Short-Term Travel Time Prediction on Freeways

by

Wenfu Wang

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Civil Engineering

Waterloo, Ontario, Canada, 2019

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Dr. Tony Qiu Associate Professor, Dept. of Civil and Environmental Engineering, University of Alberta
Supervisor(s)	Dr. Bruce Hellinga Professor, Dept. of Civil and Environmental Engineering
Internal Member	Dr. Liping Fu Professor, Dept. of Civil and Environmental Engineering
Internal Member	Dr. Chris Bachmann Assistant Professor, Dept. of Civil and Environmental Engineering
Internal-external Member	Dr. Peter van Beek Professor, Cheriton School of Computer Science, University of Waterloo

AUTHOR'S DECLARATION

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public

Statement of Contributions

Chapter 3 of this thesis contains materials from a paper submitted for publication co-authored by myself, my colleague Trevor Luke Vanderwoerd, and my supervisor Dr. Bruce Hellinga. I developed and documented the majority of the methodology, Trevor proposed the usage of Van Aerde's model, and Dr. Hellinga assisted me in the writing of the paper.

Abstract

Short-term travel time prediction supports the implementation of proactive traffic management and control strategies to alleviate if not prevent congestion and enable rational route choices and traffic mode selections to enhance travel mobility and safety. Over the last decade, Bluetooth technology has been increasingly used in collecting travel time data due to the technology's advantages over conventional detection techniques in terms of direct travel time measurement, anonymous detection, and cost-effectiveness.

However, similar to many other Automatic Vehicle Identification (AVI) technologies, Bluetooth technology has some limitations in measuring travel time information including 1) Bluetooth technology cannot associate travel time measurements with different traffic streams or facilities, therefore, the facility-specific travel time information is not directly available from Bluetooth measurements; 2) Bluetooth travel time measurements are influenced by measurement lag, because the travel time associated with vehicles that have not reached the downstream Bluetooth detector location cannot be taken at the instant of analysis.

Freeway sections may include multiple distinct traffic stream (i.e., facilities) moving in the same direction of travel under a number of scenarios including: (1) a freeway section that contain both a High Occupancy Vehicle (HOV) or High Occupancy Toll (HOT) lane and several general purpose lanes (GPL); (2) a freeway section with a nearby parallel service roadway; (3) a freeway section in which there exist physically separated lanes (e.g. express versus collector lanes); or (4) a freeway section in which a fraction of the lanes are used by vehicles to access an off ramp. In this research, two different methods were proposed in estimating facility-specific travel times from Bluetooth measurements. Method 1 applies the Anderson-Darling test in matching the distribution of real-time Bluetooth travel time measurements with reference measurements. Method 2 first clusters the travel time measurements using the K-means algorithm, and then associates the clusters with facilities using traffic flow model. The performances of these two proposed methods have been evaluated against a Benchmark method using simulation data. A sensitivity analysis was also performed to understand the impacts of traffic conditions on the performance of different models. Based on the results, Method 2 is recommended when the physical barriers or law enforcement prevent drivers from freely switching between the underlying facilities; however, when the roadway functions as a self-correcting system allowing vehicles to freely switching between underlying facilities, the Benchmark method, which assumes one facility always operating faster than the other facility, is recommended for application.

The Bluetooth travel time measurement lag leads to delayed detection of traffic condition variations and travel time changes, especially during congestion and transition periods or when consecutive Bluetooth detectors are placed far apart. In order to alleviate the travel time measurement lag, this research proposed to use non-lagged Bluetooth measurements (e.g., the number of repetitive detections for each vehicle and the time a vehicle spent in the detection zone) for inferring traffic stream states in the vicinity of the Bluetooth detectors. Two model structures including the analytical model and the statistical model have been proposed to estimate the traffic conditions based on non-lagged Bluetooth measurements. The results showed that the proposed RUSBoost classification tree achieved over 94% overall accuracy in predicting traffic conditions as congested or uncongested. When modeling traffic conditions as three traffic states (i.e., the free-flow state, the transition state, and the congested state) using the RUSBoost classification tree, the overall accuracy was 67.2%; however, the accuracy in predicting the congested traffic state was improved from 84.7% of the two state model to 87.7%.

Because traffic state information enables the travel time prediction model to more timely detect the changes in traffic conditions, both the two-state model and the three-state model have been evaluated in developing travel time prediction models in this research. The Random Forest model was the main algorithm adopted in training travel time prediction models using both travel time measurements and inferred traffic states. Using historical Bluetooth data as inputs, the model results proved that the inclusion of traffic states information consistently lead to better travel time prediction results in terms of lower root mean square errors (improved by over 11%), lower 90th percentile absolute relative error ARE (improved by over 12%), and lower standard deviations of ARE (improved by over 15%) compared to other model structures without traffic states as inputs. In addition, the impact of traffic state inclusion on travel time prediction accuracy as a function of Bluetooth detector spacing was also examined using simulation data. The results showed that the segment length of 4~8 km is optimal in terms of the improvement from using traffic state information in travel time prediction models.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Professor Bruce Hellinga for his guidance, inputs, and support throughout my PhD studies. It has been my privilege to work with him and to learn from his professionalism and personality.

I would also like to thank my committee members, Dr. Tony Qiu, Dr. Peter van Beek, Dr. Chris Bachmann, and Dr. Liping Fu for their time and invaluable comments toward the improvement of this research.

I am greatly indebted to my colleagues and friends at the University of Waterloo, Amir Zarinbal, Reza Golshan, Reza Noroozi, Ehsan Bagheri, Yaxin Hu, Soroush Salek Moghaddam, Akram Nour, Babak Mehran, Tae J. Kwon, Fei Yang, Trevor Vanderwoerd, Ibrahim Almohanna, Jordan Hart-Bishop, Sahar Tolami-Hemmati, Alan Xaykongsa, Jingnan Liu, Sigong Zhang, Zhen Cai, Cheng Zhang, Yang Zhou, Donghui Lv, and Wei Yu, for their friendship and advice.

I own my deepest gratitude to my parents, Chuanjia Wang and Ying Zhou for their unconditional love and support.

Finally, I would like to extend my loving gratitude to my wife, Michelle Qu, for her love and support during my PhD studies. This research would not have been possible without her encouragement, patience, and sacrifices. Also my dear Maple, I am grateful for having you as my daughter, thank you!

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

Traffic congestion is now a major challenge faced by humanity across the world. In Canada, the estimated annual congestion cost in 2012 reached $4.6 billion in terms of social-economical prices such as time loss, vehicle wear, fuel consumption, air pollution, and greenhouse gas emissions (Force, 2012). However, alleviating the congestion status in Canada can be hardly achieved by constructing new roadway infrastructures: (1) the investments required are beyond the capability of governments; (2) the space needed for construction is constrained, especially for large cities. Applying Intelligent Transportation System (ITS) technologies to actively manage and control traffic operations can effectively enhance traffic mobility, efficiency, and safety, whereby cost-effectively address the traffic congestions in Canada.

Travel time is an important performance index for traffic operations, and the accurate estimations and predictions for the travel time are critical to the success of ITS management and control strategy implementations. ITS management and control strategies generally fall into three categories, namely offline, reactive, and proactive strategies. All three categories involve travel time information, but the travel time prediction is most relevant when implementing the proactive strategies, which relies on the predicted future travel time in determining the corresponding strategies to alleviate or eliminate traffic congestion.

This research focuses on developing methodologies for short-term travel time predictions applications on freeways. The main source of data adopted in this research is the Bluetooth detector data, yet the proposed methods are compatible with other sources of fixed/mobile sensor data when the traffic states and travel time can be estimated. The proposed methods account for the travel time measurement lag, which arises from the fact that the travel time measurements cannot be taken until vehicles have traversed the entire target roadway section and detected at consecutive Bluetooth detector locations. The impacts of lag are most prominent when traffic conditions are changing or when consecutive Bluetooth detectors are placed further apart. To alleviate the impact of lags, this research proposed methods for inferring traffic states in the vicinity of Bluetooth detectors based on non-lagged measurements. When predicting the near-future travel time, these inferred traffic states provide more timely information about the changes in traffic conditions and are used jointly with travel time measurements to improve the quality of travel time prediction results. The results of this research can be integrated into a traveler information system and support the development of ITS control and management strategies.

1.1 Travel Time Information

1.1.1 Travel Time Definition

In transportation engineering, travel time is defined as the amount of time travelers take to traverse a target roadway segment. The travel times for individual travelers are the result of complicated interactions among roadway, drivers, vehicles, and environmental conditions. Even for vehicles traveling on the same roadway segment during the same time period, individuals' travel times differ from one another (Tu, 2008). Therefore, instead of individual travel time, the mean travel time aggregated from a sample of vehicles on the target roadway should be used to characterize the traffic conditions and to support the development of traffic management and control strategies.

Two types of travel time, namely the Arrival-Based Travel Time (ABTT) and the Departure-Based Travel Time (DBTT), are defined in this research:

Arrival-Based Travel Time (ABTT): the travel time measured after vehicles have traversed the roadway segment of interest. This travel time is available for each interval with valid observations, and the mean ABTT can be computed as

$$\overline{ABTT}_k = \frac{1}{n_A}\sum_m t_m \qquad (1\text{-}1)$$

Where k is the time interval index, n_A is the total number of vehicles passing the downstream boundary within the interval k, m represents each individual vehicle, and t_m is the travel time experienced by vehicle m.

Departure-Based Travel Time (DBTT): the estimated or predicted travel times that vehicles will experience if they enter the road segment in the current time interval or in some specified future time interval. DBTT is the variable of interest for traveler information and traffic management systems, yet we cannot measure DBTT in real-time because of the time lag associated with the time required for the vehicles to traverse the roadway segment. For offline applications, the mean DBTT can be calculated as

$$\overline{DBTT}_k = \frac{1}{n_D}\sum_m t_m \qquad (1\text{-}2)$$

Where n_D is the total number of vehicles passing the upstream boundary within the target time interval; m and k are defined the same as above.

The following diagram explains the definition of ABTT and DBTT using some hypothetical vehicle trajectories.

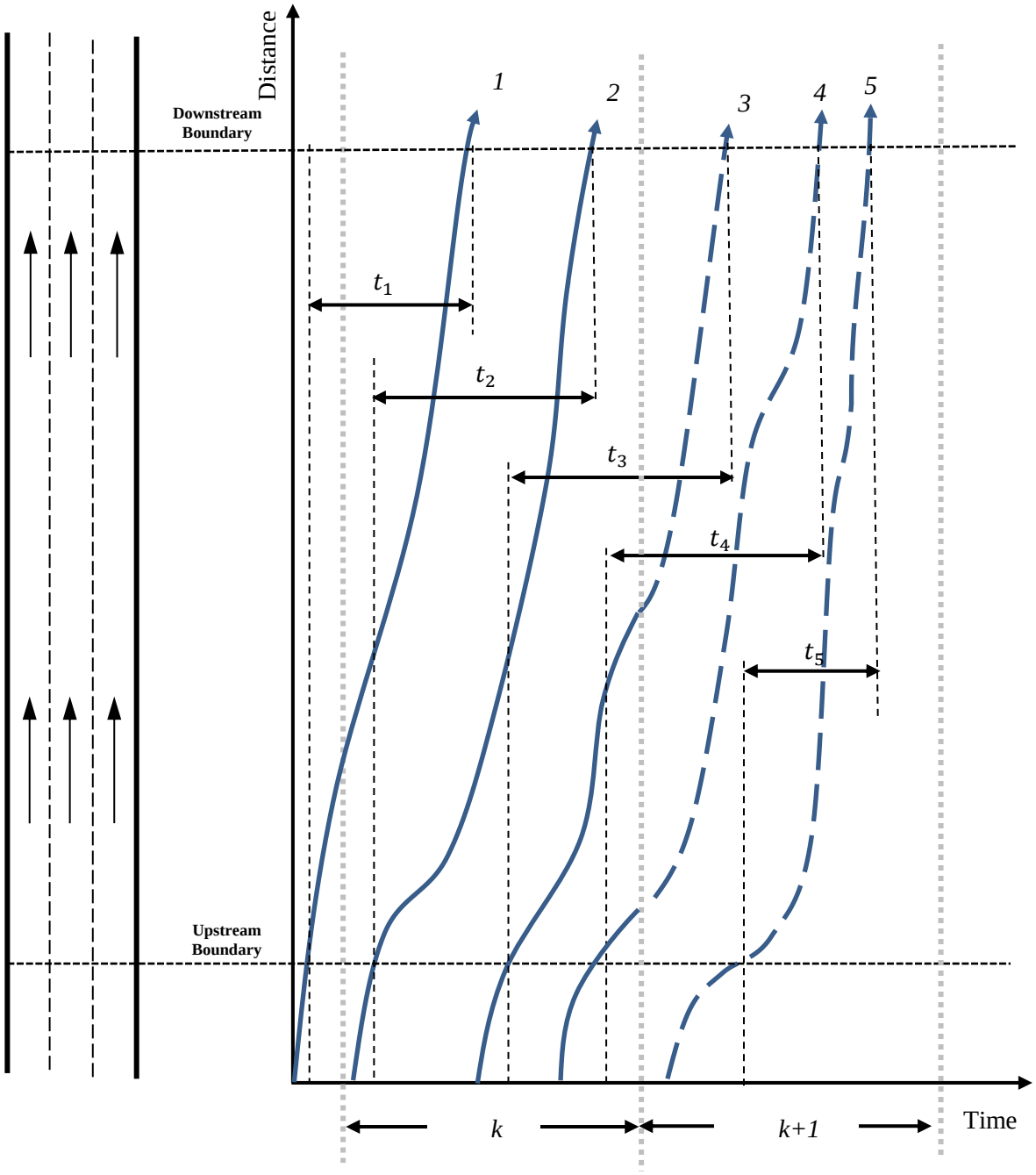


Figure 1-1 Time space diagram for vehicle trajectories

For the mean Arrival-Based Travel Time $\overline{ABTT}_k$ defined in equation (1-1), in the time step k of Figure 1-1, m includes vehicles 1 and 2, and n_A equals 2, $\overline{ABTT}_k = (t_1 + t_2)/2$. For the mean Departure-Based Travel Time $\overline{DBTT}_k$ defined in equation (1-2), in time step k, m incudes vehicles 2, 3, and 4, and n_D equals 3, $\overline{DBTT}_k = (t_2 + t_3 + t_4)/3$. It should be noted $\overline{DBTT}$ can only be calculated when all vehicles passing the upstream boundary in the time step of interest have reached the downstream boundary. In the

above diagram, at the end of time interval k, the travel times for vehicle 3 and vehicle 4 are not available, thus $\overline{DBTT_k}$ cannot be estimated, resulting in a lag in travel time estimation. The lag reduces the accuracy of travel time estimation and prediction results because the measurements used in calculating these results are not up-to-date.

1.1.2 Travel Time Estimation and Prediction

Based on definitions in existing studies (Tu, 2008; H. van Lint, 2004), travel time estimation is defined as the calculation of mean travel time for the roadway of interest from a sample of measured individual travel times in this research. Calculating the mean values for ABTT and DBTT defined in the previous section are both travel time estimations.

Travel time prediction is defined as projecting the travel time for future time intervals (i.e., when traffic conditions are unknown) on the roadway segments of interest. In general, travel time prediction studies can be categorized into three groups based on the prediction horizon:

- Online travel time prediction (i.e., real-time travel time prediction): the prediction horizon equals zero for this group of studies, that is, the prediction focuses on travel time for the vehicles entering the segment in the current time interval that have not yet reached the downstream boundary at the time of the study.
- Short-term travel time prediction: the prediction horizon is larger than zero but smaller than one hour (60 minutes). This current research also falls under this group.
- Long-term travel time prediction: the prediction horizon is larger than one hour, and this group of studies mainly relies on assumptions and aggregated data for understanding the historical trend, while the real-time and near-recent measurements are less used (Y. Hu, 2013).

1.2 Bluetooth Technology

Different types of detection systems based on technologies such as inductive loop detectors (ILD), automatic license plate recognition (i.e., ALPR), electronic toll tag readers, Bluetooth, and GPS based probe vehicles have been applied to collect traffic data including speed, density, occupancy, and travel time. In recent years, Bluetooth-based travel time information systems are gaining popularity in transportation projects due to the prevalence of Bluetooth enabled devices such as in-vehicle Bluetooth systems, Bluetooth enabled GPS devices, and Bluetooth enabled cell phones.

Bluetooth technology is a wireless communication standard enabling data exchange over a short distance. Bluetooth technology operates in the unlicensed industrial, scientific and medical (ISM) band at a range between 2.4 GHz and 2.485 GHz, and this range is available in most countries for free (Bluetooth SIG,

2014). To reduce interferences from other non-Bluetooth transceivers in the same range, Bluetooth technology introduced fast frequency hopping (FFH) mechanism. FFH following a pseudo-random hopping order determined by the master in the piconet in data transmission: Bluetooth transceivers hop among available channels at the rate of 1600 times per second with the associated time slot size of 0.625 milliseconds ($1\ \text{second}/1600 = 0.625$ millisecond) (T. Wang, 2002). Depending on the power class and the device implementation, Bluetooth transceivers can communicate with each other across distances (communication radius) up to 100 meters:

- Class 3 radio: communication radius up to 1 meter.
- Class 2 radio: communication radius up to 10 meters, most consumer electronic devices such as cell phones are equipped with this radio class.
- Class 1 radio: communication radius up to 100 meters (commonly used for industrial purposes, and Bluetooth detectors in traffic surveillance systems mainly use radio of this class).

1.2.1 Bluetooth Non-Lagged Measurements

Each Bluetooth enabled device has a unique 48-bit electronic identifier called the media access control address (MAC address). When vehicles carrying observable Bluetooth devices enter the detection range of a Bluetooth detector, the unique Bluetooth MAC addresses for devices in the vehicle can be detected and the corresponding detection time can be recorded in the database. Figure 1-2 illustrates Bluetooth detector measurements for some hypothetical vehicles traveling through the detection zone of a Bluetooth detector.

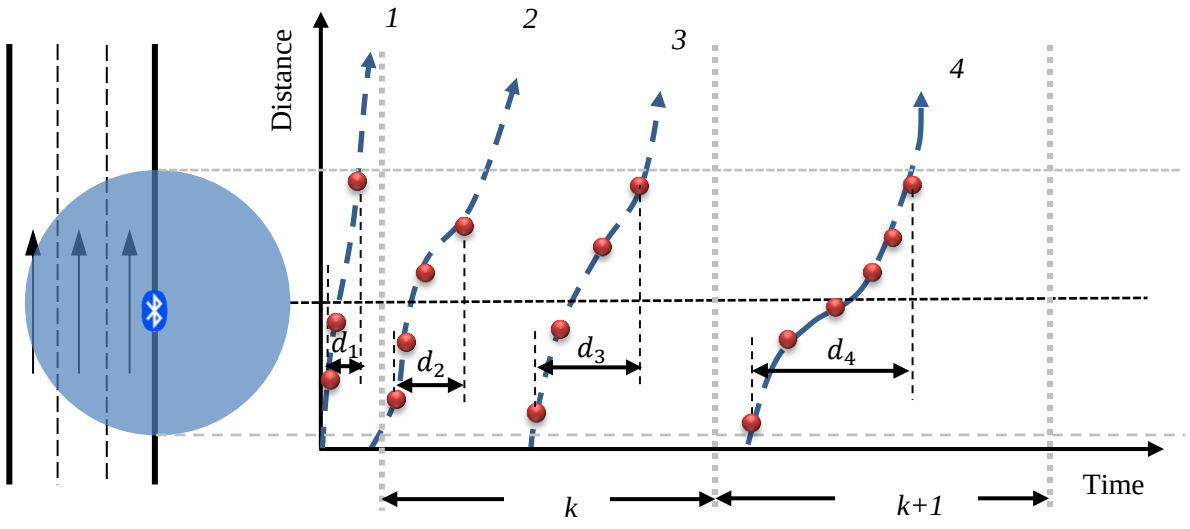


Figure 1-2 Hypothetical vehicle trajectories for defining the non-lagged Bluetooth measurements

This research refers to each observation of a vehicle by the Bluetooth detector as a "Hit" and these are represented in Figure 1-2 by solid round dots. The Bluetooth detector obtained four hits for vehicles 2 and 3 each in time interval k, and 6 hits for vehicle 4 in time interval $k+1$. Each hit measurement is associated with a time stamp and therefore this research also define "Dwell Time" representing the length of time that a vehicle was exposed to the Bluetooth detector (i.e., the time difference between the last hit and the first hit for a vehicle at one Bluetooth detector location). The Dwell Time associated with vehicles 1, 2, 3, and 4 in Figure 1-2 are noted as d_1, d_2, d_3, and d_4, respectively. Note that both hits and dwell time can be measured and reported at a fixed time interval and there is no lag waiting for vehicles to traverse a roadway segment before aggregating these measurements; in addition, the number of hits and the dwell time for a vehicle at a single detector depends on the associated vehicle speed.

1.2.2 Bluetooth Travel Time Measurement

If the same vehicle is detected at two consecutive Bluetooth detectors, the vehicle travel time can be calculated by comparing the time stamp difference at the downstream detector location with that at the upstream detector location. This process can be explained with the following diagram.

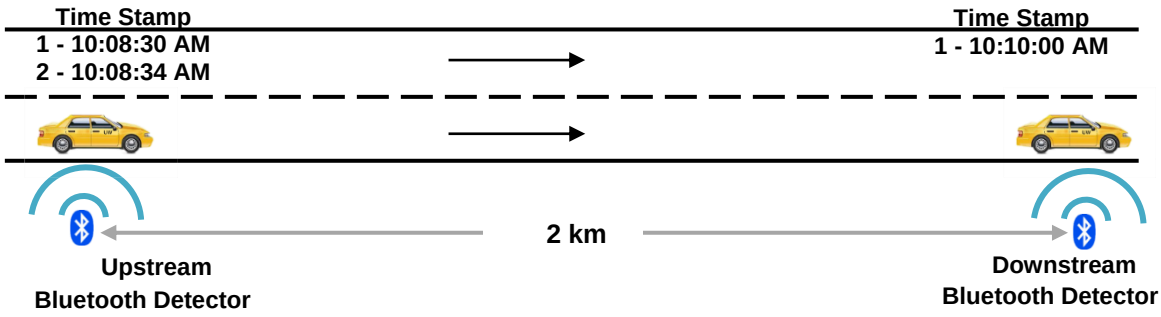


Figure 1-3 Bluetooth technology vehicle travel time detection process

In Figure 1-3, two hits are obtained for a device at the upstream detector at 10:08:30 AM and 10:08:34 AM and one hit is obtained later for the same device at the downstream detector at 10:10:00 AM. If we only consider the time stamp for the first hit of the unique MAC address at each Bluetooth detector, then the travel time for this vehicle can be estimated as 1.5 minutes. We refer to each travel time measurement as a "detection." The distance between these two consecutive detectors is known to be 2 km (measured at the time of Bluetooth detection system installations), and the average speed of the vehicle over the road segment can be computed as 80km/h. When the travel times for a group of Bluetooth equipped vehicles are detected, the prevailing traffic conditions can be estimated based on the average travel time of these vehicles.

Compared to other traffic surveillance technologies, Bluetooth technology is attractive for transportation projects for the following reasons:

- Bluetooth detection allows the direct measurement of both Arrival-Based Travel Time (ABTT) and Departure-Based Travel Time (DBTT). Technologies such as the inductive loop detector (ILD) based systems are not able to measure travel time directly.
- Unlike some other technologies, such as ANPR or toll-tag based system, Bluetooth detection offers anonymous detection. Although the MAC address is unique for each device, there is no database by which to associate a specific Bluetooth MAC address with an individual.
- A single Bluetooth detector can detect vehicles from lanes in both directions of travel and on different facility types.
- The installation of the Bluetooth detection system does not damage roadway pavements nor interrupt traffic flow (i.e., not necessarily involves lane closure during installation).
- Bluetooth detection offers cost-efficient traffic data collection.

As with all systems that measure vehicle travel times, Bluetooth travel time measurements are subject to a time-lag because the travel time cannot be measured until the vehicle has traversed the road segment and been observed at consecutive Bluetooth detectors. This research proposes methods for alleviating the impact of travel time measurement lag by incorporating non-lagged measurements in travel time predictions.

1.3 Problem Statement

Bluetooth technology is an attractive solution for measuring roadway travel times because Bluetooth detectors measure travel times directly, Bluetooth detection is anonymous, and the solution is low cost and easy to install. However, similar to many other Automatic Vehicle Identification (AVI) Systems such as automatic license plate recognition (i.e., ALPR) and electronic toll tag readers, Bluetooth detections are influenced by measurement lag which negatively impacts the travel time estimation and prediction accuracy. To improve travel time estimation and prediction, we propose to use non-lagged Bluetooth measurements such as Hits and Dwell Time, which are not subject to travel time measurement lag.

The non-lagged Bluetooth measurements can be used to estimate the traffic stream states in the vicinity of the Bluetooth detector. These states information can more timely reflect the changes in traffic conditions and can be used to improve the travel time prediction accuracy. Furthermore, this research proposes methods to associate Bluetooth travel time measurements with traffic streams traveling on different roadway facilities. In the end, traffic states inferred from non-lagged Bluetooth detector measurements,

Bluetooth travel time measurements, and historical travel time data were jointly used as inputs into the proposed travel time prediction models.

In summary, this research proposes to explore the usage of non-lag measurements and examine to what extent these measurements can improve travel time estimations and predictions. Detailed research questions and goals are provided in the next section.

1.4 Research Goals and Objectives

This thesis focuses on developing travel time prediction methods while addressing the facility association for Bluetooth travel time measurements. In order to effectively incorporate non-lag measurements to improve the quality of Bluetooth data based freeway travel time estimations and predictions, the following research questions were investigated, and the corresponding research objectives are defined:

Research Question 1: The non-lagged Bluetooth detector measurements such as the Dwell Time and Hits can be aggregated at a fixed time interval. Can these non-lagged measurements be used to estimate the speeds or states of the traffic stream in the vicinity of Bluetooth detectors?

Research Objective 1a: Investigate feature variables that can be extracted from Bluetooth measurements and are not influenced by travel time measurement lag.

The duration of a Bluetooth enabled vehicle's exposure to a Bluetooth detector (Dwell Time), and the number of repetitive detections for one Bluetooth MAC address (Hit) can be derived from Bluetooth detection database and are not suffering from travel time measurement lag. Apart from these two variables, other non-lag measurements such as the standard deviation of Dwell Time and Hit, the difference in values between consecutive Dwell Time and Hit measurements are also considered.

Research Objective 1b: Investigate the relationship between non-lag measurements (represented by Hit, Dwell Time, and other variables) and the traffic state conditions using both an analytical model and statistical models.

This thesis proposes to infer traffic states in the vicinity of each Bluetooth detectors from non-lag information; afterward, the inferred traffic states are incorporated into travel time prediction models to improve the prediction accuracy. Both analytical model and statistical models are developed to understand the relationship between non-lag measurements and traffic stream states.

Research Question 2: How can Bluetooth travel time measurements be attributed to different facility types (examples of different mixtures of facilities include expressways and collectors, general purpose lanes and high occupancy vehicle lanes.) which pass through the Bluetooth detection zone?

Research Objective 2: Develop methods to distinguish vehicle travel times from different facility types and estimate facility specific travel times in real-time from Bluetooth measurements.

Bluetooth-based systems estimate travel times by pairing time stamps associated with each vehicle at the downstream and the upstream detectors. However, the travel time measurements cannot be directly associated with the vehicles' physical lane of choice along the roadway segment. Since traffic conditions in two different facilities on the same roadway segment are not always the same, to correctly estimate travel time for each facility, it is necessary to develop methods and allocate detections into the correct facility group.

Research Question 3: To what extent can the information inferred from non-lagged measurements improve the accuracy of travel time predictions?

Research Objective 3a: Calibrate and evaluate the data-driven freeway travel time prediction models with near-recent inferred traffic states, near recent travel time measurements, and control conditions (e.g., location, facility, time of day, day of week) as inputs. Perform travel time predictions based on identifying historical data exhibiting similar traffic patterns as the target time periods.

The historical Bluetooth database provides knowledge of the changes in travel time under different prevailing conditions, which then allows us to predict the short-term travel time in real-time.

Research Objective 3b: Evaluate the performance of different model structures under different Bluetooth detector spacing and different traffic conditions to understand the impact of inferred traffic states. Make recommendations for the implementation of the proposed method.

The relationships among the above research objectives are summarized as follows:

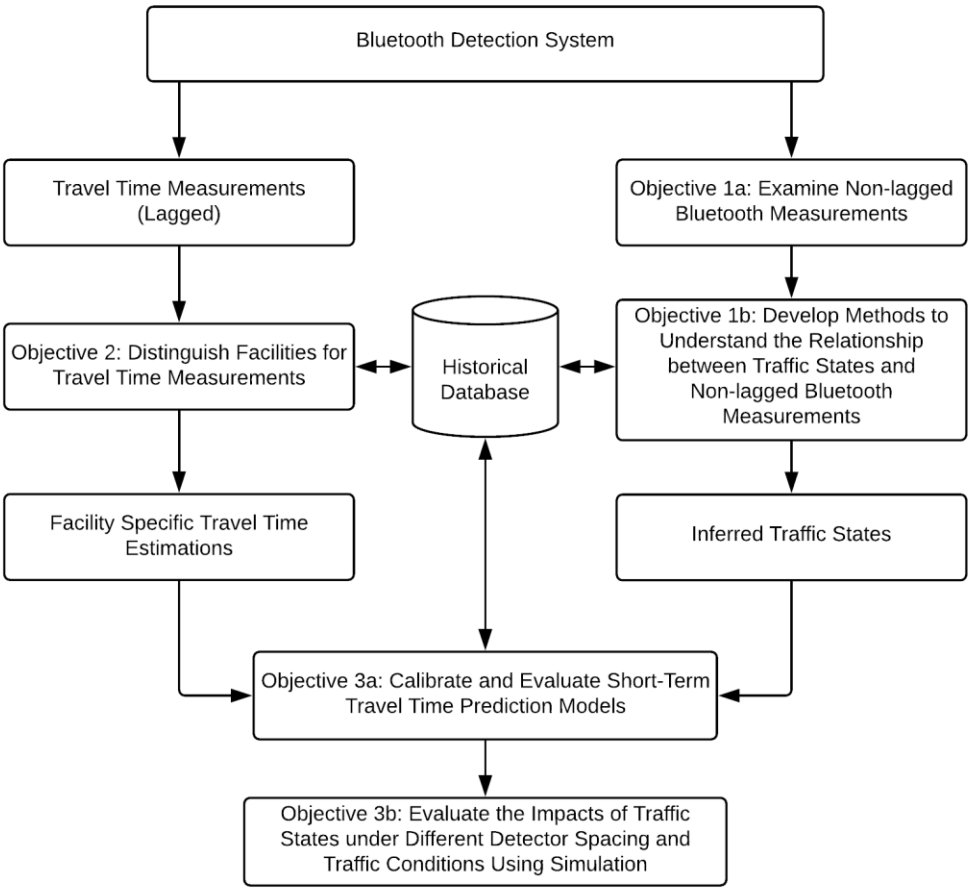


Figure 1-4 Schematic relationship for research objectives

1.5 Thesis Outline

This thesis is organized into six chapters.

Chapter 2 reviews the literature relevant to freeway travel time prediction with a special focus on methods applicable to Bluetooth-based travel time detection systems.

Chapter 3 discusses the methods for estimating facility specific travel times with the example of separating general purpose lanes and high occupancy vehicle lanes.

Chapter 4 presents the analytical and statistical model for estimating traffic states from non-lagged Bluetooth measurements; the inferred traffic states are later input into travel time prediction models.

Chapter 5 describes the calibration and validation of the developed travel time prediction model with different model structures using empirical data. Also, Chapter 5 presents the examination of the impacts of traffic state information on travel time prediction results under different detector spacing settings using simulation data.

Chapter 6 summarizes the conclusions and the thesis contributions, also recommend further work directions.

Chapter 2
Literature Review

2.1 Bluetooth Measurements

In Chapter 1, the theoretical background of Bluetooth technology and the procedures in using Bluetooth technology for travel time surveillance have been discussed. This chapter describes some implementations of Bluetooth based traveler information system and reviews some methodologies applied in existing research for predicting short-term travel time.

2.1.1 Bluetooth Technology Applications

Bluetooth technology has been applied in some transportation projects for different purposes; this section provides a summary of some real-world Bluetooth technology applications.

Started in 2008, the City of Houston conducted pilot tests with Bluetooth detectors to estimate travel time; attracted by the viability and cost-effectiveness of Bluetooth technology, now Bluetooth detectors have become indispensable parts of Houston TranStar system for travel time and speed data collection purposes (Puckett and Vickich, 2009).

In 2009, researchers from the University of Maryland used Bluetooth detectors to collect travel time data in order to validate the data quality of a private data vendor INRIX; in this study, some temporary Bluetooth detectors were placed along I-95 to serve as the source of ground truth travel time information (Haghani et al., 2009).

In 2010, the City of Calgary conducted a pilot project using Bluetooth technology to estimate travel time along a freeway segment and distributed information to the public using Variable Message Signs; now this pilot project has evolved into a Bluetooth traveler information system covering three roadway segments involving 26 Bluetooth sensors (Gray and Wong, 2013).

In 2012, researchers from Purdue University performed a series of case studies using Bluetooth technologies including assessment of winter freeway operations, traffic signal re-timing, and OD estimation under special events (Day et al., 2012).

In 2015, the York Region of Ontario, Canada performed a pilot study with Bluetooth-based travel time monitoring system (BTTMS) (Nause, 2015). Since then, the Bluetooth-based system has been greatly expanded on the Region's arterial network because of the initial favorable results from BTTMS. By the end of 2018, the York Region was operating over 300 Bluetooth detectors and planned to expand the system further to around 400 detectors in 2019. Using the outputs from BTTMS, the York Region was able to access the traffic conditions on the region's road network in real-time; the travel time information enables

the region to understand the key performance indicators in the region and support the decision-making for various projects.

2.1.2 Outlier Detection Algorithms

Data collected from traffic surveillance systems contain outliers, and this is also the case for Bluetooth based systems. The outliers associated with Bluetooth systems are from several sources: en route stop/detour of vehicles during their trip; detected Bluetooth devices are not from regular automobiles but are from buses, emergency vehicles, or nearby cyclists and pedestrians; detected Bluetooth devices are not from the main lanes of freeways but are from on/off ramps (Y. Hu, 2013). Researchers have developed different filtering algorithms to separate outliers from legitimate detections; some successfully implemented algorithms include:

- Rolling average algorithm (TranStar system)
- Deviation test algorithm (Chorus et al., 2006)
- Low-pass adaptive filtering algorithm (Dion and Rakha, 2006)
- Proactive adaptive filtering algorithm (Moghaddam, 2014)

The main motivation for adopting the filtering algorithm is to improve the quality of travel time information. However, the sophisticated filtering algorithm can introduce additional complexities and constraints for applications in real-time systems. Therefore, it is also favorable to implement simple yet efficient outlier detection algorithms for real-time systems (Moonam et al., 2019).

2.1.3 Non-Lagged Bluetooth Measurements

Bluetooth based travel time measurements are influenced by lags resulting from vehicles with unavailable travel time information: some vehicles were detected at the upstream Bluetooth detector on the target roadway segment but have not reached the consecutive downstream Bluetooth detector at the instant of analysis. Lags induce differences between estimated travel times and actual travel times for investigated time intervals, especially during congestion and transition periods (i.e., from free flow state to congested state or vice versa) or when consecutive Bluetooth detectors are placed far apart. Moreover, lags increase travel time prediction horizon and reduce travel time prediction accuracy.

To address the time lag, Hu (Y. Hu, 2013) recommends that information that does not suffer from lag effect can be used in travel time predictions; Moghaddam (Moghaddam, 2014) observed that the inclusion of sum of hits decreases travel time prediction errors on urban arterials. In addition, Schneider et al. (Schneider et al., 2012) found that Bluetooth hits might help detect incidents on freeways because the percentage of unique Bluetooth hits were found to decrease during incident periods, but these researchers didn't quantify the relationship between travel time and hit counts. Junghans and Leich (Junghans and Leich, 2016) assumed constant Bluetooth polling interval and fixed Bluetooth detector range in Bluetooth

detections, and estimated the speed of individual vehicles based on these assumptions. However, because the assumption introduced, the speed estimation resolution was low. To the best knowledge of the author, limited efforts have been devoted to using non-lagged information aside from these above researchers. Therefore, this research proposes to incorporate non-lagged measurements from each Bluetooth detectors in travel time predictions.

2.2 Short-Term Travel Time Prediction Models

With the development of Intelligent Transportation System (ITS) technologies and advancement in computation and data processing capabilities, techniques for short-term travel time prediction are now evolving from analytical approaches to data-driven approaches (Vlahogianni et al., 2014). This research categorizes the existing travel time prediction modeling into two broad categories, namely traffic flow theory based modeling and data-driven modeling, and some of the previous research efforts on both categories are reviewed.

2.2.1 Traffic Flow Theory Based Modeling

Traffic flow theory based modeling performs mathematical formulation of interactions between vehicles and roadway segments. Based on the levels of details considered in the modeling process, the adopted modeling techniques can be divided into macroscopic, microscopic, and mesoscopic modeling groups. These three groups of model-based approaches are reviewed in this research.

2.2.1.1 Macroscopic Modeling

Macroscopic traffic flow models explain traffic operations in the form of continuum flows and analyze the aggregated behavior of vehicles with variables such as vehicle flow rate, average density, and average speed (van Wageningen-Kessels et al., 2014).

LWR Model (Kinematic Wave Model)

One of the first macroscopic models was proposed by Lighthill, Whitham, and Richards (Lighthill and Whitham, 1955; Richards, 1956) and was named the LWR model. By assuming the conservation of vehicles, the LWR model describes traffic flow using the following differential equation:

$$\frac{\partial q(x,t)}{\partial x} + \frac{\partial \rho(x,t)}{\partial t} = 0 \qquad (2\text{-}1)$$

Where $q(x,t)$ is the traffic flow rate for instant t at location x, $\rho(x,t)$ is the density of the traffic stream for instant t at location x. By introducing Greenshields' speed density relation (Greenshields et al., 1935), flow rate the can be calculated from density as $q(x,t) = \rho(x,t) \cdot V[\rho(x,t)]$, with $V[\rho(x,t)]$ is the speed of the traffic stream when the density is $\rho(x,t)$. Following the conservation principles of vehicles,

researchers have estimated freeway traffic states (Nam and Drew, 1998) and predicted freeway travel times (J. Oh et al., 2002) with the LWR model.

The main challenges associated with the implementation of the LWR model are as follows:

- When modeling congested traffic states, the generated shockwave can cause rapid density change in the LWR model, which can lead to infinite acceleration/deceleration rates for vehicles in the flow.
- When modeling vehicle platoons, LWR model may generate higher speed values for the tail of platoon (than the head of the platoon) based on lower density at the tail, that is, model results may not always be consistent with real-world conditions.
- Significant side calculations are needed to keep track of the shockwaves in the LWR model (Daganzo, 1994).

Cell Transmission Model

Cell transmission model (CTM) applies discrete numerical equations to approximate the differential equation in LWR model: the whole modeling periods are discretized into unit time intervals based on the resolution needed; each roadway (i.e., link in CTM) is divided into small segments (i.e., cell in CTM) of the same length, and the length of each segment equals to the distance a vehicle can travel within a unit time under free-flow conditions; cells update their traffic states at the start of each unit time, so that the time-space evolution of traffic is captured in CTM model (Daganzo, 1994).

CTM can effectively describe different traffic phenomena such as shockwave, stop-and-go, and queuing; also, the computational efficiency of CTM is better compared to other higher order macroscopic traffic flow models, because discrete equations are applied to capture traffic dynamics. Considering these advantages, CTM models have been applied to estimate traffic states (Muñoz et al., 2003; Tampère and Immers, 2007), to evaluate signal control performance (Lo, 2001), and to predict travel time (Juri et al., 2007).

The main challenge associated with CTM model is that the model implementation requires accurate estimations for boundary conditions (such as OD matrix and flow rate at on/off ramps and on freeways). For Bluetooth based traffic surveillance systems, the task of estimating flow rates and other boundary conditions are very difficult to perform because the sampling rate is unknown and the Bluetooth detectors do not measure traffic volume.

Higher-Order Macroscopic Models

Both the LWR model and basic CTM model are low-order macroscopic models; some higher-order models have also been developed to describe the spatiotemporal evolution of traffic on roadways. Typical higher-order macroscopic models include an equation describing the speed changing dynamics (i.e., acceleration

or deceleration) toward the equilibrium speed; and the equilibrium speed is prescribed by the selected fundamental relation such as the Greenshields model (van Wageningen-Kessels et al., 2014). Some widely used higher-order models include Payne model (Payne, 1971) and Papageorgiou model (Messner and Papageorgiou, 1990; Y. Wang et al., 2006). Based on the developed equations and formulated relationships between different traffic flow variables, macroscopic simulator software has been built for predicting the traffic conditions in large networks, an example of such simulator is METANET (Papageorgiou et al. 2010). Many studies have adopted these higher-order models to perform/estimate traffic states and travel time (Nanthawichit et al., 2003; Xu et al., 2007).

2.2.1.2 Microscopic Modeling

Microscopic models describe individual vehicle behaviors in terms of car-following and lane-changing (i.e., lateral behaviors) maneuvers (van Wageningen-Kessels et al., 2014); and many microscopic models have now been integrated into simulation environments, some well-known microscopic simulation software package including VISSIM (PTV, 2011), PARAMICS (Duncan, 1997), CORSIM (Federal Highway Administration, 2006), etc.

Using microscopic models, researchers have performed different types of studies including traffic demand prediction (Chrobok et al., 2004), travel time reliability analysis (Ravi Sekhar et al., 2013), travel time estimation and prediction (M. E. Ben-Akiva et al., 2002; T. Hu, 2001).

Since microscopic models can model vehicle dynamics in details, these models can allow accurate representations and estimations for different traffic conditions. However, the following factors should be considered when using microscopic models:

- A large number of parameters are associated with each microscopic model, and the values for these parameters should be determined through a calibration process. If the model calibration has not been performed properly, the results from microscopic models will not be representative of real-world conditions and will not lead to reliable analysis results.
- Many microscopic models may encounter difficulties for real-time applications: not only the real-time OD matrix is hard to obtain, but these models tend to be computationally intensive, which may pose a problem when attempting to perform these simulations in real-time.

2.2.1.3 Mesoscopic Modeling

Mesoscopic models integrate the microscopic level vehicle interactions with macroscopic level supply simulations (S. Oh et al., 2018): the individual vehicles' route-choice behaviors on the road network are simulated as in microscopic models, whereas the traffic demands are considered in the aggregation stage with macroscopic level details. The mesoscopic modeling fills the gap between the macroscopic modeling and the microscopic modeling and has successful applications in predicting traffic demand (Balakrishna,

2006) and traffic state (Mahmassani et al., 2005). Two of the most commonly applied mesoscopic models includes the DynaMIT-R (M. Ben-Akiva et al., 1997) and the DYNASMART-X (Jayakrishnan et al., 1993) are reviewed in this research.

DynaMIT-R

DynaMIT has two versions: an off-line version for planning applications named DynaMIT-P(M. Ben-Akiva et al., 2001), and a simulation-based real-time traffic assignment model system for traffic state estimation and prediction applications named DynaMIT-R. Both historical data and real-time sensor data are input into the DynaMIT-R system to understand the interactions between traffic demand and supply. The outputs from the DynaMIT-R system supports the operation of advanced traveler information systems.

DYNASMART-X

Similar to DynaMIT, the DYNASMART also has two versions: an off-line version suitable for operation planning named DYNASMART-P; an online traffic estimation and prediction system named DYNASMART-X. The DYNASMART-X combines the input of real-time sensor data with a model-based representation of the traffic state on the network (Mahmassani et al., 2005). Historical data are also considered in DYNASMART-X to estimate the demand variations. The output from DYNASMART-X includes the estimations and predictions of traffic states on the network, based on these outputs; travelers can make proper route choice decisions.

The mesoscopic modeling considers a limited level of details in the modeling process; thus it is not suitable for understanding the interactions among drivers or between drivers and the roadway. When the detailed lane-changing and car-following behavior cannot be represented because of the network size is too large or when building the network is too resource consuming, the mesoscopic modeling is recommended for applications under these scenarios.

2.2.2 Data-Driven Modeling

2.2.2.1 Artificial Neural Network Modeling

Artificial Neural Network (ANN) modeling, whose development was inspired by the operation of biological nervous systems, includes a group of algorithms capable of deriving relationships from complicated non-linear data (Stergiou and Siganos, 1996). ANN-based models have been adopted in many travel time estimation and prediction studies, the selected neural network structures include state-space neural network (J. Van Lint, 2006; J. Van Lint et al., 2005), spectral basis neural network (Park et al., 1999; Rilett and Park, 2001), etc. In the literature, most ANN-based models have achieved similar or better prediction accuracies compared to other best performing data-driven models. However, ANN models have the following disadvantages:

- The numbers of neurons in hidden layers can have a significant impact on the performance of the model; however, there do not appear to be systematic methods by which to determine the optimal network structure.
- ANN models are typically developed for a specific application, and therefore the transferability of the model to other roadways or even other locations on the same roadway may be limited.
- It is difficult to interpret the meaning of the ANN and as such ANN models are often considered as a "black box."

2.2.2.2 K-Nearest Neighbor Modeling

K-Nearest Neighbor (k-NN) is a non-parametric instance-based learning (IBL) algorithm (some other common names for IBL include case-based learning and example-based learning) that can be used for classification and regression purposes. For k-NN classification, the algorithm assigns unclassified instances into the groups based on the similarity between unclassified instances and k nearest instances in the training dataset. For k-NN regression, the algorithm calculates the value of state vectors for objective instances based on the average values from their k nearest instances in the dataset (Peterson, 2009). The key parameters involved in the k-NN algorithm include: (1) the feature vector; (2) the distance metric; (3) the number of nearest neighbors (K in the above example); and (4) local estimation methods. Moghaddam (Moghaddam, 2014) explored the problem of travel time prediction on arterial roadways using Bluetooth data and found that the k-NN model performed the best.

In transportation engineering, the k-NN model has been used to perform traffic state prediction (Clark, 2003), demand prediction (Smith and Demetsky, 1996), travel time estimation (Robinson and Polak, 2005), travel time prediction (Chung and Kuwahara, 2004; You and Kim, 2000), and many other tasks. Some reported advantages of k-NN models include intuitive model formulation, easy implementation, and efficient computation for multivariate problems. However, a large historical/known instance dataset should be stored for a k-NN model to operate and the calculation time increases exponentially with the increase in the number of parameters (within feature vectors).

2.2.2.3 Regression Modeling

Regression models can derive the relationship between the dependent variable(s) and independent variables. In addition to parametric regression models (e.g., linear regression and non-linear least-squares regression), some non-parametric techniques such as kernel regression (Nadaraya, 1964) and locally weighted regression (Cleveland and Devlin, 1988) are also categorized as regression models.

The calibration of regression models can be performed efficiently with various software packages (e.g., SAS, SPSS, and MATLAB), and the obtained model formats allow straightforward interpretation of the relationship between input data and the output. In fact, many researchers have used regression models to

tackle different problems; in transportation engineering, regression models have been widely adopted to perform travel time predictions (Van Hinsbergen and Van Lint, 2008; Nikovski et al., 2005; Kwon and Petty, 2005; Rice and Van Zwet, 2004; Sun et al., 2003; Zhang and Rice, 2003).

The following aspects should be considered when performing regression analysis:

- Ensure the probability distribution of input variables are consistent with the underlying assumptions of the selected regression technique(s). Otherwise, the established regression models cannot lead to reliable conclusions.
- Correlation does not equal to causality; that is, the strong correlation between input variables and regression model outputs do not guarantee that the output is directly caused (influenced) by factors in the input. We should derive our conclusions based on regression results and our prior knowledge about the process under evaluation.

2.2.2.4 Time Series Modeling

Time series models infer the internal relationship (e.g., autocorrelation and cyclical/seasonal variations) within sequentially collected values from a variable of interest. Some commonly applied time series models include autoregressive integrated moving average (ARIMA) models, autoregressive moving average vector (ARMAV) models, and exponential smoothing models (NIST/SEMATECH, 2012).

Time series models have been applied to conduct traffic demand prediction (Lee and Fambro, 1999; Williams et al., 1998; Nihan and Holmesland, 1980) and travel time prediction on both freeways (Ishak and Al-Deek, 2002; D'Angelo et al., 1999) and urban arterials (Billings and Yang, 2006).

The main disadvantage associated with time series modeling is that the technique might not be able to respond effectively to dramatic changes within a short period, especially when some changes are significantly diverging from the historical trend. In travel time prediction studies, this disadvantage means that time series models might not be able to predict travel time under non-recurrent congestion conditions accurately.

2.2.2.5 Kalman Filter Modeling

Kalman filter recursively apply Bayesian theorem to obtain least square estimates for states from a noisy system (i.e., influenced by errors). Since its introduction in 1960 (Kalman, 1960), the Kalman filter has been applied to solve problems in different engineering realms. In transportation engineering the Kalman filter has achieved success in traffic state estimation (Y. Wang and Papageorgiou, 2005), traffic demand prediction (Guo et al., 2014; Xie et al., 2007; Okutani and Stephanedes, 1984), travel time prediction (Chien and Kuchipudi, 2003; Xia et al., 2011), etc.

When the state vector under evaluation is generated from a non-linear system or when the relationship between measurements and the state vector is non-linear, it is required to linearize the state vector first before applying Kalman filter to derive the optimal estimates. The Kalman filter algorithm with local linearization is referred to as the Extended Kalman Filter (EKF). Some general introduction for Kalman filter and EKF can be found in the paper from Welch and Bishop (Welch and Bishop, 1997).

Some challenges associated with the application of Kalman filters and EKF are summarized as follows:

- The covariance for measurement errors and process noises are hard to obtain, especially when the process under investigation is from time-varying systems. To deal with this difficulty, some researchers select constant values for the covariance by engineering judgment or by analyzing historical data (i.e., off-line calculation). For time-varying systems states (e.g., travel times for a freeway corridor), the use of constant covariance values may not effectively reflect the system dynamics and may deteriorate the performance of Kalman filters; under these situations, some adaptive variance estimation methods should be applied (Myers and Tapley, 1976).
- The application of EKF requires local linearization for the system under evaluation. However, the transformation changes the normal distribution for variables (in the system). Although EKF has been successfully applied for traffic estimation purposes in some previous research (Y. Wang et al., 2003), it is still necessary to examine the effect of changing variables' underlying distributions.

2.2.2.6 Random Forest Modeling

Random Forest (Breiman, 2001) is an ensemble machine learning algorithm that can perform classifications and regressions based on outputs from multiple decision trees. Each 'tree' in the 'forest' is developed using a subset of feature variables (i.e., random feature variable selection) from a randomly selected subset of training samples (i.e., bootstrap sampling); in the end, the collective knowledge from individual trees is used to generate results. Random Forest has been applied to perform long term bus travel time prediction (Mendes-Moreira et al., 2012), freeway travel time prediction (González-Brenes and Cortés, 2011), traffic flow prediction (Hamner, 2010), and harmonic average speed prediction (B. Hamner, 2010), etc. Compared to other data-driven algorithms, the performance of Random Forest (RF) models is superior if not similar in these existing studies. The advantages of RF models over many statistical models and data-driven algorithms are as follows:

- No prior assumptions regarding the distribution of input data.
- Efficient process a large number of variables with internal interactions in feature vectors.
- Inherent mechanism to reveal the importance of variables in the feature vector.
- Simultaneous process both numerical variables and categorical variables simultaneously.

- Reliable results for input data suffering from missing values or containing outliers.

The general procedures for performing classifications and regressions with RF algorithms are as follows (Liaw and Wiener, 2002):

1. Random sampling with replacement from the original dataset for n_{sub} times, generate n_{sub} sets of bootstrap samples.
2. Develop one unpruned classification or regression tree for each of these n_{sub} bootstrap sample sets. The tree should be developed with randomly selected p_{sub} feature variables from all the p feature variables, and the best split from among each p_{sub} feature variables should be adopted. In the end, n_{sub} trees are established.
3. Generate results from all n_{sub} trees: adopt the majority votes from all n_{sub} trees as the classification results, calculate the mean value from all n_{sub} trees as the regression results.

Some challenges associated with the application of RF algorithms are summarized as follows:

- The results can be hard to interpret for each tree in the forest.
- Regression results cannot predict beyond variable ranges in the training dataset and regression results trend to underestimate large values and overestimate small values because results are generated from average values from all 'trees'(Horning, 2013).
- There are no direct rules for determining the number of 'trees' in the forest, and the trial-and-error method for 'tree' size determination is not always reliable.

2.3 Summary

In this chapter, we reviewed case studies with Bluetooth technologies and some methodologies applicable to traffic time predictions. The literature review results reveal that data-driven models are gaining popularity in travel time prediction studies. Compared to the traffic flow model-based approaches, many data-driven approaches offer efficient computational solutions for real-time application. However, parametric data-driven approaches might have pre-requisites for input data (e.g., assuming unimodal distribution), which many transportation measurements fail to meet (e.g., travel times); non-parametric data-driven approaches are often capable of addressing different tasks as long as proper model calibration and training are performed. This research adopted data-driven models for short-term travel time prediction, and the proposed models are capable of handling travel time measurements, non-lagged measurements, and historical measurements simultaneously for travel time prediction purposes.

Chapter 3

Real-Time Estimation of Facility Specific Travel Times for Bluetooth Measurements on Freeways

Bluetooth detection systems are now increasingly used for measuring travel times on freeways due to the cost-effectiveness of Bluetooth technology, and the prevalence of Bluetooth enabled devices. However, Bluetooth detectors are not able to identify the specific location (lane) of a detected vehicle and therefore they are not able to directly provide separate travel time estimates for multiple traffic streams when these traffic streams are moving in the same direction. Freeway sections may include multiple distinct traffic stream under a number of scenarios including: (1) a freeway section that contain both a High Occupancy Vehicle (HOV) or High Occupancy Toll (HOT) lane and several general purpose lanes (GPL); (2) a freeway section with a nearby parallel service roadway; (3) a freeway section in which there exist physically separated lanes (e.g. express versus collector lanes); or (4) a freeway section in which a fraction of the lanes are used by vehicles to access an off ramp. For the purposes of traffic management and traveler information, it is often desirable to be able to estimate travel times separately for these different traffic streams. The remainder of this chapter explores how to provide these estimates. It does so in the context of HOV lanes versus general purpose lanes; however, the techniques and methods are also applicable for other situations in which multiple traffic streams exist.

In this chapter, three different approaches are proposed for estimating the facility-specific travel times based on Bluetooth measurements. Simulation data have been compiled to evaluate the performance of the proposed methods against a benchmark method under different traffic conditions. The proposed methods can extend the applicability of Bluetooth detection systems for freeway traffic conditions monitoring.

3.1 Background

Consider a freeway section consisting of multiple lanes in each direction and two facilities/traffic streams in each direction. For illustrative purposes, we can consider the specific case (Figure 3-1) in which one facility is an HOV lane, and the other facility is the general purpose lanes. The freeway section may contain on or off ramps. A Bluetooth detector is located at the upstream and downstream boundaries of the section.

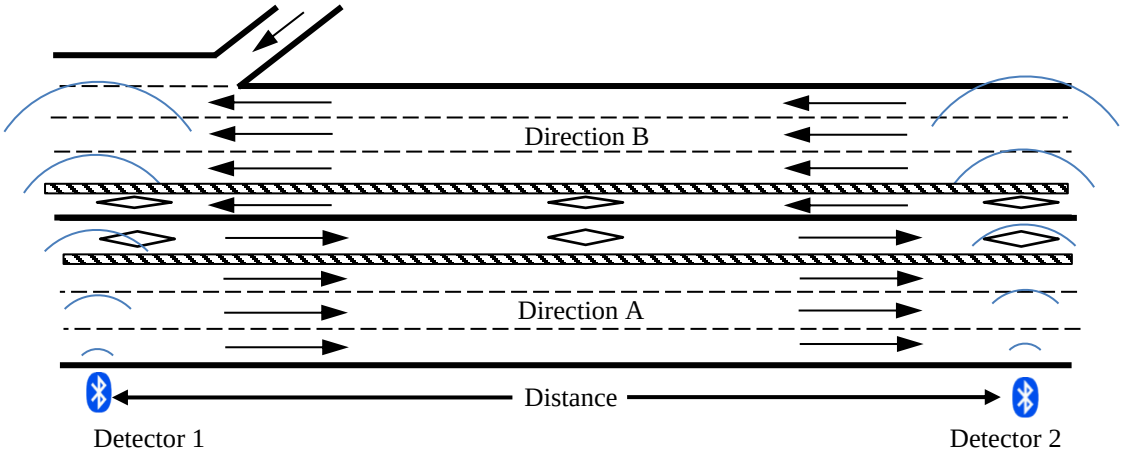


Figure 3-1 A freeway segment with both HOV and GPL facilities

When vehicles are detected at both detectors, then the vehicle travel time can be determined, and the direction of travel of the vehicle is known.

Suppose during the recent past *M* minute interval, the Bluetooth system detected *N* valid travel time measurements $t_1, t_2, t_3, \ldots t_N$ for direction A. These *N* travel times are samples of vehicles from the two facilities under surveillance. If the facility information is available for each of the *N* travel times, the mean travel times for the two facilities can be calculated as follows:

$$\overline{T}_{HOV} = \frac{\sum_{i=1}^{N_{HOV}} t_{i,HOV}}{N_{HOV}}, i = 1,..N_{HOV}$$

$$\overline{T}_{GPL} = \frac{\sum_{i=1}^{N_{GPL}} t_{i,GPL}}{N_{GPL}}, i = 1,..N_{GPL}$$

Where N_{Hov} is the number of travel times within *N* that was collected from the HOV, and N_{GPL} is the number of travel times from GPL, $N_{Hov} + N_{GPL} = N$, i is the index for measurements; $\overline{T}_{HOV}$ is the mean travel time for HOV, and $\overline{T}_{GPL}$ is the mean travel time for GPL.

However, in practice, the facility associated with each measured travel time is not known, and therefore, it is not possible to determine the mean travel time for each facility. Consequently, the objective of this chapter is to develop methods for real-time estimation of the mean travel times on each of the two facilities $\overline{T}_{HOV}$ and $\overline{T}_{GPL}$ using Bluetooth measurements.

There are essentially two problems that need to be overcome.

The first is to separate the *N* measured travel times into two subsets, N_1 and N_2 where each subset contains the travel times associated with one facility/traffic stream. We call this problem the *clustering*

problem. Once clustering is completed, the mean travel time for each subset (each cluster) can be readily computed.

The second problem is to determine which cluster mean is associated with which specific facility. In other words, we need to assign the clusters to specific facilities. We call this the *assignment* problem.

The next section describes previous work that has been done to address one or both of these problems.

3.2 Existing Methods

Although Bluetooth technology is gaining popularity in freeway travel time studies and many freeway segments consists of both HOV and general purpose lanes, very limited research efforts have been devoted to developing methods for estimating facility specific travel times on freeways using Bluetooth technology.

Salem et al. (Salem et al., 2015) used the support vector machine (SVM) classification algorithm to construct the two clusters but did so using Bluetooth data collected at a single location. They conducted a controlled experiment in which two probe vehicles carrying four Bluetooth devices (with known MAC ID) were traversing the antenna coverage range of a single Bluetooth detector repeatedly at a pre-determined speed. Because they intend to set up Bluetooth detectors for applications at single locations, they used the high gain antenna to improve the Bluetooth detection range and achieved a coverage range of around 1600 ft (490 meters), which is about 5 times the normal Bluetooth detector coverage range. In a single location context, they were not able to use travel times and instead used attributes such as the appearance time (defined as the time difference between the last and initial detection of a Bluetooth device), the mean of received signal strength indicator (RSSI), and the variance of RSSI. They attempted to classify the data into two clusters using SVM such that one was associated with vehicles having speeds > 40 mph (64 km/h) and the other with vehicles having speeds < 35 mph (56 km/h). They addressed the assignment problem by assuming observations in the low-speed cluster were associated with the general purpose lanes, and the higher-speed cluster was associated with the HOV lane. The authors declared that their approach achieved 79% classification accuracy, where 79% of the test data were correctly classified to the associated speed group. However, it is not always the case that HOV lanes operate faster than GP lanes especially with ongoing constructions or after incidents.

Haghani et al. (Haghani et al., 2014) observed separate travel time patterns within travel time measurements collected using Bluetooth detectors on segments with co-existence of HOV lanes and GP lanes. These authors highlighted the importance of distinguishing individual travel times between HOV lanes and general purpose lanes.

Zoto et al. (Zoto et al., 2012) proposed methods to estimate average speeds on different facilities (e.g., HOV lanes and regular lanes) using Bluetooth travel times. These authors addressed the clustering problem

by determining the number of distributions of individual vehicle speeds within the collected data using the Akaike information criterion. In this work, the speed of each vehicle is the average speed over a section, calculated by dividing the distance between a pair of Bluetooth sensors by the measured Bluetooth travel time. Then, they applied the expectation-maximization algorithm to approximate the mean speeds of vehicles and the percentage of vehicles traveling on different facilities. However, because large sample sizes are required for identifying distributions, the data aggregation interval duration in the study was long (30 to 60 minutes) making their method unsuitable for real-time estimation of traffic conditions. They addressed the assignment problem by assuming a vehicle in the HOV lane was always traveling at a faster speed than vehicles in the general purpose lanes. Finally, they tested their approach using synthesized data, but their evaluation dataset only contained conditions in which the HOV was traveling faster than the GPL.

In summary, it appears that there have been several proposed methods for addressing the clustering problem. Methods that require a large number of travel time observations are not suitable for real-time application. The majority of existing methods make the assumption that one of the facilities is always faster than the other (e.g., HOV is always faster than GPL). Though one facility frequently may be faster than the other, it is precisely when this is not true that real-time information is the most valuable. Therefore, using an assignment method that assumes one facility is always faster than the other seems counter-productive. Consequently, in the next sections we develop methods that meet the following objectives:

- The proposed models objectively determine the traffic conditions on each facility without assuming that one of the facilities always operating at higher speeds than the other facility;
- The proposed models are intended to provide real-time estimates and therefore should be able to provide information updates every 1-minute. The true mean travel time is associated with the traffic conditions over the past 5 minutes;
- The proposed models are evaluated using simulation data in order to permit the identification of ground truth travel times for each facility.

3.3 Proposed Methods

As described in the previous section, estimating travel times for each facility separately on the basis of Bluetooth detector measurements requires solving two sub-problems, namely the *clustering* problem and the *assignment* problem. In this section, we describe several approaches, each of which makes use of different methods to solve the clustering problem, or the assignment problem, or both.

We start by defining the Benchmark Method, which, as its name implies, is the method against which the other proposed methods are compared. We then describe two novel proposed methods. These are then subsequently evaluated and compared to the Benchmark Method.

In general, the inputs to the Benchmark Method and the proposed models can either be the Bluetooth measurements obtained over some past fixed time interval (in which case the number of Bluetooth measurements will vary) or can be the past N Bluetooth measurements (in which case the time period over which these N measurements were made will vary). Because the number of Bluetooth travel time measurements in each 5-minute interval varies due to randomness, traffic conditions, the level of Bluetooth market penetration rate, and the Bluetooth detection process (Moghaddam and Hellinga, 2013), this study decided to select N measurements as model inputs. The sensitivity of model performances to the value of N are discussed in section 3.4.3.

3.3.1 Benchmark Method

The Benchmark Method is developed based on the assumption that one of the facilities (HOV) always operates faster than the other facility (GPL). This is the same assumption made by some other researchers (Zoto et al., 2012) in estimating facility specific mean speeds using Bluetooth data. For the Benchmark Method, the recent past travel time measurements were first clustered into two groups, then the cluster with a faster speed (smaller travel time) was assigned to the HOV facility. The detailed operational steps for the Benchmark Method are summarized as follows:

1. Cluster a set of Bluetooth data B_k containing N most recent Bluetooth measurements into two groups according to speeds of each vehicle using the K-means algorithm (MacQueen, 1967). The K-means clustering algorithm determines the cluster for each travel time by solving the following optimization problem:

$$\arg\min_{c} \sum_{m=1}^{2} \sum_{b_i \in c_m} Distance(b_i, u_m)$$

   where $m = 1,2$ is the cluster index, c_m is the set of detections belonging to cluster m, and u_m is the centre for cluster m, b_i is the feature variable(s) associated with the i^{th} travel time measurement, and the speed of each measurement calculated from dividing the length between two Bluetooth detectors by the measured travel time is chosen as the feature variable, $Distance(\cdot)$ is a distance metric, with Euclidean distance selected for this study.

2. After two clusters of Bluetooth measurements are established, the space mean speed is calculated for each cluster. The faster cluster (i.e., the cluster associated with lower mean travel time) is assigned to the HOV facility, whereas the slower cluster is assigned to the GPL facility.

The value of N for Benchmark Method were determined through model calibration: the value of N achieving the best model result in the training dataset were chosen (described in more detail in Section 3.4.3).

3.3.2 Method 1 - Distribution Matching

The Distribution Matching Method proposes to compare the real-time Bluetooth measurements against reference mixture measurements of HOV and GPL facilities. Because the traffic conditions for the reference mixture measurements were known beforehand, after matching real-time measurements to the reference data, the mean travel times for the two facilities from the matched reference data are used as travel time estimates for the real-time measurements.

The reference data includes a full range of traffic conditions: different combinations of congested and uncongested traffic conditions on the two facilities. The speed of each Bluetooth measurement is calculated and referred to as measurement speed. When matching the distributions, it was decided to compare the distributions of measurement speeds as opposed to point estimations such as the mean of measurement speeds or the variation of measurement speeds, because distribution allows for more differentiation between different traffic condition mixtures. Since there has been no consensus in transportation community regarding the exact distributions for individual vehicle travel times (or associated speeds) measured from freeway segments under different traffic conditions and two facility types (two populations) are involved in the comparison samples of this study, non-parametric tests were adopted for comparing the empirical distribution of real-time measurement speeds against the distribution of reference measurement speeds. The Anderson-Darling (AD) test (Anderson and Darling, 1954) was adopted in this study for determining the closest matching distribution because the AD test can function without estimating parameters of the empirical distributions under comparison (Razali and Wah, 2011) and Arshad et al. (Arshad et al., 2003) showed the superior performance of AD test in comparing empirical distribution samples. The squared difference between the empirical distribution functions (EDF) of the two comparison samples is used in the AD test and the testing statistic for the null hypothesis $H_0: E(x) = R(x)$ is as follows:

$$A_{nm}^2 = \frac{nm}{n+m} \int_{-\infty}^{\infty} [E_n(x) - R_m(x)]^2 \omega\left(H_{n+m}(x)\right) dH_{n+m}(x)$$

Where $\omega(x)$ is a nonnegative weight function with the form of $\omega(x) = [u(1-u)]^{-1}$. In this study, $\omega(H_{n+m}(x))$ is inversely proportional to the estimator of the variance of $E_n(x) - R_m(x)$. $E_n(x)$ is the EDF estimated from the real-time measurements, $R_m(x)$ is the EDF estimated from the reference distribution mixture measurement.

The traffic conditions on HOV and GPL in the reference data were each categorized into 6 speed bins: 0-45 km/h, 45-60 km/h, 60-75 km/h, 75-90 km/h, 90-105 km/h, and over 105 km/h. The binning results in 36 categories in the reference measurements, referred to as $S_{i,j}$ where i is the GPL bin index ($i = 1, 2, \dots, 6$), j is the HOV bin index ($j = 1, 2, \dots, 6$). Data from a 5-minute interval are added to a bin if the harmonic mean speeds from the two facilities falls within the chosen bin's speed ranges. For example, if a 5-minute time interval of measurements had a harmonic mean speed of 63 km/h on the GPL and 78 km/h on the HOV, all speed measurements from that 5-minute interval would be placed into category $S_{3,4}$ in the reference data. It is important to recognize that a trade-off exists between the bin setting and the distribution matching accuracy: more bins and narrower bin ranges results in lower variability within the bin, but can cause the bins to be too similar to one another and increase the difficulty in determining the closest matching bin, and vice versa. The choice of 6 bins and the bin speed limits in this study were determined based on analyzing empirical loop data on freeways in Ontario, Canada. The operational steps of Method 1 are summarized as follows.

1. Compare the N most recent Bluetooth measurements B_k (k is the time interval index) to each $S_{i,j}$ for $i = 1,2, \dots, 6$ and $j = 1,2, \dots, 6$ (each of the 36 reference traffic conditions).
2. For each comparison, use the AD test to find the similarity of B_k to the data in the reference category $S_{i,j}$.
3. Once all the comparisons are completed, define i^* and j^* as the bins such that S_{i^*,j^*} is the reference data set with the best similarity according to the AD test.
4. The mean speeds of the GPL facility (s_{GPL}) and HOV facility (s_{HOV}) for the current time interval is estimated to be equal to the mean speed of GPL in the bin i^* and the mean speed of HOV in the bin j^*.

The optimal value of N for Method 1 was determined through model calibration (described in more detail in Section 3.4.3).

3.3.3 Method 2 - Clustering with Traffic Flow Model

Method 2 applied a two-step approach similar to the benchmark method in estimating the facility-specific mean travel time. The first step uses the same K-Means clustering algorithm as the Benchmark Method (MacQueen, 1967). However, the assignment method is novel and does not make the naïve assumption that the HOV facility is always faster than the GPL facility. Instead, the assignment makes use of the physical characteristics of the two facilities (i.e., number of lanes) and the characteristics of the macroscopic speed, flow, density relationship.

We describe the method using an illustrative example. First, we make use of Van Aerde's macroscopic traffic flow model (Van Aerde, 1995) to describe the relationship between speed, flow, and density for each facility. Figure 3-2 shows two such relationships for speed as a function of flow rate for the HOV and GPL. The flow rate is for the section (not per lane), and if we assume there are 3 GP lanes and 1 HOV lane, then the capacity flow for the GPL is approximately three times that of the HOV. For this example, assume the speeds associated with the two clusters are 100 km/h and 60 km/h, and therefore there are two possible outcomes; namely (A) HOV speed = 100 km/h and GPL speed = 60 km/h; or (B) HOV speed = 60 km/h and GPL speed = 100 km/h. These two outcomes are labeled by Option A and Option B in Figure 3-2.

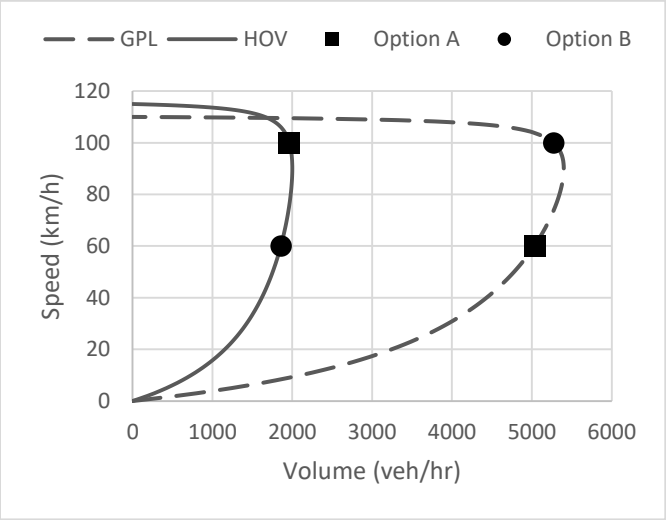


Figure 3-2 Van Aerde model with the two assignment options

For each of these options in Figure 3-2, the volumes predicted by Van Aerde's model were compared to the number of vehicles in the associated clusters (i.e., Bluetooth measurements) based on the root mean square deviation of the two facilities. Van Aerde's model provides vehicles per hour, while the recent past observations give the proportion of measurements representing the most recent 5-minute. To ensure that the two predicted volumes are compared on the same magnitude, the data were scaled so that the minimum is 1 for both the Van Aerde predicted volume and the number of observations. The formula for calculating the root mean square deviation are as follows:

$$RMSD_A = \sqrt{\frac{\left(\frac{v_g(\overline{s_1})}{\min(v_g(\overline{s_1}), v_h(\overline{s_2}))} - \frac{n_1}{\min(n_1, n_2)}\right)^2 + \left(\frac{v_h(\overline{s_2})}{\min(v_g(\overline{s_1}), v_h(\overline{s_2}))} - \frac{n_2}{\min(n_1, n_2)}\right)^2}{2}}$$

$$RMSD_B = \sqrt{\frac{\left(\frac{v_h(\overline{s_1})}{\min(v_h(\overline{s_1}), v_g(\overline{s_2}))} - \frac{n_1}{\min(n_1, n_2)}\right)^2 + \left(\frac{v_g(\overline{s_2})}{\min(v_h(\overline{s_1}), v_g(\overline{s_2}))} - \frac{n_2}{\min(n_1, n_2)}\right)^2}{2}}$$

where $\overline{s_1}$ and $\overline{s_2}$ represent the harmonic mean speeds of the two clusters as determined by the K-means clustering, v_g and v_h are the volumes calculated using Van Aerde's model for the GPL and HOV facilities, respectively, and n_1 and n_2 are the number of recent past Bluetooth measurements in each of the two clusters.

The assignment option (A or B) associated with the smaller $RMSD$ value is chosen. The parameters for Van Aerde models representing the two facilities were calibrated based on empirical loop detector data collected from Highway 403 in Ontario. In addition, the optimal value of N for Method 2 was determined through model calibration (described in more detail in Section 3.4.3).

3.4 Evaluation of the Proposed Methods

3.4.1 Simulation Data Collection

In this study, simulation data were generated for compiling reference distribution data and evaluating the performance of different proposed methods. Simulation-based studies allow users to access the ground truth information under different traffic conditions for different facility types. The microscopic traffic simulation software VISSIM version 7 (Vissim, 7) and Bluetooth detection simulation tool *BlueSynthesizer* (Masouleh and Hellinga, 2018) were used.

A hypothetical test network consisting of one direction of a 4-lane freeway segment containing both HOV (one lane) and GPL (three lanes) facilities with two on-ramps and two off-ramps was defined (Figure 3-3).

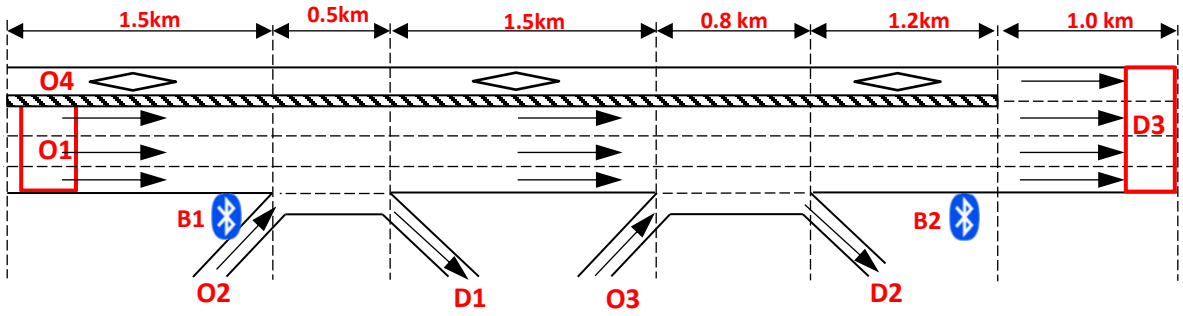


Figure 3-3 Simulation roadway geometry

The above network was modeled within VISSIM. The VISSIM model parameters were calibrated based on vehicle detector station (VDS) loop data containing both HOV lane and GPL lanes compiled from Highway 403 in Ontario for 15 weekdays in August 2015. The raw VDS data were collected from the Ministry of Transportation of Ontario (MTO) with speed, occupancy, and volume information of each lane for every 20 seconds. These raw data were aggregated across lanes at each VDS over 5-minute aggregation interval. The Van Aerde macroscopic speed-flow-density relationships were calibrated with these empirical data following the procedures developed by Rakha and Arefeh (Rakha and Arafeh, 2010). The calibration results in estimates of the free-flow speed (v_f), the speed at capacity (v_c), the capacity flow rate (q_c), and the jam density (k_j) for each facility as summarized in Table 3-1.

Table 3-1 Van Aerde model calibration results

Model Parameter	GPL	HOV
free flow speed, s_f (km/h)	110	116
speed at capacity, s_c (km/h)	90	90
jam density, k_j (veh/km/lane)	100	100
capacity, q_c (veh/h/ lane)	1820	1855

The VISSIM parameter calibration tool (http://vissim.waterlootraffic.com/) was used to calibrate the Wiedemann 99 car-following model (freeway model) parameters to produce freeway traffic streams with the characteristics noted in Table 3-1. After calibration, the chosen VISSIM parameters are summarized in Table 3-2.

Table 3-2 Calibrated Wiedemann 99 car-following model parameters

VISSIM Parameters	Value and Unit
CC0 (Standstill Distance)	4.81 meter
CC1 (Headway Time)	1.16 second
CC2 ("Following" Variation)*	4.00 meter
CC3 (Threshold for Entering "Following")*	-8.00
CC4 (Negative "Following" Threshold)	-0.51
CC5 (Positive "Following" Threshold)	0.41
CC6 (Speed Dependency of Oscillation)*	11.44
CC7 (Oscillation Acceleration)	0.38 m/s^2
CC8 (Standstill Acceleration)*	3.50 m/s^2
CC9 (Acceleration with 80 km/h)	1.79 m/s^2

*default value from VISSIM was kept

Traffic demands at each of the four origins to the network (O1, O2, O3, and O4 in Figure 3-3) were varied as a function of the simulation time. In addition, accidents were simulated in VISSIM using reduced speed areas on the two facilities. In the end, a full range of traffic conditions (including both congested and uncongested conditions) have been created on both facilities for this study. A sample vehicular demand and OD setting for one simulation run are provided in Table 3-3 and Table 3-4 respectively.

Table 3-3 VISSIM simulation input setting for one simulation run

Time Interval (seconds)	O1 Inputs* (veh/h)	O1 V/C Ratio	O2 inputs* (veh/h)	O2 V/C Ratio	O3 Inputs* (veh/h)	O3 V/C Ratio	O4 Inputs* (veh/h)	O4 V/C Ratio
0-900**	2750	0.50	767	0.35	657	0.30	821	0.38
901-1800	4400	0.80	1314	0.60	1095	0.50	1314	0.60
1801-2700	4675	0.85	1862	0.85	1643	0.75	1396	0.64
2701-3600	4675	0.85	1533	0.70	1752	0.80	1396	0.64
3601-4500	2200	0.40	1095	0.50	438	0.20	657	0.30
4501-5400	4125	0.75	1533	0.70	1424	0.65	1232	0.56
5401-6300	2750	0.50	876	0.40	438	0.20	821	0.38
6300-7200	3850	0.70	657	0.30	1533	0.70	1150	0.53
7200-8100	2200	0.40	438	0.20	438	0.20	657	0.30

*The percentages of HGV (heavy duty truck) are set as 10% for O1, O2, and O3; 2% for O4

**The initial 900-second is the warm-up period for the simulation and was not included in the analysis.

Table 3-4 OD matrix for VISSIM input (Fraction of demand going to each destination)

Origin \ Destination	D1	D2	D3
O1	0.1	0.25	0.65
O2	0.1	0.2	0.7
O3	-	0.15	0.85
O4	-	-	1

VISSIM was configured to output the trajectories and the associated travel times of all simulated vehicles. The vehicles trajectories were input into *BlueSynthesizer* for simulating Bluetooth detections, and two Bluetooth detectors covering a segment length of 4.0 km were modeled at the two ends of the study segment (represented with B1 and B2 in Figure 3-3). The Bluetooth level of market penetration rate was assumed to be 10%. Therefore, each vehicle simulated in VISSIM has 0.1 possibilities of carrying a detectable Bluetooth device. Vehicles determined to be carrying detectable Bluetooth device were assumed to carry only one device. In the end, the matched Bluetooth sampling rate was found to be 6.1%, and this value is consistent with Bluetooth sampling rates reported by other researchers (Erkan and Hastemoglu, 2016; Sharifi et al., 2011).

In the next sections, the generated simulation data are used for calibrating model parameters and testing the performance of the proposed methods against the Benchmark method. All the simulation data are used assuming they were collected in real-time, and the evaluation results can realistically represent the proposed model performances in real-world applications.

3.4.2 Measure of Performance

The performance measurements in this study include the absolute relative error (ARE) of the estimated travel times for each facility type and the root mean square error (RMSE) for travel times; the ARE reveals the error as a proportion of the true travel time, and the RMSE reflects the absolute magnitude of the error. The definitions for these two measurements are provided as follows:

$$ARE_{d,m,k} = 100 * \frac{\left|T_{m,k} - \hat{T}_{d,m,k}\right|}{T_{m,k}}$$

$$RMSE_d = \sqrt{\frac{1}{2n}\sum_{k=1}^{n}\left[\left(T_{h,k} - \hat{T}_{d,h,k}\right)^2 + \left(T_{g,k} - \hat{T}_{d,g,k}\right)^2\right]}$$

Where k is interval index ($k = 1,2,\ldots,n$), $T_{m,k}$ is the ground truth travel time for facility m in interval k calculated from all vehicles simulated in VISSIM, d stands for the Method Benchmark, Method 1, or Method 2, $\hat{T}_{d,m,k}$ is the estimated facility specific mean travel time for the time interval k using method d on facility m, $ARE_{d,m,k}$ is the absolute relative error using method d on facility m in interval k; n is the total number of intervals under comparison, $T_{h,k}$ is the ground truth travel time for HOV in interval k, $\hat{T}_{d,h,k}$ is the estimated travel time for the time interval k using method d on HOV, and $T_{g,k}$ is the ground truth travel time for GPL in interval k, $\hat{T}_{d,g,k}$ is the estimated travel time on GPL for the time interval k using method d, m represents HOV or GPL facility, $RMSE_d$ is the root mean square error associated with method d.

Note that given the definition of the true travel time, the measures of error include the following sources of errors:

1. Sampling error - $\hat{T}_{d,m,k}$ is computed on the basis of the travel times obtained from a sample of Bluetooth enabled vehicles in the traffic stream
2. Measurement error – the travel times obtained from the simulated Bluetooth equipped vehicles (i.e., by *BlueSynthesizer*) contain measurement error due to the nature of the Bluetooth detection process.
3. Model estimation error – each model has a different method of computing $\hat{T}_{d,m,k}$ and therefore has associated different errors.

3.4.3 Model Parameter Calibration and Performance Analysis

A total of 12 hours simulation data have been compiled for calibrating the parameters of each model and for generating the reference dataset in Method 1. Using $RMSE$ as the measure of performance, the optimal value of N for each of the three models was determined. The model performances associated with different values of Ns together with the chosen N for each model are summarized in the following diagram and table:

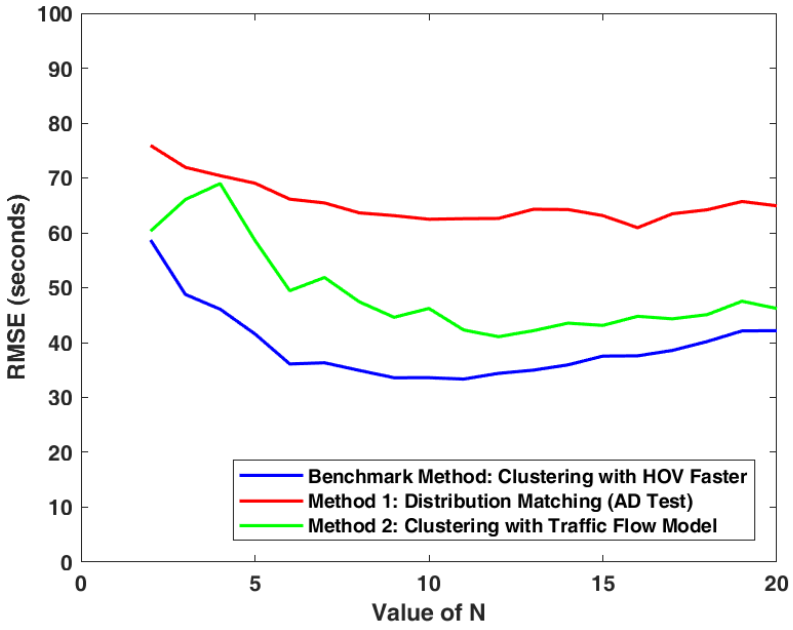


Figure 3-4 Sensitivity of model performance as a function of parameter *N*

Table 3-5 Selected optimal model parameter *N* Values

Model	*N*
Benchmark Method: Clustering with HOV Operating Faster	11
Method 1: Distribution Matching	16
Method 2: Clustering with Traffic Flow Model	12

Another 12 hours of simulation data were compiled as the testing dataset. Using the calibrated model parameters in Table 3-5 the cumulative distribution function (CDF) for ARE are compiled on the testing dataset for different methods to evaluate the model performances under different conditions: CDF for intervals when HOV lane is operating faster than GPL (i.e., vehicles using HOV lane have shorter travel

times than those using GPL facility); CDF for intervals when GPL lanes are operating faster than HOV. Figure 3-5 provides ARE results associated with intervals when HOV is faster than GPL.

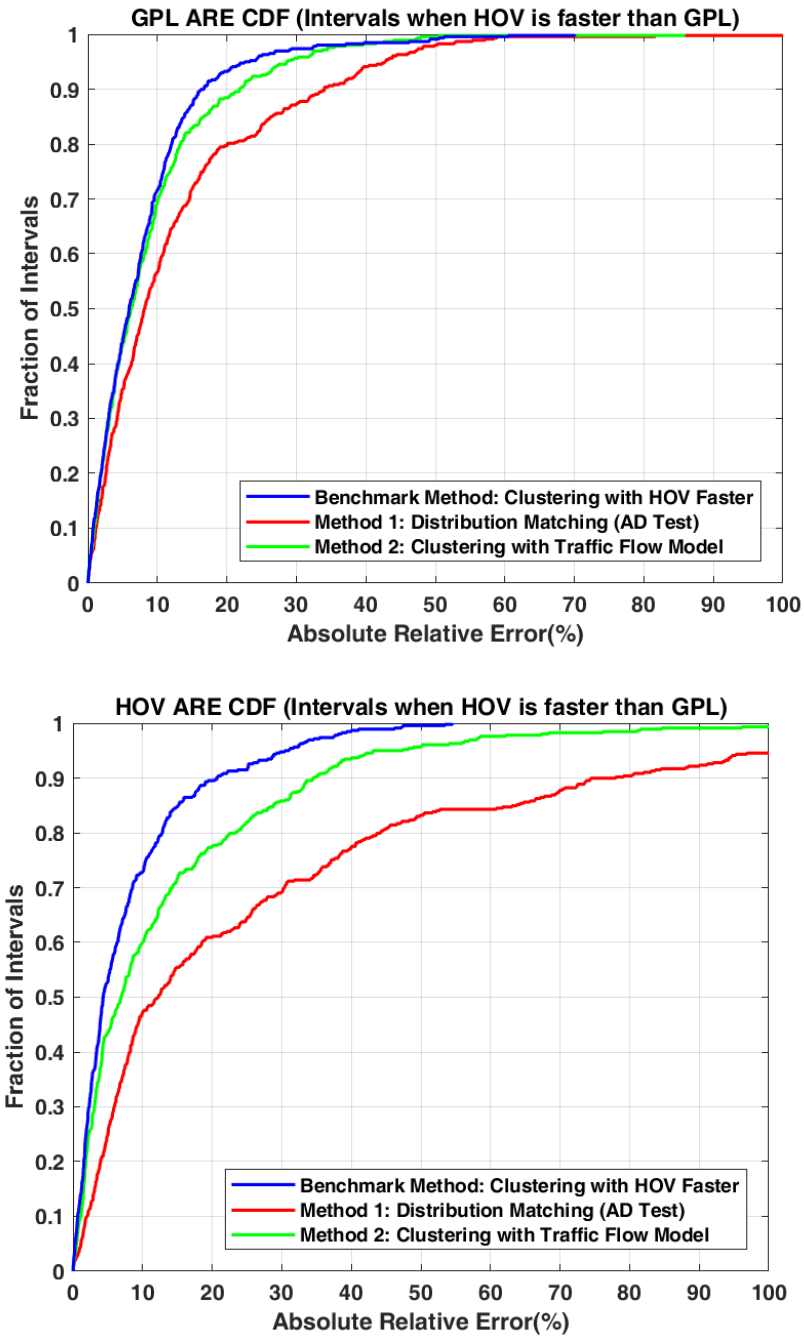


Figure 3-5 CDF of ARE for intervals when HOV is faster than GPL

Figure 3-5 shows that for intervals when HOV is faster than GPL, the Benchmark Method performs the best, especially when estimating HOV travel times. Because the Benchmark Method assumes that the HOV facility is always faster than GPL, when this assumption is the case, the algorithm naturally performs well.

At the same time, the proposed Method 2 (Clustering with traffic flow model) achieved comparable performance as the Benchmark method. In contrast, Method 1 did not perform as well as the Benchmark Method.

Figure 3-6 shows ARE results for intervals when HOV is slower than GPL.

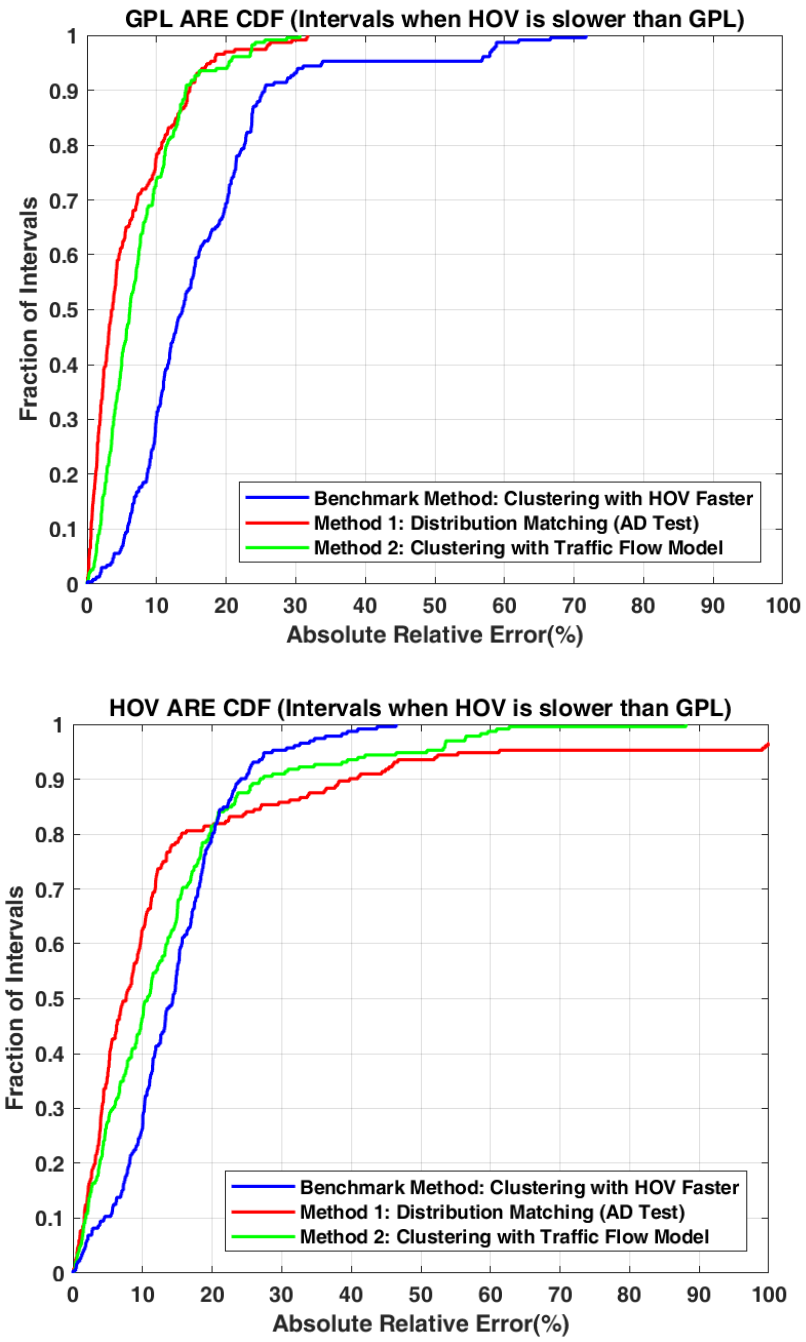


Figure 3-6 CDF of ARE for intervals when HOV is slower than GPL

Figure 3-6 indicates that the two proposed Methods 1 and 2 perform consistently better than the Benchmark Method for both the HOV and GPL facilities. In addition, the proposed Method 2 consistently outperformed the Benchmark by a significant margin: Method 2 achieved less than 10% error around 80% of the time for GPL, whereas the Benchmark method makes less than 10% error only about 30% of the time; Method 1 achieved less than 10% error around 65% of the time for HOV, whereas the Benchmark method makes less than 10% error only about 30% of the time.

The $RMSE$ results for different model structures under different traffic conditions on the testing dataset are summarized in the following table.

Table 3-6 $RMSE$ results for different model structures under different traffic conditions

Model	All intervals (seconds)	Intervals when HOV is slower than GPL (seconds)	Intervals when HOV is faster than GPL (seconds)
Benchmark Method	31.8	29.0	33.1
Method 1: Distribution Matching	68.5	41.9	78.5
Method 2: Clustering with Traffic Flow Model	39.1	24.9	44.6

The results in Table 3-6 are consistent with the results in Figure 3-5 and Figure 3-6. However, it is noticeable that the $RMSE$ value associated with the Benchmark Method is smaller than Method 1 (overall and for intervals when HOV is faster than GPL). This result is caused by the simulation inputs design assumption in this study: the two facilities of HOV and GPL work together as a self-correcting system, and some drivers with the option of travelling on both HOV and GPL will switch from HOV to GPL when HOV is operating slower than GPL. In the design of traffic input patterns, we ensured that when the traffic stream speed on HOV is slower than GPL, the speed difference is at a maximum of around 20 km/h, however, when HOV is faster than GPL, the speed difference can be as large as 50km/h. The boundary values of 20 km/h and 50 km/h were determined from the empirical data collected on Highway 403, and are corresponding to 1 percentile and 99 percentile of the speed differencials between HOV and GPL, respectively. As a result, when the Benchmark Method is not assigning the facilities correctly under the condition when HOV is actually slower, the model error is relatively small ($RMSE$ of 25.6 seconds) and such errors only occur for a small percentage of intervals (around 30% of all intervals and value determined from empirical data). However, when the speeds associated with HOV facility is faster, the Benchmark Method model is always correct. In contrast, the proposed Method 2 is sometimes wrong in assigning

facilities when HOV is actually faster, and such wrong assignment can then leads to quite large errors ($RMSE$ of 48.7 seconds).

The above results imply that the relative frequency of intervals in the input data for which one facility is slower/faster than the other facility and the relative speed differences (i.e., the underlying traffic conditions) have a significant impact on the performance of different models. Therefore, a sensitivity analysis is performed in the next section to understand the relationship between the traffic condition and the model performances. When using the proposed methods in this research, users can decide the appropriate model structures for application based on the prevailing traffic condition on the target roadway.

3.4.4 Sensitivity Analysis of Model Performances under Different Traffic Conditions

In the previous section, it was noticed that the proposed models perform differently under different traffic conditions. Therefore, a sensitivity analysis has been performed in this section. By randomly sampling one hundred 5-minute intervals meeting the pre-defined traffic conditions, the results provide insights into the quantitative relationship between different traffic conditions and model performance. The different traffic conditions are defined based on varying the relative frequency of intervals for which the HOV lane is slower than the GPL. Note that within each traffic conditions (e.g., 30% of intervals having HOV lane operating slower than GPL), the relative travel time (speed) differences between the two facilities are not controlled, because we desire the proposed models to work under all possible conditions not just one particular relative speed difference between the two facilities. The $RMSE$ results associated with the different percentage of intervals having the HOV lane operating slower than GPL are summarized in the following table and figure.

Table 3-7 Sensitivity analysis of *RMSE* results for different model structures

Percentage of Intervals with HOV Operating Slower than GPL	RMSE (seconds)		
	Benchmark Method (Clustering with HOV Faster Assignment)	Method 1 (Distribution Matching with AD Test)	Method 2 (Clustering with Traffic Flow Model Assignment)
0%	37.4	84.6	46.1
10%	30.4	65.3	50.6
20%	37.2	91.6	40.3
30%	27.4	61.7	33.0
40%	30.7	58.3	36.8
50%	32.5	63.5	34.8
60%	27.6	56.1	29.9
70%	30.7	52.0	30.8
80%	30.3	51.6	27.2
90%	29.6	44.0	27.4
100%	29.5	39.8	24.6

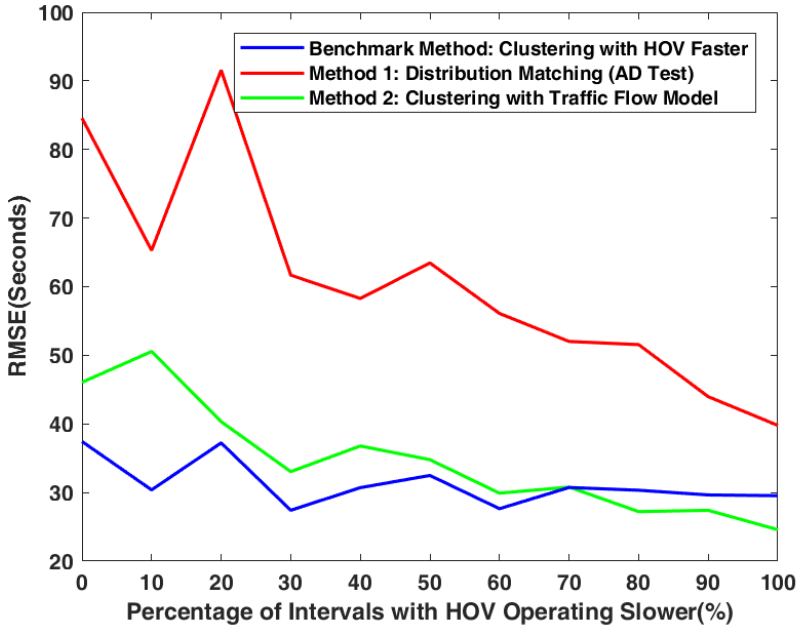


Figure 3-7 RMSE results for different model structures under different conditions

The results in Table 3-7 and Figure 3-7 reveal the impacts of changing traffic conditions on the performance of the different methods. The following observations can be made based on the sensitivity analysis results.

1. The Benchmark Method achieves better performance than Method 1 or Method 2 in terms of overall facility specific travel time $RMSE$ when the assumption of the Benchmark Method (i.e., that the HOV lane is faster than the GPL) is normally true.
2. Both the proposed Method 2 and the Benchmark Method outperforms the proposed Method 1 in terms of overall facility specific travel time $RMSE$ across all the traffic conditions investigated.
3. The performance of Method 2 improves with the increase in the percentage of prevailing conditions having HOV operating slower than GPL.

We can also note from Figure 3-7 that Method 2 provides better performance than the Benchmark Method when 70% of the intervals are associated with HOV speed less than the GPL speeds. However, it should be noted that the above sensitivity analysis results are still subject to one underlying assumption, namely that no physical barriers exist between the HOV lane and GPL lane, and most drivers having the option of switching between HOV and GPL will switch out of HOV when the HOV lane(s) is operating slower than GPL. Consequently, the speed differential corresponding to when the HOV is operating faster than the GPL is generally larger than the differential when the HOV is slower than the GPL. This assumption of no physical barriers between HOV and GPL is consistent with the existing geometric configurations in Ontario freeway systems. However, if physical barriers do exist between the two underlying facilities such that we can expect speed differentials to be equal, the sensitivity analysis results would be expected to be different from the above discussed results.

3.5 Summary and Conclusions

In this chapter, two different methods are proposed in estimating facility specific travel times from Bluetooth measurements, and the performance of these methods have been evaluated against a Benchmark method using simulation data. The Benchmark method assumes HOV always operating faster than GPL, and it first clusters the recent past travel time measurements using the K-means algorithm, and then associate the faster cluster with HOV and the slower cluster with GPL. Method 1 applied the Anderson-Darling test in matching the distribution of real-time Bluetooth travel time measurements with reference measurements. Method 2 first cluster the recent past travel time measurements using the K-means algorithm, and then associate the clusters with facilities using Van Aerde traffic flow model. A sensitivity analysis has been conducted to understand the impact of traffic conditions on the performances of these different methods. The results showed that:

1. Benchmark Method performs well when the assumption of the HOV is faster than GPL is normally true. The Benchmark Method is recommended for applications in separating two facilities when the roadway under analysis function as a true self-correcting system where vehicles have the ability of freely switching back to GPL from HOV whenever desired.
2. Method 1 – Distribution Matching cannot accurately provide facility-specific travel times as is evident from the $RMSE$ results: Method 1 has the largest facility specific travel time $RMSE$ for all conditions, and thus is not recommend for practical applications.
3. Method 2 – Clustering with Traffic Flow Model assigns facilities to clusters based on Van Aerde traffic flow models and has a strong theoretical foundation, and the performance of the Method 2 improves with the increase in the percentage of prevailing conditions having HOV operating slower than GPL. Therefore, the Method 2 is recommended for applications when physical barriers (or traffic enforcement) prevent drivers from freely switching between underlying facilities or when the speed differentials corresponding to when the HOV is operating faster than the GPL is generally equal or smaller than the differential when the HOV is slower than the GPL.

Chapter 4
Estimating Traffic Conditions from Non-Lagged Bluetooth Measurements

Bluetooth detection systems offer many comparative advantages over conventional detection system (e.g., inductive loop detector based systems) in terms of direct travel time measurements, flexible installation, and cost-effectiveness collection of data over long periods of time. However, similar to many other Automatic Vehicle Identification (AVI) systems such as automatic license plate recognition and toll tag reader based systems, travel time measurements obtained using Bluetooth detectors are influenced by measurement lag because the travel time cannot be measured until vehicles have traversed the entire roadway segment and have been identified by the downstream Bluetooth detectors. Consequently, travel time information are not available from vehicles that have not yet reached the downstream Bluetooth detector at the instant of analysis.

In order to alleviate travel time measurement lag, this chapter proposes methodologies for estimating the traffic conditions in the vicinity of Bluetooth detectors using non-lagged measurements. Examples of non-lagged measurements include the number of repeated detections associated with the same unique MAC address (Hits) and the duration of exposure to the Bluetooth detector by each unique MAC address (Dwell Time). These non-lagged measurements can be extracted from Bluetooth measurement systems at fixed time intervals (e.g., every minute) and consequently, the estimated traffic conditions can more timely reflect the spatial and temporal variations in traffic conditions and can contribute to improving the travel time estimation and prediction accuracy.

4.1 Literature Review

Travel time measurement lags induce difference between estimated travel times and the actual travel times (Van Lint and Hoogendoorn, 2010) and deteriorate the accuracy of travel time estimations (Bhaskar, 2009). To address the measurement lag, Hu (Hu, 2013) recommended that information not suffering from the measurement lag should be used in travel time studies.

Bhaskar and Chung (Bhaskar and Chung, 2013) provided detailed reviews on the technical characteristics of Bluetooth technology for transportation applications. In general, Bluetooth detections are influenced by factors including the radio signal strength of the transmitting device, the type, and placement of Bluetooth antenna (Malinovskiy et al., 2010), the installation location of Bluetooth scanners (Quayle et al., 2010), etc. However, with fixed Bluetooth hardware and protocol implementation, the detections at each detector are then mostly determined by the speeds of vehicles carrying observable Bluetooth devices, which in turn

determine the durations vehicles spent within the Bluetooth detection zone (Young et al., 2012). In addition, each unique Bluetooth MAC address can be detected multiple times during the exposure to the Bluetooth scanner, and many researchers treat multiple hits as detection ambiguity (Araghi et al., 2013; Díaz et al., 2016). However, the duration of exposure to the Bluetooth detector by each MAC address and the number of repetitive detections for the same MAC address can all be calculated at fixed durations and are less vulnerable to travel time measurement lag. These measurements are referred to as non-lagged measurement in this study.

Very limited research efforts have been devoted to examining non-lagged information from Bluetooth measurements. Schneider et al. (Schneider et al., 2012) reported that the fraction of recorded MAC addresses for which only a single hit was obtained decreases during incident periods, and therefore these authors recommended that Bluetooth technology can be used for incident detection purposes. However, this research didn't quantify the relationship between traffic speeds and hit counts. Moghaddam (Moghaddam, 2014) used the total number of Bluetooth detections at intersections as surrogate measures for intersection congestion level and found models incorporating non-lagged measurements as inputs achieved better travel time prediction results on urban arterials compared to models without such inputs. Junghans and Leich (Junghans and Leich, 2016) proposed to estimate the travel speeds of individual Bluetooth devices by assuming constant Bluetooth polling intervals of 2.56 seconds and fixed Bluetooth detector detection range; then, these authors compute the individual vehicle speed as equal to the detector range divided by the dwell time (i.e., time difference between the first and last hit of each vehicle detected). Note that these authors recognized that their Bluetooth speed estimation results were "rather rough," and they conducted data fusion of Bluetooth speed data with V2X (vehicle to infrastructure and vehicle to vehicle) speed measurements for better accuracy.

This chapter proposes methods and strategies for estimating traffic stream states from non-lagged Bluetooth measurements. The inferred traffic states could more timely reflect the variations in traffic conditions than travel time measurements because travel time can only be measured after vehicles have traversed the entire roadway and vehicles have been detected at both upstream and downstream Bluetooth detectors. In Chapter 5, the estimated traffic states from models developed in this chapter are incorporated into travel time estimation and prediction framework for improving the accuracy of travel time information.

4.2 Problem Formulation

Consider Figure 4-1 which depicts a time-space with idealized trajectories of vehicles traversing a hypothetical freeway segment. The present time is indicated by t_0.

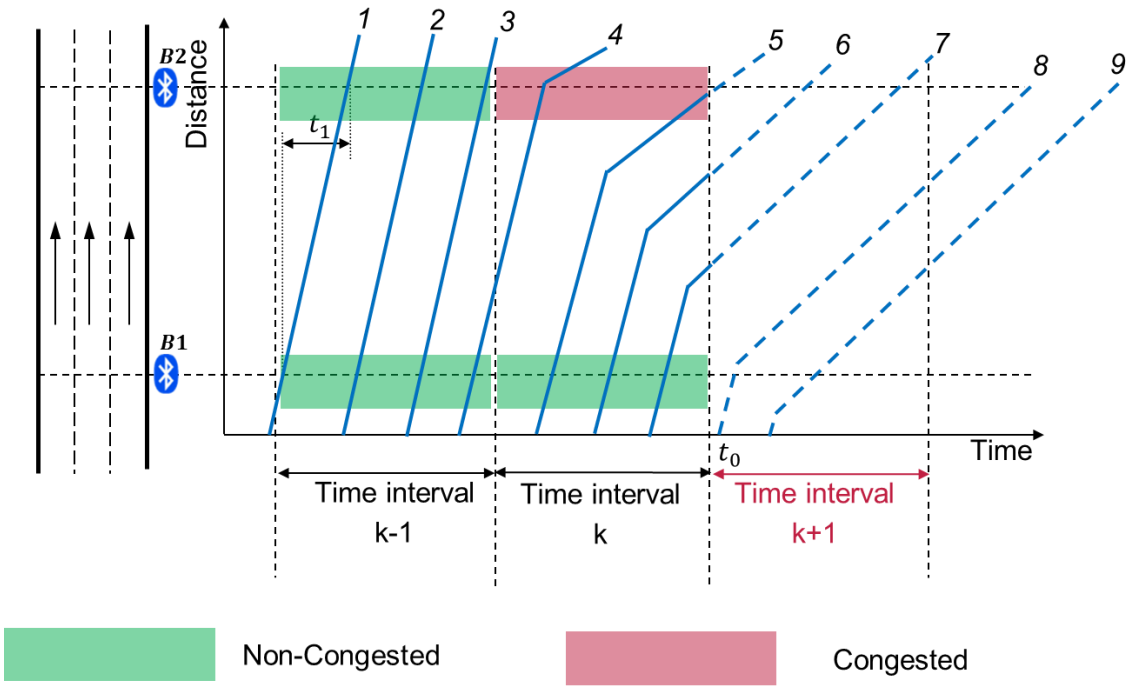


Figure 4-1 Congestion forming on a freeway segment

In Figure 4-1, the congested traffic conditions occurred at the downstream location of the target freeway segment (downstream of B2) resulting in longer travel times (compared to interval $k-1$) for vehicles entering the segment within time interval k. Because no vehicle entering the target freeway segment within interval k has traversed the entire segment by the end of the interval k, travel time measurement is not available to reflect the change in traffic conditions by the instant t_0. However, the traffic states in the vicinity of B2 is readily showing congestion by the end of time interval k (congestion represented by the red block), and considering the congested traffic states around Bluetooth detector B2 in time interval k is hypothesized to improve the accuracy of travel time estimation and eventually can lead to better travel time prediction results. The proposed methodology in this chapter estimates traffic states based on non-lagged Bluetooth measurements.

Non-lagged measurements including Hits and Dwell Time are considered in this study because they can be obtained from an individual Bluetooth detector (thus avoiding the time lag problem associated with travel times) and they are related to the traffic conditions on the roadway in the detection area.

- Hit: each detected unique MAC address (representing a vehicle) can be detected for single or multiple times while traversing the detection zone of a Bluetooth detector, each of these detections is defined as a Hit;

- Dwell time: the amount of time that each unique MAC address spends in the Bluetooth detection zone is defined as the Dwell Time, which can be calculated as the time difference between the time stamps of the last Hit and the first Hit for the MAC (for a MAC address with only one Hit, the Dwell Time is not defined).

The relationships between traffic stream states and these two measurements are analytically formulated by examining the Bluetooth detection process.

Bluetooth specification divides the available bandwidth into 79 sub-channels, and 32 of these sub-channels are allocated for the discovery of Bluetooth devices. In typical Bluetooth detection systems, Bluetooth detectors (referred to as master devices) constantly operate in inquiry sub-state while discoverable Bluetooth devices (referred to as slave devices) operate in inquiry scan sub-state. The operational behaviors of the masters and the slaves are different.

- Master (Bluetooth Detector)

For a master device in inquiry state, the available 32 sub-channels are divided into train A and train B; each train consists of 16 channels. The master device inquires on each channel for 0.625 milliseconds and consequently requires 10 milliseconds to steps through all 16 channels in a train. The Bluetooth specification dictates that a master device should repeatedly make inquiries on a train (A or B) for at least 256 times before switching to frequencies in the other train. Since it takes 10 milliseconds to complete inquiry one train, a master device spends 2.56 seconds (256 * 10 $milliseconds$) on one train before moving to the other train. It should be noted that different Bluetooth stacks (software implementing Bluetooth functionality) implement different mechanism for reporting newly discovered slave devices (Bakula et al., 2011): some stacks report immediately after device discovery, while other stacks might save the discovered devices and report after the end of the current inquiry period (reported at fixed pulling interval).

- Slave (Discoverable Devices)

For a slave in the inquiry scan state, it is listening on one of the 32 sub-channels for 11.25 milliseconds every 1.28 seconds (Welsh et al., 2002) following a pre-determined order. Because the master in inquiry state scans available channels at much higher rate (scans 16 channels every 10 $millisecond$) compared to the slaves in inquiry scan state (listen to 1 channel for 11.25 ms), if the channel that a slave is listening on is among the 16 channels within train A or B that the master device is repeatedly calling out at, the slave has very high likelihood of being detected and would respond with its Bluetooth MAC address to the master device. Upon recipient of the slave MAC address by the master with a low transmission error probability, the received MAC address is then time stamped and the discovery is completed (Bluetooth Special Interest Group, 2009).

Considering the behaviors of Bluetooth detectors (masters) and discoverable devices (slaves) mentioned above, the average number of hits per Bluetooth equipped vehicle can be estimated as a function of the time a vehicle spends within the Bluetooth detection zone; the Bluetooth polling interval; and the probability that a Bluetooth device within the detection zone can be detected.

$$AvgHit = 3600 \times (L/V_{flow}) \times (1/T_R) \times [1 - (1 - AvgP)^n] \quad (4\text{-}1)$$

Where $AvgHit$ is the average number of hits for vehicles traversing a Bluetooth detection zone; L is the detection zone length (km), which equals twice the size of Bluetooth detection radius; V_{flow} is the speed of the traffic stream (km/h); T_R is the Bluetooth detector reporting interval (second), equals $1.28n$ seconds with integer $n > 1$, typical values for T_R are 3.84 or 5.12 seconds with n equals 3 or 4, the value of T_R should be larger than 2.56 seconds or the duration for scanning one train; $AvgP$ is the average probability that a Bluetooth equipped vehicle can be detected during its exposure to one Bluetooth detector, the value can be estimated considering the decay of Bluetooth signal strength over distance.

Assuming the Bluetooth detection possibility follows a trapezoid shape (Moghaddam and Hellinga, 2013) as follows:

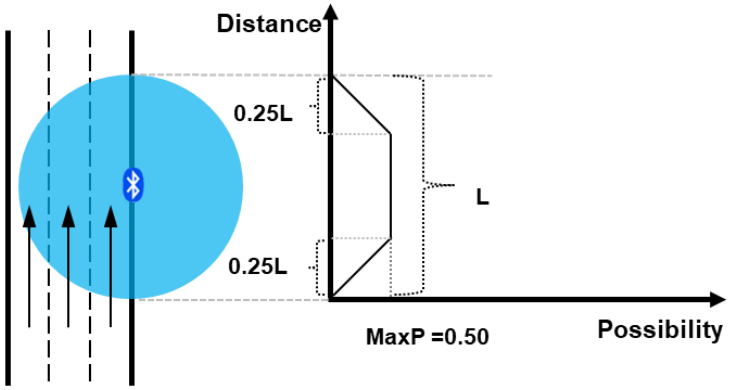


Figure 4-2 Bluetooth detection possibility distribution

$AvgP$ can be estimated using the following equation:

$$AvgP = (0.25L \cdot 0.5 \cdot MaxP + 0.5L \cdot MaxP + 0.25L \cdot 0.5 \cdot MaxP)/L = 0.75\ MaxP \quad (4\text{-}2)$$

Where $MaxP$ is the maximum probability that a discoverable Bluetooth device can be detected during each of its inquiry scan period. At any time instant the master is scanning a single train consisting of 16 channels out of the 32 sub-channels, and therefore the maximum detection probability ($MaxP$) for a slave is $16/32 = 0.5$. Substituting this value into Equation (4-2) then $AvgP = 0.375$.

Substituting $AvgP$ into equation (4-1) and rearranging the equation, the formula for V_{flow} is as follows:

$$V_{flow} = \begin{cases} \dfrac{2812.5 \cdot L\,[1 - 0.625^{n}]}{AvgHit \cdot n} & if\ AvgHit > 1 \\ V_{free} & if\ AvgHit = 1 \end{cases} \quad (4\text{-}3)$$

The average dwell time can be calculated as a function of the average number of Hits:

$$AvgDwell = T_R \cdot (AvgHit - 1) = 1.28n \cdot (AvgHit - 1) \quad (4\text{-}4)$$

Where, AvgDwell is the average Dwell Time in seconds, and other variables are defined the same as above. Substituting Equation (4-4) into (4-3), the following formula describing the relationship between traffic stream speeds and average Dwell Time can be derived:

$$V_{flow} = \begin{cases} \dfrac{3600 \cdot L\,[1 - 0.625^{n}]}{AvgDwell + 1.28 \cdot n} & if\ AvgDwell > 0 \\ V_{free} & if\ AvgDwell = 0 \end{cases} \quad (4\text{-}5)$$

The Equations (4-3) and (4-5) reveal that traffic stream speeds and non-lagged Bluetooth measurements follow non-linear relationships. Note that both equations (4-3) and (4-5) are examining the aggregated non-lagged measurements instead of individual non-lagged measurements, because measurements associated with individual vehicles are influenced by factors including the driver behavior characteristics, the implementations of Bluetooth stack within the discoverable device (Bakula et al., 2011), the existence of obstacles, the Doppler effect from the moving of vehicles, etc. The aggregated non-lagged measurements allow a better understanding of the traffic stream conditions.

The next section describes the set of empirical data that have been used to calibrate and evaluate the proposed models (including Equations (4-3) and (4-5)) for estimating traffic conditions.

4.3 Field Site and Empirical Data

This section describes the set of field data that was used to calibrate and evaluate the model proposed in Equations (4-3) and (4-5). The data were collected from the Don Valley Parkway (DVP), an urban freeway in the City of Toronto. The posted speed limit on the DVP is 90 km/h. Data were collected from this location because it is instrumented with both vehicle detection stations (VDS) and Bluetooth detectors. Vehicle detector stations (VDS) consist of dual loop detectors were deployed in each vehicle lane, and VDS detectors record traffic volume, detector occupancy, and average speed every 20 seconds. At some locations, the VDS stations are located within 100 meters of Bluetooth detectors.

The following section describes the VDS data and methods used for data processing and aggregation.

4.3.1 Vehicle Detector Station (VDS) Loop Data Aggregation

As indicated in Figure 3-5, the VDS located close to the DVP/Bloor junction and close to DVP/Danforth junction were chosen as the study targets, because Bluetooth detectors were placed close to these locations and both loop and Bluetooth data can be compiled for model evaluation.

Figure 4-3 Location of chosen Bluetooth and VDS on DVP (Image source: Google Earth; red balloon represents Bluetooth detector, yellow pin represents VDS detectors)

The raw (20-second) VDS data were processed by applying three sequential filtering methods and then aggregated across all lanes and for 1-minute periods. Only data between 5 AM to 10 PM were compiled in this study because traffic conditions are more likely to change within the 5 AM to 10 PM daily duration.

4.3.1.1 Filter 1:

The first filter determines that if a sufficiently large proportion of data from an entire day are dubious, then all data from that day are removed from further consideration. This filter attempts to address issues that arise from VDS malfunction and/or data communication failures. This filter is developed based on the method proposed by Jain and Coifman (Jain and Coifman, 2005) and consists of the following examinations:

A. Determine the fraction of 20-second raw observations in each day between 5 AM and 10 PM for which the volume = 0 and the occupancy = 0.
B. Determine the fraction of 20-second raw observations in each day between 5 AM and 10 PM for which the volume = 0 but the occupancy > 0.
C. Determine the fraction of 20-second raw observations in each day between 5 AM and 10 PM for which the volume > 0 but the occupancy = 0.
D. Determine in each day between 5 AM and 10 PM the maximum duration for which the consecutive raw 20-second occupancy values remain unchanged.

If for any day, the fraction of rule A is over 50% or the fraction of rule B is over 25% or the fraction of rule C is over 25% or the duration of rule D is over 2 hours, all the data from this day was excluded from further analysis.

4.3.1.2 Filter 2:

The second filter attempts to eliminate outliers by applying a set of thresholds to each of the 20-second volume, speed, and occupancy measurements (Table 3-6). If a measurement from a detector for a given time interval is invalid, then all three measurements (volume, speed, and occupancy) for that detector for that time interval were marked as invalid:

Table 4-1 Validity range for measurements at the 20-second level

VDS Measurements	Valid Value Ranges	Unit
Volume	$0 \leq \text{Volume} < 20$	$vehicle\ per\ 20\ second\ per\ lane$
Speed	$0 \leq \text{Speed} < 135$	km/h
Occupancy	$0 \leq \text{Occupancy} \leq 100$	%

4.3.1.3 Filter 3:

The third was formulated to identify data in which the different measurements from a detector for a given time interval were inconsistent with each other. Three conditions were defined as follows:

a) Volume equals 0, but occupancy is larger than 0
b) Volume is larger than 0, but occupancy equals 0
c) Volume is larger than 0, but speed equals 0

If any of these conditions were observed, the then all three measurements (volume, speed, and occupancy) for that detector for that time interval were marked as invalid.

4.3.1.4 Aggregation

Data that had passed the three filters were aggregate into 1-minute level across lanes at each station, using the following equations:

$$v^{(n)} = \frac{\sum_{i=1}^{3}\sum_{k=1}^{K} N_{i,k}}{\sum_{i=1}^{3}\sum_{k=1}^{K} \frac{N_{i,k}}{v_{i,k}}}$$

$$q^{(n)} = \frac{3K}{M}\sum_{i=1}^{3}\sum_{k=1}^{K} N_{i,k} \qquad (4\text{-}6)$$

$$o^{(n)} = \frac{\sum_{i=1}^{3}\sum_{k=1}^{K} o_{i,k}}{M}$$

Where

$v^{(n)}$ = the volume weighted average speed for the target VDS station at 1-minute level.

i = the index of the 20-second interval with valid detections.

k = the index of underlying freeway lanes at the target VDS station with valid detections, the maximum value equals the total number of lanes K.

$N_{i,k}$ =raw number of vehicle observations in 20-second interval i from lane k.

$v_{i,k}$ = raw speed measurements in 20-second interval i from lane k.

$q^{(n)}$= aggregated vehicle volume for the target VDS station at 1-minute level.

K= the total number of freeway lanes at the target VDS station. Therefore, the total number of raw VDS records under aggregation equals $3K$, which is the number of 20-second intervals within the aggregation time duration (3 for one 1-minute interval) multiplied by the total number of freeway lanes.

M = the number of time intervals with valid observations at the target VDS station.

$o^{(n)}$= aggregated occupancy for the target VDS station at 1-minute level.

$o_{i,k}$= occupancy measurements for 20-second interval i from freeway lane k.

In the data aggregation process, measurements for some time intervals in some lanes were missing or labeled as invalid (because they were excluded in the filtering process). For any 1-minute interval with less than 3 records (33% of 9 complete records, where 9 equals 3 intervals times 3 lanes), a linear imputation technique was used to estimate the aggregated values based on the aggregated measurement from the previous and the following valid 1-minute intervals. Figure 4-4 summarizes the data filtering and aggregation procedures.

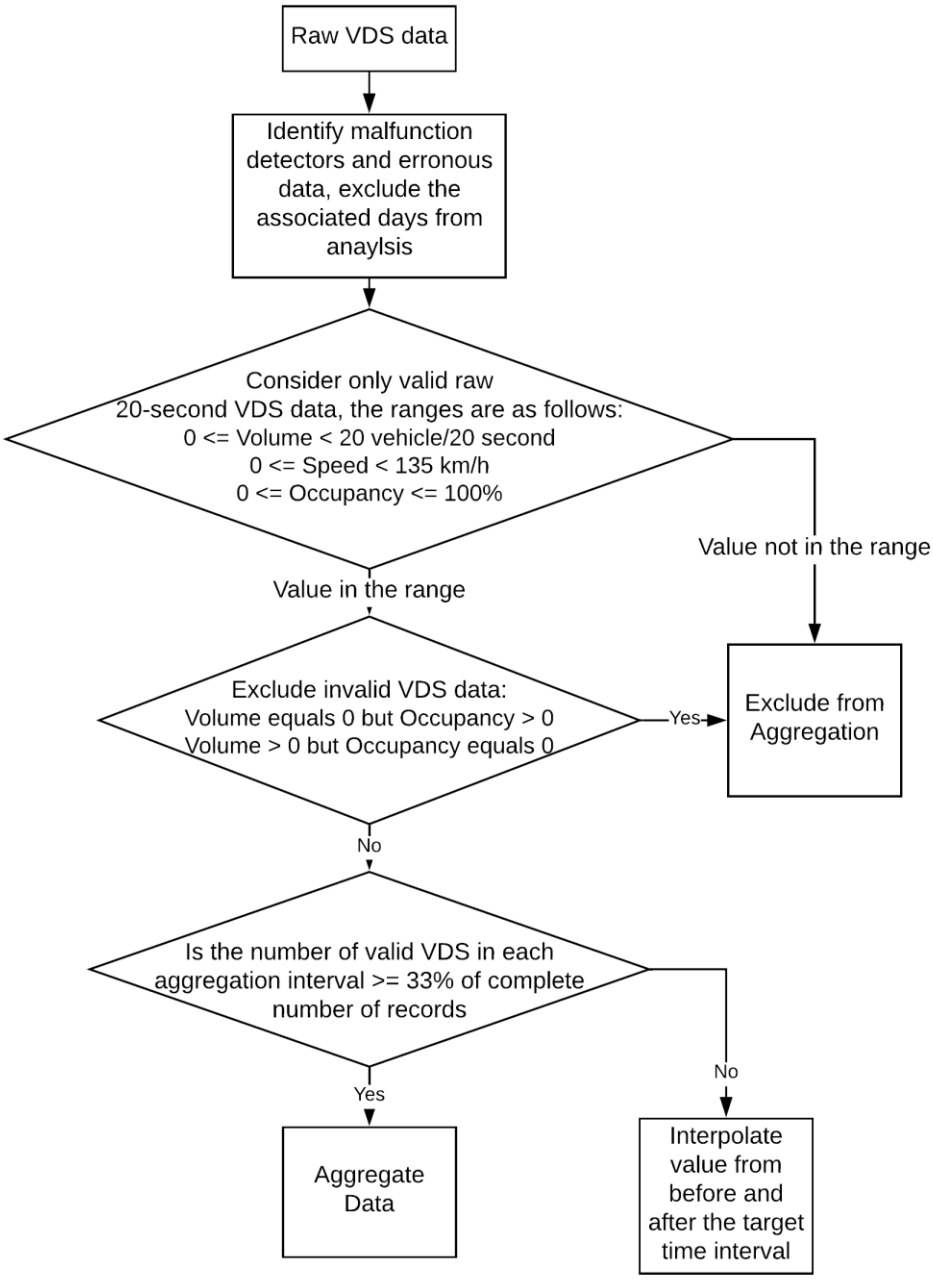


Figure 4-4 Procedures for VDS loop data filtering and aggregation

4.3.1.5 Final Empirical Loop Detector Data Set

The final data set consisted of loop detector measurements aggregated across all three lanes and over 1-minute time intervals. Data were available for five detector stations including DVP/Bloor northbound, DVP/Bloor southbound, Midblock southbound, DVP/Danforth northbound, and DVP/Danforth southbound. Data were obtained for the calendar period from February 1, 2017 to March 31, 2017 including both

weekdays and weekends. During this period, valid data were obtained from 45 days. Given that the focus for this research work is to predict travel times and the situations of interest are those times of day when travel times are most likely to change, only data from 5 AM to 10 PM were considered. Figure 4-5 illustrates the time series of aggregated loop detector measurements for detector station DVP/Danforth southbound on March 3, 2017.

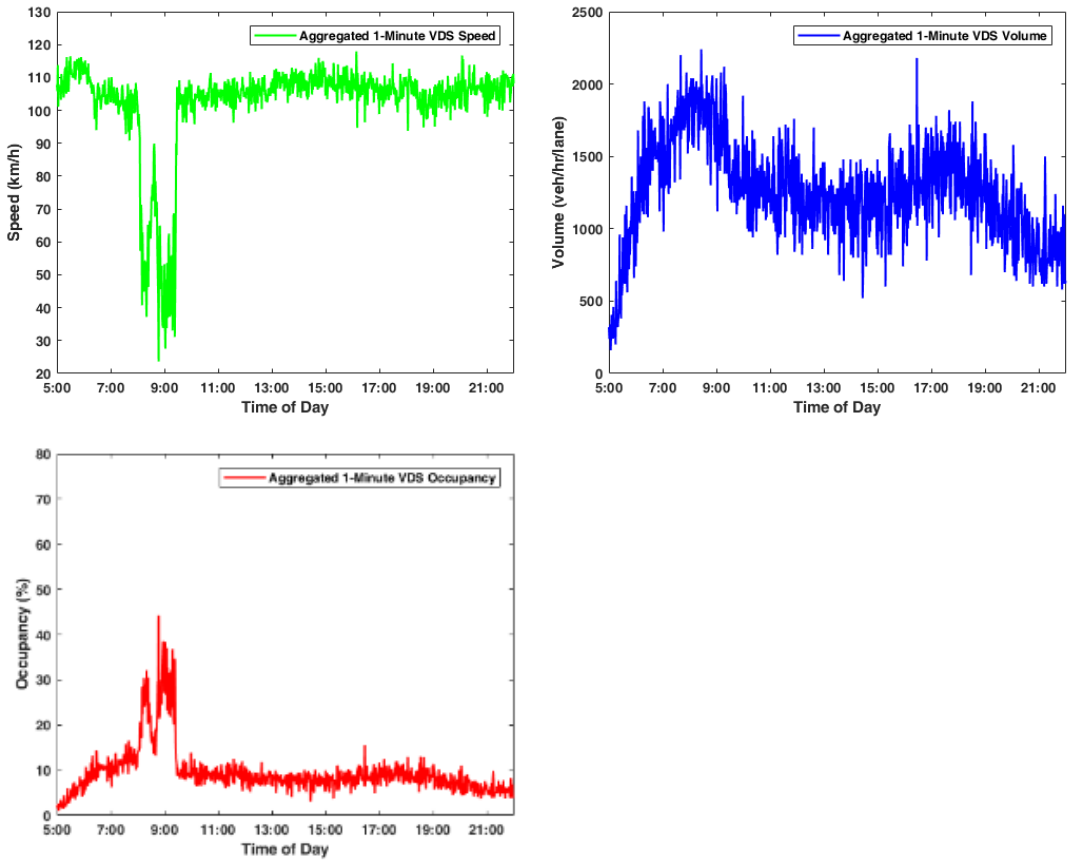


Figure 4-5 Aggregated VDS measurements for a sample day at DVP/Danforth southbound location

4.3.2 Bluetooth Data

A vendor has deployed Bluetooth detectors on the DVP under contract with the City of Toronto. Bluetooth detector data were obtained from the vendor for this study. The data consisted of the individual vehicle travel times as well as the raw Bluetooth measurements obtained at each individual Bluetooth detector. Because the raw Bluetooth measurements contain detections from both directions of travel (no directional antenna installed with Bluetooth detectors at the data collection sites), using the raw Bluetooth travel time measurements as references, the Bluetooth detections associated with only the directions of interest were extracted.

Similar to the treatment of raw 20-second VDS data, the raw Bluetooth detections need to be filtered to retain only valid measurements before aggregation. The filtering algorithm implemented by the vendor is used for removing outliers in the Bluetooth measurements (both the individual vehicle travel times and the raw Bluetooth measurements) because the algorithm has been validated before implementation for the Don Valley Parkway traveler information systems. For all the collected raw Bluetooth measurement, the filtering algorithm labels each individual measurement as valid or invalid. After applying the filtering algorithm, around 8% of raw measurements were labeled as invalid and were disregarded from further analysis.

The relationships between traffic stream speeds and two non-lagged Bluetooth measurements namely average Hits and average Dwell Time were analytically formulated in section 4.2. Aside from Hit and Dwell Time, other non-lagged measurements such as the standard deviation of Dwell Time and the standard deviation of Hit over different aggregation interval durations were also extracted from Bluetooth measurements.

The average Hits and average Dwell Time were aggregated into the 1-minute level. Considering the following diagram for explaining the aggregation of non-lagged Bluetooth measurements.

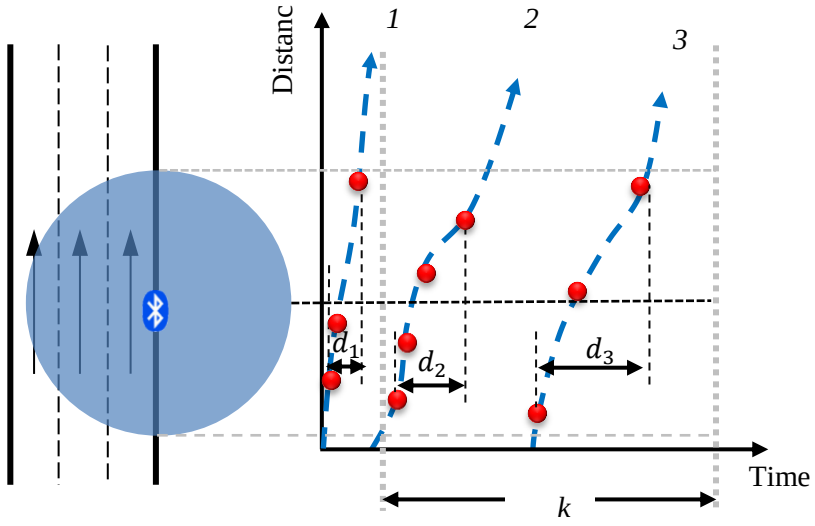


Figure 4-6 Aggregation of non-lagged Bluetooth measurements

In Figure 4-6, each solid round dot represents one hit (one observation) of the associated MAC address. In time interval k, the Bluetooth detector obtained 4 hits for vehicles 2 and 3 hits for vehicle 3, and the dwell time for vehicle 2 and 3 are d_2 and d_3, respectively. The average Hits for time interval k is 3.5, and the average Dwell Time for time interval k is $(d_2 + d_3)/2$.

After filtering, the data set consisted of Bluetooth measurements from 59 days between February 1, 2017 and March 31, 2017 with both weekdays and weekends included. Since this date range overlaps with the

VDS date range, only 45 days associated with valid data from both Bluetooth and VDS detectors were retained. In addition, only Bluetooth measurements obtained from 5 AM to 10 PM were considered. Figure 4-7 illustrates the time series of aggregated Bluetooth measurements including Average Hits and average Dwell Time from Bluetooth detector station close to DVP/Danforth southbound on March 3, 2017.

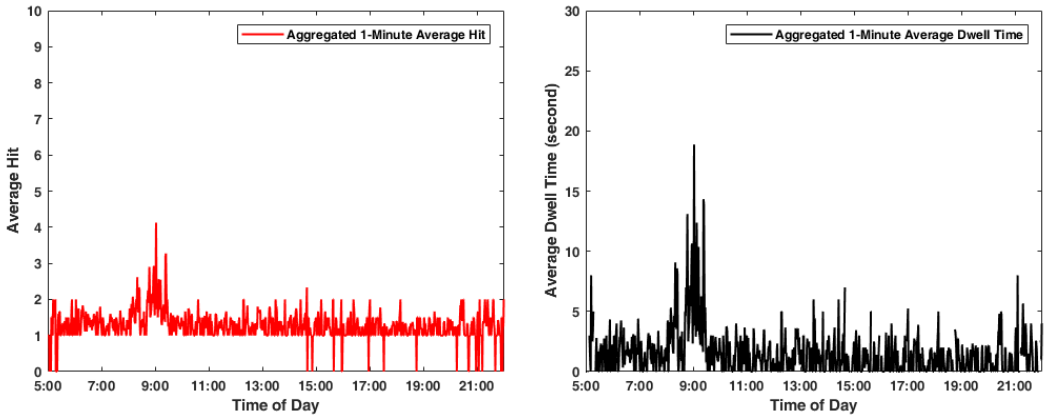


Figure 4-7 Aggregated Bluetooth measurements for a sample day at DVP/Danforth southbound location

4.4 Approach 1: Analytical Model to Estimate Speed

This section calibrates and evaluates the analytical model from section 4.2 for estimating traffic stream speeds.

The two analytical models including Equations (4-3) and Equation (4-5) for estimating traffic states are re-summarized as follows:

$$V_{flow} = \begin{cases} \dfrac{2812.5 \cdot L\,[1-0.625^{n}]}{AvgHit \cdot n} & if\ AvgHit > 1 \\ V_{free} & if\ AvgHit = 1 \end{cases} \quad (4\text{-}3)$$

$$V_{flow} = \begin{cases} \dfrac{3600 \cdot L\,[1-0.625^{n}]}{AvgDwell + 1.28 \cdot n} & if\ AvgDwell > 0 \\ V_{free} & if\ AvgDwell = 0 \end{cases} \quad (4\text{-}5)$$

Where L is the detection zone length (km), which equals twice the length of Bluetooth detection radius; n is an integer which defines the Bluetooth detector reporting interval (equals $1.28n$), the typical value of n is 3 or 4; $AvgHit$ is the average number of hits for vehicles traversing a Bluetooth detection zone; $AvgDwell$ is the average dwell time in seconds. V_{flow} is the estimated traffic stream speed.

4.4.1 Model Parameter Calibration

For the DVP data provided by the vendor, the Bluetooth detector scanning length is 4 seconds. Therefore, the value of n is set as 3 and the estimated Bluetooth detector reporting interval is 3.84 seconds ($\approx$ 4 seconds). The value of effective detection zone length L is then calibrated with aggregated 1-minute level VDS data and Bluetooth data collected between February 13 and February 23, 2017 on DVP.

The impact of different detection zone length L is measured using the root mean square error (RMSE) of speed estimation results. The definition of the speed estimation RMSE is as follows:

$$RMSE_d = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left(v_k - \hat{v}_{d,k}\right)^2} \tag{4-7}$$

Where k is interval index ($k = 1,2,\ldots,n$), v_k is the aggregate 1-minute level VDS speed measurements for interval k, $\hat{v}_{d,k}$ is the estimated speed for the time interval k using analytical model d, $RMSE_d$ is the root mean square error associated with analytical model d.

The calibration of length L for the two models are provided in the following diagram.

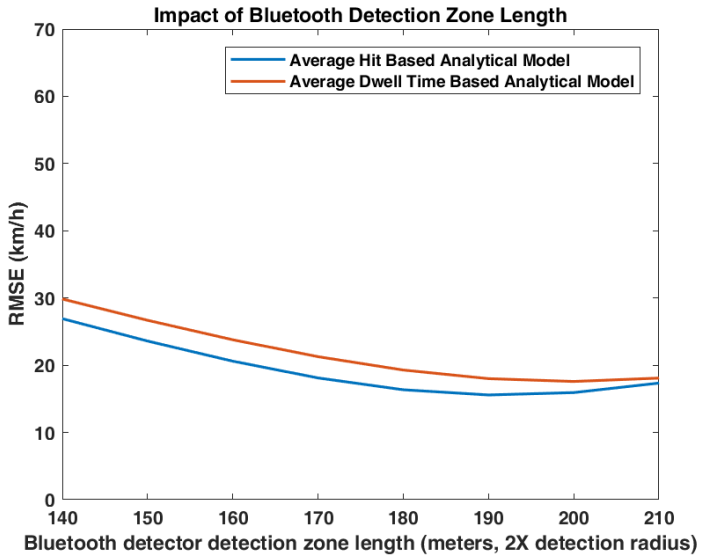


Figure 4-8 Calibration of Bluetooth detection zone length for analytical models

The smallest RMSE is achieved by the average Hits based analytical model (Equation (4-3)) when L equals 190 meters. The associated lowest RMSE value is 15.6 km/h. Therefore, the chosen value for n is 3 and the chosen value of Bluetooth detection zone length is 190 meters. In the next section, the performance of the average Hits based analytical model is evaluated using a different dataset.

4.4.2 Model Evaluation

In order to evaluate the performance of the calibrated model, additional 1-minute level VDS data and Bluetooth data were compiled between February 24 and March 6, 2017 on DVP. Using the new data set, the RMSE of the estimated speeds from the calibrated average Hits based analytical model are summarized as follows:

Table 4-2 RMSE of the calibrated average Hit based analytical model

Date	RMSE of Average Hit based Analytical Model (km/h)
24-Feb-2017	16.1
25-Feb-2017	15.9
26-Feb-2017	17.3
3-Mar-2017	15.9
4-Mar-2017	16.1
5-Mar-2017	19.1
6-Mar-2017	15.5

The results in Table 4-2 reveal that the RMSE value in the testing data set is in the same range as the calibration dataset (RMSE value of 15.6 km/h). In addition, the relationship between the aggregated 1-minute level VDS measurement speed and the model estimation speeds and the relationship between the aggregated 1-minute level VDS measurement speed and the absolute speed estimation errors for the testing data set are provided in the following diagram:

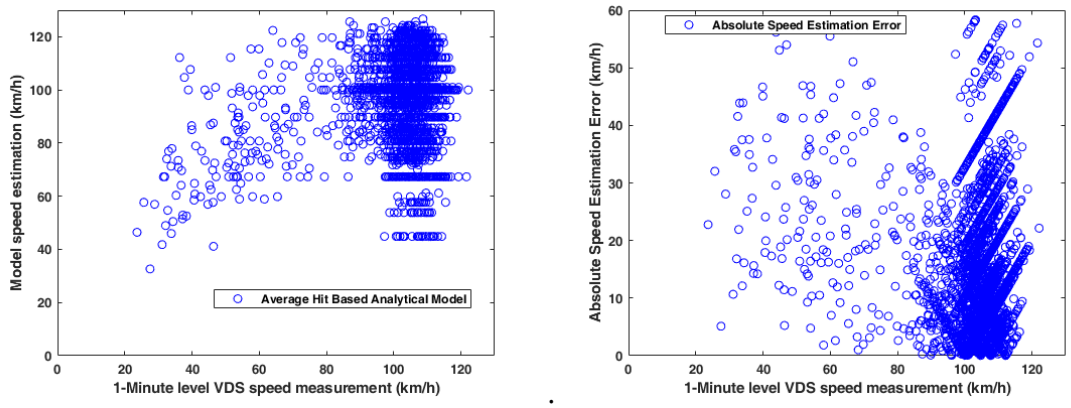


Figure 4-9 Relationship between the VDS speed measurements and the analytical model estimation results

The results in Figure 4-9 show that the majority of time intervals in the testing data set is non-congested. In addition, regardless of the traffic condition, the errors of the analytical model can be as large as 60 km/h for some intervals, and it is evident that the analytical model is not able to consistently estimate traffic

stream speeds accurately. Therefore, in the next section, statistical models for establishing the relationship between traffic stream states and non-lagged measurements are proposed.

4.5 Approach 2: Statistical Model to Estimate Traffic State

The previous section describing modeling approach 1 has shown that the non-lagged measurements available from the DVP Bluetooth detectors are not sufficient to estimate traffic stream speed with a level of accuracy that is considered necessary for improving travel time predictions. However, it is speculated that the most important information for improving travel time prediction is the knowledge of whether or not there is congestion at the downstream or upstream boundary of the road section and not the traffic stream speed itself. Consequently, in this section, we present a model that attempts to estimate the traffic state, rather than the traffic stream speed. In the next section, we define traffic states and then describe how the true traffic states were determined from the loop detector data and the DVP field data. Then the proposed state estimation model is described. We propose two models; one that considers two traffic states and one that considers three traffic states. These models are calibrated and then evaluated .

4.5.1 Defining the Traffic States

Traffic states represent a region on the speed/volume/density space. Given that we are primarily interested in knowing if congestion exists at the Bluetooth detector location, we initially define three traffic states, namely the Uncongested state, the Transition state, and the Congested traffic states and need to label the true traffic states for each time interval from the empirical data set. We carry out this labeling as follows.

We calibrated the Van Aerde macroscopic traffic flow model to the empirical loop detector data set following similar procedures as in Chapter 3. The Van Aerde traffic flow model defines the relationship between volume, speed, and density. Because loop detectors do not directly measure density, the density value of each interval was estimated from the occupancy and the effective vehicle length using the method proposed by Cassidy and Coifman (Cassidy and Coifman, 1997). Using two weeks of aggregated 1-minute level VDS detector data compiled from the DVP/Bloor northbound and the DVP/Danforth southbound stations as inputs, the calibration results of Van Aerde model for the free-flow speed (v_f), the speed at capacity (v_c), the capacity (q_c), and the jam density (k_j) are provided in Table 4-3.

Table 4-3 Van Aerde model calibration results at DVP/Bloor NB

Model Parameter	Value and Unit
Free-flow speed, v_f	100.1 (km/h)
Speed at capacity, v_c	79.0 (km/h)
Jam density, k_j	95.0 (vehicle/km/lane)
Capacity, q_c	1550 (vehicle/h/lane)

Figure 4-10 shows the calibrated Van Aerde model diagram together with VDS loop measurements. The data are obtained from a two week period from two sample VDS stations.

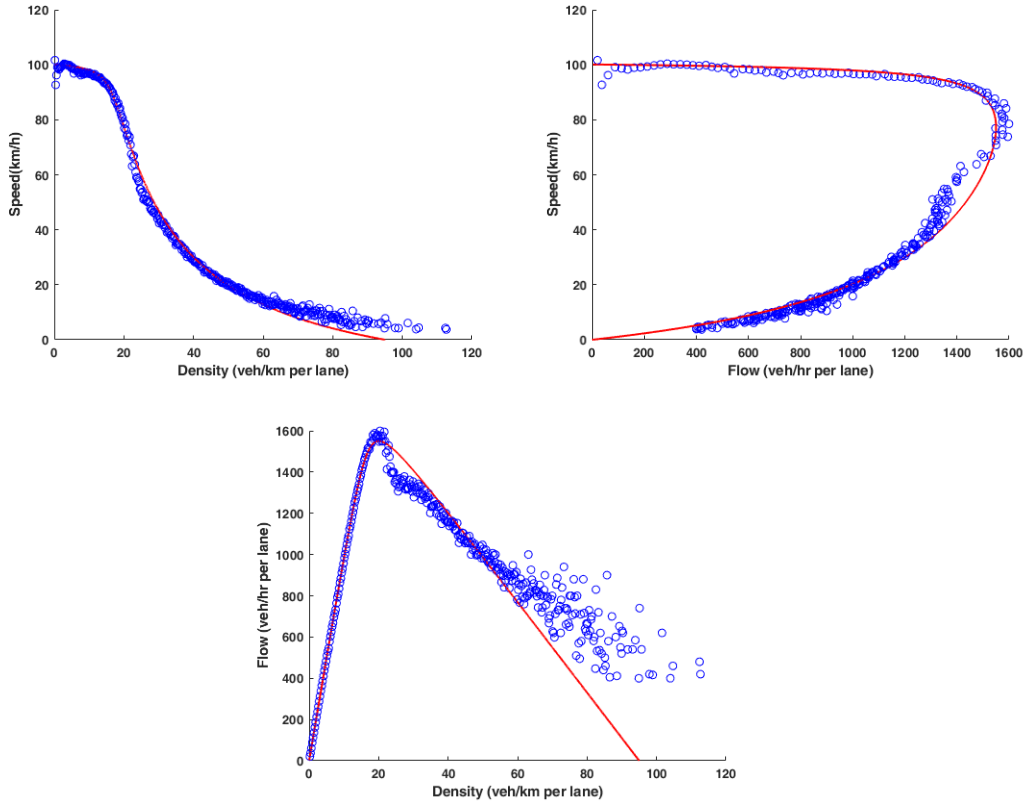


Figure 4-10 Van Aerde model calibration results

The distribution of VDS measurements in Figure 4-10 clearly indicate that the aggregated DVP data provides both congested and non-congested conditions. For each 1-minute interval, the associated traffic states were determined using the definition proposed by Reza Noroozisanani (Noroozisanani, 2017) as follows:

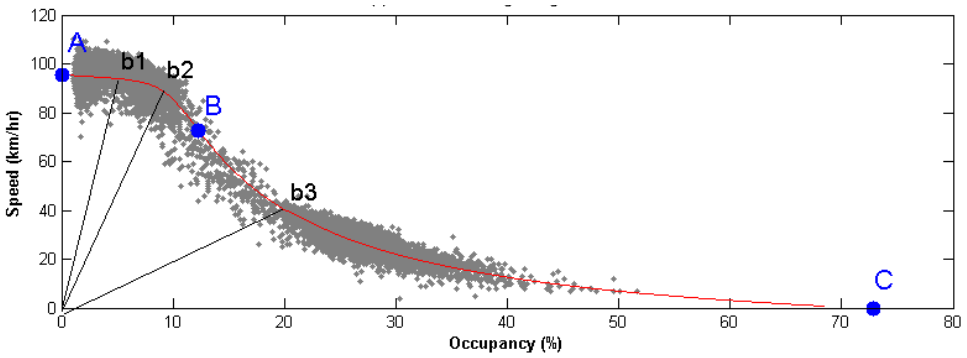


Figure 4-11 Determination of traffic states on Van Aerde's model (Figure by Reza Noroozisanani, 2017)

In Figure 4-11, point A represents the free-flow condition, point B represents the capacity, and point C represents the congested condition (jam density). Additional points b1 and b1 divides the occupancy (traffic density in this research) into three equal bins between the free-flow (point A) and the capacity (point B), point b3 divides the speed into two equal bins between the capacity (point B) and the jam density (point C). Then, the section on Van Aerde's model between A and b2 represents the free-flow state, the section between b2 and b3 represents the transition state, and the section between b3 and C represents the congested state. This state definition method accounts for the fact that the traffic conditions are not stable around the capacity (Noroozisanani, 2017), and were consistent with the distribution of the aggregated VDS measurements.

However, not all of the VDS measurements locate directly on the calibrated Van Aerde model curve in Figure 4-10 (the fitted curved represented with the red line and VDS observations represented with blue circle). In order to label the ground-truth traffic stream states using the proposed definition, each observation (i.e., the raw aggregated 1-minute VDS loop measurements) was mapped to the nearest point on Van Aerde's model using the following distance metric:

$$Distance_{i,m} = \min\left(\sqrt{\left(\frac{v_i - v_m}{v_f}\right)^2 + \left(\frac{k_i - k_m}{k_j}\right)^2 + \left(\frac{q_i - q_m}{q_c}\right)^2}\right) \quad (4\text{-}8)$$

Where v_i , k_i and q_i are the speed, density, and volume for observation i that is to be mapped; v_m, k_m, and q_m are associated with the a data point m on the fitted Van Aerde's model, the speed resolution is 0.1 km/h for generating these data points m; $Distance_{i,m}$ is the minimum distance value for observation i obtained at point m on Van Aerde's curve. The values of speed, density, and volume of point m (closest to the individual measurement i) on Van Aerde's curve are used in labeling the traffic states for observation i.

4.5.2 Feature Variables

A total of six metrics was computed across 5 different temporal aggregations periods (1-minute, 2-minute, 3-minute, 4-minute, and 5-minute interval) to create a set of 30 feature variables:

1. The average duration of a Bluetooth enabled vehicle's exposure to a Bluetooth detector (Average Dwell Time)
2. The average number of repetitive detections for Bluetooth MAC addresses (Average Hit)
3. The standard deviation of Hits (Standard Hit)
4. The standard deviation of Dwell Time (Standard Dwell Time)
5. The number of unique MACs within the aggregation interval duration
6. The number of total Hits within the aggregation interval duration

Aggregating data over different interval durations incorporate the changes in non-lagged measurements over near recent time intervals (representing the changes in traffic conditions) into the chosen classifier. In addition to the 30 feature variables identified above, two additional feature variables namely (1) An index to represent time of day (an integer variable defining the 1-minute time index where the value of this variable ranged from 300 (for 5 AM) to 1319 (for 10 PM), and (2) a binary variable representing weekday (value = 0) or weekend (value = 1) were also adopted.

4.5.3 Proposed Two State Model

The determination of traffic streams states from non-lagged measurements is modeled as a classification problem. The RUSBoost (Seiffert et al., 2010) classification tree algorithm is chosen in this study in examining the relationship between non-lagged measurements and traffic stream states.

The reasons for choosing the classification tree based ensemble algorithm RUSBoost are as follows:

- The analysis in the previous sections reveals that the traffic stream states (represented by speeds) and non-lagged measurements follow non-linear relationships (Equations (4-3) and (4-5)), and the ensemble of classification trees are capable of extracting the non-linear relationship between the input feature variables and the output class labels.
- No prior assumptions about the distribution of input feature variables are made in tree-based classifiers.
- The standard deviations in the input feature variables are calculated across different time interval durations, and it is possible some interval durations are associated with no or only one detection, thus the standard deviation cannot be calculated. Tree-based algorithms can properly handle the missing feature variable values in the input data by performing classification without using these missed features.

- The ratio between congested and uncongested intervals is 1:5 for the DVP data set. Such imbalanced classes in the dataset can be properly handled by the RUSBoost algorithm.
- The RUSBoost algorithm has achieved superior accuracies compared to many other ensemble techniques in real-world applications (Seiffert et al., 2010).

The input into the RUSBoost classification tree includes all the 32 feature variables defined in section 4.5.2, and the trained classifier predicts two classes including non-congested state and congested state. The free-flow state and the transition state defined in section 4.5.1 are combined into the non-congested state, and the definition of the congested state is the same as in section 4.5.1.

4.5.3.1 Model Parameter Calibration

Bluetooth measurement and aggregated VDS data in February were compiled from two locations on DVP for training the classification model. After removing the invalid rows, 9690 1-minute intervals were included in the training dataset.

The RUSBoost trains decision trees consecutively. When training N decision trees within the RUSBoost tree space, a total of N iterations is required, and the weights of input samples are modified in each iteration: after the training of one classification tree t in iteration t, the weight of correctly classified samples are reduced for iteration $t + 1$, while the weights of wrongly classified samples are increased for iteration $t + 1$. It is important to determine the number of iterations (trees) for the RUSBoost tree algorithm.

The misclassification rate, which is defined as the percentage of training samples that were incorrectly classified, is used as the performance measurement, the 10-fold cross-validation results associated with a different number of trees in the model are summarized as follows.

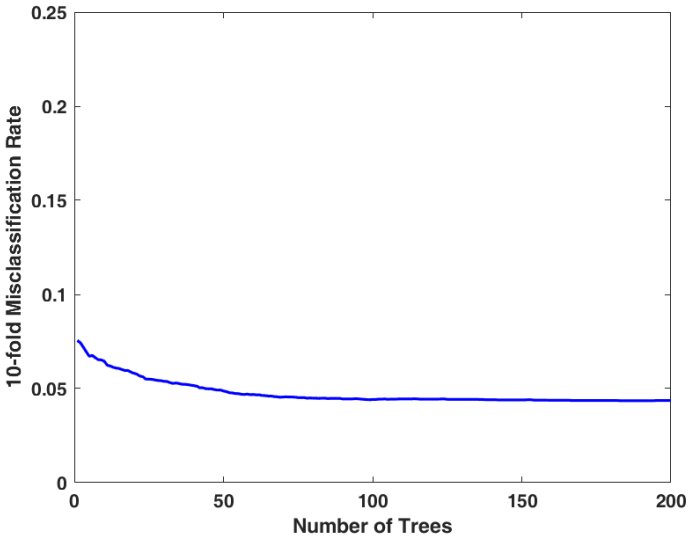


Figure 4-12 Determining the number of iterations (trees) for two states RUSBoost classification tree algorithm

From the results in Figure 4-12, the model misclassification rate remains constant at approximately 4% after 100 iterations. It is thus determined to choose 100 iterations (trees) for the two state RUSBoost classification tree in this study.

The results of the calibrated RUSBoost classification tree on the training dataset are summarized as follows:

Table 4-4 The two state RUSBoost classification tree results on the training dataset

Classification Results Category	Number of Intervals	Classification Accuracy (%)
Non-Congested State	8892	96.1%
Congested State	798	91.6%
Overall	9690	95.7%

The results in Table 4-4 show that the calibrated classifiers achieved overall 95.7% accuracy in predicting the non-congested and congested traffic states for the training dataset.

4.5.3.2 Model Evaluation

A different week of Bluetooth and VDS measurement were compiled from the same locations on DVP for evaluating the trained model performances. A total of 7650 intervals was included in the testing dataset. The calibrated classification model was applied to these data with the following results:

Table 4-5 The two state RUSBoost classification tree results on the testing dataset

Classification Results Category	Number of Intervals	Classification Accuracy (%)
Non-Congested State	6818	96.1%
Congested State	832	84.7%
Overall	7650	94.9%

The results in Table 4-5 show that the chosen classifiers achieved overall 94.9% accuracy in estimating traffic states for the testing dataset. Table 4-6 breaks down the model results in more detail.

Table 4-6 Confusion table for RUSBoost classification tree results on the testing dataset

Predicted / Ground Truth	Number of Intervals Classified as Non-Congested State	Number of Intervals Classified as Congested State	Total Number of Intervals
Number of Intervals with Ground Truth Non-Congested State	6555 (96.1%)	263 (3.9%)	6818
Number of Intervals with Ground Truth Congested State	127 (15.3%)	705 (84.7%)	832

From the results in Table 4-5 and Table 4-6, the accuracy of the RUSBoost classification tree on predicting the non-congested intervals in the testing dataset is higher compared to that on congested intervals, but overall, the RUSBoost model achieved comparable accuracies for the two traffic states.

4.5.4 Proposed Three State Model

The RUSBoost classification tree algorithm is used in modeling the relationship between non-lagged Bluetooth measurements and traffic stream states. The inputs into the classification model include all the feature variables defined in section 4.5.2. The developed classifier in this section predicts three states including the free-flow state, the transition state, and the congested state. The definitions of these three states are the same as defined in section 4.5.1.

4.5.4.1 Model Parameter Calibration

Both Bluetooth measurements and the aggregated VDS measurements at the 1-minute level were compiled from DVP for training the three state classification model. A total of 9690 1-minute intervals were included in the training dataset.

Using misclassification rate as the measurement of performance, the number of trees needed for the classification model was determined, and the 10-fold cross-validation results for calibrating the number of trees are summarized in the following diagram:

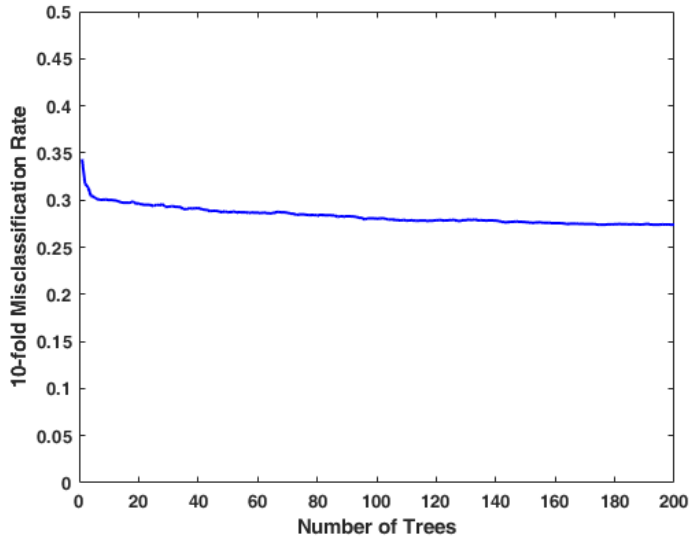


Figure 4-13 Determining the number of iterations (trees) for three states RUSBoost classification tree algorithm

Based on the classification results in Figure 4-13, it is determined to choose 150 iterations (trees) for the three state RUSBoost classification tree in this study, and the associated misclassification rate was around 28%.

Using 150 trees in the classifier, the results of the calibrated three state RUSBoost classification tree on the training dataset are summarized in the following table:

Table 4-7 The three state RUSBoost classification tree results on the training dataset

Classification Results Category	Number of Intervals	Classification Accuracy (%)
Free Flow State	5926	73.1%
Transition State	2966	62.0%
Congested State	798	90.5%
Overall	9690	71.2%

The results in Table 4-7 show that the calibrated classifier achieved overall 71.2% accuracy in predicting the traffic states for the training dataset, and such accuracy is significantly lower compared to the overall accuracy achieved by the two state RUSBoost tree classifier.

4.5.4.2 Model Evaluation

A different week of Bluetooth and VDS measurement were compiled from DVP for evaluating the trained model performances. A total of 7650 intervals are included in the testing dataset. Using the trained three state classifier, the results are as follows:

Table 4-8 The three state RUSBoost classification tree results on the testing dataset

Classification Results Category	Number of Intervals	Classification Accuracy (%)
Free Flow State	4513	65.7%
Transition State	2305	62.7%
Congested State	832	87.7
Overall	7650	67.2%

The results in Table 4-8 show that the chosen classifiers achieved overall 67.2% accuracy in estimating traffic states for the testing dataset. Table 4-9 breaks down the model results.

Table 4-9 Confusion table for RUSBoost classification tree results on the testing dataset

Predicted / Ground Truth	Number of Intervals Classified as Free-Flow State	Number of Intervals Classified as Transition State	Number of Intervals Classified as Congested State	Total Number of Intervals
Number of Intervals with Ground Truth Free-Flow State	2963 (65.7%)	1465 (32.5%)	85 (1.9%)	4513
Number of Intervals with Ground Truth Transition State	635 (27.6%)	1445 (62.7%)	225 (9.8%)	2305
Number of Intervals with Ground Truth Congested State	49 (5.9%)	53 (6.4%)	730 (87.7%)	832

From the results in Table 4-9, the accuracy of the RUSBoost classification tree on predicting the congested intervals in the testing dataset is higher compared to those on the other two states (free-flow state and transition state). However, the overall accuracy of the three state RUSBoost tree model is lower compared to the two state RUSBoost classification tree model developed in section 4.5.3.

4.6 Summary and Conclusions

This chapter focused on developing models to determine the traffic conditions (speeds or traffic states) on the basis of non-lagged Bluetooth measurements. The results suggest that:

1. The performance of the analytical model for estimating traffic stream speed on the basis of non-lagged Bluetooth detector measurements does not provide the level of accuracy that is expected to be needed to improve travel time estimation and is not recommended for estimating traffic speeds.
2. The performance of the RUSBoost classification tree achieved good overall accuracy (over 94%) in predicting traffic conditions as congested or non-congested (the two state model) based on non-lagged Bluetooth measurements. In addition, the RUSBoost classification tree achieved comparable accuracies in predicting the congested intervals and the non-congested intervals.
3. When modeling traffic conditions as three traffic states, the accuracies of the RUSBoost classification tree is lower overall compared to that of modeling traffic conditions as two traffic states. However, the three traffic state model has higher resolution in determining traffic states, as the transition state can be predicted in the three traffic state model. In addition, the three state model has higher accuracy in predicting the congested traffic state.

The next chapter examines how traffic state estimation improves the accuracy of travel time prediction models. The choice of the two state model versus the three state model is re-evaluated in the next chapter based on each model's impact on travel time prediction model results.

Chapter 5
Short-Term Travel Time Prediction

This chapter proposes methods for predicting the near future travel times on freeways using Bluetooth measurements. To address the time lag between the instant a vehicle entering the target freeway section and the actual instant the vehicle's travel time can be measured, the two traffic state estimation models developed in Chapter 4, namely the two-state model and the three-state model, are used in this Chapter. The information about the traffic conditions at locations near Bluetooth detectors enables the prediction model to more timely detect the changes in traffic conditions, thus achieving better prediction performance. In section 5.1, the Random Forest algorithm is adopted for training a prediction model based on historical Bluetooth measurements including both travel times and non-lagged measurements collected from an urban freeway; the objective of section 5.1 is on quantifying the impact of non-lagged Bluetooth measurements on travel time prediction results. In section 5.2, simulation data is generated for generalizing the impacts of non-lagged measurements on travel time prediction results as a function of Bluetooth detector spacing; the goal of section 5.2 is on recommending the scope of application for the proposed method.

5.1 Proposed Travel Time Prediction Method

The proposed travel time prediction model in this research is based on the Random Forest (Breiman, 2001) algorithm, which is an ensemble machine learning algorithm that can perform regressions based on outputs from decision trees within the trained model space. Random Forest has many applications in transportation engineering (Sun et al., 2018; Hou et al., 2017; Guo et al., 2017; Mendes-Moreira et al., 2012; Hamner, 2010), and some researchers applied Random Forest in performing highway and freeway travel time predictions (Guo et al., 2018; Fan et al., 2018; Semanjski, 2015; González-Brenes and Cortés, 2011). As indicated in these studies, the performance of the Random Forest (RF) models has been found to be comparable or superior to other machine learning algorithms.

The general procedures for performing travel time predictions using RF algorithms in this research are as follows:

1. Randomly select n instances from the historical data (size N) as the training set, $n < N$. Each instance $d_i$involves q feature variables $\{x_{i1}, x_{i2}, \dots, x_{iq}\}$ and the mean departure based travel time $DBTT_i$ associated with the prediction horizon. The candidate feature variables include historical mean arrival based travel times, historical upstream and downstream traffic states, day of week, and time of day.

2. Develop an unpruned regression tree on each group of selected n instances. This regression tree is grown with randomly selected q_{sub} feature variables from the original q variables in the feature vector, $q_{sub} \leq q$, many developed RF toolkit use the square root of q as the default value for q_{sub}. The split at each node is determined based on the best split from within q_{sub} variables, and the definition of the best split is the split associated with the minimum entropy.

3. Repeat above steps 1 and 2 by N_{tree} times. In the end, N_{tree} regression trees are established with each tree grown to the largest possible extent.

4. Perform travel time predictions by inputting the near recent mean arrival based travel times, near recent upstream and downstream traffic states, the day of week, and the time of day into N_{tree} regression trees, and the average value from all N_{tree} trees is the predicted travel time for the prediction horizon.

5.1.1 Model Structures

In order to quantify the impact of non-lagged measurements, models with and without non-lagged measurements as inputs were compared. The two-state and three-state classification models developed in Chapter 4 were used for inferring traffic conditions from non-lagged Bluetooth measurements.

Consider the following diagram for explaining the chosen feature variables for different models:

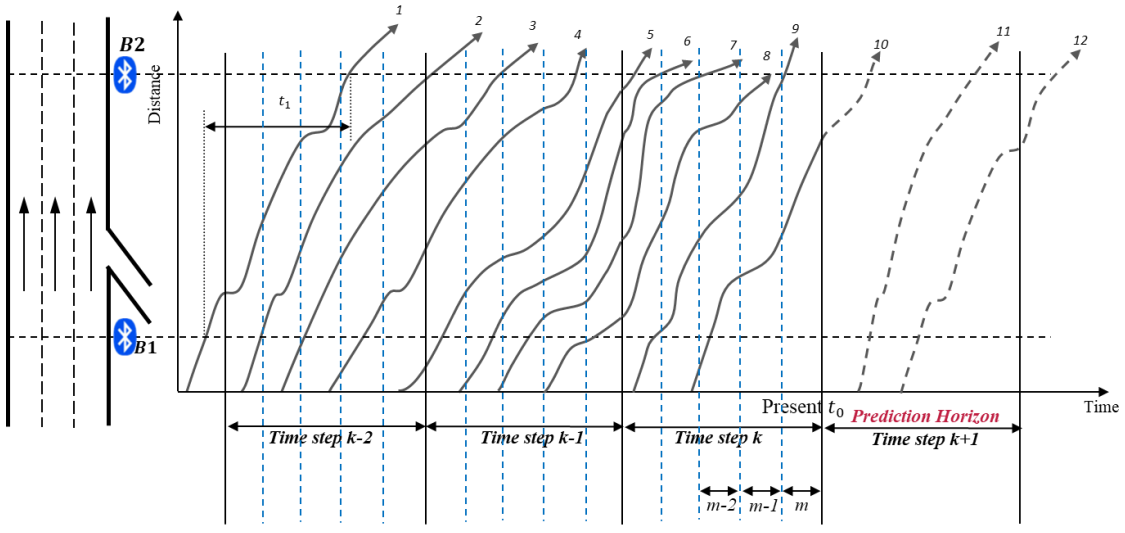


Figure 5-1 Time space diagram for defining the feature variables in different model structures

Assume each time step (e.g., $k, k-1, k-2$) is 5 minutes in Figure 5-1. The three groups of feature variables are summarized as follows:

- Group 1: Three most recent mean arrival based travel time measurements $\overline{ABTT}_k$, $\overline{ABTT}_{k-1}$, $\overline{ABTT}_{k-2}$, the time index m (measured as the number of minutes since midnight and therefore has a value ranging from 300 to 1319 or 5 AM to 10 PM) representing time of day and the binary variable representing weekday or weekend (0 for weekday, 1 for weekend).
- Group 2: All feature variables in Group 1, plus features that capture the current and near recent (estimated) traffic state at the downstream boundary of the roadway section; i.e. $DownState_m$, $DownState_{m-1}$, $DownState_{m-2}$, $DownState_{m-3}$, $DownState_{m-4}$.
- Group 3: All feature variables in Group 2, plus features that capture the current and near recent (estimated) traffic state at the upstream boundary of the roadway section; i.e. $UpState_m$, $UpState_{m-1}$, $UpState_{m-2}$, $UpState_{m-3}$, $UpState_{m-4}$

Because two traffic state estimation models, namely the two-state classification model and the three-state classification model, have been developed in Chapter 4, these two different traffic state estimation models are compared in section 5.1.4 to understand the impact of each model on travel time prediction performances. In addition, a naïve prediction model which uses the measurements from the near recent time interval as the prediction was also developed as a benchmark.

5.1.2 Field Data

Bluetooth measurements were collected from the Don Valley Parkway (DVP) in the City of Toronto for model calibration and evaluation. A 2.5-kilometer roadway section between the DVP/Beechwood_SB and the DVP/Danforth was chosen as the target roadway section (Figure 4-12).

Figure 5-2 Location of the chosen DVP roadway section (Image from Google Earth, red balloon represents the Bluetooth detector)

Data were obtained over the period from February 1, 2017 to March 31, 2017 with 59 valid days. The data obtained from the Bluetooth system consisted of raw detections and paired individual vehicle travel times. The vendor applies an outlier detection algorithm and tags all travel time observations deemed as outliers. The details of the filtering algorithm have not been made publically available. Only data that passed

the vendor's outlier detection algorithm were used. The resulting set of data contained 108,452 records of travel time observations.

5.1.3 Defining Truth

The objective of the travel time prediction model is to predict the average travel time that the population of vehicles in the traffic stream will experience. For model calibration and validation, it is necessary to have a measure of truth, where truth is the travel time without measurement error of all vehicles (i.e., no sampling error) in the traffic stream. However, the only measures of travel time that are available are those obtained from the Bluetooth detectors, and therefore they contain both measurement errors and sampling errors. In an attempt to reduce both of these errors, a moving median technique has been applied for establishing the true travel times for this study. The schematic representation of the moving median smoothing window is as follows:

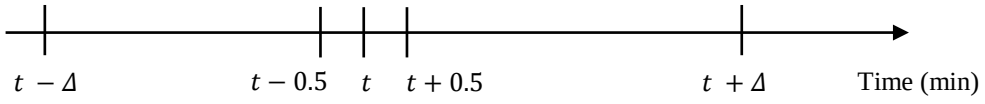


Figure 5-3 Schematic representation of the definition of the smoothing window

In Figure 5-3, the establishment of the ground truth for the 1-minute interval between $t - 0.5$ minute and $t + 0.5$ minute is based on the median value of all travel time measurements associated with vehicles entering the target segment (departure based travel time measurements) between $t - \Delta$ minute and $t + \Delta$ minutes, where the candidate values of Δ include 2.5, 5, or 7.5 minutes such that the time window over which the median is computed is 5, 10, or 15 minutes.

The impact of Δ is examined in the next section.

5.1.4 Model Calibration

For model structures defined in section 5.1.1, the number of decision trees needs to be determined for each model structure (model with different feature variables as inputs). In addition, the choice of using the two-state or three-state traffic state estimation model is also decided in this section based on travel time prediction model performances.

Two weeks of Bluetooth measurements from the chosen DVP site have been aggregated for model calibration. Since the true travel time is determined using the moving median technique, this study first examined the impact of the choice of Δ on the distribution of the number of raw Bluetooth travel time measurements. The cumulative distribution function of the number of travel time measurements in each

smoothing time window (5-minute window, 10-minute window, and 15-minute window) is summarized as follows:

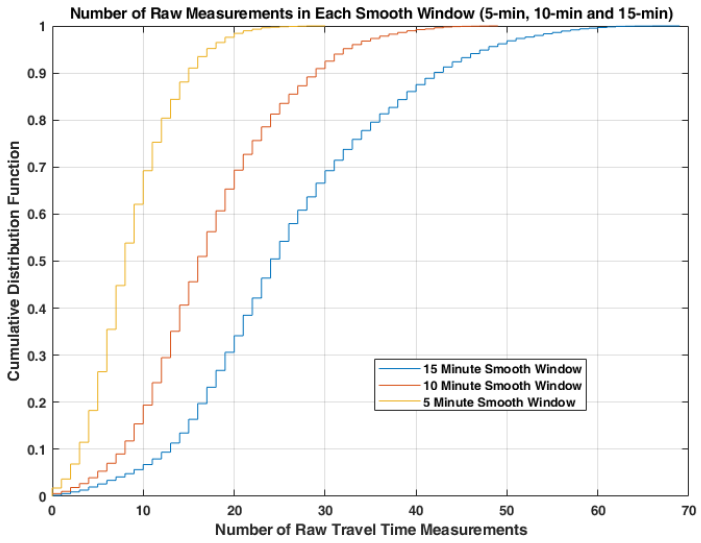


Figure 5-4 CDF plot for the number of raw travel time measurements as a function of the moving median smoothing window

In addition, the plot associated with individual vehicle travel times and the median travel time values over different smoothing time windows for one sample day between 7:00 AM and 12:00 PM is as follows:

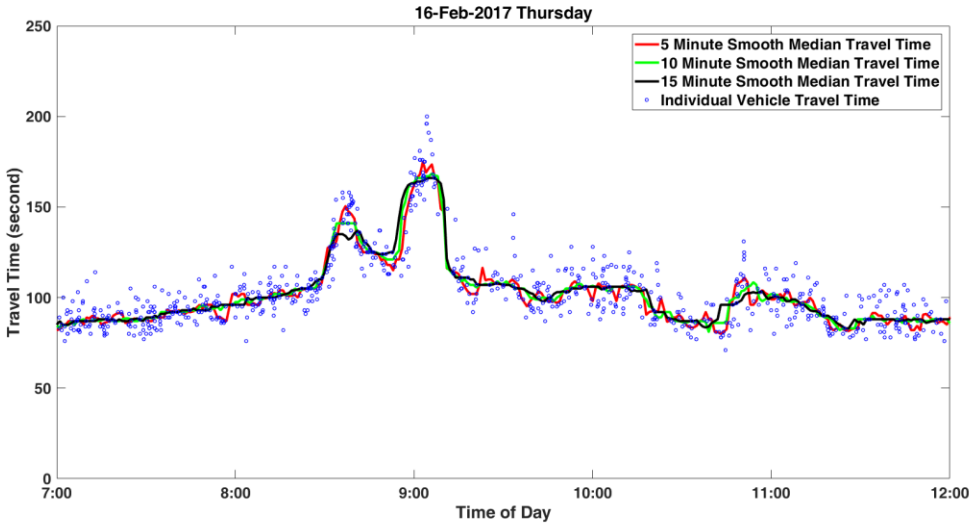


Figure 5-5 A sample plot with raw travel time measurements and smoothed median travel times

It can be observed from Figure 5-4 that as the smoothing window size increases, the number of observations increases. The main reason for adopting the moving median technique in determining the truth is to address the sampling error in Bluetooth measurements: the median calculated over a longer time

window increases the number of samples and reduces the sampling error. Figure 5-5 shows that increasing the smoothing time window duration leads to smoother travel time values between consecutive time intervals. Because over 85% of intervals contain more than 10 valid travel time measurements when the smoothing time window size is set at 10-minute (Figure 5-4), and no significant lag is observed with the 10-minute window compared to the 5-minute window (Figure 5-5), the 10-minute window is chosen for establishing the ground-truth of each prediction horizon.

The performances of different model structures are calibrated using the root mean square error (RMSE) of travel time prediction results. The definition of the measurements is as follows:

$$RMSE_d = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left(T_k - \hat{T}_{d,k}\right)^2} \tag{5-1}$$

Where k is interval index ($k = 1,2,\ldots,n$), T_k is the true travel time defined from the median value associated with interval k, $\hat{T}_{d,k}$ is the predicted travel time for the time interval k using model structure d, d represents the travel time prediction models with different feature variables as inputs as well as the naïve prediction model, $RMSE_d$ is the root mean square error associated with the prediction model structure d.

Using different feature variables as inputs, the model performances associated with different model structures using different numbers of trees are summarized in the following diagram.

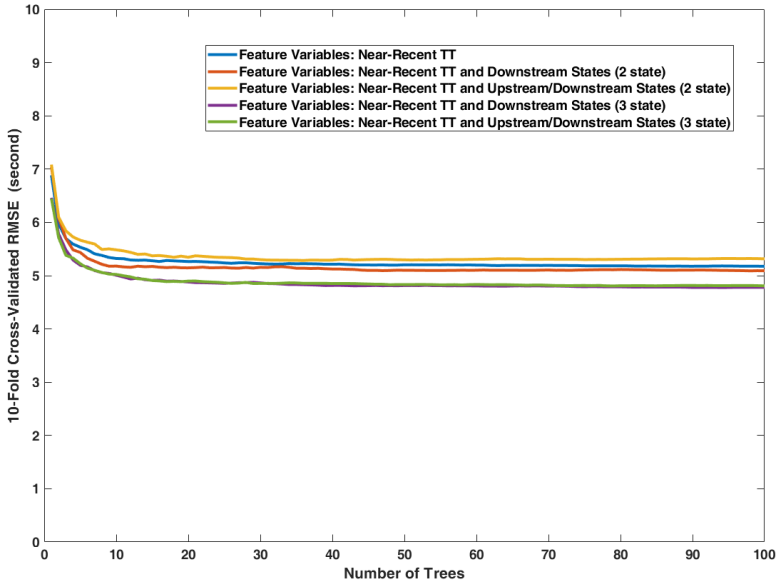


Figure 5-6 Prediction error as a function of the number of trees for the Random Forest model with different input feature variables and traffic state classification models

From the results in Figure 5-6, we can observe that RMSE values stabilized after around 50 trees and therefore 50 trees are chosen for all model structures in this study. In addition, the best cross-validation result is achieved using the three state estimation model outputs in the travel time prediction method. Therefore, the three state traffic state estimation model developed in Chapter 4 is chosen for estimating traffic states. Note that the model structure using the two state model and group 2 feature variable inputs (i.e., near recent travel time and downstream traffic states) achieved similar results as the model structure without traffic state as inputs. For the other model structure using two state model with group 3 feature variables as inputs (i.e., near recent travel time, upstream estimated traffic states and downstream traffic states), the associated RMSE value is even higher compared to the model without states as inputs. One possible explanation is that the traffic state estimation results using the two state model was not able to adequately reflect the changes in traffic conditions (and the state estimation results were incorrect for some intervals). This result highlights the importance of model calibration in determining the proper form of travel state inclusion for travel time prediction models.

Note that the importance of all the feature variables considered in different model structures have been examined using the permutation of the feature variable values on out-of-bag (OOB) samples: for each trained tree $tree_i$, only samples not used in training the corresponding $tree_i$ were evaluated to understand feature variable importance (Breiman, 1996). Since all feature variables were found to contribute to the improvement of travel time prediction, all proposed feature variables for each model structure were retained in the final model structures.

5.1.5 Model Evaluation - Empirical Data

A different two weeks of DVP data have been compiled for evaluating the performance of different model structures. The root mean square error (RMSE) and the absolute relative error (ARE) of travel time prediction results are used in evaluating model performances; the definition of ARE is provided as follows.

$$ARE_{d,k} = 100 * \frac{\left|T_k - \hat{T}_{d,k}\right|}{T_k} \tag{5-2}$$

Where $ARE_{d,k}$ is the absolute relative error of prediction model structure d in time interval k. The definitions for other notations are the same as defined for the equation (4-1). Note that the standard deviation of ARE and the 90th percentile of ARE are used for evaluating the variability and extremes of errors for different model structures, respectively.

The travel time prediction results for the naïve prediction model, the proposed different model structures over all time intervals and the changing periods in the testing dataset are summarized in the following table.

The definition of the "traffic condition is changing" is when either downstream or upstream states is determined to be in the transition state.

Table 5-1 Performance comparison of different model structures over different time periods

Time Periods	Description		RMSE	90th Percentile ARE	Standard Deviation of ARE
All Day	Naïve Prediction Model		6.0 second	7.7%	4.2%
	Group 1: Model with only travel time as inputs		5.6 second	6.8%	3.8%
	Group 2: Model with travel time and downstream traffic states		5.5 second	6.8%	3.7%
	Group 3: Model with travel time, upstream and downstream traffic states		5.5 second	6.7%	3.7%
	Comparisons	Difference between Group 1 and Naïve Model	6.1%	11.7%	9.1%
		Difference between Group 2 and Naïve Model	8.1%	11.7%	11.7%
		Difference between Group 3 and Naïve Model	8.1%	12.1%	11.9%
Traffic Conditions Changing	Naïve Prediction Model		9.1 second	9.3%	6.2%
	Group 1: Model with only travel time as inputs		8.4 second	8.3%	5.5%
	Group 2: Model with travel time and downstream traffic states		8.0 second	8.3%	5.2%
	Group 3: Model with travel time, upstream and downstream traffic states		8.0 second	8.2%	5.2%
	Comparisons	Difference between Group 1 and Naïve Model	8.3%	9.9%	11.3%
		Difference between Group 2 and Naïve Model	11.7%	10.7%	15.6%
		Difference between Group 3 and Naïve Model	11.8%	11.2%	15.6%

It can be observed from results in Table 5-1 that the model structures with traffic states (model group 2 and model group 3) as inputs achieve better results in terms of lower root mean square errors, lower 90th percentile ARE, and lower standard deviations of ARE compared to other model structures without traffic states as inputs. In addition, during time periods when traffic conditions are changing the models including traffic states as inputs achieved larger improvements compared to the naïve model; this result confirmed the hypothesis that the traffic states information are most relevant for travel time prediction system when traffic conditions are changing.

It is also observed that the overall RMSE is relatively small for all model structures (around 6 seconds). It might be tempting to conclude that though the proposed models provide improvements over the naïve prediction model, the absolute magnitude of these improvements is sufficiently small as to be of no practical consequence. However, this roadway section is relatively short (2.5 km), and the RMSE results are average

over many time intervals. The errors in individual time intervals can be much larger. Figure 5-7 illustrates the true and predicted travel times for one sample day between 7:00 AM and 12:00 PM. Additional travel time prediction results are provided in Appendix A.

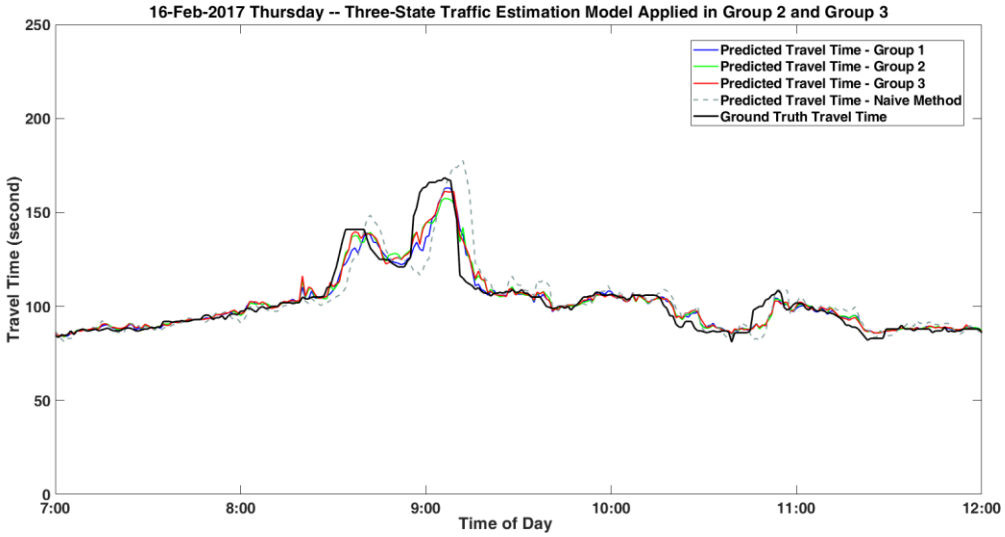


Figure 5-7 Travel time prediction results for February 16, 2017 (7:00 AM to 12:00 PM)

In the next section, a simulation model is built for understanding the impacts of traffic states on travel time prediction performances under roadway sections of varying lengths.

5.2 Evaluating the Impact of State Estimation on Travel Time Prediction for Different Bluetooth Detector Spacing Using Simulation

In the previous section, the empirical Bluetooth data has been collected for calibrating and evaluating travel time prediction models. The results in the previous section showed that the models with traffic states information as inputs could achieve better travel time prediction results compared to model structures without such inputs, but the absolute magnitude of these improvements was small. Consequently, the extent to which traffic states can contribute to the improvement of the travel time prediction model is still not clear. In addition, the empirical data are always influenced by several sources of errors including measurement error, system error, and sampling error, which makes "ground-truth" travel time hard to evaluate.

To overcome these limitations, this section proposes to use simulation data for evaluating the impact of including traffic states on travel time prediction models under different Bluetooth detector spacing. Compared to the empirical data, the geometrical configurations in the simulation can be set up to enable the investigation of Bluetooth detector spacing. More importantly, the simulation data do not contain

uncontrolled outliers or missing data, and the ground-truth travel times and traffic states can be accurately estimated from the simulation. It is hypothesized that the impact of including traffic states in travel time predictions is larger for freeways sections where consecutive Bluetooth detectors are placed further apart but not too far because traffic state conditions can lead to more timely detection of changes in traffic condition.

5.2.1 Simulation Data Collection

Similar to the simulation data collection in Chapter 3, the microscopic traffic simulation software VISSIM version 7 (Vissim, 7) and Bluetooth detection simulation tool *BlueSynthesizer* (Masouleh and Hellinga, 2018) were used for generating the simulation data in this section.

In order to prepare the simulation network geography that is representative of the real world roadway conditions, we first examined a 48-kilometer Highway 401 WB section between Kitchener and Woodstock. In contrast to urban freeways such as the DVP examined in the previous section, the chosen Highway 401 section connects suburban areas with lower population densities. It is expected that when installing Bluetooth detection systems for the chosen Highway 401 section, the detector placement density should be much smaller compared to that on urban freeways. The roadway geographic characteristics of the chosen Highway 401 section is summarized as follows:

Table 5-2 Highway 401 WB between Kitchener and Woodstock

Index	Location	Type	Distance From the Nearest Ramp (km)	Distance Between Consecutive On-Ramps (km)	Distance Between Consecutive Off-Ramps (km)
1	401 WB and King on-ramp	On-Ramp	--	--	--
2	401 WB and New Dundee Rd	Off-Ramp	2	--	--
3	401 WB and Executive Place	On-Ramp	1.2	3.2	--
4	401 WB to Cedar Creek	Off-Ramp	6.9	--	8.1
5	Cedar Creek to 401 WB	On-Ramp	0.6	7.5	--
6	401 WB to Oxford Rd 29	Off-Ramp	16.3	--	16.9
7	Oxford Rd 29 to 401 WB	On-Ramp	0.4	16.7	--
8	401 WB to Dundas St	Off-Ramp	11.6	--	12
9	Dundas St to 401 WB	On-Ramp	1.1	12.7	--
10	401 WB to Towerline Rd	Off-Ramp	0.6	--	1.7
11	Towerline Rd to 401 WB	On-Ramp	1.1	1.7	--
12	403 WB to 401	On-Ramp	1.6	1.6	--
13	401 WB to Norwich Ave	Off-Ramp	1.5	--	4.2
14	Norwich Ave to 401 WB	On-Ramp	1	2.5	--
15	401 WB to Mill St	Off-Ramp	1.1	--	2.1
16	Mill St to 401 WB	On-Ramp	1.1	2.2	--

It can be observed from the above table that the distance between consecutive access points (on/off ramps) can vary significantly. When the roadway is close to cities, the distance between consecutive on/off ramps are in the range of 1.6 km to 2.5 km, such as the ramps with the index from 10 to 16. However, for Highway 401 sections traversing the rural areas, the distance between access points can increase to 8 km or even over 16 km. After rounding the distances in the above table to the nearest integer, the roadway lengths of 1 km, 2 km, 4 km, 6km, 8km, 12 km, and 16 km are examined in the simulated network.

A hypothetical roadway network consisting of one direction of a 3-lane freeway segment was simulated in VISSIM, the geographic configuration of the simulated network is as follows.

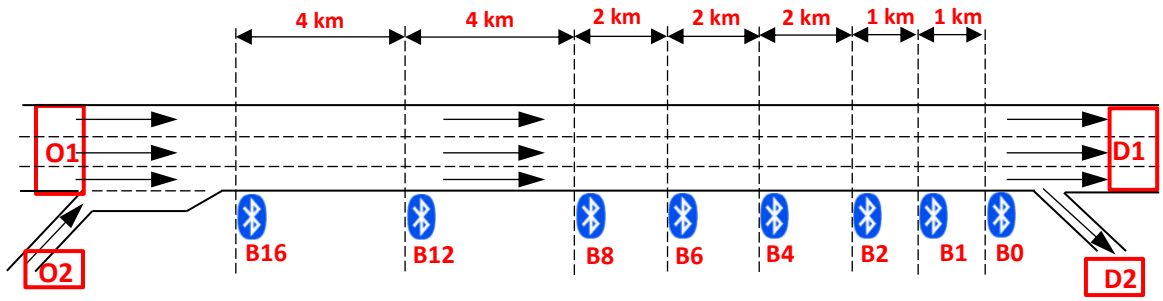


Figure 5-8 Geometric configuration for the simulation freeway roadway network

The VISSIM parameters were calibrated following the same procedures as proposed in section 3.4.1 of Chapter 3. Using two weeks of aggregated vehicle detector station (VDS) data compiled from Highway 401, VISSIM parameters for the simulation network were calibrated, and the values of these parameters are summarized in the following table.

Table 5-3 Calibrated Wiedemann 99 car-following model parameters

VISSIM Parameters	Value and Unit
CC0 (Standstill Distance)	5.93 meter
CC1 (Headway Time)	1.21 second
CC2 (“Following” Variation)*	4.00 meter
CC3 (Threshold for Entering “Following”)*	-8.00
CC4 (Negative “Following” Threshold)	-0.51
CC5 (Positive “Following” Threshold)	0.37
CC6 (Speed Dependency of Oscillation)*	11.44
CC7 (Oscillation Acceleration)	0.31 m/s^2
CC8 (Standstill Acceleration)*	3.50 m/s^2
CC9 (Acceleration with 80 km/h)	1.79 m/s^2

*Default value from VISSIM was retained

Roadway segments of different lengths are considered by pairing an upstream Bluetooth detector (B1, B2, …, B16) with Bluetooth detector B0. A range of traffic conditions is simulated by varying the traffic demands over time and by introducing incidents (simulated with the *reduced speed area* in VISSIM). Incidents only occur downstream of Bluetooth detector B0.

Traffic demands at the origins O1 and O2 were set to vary as a function the simulation time, and the duration of accidents vary between 15 to 50 minutes. In order to ensure that all congestion can dissipate before the end of each simulation run, each simulation run is set to last 3 hours without incidents occurring in the last 1 hour of each of the 3-hour simulation run. The following table provides the vehicle demands and simulation durations for one sample simulation run.

Table 5-4 Sample traffic demands and incidents duration setting for one simulation run

Time Interval (seconds)	O1 Inputs* (veh/h)	O1 V/C Ratio	O2 inputs* (veh/h)	O2 V/C Ratio
0-900**	2823	0.52	899	0.50
901-1800	2892	0.54	921	0.51
1801-2700	2229	0.41	710	0.39
2701-3600	4254	0.79	1310	0.73
3601-4500	4243	0.79	1304	0.72
4501-5400	4307	0.80	1343	0.75
5401-6300	5131	0.95	1906	1.06
6300-7200	4437	0.82	1426	0.79
7200-8100	2300	0.43	733	0.41
8100-9000	4490	0.83	1460	0.81
9000-9900	2935	0.54	935	0.52
9900-10800	2625	0.49	836	0.46
10800-11700	2963	0.55	944	0.52

Incident Index	Incident Durations+
1	2806 second to 5120 second
2	5512 second to 7758 second

*The percentages of HGV (heavy duty truck) are set as 10% for O1 and O2

**The initial 900-second is the warm-up period for the simulation and was not included in the analysis.

+Incidents decrease the traffic stream speeds to 15 km/h at the location of occurrence, which is downstream of B0

In the end, a full range of traffic operation conditions including both congested and uncongested conditions have been created in this study. The outputs from VISSIM including vehicle trajectories and vehicle travel times of all vehicles were collected. The vehicle trajectories were then inputted into *BlueSynthesizer* for simulating Bluetooth detections. The Bluetooth device penetration rate is set as 10%, because the associated Bluetooth sampling rate (measured by the percentage of vehicle population with valid Bluetooth travel time measurements) is around 6% under the 10% Bluetooth penetrate rate setting, and the value of 6% is consistent with Bluetooth sampling rate in the literature (Erkan and Hastemoglu, 2016; Sharifi et al., 2011).

5.2.2 Model Structures

Similar to section 5.1.1, three different groups of model structures, each of which involves different feature variables are adopted in this section. The Random Forest algorithm is chosen as the prediction algorithm for different groups of feature variables (representing model structures). These three groups of feature variables were aggregated for each of the candidate Bluetooth detector spacing (including roadway section lengths of 1 km, 2 km, 4 km, 6km, 8km, 12 km, and 16 km). Feature variables included in these three groups of model structures are summarized as follows:

- Group 1: Three most recent mean arrival based travel time measurements
- Group 2: All feature variables in Group 1, plus estimated near-recent downstream traffic states at 1-minute aggregation level for each of the last 5 minutes.
- Group 3: All feature variables in Group 2, plus estimated near-recent upstream traffic states at 1-minute aggregation level for each of the last 5 minutes.

In addition, a naïve prediction method, which takes the aggregated travel time measurements from the most recent 5-minute interval as the prediction for the future 5-minute interval, is used as the benchmark approach in this study.

The ground-truth travel times are estimated by averaging the travel time measurements from all the simulated vehicles entering the target roadway section in each time interval of interest. This allows accurate evaluations of different travel time prediction model structure performances.

5.2.3 Model Calibration

In order to calibrate the parameters for traffic state estimation model and for travel time prediction models (for different roadway section lengths), 60 hours of simulation data have been collected from the simulation network on all candidate roadway section lengths.

The three-state traffic state estimation model based on non-lagged Bluetooth measurements developed in Chapter 4 is re-calibrated in this section. It should be noted that the re-calibration of the traffic state estimation model is only necessary when working with different Bluetooth detector implementations (e.g., different antenna, different sensor, or different Bluetooth software stack). The ground truth traffic speeds are collected from VISSIM simulation network, and the same method proposed in section 4.5.1 of Chapter 4 is used to label the Uncongested state, the Transition state, and the Congested states for each time interval of interest.

Simulated Bluetooth detections together with the VISSIM speed measurements were compiled from the locations close to detector B0, B1, B2, B4, B6, B8, B12, and B16 in Figure 5-8. The same (as in the section 4.5.2 of Chapter 4) set of 30 feature variables computed across five different temporal durations (1-minute to 5-minute) was aggregated. However, because simulation data is examined in this section, two of the previously used variables including the time of day and the day of week cannot be compiled. A total of 3600 1-minute intervals of data were compiled for calibrating the traffic state estimation model.

Using the misclassification rate as the performance measurements, the 10-fold cross validation misclassification results associated with a different number of iterations (trees) in the RUSBoost tree model, which is the same model used in examining the relationship between non-lagged measurements and traffic stream states in section 4.5.4 of Chapter 4, are summarized as follows:

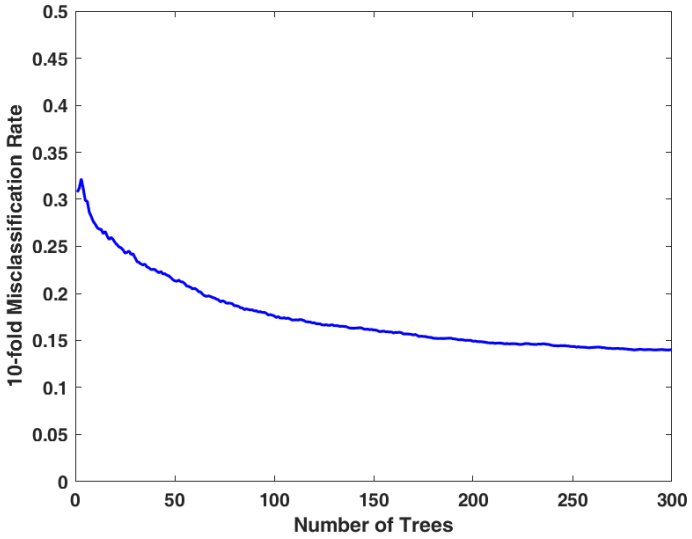


Figure 5-9 Calibrating the number of iterations for the three-state RUSBoost classification tree algorithm

Based on the results in Figure 5-9, 200 iterations are chosen for the three state model, as the performances stabilized after around 200 trees are used in the RUSBoost algorithm. After determining the number of iterations, a RUSBoost classification tree has been re-trained for estimating traffic states.

Aside from calibrating the traffic state estimation model, the number of trees needed in the Random Forest algorithm for travel time prediction is also calibrated for different section lengths. For each section length, three groups of feature variables (representing three model structures) have been examined. Using the root mean square error (RMSE) of travel time prediction results as the performance measurement, the maximum number of trees needed for the RMSE to stabilize across all section lengths under different model structures were determined to be 50. The decision of choosing a relatively large number of trees is because increasing the number of trees in the Random Forest algorithm can actually reduce the correlations among input feature variables while not negatively influence model performances. The following diagram summarizes the impacts of the number of trees on travel time prediction results of the 4-km section using different groups of feature variables.

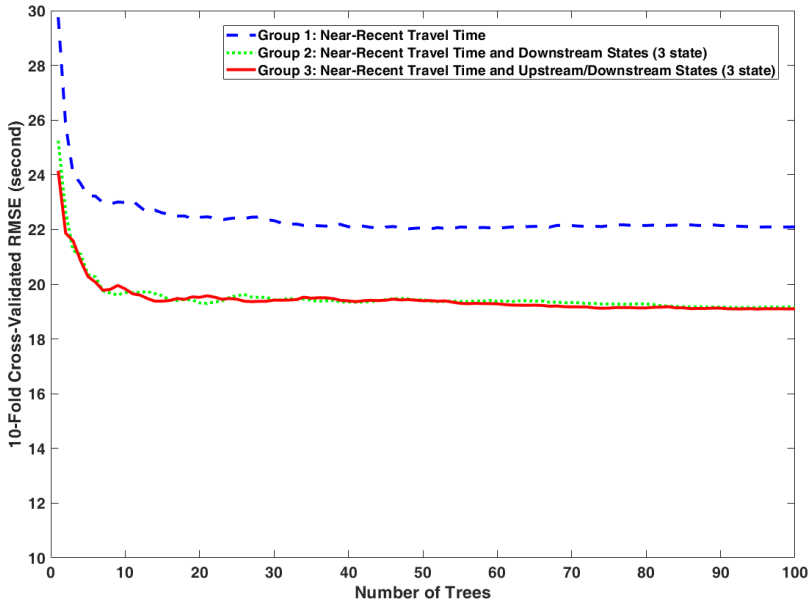


Figure 5-10 Calibrating the number of trees for the Random Forest algorithm

5.2.4 Model Evaluation - Simulation Data

In order to evaluate the impacts of traffic state estimations on travel time prediction results under different Bluetooth detector spacing. An additional 30 hours of data have been compiled for each section length to evaluate the performance of the proposed methods.

The following diagrams are the prediction results associated with one testing simulation run for different traffic state prediction model structures under different segment lengths. The plots in Figure 5-11 show (i) the predicted travel time for each of the four prediction models; (ii) the speed at the upstream and downstream segment boundaries as reported by the simulation model; and (iii) the traffic states estimated by the three state estimation model for the upstream and downstream boundary.

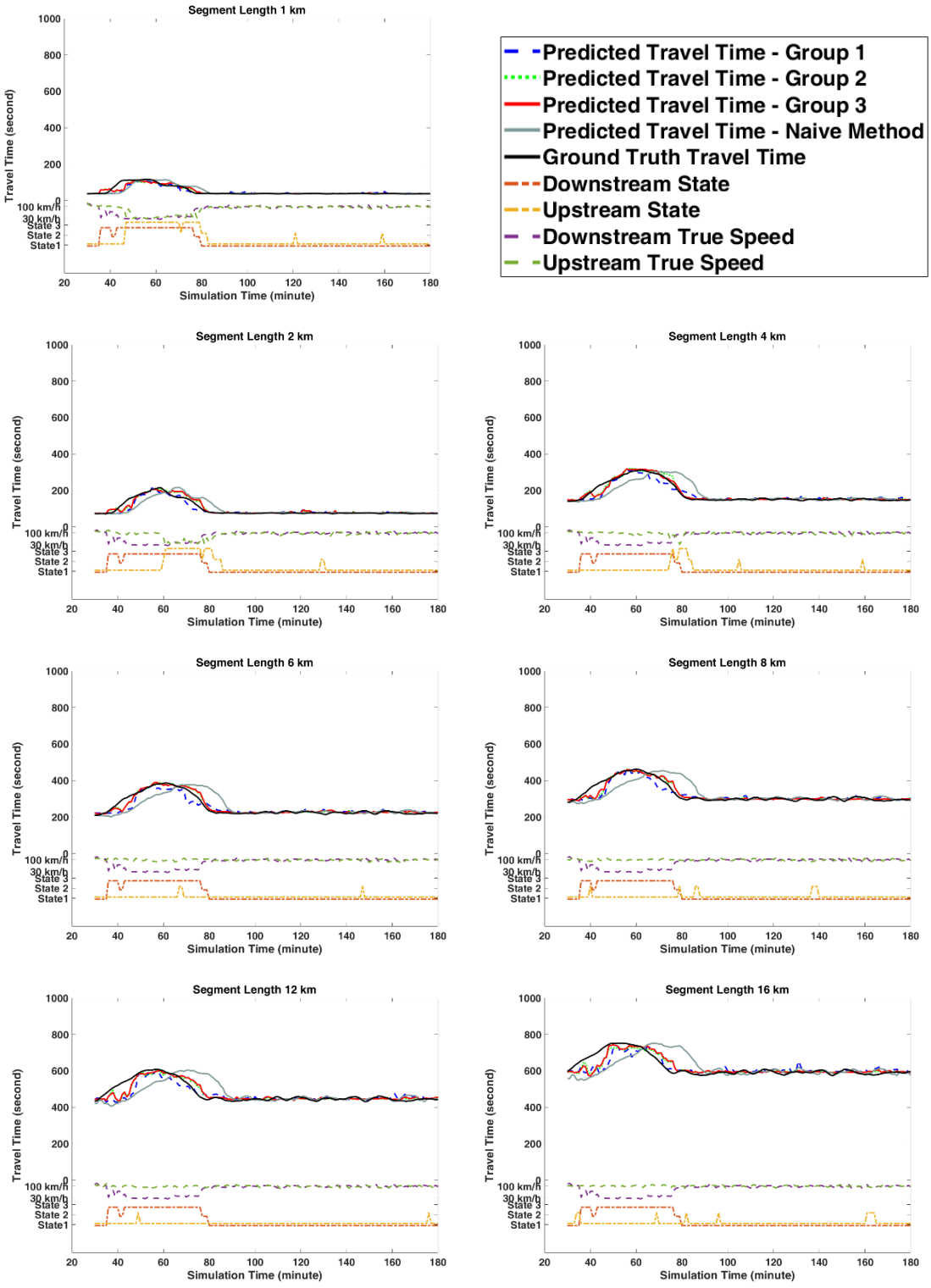


Figure 5-11 Travel time prediction results for one simulation run associated with different section lengths

Two main observations can be made from Figure 5-11. First, as the length of the freeway segment increases, the difference between the naïve model result and the ground truth became increasingly large, especially for periods when traffic condition are changing or are congested (e.g., between 40 to 80 simulation minutes); during these time periods, the proposed model structures (group 1, group 2, group 3) can better predict the travel time. Second, compared with the ground-truth speed measurements at locations close to Bluetooth detectors, the estimated traffic state results using the proposed three state model based on the non-lagged Bluetooth measurements is able to reflect the traffic conditions changes at the corresponding locations properly. For model structures using the state information including group 2 and group 3, some of the better travel time prediction results against the group 1 model could be ascribed to the use of state information, as other inputs into the Random Forest model are the same among these three model structures. Examples of such periods include on the segment of 4 km, 6km, 8km, and 12km between 60 to 80 simulation minutes. Additional prediction results are provided in Appendix B.

The root mean square error (RMSE), the 90^{th} percentile ARE, and the standard deviation of ARE were used in evaluating the model performances in terms of mean errors, extreme errors, and the reliability of the results, respectively. The travel time prediction results across different section length using different feature variables as inputs against the benchmark method (i.e., naïve predictions) for all time intervals as well as for when traffic conditions are changing are summarized in the following table. Note that the definition of the "changing periods" is based on the VISSIM speed measurements: either at the downstream or the upstream location, the associated traffic states is in the transition state.

Table 5-5 Travel time prediction results for the testing dataset using different model structures - All intervals

Performance Measurement		1 km	2 km	4 km	6 km	8 km	12 km	16 km
RMSE Naïve Model (second)		15.9	29.6	36.9	40.7	44.6	52.1	59.5
RMSE Group 1 (second)		11.7	17.4	23.7	26.6	30.2	35.5	42.1
RMSE Group 2 (second)		9.6	13.7	17.0	19.6	22.7	27.7	33.6
RMSE Group 3 (second)		9.7	13.9	17.0	19.3	23.0	27.9	33.8
RMSE Comparisons*	Group 1 vs Naïve	26.5%	41.0%	35.8%	34.7%	32.4%	31.8%	29.2%
	Group 2 vs Naïve	**39.5%**	**53.5%**	53.8%	51.8%	**49.2%**	**46.8%**	**43.5%**
	Group 3 vs Naive	39.0%	52.8%	**54.0%**	**52.6%**	48.4%	46.3%	43.2%
90th Percentile ARE Naïve Model (%)		38.0%	36.1%	30.1%	25.1%	22.1%	19.2%	17.6%
90th Percentile ARE Group 1 (%)		28.3%	24.1%	18.7%	16.1%	15.2%	12.7%	11.6%
90th Percentile ARE Group 2 (%)		20.4%	17.4%	12.0%	11.0%	10.6%	9.6%	9.2%
90th Percentile ARE Group 3 (%)		21.5%	17.2%	12.2%	10.9%	10.9%	9.4%	9.2%
90th Percentile ARE Comparisons*	Group 1 vs Naïve	25.5%	33.3%	38.0%	36.0%	31.2%	33.9%	34.3%
	Group 2 vs Naïve	**46.2%**	51.8%	**60.0%**	56.2%	**51.9%**	50.1%	**47.6%**
	Group 3 vs Naive	43.6%	**52.4%**	59.3%	**56.6%**	50.7%	**50.9%**	47.5%
Standard Deviation of ARE Naïve Model (%)		23.5%	22.3%	15.1%	11.9%	10.2%	8.1%	7.0%
Standard Deviation of ARE Group 1 (%)		13.3%	10.5%	8.5%	7.0%	6.1%	4.9%	4.5%
Standard Deviation of ARE Group 2 (%)		12.2%	8.3%	6.0%	5.1%	4.6%	4.0%	3.7%
Standard Deviation of ARE Group 3 (%)		12.6%	8.6%	5.9%	5.0%	4.7%	4.0%	3.8%
Standard Deviation of ARE Comparisons*	Group 1 vs Naïve	43.4%	52.7%	43.9%	41.1%	39.8%	39.5%	35.8%
	Group 2 vs Naïve	**48.2%**	**62.8%**	60.4%	57.2%	**54.8%**	**51.1%**	**46.6%**
	Group 3 vs Naive	46.5%	61.4%	**60.7%**	**58.2%**	54.1%	50.9%	46.6%

*The best comparison results are highlighted in bold

Table 5-6 Travel time prediction results for the testing dataset using different model structures-Transition state intervals

Performance Measurement		1 km	2 km	4 km	6 km	8 km	12 km	16 km
RMSE Naïve Model (second)		20.3	37.0	44.3	48.0	52.9	60.0	71.2
RMSE Group 1 (second)		12.5	18.5	26.3	29.0	31.4	33.0	39.7
RMSE Group 2 (second)		10.6	14.1	17.0	21.1	23.7	25.7	32.6
RMSE Group 3 (second)		10.8	14.4	16.9	20.7	23.6	25.6	32.6
RMSE Comparisons*	Group 1 vs Naïve	38.5%	49.9%	40.7%	39.7%	40.8%	45.1%	44.2%
	Group 2 vs Naïve	**48.1%**	**61.9%**	61.6%	56.1%	55.2%	57.2%	54.1%
	Group 3 vs Naïve	46.9%	61.1%	**62.0%**	**57.0%**	**55.4%**	**57.4%**	**54.2%**
90th Percentile ARE Naïve Model (%)		65.7%	70.6%	47.5%	36.8%	31.2%	26.3%	22.0%
90th Percentile ARE Group 1 (%)		39.6%	28.4%	22.9%	18.4%	16.7%	12.8%	11.5%
90th Percentile ARE Group 2 (%)		29.9%	19.0%	12.7%	13.4%	10.8%	9.2%	9.7%
90th Percentile ARE Group 3 (%)		30.9%	18.6%	12.7%	12.2%	10.9%	9.2%	9.2%
90th Percentile ARE Comparisons*	Group 1 vs Naïve	39.7%	59.8%	51.7%	49.9%	46.5%	51.3%	47.8%
	Group 2 vs Naïve	**54.5%**	73.1%	**73.3%**	63.6%	**65.3%**	**65.1%**	56.0%
	Group 3 vs Naive	53.0%	**73.6%**	73.2%	**66.9%**	65.0%	65.0%	**58.1%**
Standard Deviation of ARE Naïve Model (%)		32.9%	31.0%	19.3%	14.6%	12.5%	9.9%	8.6%
Standard Deviation of ARE Group 1 (%)		16.0%	12.3%	9.7%	7.8%	6.7%	4.8%	4.4%
Standard Deviation of ARE Group 2 (%)		15.4%	9.7%	6.5%	5.8%	5.1%	4.0%	3.8%
Standard Deviation of ARE Group 3 (%)		15.9%	10.2%	6.5%	5.7%	5.0%	4.0%	3.8%
Standard Deviation of ARE Comparisons*	Group 1 vs Naïve	51.4%	60.2%	49.8%	46.5%	46.4%	51.1%	48.8%
	Group 2 vs Naïve	**53.4%**	**68.7%**	66.1%	60.1%	59.5%	59.9%	56.0%
	Group 3 vs Naive	51.6%	67.1%	**66.4%**	**61.1%**	**59.7%**	**60.2%**	**56.1%**

*The best comparison results are highlighted in bold

The following observations can be made from Table 5-6 and Table 5-6.

1) Overall, all the proposed model structures (model group 1, 2, and 3) achieved better travel time prediction results compared with the naïve prediction model in terms of RMSE and 90th percentile ARE;

2) Overall, model structures including traffic states (i.e., feature group 2 and 3) achieve better travel time prediction results compared to models without traffic states (feature group 1) in terms of RMSE and 90th percentile ARE;

3) Compared to results associated with all the time intervals, the inclusion of traffic state information in the prediction model leads to larger travel time improvements when traffic conditions are changing.

The following plots show the impact of including the traffic state information on the accuracy of travel time prediction results as a function of the segment length based on RMSE for all intervals and for transition state intervals. The RMSE results have been normalized to be able to compare across different segment lengths by dividing the RMSE by the associated section length. This provides the measure of accuracy as the RMSE per kilometer (Figure 5-12 subplots (a) and (b)). Also, the percentage improvement in RMSE per kilometer between the group 2 model and the group 1 model, and between the group 3 model and the group 1 model are also provided in Figure 5-13. These results provide insights into the optimal segment length in terms of the improvement that traffic state information can contribute.

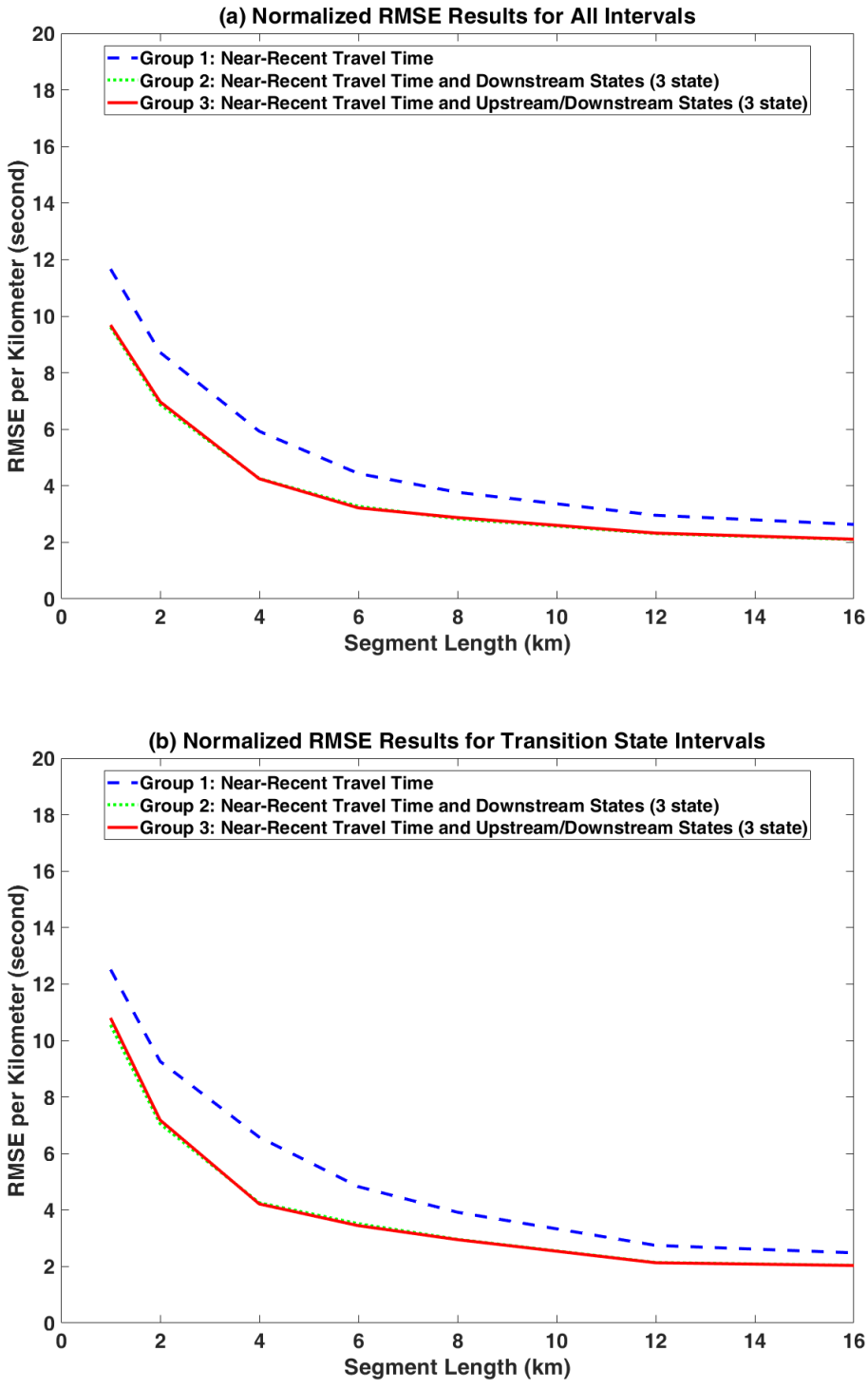


Figure 5-12 The impact of traffic state information inclusion on the normalized RMSE results as a function of the segment length

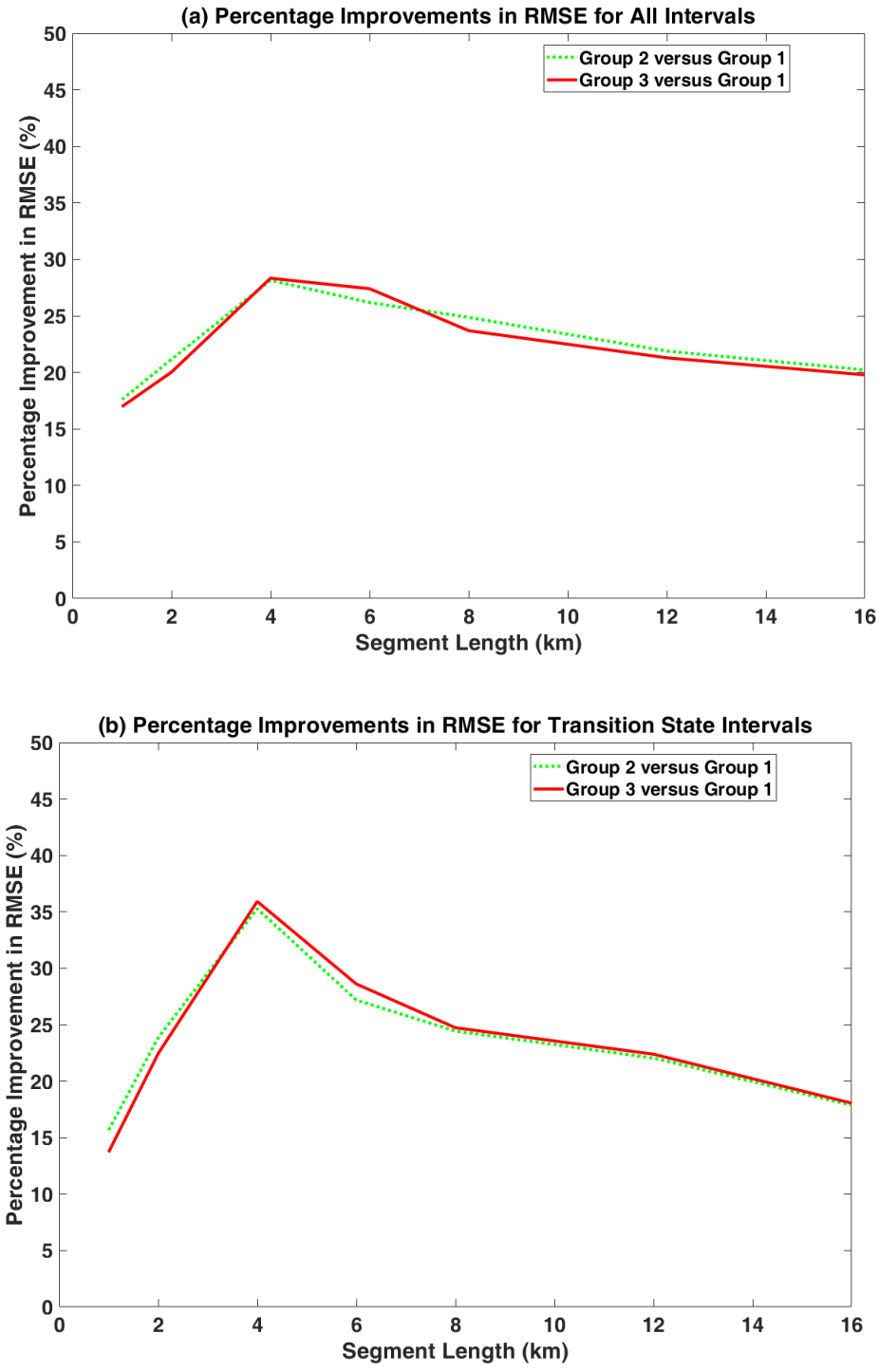


Figure 5-13 The impact of traffic state information inclusion on the percentage improvements of the RMSE results as a function of the segment length

In Figure 5-12, the performance difference between model structures including traffic states (i.e., model structure group 2 and 3) and the model structure without traffic states (i.e., group 1) are presented; the results show that the inclusion of traffic states in the travel time prediction model provides increased prediction accuracy for all segment lengths. Figure 5-13 shows that the maximum benefit associated with including the traffic states is associated with segments that are in the range of 4 to 8 km in length. These results are consistent with our expectations. When segments are short, the vehicle travel times will also tend to be short. Therefore, the time lag between conditions on the roadway changing and observing this through measured vehicle travel times will also be short. Consequently, we don't anticipate that having non-lagged information (i.e., estimated traffic states at the segment boundaries) will provide much additional value for the travel time prediction. Furthermore, when segments are very long, knowledge of the traffic states at the boundaries of the segment provides only limited information about the conditions that exist within the segment itself.

Considering the results in Table 5-6, Table 5-6, Figure 5-12 and Figure 5-13, the group 3 model structure, which includes near recent travel time, upstream traffic states, and downstream traffic states, is recommended for application. The reasons for recommending the group 3 model structure over other model structures include: 1) consistent better results compared to the Naïve model and the group 1 model structure under all intervals as well as the transition state intervals; 2) better stability when compared to the group 2 model structure in terms of lower standard deviation of ARE under transition intervals.

5.3 Summary and Conclusions

The conclusions of this chapter are summarized as follows.

First, this chapter quantifies the impacts of non-lagged measurements on travel time prediction model performances under different prevailing traffic conditions with empirical data. The results clearly indicate that the inclusion of traffic states can lead to better travel time prediction results under all time intervals; in addition, more significant improvements in travel time prediction accuracy were observed when the traffic conditions were changing.

Second, on the basis of the analysis described in this chapter, and in particular considering the travel time prediction accuracy and model stability, it is recommended that the developed three state estimation model combined with group 3 model structure be used.

Third, this chapter presented an evaluation of the impact of state estimation (on the basis of non-lagged Bluetooth detector measurements) on travel time prediction accuracy as a function of Bluetooth detector spacing. Simulation data was used for understanding the impact of Bluetooth detector spacing, as simulation methods allow the full control of Bluetooth detector location settings and provide the ground truth travel

time information for reliably evaluating model performances. The results showed that compared to shorter roadway segments, traffic state information generally leads to larger travel time improvements for longer (4 kilometers to 8 kilometers in length) freeway sections, which is likely caused by the fact that longer segments are more often influenced by larger travel time measurement lag. Also, the results showed that the improvements associated with using traffic state information are smaller on segments over 10 km in length compared to on segments of 4 km to 8 km in length. The possible reason for the reduced benefits on longer roadway segments is that the time it takes to detect shockwaves at Bluetooth detector locations (i.e., the boundary of the freeway segment) is comparable to the lag in obtaining the Bluetooth travel time measurements. Therefore, the segment length of 4~8 km is optimal in terms of the improvement from using state estimation in travel time prediction models.

Chapter 6
Conclusions and Recommendations

6.1 Conclusions

Short-term travel time predictions not only support passengers' route choice decisions but also serve as the basis for the implementation of proactive traffic management and control strategies. Over the last decade (2008 to present), Bluetooth technology has become increasingly popular in transportation projects as a viable source for travel time information. However, similar to many other automatic vehicle identification (AVI) technologies, Bluetooth technology has some limitations in measuring travel times including 1) Bluetooth detectors cannot distinguish the lane of travel for detected devices, therefore, Bluetooth detectors cannot provide facility-specific travel time information when measurements are collected from different traffic streams moving in the same direction; 2) the travel time measurements are associated with measurement lag, because the measurements associated with vehicles that have not yet reached the downstream Bluetooth detector cannot be taken at the instant of analysis.

In this thesis, methods have been developed to address the above challenges for the real-time prediction of short-term travel times on freeways. Methods were developed to estimate facility specific travel time in real-time; non-lagged Bluetooth measurements were used to infer traffic states at locations close to each Bluetooth detector; the inferred traffic states were used as inputs into travel time prediction models to allow more timely understanding of traffic condition changes on the target roadway. Bluetooth measurements are the primary source of data used in this research, but the proposed method can work with other data sources as long as the travel time and traffic state data can be estimated in real-time.

In this chapter, the summary of the main contributions from this research together with recommendations for further studies on the same subject are presented.

6.2 Major Contributions

The major contributions of this research concern five items relevant to the accurate prediction of short-term travel times on freeways. The details of these contributions are summarized as follows:

1. Developed two different methods for real-time estimation of the facility-specific travel times based on Bluetooth measurements. In this research, two new methods have been proposed for estimating travel times for each facility. The novelty of the proposed methods is using algorithms to determine the traffic conditions on each facility type instead of assuming that one of the facilities always operates at higher speeds than the other facility. Using an innovative simulation framework, the performance of the proposed methods were quantified against a benchmark method, which assumed

the HOV is always faster than the GPL. A sensitivity analysis was further performed to understand the performance of the different methods under different traffic conditions.

Based on the model results, one of the proposed method, which first clusters the travel time measurements using the K-means algorithm and associates the clusters with facilities using traffic flow model is recommended when the physical barriers or law enforcement prevent drivers from freely switching between the underlying facilities. However, when the roadway functions as a self-correcting system allowing vehicles to freely switching between underlying facilities, the benchmark method is recommended for application.

2. Developed algorithms for estimating the state of the traffic stream based on non-lagged Bluetooth measurements. This research identified a set of new non-lagged Bluetooth measurements that can be aggregated at fixed time intervals from Bluetooth detectors. Based on the understanding of Bluetooth detection process, an analytical model and statistical models have been proposed and calibrated for estimating traffic conditions close to each Bluetooth detectors using non-lagged Bluetooth measurements. The performance of the proposed models was demonstrated using field data, and the RUSBoost classification tree was recommended over the analytical model for estimating traffic states based on model results.

   Two variants of traffic state estimation models using the RUSBoost classification tree have been proposed. The results showed that the two-state model has higher overall accuracy in predicting traffic conditions as congested or uncongested; The three-state model has lower overall accuracy when compared to the two-state model, but has a higher resolution in determining traffic states, especially in determining the congested state.

3. Developed different model structures for the real-time prediction of short-term travel times on freeways. This research developed new data-driven travel time prediction models by merging outputs of the proposed traffic state estimation model (based on non-lagged Bluetooth measurements) with near-recent travel time measurements. The main innovation behind the proposed model structures is using traffic states inferred from non-lagged measurements to address the travel time measurement lag because the traffic states information enables the travel time prediction model to more timely detect changes in traffic conditions. The impact of traffic state inclusion on travel time prediction performances was quantified using filed data.

   After evaluating three groups of short-term travel time prediction model structures with different feature variable inputs based on the Random Forest algorithm, the overall model results showed that the model structures with traffic states as inputs consistently outperformed other model structures in term of better travel time prediction accuracies.

4. Quantified the impact of traffic states inclusion on travel time prediction accuracies as a function of roadway segment lengths. This research developed an innovative simulation framework, which allows the objective evaluation of model performances under different traffic condition with different geometric configurations. Using the simulation framework, the evaluation results quantitatively verified the contribution of including the inferred traffic state information in travel time prediction models under different segment lengths. The results showed clear evidence that the information of traffic states contributes to more significant improvements for longer freeway segments (4 km to 8 km) and during periods when traffic conditions are changing.
5. Recommended the model structure for incorporating non-lagged Bluetooth measurements in travel time prediction models. Based on the travel time prediction model results, this research recommends the usage of the three state estimation model over the two state estimation model in the travel time prediction model. In addition, this research recommends the Random Forest model based travel time prediction model structure with inputs of the near recent travel time measurements, the inferred three state results at the upstream detector location, and the inferred three states at the downstream detector location (if applicable, the time of day and day of week should also be included). When compared with other model structures examined, the recommended model structure achieved consistently better results in terms of mean errors, extreme errors, and model stability under all intervals.

6.3 Future Research

Based on the proposed methodologies and evaluation results, the following topics are recommended for improving and complementing this research:

1. The proposed methods for estimating the facility-specific travel time based on Bluetooth measurements were developed using simulation data because simulation data allows the access of ground truth travel time information under different traffic conditions for different facility types. Since the proposed methods are applicable to other situations in which multiple traffic streams co-exist, it is recommended that further research test the performance of the proposed methods for estimating facility specific travel times in a different context, other than the estimation for HOV and GPL travel time presented in this research.
2. Among the two proposed statistical model based traffic state estimation models, the three-state model was found to achieve more significant improvement in travel time prediction accuracies. The higher resolution of the three-state model in determining traffic conditions is likely the main reason for the better travel time prediction results. Therefore, it is recommended that further research consider integrating other sources of high-resolution data (e.g., traffic stream speeds

provided by connected vehicle technology) into travel time prediction models. In addition, it is recommended that further studies on inferring traffic states from non-lagged measurements consider other aggregation interval durations or other feature variables (e.g., the received signal strength indicator RSSI and its variants) not used in this research.

3. The calibrated traffic state estimation model is transferable to other sites using similar Bluetooth detector implementations. However, the re-calibration of the proposed traffic state estimation model is needed when the proposed methods are to be implemented with different Bluetooth detector software/hardware configurations.
4. The proposed travel time prediction models provide travel time prediction updates every 1-minute on each freeway segment (bounded by two consecutive Bluetooth detectors) with 5-minute prediction horizon. When predicting travel time for a longer route consisting of multiple shorter segments, it is recommended that a trajectory method (Izadpanah, 2010) be implemented. Note that with the implementation of the trajectory method, multiple prediction horizons are also likely involved; separate travel time prediction models should be calibrated for each of the freeway segment length and prediction horizon combination.
5. The performance of the proposed travel time prediction methods is empirically validated on a relatively short roadway section of 2.5 km because the chosen roadway section provides both Bluetooth measurements and VDS speed measurements, both of which are needed for model calibration purposes. It is recommended that the proposed methods be calibrated and validated on a different longer freeway section.

Bibliography

Anderson, T. W., & Darling, D. A. (1954). A test of goodness of fit. *Journal of the American Statistical Association, 49*(268), 765-769.

Araghi, B. N., Christensen, L. T., Krishnan, R., Olesen, J. H., & Lahrmann, H. (2013). *Use of low-level sensor data to improve the accuracy of bluetooth-based travel time estimation.*

Arshad, M., Rasool, M., & Ahmad, M. (2003). Anderson darling and modified anderson darling tests for generalized pareto distribution. *Pakistan Journal of Applied Sciences, 3*(2), 85-88.

B. Hamner. (2010). Predicting travel times with context-dependent random forests by modeling local and aggregate traffic flow. *2010 IEEE International Conference on Data Mining Workshops,* 1357-1359. doi:10.1109/ICDMW.2010.128

Bakula, C., Schneider IV, W. H., & Roth, J. (2011). Probabilistic model based on the effective range and vehicle speed to determine bluetooth MAC address matches from roadside traffic monitoring. *Journal of Transportation Engineering, 138*(1), 43-49.

Balakrishna, R. (2006). *Off-Line Calibration of Dynamic Traffic Assignment Models.*

Ben-Akiva, M. E., Bierlaire, M., Burton, D., Koutsopoulos, H. N., & Mishalani, R. (2002). Network state estimation and prediction for real-time transportation management applications. *Transportation Research Board 81st Annual Meeting.*

Ben-Akiva, M., Bierlaire, M., Bottom, J., Koutsopoulos, H., & Mishalani, R. (1997). Development of a route guidance generation system for real-time application. *IFAC Proceedings Volumes, 30*(8), 405-410.

Ben-Akiva, M., Koutsopoulos, H. N., & Walker, J. (2001). DynaMIT-P: Dynamic assignment model system for transportation planning. *WCTR Conference Seoul, Korea.*

Bhaskar, A. (2009). A methodology (CUPRITE) for urban network travel time estimation by integrating multisource data.

Bhaskar, A., & Chung, E. (2013). Fundamental understanding on the use of bluetooth scanner as a complementary transport data. *Transportation Research Part C: Emerging Technologies, 37*, 42-72.

Billings, D., & Yang, J. (2006). Application of the ARIMA models to urban roadway travel time prediction-a case study. *Systems, Man and Cybernetics, 2006. SMC'06. IEEE International Conference On, 3* 2529-2534.

Bluetooth SIG. (2014). A look at the basics of bluetooth technology. Retrieved 11/02, 2014, from http://www.bluetooth.com/Pages/Basics.aspx

Bluetooth Special Interest Group. (2009). Specification of the bluetooth system. core, version 4.0. Retrieved January/01, 2016, from http://www.bluetooth.com

Breiman, L. (1996). Bagging predictors. *Machine Learning, 24*(2), 123-140.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32.

Cassidy, M., & Coifman, B. (1997). Relation among average speed, flow, and density and analogous relation between density and occupancy. *Transportation Research Record: Journal of the Transportation Research Board*, (1591), 1-6.

Chien, S. I., & Kuchipudi, C. M. (2003). Dynamic travel time prediction with real-time and historic data. *Journal of Transportation Engineering, 129*(6), 608-616.

Chorus, C. G., Molin, E. J., & van Wee, B. (2006). Travel information as an instrument to change car-drivers' travel choices: A literature review. *European Journal of Transport and Infrastructure Research*, *6*(4), 335-364.

Chrobok, R., Pottmeier, A., Hafstein, S. F., & Schreckenberg, M. (2004). Traffic forecast in large scale freeway networks. *International Journal of Bifurcation and Chaos*, *14*(06), 1995-2004.

Chung, E., & Kuwahara, M. (2004). An adaptive travel time prediction model based on pattern matching.

Clark, S. (2003). Traffic prediction using multivariate nonparametric regression. *Journal of Transportation Engineering*, *129*(2), 161-168.

Cleveland, W. S., & Devlin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*, *83*(403), 596-610.

Daganzo, C. F. (1994). The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, *28*(4), 269-287.

D'Angelo, M. P., Al-Deek, H. M., & Wang, M. C. (1999). Travel-time prediction for freeway corridors. *Transportation Research Record: Journal of the Transportation Research Board*, *1676*(1), 184-191.

Day, C. M., Brennan, T. M., Hainen, A. M., Remias, S. M., & Bullock, D. M. (2012). Roadway system assessment using bluetooth-based automatic vehicle identification travel time data.

Díaz, J. J. V., González, A. B. R., & Wilby, M. R. (2016). Bluetooth traffic monitoring systems for travel time estimation on freeways. *IEEE Transactions on Intelligent Transportation Systems*, *17*(1), 123-132.

Dion, F., & Rakha, H. (2006). Estimating dynamic roadway travel times using automatic vehicle identification data for low sampling rates. *Transportation Research Part B: Methodological, 40*(9), 745-766.

Duncan, G. (1997). Paramics technical report: Car-following, lane-changing and junction modelling. *Edinburgh, Scotland: Quadstone, Ltd, 5*(1), 5.2.

Erkan, I., & Hastemoglu, H. (2016). Bluetooth as a traffic sensor for stream travel time estimation under bogazici bosporus conditions in turkey. *Journal of Modern Transportation, 24*(3), 207-214.

Fan, S. S., Su, C., Nien, H., Tsai, P., & Cheng, C. (2018). Using machine learning and big data approaches to predict travel time based on historical and real-time data from taiwan electronic toll collection. *Soft Computing, 22*(17), 5707-5718.

Federal Highway Administration. (2006). *CORSIM users guide version 6.0.*

Force, U. T. T. (2012). The high cost of congestion in canadian cities. *Council of Ministers Responsible for Transportation and Highway Safety.*

González-Brenes, J. P., & Cortés, G. M. (2011). Using ensemble of decision trees to forecast travel time. Retrieved 01/26, 2015, from http://blog.kaggle.com/wp-content/uploads/2011/03/team_irazu_screencast.pdf.

Gray, M., & Wong, Y. (2013). Calgary's traveller information system. *2013 Conference and Exhibition of the Transportation Association of Canada-Transportation: Better-Faster-Safer.*

Greenshields, B., Channing, W., & Miller, H. (1935). A study of traffic capacity. *Highway Research Board Proceedings, 1935.*

Guo, F., Krishnan, R., & Polak, J. (2017). The influence of alternative data smoothing prediction techniques on the performance of a two-stage short-term urban travel time prediction framework. *Journal of Intelligent Transportation Systems, 21*(3), 214-226.

Guo, F., Polak, J. W., & Krishnan, R. (2018). Predictor fusion for short-term traffic forecasting. *Transportation Research Part C: Emerging Technologies, 92*, 90-100.

Guo, J., Huang, W., & Williams, B. M. (2014). Adaptive kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification. *Transportation Research Part C: Emerging Technologies, 43, Part 1*(0), 50-64. doi:http://dx.doi.org/10.1016/j.trc.2014.02.006

Haghani, A., Hamedi, M., & Sadabadi, K. (2009). *I-95 corridor coalition vehicle probe project: Validation of INRIX data july-september 2008.*

Haghani, A., Zhang, Y., & Hamedi, M. (2014). *Impact of data source on travel time reliability assessment.* (No. No. UMD-2013-01).

Hamner, B. (2010). Predicting future traffic congestion from automated traffic recorder readings with an ensemble of random forests. *Data Mining Workshops (ICDMW), 2010 IEEE International Conference On*, 1360-1362. doi:10.1109/ICDMW.2010.169

Horning, N. (2013). Introduction to decision trees and random forests. *American Museum of Natural History's.*

Hou, Y., Edara, P., & Chang, Y. (2017). Road network state estimation using random forest ensemble learning. *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 1-6.

Hu, T. (2001). Evaluation framework for dynamic vehicle routing strategies under real-time information. *Transportation Research Record: Journal of the Transportation Research Board*, *1774*(1), 115-122.

Hu, Y. (2013). *Short-term prediction of freeway travel times using data from bluetooth detectors*, University of Waterloo.

Ishak, S., & Al-Deek, H. (2002). Performance evaluation of short-term time-series traffic prediction model. *Journal of Transportation Engineering*, *128*(6), 490-498.

Izadpanah, P. (2010). Freeway travel time prediction using data from mobile probes.

Jain, M., & Coifman, B. (2005). Improved speed estimates from freeway traffic detectors. *Journal of Transportation Engineering*, *131*(7), 483-495.

Jayakrishnan, R., Cohen, M., Kim, J., Mahmassani, H. S., & Hu, T. (1993). A simulation-based framework for the analysis of traffic networks operating with real-time information.

Junghans, M., & Leich, A. (2016). Traffic state estimation with bayesian networks at extremely low V2X penetration rates. *Information Fusion (FUSION), 2016 19th International Conference On*, 1975-1982.

Juri, N. R., Unnikrishnan, A., & Waller, S. T. (2007). *Integrated traffic simulation-statistical analysis framework for online prediction of freeway travel time.*

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, *82*(1), 35-45.

Kwon, J., & Petty, K. (2005). Travel time prediction algorithm scalable to freeway networks with many nodes with arbitrary travel routes. *Transportation Research Record: Journal of the Transportation Research Board*, (1935), 147-153.

Lee, S., & Fambro, D. B. (1999). Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting. *Transportation Research Record: Journal of the Transportation Research Board, 1678*(1), 179-188.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News, 2*(3), 18-22.

Lighthill, M. J., & Whitham, G. B. (1955). On kinematic waves. II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London.Series A.Mathematical and Physical Sciences, 229*(1178), 317-345.

Lo, H. K. (2001). A cell-based traffic control formulation: Strategies and benefits of dynamic timing plans. *Transportation Science, 35*(2), 148-164.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1*(14) 281-297.

Mahmassani, H. S., Fei, X., Eisenman, S., Zhou, X., & Qin, X. (2005). DYNASMART-X evaluation for real-time TMC application: CHART test bed. *Maryland Transportation Initiative, University of Maryland, College Park, Maryland,* 1-144.

Malinovskiy, Y., Wu, Y., Wang, Y., & Lee, U. K. (2010). Field experiments on bluetooth-based travel time data collection. *Transportation Research Board 89th Annual Meeting,* (10-3134).

Masouleh, A. Z., & Hellinga, B. R. (2018). Simulation of the bluetooth inquiry process for application in transportation engineering. *IEEE Transactions on Intelligent Transportation Systems,* (99), 1-11.

Mendes-Moreira, J., Jorge, A. M., de Sousa, J. F., & Soares, C. (2012). Comparing state-of-the-art regression methods for long term travel time prediction. *Intelligent Data Analysis, 16*(3), 427-449.

Messner, A., & Papageorgiou, M. (1990). METANET: A macroscopic simulation program for motorway networks. *Traffic Engineering & Control, 31*(8-9), 466-470.

Moghaddam, S. S. (2014). *Performing short-term travel time prediction on arterials*, University of Waterloo.

Moghaddam, S. S., & Hellinga, B. (2013). Quantifying measurement error in arterial travel 6 times measured by bluetooth detectors. *Transportation Research Board 92nd 7 Annual Meeting,*

Moonam, H. M., Qin, X., & Zhang, J. (2019). Utilizing data mining techniques to predict expected freeway travel time from experienced travel time. *Mathematics and Computers in Simulation, 155*, 154-167.

Muñoz, L., Sun, X., Horowitz, R., & Alvarez, L. (2003). Traffic density estimation with the cell transmission model. *American Control Conference, 2003. Proceedings of the 2003, 5* 3750-3755.

Myers, K., & Tapley, B. D. (1976). Adaptive sequential estimation with unknown noise statistics. *Automatic Control, IEEE Transactions On, 21*(4), 520-523.

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & its Applications, 9*(1), 141-142.

Nam, D. H., & Drew, D. R. (1998). Analyzing freeway traffic under congestion: Traffic dynamics approach. *Journal of Transportation Engineering, 124*(3), 208-212.

Nanthawichit, C., Nakatsuji, T., & Suzuki, H. (2003). Application of probe-vehicle data for real-time traffic-state estimation and short-term travel-time prediction on a freeway. *Transportation Research Record: Journal of the Transportation Research Board, 1855*(1), 49-59.

Nause, P. (2015). *Initiatives to manage traffic congestion 2015 MEA workshop.*

Nihan, N. L., & Holmesland, K. O. (1980). Use of the box and jenkins time series technique in traffic forecasting. *Transportation, 9*(2), 125-143.

Nikovski, D., Nishiuma, N., Goto, Y., & Kumazawa, H. (2005). Univariate short-term prediction of road travel times. *Intelligent Transportation Systems, 2005. Proceedings. 2005 IEEE*, 1074-1079.

NIST/SEMATECH. (2012). E-handbook of statistical methods. Retrieved 10/19, 2014, from http://www.itl.nist.gov/div898/handbook/

Noroozisanani, R. (2017). Real time prediction of traffic speed and travel time characteristics on freeways.

Oh, J., Jayakrishnan, R., & Recker, W. (2002). Section travel time estimation from point detection data. *Center for Traffic Simulation Studies.*

Oh, S., Byon, Y., Jang, K., & Yeo, H. (2018). Short-term travel-time prediction on highway: A review on model-based approach. *KSCE Journal of Civil Engineering, 22*(1), 298-310.

Okutani, I., & Stephanedes, Y. J. (1984). Dynamic prediction of traffic volume through kalman filtering theory. *Transportation Research Part B: Methodological, 18*(1), 1-11.

Papageorgiou, M., Papamichail, I., Messmer, A., & Wang, Y. (2010). Traffic simulation with METANET. Fundamentals of traffic simulation (pp. 399-430) Springer.

Park, D., Rilett, L. R., & Han, G. (1999). Spectral basis neural networks for real-time travel time forecasting. *Journal of Transportation Engineering, 125*(6), 515-523.

Payne, H. J. (1971). Models of freeway traffic and control. *Mathematical Models of Public Systems.*

Peterson, L. E. (2009). K-nearest neighbor. *Scholarpedia, 4*(2), 1883. doi:10.4249/scholarpedia.1883

PTV, A. (2011). Vissim 5.40-01 user manual. *Karlsruhe, Germany.*

Puckett, D., & Vickich, M. (2009). Anonymous wireless address matching (AWAM) for travel time data collection. *Proceedings of the 18th World Congress on Intelligent Transport Systems,* 16-20.

Quayle, S., Koonce, P., DePencier, D., & Bullock, D. (2010). Arterial performance measures with media access control readers: Portland, oregon, pilot study. *Transportation Research Record: Journal of the Transportation Research Board,* (2192), 185-193.

Rakha, H., & Arafeh, M. (2010). Calibrating steady-state traffic stream and car-following models using loop detector data. *Transportation Science, 44*(2), 151-168.

Ravi Sekhar, C., Madhu, E., Kanagadurai, B., & Gangopadhyay, S. (2013). Analysis of travel time reliability of an urban corridor using micro simulation techniques. *Current Science, 105*(3), 319-329.

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of Statistical Modeling and Analytics, 2*(1), 21-33.

Rice, J., & Van Zwet, E. (2004). A simple and effective method for predicting travel times on freeways. *IEEE Transactions on Intelligent Transportation Systems, 5*(3), 200-207.

Richards, P. I. (1956). Shock waves on the highway. *Operations Research, 4*(1), 42-51.

Rilett, L. R., & Park, D. (2001). Direct forecasting of freeway corridor travel times using spectral basis neural networks. *Transportation Research Record: Journal of the Transportation Research Board, 1752*(1), 140-147.

Robinson, S., & Polak, J. W. (2005). Modeling urban link travel time with inductive loop detector data by using the k-NN method. *Transportation Research Record: Journal of the Transportation Research Board, 1935*(1), 47-56.

Salem, A., Nadeem, T., Cetin, M., & El-Tawab, S. (2015). DriveBlue: Traffic incident prediction through single site bluetooth. *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 725-730.

Schneider, W. H., Holik, W., & Bakula, C. (2012). Application of bluetooth technology to rural freeway speed data collection.

Seiffert, C., Khoshgoftaar, T. M., Van Hulse, J., & Napolitano, A. (2010). RUSBoost: A hybrid approach to alleviating class imbalance. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, *40*(1), 185-197.

Semanjski, I. (2015). Potential of big data in forecasting travel times. *Promet-Traffic & Transportation*, *27*(6), 515-528.

Sharifi, E., Hamedi, M., Haghani, A., & Sadrsadat, H. (2011). Analysis of vehicle detection rate for bluetooth traffic sensors: A case study in maryland and delaware. *18th ITS World Congress*.

Smith, B. L., & Demetsky, M. J. (1996). Multiple-interval freeway traffic flow forecasting. *Transportation Research Record: Journal of the Transportation Research Board*, *1554*(1), 136-141.

Stergiou, C., & Siganos, D. (1996). Neural networks. Retrieved 10/28, 2014, from http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html

Sun, H., Liu, H. X., Xiao, H., He, R. R., & Ran, B. (2003). Short term traffic forecasting using the local linear regression model. *82nd Annual Meeting of the Transportation Research Board, Washington, DC*.

Sun, X., Zhang, H., Tian, F., & Yang, L. (2018). The use of a machine learning method to predict the real-time link travel time of open-pit trucks. *Mathematical Problems in Engineering*.

Tampère, C. M., & Immers, L. (2007). An extended kalman filter application for traffic state estimation using CTM with implicit mode switching and dynamic parameters. *Intelligent Transportation Systems Conference, 2007. ITSC 2007. IEEE,* 209-216.

Tu, H. (2008). *Monitoring travel time reliability on freeways* Netherlands TRAIL Research School.

Van Aerde, M. (1995). Single regime speed-flow-density relationship for congested and uncongested highways. *74th TRB Annual Conference*.

Van Hinsbergen, C., & Van Lint, J. (2008). Bayesian combination of travel time prediction models. *Transportation Research Record: Journal of the Transportation Research Board*, (2064), 73-80.

van Lint, H. (2004). *Reliable travel time prediction for freeways* Netherlands TRAIL Research School.

Van Lint, J. (2006). Reliable real-time framework for short-term freeway travel time prediction. *Journal of Transportation Engineering, 132*(12), 921-932.

Van Lint, J., & Hoogendoorn, S. P. (2010). A robust and efficient method for fusing heterogeneous data from traffic sensors on freeways. *Computer-Aided Civil and Infrastructure Engineering, 25*(8), 596-612.

Van Lint, J., Hoogendoorn, S., & van Zuylen, H. J. (2005). Accurate freeway travel time prediction with state-space neural networks under missing data. *Transportation Research Part C: Emerging Technologies, 13*(5), 347-369.

van Wageningen-Kessels, F., van Lint, H., Vuik, K., & Hoogendoorn, S. (2014). Genealogy of traffic flow models. *EURO Journal on Transportation and Logistics*, 1-29.

Vissim, P. (7). 7–user manual, 2015. *Karlsruhe, Njemačka.*

Vlahogianni, E. I., Karlaftis, M. G., & Golias, J. C. (2014). Short-term traffic forecasting: Where we are and where we're going. *Transportation Research Part C: Emerging Technologies*.

Wang, T. (2002). *A review of bluetooth technology fundamentals and its application to scatternets.* (Master, Concordia University).

Wang, Y., & Papageorgiou, M. (2005). Real-time freeway traffic state estimation based on extended kalman filter: A general approach. *Transportation Research Part B: Methodological, 39*(2), 141-167. doi:http://dx.doi.org/10.1016/j.trb.2004.03.003

Wang, Y., Papageorgiou, M., & Messmer, A. (2003). Motorway traffic state estimation based on extended kalman filter. *Proceedings of the European Control Conference ECC, 3.*

Wang, Y., Papageorgiou, M., & Messmer, A. (2006). RENAISSANCE–A unified macroscopic model-based approach to real-time freeway network traffic surveillance. *Transportation Research Part C: Emerging Technologies, 14*(3), 190-212.

Welch, G., & Bishop, G. (1997). An introduction to the kalman filter.

Welsh, E., Murphy, P., & Frantz, P. (2002). Improving connection times for bluetooth devices in mobile environments. *International Conference on Fundamentals of Electronics Communications and Computer Sciences (ICFS).*

Williams, B. M., Durvasula, P. K., & Brown, D. E. (1998). Urban freeway traffic flow prediction: Application of seasonal autoregressive integrated moving average and exponential smoothing models. *Transportation Research Record, 1644*(1), 132-141.

Xia, J., Chen, M., & Huang, W. (2011). A multistep corridor travel-time prediction method using presence-type vehicle detector data. *Journal of Intelligent Transportation Systems, 15*(2), 104-113.

Xie, Y., Zhang, Y., & Ye, Z. (2007). Short-term traffic volume forecasting using kalman filter with discrete wavelet decomposition. *Computer-Aided Civil and Infrastructure Engineering, 22*(5), 326-334.

Xu, T., Hao, Y., & Sun, L. (2007). Travel time prediction of urban expressway in unstable traffic flow. *International Conference on Transportation Engineering 2007, ICTE 2007*, 2205-2210.

You, J., & Kim, T. J. (2000). Development and evaluation of a hybrid travel time forecasting model. *Transportation Research Part C: Emerging Technologies, 8*(1–6), 231-256. doi:http://dx.doi.org/10.1016/S0968-090X(00)00012-7

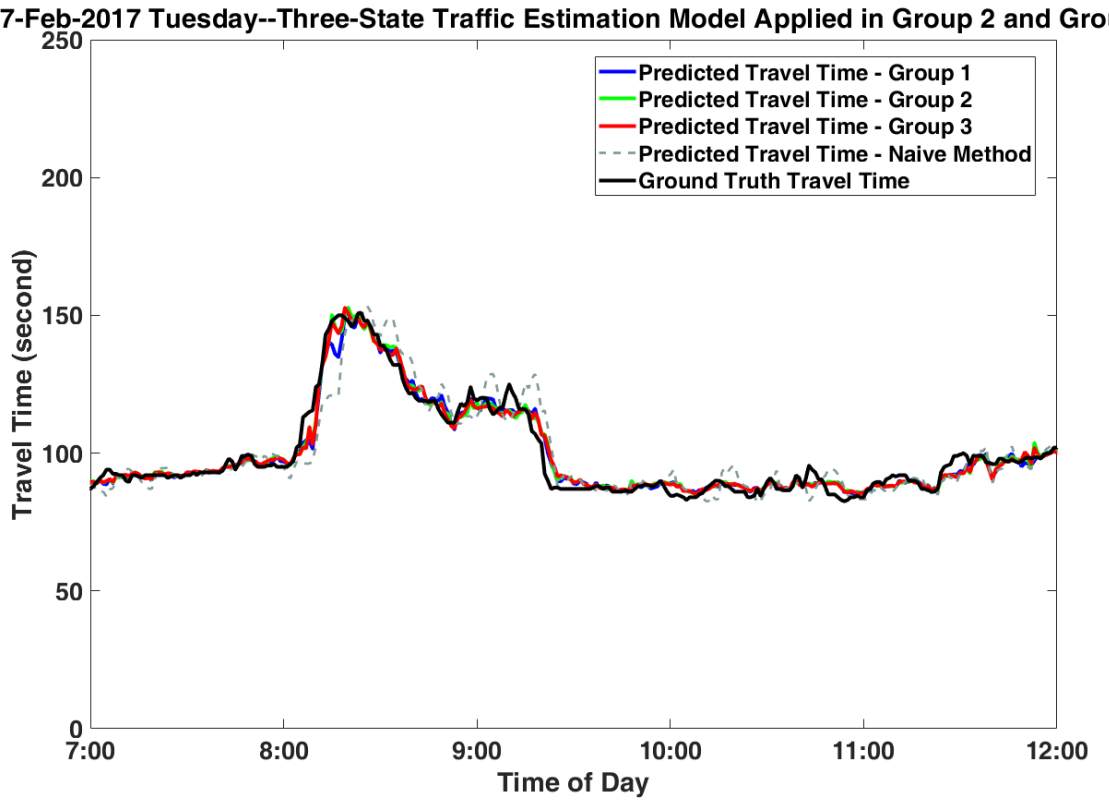
Young, S. E., Sharifi, E., Sadrsadat, H., Serulle, N. U., & Sadabadi, K. F. (2012). Detection probability models for bluetooth re-identification technology.

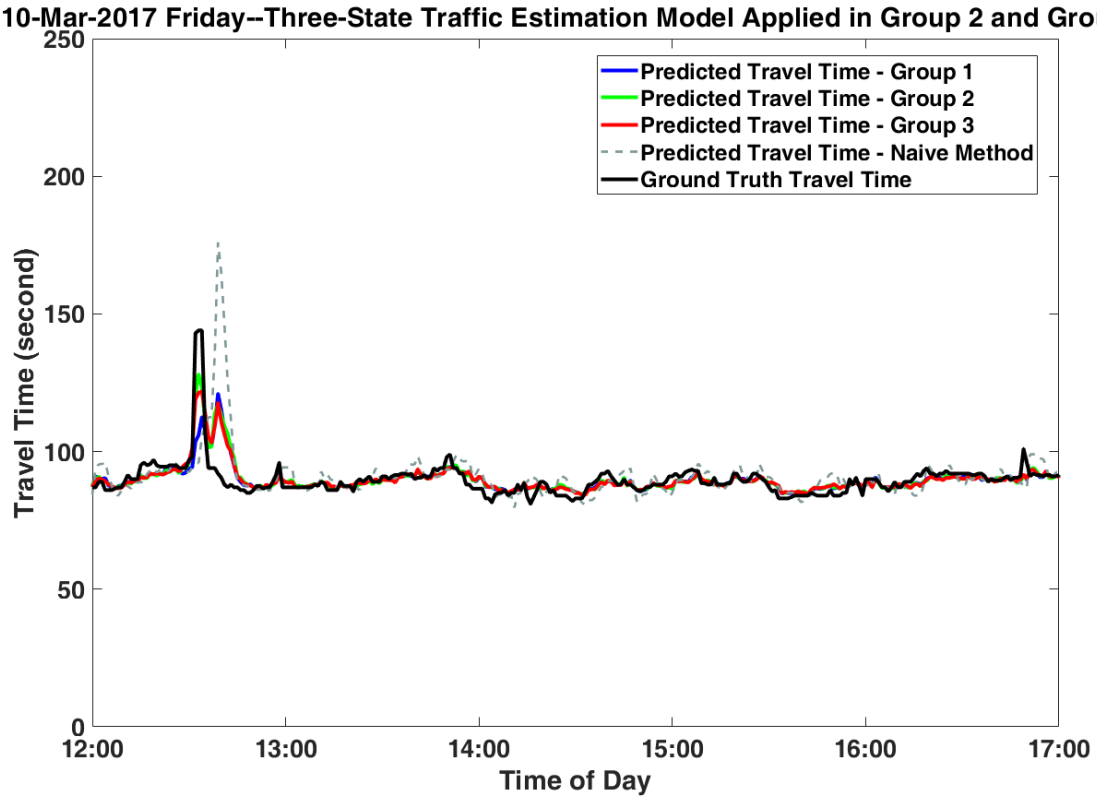
Zhang, X., & Rice, J. A. (2003). Short-term travel time prediction. *Transportation Research Part C: Emerging Technologies, 11*(3), 187-210.

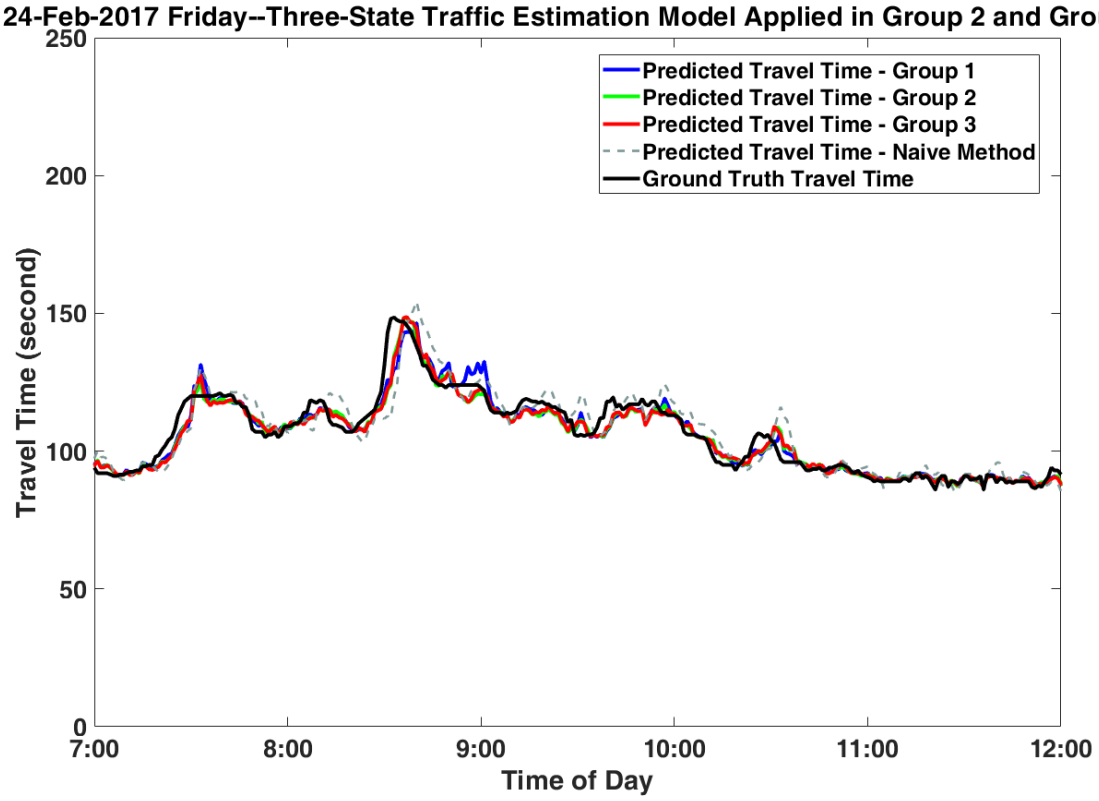
Zoto, J., La, R. J., Hamedi, M., & Haghani, A. (2012). Estimation of average vehicle speeds traveling on heterogeneous lanes using bluetooth sensors. *Vehicular Technology Conference (VTC Fall), 2012 IEEE*, 1-5.

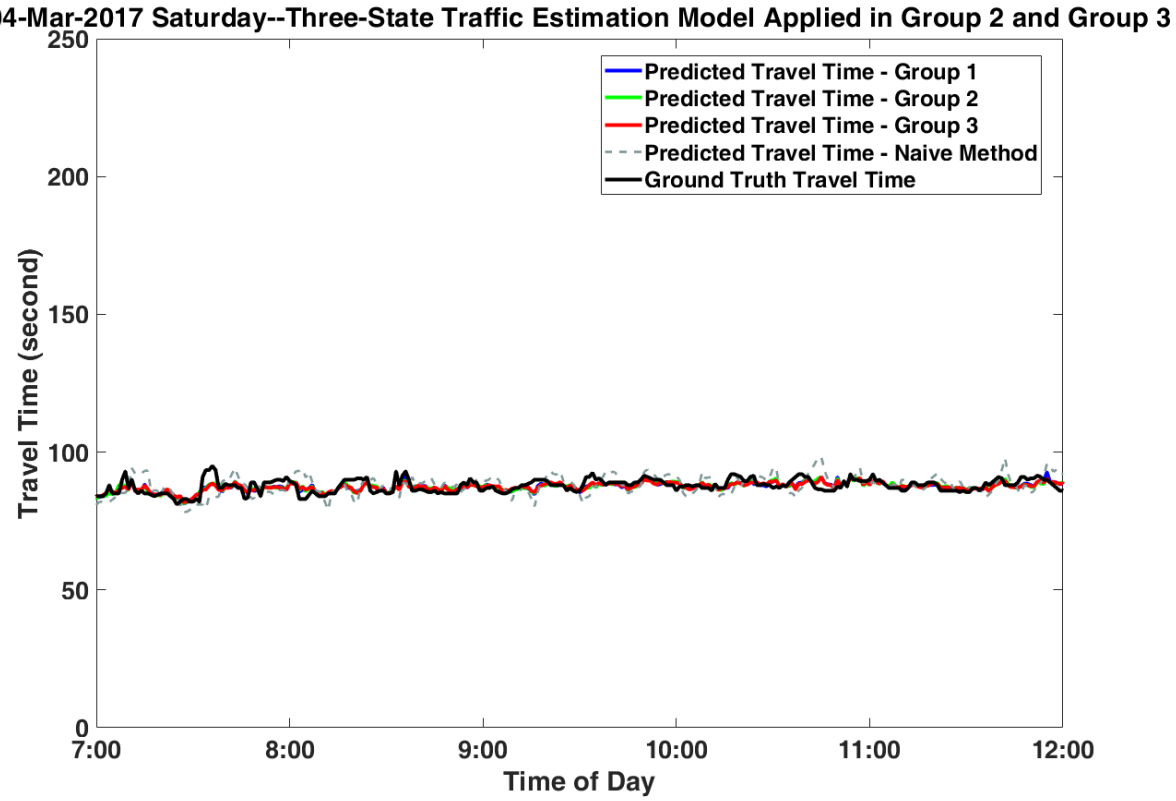
Appendix A

Additional Travel Time Prediction Results - Empirical Data




10-Mar-2017 Friday--Three-State Traffic Estimation Model Applied in Group 2 and Group 3
Predicted Travel Time - Group 1
Predicted Travel Time - Group 2
Predicted Travel Time - Group 3
Predicted Travel Time - Naive Method
Ground Truth Travel Time
Travel Time (second)
Time of Day
250
200
150
100
50
0
12:00
13:00
14:00
15:00
16:00
17:00


24-Feb-2017 Friday--Three-State Traffic Estimation Model Applied in Group 2 and Group 3
Predicted Travel Time - Group 1
Predicted Travel Time - Group 2
Predicted Travel Time - Group 3
Predicted Travel Time - Naive Method
Ground Truth Travel Time
Travel Time (second)
Time of Day
250
200
150
100
50
0
7:00
8:00
9:00
10:00
11:00
12:00


04-Mar-2017 Saturday--Three-State Traffic Estimation Model Applied in Group 2 and Group 3
Predicted Travel Time - Group 1
Predicted Travel Time - Group 2
Predicted Travel Time - Group 3
Predicted Travel Time - Naive Method
Ground Truth Travel Time
Travel Time (second)
250
200
150
100
50
0
7:00
8:00
9:00
10:00
11:00
12:00
Time of Day

Appendix B

Additional Travel Time Prediction Results - Simulation Data