10-4-2017

Data Support of Advanced Traveler Information System Considering Connected Vehicle Technology

Md Shahadat Iqbal
miqba005@fiu.edu

DOI: 10.25148/etd.FIDC004062

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

DATA SUPPORT OF ADVANCED TRAVELER INFORMATION SYSTEM CONSIDERING CONNECTED VEHICLE TECHNOLOGY

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

CIVIL ENGINEERING

by

Md Shahadat Iqbal

2017

To: Dean John Volakis
College of Engineering and Computing

This dissertation, written by Md Shahadat Iqbal, and entitled Data Support of Advanced Traveler Information System Considering Connected Vehicle Technology, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

Albert Gan

Zhenmin Chen

L. David Shen

Xia Jin

Yan Xiao

Mohammed Hadi, Major Professor

Date of Defense: October 4, 2017

The dissertation of Md Shahadat Iqbal is approved.

Dean John Volakis
College of Engineering and Computing

Andrés G. Gil
Vice President for Research and Economics Development
and Dean of the University Graduate School

Florida International University, 2017

DEDICATION

I dedicate this dissertation to my parents, Mozammel Haque and Tahera Haque, and to my wife, Sharifa Begum, for their unconditional love and support.

ACKNOWLEDGMENTS

I would like to take this opportunity to acknowledge everyone who supported me to reach my dream of achieving the doctoral degree. First and foremost, I would like to express my gratitude to my advisor and mentor, Dr. Mohammed Hadi, for his continuous guidance, constructive criticism, and directive suggestions at every stage of the dissertation. I am thankful to him not only for his thoughtful insights provided in completing the dissertation but also for his care in every other aspect during the time of my doctoral study at Florida International University. I consider myself fortunate enough to have been one of the students of Dr. Hadi. I will be grateful to him throughout my life.

I am also grateful to Dr. Yan Xiao for being supportive from the very beginning of the research and guided me with critical problem-solving ideas till the completion of the processes. Without Dr. Yan's consistent instructions in this research, the dissertation would become impossible.

I would also like to thank all my other committee members, Dr. Albert Gan, Dr. L. David Shen, Dr. Xia Jin and Dr. Zhenmin Chen for their interest in my research. I sincerely appreciate their invaluable time for reading the dissertation and providing comments and suggestions to improve it. Special thanks to Dr. Chen for his guidance over the statistical problems.

I would also like to thank all my colleagues at Lehman Center for Transpiration Research (LCTR) for collaborating with me at various stages of my dissertation and helping me pass a good time these years. Special thanks to Fatema Haque Farzana for helping me to run some of the simulation analysis. Furthermore, I would like to acknowledge Federal Highway Administration (FHWA) and Florida Department of Transportation (FDOT) Research Center for providing financial support for this research.

I would like to convey a very special thanks to all my family members for their boundless love, support, and encouragement in pursuing my doctoral degree.

ABSTRACT OF THE DISSERTATION

DATA SUPPORT OF ADVANCED TRAVELER INFORMATION SYSTEM CONSIDERING CONNECTED VEHICLE TECHNOLOGY

by

Md Shahadat Iqbal

Florida International University, 2017

Miami, Florida

Professor Mohammed Hadi, Major Professor

Traveler information systems play a significant role in most travelers' daily trips. These systems assist travelers in choosing the best routes to reach their destinations and possibly select suitable departure times and modes for their trips. Connected Vehicle (CV) technologies are now in the pilot program stage. Vehicle-to-Infrastructure (V2I) communications will be an important source of data for traffic agencies. If this data is processed properly, then agencies will be able to better determine traffic conditions, allowing them to take proper countermeasures to remedy transportation system problems under different conditions.

This research focuses on developing methods to assess the potential of utilizing CV data to support the traveler information system data collection process. The results from the assessment can be used to establish a timeline indicating when an agency can stop investing, at least partially, in traditional technologies, and instead rely on CV technologies for traveler information system support. This research utilizes real-world vehicle trajectory data collected under the Next Generation Simulation (NGSIM) program and simulation modeling to emulate the use of connected vehicle data to support

the traveler information system. NGSIM datasets collected from an arterial segment and a freeway segment are used in this research. Microscopic simulation modeling is also used to generate required trajectory data, allowing further analysis, which is not possible using NGSIM data.

The first step is to predict the market penetration of connected vehicles in future years. This estimated market penetration is then used for the evaluation of the effectiveness of CV-based data for travel time and volume estimation, which are two important inputs for the traveler information system. The travel times are estimated at different market penetrations of CV. The quality of the estimation is assessed by investigating the accuracy and reliability with different CV deployment scenarios. The quality of volume estimates is also assessed using the same data with different future scenarios of CV deployment and partial or no detector data. Such assessment supports the identification of a timeline indicating when CV data can be used to support the traveler information system.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

AADT	Annual Average Daily Traffic
ALPR	Automatic License Plate Reader
APD	Absolute Percentage Deviation
AR	Augmented Reality
ATIS	Advanced Traveler Information System
AVL	Automatic Vehicle Location
BEA	Bureau of Economic Analysis
BSM	Basic Safety Message
CAM	Cooperative Awareness Messages
CCTV	Closed Circuit Television Camera
CFR	Code of Federal Regulations
CV	Connected Vehicle
CVRIA	Connected Vehicle Reference Implementation Architecture
DMA	Dynamic Mobility Applications
DMS	Dynamic Message Sign
DSRC	Dedicated Short-Range Communication
ETC	Electronic Toll Collection
FCC	Federal Communications Commission
FDOT	Florida Department of Transportation
FHWA	Federal Highway Administration
FITSEVAL	Florida Intelligent Transportation Systems Evaluation Tool
GIS	Geographic Information System

GPS	Global Positioning System
HAR	Highway Advisory Radio
HCM	Highway Capacity Manual
ID	Identification
ITIS	Intelligent Traveler Information Systems
ITS	Intelligent Transportation System
LDT	Light-Duty Trucks
LDV	Light-Duty Vehicle
LPR	License Plate Reader
LTEE	Link Travel Time Estimation Error
MAC	Media Access Control
MAPD	Mean Absolute Percentage Deviation
MAPE	Mean Absolute Percentage Error
MORPC	Mid-Ohio Regional Planning Commission
MP	Market Penetration
NGSIM	Next Generation Simulation
NHTSA	National Highway Traffic Safety Administration
OBD-II	OnBoard Diagnostic port II
OBU	On Board Unit
OD	Origin-Destination
ODOT	Ohio Department of Transportation
PDM	Probe Data Message
PSN	Probe Segment Number

RMSE	Root Mean Squared Error
RSU	Roadside Unit
RTEE	Route Travel Time Estimation Error
RTSMIP	Real-Time System Management Information Program
SAE	Society of Automotive Engineers
SERPM	Southeast Florida Regional Planning Model
TCA	Trajectory Conversion Algorithm
TCS	Traction Control System
TIS	Traveler Information System
TV	Television
USDOT	United States Department of Transportation
VDOT	Virginia Department of Transportation
V2I	Vehicle to Infrastructure
V2V	Vehicle to Vehicle
WWW	World Wide Web

CHAPTER I

INTRODUCTION

1.1 Problem Statement

Traveler information systems play a significant role in most travelers' daily trips. The systems assist the travelers in choosing the best route to reach their destination and possibly select a suitable departure time and mode for their trips (Khattak et al., 1995; 1996a; Khattak and De Palma, 1996; Mahmassani and Liu, 1999; Al-Deek et al., 1998; Jou, 2001). With the availability of information, travelers can change their routes if there is non-recurrent congestion on the roadway due to various reasons such as incidents, special events, or bad weather conditions (Sen et al., 2001; Levinson, 2003; 1997; Dia, 2002; Zhang et al., 2008; Adler, 2001; Khattak et al., 1995; Jou, 2001). Traveler information may also help to reduce secondary crash probabilities (Al-Deek et al., 1998). It has been reported that traveler information is not only important for a traveler's regular trips, but for other types of trips such as pleasure (Gandy and Meitner, 2007) and shopping trips (Kraan et al., 2000).

The traveler information dissemination systems continue changing with as technology advances. Before the 1990s, the typical traveler information dissemination techniques were media outlets and very limited field device deployments. Starting in the first half of the 1990s, there has been an increasing use of infrastructure devices and other dissemination methods. The common media outlets are television, radio, and newspapers. The latest field devices include Dynamic Message Signs (DMS), and Highway Advisory

Radio (HAR). Additional dissemination methods are websites, mobile apps, in-vehicle dynamic navigation systems, social media, and 511 phone call systems.

In terms of data acquisition, the technologies used to collect traffic data have also been evolving. Initially, data collection methods were limited to manual data collection such as airplane monitoring, motorist calls, and police calls. Data collection techniques have also relied on detector data and Closed Circuit Television Camera (CCTV) installations. More recently, several promising technologies have been used to collect traffic data, including vehicle re-identification technologies based on Bluetooth, Wi-Fi, and Magnetometer technologies. Apart from these, third-party vendors are also playing a vital role in traffic data collection. Examples of such vendors include HERE, INRIX, TomTom, Waze, and Google.

Connected Vehicle (CV) technologies are now in the pilot program stage. The benefits from connected vehicle system deployments are multidimensional. The associated Vehicle to Vehicle (V2V) and Vehicle to Infrastructure (V2I) communications will improve traffic safety, mobility, and environmental impacts. The V2I communication will be an important source of data for traffic agencies. If this data is processed properly, then agencies will be able to better determine traffic conditions, allowing them to take proper countermeasures under different conditions. The derived traffic conditions could then be sent back to the CVs as traveler information. This two-way communication will certainly provide a new opportunity for transformative advancements in the fields of traveler information systems and traffic management.

This research focuses on assessing the possibilities of utilizing CVs to support the collection of data as part of traveler information systems. The results from this research can be used to establish a timeline indicating when an agency can stop investing in traditional technologies and place emphasis on emerging technologies based on CVs.

1.2 Research Goal and Objectives

The main goal of this research is to develop a method to identify a timeline indicating when CV-based technologies can be used as a more effective alternative to support traveler information systems, compared to other existing technologies. The main objectives of this research are:

- Develop a method for the prediction of the future market penetration of connected vehicles.
- Develop a method to evaluate the use of connected vehicle data to provide accurate and reliable travel time.
- Develop a method to evaluate the use of connected vehicle data to provide accurate and reliable volume estimates.

1.3 Research Organization

This research includes a total of six chapters. Chapter II provides a literature review on traveler information system topics that are related to this research. Chapter III presents the methodology used to achieve the stated objectives. Chapter IV describes the data sources used in this research. Chapter V provides the details of the model application results, and finally, Chapter VI summarizes the full research and provides recommendations for future studies.

CHAPTER II

LITERATURE REVIEW

A Traveler Information System (TIS) consists of three processes. The first process is to acquire traffic data from multiple sources. The second process analyzes the data to derive the performance measurements that need to be disseminated to travelers. The final process disseminates a wide range of transportation information to the travelers. The main objective of this system is to make a traveler's journey more convenient, more efficient and safer. TIS collects traffic data using various surveillance technologies to determine the conditions of the network in real time. Next, the data are processed utilizing a central software system and are disseminated to travelers both prior to their trips and en-route. Adlar and Blue (1998) grouped the development of traveler information systems into two distinct generations. The first generation is simply called "traveler information system," which arose from the emergence of traffic surveillance, traffic control and computer technologies in the late 1960s and early 1970s. The Advanced Traveler Information System (ATIS) is considered the second generation (Watkins et al., 2013, Adlar and Blue, 1998), allowing some level of customized or interactive information dissemination (Adlar and Blue, 1998; Khattak et al., 2008). The information could be descriptive, providing measures such as travel time to allow travelers to make their own trip modification decisions or prescriptive by providing specific guidance to travelers in selecting their routes, trip time, or mode of choices.

Third generation or next generation traveler information systems (Watkins et al., 2013) are in the research stage and expound on the existing ATIS. This chapter first describes the state of the art of ATIS, followed by a description of the next generation of ATIS.

2.1 ATIS Purposes

The Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU) legislation mandated the collection and distribution of real-time traveler information (Fischer and Meitner, 2007). The main objective of this mandate is to "provide the capability to monitor, in real-time, the traffic and travel conditions of the major highways of the United States." The main objective of ATIS is to deliver accurate, reliable, and timely transportation information to travelers so that they can make an informed decision about their daily trips. A study done for the Mid-Ohio Regional Planning Commission (MORPC, 2009) divided ATIS into three categories: pre-route planning, en route information sharing, and operational improvements. Adler and Blue (1998) pointed out that the ATIS has an implicit goal of supporting transportation management and operations. As such, ATIS has been found effective in reducing congestion (Gan et al., 2006), increasing system throughput (Yu et al., 2010), and increase safety (Al-Deek et al., 1993). Khattak et al. (1996b) showed that travelers usually switch to available alternative routes if they are aware of the delay on their current route. A survey conducted by Robinson et al. (2012) ranked the purpose of traveler information according to the responses by the agencies. The result of the survey is shown in Table 2-1.

Table 2-1: Reasons for Disseminating Traveler Information

No	Reason	Percentage voted for the reason
1	It allows the traveling public to make better travel decisions	95%
2	Federal Highway Administration (FHWA) guidance and encouragement	70%
3	Other agencies like ours are providing it	68%
4	It is a part of our agency’s overall demand management strategy	65%
5	The literature and past research indicate that it is an effective demand management strategy	57%
6	FHWA notice of the final rule on real-time system management information program	51%
7	We have an ongoing program for evaluating the provision of TI	40%
8	We have evaluation data that demonstrates its benefits	28%

Fernandez et al. (2009) found that ATIS is an effective tool used to reduce the effects of incidents in a transportation network, especially those incidents that cause significant deterioration of operational conditions of a network. The study found that 90% of the benefits of ATIS are achieved with 50% market penetration of ATIS users and the rate of the gain in benefits with additional increases in the market penetration is always decreasing. Balakrishna et al. (2005) also found that the average travel time-saving increases with the increase in the market penetration of ATIS; however, after about 50% market penetration, the saving starts to decrease.

ATIS also has a significant impact on on-time reliability. Wunderlich et al. (2001) found that during peak periods, commuters who use ATIS are six times less likely to arrive late, compared to commuters who do not use ATIS. In another report of the same project, Jung et al. (2002) showed that ATIS could reduce the travel disutility by up to 15% by reducing the late and early arrivals for travelers who are familiar with the network. This reduction was valued at $0.41 per trip. For the drivers who are unfamiliar with the network, this reduction was estimated to be higher, ranging from 25% to 34%.

Kuhn et al. (2014) also conducted a detailed study on the effects of traveler information on travel time reliability. They found that the provision of accurate reliability information will result in improved on-time performance by 9% to 21%.

Several studies have estimated the travel time savings due to the implementation of ATIS. A list of the findings from different studies is provided in Table 2-2.

Table 2-2: Maximum Travel Time Saving (%) Due to ATIS Implementation

Literature	Maximum Travel Time Saving (%)		
	Informed Traveler	Uninformed Traveler	All
Tsuji et al. 1985	11.5	-	-
Al-Deek et al., 1989	37	-	-
Koutsopoulos and Lotan, 1990	-	-	7.1
Kanafani and Al-Deek, 1991	-	-	4
Mahmassani and jayakrishnan, 1991	12	3	15
Mahmassani and Chen, 1991, 1993	40	10	20
Al-Deek and Kanafani, 1993	-	-	40
Gardes and May, 1993	6.2	-	5.3
Emmerink et al. 1995	30	25	25
Inman et al. 1995	15	-	-
Emmerink et al., 1996a	16	4	5
Hadj-Alouane et al., 1996	5.9	1.7	2.1
Schofer et al., 1996	1.6	-	-
Yang et al., 1996	4	3	3
Van Aerde and Rekha, 1996	18	8	10
Wunderlich, 1998	12.9	-	-
Adler et al., 1999	6	2	3
Bottom et al., 1999	-	-	20
Carter et al. 2000	8.1	-	-
Shah et al., 2001, 2003	4.8	-	-
Wunderlich et al, 2001	2.2	-	-
Anderson and Souleyrette, 2002	-	-	19.4
Jung et al., 2002	5.4	-	-
Lo and Szeto, 2002	-	-	5.4
Levison, 2003	40	30	-
Abdulla and Abdel-Aty, 2004	12.4	-	-
Vasudevan et al., 2004	0.7	-	-

2.2 Provided Information

The Society of Automotive Engineers (SAE) J2354 standards (SAE International, 2004) provide definitions of the message sets for ATIS to be used for delivering information to vehicles. According to the standards, the message set is divided into seven major groups of ATIS applications, as follows:

1. ***Traveler information:*** Traffic, incidents, events, weather, environmental conditions, and public transit schedules.
2. ***Trip guidance:*** Route a plan to a specific destination, including the mode of transportation, points of interest, etc.
3. ***Directory services:*** Electronic "Yellow Pages," possibly location-based.
4. ***Parking:*** Parking lot and space availability.
5. ***Settings:*** Traveler's personal preferences for format and content of traveler information.
6. ***Mayday:*** Emergency information, including requests for assistance and vehicle information.
7. ***Reduced Bandwidth:*** Streamlined version of certain data elements to accommodate bandwidth restricted media.

Part 511 of Title 23 of the Code of Federal Regulations (CFR) also specified the requirements for state traveler information programs. The regulation states that the provided information "shall include traffic and travel condition information for, as a minimum, all the Interstate highways operated by the states." Traffic and travel conditions are defined as those impacting traveler experiences, including road or lane closures, roadway weather or other environmental conditions, and travel times in metropolitan areas that experience recurring congestion. However, the 23 CFR 511 regulation does not require the dissemination of real-time information in any particular manner or technology, or to utilize any business model for collecting, processing and disseminating the information (ConSysTec and Cambridge Systematics 2013).

A study for Mid-Ohio Regional Planning Commission (MORPC, 2009) has identified twelve important functional areas that ATIS should cover. They are travel times, tourism, parking, rideshare, bike/pedestrian, transit, events, alerts, weather, construction, air travel, and maps. Different agencies could choose different functions and provide supporting information for the selected functions. Robinson et al. (2012) listed eleven distinct categories for traveler information, as follows:

1. Alternative routes
2. Live traffic cameras
3. Parking availability
4. Public safety information (e.g., amber alerts, silver alerts)
5. Roadwork/construction zones
6. Safety information (e.g., Buckle up, signal when changing lanes)
7. Special events
8. Travel times
9. Weather information
10. Other information types

Other similar research (Lappin, 2000) reported that experienced travelers seek the following information:

- Travel time between traveler's origin and destination

- Speed information for each highway segment
- Camera views that portray road conditions
- Incident information
- Coverage of all freeways and arterials
- En route access to good traffic information

Agencies are interested in the provision of information desired by most users and also in potentially modifying traveler trip decisions that improve system performance. Ng et al. (1995) analyzed a survey result which shows that the most common reason for choosing an alternative route is an accident (86% of the respondents to the survey), followed by road construction (79%), high traffic demands (71%), and time-saving by rerouting (49%). A survey conducted in the state of Michigan (Streff and Wallace, 199 3) showed that most drivers like to use more than one type of direction for driving in an unfamiliar area, with the highest percentage of travelers favoring verbal, written and maps (25.5%), followed by written and maps (25.4%), and followed by verbal and written (12.7%). Drivers use mostly maps (86.4%) for trip information services, followed by traveler information services (81.1%), and roadside signs (76.7%). The study suggested that "no single ATIS will be best or appropriate for all users." Based on the information above, guidelines for ATIS design and deployment were presented in the study, including:

- ATIS should provide more traffic information than just showing the path on a map.
- ATIS with voice and text supplements are more attractive.

- Timeliness is an important factor for the success of ATIS.
- There are large varieties of users who prefer different types of traveler information. Therefore, the public and private entities should have a wider consideration in ATIS deployment.

In another survey conducted by Hobeika et al. (1996), 70% of the respondents mentioned that weather, construction and traffic conditions are the top priority needs for pre-trip planning, and 80% of the respondents reported that alternative routes, construction, weather, and traffic are the top priority information for en route planning.

Therefore, ATIS should provide a wide range of information to fulfill its purposes. However, sometimes ATIS can overwhelm the traveler with unnecessary information. Therefore, the ATIS should provide preference options for the user to select the specific type of information they are seeking. For example, travelers driving in a familiar network may only seek minimum information about traffic such as incident details; however, travelers with an unfamiliar network normally seek more detailed information such as delay times and alternative routes (Yang et al., 1998). Also, elderly drivers need longer attention time for traveler information perception or cognition (Liu, 2000; Temple, 1989; Dingus et al., 1997). Thus, for elderly drivers, the information should be concise and simple (Liu, 2000).

Although different travelers may seek different information, their focus remains on how long the trip will take. Travelers want to know the impacts of incidents, congestion, special events, construction, and any other influencing factors on travel time. Travel time plays the most significant role in the traveler decision-making process, as shown in

Figure 2-1, which is called traffic (or traveler) information clock (EasyWay, 2012). The traffic information clock illustrates how a traveler sets the departure time and route depending on traffic conditions assessed based on historical travel time patterns, real-time travel time measurements, and travel time prediction. For this reason, the focus of this study is on travel time estimation using existing and emerging technologies.

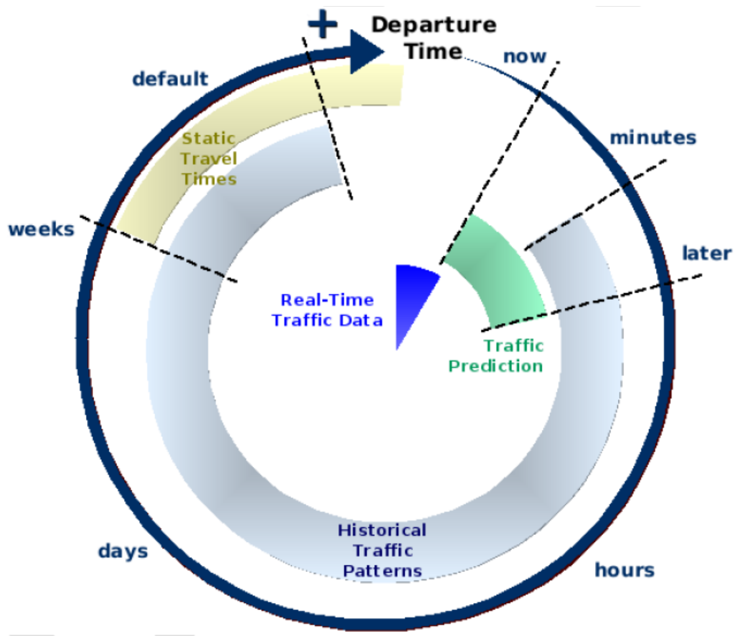


Figure 2-1: Traffic Information Clock (EasyWay, 2012)

2.3 ATIS Elements

The traveler information system is composed of several elements. In a Virginia Department of Transportation (VDOT) whitepaper (2002), traveler information was divided into three elements: data collection, data fusion and analysis, and information dissemination. These three elements work interconnectedly to deliver the required functionality of ATIS. Mouskos and Greenfeld (1999) mentioned four principle elements of ATIS: the types of users and their needs, transportation network surveillance, communications, and data processing. Before providing information, the developer of the

ATIS needs to know who will be using the information, along with their actual needs. Transportation network monitoring is the data collection process, which estimates the conditions of the network. Communication is the connection between the traffic monitoring system, traffic information system, and the users of the systems. Data processing includes different algorithms and processes such as data reduction, data filtering, data fusion, travel time estimation and prediction, traffic assignment, route planning, incident detection, and user interface. Figure 2.2 shows how those elements are interconnected. In this section, the whole ATIS elements are described in three subsections: data collection, data analysis, and information dissemination.

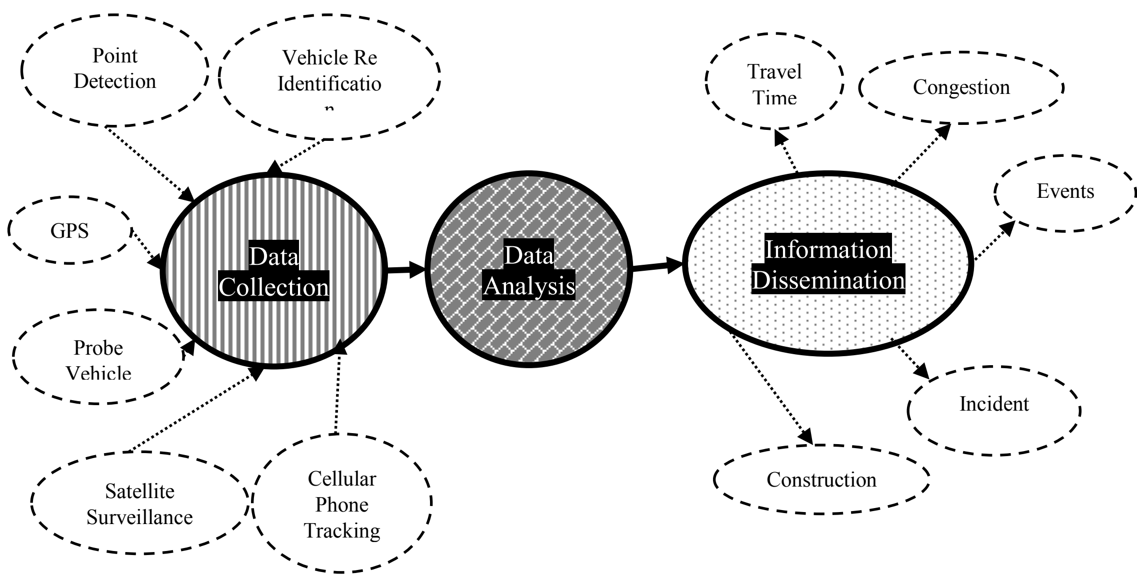


Figure 2-2: ATIS Elements as Outlined in the Literature

2.3.1 Data Collection

The traveler information system starts with the data collection process. This process includes collecting, retrieving, gathering and/or collating of real-time status information about the transportation network, which is served by ATIS. Data collection could solely be done by the public agencies or perhaps through acquiring data from other public and/or private agencies and providers. Some data are solely collected for traveler information systems, and some are bi-products of other data collection systems. There are several technologies available to collect the data; the main technologies utilized are described below.

Point Detectors

Point detection technology is the most commonly used technology for traffic monitoring. The use of point detection technology is widespread. This technology has been used for different applications, including transportation planning, traffic management, emergency management, evacuation monitoring, and traveler information systems. There are two types of detectors (Klein et al., 2006): pavement intrusive detectors and non-pavement intrusive detectors. Pavement intrusive detectors are placed underneath the pavement. Examples of this type of detectors are the inductive loop, magnetometer, and magnetic detectors. Non-pavement intrusive detectors include microwave radar, active infrared, ultrasonic, acoustic, and video image processor technology. In general, point detectors provide volume counts, point speeds, occupancy, and in some cases, vehicle classification. Travel time is estimated for segments based on point detector measurements using methods with different levels of sophistication.

Vehicle Re-identification Technology

There has been an increase in the use of vehicle re-identification technologies in recent years. This technology performs travel time estimation and origin-destination

(OD) matrix estimation. There are several types of vehicle re-identification technologies (Jeng, 2007; Oh et al., 2003).

Traditionally, vehicle re-identification technology has included Electronic Toll Collection (ETC) and License Plate Reader (LPR). ETC readers collect the vehicle tag identification at different locations for the estimation of traffic parameters, such as travel time and OD matrices. LPR systems, on the other hand, are based on license plate reading with the use of image processing technologies.

Due to their lower cost and installation flexibility, Bluetooth and Wi-Fi readers have become the most widely used vehicle technologies in recent years. With these systems, vehicles traveling with Bluetooth or Wi-Fi devices are identified by a reader at upstream and downstream locations. The vehicles with matched Media Access Control (MAC) addresses are used to measure travel time.

Vehicle signature-based systems have been proposed that identify and use vehicle signatures based matching. Loop detectors and magnetometers have been used to provide the inputs to the vehicle signature recognition with a magnetometer-based system, which is currently commercially available.

Tracking-based Probe Vehicle Technologies

Vehicle tracking-based probe vehicle techniques estimate travel time and OD matrices. These techniques utilize Automatic Vehicle Location (AVL) technologies such as the Global Positioning System (GPS) and cellular probe. AVL is a system that automatically determines and transmits the geographic location of a vehicle. The most common way of determining the location is the use of GPS, which utilizes satellite and land communications to determine a vehicle's location, status, heading, and speed. Another way to collect probe vehicle data is to track cell phones within the cellular network (Qui and Cheng, 2007). This data is then used to determine the travel time and other traffic-related information.

Private Sector Data

Private-sector data are generally collected from tracking vehicles and/or mobile devices. Some of the vendors combine the mobile data with data from other sources such as point detectors in their estimation processes. The data are used to estimate travel time. Trip O-D tables are also available from some vendors. Sufficient sample sizes of the collected data are necessary to provide the required accuracy. This should not be a problem for urban freeways, particularly in time periods with sufficient demands. However, it could be an issue for some urban streets and rural freeways, particularly during off-peak periods. Nevertheless, there has been a significant increase in the number of tracked devices by the private sector. This trend is expected to continue, resulting in further improvements in the quality of data. Examples of private sector vendors are INRIX, HERE, Google/WAZE, TomTom, and Airsage. A 2012 public sector survey (Crowson and Deeter 2012) indicated that 18% of the respondents were using third-party private sector data to support their processes.

Connected Vehicles

Connected Vehicle (CV) technology is increasingly being considered by agencies for possible implementation in the next few years. The CV requires the use of wireless communication for Vehicle to Vehicle (V2V) and Vehicle to Infrastructure (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 and documented in the Society of Automotive Engineers (SAE) J2735 standards (SAE International, 2016).

The BSM contains vehicle safety-related information broadcasted by vehicles 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, which is expected to be mandated and broadcasted by the National Highway Traffic Safety Administration (NHTSA) ruling, will be sent in every BSM message and broadcasted ten times per second. 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. These data could be a potential source of traveler information.

2.3.2 Data Analysis

For best utilization, data collected from multiple sources and multiple technologies are considered together for data analysis. This data collection effort poses significant issues, such as the need to certify the quality and timeliness of the data, and coordinating data collection efforts throughout the region.

The data collected using single or multiple technologies described in the previous section should be combined and processed to obtain the required traffic information. This data fusion can be conducted as a centralized or distributed data fusion. In the centralized data fusion process, data from all participating regional partners are collected by one regional entity that will store and maintain the information. On the other hand, in the distributed data fusion process, each agency will collect and maintain the data. The agencies work together and share their information to have a consistent view of the transportation network throughout the region.

Traveler Information Quality

The European Union (EasyWay, 2012) defined a standard for the levels of quality of services provided by traveler information services. The core criteria for this quality assessment are as follows:

1. ***Accessibility*** is the percentage of area covered by the service.
2. ***Availability*** is the percentage of time the service is available.
3. ***Timeliness*** is the time delay between the event detection and the information distributed to the users.
4. ***Update frequency*** is the frequency of updating information or data update interval.
5. ***Quality assurance*** is the application of standards.
6. ***Reliability or cross-verified*** is the degree of certainty of the information.
7. ***Accuracy*** is the error of the estimation.

The following measures have been recommended by different studies (Cambridge Systematics 2011, FHWA 2014) to assess the quality of data, which can be also used for the purpose of this study:

- **Accuracy** is defined as the degree of agreement between a data value or set of values and a source assumed to be correct.
- **Completeness**, also referred to as "Availability," measures how much data is available, compared to how much data should be available, and is typically described in terms of percentages or number of data values. Completeness can refer to both the temporal and spatial data availability.
- **Validity** can be expressed as the percentage of data values that pass or fail data validity checks.
- **Timeliness** sometimes referred to as "latency," reflects the latency that the data are provided at the time required.
- **Coverage** is the degree to which a sample of the data accurately represents the entire population.
- **Accessibility** reflects the relative ease with which data can be retrieved and manipulated by data users.

A survey conducted as part of the Enterprise pool-funded study (Crowson and Deeter 2012) asked agencies to describe any accuracy requirements that they may have. Their responses as references to Federal Regulation 23 CFR 511 (Table 2-3) showed that 95% or greater accuracy and latency were expected.

Table 2-3: Information Quality Criteria Presented in Federal Regulation 23 CRF 511

	Travel Time	Weather Information	Construction Activities	Traffic Incidents
Accuracy	>85%	N/A	>85%	>85%
Availability	>90%	>90%	>90%	>90%
Timeliness	<10 min.	<20 min.	<10 min. urban <20 min. rural	<10min. urban <20 min. rural

Deeter (2009) pointed out that travelers expect accuracy, timeliness, reliability, convenience (ease of access and speed), and safety (of operation) with the provided traveler information system. A study (Fox and Boehm-Davis, 1998) was conducted utilizing a driving simulator and concluded that people will trust the ATIS if its accuracy is at least 60%.

Travel Time Estimation

As this research focuses on travel time and volume information of ATIS, this section describes the related existing literature on travel time estimation and the associated accuracy. The subsequent section will describe the existing literature on volume estimation and the associated accuracy. Turner et al. (1998) recommended criteria to compare travel time data collection techniques, including initial cost; operating efficiency; required skill or knowledge level; data reduction and/or processing; route flexibility; accuracy; and sampling rate over time, space, and vehicles.

Point detectors have been used to measure the traffic speeds on uninterrupted facilities like freeways. These speed measurements are utilized to estimate the travel times between segments. A Nebraska study's assessment (Grone, 2012) of speed measurements found 95% accuracy for two widely used true presence microwave detectors tested in the study, 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 accuracy of less than 1 mph. It should be mentioned, however, that the estimated travel time based on these speeds will have higher errors than the speed measurement errors at the point detection locations. This error depends on the errors in point detector speed measurements, distances between detectors, the method of travel time estimation from point detection, the congestion conditions, and the speed of the movement of the congestion shockwave. In research conducted by this project researcher (Xiao, 2011), based on simulation analysis, it was found that the mean absolute percentage errors when there is a random error of 90% are 1.3% and 1.6 % for uncongested conditions when the detector spacing was 0.3 and 0.6 mile, respectively. The corresponding values for congested conditions during lane blockage incidents were 10% and 20%, respectively.

Estimating travel times based on point detectors is only possible for uninterrupted facilities. There has been an increasing use of automatic vehicle re-identification techniques and third-party vendors for travel time estimation in recent years. A study conducted in the state of Washington evaluated the accuracy of these methods to measure the travel time of an urban arterial and a rural freeway (Wang et al., 2014). The tested commercial products were an Automatic License Plate Reader (ALPR) system matching based on vehicle signatures using magnetometers, a Bluetooth and Wi-Fi-based systems, and a third-party feed. The ALPR system used a ground truth system since it had been previously evaluated and proven to be accurate enough to serve as the source of ground truth data. The study found that the sample size of data from a Bluetooth device can be as low as 5-7% of the total volume. However, other devices have higher penetration rates (11% to 13%), and when combined with Wi-Fi data, can reach a sample size of 26-32%. The sample size of the magnetometer-based sensors was 82%-100% of the total volume. The Mean Absolute Percentage Error (MAPE) for Bluetooth devices and Bluetooth plus Wi-Fi devices compared to ALPR ranged between 13% and 20%. The MAPE of the magnetometer-based measurements ranged between 18% and 25%. The third-party vendor measurements ranged from 15% to 48%. It was also reported that the third-party vendor data significantly underestimated the travel time on urban arterials and could be less responsive to traffic conditions due to smoothing.

There are several challenges when estimating travel times for urban arterials, including lower volumes (thus lower sample sizes), interrupted flow operations that cause variations in travel times in time and space, driveways, and adjacent land uses, and activities that may affect the data collection efforts. An evaluation of probe data for the I-95 coalition in 2015 (Young et al., 2015) found that third-party probe data have a reasonably good accuracy rate on arterial streets when the number of signalized intersections per mile is less than or equal to 1 on principal arterials with an Annual Average Daily Traffic (AADT) of 40,000 vehicles per day or more. However, this evaluation increasingly underestimates congestion as the number of intersections increase due to the increase in the variation in travel times and the volume decrease due to the small sample size. Specifically, the study (Young et al., 2015) found that:

- Probe data is recommended for AADT > 40,000 vehicles with two or more lanes in each and one or fewer signals per mile.

- Probe data should be examined for AADT between 20,000 and 40,000 with two or more lanes in each and one to two signals per mile; some segments produce satisfactory results, while others may not.
- Probe data are not recommended for AADT less than 20,000 vehicles per day, more than or equal to two signals per mile.

An FHWA project conducted by Toppen and Wunderlich (2003) reported the required accuracy of travel time data collection for the measurement of ATIS. The results showed a utility benefit curve at different ATIS travel time errors for Los Angeles. Figure 2-3 shows that in order to get positive utilities of using travel time as ATIS, the error should be less than 17% for all trips.

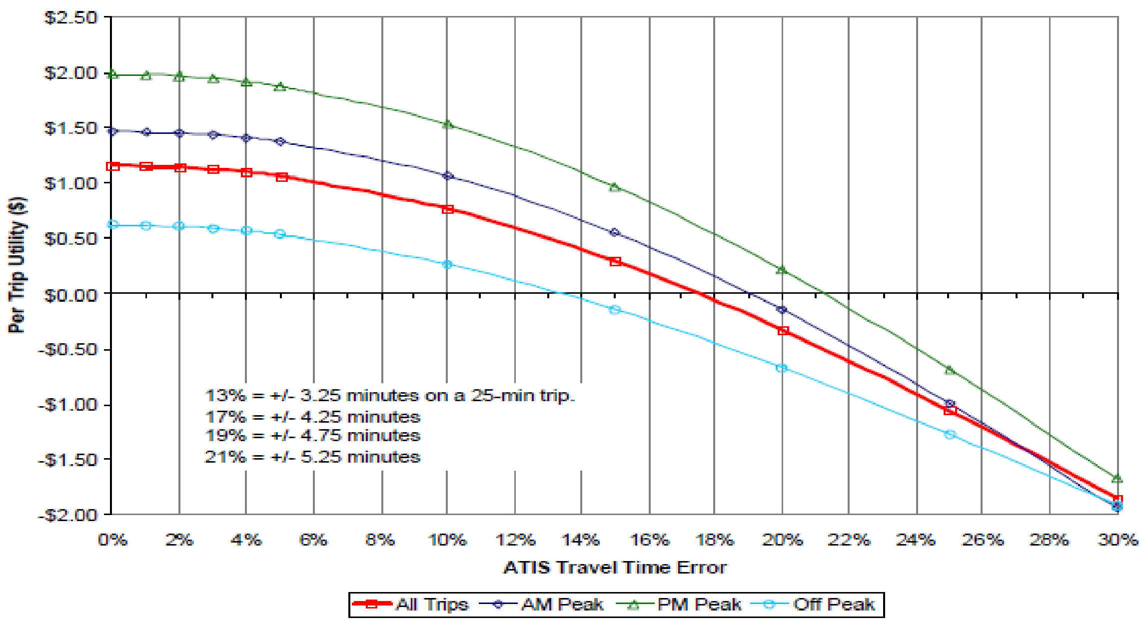


Figure 2-3: Benefit-Accuracy Relationship for Los Angeles (Source: Toppen and Wunderlich, 2003)

There are several studies that focus on the potential of travel time estimation using connected vehicle data. Doan et al. (2010) proposed a method to determine travel time using the vehicle to infrastructure communication data. They assumed that every connected vehicle will have a fixed identification when estimating travel time. It should be mentioned that current CV standards specify maintaining CV Identification (ID) for about five minutes to protect privacy. They used the VISSIM simulation software in their assessment. Zou et al. (2010) estimated the link travel time based on connected vehicle probe data. They also used the VISSIM software to simulate a real arterial with five links and six signalized intersections. The average error percentages found in their study were 27.6%, 12.5%, and 8.2% for 1%, 5%, and 10% market penetration, respectively. Vasudevan et al. (2015) proposed a method to calculate travel time and queue length estimates system using CV data message sets. Their methodology did not use the vehicle ID to determine travel time. However, this was an initial study which did not examine in-depth the accuracy of the estimation and the influencing factors. Tian et al. (2015) proposed a method to estimate dynamic travel time based on connected vehicle data.

Based on a simulation, Oh and Jayakrishnan (2002) found that travel time accuracy is a function of the increase in the market penetration of probe vehicles. They found that the Link Travel Time Estimation Error (LTEE) decreases as the market penetration of probe vehicle increases; however, for longer update intervals (three to five minutes), the LTEE does not decrease much beyond a certain market penetration (30%-40%). The Route Travel Time Estimation Error (RTEE) also decreases with the increase in market penetration up to 20%. After 20%, the RTEE remains constant with shorter update intervals but increases with longer update intervals (three to five minutes).

Hellinga and Gudapati (2000) examined travel time accuracy using loop detectors, probe vehicles, and driver reports. The study used simulation in the evaluation. This study found that the use of probe vehicles produces the most accurate estimation among these methods, with 26% RMSE of the mean link delay at 10% market penetration.

Volume Estimation

Roadway traffic volume is also a critical input for the traveler information system processes. Travel time, along with volume data, is an indicator of roadway congestion and a traffic incident. As described in Section 2.3.1, 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 network handover data to estimate the volume. They used available citywide traffic counter data combined with 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 produced a mean absolute percentage error (MAPE) less than 17%. Nantes et al. (2013) estimated traffic volume using Bluetooth data. They proposed a Bayesian network to estimate the volume.

Vehicle trajectory data collected utilizing GPS data has also been used to estimate the traffic volume. Zhan et al. (2017) used a hybrid framework that combines machine learning techniques and a traffic flow theory to estimate traffic volumes. They evaluated the proposed methodology utilizing GPS datasets from 33,000 Beijing taxis and volume ground truth data obtained from 4,980 video clips. Anuar and Cetin (2017) used probe vehicle trajectory data and a machine learning technique that had been combined with a shockwave theory to estimate the volumes and found an average error of 5% in the volume estimation when using this method.

Studies have also been conducted 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 (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).

Regarding 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 a 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 CVs 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.

2.3.3 Information Dissemination

After collecting and processing the traffic data, the extracted traveler information needs to be disseminated to travelers. Public agencies use multiple outlets to disseminate the information. To disseminate information, it is important to know the potential users and their characteristics. Mehndiratta et al. (2000) divided the potential ATIS users into several segments depending on how much information they want to know about their trips, how comfortable they are with updated technologies, how much they are willing to pay for information, and what level of convenience they demand in receiving the information.

Toledo and Beinhaker (2006) compared the potential benefits of the ATIS system with different types of routing information. The different types of routing information examined in the study were static, historic, instantaneous, and predictive travel time. The instantaneous travel time is travel time measurement for the current interval (when the information is provided). The static travel time is calculated as the length of the link divided by the speed limit. The historical travel time is based on time-of-day travel times from previous days. The study found that routing could save travel time by up to 14% and reduce the travel time variability by up to 50%, depending on the type of information. It also found that the use of predictive travel time for routing could provide the largest savings, followed by instantaneous en route, instantaneous pre-trip, and historical information routing.

A study by Sen et al. (2001) showed that travelers made diversion decisions based on mean travel and travel time reliability. It was reported that the traveler may take a route with a lower travel time variability, although it may not be the route with the lowest mean travel time.

Kim et al. (2009) simulated wireless V2V and V2I communications to investigate its effects on dynamic advanced traveler information systems. Both centralized and decentralized ATIS architectures were tested in this study. Travel time was the key consideration in the evaluations. The results showed that both architectures can save travel time during incidents by providing information to the connected vehicles with alternative routes. However, the study found that the centralized architecture saves more travel time at lower market penetrations; however, the decentralized architecture works better at higher market penetrations.

A study done by Robinson et al. (2012) examined agencies' utilization of information dissemination methods, as shown in Figure 2-4.

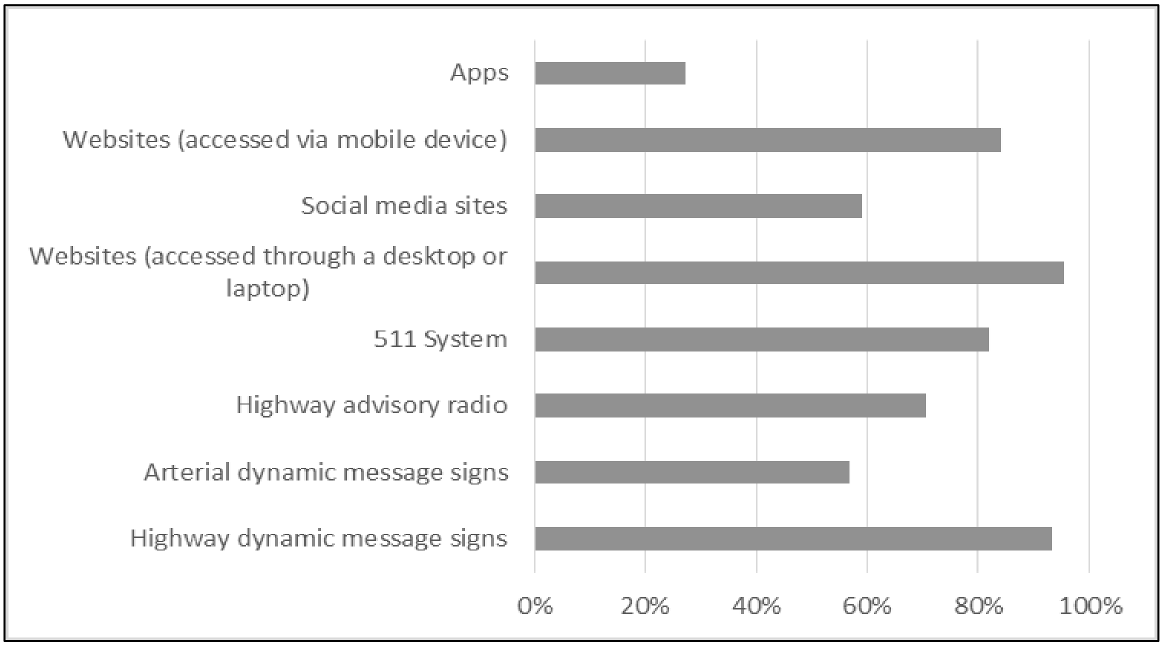
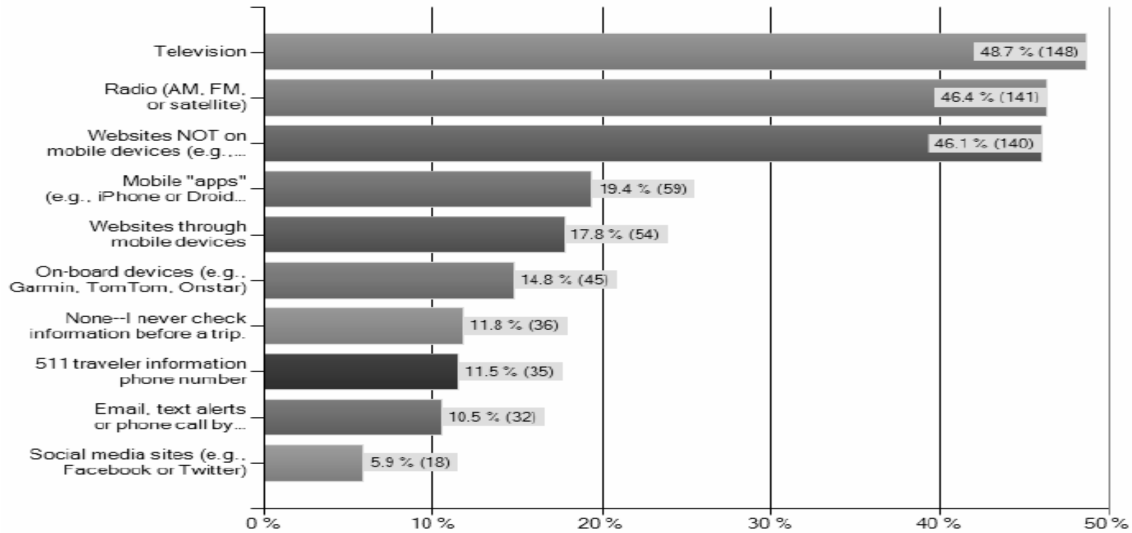


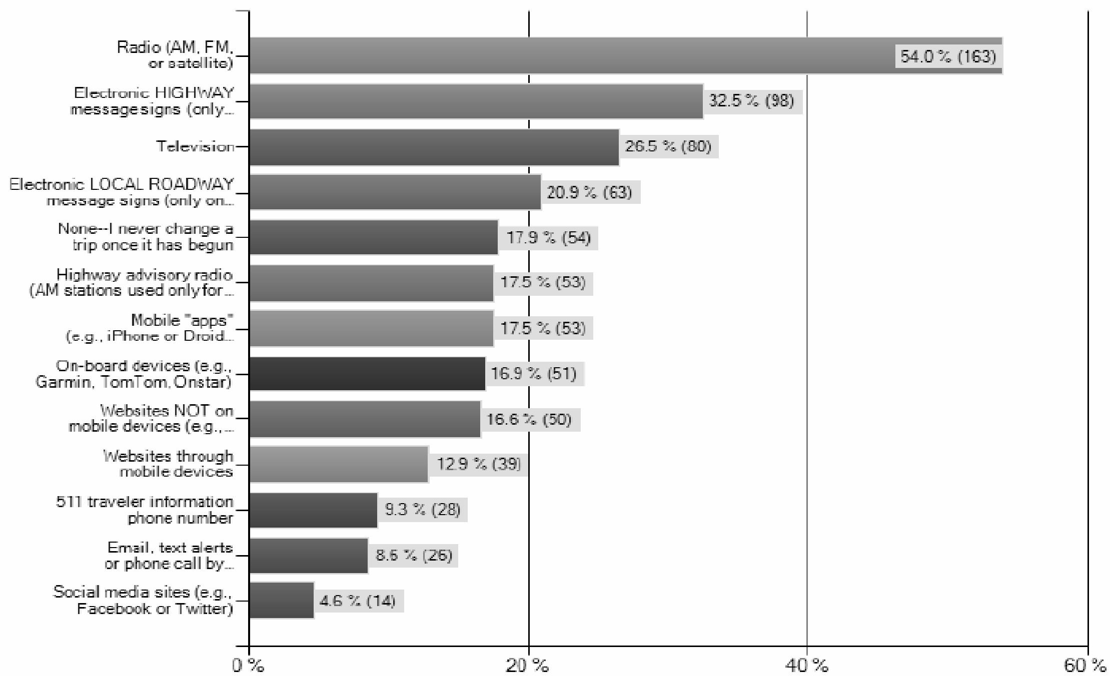
Figure 2-4: Percentage of Agencies Using Different ATIS Dissemination Methods (Source: Robinson et al., 2012)

Figure 2-4 shows that the most popular means of traveler information dissemination at the time of the study (2012) were DMS, websites, 511, media outlets, and HAR. This study also conducted a survey of travelers to determine the source of traveler information used. Figure 2-5 summarizes the results.

The same study also did an on-road naturalistic driving study by collecting traveler logs or diaries and supplemental surveys. The results showed that traffic incident information is the most popular traveler information type used to change a trip. The results are shown in Figure 2-6.



(a) For pre-trip planning



(b)For en route planning

Figure 2-5: Percentage of Travelers Using Different ATIS Dissemination Methods (Source: Robinson et al., 2012)

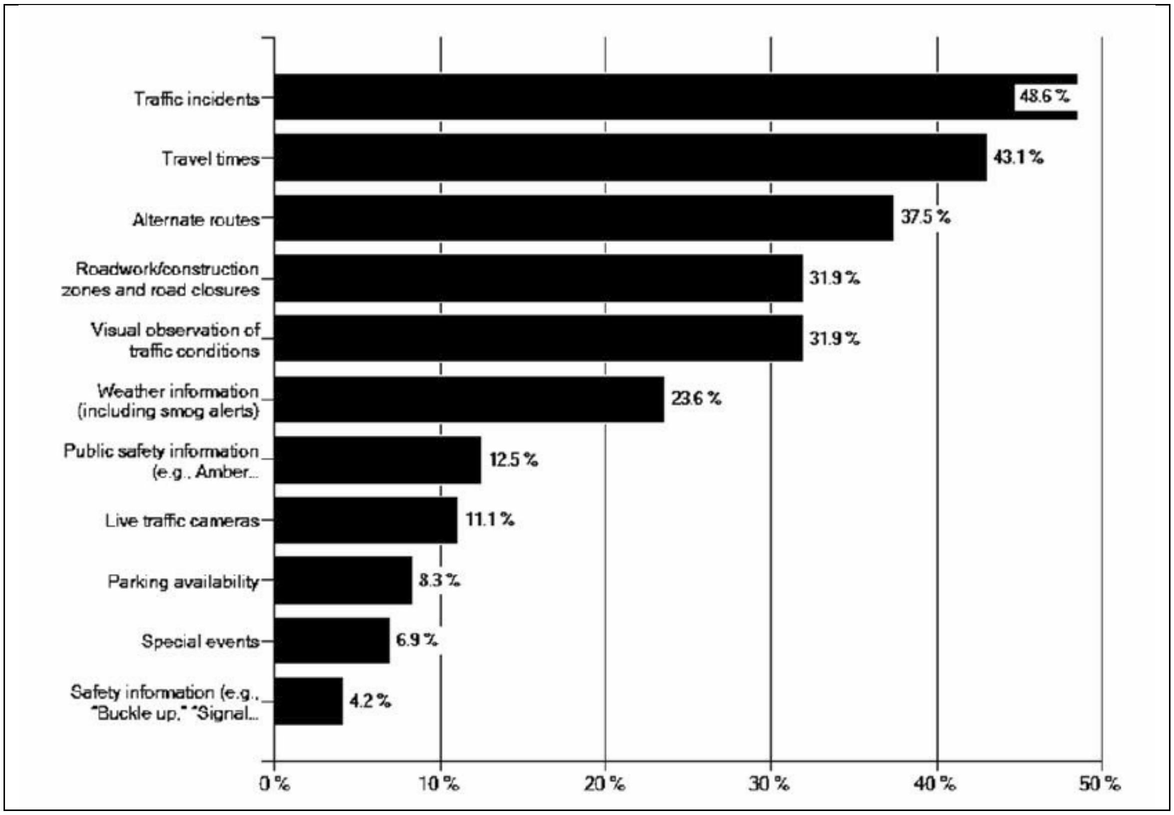
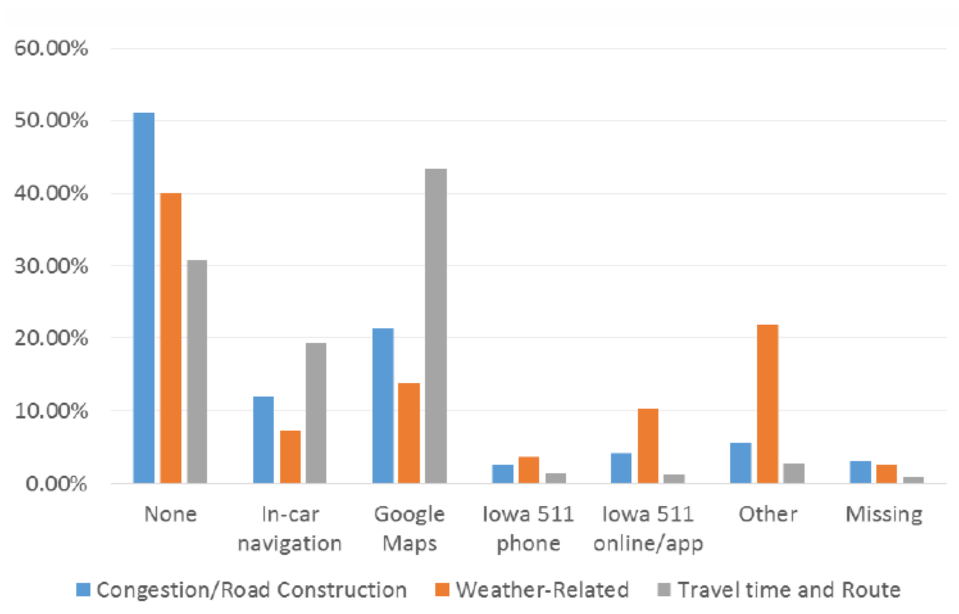


Figure 2-6: Percentage of Travelers Using Different Traveler Information Types to Change a Trip (Source: Robinson et al., 2012)

Zhang and Levinson (2008) conducted a field experiment to determine the value of traveler information. They found that information accuracy, a positive attitude toward information services, commute time, household vehicle ownership, and ownership of computers have a positive effect on the use of traveler information. The elderly population older than 55 years was found to use traveler information less often than others. In another study, Chen and Jovanis (2003) found that freeway advice, turning advice, congestion occurrence, incident occurrence, subjects' spatial experience, temporal experience, and education level are significant factors that impact drivers' en route guidance compliance.

A survey conducted for Iowa Department of Transportation (Sharma et al., 2015) also identified the relative use of different traveler information services, as shown in Figure 2-7.



Other: TV/News (2%, 11.18%, 0.59%); Radio (1.88%, 2.59%, 0.12%); Phone apps other than 511 (0.82%, 4.24%, 1.06%)

Figure 2-7: Traffic Information Services Used by the Survey Respondents (Source: Sharma et al., 2015)

Existing Public-Sector Roadside Subsystems

Existing roadside subsystems managed by the public sector includes DMS and HAR. These subsystems mainly influence a traveler's en route decisions. Yu et al. (2010) concluded that the DMS subsystem with travel time information provision can increase the system throughput up to 8-10%. Another study done by Gan et al. (2006) indicated that 87% of users always notice DMS, which can increase the free-flow duration time by 20% and reduce the congestion time by 20%. A survey conducted by Robinson et al. (2012) found that 27.5% respondents rated the information on the DMS as extremely accurate. More than 32% of respondents used this medium of traveler information to make decisions in changing their trip routes.

The HAR also had a good number of users in 2012 (~18%), but users complained about its sound quality, usefulness, and timeliness of information (Robinson et al., 2012). A study conducted by Al-Deek et al. (2016) for the Florida Department of Transportation (FDOT) assessed the impacts and usefulness of HAR in providing traveler information. They conducted a survey through the use of the phone, the Web, and face-to-face meetings. This study found that only 57% of travelers were aware of HAR, and only 24% had used it. During emergencies, the percentage of use is higher (87%). Hence, 87% of travelers and 70% of the local agencies said HAR should be continued in the case of emergency situations. The benefit-cost analysis assumed that a 10% diversion was caused by HAR messages. This was estimated to result in a benefit-cost ratio of 1.19, and this value was estimated to be 11.91 if the diversion is 100%. This study pointed out that travelers are increasingly favoring the use of smartphones to receive traveler information. Thus, it recommended integrating the HAR in the smartphone.

Media Outlets

Commercial radio, commercial TV, and newspapers are information dissemination outlets that reach the largest audience at the least cost. People obtain weather-related information mostly from media outlets such as TV and radio (Sharma et al., 2015). An earlier study in 1998 conducted by Al-Deek et al. (1998) showed that radio broadcasts were more beneficial in terms of reducing the average travel time of a corridor during an incident, compared to public sector traffic advisory infrastructure. Another earlier study by Ng et al. (1995) analyzed survey data and found that 57% of private vehicle users choose alternative route after obtaining information from commercial radio, followed by observation of traffic conditions (54%) and commercial TV (14%). Khattak et al. (1996b) conducted a survey which shows that the sources of unexpected traffic congestion are radio (72%), television (30%), own observation (22%). Robinson et al. (2012) found that approximately 55% of travelers use the radio for their en route decisions, which is the most widely used in trip decisions.

Use of Internet by Public Agencies

The introduction of the World Wide Web (WWW) to the public in the early to mid-1990s influenced the creation of a new generation of traveler information systems. The use of the Web is considered the first paradigm shift in the field of traveler information (Deeter, 2009), providing an effective way to disseminate information to a wide range of travelers. According to a study done by Deeter (2009), the internet affects traveler information in two distinct ways:

a) For public agencies, the introduction of the internet provides a large amount of traveler information for a state by creating a single website. People who have internet access can visit the web page and find relevant information according to his/her needs.
b) It allows private sector information service providers to create local, regional, or nationwide traveler information systems and reaches travelers nationwide.

In 2013, the internet was the second most popular type of technology used by the public for news and other information (NTIA Report, 2013). A 2008 study done by Khattak et al. (2008) also showed that the internet was the most influential type of technology in changing travel patterns (e.g., time, mode, route, trip cancellation). The use of the internet could be through an agency's website or the use of smartphone applications or even social media. With the increasing popularity of using smartphone applications and social media, agencies are now creating their own smartphone applications and social media pages to disseminate traveler information to travelers. In some places, travelers can subscribe to alert services for information about roadway conditions and emergencies via email and/or text messages. The smartphone applications are also increasingly popular. Several public agencies have created mobile apps for the dissemination of traveler information. Figure 2-8 shows the number of users of different websites and mobile apps in Iowa (Sharma et al., 2015). The study also showed that the use of every media increases during a severe winter storm.

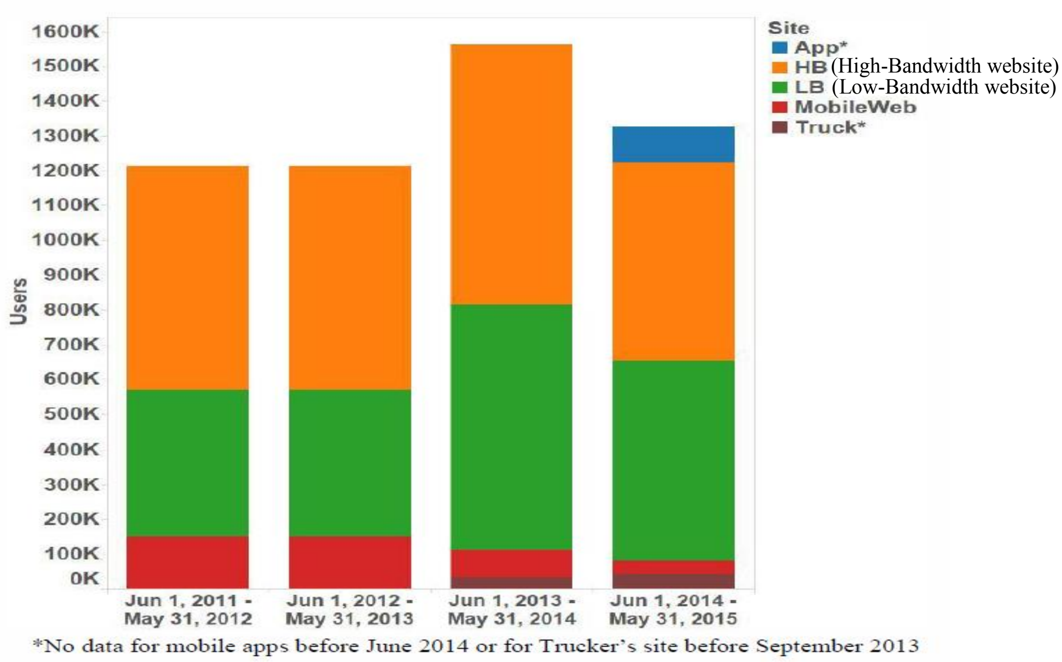
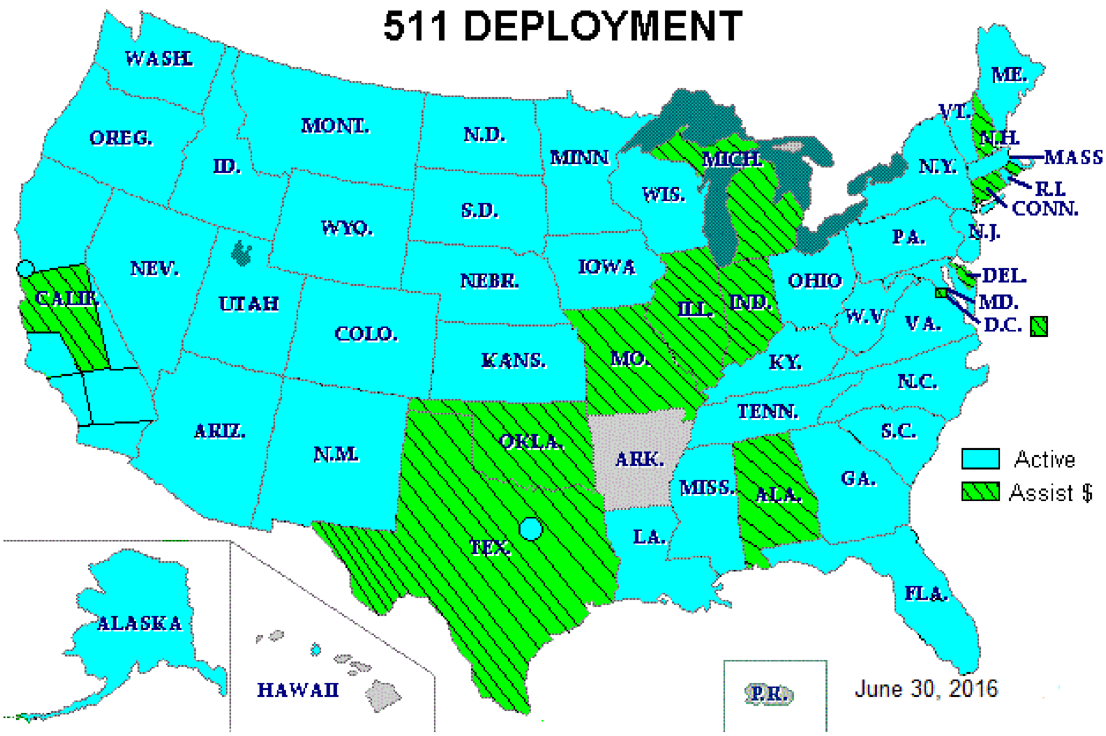


Figure 2-8: Users of Iowa 511 Websites and Mobile App (Source: Sharma et al., 2015)

Telephone Delivery Systems/511

The 511 or telephone delivery system was the second largest paradigm shift in the traveler information industry (Deeter, 2009). The Federal Communications Commission (FCC) reserved the 511 service as a phone number for nationwide traveler information services, which began in July of 2000. The 511 service has expanded for use in most of the states. The phone call system includes a combination of a live operator, is touch-tone activated, contains a multi-language option (English/Spanish), and has a phone-based interactive voice recognition system (IVR), although most current systems use IVR.

The 511 deployment coalition (FHWA, 2005) defined its vision as "511 will be a customer-driven, multi-modal travel information service, available across the United States, accessed via telephones, and other personal communications devices, realized through locally deployed interoperable systems, enabling a safer, more reliable and efficient transportation system." By June 30, 2016, the 511 phone service was active in 46 locations throughout the United States (FHWA website). Figure 2-9 shows the deployment status of 511 throughout the country.



* *Active locations mean the 511 service is available and Assist locations means those have received funding under the 511 planning assistance program*

Figure 2-9: 511 Deployment Status (Source: FHWA website)

Although the 511 phone system is deployed in most states, it is still lacking publicity at an expected level. In a survey conducted in 2012, it was found that most of the respondents were unaware of the 511 system (Robinson et al., 2012). The 511 systems were encouraged to be used for pre-trip planning rather than en route decisions due to safety and legal issues of using a cell phone while driving. Another drawback of this system is that it normally covers only major roadways.

Private Sector Smartphone Applications

With the enormous popularity of smartphones nowadays, smartphone transportation system applications have possibly become the most popular means of disseminating information. Many private sector information providers are adding more flexibility and options to plan their trips with real-time traffic information using such apps. According to Robinson et al. (2012), the apps should have the following functions:

- Have a voice option for use while driving
- Are interactive and allow customization
- Are GPS-enabled and provide local, relevant information to a traveler
- Provide real-time information

Most of private sector smartphone applications provide these functions. A survey conducted by Sharma et al. (2015) found that a high portion of travelers uses Google Maps for travel time information (about 42%) and congestion, road construction and closure information (about 24%). Other popular apps include Apple Map, WAZE, TomTom, and Inrix.

Traditional In-vehicle Systems

The in-vehicle traveler information system may include a navigation system, voice service, and/or emergency system. A survey conducted by Khattak (1993) found that 46.4% of the participants reported that the inclusion of construction and incident information will increase the willingness to use in-vehicle information systems. A large portion of users found in-vehicle information systems to be reliable in providing alternative routes and traffic congestion information (Robinson et al., 2012). Recently, there has been an increase in the internet connectivity of onboard devices to provide real-time information about traffic. The in-vehicle information system could be visual, auditory or both. It was found that people are less prone to errors when there are both visual and auditory systems for traveler information or navigation (Liu, 2000; 2001;

Yang et al., 1998). Connected vehicles will be equipped with the next-generation dynamic in-vehicle navigation system.

2.4 Future of Traveler Information System

With recent and expected advancements in technology, there have been tremendous changes and potential changes in traveler information systems. Moreover, various new features will play a vital role in changing the system in the future. Robinson et al. (2012) identified HAR, social media, public agency apps and 511 call systems as less effective according to both agencies and the public. However, these systems are still being used by transportation agencies for information delivery. The main drawback with the HAR is the sound quality. A very limited number of people use social media for traveler information, although the rate of using the social media itself is very high. Public agency applications will be most challenged by private sector applications, which provide more user-friendly options and flexibility to customize public needs. In the case of the 511 phone system, although agencies are investing to improve the system and make it more popular, the main hindrance to its implementation is the limited awareness of its existence and the ban on cell phone use while driving in some states.

With an extensive survey and research, Robinson et al. (2012) provided a prediction about future dissemination technologies that may come to the market, as shown in Figure 2-10.

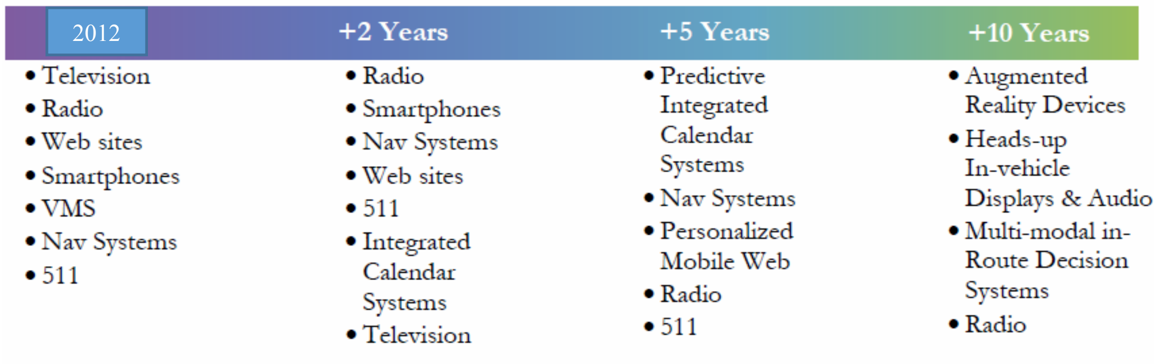


Figure 2-10: 511 Prevailing Dissemination Technologies (Source: Robinson et al., 2012)

There are several smartphone applications that provide navigation with real-time traffic information. The percentage of smartphone use is increasing, thus, the use of navigation will increase. Some of the new vehicles now have in-vehicle navigation systems, and this feature is expected to become more popular with the introduction of CV. The modern in-vehicle navigation system could connect to the internet using cellular connectivity and could also have live traffic information. DSRC communication will also provide information to in-vehicle devices from roadside units and possibly other vehicles.

It is anticipated that next-generation ATIS will include the use of artificial intelligence that learns a traveler's regular travel patterns and suggests the best routes for travel and times for starting the trip. Augmented Reality (AR) devices could be also useful in the future to support ATIS. AR is a live, direct or indirect view of a physical, real-world environment whose elements are augmented by computer-generated sensory input such as sound, video, graphics, or GPS data (Robinson et al., 2012). AR is now being used as part of various smartphone applications. Google, Yelp, and many other companies are now using AR to guide people in their tour plan. For example, the iPhone application called "Yelp Monocle" uses the smartphone's GPS and compass to display

AR markers for nearby restaurants, bars, and other businesses. The application can additionally provide directions to nearby AR marker places. The San Francisco transit agency is using an AR application to help their transit travelers navigate their way to stations by holding their camera phone and pointing it in a direction, and then seeing the transit station entrances highlighted on their screens.

The future of ATIS also depends on the rate of technology adoption by travelers. Goulisas et al. (2004) found that elderly individuals are less likely to be aware of the traveler information availability offered by emerging technologies. On the other hand, people with professional employment, and people who increased in car ownership and in the number of children were found to be highly correlated with a higher use of advanced technology. Persons who work five days a week were found to be more likely to use ATIS services.

Adler and Blue (1998) mentioned that the third-generation traveler information system would be an Intelligent Traveler Information Systems (ITIS), which would be the result of the integration of Artificial Intelligence with ATIS. ATIS will learn a user's travel characteristics and preference from his/her day-to-day travel and input interactions. This learning will help the ATIS suggest the user's preferred trip information automatically. For example, the vehicle could suggest trip recommendation, depending on the time of day and day of the week.

Connected vehicles can be an outlet for traveler information dissemination. An ENTERPRISE pooled-funded study (2015) on V2I Pre-Deployment planning provided a guideline for ATIS that provide en route in-vehicle dynamic signing to inform travelers of traffic conditions. The main purpose of this signing would be to provide current traffic status information (incidents, congestion, travel time, and road work) to drivers so that they can choose to divert or avoid a potentially unsafe situation, reduce driver anxiety, and reduce crashes involving drivers encountering unexpected stopped traffic.

The Connected Vehicle Reference Implementation Architecture (CVRIA) developed by the USDOT includes the traffic condition information dissemination function within the “Provided Driver and Traveler Services” process (Iteris website, 2016). Kim (2010) developed an advanced traveler information system using V2V and V2I communications. The study assessed the system using simulation on a non-signalized simple network, signalized simple network, and signalized grid network with different communication systems. It was found that the average travel time saving per vehicle reached the marginal travel time-saving at 60% penetration of ATIS. Travel time benefits increase with the increase in the market penetration, traffic demand, and communication radio range. Vehicle re-routing could produce travel time savings up to 20% and 30% in the signalized traffic network. The research also found that at lower market penetrations, the use of V2I works better. However, at higher market penetrations, V2V communications can provide promising results.

Enabling Advanced Traveler Information System (EnableATIS) refers to the traveler information component of the USDOT Dynamic Mobility Application (DMA) program, which is a part of the connected vehicle. This program has proposed two operational concepts, as listed below (Adler 2014, Burgess 2012).

The first concept (Figure 2-11), referred to as the laissez-faire operational scenario, is a continuation of current advancements in ATIS with an incremental enhancement over time, assuming an increasing level of data and data processing and use, as well as continued innovation in delivery mechanisms.

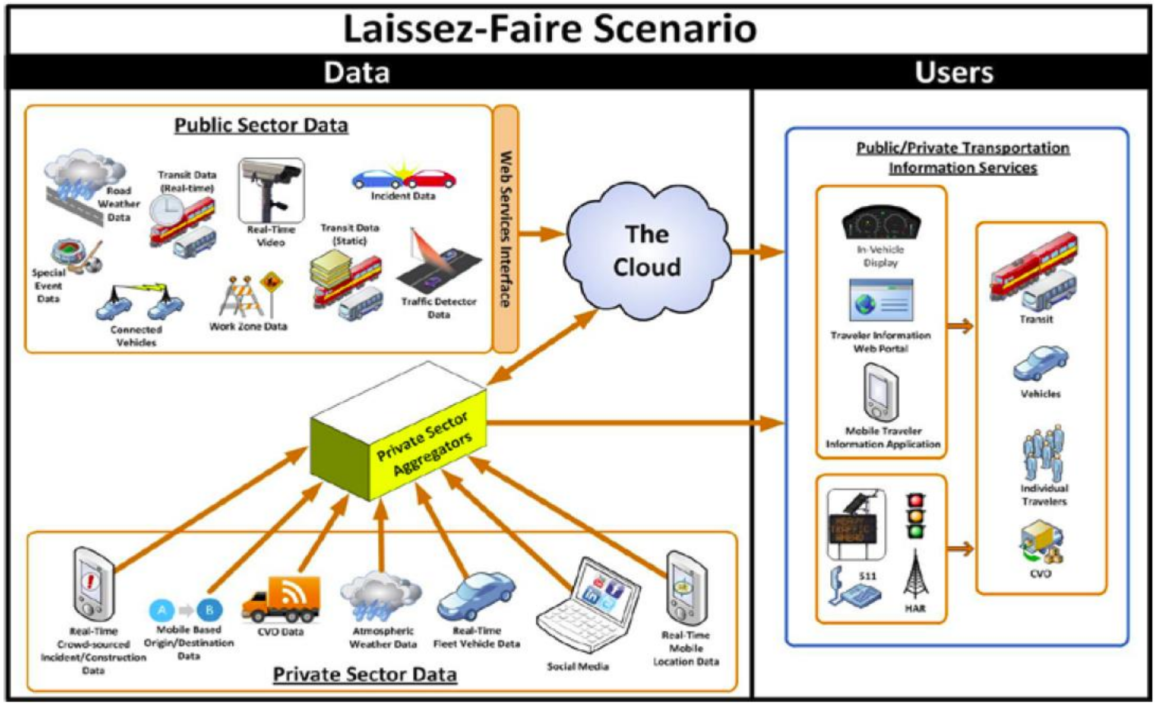


Figure 2-11: 511 Laissez-Faire EnableATIS Operational Scenario (Source: Burgess 2012)

The second concept (Figure 2-12), the preferred ultimate scenario according to the EnableATIS documents, is the robust operational scenario. This scenario assumes public and private sector leadership in delivering a comprehensive, multisource and multimodal data environment that offers advanced traveler information services that consider achieving system optimal conditions, as well as user-optimal conditions.

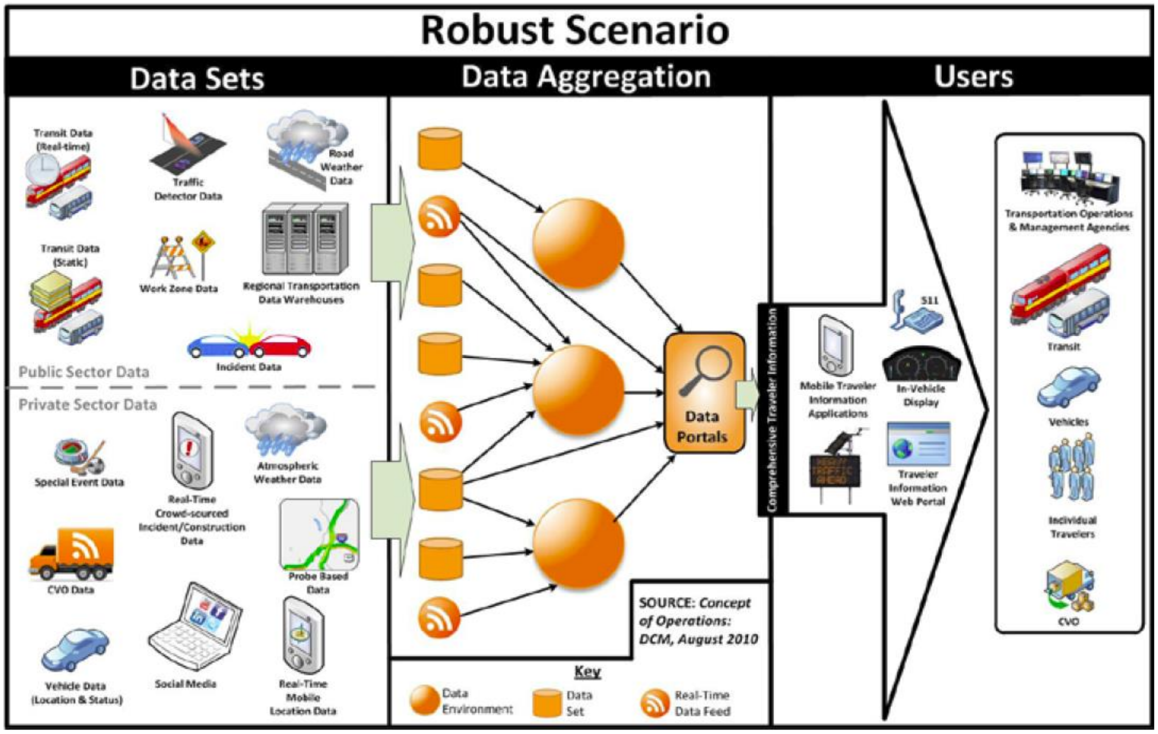


Figure 2-12: 511 Robust EnableATIS Operational Scenario (Source: Burgess 2012)

Over the recent years, the use of the smartphone has increased. A recent study conducted by the Pew Research Center (Anderson, 2015) has shown that 68% of the U.S. adults have a smartphone, compared to only 35% in 2011. The rate of adaptation of smartphones is higher than the rate of the adaption of cellphones (Figure 2-13). In the near future, all the cellphones are expected to be smartphones. The study also shows the percentage of smartphone usage over different demographic groups. It shows that only 30% of elderly people (65 and older) use smartphones, whereas for young people, this percentage is 86%, as shown in Table 2-4. The percentage of smartphone use is almost the same for urban and suburban areas (70% and 72%, respectively), but is less in rural areas (52%). In the case of tablet users, around 45% of U.S. adults own a tablet, where it was only 4% in 2010 (Anderson, 2015).

Table 2-4: Smartphone Usage by Different Socioeconomic Characteristics (Source: Anderson, 2015)

Age group	Percentage
18-29	86
30-49	83
50-64	58
65+	30
Household income	Percentage
<$30K	52
$30K-$49,999	69
$50K-$74,999	76
$75K+	87
Community type	Percentage
Urban	72
Suburban	70
Rural	52

Separate research conducted by the Pew Research Center (Poushter, 2016) found that about 89% of smartphone users use the internet, at least occasionally. This rate is higher (99%) for young people (age 18-34) and lower (85%) for older people (age 35 and older).

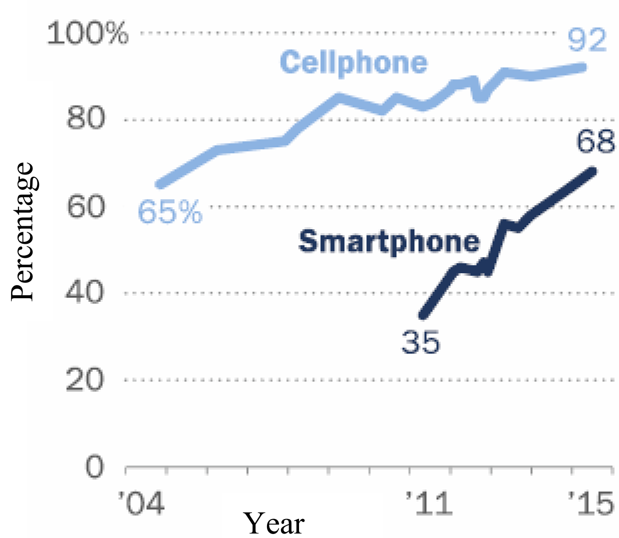


Figure 2-13: Smartphone Usage Increase Over the Years (Source: Anderson, 2015)

2.5 Summary

Connected vehicles will play an important role in the ATIS system in the near future. The National Highway Traffic Safety Administration (NHTSA) has published an advanced notice of proposed rulemaking on Vehicle-to-Vehicle (V2V) communications utilizing connected vehicles (CV). It is expected that the NHTSA will mandate CV technologies for all new vehicles, which will cause a significant increase in the CV market penetration in the coming years. Therefore, the CV data could start supporting the ATIS system within a few years.

Although ATIS requires a wide range of information to support the system, the main input of ATIS is the travel time, as described in Section 2.2. Different travelers may seek different information, but their focus remains on the trip's travel time. Travel time, along with volume information, could be used to determine other performance measurements such as roadway congestion, incident locations, and impacts. Therefore, in this study, travel time estimation and volume estimation is considered a part of the assessment of traveler information systems. The important consideration for such data utilization is the accuracy and reliability of the estimated measurements. Thus, a methodology is developed to determine the accuracy and reliability of the estimated travel time and volume. Finally, a methodology is developed and applied to identify a timeline indicating when CV data could be used for ATIS purposes. The following chapters will describe the details of the data utilized, the methodology, and the model application.

CHAPTER III

METHODOLOGY DEVELOPMENT

This chapter describes the methodology used to support the goal and objectives of this research. The first section, the methodological framework, provides an overview of the method of this research. A detailed description of each step is then presented in the subsequent sections.

3.1 Methodological Framework

As mentioned in the previous chapter, the main objective of this research is to identify the timeline when CV data can be used to support the ATIS data collection process. This identification can be used to support decisions regarding the use of the connected vehicle and other potential technologies for this purpose. The framework of the method is shown in Figure 3-1. The first step of this research is to predict the market penetration of connected vehicles for future years. This estimated market penetration is then used for the assessment of travel time and volume estimation for different years after the CV implementation. Finally, a timeline is created to determine when the existing technology should be replaced by the CV technology.

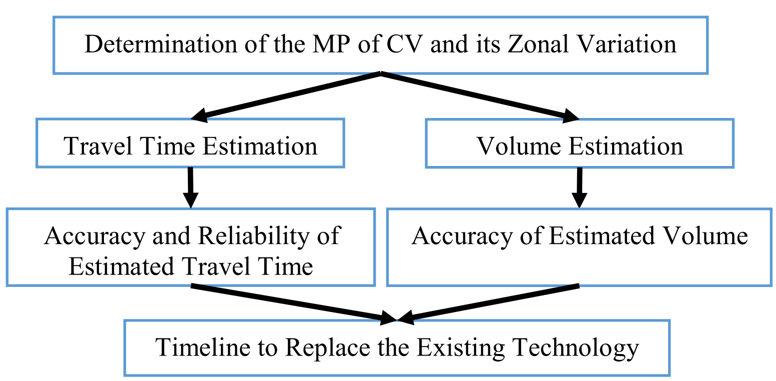


Figure 3-1: Proposed Framework of the Methodology

3.2 Determination of the Market Penetration of CV

The process of determining the Market Penetration (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 socio-economic characteristics of these zones. The third part is to determine the variations of MP on different links in the region by utilizing the traffic assignment procedures incorporated in the regional demand forecasting models.

3.2.1 Assumption of a Scenario for CV Implementation

As stated in the literature review section, the NHTSA is expected to mandate the implementation of connected vehicle technologies for all new vehicles. Apart from this mandate, after-market plug-in equipment will be available for installation on older cars, which 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 a vital role in the determination of the market penetration of CVs. The after-market installations in this research are not considered to be on the conservative side in estimating market penetrations. The USDOT (2008) conducted a research to estimate the benefits and costs of CV implementations. For that purpose, the research predicted the probable market penetrations of CV in future years. In the estimation, the analysts considered a scenario where only new vehicles will use the CV technology, with the following assumptions in percentages that new vehicles will have connectivity: in the first year, 25%; in the second year, 50%; in the third year, 75%; and afterward, 100%. 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 scenario is called the "5-year mandate" scenario, in which automobile companies would include OBUs in new vehicles over a five-year period. The best-case scenario is the "1-year mandate" scenario, where all 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 research could be applied to any of the above and other scenarios of CV implementations, following the same procedure.

3.2.2 Determination of Zone Specific MP

Variations of CV on different links in a region are expected to occur due to the variations of the percentage of CV between zones in the region reflecting the associated socio-economical 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. County-by-county 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 3-2 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.

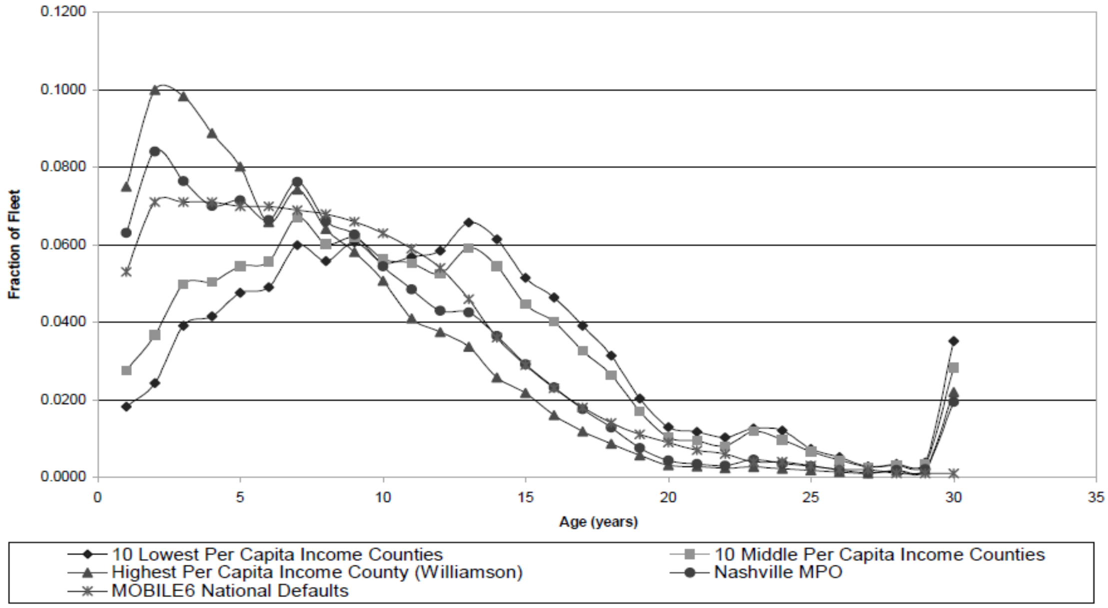


Figure 3-2: LDV Age Distribution for Tennessee Counties (Source: Miller et al., 2002)

Figure 3-2 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 research uses the results from Miller et al. (2002) to produce cumulative percentage distributions of CV for regions or zones with the highest and lowest income areas based on the results in Figure 3-2. Figure 3-3 shows the resulting distributions.

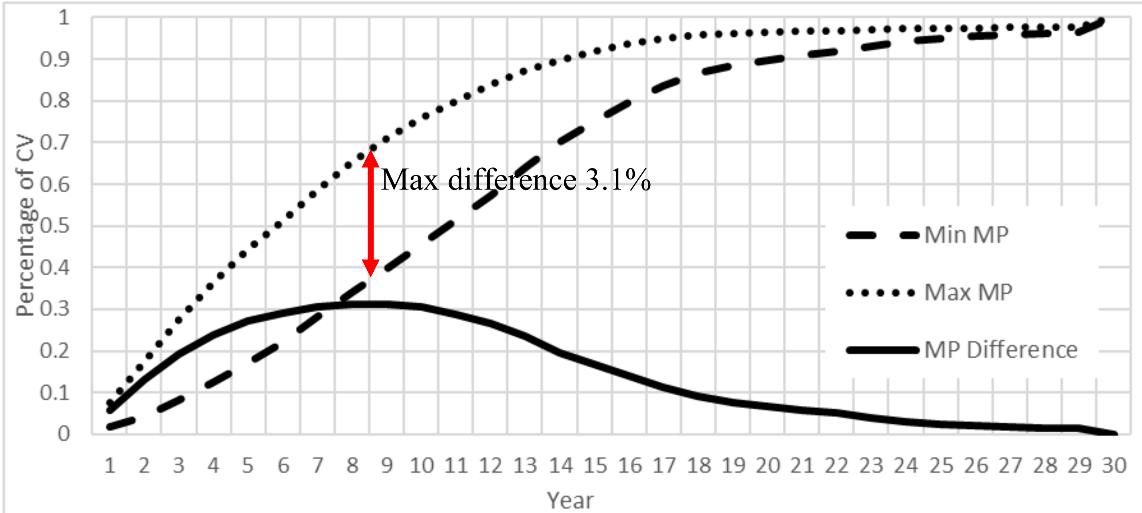


Figure 3-3: Variation of the CV Market Penetration in Different Areas Based on the Information Presented in Figure 4-2

Figure 3-3 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 4-3 shows the difference between the maximum and minimum MP. An important observation from this graph is that the variability of CV market penetration between different areas increases for the first few years and reaches the maximum point around year 8 (3.1% as shown in the Figure 4-2). After that, the variability decreases and eventually becomes very low, as expected.

3.2.3 Determination of the Variation of MP between Links

This research uses the cumulative MP of CV distribution described above to determine the MP of CV for each zone in a region based on the socio-economic characteristics of that region. The determined MP is associated with each zone in a demand-forecasting model. Then, the assignment procedure of the model is 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 estimate the Market penetration of CV on the links, this research utilizes the assignment step of the four-step demand-forecasting model of the southeast Florida region, which is referred to as the Southeast Florida Regional Planning Model version 6 (SERPM6). A brief description of the model network is provided in Section 3.4. 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 in this research was collected using the American Community Survey (ACS) 5-year estimates (2010-2014). This data is available for download from the ACS website (Accessed May 23, 2016) as a Geographic Information System (GIS) database shapefile format. The income per capita in each census tract level is used in the analysis for this research. The income data downloaded at the census tract level is associated with the zone data using the ArcGIS software. This association is performed with the average income of all of the tracts within a certain zone. There are a total 4,106 zones within the SERPM6 model. Using the income of each zone and the variation shown in Figure 4-3, the MP of a zone will be calculated. The minimum MP is assigned for the lower income zone and maximum MP is assigned to the higher income zone. For other zones, the MP of CV linearly varies with the change of per capita income within the maximum-minimum limit.

Once the MP for each zone is identified, as described above, each OD matrix in the demand model is divided into two matrices. The first OD matrix is for the connected vehicles, and the second 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 obtain the number of trips made by the CV. The remaining trips are considered non-equipped (non-connected) vehicle trips. The two types of OD 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 research then uses the link level CV proportions, as determined above, to identify the statistical distributions of these proportions. Analysts can use these type of distributions in lieu of using fixed CV proportions when assessing the performance of applications based on CV technologies.

3.3 Travel Time Estimation

The second objective of this research is the travel time estimation. To obtain the travel time, the emulated BSM data is further processed to estimate the segment travel times under different market penetration scenarios considering the randomness in CV identifications on the links. The results are 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.

3.3.1 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 travel time estimation largely depends on which vehicles are considered connected vehicles. The consideration of this stochasticity is important since it results in higher variations in travel time estimates.

An additional source of variation is the time at which the temporary ID for each vehicle changes within the RSU zone. The SAE J2735 standards (SAE International, 2016) specify that the BSM data contains a temporary ID that 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 conducted for a smaller segment, as is the case in this research. The TCA assigns a temporary ID to each vehicle at the beginning of the simulated segment for all vehicles. The ID for each vehicle is re-assigned after five minutes. In this research, the travel time of the simulated segment is less than five minutes; therefore, the ID could 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 research so that the Vehicle ID on the simulated link changes at random locations within the segment. This is possible since the TCA is an open-source software.

The sources of uncertainty are accounted for in this research using the Monte Carlo simulation, which allows for random selection of the CV vehicles and the time the vehicle ID changes. This requires multiple sampling of the data, each with the different random selection. The travel time estimation is conducted for each 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. In this research, 500 estimates were conducted for both the freeway and arterial segments.

3.3.2 Travel Time Determination

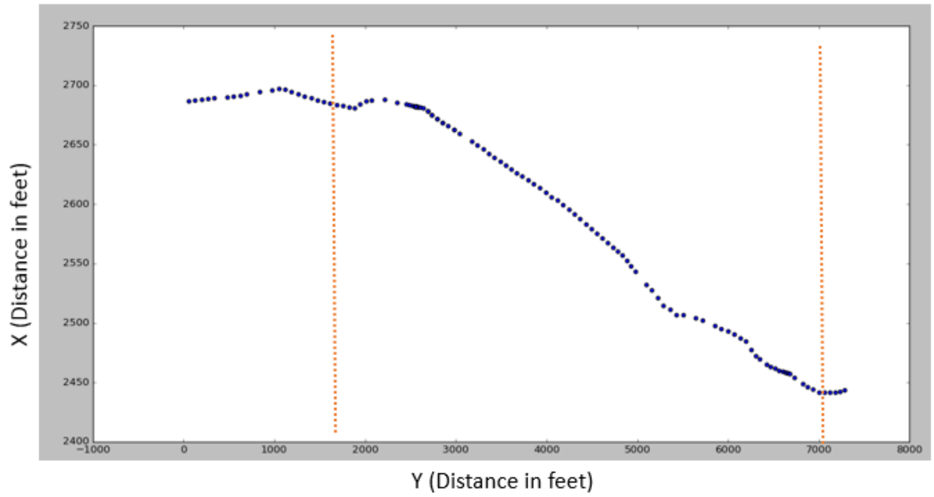
Travel time is calculated utilizing the partial trajectories of the CV with unique IDs. The process of analyzing the BSM data to estimate the travel time is described below.

1. Emulated BSM data resulting from processing vehicle trajectories using the TCA tool contains different information, including vehicle ID, time, speed, location, and other parameters. The first step of the analysis is to separate the required information (fields) from the raw BSM data to reduce the data size. To estimate the travel time, this research utilized the following fields: vehicle ID, which is defined as Probe Segment Number (PSN), time (localtime), the speed of the vehicle (spd), and the horizontal (y) and vertical (x) location of each equipped vehicle.
2. After filtering out the data as described in the first step, the BSM data is sorted by the vehicle ID (which is denoted by PSN) (Figure 3-4).

PSN	localtime	spd	x	y
35	1866	7.12	2720.84856	2831.67320
35	1853	0.00	2720.58288	2836.88840
35	1853	0.00	2720.58288	2836.88840
35	1894	45.97	2756.81704	1245.91456
35	1852	0.00	2720.58288	2836.88840
35	1896	45.80	2758.82112	1111.40832
35	1833	0.00	2720.58288	2836.88840
35	1833	0.00	2720.58288	2836.88840
35	1809	0.00	2720.58288	2836.88840
35	1809	0.00	2720.58288	2836.88840
35	1897	45.68	2759.70672	1044.39464
35	1823	0.00	2720.58288	2836.88840
35	1804	0.00	2720.58288	2836.88840
35	1804	0.00	2720.58288	2836.88840
35	1823	0.00	2720.58288	2836.88840
35	1851	0.00	2720.58288	2836.88840
35	1851	0.00	2720.58288	2836.88840
35	1819	0.00	2720.58288	2836.88840
35	1819	0.00	2720.58288	2836.88840
35	1898	45.33	2760.59232	977.72208
35	1899	44.97	2761.14664	911.56776
35	1850	0.00	2720.58288	2836.88840
35	1850	0.00	2720.58288	2836.88840

Figure 3-4: Sorted BSM Data

3. Each vehicle PSN is checked to determine whether it is within the boundary locations of a segment (the starting and end points) for which the travel time is being estimated, as shown in Figure 3-5. In order to obtain a homogeneous sample, this step eliminates the vehicles traveling to and from side roads (arterial) or on-ramp to off-ramp (freeway). Only the data points identified for the segments are used in estimating the travel time.



**Red dot line is the segment Boundary*

Figure 3-5: Vehicle Trajectory with Segment Boundary

4. The TCA tool assigns PSNs randomly from a set of values. A PSN, which is assigned to a specific vehicle, can be assigned to another vehicle at a later time when the previous one is not on the network. To eliminate the effect of this feature, the trajectories of each PSN are checked to determine if there is a larger time gap between two data points. If the time gap between two consecutive data points is larger than a certain threshold (100 seconds), then the PSN is divided into two different IDs.
5. As mentioned earlier, the vehicle's temporary ID changes periodically. Therefore, a vehicle may not pass the whole section with the same ID. The test sections in this research have shorter average travel times, compared to the ID changeover period (300 seconds). Thus, the vehicles may change their ID once within the section. In this step, each vehicle ID (PSN) is divided into two parts (PSN_1 and PSN_2) to emulate such change in vehicle ID. A random

breakpoint is selected to change the ID and consider the 30-second silent period in between the two parts, according to the SAE standards [Figure 3-6].

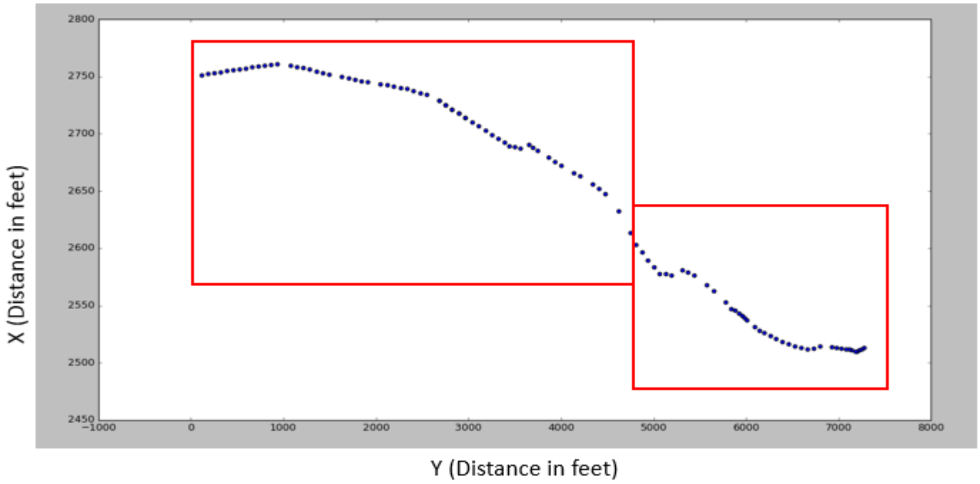


Figure 3-6: Emulating the Change of the Vehicle ID (PSN)

6. If the distance traveled by any of the two IDs is less than 1,000 feet, then this ID is not considered for travel time calculation.

7. The data points of each ID (PSN) are sorted by time. The difference between their maximum and minimum times is the travel time of that PSN and the associated Y points, which provide the distance traveled during a specific amount of time.

The summation of all the valid travel times and distances are used in the estimation of the total travel time and total distance. The link average travel time for each time period is calculated using the following equation:

$$T_{av} = \frac{\sum_i^n tt_i}{\sum_i^n d_i} * L_s \qquad (3\text{-}1)$$

where T_{av} is the average travel time of the segment, which has a length of L_s, and tt_i and d_i are the travel time and distance traveled by the i^{th} PSN. This full process is repeated multiple times (500 times) with the Monte Carlo simulation to obtain the distribution of travel times and the associated error for each market penetration.

3.3.3 Assessment of the Accuracy and Reliability of Travel Time Estimation

The Monte Carlo simulation process described above is performed with selected MP of CV. For this research, the market penetrations are 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 time estimates for each set of randomly selected vehicles and their changing IDs. Each of these selections represents a single day of operations. These estimates are compared with ground truth (base) travel time, which are estimated based on the complete NGSIM vehicle trajectory data using several accuracy and reliability measures. Toppen and Wunderlich (2003) reported four different measures to investigate the estimation error. The details of these four measurements are provided in Table 3-1.

Table 3-1: Travel Time Accuracy Measures

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

$$E_{rr} = \beta_0 e^{f(x)} \tag{3-7}$$

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

where f(x) is a polynomial function, x is the market penetration of CV, E_{rr} is the error value, and β are the coefficients. The degree of the polynomial function and the coefficient values vary depending on the type of error and the type of roadway. The statistical software R was used to fit the regression equation and validate it with proper statistics.

3.3.4 Effect of Demand Variation on the Accuracy of Estimated Travel Time

The accuracy of the estimated travel time is also a function of the roadway traffic volume. At higher traffic volumes, there will be more CVs. Therefore, the accuracy of the estimated travel time is expected to increase. However, at higher traffic volumes, there is a higher variation of travel time between the vehicles due to the increased interaction between vehicles during congested conditions. This results in a lower level of accuracy in the estimated travel time. The effect of this demand variation is also investigated utilizing the simulated vehicle trajectories.

The same methodology mentioned in the previous sections (Section 3.3.1 to Section 3.3.3) was applied in the simulation test sections to investigate the effects of demand variation on the accuracy of estimated travel time. This research develops models of the relationship between the performance of the travel time estimation using CV data and two independent variables: traffic demand and CV proportion. In this research, multiple linear regression is used to derive the relationships. The derived model for travel time estimation is presented in Equation 3-9. The independent variable in the regression is the error in travel time estimation, and the dependent variables are the MP of CV (CP) and the traffic flow measurement (S). 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 the following equation:

$$E_{err} = \beta_0 + \beta_1 * \cos(S) + \beta_2 * \log(CP) + \beta_3 * \mathrm{S} + \beta_4 * \mathrm{CP} + \varepsilon \qquad (3\text{-}9)$$

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 four types mentioned in Table 3-1. 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.

3.4 Volume Estimation

The third objective of this research is to estimate the traffic volume utilizing the CV data. This section proposes a methodology to estimate the traffic volume utilizing CV data. With the increase in the CV market penetration, at some point in time, there will be sufficient CV data that allows the removal of at least a subset of the midblock detectors on arterial streets. 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, and 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 research, 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 conducted 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 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 initial 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. Following is a description of the base scenario and the three scenarios listed above.

3.4.1 Base (Existing) Scenario

Due to a limited budget, it is not possible to place detectors in every location of a highway corridor. Different studies have been conducted to find 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 research follows this concept of selecting the initial list of links for instrumentation to represent the base (existing) scenario based on the differences in volume measurements between links. For this purpose, three measurements are used to check the variability in volumes between the locations, as follows:

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

  $$MAPD = \frac{1}{n}\sum_{i=1}^{n}\frac{|x_i - y_i|}{|x_i|} \quad (3\text{-}10)$$

  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 observations.

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

- The Sample Correlation coefficient (r_{xy}) is calculated using the following equation:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \qquad (3\text{-}11)$$

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

3.4.2 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, as discussed in the previous section. Therefore, it is expected that the ratio of the actual volume and the CV volume will be close to 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 have low variability in volume, compared with the instrumented location, to find the actual volume ($V_{j,t}$) based on the CV volume ($CV_{j,t}$) at that location and at that time interval (time interval t). This calculation can be expressed by the following equations:

$$C_{f,t} = \frac{V_{i,t}}{CV_{i,t}} \tag{3-12}$$

$$V_{j,t} = C_{f,t} * CV_{j,t} \tag{3-13}$$

3.4.3 Future Scenario 2

In the Base Scenario and Future Scenario 1, the links are selected to be non-instrumented (no 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 the CV, it will be possible to estimate 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 for 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 the CV data of all locations. A general form of the derived regression equation is:

$$V_i^{est} = \alpha + \sum_{j=1}^{n} \beta_j * CV_j + \sum_{j=1}^{n-1} \gamma_k * V + \varepsilon \tag{3-14}$$

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 of 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 4-15. The optimization problem is solved using the Bisection method and utilizing the R software.

Minimize, $f = \sum_{i=1}^{n} \sum_{t=1}^{p} \left| V_{i,t}^{est} - V_{i,t}^{act} \right|$ (3-15)

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

3.4.4 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 penetration, 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 3-16 and 3-17, in which an average expansion factor (C_f) is calculated from the temporary detector counts and the partial counts based on the 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.

$$C_f = \frac{1}{p}\left(\sum_{t=1}^{p}\frac{V_t^{temp}}{CV_t}\right) \tag{3-16}$$

$$V = C_f * CV \tag{3-17}$$

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 (C_f), and Equation 3-17 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.

3.4.5 Accuracy Measurements of Volume Estimation

In all of the above-mentioned scenarios, the accuracy of the estimation is assessed in this research by calculating the Absolute percentage error (E_{rr}) of the estimated volume (V^{est}), compared to the actual volume (V^{act}), using the following equation:

$$E_{rr} = \frac{\left|V^{act} - V^{est}\right|}{V^{act}} * 100 \tag{3-18}$$

This research 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 an agency's policy or existing guidelines, such as those presented in Reference (FHWA, 2014). This selection is used to develop a timeline indicating 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.

3.5 Summary

This research focuses on determining the potential of utilizing connected vehicle data to support the traveler information system. In the first step, the market penetration of CV, along with its variation between zones, is estimated for future years after the expected CV mandate becomes effective. The traffic assignment step of the regional demand models is used to obtain the MP variation. In the next step, the travel time is estimated at different MPs considering the zonal variation. The accuracy and reliability of the estimated travel time is investigated utilizing different statistical measures. The volume estimation is also performed and examined for different probable future scenarios and setups.

CHAPTER IV

DATA PREPARATION AND PROCESSING

This chapter describes the data sources that have been used in this research to demonstrate the methodology described in the previous chapter. A trajectory conversion tool is used in this research to emulate the BSM data from vehicle trajectories. The first section explains how vehicle trajectory data is collected. The second section describes the tools used in this research for the conversion. The third section describes the process to emulate the CV BSM data. The last section describes the roadway network that is utilized in the case study to determine the CV market penetration. The complete process of CV data generation is shown in Figure 4-1.

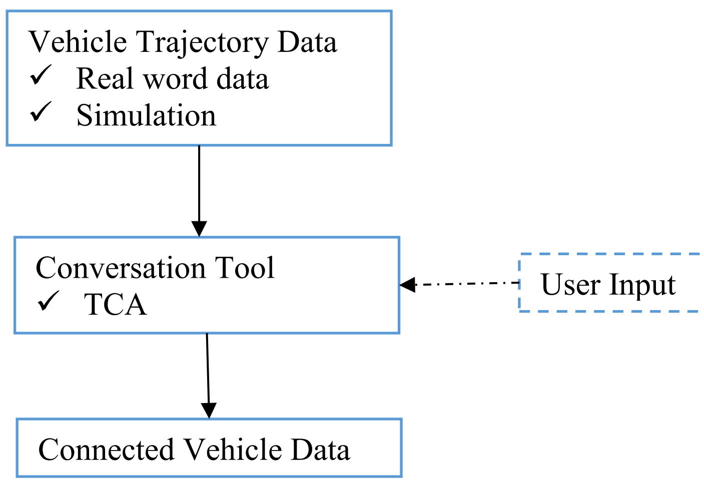


Figure 4-1: Connected Vehicle Data Emulation Process

4.1 Vehicle Trajectory Data

Currently, there is no real-world connected vehicle data that can be used to demonstrate the use of CV with different market penetrations for traveler information processes. Connected vehicle data became available after a few testbeds and pilot projects were conducted around the United States. However, the CV proportions associated with these implementations are very low. Therefore, this research utilizes an emulation of connected vehicle data by inputting vehicle trajectory data into a tool named Trajectory Conversion Algorithm (TCA) [Version 2.3] (OSADAP, 2015), which was developed by the Federal Highway Administration (FHWA). The tool converts the trajectory data to CV data, according to the SAE J2735 standards (SAE International, 2016). This research utilizes both real-world vehicle trajectories and simulation model generated trajectories for input into the TCA tool.

A vehicle trajectory data is a collection of snapshots on vehicle statuses. Vehicle status is defined by various types of information, such as vehicle location, speed, acceleration, and timestamp. Such information is recorded at a regular time interval (e.g., 0.1 seconds) to obtain the whole travel history of a vehicle. A collection of all of a vehicle's travel history is called vehicles trajectory data.

4.2.1 NGSIM Data

For real-world vehicle trajectories, this research utilizes data collected under the Next Generation Simulation (NGSIM) program of the FHWA as a base. The data is processed to emulate the BSM data. The NGSIM program collected high-quality traffic and vehicle trajectory data. This data are available for researchers to download (NGSIM, accessed July 2015). The data was collected from four different locations within the USA. Two of the locations are arterial street segments, and two locations are freeway segments. In this research, one of the arterials (Peachtree Street, Atlanta) and one of the Freeways (U.S. Highway 101, Los Angeles, California) are used as the test locations in this research.

Arterial Location Data

The data was collected for the PM peak interval (4:00 PM and 4:15 PM on November 08, 2006) from Peachtree Street. The collected video data was processed to produce trajectories of vehicles and an aggregated summary of traffic flow characteristics. The total length of the test link is approximately 2,100 feet with five intersections, as shown in Figure 4-2. It has two to three through lanes in each direction of travel. The trajectory file contains twenty-four different fields, as follows:

1. Vehicle identification number
2. Frame identification number
3. Total frames
4. Global time (epoch time) in milliseconds
5. Local X coordinate of the front-center of the vehicle with respect to the left-most edge of the section in the direction of travel
6. Local Y coordinate of the front-center of the vehicle with respect to the left-most edge of the section in the direction of travel
7. Global X coordinate of the front-center of the vehicle based on CA State Plane III in NAD83
8. Global Y coordinate of the front-center of the vehicle based on CA State Plane III in NAD83
9. Vehicle length
10. Vehicle Width
11. Vehicle class (motorcycle/auto/truck)
12. Instantaneous velocity of the vehicle

13. Instantaneous acceleration of the vehicle
14. Lane identification; lane numbering is the increment from the left-most lane, except for locations where a left-turn or right-turn bay exists. Left-turn bays are numbered starting from 11 and are incremented from the left-most left-turn bay.
15. Origin zone of the vehicle; numbered from 101 through 123
16. Destination zone of the vehicle; numbered from 201 through 223
17. Intersection in which the vehicle is traveling; numbered from 1 to 5
18. Section in which the vehicle is traveling; the street section is divided into six sections
19. Moving direction of the vehicle; east-bound, west-bound, north-bound, south-bound
20. Movement of the vehicle; through, left-turn, right-turn
21. Identification number of the preceding vehicle in the same lane, and "0" represents no preceding vehicle
22. Identification number of following vehicle in the same lane, and "0" represents no following vehicle
23. Spacing between the front-center of a vehicle to the front-center of the preceding vehicle
24. Time headway between the front-center of a vehicle to the front-center of the preceding vehicle. Among the above fields, Vehicle Id, Local X, Local Y, Vehicle velocity and acceleration were used for analysis.

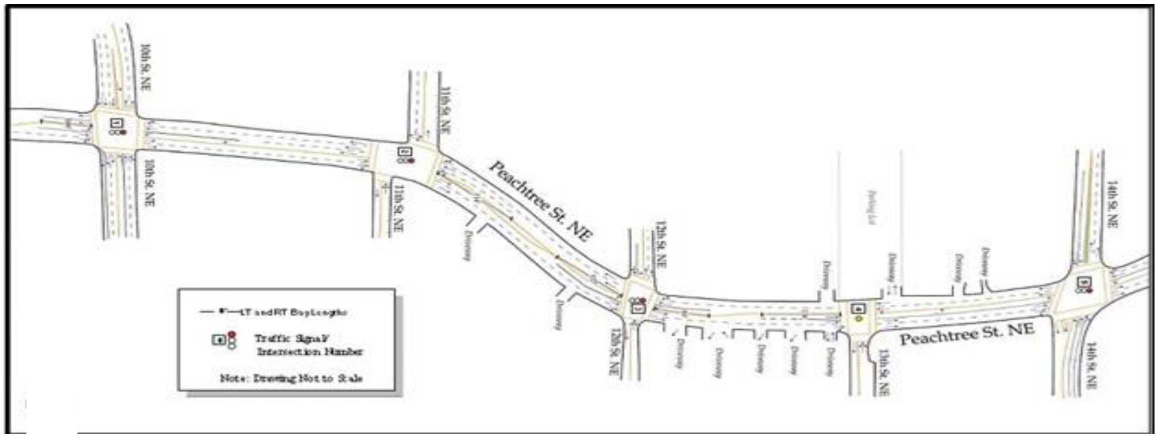


Figure 4-2: Test Link at Peachtree Street, Atlanta, Georgia (Source: NGSIM)

Freeway Location Data

The freeway location, for which data is available and used in this research, is U.S. Highway 101 (Hollywood Freeway), Los Angeles, California. The data was collected for the AM peak period (between 7:50 AM and 8:05 AM on June 15, 2005) in the southbound direction. The processed video dataset contains the trajectories of vehicles, as well as an aggregated summary of the traffic flow and speed of the vehicles. The total length of the link is 2,100 feet. Figure 3-3 shows the freeway test location. It has five mainline lanes and an auxiliary lane in between the on-ramp and off-ramp.

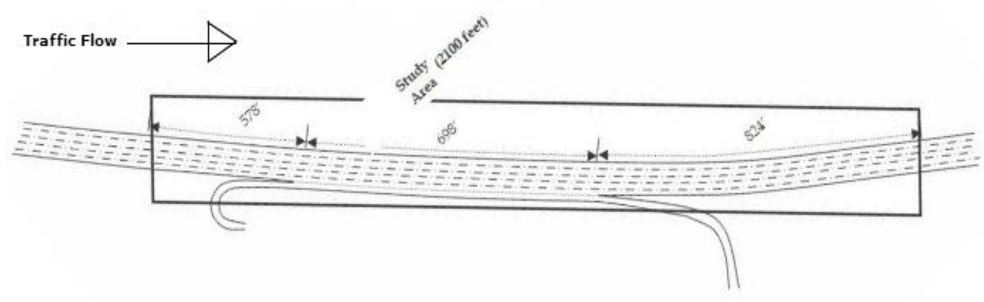


Figure 4-3: Test Link at US 101, Los Angeles, California (Source: NGSIM)

4.2.2 Simulation Data

In this research, trajectories are also generated utilizing the simulation software VISSIM. The reason for using the simulation is to have the flexibility to change the demands to examine the impact of demands on the assessment results. An arterial segment and a freeway segment were coded and calibrated in VISSIM for this purpose. The details of these two networks are provided below.

Freeway Segment

A freeway weaving segment was modeled in this research to examine the effect of CV market penetration on travel time estimation at different traffic demands. The coded freeway segment was calibrated to ensure it produces measures (maximum flow rate and speed) similar to the results from the HCM analysis. The total length of the segment is one mile (5280 feet). The on-ramp and off-ramp segments are separated by a 1,000 feet freeway segment. As a typical Road-Side Unit (RSU) has a range of one mile, a one-mile test section is selected as the test study section.

As stated above, the network is calibrated to produce the measures (maximum flow rate and speed) estimated using the HCM (2010) procedure. According to the definition of the HCM, the middle 1,000 feet section is considered a weaving section, and the beginning and end part (2,140ft each) is considered a basic section (see Figure 4-4).

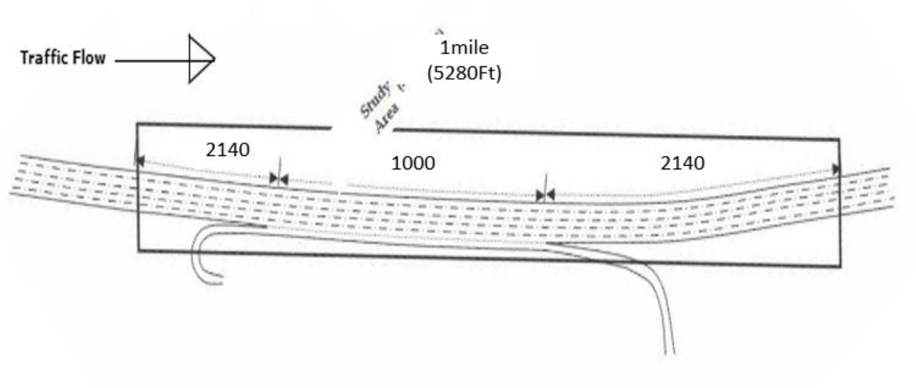


Figure 4-4: Coded Freeway Section

A document (Mai et al., 2011) prepared by the Ohio Department of Transportation (ODOT) showed a detailed guideline for the calibration of a freeway network in VISSIM. The document suggested the use of Wiedemann 99 as the car-following model for a freeway segment. There are ten different parameters in the Wiedemann 99 model. Among these ten parameters, the document suggested adjusting the first three parameters: Standstill distance (CC0), headway time (CC1), and following variation (CC2). The suggested ranges and selected values for these parameters in this research are presented in Table 4-1.

Table 4-1: Parameters Selection for Freeway Car-Following Model in VISSIM

				Suggested Range		Selected Value	
		Default	Unit	Basic Segment	Merging/ Diverging	Basic Segment	Merging/ Diverging
CC0	Standstill Distance	4.92	feet	4.5-5.5	>4.92	4.5	4.92
CC1	Headway time	0	second	0.85-1.05	0.9-1.5	0.85	0.9
CC2	Following Variation	13.12	feet	6.56-22.97	13.12-39.37	6.56	13.12

For the lane-changing parameters, the ODOT document (Mai et al., 2011) also provided a range to follow. This research manually adjusted the value within the suggested range and identified the best combination to generate the HCM suggested capacity. Table 4-2 provides the suggested ranges and the selected values for the lane-changing parameters.

Table 4-2: Parameters Selection for Freeway Lane Changing Model in VISSIM

	Default	unit	Suggested	Selected
Maximum deceleration (own)	-13.2	Ft/s2	-15 to -12	-12
Maximum deceleration (Trailing vehicle)	-9.84	Ft/s2	-15 to -8	-12
-1 ft/s2 per distance (own)	200	Ft	150 to 250	200
-1 ft/s2 per distance (Trailing vehicle)	200	Ft	150 to 250	200
Accepted deceleration (own)	-3.28	Ft/s2	-2.5 to -4	-4
Accepted deceleration (Trailing vehicle)	-1.64	Ft/s2	-2.5 to -2.5	-2.5
Waiting time before diffusion	60	s	200	200
Min. headway	1.64	Ft	1.5 to 2	1.5
Safety distance reduction factor	0.6	-	0.25 to 1	0.25
Maximum deceleration for cooperative braking	-9.84	Ft/s2	-8 to -15	-9.84

Arterial Network

For the arterial network, this research selected Glades Road, located in the City of Boca Raton, Florida, for the demonstration of the effects of demand variation on travel time and volume estimation. 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 miles 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 is coded in VISSIM, followed by the calibration process, which is performed in two stages based on real-world measurements.

In the first stage, an initial fine-tuning of the model parameters is conducted to produce saturation flow rates that are in agreement with the Highway Capacity Manual 2010 (HCM 2010) procedures and previous observations from the field in South Florida. The target saturation flow rate for the purpose of this calibration is set to 1,850 to 1,900 passenger cars per hour per lane. The parameters of the VISSIM model are fine-tuned to produce this value. Two of the parameters of the urban driver car-following model in VISSIM (the Wiedemann 74 driver behavior model) are fine-tuned in this research in order to obtain the target saturation flow rate according to the recommendation in previous studies (VISSIM, 2014; Kim, 2005; Al-Nuaimi, 2013). The parameters are:

1. The additive part of desired safety distance (bx_add).
2. The multiplicative part of desired safety distance (bx_mult).

These parameters determine the target desired safety distance, which has a direct impact on the saturation flow rate (ODOT, 2011). The most appropriate combination of the two parameters is found to be 2.4 feet for the additive part of the desired safety distance, and 3.4 for the multiplicative part of the desired safety distance. The resulting saturation flow rate for the simulated through movement is 1,854 passenger cars per hour per lane, based on the average of ten simulation runs with different random seeds, which is close to the target saturation flow rate of the calibration.

The second part of the calibration process is to compare the traffic flow performance according to the Glades Road simulation model with real-world traffic data. Three traffic data types were used in the calibration, including:

1. Signal timing plans of all intersections.
2. Historical turning movement counts of all intersections.
3. Data from permanently installed magnetometers (from SENSYS®).

These data types provide point measurements of speed, occupancy, and volume, as well as travel time measurements based on automatic vehicle re-identification. Virtual detectors are included in the simulation at the same locations of the detectors in the field.

The measured and simulated parameters based on ten simulation runs with different seed numbers are examined to determine if further fine-tuning of the parameters is needed. It is determined that the Root Mean Square Percentage Errors (RMSPE) for the simulated versus the measured volume and speed values are below 15% for both directions of travel.

The Glades Road segment coded in the simulator is 17,000 feet long (3.25 miles) and has nine signalized intersections. However, the segment used in the travel time estimation assessment is one mile, as this is the maximum coverage by one RSU. The section selected for analysis is shown in Figure 3-5. The simulation was run for one hour. The initial 30-minute period is considered a warm-up period.

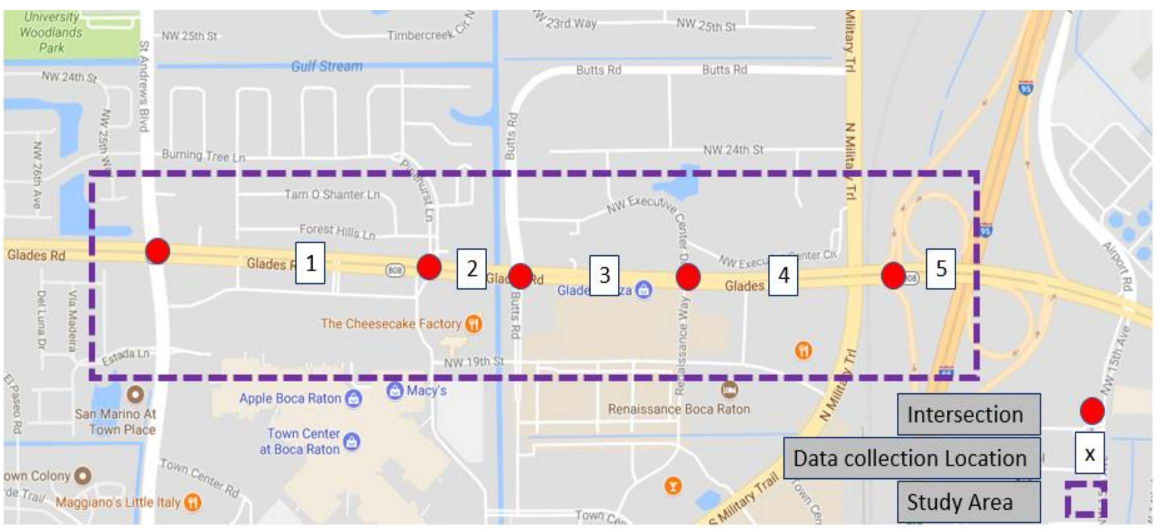


Figure 4-5: Coded Arterial Section

4.2 Trajectory Conversion Tool

For CV data generation, the Trajectory Conversion Algorithm (TCA) is used in this research. The TCA tool [Version 2.3] (OSADAP, Accessed July 13, 2015) was developed by the FHWA to test different strategies for producing, transmitting, and storing Connected Vehicle Information. This software reads and uses vehicle trajectory information, assumed Roadside Equipment (RSE) location information, and other information to produce CV data according to the J2735 standards. The TCA is an open source software written on the Python platform. The following TCA input files are configured for the purpose of this research.

Control File

The control file is an XML format file that contains all of the information about the names of all other input files and the percentage of equipped vehicles.

Vehicle Trajectory File

The vehicle trajectory file is a comma-delimited file (CSV) that provides all of the vehicles trajectory information. This file contains the following trajectory attributes: vehicle ID, time in seconds from the beginning of the trajectory, the speed of the vehicle in mph, and the x and y coordinates of the vehicle location in feet. The data point in the file must be sorted by time.

RSE location File

The RSE location file is also a CSV file that contains the geographical location information of the assumed RSEs. The attributes of the files are x and y coordinates of the RSE in meters, latency value associated with the RSE in seconds, and the loss rate percentage associated with that RSE. All of these parameters are user inputs. One RSE is considered in this research to collect the CV data, which is placed in the middle of the test section.

Strategy File

The strategy file is an XML-based file that stores all of the information for controlling the snapshots generated by the TCA. This file contains the settings of the J2735 standard and is set to the default values in this research. The user can alter some of the settings within the range provided by the standard or can create several variations of the standard to test their impacts.

4.3 Emulation of BSM data

The NGSIM data are processed to emulate BSM Type 1 data collected at 1/10th second intervals. The downloaded NGSIM trajectory files, described earlier, are processed in this research by filtering out the required variables and performing unit conversions. Using the trajectories as inputs, the TCA tool generates connected vehicle BSM Type 1 data following the SAE J2735 standards (SAE International, 2016). The user of the TCA tool can select the market penetration, message type, and the communication type (DSRC or cellular). The user can also specify the RSU locations. Since the roadway section in the NGSIM data is shorter in length than the coverage area of an RSU, which is about 2,600 feet to 4,000 feet (Andrews and Cops, 2009; McGurrin, 2012), it can be assumed that only one RSU is placed halfway in the section, providing 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 research, a 12% transmission loss is specified as an input into the TCA tool. It is expected that the actual loss is a function of the availability of the line of site in the coverage area, which should be considered with the RSU siting. The

TCA tool changes the vehicle ID every five minutes, with a 30-second buffer window, as specified by the J2735 standards to protect privacy. 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.

4.4 Network Model Data

The method developed in this research to estimate the CV proportion on each link requires running the assignment component of the regional demand model. This research has used the Southeast Florida Regional Planning Model (SERPM) network for this purpose. SERPM, a CUBE software-based model, consists of the region covered by FDOT District 4 and District 6, which consists of Broward, Palm Beach, and Miami-Dade counties. A snapshot of the SERPM model network is provided in Figure 4-6.

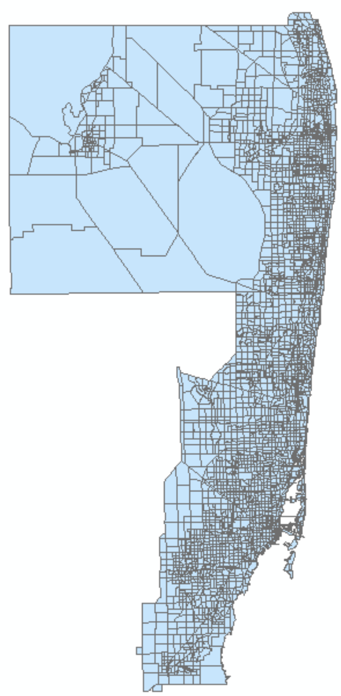


Figure 4-6: SERPM Model Network Used as Part of the Estimation of Link-Based Proportions of CV

4.5 Summary

In this research, vehicle trajectory data is collected from the NGSIM program and from the simulation. The Trajectory Conversation Algorithm (TCA) tool is used to emulate the CV data utilizing the vehicle trajectories. Finally, the CV data is further analyzed to fulfill the research objectives. The proposed methodology of this research, described in Chapter III, is demonstrated utilizing the data mentioned in this chapter, and the results are listed in Chapter V.

CHAPTER V

MODEL APPLICATION AND RESULTS

The methodology described in the Chapter IV was implemented in the test data described in Chapter III to demonstrate the proposed methodology. This chapter discusses the details of the application and the results.

5.1 Predicting the Future CV Proportion

For the Southeast Florida Case Study (Section 4.4) utilized in this dissertation, there are about 45,000 links, each of which the CV percentages are identified based on the methodology developed in this research. As described in Section 4.2, this estimation was performed using the assignment model of the regional demand model. The first step of model application is to determine the distribution of these proportions. First, this research investigated whether the distribution of the CV market penetrations on the links resulting from the application of the methodology is normal or not. For a large sample, the most widespread practice of the normality confirmation is to check the normality test plots, as shown in Figure 5-1. 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 on a normal distribution than the original data.

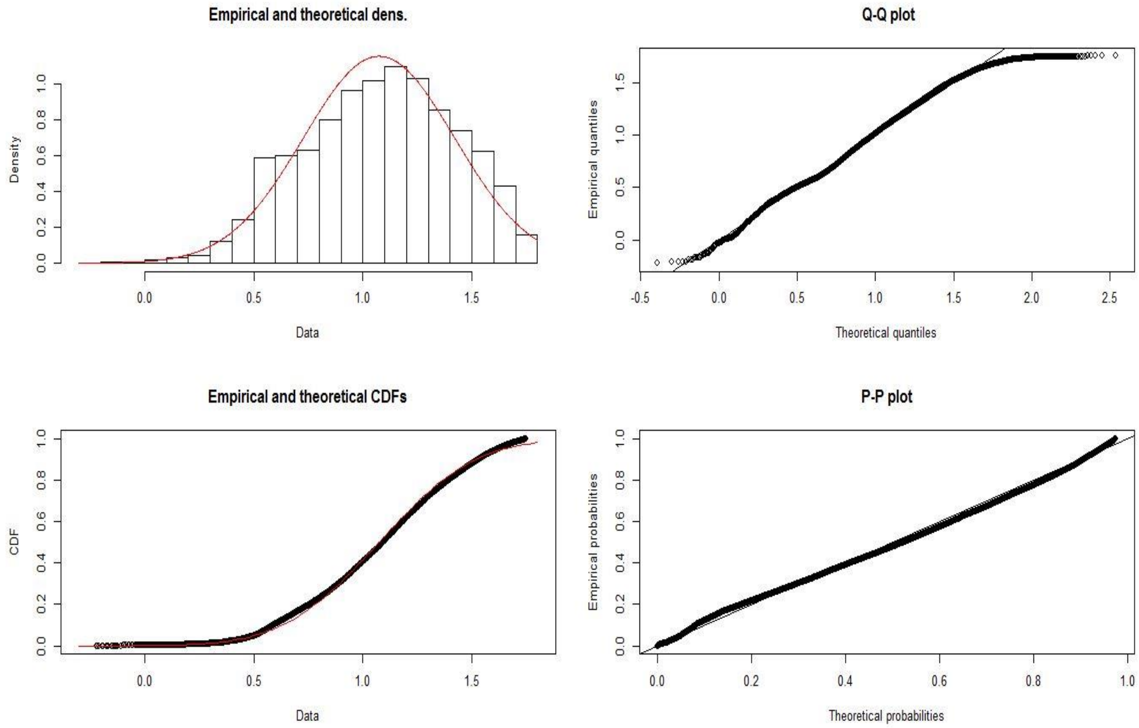


Figure 5-1: Different Plots to Check the Distribution of the Links CV Percentages [Year 1]

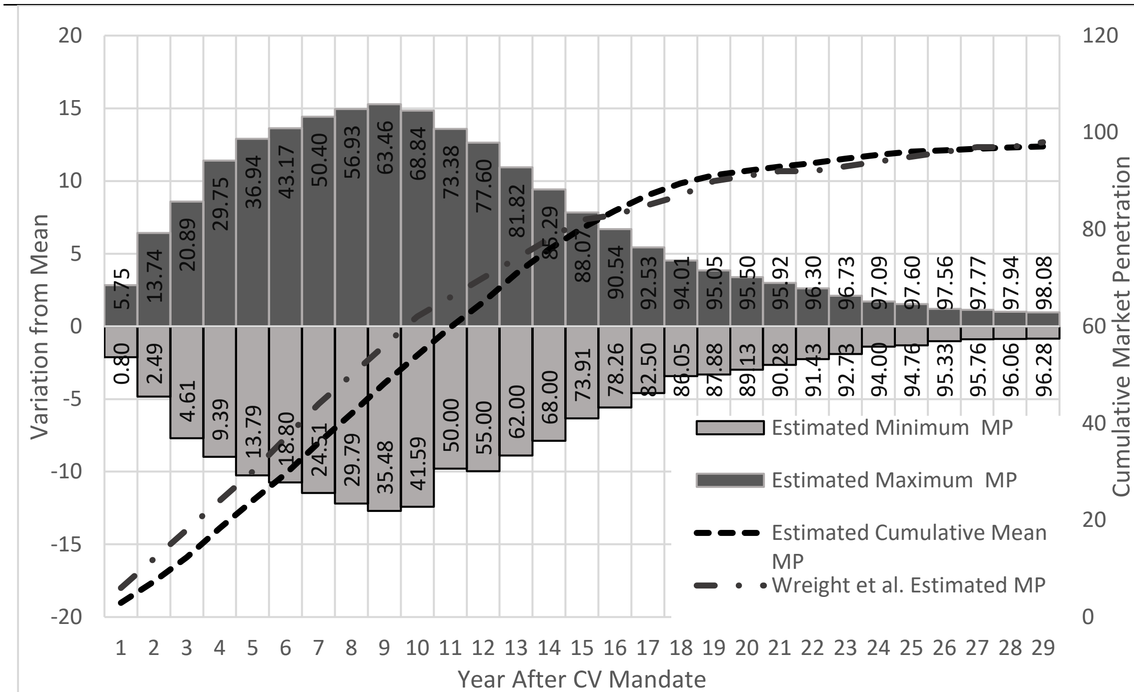
Figure 5-1 shows the plots utilized to check the normality of the logarithmic value of the CV percentages for different links in year 1, after the mandate of the CV. It should be noted that the results presented in this section are for the PM peak period. The analysis can be repeated for different peak periods to show the variations in MP by a specific period of the day. This will account for the difference in the distribution of the trips by time of day of users with different income levels.

The plots show that the logarithm of the CV percentages is normally distributed. This same test is repeated for all future years. For all years, the distributions are found to be lognormal. Table 5-1 shows the mean and standard deviation of each year's distribution.

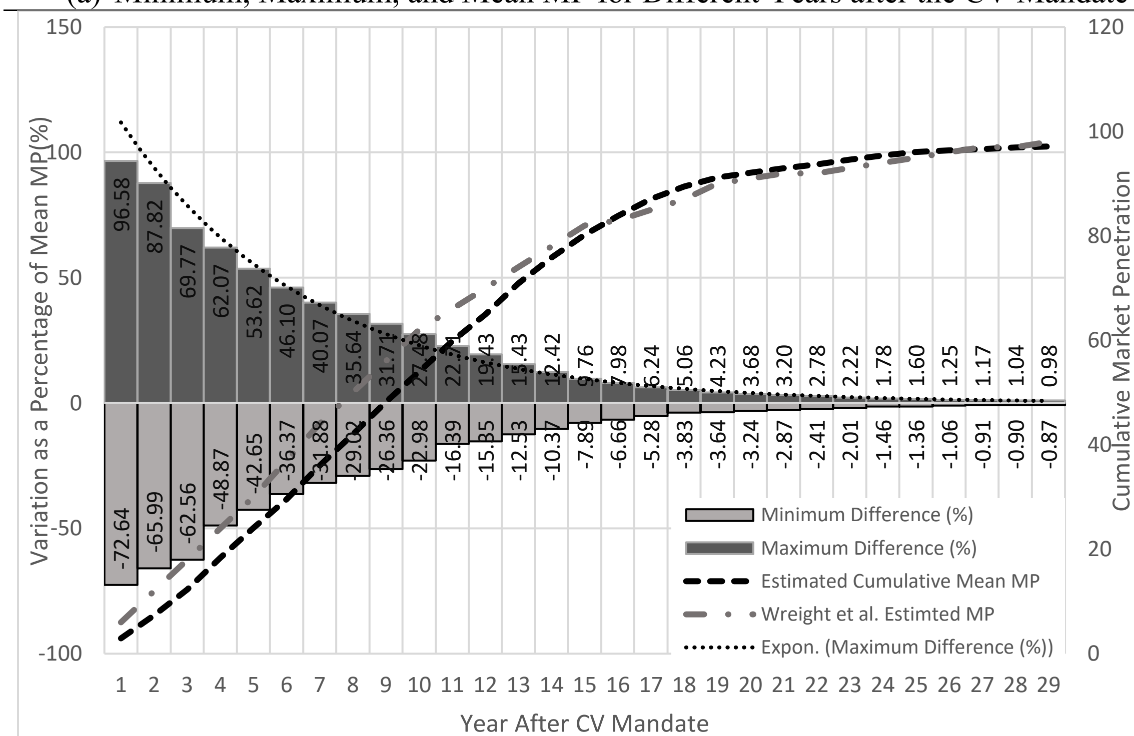
Table 5-1: Mean and Standard Deviation (SD) of Link-level MP Distribution by Year

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				
*Year after Mandating CV on new vehicles							

After analyzing the data for all of the years following the methodology, the mean market penetration on each link and the variation for each future year after the CV mandate are presented in Figure 5-2. In Figure 5-2, 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 values of MP among the links in a region during a particular year of the analysis. The two-dotted line represents the cumulative market penetration calculated in this research and the cumulative MP calculated by Wright et al. for comparison purposes (Wright et al., 2014).



(a) Minimum, Maximum, and Mean MP for Different Years after the CV Mandate



(b) Variation as a Percentage of the Mean MP for Different Years after the CV Mandate

Figure 5-2: CV Proportion Analysis Results

Figure 5-2(a) shows the actual variation, and Figure 5-2(b) shows the variation as a percentage of the mean value. For lower market penetrations, the variations are lower; however, the percentage variations are higher. An exponential function that is fit to the data, as shown in Figure 5-2(b), shows that the percentage of the MP variation decreases exponentially.

The average percentage increase of CV for each year is presented in Figure 5-3. Figure 5-3 shows that the rate of the MP increase grows for the first several years, and then 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.

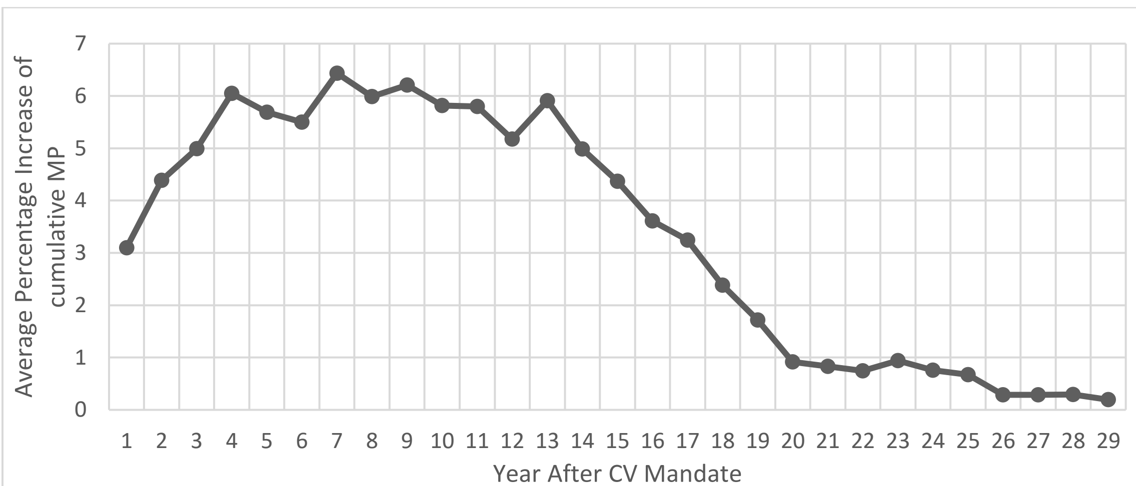


Figure 5-3: Average Percentage Increase of Cumulative MP of CV by Year

This research also investigated the MP variations on different facility types (Figure 5-4). Figure 5-4 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 a lesser extent, arterials. Thus, it is recommended that these variations are considered separately by facility type, with a different distribution identified for each type.

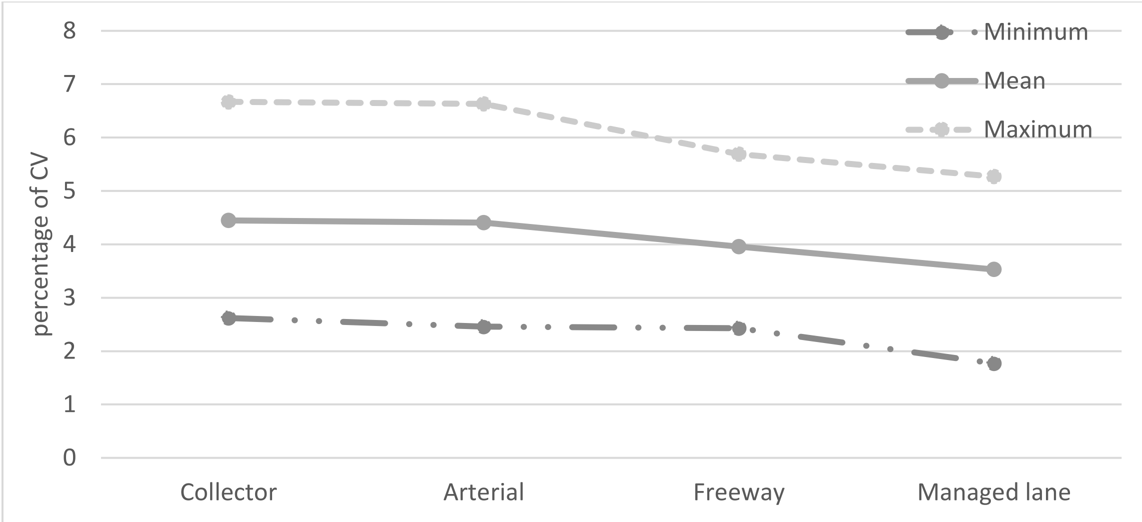


Figure 5-4: Variation of MP by Facility Type (Year 1)

5.2 Travel Time Estimation

The methodology described in Section 4.3 is applied to both the NGSIM data and the simulation data. The NGSIM data is available for certain demand levels and does not provide provision to investigate the effects of demand variation. Therefore, simulation data is further used to demonstrate the demand variation on travel time estimation.

5.2.1 Utilizing Real World Vehicle Trajectory Data

The assessment of travel time estimation using CV, as described in the methodology section, is applied to the freeway and arterial segments of NGSIM data used for the purpose of this research. Since the travel times of all vehicles are available based on the NGSIM data, the ground truth travel time can be easily calculated based on these travel times. The ground truth travel time of the freeway segment is 41.3 seconds, and for the arterial segment, 79.1 seconds. The resulting equations for the two case studies are presented in Table 5-2.

Table 5-2: Regression Analysis Results for Travel Time Estimation Error

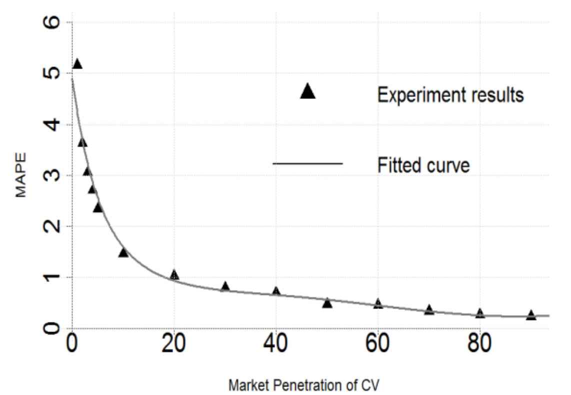
	Error	β0	β1 (*10^{-02})	β2 (*10^{-03})	β3 (*10^{-05})	β4 (*10^{-07})	R-squared value	Adjusted R-squared value	Shapiro-Wilk normality test for model residual (p-value)	Mean of the residual (*10^{-18})
Freeway	**MAPE**	4.963	-15.1	4.52	-6.18	2.89	0.991	0.986	0.998	1.33
	SDPE	6.328	-15.7	4.78	-6.53	3.04	0.993	0.989	0.767	-2.67
	RMSE	2.611	-15.5	4.67	-6.29	2.87	0.993	0.989	0.612	1.34
	MAD	2.064	-15.5	4.75	-6.60	3.13	0.991	0.987	0.100	-7.47
	95%**	10.24	-9.80	1.61	-9.99	-	0.990	0.987	0.931	-1.33
	85%***	8.654	-14.6	4.35	-6.02	2.85	0.988	0.981	0.626	-4.54
Arterial	**MAPE**	13.957	-9.70	1.79	-1.27	-	0.995	0.994	0.153	0.70
	SDPE	18.954	-10.6	1.97	-1.38	-	0.994	0.993	0.114	0.09
	RMSE	14.984	-10.4	1.92	-1.33	-	0.994	0.993	0.194	1.11
	MAD	11.012	-9.62	1.76	-1.24	-	0.995	0.994	0.198	0.22
	95%**	38.590	-10.5	1.85	-1.24	-	0.989	0.986	0.401	3.97
	85%***	26.523	-11.1	2.14	-1.51	-	0.993	0.990	0.120	4.21

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

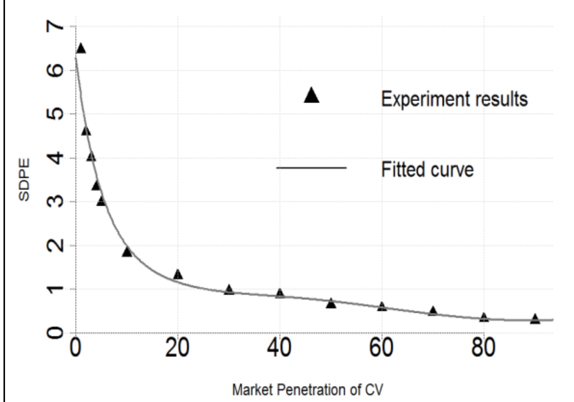
** 95% absolute percentage error curve represents the 5% highest errors among the travel time estimates

*** 85% absolute percentage error curve represents the 15% highest errors among the travel time estimates

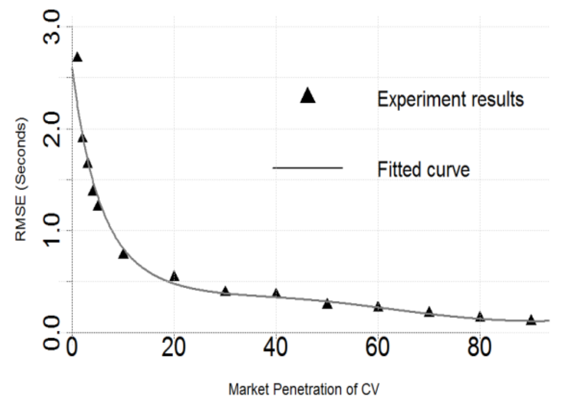
Different types of error measures and the fitted curves to these measures based on regression analysis are shown in Figures 5-5 and 5-6 for the freeway and arterial segments, respectively. The horizontal axis of the plots represents the MP of CV. and the vertical axis represents the error. The regression analysis results are presented in Table 5-2. The exponential equation is transformed into a linear form, and a simple linear regression analysis is performed for this research. All of the β coefficients are considered acceptable if they pass the t-statistics at a 95% confidence interval. For the transformed regression equation, it is recommended that the normality of the model residual and the mean of the residual be checked. For an acceptable transformed linear regression model, the model residual will have a zero mean and a normal distribution. Hence, a Shapiro-Wilk normality test is also conducted on the model residual. The derived equations are used to predict the error at certain market penetrations.



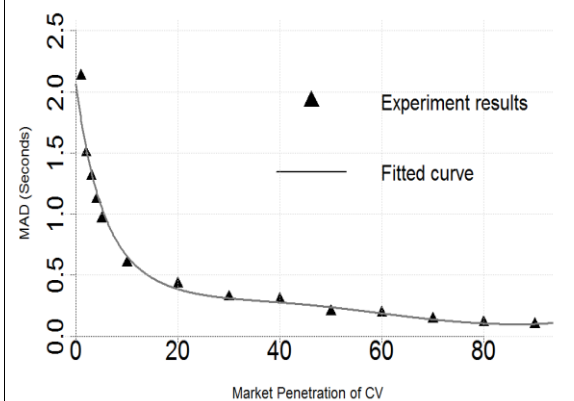
MAPE (%)



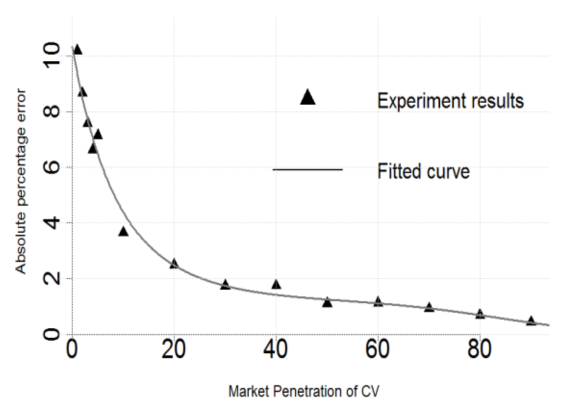
SDPE (%)



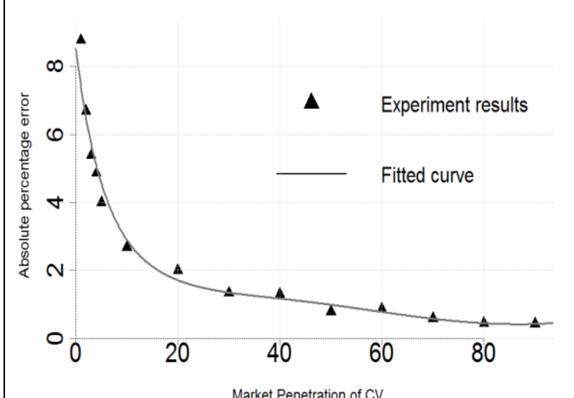
RMSE (Seconds)



MAD (Seconds)



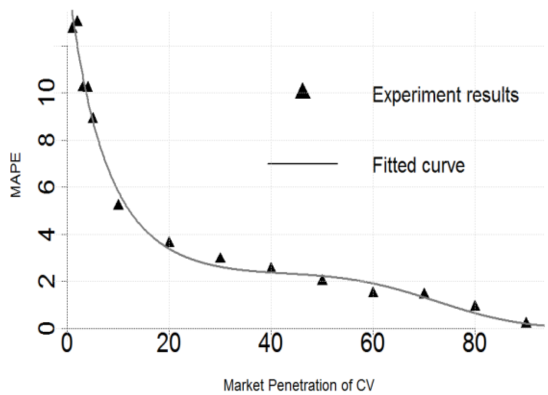
95% *

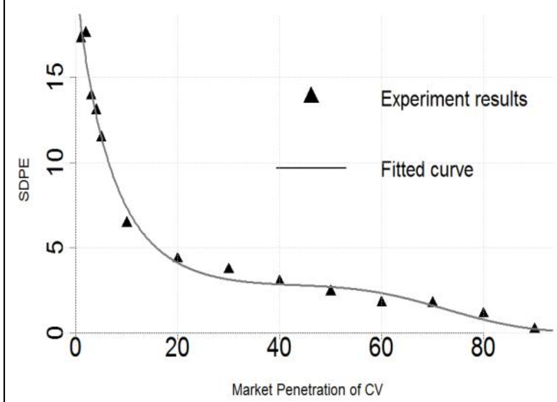


85% *

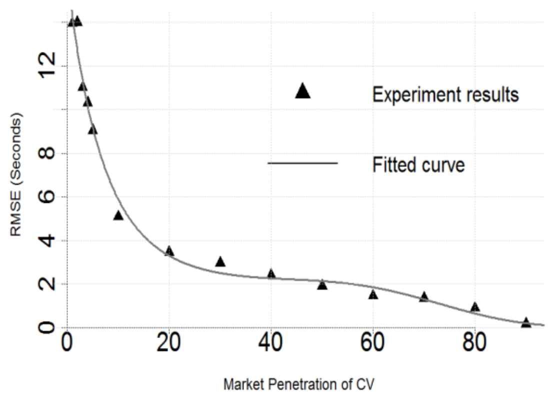
**95% and 85% absolute percentage error curve represents the 5% and 15% highest errors among the travel time estimates*

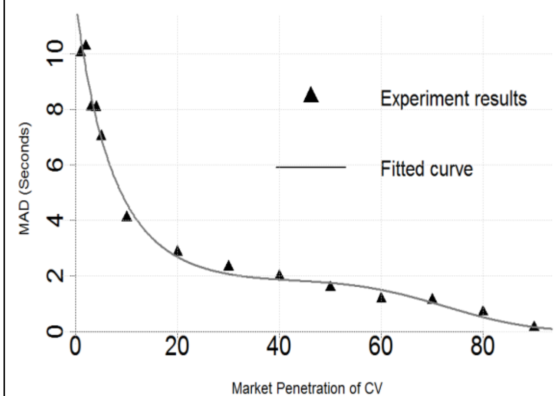
Figure 5-5: Travel Time Accuracy Measures for the Freeway Segment Examined in this Research



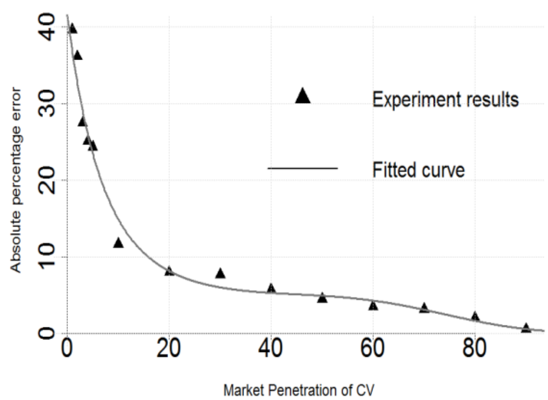


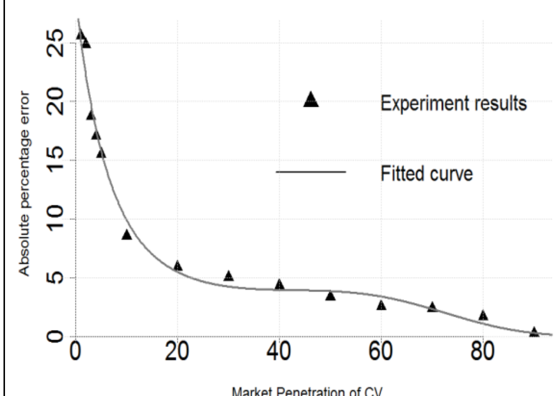
MAPE (%)	SDPE (%)





RMSE (Seconds)	MAD (Seconds)





95% *	85% *

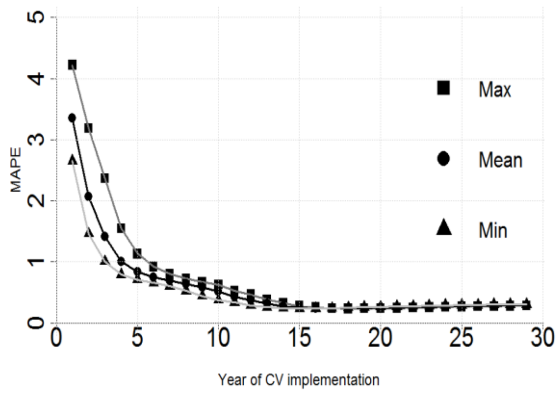
**95% and 85% absolute percentage error curve represents the 5% and 15% highest errors among the travel time estimates*

Figure 5-6: Travel Time Accuracy Measures for the Arterial Segment Examined in this Research

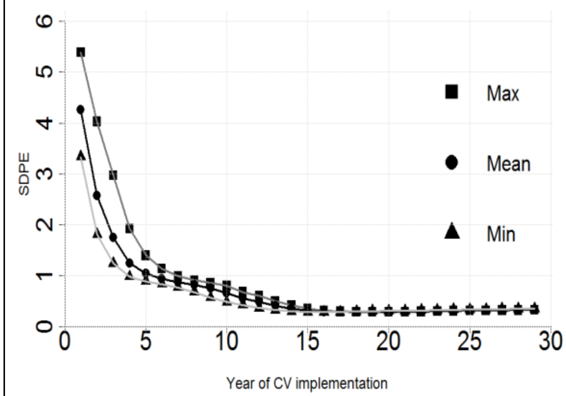
At lower market penetrations, the travel time measurements can vary largely due to the small sample size. At these penetrations, the average of the error represented by the MAPE curve in Figures 5-5 and 5-6 can be acceptable. However, the individual estimates for samples identified based on the Monte Carlo simulation can have large errors for the examined urban arterial when examining the 95% or 85% error curves in Figures 5-5 and 5-6 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 5-5 and 5-6 show that the low market penetrations of 1% on the freeway and about 3%-4% on urban streets are adequate to produce sufficient data quality for planning purposes when averaged over multiple days. A low market penetration (1%-2%) is generally sufficient to produce an error that is lower than 10% for operational use for almost all days for the examined high-demand freeway segment, as indicated in Figure 5-5. However, for the urban street segments (Figure 5-6) this data quality cannot be achieved until the market penetration of CV exceeds 10%-15%. 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 (which was investigated in Section 5.2.2). It was reported that for transportation operation purposes, the estimated travel time accuracy should be less than 10-15% RMSE (Turner et al., 1998). According to this criterion, the desired accuracy (10% RMSE is 4 seconds) for the freeway could be achieved at 1% MP. However, for the urban street segment, the 10% RMSE (8 seconds) can be achieved when the MP increases to 5%. Toppen et al. (2003) reported that if the

SDPE is greater than 12%, then there are no user benefits to travel time use in advanced traveler information systems. As the highest SDPE for the research, the freeway location is about 4.5% at 1% MP, and the CV data could be a reliable source of travel time from the very beginning of the CV mandate implementation. In the case of urban street segments, the reliable travel time could be achieved at an MP of 10%.

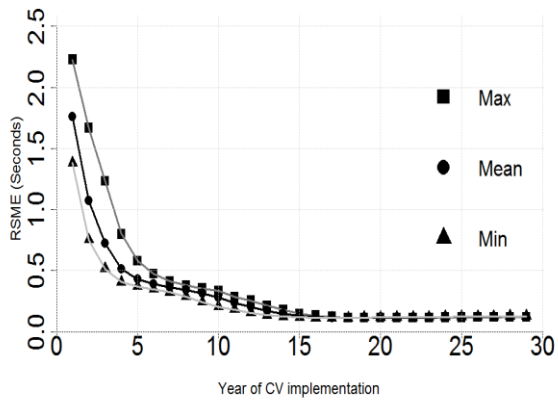
The results in Figures 5-5 and 5-6 present the data quality measures with different market penetrations. The next step is to determine the data quality for each year after the NHTSA mandate. This accuracy is determined for the corridor in this research, with the assumption that the CV market penetration is equal 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 socio-economical characteristics of the users of the facility. Using the derived regression equations developed in Table 5-2 and the CV market penetration variation of Figure 5-2, the different error types for the three above-mentioned scenarios in different years after the CV mandate implementation are calculated and presented in Figures 5-7 and 5-8. 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 and consideration of the variations in the socio-economical characteristics in the region.



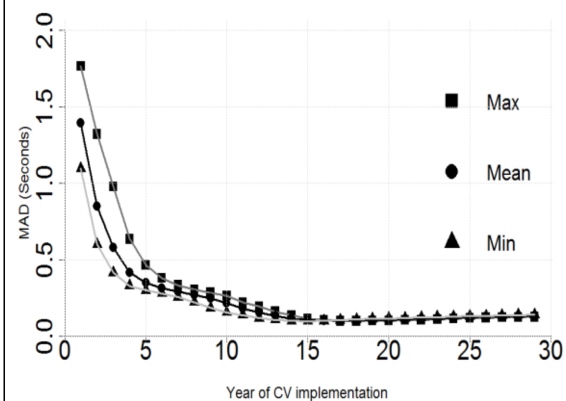
MAPE (%)



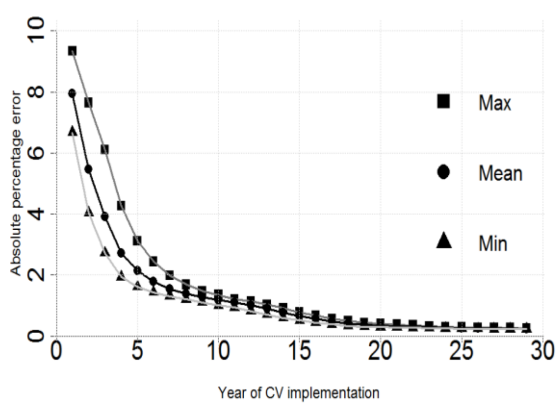
SDPE (%)



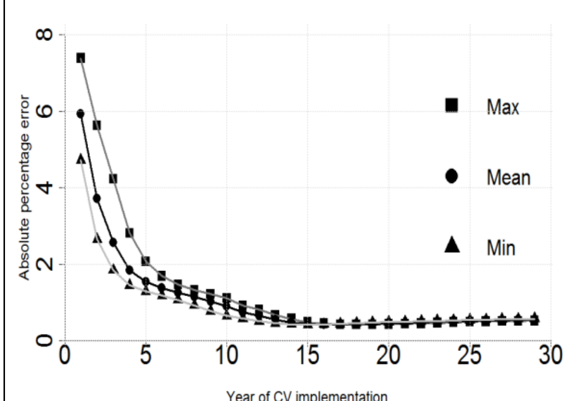
RMSE (Seconds)



MAD (Seconds)



95%**

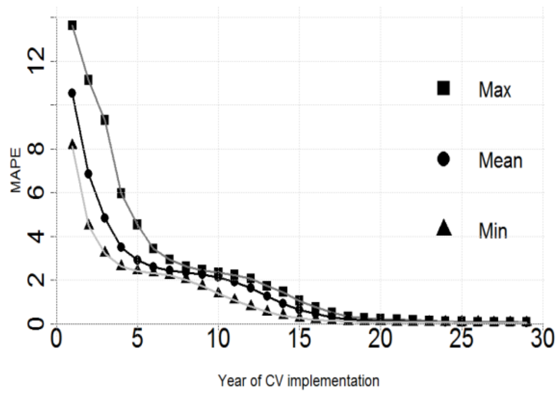


85%**

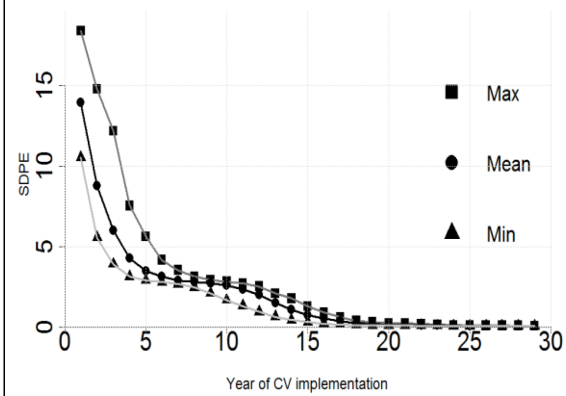
** Max = Maximum, Min = Minimum*

***95% and 85% absolute percentage error curve represents the 5% and 15% highest errors among the travel time estimates*

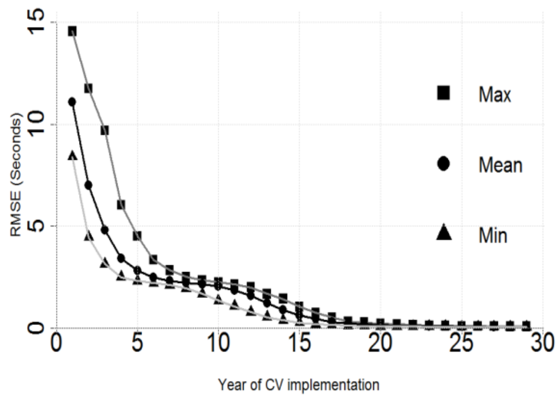
Figure 5-7: Probable Travel Time Accuracy Measures by Year for the Freeway Segment Examined in this Research



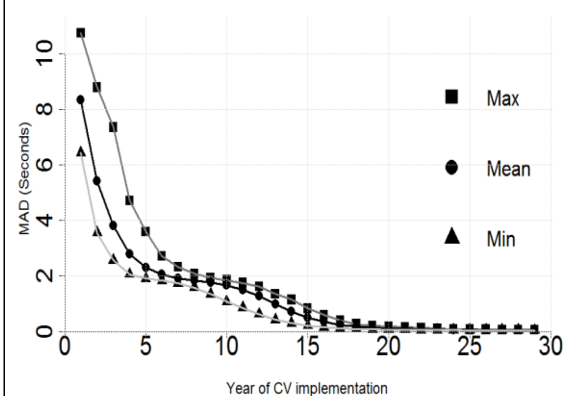
MAPE (%)



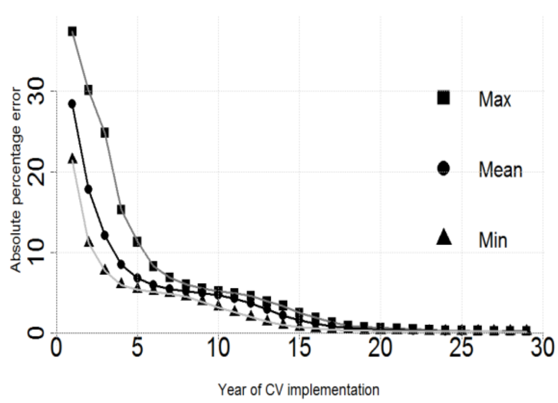
SDPE (%)



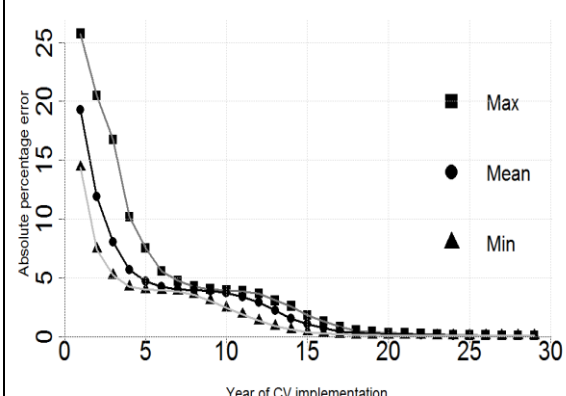
RMSE (Seconds)



MAD (Seconds)



95%**



85%**

** Max = Maximum, Min = Minimum*

***95% and 85% absolute percentage error curve represents the 5% and 15% highest errors among the travel time estimates*

Figure 5-8: Probable Travel Time Accuracy Measures by Year for the Arterial Segment Examined in this Research

5.2.2 Effect of Demand Variation on the Accuracy of Estimated Travel Time

The same methodology is also applied on the simulated test sections. The Mean Absolute Percentage Error (MAPE) and Standard Deviation Percentage Error (SDPE) of the estimated travel time are 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 saturation are 0.3, 0.6, 0.7, 0.8, and 0.9. For the freeway segment, the investigated v/c ratios are 0.36, 0.5, 0.72, 0.86, and 1.01. The results from the Monte Carlo runs are used to develop the regression equations, which are then used to calculate the error percentage at a certain CV proportion and demand level combination. The developed equations and their statistics are provided in Table 5-3.

The developed equations of Table 5-3 are later used to determine the quality of CV-based data in terms of travel time measurement accuracy (in terms of MAPE) and travel time measurement reliability (in terms of SDPE) 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 for different demand levels (time of day). The CV proportion in the traffic stream will increase as new vehicles join the fleet every year. The prediction model developed in Section 5.1 is used to determine the accuracy of performance measurements at different demand levels and for different years after the CV mandate becomes effective.

Table 5-3: Estimated Travel Time Error Equations for Simulation Data

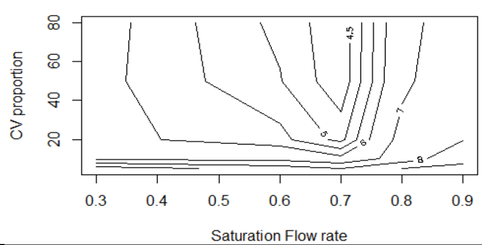
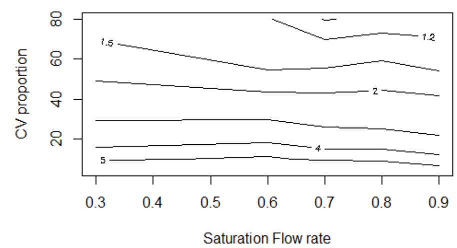
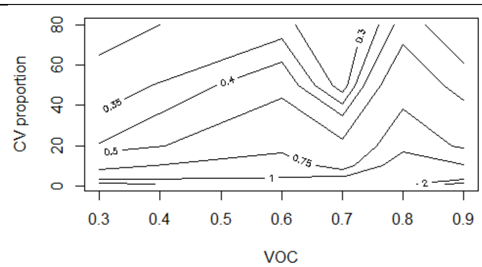
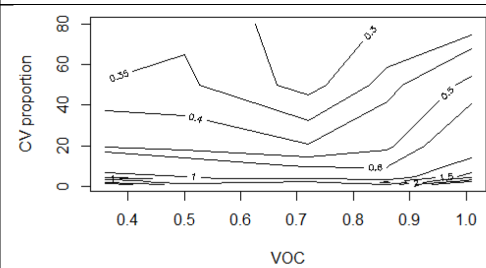
<table>
<tr><th></th><th colspan="4">MAPE</th><th colspan="4">SDPE</th></tr>
<tr><td colspan="9">Arterial Street</td></tr>
<tr><td>Equation</td><td colspan="4">$E_{MAPE}=68.79-52.9*\cos(S)-0.996*\log(CP)-27.3*S$ (5-1)</td><td colspan="4">$E_{SDPE}=9.6+3.04*\cos(S)-3.14*\log(CP)+0.038*CP$ (5-2)</td></tr>
<tr><td>Contour Map</td><td colspan="4">
</td><td colspan="4">
</td></tr>
<tr><td rowspan="7">Coefficients Statistics</td><td>Coefficients</td><td>Std. Error</td><td>t value</td><td>Pr(>|t|)</td><td>Coefficien ts</td><td>Std. Error</td><td>t value</td><td>Pr(>|t|)</td></tr>
<tr><td>β_0</td><td>11.55</td><td>5.96</td><td>6.54e-06</td><td>β_0</td><td>1.09</td><td>8.8</td><td>1.72e-08</td></tr>
<tr><td>β_1</td><td>5.62</td><td>-4.86</td><td>8.41e-05</td><td>β_1</td><td>1.05</td><td>2.88</td><td>0.0089</td></tr>
<tr><td>β_2</td><td>10.46</td><td>-5.2</td><td>3.71e-05</td><td>β_2</td><td>0.36</td><td>-8.62</td><td>2.45e-08</td></tr>
<tr><td>β_3</td><td>0.15</td><td>-6.46</td><td>2.12e-06</td><td>β_3</td><td>0.013</td><td>2.88</td><td>0.009</td></tr>
<tr><td colspan="4">Multiple R-squared: 0.78</td><td colspan="4">Multiple R-squared: 0.94</td></tr>
<tr><td colspan="4">Adjusted R-squared: 0.75</td><td colspan="4">Adjusted R-squared: 0.93</td></tr>
<tr><td colspan="9">Freeway</td></tr>
<tr><td>Equation</td><td colspan="4">$E_{MAPE}=12.05-9.18*\cos(S)-0.52*\log(CP)-4.81*S+0.0085*CP$ (5-3)</td><td colspan="4">$E_{SDPE}=9.59+3.04*\cos(S)-3.14*\log(CP)-0.038*CP$ (5-4)</td></tr>
<tr><td>Contour Map</td><td colspan="4">
</td><td colspan="4">
</td></tr>
<tr><td rowspan="8">Coefficients Statistics</td><td>Coefficients</td><td>Std. Error</td><td>t value</td><td>Pr(>|t|)</td><td>Coefficien ts</td><td>Std. Error</td><td>t value</td><td>Pr(>|t|)</td></tr>
<tr><td>β_0</td><td>3.42</td><td>3.56</td><td>0.001</td><td>β_0</td><td>1.6</td><td>8.8</td><td>1.72e-08</td></tr>
<tr><td>β_1</td><td>2.9</td><td>-3.164</td><td>0.003</td><td>β_1</td><td>1.05</td><td>2.88</td><td>0.0089</td></tr>
<tr><td>β_2</td><td>0.075</td><td>-6.95</td><td>2.2e-08</td><td>β_2</td><td>0.36</td><td>-8.61</td><td>2.45e-08</td></tr>
<tr><td>β_3</td><td>1.81</td><td>-2.66</td><td>0.01</td><td>β_3</td><td>0.01</td><td>2.88</td><td>0.0089</td></tr>
<tr><td>β_4</td><td>0.0085</td><td>0.004</td><td>0.04</td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="4">Multiple R-squared: 0.8</td><td colspan="4">Multiple R-squared: 0.94</td></tr>
<tr><td colspan="4">Adjusted R-squared: 0.78</td><td colspan="4">Adjusted R-squared: 0.93</td></tr>
</table>

Table 5-4: Minimum Accuracy for Different Years after the CV Mandate

v/c	Year	Travel Time Accuracy (%)			
		Freeway		Arterial	
		MAPE	SDPE	MAPE	SDPE
0.1	1	2.66	13.2*	13.60*	13.2
	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	1.23	9.28	7.74	9.29
	3	0.97	7.74	7.23	7.75
	4	0.61	5.5	6.45	5.5
	5	0.46	4.53	6.10	4.53
	10	0.11	2.12	4.99	2.11
	15	0.07	1.56	4.43	1.54
	20	0.09	1.55	4.24	1.53
	25	0.11	1.57	4.18	1.55
1	1	2.5	11.82	13.08*	11.85
	2	1.92	8.26	11.93*	8.29
	3	1.66	6.72	11.42*	6.74
	4	1.3	4.47	10.65*	4.5
	5	1.15	3.50	10.29*	3.52
	10	0.8	1.09	9.18	1.1
	15	0.76	0.53	8.62	0.54
	20	0.78	0.53	8.43	0.52
	25	0.8	0.55	8.37	0.54
* Error more than the minimum acceptable threshold					

Following the guidelines and practices for acceptable errors mentioned in Section 5.2.1, Table 5-4 shows that the accuracy of the travel time is also affected by the v/c ratio. At lower v/c ratios, 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.

The results show that the travel time data estimation on freeways could be done starting from the second year after the CV is mandated for all demand levels. For the arterial facility, it will take four to five years to obtain acceptable accuracy levels of travel time estimation.

5.3 Volume Estimation

To fulfill the third objective, the methodology proposed in Section 4.4 is applied on the simulated arterial section. The results of different scenarios are presented in the subsequent sections.

5.3.1 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 5-5.

Table 5-5: 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

** Bold values have low variation: MAPD <5 (95%APD <12)*

Table 5-5 shows that there is a relatively low variation between the volume counts of Locations 1 and 2, and between the volume counts of Locations 4 and 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 Locations 2 and 4. In this case, the traffic measurements at Location 1 are expected to reflect the traffic pattern at Location 2 with a MAPD of 4.51% and 95%APD of 9.55%. Location 5 is expected to reflect the traffic pattern at Location 4 with a MAPD of 4.84% and a 95% APD of 10.87%, based on the results presented in Table 5-6.

Table 5-6: Volume Estimation Error with Different Scenarios and Associated Methods

Year After CV Mandate is effective	Mean CV Proportion (Iqbal et al. (2017))	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

5.3.2 Future Scenario 1

Given the above discussion regarding the Base Scenario, this research 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 Locations 2 and 4, according to the Scenario 1 methodology. The resulting MAPD and 95% APD are presented in this paper for estimating the volume at Location 2 utilizing an expansion factor calculated based on Location 1 volume data. As can be seen in Table 5-6, it will take four years after the mandate on all new vehicles becomes effective for CV data to become beneficial in supplementing existing detectors. After four years, the use of CV data will provide volume estimates on the non-instrumented vehicles, which 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.

5.3.3 Future Scenario 2

Scenario 1 involves installing permanent detectors at Locations 1, 3, and 5. In Scenario 2, the research 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, the two methods for the utilization of the partial volume counts from CV in the estimation are the 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 develop the regression model. Table 5-7 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 non-instrumented locations.

Table 5-7: Regression Equations Developed for Estimating Volume at Location 3

Year 1				
Equation	$V_3^{est} = 0.071 + 0.469 * V_1^{act} + 0.447 * V_5^{act} - 10.316 * CV_2 + 9.323 * CV_3$ (5-5) where V_1^{act}, V_5^{act} are the actual volumes at locations 1 and 5 CV_2, CV_3 are the CV volumes at locations 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	$V_3^{est} = 11.6647 + 0.741 * V_5^{act} + 3.6684 * CV_3 - 2.9492 * CV_5$ (5-6) where $V_5^{act}$are the actual volume at location 5; CV_3, CV_5 are the CV volumes at locations 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			

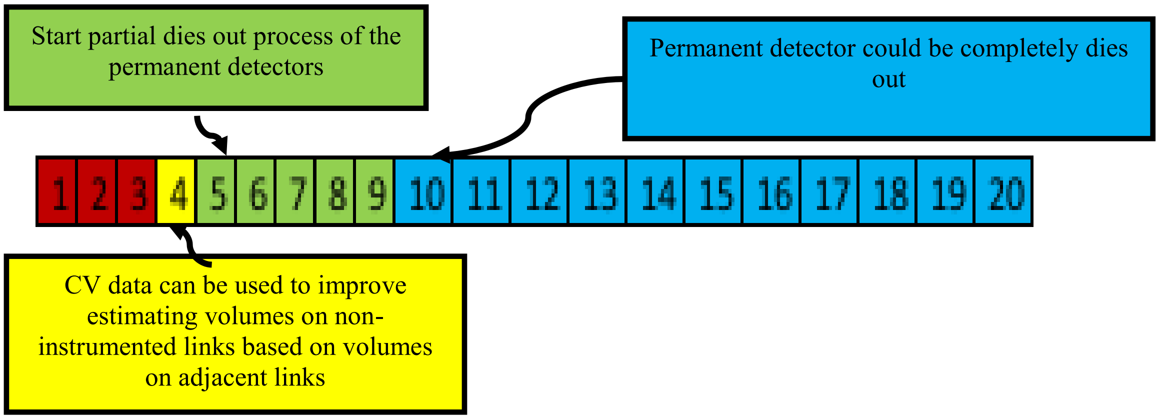
The main drawback of Method 1 is that it needs a temporary detector to develop the model. Furthermore, the model would need to be updated at a regular interval to reflect changes in traffic patterns. This would increase the cost of the data collection. To overcome this drawback of Method 1, the optimization 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 Table 5-6. As can be seen from the table, utilizing both methods to estimate the volumes for Location 3 results in a MAPD of less than 10% in 4-5 years, and Method 2 can produce a 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, which does not require temporary counts, actually produced better results than Method 1 in future years (beyond Year 7), which have increased CV market penetration, although Method 1 performed better in the first few years. The performance of Method 1 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.

5.3.4 Future Scenario 3

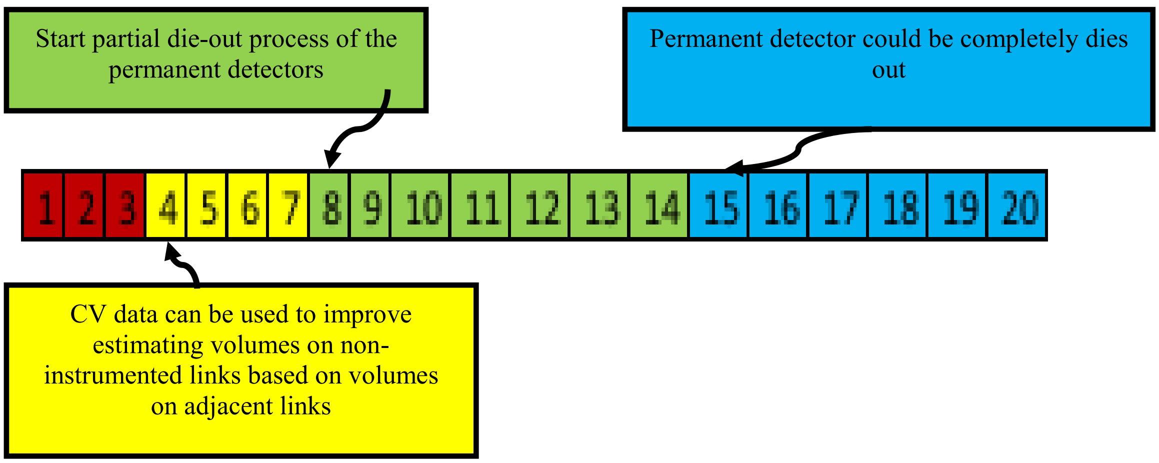
The last scenario (Scenario 3) involves removing all detectors. At the early stage of this scenario, the volume estimation of a location could be supported by temporary detectors to calculate the expansion factor, as mentioned in the Methodology section. The accuracy of this scenario is demonstrated for Location 3, as shown in Table 5-6. The results in the table indicate that in order 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.

5.3.5 Identification of Detector Die-Out Timeline

The results obtained based on the analysis of the previous section can be used to prepare a probable timeline for the detector die-out 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 is shown in Figures 5-13(a) and 5-13(b), for 5% and 10% volume estimation accuracy requirements, respectively.



(a) 10% Acceptable Error



(b) 5% Acceptable Error

Figure 5-9: Timeline for Detector Die-Out from the Study Location

5.4 Summary

The application of the methodology presented in this chapter shows that the market penetration of CV is lognormally distributed, and the variation of MP decreases exponentially with the increase of the MP in future years. The accuracy of the travel time also follows an exponential function. The accuracy of estimated travel time is lower at lower market penetrations and increases exponentially with the increase of the MP. Connected vehicle data could also be used to calculate the traffic volume, either in conjunction with detector data or independently with varying accuracy, depending on the MP and deployment scenario and setup.

The application result shows that the CV data could be used for travel time estimation starting the second year of the implementation of the CV mandate for freeways. For an arterial, it would take up to five years to reliably estimate the travel time. For volume estimation, the partial die-out of detectors could be started five years after the CV mandate and could be completely removed after ten to fifteen years.

CHAPTER VI

CONCLUSION AND RECOMMENDATIONS

6.1 Summary and Conclusions

The traveler information system is defined as a system that collects traffic data, and analyzes and disseminates the information to travelers to make their trips more convenient, more efficient, and safer. With the availability of information, travelers can choose the best routes to reach their destinations and possibly select a suitable departure time and mode for their trips. Over the course of time, the technologies to support the traveler information system have been evolving. At the early stage of the traveler information system, the data collection was limited to manual data collection. Later, the data collection techniques also relied on detector data and Closed Circuit Television Camera (CCTV) installations. More recently, there have been several promising technologies to collect traffic data, including vehicle re-identification based on Bluetooth, Wi-Fi, and Magnetometer technologies. Apart from that, third-party vendors are also playing a vital role in traffic data collection. The introduction of Connected Vehicle (CV) technology will provide a potential alternative source to collecting traffic data for the traveler information system.

The main goal of this research was to develop a method to identify a timeline indicating when CV-based technologies can be used as a more effective alternative to support the traveler information systems data collection process, compared to other existing technologies. The specific objectives of this research were:

- Develop a method for the prediction of the future market penetration of connected vehicles.
- Develop a method to evaluate the use of connected vehicle data to provide accurate and reliable travel time.
- Develop a method to evaluate the use of connected vehicle data to provide accurate and reliable volume estimates that can support traveler information systems by indicating traffic statuses such as congestion, incident, and flow.

The data collection and preparation process was the key component in accomplishing the research objectives. The major hindrance in the data preparation process was the unavailability of connected vehicle data. The CV data is now available from a few testbeds and pilot projects around the United States. However, the CV proportions in the traffic stream associated with these implementations are very low. Therefore, this research utilized an emulation of connected vehicle data by inputting vehicle trajectory data collected in a previous national effort and was input into a tool called the Trajectory Conversion Algorithm (TCA), which was developed by the Federal Highway Administration (FHWA). The tool converts the trajectory data into CV data, according to the SAE J2735 standards (SAE International, 2016). This research utilized real-world trajectory data collected under the NGSIM program. In addition, simulation modeling was used to generate trajectories for input into the TCA tool. The reason for using the simulation was to have the flexibility to change the demand and proportion of CV and examine their impacts on the assessment results. An arterial segment and a freeway segment were coded and calibrated in the microscopic simulation for this purpose.

The first step of this research was to predict the market penetration of connected vehicles for future years. Past efforts have assumed the growth in CV market penetrations without considering the variations in the socio-economical characteristics between regions and zones within a region. This research proposed a methodology to determine the variation of CV market penetration between regions, zones within a certain region, links within the region, and time of day. The developed methodology can be implemented with various CV implementation scenario assumptions, and considers the variations in the socioeconomic characteristics of travelers to a region.

The results of the research indicated that the distribution of the link-specific CV market penetration 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 market penetration variations between links are the highest on collectors, followed by arterials, and followed by freeways. The results also showed that the average percentage increase in the CV market penetration grows in the first several years, and then remains almost constant before dropping sharply.

The second objective of this research was to determine the accuracy and reliability of travel time estimation utilizing CV data. This research assessed the quality of travel time estimates based on CV data on the freeway and urban street segments. The data quality was examined under different market penetration scenarios considering the randomness in the presence of CV son 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 case study on real-world scenarios, it can be stated that the CV market penetration will be sufficient for use in the planning and real-time operations of the freeway segment in the first year after the expected mandate to install CV technology into all new vehicles becomes 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 proportion of the CV, considering the variations in the socio-economic characteristics in the region. Another case study based on simulation results shows that the accuracy of the travel time is also affected by the demand variation. At lower v/c ratios, 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 the travel time between vehicles due to congestion, which results in a lower accuracy of the estimated travel time.

The final objective of this research was to evaluate the accuracy of the estimation of traffic volume utilizing CV data. 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, and combined with partial volume data collected utilizing the CV. Eventually, at high market penetration levels, it may be possible to remove most, if not all, of the detectors. This research considered three different scenarios to estimate the volume utilizing the CV data. The results from applying the methodology to a case study indicated that after four years of the application of the mandate of CV into 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 be possible to start removing some of the detectors within 5 to 8 years after the CV mandate. The agencies can remove all detectors within 10 to 15 years after the CV mandate, depending on the accuracy threshold.

6.2 Research Contribution

The connected vehicle is an emerging technology that promises to provide detailed transportation system data for use in different planning and operation applications. This research focused on utilizing the CV data to support traveler information system applications. This dissertation provided a unique methodology to determine the market penetration of CV for future years after applying the expected CV mandate, considering the socioeconomic variation between different zones.

This research also provided a methodology to estimate the travel time and volume based on CV data and their accuracy and reliability throughout different years after the CV mandate. Results of this research can be used by transportation system management and operations (TSM&O) programs and agencies to plan their investment in data collection alternatives. They can also be used to support the preparation of a timeline to replace the existing data collection systems with CV technology, depending on the traffic demand of a specific roadway.

6.3 Recommendation for Future Research

Future studies to extend this dissertation research could include the following:

1. This research focused on the estimation of the two main performance measurements of a roadway: travel time and volume. Other performance

measurements such as occupancy, density, headway, and possible new measures could also be estimated utilizing the CV data. Further studies should be conducted on the estimation of these performance measurements based on CV data.

2. This research utilized only the BSM part I data, which is expected to be available in the initial phase of CV implementation. Other data types, as specified in the J2735 standards, could also be used to derive additional measures and improve the accuracy and reliability of the performance measurements.
3. CV technology can also be used to support traveler information dissemination, which can provide additional benefits, compared to the use of existing technology such as dynamic message signs, highway advisory radio, and the 511 phone system. Further studies should be conducted to assess those benefits.

REFERENCES

Abdalla, M. F., & Abdel-Aty, M. (2006). Modeling travel time under ATIS using mixed linear models. *Transportation*, 33(1), 63-82.

Abdel-Aty, M. A., Kitamura, R., & Jovanis, P. P. (1997). Using stated preference data for studying the effect of advanced traffic information on drivers' route choice. *Transportation Research Part C: Emerging Technologies*, 5(1), 39-50.

ACS website (2016). *American Community Survey 5-year estimate – Geodatabase format*. Retrieved on May 23, 2016, from https://www.census.gov/geo/maps-data/data/tiger-data.html.

Adler, J. L. (2001). Investigating the learning effects of route guidance and traffic advisories on route choice behavior. *Transportation Research Part C: Emerging Technologies*, 9(1), 1-14.

Adler, J. L., & Blue, V. J. (1998). Toward the design of intelligent traveler information systems. *Transportation Research Part C: Emerging Technologies*, 6(3), 157-172.

Adler, J., Blue, V., & Wu, T. L. (1999). Assessing network and driver benefits from bi-objective in-vehicle route guidance. In *78th Annual Meeting of the Transportation Research Board, Washington, DC*.

Adler, J., J. Bottom, C. Nelson, & K. Wunderlich. (2014). *EnableATIS strategy assessment* (FHWA-JPO-14-113). Washington, D.C.

Al-Deek, H., Ishak, S., & Radwan, A. E. (1993). The potential impact of advanced traveler information systems (ATIS) on accident rates in an urban transportation network. In *Vehicle Navigation and Information Systems Conference, Proceedings of the IEEE-IEE* (pp. 633-636). IEEE.

Al-Deek, H. & Kanafani, A. (1993). Modeling the benefits of advanced traveler information systems in corridors with incidents, *Transportation Research Part C*, 1, 303–324.

Al-Deek, H., Khattak, A. J., & Thananjeyan, P. (1998). A combined traveler behavior and system performance model with advanced traveler information systems. *Transportation Research Part A: Policy and Practice*, 32(7), 479-493.

Al-Deek, H., Martello, M., May, A. D., & Sanders W. (1989*). Potential benefits of in-vehicle information systems in a real-life freeway corridor under recurring and incident induced congestion* (UCB-ITS-PRR-88-2). The University of California.

Al-Deek, H., Sandt, A., Alomari, A., Rogers, J., & Muhaisen, N. (2016). *Evaluating the impact and usefulness of highway advisory radio (HAR) and citizens' band radio advisory systems (CBRAS) in providing traveler information and improving the user experience on the Florida turnpike enterprise's toll road network and the Florida interstate highway (FIH) system* (BDV24-977-11). Florida Department of Transportation. Retrieved on January 10, 2017, from http://ntl.bts.gov/lib/57000/57100/57194/FDOT-BDV24-977-11-rpt.pdf.

Al-Nuaimi, H., Ayers, R, & Somasundaraswaran, K. (2013). Traffic flow modeling for urban arterial road in BRISBANE city. *Presented at 14th REAAA Conference*.

Anderson, M. (2015). *Technology device ownership: 2015*. Pew Research Center: Global Attitudes & Trends.

Anderson, M., & Souleyrette, R. (2002). Pseudo-dynamic travel model application to assess traveler information, *Transportation*, 22, 307-319.

Andrews, S., & Cops, M. (2009). *Final report: Vehicle infrastructure integration proof of concept technical description–Vehicle. The VII Consortium* (FHWA - JPO-09-043). U.S. Department of Transportation, Washington, D.C.

Anuar, K., & Cetin, M. (2017). Estimating freeway traffic volume using shockwaves and probe vehicle trajectory data. *Transportation Research Procedia*, 22, 183-192.

Balakrishna, R., Koutsopoulos, H., Ben-Akiva, M., Fernandez Ruiz, B., & Mehta, M. (2005). Simulation-based evaluation of advanced traveler information systems. *Transportation Research Record: Journal of the Transportation Research Board*, 1910, 90-98.

Bottom, J., Ben-Akiva, M., Bierlaire, M., Chabini, I., Koutsopoulos, H., & Yang, Q. (1999). Investigation of route guidance generation issues by simulation with DynaMIT. In *Proceedings of the 14th International Symposium on Transportation and Traffic Theory, ed. Avishai Ceder, 577–600. Pergamon*.

Burgess, L. Toppen A., & Harris, M. (2012). *Vision and operational concept for enabling advanced traveler information systems (EnableATIS)* (FHWA-JPO-12-052). U.S. Department of Transportation, Washington, D.C.

Caceres, N., Romero, L. M., Benitez, F. G., & del Castillo, J. M. (2012). Traffic flow estimation models using cellular phone data. *IEEE Transactions on Intelligent Transportation Systems*, 13(3), 1430-1441.

Cambridge Systematics (2011). *Real-time data capture and management evaluation and performance measures - evaluation framework* (FHWA-JPO-11-136). U.S. Department of Transportation, Washington, D.C.

Carter, M., Cluett, C., DeBlasio, A. Dion, F., Hicks, B., Lappin, J., Novak, D., Rakha, H., Riley, J., St-Onge, C., & Van Aerde, M. (2000). *Metropolitan model deployment initiative San Antonio evaluation report* (FHWA-OP-00-017). McLean, VA: Federal Highway Administration.

Chen, W. H., & Jovanis, P. (2003). Driver en route guidance compliance and driver learning with advanced traveler information systems: analysis with travel simulation experiment. *Transportation Research Record: Journal of the Transportation Research Board*, (1843), 81-88.

ConSysTec & Cambridge Systematics (2013). *Real-time system management information program data exchange format specification* (FHWA-HOP-13-047). U.S. Department of Transportation, Washington, D.C.

Crowson, G. & Deeter, D. (2012). *Understanding utilization of third party data and information* (ENTERPRISE Pooled Fund Study TPF-5(231) ENT-2012-1). Lansing, MI.

Deeter, D. (2009). *Real-time traveler information systems* (NCHRP Synthesis 399). Transportation Research Board, Washington DC.

Demissie, M., Correia, G. H. & Bento, C. L. (2013). Traffic volume estimation through cellular networks handover information. *13th World Conference on Transportation Research*, Rio de Janeiro, Brazil.

Dia, H. (2002). An agent-based approach to modelling driver route choice behaviour under the influence of real-time information. *Transportation Research Part C: Emerging Technologies*, 10(5), 331-349.

Doan, Q. C., Mouzna, J., & Berradia, T. (2010). Travel time measurement system using vehicle to infrastructure communication. *ESIGELEC Saint Etienne du Rouvray*, 76801.

EasyWay (2012). Traveller information services reference document: tis deployment guideline (TIS-DG07). European Commission. Retrieved on January 10, 2017, from http://www.transport.gov.mt/admin/uploads/media-library/files/comodaltravelerinformationservices.pdf.

Emmerink, R., Axhausen, K., Nijkamp, P., & Rietveld, P. (1995). The potential of information provision in a simulated road transport network with non-recurrent congestion. *Transportation Research Part C,* 3, 183–309.

Emmerink, R., Axhausen, K., & Rietveld, P. (1996). Effects of information in road transport networks with recurrent congestion. *Transportation*, 22, 21–53.

Enterprise pooled funded study (2015). *Connected vehicle pre-deployment guidance examples for vehicle to infrastructure (V2I) applications: reduced speed zone warning and advanced traveler information system (ATI*S). Draft Report, FHWA, Washington, D.C.

Fernandez L., J.E, de Cea Ch, J., & Valverde, G. G. (2009). Effect of advanced traveler information systems and road pricing in a network with non-recurrent congestion. *Transportation Research Part A: Policy and Practice*, 43(5), 481-499.

FHWA, (2005). *America's travel information number implementation and operational guidelines for 511 services*, Published by the 511 Deployment Coalition, version 3.0. Retrieved on September 4, 2016, from https://ops.fhwa.dot.gov/511/resources/publications/511guide_ver3/511guide3.htm

FHWA. (2014). *Traffic data quality measurement* (FHWA Report 14058), Washington, D.C.

FHWA Website. *511: America's traveler information telephone number*. Retrieved on September 4, 2016, from http://www.fhwa.dot.gov/trafficinfo/511.htm.

Fischer, J. W. (2005). *Safe, accountable, flexible, efficient transportation equity act-a legacy for users (SAFETEA-LU or SAFETEA): Selected Major Provisions*. vol. 307, no. c. 2005, pp. 1–314.

Fox, J., & Boehm-Davis, D. (1998). Effects of age and congestion information accuracy of advanced traveler information systems on user trust and compliance. *Transportation Research Record: Journal of the Transportation Research Board*, (1621), 43-49.

Gan, H., Sun, L., Chen, J., & Yuan, W. (2006). Advanced traveler information system for metropolitan expressways in Shanghai, China. *Transportation Research Record: Journal of the Transportation Research Board*, (1944), 35-40.

Gandy, R., & Meitner, M. J. (2007). The effects of an advanced traveler information system on scenic beauty ratings and the enjoyment of a recreational drive. *Landscape and urban planning*, 82(1), 85-93.

Gardes, Y. & May, A. (1993). *Simulation of IVHS on the Santa Monica freeway corridor using the INTEGRATION model. Phase 2: preliminary ATIS and ATMS experiments* (UCB–ITS– PWP–93–6). Berkeley: University of California.

Goulias, K. G., Kim, T. G., & Pribyl, O. (2004). A longitudinal analysis of awareness and use for advanced traveler information systems. *Intelligent transportation systems*, 8(1), 3-17.

Grone, B. W. (2012). *An evaluation of non-intrusive traffic detectors at the NTC/NDOR detector test bed* (Doctoral dissertation).The University of Nebraska, Lincoln, Nebraska.

Hadi, M., Xiao, Y., Wang, T., Qom, S.F., Azizi, L., Jia, J., Massahi, A., & Iqbal, M. S. (2016). *Framework for multi-resolution analyses of advanced traffic management strategies* (BDV29-977-19). FDOT.

Hadj-Alouane, A., Hadj-Alouane, N., Juma, O., Sarathy, G., & Underwood, S. (1996). The Ali-Scout route guidance simulation. *FAST-TRAC phase IIB deliverable*.

HCM (2010). Highway capacity manual, The Transportation Research Board of the National Academies, Washington D.C.

Hellinga, B., & Gudapati, R. (2000). Estimating link travel times for advanced traveler information systems. In *Proceedings of the Canadian Society of Civil Engineers, 3rd Transportation Specialty Conference*. London, Ontario, Canada.

Hobeika, A., Sivanandan, R., Jehanian, K., & Ameen, M. (1996). Advanced traveler information system users' needs in I-95 Northeast Corridor. *Transportation Research Record: Journal of the Transportation Research Board*, 1537, 55-62.

Inman, V., Sanchez, R, Porter, C., & Bernstein, L. (1995). *TravTek evaluation yoked driver study* (FHWA–RD–94–139). McLean, VA: Federal Highway Administration.

Iteris website. *Provide driver and traveler services, connected vehicle reference implementation architecture (CVRIA)*. Retrieved on September 4, 2016, from http://www.iteris.com/cvria/html/pspecs/dfd6211.html.

Jeng, S. T. (2007). *Real-time vehicle reidentification system for freeway performance measurements* (Doctoral dissertation), University of California, Irvine. Retrieved on September 4, 2016, from https://escholarship.org/uc/item/2fh092zk#page-1.

Jou, R. C. (2001). Modeling the impact of pre-trip information on commuter departure time and route choice. *Transportation Research Part B: Methodological*, 35(10), 887-902.

Jung, S., Larkin, J., Shah, V., Toppen, A., Vasudevan, M., & Wunderlich, K. (2002). *On-time reliability impacts of advanced traveler information services (ATIS), Volume II: extensions and applications of the simulated yoked study concept* (FHWA-JPO-04-074). U.S. Department of Transportation, Washington, DC.

Kanafani, A. & Al-Deek, H. (1991). A simple model for route guidance benefits, *Transportation Research Part B*, 25, 191–202.

Kandarpa, R., Chenzaie, M., Anderson, J., Marousek, J., Weil, T., Perry, F., Schworer, L., Beal, J., & Anderson, C. (2009). *Vehicle infrastructure integration proof-of-concept results and findings - infrastructure, volume 3b, Final Report* (FHWA-JPO-09-057). U.S. Department of Transportation, Washington, DC.

Khattak, A. J. (1993). *Behavioral impacts of recurring and incident congestion and response to Advanced Traveler Information Systems in the Bay Area: An overview* (*UCB-*ITS-PWP-93-12). California Partners for Advanced Transit and Highways (PATH).

Khattak, A. J., & De Palma, A. (1997). The impact of adverse weather conditions on the propensity to change travel decisions: a survey of Brussels commuters. *Transportation Research Part A: Policy and Practice*, 31(3), 181-203.

Khattak, A., Polydoropoulou, A., & Ben-Akiva, M. (1996a). Modeling revealed and stated pre trip travel response to advanced traveler information systems. *Transportation Research Record: Journal of the Transportation Research Board*, (1537), 46-54.

Khattak, A., Polydoropoulou, A., & Ben-Akiva, M. (1996b). *Commuters' normal and shift decisions in unexpected congestion:* pre-*trip response to advanced traveler information systems* (UCB-ITS-PRR-96-7). California Partners for Advanced Transit and Highways (PATH).

Khattak, A. J., Schofer, J. L., & Koppelman, F. S. (1995). Effect of traffic information on commuters' propensity to change route and departure time. *Journal of Advanced Transportation*, 29(2), 193-212.

Khattak, A., Pan, X., Williams, B., Rouphail, N., & Fan, Y. (2008). Traveler information delivery mechanisms: Impact on consumer behavior. *Transportation Research Record: Journal of the Transportation Research Board*, (2069), 77-84.

Kim, H. K. (2010). *Development and evaluation of advanced traveler information system (ATIS) using vehicle-to-vehicle (V2V) communication system* (Doctoral dissertation, Georgia Institute of Technology, Atlanta, Georgia). Retrieved from https://smartech.gatech.edu/bitstream/handle/1853/33828/kim_hoekyoung_201005_phd.pdf?sequence=1&isAllowed=y.

Kim, H. K., Hunter, M. P., & Fujimoto, R. M. (2009). A simulation-based investigation of a dynamic advanced traveler information system. In *Proceedings of the 2009 Winter Simulation Conference (WSC)* (pp. 1124-1135). IEEE.

Kim, S.J., Wonho K., & Rilett L. (2005). Calibration of microsimulation models using nonparametric statistical techniques. *Transportation Research Record: Journal of the Transportation Research Board,* 1935,111-119.

Klein, L. A., Mills, M. K., & Gibson, D. R. (2006). *Traffic detector handbook: -volume ii* (No. FHWA-HRT-06-139). Retrieved from http://www.fhwa.dot.gov/publications/research/safety/ip90002/index.cfm.

Koutsopoulos, H., & Lotan, T. (1990). Motorist information systems and recurrent traffic congestion: Sensitivity analysis of expected results, *Transportation Research Record: Journal of the Transportation Research Board*, 1281, 148-158.

Kraan, M., Mahmassani, H., & Huynh, N. (2000). Traveler responses to advanced traveler information systems for shopping trips: Interactive survey approach. *Transportation Research Record: Journal of the Transportation Research Board*, 1725, 116-123.

Kuhn, B., Higgins, L., Nelson, A., Finley, M., Ullman, G., Chrysler, S., Wunderlich, K., Shah, V., & Dudek, C. (2014). *Effectiveness of different approaches to disseminating traveler information on travel time reliability* (S2-L14-RW-1). Retrieved from http://onlinepubs.trb.org/onlinepubs/shrp2/SHRP2prepubL14.pdf.

Lappin, J. E. (2000). What Have We Learned About Advanced Traveler Information Systems and Customer Satisfaction? *Chapter 4, Publication of Island Press*. Retrieved from http://ntl.bts.gov/lib/jpodocs/repts_te/13320.pdf.

Levinson, D. (2003). The value of advanced traveler information systems for route choice. *Transportation Research Part C: Emerging Technologies*, 11(1), 75-87.

Liu, Y. C. (2000). Effect of advanced traveler information system displays on younger and older drivers' performance. *Displays*, 21(4), 161-168.

Liu, Y. C. (2001). Comparative study of the effects of auditory, visual and multimodality displays on drivers' performance in advanced traveller information systems. *Ergonomics*, 44(4), 425-442.

Lo, H. & Szeto, W. (2002). A methodology for sustainable traveler information services, *Transportation Research Part B*, 36, 113–130.

Mahmassani, H. & Chen, S.T. (1991). Comparative assessment of origin-based and en route real-time information under alternative user behavior rules, *Transportation Research Record: Journal of the Transportation Research Board*, 1306, 69–81.

Mahmassani, H. & Chen, S.T. (1993). An investigation of the reliability of real-time information for route choice decisions in a congested traffic system, *Transportation*, 20, 157–178.

Mahmassani, H., & Jayakrishnan, R. (1991). System performance and user response under real-time information in a congested traffic corridor, *Transportation Research Part A*, 25, 293–307.

Mahmassani, H. S., & Liu, Y. H. (1999). Dynamics of commuting decision behaviour under advanced traveller information systems. *Transportation Research Part C: Emerging Technologies*, 7(2), 91-107.

Mai, C., McDaniel-Wilson, C., Norval, D., Upton, D., Auth, J., Schuytema, P., Abbott, S. & Delahanty, R. (2011). *Protocol for VISSIM simulation*, Oregon Department of Transportation.

McGurrin, M. (2012). *Vehicle information exchange needs for mobility applications exchange* (FHWA-JPO-12-02). Washington. D.C.

Mehndiratta, S., Kemp, M., Lappin, J., & Nierenberg, E. (2000). Likely users of advanced traveler information systems: Evidence from the Seattle Region. *Transportation Research Record: Journal of the Transportation Research Board*, (1739), 15-24.

Miller, T., Davis, W., Reed, G., Doraiswamy, P., & Tang, A. (2002). Effect of county-level income on vehicle age distribution and emissions. *Transportation Research Record: Journal of the Transportation Research Board*, 1815, 47-53.

Minge, E., Kotzenmacher, J., & Peterson, S. (2010). *Evaluation of non-intrusive technologies for traffic detection* (No. MN/RC 2010-36). Minnesota Department of Transportation, Research Services Section.

MORPC (2009). *Advanced traveler information system study*, Final report, Mid-Ohio Regional Planning commission. Retrieved from http://www.morpc.org/trans/AdvancedTravelerInfo_FinalReport.pdf.

Mouskos, K. C., & Greenfeld, J. (1999). A GIS-based multimodal advanced traveler information system. *Computer-Aided Civil and Infrastructure Engineering*, 14(4), 267-279.

Nantes, A., Billot, R., Miska, M., & Chung, E. (2013). Bayesian inference of traffic volumes based on Bluetooth data. In *92nd Annual Meeting of the Transportation Research Board, Washington, DC*.

Ng, L., Barfield, W., & Mannering, F. (1995). A survey-based methodology to determine information requirements for advanced traveler information systems. *Transportation Research Part C: Emerging Technologies*, 3(2), 113-127.

NGSIM, *Home of Next Generation Simulation community*. Retrieved on July 10, 2015, from http://ngsim-community.org.

Nihan, N. L., Wang, Y., & Zhang, X. (2002). *Evaluation of dual-loop data accuracy using video ground truth data* (WA-RD 535.1). Transportation Northwest, Department of Civil Engineering, University of Washington.

NTIA Report (2013). *Exploring the digital nation: America's emerging online experience.* Economics and Statistics Administration and National Telecommunications and Information Administration Report, Washington, D.C.

ODOT (June 2011). *Protocol for VISSIM simulation*. Oregon Department of Transportation

Oh, C., & Ritchie, S. (2003). Anonymous vehicle tracking for real-time traffic surveillance and performance on signalized arterials. *Transportation Research Record: Journal of the Transportation Research Board*, (1826), 37-44.

Oh, J. S., & Jayakrishnan, R. (2002). Emergence of private advanced traveler information system providers and their effect on traffic network performance. *Transportation Research Record: Journal of the Transportation Research Board*, 1783, 167-177.

Oh, C., & Ritchie, S. (2003). Anonymous vehicle tracking for real-time traffic surveillance and performance on signalized arterials. *Transportation Research Record: Journal of the Transportation Research Board*, 1826, 37-44.

OSADAP. *Open source application development portal*. Retrieved on July 13, 2015, from http://www.itsforge.net/78-news/140-tca-2-3-release.

Preisen, L., & Deeter, D. (2014). *Next generation traffic data and incident detection* (ENT-2014-2). ENTERPRISE Pooled Fund Study TPF-5(231).

Poushter, J. (2016). *Smartphone ownership and internet usage continues to climb in emerging economies*. Pew Research Center: Global Attitudes & Trends.

Qiu, Z., & Cheng, P. (2007). State of the art and practice: cellular probe technology applied in advanced traveler information system. In *86th Annual Meeting of the Transportation Research Board, Washington, DC (No. 0223).*

Robinson, E., Jacobs, T., *Frankle*, K., Serulle, N., Pack, M., & Rockville, M. D. (2012). *Deployment, use, and effect of real-time traveler information systems* (NCHRP Project 08-82). DOI 10.17226/22664

Roess, R. P., Prassas, E. S. & McShane W. R. (2011). Traffic engineering. Pearson.

SAE International (2004). *Message sets for advanced traveler information system (ATIS)*, SAE J2354.

SAE International (2016). *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.

Schofer, J., Koppelman, F., Webster, R., Berka, S., & Peng, T. (1996). *Field test of the effectiveness of ADVANCE dynamic route guidance on a suburban arterial street network*. The ADVANCE Project—Formal Evaluation of the Targeted Deployment, 2.

Shah, V.P., Wunderlich, K., & Larkin, J. (2001). Time management impacts of pre trip advanced traveler information systems: Findings from a Washington D.C. case study, *Transportation Research Record*, 1826, 7–15.

Shah, V. P, Wunderlich, K., Toppen, A., & Larkin, J. (2003). Potential of advanced traveler information system to reduce travel disutility: Assessment in Washington D.C. region. *Transportation Research Record*, 1774, 36–43.

Sharma, A., M. Zhao, N. Hawkins, P. Savolainen, & S. Knickerbocker. (2015). *Iowa 511 traveler information system user analysis*. Produced for Iowa Department of Transportation by Iowa State University, Ames, IA.

Shekhar, S., Fetterer, A., & Liu, D. R. (1996). Genesis: an approach to data dissemination in advanced traveler information systems. *Data Engineering*, 19(3), 40-47.

Sen, S., Pillai, R., Joshi, S., & Rathi, A. K. (2001). A mean-variance model for route guidance in advanced traveler information systems. *Transportation Science*, 35(1), 37-49.

Streff, F., & Wallace, R. R. (1993). Analysis of driver's information preferences and use in automobile travel: implications for advanced traveler information systems. In *Vehicle Navigation and Information Systems Conference*, Proceedings of the IEEE-IEE (pp. 414-418), IEEE.

Temple, L. L. (1989). *Perceptual and cognitive factors affecting driving ability of young and old drivers*. Department of Psychology, University of Nevada.

Tian, D., Yuan, Y., Qi, H., Lu, Y., Wang, Y., Xia, H., & He, A. (2015). A dynamic travel time estimation model based on connected vehicles. *Mathematical Problems in Engineering, 2015*.

Toledo, T., & Beinhaker, R., (2006). Evaluation of the potential benefits of advanced traveler information systems. *Journal of Intelligent Transportation Systems*, 10(4), 173-183.

Toppen, A., & Wunderlich, K. (2003). *Travel time data collection for measurement of advanced traveler information systems accuracy*. Mitretek Systems.

Tsuji, H., Takahashi, R., & Kawashima, H. (1985). A stochastic approach for estimating the effectiveness of a route guidance system and its related parameters, *Transportation Science*, 19, 333-351.

Turner, S. M., Eisele, W. L., Benz, R. J., & Holdener, D. J. (1998). *Travel time data collection handbook* (FHWA-PL-98-035). U.S. Department of Transportation, Washington, D.C.

USDOT (2008). *Vehicle-infrastructure integration (VII) initiative benefit-cost analysis, Version 2.3 (Draft)*. U.S. Department of Transportation, Washington, D.C. Retrieved from http://www.its.dot.gov/research_docs/pdf/7vii-benefit-cost-analysis.pdf.

Van Aerde, M., & Rekha, H. (1996). *TravTek evaluation modeling study* (FHWA-RD-95-090), McLean, VA: Federal Highway Administration.

Vasudevan, M., & O'Hara, J. (2015). *Measure estimation using connected vehicle data*, TRB HCQS Committee Mid-Year Meeting Connected Vehicle Data Workshop.

Vasudevan, M., Wunderlich, K., Toppen, A., & Larkin, J. (2004). Making the most of limited data in evaluating advanced traveler information systems by experimental resampling, *Transportation Research Record*, 1897, 1-8.

VISSIM (2014). *VISSIM version 6 manual*. PTV Planning Transport, Verkehr AG, Karlsruhe, Germany.

VDOT (May 2002). *Advanced traveler information systems (ATIS) white paper for the metropolitan Washington region*. Virginia Department of Transportation.

Wang, Y., Araghi, B. N., Malinovskiy, Y., Corey, J., & Cheng, T. (2014). *Error assessment for emerging traffic data collection devices* (Report No. WA-RD 810). Department of Transportation by University of Washington, Seattle, Washington.

Watkins, K. E., Wong, J., & Roell, M. (2013). *Technology scan of future traveler information systems and applications in Georgia* (FHWA-GA-13-11-25). Retrieved from http://www.dot.ga.gov/BuildSmart/research/Documents/11-25.pdf.

Wright, J., Garrett, J.K., Hill, C. J., Krueger, G.D., Evans, J. H., Andrews, S., Wilson, C. K., Rajbhandari, R. & Burkhard, B. (June 2014). *National Connected Vehicle Field*

Infrastructure Footprint Analysis (FHWA-JPO-14-125). US Department of Transportation, Washington, D.C.

Wunderlich, K. (1998). A simulation-based assessment of route guidance benefits under variable network congestion conditions, *Mathematical and Computer Modelling*, 27, 87-101.

Wunderlich, K., Hardy, M., Larkin, J., & Shah, V. (2001). *On-time reliability impacts of Advanced Traveler Information Services (ATIS): Washington, DC case study*. US Department of Transportation, Washington, D.C.

Xiao, Y. (2011). *Hybrid approaches to estimating freeway travel times using point traffic detector data*. Ph.D. Dissertation, Florida International University.

Yang, C. D., Fricker, J., & Kuczek, T. (1998). Designing advanced traveler information systems from a driver's perspective: Results of a driving simulation study. *Transportation Research Record: Journal of the Transportation Research Board*, 1621, 20-26.

Yang, Q., Koutsopoulos, H., & Ben-Akiva, M. (2000) Simulation laboratory for evaluating dynamic traffic management systems, *Transportation Research Record*, 1710, 122–130.

Young, S. E. (2008). *Monitoring and assessing arterial traffic performance*, University of Maryland. Retrieved on September 3, 2016, from http://itsmd.org/wp-content/uploads/Session-1B-Stan-Young-Arterial-Signal-Operations.pdf.

Young, S.E., Hamedi, M., Sharifi, E., Juster, R. M., Kaushik, K., & Eshragh, S. (2015), *I-95 corridor coalition vehicle probe project validation of arterial probe data*. Summary Report, Prepared for I-95 Corridor Coalition.

Yu, J., Park, S., & Chang, G. L. (2010). Advanced traveler information system for guiding route choice to ocean city, Maryland. *Transportation Research Record: Journal of the Transportation Research Board*, 2189, 56-67.

Zhan, X., Zheng, Y., Yi, X., & Ukkusuri, S. V. (2017). Citywide traffic volume estimation using trajectory data. *IEEE Transactions on Knowledge and Data Engineering*, 29(2), 272-285.

Zhang, L., & Levinson, D. (2008). Determinants of route choice and value of traveler information: a field experiment. *Transportation Research Record: Journal of the Transportation Research Board*, 2086, 81-92.

Zheng, J., & Liu, H. X. (2017). Estimating traffic volumes for signalized intersections using connected vehicle data. Transportation Research Part C: Emerging Technologies, 79, 347-362.

Zhou, X., & Taylor, J. (2014). DTALite: A queue-based mesoscopic traffic simulator for fast model evaluation and calibration. *Cogent Engineering*, 1(1), 961345.

Zou, Z., Li, M., & Bu, F. (2010). Link travel time estimation based on vehicle infrastructure integration probe data. In *ICCTP 2010: Integrated Transportation Systems: Green, Intelligent, Reliable* (pp. 2266-2276). *ASCE*.

VITA

MD SHAHADAT IQBAL

Born, Bangladesh

2006 - 2011	B.Sc., Civil Engineering Bangladesh University of Engineering and Technology, Dhaka, Bangladesh
2011 - 2013	M.Sc., Civil Engineering (Transportation) Bangladesh University of Engineering and Technology, Dhaka, Bangladesh
	Research Engineer Bangladesh University of Engineering and Technology, Dhaka, Bangladesh
2013 - Present	PhD. Student Department of Civil and Environmental Engineering Florida International University Miami, Florida
2013 - 2017	Graduate Research Assistant Department of Civil and Environmental Engineering Florida International University Miami, Florida
2013 – 2014	Vice President, ITE Student Chapter at FIU
2014 – 2015	President, ITE Student Chapter at FIU
2015 – 2016	Vice President, Bangladesh Student Organization at FIU
2015	Anne S. Brewer Scholarship, ITS Florida
2017	Dissertation Year Fellowship, University Graduate School, FIU

PUBLICATIONS AND PRESENTATIONS

Iqbal, M. S., Farzana, F. H., & Hadi, M. (2018). Identifying a timeline for future utilization of connected vehicle data to support traffic volumes estimation on urban streets. *Accepted for the 97th Annual Meeting of the Transportation Research Board*, Washington DC, USA

Iqbal, M. S., Kharaeian, S., & Hadi, M. (2018). A methodology to assess the quality of travel time estimation and incident detection based on connected vehicle data. *Accepted for the 97th Annual Meeting of the Transportation Research Board*, Washington DC, USA

Azizi, L. & Iqbal, M. S., & Hadi, M., (2018). Estimation of freeway platooning measures using surrogate measures based on connected vehicle data, *Accepted for the 97th Annual Meeting of the Transportation Research Board*, Washington DC, USA

Iqbal, M. S., Hadi, M., & Xiao, Y. (2017). Assessment of link level variation of connected vehicle market penetration. *Proceedings of the 96th Annual Meeting of the Transportation Research Board*, Washington DC, USA.

Iqbal, M. S., Hadi, M., & Xiao, Y. (2017). Accuracy and reliability of estimated travel time using basic safety message (bsm) data collected from connected vehicles. *Proceedings of the 96th Annual Meeting of the Transportation Research Board*, Washington DC, USA.

Hadi, M., Xiao, Y., Wang, T., Qom, S.F., Azizi, L., Jia, J., Massahi, A., & Iqbal, M. S. (2016). *Framework for multi-resolution analyses of advanced traffic management strategies*, Final Report No. BDV29-977-19, Prepared for FDOT.

Iqbal, M. S., Hadi, M., & Xiao, Y. (2015). Predicting arterial breakdown probability: a data mining approach. *Journal of Intelligent Transportation System*, 21(3), pp. 190-201.

Hadi, M., Xiao, Y., Wang, T., Iqbal, M. S., Massahi, A., Jia, J., Chen, X., & Fartash, H. (2015*). Decision support systems for transportation system management and operations (TSM&O).* Final Report, Prepared for FDOT.

Iqbal, M. S., Choudhury C.F., Wang P., & González M.C. (2014). Development of origin-destination matrices using mobile phone call data. *Transportation Research Part C: Emerging Technologies*, 40, pp.63-74.

Iqbal, M. S., Choudhury, C.F., Wang P, & González, M.C., (2014). Development of origin-destination matrices using mobile phone call data: a simulation based approach. *Proceedings of the 93rd Annual Meeting of the Transportation Research Board*, Washington DC, USA.