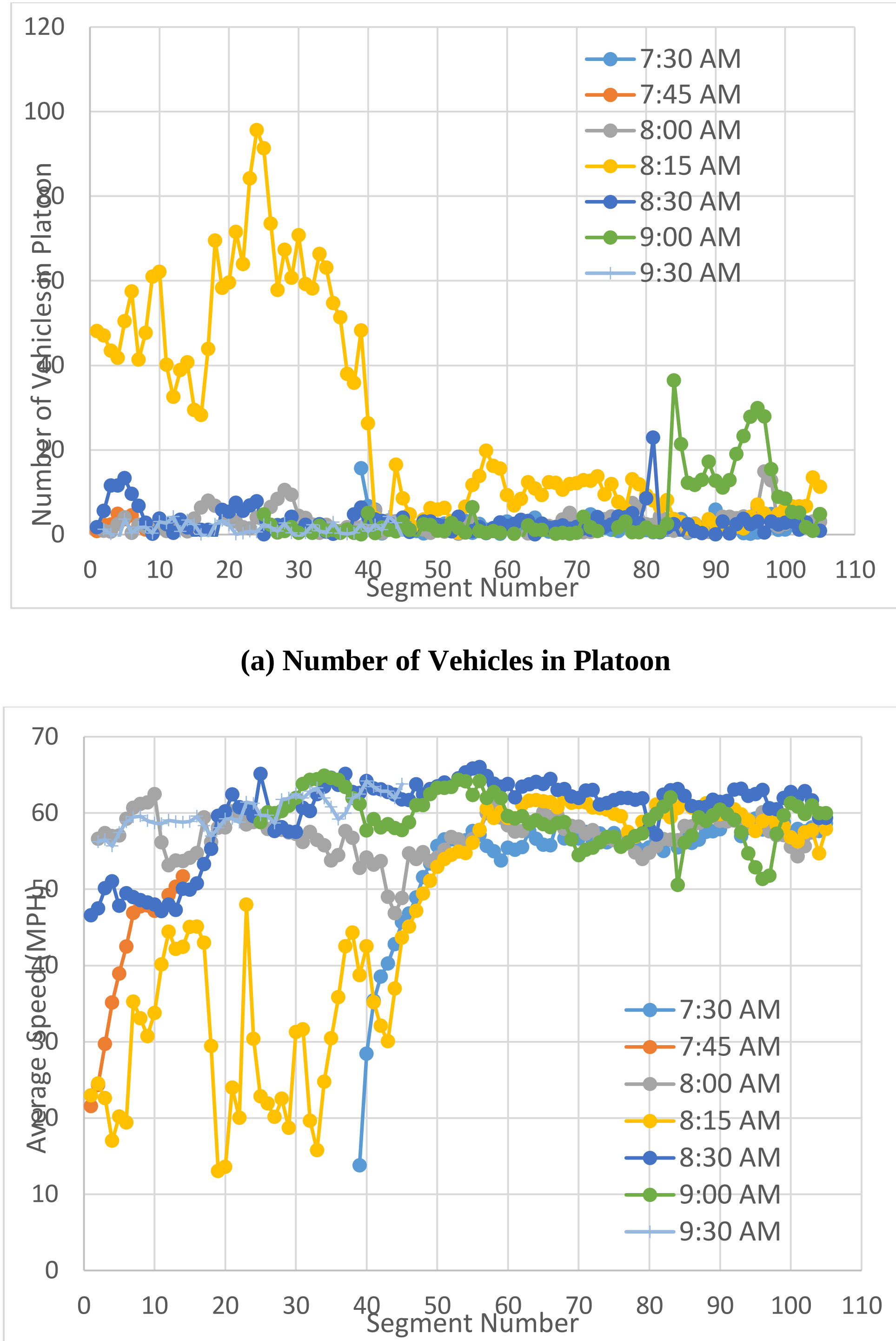


Figure 5-23 Distribution of Fuel Rate in Each 300-ft Segment along the I-4 EB Mainline

Additional performance measures can also be derived from the BSM data other than those directly output of the CAN measurements, for example, platoon size that allows the determination of the stability and congestion level of the traffic stream. In this study, the method proposed by the authors of this study in another project (Azizi et al. 2017) was applied to estimate the number of vehicles in platoon in the sections that experienced breakdown conditions. Equation 5-1 shows the utilized expression.

$$\text{NVP} = -94.29(\ln(10.56 - 0.3048SD_v) - \ln(10.4)) \tag{5-1}$$

Where SD_V is the standard deviation of the speed of each vehicle in ft/s and NPV is the number of vehicles in the platoon. Figure 5-24 shows the results of the estimated number of vehicles in the platoon along the study segments after breakdown. It is seen from this figure that the number of vehicles in platoon is relatively small around 7:45 am. The platoon size reaches the maximum value at 8:15 to 8:30 am and after that the platoon size is reduced.



(a) Number of Vehicles in Platoon

(b) Average Speed

Figure 5-24 Number of Vehicles in Platoon and Average Speed along the I-4 EB Mainline

On-Ramp Analysis Results

Unlike freeway mainline segments that are normally fully equipped with detectors in Florida, detectors are normally not equipped with permanent detectors, unless they are metered. Connected vehicle data will provide an alternative source for ramp data. Even though there are three on-ramps located within the study corridor, due to the specific route that road ranger vehicles follow, CV data was available only for the on-ramp from Osceola Polk Line Rd. Figure 5-25 shows the speed distribution on each 300-ft segment, starting from the Osceola Polk Line Rd EB on-ramp entrance (upstream) to the merging point with freeway mainline (downstream), based on the collected BSM data. The results in this figure show that the vehicle speed gradually increases from the on-ramp entrance to the freeway mainline until Segment 6 (1,800 ft from the upstream) indicating that this part of the ramp is not congested. However, starting from Segment 7 until the end of the ramp at Segment 9, there is some drop in the AM peak average speed and more importantly high variations between the minimum and maximum speeds on the ramp in the AM peak indicating periods of congestion on the last 900 ft on the on-ramp due to the congestion at the merging point. It can also be seen that the speed variation is relatively small at the entrance of the on-ramp but the variation increases at the end of the on-ramp because of merging with mainline vehicles.

.

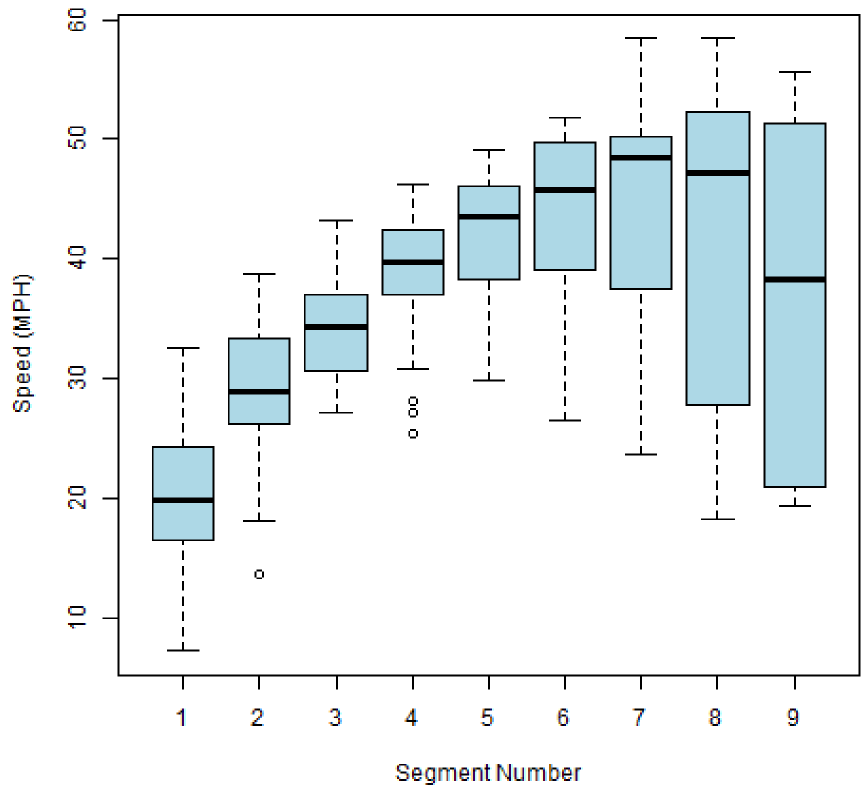


Figure 5-25 Speed Distribution along the Onramp at Osceola Polk Line Rd

The corresponding longitudinal acceleration distribution is shown in Figure 5-26. Note that these are instantaneous individual vehicle accelerations. Depending on the interaction with the vehicles in front, the vehicle may accelerate or decelerate at the on-ramp, resulting in more unexpected fluctuations shown in Figure 5-26. The results in this figure show that the longitudinal acceleration rate is between about -25 ft/s^2 and 20 ft/s^2 with a median value of zero for some of onramp segments. Still, as with the acceleration data on the mainline, the results do not give a meaningful trend. This will be investigated in future studies.

Figure 5-27 presents the distribution of the RPM along the on-ramp segments. As shown in this figure, the RPM is lower when travel speed is low at the on-ramp entrance compared to other ramp segments.

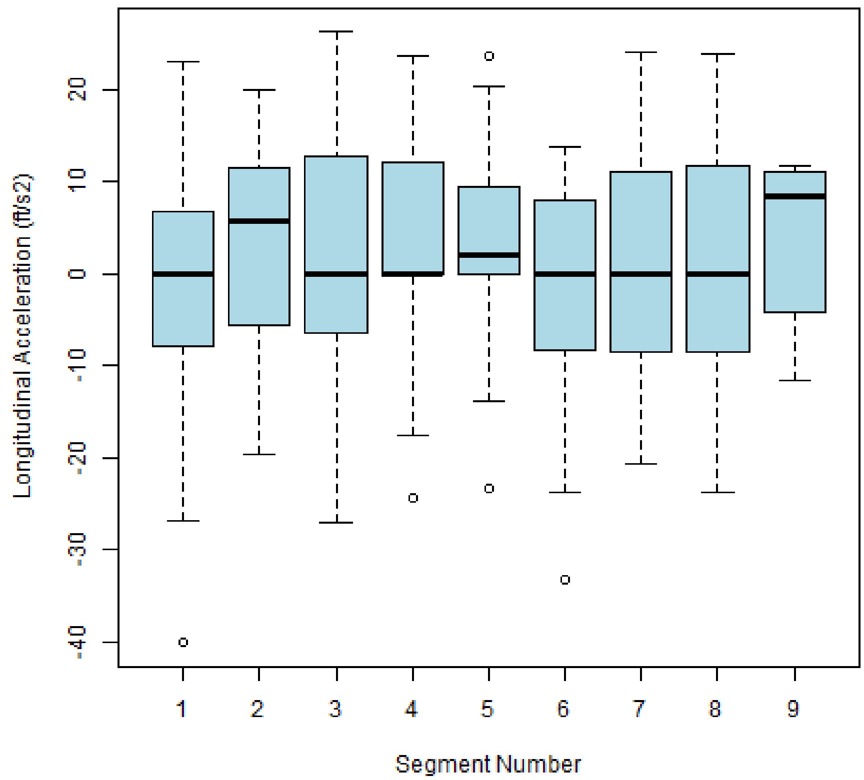


Figure 5-26 Longitudinal Acceleration Distribution along the Onramp at Osceola Polk Line Rd

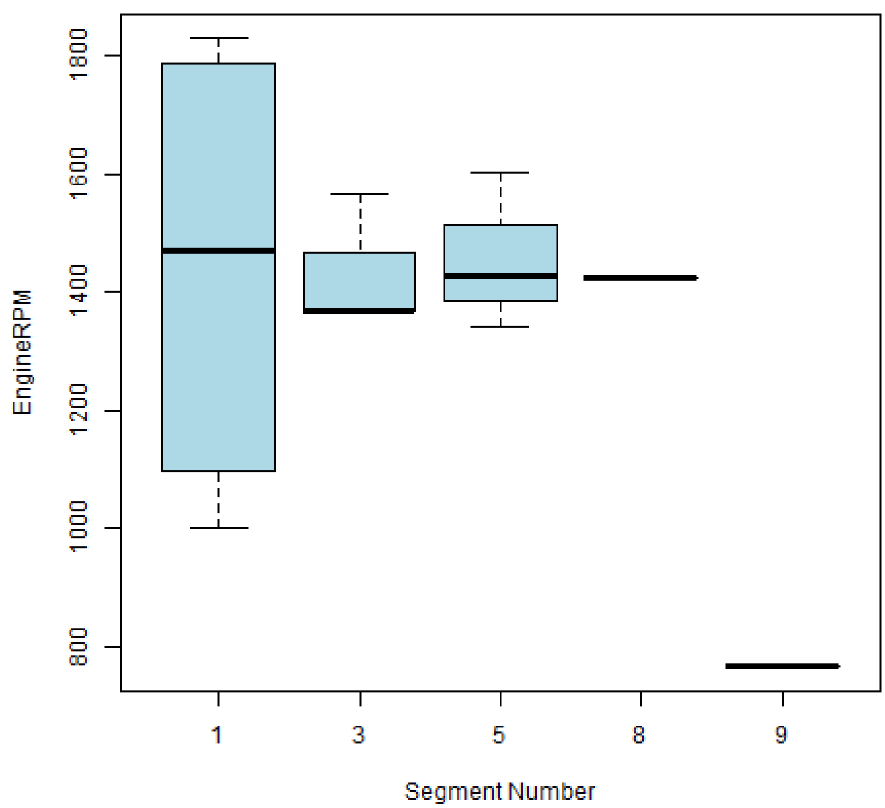


Figure 5-27 Engine RPM Distribution along the Onramp at Osceola Polk Line Rd

6. ASSESSMENT OF LINK LEVEL VARIATION OF CONNECTED VEHICLE MARKET PENETRATION

Connected vehicle (CV) technologies promise significant safety, mobility, and environmental benefits. However, the benefits of these technologies largely depend on their market penetrations (MP) in the coming years. With regard to CV, the United States Department of Transportation's (USDOT) National Highway Traffic Safety Administration (NHTSA) released an advance notice of proposed rulemaking (ANPRM) and a supporting comprehensive research report on vehicle-to-vehicle (V2V) communication technology in 2014 (National Transportation Safety Board, 2016). The draft Federal Highway Administration (FHWA) V2I Deployment Guidance (FHWA, 2015b) encourages V2I deployments, but it states that the USDOT will not require public agencies to implement V2I technology or applications, and recommends that these implementations should be done based on agency assessments. General Motors (GM) announced that it will release vehicle-to-vehicle (V2V)-equipped Cadillac by 2017 (U.S. News & World Report, 2016). Other car manufacturers are expected to equip their vehicles soon afterward. Thus, it is obvious that in the coming few years there will an increase in the proportions of cars with connectivity features.

To assess the impacts of CV technologies, there is a need to estimate the MP of the vehicles in the coming years. The estimation of the CV market penetration, in the past, has provided an average estimate of the growth of CV for the nation. If the overall average market penetration of CV is at a given level, there is no guarantee that a certain link within a region will have that market penetration at any given period. This is because of the variations in the socioeconomic characteristics of the regions compared to other regions, between the zones within a region, and the trip making characteristics of travelers with different socioeconomic characteristics. In particular, the higher the income and new car ownership characteristics (vehicle age distribution) in a region or a zone, the higher is the ratio of new vehicles introduced into the traffic stream and thus the higher is the market penetration of connected vehicles. The market penetration of CV also varies between links as a result of the variation of inter-zonal socioeconomic characteristics. This variation could be determined using the assignment of trips to the network and analyzing the output. The objective of this study is to develop a method to find the variations of the market penetration of CV between regions, zones within the region, and network links in a region, considering an average market penetration for the region. Although the focus of this study is on CV, the method presented in this document is also applicable to the estimation of market penetration of different levels of vehicle automation.

6.1. Application of Vehicle-to-Infrastructure (V2I) Communications

Connected vehicle data and the provision of messages to vehicles will play a major role in supporting the planning, operation, and management of the transportation system. The connected vehicle data and disseminated information will need to be transmitted using standardized messages utilizing dedicated short-range communication (DSRC) or other communication technologies, such as cellular, Wi-Fi, and WiMAX. A connected vehicle will be equipped with an Onboard Unit (OBU), which consists of several components such as computer modules, display units, and a wireless communication module (either DSRC or cellular). The roadside infrastructure will be equipped with Roadside Units (RSU) that communicate with the OBUs and a central location when utilizing the DSRC option for communication. The connected vehicle (CV) message types and components are specified in the Society of Automotive Engineers (SAE) J2735 standards (SAE, 2009).

The transmitted CV data could be used for a large number of applications (U.S. Department of Transportation, 2016) including safety, dynamic mobility, road weather management, and environment applications. The performances of these applications largely depend on the MP of CV. The estimated MP of CV is an important parameter for analyzing the impacts of connected vehicles on safety, mobility, and the environment (Olia et al., 2014). Previous studies have utilized the MP to assess the impacts of CV applications on mode choice (Minelli, 2015), traffic signal control (Smith et al., 2010; He, et al., 2012; Priemer and Friedrich, 2009), Freeway incident detection (Barria and Thajchayapong, 2011), lane-level speed estimation (Rim et al., 2011), arterial performance measurement (Li et al., 2008; Argote et al., 2011; Argote et al., 2012), intersection analysis (Ban et al., 2009b; Ban et al., 2011; Hao et al., 2012), vehicle position detection (Goodall et al., 2016), transportation operation (Smith et al., 2007), and arterial queue spillback (Christofa et al., 2013). However, these studies have not considered the variations of the market penetration of CV between regions or in a region, as stated earlier.

6.2. Methodology

The process of determining the MP distribution consists of three parts. The first part is to assume a scenario for CV implementation. The second part is to determine the MP of CV in different zones in a region depending on the socioeconomic characteristics of these zones. The third part is to determine the variations of MP on different links in the region utilizing the traffic assignment procedures incorporated in the regional demand forecasting models.

6.2.1. Determination of Zone-Specific MP

The main objective of this study is to determine the variations of CV on different links in a region. These variations will occur due to the variations of the percentage of CV between zones in the region reflecting the associated socioeconomic characteristics of the trip makers from/to these zones. Miller et al. (2002) showed that the vehicle age distribution is related to the per capita income in a county. They used the data from the state of Tennessee to illustrate the relationship. Country-by-country vehicle registration data and per capita income were used in their analysis. The per capita personal income information was collected from the United States Department of Commerce, Bureau of Economic Analysis (BEA) website. The vehicle age distributions for different income categories were developed for two vehicle types: light-duty vehicles (LDVs) referencing passenger cars and light-duty trucks (LDTs). The obtained age distributions of LDVs are shown in Figure 6-1 for counties with different income levels. The horizontal axis is the vehicle age and the vertical axis is the fraction of vehicles out of the total vehicles that have a certain age. The vehicles that have an age of thirty years or more are placed in the 30 years' age group.

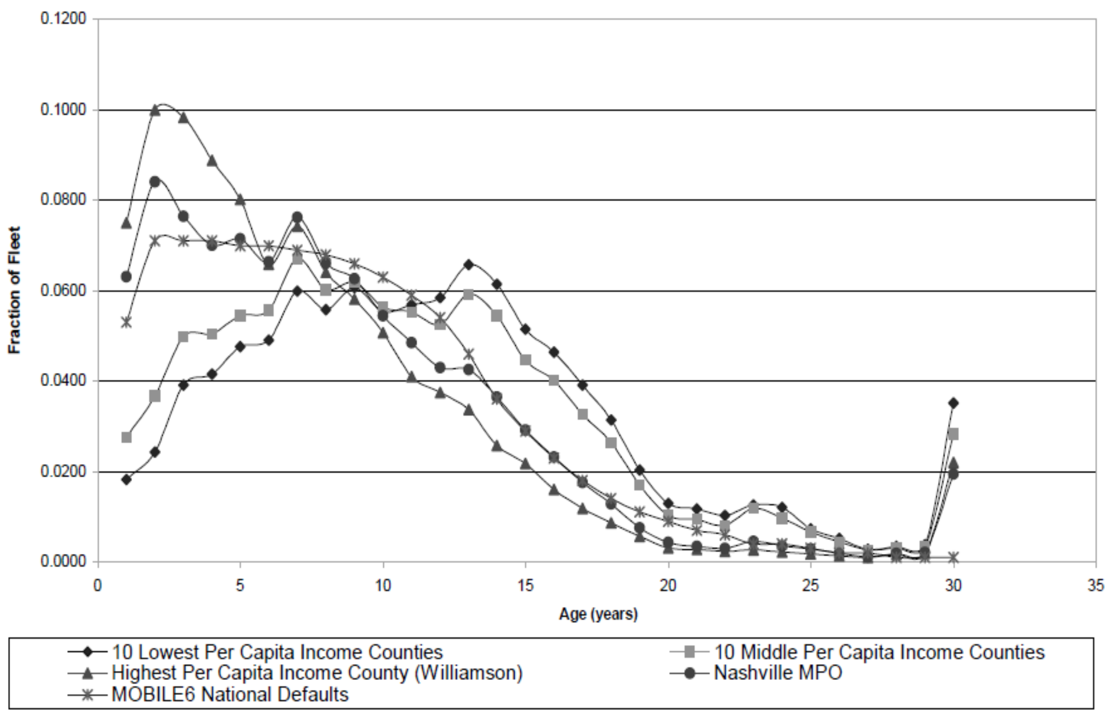


Figure 6-1 LDV Age Distribution for Tennessee Counties (Source: Miller et al., 2002)

Figure 6-2 shows that the fraction of one-year-old vehicles varies from 1.8% to 7.5% depending on the per capita income of an area. This means that if the connected vehicles are mandated for all new vehicles, then at the end of the first year the MP of CV will vary from 1.8% to 7.5% in a given area with an average value of 4.65%. This percentage will cumulatively increase each year as new vehicles are introduced in the market. This study uses the results from Miller et al. (2002) to produce cumulative percentage distributions of CV for regions or zones with the highest income area and lowest income area based on the results of Figure 6-2. Figure 6-3 shows the resulting distributions. Note that the results presented in this study are based on data from Tennessee due to the availability of this data to the researchers. However, the methodology is applicable to similar data, if available from other states and regions.

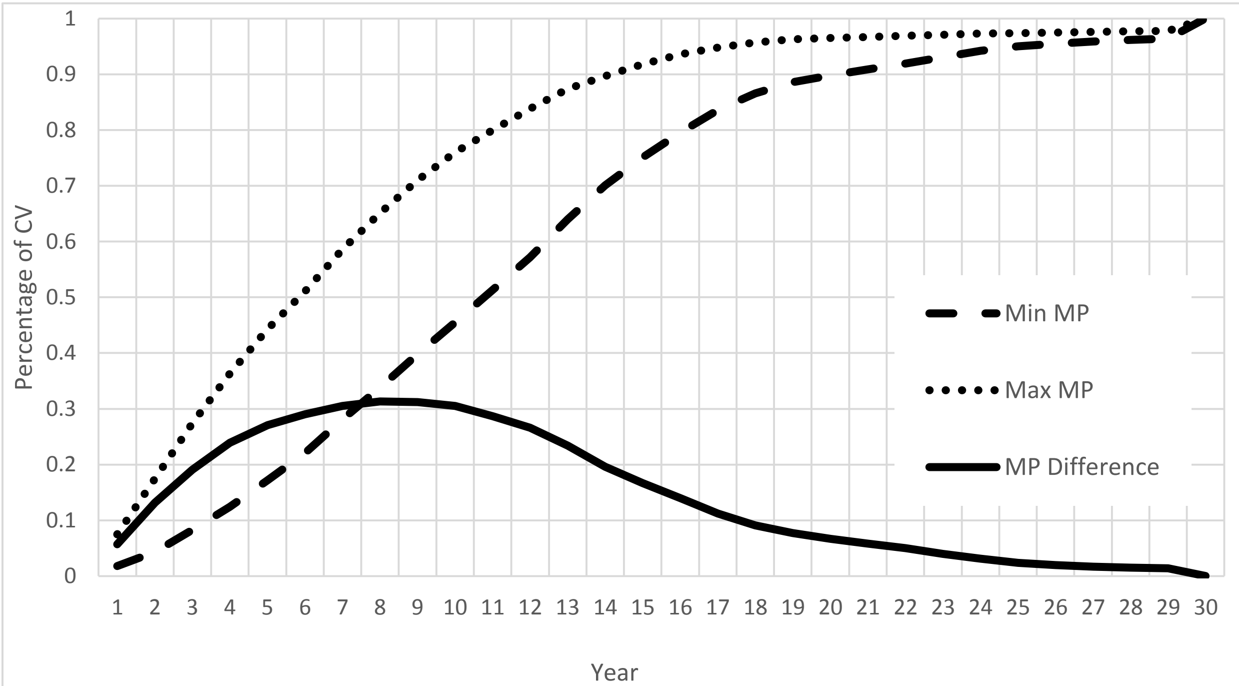


Figure 6-2 Variation of the CV Market Penetration in Different Areas Based on the Information Presented in Figure 6-1

Figure 6-2 shows the cumulative increase of CV each year for both the highest income area (Max MP) and the lowest income area (Min MP). The trend line shows that the increase in the CV market penetration is higher in the early stages and it slows down significantly after 12 to 18 years due to the market getting closer to saturation. The solid line in Figure 6-2 shows the difference between the maximum and minimum MP. An important observation from this graph is that the variability between areas increases for the first few years and reaches the maximum point around year 8 (3.1%). After that, the variability decreases and eventually becomes very low, as expected.

6.2.2. CV Implementation Scenario

As stated in the previous section, NHTSA is expected to mandate connected vehicle technologies on all new vehicles. Apart from this mandate, after-market plug-in equipment will be available for installation on older cars but this is not expected to be mandated. However, it is not certain how many people will buy the after-market devices. Thus, the connectivity of the new cars will play the vital role in the determination of the market penetration of the CV and the after-market installations are not considered in this study to be on the conservative side in estimating the market penetrations. The USDOT (2008) conducted a study to estimate the benefits and costs of CV implementations. For that purpose, the study predicted the probable market penetrations of CV in future years. In the estimation, the analysts considered the scenario where only new vehicles will use the CV technology with the assumption that in the first year 25%, second year 50%, third year 75% and afterward 100% of the new vehicles will have the connectivity. Wright et al. (2014) suggested three different scenarios for probable CV implementations. The most conservative scenario among the three is called the “15-year organic” scenario, which assumes that the CV will come into the fleet as organic sales of the new capability. The moderate one is called the “5-year mandate” scenario, in which manufacture would include OBUs into the new vehicles over a five-year period. The best-case scenario is the “1-year mandate” scenario where all the new vehicles will be equipped with OBU starting from the year that the CV is mandated. In this research, the “1-year mandate” ‘scenario is assumed when producing the results. However, the methodology of this study could be applied to any of the above and other scenarios of CV implementations.

6.2.3. Determination of the Variation of MP between Links

This study used the cumulative MP of CV distribution described above to determine the MP of CV for each Zone in a region based on the socioeconomic characteristics in the region. The determined MP was associated with each zone in a demand-forecasting model. Then, the assignment procedure of the model was exercised and the percentage of CV on each link is determined for each time period of the day based on the total volumes and CV volumes resulting from the assignment.

To demonstrate the application of the methodology to estimate the CV MP on the links, this study utilized the assignment step of four step demand-forecasting model in the southeast Florida region, which is referred to as the Southeast Florida Regional Planning Model (SERPM6). This model covers Miami-Dade, Broward, and Palm Beach in southeast Florida. Unfortunately, the SERPM6 model does not include zone specific per capita income data. This information is available in other regional demand forecasting models but not in SERPM. Thus, the first step in this process is to identify the Southeast Florida income data per zone.

The income data is collected in this study using the American Community Survey (ACS) 5-year estimates (2010-2014). This data is available for downloaded from the ACS website (United States Census Bureau, 2016) as a Geographic Information System (GIS) database shape file format. The income per capita in each census tract level is used in the analysis for this study. The income data downloaded at the census tract level is associated with the zone data using the ArcGIS software. There are a total 4,106 zones within the SERPM6 model. The minimum per capita income in the ACS database for southeast Florida is $15,500 and the maximum per capita income is $1,44,900. Using the income of each zone and the variation shown in Figure 6-2, the MP of a zone is calculated.

Once the MP for each Zone is identified, as described above, each O-D matrix in the demand model is divided into two matrices. The first O-D matrix is for the connected vehicles and the second one is for the non-connected vehicles. The total number of trips originating from each zone is multiplied by the MP of the CV associated with that zone to get the number of trips that are made by the CV. The remaining trips are considered as non-equipped vehicle trips. The two types of O-D matrices are then used as inputs to the trip assignment process. After the trip assignment, the link-level traffic volumes are analyzed and the percentage of CV for each link is calculated.

The study then uses the link level CV proportions in the region as determined above to identify the statistical distributions of these proportions. Analysts can use this type of distributions in lieu of using fixed CV proportions when assessing the performance of applications that are based on CV technologies.

6.3. Case Study Results

For the Southeast Florida Case Study, there are about 45 thousand links, for each of which the CV percentages are identified based on the analysis of this study. As mentioned above, the next step is to determine the distribution of these proportions. First, this study investigated whether the distribution of the CV market penetrations on the links is normal or not. For a large sample, the most common practice of the normality confirmation is to check the normality test plots, as shown in Figure 6-3. After analyzing the data, it was found that the percentage of CV on different links follows a lognormal distribution rather than a normal distribution. The density plot of this variable is skewed to the right and it was found that the logarithm of the link CV percentage fits better a normal distribution than does the original data. Figure 6-4 shows the plots to check the normality of the logarithmic value of the CV percentages for different links in year 1. Please, note that the results presented in this section is for the PM peak period. The analysis can be repeated for different peak periods to show the variations in MP by a period of the day. This will account for the difference in the distributions of the trips by time of day of users with different income levels.

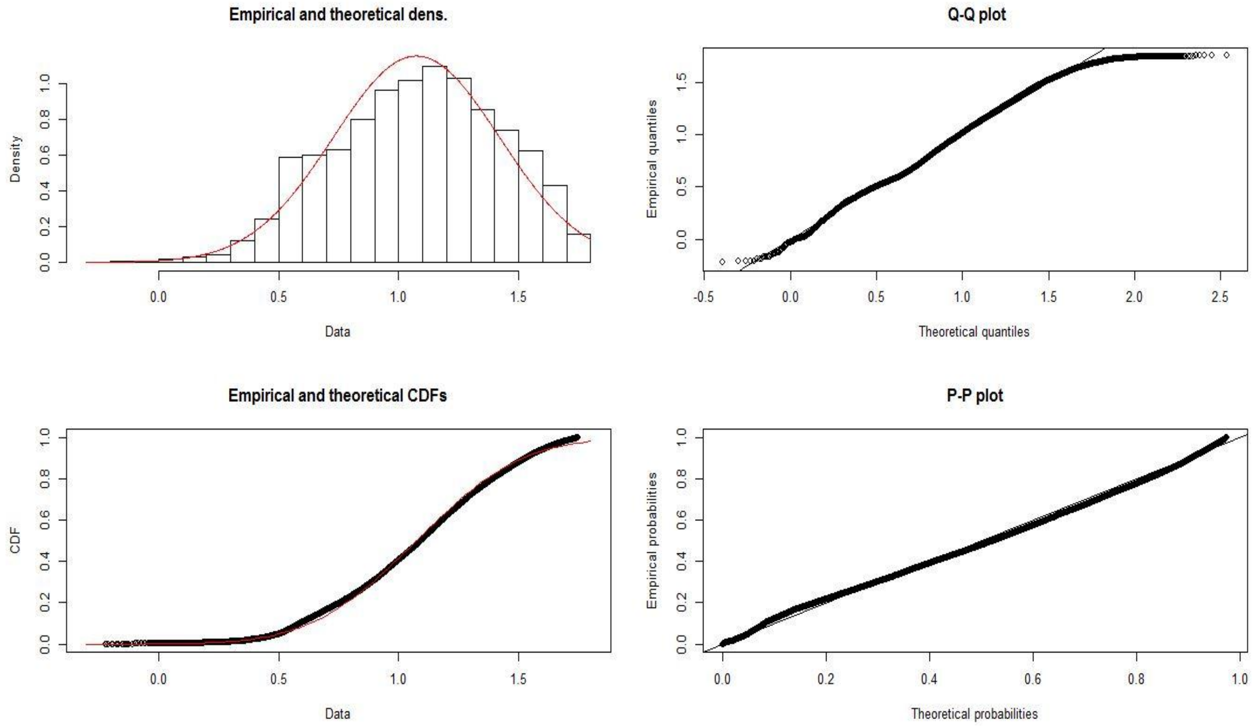


Figure 6-3 Different Plots to Check the Distribution of the CV Percentages for Various Links [Year 1]

The plots show that the logarithm of the CV percentages is normally distributed. This same test was repeated for all future years. For all years, the distributions were found to be lognormal. Table 6-1 shows the mean and standard deviation of each year distribution.

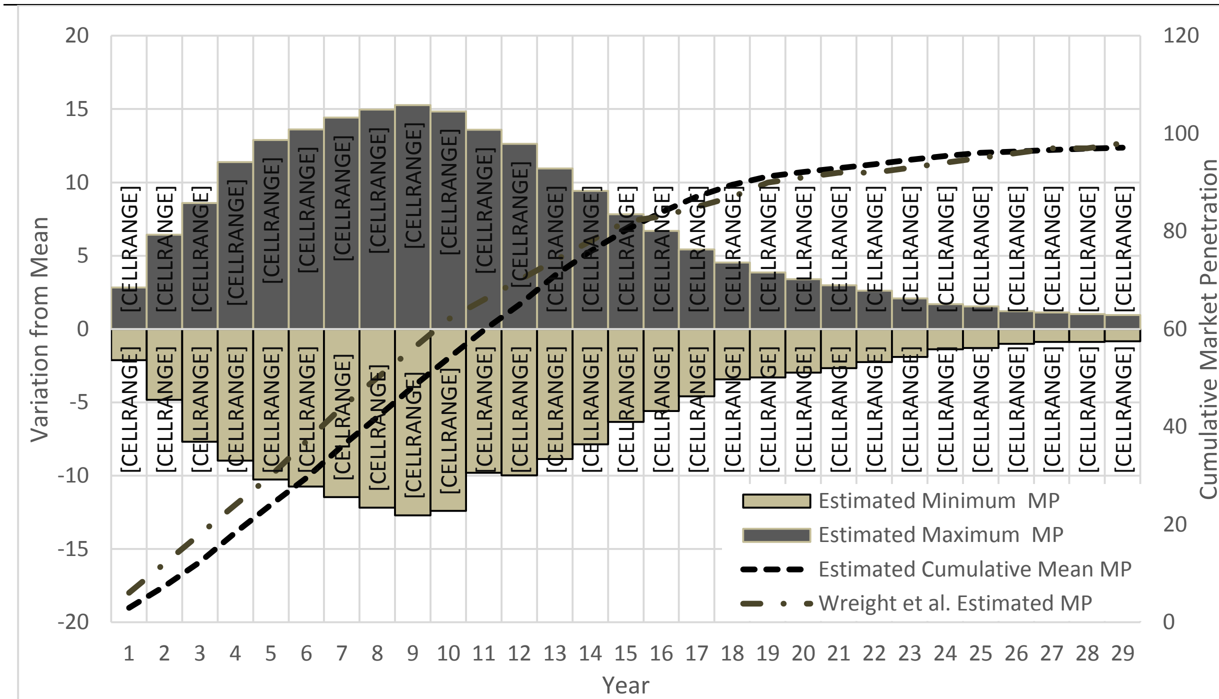
After analyzing the data for all the years following the methodology presented in the previous section, the mean market penetration and the variation level at each year are presented in Figure 6-4. In Figure 6-4, the left vertical axis shows the variation of the MP between links for each year and the right vertical axis shows the cumulative average MP of CV. The number inside each bar chart provides the maximum and minimum value of MP among the links in a region at a particular year of the analysis. The two dotted line represents the cumulative market penetration calculated in this study and the cumulative MP calculated by Wright et al. (2014).

Table 6-1 Mean and Standard Deviation (SD) of Link-level MP Distribution by Year

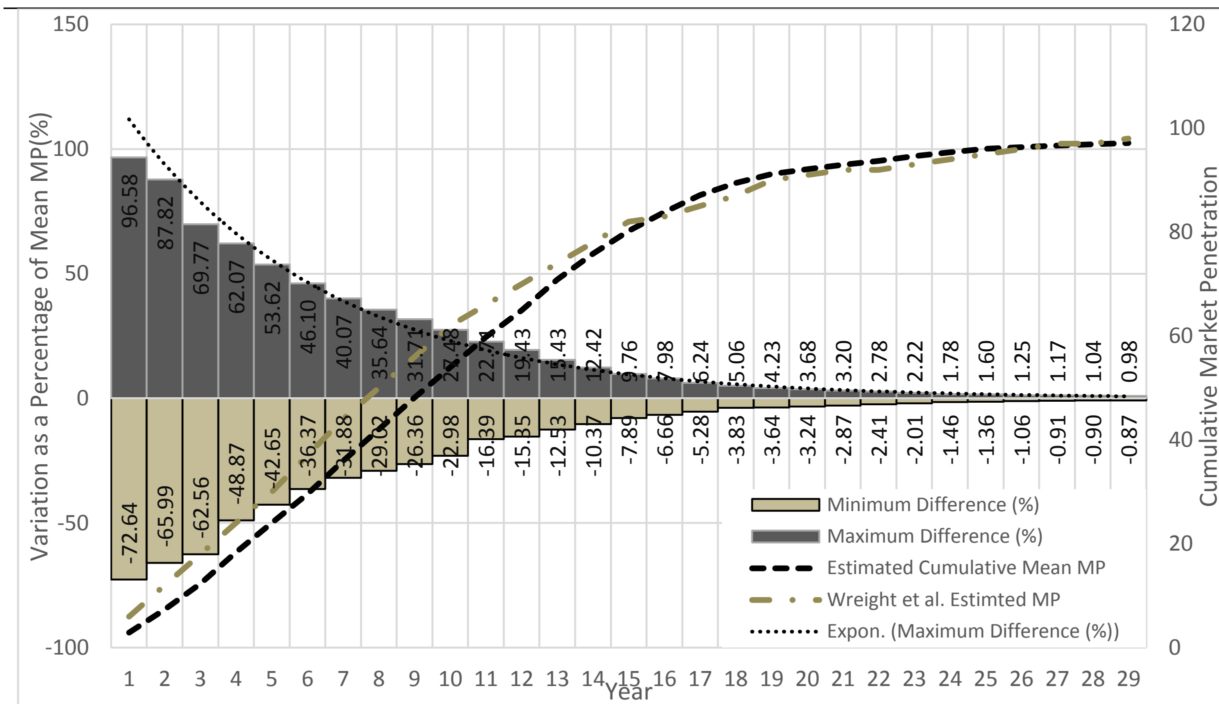
Year	Mean (log)	Mean (Actual)	SD (log)	Year	Mean (log)	Mean (Actual)	SD (log)
1	1.074	2.9	0.5460	16	4.429	83.8	0.0310
2	1.990	7.3	0.3230	17	4.467	87.1	0.0240
3	2.510	12.3	0.3000	18	4.494	89.5	0.0195
4	2.910	18.4	0.2300	19	4.513	91.2	0.0170
5	3.180	24.0	0.1960	20	4.523	92.1	0.0146
6	3.386	29.5	0.1720	21	4.532	92.9	0.0130
7	3.583	36.0	0.1480	22	4.540	93.7	0.0110
8	3.737	42.0	0.1330	23	4.550	94.6	0.0088
9	3.875	48.2	0.1190	24	4.558	95.4	0.0068
10	3.989	54.0	0.1060	25	4.565	96.1	0.0061
11	4.091	59.8	0.0860	26	4.568	96.4	0.0049
12	4.174	65.0	0.0740	27	4.571	96.6	0.0045
13	4.261	70.9	0.0600	28	4.574	96.9	0.0042
14	4.329	75.9	0.0470	29	4.576	97.1	0.0040
15	4.385	80.2	0.0380				

Figure 6-4(a) shows the actual variation and Figure 6-4(b) shows the variation as a percentage of the mean value. For lower market penetrations, the variations are lower but the percentage variations are higher. An exponential function that is fit to the data as shown in Figure 6-4(b) shows that the percentage of the MP variation decreases exponentially.

The average percentage increase of CV for each year is presented in Figure 6-5. Figure 6-5 shows that the MP increase rate grows up for the first several years and then it remains almost constant for the next few years before decreasing at a steep slope and finally becoming flat at a low value, due to reaching the oversaturation level.



(a) MP variation values



(b) Variation as a percentage of the mean MP

Figure 6-4 Variation of CV Market Penetration

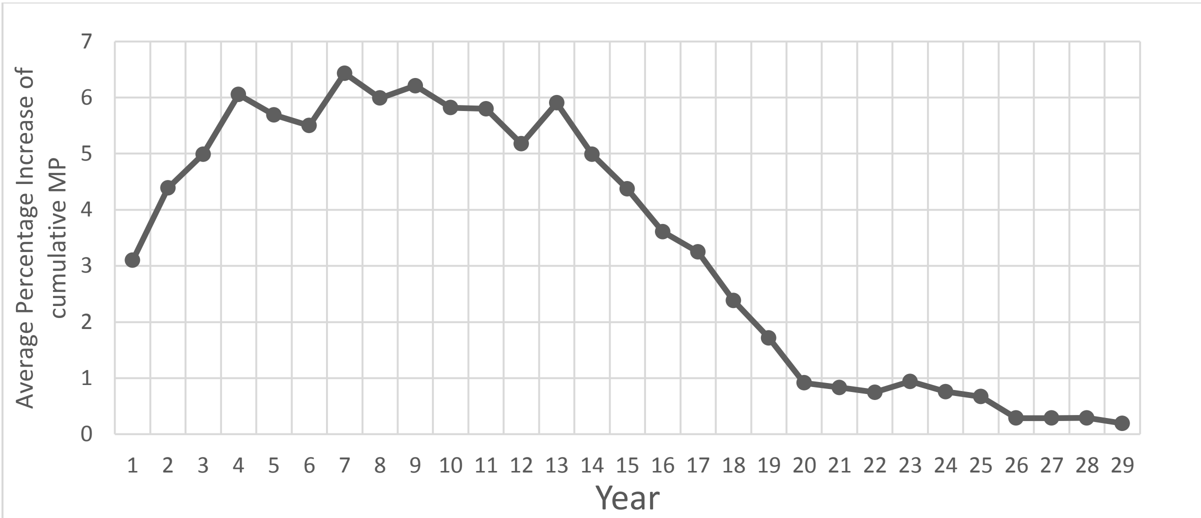
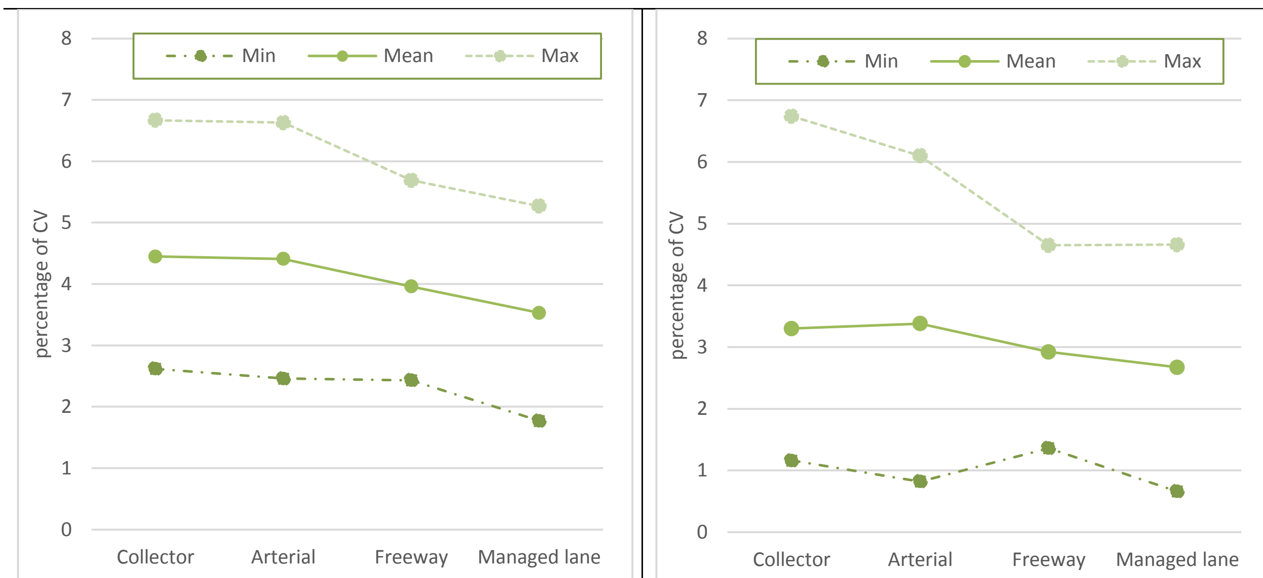


Figure 6-5 Average Percentage Increase of Cumulative MP of CV by Year

This study also investigated the difference of MP variations on different facility types (Figure 6-6). Figure 6-6 shows that the variability decreases when moving from collector to arterial and from arterial to freeway and managed lane facilities. This is due to the mix of traffic from various zones that normally use freeways and to less extent arterials. Thus, it is recommended that these variations are considered separately by facility type with a different distribution is identified for each type.



(a) Year 1	(b) Year 10

Figure 6-6 Variation of MP by Facility Type

6.4. Summary

This study proposed a methodology to determine the variation of CV market penetration between regions, zones within a certain region, and links within the region. The methodology can be implemented with various CV implementation scenario assumptions and considers the variations in the socioeconomic characteristics between regions and zones. The analysis can be repeated for different peak periods to show the variations in MP by period of the day. This will account for the difference in the distributions of the trips by time of day of users with different income levels. Applying the methodology of this study to a case study indicates that the distribution of the link-specific CV MP follows a lognormal distribution. The percentage variation in the market penetration is shown to be the highest in the first year of CV implementation and decreases exponentially with the number of years passing since the implementation. The MP variations between links are the highest on collectors followed by arterials followed by freeways. It is recommended that the variations in MP are considered separately by facility type. The study also shows that the average percentage increase in the CV MP grows up in the first several years then remain almost constant before dropping sharply.

This study concludes that analyzing the impacts of CV implementation utilize a fixed forecasted MP for the whole nation will not be able to reflect the real-world conditions that involve variations in socioeconomic characteristics between regions and zones. It is recommended that the methodology developed in this study to forecast the CV market penetrations and their variations between regions, zones, and links are used when evaluating CV impacts.

7. ACCURACY AND RELIABILITY OF ESTIMATED TRAVEL TIME USING BASIC SAFETY MESSAGE (BSM) DATA COLLECTED FROM CONNECTED VEHICLES

7.1. Introduction

Travel time is a critical performance measure of the transportation system. Other important parameters of the transportation system such as speed, travel time reliability measures, level of service (LOS), incident detection, and back of queue identification can also be derived based on this measure. In addition to its use in the management and operations of the transportation system, it is a critical component of the information provided by traveler information systems. The accuracy and reliability of the estimated travel time are key considerations when collecting the data required for its estimation and utilization for different purposes. Before selecting a technology for collecting the required data, it is important to know the quality of the travel times that can be estimated under different conditions and scenarios. Automatic data collection technologies have included point detector technologies (Kwon et al., 2000; Coifman, 2002; Sisiopiku et al., 1994; Coifman, 1998a; Bhaskar et al., 2011), Automatic Vehicle Identification (AVI) (Kanayama et al., 1991; Turner et al., 1998; Li et al., 2006; Sherali et al.; 2006; Tam and Lam, 2008; Kwong et al., 2009; Sun et al., 2003; Cheung et al., 2005; Ndoye et al., 2008; Coifman, 1998b; Oh et al., 2004; Coifman and Cassidy, 2002) also known as vehicle re-identification technologies, and vehicle tracking based on automatic vehicle location (AVL) technologies (Herrera et al., 2010; Bar-Gera, 2007; Ygnace et al., 2000). These technologies vary in their ability to support accurate and reliable travel times and also in their costs and constraints of the implementations. The expected implementations of connected vehicles in the next few years will provide a promising alternative to existing technologies in providing data for accurate and reliable travel time estimates.

The United States Department of Transportation's (USDOT) National Highway Traffic Safety Administration (NHTSA) released an advance notice of proposed rulemaking (ANPRM) and a supporting comprehensive research report on vehicle-to-vehicle (V2V) communication technology in 2014 (National Transportation Safety Board, 2016). NHTSA is expected to mandate connected vehicle (CV) technologies on all new vehicles. Vehicle to Vehicle (V2V) along with Vehicle to Infrastructure (V2I), facilitated by CV technology, will improve the traffic safety, mobility and environmental impacts of the transportation system. At the same time, the V2I communications will be an important source of traffic data for use by transportation system agencies. If this data is processed properly then the agencies will be able to better determine the traffic conditions allowing them to take proper countermeasures to address different conditions and for dissemination to travelers and other agencies. However, it is important to assess the data quality provided by this new technology compared to existing technology. This study utilizes a methodology to determine the travel time using CV data and examines the quality of the resulting travel time under different set-ups and scenarios.

7.2. Background

CV technologies are increasingly being considered by agencies for possible implementation in the next few years. The CV requires the use of wireless communication for V2V and V2I transmission of data. This can be based on the Dedicated Short Range communication (DSRC) and/or a cellular technology. A connected vehicle will also be equipped with an Onboard Unit (OBU), which consists of several components such as computer modules, display units, and a wireless communication module (either DSRC or cellular). The roadside infrastructure will be equipped with a Roadside unit (RSU) which communicates with the OBU if the DSRC option is utilized. CVs equipped with OBU can generate and transmit Probe Data Messages (PDMs), Basic Safety Messages (BSM), ITS Spot messages, and/or European Cooperative Awareness Messages (CAM). These different message sets have been standardized as documented in the Society of Automotive Engineers (SAE) J2735 standards (SAE, 2009).

The BSM contains vehicle safety-related information broadcasted by the vehicles to the surrounding vehicles, but can be also sent and/or captured by the infrastructure. The BSM, as defined in the J2735 standards, consists of two parts. Part 1 is sent in every BSM message broadcasted 10 times per second and is expected to be mandated to be broadcasted by the anticipated NHTSA ruling. It contains core data elements, including vehicle position, heading, speed, acceleration, steering wheel angle, and vehicle size. BSM Part 2 consists of a large set of optional elements. The BSM Part 1 elements are sufficient to estimate the average travel times, as done in this study.

There are few studies that start to appear in the literature regarding travel time estimation using connected vehicle data. Zou et al. (2010) estimated travel time based on CV and found an average error percentage 27.6%, 12.5%, and 8.2% for 1%, 5%, and 10% market penetrations, respectively. These estimates were based on traffic simulation of a hypothetical network that simulates vehicles broadcasting PDM data according to J2735 standards. Vasudevan and O’Hara (2015), in a workshop presentation about an unpublished work, summarized a method to estimate the travel time and back of queue location using BSM and PDM data. Their methodology does not use the vehicle ID and thus the trajectories of the vehicles. The BSM include temporary vehicle IDs but does not keep the same ID for a long period of time to protect privacy. Argote et al. (2012) estimated measures of effectiveness based on real-world vehicle trajectories. These measures of effectiveness include queue length, speed, number-of-stops, acceleration noise, and average delay per unit distance. They used the Next Generation Simulation (NGSIM) data for the testing purpose. A drawback of this study is that it uses the vehicle ID but does not consider the change of vehicle ID during its course of travel, as specified in the J2735 standards.

The above discussion indicates that there is no detailed study, which focuses on estimating travel time and assessing the accuracy and reliability under different CV market penetration scenarios in the coming years. The objective of this study is to assess the accuracy and reliability of travel time estimation based on BSM data collected utilizing DSRC technology under different market penetrations of CV, considering the randomness in CV presence on the links, vehicle ID change, and the variation in the market penetration between links in the same region due to the variation in the socioeconomic characteristics of the regions and the zones within the region. The quality of the collected data is assessed for both real-time uses based on one-day data, as well as for average off-line use based on a historical average of many days.

7.3. Utilized Data

This study utilized trajectory data collected under the Next Generation Simulation (NGSIM) program as a base and process the data to emulate BSM data. The NGSIM program was initiated by FHWA and collected high-quality traffic and vehicle trajectory data. The NGSIM data are available for the researcher to download from a website (Next Generation Simulation Community, 2015). The NGSIM data are for four different locations. Two of the locations are arterial roadway sections and two locations are freeway sections. In this research, one of the arterials (Peachtree Street, Atlanta, GA) and one of the freeways (U.S. Highway 101, Los Angeles, CA) are used for the purpose of this study.

The data was collected using video image processing by the NGSIM program. For the Peachtree Street, the processed data contains the trajectories of vehicles and an aggregated summary of traffic flow and speed, number of lane changes, headway and gap analysis, and input-output analysis of flows. The total length of the test section is approximately 2,100 feet with five intersections. A 1,500 feet segment in the northbound direction of the test section is used in the analysis of this study. Traffic data was collected for 15 minutes during the PM peak (4:00 PM to 4:15 PM) with five intersections (average intersection spacing of 500 ft) and an average speed of 18.5 mph in the northbound direction. The processed data for the U.S. Highway 101 freeway segment includes the trajectories of the vehicles and an aggregated summary of traffic flow and speed of the vehicles aggregated by time, distance (100 feet intervals), and lane. The total length of the test link is 2,100 feet. For the purpose of this project, a 1,500 feet section is used. The southbound direction data was collected for the AM peak (From 7:50 AM to 8:05 AM) with an average flow of 1,630 vphpl and an average speed of 28 mph.

7.4. Methodology

As mentioned earlier, the main objective of this study is to determine the accuracy and reliability of estimated travel time using BSM data. For this purpose, first, the trajectory data collected using the NGSIM program was processed to emulate the BSM data. Then the BSM data was further processed to estimate the segment travel times under different market penetration scenarios considering the randomness in CV identifications on the links, and the variation in the market penetration between links in the same region due to the variation in the socioeconomic characteristics of the zones in the region. The results were then compared with the ground truth travel time estimates, which are obtained based on all vehicle trajectories before converting to CV BSM data. The details of the process are described below.

7.4.1. Data Processing to Emulate BSM

The NGSIM data was processed to emulate BSM Type 1 data collected at 1/10th second. The study utilized the Trajectory Conversion Algorithm (TCA) [Version 2.3] (OSADP, 2015) developed by the Federal Highway Administration (FHWA). Vehicle trajectory data, either generated from a microscopic simulation software or collected from the real world can be used as an input to the TCA. The input trajectory file to the TCA should contain vehicle ID, time, speed, position coordinates, and acceleration data of the vehicles. The downloaded NGSIM trajectory files were processed in this study by filtering out the required variables and performing unit conversions. Using the trajectories as inputs, the TCA generates connected vehicle BSM Type 1 data following the SAE J2735 standards. The user of the TCA tool can select the market penetration, message type (e.g., BSM or PDM) and the communication type (DSRC or cellular). As stated earlier, in this study, only the BSM Type 1 data transmission with DSRC communications is considered. The user can also specify the Roadside Unit (RSU) locations. Since the roadway section in this study is shorter in length than the coverage area of a RSU, which is about 2,600 feet to 4,000 feet (Andrews and Cops, 2009; McGurrin, 2012), it is assumed that only one RSU is placed halfway of the section providing a full coverage of the section. The tool allows the specification of data transmission loss during the data transfer between the OBU and the RSU. It has been reported that the loss rate with DSRC communications varies between 10% and 20% with an average of 12% (Kandarpa et al., 2009). In this study, 12% transmission loss is specified as input to the TCA tool. It is expected that the actual loss is a function of the availability of line of site in the coverage area and this should be considered with RSU siting. The TCA tool changes the vehicle ID every five minutes with a 30 seconds buffer window. This means that the ID will change over a period that ranges from 5:00 to 5:30 minute, as is commonly used in current DSRC-based CV implementations.

7.4.2. Sources of Stochasticity

For a certain market penetration of CV, there is a need to identify the specific vehicles on the link that are equipped with CV devices. This is important particularly for links with high variations in speed between vehicles such as on urban arterials and between lanes on freeway segments with large variations between lanes due to weaving, merging, and lane drops, variations due to signal control, and when the sample size is smaller. In these cases, the accuracy of the travel time estimation largely depends on which vehicles are considered as connected vehicles. The consideration of this stochasticity is important since it results in higher variations in travel time estimates.

An additional source of variations is the time at which the temporary ID for each vehicle changes within the RSU zone. This also affects the results as described later in this section. These two sources of inaccuracy are accounted for in this study using Monte Carlo simulation that allows random selection the CV vehicles and the time the vehicle ID changes. This requires multiple sampling of the data, each with different random selection. As stated earlier, the SAE J2735 standards (SAE, 2009) specify that the BSM data contains a temporary ID which periodically changes to ensure the overall anonymity of the vehicle. One of the limitations of the TCA is in its assignment of the changeover ID, particularly, if the simulation is done for a smaller segment as is the case in this study. The TCA assigns a temporary ID to each vehicle at the beginning of the simulated segment for all vehicles. Thus the ID for all vehicles change at the same time and if the travel time of the simulated segment is less than 5 minutes then the ID will not change during the simulation. In reality, different vehicles will come from different origins and thus the ID will change at different times within the selected link. To overcome this limitation, the TCA code in this study was modified in this study so that the Vehicle ID on the simulated link changes at random times. This was possible since the TCA is an open source software.

The travel time estimation is done for each selected CV selection and ID changing scenarios in the Monte Carlo simulation and both the average and the individual selection results are used in the assessment of the accuracy of the estimation. The average of the runs represents the real-world use of data from multiple days for off-line analysis. An individual run represents using the results for the real-time operation of one day. It is expected that a lower market penetration is acceptable when averaging across multiple days but may not be acceptable when using the data for a single day real-time applications. In this study, 500 estimates were done for both the freeway and the arterial segment.

7.4.3. Travel Time Determination

As mentioned earlier that the vehicle temporary ID changes periodically, so a vehicle may not pass the whole section with the same ID. The test sections have shorter average travel time (41 seconds for freeway, 79 seconds for arterial) compared to ID changeover period (300 seconds).

Thus, the vehicles may change their ID once within the section. For each of the vehicles IDs, the travel time and distance traveled by the ID are calculated. If an ID does not travel a minimum distance (e.g., 100ft) then that ID is not considered in the travel time calculation. The summations of all the valid travel times and distances are used in the estimation of the total travel time (TT) and total distance (D). The link average travel time for each time period is calculated using the following equation:

$$T_{av} = (TT/D)*L_s \qquad (7\text{-}1)$$

where, T_{av} is the average travel time of the segment which has a length of L_s. This whole process is repeated multiple times (500 times) to get the distribution of travel times and the associated error at each market penetration.

7.4.4. Assessment of the Accuracy and Reliability of Travel Time Estimation

Following the above procedure, the Monte Carlo simulation process described above is performed with selected MP of CV. For this study, the considered MP includes 1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 60, 70, 80, and 90 percent. For each MP, the results from the Monte Carlo simulation include travel times estimates for each random selection of the vehicles and their changing IDs. Each of these selections represent a single day of operations. These estimates are compared with the ground truth travel time, estimated based on the complete NGSIM vehicle trajectory data, using a number of accuracy and reliability measures. Toppen and Wunderlich (2004) reported four different measures to represent the error. The details of these four measurements are provided in Table 7-1.

Table 7-1 Travel Time Accuracy Measures

Name	Description	Equation
Mean Absolute Percent Error (MAPE)	Average absolute percentage difference between the estimate and ground truth	$\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{y_i-y}{y_i}\right\rvert$ (7-2)
Mean Absolute Deviation (MAD)/Mean Absolute Error	Average of errors	$\frac{1}{n}\sum_{i=1}^{n}\lvert y_i - y\rvert$ (7-3)
Root Mean Squared Error (RMSE)	Square root of the average of the squared error	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y)^2}$ (7-4)
The Standard Deviation of Percentage Error (SDPE)	Square root of the average of the squared percentage errors	$\sqrt{\frac{1}{n-1}\left(\sum_{i=1}^{n} w_i^2 - n\overline{w}^2\right)}$; (7-5) $w_i = \frac{y_i-y}{y_i}$ (7-6)
* y_i is the estimated travel time of i^{th} iteration, y is the ground truth travel time, n is the total number of iterations and $\overline{w}$ is the average of all the w_i		

The measures, described in Table 7-1, are assessed and the results are plotted against the increase in MP. The data is also used to derive an equation that relates the values of the measures to the market penetration. Different forms are tried for this relationship when fitting an equation based on the simulation results. It is found that exponential curves can represent the relationship with a significant confidence level. As it is not possible to run the simulation for all possible MP of CV to calculate the associated error, this developed equation can provide the expected error at any given MP of CV. The fitted exponential function is provided in Equations 7-7 and 7-8.

$$E_{rr} = \beta_0 e^{f(x)} \tag{7-7}$$

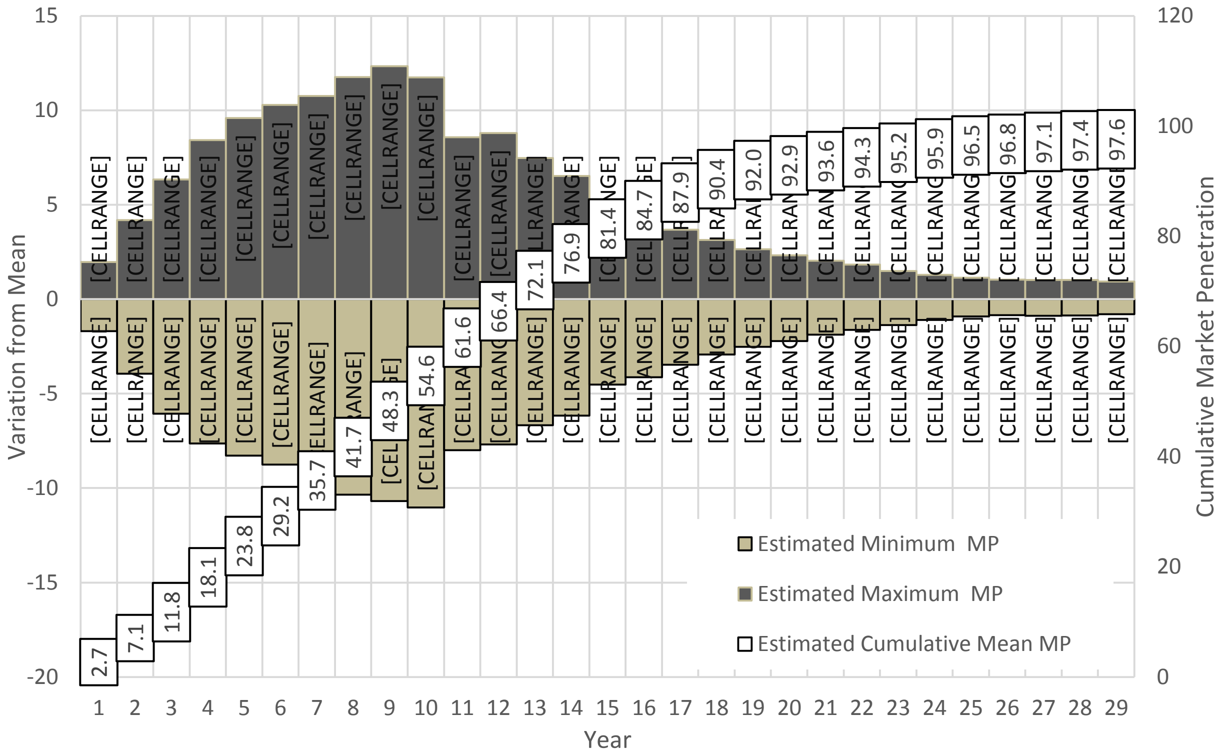
$$f(x) = \beta_1 * x + \beta_2 * x^2 + \beta_3 * x^3 + \beta_4 * x^4 + \ldots\ldots \tag{7-8}$$

where f(x) is a polynomial function, x is the market penetration of CV, E_{rr} is the error value/percentage, and β are the coefficients. The degree of the polynomial function and the coefficient values can vary depending on the type of error and the type of roadway. The statistical software R has been used to fit the regression equation and to validate it with proper statistics. The resulting equations for the two case studies are presented in Table 7-2.

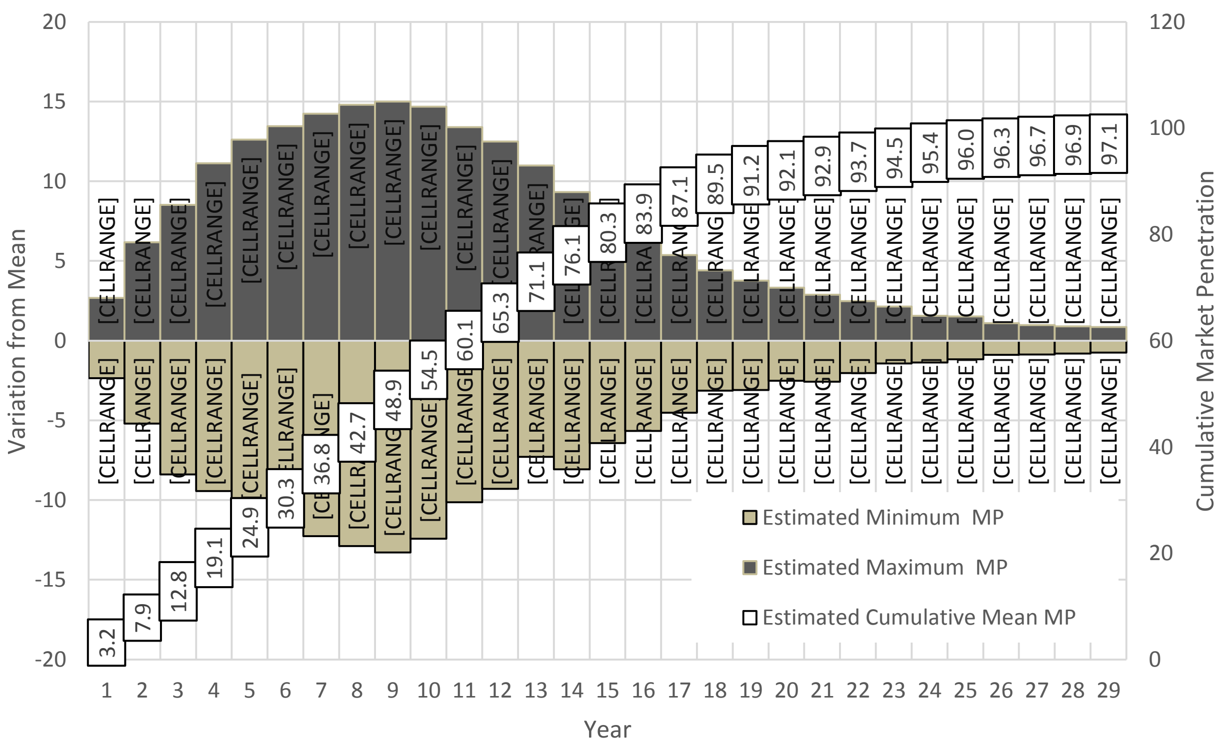
Table 7-2 Regression Analysis Results

	Equation for	β_0	β_1	β_2	β_3	β_4	R-squared value	Adjusted R-squared value	Shapiro-Wilk normality test for model residual (p-value)	Mean of the residual
Freeway	**MAPE**	4.963	$-1.514e^{-01}$	$4.525e^{-03}$	$-6.178e^{-05}$	$2.886e^{-07}$	0.991	0.986	0.998	$1.335e^{-18}$
	SDPE	6.328	$-1.567e^{-01}$	$4.776e^{-03}$	$-6.533e^{-05}$	$3.036e^{-07}$	0.993	0.989	0.767	$-2.668e^{-18}$
	RMSE	2.611	$-1.552e^{-01}$	$4.670e^{-03}$	$-6.286e^{-05}$	$2.868e^{-07}$	0.993	0.989	0.612	$1.335e^{-18}$
	MAD	2.064	$-1.549e^{-01}$	$4.753e^{-03}$	$-6.604e^{-05}$	$3.125e^{-07}$	0.991	0.987	0.100	$-7.473e^{-18}$
	95%	10.24	$-9.807e^{-02}$	$1.611e^{-03}$	$-9.985e^{-06}$	-	0.990	0.987	0.931	$-1.335e^{-18}$
	85%	8.654	$-1.462e^{-01}$	$4.351e^{-03}$	$-6.016e^{-05}$	$2.852e^{-07}$	0.988	0.981	0.626	$-4.535e^{-18}$
Arterial	**MAPE**	13.957	$-9.702e^{-02}$	$1.794e^{-03}$	$-1.265e^{-05}$	-	0.995	0.994	0.153	$6.971e^{-19}$
	SDPE	18.954	$-1.056e^{-01}$	$1.975e^{-03}$	$-1.378e^{-05}$	-	0.994	0.993	0.114	$9.162e^{-20}$
	RMSE	14.984	$-1.041e^{-01}$	$1.920e^{-03}$	$-1.328e^{-05}$	-	0.994	0.993	0.194	$1.117e^{-18}$
	MAD	11.012	$-9.624e^{-02}$	$1.765e^{-03}$	$-1.240e^{-05}$	-	0.995	0.994	0.198	$2.168e^{-19}$
	95%	38.590	$-1.053e^{-01}$	$1.855e^{-03}$	$-1.239e^{-05}$	-	0.989	0.986	0.401	$3.965e^{-18}$
	85%	26.523	$-1.106e^{-01}$	$2.148e^{-03}$	$-1.511e^{-05}$	-	0.993	0.990	0.120	$4.211e^{-18}$

*All coefficients (β) are significant at 95% confidence interval of t-test



(a) Freeway



(b) Arterial

Figure 7-1 MP Variations for Arterials and Freeways

Different accuracy measurements have different significance and thus are useful for different purposes. SDPE is a measure of reliability of the travel time estimates. Higher SDPE values mean that there is a higher variability within the data and that the data may not be reliable to use. MAPE is the average error of all days. RMSE put more emphasis on measurements with higher errors. MAD is the mean of the actual error in seconds. Along with these measures, two percentile measures are also calculated in this study to represent the variations of error between days. The 95% error represents the error that is exceeded 5 days out of 100 days. The 85% percentile measure represent the errors that is exceeded 15 days out of 100 days

7.4.5. Market Penetration Variations between Zones and Links

This study examines the effect of market penetration on travel time accuracy. To be useful to agencies, it is important to estimate the future year, at which each market penetration will be realized. It should be noted that if the overall average market penetration of CV in the nation/region is at a given level, there is no guarantee that a certain link within a region at a certain time will get that market penetration. This is because of the variation in the socioeconomic characteristics of the regions and between the zones of the regions. The higher are the income and thus the new car ownership that determine the vehicle age distribution in a region, the higher is the rate of new vehicles introduction into the traffic stream and thus the higher is the market penetration of CVs.

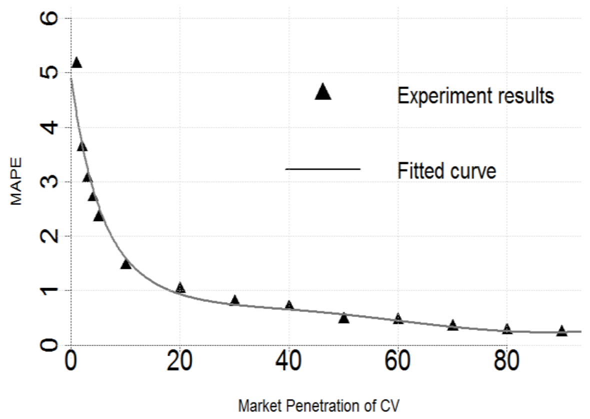
The research of this study (Iqbal, 2017) developed a method to account for this variation for each future year after the NHTSA mandate of CV on all the new vehicles become effective. The method allows the calculation of the variations in the MP between different links in a region. Taking the Southeast Florida region, as a case study test location, Figure 7-1 shows the estimated cumulative mean MP of CV for each year and also the variations of the MP between the links for both the arterials and the freeways. The method described above can be used to estimate the market penetration on specific links when assessing the accuracy of travel time estimation based on connected vehicles in future years.

7.5. Results

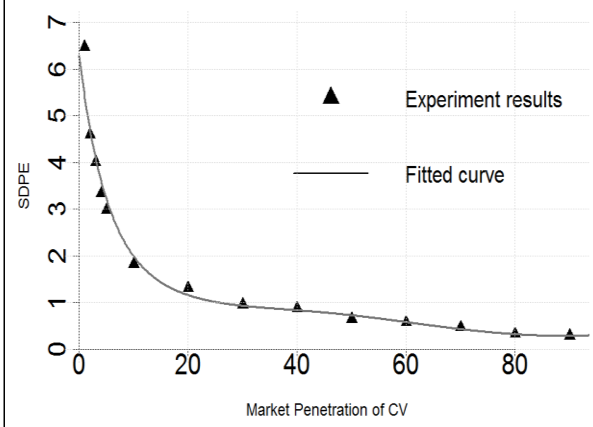
The assessment of travel time estimation using CV, as described in the methodology section was applied to the freeway and arterial segments used for the purpose of this study. Since the travel times of all the vehicles are available based on the NGSIM data, the ground truth travel time can be easily calculated based on these travel time. The ground truth travel time of the freeway segment is 41.3 seconds and for the arterial segment is 79.1 seconds.

The different types of error measures and the fitted curves to these measures based on regression analysis are shown in Figure 7-2 and Figure 7-3 for the freeway and arterial segments, respectively. The horizontal axis of the plots represents the MP of CV and the vertical axis represents the error. The regression analysis results are presented in Table 7-2. The exponential equation is transformed to a linear form and a simple linear regression analysis was performed for this study. All the β coefficients are considered acceptable if it passes the t-statistics at 95% confidence interval. For transformed regression equation it is recommended to check the normality of the model residual and the mean of the residual. For an acceptable transformed linear regression model, the model residual will have a zero mean and a normal distribution. So, a Shapiro-Wilk normality test has also been done on the model residual. The derived equations are useful to predict the error at certain market penetration.

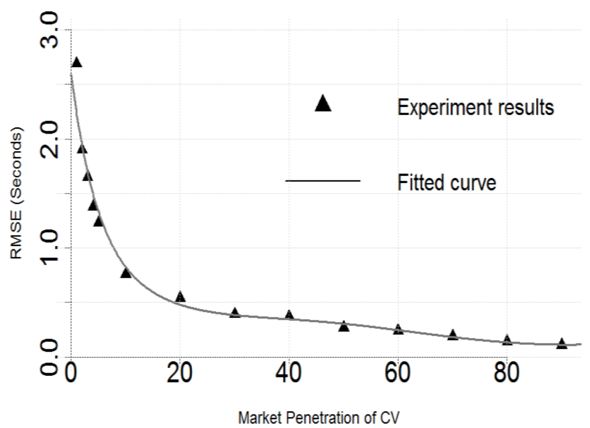
At lower market penetrations, the travel time measurement can vary largely due to the small sample size. At these penetrations, the average of the error represented by the MAPE curve in Figures 7-2 and 7-3, can acceptable. However, the individual estimates for samples identified based the Monte Carlo simulation can have large errors for the examined urban arterial, when examining the 95% or 85% error curves in Figures 7-2 and 7-3 that represent the 5% and 15% highest errors among the travel time estimates. As stated earlier, the individual run results represent travel time measurement on a single day, while the average represents averaging of travel time over multiple days. Figures 7-2 and 7-3 show that a low market penetration of 1% on freeway and about 3%-4% on urban streets are sufficient to produce sufficient data quality for planning purposes, when averaged over multiple days. A low market penetration (1%-2%) is sufficient of is generally sufficient to produce an error that is lower than 10% for operational use in almost all days for the examined high-demand freeway segment, as indicated in Figure 7-2. For the urban street segments (Figure 7-3), however, this data quality cannot be achieved until the market penetration of CV exceeds 10%-15%. Please, note that these results can be different for segments with different demands and configurations, particularly with regard to the congestion level and average spacing of intersections on the urban street segments since these contribute largely to travel time estimate variations. It has been reported that for transportation operation purposes the estimated travel time accuracy should be less than 10-15% RMSE (Turner et al., 1998). So for the freeway, according to this criterion, the desired accuracy (10% RMSE is 4 seconds) could be achieved at 1% MP. However, for the urban street segment, the 10% RMSE (8 seconds) can be achieved when the MP increase to 5%. Toppen and Wunderlich (2004) reported that, if the SDPE is greater than 12% then there is no user benefits when using that travel time use in advanced traveler information systems. As the highest SDPE for the study freeway location is about 4.5% at 1% MP, the CV data could be a reliable source of travel time from the very beginning of the CV mandate implementation. In case of the urban street segments, the reliable travel time could be achieved at a MP of 10%.



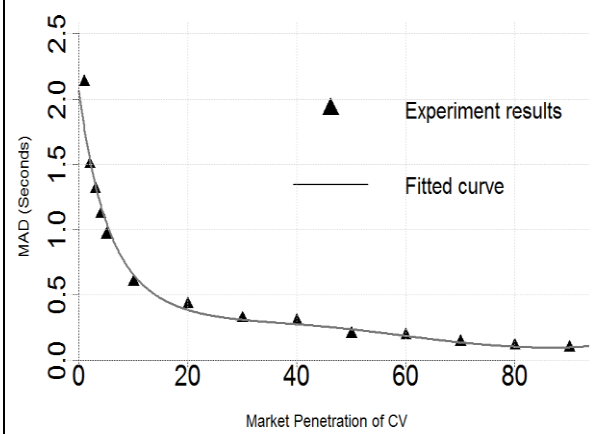
MAPE (%)



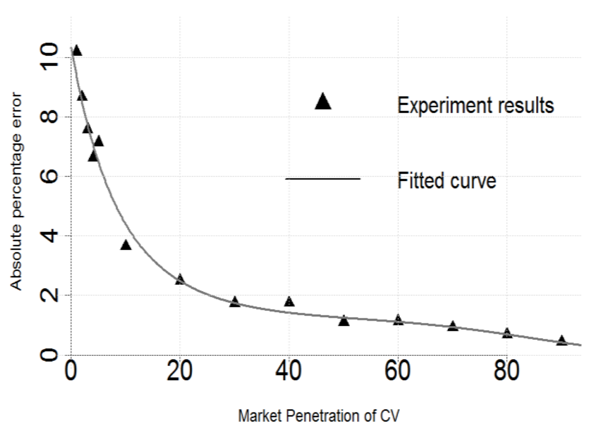
SDPE (%)



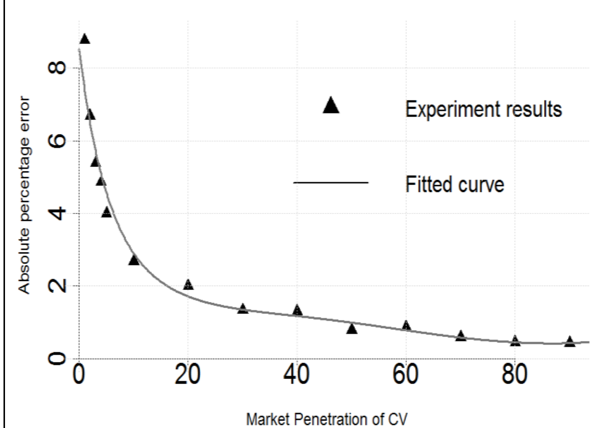
RMSE (Seconds)



MAD (Seconds)



95% (Absolute percentage error)



85% (Absolute percentage error)

Figure 7-2 Travel Time Accuracy Measures for the Freeway Segment Examined in This Study

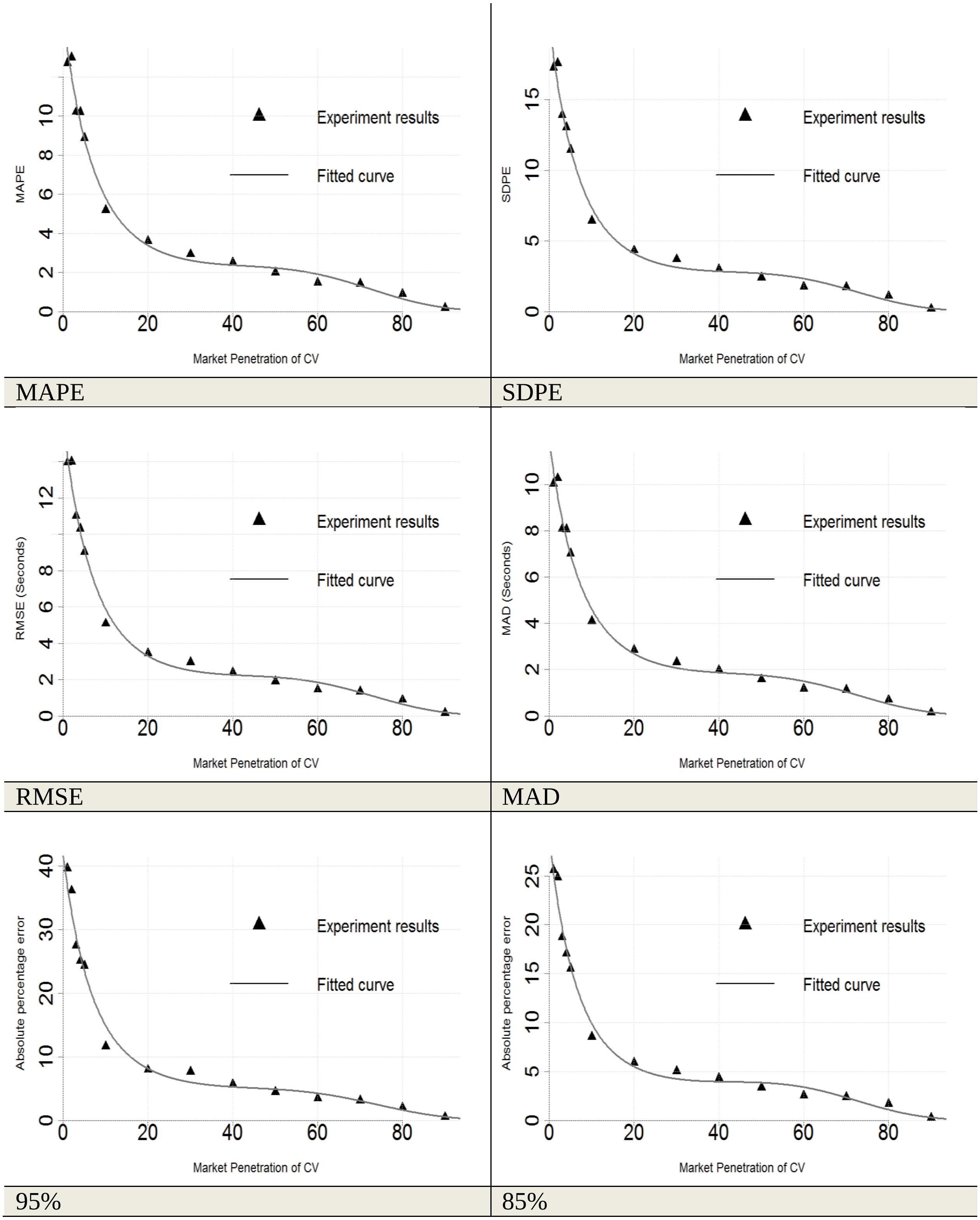
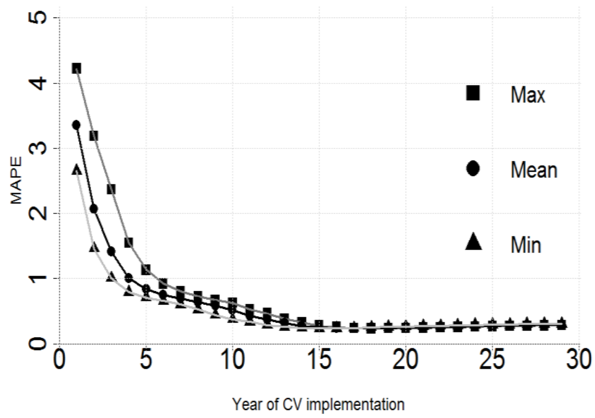
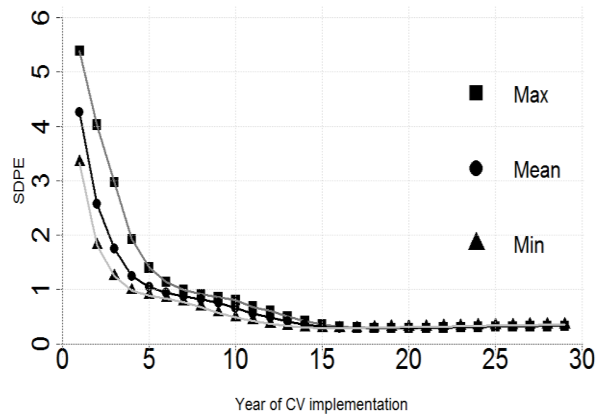


Figure 7-3 Travel Time Accuracy Measures for the Arterial Segment Examined in This Study

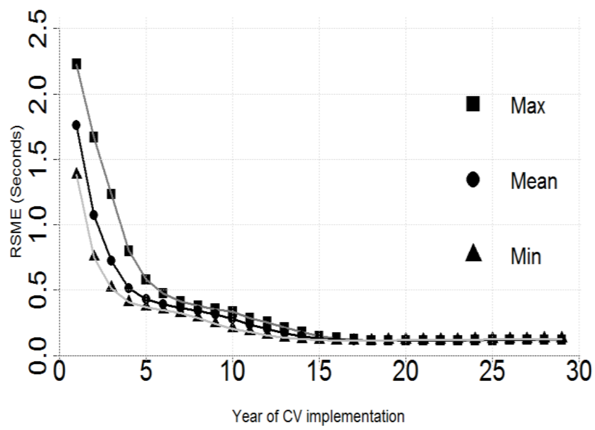
The results in Figures 7-2 and 7-3 present the data quality measures with different market penetrations. The next step is to determine the data quality for each year after the NHTSA mandate. This accuracy is determined for the corridor in this study with the assumption that the CV market penetration on it equals to the mean MP in the region, the minimum MP in the region, and the maximum MP in the region; which as described earlier reflects the socioeconomic characteristics of the users of the facility. Using the derived regression equations developed in Table 7-2 and the CV market penetration variation of Figure 7-1, the different error types for the three above mentioned scenarios in different years after the CV mandate implementation are calculated and presented in Figure 7-4 and Figure 7-5. The figure shows that the differences between the three scenarios are higher in the initial years and decreases gradually in the following years. Also the differences are higher on the urban street compared to the freeway. The variation curve shows that the CV data could be used for both planning and operation purposes from the very first year of CV implementation on the examined freeway segment. However, for the urban street, it will take one to three years for the data quality to be sufficient for use or planning purposes and three to six years for operation purposes depending on the MP of the CV considering the variations in the socioeconomic characteristics in the region.



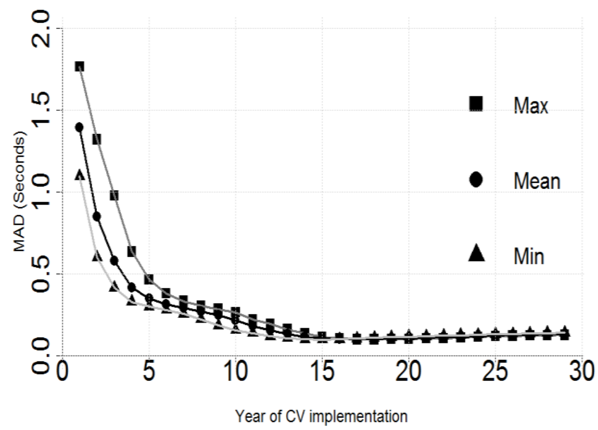
MAPE



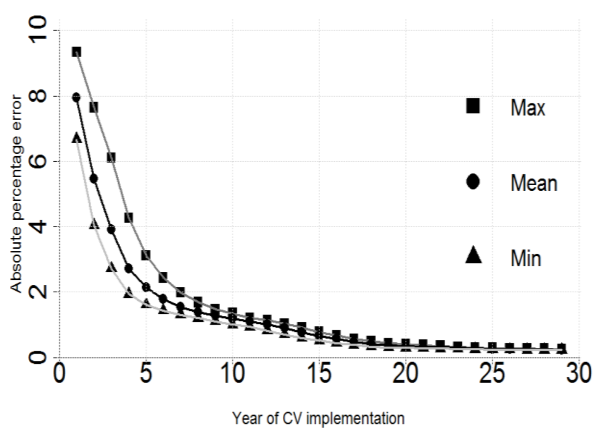
SDPE



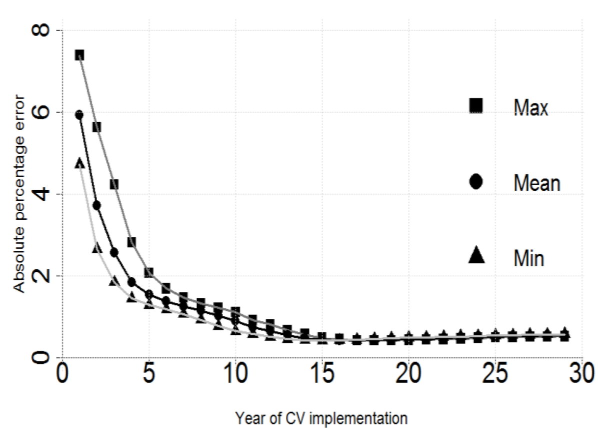
RMSE



MAD



95%



85%

Figure 7-4 Probable Travel Time Accuracy Measures by Year for the Freeway Segment Examined in This Study

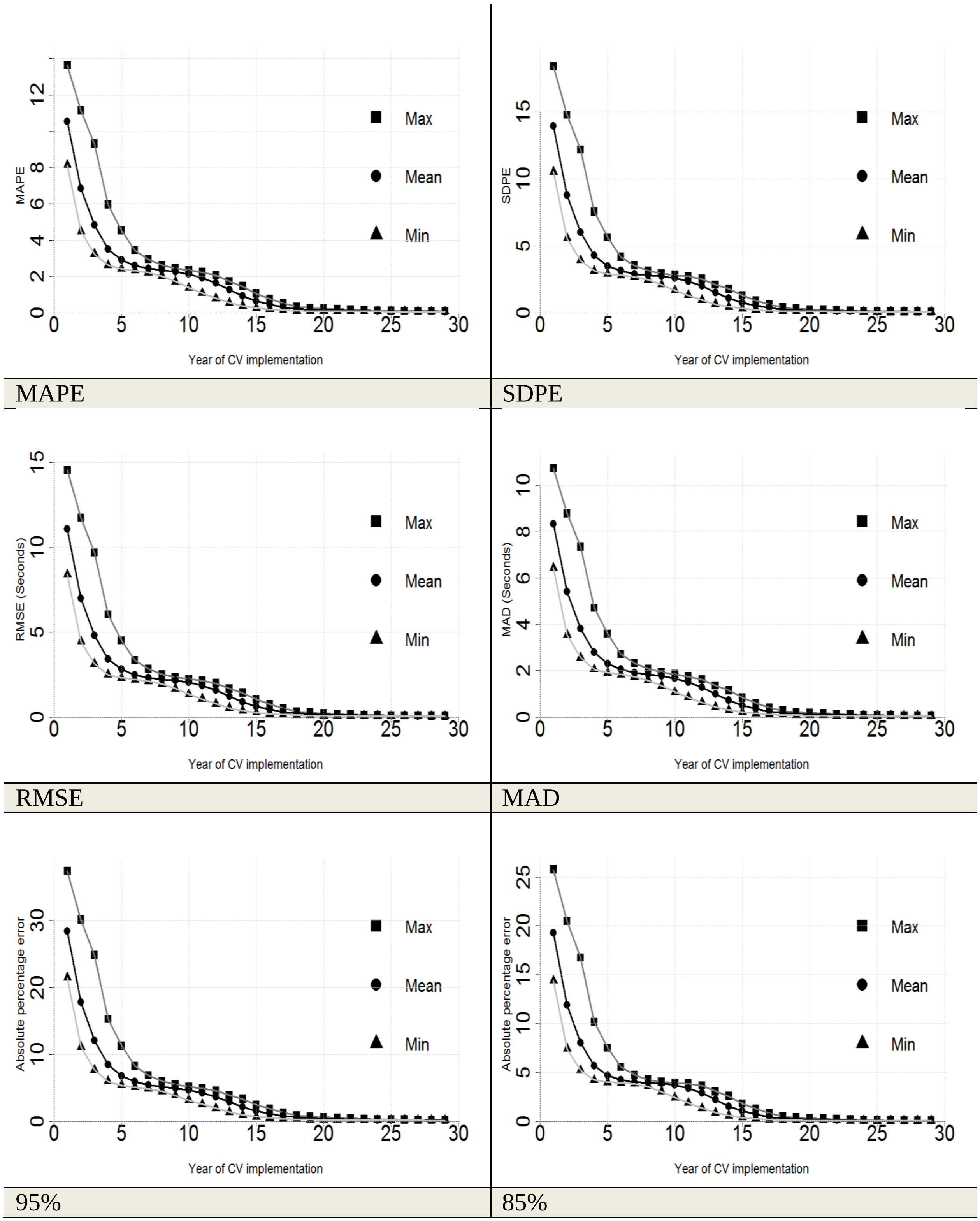


Figure 7-5 Probable Travel Time Accuracy Measures by Year for the Arterial Segment Examined in This Study

7.6. Summary

This study assesses the quality of travel time estimates based on CV data on freeway and urban street segments with high demands and a high density of signalized intersections on the urban street segment. The data quality is examined under different market penetration scenarios considering the randomness in CV presence on the links and the variation in the market penetration between links in the same region due to the variation in the socioeconomic characteristics of the zones in the region. The results show that a low market penetration of 1% on freeway and about 3%-4% on urban streets are sufficient to produce sufficient data quality for planning purposes, when averaged over multiple days. A low market penetration (1%-2%) is sufficient of is generally sufficient to produce an error that is lower than 10% for operational use in almost all days for the examined high-demand freeway segment, as indicated in Figure 5-2. For the urban street segments, however, this data quality cannot be achieved until the market penetration of CV exceeds 10%-15%. It should be noted that these results can be different for segments with different demands and configurations, particularly with regard to the congestion level and average spacing of intersections on the urban street segments since these contribute largely to travel time estimate variations.

Based on the above and in comparison with the expected average market penetration growth, it can be stated that the CV market penetration will be sufficient for use in planning and real-time operation of the investigated freeway segment in the first year after the expected mandate for installing CV technology on new vehicles become effective. However, for the urban street, it will take one to three years for the data quality to be sufficient for use or planning purposes and three to six years for operation purposes depending on the MP of the CV considering the variations in the socioeconomic characteristics in the region.

8. ASSESSMENT OF THE BENEFITS OF QUEUE WARNING IN A CONNECTED VEHICLE ENVIRONMENT BASED ON SURROGATE SAFETY MEASURES

8.1. Introduction

Rear-end collisions are a main safety concern on freeways generally caused by slow or stopped traffic. About one-third of freeway crashes are rear-end crashes (National Transportation Safety Board, 2001a). Recurrent bottlenecks, incidents, and work zones are three main causes of slow queued traffic conditions and consequently rear-end collisions. Queue warning systems are designed to inform drivers about the queued traffic ahead so that they can react in a timely manner. Daimler-Benz (National Transportation Safety Board, 2001b) concluded that 60% of the rear-end collision can be prevented if the drivers had an extra half a second to react to slow traffic ahead. They also concluded that 90% of the rear end collisions could be prevented, if an additional second is given to the drivers. Findings from the evaluation of a number of queue warning systems (QWS) showed reductions in rear-end crashes that range from 14% to 44% (S. T. Team, 1999; Wiles et al., 2003; Ullman, 2016).

Petersen et al. (2013) investigated the accuracy and latency of the queue warning and travel time estimation systems for a Minnesota I-94 smart work zone project. Portable DMSs were used to display the distance to the stopped/slow traffic based on traffic point detector data. The results showed that the accuracy of the distances to the queue were within one mile across different queue lengths. The detectors were spaced at one mile apart (ARTBA Work Zone Safety Consortium, 2015, 2015). The maximum queuing detector error is expected to equal the detector spacing distance.

Existing QWS rely on fixed traffic sensors to detect the back of queue. Thus, the location of back of queue cannot be detected exactly, as indicated by the study of Petersen et al. mentioned above (Petersen, 2013). It is expected that if the transmitted messages from the connected vehicles (CV) are utilized for this purpose, the detection can be faster and more accurate. In addition, with connected vehicles, the delivery of the messages can be done using onboard units instead of DMS, providing more flexibility on how far upstream of the queue the messages are delivered. In the future, the responses of vehicles to the queue warning can be automated when connected automated vehicles become available. This study investigates the accuracy and benefits of the QWS based on connected vehicle data. The study evaluates the safety benefits of the QWS under different market penetrations of CV in future years based on safety surrogate measures estimated using simulation modeling combined with the Surrogate Safety Assessment Model (SSAM) tool.

8.2. Example of Queue Warning Systems

Traffic point detection technologies have been used to detect queue length. A video-based queue detection was implemented on the E313 freeway in Belgian. The detection algorithm read the speed and occupancy measured by the video detection camera and if the occupancy is more than 50% and speed is less than 31 mph, a warning message is sent to the dynamic message signs (DMS) upstream of the incident (Traficon, 1998). The QWS implemented in Toronto on the Catharine's Queen Elizabeth Way use data from microwave sensor stations data and warn the drivers about slow/stopped traffic. If the differential speed between two microwave stations is high, drivers are advised through DMS (Alexander and Chamberlain, 2002). In a city in Denmark, a QWS is activated when the speed obtained from traffic sensors is below 31 mph. After the system activation, speed limits of 56, 44 and 31 mph are shown on the successive DMS upstream of the queue tail (Wiles et al., 2003). The QWS in Oslo, Norway utilizes video detectors to detect the back of queue. If the speed is below 19 mph and occupancy is higher than 30%, and a speed threshold is exceeded for more than 15 sec, the warning is sent to drivers through DMS.

Pesti et al. (2013) proposed a point detector-based queue detection algorithm and used VISSIM micro simulation to evaluate the algorithm. Different system parameters such as speed thresholds, aggregation interval, detector spacing, and portable DMS location were examined. A speed threshold of 35 mph, aggregation interval of 5 minute, and message update interval of 1 or 5 minute were recommended by the study (Pesti et al., 2013).

8.3. Previous Research on CV-Based Queue Warning

Balke et al., (2014) developed a speed harmonization (SPD-HARM) and queue warning (Q-WARN) algorithms, as part of the Intelligent Network Flow Optimization (INFLO) prototype. Three types of queue warning algorithms were proposed for the prototype: traffic management entity (TME)-based, cloud-based, and vehicle-based. The TME queue warning algorithm fuses the data obtained from the traffic sensors and connected vehicles to detect the back of queue (BOQ) and generate queue warning messages through both infrastructure signs and connected vehicles. The process of determining the BOQ is repeated each 5 second. In the cloud-based queue warning algorithm, connected vehicles send BSM messages, queued state (Y or N) and mile marker (MM) location of the vehicle to the cloud using cellular communication. The information is analyzed in the cloud and BSM messages are assigned to sub-links based on the vehicles locations. Then, the back of queue information is sent to the vehicles based on their locations using cellular communications. In the vehicle based algorithm, each vehicle sends its mile marker location and queued state (Y or N) each 1/10th of second to surrounding vehicles using dedicated short range communications (DSRC).

Dowling et al. (2015) evaluated the impact of the prototypes of SPD-HARM and Q-WARN developed by Balke et al. (2014), as described above. The SPD-HARM and Q-WARN prototypes were written in the VISSIM COM interface. VISSIM microscopic simulation was used to model an 8.5 miles of the US-101 freeway in San Mateo, CA. However the Q-WARN application could not be assessed in the microscopic simulation due to the lack of information on how drivers would react to the queue warning messages, according to the authors.

A small-scale demonstration was conducted by the Texas Transportation Institute (TTI), Battelle and Washington State Department of Transportation (Stephens et al., 2015) in order to equip 21 vehicles with CV technologies, traveling on a 23 mile corridor of I-5 in downtown Seattle during the week of January 12, 2015. The connected vehicle data was transmitted and gathered using bot cellular phone and dedicated short range communication (DSRC). The study concluded that no loss of BSM data was observed and there was no disruption in the algorithm due to the loss of BSM data. The data capture, processing and delivering messages to the drivers took less than 10 second. This guarantees that drivers receive the queue warning message 1 mile in advance of the back of queue. The Q-WARN was found to detect the back of queue 3 min sooner and could locate the back of queue more accurately (0.5 to 1.5 miles farther upstream) than the road loop detectors with one mile spacing.

8.4. Surrogate Safety Assessment Model (SSAM)

The implemented queue warning system is evaluated in this study using the SSAM tool. SSAM is a tool developed by the FHWA for performing safety analysis of vehicle trajectory data generated by micro-simulation software to calculate the frequency of different types of conflicts including crossing, lane changing and rear-end conflicts and severity indicators such as time to collision (TTC) and post encroachment time (PET). The tool computes safety of the surrogate measures for each conflict and then generates a summary of the surrogate measures including the mean, maximum, minimum and variance. A conflict is a situation in which two vehicles approach each other so that there is a risk of collision if they keep their movements. TTC is estimated using speed, position, and future trajectory of the two vehicles. If the estimated TTC is less than a threshold, the movement is identified as a conflict. The conflict type is another output of the SSAM tool that determines whether the conflict is a result of rear-end, lane change, or crossing movement (Gettman et al., 2008).

FHWA studies indicate that among the SSAM outputs, the frequency measures (number of conflicts) are more reliable than the severity indicator measures such as TTC (Time To Collision) and PET (Post Encroachment Time (Gettman et al., 2008; Gettman and Head, 2003). Sabra et al. (2010), investigated the relationship between signal timing parameters and the frequency of rear-end, angel and lane change conflicts resulting from SSAM. The study reported the number of conflicts only.

8.5. Methodology

Microscopic simulation modeling using the VISSIM microscopic simulation model is used for the purpose of this research. This study compares the accuracy of queue detection with point detectors to the detection using CV. Thus, this section first describes how the queue is detected using these technologies in this study. Vehicle's speed is used as the measure to detect the queue formation and the back of queue. The speeds are estimated based on point detector and connected vehicle data generated using simulation modeling. To simulate the connected vehicles messages, the Trajectory Conversion Algorithm (TCA) tool, produced by the Federal Highway Administration (FHWA) (2014) was used. The TCA can read vehicle trajectories from simulation or real-world data and emulate the transmission of the Basic Safety Messages (BSM), Probe Data Messages (PDMs) and European Cooperative Awareness Messages (CAM). The method of the study assumes that the speed of the vehicles will be estimated based on data collected from connected vehicles using the BSM specified in the Society of Automotive Engineers (SAE) J2735 standards (SubCarrier Systems Corp, 2009) . It is assumed that the BSM messages will be communicated to the infrastructure using either the Dedicated Short Range Communication (DSRC) or low latency cellular communication technology

MATLAB is used in this study to implement the bottleneck location and BOQ detection algorithms based CV and point detector data generated by the simulation model, as described below. The network is decomposed to *m* segments, with each segment having a 100 ft length. The average speed of each segment is calculated for each analysis time interval, taken as one minute in this study. First, the bottleneck location is identified using the speed measurements, generated by the TCA, by comparing the downstream average speed with the upstream average speed. If the difference is more than 30 mph the location is identified as a bottleneck. To identify the back of queue, the segments are sorted based on their position compared to the bottleneck location. If the segment average speed is below a threshold, the segment is considered queued. The algorithm continues while the segment is queued and the first unqueued segment upstream of the bottleneck is declared as the back of queue. Finally the performance of the connected vehicle-based BOQ detection is compared with the ground truth queue based on VISSIM results and also with the queue estimated based on point detection. The bottleneck location based on point detectors is identified based on comparing the upstream and downstream detector speeds. The detector-based BOQ algorithm is taken from the Pesti et al. (2013) study which estimates the location of queue using equation 8-1.

$$X = X_{DET}(i) + \frac{1}{2}\Delta X_{DET} \tag{8-1}$$

where $X_{DET}(i)$ is the distance from the bottleneck location to the speed detector i, which is farthest detector from the bottleneck location detecting the queue conditions from the bottleneck location and ΔX_{DET} is detector spacing.

The queue warning system is activated when the bottleneck is detected and the queue starts growing. In this study, the queue warning impact is modeled by changing a certain percentage of vehicle's speed upstream of the queue using the VISSIM COM interface. It is assumed that the back-of-queue is detected by the connected vehicle data and the queue warning message is shown dynamically at specific location upstream of the back-of-queue using a DMS upstream of the bottleneck location or connected vehicles onboard units (OBUs). The proportions of vehicles changing speeds in response to the messages reflect the number of connected vehicles equipped with OBU and driver acceptance of the message advisory. In the future, with the introduction of connected vehicle automation, the response to queue warning messages will be set automatically by the vehicle and the driver acceptance will become less of a factor.

The vehicle's trajectories produced by VISSIM are fed to the TCA tool to emulate BSM messages generating from the simulated vehicles. The generated BSM messages are input to the bottleneck and BOQ detection algorithms utilized in this study to investigate their performances. The trajectories of the simulated vehicles were also input to the SSAM tool to obtain safety surrogate measures to analyze the benefits of queue warning system.

8.6. Case Study

A well calibrated network in VISSIM was used to test the BOQ detection algorithms and the queue warning system of this study. The network is a 20 mile segment of the I-95 southbound corridor in Broward County, Florida. A one-lane blockage incident was introduced into the traffic stream to generate a bottleneck location for the purpose of this study. Since VISSIM cannot model incidents directly, the incident was modeled by creating a signal head in one lane of the freeway and the rubbernecking effect was modeled by changing the car following factors upstream of the incident location through the VISSIM COM interface.

The values of the car following factors for the incident link during the incident conditions were changed based on values recommended in the literature. Knoop et al. (2009) explored driver's behaviors under incident condition by looking at the vehicles trajectories collected by a helicopter from two actual incidents. The results showed that driver's reaction time, headway distribution and capacity are effected by the incident presence. A bimodal headway distribution, increase in reaction time and a 30% decrease in capacity were observed. Some researchers found that incident impacts can be modeled by changing VISSIM car following factors. A detailed description of the VISSIM car modeling factors can be found in VISSIM user manual (PTV, 2012). A study of the VISSIM calibration parameters (Woody, 2006) found that the CC0, CC1, CC2, CC4 and CC5 parameters have the most effect on the capacity of a freeway link. CC0 is the desired distance between stopped cars. CC1 is the desired time headway that the driver wants to keep to the leading vehicle. CC2 is the safety distance, which is the minimum distance a driver keeps with the leading vehicle. CC4 and CC5 are the parameters controlling the speed differences between a following and leading vehicle. Increasing CC0, CC1, and the absolute value of the CC4/CC5 ratio result in capacity reduction. Gomes et al. (2004) modified three CC1, CC2, and CC4/CC5 to model the capacity drops resulting from curvatures. The calibrated parameters for the studied freeway segment were CC1, CC2, and CC4/CC5 were 1.7, 0.9 and -2/2 respectively (22). Rahman and Mattingly (2014) found that the car following values proposed in prior studies for incident modeling produced better results in terms of macroscopic measures compared to the default values. Rompis et al., (2014) developed a methodology for VISSIM calibration based on kinematic queuing theory. . The incident was modeled by coding a traffic signal in VISSIM. Based on the reviewed study results; CC0, CC1 and CC2 values of 3.8, 1.5 sec and 7.5 m, respectively, were used to replicate the actual freeway operation under incident condition.

The vehicle measurements including vehicle's speeds are assumed to be continuously transmitted as part of BSM messages communicated to the infrastructure. These messages can be then processed by roadside units or at other infrastructure locations. As described earlier, in this study, the TCA tool was used to generate BSM messages based on vehicle trajectories from the VISSIM simulation model. To test the methods presented in this study, a total of 40 scenarios with different market penetrations and TCA seed numbers were evaluated (4 different CV market penetrations and for each one 10 TCA runs with different seed numbers was performed). The purpose of having multiple TCA runs is to randomly assign connectivity to the vehicles, since the position of the connected vehicles relative to the bottleneck location affects the accuracy of the results. To confirm that the 10 TCA runs are sufficient, the required numbers of runs at the 95% confidence level were estimated using Equation 8-2.

$$n = \frac{\sigma^2 * t^2}{e^2} \tag{8-2}$$

where σ is the standard deviation of the measurements, t is the associated t value with the 95% confidence interval (T distribution) and e is the acceptable measurement error (assumed to be 200 ft in this study). The calculated required numbers of runs for the 3% market penetration was 9 runs and for 6%, 9% and 15% was 2 runs, conforming that 10 runs were sufficient.

The generated BSMs were fed to the detection algorithms and the bottleneck location and back of queue were determined each minute. The detected back of queue is compared with the ones detected by the VISSIM full trajectory and the error percentage was calculated using Equation 8-3.

$$Error(i,k) = 100 * \frac{(BOQ\_C_{ik}) - (BOQ\_GT_i)}{BOQ\_GT_i} \quad (8\text{-}3)$$

where

Error(i,k) is the error percentage at time step i for the kth TCA run,

BOQ_C_{ik} is the back of queue detected by connected vehicle data at time step i for the kth TCA run, and

BOQ_GT_i is the back of queue detected by full VISSIM trajectory at time step i

The average error for each run k was calculated using Equation 8-4 and is shown in Figure 8-1. As shown in this figure, the average error decreased with increasing the market penetration. The median error for the 3%, 6%, 9% and 15% market penetrations were 6.52%, 3.72%, 2.94% and 2.36% respectively. The error range for each market penetration is shown in Table 8-1. The detector-based error was calculated to be 49.53% compare to the VISSIM ground truth.

$$Average\_Error(k) = \frac{\sum_{i=1}^{10} Error(i,k)}{10} \quad (8\text{-}4)$$

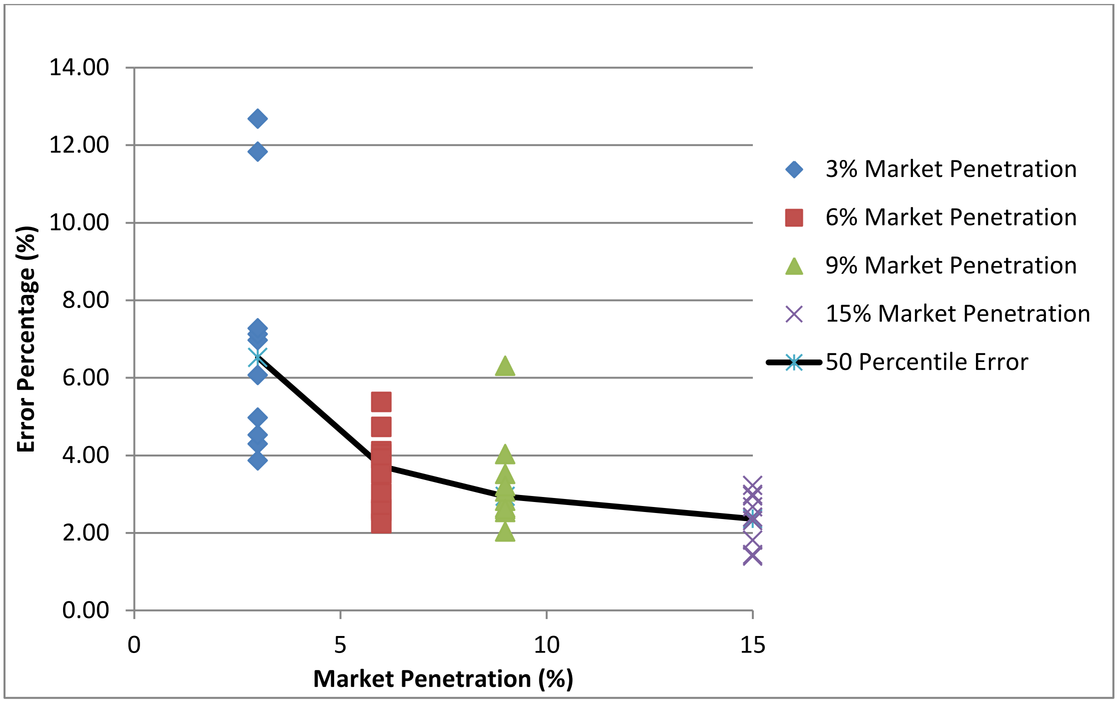


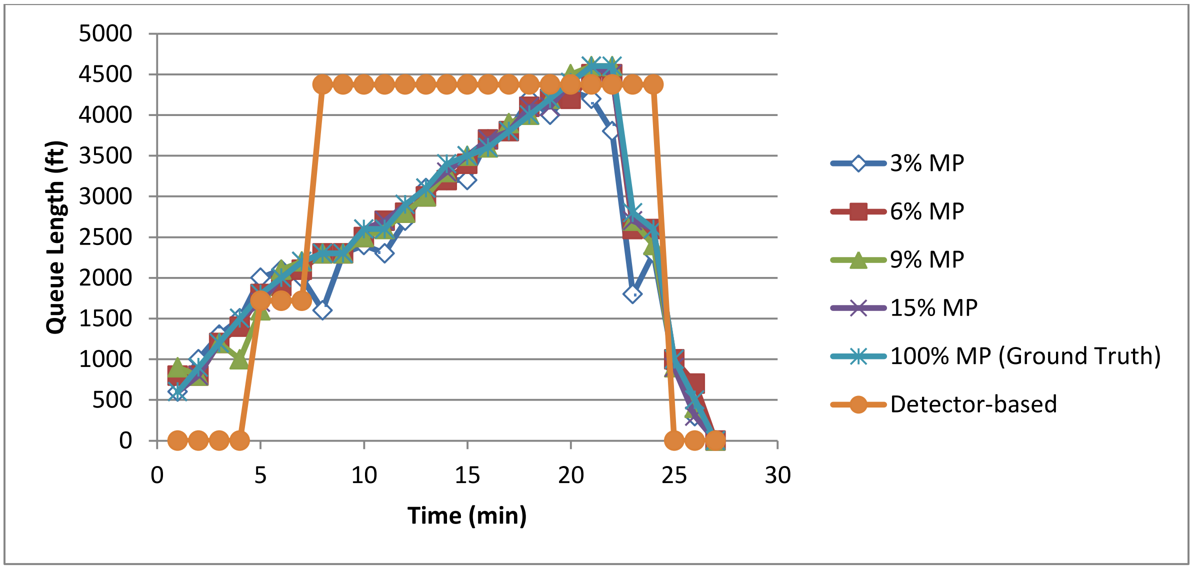
Figure 8-1 Average Error Percentages for Each Run

Table 8-1 Error Percentage for Each Market Penetration Rate

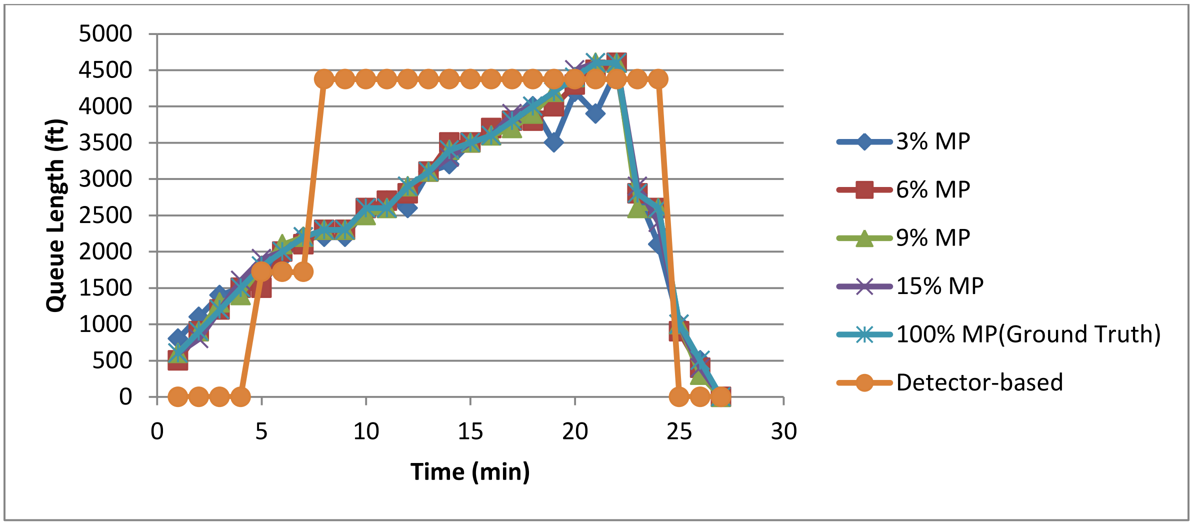
Market Penetration Rate (%)	Error Range (%)	Median Error (%)
3	3.86 to 12.68	6.52
6	2.25 to 5.37	3.72
9	2.03 to 6.3	2.94
15	1.36 to 3.22	2.4

Figures 8-2(a) and 8-2(b) shows the estimated queue lengths associated with the worse and median runs of the 10 TCA runs for the four different connected vehicle market penetrations versus the ones estimated by detector data and ground truth trajectory. As can be seen from Figure 8-2, even the worst estimated queue lengths of the ten TCA runs are better that the ones estimated by the point detectors compared to ground truth. With low market penetration (3%), however, although the median of the runs produce accurate results at all time intervals, the worst TCA runs produced errors at the beginning of the formation and dissipation of the queue, as shown for the 8^{th} and 23^{th} minute in Figure 8-2(a). This reflects the lower number of vehicles in shorter queues. It should be mentioned however, that with the expected National Highway Transportation Safety Administration (NHTSA) mandating CV technologies on all new vehicles, it is forecasted that the CV market penetration will be around 5-7% in this first year that this mandate becomes effective, considering the rate, at which new vehicles are introduced in the traffic stream [20]. The result from this study was also examined to determine how fast the bottleneck that starts the queue buildup is detected. It was found that the bottleneck was detected 4 minutes sooner with connected vehicle data compared to point detectors at all the market penetration rates, as can be seen from Figure 8-2(a).

Once the queue is detected by the developed algorithm, the queue warning system is assumed to be activated by delivering messages to the drivers ahead of the bottleneck. In order to implement the queue warning system, driver’s response to the queue warning messages needs to be known or estimated. Li et al. (2010) conducted a study of the effectiveness of PCMS messages (‘Road Work Ahead’) on reducing vehicle speed in rural highway work zones. The work zone was located on US-36 and US-73 in Kansas. The study concluded that the vehicles speeds were reduced by 4.7 mph over a distance of 500 ft when the PCMS was turned on. Richards and Dudek (1986) found, based on a field study, that when using DMS, as a speed control device in work zones, vehicle speeds were reduced by 7 mph on average. Zech et al. (2008) evaluated the effectiveness of three DMS messages on vehicle speed and variance reduction in highway work zones. The effective message reduced the speed by 3.3 mph-6.7 mph. Dixon and Wang (2002) found a 6-7 mph reduction in vehicle’s speed in the immediate vicinity of DMS signs providing advanced warning of DMS. Garber and Patel (29) concluded that vehicle speeds reduced by 6 mph due to DMS posted ahead of work zones. A similar study by McCoy et al. (1995) reported that DMS activation resulted in 4 to 5 mph reduction in vehicle speed and the percentage of drivers exceeding the speed limit (45 mph) was reduced by 20 to 40 percent. This study assumed that the messages are delivered using OBUs or DMS located about one mile ahead of the maximum queue length. The complying drivers were assumed to reduce their speed by 10 mph and as a result, the average speed of the warning zone reduced by 3 to 10 mph depending on the compliance rate with the QWS.



(a) Estimated Queue Lengths Associated with the Worst Run vs Detector-Based and Ground Truth



(b) Estimated Queue Lengths Associated with the Worst Run vs Detector-Based and Ground Truth

MP: Market Penetration of connected vehicles in the traffic stream

Figure 8-2 Estimated Queue Lengths by Connected Vehicle Data vs. Detector-Based and Ground Truth-Results Associated with (a) Worst TCA Run (b) Median TCA Run

The next step was to examine the safety impacts of the QWS. The trajectory file was imported to the SSAM tool to estimate the surrogate safety performance measures. In this study, the trajectories with and without queue waning were imported to the SSAM tool to investigate the effect of the implemented queue warning system on the number of rear-end conflicts. The tested compliance rates were 3%, 5%, 10%, 15%, 20%, 30%, 50% and 70%. The compliance rate with the QWS was assumed to be the combination of compliance with the DMS signs and the OBUs. 100 SSAM runs were performed for each compliance rate and a t-test with a significance level of 0.05 was conducted for each scenario. The results of the t-test are shown in Table 8-2. Number of rear-end conflicts for each scenario was compared with the conflicts of no queue warning scenario and the significance of the reduction in the conflicts with different compliance rates were determined. As it can be seen from Table 8-2, the reduction in the number of the rear-end conflicts becomes significant when the compliance with the queue warning messages is more than 15%. Most of the expected safety benefits are expected to be achieved when reaching a 50% compliance rate, as the difference in the percentage reduction in rear-end conflicts between 50% and 70% compliance rates is very small.

Table 8-2 T-test Conducted for Determining the Effect of Queue Warning System on the Number of Rear-end Conflicts Resulted from the SSAM

SSAM Measure	Compliance with the queue warning System	Mean	Replications	t-value	t-critical	Significance	Mean difference	Difference Percentage (%)
Rear-end Conflicts	0% (No Queue Warning)	6589.164	100	NA	NA	NA	NA	NA
	2%	6589.164	100	0.039	1.66	No	6.492	0.10
	5%	6136.208	100	1.321	1.66	No	226.478	3.56
	10%	6332.944	100	1.472	1.66	No	256.22	3.89
	15%	6238.954	100	2.201	1.66	Yes	350.21	5.31
	20%	6235.652	100	2.212	1.66	Yes	353.512	5.37
	30%	6228.446	100	2.331	1.66	Yes	360.718	5.47
	50%	6097.12	100	2.359	1.66	Yes	492.164	7.47
	70%	6095.17	100	2.37	1.66	Yes	493.992	7.50

NA: Not Applicable

8.7. Summary

The results from this study indicate that a relatively low market penetration, around 3% to 6%, for the congested freeway examined in this study, is sufficient for accurate and reliable estimation of the queue length. It can be concluded that having 6% connected vehicles in the traffic stream allows the estimation of the back of queue location with less than 4% error on average and a maximum error of about 5.4%. Even at 3% market penetration, the CV-based estimation of back of queue identification is more accurate than that based on detector measurements. It is also found that CV data allows faster detection of the bottleneck and queue formation. The connected vehicle-based algorithm can detect the start of queue 4 minute sooner than the detector-based algorithm. Further, it is concluded that the QWS improved the safety conditions of the network by reducing the number of rear-end conflicts. The safety impacts become significant when the compliance with the queue warning messages is more than 15%. However, most of the expected safety benefits are expected to be achieved when reaching a 50% compliance percentage and the improvements with increasing the compliance percentage above 50% is small.

9. A METHODOLOGY TO ASSESS THE QUALITY OF TRAVEL TIME ESTIMATION AND INCIDENT DETECTION BASED ON CONNECTED VEHICLE DATA

9.1. Introduction

Connected vehicles technologies are expected to have a significant influence on transportation system management and operations (TSM&O), starting in the next few years. Thus, transportation agencies need to start preparing for the next generation of advanced traffic management strategies that utilize connected vehicle to Infrastructure (V2I) technologies. The draft Federal Highway Administration (FHWA) V2I (Vehicle to Infrastructure) Deployment Guidance (FHWA, 2015) encourages V2I deployments, but it states that the United States Department of Transportation (USDOT) will not require public agencies to implement V2I technology or applications, and recommends that this implementation should be done based on agency assessments.

One of the potential applications is the use of CV data to support various TSM&O processes. The National CV Field Infrastructure Footprint Analysis document produced by the American Association of State Highway and Transportation Officials (AASHTO) (Wright et al., 2014) stated that "Public agencies will assess and trade-off the opportunities to use connected vehicle probe data aggregation and processing versus the continued deployment, operations, and maintenance of traditional ITS vehicle detection versus purchasing commercial traffic information services." Given the accelerated advancements in technologies, there will be an increasing need to determine the quality of the data collected using CV, in comparison to the quality of data collected from existing data sources.

Researchers have investigated utilizing connected vehicle data assessing various performance measurements. These performance measurements can be used to support planning, planning for operations, and operations. However, the quality of the estimation is expected to be a function of the sample size of the CV. The sample size is the multiplication of the percentage of CV in the traffic stream and the total traffic volume. The CV percentage is a function of the CV market penetration and is expected to grow with time. This study investigates the use of CV data as an alternative to existing data acquisition techniques in providing two critical functions to support TSM&O: travel time estimation and incident detection. In support of this investigation, the study develops regression models to estimate travel time measurement accuracy, travel time measurement reliability, and incident detection latency as functions of the traffic demand level and the CV proportion in the traffic stream. The developed regression models are used in conjunction with a prediction of CV proportions in future years to determine when the CV technology can provide sufficient data quality to replace existing data acquisition systems. The results can be used by TSM&O programs and agencies to plan their investment in data acquisition alternatives in future years. Although a wide range of performance measurements could be derived using connected vehicle data, this study focuses on the above two performance measurements to demonstrate a methodology that can be used for other measures.

9.2. Connected Vehicle Data

The data elements generated from CV can be used to support various mobility, safety, and environmental applications. Although some of these data elements such as speed and acceleration and deceleration can be obtained without the need to connect to a vehicle's onboard diagnostic port (OBD-II), obtaining many other useful data elements requires connection to the OBD-II. The data will then need to be transmitted using connected vehicle messages utilizing dedicated short-range communication (DSRC) or other communication technologies such as cellular communications.

The CV message types and components are specified in the Society of Automotive Engineers (SAE) J2735 standards (SAE, 2016). The J2735 standards specify the number of message types. The basic safety message (BSM) is one of these message types that will be used for vehicle-to-vehicle communication. The BSM contains vehicle safety-related information broadcasted to surrounding vehicles but can be also sent and/or captured by the infrastructure. The BSM, as defined in the J2735 standards, consists of two parts. Part 1 is sent in every BSM message broadcasted 10 times per second and will be mandated to be broadcasted by the NHTSA ruling. It contains core data elements, including vehicle position, heading, speed, acceleration, steering wheel angle, and vehicle size. BSM Part 2 consists of a large set of optional elements such as precipitation, air temperature, wiper status, light status, road coefficient of friction, Antilock Brake System (ABS) activation, Traction Control System (TCS) activation, and vehicle type. BSM Part 2 elements are sent based on criteria that are not specified in the J2735 standards. However, not all of these parameters are currently available from vehicles and they will not be mandated by the USDOT. In this study, Basic Safety Message (BSM) part I data is considered in the evaluation, as it is expected to be available in the first phase of CV implementation and is sufficient for travel time estimation and incident detection.

9.3. Findings from Previous Research

There are a number of studies in the literature focusing on travel time estimation utilizing connected vehicle data. Zou et al. (2010) estimated travel time based on CV data using traffic simulation of a hypothetical network that simulates vehicles broadcasting Probe Data Messages (PDM) data according to the J2735 standards (SAE, 2016). They found an average error percentage of 27.6%, 12.5%, and 8.2% for 1%, 5%, and 10% CV proportions, respectively. However, that study did not consider the demand level in this investigation, which as stated earlier is one of the two factors that determine the sample size. Argote et al. (2012) estimated measures of effectiveness based on real-world vehicle trajectories collected by the Next Generation Simulation (NGSIM) program. These measures of effectiveness include queue length, speed, number-of-stops, acceleration noise, and the average delay per unit distance. A drawback of this study was that the method used in the calculation of travel time does not consider the change of vehicle ID during its course of travel, as specified in the J2735 standards to protect traveler privacy. Iqbal et al. (2017a) analyzed the BSM data to estimate the travel time utilizing the NGSIM data.

Connected vehicle data has been also used by few researchers to detect incidents. Utilizing microscopic simulation, Crabtree and Stamatiadis (2007) examined the use of CV data communicated using DSRC to detect freeway incidents. The utilized algorithm was based on comparing measured travel time and "normal" travel time estimated based on no-incident condition data. The results showed that the proposed algorithm can rapidly and reliably detect incidents. The study concluded that for a CV proportion of 30% (25% of the connected vehicles were trucks and 5% were cars), the mean time to detect (MTTD) ranges from 2 minutes to 4 minutes for a DSRC roadside unit (RSU) spacing of 2.0 miles and 2.5 minutes to 14 minutes for a reader spacing of 10 miles.

None of the above studies have investigated the effect of the demand variation on either the travel time estimation or incident detection. This is important since as stated earlier, the accuracy of travel time estimation and the latency of incident detection are expected to be a function of the sample size, which is a function of both the percentage of CV, as well as the demand level. This study focuses on examining the effect of demand variation on travel time estimation and also the timeliness of incident detection under different CV. The results of this examination, combined with forecasting the proportions of CV in future years, will allow agencies to determine for a given future year, the freeway and urban street segments by time-of-days, where CV provides acceptable travel time estimation and incident detection.

9.4. Methodology

The first step to achieve the objectives of this study is to emulate the CV BSM data utilizing vehicle trajectories generated from the simulation of a freeway network and an arterial street network with different levels of traffic demand. Next, the generated BSM data is used to estimate the travel time and for incident detection under different CV proportion assumption. Then, different regression equations are developed to estimate the effect of demand and the CV proportion on the performance of the travel time estimation and incident detection. These regression equations are then used to predict the performance of travel time estimation and incident detection for future years of CV implementation to determine the year when the CV data utilization produces acceptable accuracy and timeliness. The subsections below present more detailed discussion of the methodology.

9.4.1. CV Data Emulation

CV data has started to become available for research from pilots, testbeds and initial CV deployments across the country. However, the CV proportions in the traffic stream at these locations are very limited. In this study, CV data is emulated using trajectories from microscopic simulation models that are used to as inputs to the trajectory Conversion Algorithm (TCA) [Version 2.3] (22) tool developed by the Federal Highway Administration (FHWA). The TCA tool then generates data that emulate CV data following the J2735 standard (OSADP, 2015). The vehicle trajectories are generated utilizing the VISSIM microscopic simulation modeling tool.

The output from the TCA tool to produce emulated BSM Type 1 data at 1/10th-second resolution utilizing trajectory data obtained from the simulation models. The input trajectory file to the TCA contains the identifications (ID), time stamps, speeds, position coordinates, and acceleration data of the vehicles. Using the trajectories as inputs, the TCA generates connected vehicle BSM Type 1 data following the SAE J2735 standards. The user of the TCA tool can select the CV proportion, message type (e.g., BSM or PDM) and the communication type (DSRC or cellular). This study focuses on BSM Type 1 data that is expected to be captured by DSRC RSUs. The coverage area of an RSU is about 2,600 feet to 4,000 feet (Andrews and Cops, 2009; McGurrin, 2012). The TCA tool allows the specification of data transmission loss during the data transfer between the OBU and the RSU. It has been reported that the loss rate with DSRC communications varies between 10% and 20% with an average of 12% (Kandarpa et al., 2009). In this study, 12% transmission loss is specified in the input to the TCA tool. It is expected that the actual loss is a function of the availability of line of site in the coverage area and this should be considered with RSU siting. The TCA tool changes the vehicle ID every five minutes with a 30-second buffer window. This means that the ID will change over a period that ranges from 5 to 5.5 minutes, as is commonly used in current DSRC-based CV implementations.

There are two sources of stochasticity that are considered when assessing the utilization of the CV data. First, the accuracy of travel time estimation largely depends on the selection of vehicles on the link that are CV-equipped. Selecting different vehicles could result in variations in travel time particularly at lower CV proportion when there is a higher variation in travel time between vehicles. The variation in speed/travel time increases due to weaving, merging, diverging, and lane drops on freeways and due to signal control on arterial streets. The second source of the stochasticity is the time at which the temporary ID for each vehicle changes within the RSU zone. As stated earlier, a CV vehicle ID changes every 5 minutes at different times in the simulation. A Monte Carlo analysis is used in this study to account for these two sources of stochasticity by randomly selecting the CV vehicles from the traffic stream and the time the vehicle ID changes. This requires multiple sampling of the data, each with the different random selection. As stated earlier, the SAE J2735 standards (SAE, 2016) specify that the BSM data contains a temporary ID which periodically changes to ensure the overall anonymity of the vehicle. One of the limitations of the TCA is in its assignment of the changeover ID, particularly, if the simulation is done for a smaller segment as is the case in this study. The TCA assigns a temporary ID to each vehicle at the beginning of the simulated segment for all vehicles. Thus, the ID for all vehicles change at the same time and if the travel time of the simulated segment is less than 5 minutes then the ID will not change during the simulation. In reality, different vehicles will come from different origins and thus the ID will change at different times within the selected link. To overcome this limitation, the TCA code is modified in this study so that the Vehicle ID on the simulated link changes at random times. This is possible since the TCA is an open source software.

9.4.2. Travel Time Estimation

The travel time estimation is done for each CV selection and ID changing scenario resulting from the Monte Carlo analysis and both the average and the individual selection results are used in the assessment of the accuracy of the estimation. The average of the runs represents the real-world use of data from multiple days for off-line analysis. An individual run represents using the results for the real-time operation of one day. It is expected that a lower CV proportion is acceptable when averaging across multiple days but may not be acceptable when using the data for a single day real-time applications. In this study, 500 runs are done for both the freeway and the arterial street segments.

The travel time estimation method needs to consider the vehicle ID change requirements, according to SAE J2735 standards. As mentioned earlier, the vehicle temporary ID changes periodically, so a vehicle may not pass the whole section with the same ID and a method that estimates the travel time based on full vehicle trajectories is not possible. The vehicles may change their ID once within the section. The utilized method is described next. For each of the vehicles IDs, the travel time and distance traveled by the ID are calculated. If an ID does not travel a minimum distance (e.g., 100ft) then that ID is not considered in the travel time calculation. The summations of all the valid travel times and distances are used in the estimation of the total travel time (TT) and total distance (D). The link average travel time for each time period is calculated using Equation 9-1.

$$T = \frac{\sum_{i}^{n} t_i}{\sum_{i}^{n} d_i} D_s \tag{9-1}$$

where, T is the average travel time of the segment which has a length of D_s; t_i, and d_i are the travel time and distance traveled by the ith ID; n is the total number of ID's. A Monte Carlo analysis is done with multiple runs (500 times) to get the distribution of travel times and the associated error at each CV proportion and demand level.

9.4.3. Incident Detection

This study examines freeway incident detection utilizing CV data. Vehicle's speed estimated based on CV BSM data is selected as the base measurement for incident detection. The utilized method which is based on hypothesis testing of the occurrence of the incident in a segment. The simulated network is decomposed to 300 feet long segments and the average speed of each segment is calculated at each time step based on CV data. This method simply takes the average of the measurements in each segment during a time period (45 seconds in this study) and if the difference between the average speeds of the downstream and upstream segments is more than a threshold, the location is identified as an incident location. A certain percentile of speed differences distribution under no incident conditions (99 percentile in this study) is selected as the threshold for testing the hypothesis that there is an incident on the subject segment. In other words, the probability of false alarm is set to 1% and the associated threshold is calculated.

As with travel time estimation, the method is tested using emulation of CV data generated by the TCA tool based on trajectories produced by the VISSIM microscopic simulation tool. VISSIM is used to emulate incident occurring in a mixed connected vehicle and not connected vehicles in a traffic stream, as described earlier.

9.4.4. Regression Analysis

This study develops models of the relationship between the performance of the travel time estimation and incident detection using CV data and two independent variables: traffic demand and CV proportion. In this study, multiple linear regression is used to derive the relationships. In the derived model for travel time estimation, the independent variable in the regression is the error in travel time estimation and the dependent variables are the CV proportion and the traffic flow rate. Different transformations of the independent variables are tried to identify the best fit of the regression model. It is found that the best form is the one shown in Equation 9-2.

$$E_{err} = \beta_0 + \beta_1 * \cos(S) + \beta_2 * \log(CP) + \beta_3 * S + \beta_4 * CP + \varepsilon \qquad (9\text{-}2)$$

where β_i are the regression coefficients, S is the flow rate measure and CP is the CV proportion, E_{err} is the error of type 'err', which is one of the two types discussed next. Two types of error are considered in this study: the Absolute Mean Percentage Error (AMPE) and the Standard

Deviation Percentage Error (SDPE). The AMPE is the average error of all days. The AMPE is useful when assessing the average accuracy, and is possibly appropriate for the assessment when using data for planning purposes. The Standard Deviation of Percentage Error (SDPE) is a measure of the reliability of travel time estimates with higher values, indicating a higher variability in the quality of the data. For the freeway, the flow rate measurement is represented by the volume/capacity (v/c) ratio. For the arterial street, the flow rate measurement is represented by the degree of saturation (volume/saturation flow ratio) of the critical intersection along the test section. The statistical software R is used in the analysis.

The MTTD incident is also estimated as a function of the v/c ratio and CV proportion for freeway segments. This study has found that an exponential regression equation could represent the relationship well as follows:

$$MTTD = \beta_0 + \beta_1 * e^{-(v/c)*CP} \tag{9-3}$$

where β are the regression coefficients, CP is CV Proportion (%), v is volume (vph) and c is the capacity (vph).

9.5. Test Locations

To demonstrate the proposed methodology, simulation models were developed and calibrated for a freeway network and an arterial network utilizing the VISSIM microscopic model. The details of these networks are provided below.

9.5.1. Freeway Segment

The freeway segment utilized this study is a 15-mile segment of the I-95 southbound corridor in Broward County, FL coded and calibrated in VISSIM. The segment that is selected for the assessment in this study consists of an on-ramp, an off-ramp, and a weaving segment with a total length of one mile.

A one-lane blockage incident was introduced based on real-world incident data to generate a bottleneck location for the purpose of assessing incident detection timeliness. Since VISSIM cannot model incidents directly, the incident was modeled by creating a signal head in one lane of the freeway and the rubbernecking effect was modeled by changing the car following factors upstream of the incident location utilizing the VISSIM COM interface.

The values of the car following factors for the incident link during the incident conditions were changed using VISSIM COM based on values recommended in the literature. Knoop et al. (2009) investigated driver’s behaviors under incident conditions by analyzing the vehicles trajectories collected by a helicopter from two actual incidents. The results showed that driver's reaction time, headway distribution, and capacity are affected by the incident presence. That study observed a bimodal headway distribution, an increase in reaction time, and a 30% decrease in capacity during incident conditions. Some researchers found that incident impacts can be modeled by changing VISSIM car following factors. A detailed description of the VISSIM car following factors can be found in VISSIM user manual (PTV, 2012). A study of the VISSIM calibration parameters (Woody, 2006) found that the CC0, CC1, CC2, CC4 and CC5 parameters have the most effect on the capacity of a freeway link. CC0 is the desired distance between stopped cars. CC1 is the desired time headway that the driver wants to keep to the leading vehicle. CC2 is the safety distance, which is the minimum distance a driver keeps with the leading vehicle. CC4 and CC5 are the parameters controlling the speed differences between a following and leading vehicle. Increasing CC0, CC1, and the absolute value of the CC4/CC5 ratio result in a capacity reduction. Gomes et al. (2004) modified three CC1, CC2, and CC4/CC5 to model the capacity drops resulting from curvatures. The calibrated parameters for the studied freeway segment are CC1, CC2, and CC4/CC5 were 1.7, 0.9 and -2/2 respectively (Gomes et al., 2004). Rahman and Mattingly (2014) found that the car following values proposed in prior studies for incident modeling produced better results in terms of macroscopic measures compared to the default values. Rompis et al. (2014) developed a methodology for VISSIM calibration based on kinematic queuing theory. The incident was modeled by coding a traffic signal in VISSIM. Based on the reviewed study results, CC0, CC1 and CC2 values of 3.8, 1.5 sec and 7.5 m, respectively, are used to replicate the actual freeway operation under incident condition.

9.5.2. Arterial Street Segment

Travel time estimation is more of a challenge for urban arterials due to lower volumes at times of the day and the high travel time variations, which increase with the increase in the intersection density. The assessment of the estimation of travel time based on CV data is also done for an arterial street: Glades Road, located in the City of Boca Raton, FL. The arterial network consists of three sections. The first section is 0.64 miles long between the Renaissance Way and Airport Road intersection. The second section is 0.76-mile long between Renaissance Way and St. Andrews Boulevard, and the third section is a 1.09-mile long link between East University Drive and Airport Road. This arterial network was modeled in VISSIM and then a manual calibration process is performed based on real-world measurements. The first step was to perform an initial fine-tuning of the model parameters to produce saturation flow rates that are in agreement with the Highway Capacity Manual 2010 procedures (TRB, 2010) and previous observations from the field in South Florida. The target saturation flow rate for the purpose of this calibration was set to 1,850 to 1,900 passenger cars per hour per lane. The parameters of the VISSIM model were fine-tuned to produce this value. The second part of the calibration process was to compare the traffic flow performance according to the Glades Road simulation model with the performance assessed based the real-world traffic data. The VISSIM model parameters that were fine-tuned are the additive part (bx_add) and the multiplication part (bx_mult) of the desired safety distance in the Wideman 74 driver behavior model, which is used for urban arterial street modeling. The most appropriate combination of the two parameters was found to be 2.4 feet and 3.4 respectively. It was determined that the Root-Mean-Square-Percentage-Errors (RMSPE) for the simulated versus measured volume and speed values are below 15% for both directions of travel.

9.6. Results

The proposed methodology was applied to the test networks. The travel time estimation methodology was applied on both the freeway and arterial segments. The incident detection methodology was applied only to the freeway segment.

9.6.1. Travel Time Estimation Results

The Absolute Mean Percentage Error (AMPE) and Standard Deviation Percentage Error (SDPE) of the estimated travel time were calculated for each of the 500 Monte Carlo runs performed for each CV proportion and flow rate combinations. For the arterial street segment, the investigated degrees of saturations were 0.3, 0.6, 0.7, 0.8, and 0.9. For the freeway segment, the investigated v/c ratios were 0.36, 0.5, 0.72, 0.86, and 1.01. For both freeways and arterial streets, the test CV proportions were 5, 10, 20, 50, and 80 percent. The results from the Monte Carlo runs were used to develop the regression equations to allow the calculation of the error percentage at a certain CV proportion and demand level combination. The developed equations and their statistics are provided in Table 9-1.

Table 9-1 Estimated Travel Time Error Equation

Arterial Street								
	AMPE				**SDPE**			
Equation	$E_{AMEP} = 68.79 - 52.9 * \cos(S) - 0.996 * \log(CP) - 27.3 * S$				$E_{SDPE} = 9.6 + 3.04 * \cos(S) - 3.14 * \log(CP) + 0.038 * CP$			
Coefficients Statistics	**Coefficients**	**Std. Error**	**t value**	**Pr(>|t|)**	**Coefficients**	**Std. Error**	**t value**	**Pr(>|t|)**
	β_0	11.55	5.96	6.54e-06	β_0	1.09	8.8	1.72e-08
	β_1	5.62	-4.86	8.41e-05	β_1	1.05	2.88	0.0089
	β_2	10.46	-5.2	3.71e-05	β_2	0.36	-8.62	2.45e-08
	β_3	0.15	-6.46	2.12e-06	β_3	0.013	2.88	0.009
	Multiple R-squared: 0.78				Multiple R-squared: 0.94			
	Adjusted R-squared: 0.75				Adjusted R-squared: 0.93			

Freeway								
	AMPE				SDPE			
Equation	$E_{AMEP} = 12.05 - 9.18*\cos(S) - 0.52*\log(CP) - 4.81*S + 0.0085*CP$				$E_{SDPE} = 9.59 + 3.04*\cos(S) - 3.14*\log(CP) - 0.038*CP$			
Coefficients Statistics	**Coefficients**	**Std. Error**	**t value**	**Pr(>|t|)**	**Coefficients**	**Std. Error**	**t value**	**Pr(>|t|)**
	β_0	3.42	3.56	0.001	β_0	1.6	8.8	1.72e-08
	β_1	2.9	-3.164	0.003	β_1	1.05	2.88	0.0089
	β_2	0.075	-6.95	2.2e-08	β_2	0.36	-8.61	2.45e-08
	β_3	1.81	-2.66	0.01	β_3	0.01	2.88	0.0089
	β_4	0.0085	0.004	0.04				
	Multiple R-squared: 0.8				Multiple R-squared: 0.94			
	Adjusted R-squared: 0.78				Adjusted R-squared: 0.93			

9.6.2. Incident Detection Result

As mentioned in the methodology section, Monte Carlo simulation was done to randomly assign connectivity to the vehicles since the position of the connected vehicles relative to the bottleneck location affects the results. Table 9-2 shows the 80 and 95 percentiles of the resulted MTTD in minutes for each of the scenarios. The result shows that the MTTD decreases as the V/C ratio and CV proportions (CP) increase. It should be noted that the equations in Table 9-2 are applicable to one lane blockage incidents on four lane freeways. The equations are also applicable when the v/c ratio is greater than 0.4. This is because the incident is not detectable based on speed for v/c less than 0.4, as there is no queue formed.

9.7. Quality of CV Utilization for Different Future Years

This section discusses the application of the models developed, as discussed earlier in this report, to determine the quality of CV-based data in terms of travel time measurement accuracy (in terms of AMPE), travel time measurement reliability (in terms of SDPE), and incident detection latency (in terms of MTTD) for different years in the future. Such analysis is critical for agencies to make investments in CV versus another mode of data acquisition in future years. The CV proportion in the traffic stream will increase as new vehicles join the fleet every year. Different studies have predicted the future CV proportion after the CV technology mandate by the USDOT becomes effective. In this study, the prediction model developed by Iqbal et al. (2017b) was used to determine the accuracy of performance measurement at different demand levels and at different years after the CV mandate becomes effective.

Table 9-2 MTTD for Different CV proportions and Different V/C Ratios

	Mean Time to Detect (MTTD) (80%)			
Equation	$MTTD = \begin{cases} 0.75 & CP > 6\% \\ -51.649 + 54.563 * e^{-(v/c)*CP} & CP \le 6\% \end{cases}$			
Coefficients Statistics	**Coefficients**	**Std. Error**	**t value**	**Pr(>|t|)**
	β_0	7.447	-6.936	0.000
	β_1	7.607	7.173	0.000
	Multiple R-squared: 0.798			
	Adjusted R-squared: 0.783			
	Mean Time to Detect (MTTD) (95%)			
Equation	$MTTD = \begin{cases} 0.75 & CP > 6\% \\ -102.892 + 108.192 * e^{-(v/c)*CP} & CP \le 6\% \end{cases}$			
Coefficients Statistics	**Coefficients**	**Std. Error**	**t value**	**Pr(>|t|)**
	β_0	20.864	-4.931	0.000
	β_1	21.313	5.076	0.000
	Multiple R-squared: 0.665			
	Adjusted R-squared: 0.639			

The results of the analysis are shown in Table 9-3. Previous guidelines and practices for acceptable error of travel time point to an AMPE of less than 10% absolute (Turner et al., 1998; Fries et al., 2015). Toppen and Wunderlich (2004) reported that, if the SDPE is greater than 12%, then there are no user benefits when using the travel time estimates in advanced traveler information systems. Following those guidelines, Table 9-3 shows that the travel time data estimation on freeways starting the second year after the CV is mandated. For the arterial facility, it will take four to five years to get an acceptable accuracy of travel time estimation. The accuracy of the travel time is also affected by the v/c ratio. At lower v/c ratio, the accuracy is lower because of the smaller sample size. The accuracy increases with the increase of the sample size up to a certain point and then it starts to decrease. At higher v/c ratios, there are higher variations in travel time between vehicles due to congestion, which results in a lower accuracy in the estimated travel time. CV data could be used for incident detection from the very first year of CV implementation on freeways with a latency less than 5 minutes. Martin et al. (2009) identified the detection latency to be between 0.85 to 4 minutes. Therefore, the existing detectors could be completely replaced by the CV technology for the purpose of incident detection at year 4.

Table 9-3 Minimum Accuracy at Different Year after the CV Mandate Becomes Effective

v/c	Year	Travel Time Accuracy (%)				Incident Detection Time (Minutes)	
		Freeway		Arterial		Freeway	
		AMPE	SDPE	AMPE	SDPE	80% MTTD	95% MTTD
0.1	1	2.66	13.2	13.60	13.2	Accident cannot not be detected	
	2	2.07	9.64	12.45	9.64		
	3	1.82	8.1	11.94	8.09		
	4	1.46	5.86	11.16	5.84		
	5	1.30	4.89	10.80	4.87		
	10	0.95	2.47	9.70	2.45		
	15	0.91	1.92	9.14	1.88		
	20	0.94	1.91	8.95	1.87		
	25	0.95	1.93	8.89	1.89		
0.5	1	1.81	12.84	8.89	12.86	2.94	4.76
	2	1.23	9.28	7.74	9.29	2.19	3.86
	3	0.97	7.74	7.23	7.75	1.71	2.91
	4	0.61	5.5	6.45	5.5	0.75	0.75
	5	0.46	4.53	6.10	4.53	0.75	0.75
	10	0.11	2.12	4.99	2.11	0.75	0.75
	15	0.07	1.56	4.43	1.54	0.75	0.75
	20	0.09	1.55	4.24	1.53	0.75	0.75
	25	0.11	1.57	4.18	1.55	0.75	0.75
1	1	2.5	11.82	13.08	11.85	2.37	4.22
	2	1.92	8.26	11.93	8.29	1.47	2.45
	3	1.66	6.72	11.42	6.74	0.75	0.75
	4	1.3	4.47	10.65	4.5	0.75	0.75
	5	1.15	3.50	10.29	3.52	0.75	0.75
	10	0.8	1.09	9.18	1.1	0.75	0.75
	15	0.76	0.53	8.62	0.54	0.75	0.75
	20	0.78	0.53	8.43	0.52	0.75	0.75
	25	0.8	0.55	8.37	0.54	0.75	0.75
	Error more than the minimum acceptable threshold						

9.8. Summary

This study investigated the use of CV data as an alternative to existing data acquisition techniques in providing two critical functions to support transportation system management and operations: travel time estimation and incident detection. In support of this investigation, the study examined the measurement accuracy and measurement reliability of the estimated travel time and the latency of incident detection under different demands and CV proportion in the traffic stream. Regression models were developed to estimate measures of travel time accuracy (AMPE), reliability (SDPE), and latency (MTTD). Significant relationships were found between these measures and the demand level and CV proportions. The developed regression models were used in conjunction with a prediction of CV proportion in future years to estimate the accuracy, reliability, and latency measures for each future year after the CV USDOT mandate becomes effective. The results show that the CV data can be used for freeway travel time estimation from the very first year of the CV implementation. For arterials, it may take four to five years to reach an acceptable accuracy. The use of CV data for incident detection on freeways will provide a latency of 5 minutes or less for the first year after implementation when the V/C ratio is higher than 0.4. The latency is expected to be lower than the existing detectors for all incidents four years after the CV mandate becomes effective.

Results of this study can be used by transportation system management and operations (TSM&O) programs and agencies to plan their investment in data collection alternatives. They can be used to support the preparation of a timeline to replace the existing data collection systems with the CV technology, depending on the traffic demand of a specific roadway. This should be done, however, with the consideration that some point detectors may still be necessary to provide other parameters such as traffic volume.

10. IDENTIFIYING A TIMELINE FOR FUTURE UTILIZATION OF CONNECTED VEHICLE DATA TO SUPPORT TRAFFIC VOLUME ESTIMATION ON URBAN STREETS

10.1. Introduction

Transportation agencies have increasingly invested in traffic monitoring systems to support their planning, planning for operations, and operations and management. These systems have allowed the estimation of critical measures such as volume, speed, travel time, travel time reliability, delays, and queue lengths. These measures are currently obtained utilizing various detection technologies and from third party data vendors. The introduction of connected vehicles (CV) in the coming years will provide an important source of data to complement, and in some cases, replace existing traffic detection technologies. The National Highway Traffic Safety Administration (NHTSA) has published an advance notice of proposed rulemaking on Vehicle-to-Vehicle (V2V) communications utilizing connected vehicles (CV) (Zhan et al., 2017). It is expected that NHTSA will mandate CV technologies on all new vehicles, allowing a significant increase in CV market penetration in the coming years.

It will be possible to estimate some of the performance measures utilizing data from a relatively low CV market penetration. These include measures such as speed, travel time, delay, and number of stops (Argote et al, 2012; Iqbal et al., 2017a; Garelli et al., 2011; Doan et al., 2009). However, a high CV proportion in the traffic stream is expected to be needed to get an accurate estimate of parameters such as traffic volume and density (Zhou et al., 2015; Khan, 2015). Traffic detectors count every vehicle to allow the estimation of traffic volumes. When not all vehicles are equipped with CV technology, a method will be needed to estimate the volumes based on the partial vehicle counts, provided by CV. If such a method is developed and proven effective, it will be possible to use CV data to support the estimation of volumes on links with no detectors based on volumes measured with existing detectors and eventually removing at least some of the existing system detectors. The objective of this study is to develop a method to determine the approximate time in the future when connected vehicle data can replace or complement existing detectors on urban streets in estimating segment traffic volumes.

10.2. Literature review

Roadway traffic volume is a critical input for transportation system planning, planning and operation, and operation processes. Transportation agencies have collected volume data using traffic sensor technologies such as inductive loops, microwave, video image processing, infrared, and magnetic detectors, and more recently video analytic products utilizing CCTV camera images. Apart from direct measurements, researchers have investigated deriving traffic volume based on partial volume counts available from various data sources. Demissie et al. (2013) used cellular networks handover data to estimate the volume. They used available citywide traffic counter data combined with the cellular data to estimate traffic volumes based on a developed regression equation. The model evaluation showed a mean absolute percentage error (MAPE) of 46.8%. Caceres et al. (2012) also performed a similar study and found that their method produces a mean absolute percentage error (MAPE) less than 17%. Nantes et al. (2013) estimated the traffic volume using Bluetooth data. They proposed a Bayesian network to estimate the volume.

Vehicle trajectory data collected utilizing GPS data has also been also used to estimate the traffic volume (Zhan et al., 2017; Anuar and Cetin, 2017). Zhan et al. (2017) used a hybrid framework that combines machine learning techniques and traffic flow theory to estimate traffic volumes. They evaluated the proposed methodology utilizing GPS dataset from 33,000 Beijing taxis and volume ground truth data obtained from 4980 video clips. Anuar and Cetin (2017) used probe vehicle trajectory data and a machine learning technique had been combined with shockwave theory to estimate the volumes and found an average error of 5% in the volume estimation when using this method.

There are a number of studies conducted to help in selecting optimal sensor locations for volume measurements (Bao et al., 2016; Ban et al., 2009a; Ivanchev et al., 2016; Ye and Wen, 2017; Shao et al., 2016; Mitsakis et al., 2017; Fei et al., 2007) for different purposes and utilizing different algorithms. Bao et al. (2016) optimized sensor locations for traffic network using different special distributions of traffic information. Shao et al. (2016) proposed an optimization methodology to find the minimum number and optimal locations of the sensors based on turning ratios at the intersections. Mitsakis et al. (2017) proposed a quadratic programming model to determine the optimal locations of traffic sensors. They implemented the model in the urban road network of the city of Thessaloniki (Greece). The result showed that the optimized sensor locations could represent 87% of the traffic flow along the major paths.

Studies have been also done to identify the accuracy of different detection technologies to collect traffic volume data. Point detectors like the inductive loop, video image detectors, and microwave detectors were found to produce acceptable volume count accuracy, although they are subject to errors, particularly during congested conditions when the proximity of vehicles to each other can result in counting more than one vehicle as one vehicle. A Minnesota Department of Transportation study (Minge et al., 2010) found that four tested non-intrusive detection technology products produced a volume accuracy comparable to loops (typically within 1.6 percent), during both free-flow and congested conditions. However, a per-vehicle analysis revealed some occlusion when slow moving trucks in the lane nearest to the sensor blocked subsequent lanes, resulting in undercounting of about 20% in the occluded lanes in periods of heavy congestion and short counting intervals. This is expected to be a function of the number of lanes and trucks on the freeway. A study in Nebraska (Grone, 2012) found an error in a one-minute traffic count ranging from 5.5% to 8.2% for four widely used non-intrusive point detectors. However, this error dropped at higher aggregation levels (5 minutes or 15 minutes). Nihan et al. (2002) found an error of just 1-3% in volume measurements using loop detectors when aggregated at the 60-minute levels. However, when examined at the 20-second level, 22.1% of the intervals had incorrect values. A study of loop detectors in Arizona showed an average error in 5-minute counts of 3% to 6%, with an error range of 1-20% (Samuelson, 2011).

With regard to video analytics based on existing CCTV cameras, an ENTERPRISE pool funded study (Preisen and Deeter, 2014) found average traffic volume errors of 9% during daytime conditions and 17% for nighttime conditions. The study found a 14% average error for the AM peak period and 9% average error for the PM peak period.

Zheng and Liu (2017) proposed a methodology to estimate traffic volumes utilizing data from a small number of CV with high-resolution signal controller data. Their results showed that the MAPE is 9-12% compared to manual and detector data. The main drawback of this study is that it is only applicable to signalized intersections with high-resolution signal data availability.

10.3. Estimation of Connected Vehicle Market Penetration

The derivation of the timeline for the usability of CV data for traffic volume estimation requires the estimation of CV proportions in future years. The United States Department of Transportation (USDOT) (2008) predicted the probable market penetrations of CV in future years to estimate the benefits and costs of CV implementations. In the estimation, the analysts considered the scenario where only new vehicles will use the CV technology with the assumption that in the first year 25%, second year 50%, third year 75% and afterward 100% of the new vehicles will have the connectivity. Wright et al. (2014) suggested three different scenarios for probable CV implementations. The most conservative scenario among the three is called the "15-year organic" scenario, which assumes that the CV will come into the fleet as organic sales of the new capability. The moderate one is called the "5-year mandate" scenario, in which manufacturer would include OBUs into the new vehicles over a five-year period. The best-case scenario is the "1-year mandate" scenario where all the new vehicles will be equipped with OBU starting from the year that the CV is mandated. Iqbal et al. (2017b) assumed the "1-year mandate" scenario and considered the variation in the socioeconomic characteristics between different zones to determine the future CV proportions of vehicles generated from different zones and the resulting roadway segment level variations in CV proportions. Rather than providing a constant CV proportion in a future year, in this study, we assume a range of CV proportion for each future year when estimating volume accuracy based on the results from the Iqbal et al. (2017b) study.

10.4. Methodology

As stated earlier, the objective of this study is to determine the approximate time when connected vehicle data could replace or at least partially replace existing midblock detectors to estimate traffic volumes on urban streets. With the increase of the CV market penetration, at some point in time, there will be sufficient CV data that allows; the removal of one or more of the midblock detectors on an arterial street. This can be achieved by estimating the volumes for the segment, from which the detector is removed, as a function of detector measurements installed at other locations, combined with partial volume data collected utilizing the available sample size of CV. Eventually, at high market penetrations, it may be possible to remove most, if not all, of the detectors.

In this study, three different possible scenarios are considered to potentially occur in future years, with the increase in CV market penetrations, as follows:

- In the first scenario (Scenario 1), the detectors on arterial links are kept on a subset of the links on the urban arterial, as would be done in current applications. However, an investigation is done to determine if utilizing partial volume counts based on CV data can improve the estimation of volumes on the links with no detectors (non-instrumented links). The traffic volumes on non-instrumented links are calculated utilizing a factor obtained based on CV data and the volumes measured by detectors at an adjacent instrumented link, that has the lowest percentage difference in volume measurements compared to the non-instrumented link volumes.

- The second scenario (Scenario 2) involves removing some of the detectors in Scenario 1. This will result in additional non-instrumented links that have less correlation between their volumes and the volumes of the instrumented links. Regression analysis is derived and used to estimate the volumes for these non-instrumented links based on the volumes of the instrumented links and CV data.

- Scenario 3 involves estimating the traffic volumes on the arterial links utilizing only the CV data without permanently instrumenting any link on the arterial segment with detectors.

Scenario 1 is expected to be applicable to the early stages of the CV introduction, with low CV proportions in the traffic stream. With the increase of CV proportions, Scenario 2 and then Scenario 3 may become feasible. Below is a description of the base scenario and the three scenarios listed above.

10.4.1. Base (Existing) Scenario

Due to a limited budget, it is not possible to place detectors at every possible location of a highway corridor. Different studies have been done to find out the optimal locations of a segment to install the detectors, depending on the variability of the traffic flow, geometric condition, and network complexity; as indicated in the review of literature section. Roess et al. (2011) described a recommended practice that identifies links not to be instrumented with detectors when they have less than 10% mean absolute difference in volume measurements compared to the measurements at an instrumented link. This study follows this concept in selecting the initial list of links for instrumentation to represent the base (existing) scenario based on the differences in the volume measurements between links. For this purpose, three measurements are used to check the variability in the volumes between the locations, as follows:

- The Mean Absolute Percentage Deviation (MAPD) is calculated taking the average of absolute percentage difference of all the measurements at two links. The equation of the MAPD is presented below.

$$MAPD = \frac{1}{n}\sum_{i=1}^{n}\frac{|x_i - y_i|}{x_i} \tag{10-1}$$

  where, x_i and y_i are the volumes at a time interval (observation) i for Location X and Y, respectively, and n is the total observation.

- 95% Absolute Percentage Deviation (95APD) is the 95% of the absolute percentage difference between x_i and y_i.

- Sample correlation coefficient (rxy) is calculated using Equation 10-2.

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{10-2}$$

  where, $\bar{x}$ and $\bar{y}$ are the mean of all the observations at location X and location Y, respectively.

10.4.2. Future Scenario 1

With this scenario, the volumes at the non-instrumented locations identified in the Base Scenario are estimated by applying an expansion factor to the partial CV volume counts. The expansion factor is derived based on measurements at an adjacent detector location. The variabilities between the volume patterns at the instrumented and non-instrumented locations are low since this variability is used as a criterion for selecting links for instrumentation in setting the Base Scenario, discussed in the previous section. Therefore, it is expected that the ratio of the actual volume and the CV volume will be close for the instrument location and the associated non-instrumented locations. An expansion factor ($C_{f,t}$) is calculated as the ratio of the actual volume ($V_{i,t}$) and the CV volume ($CV_{i,t}$) at the instrumented link location (location i) at a certain time interval (time interval t). The expansion factor ($C_{f,t}$) is then used to expand the CV volumes at a non-instrumented location (location j) that has low variability in volume compared with the instrumented location, to find out the actual volume ($V_{j,t}$) based on the CV volume ($CV_{j,t}$) at that location at the same time interval (time interval t). This calculation can be expressed by Equation 10-3 and Equation 10-4 below.

$$C_{f,t} = \frac{V_{i,t}}{CV_{i,t}} \tag{10-3}$$

$$V_{j,t} = C_{f,t} * CV_{j,t} \tag{10-4}$$

10.4.3. Future Scenario 2

In the Base Scenario and Future Scenario 1, the links are selected to be not instrumented with detectors utilizing a criterion that the MAPD between the non-instrumented link and an adjacent instrumented link is below a certain value. In Scenario 2, this criterion is relaxed, allowing higher variability between the volumes of the instrumented and non-instrumented links. The idea is that with the availability of partial counts from CV, it will be possible to estimates the volumes on the non-instrumented links even with this higher variability. Two methods are explored for this estimation, as described below.

Method 1

The first method is to perform counts using temporary detectors at the non-instrumented locations to allow the derivation of regression equations to estimate the volumes at the non-instrumented location, as a function of the volume data measured using detectors installed at other instrumented locations and partial counts obtained using CV data of all locations. Equation 10-5 represents a general form of the derived regression equation.

$$V_i^{est} = \alpha + \sum_{j=1}^{n} \beta_j * CV_j + \sum_{j=1}^{n-1} \gamma_k * V \tag{10-5}$$

where, V_i^{est} is the estimated volume at location i; α, β, γ are the regression coefficients; CV_j and V_j are the CV volume and the actual volume at location j, n is the total number of locations in the segment. Temporary detectors could be placed at the study locations for a short period of time to recalibrate the regression equations, possibly every few years, to capture changes in traffic patterns.

Method 2

There is a cost associated with conducting the temporary counts required in Method 1. Therefore, a second method is explored to estimate the volumes based only on the available permanent detector data throughout the segment and the partial CV counts, without the need for temporary detectors. An optimization procedure is used to find a generalized expansion factor for the whole segment. This expansion factor is then used to calculate the volumes for the non-instrumented locations. The objective function (f) of this optimization problem is to minimize the sum of the difference between the estimated volume ($V_{i,t}^{est}$) and the actual volume ($V_{i,t}^{act}$) at all the detector locations (n) at a certain time interval t (there is a total of p intervals). The objective function in the optimization is presented in Equation 10-6. The optimization problem is solved using the Bisection method utilizing the R software.

$$\text{Minimize, } f = \sum_{i=1,t=1}^{n,p} \left| V_{i,t}^{est} - V_{i,t}^{act} \right| \tag{10-6}$$

$$\text{Such that, } V_{i,t}^{est} = \alpha * CV_{i,t}$$

10.4.4. Future Scenario 3

Scenario 3 represents the ultimate future scenario, in which all permanent detectors are removed and the volume estimation is solely made based on CV data. This scenario may become feasible at high CV market penetrations, which will occur several years in the future. Expansion factors could be applied based on the estimated national or regional CV market penetrations or may be calculated based on data from temporary detectors installed for a short period of time. In this scenario, the calculation of the expansion factor can be done according to Equations 10-7 and 10-8, in which an average expansion factor (C_f) is calculated from the temporary detector counts and the partial counts based on CV at a study location over a time period of ‘p’ intervals. Different expansion factors could be developed for different times of the day to improve the results, considering that the CV proportion in the traffic stream may vary by time of day depending on the socioeconomic characteristics of drivers at different times of the day.

$$Cf = \frac{1}{t}\left(\sum_{t=1}^{p} \frac{V_t^{temp}}{CV_t} \right) \tag{10-7}$$

$$V = Cf * CV \tag{10-8}$$

If the regional/national average market penetration of the connected vehicles for the year under consideration is used to calculate the expansion factor, then no temporary detector is needed to calculate the expansion factor (Cf) and Equation 10-8 could be directly used to estimate the volumes. This scenario is expected to produce less accurate results compared to what is possible with the other scenarios. However, it requires less investment in instrumenting a subset of the network with detectors.

10.4.5. Accuracy Measurements

In all of the above-mentioned scenarios, the accuracy of the estimation is assessed in this study by calculating the absolute percentage error (Err) of the estimated volume (Vest) compared to the actual volume (Vact), using Equation 10-9.

$$E_{rr} = \frac{\left|V^{act} - V^{est}\right|}{V^{act}} * 100 \tag{10-9}$$

This study determines the accuracy of the estimated traffic volumes with the different scenarios and associated methods mentioned above, for each year of the CV implementation. An acceptable error percentage in volume estimation should be selected based on the agency policy or existing guidelines such as those presented in Reference (FHWA, 2014). This selection is used to develop a timeline of when an agency can remove at least some of the detectors as a result of using the CV data when estimating traffic volumes. It should be mentioned that different agencies may set different volume accuracy requirements for different applications. For example, many agencies set this accuracy at 5-10% for different applications. Different accuracy requirements will result in different timelines for the usability of CV data for volume estimation.

10.5. Test Location

This study utilizes a segment of Glades Road, a major urban arterial located in the City of Boca Raton, Florida to demonstrate the accuracy of the volume estimation with the scenarios and methods, described earlier in this section. The segment is one mile long with five intersections. The network was coded in the VISSIM microscopic simulation model and the model was calibrated based on real-world measurements in terms of saturation flow rate, volume, and travel times. A general configuration of the location is provided in Figure 10-1. The circles represent the intersection locations. There are five potential locations for midblock detectors which are represented by numbers (1-5), in Figure 10-1. The actual locations for detection in the Base Scenario were selected in this study, as described in the methodology section. The simulation was done for different congestion levels. The eastbound traffic is considered for demonstrating the proposed methodology.

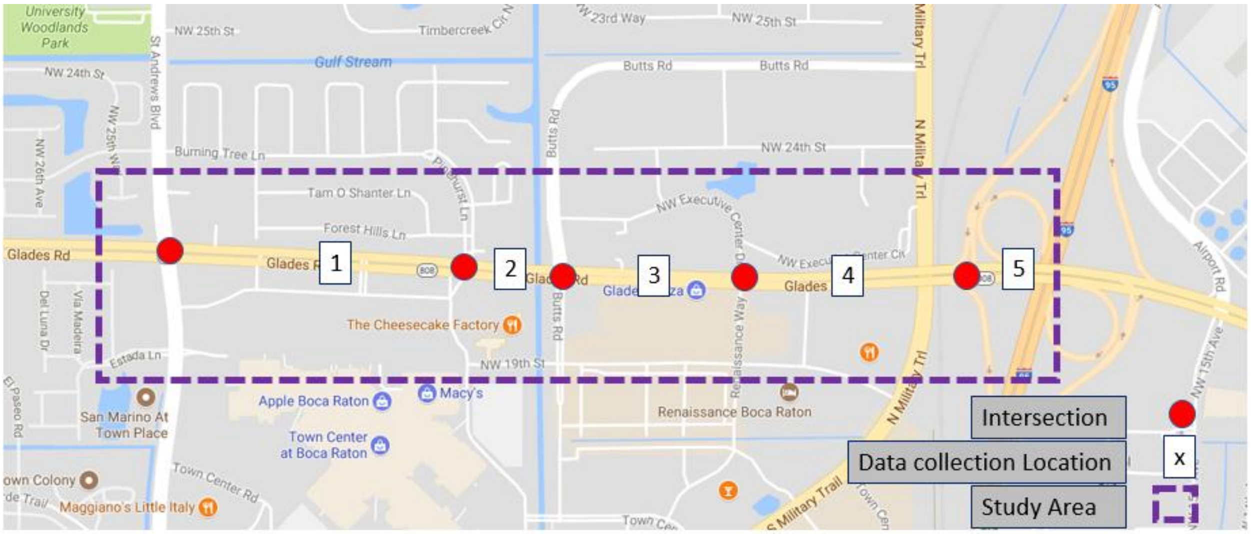


Figure 10-1 The Glades Road Study Location in Boca Raton, Florida

10.6. Emulation of CV Data

The connected vehicles will communicate with other vehicles and the infrastructure using a standardized message set that are defined by the Society of Automotive Engineers (SAE) J2735 standards (SAE, 2016). The J2735 standards specify a number of message types. The basic safety message (BSM) is one of these message types that will be used for vehicle-to-vehicle communications using dedicated short range communications (DSRC). The BSM contains vehicle safety-related information broadcasted to surrounding vehicles but can be also captured by the infrastructure utilizing roadside units (RSUs). The BSM, as defined in the J2735 standards, consists of two parts. Part 1 data elements are sent in every BSM message broadcasted at 10 times per second rate. BSM Part 1 is expected to be mandated by the United States Department of Transportation (USDOT) on all new vehicles. It contains core data elements, including vehicle position, heading, speed, acceleration, steering wheel angle, and vehicle size. BSM Part 2 consists of a large set of optional elements such as precipitation, air temperature, wiper status, light status, road coefficient of friction, Antilock Brake System (ABS) activation, Traction Control System (TCS) activation, and vehicle type. BSM Part 2 elements are sent based on criteria that are not specified in the J2735 standards. However, not all of these parameters are currently available from vehicles and will not be mandated by the USDOT. The Basic Safety Message (BSM) part I is considered for use in this study.

CV data is starting to be available from few test beds and pilot deployments. However, the CV proportions at these locations are very limited. Therefore, this study has utilized the Trajectory Conversion Algorithm (TCA) [Version 2.3] (OSADP, 2015) developed by the Federal Highway Administration (FHWA) for emulating the BSM data following the SAE J2735 standards based on vehicle trajectories generated from simulation modeling. The mentioned arterial segment is analyzed in VISSIM to generate the vehicle trajectories at different demand levels. These trajectory files are used as an input to the TCA tool to emulate connected vehicle BSM data collected at 1/10th second. The input trajectory file to the TCA should contain vehicle ID, time, speed, position coordinates, and acceleration data of the vehicles. The user of the TCA tool can select the market penetration, CV message type (in this case, BSM) and the communication type (DSRC or cellular). The user can also specify the RSU locations. The tool allows the specification of data transmission loss during the data transfer between the Onboard Units (OBU) and the RSU. It has been reported that the loss rate with DSRC communications varies between 10% and 20% with an average of 12% (Kandarpa et al., 2009). In this study, 12% transmission loss is specified as input to the TCA tool. It is expected that the actual loss is a function of the availability of a line of sight in the coverage area and this should be considered with RSU setting.

There are two sources of variation that could affect the volume estimation accuracy. The first one is the variation of the CV proportion at a different time of the day and different segments in the network. Iqbal et al. (2017b) showed that the CV proportion at a certain future year follows a lognormal distribution, as discussed further in the next section. The second source of variation is the specific vehicles in the network that are assumed to be connected. Both of these factors affect the accuracy of the volume estimation. Thus, instead of assuming these variables as deterministic variables, they are assumed to be stochastic and randomly generated from distributions in a large number of Monte Carlo analysis runs (100 runs).

The investigation is done for five different demand levels on the arterial (degree of saturation of 0.25, 0.5, 0.75, 1.0, and 1.2). The CV data is generated based on the CV proportions estimated for 13 different future year CV (yearly between Year 1 to Year 10, Year 15, Year 20, and Year 25). For each demand level and future year CV proportion, hundreds of Monte Carlo runs are performed to account for the stochasticity mentioned earlier, resulting in a total 6500 datasets to analyze.

10.7. Results

The scenarios and associated methods of this study were examined to determine their applicability in different future years. The results are presented below.

10.7.1. Base Scenario

The first step is to select links for instrumentation in the Base Scenario, by eliminating links that have low MAPD and 95%APD values and high correlations in traffic volumes, compared to other links that will be instrumented with detectors. The MAPD and 95%APD values between each pair of the five links of the case study are presented in Table 10-1.

Table 10-1 MAPD (and 95%APD) Variation of Actual Volume in Percentage

Location	1	2	3	4	5
1	0	***4.68(10.06)***	10.09(24.49)	14.80(28.79)	15.37(28.88)
2	***4.51(9.55)***	0	10.71(27.8)	12.74(24.95)	14.70(25.71)
3	8.87(19.16)	9.21(21.74)	0	17.16(29.98)	17.17(27.57)
4	18.41(40.5)	15.38(33.26)	21.72(42.86)	0	***4.84(10.87)***
5	19.17(40.82)	17.31(34.62)	22.00(38.22)	***4.86(11.21)***	0

Table 10-1 shows that there is a relatively low variation between the volume counts of Location 1 and Location 2 and between the volume counts of Location 4 and Location 5. In both cases, the MAPD is less than 5% and the 95%APD is below 10.87% for these locations. The correlation analysis also shows that these two pairs have correlation coefficients greater than 98%. Considering this, it was decided not to place detectors at Location 2 and Location 4. In this case, the traffic measurements at Location 1 is expected to be able to reflect the traffic pattern at Location 2 with an MAPD of 4.51% and 95%APD of 9.55%. Location 5 is expected to be able to reflect the traffic pattern at Location 4 with an MAPD of 4.84% and 95%APD of 10.87%, based on the results presented in Table 10-1.

10.7.2. Future Scenario 1

Given the above discussion regarding the Base Scenario, this study investigates when partial counts using CV data will be able to improve on the above-stated accuracy for the Base Scenario, when calculating the volumes at Location 2 and Location 4, according to the Scenario 1 methodology. The resulting MAPD and 95%APD are presented in this document for estimating the volume at Location 2 utilizing an expansion factor calculated based on Location 1 volume data. As can be seen in Table 10-2, it will take 4 years after the CV mandate on all new vehicles become effective for the CV data to become beneficial in supplementing existing detectors. After 4 years, using CV data will allow providing volume estimates on the non-instrumented vehicles that are more accurate than those based on existing data alone (MAPD = 4.51% and 95%APD = 9.55). The errors decreased significantly with the increase in CV market penetration in future years.

Table 10-2 Volume Estimation Error with Different Scenarios and Associated Methods

Year after CV Mandate is Effective	Mean CV Proportion (Iqbal et al. (2017b))	Scenario 1		Scenario 2 (Method 1)		Scenario 2 (Method 2)		Scenario 3 (Temporary Detector)		Scenario 3 (National Average)	
		MAPD	95%APD	MAPD	95%APD	MAPD	95%APD	MAPD	95%APD	MAPD	95%APD
1	3.2	12.48	36.08	18.45	42.24	36.02	77.09	102.1	336.6	57.15	187.3
2	7.89	7.14	18.91	15.30	34.53	24.57	48.04	31.11	83.38	28.29	67.94
3	12.84	5.44	14.62	9.90	25.29	12.11	22.60	26.25	60.63	25.87	60.57
4	19.13	4.26	10.90	10.10	24.87	10.27	19.74	19.31	47.14	19.49	47.23
5	24.89	3.58	9.06	8.94	18.23	8.89	15.89	17.26	41.54	16.93	40.53
6	30.34	3.38	8.40	5.48	12.81	7.25	14.79	17.57	42.83	14.92	35.59
7	36.78	2.58	6.70	5.77	15.02	6.83	13.24	13.09	33.46	12.55	32.54
8	42.74	2.47	6.70	6.03	15.02	5.01	10.69	11.87	30.22	11.88	30.42
9	48.87	2.19	5.73	5.71	13.31	4.25	7.16	10.34	24.66	10.37	25.02
10	54.54	1.81	4.52	9.35	23.32	2.97	6.68	9.46	23.31	9.16	21.86
15	80.34	1.08	2.72	5.12	10.84	2.55	5.09	3.96	9.08	4.20	9.67
20	92.07	0.76	2.00	1.40	3.42	1.98	4.27	2.93	6.52	3.23	6.94
25	96	0.62	1.72	3.50	6.75	1.62	3.89	3.55	6.77	3.15	6.35

10.7.3. Future Scenario 2

Scenario 1 involves installing permanent detectors at Locations 1, 3, and 5. In Scenario 2, the study explores removing one or two of these three detectors, by taking advantage of the availability of additional information from CV counts. As described in the methodology sections, two methods are used for the utilization of the partial volume counts from CV in the estimation: a regression analysis method (Method 1) and an optimization method (Method 2). For the purpose of demonstration, the detector at Location 3 is removed and the two methods associated with Scenario 2 are applied. In the regression method (Method 1), a temporary detector is assumed to be placed at Location 3 over a short period to allow the development of the regression model. Table 10-3 shows the regression model derived for Year 1 and Year 5. The developed regression model has a high R-squared value, indicating that the model has the potential to be used to estimate the volumes at the non-instrumented locations.

The main drawback of Method 1 is that it needs a temporary detector to develop the model. Furthermore, the model would need to be updated at a regular interval to reflect changes in traffic patterns. This would increase the cost of the data collection. To overcome this drawback of Method 1, the optimization method (Method 2) was used to calculate the volumes at Location 3, without the need for additional detectors. The results from both methods are presented in

Tables 10-2. As can be seen from the table, utilizing both methods to estimate the volumes for Location 3 results in MAPD of less than 10% in 4-5 years and Method 2 can produce an MAPD of less than 5% in 8 years. This indicates that removing some of the existing detectors will be possible in 5-8 years. Method 2 that does not require temporary counts actually produced better results than Method 1 in future years (beyond Year 7) that have increased CV market penetration, although Method 1 performed better in the first few years. Method 1 performance could have been improved for future years, if the regression model would have been updated in Year 7 for example, to capture the change in traffic patterns in future years.

Table 10-3 Regression Equations Developed for Estimating Volume at Location 3

Year 1				
Equation	$V_3^{est} = 0.071 + 0.469 * V_1^{act} + 0.447 * V_5^{act} - 10.316 * CV_2 + 9.323 * CV_3$			
	where, V_1^{act}, V_5^{act} are the actual volume at location 1 and 5 CV_2, CV_3, are the CV volume at location 2 and 3			
Coefficients Statistics	Coefficients	Std. Error	t value	Pr(>|t|)
	β_0	10.592	0.007	0.99477
	β_1	0.14247	3.291	0.00814
	β_2	0.09052	4.937	0.00059
	β_3	2.2952	-4.495	0.00115
	β_4	2.0738	4.496	0.00115
	Multiple R-squared: 0.9813			
	Adjusted R-squared: 0.9738			
Year 5				
Equation	$V_3^{est} = 11.6647 + 0.741 * V_5^{act} + 3.6684 * CV_3 - 2.9492 * CV_5$			
	where, $V_5^{act}$are the actual volume at 5 CV_3, CV_5, are the CV volume at location 3 and 5			
Coefficients Statistics	Coefficients	Std. Error	t value	Pr(>|t|)
	β_0	8.7386	1.335	0.209
	β_1	0.1424	5.205	0.000292
	β_2	0.6614	5.547	0.000174
	β_4	0.6554	-4.5	0.000902
	Multiple R-squared: 0.9763			
	Adjusted R-squared: 0.9698			

10.7.4. Future Scenario 3

The last scenario (Scenario 3) involves removing all detectors. At the early stage of this scenario, the volume estimation of a location could be supported by temporary detectors to calculate the expansion factor as mentioned in the methodology section. The accuracy of this scenario is demonstrated for Location 3, as shown in Tables 10-2. The results in the table indicate that to remove the last detector in the test segment, it will take 10 years to achieve a 10% accuracy of volume estimation and 15 years to achieve 5% accuracy.

10.8. Identification of Detector Removal Timeline

The results obtained based on the analysis of the previous section can be used to prepare a probable timeline for the detector removal process for a specific roadway segment. Based on the CV proportion in future years and considering the assessment of the three scenarios and associated methods, presented in the previous section, the timeline for utilizing CV data to support traffic volume estimation for the study segment are shown in Figures 10-2(a) and 10-2(b), for 5% and 10% volume estimation accuracy requirement, respectively.

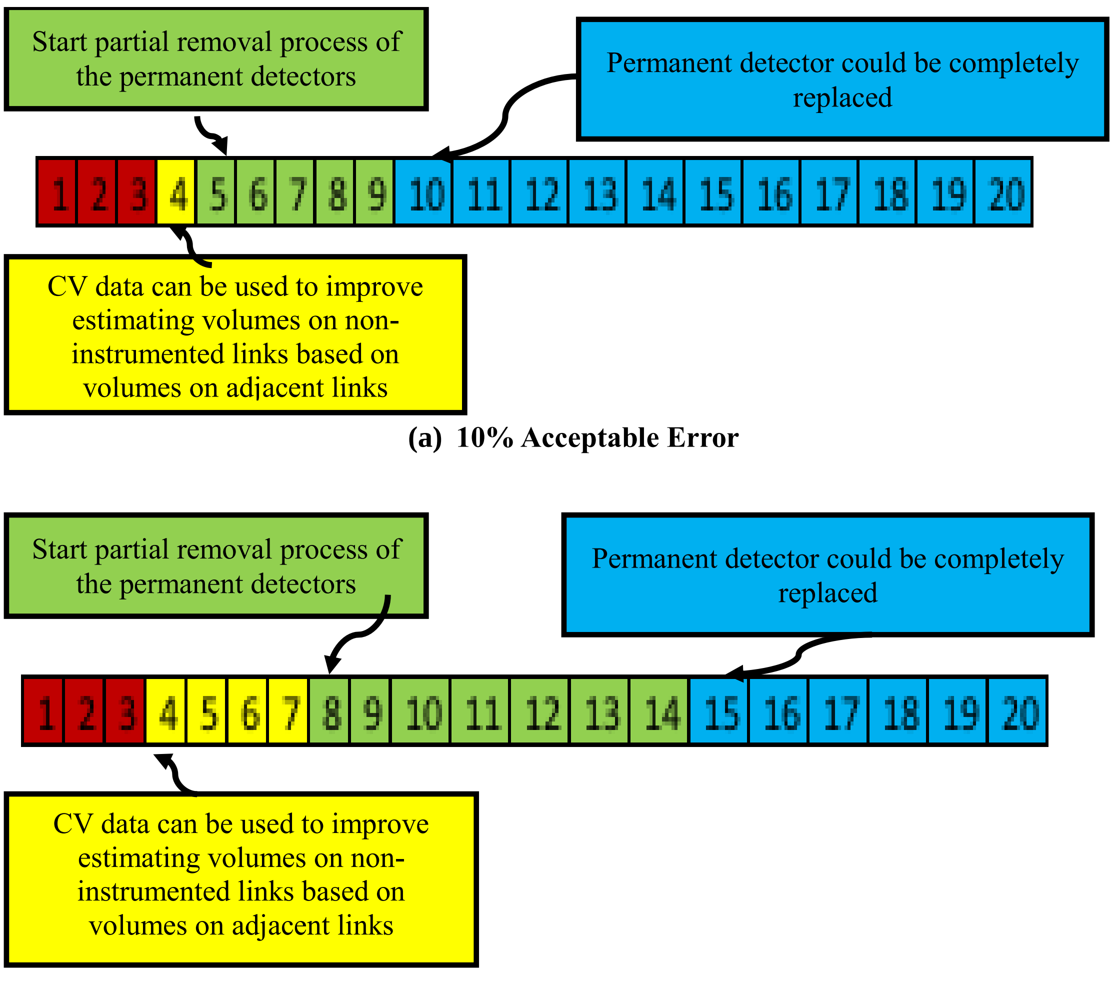


Figure 10-2 Timeline for Detector Removal from the Study Location

10.9. Summary

This study confirms that CV data can be used as a potential source to supplement, partially replace, or completely replace existing traffic detectors to estimate link volumes on urban arterial streets, depending on the CV market proportion in future years. The results from applying the methodology to a case study indicate that after four years of the mandate of CV of new vehicles, CV data can be used to improve the estimation of volumes on the street links with no detectors without removing the existing detectors on the other links. Depending on the adopted volume accuracy thresholds utilized by agencies, it will possible to start removing some of the detectors after 5 to 8 years. The agencies can remove all detectors after 10 to 15 years, depending on the accuracy threshold.

The methodology and results of this study can be used by agencies in their planning of future data collection investments. Although this study has tested the methodology on arterial segment only, it is also applicable to freeway segment. In fact, estimating freeway segment volumes in part based on CV data may reach the minimum acceptable accuracy requirements earlier than those on the arterials due to the higher sample sizes of CV expected on freeway segments compared to arterial segments.

REFERENCES

Adler, J., J. Bottom, C. Nelson, and K. Wunderlich. (2014). *EnableATIS Strategy Assessment* (FHWA-JPO-14-113), Federal Highway Administration, Washington, D.C.

Alexander, W., and M. Chamberlain. (2002). *ITS in Ontario: The COMPASS System*. Personal Communication; Paper for Land Use and Transportation graduate course, University of Toronto, Toronto.

Al-Sobky, A. S. A., and R. M. Mousa. (2016). "Traffic Density Determination and its Applications Using Smartphone." *Alexandria Engineering Journal*, Vol. 55, No. 1, pp. 513-523.

American Association of State Highway and Transportation Officials (AASHTO). (2013). *National Connected Vehicle Field Infrastructure Footprint Analysis – Design Concepts*. Federal Highway Administration, Washington, D.C.

Andrews, S., and M. Cops. (2009). *Final Report: Vehicle Infrastructure Integration Proof of Concept Technical Description - Vehicle*, Publication FHWA - JPO-09-043, FHWA, U.S. Department of Transportation.

Anuar, K., and M. Cetin. (2017). "Estimating Freeway Traffic Volume Using Shockwaves and Probe Vehicle Trajectory Data." *Transportation Research Procedia*, Vol. 22, pp. 183-192.

Argote, J., E. Christofa, Y. Xuan, and A. Skabardonis. (2011). "Estimation of Measures of Effectiveness Based on Connected Vehicle Data." In *Proceedings of the 14th International IEEE Conference on at Intelligent Transportation Systems (ITSC)*, October 5-7, 2011, Washington, D.C. IEEE, Washington, D.C., pp. 1767-1772.

Argote, J., E. Christofa, Y. Xuan, and A. Skabardonis. (2012). "Estimation of Arterial Measures of Effectiveness with Connected Vehicle Data." In *Proceedings of the 91st Annual Meeting Transportation Research Board*, Washington D.C.

ARTBA Work Zone Safety Consortium. (2015). Innovative End-of Queue Warning System Reduces Crashes Up to 45%. https://www.workzonesafety.org/publication/innovative-end-of-queue-warning-system-reduces-crashes-up-to-45/. Accessed February, 2017.

Asakura, Y., T. Kusakabe, N. Xuan Long and T. Ushiki. (2015). "Incident Detection Methods Using Probe Vehicles with On-board GPS Equipment." *Transportation Reasearch Procedia*, Vol. 6, pp. 17-27.

Azizi, L. Md S. Iqbal, and M. Hadi. (2017). Estimation of Freeway Platooning Measures Using Surrogate Measures Based on Connected Vehicle Data. Submitted for Presentation and Publication at the 97th TRB Annual Meeting.

Balke, K., H. Charara and S. Sunkari (2014). *Report on Dynamic Speed Harmonization and Queue Warning Algorithm Design*. Texas A& M Transportation Institute, Texas A&M University System, College Station, Texas, 2014.

Ban, X., R. Herring, J. D. Margulici, and A. M. Bayen. (2009a). Optimal Sensor Placement for Freeway Travel Time Estimation. *Transportation and Traffic Theory 2009: Golden Jubilee*, pp. 697-721.

Ban, X., R. Herring, P. Hao, and A.M. Bayen. (2009b). "Delay Pattern Estimation for Signalized Intersections Using Sampled Travel Times." *Transportation Research Record: Journal of the Transportation Research Board*, No. 2130, pp. 109–119.

Ban, X., P. Hao, and Z. Sun. (2011). "Real Time Queue Length Estimation for Signalized Intersections Using Travel Times from Mobile Sensors." *Transportation Research Part C: Emerging Technologies*, 19(6), pp. 1133–1156.

Bao, X., H. Li, L. Qin, D. Xu, B. Ran, and J. Rong. (2016). "Sensor Location Problem Optimization for Traffic Network with Different Spatial Distributions of Traffic Information." *Sensors*, Vol. 16, No. 11, 1790.

Barbaresso, J. and P. Johnson. (2014). *Connected Vehicle Infrastructure Deployment Considerations: Lessons Learned from the Safety Pilot Model Deployment*, Produced for USDOT.

Bar-Gera, H. (2007). "Evaluation of a Cellular Phone-Based System for Measurements of Traffic Speeds and Travel Times: A Case Study from Israel." *Transportation Research Part C: Emerging Technologies*, Vol. 15, No. 6, pp. 380-391.

Barrachina, J., J. A. Sanguesa, M. Fogue, P. Garrido, F. J. Martinez, J. C. Cano, C. T. Calafate, and P. Manzoni. (2013). "V2X-d: A Vehicular Density Estimation System that Combines V2V and V2I Communications." In *Wireless Days (WD), 2013 IFIP*, IEEE, pp. 1-6.

Barrachina, J., P. Garrido, M. Fogue, F. J. Martinez, J. C. Cano, C. T. Calafate, and P. Manzoni. (2015). "A V2I-based Real-Time Traffic Density Estimation System." *Wireless Personal Communications*, Vol. 83, No. 1, pp. 259-280.

Barria, J. A., and S. Thajchayapong. (2011). “Detection and Classification of Traffic Anomalies Using Microscopic Traffic Variables.” *IEEE Transactions on Intelligent Transportation Systems*, Vol. 12, No. 3, pp. 695–704.

Bhaskar, A., E. Chung, and A. G. Dumont. (2011). “Fusing Loop Detector and Probe Vehicle Data to Estimate Travel Time Statistics on Signalized Urban Networks.” *Computer‐Aided Civil and Infrastructure Engineering*, Vol. 26, No. 6, pp. 433-450.

BMW Group. (2007). Advance Warning of Drivers Heading in the Wrong Direction- the “Wrong-Way Driver” Information. https://www.press.bmwgroup.com/global/article/detail/T0012266EN/advance-warning-of-drivers-heading-in-the-wrong-direction-the-wrong-way-driver-information?language=en. Accessed February 5, 2017.

Booz Allen Hamilton. (2015). *Safety Pilot Model Deployment - Sample Data Environment Data Handbook Version 1.3*. Prepared for US Department of Transportation Research and Technology Innovation Administration Intelligent Transportation Systems – Joint Program Office.

Caceres, N., L. M. Romero, F. G. Benitez, and J. M. del Castillo. (2012). “Traffic Flow Estimation Models Using Cellular Phone Data.” *IEEE Transactions on Intelligent Transportation Systems*, Vol. 13, No. 3, pp. 1430-1441.

Cheung, S. Y., S. C. Ergen, and P. Varaiya. (2005). Traffic Surveillance with Wireless Magnetic Sensors. Presented at 12th ITS World Congress, San Francisco.

Christofa, E., J. Argote, and A. Skabardonis. (2013). “Arterial Queue Spillback Detection and Signal Control Based on Connected Vehicle Technology.” *Transportation Research Record: Journal of the Transportation Research Board*, No. 2356, pp. 61-70.

Chue, R., H. Qi and D. Lee. (2002). “Mobile Sensor and Sample-Based Algorithm for Freeway Incident Detection.” *Transportation Research Record: Journal of the Transportation Research Board*, No. 1811, pp. 12-20.

Coifman, B. (1998a). “Vehicle Re-identification and Travel Time Measurement in Real-Time on Freeways Using existing Loop Detector Infrastructure.” *Transportation Research Record: Journal of the Transportation Research Board*, No. 1643, pp. 181-191.

Coifman, B. (1998b). “A New Algorithm for Vehicle Reidentification and Travel Time Measurement on Freeways.” *Applications of Advanced Technologies in Transportation*, American Society of Civil Engineers, pp. 167-174.

Coifman, B. (2002). “Estimating Travel Times and Vehicle Trajectories on Freeways Using Dual Loop Detectors.” *Transportation Research Part A: Policy and Practice*, Vol. 36, No. 4, pp. 351-364.

Coifman, B., and M. Cassidy. (2002). “Vehicle Reidentification and Travel Time Measurement on Congested Freeways.” *Transportation Research Part A: Policy and Practice*, Vol. 36, No. 10, pp. 899-917.

ConSysTec and Cambridge Systematics. (2013). *Real-Time System Management Information Program Data Exchange Format Specification*. Final Report, FHWA-HOP-13-047, Washington, D.C.

Crabtree, J. and N. Stamatiadis. (2007). “Dedicated Short-Range Communications Technology for Freeway Incident Detection.” *Transportation Research Record: Journal of the Transportation Research Board*, No. 2000, pp. 59-69.

Cronin, B. (2012). Vehicle Based Data and Availability. USDOT Webinar, Washington, D.C.

Davis, L. C. (2004). “Effect of Adaptive Cruise Control Systems on Traffic Flow.” *Physical Review E*, Vol. 69, No. 066110, pp. 1–8.

Demissie, M., G. H. Correia, and C. L. Bento. (2013). Traffic Volume Estimation through Cellular Networks Handover Information. 13th World Conference on Transportation Research.

Dixon, K. , and C. Wang. (2002). *Development of Speed Reduction Strategies for Highway Work Zones*. Final Report FHWA-GA-02-9810. FHWA, U.S. Department of Transportation.

Doan, Q. C., T. Berradia, and J. Mouzna. (2009). “Vehicle Speed and Volume Measurement Using Vehicle-to-Infrastructure Communication.” *WSEAS Transactions on Information Science and Applications*, Vol. 6, No. 9, pp. 1468-1477.

Dowling, R., A. Skabardonis, J. Barrios, A. Jia and B. Nevers. (2015). *Impacts Assessment of Dynamic Speed Harmonization with Queue Warning*. U.S. Department of Transportation, Federal Highway Administration (FHWA), Washington D.C.

Elefteriadou, L., C. Letter, and E. Mintsis. (2012). *Managed Lane Operations – Adjusted Time of Day Pricing vs. Near-Real Time Dynamic Pricing, Volume II: Ramp Signaling and Variable Speed Limits (VSL)*. Final Report, No. UF Project 00081551, Prepared for FDOT by the University of Florida Transportation Research Center, Gainesville, FL.

Enterprise. (2014). *Queue Warning Systems in WZs*. http://enterprise.prog.org/Projects/2010_Present/iwz/QueueWarning_Summary_FINAL_June2014.pdf. Accessed June, 2014.

European Commission. (2016). *Cooperative ITS (C-ITS) Platform*. European Commission Final Report.

Federal Highway Administration (FHWA). (2016). *PIKALERT-4.5*. https://www.itsforge.net/index.php/community/explore-applications#/40/95. Accessed February 5, 2017.

Federal Highway Administration (FHWA). (2015). *Vehicle-to-Infrastructure Deployment Guidance and Products*. Draft Report, FHWA-HOP-15-015, Washington, D.C.

Federal Highway Administration (FHWA). (2014). *Traffic Data Quality Measurement*. FHWA Report 14058, Washington, D.C.

Fei, X., H. Mahmassani, and S. Eisenman. (2007). "Sensor Coverage and Location for Real-Time Traffic Prediction in Large-Scale Networks." *Transportation Research Record: Journal of the Transportation Research Board*, No. 2039, pp. 1-15.

Finley, M. D., K. N. Balke, R. Rajbhandari, S. T. Chrysler, C. S. Dobrovolny, N. D. Trout, P. Avery, D. Vickers, and C. Mott. (2016). *Conceptual Design of a Connected Vehicle Wrong-Way Driving Detection and Management System*. Technical Report. FHWA/TX-16/0-6867-1.

Florida Department of Transportation (FDOT). (2016). *Regional Traffic Management Centers*, http://www.dot.state.fl.us/trafficoperations/ITS/Projects_Deploy/RTMC.shtm. Accessed February 2016.

Florida Department of Transportation (FDOT). (2014). *FDOT SunGuide Software: Florida Advanced Traffic Management System Software*. http://www.dot.state.fl.us/trafficoperations/its/projects_arch/pdf/2014-SG-Brochure.pdf, Accessed February 2016.

Fries, R., P. Yousefzadehfard, K. Bahaaldin, A. Fadoul, K. Ghale. M. Atiquzzaman, M.A. Chowdhury, S. Petersen, S., and E. Minge. (2015). *Real-Time Information Dissemination Requirements for Illinois per New Federal Rule*. Produced for Illinois Department of Transportation, Springfield, IL.

Garber, N. J. , and S. T. Patel. (1995). “Control of Vehicle Speeds in Temporary Traffic Control Zones Using Changeable Message Signs with Radar.” *Transportation Research Record: Journal of the Transportation Research Board*, No. 1509, pp. 73-81.

Garelli, L., C. Casetti, C.F. Chiasserini, and M. Fiore. (2011). Mobsampling: “V2V Communications for Traffic Density Estimation.” In *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, IEEE, pp. 1-5.

Gazis, D. C., and C. H. Knapp. (1971). “On-Line Estimation of Traffic Densities from Time-Series of Flow and Speed Data.” *Transportation Science*, Vol. 5, No. 3, pp. 283–301.

Gettman, D. , and L. Head. (2003). *Surrogate Safety Measures From Traffic Simulation Models*. FHWA-RD-03-050. Federal Highway Administration, Washington, DC.

Gettman, D., L. Pu, T. Sayed and S. Shelby. (2008). *Surrogate Safety Assessment Model and Validation: Final Report*. FHWA-HRT-08-051. Federal Highway Administration, Washington, DC.

Gomes, G. , A. May and R. Horowitz. (2004). *Calibration of VISSIM for a Congested Freeway. California PATH Program*. Institute of Transportation Studies, University of California, Berkeley, California.

Goodall, N. J., B.L. Smith, and B.B. Park. (2016). “Microscopic Estimation of Freeway Vehicle Positions from the Behavior of Connected Vehicles.” *Journal of Intelligent Transportation Systems*, Vol. 20, No. 1, pp. 45-54.

Grone, B. W. (2012). *An Evaluation of Non-Intrusive Traffic Detectors at the NTC/NDOR Detector Test Bed*. Civil Engineering Theses, Dissertations, and Student Research. University of Nebraska, Lincoln, Nebraska.

Hadi, M., Y. Xiao, S. Fakharian Qom, E. Cabrera, and H. Al-Deek. (2014). *Performance Measurements of Transportation Systems Based on Fined-Grained Data Collected by AVI and AVL Systems*. Final Research Report, National Center for Transportation Systems Productivity and Management.

Hammit, B., and R. Young. (2015). *Connected Vehicle Weather Data for Operation of Rural Variable Speed Limit Corridors*. Mountain-Plains Consortium.

Hao, P., X. Ban, K.P. Bennett, Q. Ji, and Z. Sun. (2012). "Signal Timing Estimation Using Sample Intersection Travel Times." *IEEE Transactions on Intelligent Transportation Systems*, Vol. 13, No. 2, pp. 792–804.

Harbord, B. (1995). *Application of SISTM to Dynamic Control on the M25*. IEE Colloquium on Dynamic Control of Strategic Inter-Urban Road Networks. http://digital-library.theiet.org/content/conferences/10.1049/ic_19950246, Accessed November 25, 2012.

He, Q., K. L. Head, and J. Ding. (2012). "PAMSCOD: Platoon-Based Arterial Multi-Modal Signal Control with Online Data." *Transportation Research Part C: Emerging Technologies*, Vol. 20, No. 1, pp. 164-184.

Herrera, J. C., D. B. Work, R. Herring, X. J. Ban, Q. Jacobson, and A. M. Bayen. (2010). "Evaluation of Traffic Data Obtained via GPS-enabled Mobile Phones: The Mobile Century Field Experiment." *Transportation Research Part C: Emerging Technologies*, Vol. 18, No. 4, pp. 568-583.

Hong, Q., R. Wallace, G. and Krueger. (2014). *Connected V. Automated Vehicles As Generators Of Useful Data*, Report Produced for Michigan Department of Transportation (MDOT) by the Center for Automotive Research (CAR).

Iqbal M.S, M. Hadi, and Y. Xiao. (2017a). Accuracy, and Reliability of Estimated Travel Time Using Basic Safety Message (BSM) Data Collected from Connected Vehicles. Presented at 96th Annual Meeting of the Transportation Research Board, January, Washington, D.C.

Iqbal, M.S., M. Hadi, and Y. Xiao. (2017b). Assessment of Link Level Variation of Connected Vehicle Market Penetration, Presented at the 96th *Annual* meeting of Transportation Research Board, Washington D.C., USA.

Iteris. (2016). *Connected Vehicle Reference Implementation Architecture*. http://local.iteris.com/cvria/html/applications/applications.html. Accessed February 5, 2017.

Ivanchev, J., H. Aydt, and A. Knoll. (2016). "Information Maximizing Optimal Sensor Placement Robust against Variations of Traffic Demand Based on Importance of Nodes." *IEEE Transactions on Intelligent Transportation Systems*, Vol. 17, No. 3, pp. 714-725.

Kanayama, K., Y. Fujikawa, K. Fujimoto, and M. Horino. (1991). "Development of Vehicle-License Number Recognition System Using Real-Time Image Processing and Its Application to

Travel-Time Measurement." In *Vehicular Technology Conference, Gateway to the Future Technology in Motion, 41st IEEE*, IEEE, pp. 798-804.

Kandarpa, R., M. Chenzaie, J. Anderson, J. Marousek, T. Weil, F. Perry, L. Schworer, J. Beal, and C. Anderson. (2009). *Vehicle Infrastructure Integration Proof-of-Concept Results and Findings - Infrastructure, Volume 3B*, Final Report. Publication *FHWA*-JPO-09-057. Research and Innovative Technology Administration (RITA), U.S. Department of Transportation, Washington, DC.

Kattan, L., and S. Saidi. (2013). "Comparative Analysis of Probe-Based Ramp Metering with Detector-Based and Pre-timed Ramp Metering." *Journal of Advanced Transportation*, Vol. 47, pp: 61-78.

Khan, S. M. (2015). *Real-Time Traffic Condition Assessment with Connected Vehicles*. Master Thesis, Clemson University.

Khondaker, B., and L. Kattan. (2015). "Variable Speed Limit: A Microscopic Analysis in a Connected Vehicle Environment." *Transportation Research Part C: Emerging Technologies*, Vol. 58, pp. 146-159.

Knoop, V. L. , H. J. Van Zuylen and S. P. Hoogendoorn. (2009). Microscopic Traffic Behavior Near Incidents. In Second Sino-Dutch Joint Workshop in Transportation and Traffic Study, Shanghai, China.

Kurkjian, A., S. B. Gershwin, P. K. Houpt, A. S. Willsky, E. Y. Chow, and C. S. Greene. (1980). "Estimation of Roadway Traffic Density on Freeways Using Presence Detector Data." *Transportation Science*, Vol. 14, No. 3, pp. 232–261.

Kwon, J., B. Coifman, and P. Bickel. (2000). "Day-to-Day Travel-Time Trends and Travel-Time Prediction from Loop-Detector Data." *Transportation Research Record: Journal of the Transportation Research Board*, No. 1717, pp. 120-129.

Kwong, K., R. Kavaler, R. Rajagopal, and P. Varaiya. (2009). "Arterial Travel Time Estimation Based on Vehicle Re-identification Using Wireless Magnetic Sensors." *Transportation Research Part C: Emerging Technologies*, Vol. 17, No. 6, pp. 586-606.

Lee, C., B. Hellinga, and F., Saccomanno. (2004). "Assessing Safety Benefits of Variable Speed Limits." *Transportation Research Record: Journal of the Transportation Research Board*, Vol.1897, pp. 183-190.

Lee, J., B. Park, and I. Yun. (2013). “Cumulative Travel-Time Responsive Real-Time Intersection Control Algorithm in the Connected Vehicle Environment.” *Journal of Transportation Engineering*, Vol. 139, No. 10, pp. 1020-1029.

Lenz, H., R. Sollacher, and M. Lang. (1999). “Nonlinear Speed-Control for a Continuum Theory of Traffic Flow.” In *Proceedings of the 14th World Congress of IFAC*, Beijing, China.

Li, R., G. Rose, and M. Sarvi. (2006). “Using Automatic Vehicle Identification Data to Gain Insight into Travel Time Variability and Its Causes.” *Transportation Research Record: Journal of the Transportation Research Board*, No. 1945, pp. 24-32.

Li, J., K. Zhou, S, Shladover, and A. Skabardonis. (2013). “Estimating Queue Length under the Connected Vehicle Technology: Using Probe Vehicle, Loop Detector, and Fused Data.” In *Proceedings of the 92nd Annual Meeting of Transportation Research Board*, Washington, D.C.

Li, M., Z. Zou, F. Bu, and W.B. Zhang. (2008). Application of Vehicle-Infrastructure Integration Data on Real-Time Arterial Performance Measurements. Presented at 87th Annual Meeting of Transportation Research Board, Washington D.C.

Li, Y., Y. Bai and U. Firman. (2010). “Determining the Effectiveness of PCMS on Reducing Vehicle Speed in Rural Highway Work Zones.” In *Proceedings of the 89th Annual Meeting of Transportation Research Board*, Washington, D.C.

Mao, R., and G. Mao. (2013). “Road Traffic Density Estimation in Vehicular Networks.” In *Wireless Communications and Networking Conference (WCNC)*, IEEE, pp. 4653-4658.

Martin, P., J. Perrin, B. Hansen, R. Kump and D. Moore. (2009). *Incident Detection Algorithm Evaluation*. Utah Department of Transportation, Utah.

McCoy, P. T. , J. A. Bonneson and J. A. Kollbaum. (1995). “Speed Reduction Effects of Speed Monitoring Displays with Radar in Work Zones on Interstate Highways.” *Transportation Research Record: Journal of the Transportation Research Board*, No. 1509, pp. 65-72.

McGurrin, M. (2012). *Vehicle Information Exchange Needs for Mobility Applications Exchange*. FHWA Report FHWA-JPO-12-021, Washington. D.C.

Michigan Department of Transportation (MDOT). (2012). *Connected Vehicle Technology Local Government Delphi Study*. Report Produced for Michigan Department of Transportation (MDOT) by the Center for Automotive Research (CAR).

Miller, T., W. Davis, G. Reed, P. Doraiswamy, and A. Tang. (2002). “Effect of County-Level Income on Vehicle Age Distribution and Emissions.” *Transportation Research Record: Journal of the Transportation Research Board*, No. 1815, pp. 47-53.

Minge, E., J. Kotzenmacher, and S. Petersen. (2010). *Valuation of Non- Intrusive Technologies for Traffic Detection*. Produced for Minnesota DOT by SRF Consulting Group, Minneapolis, MN.

Minelli, S., P. Izadpanah, and S. Razavi. (2015). “Evaluation of Connected Vehicle Impact on Mobility and Mode Choice.” *Journal of Traffic and Transportation Engineering (English Edition)*, Vol. 2, No. 5, pp. 301-312.

Mitsakis, E., E. Chrysohoou, J. M. Salanova Grau, P. Iordanopoulos, and G. Aifadopoulou. (2017). “The Sensor Location Problem: Methodological Approach and Application.” *Transport*, Vol. 32, No. 2, pp. 113-119.

Mouskos, K. C., E. Niver, L. Pignataro, S. Lee, N. Antoniou and . L. Papadopoulos. (1998). *Evaluation of the TRANSMIT system, Institute for Transportation*, New Jersey Institute of Technology.

Nam, D. H., and D. R. Drew. (1999). “Automatic Measurement of Traffic Variables for Intelligent Transportation Systems Applications," *Transportation Research Part B*, Vol. 33, pp. 437-457.

Nantes, A., R. Billot, M. Miska, and E. Chung. (2013). “Bayesian Inference of Traffic Volumes Based on Bluetooth Data.” In *Proceedings of the 92nd Annual Meeting of Transportation Research Board*, Washington, D.C., pp. 13-17.

National Transportation Safety Board. (2012). *Highway Special Investigation Report: Wrong Way Driving*. Report NTSB/SIR-12/01., Washington, D.C.

National Transportation Safety Board. (2016). *U.S. Department of Transportation Issues Advance Notice of Proposed Rulemaking to Begin Implementation of Vehicle-to-Vehicle Communications Technology*. http://www.nhtsa.gov/About+NHTSA/Press+Releases/NHTSA-issues-advanced-notice-of-proposed-rulemaking-on-V2V-communications. Accessed May 5, 2016.

National Transportation Safety Board. (2001a). *Vehicle-and Infrastructure-Based Technology for the Prevention of Rear-End Collisions*. Washington, D.C.

National Transportation Safety Board. (2001b). *Safety Recommendation H-01-10 and H-01-11*.

Ndoye, M., V. Totten, B. Carter, D. M. Bullock, and J. V. Krogmeier. (2008). Vehicle Detector Signature Processing and Vehicle Reidentification for Travel Time Estimation. Presented at 87th Annual Meeting of Transportation Research Board.

Nemsky, T. P. (2014). *ITS Work Zone Experiences in Southern Illinois. Webinar Presentation on Work zone Intelligent Transportation Systems*. Sponsored by FHWA, U.S. Department of Transportation.
http://www.ops.fhwa.dot.gov/wz/webinars/itsoverview013014/nemsky/index.htm.

Nihan, N. L., Y. Wang, and X, Zhang. (2002). *Evaluation of Dual-Loop Data Accuracy Using Video Ground Truth Data*. Research Report, Wa-RD 535.1, Transportation Northwest, Department of Civil Engineering, University of Washington.

Oh, C., S. G. Ritchie, and S. T. Jeng. (2004). Vehicle Re-identification Using Heterogeneous Detection Systems. Presented at the 83rd Annual Meeting of the Transportation Research Board. Washington, D.C.

Olia, A., H. Abdelgawad, B. Abdulhai, and S. N. Razavi. (2014). Assessing the Potential Impacts of Connected Vehicles: Mobility, Environmental, and Safety Perspectives. Presented at the 93rd Annual Meeting of Transportation Research Board.

Open Source Application *Development Portal (OSADP) (2015)*. http://www.itsforge.net/78-news/140-tca-2-3-release. Accessed July 13, 2015.

Panichpapiboon, S., and W. Pattara-Atikom. (2011). "Exploiting Wireless Communication in Vehicle Density Estimation." *IEEE Transactions on Vehicular Technology*, Vol. 60, No. 6, pp. 2742-2751.

Park, H. (2008). *Investigating Potential Benefits of Vehicle Infrastructure Integration- Enabled Ramp Metering*. Final Report No. UVACTS-13-0-114. Center for Transportation Studies, University of Virginia, Charlottesville.

Park, H., and B. L. Smith. (2012). "Investigating Benefits of IntelliDrive in Freeway Operations: Lane-Changing Advisory Case Study." *Journal o Transportation Engineering*, pp.1113-1122

Park, H., C. Sarma Bhamidipati, and B. Lee Smith. (2011). "Development and Evaluation of Enhanced IntelliDrive-Enabled Lane Changing Advisory Algorithm to Address Freeway Merge

Conflict.” *Transportation Research Record: Journal of the Transportation Research Board*, No. 2243, pp. 146–157.

Parkany, E. and D. Bernstein. (1995). “Design of Incident Cetection Algorithms Using Vehicle-to Roadside Communication Sensors.” *Transportation Research Record: Journal of the Transportation Research Board*, No. 1494, pp. 67-74.

Pesti, G., C.-L. Chu, H. Charara, G. Ullman and K. Balke. (2013). “Simulation Based Evaluation of Dynamic Queue Warning System Performance.” In *Proceedings of the 9ndh Annual Meeting of Transportation Research Record*. Washington, D.C.

Preisen, L., and D. Deeter, (2014). *Next Generation Traffic Data and Incident Detection. ENTERPRISE Pooled Fund Study TPF-5(231)*, Report No. ENT-2014-2, Lansing, MI.

Petersen, S., E. Minge and C. Iverson. (2013). “Queue Warning and Travel Time Estimation near a Work Zone.” In *Proceedings of the 93rd Annual Meeting of Transportation Research Record*, Washington, D.C.

Piao, J., and M. McDonald. (2008). “Safety Impacts of Variable Speed Limits – A Simulation Study.” In *Proceedings of the 11th International IEEE, Conference on Intelligent Transportation Systems*, Beijing, China.

Priemer, C., and B. Friedrich. (2009). “A Decentralized Adaptive Traffic Signal Control Using V2I Communication Data.” In *2009 12th International IEEE Conference on Intelligent Transportation Systems*, St. Louis, MO, pp. 1–6.

Pueboobpaphan, R., F. Liu, and B. van Arem. (2010). “The Impacts of a Communication based Merging Assistant on Traffic Flows of Manual and Equipped Vehicles at an On-Ramp Using Traffic Flow Simulation.” In *Proceedings of the 13th International IEEE Conference on Intelligent Transportation systems*.

PTV. (2012). *VISSIM 5.40 - User Manual*. Karlsruhe, Germany.

Qiu, T. Z., X.Y. Lu, and A. H. Chow. (2009). “Real-Time Density Estimation on Freeways with Loop Detector and Probe Data.” *IFAC Proceedings*, Vol. 42, No. 15, pp. 298-303.

Qiu, T., X. Y. Lu, A. H. Chow, and S. Shladover. (2010). “Estimation of Freeway Traffic Density with Loop detector and Probe Vehicle Data.” *Transportation Research Record: Journal of the Transportation Research Board*, No. 2178, pp 21-29.

Rahman, Z. , and S. P. Mattingly. (2014). “Macroscopic Evaluation of Incident-Induction Deriver Behavior Changes.” In *Symposium Celebrating 50 Years of Traffic Flow Theory*, Portland, Oregon.

Randolph, L. (2016). Detecting Wrong Way Drivers. The 2016 National Rural ITS Conference. Chattanooga, Tennessee.

Richards, S. A. , R. C. Wunderlich and C. L. Dudek. (1986). “Field Evaluation of Work Zone Speed Control Techniques.” *Transportation Research Record: Journal of the Transportation Research Board*, No. 1035, pp. 66-78.

Rim, H., C. Oh, K. Kang, and S. Kim. (2011). “Estimation of Lane-Level Travel Times in Vehicle-to-Vehicle and Vehicle-to-Infrastructure-Based Traffic Information System.” *Transportation Research Record: Journal of the Transportation Research Board*, No. 2243, pp. 9–16.

Roess, R. P., E. S. Prassas, and W. R. McShane. (2011). *Traffic Engineering*. Pearson.

Rompis, S. Y. , F. G. Habtemichael and M. Cetin. (2014). A Methodology for Calibrating Microscopic Simulation for Modeling Traffic Flow under Incidents. In IEEE 17th International Conference on Intelligent Transportation Systems (ITSC), Qingdoa, China.

Sabra, Z. A. , D. Gettman, R. D. Henry and V. Nallamothu. (2010). *Balancing Safety and Capacity in an Adaptive Signal Control System—Phase 1*. FHWA-HRT-10-038, Federal Highway Administration, Mclean, VA.

SAE International. (2016). *Dedicated Short Range Communications (DSRC) Message Set Dictionary, SAE J2735*.

SAE International. (2009). *Dedicated Short Range Communications (DSRC) Message Set Dictionary, SAE J2735*.

Samuelson, J. P. (2011). *Experience in Data Quality Assessment on Archived Historical Freeway Traffic Data*. Master of Science Thesis, Arizona State University.

Sanguesa, J. A., J. Barrachina, M. Fogue, P. Garrido, F.J. Martinez, J.C. Cano, and P. Manzoni. Sensing (2015). “Traffic Density Combining V2V and V2I Wireless Communications.” *Sensors*, Vol. 15, No. 12, pp. 31794-31810.

Scarinci, R., A. Hegyi, and B. G. Heydecker. (2013). Cooperative Ramp Metering; A Study of the Practicality of a Ramp Metering Development Using Intelligent Vehicles. Presented in 45th Annual UTSG Universities' Transport Study Group Conference.

Shao, M., L. Sun, and X. Shao. (2016). *Sensor Location Problem for Network Traffic Flow Derivation Based on Turning Ratios at Intersection*. Mathematical Problems in Engineering.

Sherali, H. D., J. Desai, and H. Rakha. (2006). "A Discrete Optimization Approach for Locating Automatic Vehicle Identification Readers for the Provision of Roadway Travel Times." *Transportation Research Part B: Methodological*, Vol. 40, No. 10, pp. 857-871.

Sisiopiku, V. P., N. M. Rouphail, and A. Santiago. (1994). "Analysis of Correlation between Arterial Travel Time and Detector Data from Simulation and Field Studies." *Transportation Research Record: Journal of the Transportation Research Board*, No. 1457, pp. 166-173.

Smith, B. L., B. B. Park, H. Tanikella, and G. Zhang. (2007). *Preparing to Use Vehicle Infrastructure Integration in Transportation Operations: Phase I.* Publication VTRC 08-CR1, Virginia Department of Transportation.

Smith, B. L., H. Park, R. Venkatanarayana, N. J. Goodall, J. Datesh, and C. Skeritt. (2010). *Development of New Traffic Signal Control Algorithms under IntelliDrive (No. 2)*. IntelliDrive Pooled Fund Study.

Stephens, D., T. Timcho, J. Schroeder, J. Brown, P. Bacon, T. Smith, K. Balke, H. Charara and S. Sunkari. (2015). *Intelligent Network Flow Optimization (INFLO) Prototype Seattle Small-Scale Demonstration Report.* United States Department of Transportation, ITS Joint Program Office, Research And Innovative Technology Administration (RITA), Washington, D.C.

S. T. Team. (1999). *Innovative Traffic Control Technology and Practice in Europe*. Federal Highway Administration, International Technology Exchange Program.

SubCarrier Systems Corp (SCSC). (2009*). Surface Vehicle Standard, SAE 2735*. for Use by the DSRC committee of the SAE.

Sun, C., G. Arr, and R. Ramachandran. (2003). "Vehicle Reidentification as Method for Deriving Travel Time and Travel Time Distributions: Investigation." *Transportation Research Record: Journal of the Transportation Research Board*, No. 1826, pp. 25-30.

Szczesny, Joseph. (2013). *Daimler Debuts Alert System for Wrong-Way Drivers*.

http://www.thedetroitbureau.com/2013/02/daimler-debuts-alert-system-for-wrong-way-drivers/. Accessed February 5, 2017.

Tam, M. L., and W. H. Lam. (2008). "Using Automatic Vehicle Identification Data for Travel Time Estimation in Hong Kong." *Transportmetrica*, Vol. 4, No. 3, pp. 179-194.

Thompson, D. (2013). Real-Time Data Capture and Management (DCM) Program Update on Program Status. Winter USDOT Webinar Series, Washington, D.C.

Toth, C., M. Waiseley, J. Schroeder, M. Omay, C. Castle, and S. Cook. (2016). *Michigan Department of Transportation (MDOT) Weather Responsive Traveler Information (Wx-TINFO) System – Final Report*. FHWA-JPO-16-323.

Toppen, A., and K. Wunderlich. (2004). Travel Time Data Collection for Measurement of Advanced Traveler Information Systems Accuracy. Presented at 14th Annual Meeting and Exposition of ITS America: At the Crossroads: Integrating Mobility Safety and Security, San Antonio, Texas.

Traficon. (1998). *E313 Antwerpen to Hasselt, Belgium Case Study*. http://www.traficon.com/inhoud_1_4_3_1.lasso?Case_ID=58. Accessed August 22, 2002.

Transportation Research Board (TRB). (2010). *Highway Capacity Manual*. National Research Council. Washington, D.C.

Turner, S. M., W. L. Eisele, R. J. Benz, and D. J. Holdener. (1998). *Travel Time Data Collection Handbook*. Publication FHWA-PL-98-035, FHWA, U.S. Department of Transportation.

Ullman, G. L., V. Iragavarapu and R. E. Brydia. (2016). "Safety Effects of Portable End-of-queue Warning System Deployments at Texas Work Zones." *Transportation Research Record, Journal of Transportation Research Board.* No. 2555, pp. 46-52.

United States Census Bureau. (2016). *American Community Survey 5-Year Estimate – Geodatabase format*. https://www.census.gov/geo/maps-data/data/tiger-data.html. Accessed May 23, 2016.

U.S. Department of Transportation (USDOT). (2017). *Road Weather Connected Vehicle Applications: Progress and Insights.* http://www.its.dot.gov/research_archives/road_weather/road_weather_progress.htm. Accessed February 5, 2017.

U.S. Department of Transportation (USDOT). (2016). *Connected Vehicle Technology*. http://www.its.dot.gov/connected_vehicle/connected_vehicle_tech.htm. Accessed May 5, 2016.

U.S. Department of Transportation (USDOT). (2015a). *ITS Standards Fact Sheets.* https://www.standards.its.dot.gov/Factsheets/Factsheet/71, Accessed December 30, 2015.

U.S. Department of Transportation (USDOT). (2015b). *Connected Vehicle Reference Implementation Architecture Website, Connected Vehicle Reference Implementation Architecture*, http://www.iteris.com/cvria/, Accessed December 30, 2015.

U.S. Department of Transportation (USDOT). (2014). *Trajectory Coversion Algorithm-VISSIM Software User Manual*.

U.S. Department of Transportation (USDOT). (2008). *Vehicle-Infrastructure Integration (VII) Initiative Benefit-Cost Analysis, Version 2.3 (Draft)*. Report for ITS Joint program office, http://www.its.dot.gov/research_docs/pdf/7vii-benefit-cost-analysis.pdf. Accessed May 23, 2016.

U.S. News & World Report. (2016). *Car-to-Car Communication May Hit Roads Soon*. http://www.usnews.com/news/articles/2015/03/04/car-to-car-communication-may-hit-roads-soon. Accessed May 5, 2016

Vanajakshi, L. D., B. M. Williams and L. R. Rilett. (2009). "Improved Flow-Based Travel Time Estimation Method from Point Detector for Freeways." *Journal of Transportation Engineering*, pp. 26-36.

Vasudevan, M., and J. O'Hara. (2015). Measure Estimation Using Connected Vehicle Data. TRB HCQS Committee Mid-Year Meeting Connected Vehicle Data Workshop.

Wiles, P. B. , S. A. Cooner, C. H. Walters and E. J. Pultorak. (2003). *Advance Warning of Stopped Traffic On Freeways: Current Practices And Field Studies of Queue Propagation Speeds*. Texas Transportation Institute, The Texas A&M University System, College Station, Texas.

Woody, T. (2006). *Calibrating Freeway Simulation Models in VISSIM*. University of Washington, Seattle, WA.

Wright, J., J.K. Garrett, C. J. Hill, G.D. Krueger, J. H. Evans, S. Andrews, C. K. Wilson, R. Rajbhandari, and B. Burkhard. (2014). *National Connected Vehicle Field Infrastructure Footprint Analysis*. Final Report, FHWA-JPO-14-125, Washington, D.C.

Xiao, Y. (2011). *Hybrid Approaches to Estimating Freeway Travel Times Using Point Traffic Detector Data*. Ph.D. Dissertation, Florida International University.

Ye, P., and D. Wen. (2007). "Optimal Traffic Sensor Location for Origin–Destination Estimation Using a Compressed Sensing Framework." *IEEE Transactions on Intelligent Transportation Systems*, Vol. 18, No. 7, pp. 1857-1866.

Ygnace, J. L., C. Drane, Y. B. Yim, and R. De Lacvivier. (2000). *Travel Time Estimation on the San Francisco Bay Area Network Using Cellular Phones as Probes*. Publication UCB-ITS_PWP-2000-18, California Partners for Advanced Transit and Highways (PATH).

Young, S. E., M. Hamedi, E. Sharifi, R. M. Juster, K. Kaushik, and S. Eshragh. (2015). *I-95 Corridor Coalition Vehicle Probe Project Validation of Arterial Probe Data. Summary Report*. Prepared for I-95 Corridor Coalition.

Zech, W. C. , S. B. Mohan and J. Dmochowski. (2008). "Evaluation of Messages on Changeable Message Signs as a Speed Control Measure in Highway Work Zones." *Practical Periodical on Structural Design and Construction*, Vol. 13, No. 1, pp. 11-18.

Zeng, X., K. Balke, and P. Songchitruksa. (2012). *Potential Connected Vehicle Applications to Enhance Mobility, Safety, and Environmental Security*. Report 161103-1.

Zhan, X., Y. Zheng, X. Yi, and S. V. Ukkusuri. (2017). "Citywide Traffic Volume Estimation Using Trajectory Data." *IEEE Transactions on Knowledge and Data Engineering*, Vol. 29, No. 2, pp. 272-285.

Zhang, W. (2006). *Freeway Travel Time Estimation Based on Spot Speed Measurements*. Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Zheng, J., and H.X. Liu. (2017). "Estimating Traffic Volumes for Signalized Intersections Using Connected Vehicle Data." *Transportation Research Part C: Emerging Technologies*, Vol. 79, pp. 347-362.

Zhou, Y., S. Chen, Y. Zhou, M. Chen, and Q. Xiao. (2015). "Privacy-Preserving Multi-Point Traffic Volume Measurement through Vehicle-to-Infrastructure Communications." *IEEE Transactions on Vehicular Technology*, Vol. 64, No. 12, pp. 5619-5630.

Zou, Z., M. Li, and F. Bu. (2010). Link Travel Time Estimation Based on Vehicle Infrastructure Integration Probe Data. ICCTP 2010: Integrated Transportation Systems: Green, Intelligent, Reliable. ASCE, pp. 2266-2276.

Appendix. Descriptions of Safety Pilot Model Deployment Data from the RDE Database

Table A-1 Data Files Included in the Safety Pilot Model Deployment (Source: Booz Allen Hamilton, 2015)

SPMD Environment				
Driving Data		**BSM**	**RSE**	**Contextual**
DAS1	**DAS2**	**BSM**	**RSE**	**Weather**
AudioTimes	HV_Radar	BrakeByte1Events	BSM	Weather/climatic data
DataFrontTarget	HV_Primary	BrakeByte2Events	Geometry	
DataLane	DAS2_Trip_Summary	BsmP1	Lane	
DataWsu		ExteriorLightsEvents	LaneNode	
DAS1_Trip_Summary		PosAccurByte1Events	MAP	
		PosAccurByte2Events	Packet	
		PosAccurByte3Events	PCAPFile	
		PosAccurByte4Events	SPAT	
		SteerAngleEvents	SPATMovement	
		ThrottlePositionEvents	TIM	
		TransStateEvents	TIMRegion	
		WiperStatusFrontEvents	TIMRegionNode	
		BSM_Trip_Summary		

Table A-2 DAS1 Data Set (Source: Booz Allen Hamilton, 2015)

File Number	File	Description	Sample Rate
1	DataFrontTargets	Log of the data collected by the Mobileye sensor which is a part of the DAS; largely includes data about the (vehicle) object that is in front of the host vehicle	10Hz
2	DataLane	Logs quality of the lane markings next to the host vehicle as well as the distances between each side of the vehicle and each lane line	10Hz
3	DataWsu	Log of GPS and CAN bus data obtained via the onboard WSU	10Hz
4	DAS1_Trip_Summary	A list of summary measures for each trip completed by a vehicle equipped with DAS1	1 per trip

Table A-3 DAS2 Data Set (Source: Booz Allen Hamilton, 2015)

Number	File	Description	Sample Rate
1	HV_Primary	Main log file for the data acquisition system that logs vehicle position and motion data	10Hz
2	HV_Radar	Registered information from the host vehicle's radar unit	10Hz
3	DAS2_Trip_Summary_File	A list of summary measures for each trip completed by a vehicle equipped with DAS2	1 per trip

Table A-4 BSM Data Set (Source: Booz Allen Hamilton, 2015)

File Number	File	Description	Sample Rate
1	BrakeByte1Events	Status of the vehicle's primary brake system	On Event
2	BrakeByte2Events	Status of the vehicle's system control/advance breaking features (e.g., antilock brake system, stability control)	On Event
3	BsmP1	Part I of the BSM, primarily vehicle position and motion data	10 Hz
4	ExteriorLightsEvents	Provides the status of all exterior lights on the vehicle	On Event
5	PosAccurByte1Events	Accuracy of the positional determination with respect to each given Axis – semi-major	On Event
6	PosAccurByte2Events	Accuracy of the positional determination with respect to each given Axis – semi-minor	On Event
7	PosAccurByte3Events	Accuracy of the positional determination with respect to each given Axis – orientation of semi-major axis	On Event

File Number	File	Description	Sample Rate
8	PosAccurByte4Events	Accuracy of the positional determination with respect to each given Axis – orientation of semi-minor axis	On Event
9	SteerAngleEvents	The angle of the steering wheel (signed value)	On Event
10	ThrottlePositionEvents	Throttle position, expressed in units of 0.5 percent of range of travel	On Event
11	TransStateEvents	Current state of a vehicle's transmission	On Event
12	WiperStatusFrontEvents	Current state of the wiper system at the front of the vehicle	On Event
13	BSM_Trip_Summary_File	A list of summary measures for each trip per transmitted BSMs	Per trip

Table A-5 RSE Data Set (Source: Booz Allen Hamilton, 2015)

File Number	File	Description	Sample Rate
1	BSM	BSM data including motion and location elements	10 Hz
2	Geometry	Describes intersection detail at locations where RSE were placed	N/A
3	Lane	Describes lane attributes in the vicinity of the RSE, which is often located at an intersection with multiple (lane) approaches	N/A
4	LaneNode	Describes lane descriptors as they relate to nodes (ground reference points), usually near an intersection	N/A
5	Map	Wrapper object for map data. Includes complex intersection descriptions, high speed curve outlines, and segments of roadway. Sometimes referred to as the GID layer	N/A
6	Packet	Provides details for every packet transmission. Packet includes fileIDs, sources, and time stamps	10 Hz
7	PCAP	Describes packet capture header information and specifies listening setup used to capture all vehicle to vehicle communications	10 Hz
8	SPAT	Contains basics of a SPAT message including intersection details	10 Hz
9	SPATMovement	Describes signal and timing information for movements at intersections	10 Hz
10	TIM	Contains Traveler Information Message information which transmits advisory and road sign messages to vehicles	10 Hz

File Number	File	Description	Sample Rate
11	TIMRegion	Specifies types of regions to which TIMs apply	N/A
12	TIMRegionNode	Specifies types of regions to which TIMs apply in terms of offsets from a give node	N/A
13	TIMRegionXRef	Maps TIMs to the regions in which messages are applicable	N/A
14	WeatherData	Specifies surface weather data at stations of interest	Varies