Evaluation of Signal Retiming Measures Using Bluetooth Travel Time Data

by

Cameron Berko

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Applied Science

in

Civil Engineering

Waterloo, Ontario, Canada, 2015

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any final required revisions, as accepted by my examiners.

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

Abstract

Signal retiming is an appealing strategy for improving network performance because it does not require the addition of new roadway capacity. The emergence of Bluetooth technology presents an alternative method of collecting travel time data to evaluate the implementation of signal retiming measures along an arterial corridor via a before-and-after study. As opposed to the industry standard of collecting a limited number of travel times via dedicated travel time runs using vehicles equipped with GPS data loggers, Bluetooth technology allows for a much greater number of travel times to be collected from a wider range of vehicles and drivers. However, the need persists for a practitioner-ready methodology that details how data collected in this manner should be used to evaluate signal retiming measures. This need formed the basis upon which this investigation was conducted.

Both field and simulated arterial corridors were examined in this research. The field corridor consisted of a 15.1-kilometre long section of Victoria Park Avenue located in Toronto, Ontario that contained 37 signalized intersections. Seven Bluetooth detectors were deployed to collect data, meaning that the corridor was divided into six links. GPS probe runs were also available for comparison. The simulated corridor consisted of a 4.8-kilometre long section of Hespeler Road in Cambridge, Ontario that contained 12 signalized intersections. Three Bluetooth detectors were deployed to collect data, meaning that the corridor was divided into two links. GPS probe runs were also simulated.

Bluetooth travel times were available at the path level (i.e. travel times of vehicles that traversed the entire length of the arterial corridor) and at the link level (i.e. travel times of vehicles that traversed only part of the corridor). To develop measures of effectiveness for evaluating signal retiming measures, the merits of each of these data sets for this purpose were first identified. Through statistical testing, it was found to be infeasible to use the travel times of vehicles that traversed the entire corridor for signal retiming evaluation due to the small number of travel times collected. Instead, a corridor should be subdivided into links through the placement of multiple Bluetooth detectors to increase the number of travel times collected.

Next, recommendations regarding the characteristics of a signal retiming study were proposed. A regression model was developed using the field data to allow a practitioner to estimate the

duration of the data collection period based on the characteristics of the corridor. Using the results produced by applying this regression model to the field data, recommendations were provided for the spacing of detectors.

Next, measures of effectiveness to assess the impacts of signal retiming were developed. The recommended measures incorporated the difference in the means of the Before and After travel time data, the number of vehicles that traversed each link of the corridor, and statistical significance of the difference in the means. These measures provide a practitioner with an idea of the travel time savings or losses produced for the corridor, the degree to which these savings or losses were experienced by vehicles that traversed the corridor, and whether or not these savings or losses were statistically significant.

The proposed measures were applied to both the field and simulated Bluetooth travel time data. These results were then compared to the results obtained by applying these measures to the GPS probe runs and to the true changes in travel time for the simulated corridor.

Since one commonly cited weakness of Bluetooth travel time data is the presence of outliers in the measured travel times, the sensitivity of the proposed measure of effectiveness to the presence of outliers (i.e. travel times whose magnitudes were not representative of the traffic stream for which signal retiming was intended) was examined using the field Bluetooth travel time data. It was demonstrated that the developed measures are not significantly influenced by the presence of outliers.

This investigation provides a practitioner with guidance on how to perform a before-and-after evaluation of a signal retiming study using Bluetooth travel time data. This investigation demonstrated that the division of an arterial corridor into smaller segments produces enough data to be able to statistically differentiate between travel times collected before and after signal retiming measures have been implemented. Guidance is also provided regarding the duration of the data collection period and how to divide the corridor into links through detector spacing. Finally, the developed measures of effectiveness provide concise evidence of the success or failure of signal retiming that a practitioner can present to stakeholders and policymakers with ease.

Acknowledgements

This research project could not have been completed without the contributions of the following people.

I would like to thank my supervisor, Dr. Bruce Hellinga, for his scholarly input, continuous encouragement and financial support. Every meeting that we had together was incredibly informative and thought-provoking. I thoroughly enjoyed the opportunity to work under your guidance and expertise.

I would also like to thank the members of my reading committee, Dr. Liping Fu and Dr. Nandita Basu, for taking the time to provide their input to improve this research project. I am also thankful to my colleagues Dr. Soroush Salek Moghaddam, Dr. Reza Noroozi and Dr. Akram Nour for their support and patience in answering my questions while I navigated the MATLAB environment. Special thanks go to fellow Master's student Jordan Hart-Bishop for developing the simulated model and for always being willing to listen and provide input on my ideas.

I would also like to thank my fellow graduate students, friends and roommates Joseph Stoner and Kevin Goorts. Your willingness to listen and provide input when I encountered obstacles over these past two years was critical to the timely completion of this research project. You provided me with the input needed to keep moving forward in my research. I am eternally humbled by your work ethic, intellect and dependability.

I owe great thanks to my parents, Pam and Walter Berko, and the rest of my family for providing me with endless support, generosity and motivation throughout my undergraduate and graduate studies at the University of Waterloo. I never would have been able to complete my studies without your guidance. Finally, I would like to thank my girlfriend Shannon Scanlon and her family for their support. Shannon, your patience and understanding are contagious and have helped me become a better person. Thank you for being by my side.

Table of Contents

List of Tables

List of Figures

1 Introduction

Traffic congestion has been steadily increasing as a result of the growth of traffic volumes in urban areas. In a study of the impacts of traffic congestion in the Greater Toronto Area, it was found that on an annual basis congestion cost commuters $3.3 billion and cost the local economy $2.7 billion (HDR Corporation, 2008). Without remedial measures, population growth will further increase traffic demand, thus increasing the degree of congestion on the road network and increasing the costs incurred by commuters and the local economy.

The mitigation of network congestion is impeded by constrained municipal budgets, lack of available space for expansion of the network in urban areas, and a desire to appropriate space to transit, pedestrian and cyclist facilities. Faced with these obstacles, transportation agencies must explore measures that more efficiently utilize the existing road network in a cost-effective manner.

The performance of arterial roadways has a significant impact on the economy and on quality of life (Federal Highway Administration, 2015a). Arterial performance must therefore be a priority for all urban transportation agencies looking to improve traffic operations on their facilities. Delcan and Lura Consulting (2013) cited poor traffic signal coordination as one of the major causes of poor arterial performance in the City of Toronto and recognized signal retiming to be vital to mitigating congestion. The City's 2012-2013 Corridor Retiming Program involved the retiming of 112 intersections along three arterial corridors. The projected benefits of this program included annual reductions in traveller delay by 380,000 hours, vehicle stops by 33 million, fuel consumption by 2.1 million litres and greenhouse gases by 55 tonnes. These benefits contributed to a benefit/cost ratio of 66:1, which highlights the magnitude of the potential benefits of signal retiming measures (Delcan and Lura Consulting, 2013).

Having established the potential benefits of signal retiming measures, the remainder of this chapter presents the concepts central to the evaluation of signal retiming measures. First, the concepts of signal retiming and signal coordination are introduced. Next, the technologies used to collect travel time data are reviewed with a focus on Bluetooth technology. Comparisons of floating car probe and Bluetooth technologies are then reviewed and the major advantages and disadvantages of Bluetooth technology are highlighted. The problem statement, research questions and research objectives of this investigation are then presented and an outline of this thesis is provided.

1.1 Background

1.1.1 Signal Retiming and Signal Coordination

While local and collector roads provide access to land use and highways and freeways provide high-speed mobility, arterials are unique in their ability to provide access to land use as well as mobility (Federal Highway Administration, 2015b). Like local and collector roads, arterials employ traffic control devices and allow for midblock access and egress. These devices are impedances to progression along an arterial corridor and introduce variation to the amount of time it takes travellers to traverse the corridor. These devices do not exist on highway or freeway facilities, making it important to the functionality of an arterial that these sources of travel time variation are controlled for.

Traffic signals represent significant impedances to progression, which to a degree can be controlled for through signal retiming. Signal retiming is the practice of adjusting the signal timing plan of an intersection to increase the number of vehicles that the intersection is able to serve in a given time period and to reduce the delays that vehicles traversing the intersection experience. A signal timing plan consists of a sequence of green, amber and red phases that collectively dictate which vehicle movements have the right of way at an intersection at a certain time. Signal retiming is performed to accommodate the changing hourly, daily, weekly and/or seasonal variations in demand at an intersection. Therefore, the performance of an intersection will degrade over time if its signal timing plan remains fixed while the traffic patterns at the intersection change. As a result of changing traffic patterns, signal timing plans are recommended to be reviewed on an annual basis (Institute of Transportation Engineers, 2015).

Signal coordination consists of the retiming of a series of signals to facilitate the progression of vehicles along a corridor of interest. Figure 1.1 illustrates the basic components of signal coordination for the northbound direction (assuming “up” is north) in a space-time diagram. The y-axis represents the distance travelled along the corridor and the x-axis represents time. The basic components of signal coordination are labelled in the diagram and are as follows:

- Item A is the **cycle length**, which represents the amount of time it takes for a signal to complete a single iteration of its phases. To achieve coordination, all intersections involved should have the same cycle length.
- Items B and C are the **red and green intervals**, which represent the state of a signal at an intersection at a given time. The amber interval duration is divided into two components; the first part of the amber in which vehicles continue to discharge is considered as part of the green interval and the second part when vehicles stop is considered as part of the red interval.
- Item D is the **master intersection**, which is the intersection that is used as a reference point in space and time for signal coordination.
- Item E indicates the **offsets** of the non-master intersections. The offset is the amount of time that the signal timing plan at an intersection is shifted in time relative to the master intersection in order for coordination to be achieved. For a given intersection, the offset is measured from the start of the green phase at the master intersection to the start of the green phase at the intersection of interest.
- Item F is the **green band**, which is the area in space and time in which a vehicle desires to be while traversing the corridor so that they experience zero delay. The width of this area is a measure of effectiveness used to assess the success of signal coordination.
- Item G is the **platoon speed**, which is the slope of the green band. This represents the typical speed of a group of vehicles that formed at the stop line during the red phase, thus creating a platoon. For the purposes of a space-time diagram, the platoon speed is typically estimated using the speed limit, the geometry of the roadway and historical vehicle speed data (Day et al., 2014).

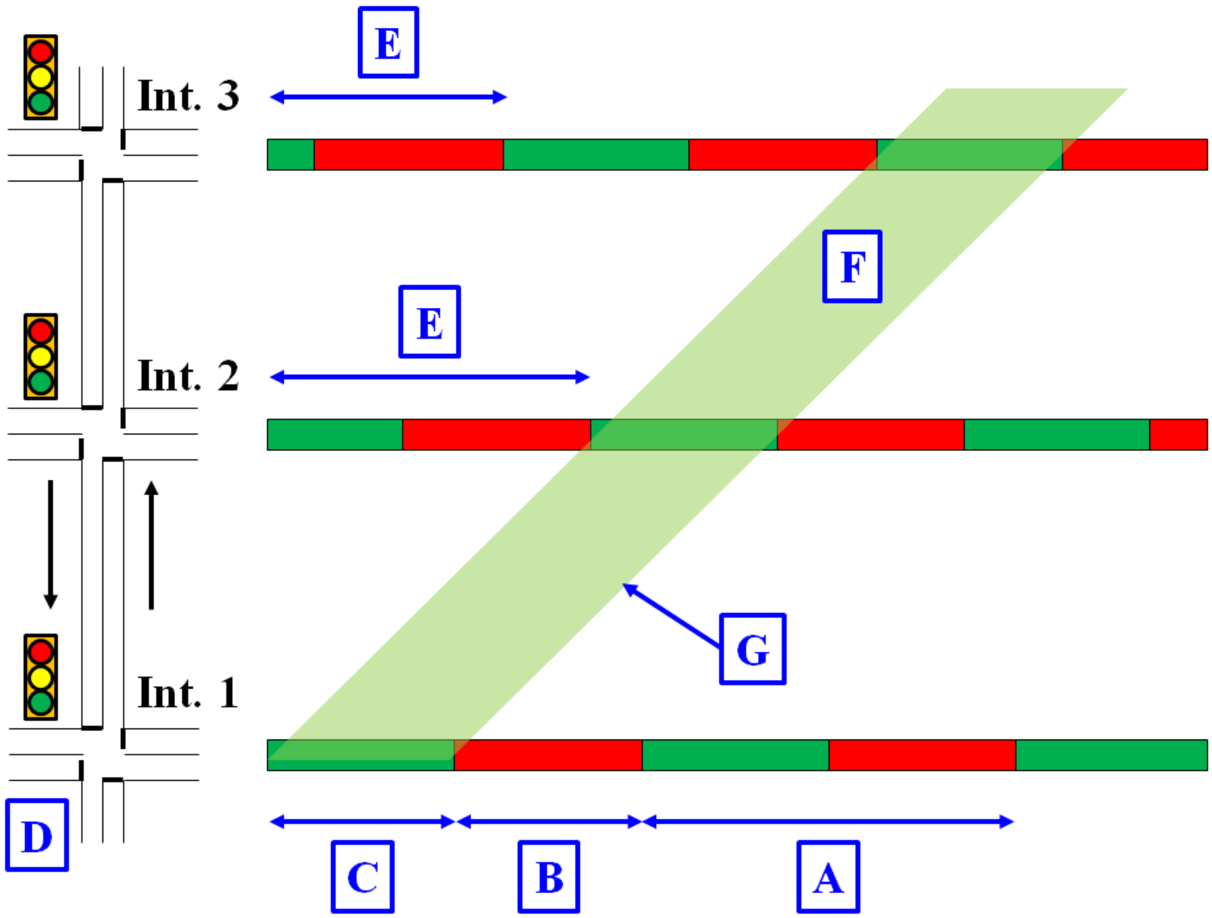

Figure 1.1: Basic components of signal coordination

Signal retiming measures primarily consist of changes to the green splits, offsets and cycle length of each intersection (Day et al., 2010). Changes to these components make up the most basic of signal retiming strategies for corridors with fixed signal timing systems (i.e. the durations of the red and green intervals remain constant over time). To accommodate peak period traffic, a library of fixed signal timing plans tailored to each time period can be created and drawn from so that the appropriate signal timing plan can be activated at the appropriate time. Facilitated by technology such as loop detectors embedded in the pavement or cameras, actuation can be used to extend a phase or terminate a phase early. While actuation adds complexity to the system by changing the length of the phases and the cycle, actuation allows for more efficient utilization of the capacity of the intersection. Signals can be semi-actuated (detectors are present only on minor street approaches of an intersection) or fully actuated (detectors are present on all approaches of an intersection).

Due to the complexity involved with changing multiple signal timing plans in response to changes in traffic demand, software packages that utilize certain algorithms are employed to determine optimal signal timing plans. These software packages can be used to maximize the width of the green band along a corridor, which is done using vehicle flow profiles (i.e. vehicle volumes as a function of time) as input. The modification of signal timing plans along a corridor is typically performed using the off-line software TRANSYT or the real-time demand-responsive software SCOOT (Imtech Traffic et al., 2013).

Signal coordination is performed for a certain direction of traffic according to time of day to accommodate peak period volumes. For example, if an arterial connects a northern suburban area to a southern business district, signal coordination will be performed for the southbound direction in the morning to accommodate traffic flow into the city and for the northbound direction in the evening to accommodate traffic flow out of the city. It is evident that the green band in Figure 1.1 allows for coordination for the northbound direction while it is difficult to define a green band for the southbound direction. Under this sequence of signal timing plans, southbound vehicles will experience difficulty progressing along the corridor as easily as northbound vehicles.

1.1.2 Bluetooth as a Travel Time Measuring Technology

Signal retiming performance can be evaluated on the basis of a variety of measures including delay, the number of vehicle stops, fuel consumption and carbon dioxide emissions (Gordon, 2010). Despite this variety of measures, travel time is the most commonly used measure in traffic studies (Gettman et al., 2013). Travel time has been recognized as a fundamental measure in characterizing transportation systems as it is understood by a variety of audiences (Turner et al., 1998). In this investigation, the travel time of a vehicle is defined to be the amount of time that a vehicle requires to traverse a given section of roadway.

To determine how effective signal retiming measures were at improving traffic operations along an arterial corridor, a “before-and-after” study is typically performed. A before-and-after study provides evidence that the signal timing plans optimized using traffic analysis software such as

Synchro produce measurable, real-world benefits to motorists. In a before-and-after study, travel times along an arterial corridor are collected before and after signal retiming measures are implemented. The travel times within these two data sets are then compared with success being defined as a reduction in the magnitudes of the travel times collected during the Before period compared to those of the travel times collected during the After period.

Travel time data collection technologies primarily consist of sensors that are either fixed, mobile, or a combination of the two (Moghaddam, 2014). Technologies such as Bluetooth, toll tag tracking and license plate matching each represent a series of fixed sensors while GPS devices, cell phones and devices connected to vehicle odometers are examples of mobile sensors. One means of collecting travel time data that has emerged in recent years is through the use of Bluetooth technology.

The term "Bluetooth" refers to a telecommunications industry specification that defines how short-range wireless communication is possible between digital devices. Provided that they are both Bluetooth-enabled, two devices can communicate with each other at a distance of up to 100 metres (Haghani and Hamedi, 2013). Each device with Bluetooth capability has a unique media access control address (commonly called a MAC address or MAC ID) assigned by the manufacturer, with each address being an alphanumeric sequence of 12 characters (for example, 00:28:3B:EF:23:67). When activated, fixed Bluetooth detectors mounted on the side of a roadway are able to continuously detect Bluetooth-enabled devices that pass through its detection zone by recording the device's MAC ID and the time at which this MAC ID was detected. This includes Bluetooth-enabled vehicles as well as other Bluetooth-enabled devices within the range of a detector. Friesen and McLeod (2014) provide a detailed description of the technical attributes of the software associated with roadside Bluetooth detectors.

When a Bluetooth-enabled vehicle passes through a roadside detector's search range, the detector is able to detect the vehicle multiple times. This series of detections is commonly known as a collection of "hits" (Moghaddam, Noroozi and Hellinga, 2014). Bluetooth detectors scan on 18 frequencies once every 1.28 seconds; since there are 36 frequencies within the Bluetooth frequency spectrum, the probability of a Bluetooth-enabled vehicle being detected while it is within 50 metres of a detector is 56.25% (Moghaddam, 2014). This probability decreases linearly due to the weakening of the signal strength beyond this 50-metre range. Moghaddam (2014) modelled the relationship between a vehicle's probability of detection and its distance to the detector as an isosceles trapezoid. This relationship is shown in Figure 1.2.

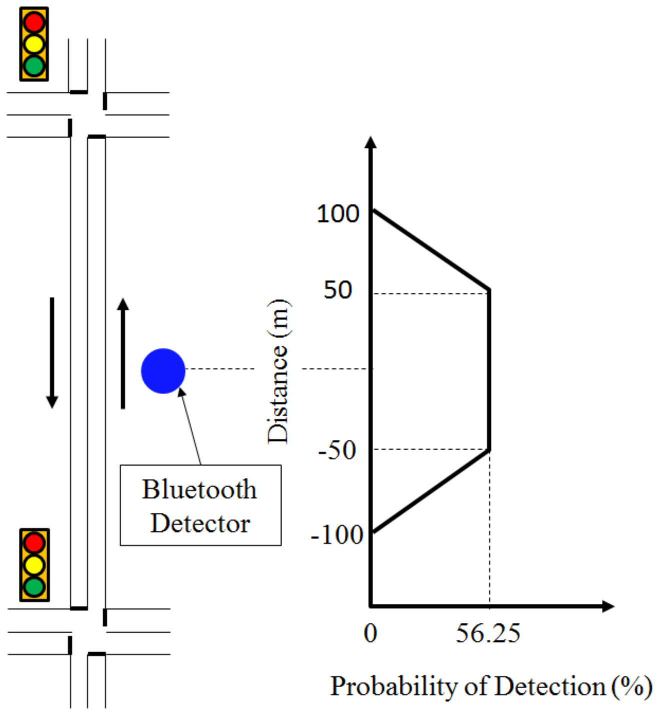

Figure 1.2: Probability of Bluetooth detection as a function of the distance to the detector (Moghaddam, 2014)

Knowing that each Bluetooth-enabled device has a unique MAC ID assigned to it, the time stamps at successive detectors can be compared to determine the amount of time it took a Bluetooth-enabled vehicle to travel from one detector to another. Wasson and Bullock (2012) state that using the first appearance of a MAC ID at a detector is the most common basis for calculating travel time. Quayle et al. (2010), Day et al. (2012) and Moghaddam and Hellinga (2012) also use "first-first" MAC ID matching. Figure 1.3 illustrates how a travel time is extracted from Bluetooth detector data.

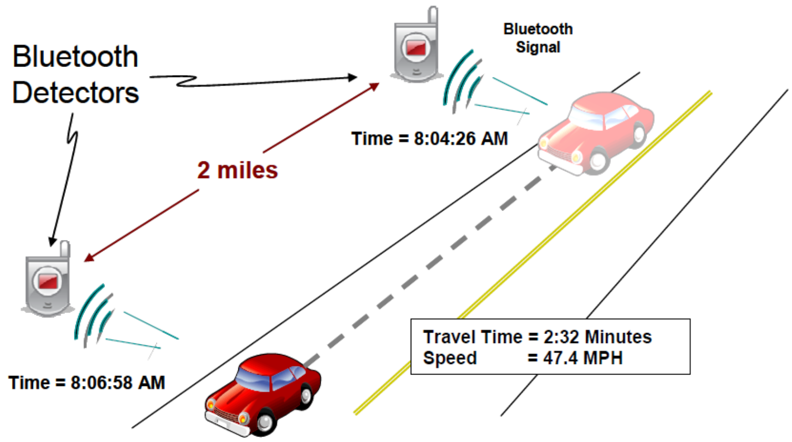

Figure 1.3: Bluetooth travel time data collection process (Young, 2008)

1.1.3 Comparison of Bluetooth and Floating Car Probe Technology

The floating car probe method is used routinely by transportation professionals to collect travel time data (Wasson, Sturdevant and Bullock, 2008). Liu, Chien and Kim (2012) referred to floating car probe data as a ground truth when evaluating the accuracy of Bluetooth travel times in a freeway environment. As described by Koonce et al. (2008), the most advanced floating car probe studies employ portable GPS devices to collect data.

Floating car probe data are highly detailed in that instantaneous vehicle position data (latitude and longitude), speed data and time data are collected over the entire span of a probe vehicle's trip, typically on a per second basis. Conversely, Bluetooth data are much less detailed in that only time stamps collected at each detector are available. This contrast is illustrated in Figure 1.4. In this figure, a Bluetooth detector is set up at each intersection. Bluetooth time stamps can only be collected within the detection range of the Bluetooth detectors while a probe vehicle equipped with a portable GPS device is able to collect data along the entire corridor.

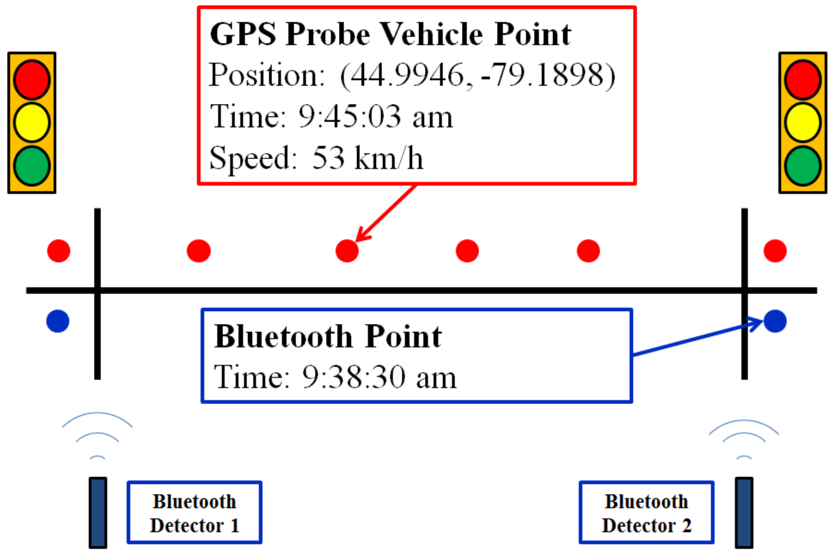

Figure 1.4: GPS probe vehicle data versus Bluetooth data

The high level of detail provided by GPS probe vehicle data allows for the creation of individual vehicle trajectories, speed-time and space-time diagrams, as illustrated for hypothetical data in Figure 1.5, Figure 1.6 and Figure 1.7. These highly detailed data can be used to determine when and where vehicles experienced significant delay along a corridor. As depicted in these figures, a vehicle travelling along a freeway may have encountered recurrent or non-recurrent congestion during its trip. This is indicated by points becoming closer together in space in the vehicle trajectory diagram and a sudden reduction in speed in the speed-time diagram. While the occurrence of such an event can be identified using Bluetooth data by an increase in travel time, the location along the corridor at which the congestion was encountered is unknown due to the absence of data between the detectors.

For the purposes of evaluating signal retiming measures, the coarse level of detail of Bluetooth travel time data is not a significant obstacle. Considering that only basic travel times are needed to evaluate signal retiming measures in a before-and-after study, a high level of detail is not required in this context.

Figure 1.5: Hypothetical GPS probe vehicle trajectory (Google Maps, 2015)

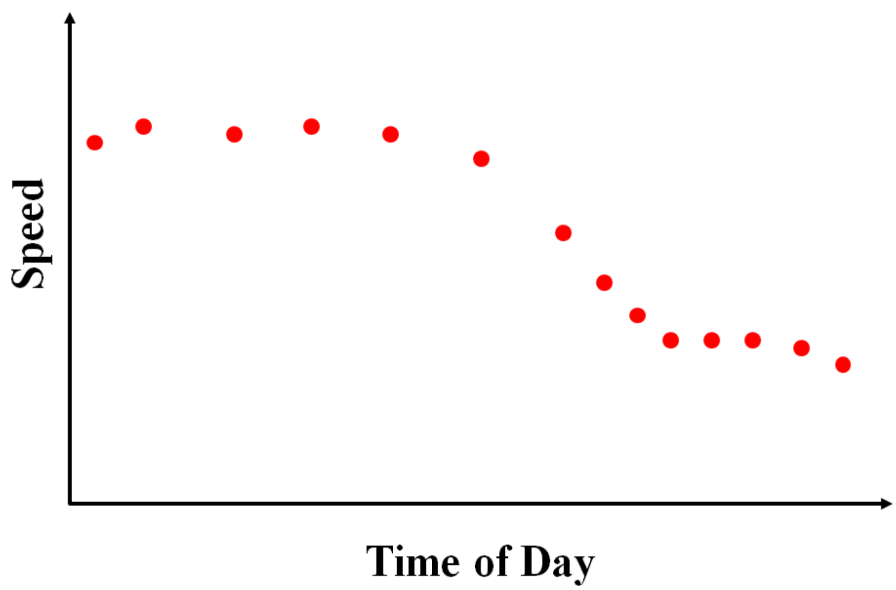

Figure 1.6: Hypothetical GPS probe vehicle speed data as a function of time of day

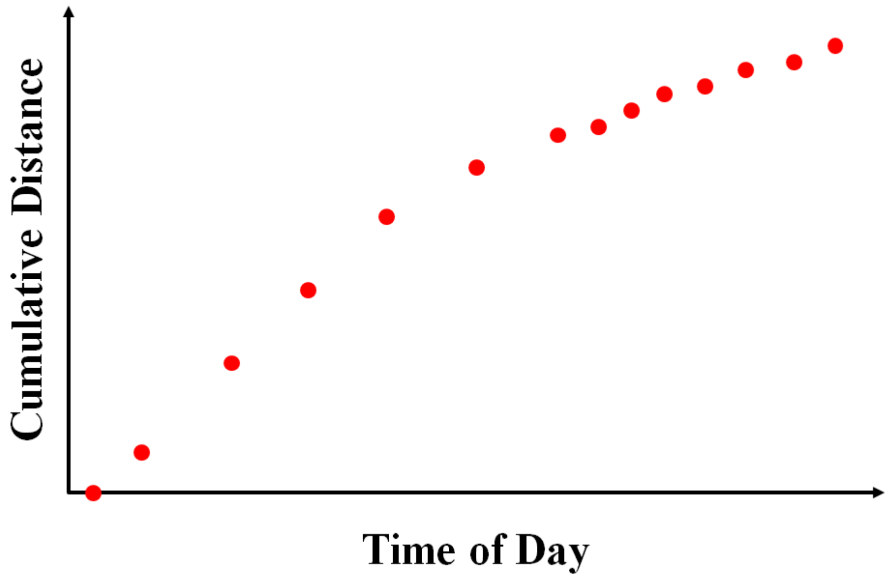

Figure 1.7: Hypothetical GPS probe vehicle cumulative distance data as a function of time of day

Despite the potential of using Bluetooth technology to collect travel time data, transportation agencies looking to implement signal retiming measures are concerned not only with the level of detail of the resulting data but also with cost. KMJ Consulting (2011) compared the costs of floating car probe and Bluetooth technologies in terms of equipment, labour and mileage costs. It was found that the cost of Bluetooth data collection was substantially higher for a road segment three miles in length and a data collection period of three days. In terms of dollar values, the sum of these three expenditures resulted in approximately $12,000 to collect data via Bluetooth versus $4,000 to collect data via probe runs for this project. However, the total cost per data point was substantially lower for Bluetooth technology due to the ability to collect a much greater number of travel times over a given time period. It was also found in the same study that Bluetooth data collection becomes more cost effective than probe vehicle data collection as the number of roadway segments involved in the study increases. Young et al. (2008) also reported that the cost per data point for Bluetooth technology is between 500 and 2500 times cheaper than that of floating car probe technology. The lower incremental costs of Bluetooth technology in comparison to floating car probe technology are reiterated by Quayle et al. (2010).

KMJ Consulting (2011) and Moghaddam (2014) summarized the advantages and disadvantages of Bluetooth technology compared to floating car probe technology. The major advantages and disadvantages of using Bluetooth detectors to collect travel times are summarized below.

The major advantages include:

- The ability to collect a large number of travel times over a given time period
- Continuous data collection (i.e. Bluetooth detectors can be deployed for the entire duration of the data collection period without interruption)
- The opportunity to repurpose the data for use in origin-destination studies
- The lack of bias associated with the data (i.e. the travel times belong to vehicles that are not inherently dedicated to collecting data for a travel time study)
- No need to acquire GPS satellite signals, therefore data can be collected in urban canyons and tunnels or other locations where GPS satellite signal obstruction occurs

The major disadvantages include:

- The high cost when limited and infrequent data collection is needed
- Sampling bias due to the detection of multiple Bluetooth devices within a vehicle
- The need to address extreme values in the data through outlier detection due to greater variability in travel times
- The learning curve associated with using a new technology for practitioners who are familiar with probe vehicle data
- The required access to infrastructure for mounting and powering the equipment (or the additional cost of a portable power source)
- The coarse level of detail of the data (i.e. data are acquired only when a vehicle is within the detection zone instead of along the entire route of the vehicle)

1.2 Problem Statement

There is a consensus among traffic engineering professionals that sub-optimal signal timings degrade the performance of arterial roadways by inefficiently allocating the available capacity. Given that traffic patterns change over time, it is necessary to re-evaluate the optimality of signal timings of key arterial corridors on a relatively frequent basis. With the emergence of Bluetooth as a viable technology for collecting travel time data, the challenge is to determine the most appropriate way to use the resulting travel time data to evaluate signal retiming measures. Although previous studies have examined this challenge, there is not yet a satisfactory solution as these studies generally suffer from at least one of the following limitations:

- They recommend multiple measures of effectiveness and it is unclear as to which one(s) should be used by practitioners to make decisions (Day et al., 2012; Gettman et al., 2013).
- They recommend methods that have sound theoretical bases but are difficult for practitioners to present to the public and policymakers (Day et al., 2012).
- They recommend methods that are attractive to decision makers but have not been proven to be statistically valid (Koonce et al., 2008; Tindale-Oliver & Associates, 2008; Quayle et al., 2010; Lee County Department of Transportation, 2008; Wang et al., 2010).

Furthermore, the following issues have also been found with respect to past studies involving signal retiming evaluation:

- Little to no justification is provided for the selection of the values of key study parameters such as the duration of the data collection period or the number of travel times required for the evaluation of signal retiming measures (Quayle et al., 2010; Wang, 2014). Furthermore, the effect that these parameters have on the resulting Bluetooth travel time data has not been adequately studied.
- The effect that abnormally high travel times (i.e. outliers) have on the measures of effectiveness used to assess signal retiming measures has not been adequately studied.

The need for a practitioner-ready methodology detailing how to use these data to evaluate signal retiming measures is clear. This research details the exploration of Bluetooth arterial travel time data and the development of a methodology to analyze these data. To ensure the robustness of this methodology, travel time data collected in the field as well as in a simulated environment are examined.

1.3 Research Questions and Objectives

To frame the investigation of Bluetooth arterial travel time data for the purposes of signal retiming evaluation, the following four research questions and four objectives are presented.

With Bluetooth travel time data available at different scales in this study, the viability of each data set must be explored. One data set contains the travel times of vehicles that traversed the entire corridor (path-level data) while the other contains the travel times of vehicles that have traversed smaller segments of the corridor (link-level data). The path-level data are attractive in that they require fewer detectors and therefore less time and resources for collection. However, far fewer travel times are typically available at this scale compared to the link-level data because a smaller number of vehicles travel the entire corridor compared to the number of vehicles that travel a single link. Taking these characteristics into account, the viability of the path-level data as a source of information for evaluating signal retiming measures must be examined in comparison to the link-level data.

Research question 1:

Does the smaller path-level travel time data set contain enough information to make conclusions regarding the effectiveness of a signal retiming project as an alternative to the much larger link-level travel time data set?

Research objective 1:

Explore the Bluetooth travel time data at each scale and determine whether or not the large-scale, small-sample path-level data are an adequate substitute for the small-scale, large-sample link-level data by identifying whether or not statistically significant conclusions can be made using the path-level data. Compare these results to the results of the application of this same methodology using the link-level data. For verification, perform this analysis on both field and simulated data.

Knowing that Bluetooth technology is a viable option for collecting link-level travel time data, the values of the parameters that have significant implications on the cost and resources associated with evaluating signal retiming measures must be defined. More specifically, the following variables must be quantified: when and for how long detectors must be deployed, the location and spacing of detectors, and the number of travel time samples required to produce conclusive results. Recommendations for these parameters are essential to any practitioner considering the use of Bluetooth technology in a signal retiming study.

Research question 2:

Where and how far apart should Bluetooth detectors be placed, what should the characteristics of the data collection period be, and how long should the data collection period be to collect enough travel time samples to assess the success or failure of a signal retiming study?

Research objective 2:

Identify values that have been used in previous studies and build upon these to provide recommendations for the placement of detectors and the characteristics of the data collection period. Using the field data, develop a model that can be used to determine the duration of the data collection period and an adequate detector spacing using the characteristics of the corridor as input.

Measures of effectiveness are vital to proving the worth of any study intended to improve traffic operations. While measures of effectiveness that employ Bluetooth travel time data have been explored, practical and definitive measures of effectiveness that can be used in confidence by practitioners has yet to be developed. The validity of any developed measures of effectiveness must also be ensured from a statistical standpoint to add credibility to the conclusion that a signal retiming study has successfully improved traffic operations.

Research question 3:

What measures of effectiveness should be used to measure the improvements facilitated by the implementation of a signal retiming study? How should the use of these measures be verified?

Research objective 3:

Develop measures of effectiveness using Bluetooth travel time data that can be used by practitioners to definitively confirm (or refute) the effectiveness of the implemented signal retiming measures. Ensure that these measures of effectiveness accounts for the number of vehicles that experience the implemented signal retiming measures and that they account for the statistical significance of the results. Test these measures of effectiveness using Bluetooth travel time data collected from the field and in a simulated environment. Verify that this measure of effectiveness produces reasonable results by applying it to the true travel time data from which the simulated Bluetooth and GPS probe vehicle data were drawn.

A major difference between floating car probe data and Bluetooth data is the higher frequency of extreme values (i.e. outliers) within the Bluetooth data that are virtually nonexistent in floating car probe data. In order for valid conclusions to be made for a signal retiming evaluation that utilizes Bluetooth travel time data, the effect that these extreme values have on the previously defined measures of effectiveness must be quantified. In the context of this investigation, an outlier is defined as a travel time that it is identified not to be representative of the traffic stream for which signal retiming is intended. In this investigation, an outlier is defined only by the magnitude of its travel time.

Research question 4:

What is the effect that outliers have on a signal retiming project that utilizes Bluetooth travel time data?

Research objective 4:

Quantify the effect that outliers have on the recommended measures of effectiveness using the field Bluetooth travel time data.

1.4 Thesis Outline

This thesis is organized as follows:

Chapter 2 contains a review of the literature related to the research objectives.

Chapter 3 examines the characteristics of the field and simulated data explored in this research.

Chapter 4 examines the viability of the path-level data in evaluating signal retiming measures from a statistical perspective. These results are then compared to the link-level data, which is examined using the same methodology.

Chapter 5 examines detector location and spacing, the characteristics of the data collection period, and the duration of the data collection period. Based on guidance from previous studies and the characteristics of the field data and the corridor from which these data were collected, recommendations are provided for each of these parameters.

Chapter 6 describes the development of path-level measures of effectiveness using link-level Bluetooth travel time data that can be used to evaluate signal retiming measures. These measures of effectiveness are then tested using field and simulated Bluetooth travel time data.

Chapter 7 examines the influence that outliers have on the results of the produced measures of effectiveness for the field Bluetooth travel time data.

Chapter 8 summarizes the main conclusions made from the results of this investigation as well as recommendations for future studies involving these results.

2 Literature Review

2.1 Introduction

This chapter presents a review of previous studies relevant to the research objectives of this investigation. First, studies that have examined the relationship between link-level travel times and path-level travel times are presented. Measures of effectiveness utilized in practice and in research for travel time studies are then explored. Studies that have examined the influence of extreme values in travel time studies are then reviewed. The location and spacing of detectors and the characteristics and duration of the data collection period in previous studies are then discussed. This chapter concludes with a summary of the major points of this literature review.

2.2 Path-Level and Link-Level Travel Time Data

Unlike freeway facilities, arterials contain multiple impedances to the progression of traffic along the corridor, including traffic signals. The variability of the travel times of vehicles that traverse a corridor increases as the length of the arterial increases due to a greater number of opportunities to encounter traffic signals or other impedances. The following studies provide insight into the influence that the segmentation of the study area has on the resulting data set.

Li, Chai and Tang (2013) examined the relationship between links along an arterial and the path to which those links belong. The travel time data in their study consisted of one day of data divided into 30-minute increments for five different paths with each path consisting of two links. For a given path or link, the mean and variance of the travel time data were calculated according to standard calculation methods, as shown in Equations 2.1 and 2.2:

$$\bar{t} = \frac{1}{w} * \sum_{v=1}^{w} t_v \tag{2.1}$$

$$s = \sqrt{\frac{1}{w} * \sum_{v=1}^{w} (t_v - \bar{t})^2} \tag{2.2}$$

where $\bar{t}$ represents the mean travel time for the given path or link, t_v represents the travel time of vehicle v belonging to the set of all w vehicles that traversed the given path or link, and s represents the standard deviation of travel times for the given path or link.

Travel times at the path level produced using the link-level data were then created, which are defined as "link-to-path" travel times. The mean link-to-path travel time was calculated as the sum of the link-level mean travel times along the path that those links belonged to. Three methods of calculating the link-to-path variance of travel times were tested and it was concluded that the most accurate method was simply the sum of the link-level variances of travel times along the path. The accuracy of each calculation method was assessed by calculating the differences between these link-to-path measures and the actual path-level measures. This comparison implies that the actual path-level variances of travel times were assumed to represent the ground truth despite having a much smaller number of travel times to draw from compared to the link-level variances of travel times. Summing link-level variances of travel times to produce the link-to-path variance of travel times was recommended because it yielded the lowest mean absolute error relative to the path-level variance of travel times in most cases.

A key assumption associated with the simple summation of variances of travel times across links is that links belonging to the same path are statistically independent. Rakha et al. (2006) reasoned that this assumption is inaccurate because traffic conditions along a path propagate across multiple links. Their study showed that the summation of link-level variances of travel times underestimates the path-level variance of travel times significantly on a freeway facility in comparison to the actual path-level variance of travel times. Consequently, the authors recommended a different calculation method, one that produced the lowest error relative to the path-level variance of travel times. The recommended methodology estimated the link-to-path variance of travel times according to Equation 2.3:

$$\sigma_{LP}{}^2 = \left(\frac{\bar{T}}{j} * \sum_{i=1}^{j} \frac{\sigma_i}{\bar{t}_i}\right)^2 \tag{2.3}$$

where $\sigma_{LP}{}^2$ represents the link-to-path variance of travel times, $\bar{T}$ represents the mean path travel time, and σ_i and $\bar{t}_i$ are the standard deviation of travel times and mean travel time for link i belonging to the set of all links j that make up the path of interest.

Day et al. (2012) examined the amount of variability present in Bluetooth travel time data while varying the length of the arterial segment for which the data were collected. This study found that the variability of the travel time data increases as the length of the arterial segment increases, which is indicated by the development of a shallower slope in the cumulative travel time distribution. In other words, a shallower slope indicates that a greater variety of travel times is present within the distribution. This is illustrated by the cumulative distributions of travel time in Figure 2.1 with “BMS” indicating a Bluetooth monitoring station (i.e. a detector). A larger difference between detector indices indicates a longer segment of the arterial (for example, BMS-1 to BMS-3 represents a longer segment than that of BMS-1 to BMS-2).

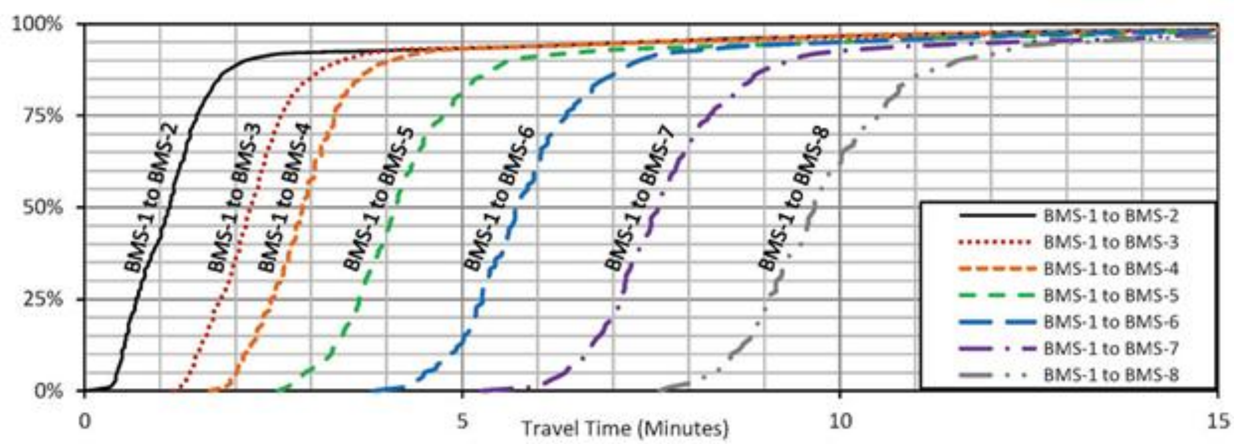

Figure 2.1: Cumulative distributions of travel time as a function of segment length (Day et al., 2012)

The comparison of variability between different arterial segments was also examined using control delay as a measure of effectiveness in this study. Control delay was defined as the difference between the travel time along a given segment and the time required to travel that segment at the free flow speed. By calculating the control delay for every travel time of a segment, a distribution of the control delay can be created for that segment. This method effectively normalizes delay so that arterial segments of varying lengths can be compared to each other. A segment that has a control delay cumulative distribution with a shallow slope contains a greater frequency of longer control delay values and can therefore be identified as a candidate for signal retiming. As shown in Figure 2.2, the cumulative distributions of control delay for segments BMS-1 to BMS-2 and BMS-4 to BMS-5 indicate greater frequencies of longer control delay values. Therefore, these segments can be identified as detrimental to the progression of traffic along the corridor that these segments belong to.

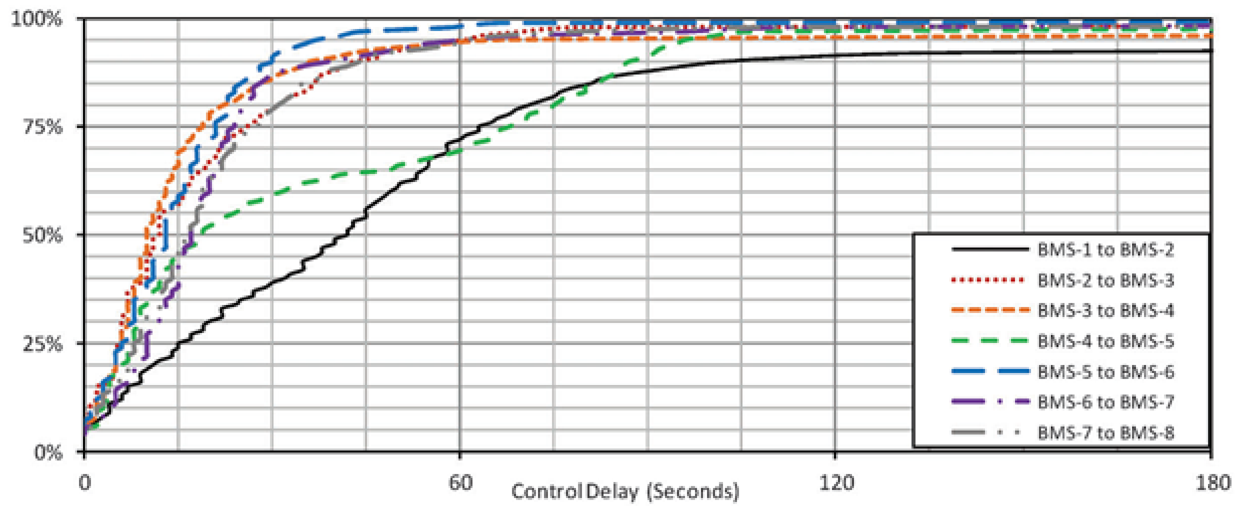

Figure 2.2: Cumulative distributions of control delay as a function of segment length (Day et al., 2012)

2.3 Detector Location and Detector Spacing

In a comprehensive study of detector location, Day et al. (2012) recommended the placement of detectors at midblock locations as opposed to at an intersection. This was recommended because a midblock detector is far enough removed from an intersection such that the resulting travel times are less affected by vehicle interactions at intersections. Minimizing the influence that vehicle interactions have on the resulting travel time data is desired to capture the travel times that vehicles experience along the corridor as a result of the quality of progression along the network, not the travel times that vehicles experience as a result of vehicle interactions. However, midblock detectors are less feasible from a practical standpoint as they require portable power sources while placing detectors at intersections allows for a connection to an existing power source.

Most studies that examine detector spacing optimization involve loop detectors and freeways, including studies by Ban et al. (2009) and Edara et al. (2008). Oh and Choi (2004) examined the placement of loop detectors in an arterial environment to estimate link speed and found link length and green time to be the primary variables that influenced detector spacing. Sherali, Desai and Rakha (2006) recommended that detectors should be placed where the geometry of the roadway changes to capture the resulting variation in travel times. Using the resulting data to predict travel time on a short-term basis on arterial facilities, Moghaddam (2014) used three Bluetooth detectors spaced 1.9 kilometres and 1.2 kilometres apart. In terms of location, two detectors were placed at minor signalized intersections and one was placed at a highway on-ramp. With the intention of optimizing Bluetooth detector spacing based on travel time prediction accuracy for freeway facilities, Hu (2013) found the optimal detector spacing to be between two kilometres and five kilometres depending on the length of the study area. In developing a method to minimize travel time estimation error for freeway facilities, Haghani et al. (2010) utilized a Bluetooth detector spacing of 1.6 kilometres. A Bluetooth detector spacing of 4.8 kilometres between the two detectors was utilized in the aforementioned study by Day et al. (2012). Provided that each segment has consistent operating characteristics and geometric cross sections, Turner et al. (1998) recommended a sampling segment length of no longer than 3.2 kilometres for principal arterials.

2.4 Data Collection Period and Sample Size

While it is logical that the Before and After data collection periods should share the same duration and cover the same days of the week, there is no rationale provided in signal retiming studies in the literature for the duration of the data collection period. For the floating car probe data collection method, Wang (2014) stated that the City of Toronto recommends that the number of trajectories required for each period is to be collected over a period of time no greater than three days for each of the Before and After periods. In another GPS probe-based study, Wang et al. (2010) used one day of data to represent each of the Before and After data sets. In a signal retiming study evaluated using Bluetooth data, Day et al. (2010) were interested in evaluating signal retiming measures for a single Saturday, so data collected for a single Saturday were utilized for each of the Before and After periods. Quayle et al. (2010) used ten days of data during the months of February and March to represent each period to evaluate signal retiming measures using Bluetooth data.

One variable that is a function of the duration of the data collection period is the amount of data that a practitioner must collect, which is typically referred to as the sample size. When evaluating signal retiming studies using data collected via the floating car probe method, sample size is often neglected due to cost (Gettman et al., 2013). Instead, practitioners abide by a fixed number of probe runs (Wang, 2014). The opportunity provided by Bluetooth technology to collect a large number of travel times allows for a more thorough assessment of the amount of data required for a signal retiming study.

Assuming normality, Turner et al. (1998) recommend the central limit theorem as a basis for estimating the required sample size. Equation 2.4 and Equation 2.5 illustrate this calculation:

$$n = \left(\frac{z_{\alpha/2} * COV}{e}\right)^2 \quad (2.4)$$

$$COV = \frac{\sigma}{\mu} \quad (2.5)$$

where n represents the required sample size, $z_{\alpha/2}$ represents the z-statistic for a given level of confidence α assigned by the user, COV represents the coefficient of variation, e represents the maximum permitted error in the estimated mean travel time (expressed as a percentage of the mean) assigned by the user, σ represents the standard deviation of travel times, and μ represents the mean travel time.

According to Turner et al. (1998), the required sample size for travel time studies that utilize license plate matching technology (which is akin to Bluetooth technology) to collect travel time data is dependent upon the coefficient of variation. The coefficient of variation is in turn dependent upon the type of roadway facility (freeway or arterial), the amount of traffic that the facility experiences and the time period over which travel time data are collected. Based on certain values of each of these variables, a range of coefficients of variation can be determined using field trials, which can then be used to determine the required sample size. KMJ Consulting

(2011) also recommended the use of the coefficient of variation to determine the required sample size.

Table 2.1 shows the coefficients of variation determined through various studies for license plate matching technology according to various amounts of traffic and time periods. For comparison, Table 2.2 and Table 2.3 show the coefficients of variation for dedicated GPS probe vehicle technology as a function of average daily traffic and traffic signal density, respectively. In comparison to the required sample size for floating car probe data, the required sample size for license plate matching data is much higher. The rationale for this is that the license plate matching technology produces a much more diverse set of travel times due to greater variability in driver behaviour (Turner et al., 1998). This idea is also applicable to Bluetooth technology due to its similar functionality as a fixed detector and due to the possibility of Bluetooth devices not associated with a vehicle being available for detection.

Table 2.1: Sample size as a function of traffic conditions and time period for license plate matching travel time data (Turner et al., 1998)

Traffic Conditions and Time Period	Average Coefficient of Variation (%)	Sample Size		
		90% Confidence, ± 10% Error	95% Confidence, ± 10% Error	95% Confidence, ± 5% Error
Low to moderate traffic, 15- to 30-minute period	10	4	5	18
Low to moderate traffic, 1- to 2-hour period	20	12	18	62
Congested traffic, 15- to 30-minute period	25	18	27	96
Congested traffic, 1- to 2-hour period	35	34	48	189

Table 2.2: Sample size as a function of average daily traffic for dedicated GPS probe vehicle travel time data (Turner et al., 1998)

Average Daily Traffic (ADT)	Average Coefficient of	Sample Size		
		90% Confidence,	95% Confidence,	95% Confidence,

(volume per lane)	Variation (%)	± 10% Error	± 10% Error	± 5% Error
Less than 15,000	9	5	6	15
15,000 to 20,000	11	6	8	21
Greater than 20,000	17	10	14	47

Table 2.3: Sample size as a function of traffic signal density for dedicated GPS probe vehicle travel time data (Turner et al., 1998)

Traffic Signal Density (signals per mile)	Average Coefficient of Variation (%)	Sample Size		
		90% Confidence, ± 10% Error	95% Confidence, ± 10% Error	95% Confidence, ± 5% Error
Less than 3	9	5	6	15
3 to 6	12	6	8	25
Greater than 6	15	9	23	37

Quiroga and Bullock (1998) examined the results of four methods of estimating the required sample size for floating car probe runs that use the central limit theorem as a basis. The method outlined in the Institute of Transportation Engineers manuals (commonly known as the ITE method) is shown in Equations 2.6 and 2.7:

$$n = \left(\frac{z_{\alpha/2} * \bar{R}}{d * \varepsilon}\right)^2 \qquad (2.6)$$

$$d = \frac{\bar{R}}{\sigma} \qquad (2.7)$$

where $z_{\alpha/2}$ represents the z-statistic for a given level of confidence α assigned by the user, $\bar{R}$ represents the average sample range, d represents the ratio of the average sample range $\bar{R}$ to the standard deviation of travel times σ, and ε represents the allowable error assigned by the user. Since d is a function of n, a value of d must be assumed using Oppenlander's (1976) table of values for d as a basis and n must be solved for iteratively. Noting that $\varepsilon = e * \mu$, it can also be shown that Equation 2.6 is equal to Equation 2.4:

$$\left(\frac{z_{\alpha/2} \ * COV}{e}\right)^2 = \left(\frac{z_{\alpha/2} * \bar{R}}{d * \varepsilon}\right)^2$$

$$\frac{z_{\alpha/2} \ * \left(\frac{\sigma}{\mu}\right)}{e} = \frac{z_{\alpha/2} * \bar{R}}{\left(\frac{\bar{R}}{\sigma}\right) * (e * \mu)}$$

$$\frac{z_{\alpha/2} \ * \sigma}{e * \mu} = \frac{z_{\alpha/2} \ * \sigma}{e * \mu}$$

The ITE method calculates the average sample range $\bar{R}$ as the sum of all of the absolute differences between consecutive travel time entries, as shown in Equation 2.8:

$$\bar{R} = \frac{1}{n-1} * \sum_{i=2}^{n} |t_i - t_{i-1}| \qquad (2.8)$$

where i represents the index of a travel time belonging to the set of n travel times and t_i represents the travel time for index i. This method was found to be biased because consecutive differences imply that each travel time entry (other than the very first and very last entries) is used twice, causing $\bar{R}$ to be underestimated. To avoid this bias, Oppenlander (1976) recommended calculating $\bar{R}$ as the difference between the maximum and minimum travel time entries for the ITE method, as illustrated by Equation 2.9:

$$\bar{R} = \max(t_i) - \min(t_i) \qquad (2.9)$$

A key assumption of this method is that the relationship between the average sample range $\bar{R}$ and the standard deviation of travel times σ is constant. However, it can be shown through an example that Equation 2.9 will always produce a larger value than Equation 2.8 and that Equation 2.9 is very sensitive to extreme values. Consider the situation in which there are four observed travel times available with durations of 3, 6, 10 and 8 minutes. From Equation 2.8, the following is obtained:

$$\bar{R} = \left(\frac{1}{3}\right) * (|6-3| + |10-6| + |8-10|) = \frac{9}{3} = 3$$

From Equation 2.9, the following is obtained:

$$\bar{R} = 10 - 3 = 7$$

The standard deviation method is presented by Quiroga and Bullock (1998) as an alternative to the ITE method and Oppenlander's method, which both consistently underestimate the required sample size. The standard deviation method is shown in Equation 2.10:

$$n = \left(\frac{t_{\alpha/2} * s}{\varepsilon}\right)^2 \tag{2.10}$$

where $t_{\alpha/2}$ represents the t-statistic for a given level of confidence α assigned by the user and s represents the standard deviation of travel times. The use of the t-statistic is more suitable for estimating floating car probe runs than the z-statistic because of the relatively small amount of probe runs that can be feasibly collected for a signal retiming study.

A drawback of this method is that the standard deviation of travel times s is regarded as a somewhat abstract statistic that is difficult for field personnel to understand, which implies that this method is not as user-friendly as methods that utilize the more intuitive average sample range $\bar{R}$ (Quiroga and Bullock, 1998). Consequently, the authors combine the use of the average sample range $\bar{R}$ and the use of the t-statistic to yield a hybrid formula, which is shown in Equation 2.11:

$$n = \left(\frac{t_{\alpha/2} * \bar{R}}{d * \varepsilon}\right)^2 \tag{2.11}$$

Although this method produces estimates of the required sample size that are very close to those estimated using the standard deviation method, this method still requires an iterative procedure to solve for n.

Ernst et al. (2012) estimated the required sample size based on the similarity of the Before and After travel time distributions. To measure this similarity, the Kullback-Liebler (KL) divergence was recommended. This parameter is a nonparametric measure of the difference between two distributions that compares the contents of each bin of the percent relative frequency plot for the Before data to those of the After data. The KL divergence, D, is calculated using Equation 2.12:

$$D = \sum_{i=1}^{N} P(i) * \log_2 \left(\frac{P(i)}{Q(i)} \right) \tag{2.12}$$

where i represents the bin of the percent relative frequency plot belonging to the entire set of N bins, $P(i)$ represents the relative frequency for entries in bin i for the first distribution, and $Q(i)$ represents the relative frequency for entries in bin i for the second distribution. According to this methodology, D will be small for two similar distributions and will be large for two dissimilar distributions. After using this formula, the required sample size n is estimated with the KL divergence parameter as input using Equation 2.13:

$$n = -\frac{\log_2(P)}{D} \tag{2.13}$$

where P represents the probability of error assigned by the user and D represents the KL divergence. Ernst, Krogmeier and Bullock (2014) demonstrated that Oppenlander's methodology is a subset of the more generalized KL divergence methodology.

One flaw in this methodology is that D and consequently n change based on which distribution is defined to be $P(i)$ and which is defined to be $Q(i)$. Therefore, the results for when $P(i)$ is defined to be the Before distribution will be different than for those when $P(i)$ is defined to be the After distribution.

A common theme among many of these methods of estimating the required sample size is that they require the collection of travel time data as input. The collection of travel time data before the required sample size is estimated is difficult to justify on a limited budget in the event that a practitioner is forced to redeploy Bluetooth sensors for additional data collection. One of the few studies to address sample size specifically for Bluetooth travel time data claims that sample size is a non-issue (Cambridge Systematics, 2012) because the use of Bluetooth detectors results in much larger samples relative to the floating car probe method and that if the penetration rate (the percentage of the vehicle population that is Bluetooth-enabled) is at least 5%, the resulting sample size will be adequate. However, it is clear that the sample size is a function of the length of time that the detectors are deployed, so this conclusion is contingent upon the duration of the data collection period.

2.5 Measures of Effectiveness for Travel Time Data

To quantify the success of a signal retiming study, reliable measures of effectiveness are required. Current practice involves the measurement of the differences between the Before and After mean travel times for a given link of a corridor that has had its signals retimed. The prevalence of this strategy was observed by Koonce et al. (2008) and by Tindale-Oliver & Associates (2008).To visualize the impacts of signal retiming measures, the Lee County Department of Transportation (2008) study utilized a diagram showing the difference and percent difference along a corridor for each link. A sample diagram from this study is shown in Figure 2.3. In this figure, the difference between mean travel times for a link is represented by the "average time savings" and the percent difference between mean travel times is represented by the "percent reduction". Link-level mean travel times were then summed to produce the link-to-path mean travel time. In this study, it was concluded that the implemented signal retiming measures were successful based on positive travel time savings and percent reductions on the majority of links.

In a slight variation, Wang et al. (2010) compared median travel times instead of mean travel times in an arterial setting, although no rationale for this deviation from typical practice was provided.

One pilot study that primarily used Bluetooth data in the evaluation of signal retiming measures was a study performed by Quayle et al. (2010). This study utilized 20 days of data (February 10 to 19, 2009 for the Before period and March 20 to 30, 2009 for the After period) that were continuously collected by roadside Bluetooth detectors. While a total of three detectors collected travel time data along a 4-kilometre arterial, it was not specified whether the travel time data used to produce final results were collected at the path level or at the link level. The difference between mean travel times was the primary measure of performance in this study, as illustrated in Table 2.4. This study found that the reductions in the mean were large enough to conclude that the signal retiming study was successful for both the morning and evening peak periods and for both directions. It must be noted that statistical testing was not used in this study to confirm that the observed changes in travel time were statistically significant.

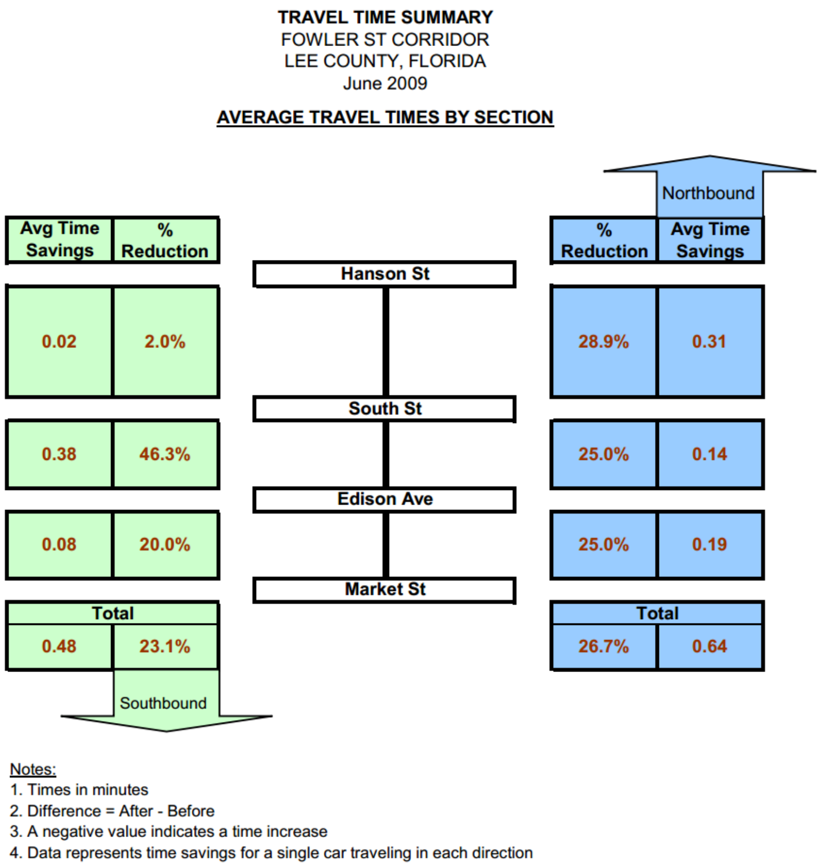

Figure 2.3: Sample measure of effectiveness diagram used in practice (Lee County DOT, 2008)

Table 2.4: Results of a sample Bluetooth signal retiming study evaluation (Quayle et al., 2010)

Direction	Period	Signal Timing	Sample Size	Average Travel Time (min:s)	Standard Deviation (min:s)	Difference in Means (s)
Westbound	7-9 a.m. peak	Before	423	8:50	6:25	-100
		After	203	7:10	7:30	
Eastbound		Before	338	8:05	5:30	-110
		After	96	6:15	4:20	
Westbound	4-6 p.m. peak	Before	374	11:30	6:55	-250
		After	172	7:20	5:25	
Eastbound		Before	398	9:45	6:50	-55
		After	105	8:50	5:30	

A study by Day et al. (2012) is the second and most recent study found in the literature that used Bluetooth travel time data to evaluate signal retiming measures. A variety of measures of effectiveness were calculated in this study including the differences between means, medians, standard deviations, percentiles, interquartile ranges and travel time distributions. The median of the travel time distribution was argued to be a better reflection of the central tendency of the travel time distribution than the mean, especially on longer corridors with greater variability.

According to Lomax et al. (2003), travel time reliability is the amount of consistency or variability in transportation service that a facility experiences over a given time period. As described by Li, Chai and Tang (2013), high variability indicates increased unpredictability and decreased reliability. Despite many travel time studies using the variance or standard deviation of travel times, Day et al. (2012) recommend the use of the interquartile range (i.e. the difference between the 75^{th} and 25^{th} percentiles) as a measure of travel time reliability. Success is indicated by a reduction in the interquartile range (i.e. a higher frequency of shorter travel times is evident for the After period compared to the Before period).

Other measures that represent alternatives to the standard deviation have also been developed. Lomax et al. (2003) recommended the "misery index", which is the difference between the mean of the top 20^{th} percentile travel times and the mean travel time. Van Lint and Van Zuylen (2005) recommended the ratio of the difference between the 90th and 50th percentiles and the difference between the 50th and 10th percentiles.

Day et al. (2012) emphasized the importance of the distribution of Bluetooth travel time data due to the larger number of travel times that Bluetooth technology provides compared to the number of floating car probe runs that can be collected over the same period of time. In other words, because more data can be collected using Bluetooth technology, a larger proportion of the traffic stream's travel time behaviour can be captured and this larger data set must be examined further. Comparing the cumulative frequency diagrams (CFDs) of the Before and After data, improvement is indicated by a steeper CFD curve for the After data (an indicator of less variability) as opposed to a shallower CFD curve for the Before data (an indicator of greater variability). This is shown in Figure 2.4 in which the thin-lined cumulative travel time distribution representing the After data has a steeper slope than the thick-lined cumulative travel time distribution representing the Before data. This indicates that there has been an increase in the frequency of shorter travel times between the Before and After periods, contributing to an overall shift in the distribution to the left from the Before period to the After period. It can also be observed in this figure that the distribution of travel times is non-normal as each distribution has a longer tail to the right than it does to the left. This is important to note when statistical testing is considered as a measure of effectiveness.

Day et al. (2012) recommended the use of statistical tests involving the travel time distribution and its components. This includes the t-test to compare the means of the Before and After travel time distributions. Nonparametric tests such as the Mann-Whitney U-test and the Kolmogorov-Smirnov test were recommended over the parametric t-test due to the assumption of normality associated with parametric tests. Nonparametric tests do not make this assumption and are therefore claimed to be more appropriate for evaluating a non-normal distribution such as that of arterial travel times. Despite the many measures of effectiveness examined by Day et al. (2012), no definitive measures were provided to allow a practitioner to make conclusions on the success of a signal retiming study.

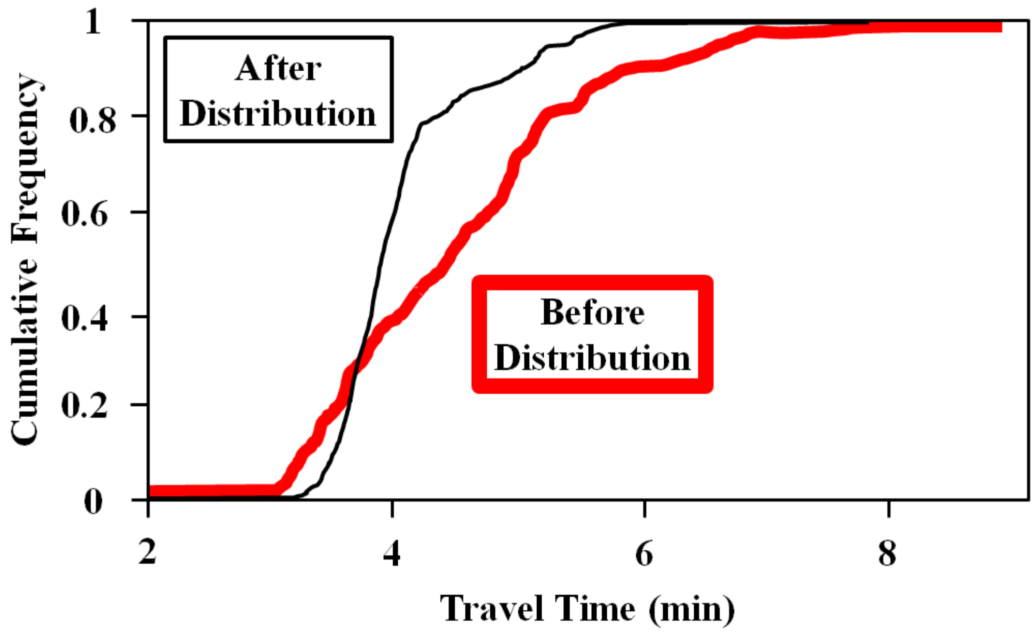

Figure 2.4: Comparison of sample cumulative travel time distributions (Day et al., 2012)

Li, Chai and Tang (2013) investigated the practicality of the assumption of normality for arterial travel time data. This study examined a day of data obtained from licence plate matching technology on an arterial facility split into 15-minute increments. The travel time distribution of each 15-minute increment for each link that made up the path was then determined. After excluding 15-minute intervals that did not have enough data for testing, four distributions were fitted to each travel time data set: normal, log-normal, gamma and Weibull. The percent differences between the results of the assumption of normality compared to the assumption of one of the other three distributions were then calculated for the mean and for the standard deviation. Based on the relatively small values of these percent differences, this study concluded that the normal distribution is acceptable for estimating the mean and standard deviation of travel time under most traffic situations. This was concluded with the caveat that the normal distribution is not recommended in situations in which high accuracy is required.

Rakha et al. (2006) stated that the assumption of normality for the distribution of travel times is inaccurate from a theoretical standpoint. While goodness-of-fit tests revealed that the normal distribution did not adequately fit the travel time distribution in this study, it was concluded that this was due to outliers in the automatic vehicle identification (AVI) data collected within the long tail to the right of the distribution. The authors also concluded that despite not being a good assumption from a theoretical standpoint, the assumption of normality is reasonable from a practical standpoint. It must be noted that that travel time data examined in this study were collected on a freeway facility.

2.6 Influence of Outliers on Fixed Detector Travel Time Data

Various outlier detection methods have been developed for roadside detector technology. Araghi et al. (2012), Quayle et al. (2010) and Iteris (2011) used static upper and lower bounds. Rakha et al. (2006) used a series of heuristic measures to filter AVI data collected on a freeway facility. To filter the data, upper and lower filtering limits were constructed using the collected sample size, the number of heuristically identified outliers, and the sequence of travel times along a segment as input. Other approaches to filtering for outliers have incorporated the median absolute deviation (Khoei, Bhaskar and Chung, 2013), the modified z-test approach (Vo, 2011), the interquartile range (Li, Chai and Tang, 2013), and time series modeling (Roth, 2010). Moghaddam, Noroozi and Hellinga (2014) described various methods of outlier detection used in the past for roadside detector technology. This included the use of percentiles, deviation tests (i.e. a validity window) and adaptive filtering (deviation tests that incorporate observed trends in past travel times). A real-time outlier detection algorithm for arterial roadways was then proposed in this study, which combined short-term and long-term trends in travel times to form a validity window for data to be filtered through.

To examine the influence that outliers have on the results of a travel time study, Boxel et al. (2011) developed a real-time filtering algorithm based on confidence intervals and compared the results of the Shapiro-Wilk normality test before and after the data were filtered for outliers. Applying this test to both freeway and arterial facilities, it was found that the travel time distribution was non-normal before filtering and normal after filtering. Other than this study, it was found that the desire to create a realistic outlier filtering algorithm exists throughout the literature but the influence that outliers have on the results of a travel time study of any nature has not been adequately studied.

2.7 Summary

In this chapter, four key aspects were examined in the literature: the use of path-level data versus link-level data, the values of key characteristics of before-and-after studies including the duration of the data collection period and detector location and spacing, measures of effectiveness used to evaluate signal retiming measures, and the influence of outliers on fixed detector travel time data. The following findings were concluded from this literature review:

- Clear rationale has been provided for the placement of detectors and a wide variety of methods for determining detector spacing have been developed for various purposes. While the characteristics of the data collection period can be identified through reasoning, no clear rationale has been provided for the duration of the data collection period. Rules of thumb are standard practice in this respect.
- Various methods of estimating the required sample size for a travel time study are available, including the use of the central limit theorem, coefficients of variation, the average sample range and the Kullback-Liebler divergence. However, these methods require the collection of travel time data before the required sample size is estimated, which is difficult to justify on a limited budget when a practitioner is forced to redeploy Bluetooth sensors for additional data collection.
- Methods with clear rationale are available for calculating measures of effectiveness at the path level using link-level data. The methodology for calculating the path-level mean travel time using link-level travel times is clear while there are several methods available for calculating the path-level standard deviation of travel times using link-level standard deviations of travel times.
- Many measures of effectiveness currently used in practice have little to no statistical basis due to their reliance upon floating car probe data, whose typical sample size renders a travel time study's cost to be too high for statistical significance to be considered.
- A wide variety of measures of effectiveness are available for assessing the impacts of signal retiming measures including the differences between means, medians, interquartile ranges and the distributions of the travel time data. The use of statistical tests is also a possibility. It

is unclear which of these measures should be relied upon for a definitive assessment of the success of signal retiming measures.

- Despite the non-normal distribution of arterial travel time data, travel time studies employ the use of parametric tests to examine arterial travel time data due to its practicality.
- While many studies exist on how to identify outliers in Bluetooth travel time data, it was found that the influence that these outliers have on the results of a travel time study has not been adequately studied.

3 Site Description and Data Characteristics

3.1 Introduction

This chapter examines the characteristics of the data available for this research as well as the sites at which these data were collected. First, the field data site and the associated Bluetooth and GPS probe vehicle travel time data are presented. Next, the simulated data site and the simulated travel time data is presented from which simulated Bluetooth and GPS probe vehicle data were drawn. Emphasis is placed upon providing an overview of the primary characteristics of the data and its capability of being used for evaluating signal retiming measures. This chapter concludes with a summary of the major observations made during the exploration of the field and simulated travel time data.

An important assumption associated with the use of Bluetooth travel time data to evaluate signal retiming measures is that the collected Bluetooth travel times are representative of the travel times of the entire population of vehicles. This assumption allows for conclusions to be made about the success of the implementation of signal retiming measures on an arterial corridor.

3.2 Data Overview

A total of nine data sets were utilized in this investigation. Each data set had three distinct characteristics: the type of data (Bluetooth, GPS, or truth), whether or not it was collected in the field or collected in a simulated environment, and the scale of the data (path or link). The true travel time data is the simulated data from which Bluetooth and GPS data were drawn through sampling. The characteristics of each data set are summarized in Table 3.1. While Bluetooth travel time data were the focus of this investigation, GPS probe vehicle travel time data were also available for comparison to the Bluetooth data. As explained in this chapter, the field data originates from one site while the simulated data originates from another.

Table 3.1: Travel time data set characteristics

Travel Time Data Set	Type	Field or Simulated?	Scale
1	Bluetooth	Field	Path
2	Bluetooth	Field	Link
3	GPS	Field	Link
4	Bluetooth	Simulated	Path
5	Bluetooth	Simulated	Link
6	GPS	Simulated	Path
7	GPS	Simulated	Link
8	Truth	Simulated	Path
9	Truth	Simulated	Link

In the context of this research, the difference between a “link” and a “path” is spatial scale. A path represents the entire arterial corridor of interest, which can then be divided into a series of links. For example, a corridor that contains four Bluetooth detectors will consist of one path whose travel time is represented by the time taken to traverse the entire corridor (i.e. from detector 1 to detector 4 or vice versa). That same corridor containing four detectors will be made up of three links (detectors 1 to 2, 2 to 3 and 3 to 4) and each link will have its own collection of travel times for each direction.

3.3 Characteristics of the Field Data

In this section, the field data collection site is examined. The aggregation of the field Bluetooth data is then described. An investigation into the primary characteristics of the field Bluetooth and GPS data sets is then presented.

3.3.1 Field Data Collection Site

The field data were collected along Victoria Park Avenue, a four-lane arterial roadway located in Toronto, Ontario, Canada. This arterial runs north-south and links residential neighbourhoods in the north to Toronto’s central business district in the south. Figure 3.1 shows the corridor of interest along with the major roadways that cross it, the distance between each major roadway, the location of each Bluetooth detector, and the distance between consecutive Bluetooth detectors. Figure 3.2 shows a map containing the seven Bluetooth detectors along Victoria Park Avenue that were used to collect travel time data. The location of each Bluetooth detector is indicated by a yellow thumbtack along with an index and the name of the facility that crosses Victoria Park Avenue where the detector is located. It must be noted that each detector deployed along this corridor was powered by a portable power source as opposed to a permanent power source at a nearby traffic signal. The significance of this detail is discussed further in Chapter 5.

For the purposes of this investigation, the limits of the corridor of interest are Gordon Baker Road in the north (just south of Steeles Avenue East) and Kingston Road in the south. These limits define a study area that is 15.1 kilometres in length.

The field study area contains a total of 37 signalized intersections and 2 exclusive pedestrian crossings that are activated by push button. Intersection turning lanes, raised medians, dedicated bus bays and on-street bus stations are present throughout the study area. The northern part of the study area is dense in retail malls and residential neighbourhoods. From Gordon Baker Road to Eglinton Avenue, the speed limit is 60 kilometres per hour and there are two lanes in each direction. Moving south, the number of high-rise residential buildings increases and retail malls transition to retail plazas. The speed limit decreases to 50 kilometres per hour from Eglinton Avenue to Dawes Road. One-and two-storey dwellings become common as the number of lanes in each direction decreases from two to one at Dawes Road until Crescent Town Road. The frequency of retail plazas increases from Crescent Town Road to Gerrard Street and there are two lanes in each direction for this section. Access to the Victoria Park subway station is provided along this corridor just north of Danforth Avenue. From Gerrard Street to Kingston Road, one lane is present for each direction and the area is entirely one- and two-storey residential buildings. Streetcar infrastructure is also present along the corridor between Meadow Avenue and Kingston Road. Table 3.2 summarizes the primary features of the field study area.

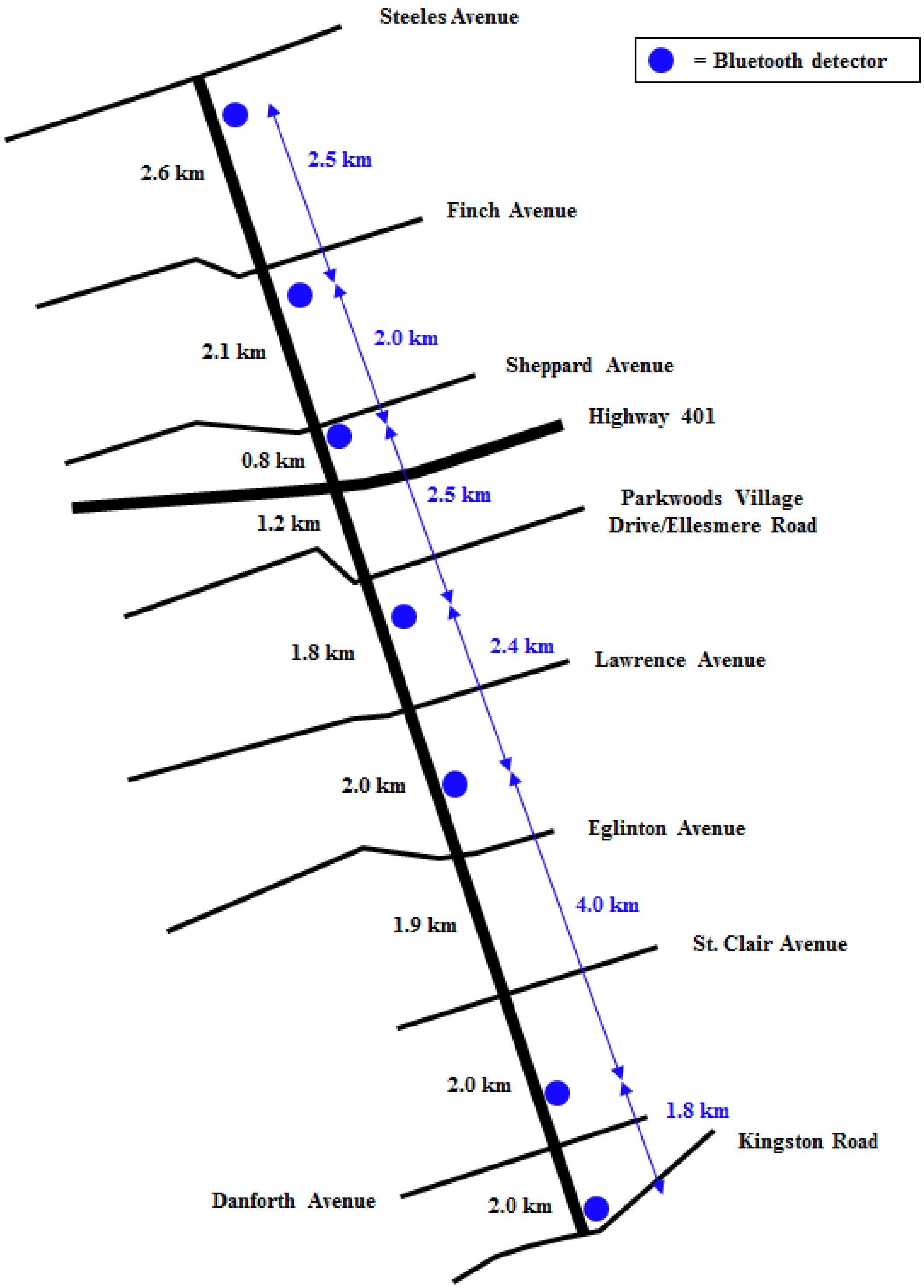

Figure 3.1: Diagram of the field study area

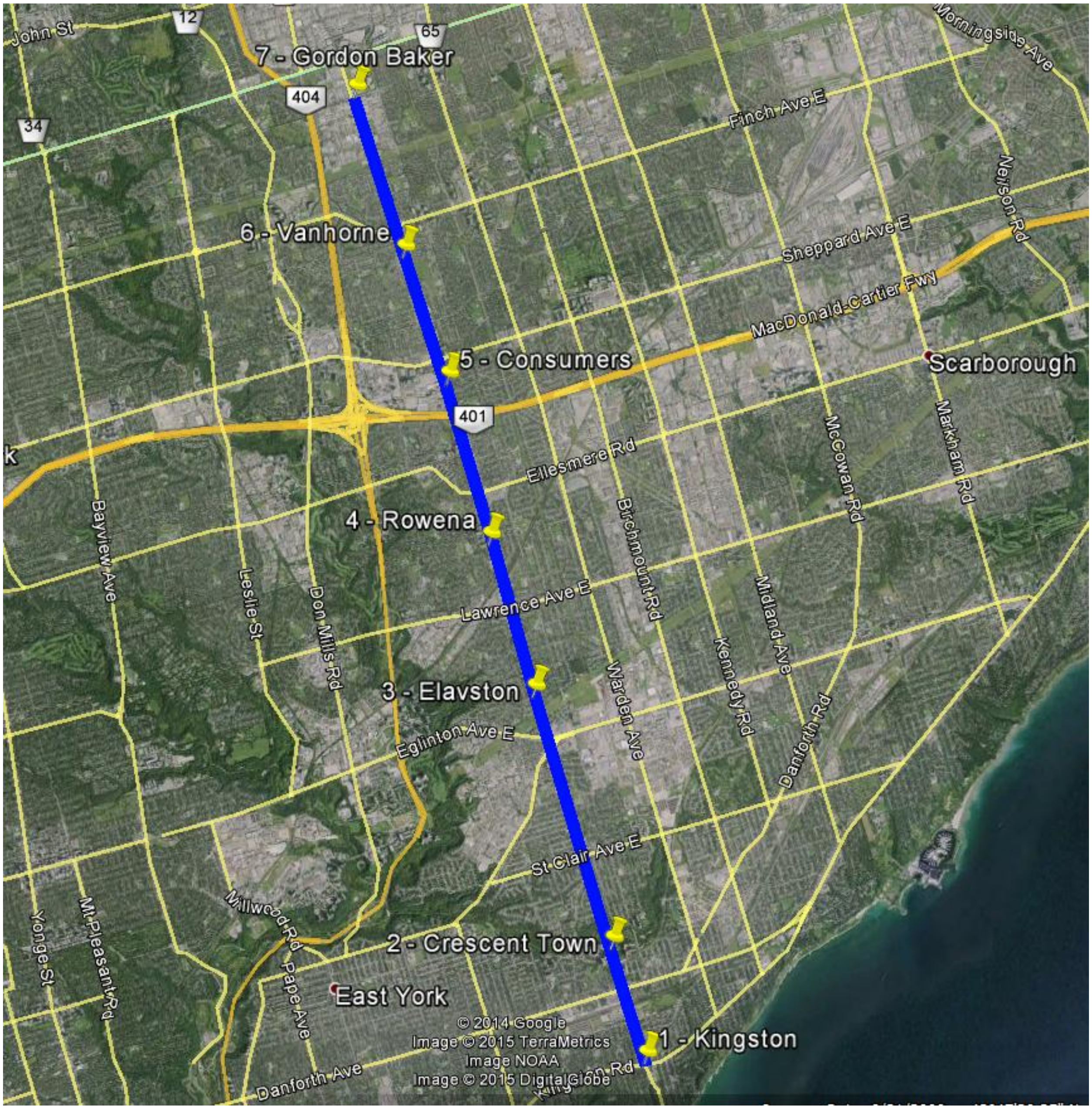

Figure 3.2: Map of the field study area (Google, TerraMetrics, National Oceanic and Atmospheric Administration, & DigitalGlobe, 2015)

Table 3.2: Primary features of the field study area

From	To	Speed Limit (km/h)	Lanes per Direction
Gordon Baker Road	Eglinton Avenue	60	2
Eglinton Avenue	Dawes Road	50	2
Dawes Road	Crescent Town Road	50	1
Crescent Town Road	Gerrard Street	50	2
Gerrard Street	Kingston Road	50	1

3.3.2 Processing of the Field Bluetooth Data

Before a formal investigation of the field data was performed, the data required processing. This was performed using MATLAB software. Since the effects of signal retiming measures on the weekday peak period traffic conditions were of interest in this investigation, weekend data were excluded unless otherwise noted.

Two time periods were defined for the data: morning peak period and evening peak period. In terms of aggregation, a travel time entry was collected only if both its start time stamp and end time stamp were within the time period of interest. The vendor of the Bluetooth detection system has also incorporated an outlier detection filter that labels observed travel times as valid or invalid. Throughout this chapter and this investigation, only valid field Bluetooth data were used unless otherwise noted.

3.3.3 Characteristics of the Field Bluetooth Data

The Bluetooth travel time data used in this study was produced by the industry partner from the raw Bluetooth hits using a middle-middle strategy. In other words, when multiple hits were obtained for a single Bluetooth-enabled device at a given detector, the hit that was chronologically in the middle of the set of hits was used to generate a travel time. This was done in an effort to emulate the GPS probe vehicle data that were also collected for this corridor.

For these data, the Before period was defined to be Saturday October 6, 2013 to Sunday October 19, 2013 and the After period was defined to be Sunday November 24, 2013 to Saturday December 7, 2013 (both two weeks in duration).

The path-level data consisted of the travel times of detected vehicles that traversed the corridor from the detector at Kingston Road to the detector at Gordon Baker Road or vice versa. The link-level data consisted of the travel times of detected vehicles that traversed the corridor between adjacent detectors. For example, travel times from Kingston to Crescent Town were part of the link-level Bluetooth data set while travel times from Kingston to Elvaston were not because those detectors were not adjacent to each other. Other than travel times between the two furthest detectors (i.e. path-level data), travel times between non-adjacent detectors were unavailable for this investigation.

Both the path-level and link-level field data were converted from raw MAC ID detections and time stamps to travel times before this investigation took place (i.e. this step in the processing of the data was beyond the scope of this investigation). Table 3.3 shows a sample of the travel time data from the field in its raw form. Each entry of the table represents a pair of matched detections of the same MAC ID that together produce a travel time. Note that MAC IDs have been anonymized to alleviate any potential privacy concerns. Each row entry contains the anonymized MAC ID matched, the start date and time of the first chronological hit, the end date and time of the last chronological hit, the travel time (which is the difference between the start and end time stamps), and a flag to indicate whether or not a travel time entry was a valid observation (flag = 0) or an outlier (flag = 1). While it was known that outliers were identified using the interquartile range via software provided by Traffax Inc., the exact nature of this interquartile range filter (i.e. how the algorithm aggregated the data before applying the filter) was not within the scope of this investigation.

A separate data file was created for each possible from-to detector combination containing only the travel time entries for that link. With there being only two possible start and end detectors for the path level, a total of four path-level data files were available (one path, two directions, two data sets). With seven Bluetooth detectors available at the link level, a total of 24 link-level data files were available (six links, two directions, two data sets). Each path-level file contained an average of approximately 100 travel times while each link-level file contained an average of approximately 7,000 travel times, highlighting the distinct difference in sample size between the two data sets.

Table 3.3: Sample Bluetooth travel time data collected in the field

Anonymized MAC ID	Begin Date	Begin Time	End Date	End Time	Travel Time (min)	Outlier?
00:25:3I:FV:43:09	11/25/13	7:18:23	11/25/13	7:39:21	20.967	No
00:18:3S:YF:34:45	11/25/13	7:40:31	11/25/13	8:20:23	39.875	No
00:36:3D:UN:82:89	11/25/13	7:58:31	11/25/13	9:34:21	95.833	Yes
00:74:3Y:RD:49:84	11/25/13	8:00:29	11/25/13	8:22:57	22.458	No
00:10:3T:EX:57:16	11/25/13	8:52:50	11/25/13	9:35:14	42.4	No

Figure 3.3 shows the field Bluetooth travel times as a function of time of day for a given link and direction for a single week (a Saturday to a Friday) during the Before period. All Bluetooth travel times (both valid observations and those identified as outliers) are included in this plot. In this figure, travel times are plotted at the time associated with the first time stamp at the upstream detector. Variations in the travel time magnitude and increases in the number of collected travel times can be observed during daylight hours. Conversely, few or no observations were collected during the late night-early morning period. These trends highlight the potential for large amounts of data being available for the peak periods for which improved traffic conditions are desired. Excessively high travel times are also evident in this figure, highlighting the need for outlier filtering.

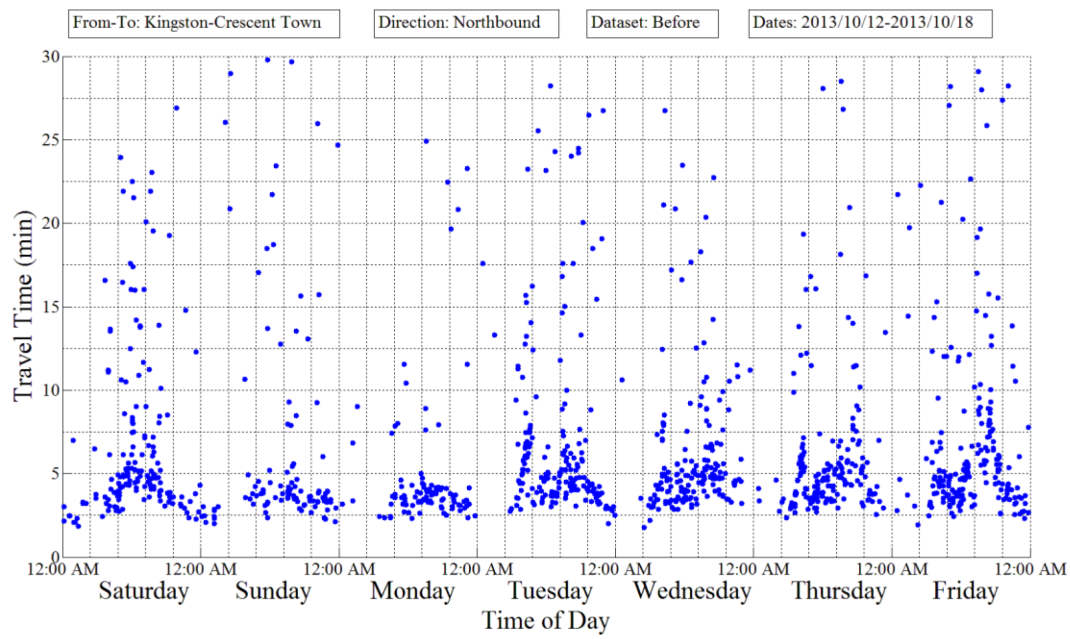

Figure 3.3: Travel time as a function of time of day for a single week for the field Bluetooth data

Figure 3.4 shows the field Bluetooth travel times as a function of time of day for a given link, direction and date (October 9, a Wednesday) during the Before period. All Bluetooth travel times (both valid observations and those identified as outliers) are included in this plot. Each travel time was placed at the time at which the first time stamp for the corresponding MAC ID was collected. The peak periods can be identified by the increases in the mean travel time at 8:30am (morning peak period) and at 5:30pm (evening peak period). Increases in the number of Bluetooth-enabled devices are also evident during the peak periods compared to the off-peak periods in which data are less frequent. Extreme values are also evident as abnormally high travel times greater than 20 minutes can be seen between 4:00pm and 6:00pm.

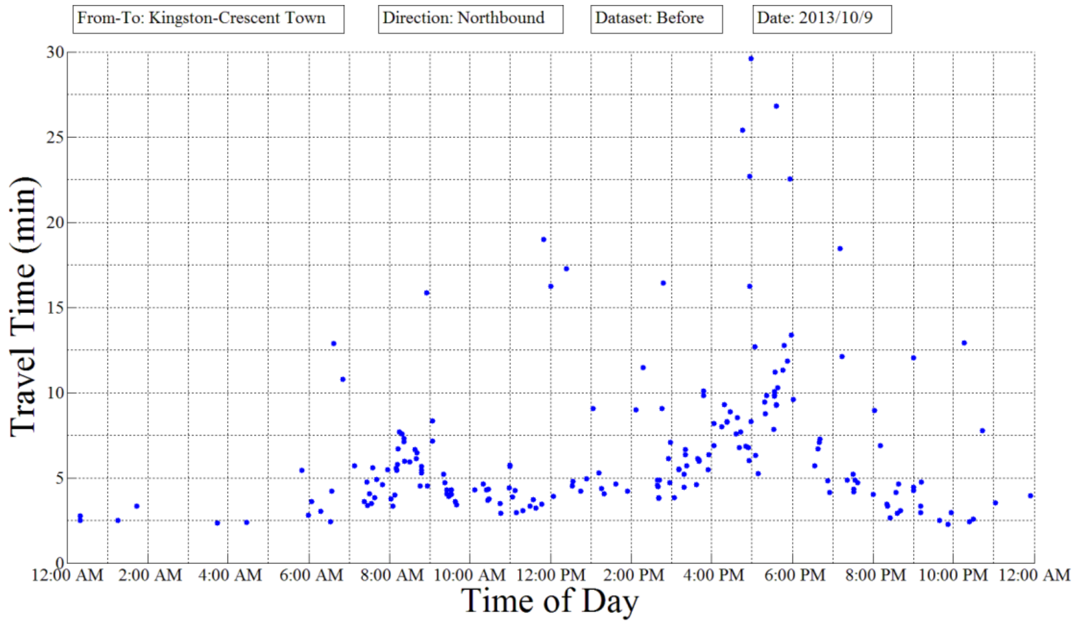

Figure 3.4: Travel time as a function of time of day for a single day for the field Bluetooth data

Figure 3.5 shows the number of field Bluetooth travel times collected as a function of time of day for a given direction and date (October 9, a Wednesday) during the Before period. Link-level data were collected for a given link to compare to path-level data under the same conditions. Travel times were aggregated on a half-hour basis and a travel time was collected if its start time stamp was within the half-hour aggregation period of interest. The contrast between the amount of travel time data available at the link level and at the path level for the same time period can be easily observed. For example, the 8:00am-8:30am time period experienced 12 travel times at the link level and zero travel times at the path level. It can also be observed that link-level travel times fluctuate according to daily variation in traffic volumes while path-level travel times are rarely available regardless of time of day.

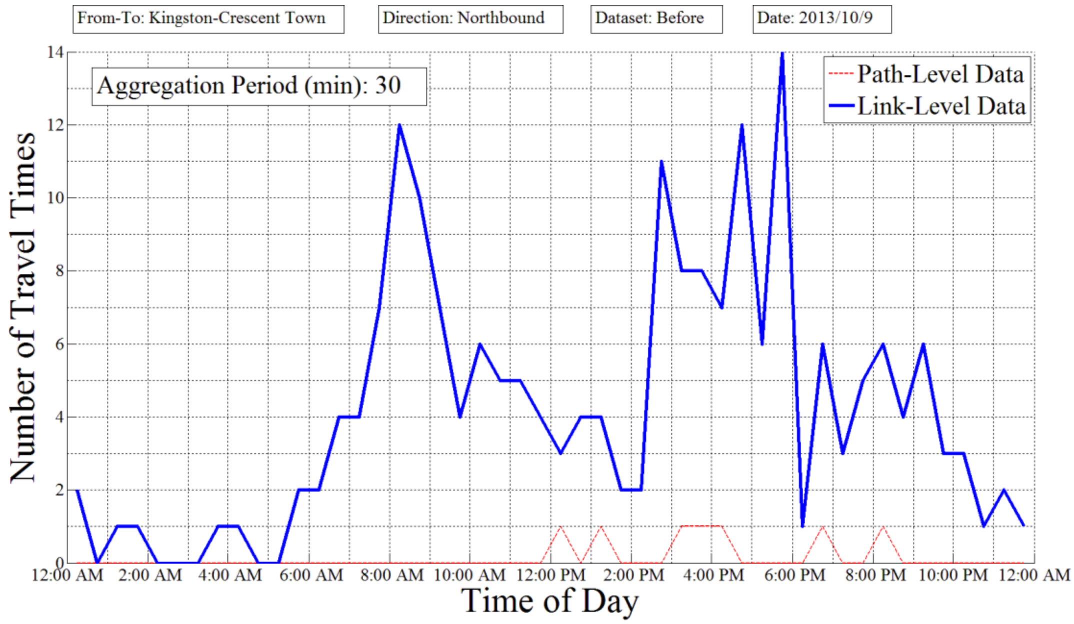

Figure 3.5: Number of travel time observations as a function of time of day for a single day for the field Bluetooth data

Figure 3.6 shows the cumulative travel time frequency distributions for the Before and After field Bluetooth data for a given link, direction and time period. As seen in this figure, travel times between two minutes and ten minutes are most common. It can also be observed visually that the distribution of the After data is made up of a greater frequency of shorter travel times compared to the distribution of the Before data as the After distribution is to the left of the Before distribution. The long tail to the right and the short tail to the left of each distribution indicate the asymmetry of the data; in other words, the data are not normally distributed.

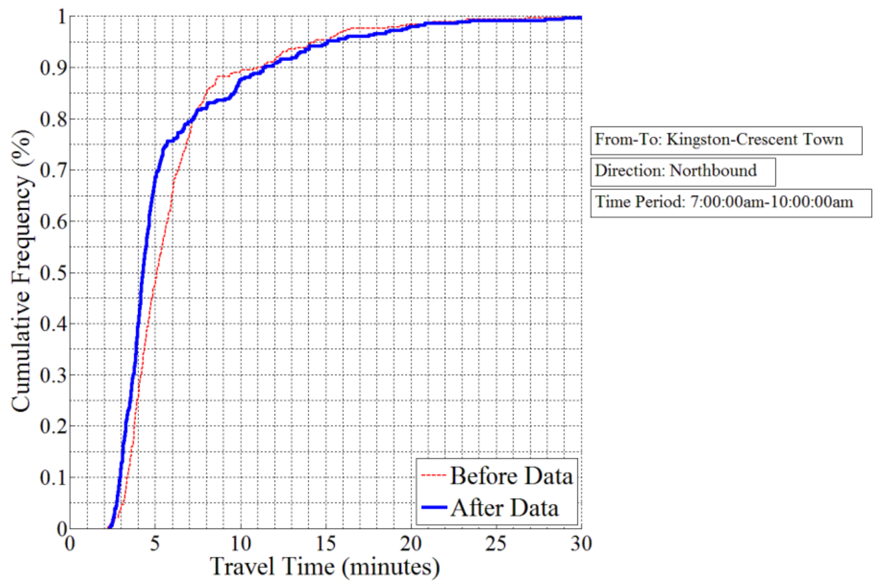

Figure 3.6: Sample cumulative travel time frequency distribution for the field Bluetooth data

Figure 3.7 shows the mean travel times on a per kilometre basis for the field Bluetooth data collected for each link for a given direction and time period. It is evident that the mean travel time per kilometre did not decrease from the Before period to the After period for all links. This shows that it is difficult to conclude whether or not the implemented signal retiming measures improved traffic conditions based on a simple comparison of link-level mean travel times. Another important observation that can be made from this figure is that the mean travel time per kilometre varies according to link. For example, the Kingston-Crescent Town link has a much higher mean travel time than the Crescent Town-Elvaston link. Possible reasons for this could be that a greater amount of impedances to progression exist on the Kingston-Crescent Town link compared to the Crescent Town-Elvaston link on a per kilometre basis.

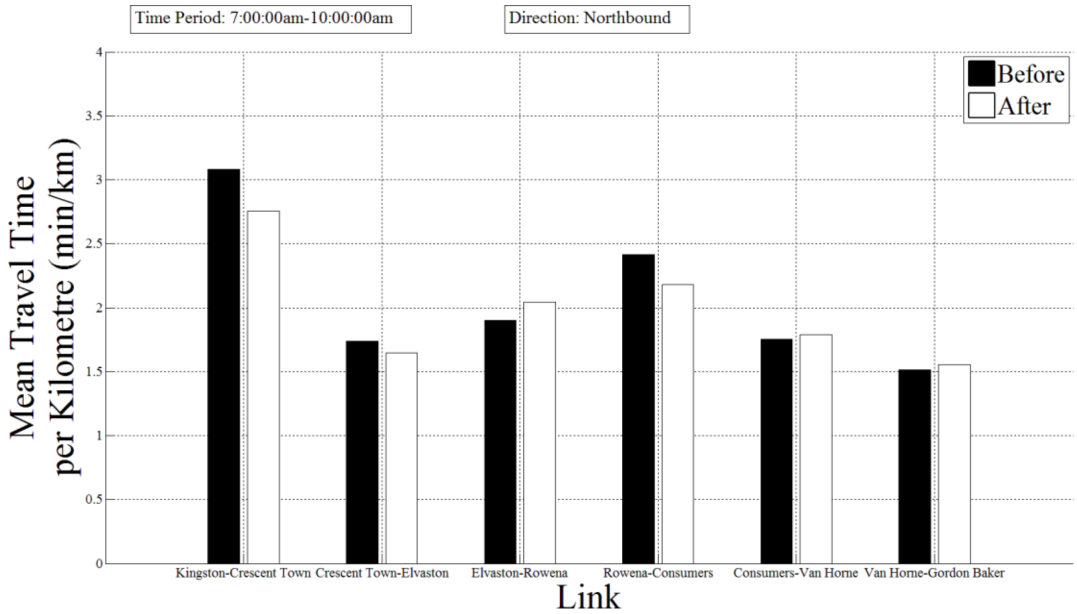

Figure 3.7: Mean travel time per kilometre of each link for the field Bluetooth data for a sample time period and direction

Figure 3.8 shows the number of field Bluetooth travel times collected on a per kilometre basis on each link for a given direction and time period. It can be seen that the number of travel time observations per kilometre obtained varies substantially across the different links. For example, approximately 200 travel times per kilometre were obtained for the Kingston-Crescent Town link and more than 700 travel times per kilometre were obtained for the Elvaston-Rowena link for each of the Before and After periods. Assuming that the Bluetooth data are representative of the overall traffic stream, it is evident that certain links experience much higher volumes than others. This could be due to some links containing access to more amenities than others.

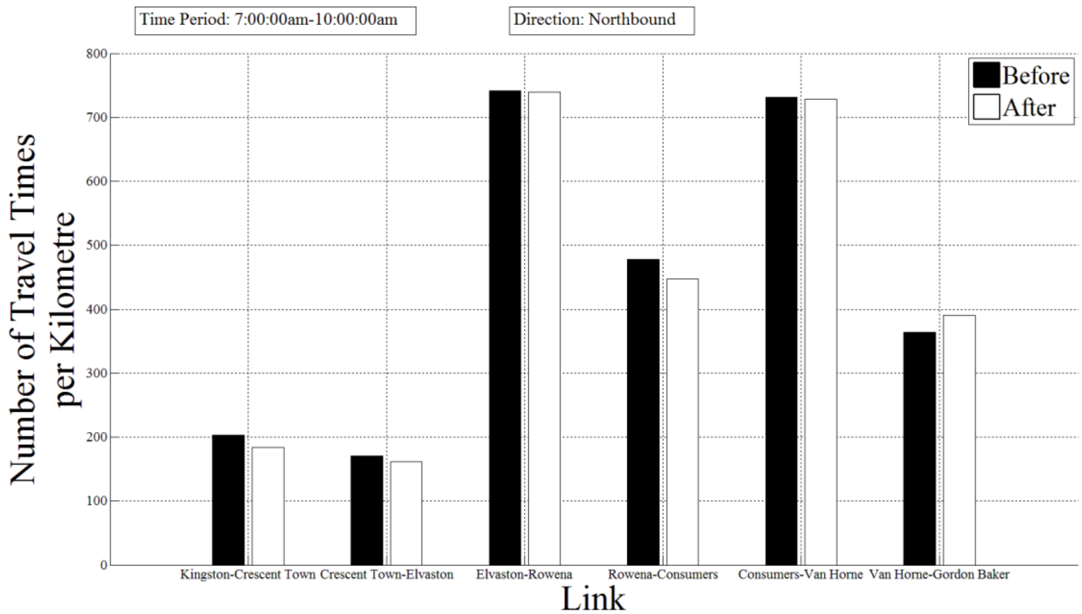

Figure 3.8: Number of observations per kilometre collected for each link for the field Bluetooth data for a sample time period and direction

Assuming that a Bluetooth-enabled vehicle possesses the same MAC ID over the entire Before or After data collection period, the frequency with which the group of collected MAC IDs appear for a given combination of link, direction, time period and Before or After data set can be identified. Figure 3.9 shows the frequency with which a MAC ID is detected for a given link, direction and time period during the Before period. All Bluetooth travel times (both valid observations and those identified as outliers) are included in this plot. It can be seen that approximately 25% of MAC IDs appear more than once. This indicates that a significant percentage of Bluetooth-enabled vehicles traverse the corridor under these conditions on a regular basis. This suggests that commuters represent a significant portion of the population of Bluetooth-enabled vehicles. Furthermore, the magnitude of this percentage was relatively consistent across all combinations of link, time period, direction and Before or After data set. One possible reason for this observation is that a portion of the detected vehicles traverse the link more than once in a given analysis period (e.g. delivery vehicles, buses, taxis, etc.). However, further examination of the data showed that this was not the case. Rather, approximately 25% of MAC IDs appear in the data set multiple times because these MAC IDs are observed over multiple days, not multiple times on the same day. This supports the notion that the travel times of commuter traffic make up 25% of all collected travel times because 25% of all collected MAC IDs appear across multiple days.

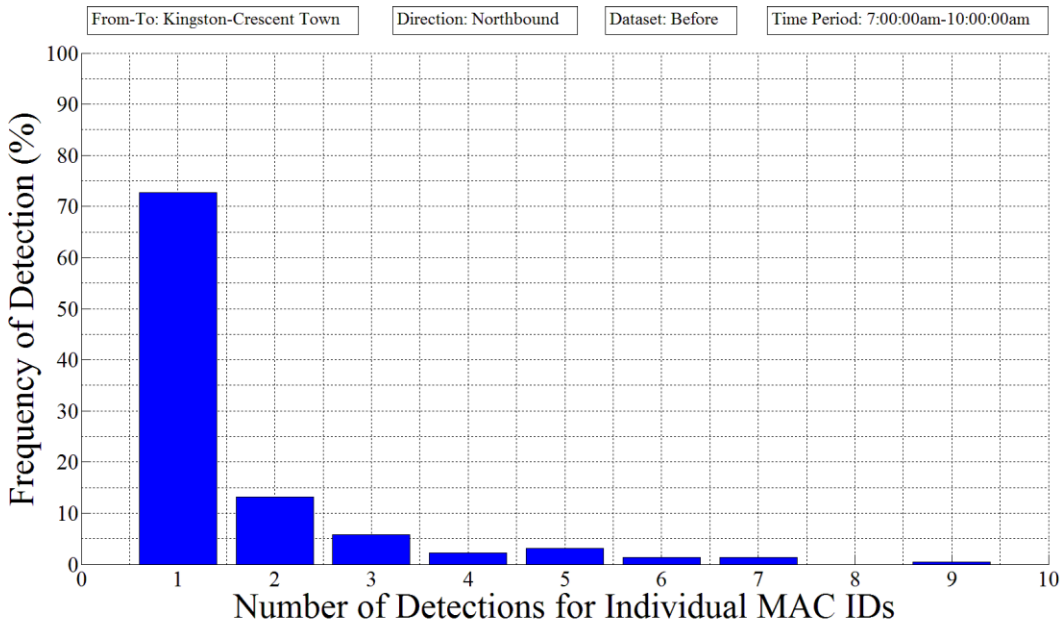

Figure 3.9: Frequency of detection for individual MAC IDs for the field Bluetooth data

Assuming that (1) a Bluetooth-enabled vehicle possesses the same MAC ID at least throughout the duration of its trip along the corridor and (2) a Bluetooth device that is detected at one detector will be detected at each of the other Bluetooth detectors passed during the trip, a vehicle's progression along the corridor can be tracked and the number of vehicles that start at the beginning of the study area but do not pass the next downstream Bluetooth detector (i.e. these vehicles exit the corridor) can be identified.

Figure 3.10 and Figure 3.11 show for a given direction the percentage of MAC IDs that appeared at the starting link and at each following link along the corridor using lines as well as the percentage of MAC IDs that appeared at consecutive links using bars. All Bluetooth travel times (both valid observations and those identified as outliers) are included in this plot. For example, for the northbound direction, approximately 45% of the MAC IDs that appeared at the Kingston-Crescent Town link also appeared at the Crescent Town-Elvaston link according to the line plots. This shows that approximately half of the Bluetooth-enabled vehicles that traversed the Kingston-Crescent Town link exited the corridor somewhere along the Crescent Town-Elvaston link and never reached the detector at Elvaston. Next, approximately 30% of the MAC IDs that appeared at the Kingston-Crescent Town link also appeared at the Elvaston-Rowena link according to the line plots. This shows that of the 45% of the vehicles that appeared at the Crescent Town-Elvaston link, approximately 30% of those vehicles also traversed the Elvaston-Rowena link. According to the bar plots, approximately 70% of the MAC IDs that appeared at the Crescent Town-Elvaston link also appeared at the Elvaston-Rowena link. This shows that the majority of vehicles that traversed the Crescent Town-Elvaston link also traversed the Elvaston-Rowena link. This could indicate that the ingress and egress points on the Crescent Town-Elvaston link are less frequently used compared to other links.

As shown in these plots, estimations can be made regarding the proportions of vehicles that traverse each link of the corridor. The ability to track where a vehicle entered and exited the corridor highlights the potential for using link-level Bluetooth travel time data to create origin-destination matrices, which are a key component in transportation planning studies.

These figures also highlight the vast difference in sample size between the path level and the link level. For example, for the northbound direction, less than 5% of the MAC IDs that traversed the Kingston-Crescent Town link also traversed the entire corridor through to the Van Horne-Gordon Baker link. Given that on average 7,000 travel times were observed for a given link and direction for the Before or After period, it can be expected that less than 350 (5% of 7,000) travel times would be available at the path level.

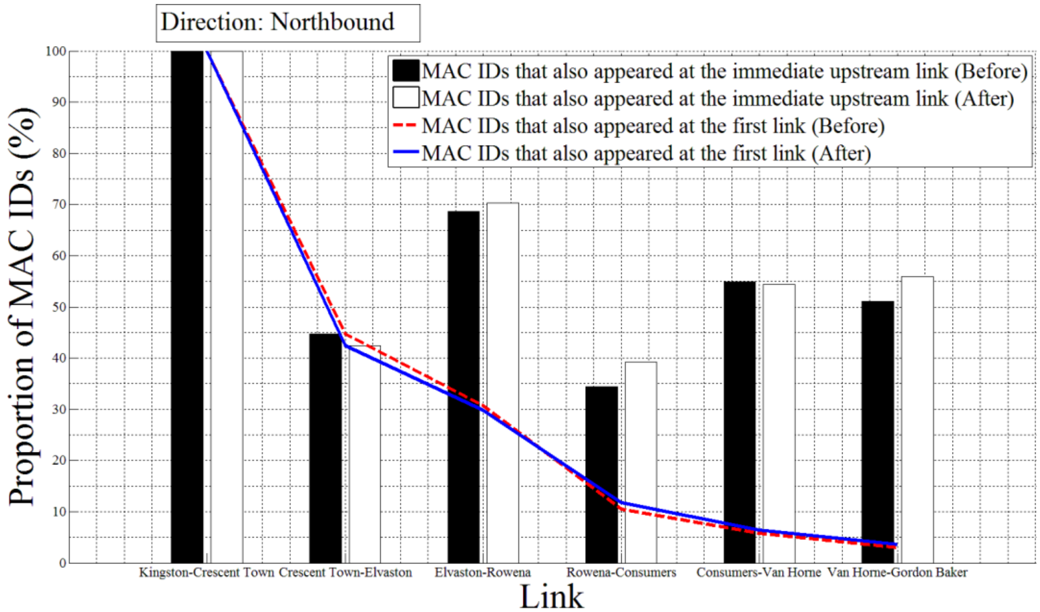

Figure 3.10: Propagation of field data MAC IDs in the northbound direction

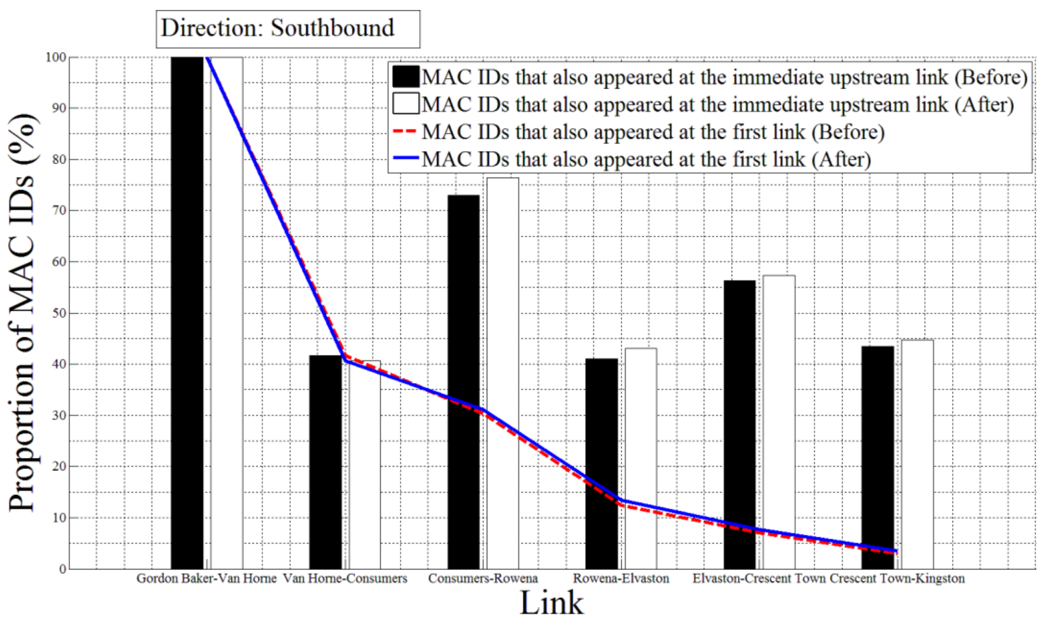

Figure 3.11: Propagation of field data MAC IDs in the southbound direction

As stated previously, the ability to infer travel patterns from the acquired Bluetooth detectors requires the assumption that a Bluetooth-enabled vehicle that is detected by a Bluetooth detector at one location will also be detected at all other Bluetooth detectors that the vehicle passes along the arterial corridor. An attempt to verify this assumption is presented below.

Using the available field data, an attempt is made to determine the probability that a vehicle detected by one Bluetooth detector passes a detector downstream of the first but is not detected. If this probability is sufficiently small, then the previously discussed assumption is valid and the previously discussed proportions are reliable estimates of vehicle ingress and egress along this corridor (assuming that the Bluetooth travel times are representative of the entire vehicle population).

To calculate the probability of a missed detection, the experiment illustrated in Figure 3.12 was constructed. In this figure, the situation in which a vehicle passes by a detector is modelled by the following series of parameters:

- The range of the Bluetooth detector was assumed to be 100 metres according to the model proposed by Moghaddam (2014).
- The distance between the detector and the road was assumed to be 10 metres to model the situation in which a detector is placed adjacent to a roadway.
- The maximum frequency of hits was assumed to be 1 potential hit every 1.28 seconds as defined by Moghaddam (2014). The resulting distance spacing between hits was a function of this frequency and the speed of the vehicle.

A vehicle was assumed to travel at a constant speed through the detection zone. It was also assumed that the first potential hit of the vehicle occurred at the instance in time when it entered the detection zone. All other possible hits along the vehicle's trajectory were located according to the spacing defined by the vehicle speed.

For a fixed vehicle speed, the time and location of each hit for the vehicle trajectory was calculated. Using the distance between each hit and the detector as input into Moghaddam's proposed model (2014), the probability of a missed detection for each hit was calculated. Assuming that these probabilities are independent, the probability of zero hits during the period that the vehicle traverses the detection zone is simply the product of each of the individual hit probabilities. Figure 3.13 shows the probability of a missed detection as a function of vehicle speed for a range of speeds from 5 kilometres per hour to 80 kilometres per hour. As expected, the probability of a missed detection increases and the number of hits decreases as vehicle speed increases. Since the greatest probability of a missed detection for this situation is less than two percent, it is clear that the probability of a missed detection is very low.

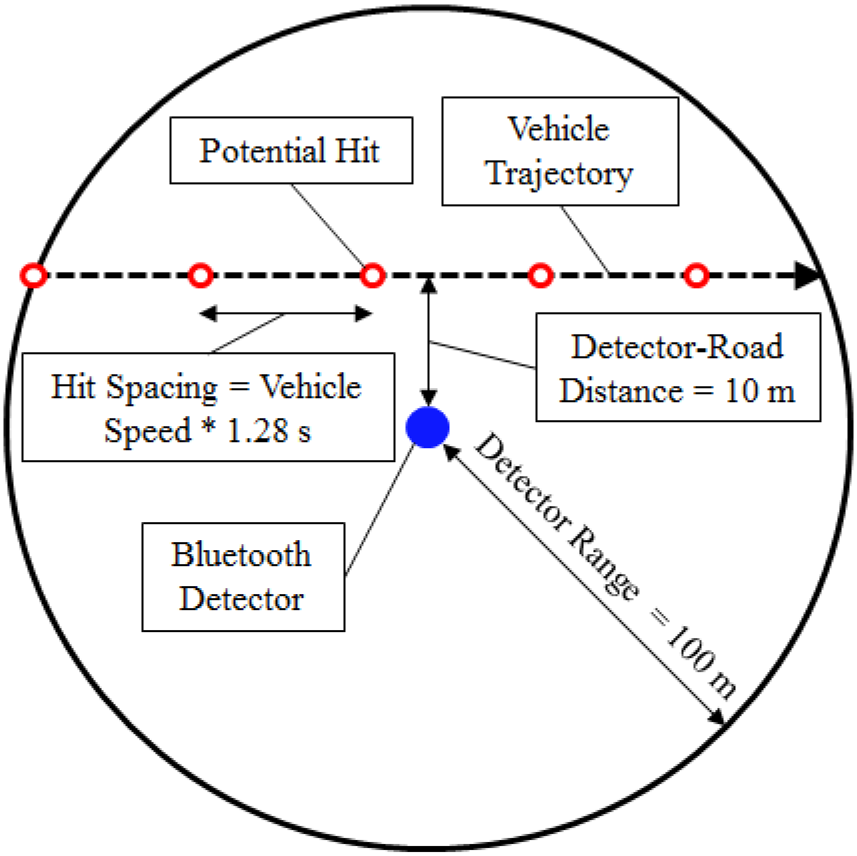

Figure 3.12: Scenario to test the probability of a missed detection

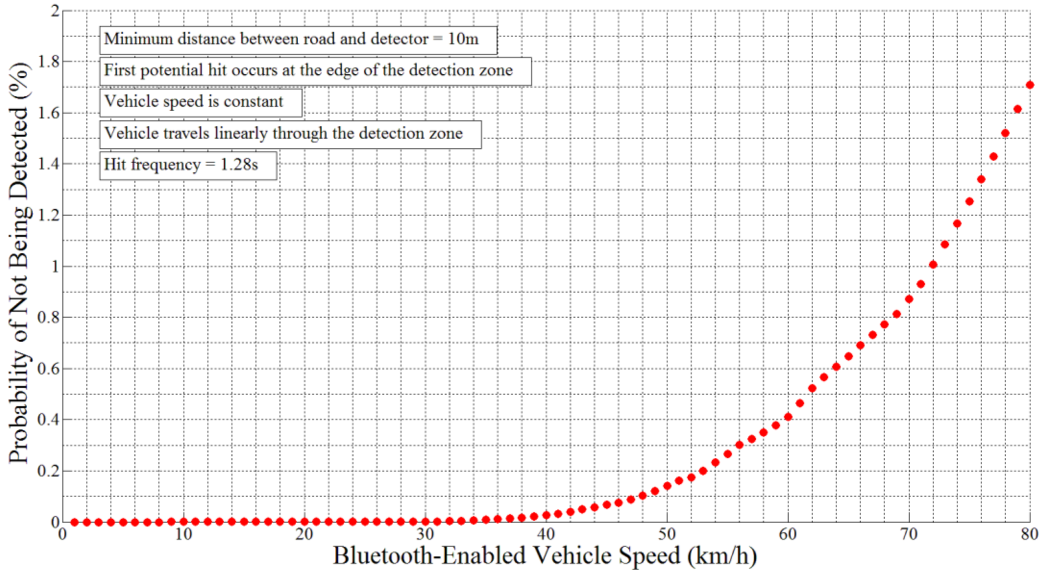

Figure 3.13: Probability of a missed detection as a function of vehicle speed

The scenario in which a vehicle is detected at some detectors but not at others was also examined. Specifically, the probability that a vehicle was successfully detected at one detector at least once, not detected at a second detector and successfully detected at a third detector at least once was calculated. Examining this scenario revealed the probability that a vehicle known to pass all three detectors in reality appears to have left the corridor and then returned to it according to the Bluetooth data. To calculate this probability, the experiment illustrated in Figure 3.14 was constructed. The same parameters and assumptions of the experiment that examined the probability of a missed detection also applied to this experiment.

The probability of a missed detection was used as a basis to calculate the probability for this experiment. As a result, the probability of being detected at least once at a detector was squared (to represent detectors 1 and 3) and multiplied by the probability of a missed detection at a detector (to represent detector 2). For example, for a speed of 80 kilometres per hour, the probability of a missed detection was calculated to be approximately 1.7%. Therefore, the probability of being detected at least once is approximately 98.3%. Squaring 98.3% and multiplying the result by 1.7% results in a probability of approximately 1.6%. Like the probability of a missed detection at a single detector, the probability of a missed detection at detector 2 while being successfully detected at detectors 1 and 3 is also very low.

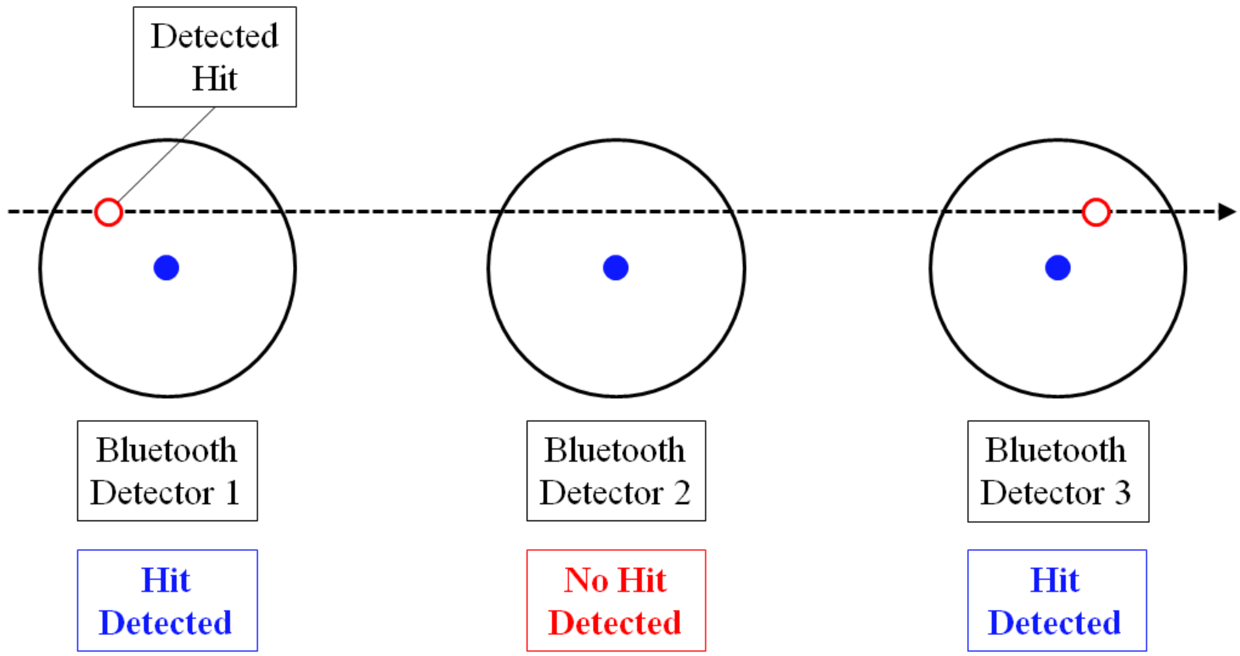

Figure 3.14: Probability of a missed detection at a detector between two other detectors where detection was successful

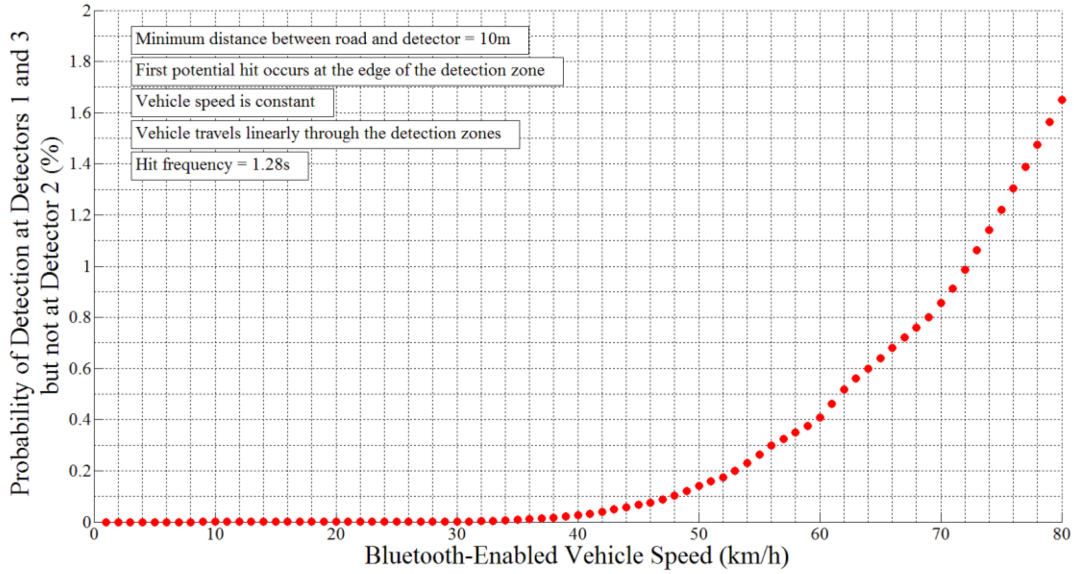

Figure 3.15: Probability of a missed detection at a detector between two other detectors where detection was successful as a function of vehicle speed

Despite proving that the probability of a missed detection is very low, there is a number of ways in which a path-level travel time is obtained for a particular MAC ID but the corresponding link-level travel times of that MAC ID are not obtained for all links along the path, including:

- The Bluetooth device being detected was turned off while the vehicle traversed one or more of the links along the corridor.
- The vehicle left the corridor and then returned to the corridor. It may be possible to identify this situation in field data when the path-level travel time is unusually high while some of the corresponding link-level travel times are absent. However, this method is not definitive.
- The vehicle traversed the link but made an on-route stop, resulting in an excessively high travel time. As a result, the algorithm used to match MAC IDs at the upstream and downstream detectors may not consider the resulting travel time to be valid. Therefore, this travel time is discarded.
- Various obstructions could result in signal interference.

The challenge associated with obtaining path-level travel times versus link-level travel times suggests that the evaluation of signal retiming studies are likely to be more feasible using link-level data than using path-level data. The feasibility of using path-level or link-level data to evaluate signal retiming measures is examined in more detail in Chapter 4.

To explore the potential for using the field Bluetooth data to evaluate signal retiming measures, a regression analysis was performed on the means and standard deviations of the travel times for each combination of link, direction and time period. If improvements were made through signal retiming, this would become evident through a positive linear relationship between the Before and After means and standard deviations of the travel time data. A statistically significant slope with a magnitude less than 1 would indicate a reduction in the independent variable (mean or standard deviation of the After data) compared to the dependent variable (mean or standard deviation of the Before data). Statistical significance was assessed based on the slope having a p-value less than 0.05 and the y-intercept having a p-value greater than 0.05, the desired level of confidence.

Figure 3.16 and Figure 3.17 respectively show the mean and standard deviation of the link-level field Bluetooth data used in these regression analyses. For this figure and all other regression figures, the dashed line is simply a $y = x$ line and the solid line is the trendline. From Figure 3.16 it can be observed that on average (across all links, directions and time periods) the mean travel time after signal retiming was 95% of the mean travel time before signal retiming. Insufficient evidence was found to support the statistical significance of the intercept so it was set to zero. The lack of significance of the intercept, the magnitude of the slope coefficient being less 1, and the relatively high R-squared value show that signal retiming was successful for this corridor according to the mean travel time.

The results for the standard deviation of travel times are less definitive because both the slope and the intercept are statistically significant. Consequently, when the standard deviation in the Before period was small (e.g. 1 minute), the regression suggests that the signal retiming resulted in an increase in the standard deviation. However, when the standard deviation in the Before period was large, the regression suggests that the signal retiming resulted in a decrease in the standard deviation. It should also be noted that the regression for the standard deviation only explains 36% of the variance in the observed data. This reinforces the notion that the impact of signal retiming on the variability of travel times is inconsistent.

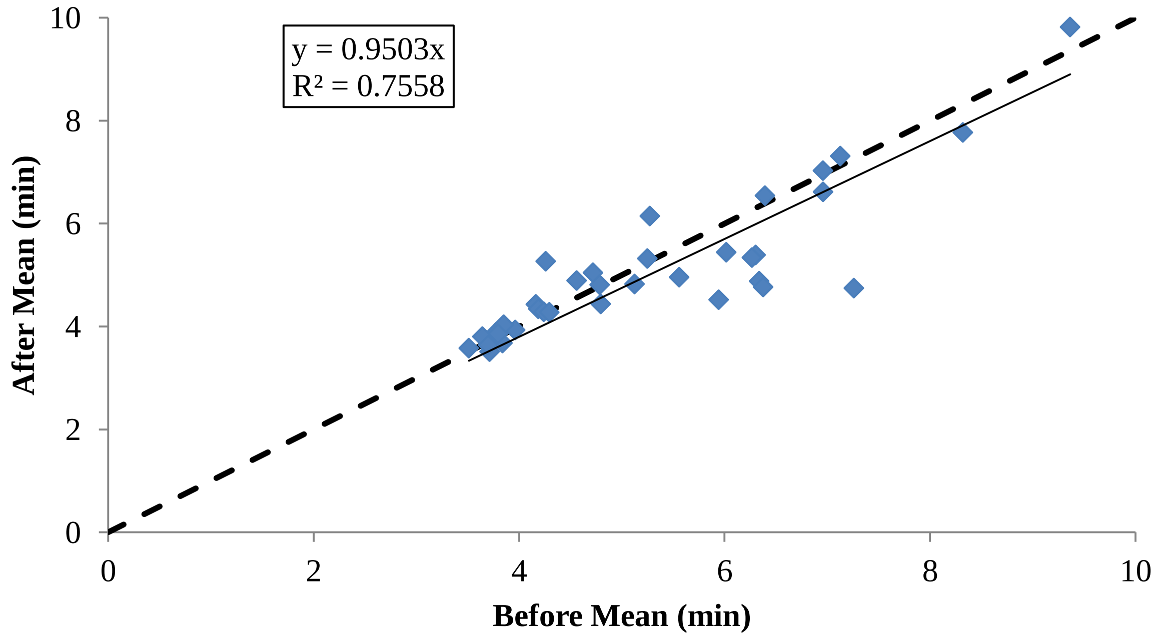

Figure 3.16: Comparison of link-level Before and After mean travel times for the field Bluetooth data

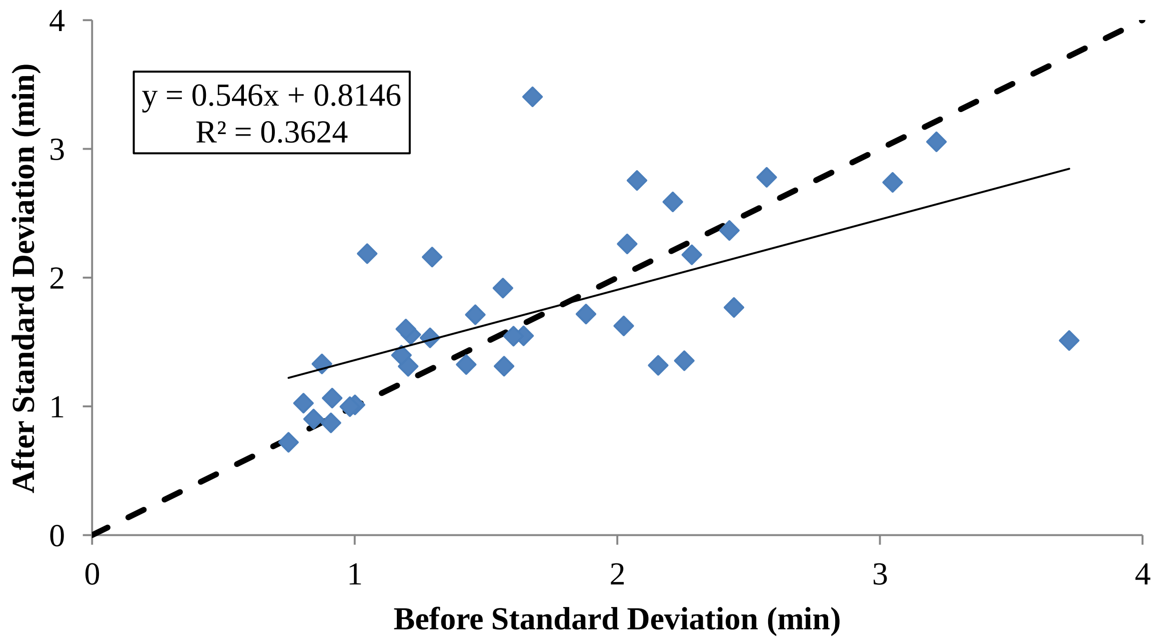

Figure 3.17: Comparison of link-level Before and After standard deviations of the travel times for the field Bluetooth data

3.3.4 Characteristics of the Field GPS Data

In addition to path-level and link-level Bluetooth data, link-level GPS mean travel times were obtained from probe vehicle runs and were also available for analysis. GPS mean travel times were only available for the Before period and the number of probe runs used to calculate each mean was unknown. This field GPS probe vehicle data is available in Appendix A. Within that Before period, data were available for two time periods with each being three hours in duration: the morning peak period (7:00am-10:00am) and the evening peak period (3:30pm-6:30pm). Assuming that these time periods were representative of the major time periods during which the state of traffic conditions were to be captured, all field data were aggregated according to these time periods.

Figure 3.18 shows the results of a regression carried out for the GPS mean travel times as a function of the Bluetooth mean travel times for the field data for all combinations of link, direction and time period for the Before period. The y-intercept was not found to be statistically significant and was therefore removed. Based on the slope coefficient having a magnitude less than 1, it is clear that the GPS mean travel time is generally less than the Bluetooth mean travel time. Possible reasons for this could include the presence of outliers in the Bluetooth data that were not removed by the filter, the manual exclusion of excessively long travel times from the GPS data (which was observed in the investigation by Wang in 2014), and the inherent bias in the GPS probe vehicle runs (for example, the dedicated driver could have avoided days in which anything "unusual" occurred along the corridor or could have missed the most congested part of the peak period).

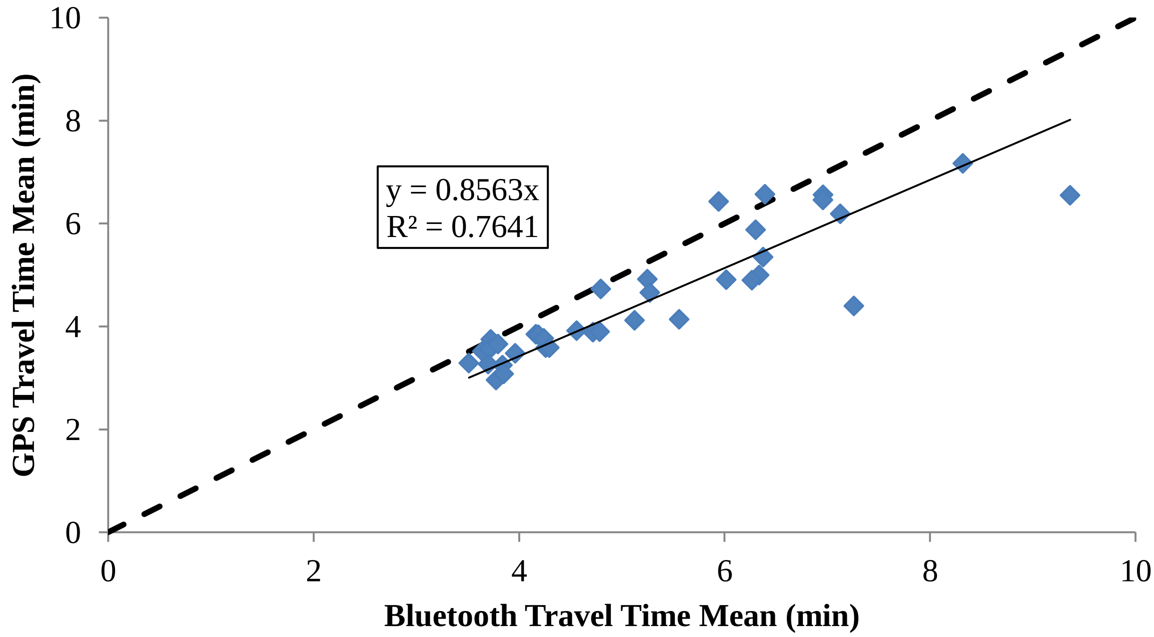

Figure 3.18: Comparison of link-level mean travel times for the field GPS and Bluetooth data

3.4 Characteristics of the Simulated Data

In this section, the simulated data collection site is examined. The generation, processing and verification of the simulated data are then described. An investigation into the primary characteristics of the simulated true travel time data is then presented. These data are presented instead of the Bluetooth and GPS travel time data because this is the data set from which simulated Bluetooth and GPS data are drawn.

3.4.1 Simulated Data Collection Site

The site at which the simulated data were collected is Hespeler Road, a six-lane arterial roadway located in Cambridge, Ontario, Canada. This arterial runs north-south through a major retail area of the city of Cambridge and connects the residential neighbourhoods south of this retail area to Highway 401 in the north. Figure 3.19 shows the corridor of interest along with the location of each Bluetooth detector and the major roadways that cross it. Figure 3.20 shows a map containing the three simulated Bluetooth detectors along Hespeler Road that were used to collect travel time data. The location of each Bluetooth detector is indicated by a yellow thumbtack along with an index and the name of the facility that crosses Hespeler Road where the detector is located.

For the purposes of this investigation, the limits of the corridor of interest were Eagle Street/Pinebush Road in the north and Todd Street in the south. These limits form a study area that is 4.8 kilometres in length.

The simulated study area contains a total of 12 signalized intersections. Intersection turning lanes, raised medians, dedicated bus bays and on-street bus stations are present throughout the study area. The study area is also dense in retail malls and plazas throughout. Three lanes exist from Eagle Street/Pinebush Road to Can-Amera Parkway. The speed limit is 60 kilometres per hour throughout this section. The speed limit then decreases to 50 kilometres per hour for the remainder of the study area. A transition area in which the northbound direction has three lanes and the southbound direction has two lanes is present between Can-Amera Parkway and the Isherwood Avenue/Munch Avenue intersection. South of this intersection, there are two lanes per direction. An at-grade rail crossing exists between the Eagle Street/Pinebush Road intersection and the Langs Drive/Sheldon Drive intersection. Table 3.4 summarizes the primary features of the simulated study area.

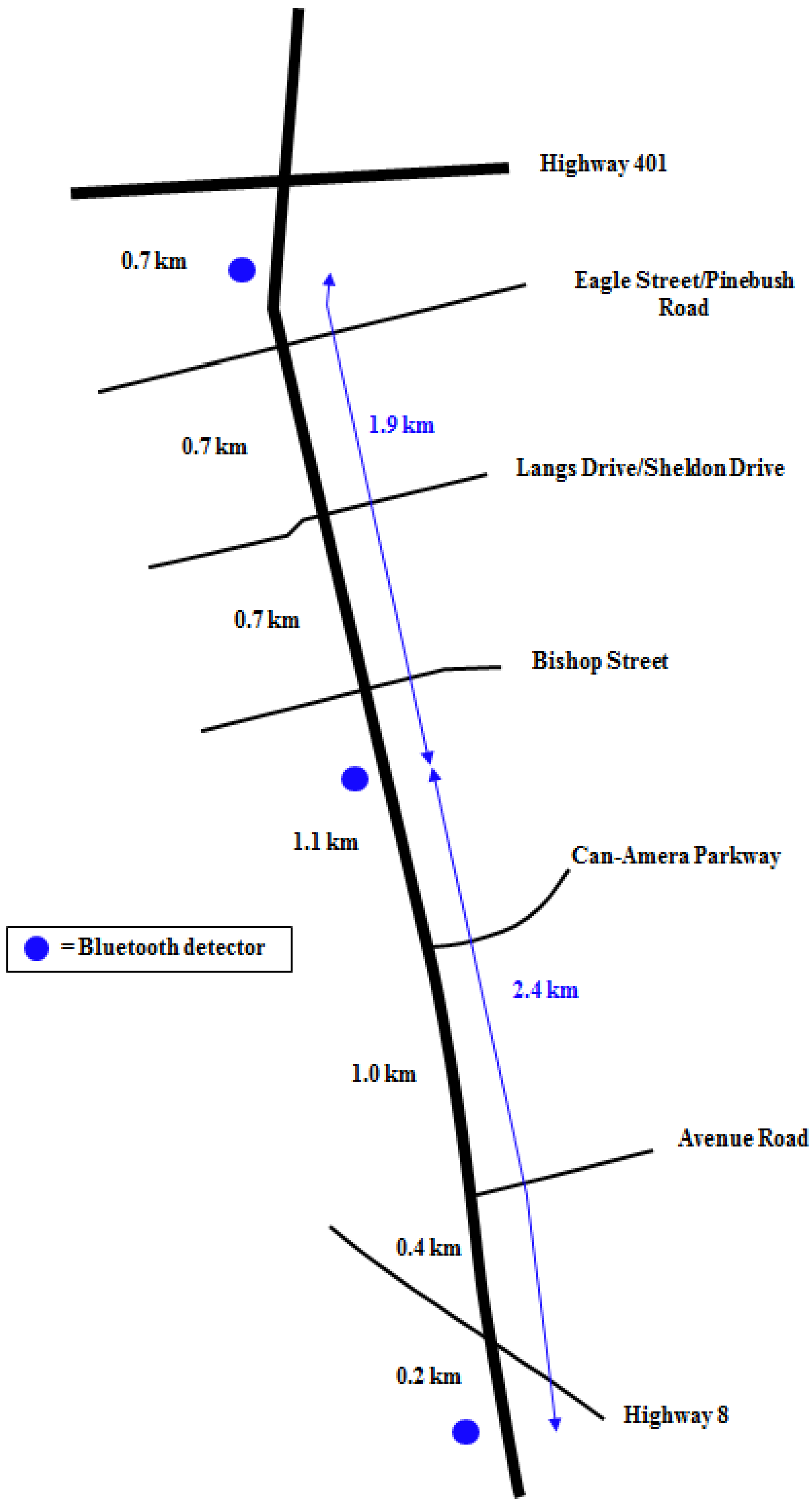

Figure 3.19: Diagram of the simulated study area

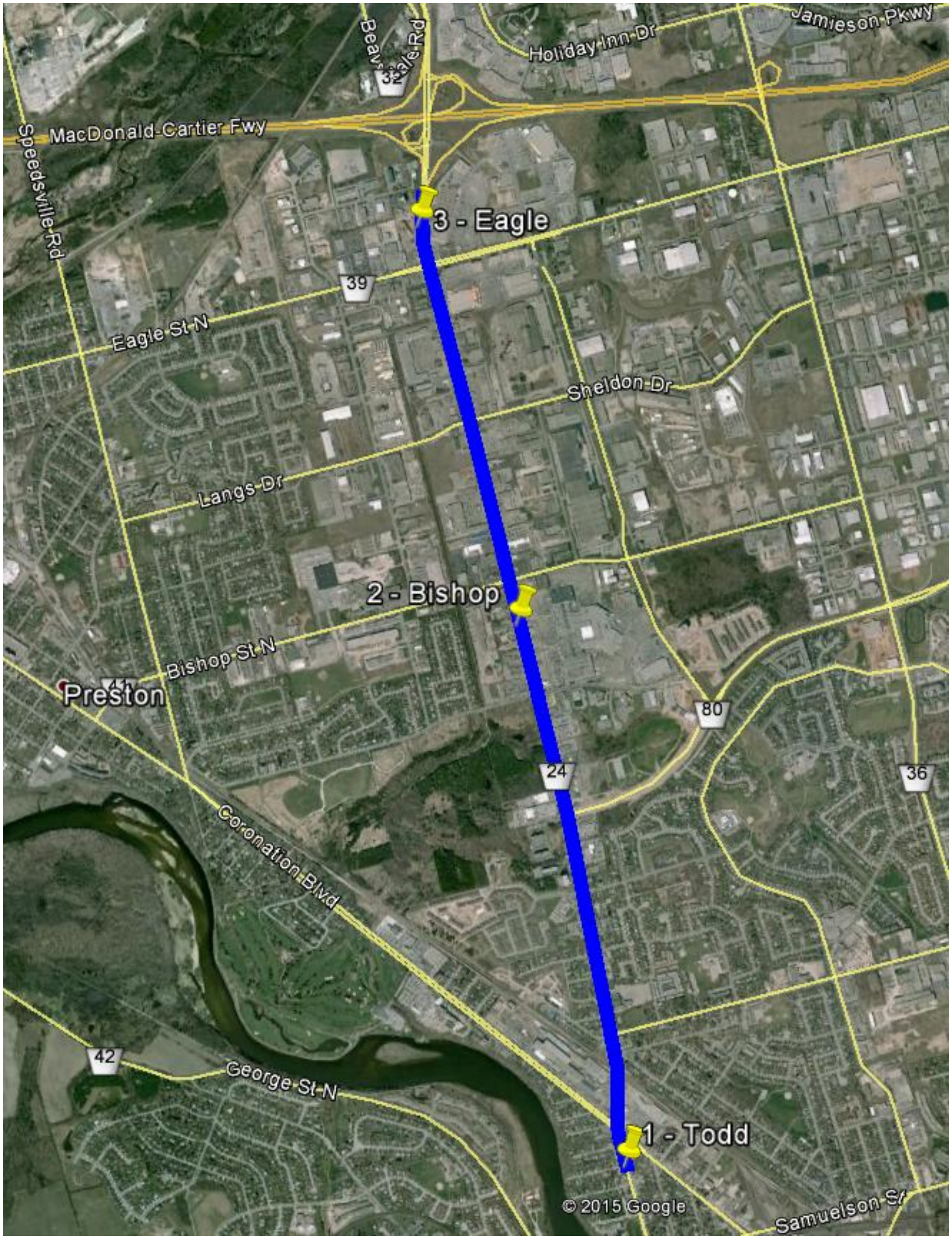

Figure 3.20: Map of the simulated study area (Google, 2014)

Table 3.4: Primary features of the simulated study area

From	To	Speed Limit (km/h)	Number of Lanes per Direction
Eagle Street/Pinebush Road	Can-Amera Parkway	60	3
Can-Amera Parkway	Isherwood Avenue/Munch Avenue	50	3 NB, 2 SB
Isherwood Avenue/Munch Avenue	Todd Street	50	2

3.4.2 Simulation Parameter Settings

The Hespeler Road corridor was simulated using the Vissim software and probe vehicle travel times along the corridor were generated. The simulation environment had the following characteristics and conditions:

- 15-minute turning movement counts for the AM and PM peak periods provided by the Region of Waterloo were used as input into the perimeter entrances of the model. These turning movement counts were collected during various weekdays during various months of the year over a span of four years (2010 to 2013). The AM period was defined to be from 7:30am to 10:30am and the PM period was defined to be from 3:00pm to 6:00pm.
- Static vehicle routing was applied, meaning that turning proportions were directly calculated from the turning movement counts. These turning proportions dictated which turning movement vehicles would make at each intersection in the network.
- Reduced speed zones were included at each intersection to model the need for vehicles to reduce their speed before making a turning movement. More specifically, a right-turning speed of 15 kilometres per hour and a left-turning speed of 30 kilometres per hour were defined.
- Co-operative driving behaviour was enabled for the network. This meant that if vehicle A wants to change lanes but vehicle B is in the way because it is currently in the lane, vehicle B will change lanes to allow vehicle A into its lane. This ensured that traffic progressed freely on the weaving section (i.e. a section in which lane changing is frequent) located on the Hespeler Road-Highway 401 overpass.

- Based on the ratio of total network counts to total network heavy vehicle counts, a heavy vehicle percentage of 5.6% was used throughout the network.
- The existing speed limits along each section of the arterial were used as input into Vissim's prescribed model for assigning a desired speed to each vehicle.
- U-turning vehicles were not modelled.
- Rail traffic was assumed to be too infrequent to have a significant impact on the operations of the corridor and was therefore not modeled.
- All buses that traverse the corridor were modeled according to the weekday fall schedule.
- Bus-bay stops and on-street bus stops were modelled for buses.
- For simplicity, the minimum possible headway was applied to all buses throughout the simulation period.
- Once a vehicle left the network, it was not allowed to return. For example, if a vehicle exited the network at a plaza entrance, it was not allowed to return to the network. This artificially excluded outliers from the resulting travel time data set (i.e. drivers that pass a Bluetooth detector, stop at a plaza, and then pass another downstream Bluetooth detector).
- A warm-up period (i.e. a time period to allow the entry of vehicles into the network) of 15 minutes was used as the average vehicle delay was found to stabilize during this time.
- Pedestrian and cyclist flows were not modelled.
- All other parameters were set to the default values already set in Vissim.

More detailed information on these and other parameters is available in the Vissim 7 User Manual (PTV AG, 2014).

3.4.3 Simulated Signal Timing Plans

Separate signal timing plans were obtained from the Region of Waterloo for each of the AM and PM peak periods. These signal timing plans were used in the simulation of the AM and PM After periods to represent the implementation of signal retiming measures. The majority of the signalized intersections within the corridor operate as fully actuated.

Field signal timings representing the Before conditions were not available for this corridor. Consequently, it was necessary to generate suitable signal timing plans for the Before period.

To produce signal timing plans for the Before period, the signal timing plans of the After period were modified to degrade the progression of traffic along the corridor. More specifically, the minor through and left-turning movements and major left-turning movements were set to minimum recall for the AM and PM peak periods. In the original signal timing plans, the signal controllers would skip these phases in the event that no vehicles made a call for service (i.e. no vehicles triggered the loop detector in the pavement on that approach to indicate the need for a phase change). The addition of minimum recall to a movement meant that even if no vehicles made a call for that movement, the phases for that movement would still run for a period of time equal to the minimum green time every cycle. This modification resulted in less green time being provided to the major street through movements, thereby inhibiting progression and increasing the amount of delay experienced by vehicles traversing the corridor.

Using this strategy, simulated travel times were generated according to four sets of conditions:

- The Before data for the morning peak period (AM data, modified AM signal timing plans),
- The After data for the morning peak period (AM data, AM signal timing plans),
- The Before data for the evening peak period (PM data, modified PM signal timing plans), and
- The After data for the evening peak period (PM data, PM signal timing plans).

A total of four simulation runs were performed, resulting in one day of data generated for each set of conditions.

3.4.4 Generation of Simulated Data

The following steps were performed to prepare the data from each simulation run so that simulated travel time data could be produced:

1. Raw simulated vehicle trajectory points were imported from Vissim into Excel and converted from .fzp files to .txt files that were then imported into MATLAB. Each row of the raw data collected represented a space-time point of the vehicle trajectory and consisted of the ID of the vehicle trajectory, a simulation time stamp (in units of integer seconds), an x-coordinate (in units of decimal degrees), a y-coordinate (in units of decimal degrees), an instantaneous speed (in units of kilometres per hour) and vehicle type (either car or heavy vehicle/bus).

Since the objective was to estimate the travel time of commuter traffic, the travel times of buses and heavy vehicles were discarded. Data points for each vehicle were collected five times per second (i.e. a temporal resolution of 0.2 seconds).

2. Typically, GPS probe vehicles collect data at a frequency of one observation per second. Consequently, GPS points were created by collecting all points with a simulation time stamp that was a multiple of 1 second.
3. From the literature, the polling frequency for Bluetooth detectors was found to be 1.28 seconds (Moghaddam, 2014). Consequently, Bluetooth points were created by collecting all points with a time stamp that was a multiple of 1.2 seconds.
4. The location of each Bluetooth detector was selected. A total of three detectors were placed for data collection based on the number of detectors per kilometre for the field data corridor. Along the field study area, seven detectors were distributed across 15.1 kilometres, yielding approximately one detector every two kilometres. Applying this rate over the simulated study area's distance of 4.8 kilometres, a total three detectors were placed: one at the southwest corner of the Todd Street-Hespeler Road intersection (Detector 1), one approximately 200 metres south of the intersection at Bishop Street on the west side of Hespeler Road (Detector 2), and one approximately 160 metres north of the intersection at Eagle Street on the west side of Hespeler Road (Detector 3). In an attempt to minimize sources of travel time variability as discussed by Day et al. (2012), these detectors were not placed directly at major intersections.
5. The location of each "true detector" was selected. True detectors were defined to be detectors used to "detect" true points and GPS points. This meant that for each Bluetooth detector used to collect Bluetooth points, there was a corresponding true detector used to collect true points and GPS points. Each true detector was placed close to the corresponding Bluetooth detector but was instead placed in the middle of the roadway to reflect the fact that true travel times and GPS travel times start and end based on the limits of the link directly on the roadway. In other words, Bluetooth travel times are produced using roadside detectors while true travel times and GPS travel times are produced using the corresponding on-road detectors.

To produce simulated true travel time data, the following steps were performed:

1. The distance from each true point to each true detector was calculated. Knowing that vehicles more than 100 metres from a Bluetooth detector cannot be detected, these points were excluded. While it is known that true travel time data do not operate on the basis of detectors, the exclusion of true points that were a certain distance away from a "true detector" reduced the time needed to process the data.
2. For a given link, trajectories that had points appear at both the start and end true detectors were identified. Travel times were then calculated using the "minimum-minimum" strategy. In other words, for a given trajectory, the time stamp of the point that was the minimum distance from each true detector was used to produce a travel time. This produced simulated true travel times at the link level.
3. To produce true travel time data at the path level, true trajectories that traversed all links were identified. The link-level true travel times of these trajectories were then used to produce path-level true travel times.

The simulated true travel time data are available in Appendix B.

To produce simulated GPS travel time data, the following steps were performed:

1. The distance from each GPS point to each true detector was calculated and points more than 100 metres from a detector were excluded.
2. For a given link, trajectories that had points pass both the start and end true detectors were identified. Travel times were then calculated using the "minimum-minimum" strategy.
3. To produce GPS travel time data at the path level, GPS trajectories that traversed all links were identified. The link-level GPS travel times of these trajectories were then used to produce path-level GPS travel times.
4. According to Wang (2014), it is common practice by the City of Toronto when evaluating signal retiming measures to collect data from five GPS runs for each of the morning and evening peak periods collected over a span of three days for each direction. Consequently, in this study, five path-level GPS travel times were randomly selected for each combination of time period (morning peak and evening peak), direction, and Before and After data set. This yielded a total of 40 selected GPS trajectories that were known to have traversed both links of the simulated corridor.

The simulated GPS probe vehicle travel time data are available in Appendix B.

To produce simulated Bluetooth travel time data, the following steps were performed:

1. A penetration rate of 10% of the vehicle population was assumed. This meant that 10% of all trajectories were able to be detected by the detectors (i.e. Bluetooth-enabled). Trajectories were randomly selected to be Bluetooth-enabled according to this percentage. This percentage was selected based on the identification of the current range of Bluetooth penetration to be between 2% and 10% (Moghaddam, 2014). The upper bound of 10% was selected to reflect the expectation that Bluetooth technology will become more prevalent among vehicles over time. It must be noted that an assumption built into this step was that a Bluetooth-enabled vehicle remained Bluetooth-enabled throughout its trajectory.
2. The distance from each Bluetooth point to each Bluetooth detector was calculated and points more than 100 metres from a detector were excluded.
3. Using the distance to the detector as input into the trapezoidal probability distribution developed by Moghaddam (2014), the probability of detection for each Bluetooth point was calculated.
4. Whether or not a Bluetooth point was actually detected was identified based on whether or not a random number between 0 and 100 (decimals included) assigned to that point was less than its probability of detection. For example, if a point had a probability of detection of 20% and it was assigned a random number of 10, that point would be classified as "detected". Conversely, if that same point had been assigned a random number of 60, that point would be classified as "not detected". All points that were classified as "detected" were then collected.
5. For a given link, trajectories that had points appear at both the start and end Bluetooth detectors were identified. Travel times were then determined for these points using a "middle-middle" strategy (i.e. if more than one point was obtained for a given vehicle trajectory at a given detector, then the middle point was selected for calculating the travel time).
6. To produce Bluetooth travel time data at the path level, Bluetooth trajectories that traversed all links were identified. The link-level Bluetooth travel times of these trajectories were then used to produce path-level Bluetooth travel times.

The simulated Bluetooth travel time data without error incorporated (which is a concept introduced in the next section) are available in Appendix B.

A key characteristic of the Bluetooth data produced using this methodology is that frequency hopping was not incorporated in the aggregation process. For Bluetooth technology, frequency hopping is used to reduce the effects of signal interference and to abide by regulatory requirements (Institute of Electrical and Electronics Engineers, 2005). This results in a "lockout period" in which a hit cannot be detected for a given MAC ID for a certain time period after it is first detected. From a data collection standpoint, a reduction in the number of hits collected for each MAC ID will occur as a result of frequency hopping. However, the goal of this investigation was not to create a detailed model of the entire Bluetooth detection process but to capture the key components of this process with respect to the generation of travel time data.

3.4.5 Incorporation of Error in the Simulated Bluetooth Data

While the lockout period seen in Bluetooth data was not modeled, the presence of error in the Bluetooth data was considered to examine the impact of outliers that were not removed by an outlier detection algorithm. Consequently, this procedure aimed to model the presence of "missed outliers" in the data.

To incorporate these "missed outliers" in the simulated Bluetooth data, the following steps were performed:

1. It was assumed that 5% of all link-level travel times would be classified as "missed outliers". The actual fraction present in the field data is unknown, but it is clearly a function of the outlier detection algorithm applied to the raw travel time data.
2. The distribution of travel time for these "missed outliers" was determined by multiplying the recorded Bluetooth travel time by an error factor, $(1 + \varepsilon)$. The parameter ε was assumed to be lognormally distributed. The lognormal distribution was selected due to its similarity in shape to the distribution of travel times for the field data. The lognormal distribution allowed for a lower bound to be defined, thereby forcing the generated errors to cause an increase in travel time for those travel times identified as missed outliers. The presence of excessively high travel times was indicated by the long tail to the right of the travel time distributions of

the field data. Equation 3.1 shows the cumulative distribution function used to produce values of ε for a given random number x between 0 and 1 (The Mathworks, Inc., 2015):

$$\varepsilon = \frac{1}{x * \sigma * \sqrt{2 * \pi}} * \int_0^x \left(\frac{\exp\left(\frac{-(\ln(t) - \mu)^2}{2 * \sigma^2}\right)}{t} \right) dt \tag{3.1}$$

3. Through trial and error, the mean μ was selected to be -2 and the standard deviation σ was selected to be 0.9 for this cumulative distribution function. To prevent the situation in which the travel time is unrealistically large (and most likely would have been identified by even a poor outlier detection algorithm), the value of ε was constrained to be less than or equal to 1. This meant that a travel time classified as a missed outlier could only have its travel time increase by a factor between 1 and 2. Figure 3.21 illustrates the distribution of ε.

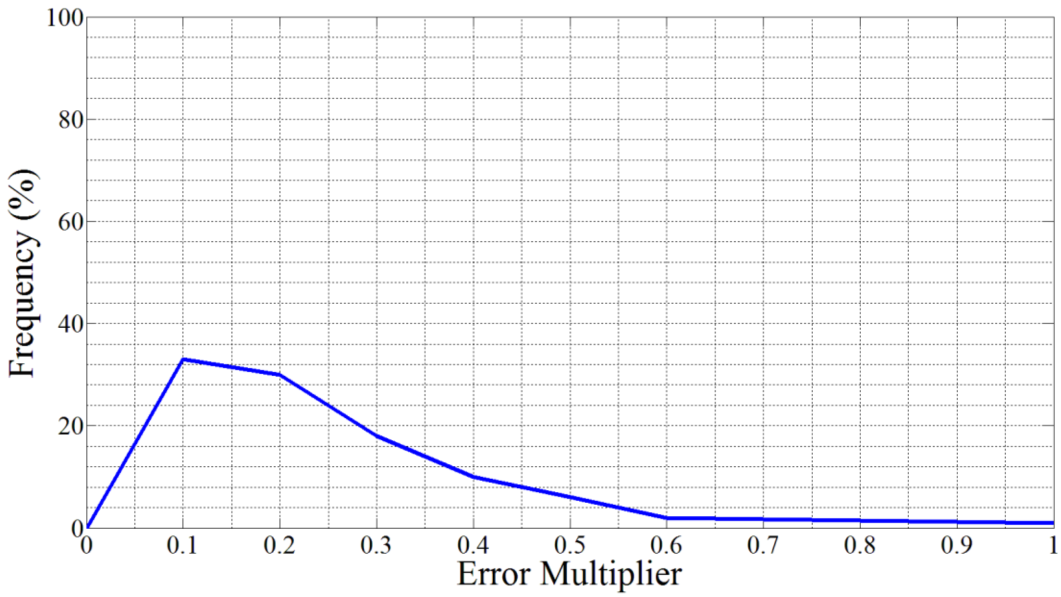

Figure 3.21: Distribution of the simulated Bluetooth travel time errors

4. To capture the impact of these outliers at the path level, the path-level Bluetooth travel time data was regenerated by summing the newly modified link-level travel times of each Bluetooth trajectory that was known to traverse the entire corridor.

The simulated Bluetooth travel time data with error incorporated is presented in Chapter 4.

3.4.6 Verification of Simulated Bluetooth and GPS Probe Vehicle Travel Time Data

Knowing that the true travel time data are available for comparison to the sampled Bluetooth and GPS probe vehicle data, a formal verification of these sampled simulated data is presented in this section.

Figure 3.22 shows the results of a regression analysis carried out between the link-level simulated Bluetooth travel time data without error (i.e. “missed outliers” were not considered) and the corresponding true travel time data. Figure 3.23 shows the results of a regression analysis carried out between the link-level simulated Bluetooth travel time data with error included and the corresponding true travel time data. Figure 3.24 shows the same results with the data separated according to whether or not a travel time received error. Figure 3.25 shows the results of a regression analysis carried out between the link-level simulated GPS travel time data and the corresponding true travel time data. In this context, the “corresponding true travel time data” means that for a given vehicle ID, each Bluetooth (or GPS) travel time was plotted as a function of the true travel time belonging to the same vehicle ID.

It can be seen in each figure that the slope coefficient is close to 1, indicating little bias in each model. R-squared values are also very close to 1 for all three models, indicating high correlations between each sampled data set and the true travel time data set. Insufficient evidence was found to support the statistical significance of the y-intercept for all three cases. It can also be observed that the slope coefficient changes from being less than 1 for the Bluetooth data without error to greater than 1 for the Bluetooth data with error. This was expected as the error incorporated in the Bluetooth resulted in travel times that had extra time assigned to them to represent “missed outliers”, resulting in an increase in the slope coefficient.

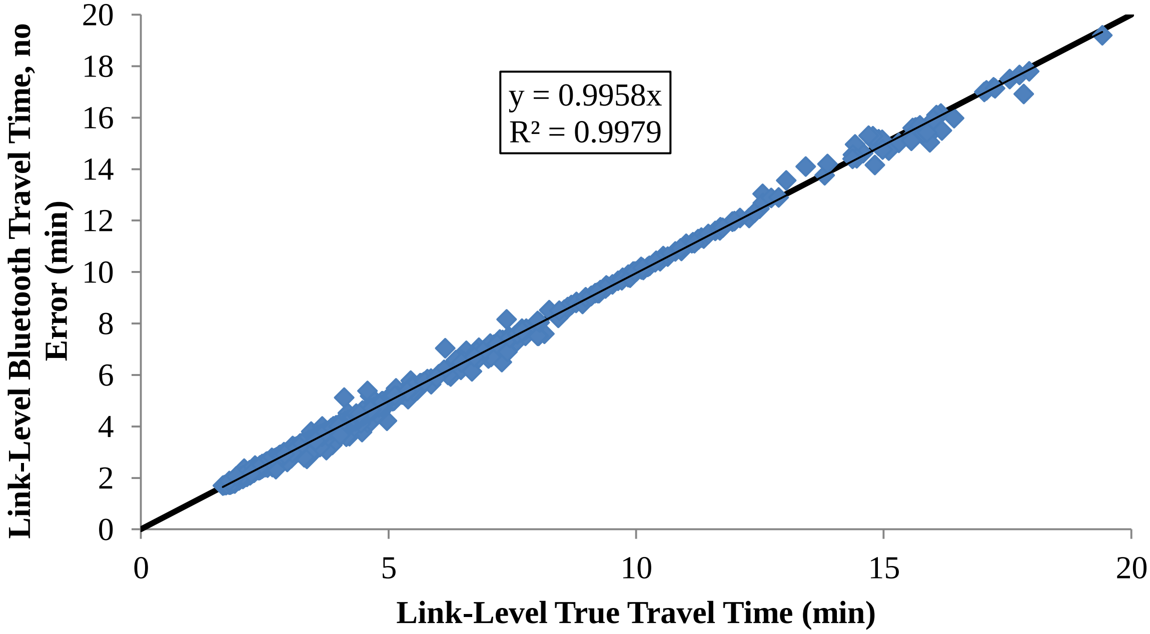

Figure 3.22: Comparison of path-level simulated Bluetooth travel time data without error and true travel time data

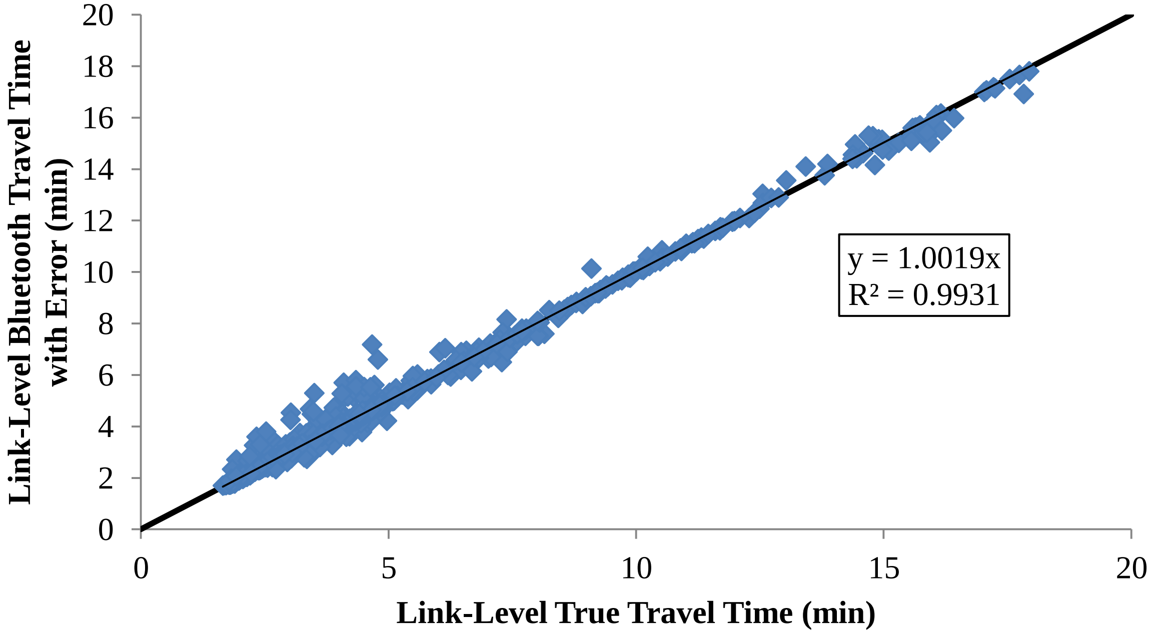

Figure 3.23: Comparison of link-level simulated Bluetooth travel time data with error and true travel time data

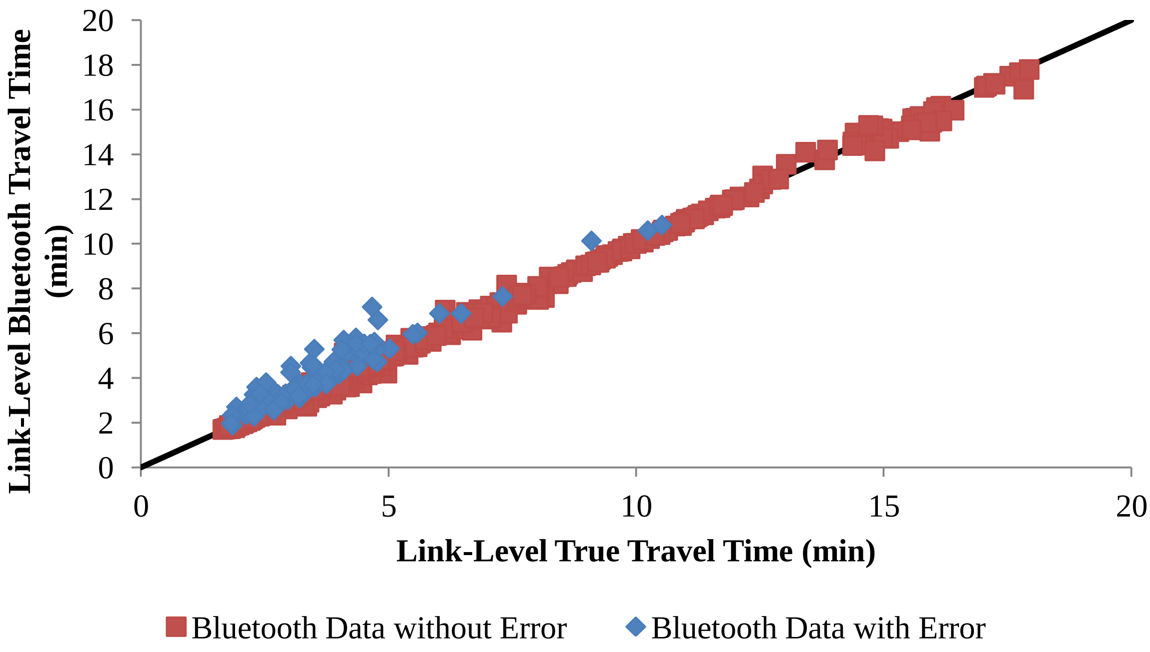

Figure 3.24: Comparison of link-level simulated Bluetooth travel time data with error and true travel time data with data separated according to error

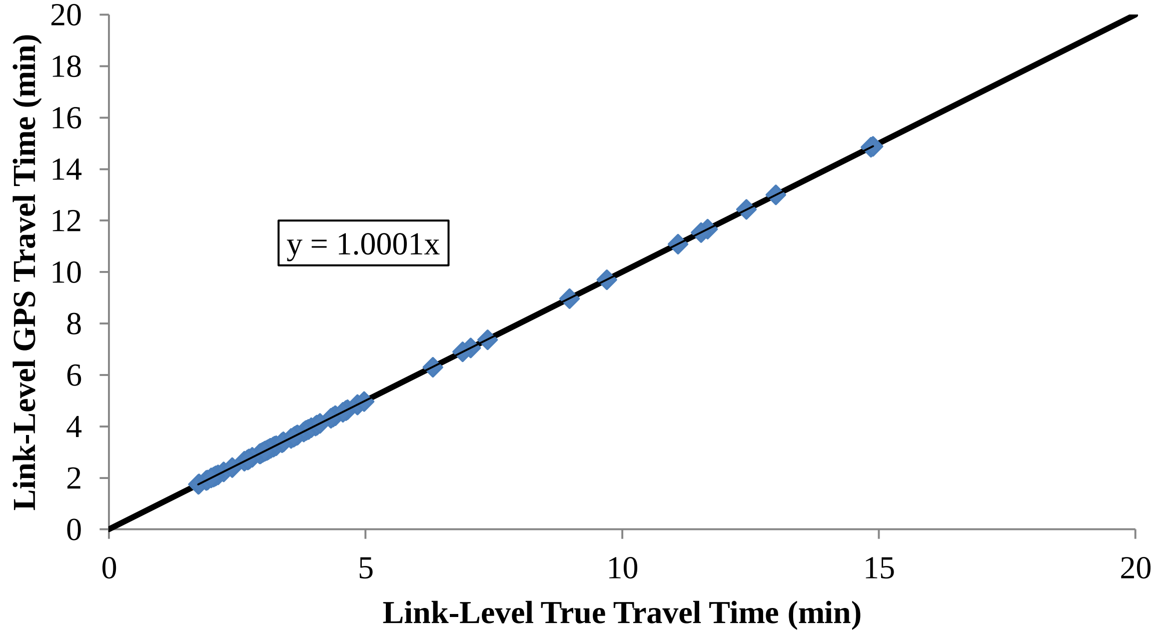

Figure 3.25: Comparison of link-level simulated GPS travel time data and true travel time data

3.4.7 Characteristics of the Simulated True Travel Time Data

Figure 3.26 shows the simulated true travel times as a function of time of day for a given link, direction and time period during the Before period while Figure 3.27 shows the true travel times under the same conditions during the After period. For this corridor, it was observed that distinct peaks developed for the southbound direction for both links but not for the northbound direction. This indicated that traffic behaviour for the northbound direction remained relatively constant for both the Before and After periods while the southbound direction experienced a "peaking" of travel times similar to what was observed for the field corridor.

In these figures, it can be observed that the maximum travel time has been reduced from the Before period to the After period from approximately 20 minutes to 13 minutes. This highlights the travel time savings produced by the removal of the "minimum recall" feature implemented during the Before period to purposely impede the progression of traffic in the simulated environment.

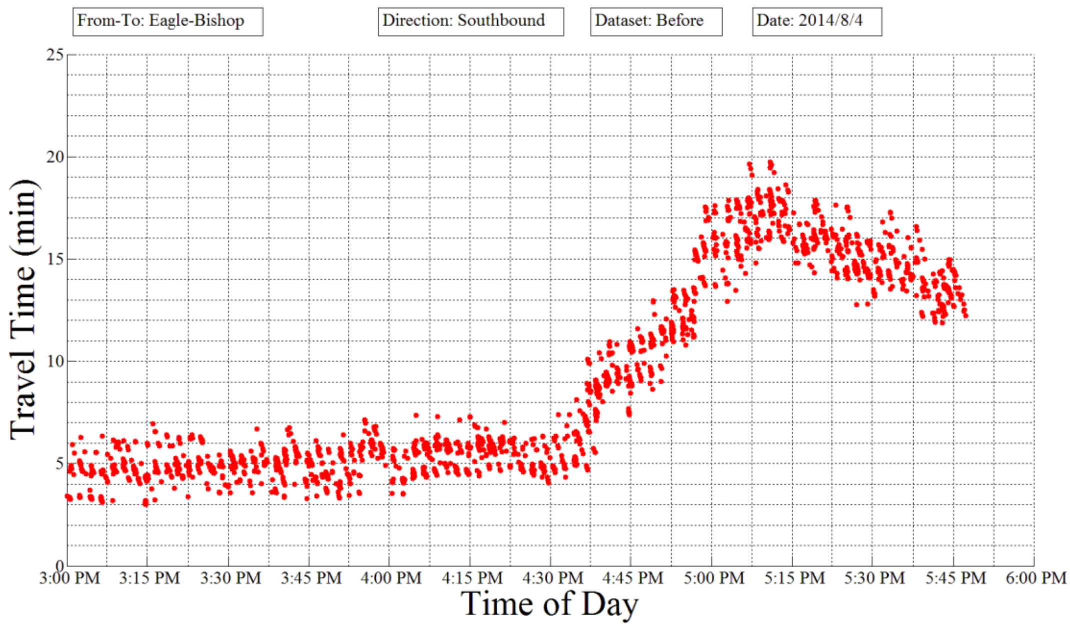

Figure 3.26: Travel time as a function of time of day for a single day for the field Bluetooth data during the Before period

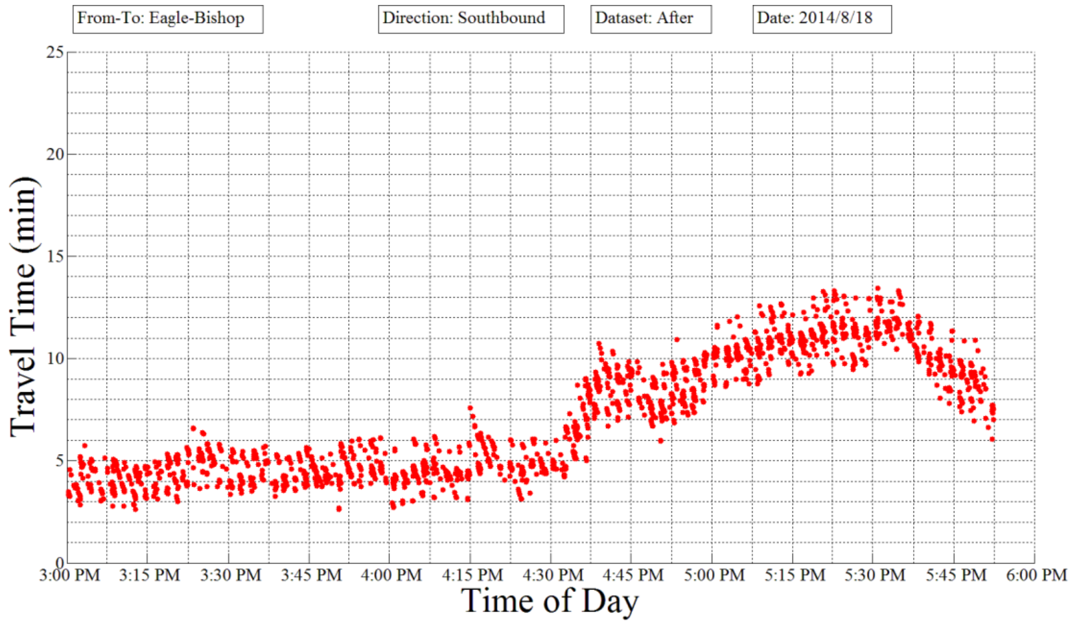

Figure 3.27: Travel time as a function of time of day for a single day for the field Bluetooth data during the After period

These differences in the peak period travel times can be attributed to increased traffic demand at several of the signalized intersections on this link. Figure 3.28 and Figure 3.29 show screenshots of the simulated environment at the intersection of Hespeler Road and Bishop Street for the evening peak period during the Before and After periods respectively. Each screenshot was taken at approximately 5:15pm in the simulation environment to capture the peak traffic behaviour observed for the southbound direction. Traffic travelling southbound along at Bishop Street was therefore observed for queuing behaviour.

It can be seen in Figure 3.28 that a relatively long queue has developed on the southbound approach, which is a result high traffic demand and poor progression. Furthermore, it can be observed that the queue from the next downstream intersection has spilled back to Bishop Streetm resulting in oversaturated conditions. These conditions persisted at this location for approximately 30 minutes and explain the distinct peak in travel times observed in Figure 3.26.

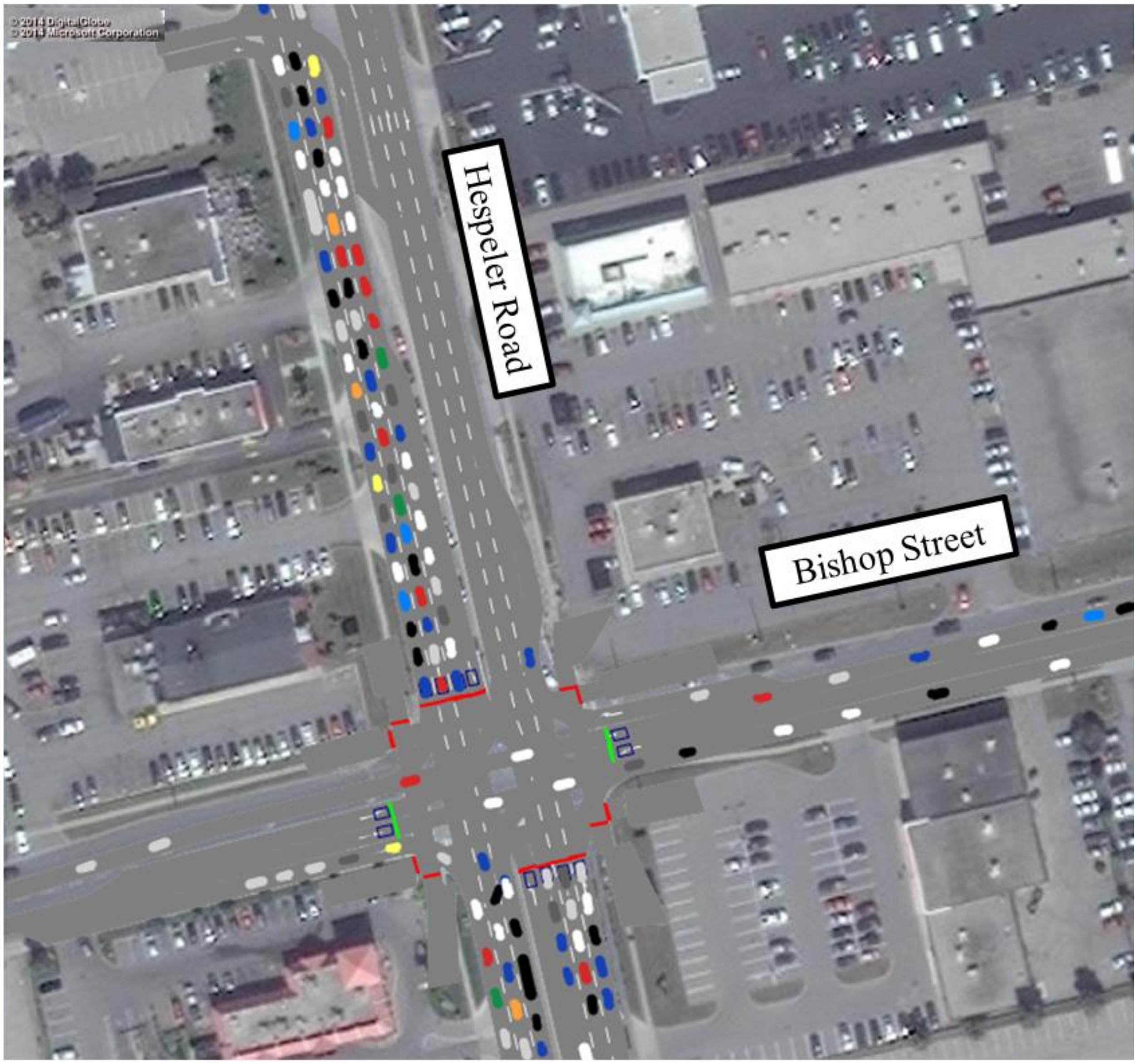

Figure 3.28: Screenshot of the simulation environment at Hespeler Road and Bishop Street for the evening peak period during the Before period

Conversely, Figure 3.29 shows a shorter queue at the southbound approach to Bishop Street and no queue spilling back to Bishop Street from the next downstream intersection. No oversaturation (i.e. queue spillback to adjacent intersections) was observed at this location during the After period. These figures highlight the differences between the Before and After periods in terms of traffic progression and provide a visual representation of the differences in travel time that were observed between the Before and After periods.

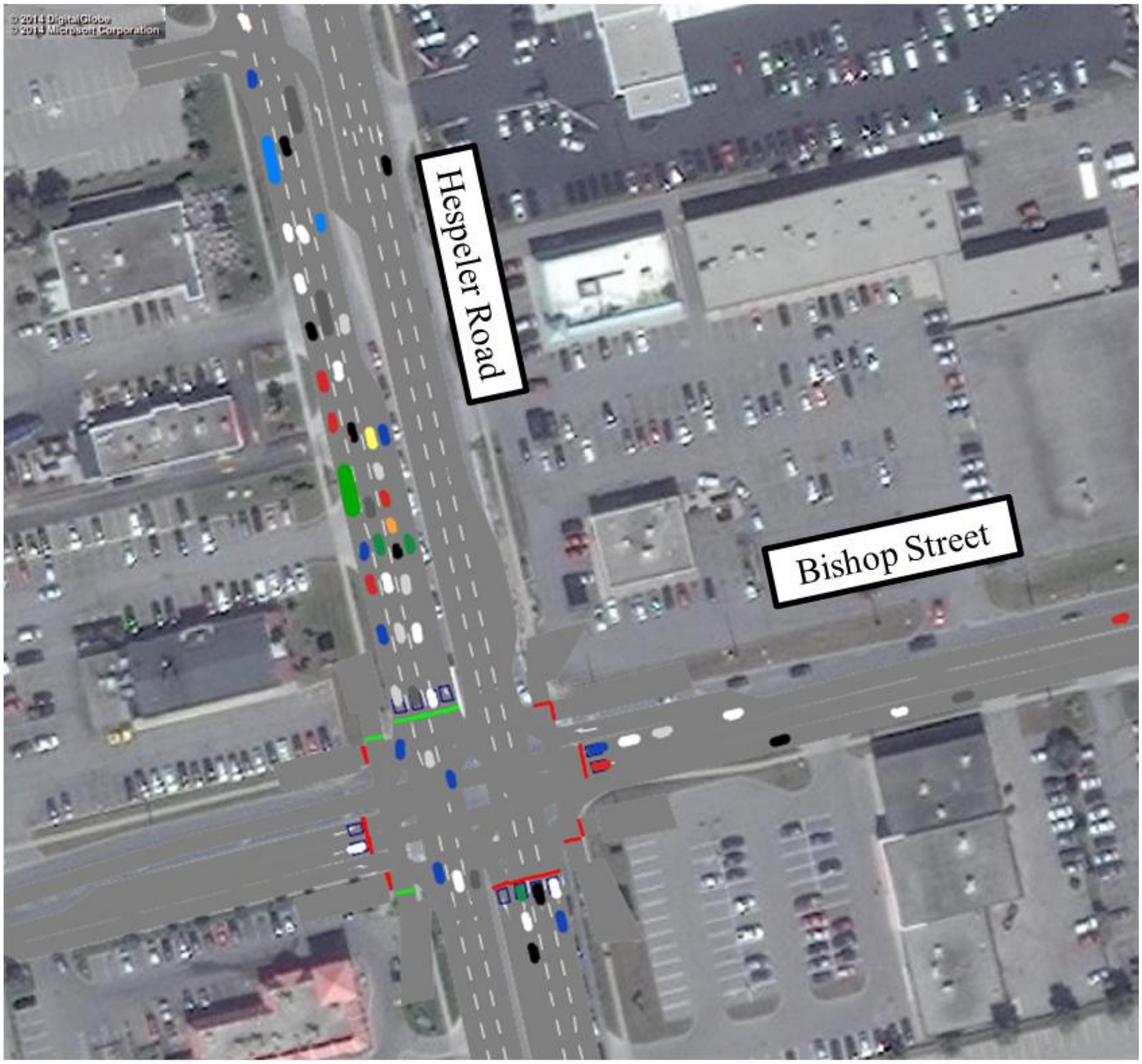

Figure 3.29: Screenshot of the simulation environment at Hespeler Road and Bishop Street for the evening peak period during the After period

Figure 3.30 shows the number of simulated true travel times collected as a function of time of day for a given link, direction and time period during the Before period. An aggregation period of 30 minutes was used. As with the field data, the contrast between the amount of travel time data available at the link level and at the path level can be easily observed. For this link a larger number of observations were obtained for the evening peak period compared to the morning peak period, which suggests that an evening commuting pattern into the city may exist. This observation is also consistent with the fact that congestion is more pronounced in the evening peak period than the morning peak period. More observations were found to have been collected for this link during the morning peak period compared to the evening peak period for the northbound direction, which suggests that a morning commuting pattern out of the city may also exist.

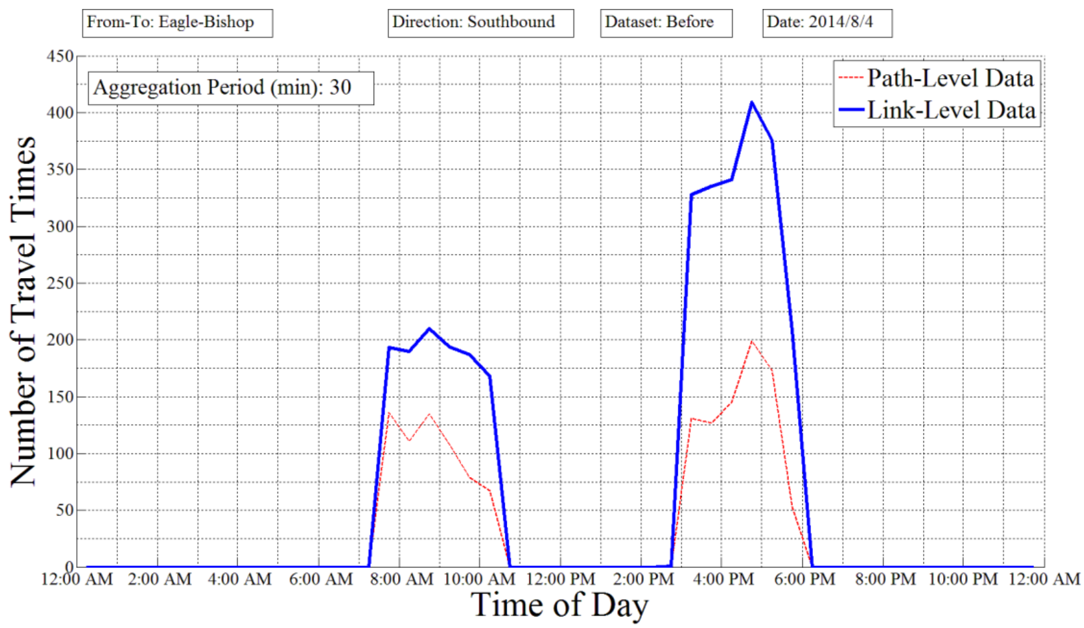

Figure 3.30: Number of travel time observations as a function of time of day for a single day for the field Bluetooth data

Figure 3.31 shows the cumulative travel time frequency distributions for the Before and After simulated Bluetooth data for a given link, direction and time period. As seen in this figure, travel times approximately four minutes in duration are most common. It can also be observed visually that the two distributions are similar in shape with the After distribution having a slightly greater frequency of shorter travel times (i.e. the After distribution is to the left of the Before distribution), indicating that a reduction in travel times has occurred.

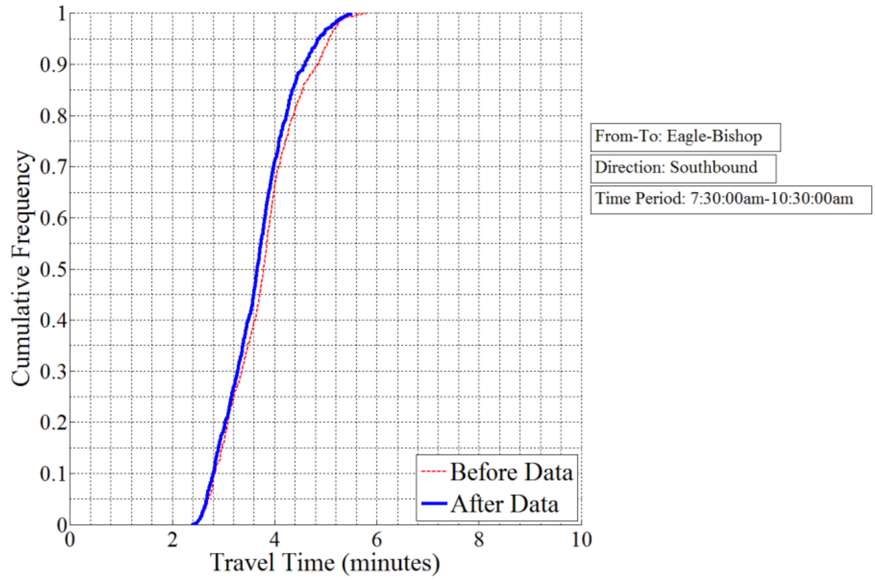

Figure 3.31: Sample travel time cumulative frequency distribution for the simulated true travel time data

Figure 3.32 shows the mean travel times per kilometre for the true travel time data detected for each link for a given direction and time period. Travel times on the northernmost link are slightly higher than those of the southernmost link, which is a reflection of the greater amount of congestion that the northernmost link experiences.

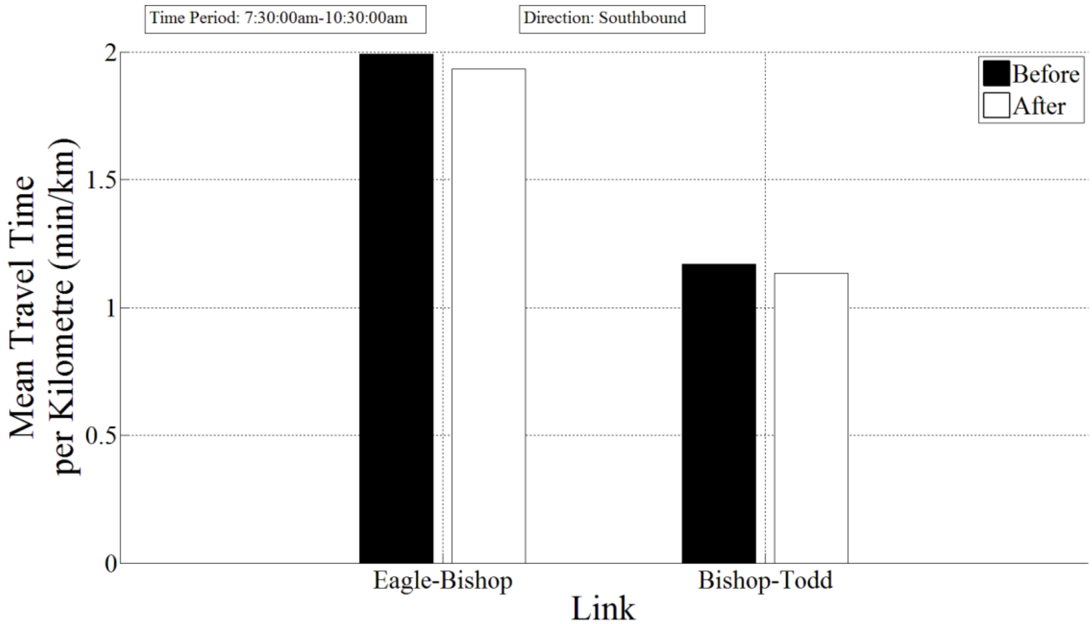

Figure 3.32: Mean travel time per kilometre of each link for the simulated true travel time data

Figure 3.33 shows the number of true travel times per kilometre collected at each link for a given direction and time period. The greater amount of congestion seen on the northernmost link is reflected by the slightly greater number of vehicles that traverse that link compared to the southernmost link.

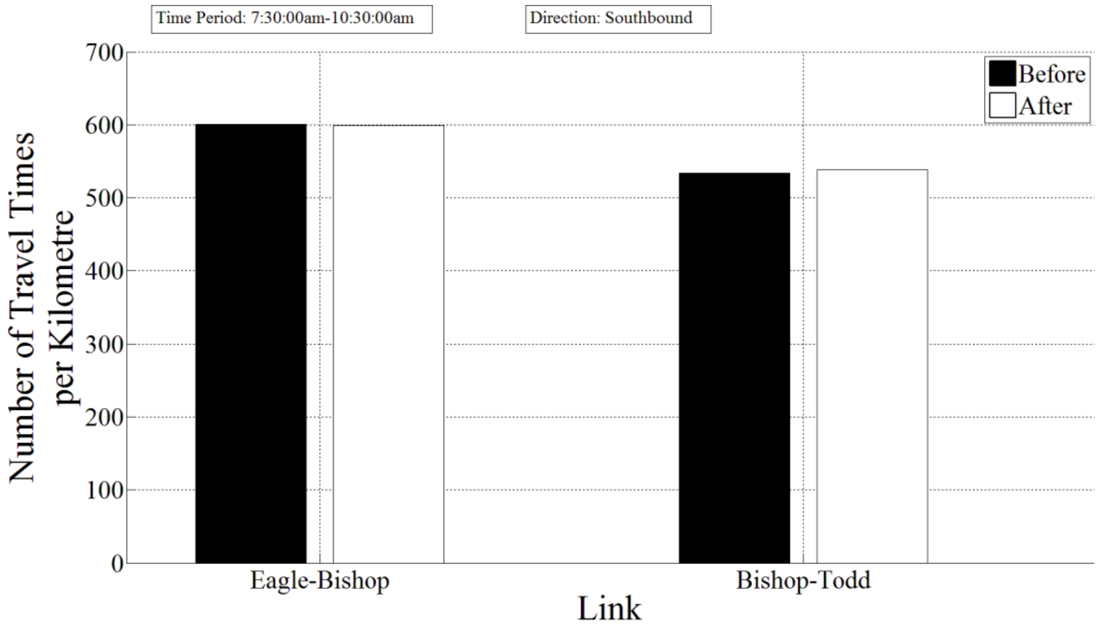

Figure 3.33: Number of observations per kilometre collected for each link for the simulated true travel time data

Figure 3.34 and Figure 3.35 show the percentage of vehicles that appeared at the starting link and at each following link along the corridor for each direction for the simulated true travel time data. For both directions, approximately 30% of the vehicles that traversed the Todd-Bishop link also traversed the Bishop-Eagle link. This indicates that the majority of vehicles exited the corridor somewhere along the Bishop-Eagle link and did not traverse the entire link. These figures show that the majority of vehicles that enter the corridor do not traverse the entire corridor.

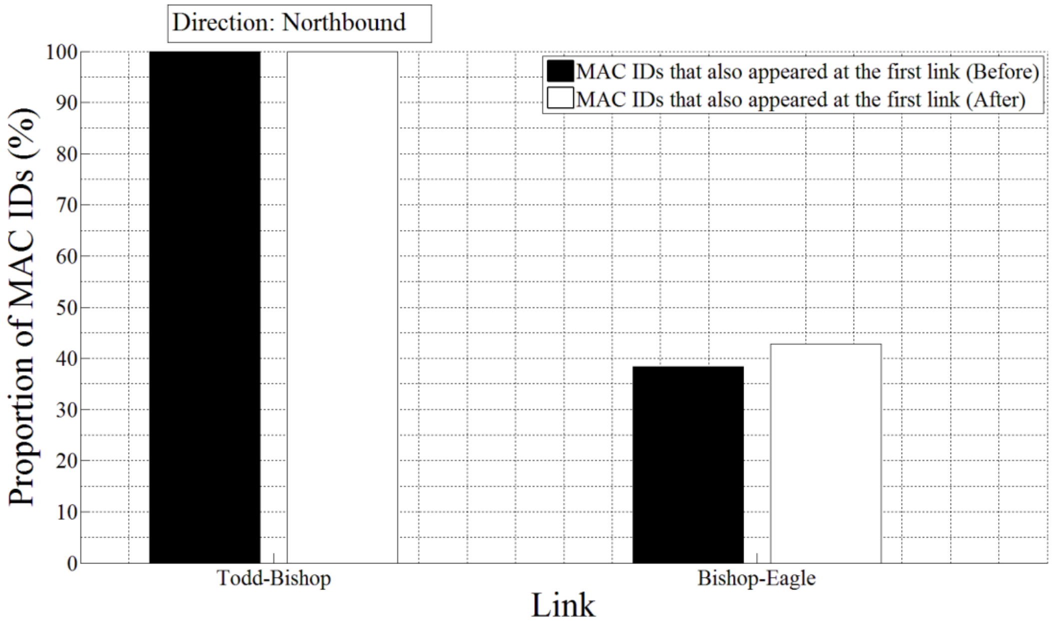

Figure 3.34: Propagation of simulated true travel time data MAC IDs in the northbound direction

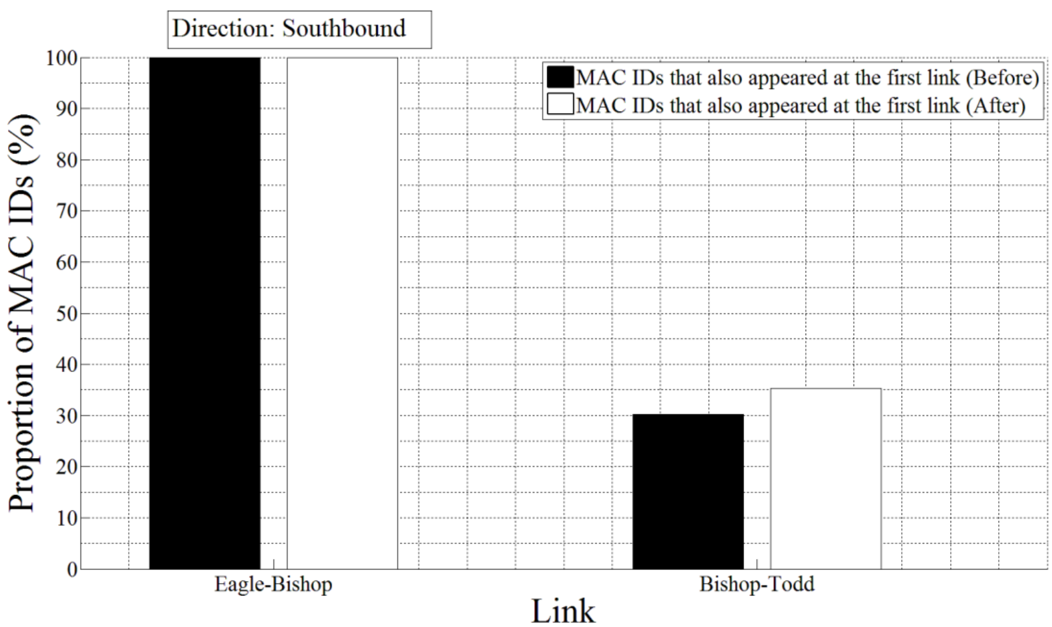

Figure 3.35: Propagation of true travel time data MAC IDs in the southbound direction

3.5 Summary

In this chapter, the data collection sites and basic characteristics of the travel time data collected in the field and in a simulated environment were described and examined in detail. The following findings were concluded from this chapter:

- Day-of-week and time-of-day variation in travel time and the number of available samples are apparent in the field Bluetooth travel time data. Time-of-day variation is less evident in the simulated true travel time data, likely due to lower traffic volumes and less variation in traffic demands during the simulated peak period.
- The number of Bluetooth travel times available at the path level is much smaller than the number of samples available at the link level.
- As seen in the field Bluetooth travel time data, links belonging to the same path exhibit significantly different mean travel times relative to each other, indicating a varying influence of impedances to traffic progression on each link. Link length could be one such impedance.
- Examining the magnitudes and frequencies of the travel times collected for the simulated site indicated that the a greater number of vehicles and a greater peak in travel times exist during the evening peak period for the southbound direction, especially along the Bishop-Eagle link leading to Highway 401. This suggests significant commuter traffic arriving onto Hespeler Road from Highway 401 and heading southbound during the evening peak period.
- As seen in the field Bluetooth travel time data, links belonging to the same path experience significantly different vehicle volumes (i.e. the number of travel times collected) relative to each other, which could be due to some links containing access to more amenities than others.
- As seen in the field Bluetooth travel time data, approximately 25% of MAC IDs appear across multiple days for a given link, direction and time period. This indicates that a significant percentage of Bluetooth-enabled vehicles traverse the corridor under these conditions on a regular (i.e. daily) basis and suggests that commuters represent a significant portion of the Bluetooth-enabled vehicles.
- The progression of MAC IDs can be tracked along the corridor, thus allowing for the creation of OD matrices using Bluetooth travel time data.

- The theoretical estimation of the probability of a missed detection produced a very low value, which gives a practitioner confidence that Bluetooth technology is a reliable source of travel time data for an arterial corridor. However, additional factors may lead to the occurrence of missed detections.
- A comparison of the field Bluetooth link-level mean travel times from the Before and After periods suggested significant differences in the means. This highlights the potential for the mean to be used to assess the effects of signal retiming measures.

4 Comparison of Path-Level and Link-Level Bluetooth Travel Time Data

4.1 Introduction

This chapter describes the investigation of whether or not Bluetooth travel time data collected at the path level is an adequate replacement for Bluetooth travel time data collected at the link level. First, the methodology used to compare the path-level data to the link-level data is introduced. This methodology is then applied to the path-level and link-level field and simulated data. This chapter concludes with a summary of the major observations made regarding the results of the link versus path comparison and a recommendation as to which should be utilized in the evaluation of signal retiming measures.

4.2 Methodology

The motivation for comparing Bluetooth data at different scales was that the collection of path-level Bluetooth travel time data requires far less time and resources due to the need for only two detectors. However, with ease of collection comes a vastly decreased sample size relative to that of link-level data. In other words, for a given period of time, the number of vehicles that travelled between two detectors spaced 15 kilometres apart will be less than if the detectors were spaced 3 kilometres apart (assuming that there are intermediate intersections and mid-block driveways along the corridor). The resulting number of observations (i.e. sample size) dictates the statistical significance of the differences between the collected Before and After travel time data, making it important to establish the magnitude of the sample sizes present at the path level in comparison to the link level. This section details the investigation of whether or not this reduction in sample size is enough to render path-level data unfit for signal retiming assessment in comparison to link-level data.

To examine whether or not path-level data were an adequate replacement for link-level data, the following questions were posed:

1. What is the relationship between the path-level data and the link-level data?

In order for the two data sets to be compared to each other, it must first be established that the two data sets are in fact related to each other. The existence of a relationship between these data sets can be verified by identifying link-level travel times that "belong" to a path-level travel time. In other words, if a MAC ID appears at the path level, then that same MAC ID must also appear at all links belonging to that path and a distinct relationship between travel times at the two levels must be evident.

To answer this question, each MAC ID that possessed a travel time at the path level also had its travel times collected at the link level. Whether or not each MAC ID appeared at all links was identified. For each MAC ID, the sum of the link-level travel times was calculated and compared to the path-level travel time. It was expected that the path-level travel time would be equal to the sum of the link-level travel times. The purpose of this first question was to confirm this expectation.

2. Are the observed differences in the Before and After means and standard deviations statistically significant?

If it can be shown that mean travel times from the After period are statistically smaller in magnitude than the mean travel times from the Before period, and the only change that has occurred along the corridor between the Before and After periods has been the change in signal timings, it can be concluded that the implemented signal retiming measures have produced significant benefits in terms of travel time reduction along the corridor. However, if no statistically significant difference is observed, then it cannot be concluded that the implemented signal retiming measures have produced significant benefits.

Similarly for the standard deviation, if it can be shown that the standard deviation has been reduced between the Before and After periods, it can be concluded that the implemented signal retiming measures have produced significant benefits in terms of a reduction in travel time variation along the corridor.

The ability of the link-level and the path-level data sets to support this type of analysis is examined in this chapter. Statistical comparisons were performed between the mean travel times and standard deviations of travel times collected during the Before and After periods using the parametric two-sample t-test for means and the f-test for standard deviations. In order to be able to use these statistical tests, the assumptions associated with these statistical tests must be satisfied.

According to Lund Research (2013), the following assumptions apply to the use of Student's two-sample t-test (rationale is also provided as to how the Bluetooth data meets these assumptions):

1. The data should be continuous.

This is true as the travel time data are continuous in nature.

2. The data should consist of two categorical, independent groups.

This is true as the Before and After time periods differentiate the two data sets.

3. Observations should be independent of each other.

The travel times of vehicles which travel together along a roadway are not strictly independent because the vehicles are subject to the same traffic signal timing delays and the behaviour of one vehicle may impact the travel time of vehicles upstream of it. However, given that Bluetooth travel time data represent a sample of the population of vehicle travel times, the probability of obtaining travel times from vehicles that are in direct interaction as they traverse the link is very low.

The possibility of acquiring multiple travel times from the same vehicle as a resulting of it making multiple trips along the link is also very low From the field data, it was found in Chapter 3 that each travel time is associated with a MAC ID that only appears once per day for a given combination of link, direction, time period and Before or After data set. For the simulated data, vehicles were prevented from re-entering the network, thus making each simulated travel time independent.

4. No significant outliers exist in the data.

It is assumed that this is true as the outlier filtering algorithm applied to the field data is assumed to remove significant outliers. In the simulated data, missed outliers were added to enhance the realism of the data but a travel time that was classified as a missed outlier was restricted from being more than twice as large as its base travel time. This ensured that the magnitudes of the resulting missed outliers were not "significant".

5. The data should be approximately normally distributed.

This assumption appears to create an issue because the distribution of the travel time data examined in this investigation was observed to be non-normal in nature in Chapter 3. Guiard and Rasch (2004) address this issue by recognizing that the t-test is robust against non-normality of the data, especially when sample sizes are large. As observed in Chapter 3, the number of observations is on the order of hundreds to thousands of observations for the link-level field data, which shows that sample size is not an issue. Furthermore, in this investigation it is the means of the Before and After data sets that are being compared, not the data sets themselves. Since it is known according to the central limit theorem that when randomly sampled a large number of times the mean of a population is normally distributed, the t-test can then be used to compare the Before and After means and standard deviations of the travel time data.

6. The variances of the two data sets must be equal. If they are not equal, Welch's t-test, which assumes inequality of variances, must be used instead.

Equality of variance can be confirmed or refuted using the f-test and Welch's t-test can be used in cases where the f-test indicates that there is evidence that the standard deviations (i.e. variances) are not equal.

According to Peng (2009), the following assumptions apply to the use of the f-test of equality of variances:

1. Observations are randomly drawn from their respective populations.

   Since the travel times collected via Bluetooth technology were not explicitly drawn from the overall population of travel times according to a defined sampling strategy, it is assumed that Bluetooth travel times represent a random sample from the population of vehicles.

2. The data should be approximately normally distributed.

This assumption is described by Peng (2009) as robust, especially when the travel time distributions have similar shapes and are equal in sample size. As observed in Chapter 3, the distributions of the Before and After data sets possess similar shapes with a long tail to the right and have sample sizes based on data collection periods of the same duration. These facts ensure the practitioner's confidence in using this test.

Based on these tests on the mean and standard deviation, it was hypothesized that the small sample sizes of the path-level data would result in many cases in which there was insufficient evidence available to statistically differentiate between the Before and After means and standard deviations. It was also hypothesized that the larger sample sizes of the link-level data would result in more cases in which there was sufficient evidence available to statistically differentiate between the Before and After means and standard deviations.

As seen in most travel time studies, a level of confidence of 95% was utilized throughout this investigation.

4.3 Relationship between Path-level and Link-level Field Data

Knowing that each path is made up of a series of links, it was hypothesized that each path-level travel time could be matched to a set of link-level travel times that, when summed, closely resembled the original path-level travel time. Figure 4.1 illustrates the expectation that a vehicle known to traverse the entire corridor will have a sum of its link-level travel times that is equal to its path-level travel time.

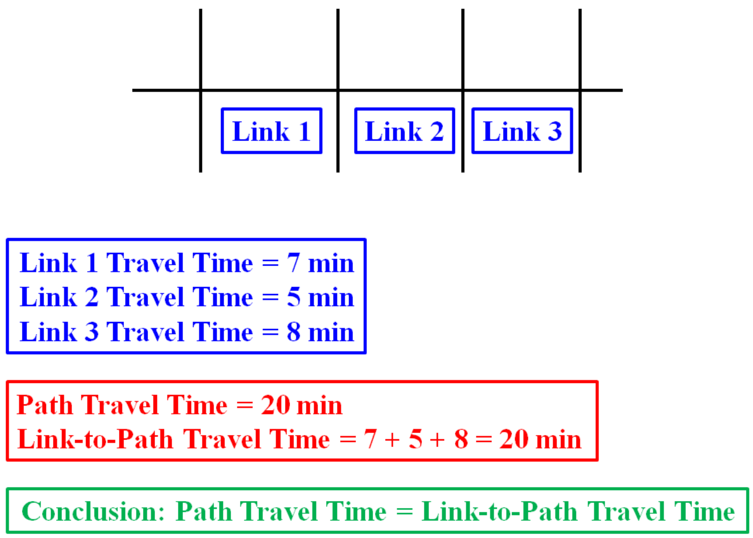

Figure 4.1: Expected relationship between path-level and link-level travel times

This hypothesis was tested on the unfiltered field Bluetooth data by collecting all of the link-level travel times that had the same MAC ID as the corresponding path-level travel time and was recorded during the time in which the time stamps of the path-level travel time were recorded. This was done separately for both the Before data and the After data for which there were 65 and 161 path-level travel times respectively that possessed six link-level travel times. A regression analysis was performed for the link-to-path travel times as a function of the path-level travel times. For both the Before and After data sets the slope of the regression line was found to be statistically significant (i.e. a rejection of the null hypothesis that the slope was zero) with a magnitude of 0.9805 for the Before data and 1.0000 for the After data. Additionally, insufficient evidence was found to support the statistical significance of the y-intercept for both the before and After data. These results confirm the expectation that the path-level travel times are approximately equal to the summation of the link-level travel times.

Since the path-level simulated Bluetooth travel times are generated using link-level travel times, the relationship between data has already been defined and does not have to be investigated.

4.4 Statistical Analysis of the Field Data

For a given from-to detector combination, direction and time period, the path and link mean travel times were calculated using the arithmetic average, which is shown in Equation 4.1 and Equation 4.2:

$$\bar{T} = \frac{1}{w} * \sum_{v=1}^{w} T_v \qquad (4.1)$$

$$\bar{t} = \frac{1}{w} * \sum_{v=1}^{w} t_v \qquad (4.2)$$

where $\bar{T}$ represents the mean travel time for a given path, T_v represents the travel time of vehicle v belonging to the set of all w vehicles that traversed the given path, $\bar{t}$ represents the mean travel time for a given link, and t_v represents the travel time of vehicle v belonging to the set of all w vehicles that traversed the given link.

The path and link standard deviations were also calculated using the standard calculation method, which is shown in Equations 4.3 and 4.4:

$$S = \sqrt{\frac{1}{w-1} * \sum_{v=1}^{w} (T_v - \bar{T})^2} \qquad (4.3)$$

$$s = \sqrt{\frac{1}{w-1} * \sum_{v=1}^{w} (t_v - \bar{t})^2} \qquad (4.4)$$

where S represents the standard deviation for a given path and s represents the standard deviation for a given link.

Knowing the mean, standard deviation and sample size for the Before and After travel time data, parametric tests were performed on the mean and standard deviation to determine whether or not they were statistically different between the Before and After periods. First, the two-sample f-test was used to determine whether or not the standard deviations were statistically different. For this test, the null hypothesis (H_0) was that the Before and After standard deviations were equal and the alternative hypothesis (H_1) was that the Before and After standard deviations were not equal. After this, the two-sample t-test was used to determine whether or not the means were statistically different. For this test, the null hypothesis (H_0) was that the Before and After means were equal and the alternative hypothesis (H_1) was that the Before and After means were not equal. It must be noted that the f-test was performed before the t-test to first determine whether or not the standard deviations were different. If they were found to be different, then a two-sample t-test assuming inequality of variances was performed. If they were not found to be different, a two-sample t-test assuming equality of variances was performed.

Table 4.1 shows the statistical analysis results for the path-level field Bluetooth data. This table contains information about the direction of traffic movement, the time period for which data were collected, the sample size, mean and standard deviation of the data, and the differences between the Before and After means and standard deviations of the travel time data. The sample sizes are on the order of tens of observations, an indicator of the small number of travel times available at this scale.

In this table, a "Y" indicates that the Before and After means or standard deviations were found to be statistically different (i.e. a rejection of the null hypothesis) and an "N" indicates that there was not enough evidence to conclude that the Before and After means or standard deviations were statistically different (i.e. a failure to reject the null hypothesis). From this table, it can be observed that the t-test results are unanimous; the mean travel times of the Before and After data are not statistically different and the improvements made thorough signal retiming cannot be statistically confirmed using these data. The results of the f-test indicate that for two out of the four data sets tested there was a statistically significant change in the standard deviation of travel times.

In summary, there are only four opportunities at the path level to test the difference in the means or standard deviations. Being unable to find statistical evidence to support such a difference for even just one test makes it difficult to provide statistical conclusions using path-level travel time data.

Table 4.1: Statistical test results for path-level field Bluetooth travel time data

Direction	Time Period	Sample Size		Mean Travel Time (minutes)		Before Mean - After Mean (minutes)	Stdev of Travel Time (minutes)		Before Stdev - After Stdev (minutes)	Stdevs Different?	Means Different?
		Before	After	Before	After		Before	After			
NB	Morning	16	48	33.734	27.870	5.865	13.085	5.649	7.436	Y	N
	Evening	19	32	40.052	34.881	5.171	10.166	11.825	-1.659	N	N
SB	Morning	11	29	40.483	29.889	10.593	18.390	4.746	13.643	Y	N
	Evening	10	35	38.712	36.949	1.763	8.704	7.175	1.528	N	N

Appendix C contains the statistical analysis results for the link-level field Bluetooth data. For these data, a total of 24 scenarios existed based on six links, two directions and two time periods. Table 4.2 compares the statistical test results of the path-level and link-level field Bluetooth data for each combination of direction and time period. Each cell in the table indicates the number of comparisons which were statistically different divided by the total number of comparisons. For all path-level results, there is only one comparison made so the results are either 0% (not statistically different) or 100% (statistically different). The corridor is divided into six links, meaning that for the link-level results there are six comparisons made.

It is clear from this table that the greater number of comparisons that can be performed at the link level allows for more opportunities to identify statistically significant differences in the mean and standard deviation. This is especially evident for the results associated with the mean travel time. The path-level analysis indicates that the signal retiming has not had a significant impact for any direction nor any time period. However, the link-level analysis indicates that for at least two thirds of the links, the signal retiming has had a statistically significant impact on the mean travel times. It should be noted that the impact of the signal retiming at the link-level is not positive (i.e. the signal retiming does not always result in a reduction in the mean travel time) for all combinations of link, direction and time period. Consequently, signal retiming can result in statistically significant savings along one link of the corridor but statistically significant losses along another link of the same corridor. When this occurs, the benefits at the path level are reduced. In the case of the field Bluetooth travel time data, the path-level benefits are reduced to the extent that there is insufficient evidence to support their statistical significance.

Relating back to the original question about whether or not the differences in the mean and standard deviation are statistically significant, these results show that link-level analysis provides two advantages over path-level analysis, namely: (1) the larger number of observations at the link level provides more reliable results regarding the change in the population mean and standard deviation; and (2) the increased spatial resolution at the link level permits allows for conclusions to be made for sections of the corridor instead of for the entire corridor, allowing for a higher level of detail in any conclusions that are made.

Table 4.2: Fraction of path-level and link-level statistical tests showing significant differences (field data)

Direction	Time Period	Differences in Standard Deviations of Travel Times		Differences in Mean Travel Times	
		Path Level	Link Level	Path Level	Link Level
NB	Morning	1/1 = 100%	3/6 = 50%	0/1 = 0%	5/6 = 83%
	Evening	0/1 = 0%	5/6 = 83%	0/1 = 0%	6/6 = 100%
SB	Morning	1/1 = 100%	4/6 = 66%	0/1 = 0%	4/6 = 66%
	Evening	0/1 = 0%	6/6 = 100%	0/1 = 0%	4/6 = 66%

4.5 Statistical Analysis of Simulated Data

Table 4.3 and Table 4.4 show the results of the data processing described earlier for the path-level and link-level simulated data respectively and the statistical test results for each. From these tables, the following observations can be made:

1. The path-level data indicate that statistically significant reductions in the standard deviation of travel times have been achieved through the signal improvements in the southbound direction for the evening peak period. The link-level data confirm the path-level results for the southbound direction.
2. The path-level data indicate a statistically significant reduction in the mean travel time in the southbound direction for the evening peak period. The link-level data confirm this finding and also indicate statistically significant reductions for the Bishop- Eagle link in the northbound direction for the morning peak period.

Similar to the field data results, the results from the simulated data show that the link-level data provide larger sample sizes and a finer spatial resolution. Based on these observations, it is recommended that the corridor of interest for a signal retiming study be subdivided into links through the placement of multiple Bluetooth detectors instead of only placing a detector at each end of the corridor of interest.

Table 4.3: Statistical test results for path-level simulated Bluetooth travel time data

Direction	Time Period	Sample Size		Mean Travel Time (minutes)		Before Mean - After Mean (minutes)	Stdev of Travel Time (minutes)		Before Stdev - After Stdev (minutes)	Stdevs Different?	Means Different?
		Before	After	Before	After		Before	After			
NB	Morning	47	58	6.722	6.513	0.209	0.883	0.856	0.026	N	N
	Evening	37	29	7.321	7.057	0.264	0.928	0.751	0.177	N	N
SB	Morning	41	42	6.740	6.472	0.268	0.604	0.708	-0.104	N	N
	Evening	58	66	14.207	10.235	3.972	5.966	2.947	3.019	Y	Y

Table 4.4: Statistical test results for link-level simulated Bluetooth travel time data

From	To	Direction	Time Period	Sample Size		Mean Travel Time (minutes)		Before Mean - After Mean (minutes)	Stdev of Travel Time (minutes)		Before Stdev - After Stdev (minutes)	Stdevs Different?	Means Different?
				Before	After	Before	After		Before	After			
Todd	Bishop	NB	Morning	136	119	3.795	3.683	0.111	0.576	0.533	0.043	N	N
Bishop	Eagle			108	127	2.881	2.689	0.192	0.615	0.661	-0.045	N	Y
Todd	Bishop		Evening	83	84	3.835	3.724	0.111	0.545	0.574	-0.029	N	N
Bishop	Eagle			105	90	3.430	3.243	0.187	0.770	0.717	0.053	N	N
Eagle	Bishop	SB	Morning	77	76	3.685	3.673	0.012	0.720	0.665	0.055	N	N
Bishop	Todd			133	122	2.830	2.734	0.096	0.534	0.458	0.075	N	N
Eagle	Bishop		Evening	157	152	9.107	6.792	2.315	4.718	2.953	1.765	Y	Y
Bishop	Todd			195	184	4.188	3.123	1.065	1.705	0.682	1.023	Y	Y

It is important to note that prior knowledge of the traffic patterns along the corridor is important when assessing the utility of path-level data relative to link-level data in terms of sample size. In the case of the field corridor (Victoria Park Avenue), it was known that vehicles very rarely traverse the entire corridor relative to smaller segments of the corridor. This was confirmed by the small number of travel times collected at the path level relative to the link level. Consequently, it is expected that the number of path-level travel times that are able to be collected will be higher for a corridor for which it is known that vehicles more frequently traverse the entire corridor during the peak commuting periods. Therefore, it is recommended that the path-level and link-level data should be examined for other corridors for which it is known that vehicles generally utilize the entirety of the corridor during the peak commuting periods.

4.6 Summary

In this chapter, the relationship between path-level and link-level Bluetooth travel time data and the adequacy of using path-level data for the purposes of signal retiming evaluation instead of link-level data were examined. The following findings were concluded from this chapter:

- There is a distinct relationship between path-level and link-level travel times that share the same MAC ID. Provided that a MAC ID's link-level travel times exist within the time stamps of the corresponding path-level travel time, the sum of link-level travel times was found to be approximately equal to the path-level travel time for that MAC ID for the field Bluetooth travel time data.
- The number of travel time observations at the path-level is much smaller than at the link-level. The smaller sample size reduces the accuracy by which the population mean and standard deviation can be estimated and therefore inhibits the ability of a practitioner to make conclusions on the impact of signal retiming measures.
- At the link level, statistical differences in the means of the Bluetooth travel time data are far more prevalent than at the path level. Furthermore, a greater amount of statistical differences were found due to a greater number of links that possess travel time data. This was found to be true for both the field and simulated Bluetooth data. Due to the larger number of

observations collected at this scale, the link-level data are more suitable for assessing the impacts of signal retiming measures.

5 Detector Location, Data Collection Period and Detector Spacing

5.1 Introduction

This chapter describes the development of methods for identifying the appropriate values of several key parameters associated with the evaluation of signal retiming measures using Bluetooth travel time data. First, the adequate location of Bluetooth detectors is investigated. Next, the characteristics of the data collection period are examined. Next, the duration of the data collection period is examined with the intention of allowing a practitioner to estimate the duration of the collection period before data collection takes place. Lastly, the relationship between detector spacing and the duration of the data collection period is examined. This chapter concludes with a summary of the major observations made regarding each parameter.

It must be noted that the analyses described in this chapter represent the fitting of models to empirical data and therefore only the field Bluetooth detector data were used to calibrate the models; the simulated Bluetooth data were not used.

5.2 Detector Location

The literature review revealed that Sherali et al. (2006) recommended that detectors must be placed where the geometry of the roadway changes, a useful measure to take to eliminate a possible source of travel time variability. Therefore, if possible, detectors must be placed at locations where the physical state of the roadway changes so that the state of the roadway between the detectors (and the resulting travel times) is relatively consistent.

Since Bluetooth detectors require a power source, detector placement at an intersection for connection to the signal cabinet is more cost-effective than placing detectors at midblock locations and having to use a portable power source. As discussed by Day et al. (2012), this measure is cost-effective but neglects the effect that vehicle interactions and traffic signals have on the resulting travel times. Though it is recommended that detectors be placed at midblock locations, this may not be practical given the additional cost and time required to secure the required permission to mount detectors at mid-block locations and either connect to existing power sources or install portable power supplies. It is frequently much easier and less costly to install Bluetooth equipment at signalized intersections. Installing detectors at minor street signalized intersections was the approach adopted for the field site examined in this study and portable batteries were used to provide power rather than connecting to city infrastructure. Bluetooth detector placement at minor signalized intersections was also applied by Moghaddam (2014).

Based on this experience, detector location depends upon the ease with which permission can be obtained and installation can be completed. It is desirable to place detectors at mid-block locations; however, when this is not feasible, then placing the detectors at minor signalized intersections is acceptable.

5.3 Characteristics of the Data Collection Period

The literature review presented in Chapter 2 revealed that the duration over which GPS probe data are to be collected is an established value (Wang, 2014). For Bluetooth travel time data, there is no rationale provided by studies that use this type of data to evaluate signal retiming measures. As the only comparable study in this respect, Quayle et al. (2010) used a period of ten days (weekends included) during the months of February and March to represent each of the Before and After periods. In comparison, the field data examined in this investigation employed a duration of 14 days (weekends included) for both the Before and After periods during the months of October, November and December. This time period was selected based on the battery life of the portable power source associated with each Bluetooth detector. This parameter is investigated further in the next section.

As discussed previously, it is desirable for a practitioner to control for as many external sources of travel time variability as possible so that the effects of the implemented signal retiming measures are made apparent by the resulting reduction in travel time durations between the Before and After periods. While achieving a closed system in an arterial environment is impossible, the careful selection of certain characteristics of the data collection period allows a practitioner to control for at least some sources of travel time variability. Some sources that are difficult to control for in a signal retiming study include incidents such as vehicle collisions or unscheduled road maintenance not only on the arterial of interest but also on adjacent facilities. To ensure that travel time variability is controlled for so that any observed benefits can be attributed to signal retiming and not to other factors, a practitioner must abide by as many of the following conditions as possible when selecting the characteristics of the data collection period:

- Data collection must occur during a time in which no scheduled maintenance activity is to take place on the arterial of interest or (if possible) on surrounding facilities.
- Data collection must occur during a time in which no large community events are scheduled to take place that could be expected to significantly affect traffic operations along the study corridor.
- Data collection should occur during a time in which severe inclement weather conditions (e.g. snow storms) are unlikely to occur.
- These above conditions come with the caveat that the primary goal of implementing signal retiming measures is to improve traffic progression during the peak periods. Therefore, minimizing the influence of these sources of travel time variability at least during the morning and evening peak periods will suffice.
- To ensure consistency in the traffic conditions experienced across the entire data collection period, the period of time between the Before and After periods should be minimized. For reference, the Before and After data collection periods for the field data used in this study were approximately one month apart from each other. Unless the signal retiming measures take a month to implement, this gap should be made much smaller to ensure that changing traffic patterns as well as weather conditions do not have the opportunity to contribute to the variability of travel times along the corridor and influence the differences between the Before and After data sets.

Due to the desire to improve traffic operations during periods in which volumes are greatest, the duration of the peak morning and evening periods must be defined. Since Bluetooth detectors are able to continuously collect data, they can simply be left out beyond these peak periods and the data can then be filtered according to the peak periods. Depending on the traffic patterns of the urban area of interest, the time and duration of the peak periods can vary. The peak periods identified for the field site used in this study were a 7:00am-10:00am morning peak and a 3:30pm-6:30pm evening peak. For the simulated data, a 7:30am-10:30am morning peak and a

3:00pm-6:00pm evening peak were defined. These times and peak period durations are likely typical for many jurisdictions.

5.4 Duration of the Data Collection Period

This section presents an investigation of the number of travel time observations required, the expected number of travel time measurements, and the duration of the data collection period required to perform a before-and-after signal retiming evaluation. Since current methods are impractical due to data collection having to occur before the number of travel times required is determined, the intention is to estimate the duration of the data collection period before data collection occurs. More specifically, methods to estimate the number of travel times required and the number of travel times collected are developed so that the duration of the data collection period can then be estimated.

5.4.1 Methodology

In this part of the investigation, models were developed to allow a practitioner to estimate:

1. The number of travel time observations that should be collected during the Before and After periods in order to be able to make statistically significant conclusions about the impact that the signal retiming study has had on travel times (i.e. the number of observations required);
2. The number of travel time observations that can be expected to be collected during the Before and After periods (i.e. the number of observations collected); and
3. The length of time over which the Bluetooth detectors should be deployed in order to collect the number of observations required (i.e. the duration of the data collection period in units of weekdays).

Regression analysis was used to develop each model because this approach permits the identification of the statistically significant independent variables that influence the dependent variable. The conceptual structure of each model is first described and the results of each regression analysis are then presented. Each model was developed for application at the link level, meaning that all independent and dependent variables apply to a given link of the field data collection site and for a given direction (northbound or southbound) and time period (morning peak period or evening peak period).

The statistical significance of each regression coefficient was assessed at a 95% level of significance (i.e. the p-value of a variable of interest must be less than 0.05 to be statistically significant). The accuracy of each model was evaluated based on the comparison between the estimated value and the actual value through regression. For example, the estimated number of observations collected that was produced using the developed regression model was compared to the actual number of observations collected.

5.4.2 Potential Independent Variables

In order to develop regression equations, potential independent variables must be identified. Following and expanding on the work by Turner et al. (1998), the following characteristics of the study site were considered as potential independent variables:

- minimum number of through lanes along the link for a given direction (L);
- link length (kilometres) ($length$);
- 24-hour AADT (average annual daily traffic measured in vehicles) ($AADT$);
- peak period AADT (vehicles) ($peak\_AADT$);
- peak period AADT per minimum number of lanes (vehicles/lane) ($peak\ AADT/L$);
- number of signalized intersections (num_sig);
- number of signalized intersections per kilometre (sig/km);
- number of unsignalized intersections (num_unsig);
- number of unsignalized intersections per kilometre ($unsig/km$);
- number of intersections (num_int);
- number of intersections per kilometre (int/km); and
- posted speed limit (kilometres per hour) ($speed$).

The potential independent variables were compiled for each combination of link, direction and time period for the field study corridor. AADT data were collected from the City of Toronto's website (City of Toronto, 2015). These data were collected in 2009 and 2013 and are available in Appendix D along with all of the other data used to develop the regression models presented in this chapter. The AADT data required aggregation because the links within the AADT maps did not match up with the links defined by the placement of Bluetooth detectors in the field. To rectify this, the AADT for all links that overlapped (in whole or in part) the link defined by the Bluetooth detectors were collected and averaged. For example, if the AADT map had four links that overlapped the link of interest in full or in part, then these four corresponding AADT values were averaged to produce the AADT of the link defined by Bluetooth detectors. AADT data were available in the form of 24-hour counts, AM peak period and PM peak period counts.

The length of each link and the number of signalized and unsignalized intersections associated with each link were estimated using Google Earth (Google, TerraMetrics, National Oceanic and Atmospheric Administration, & DigitalGlobe, 2015). It must be noted that each link length was estimated to one decimal place. The signals of the intersections defining each link were included in these counts but pedestrian crossings (of which there were two along the entire corridor) were not classified as signals and were therefore not counted.

With six links along the field data corridor, two directions and two time periods, each regression model had a total of 24 data points available.

It was expected that AADT, link length and the number of intersections (signalized, unsignalized, or both) would be significant estimators for these models. AADT was expected to be significant for the model to estimate the number of observations collected because as the vehicle volume increases the number of Bluetooth-enabled vehicles was expected to increase. Link length was expected to be a major contributor because link length represents a source of variability. In other words, a longer link presents more opportunities for a vehicle to deviate from the expected travel time along that link. The same is also true for the number of intersections; more intersections allows for more access and egress points to and from a link, thus decreasing the number of Bluetooth-enabled vehicles that are likely to traverse the entire link. This makes the number of intersections a potential variable in the model to estimate the number of observations collected. More intersections also results in more potential impedances to a vehicle's progression, making it a potential variable in the model to estimate the number of observations required.

The sequential elimination of variables for which insufficient evidence was found to support their statistical significance was used to arrive at the final regression model for each variable that was to be estimated. Emphasis was placed upon retaining link length as an independent variable for both the model to estimate the number of observations required and the model to estimate the number of observations collected. This was desired so that the relationship between detector spacing (i.e. link length) and the duration of the data collection period could be examined after these models were developed.

5.4.3 Model to Estimate the Number of Observations Required

The z-test or Student's t-test can be used to determine whether or not the difference between the Before and After mean travel times is statistically significant. Equation 5.1 shows the method to estimate the calculated t-statistic:

$$t_{calc} = \frac{\bar{x}_B - \bar{x}_A}{\sqrt{\frac{s_B^2}{n_B} + \frac{s_A^2}{n_A}}} \tag{5.1}$$

where t_{calc} is the calculated t-statistic, s_B and s_A are the standard deviations of the Before and After data respectively (in units of minutes), n_B and n_A are the number of observations required for the Before and After travel time data respectively, and $\bar{x}_B$ and $\bar{x}_A$ are the means of the Before and After travel time data respectively (in units of minutes).

It is necessary to know the Before and After mean travel time, the Before and After standard deviation of travel time and the number of travel time observations in the Before and After periods in order to use Equation 5.1. Naturally, when attempting to estimate the number of travel time observations that will be required before data collection occurs, no values are available for any of these variables. Even if the equation is rearranged to solve for n_B or n_A and assuming the rest of the terms are known, when the difference in the mean travel times is very small the equation will estimate that a very (unrealistically) large number of observations is required. Though the estimate is correct (i.e. this is the number of observations required to demonstrate that small of a difference in the mean travel times is in fact statistically significant), the result is not of practical value. When the difference in the means is very small, it is very likely that this difference is not statistically significant (i.e. the signal retiming has not resulted in a significant reduction in the mean travel time) and there is no need to expend resources to collect large numbers of observations to come to the same conclusion.

In order to develop a model that is of practical use for estimating the number of travel time observations required, the following assumptions are made:

1. The number of observations required for the Before and After periods are equal ($n_B = n_A = n$).
2. The standard deviations of the travel times of the Before and After travel times are equal ($s_B = s_A = s$).
3. The t-statistic can be replaced with the z-statistic.

Using these assumptions, Equation 5.1 can be rearranged to solve for n:

$$n = 2 * \left(\frac{s * z_{\alpha/2}}{\bar{x}_B - \bar{x}_A}\right)^2 \tag{5.2}$$

where n is the number of observations required for each of the Before and After periods and $z_{\alpha/2}$ is the critical z-statistic. A value of 1.96 was used for the z-statistic to represent a 95% level of confidence ($\alpha = 0.05$).

Though Equation 5.2 requires fewer inputs than Equation 5.1 it still requires prior knowledge of the mean travel times for both the Before and After periods. To address this issue, a new parameter β can be defined, as shown in Equation 5.3:

$$\beta = \frac{\bar{x}_B - \bar{x}_A}{\bar{x}_B} \tag{5.3}$$

where β is the minimum percent reduction in the mean travel time that a practitioner wants to statistically confirm. For example, if $\beta = 0.10$ then it is desired by the practitioner to estimate the number of travel times needed to be able to statistically confirm at least a 10% reduction in the mean. Consequently, the number of observations required will increase as β decreases. In other words, it requires a greater number of observations to statistically confirm a smaller reduction in the mean.

A sensitivity analysis of the effect that β has on the estimated number of observations required was carried out and the results are provided later in this chapter.

Finally, assuming $\bar{x}_B = \bar{x}$ and substituting Equation 5.3 into Equation 5.2 produces Equation 5.4:

$$n = 2 * \left(\frac{s * z_{\alpha/2}}{\beta * \bar{x}}\right)^2 \qquad (5.4)$$

Having developed an equation to estimate the number of observations required, regression equations can be developed to estimate the mean travel time ($\bar{x}$) and the standard deviation of travel times (s) of the Before data.

Using the sequential elimination of variables that did not have evidence to support their statistical significance, the standard deviation of the Before data was found to be a function of the peak AADT per minimum number of lanes per direction, the number of signals per kilometre and the number of unsignalized intersections per kilometre. Equation 5.5 was the resulting equation:

$$s_{est} = c_1 + c_2 * peak\_AADT/L + c_3 * sig/km + c_4 * unsig/km \qquad (5.5)$$

where s_{est} is the estimated standard deviation of the Before data (in units of minutes), $peak\_AADT/L$ is the AADT per minimum number of lanes per direction (in units of vehicles per minimum number of lanes for a given direction), sig/km is the number of signals per kilometre, $unsig/km$ is the number of unsignalized intersections per kilometre, c_1 is the y-intercept (-3.870), and c_2, c_3 and c_4 are the coefficients of the respective independent variables (0.002, 1.082 and 0.404 respectively). A drawback to this model is that the y-intercept (c_1) coefficient was found to be statistically significant, which is difficult to explain from a real-world perspective. However, as seen in Figure 5.1, a relatively high R-squared value is produced by this model, which supports the notion that the explanatory power of this model is high.

Figure 5.1 compares the actual standard deviations to the standard deviations estimated using Equation 5.5 for the Before data. For this figure and all other regression figures, the dashed line is simply a $y = x$ line and the solid line is the trendline. Performing a linear regression on the actual and estimated standard deviations indicates that there is a strong linear correlation (i.e. a relatively high R-squared value). However, there is bias present in the model in that the slope coefficient is not equal to 1 and the intercept is equal to zero.

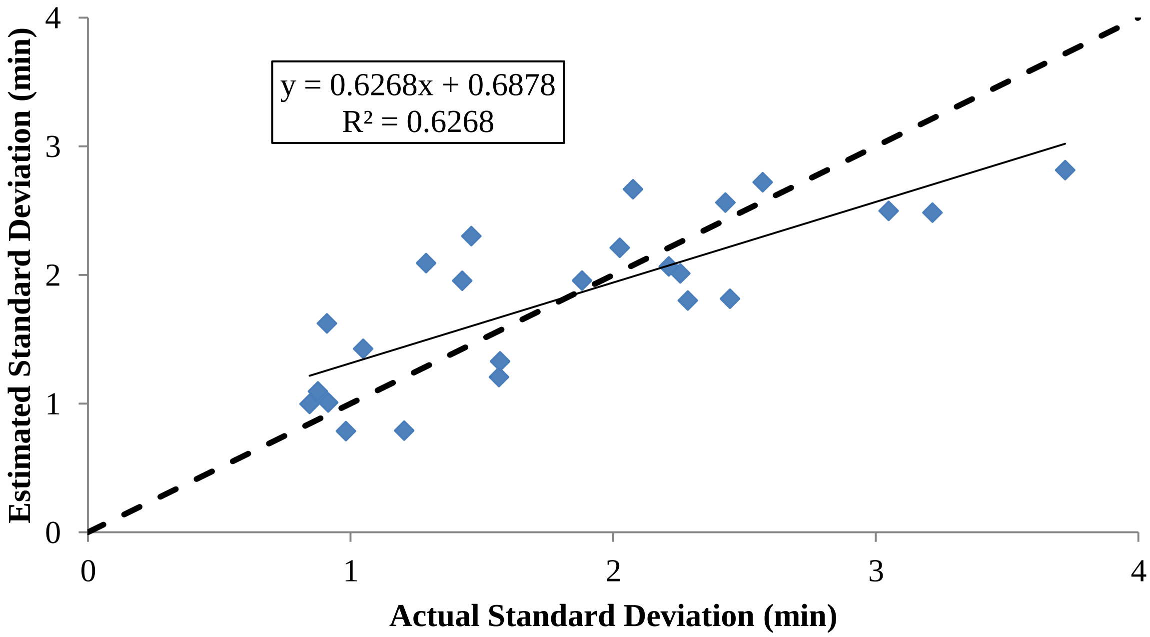

Figure 5.1: Comparison of estimated and actual standard deviations of travel times for the Before data

A similar approach was adopted to develop an expression to estimate the mean Before period travel time. The mean of the Before data was found to be a function of length and number of signals per kilometre, as shown in Equation 5.6:

$$\bar{x}_{est} = c_5 + c_6 * length + c_7 * sig/km \qquad (5.6)$$

where $\bar{x}_{est}$ is the estimated mean of the Before or After data (in units of minutes), $length$ is the link length (in units of kilometres), c_5 is the y-intercept (-6.228), and c_6 and c_7 are the slope coefficients of the respective independent variables (2.361 and 2.031 respectively). A drawback to this model is that the y-intercept was found to be statistically significant, which is difficult to explain from a real-world perspective as the mean travel time is expected to approach zero as the length of the link approaches zero. This observation reiterates the importance of only the applying the regression model to links for which the link characteristics (i.e. the independent variables) fall within the range of characteristics contained within the model calibration data set. As seen in Figure 5.2, this model also provides a relatively high R-squared value.

Figure 5.2 compares the actual means to the means estimated using Equation 5.6 for the Before data. Performing a linear regression on the actual and estimated means indicates that there is a strong linear correlation. However, there is bias present in the model in that the slope coefficient is not equal to 1 and the intercept is not equal to zero.

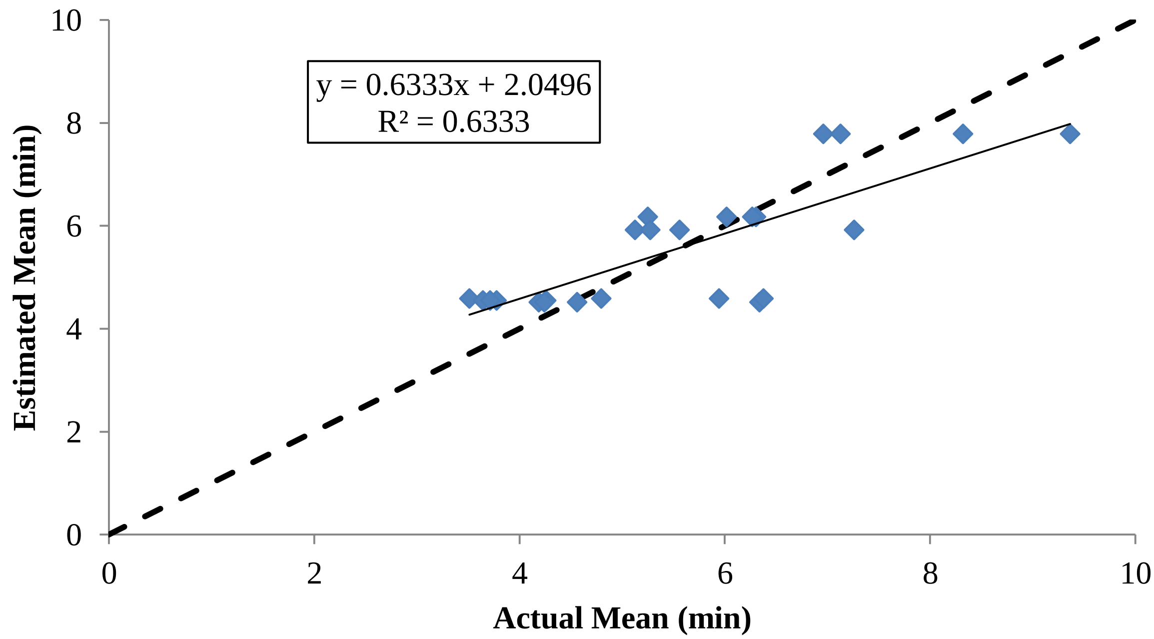

Figure 5.2: Comparison of estimated and actual means for the Before data

Placing Equation 5.5 and Equation 5.6 into Equation 5.4 produces Equation 5.7, the final model to estimate the number of observations required:

$$n_{est} = 2 * \left(\frac{(c_1 + c_2 * peak\ AADT/L + c_3 * sig/km + c_4 * unsig/km) * z_{\alpha/2}}{\beta * (c_5 + c_6 * length + c_7 * sig/km)} \right)^2 \qquad (5.7)$$

where n_{est} is the number of observations required for each of the Before and After periods.

Figure 5.3 compares the actual number of observations required (calculated using Equation 5.4 with the known values for the standard deviation and the mean) to the number of observations required estimated using Equation 5.7. Performing a linear regression on the actual and estimated number of observations required indicates that there is a strong linear correlation and that there is little bias present in the model (i.e. the slope coefficient is close to 1 and the intercept is equal to zero).

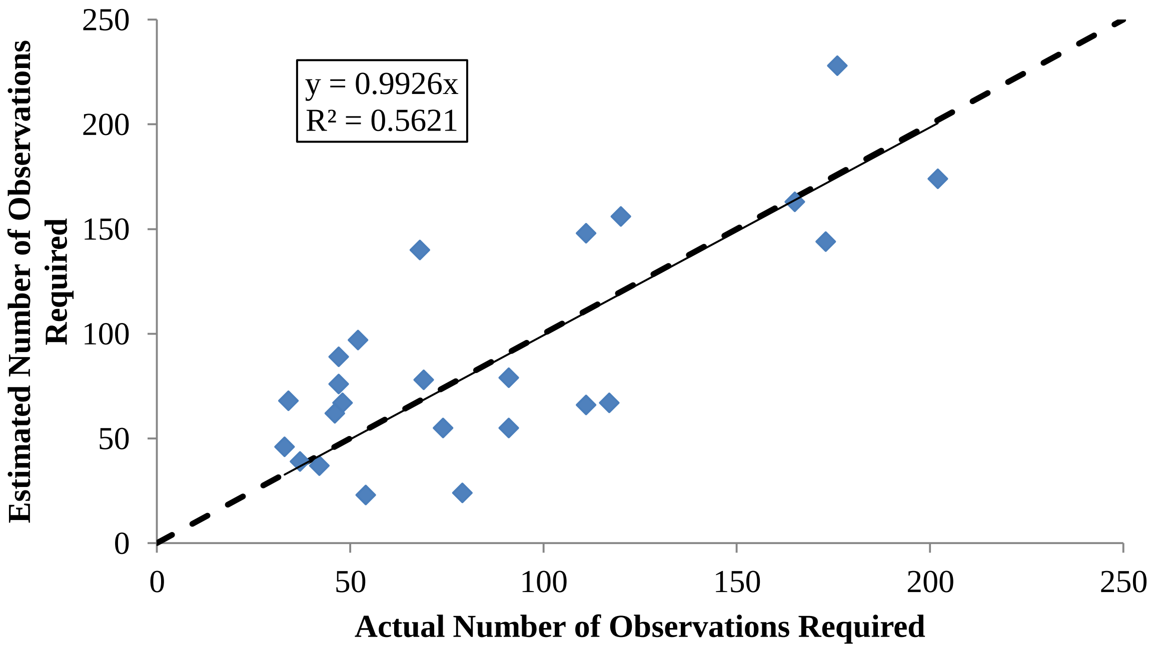

Figure 5.3: Comparison of estimated and actual number of observations required

5.4.4 Sensitivity Analysis of the Model to Estimate the Number of Observations Required

To better understand the nature of the parameter β, a sensitivity analysis of the number of observations required with respect to two variables was examined: the parameter β and the peak AADT. The relationship between the number of observations required and detector spacing (i.e. the length of a link) is examined later in this chapter.

Calculations were performed using Equation 5.3 (the minimum percent reduction in the mean travel time that a practitioner wants to statistically confirm), Equation 5.4 (the number of observations required), Equation 5.5 (the mean of the Before data) and Equation 5.6 (the standard deviation of the Before data). For this sensitivity analysis, the number of signals per kilometre was set to 3, the number of unsignalized intersections per kilometre was set to 3.5, and the minimum number of lanes per direction was set to 2.

Figure 5.4 shows the relationship between the minimum difference in the means ($\bar{x}_B - \bar{x}_A$) and β for three link lengths. As expected a positive linear relationship exists between the minimum difference in the means and β. More importantly, this figure shows that for the assumed link characteristics, selecting $\beta = 0.10$ implies that the analyst wishes to estimate the number of travel time observations required to statistically confirm that a change in the mean travel times is equal to or exceeds 0.5 minutes for a link that is two kilometres long, 0.7 minutes for a link that is three kilometres long, and 0.9 minutes for a link that is four kilometres long.

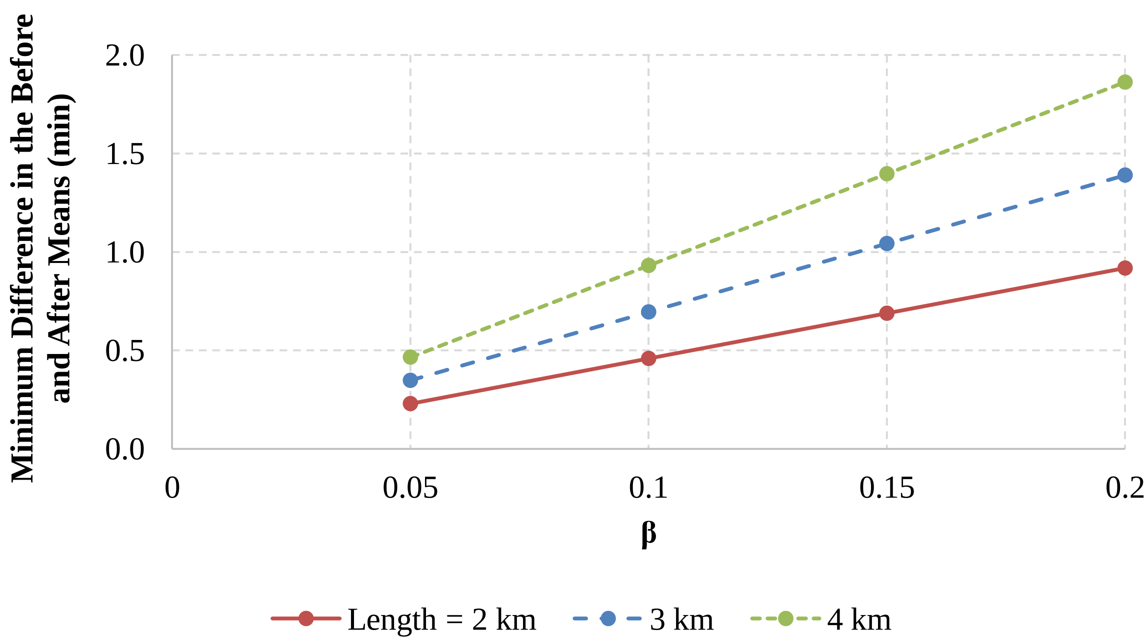

Figure 5.4: Sensitivity analysis of the difference in the means as a function of β for various link lengths

Next, the number of observations required was calculated as a function of β for three link lengths while the peak AADT was set to 1400 vehicles. Figure 5.5 shows the results of these calculations. From this figure it can be seen that an exponential decay relationship exists between the number of observations required and β. As expected, this shows that trying to identify small differences in the mean by setting β to a very small value results in a significant increase in the amount of data needed to statistically identify this difference. This highlights the fact that care must be taken in the selection of the value of β.

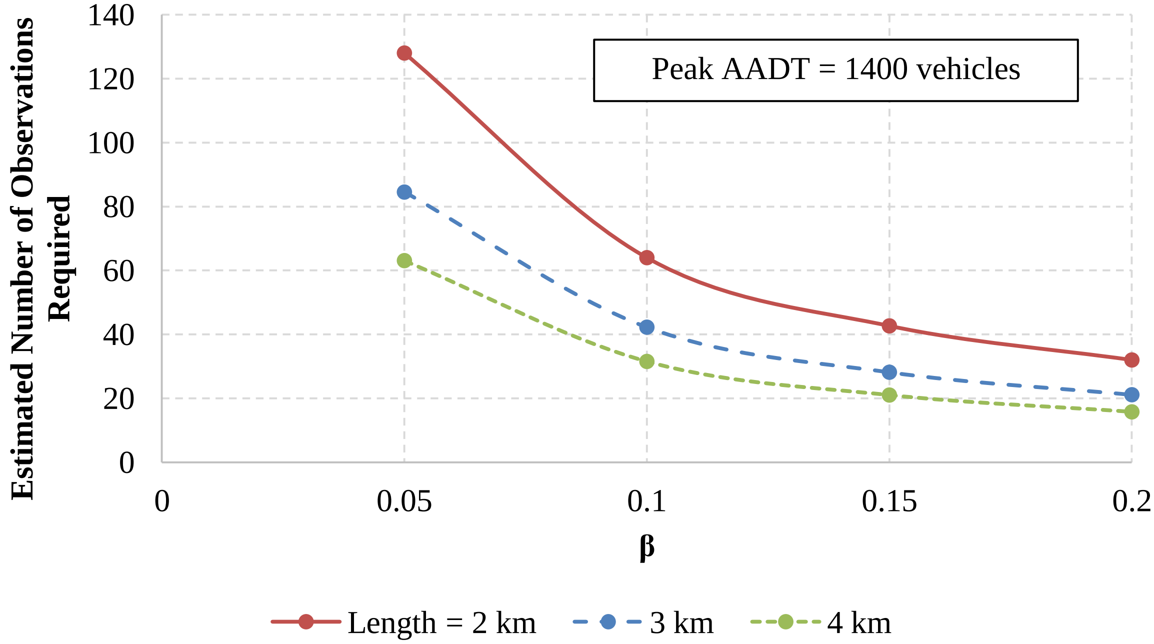

Figure 5.5: Sensitivity analysis of the estimated number of observations required as a function of β for various link lengths

Next, the number of observations required was calculated as a function of the peak AADT for four values of β while the link length was set to three kilometres. Figure 5.6 shows the results of these calculations. From this figure it can be seen that a positive and approximately linear relationship exists between the number of observations required and β. This shows that as the volume on a link increases, the number of observations required also increases. This is expected because as the AADT increases, the variance in travel times increases as a result of a greater frequency of vehicle interactions, which means that the number of observations needed to accurately represent the conditions of the overall population of travel times increases.

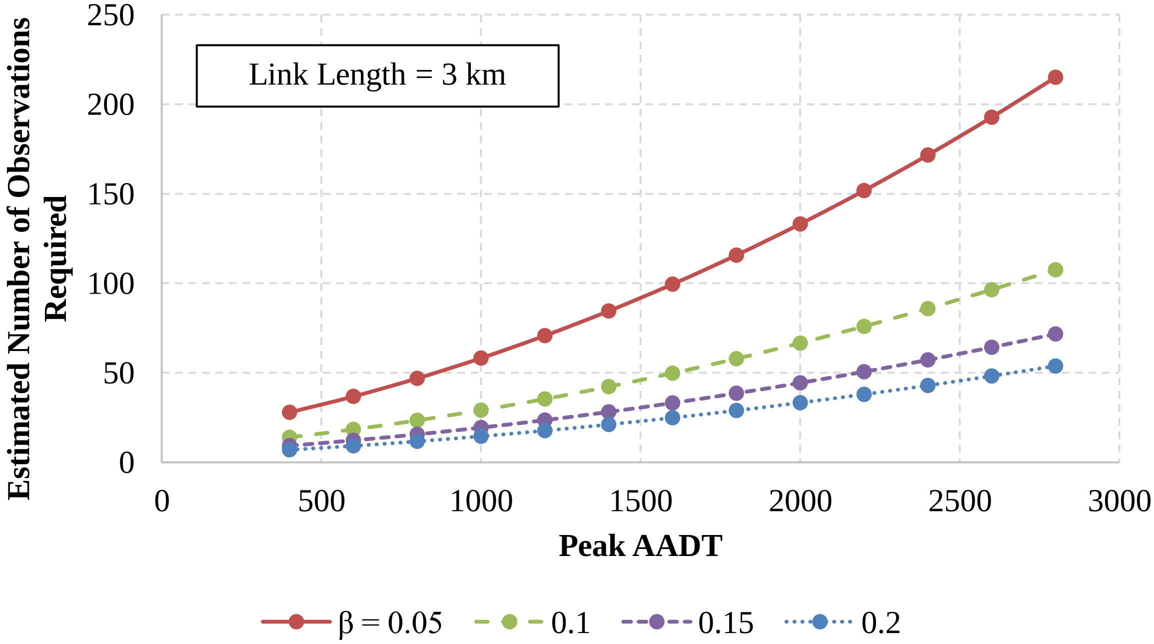

Figure 5.6: Sensitivity analysis of the estimated number of observations required as a function of peak AADT for values of β

5.4.5 Model to Estimate the Number of Observations Collected per Weekday

The number of valid Bluetooth travel time observations collected was calculated using Equation 5.8:

$$n'/weekday = \frac{n_B + n_A}{d} \tag{5.8}$$

where $n'/weekday$ is the number of observations collected per weekday, n_B and n_A are the number of observations collected during each of the Before and After periods respectively, and d is the duration of the data collection period (the sum of the Before and After periods, in units of weekdays). This model was developed on a per weekday basis due to the fact that 20 weekdays of data were collected at the field data collection site. This means that d had a value of 20 weekdays for the field data.

To estimate the number of observations collected, the same set of independent variables considered for the model to estimate the number of observations required was examined. Using the sequential elimination of variables that did not have evidence to support their statistical significance, the number of observations collected was found to be a function of length, the peak AADT and the number of signals per kilometre. Equation 5.9 was the resulting equation:

$$n'_{est}/weekday = c_8 + c_9 * length + c_{10} * peak\ AADT + c_{11} * sig/km \qquad (5.9)$$

where $n'_{est}/weekday$ is the estimated number of observations collected per weekday, $length$ is the link length (in units of kilometres), $peak\ AADT$ is the peak AADT of the period of interest (AM or PM), sig/km is the number of signals per kilometre, c_8 is the y-intercept (425.416), and c_9, c_{10} and c_{11} are the slope coefficients of the respective independent variables (0.118, -72.126 and -86.796 respectively). Figure 5.7 compares the actual number of observations collected per weekday (calculated using Equation 5.8) to the number of observations collected per weekday estimated using Equation 5.9. Performing a linear regression on the actual and estimated number of observations collected per weekday indicates that there is a strong linear correlation and that there is little bias present in the model.

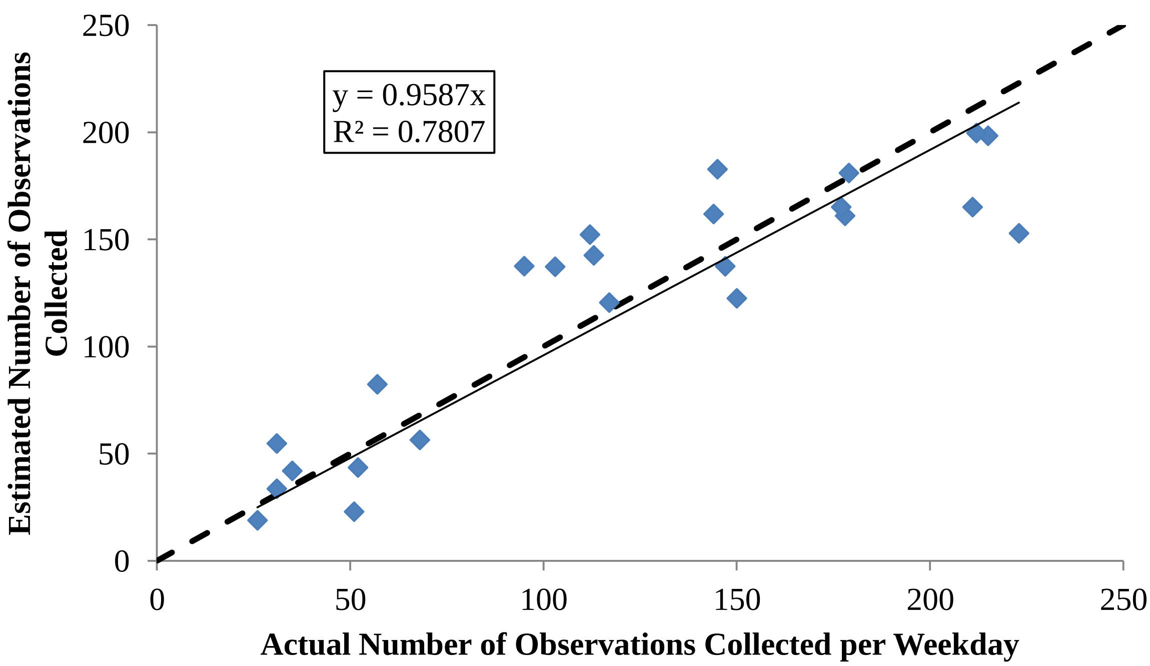

Figure 5.7: Comparison of actual and estimated number of observations collected

5.4.6 Model to Estimate the Duration of the Data Collection Period

Since it is desired that the number of observations collected is greater than or equal to the number of observations required, the resulting relationship shown in Equation 5.10 must hold:

$$n' \geq n \tag{5.10}$$

Since the number of observations collected (n') is being examined on a per weekday basis in this investigation, this can be accounted for by modifying n', as shown in in Equation 5.11:

$$n'/weekday * d \geq n \tag{5.11}$$

This equation can then be rearranged so that the duration of the data collection period (d) can be solved for, as shown in Equation 5.12:

$$d \geq \frac{n}{n'\ /weekday} \tag{5.12}$$

The equation to calculate the number of observations required (Equation 5.2) can then be substituted into Equation 5.12 to produce Equation 5.13:

$$d \geq \frac{2 * \left(\frac{s * z_{\alpha/2}}{\beta * \bar{x}}\right)^2}{n'\ /weekday} \tag{5.13}$$

The equations developed to estimate the number of observations required and the number of observations collected (Equation 5.7 and Equation 5.9 respectively) can then be substituted into Equation 5.13 to produce Equation 5.14:

$$d_{est} \geq \frac{2 * \left(\frac{(c_1 + c_2 * peak\ AADT/L + c_3 * sig/km + c_4 * unsig/km) * z_{\alpha/2}}{\beta * (c_5 + c_6 * length + c_7 * sig/km)}\right)^2}{c_8 + c_9 * length + c_{10} * peak\ AADT + c_{11} * sig/km} \tag{5.14}$$

where d_{est} is the estimated duration of the data collection period. For a given link, direction and time period, the duration of the data collection period can be estimated using the peak period

AADT, the minimum number of lanes per direction, the number of signals per kilometre, the number of unsignalized intersections per kilometre, the length of the link, and β as input.

Figure 5.8 compares the actual duration of the data collection period (calculated using Equation 5.13) to the number of observations collected estimated using Equation 5.14. Performing a linear regression on the actual and estimated durations of the data collection period indicates that there is a strong linear correlation and that there is little bias present in the model.

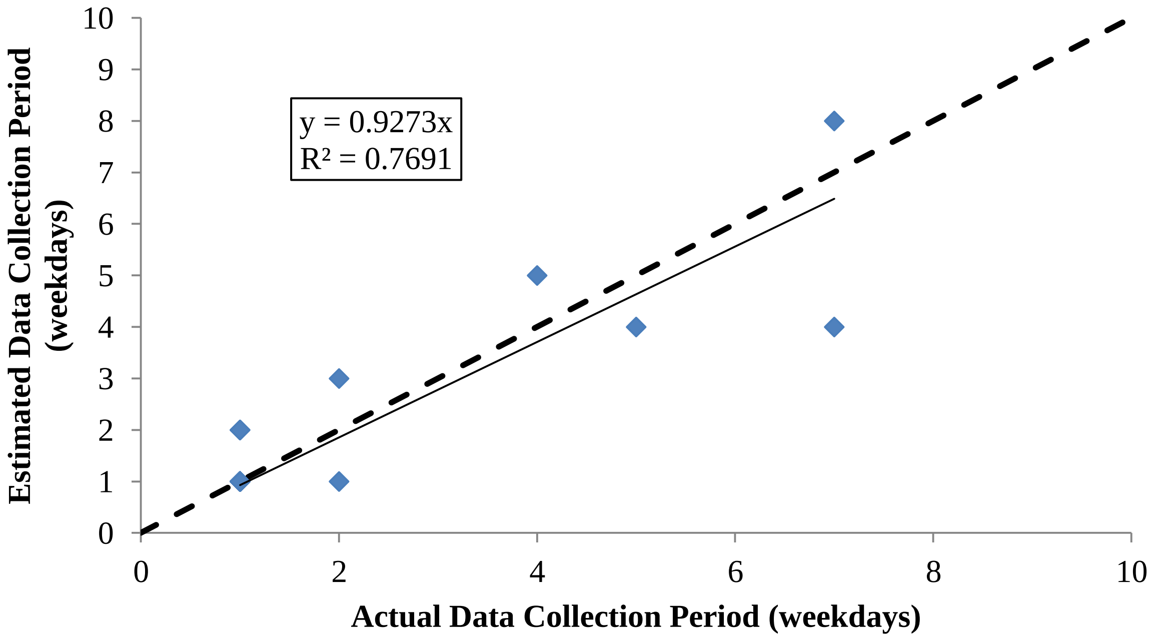

Figure 5.8: Comparison of actual and estimated data collection periods

The maximum estimated duration of the data collection period produced by applying Equation 5.14 to the four scenarios of each link (i.e. every combination of direction and time period) are shown in Table 5.1 for two different values of β for both Bluetooth and GPS technology. In other words, only the maximum of these four values is shown in this table for each link. It can be observed that the estimated duration of the data collection period sharply increases as β decreases. This table shows that for this arterial corridor and for $\beta = 0.10$ Bluetooth travel time data should be collected over a period of at least 16 weekdays (8 weekdays for the Before period and 8 weekdays for the After period). This increases to at least 31 weekdays for $\beta = 0.05$.

According to Wang (2014), it is common practice by the City of Toronto when evaluating signal retiming measures to collect data from five GPS probe runs for each of the morning and evening peak periods for each direction. This table shows the estimated duration of the data collection period based on this number (i.e. $n'_{est}/weekday = 5$). It can be seen that the very low number of observations results in a very high estimation of the data collection period. Comparatively, between 20 and 220 Bluetooth travel times were collected for a given link, direction and time period. This means that from a statistical perspective, it is difficult to justify the use of the floating car probe method to collect travel time data for a signal retiming study, making Bluetooth technology much more attractive in this respect.

Table 5.1: Maximum estimate of the data collection period for each link

Link Number	Link	Maximum Estimate of the Data Collection Period (weekdays)			
		Bluetooth		GPS	
		β = 0.05	β = 0.10	β = 0.05	β = 0.10
1	Kingston-Crescent Town	31	8	139	35
2	Crescent Town-Elvaston	10	3	62	16
3	Elvaston-Rowena	2	1	112	28
4	Rowena-Consumers	3	1	183	46
5	Consumers-Van Horne	5	2	71	18
6	Van Horne-Gordon Baker	2	1	125	32

With respect to recommending a value of β that a practitioner should select, it is difficult to do so with only a single arterial corridor available for testing. Further testing of other corridors is needed to be able to provide recommendations for a value of β.

It must be noted the developed models should only be applied when the characteristics of the links fall within the ranges of the independent variables used to calibrate these models. Since the data used to calibrate these models were obtained from only a single arterial corridor in Toronto, a similar model calibration process should be carried out with data from a larger number of arterial corridors in order to develop a wide range of values for each regression coefficient to increase the range of corridors to which the models can be applied. Only then can the reliability of the developed models truly be assessed.

5.5 Detector Spacing

Table 5.2 shows the length of each link of the corridor. Knowing the maximum estimate of the duration of the data collection period for each link for a given value of β, the lengths of the links that produced relatively low estimates of the data collection period can be used to produce a general recommendation regarding link length (since it is desired to minimize the duration of the data collection period so that operational costs are minimized).

Examining the estimates of the data collection period in Table 5.1 for each link and for two values of β, it can be seen that links 3 through 6 produce relatively low estimates. Since the lengths of these links range between 2.0 kilometres and 2.5 kilometres, this range is the recommended range for which detectors should be spaced to collect Bluetooth data to evaluate signal retiming measures.

It must be repeated that this recommended range should only be applied when the characteristics of the links fall within the ranges of the independent variables used to calibrate the developed regression models.

Table 5.2: Length of each link

Link Number	Link	Length (km)
1	Kingston-Crescent Town	1.8
2	Crescent Town-Elvaston	4.0
3	Elvaston-Rowena	2.4
4	Rowena-Consumers	2.5
5	Consumers-Van Horne	2.0
6	Van Horne-Gordon Baker	2.5

5.6 Summary

In this chapter, the primary parameters associated with a before-and-after signal retiming study associated with Bluetooth technology were examined. The following findings were concluded from this chapter:

- Detectors should be placed at midblock locations to ensure that sources of travel time variability are controlled for. If mid-block deployment is not practical, then detectors should be deployed at minor street intersections instead.
- Using the field data as input, regression models were developed to estimate the number of data observations required, the number of observations collected and the duration of the data collection period. Each of these models used the characteristics of the corridor as input.
- The parameter β was introduced to represent the minimum percent reduction in the mean travel time that a practitioner wants to statistically confirm. A sensitivity analysis was performed to examine the influence that this parameter has on the number of observations required. It was found that trying to identify small differences in the mean by setting β to a very small value results in a significant increase in the amount of data needed. Therefore, care must be taken in the selection of the value of β.
- Using these models and a β value of 0.10, the minimum data collection period was found to be eight weekdays for each of the Before and After periods for the field Bluetooth data. Decreasing β to 0.05 produced a value of 31 weekdays, indicating the importance of selecting an appropriate value of β.
- The duration of the data collection period was also estimated based on the number of travel times that are typically collected using the GPS probe vehicle data collection method. These values were much higher than those of the Bluetooth data collection method, highlighting the vast difference that sample size makes when statistical significance is desired (which the Bluetooth data collection method can provide while the GPS probe vehicle data collection method cannot).
- With respect to recommending a value of β that a practitioner should select, it is difficult to do so with only a single arterial corridor available for testing. Further testing of other corridors is needed to be able to provide recommendations for a value of β.
- Using the estimations of the data collection period for each link, it was recommended that Bluetooth detectors should be spaced apart between 2.0 kilometres and 2.5 kilometres.

6 Measures of Effectiveness

6.1 Introduction

This chapter describes the development of measures of effectiveness for the purposes of evaluating signal retiming measures on an arterial corridor. First, the methodology developed to produce conclusive measures of effectiveness is presented and the rationale associated with each step of the methodology is described in detail. This methodology is then applied to Bluetooth data collected in the field and in a simulated environment. This chapter concludes with a summary of the major observations made regarding the results of applying the newly developed measures of effectiveness to the data.

6.2 Methodology

This section introduces the four components to be included in a measure of effectiveness to be used to evaluate signal retiming measures: the base measure, the link-to-path upgrade, the volume that experienced the base measure, and the statistical test. Detailed rationale for the selection of each component is also provided.

6.2.1 Base Measure Methodology

In this context, the base measure is defined as the measure to which other components are added to produce a more robust measure. Base measures considered include the mean, median, standard deviation, interquartile range and an index of some form.

Despite their near-ubiquitous presence in published signal retiming studies, several key disadvantages exist for using the mean and the standard deviation to measure improvements. As described by Day et al. (2012), the median is a better reflection of the central tendency of the distribution than the mean, especially on longer corridors with greater variability. Day et al. (2012) also recommend the interquartile range over the standard deviation. Both of these recommendations are made because the median and the interquartile range are less susceptible to outliers as compared to the mean and standard deviation.

Other base measures of effectiveness were also explored in the literature in place of the mean and standard deviation including the comparison of the Before and After travel time distributions as well as various indices. A major disadvantage that these measures share with the median and the interquartile range is a lack of use in practice as very few of these measures have been applied in signal retiming studies. Another disadvantage is that there is no definitive measure among them as it was seen in Chapter 2 that a wide variety of measures have been recommended or used in practice. This includes the mean, standard deviation, interquartile range, and various indices.

The goal of this part of the investigation was to produce a measure of effectiveness that will be useful to practitioners who are already familiar with using GPS probe travel time data. This minimizes the opportunity for misinterpretation not only by practitioners but also by the target audience to whom the practitioner is to present the results. Therefore, the transition from the use of conventional GPS probe data to Bluetooth data in signal retiming studies must be facilitated by ensuring a level of continuity for all stakeholders involved.

The issue that Day et al. (2012) raise about the median and interquartile range being less susceptible to outliers than the mean and standard deviation is a valid concern, particularly when the sample size is small. However, link-level analysis tends to provide relatively large sample sizes and the use of outlier filters tends to limit the influence of outliers within the data. Consequently, the mean and standard deviation were retained as base measures.

Two issues that undermine the credibility of the mean and standard deviation are that (1) these measures do not account for the number of vehicles that experience the calculated benefits and (2) there is no statistical basis for the conclusions made using these measures. The first issue can be addressed by incorporating the number of vehicles that experience the calculated benefits into the calculation process. The second issue can be addressed by incorporating statistical testing into the calculation process. The following sections detail the rectification of these issues.

6.2.2 Link-to-Path Methodology

Having found in Chapter 4 that path-level data are not an adequate substitute for link-level data for the purposes of signal retiming evaluation, a method of producing a path-level measure of effectiveness using link-level data was desired. Producing a single path-level measure instead of a series of link-level measures allows a practitioner to provide a single measure of impact for the corridor when justifying the improvements produced by signal retiming. While a variety of link-level measures are desired in the event that links exhibiting poor progression must be identified, the goal of this part of the investigation was not to identify problematic links but to determine the impact of signal retiming measures on the corridor as a whole.

Having demonstrated that (1) data at the link level and at the path level are compatible with each other, (2) link-level data are of greater use than path-level data for signal retiming evaluation, and (3) a desire to produce a concise path-level measure exists, a methodology to facilitate the "upgrading" of link-level measures to the path level must be defined.

The method for upgrading the mean from the link level to the path level was demonstrated by Rakha et al. (2006) and by Li, Chai and Tang (2013) in which links were assumed to be statistically independent despite belonging to the same path. Under this assumption, link-level means are simply summed to produce a link-to-path measure, as shown in Equation 6.1:

$$\bar{t}_{LP} = \sum_{i=1}^{j} \bar{t}_i \qquad (6.1)$$

where $\bar{t}_{LP}$ represents the link-to-path mean travel time for a given path, $\bar{t}_i$ represents the mean travel time of link i belonging to the set of all j links that constitute the given path.

Studies by Rakha et al. (2006) and by Li, Chai and Tang (2013) explored methods of estimating the standard deviation of the path travel time as a function of the standard deviation of link level travel times for freeways. Their methods included a covariance term to account for the propagation of traffic conditions from one link to another, thus implying that the travel time distributions of those links are also correlated. In other words, the inclusion of a covariance term implies that if a vehicle experiences a high travel time at an upstream link then it will also experience a high travel time at the next downstream link. This correlation is difficult to assess in an arterial environment in which the many sources of travel time variability add uncertainty to this correlation compared to the freeway environment examined by Rakha et al. (2006). As discussed by Wakabayashi (2010), it is difficult to know the covariance of traffic on a highway facility, much less an arterial facility. Thus, for signal retiming studies, it is typically impractical to know in advance the value that should be used for the covariance term and therefore it is not possible to estimate the path-level standard deviation on the basis of the link-level measurements. As a result, the mean was carried forward as the sole base measure.

6.2.3 Volume Weighting Methodology

The methodology described in the previous section provides an estimate of the mean travel time savings for vehicles that traverse the entire corridor. However, it is typical that only a small portion of vehicles travel the entire corridor as it was found in Chapter 3 that most vehicles travel only a portion of the corridor. As a result, it is beneficial to account for the different traffic volumes on each link when attempting to assess the impact of the implemented signal retiming measures. Links with a higher number of trips should be weighted more than links with a smaller number of trips. This can be done by multiplying each link-level difference in the Before and After mean travel times by the number of vehicles that experienced this difference and dividing by the number of vehicles on all links of the corridor.

6.2.4 Statistical Test Methodology

As described by Day et al. (2012), statistical testing adds credibility to signal retiming measures that are meant to positively influence travel times. Any developed measure of effectiveness must be able to measure the observed changes in travel times due to signal retiming, not due to randomness; statistical testing makes this possible.

As in Chapter 4, the t-test was selected as the test to be used to indicate whether or not the means were statistically different between the Before and After periods. This test's ease of use and understanding relative to more complex statistical tests is beneficial to practitioners who aim to gain the backing of stakeholders unfamiliar with statistical testing. The resulting measure shows whether or not the travel time savings (or losses) produced for each link are experienced by a large portion of the sampled vehicles and whether or not these savings (or losses) are statistically significant.

6.3 Measures of Effectiveness

In this section, the mathematical basis for each step of the measures of effectiveness being developed is described. The three measures of effectiveness developed build upon each other as more components are added at each step.

6.3.1 MOE$_1$: Average Travel Time Savings per Corridor Trip

The first measure of effectiveness estimates the average impact of the implemented signal retiming measures for a vehicle traversing the entire corridor. It is computed as the sum of the differences between the Before and After mean travel times across all links of the corridor, as shown in Equation 6.2:

$$MOE_1 = \sum_{i=1}^{j} (\bar{t}_{Bi} - \bar{t}_{Ai}) \tag{6.2}$$

where MOE_1 represents the average travel time savings per corridor trip (in units of minutes) and $\bar{t}_{Bi}$ and $\bar{t}_{Ai}$ represent the Before and After mean travel times respectively (in units of minutes) for link i belonging to all j links of the corridor. This measure captures the average change in travel time experienced by a vehicle that traversed the entire corridor during the After period compared to the Before period. A positive value indicates a reduction in the travel time and a negative value indicates an increase in the travel time.

While a simple difference is a measure often used by practitioners and is easy to calculate, it does not account for the number of vehicles that experience these benefits on each link. Furthermore, it was found in Chapter 3 that the number of vehicles that traverse the entire corridor is very low relative to the number of vehicles that traverse a given link. Therefore, this measure applies only to a very small proportion of the overall vehicle population.

To illustrate the importance of accounting for the number of vehicles that experience the travel time savings (or losses) on each link, consider a hypothetical corridor consisting of three links in which the differences between the Before and After mean travel times for the three links are {-1, 7, 3} minutes. Using Equation 6.2, the average travel time savings per corridor trip is computed as nine minutes. Given that there are three links in the corridor, the result from Equation 6.2 can be divided by three to yield average travel time savings per link of three minutes.

This approach assumes that the same number of trips traverse each link (e.g. the traffic volumes on the three links are equal). Assuming that the volumes for these links during the analysis period are {1000, 200, 600} vehicles, the volume-weighted average travel time savings per link is computed as:

$$\frac{-1 * 1000 + 7 * 200 + 3 * 600}{1000 + 200 + 600} = 1.2\, minutes$$

These two estimates of the average travel time savings per link are quite different, which highlights the need to consider the number of vehicles that traversed each link in the evaluation of signal retiming measures. However, computing the impact in terms of the change in travel time on a per link basis is unlikely to be useful to a practitioner. First, links are likely to be different lengths. Second, these links exist only in terms of the Bluetooth detector locations for the Before and After data collections and are not the concern of a practitioner wanting to produce final results independent of the Bluetooth detector locations. Consequently, it is more useful to calculate the impact of the signal retiming by weighting by the link volume and normalizing for the link length (i.e. computing on a per kilometre basis instead of on a per link basis).

6.3.2 MOE$_2$: Volume-Weighted Average Travel Time Savings per Kilometre

To establish the effect of the number of vehicles that experienced the travel time savings on each link, the difference between the means of the Before and After travel time data sets can be weighted according to the number of vehicles that experienced these travel time savings relative to the number of vehicles on all links of the corridor. This value can then be divided by the length of the link to produce a measure on a per kilometre basis. When the volumes on each link are not known, it is recommended that the number of acquired Bluetooth travel times be used instead. This assumes that the level of market penetration of Bluetooth technology is constant across different links. Equation 6.3 and Equation 6.4 show the calculation of the second measure of effectiveness, the volume-weighted average travel time savings per kilometre:

$$MOE_2 = \frac{1}{\sum_{i=1}^{j} v_i} * \sum_{i=1}^{j} \left(\frac{(\bar{t}_{Bi} - \bar{t}_{Ai}) * v_i}{L_i} \right) \tag{6.3}$$

$$MOE_2 = \frac{1}{\sum_{i=1}^{j} (n_{Bi} + n_{Ai})} * \sum_{i=1}^{j} \left(\frac{(\bar{t}_{Bi} - \bar{t}_{Ai}) * (n_{Bi} + n_{Ai})}{L_i} \right) \tag{6.4}$$

where MOE_2 represents the volume-weighted average savings (in units of minutes per kilometre), v_i represents the volume on a given link i, n_{Bi} and n_{Ai} represent the number of travel times recorded during the Before and After periods for link i respectively, and L_i represents the length of link i. Equation 6.3 can be used when link volumes are available while Equation 6.4 can be used when they are not. This applies to all father developed measures of effectiveness that attempt to incorporate link volumes.

As an example, if this measure of effectiveness yields a value of 1.5 minutes, this value can be interpreted as the average travel time savings per kilometre that a vehicle is expected to experience between the Before and After periods. This measure is more robust in that it reflects the fact that only a small number of vehicles traverse the entire corridor.

It can be understood that increasing the number of vehicles (v_i or $n_{Bi} + n_{Ai}$) that a link experienced will mathematically magnify the travel time savings or losses experienced on that link. Conversely, decreasing the number of vehicles that a link experienced will reduce the travel time savings or losses experienced on that link. In this way, this measure of effectiveness magnifies or reduces the calculated travel time savings or losses on each link according to how many vehicles experienced those savings or losses.

Having incorporated the number of vehicles that experienced the calculated travel time savings on a per kilometre basis, the statistical validity of these savings is the final component to be accounted for. Since current practice for evaluating signal retiming measures involves the collection of a small number of GPS probe vehicle travel times, statistical significance is often neglected. Incorporating a statistical indicator into the calculation process will produce a measure that can show that the implemented signal retiming measures have resulted in statistically significant travel time savings for a large number of vehicles.

6.3.3 MOE$_3$: Volume-Weighted Statistically Significant Average Travel Time Savings per Kilometre

The third and final measure of effectiveness accounts for the statistical significance of the difference in the means for each link by modifying the volume-weighted average travel time savings per kilometre according to the results of the t-test. Equation 6.5 shows the calculation of the third measure of effectiveness, the volume-weighted statistically significant average travel time savings per kilometre:

$$MOE_3 = \frac{1}{\sum_{i=1}^{j}(n_{Bi} + n_{Ai})} * \sum_{i=1}^{j}\left(\frac{(\bar{t}_{Bi} - \bar{t}_{Ai}) * (n_{Bi} + n_{Ai}) * \gamma_i}{L_i}\right) \tag{6.5}$$

where MOE_3 represents the volume-weighted statistically significant average savings (in units of minutes per kilometre) and γ_i is a binary variable that is equal to 1 if the difference between the Before and After means is statistically significant and is equal to zero if there is insufficient evidence to conclude that the difference between the Before and After mean travel times is different from zero.

It can be understood that the travel time savings produced by this measure are reduced if the t-test does not reveal any evidence to support the statistical significance of the difference in the means for a given link (i.e. $\gamma_i = 0$ for at least one link). Conversely, if all links yield statistically significant differences in the mean (i.e. $\gamma_i = 1$ for all links), then there will be no difference between the second and third measures of effectiveness.

This third measure of effectiveness is the recommended measure to be used by practitioners to quantify the success of signal retiming measures for the following reasons:

- It is based upon a measure that is well understood by practitioners in the difference in the travel time means;
- It accounts for the number of vehicles that experience these differences through volume weighting;
- It represents a measure that can be applied to any vehicle that traverses any portion of the corridor by producing a measure that is on a per kilometre basis; and
- It accounts for the statistical significance of the difference in the means through statistical testing so that any calculated savings can be attributed to the implemented signal retiming measures and not to randomness (according to a 95% level of confidence).

While these measures of effectiveness have been developed with Bluetooth travel time data in mind, the robustness of these measures allows them to be applied to travel times collected by any type of fixed-detector technology such as license plate matching technology or Wi-Fi technology. Furthermore, with the greater prevalence of travel time data obtained through smartphones via GPS satellites, it is not infeasible to place mock detectors along a corridor and use these to collect "hits" to obtain travel times as input into these measures of effectiveness (this procedure was carried out in the development of simulated GPS probe vehicle travel time data and true travel time data, as described in Chapter 3).

6.3.4 Percent Reduction Measures of Effectiveness

As an alternative to the developed measures of effectiveness that show the absolute travel time savings along a corridor, it is likely that the reduction in the mean expressed as a percentage is another parameter desired by a practitioner. This section details the development of such a measure using each developed measure of effectiveness as a basis.

Equation 6.6 shows the calculation of the percent reduction in the mean using the average travel time savings per corridor trip (MOE_1) as a basis:

$$R_1 = \frac{MOE_1}{\sum_{i=1}^{j} \bar{t}_{Bi}} \qquad (6.6)$$

where R_1 represents the average percent reduction in the mean per corridor trip.

Taking volume weighting into account, Equation 6.7 shows the calculation of the percent reduction in the mean using the volume-weighted average travel time savings per kilometre (MOE_2) as a basis:

$$R_2 = \frac{MOE_2}{\frac{1}{\sum_{i=1}^{j}(n_{Bi} + n_{Ai})} * \sum_{i=1}^{j}\left(\frac{\bar{t}_{Bi} * (n_{Bi} + n_{Ai})}{L_i}\right)} \qquad (6.7)$$

where R_2 represents the impact of the signal retiming (in terms of the volume-weighted impacts) as a fraction of the average before travel time per kilometre.

Taking the statistical significance of the difference in the mean into account, Equation 6.8 shows the calculation of the percent reduction in the mean using the statistically significant volume-weighted average travel time savings per kilometre (MOE_3) as a basis:

$$R_3 = \frac{MOE_3}{\frac{1}{\sum_{i=1}^{j}(n_{Bi} + n_{Ai})} * \sum_{i=1}^{j}\left(\frac{\bar{t}_{Bi} * (n_{Bi} + n_{Ai})}{L_i}\right)} \tag{6.8}$$

where R_3 represents the impact of the signal retiming (in terms of the volume-weighted statistically significant impacts) as a fraction of the average before travel time per kilometre. It must be noted that the denominator is the same in both Equation 6.7 and Equation 6.8.

Similar to the third measure of effectiveness in Equation 6.4, the third percent reduction in the mean encompasses all of the components that a practitioner needs to account for to be able to definitely quantify the success of a signal retiming study. It is important to note that this measure can also provide insight into possible values of β, the minimum percent reduction in the mean that a practitioner wants to statistically confirm when identifying the duration of the data collection period (which is elaborated upon in Chapter 5).

6.4 Application of Measures of Effectiveness using Bluetooth Travel Time Data

In this section, the developed measures of effectiveness are used to evaluate the effectiveness of the implemented signal retiming measures for the field corridor as well as the simulated corridor using the collected Bluetooth travel time data.

6.4.1 Measures of Effectiveness for the Field Bluetooth Data

To illustrate the results of the developed measures of effectiveness, a diagram was developed that was similar to what the Lee County Department of Transportation used in their 2008 signal retiming study. Figure 6.1 shows the measure of effectiveness diagram for the peak morning period for the field Bluetooth data. In this figure, the northbound measures are shown on the right side of the diagram and the southbound measures are shown on the left side of the diagram. Each row of data represents the data collected for a given link defined by a pair of adjacent Bluetooth detectors with the seven Bluetooth detectors shown in the centre of the diagram.

The data for the southbound direction can be used to demonstrate the meaning of each variable, each of which are represented by a column of data. The first column contains the mean travel times of the before data ($\bar{t}_{Bi}$), which is the basis for calculating the percent reductions using Equation 6.6, Equation 6.7 and Equation 6.8. The second column contains the differences in the mean travel times of the Before and After data ($\bar{t}_{Bi} - \bar{t}_{Ai}$), which are the basis for calculating the absolute travel time savings using Equation 6.2, Equation 6.3, Equation 6.4 and Equation 6.5. Four links were found to experience travel time savings, which are indicated by cells highlighted green. The two links that experienced travel time increases are indicated by cells highlighted red. The average travel time savings per corridor trip (MOE_1) can be calculated using the values of the second column as input into Equation 6.2, yielding savings of 151 seconds. The average percent reduction in the mean travel time per corridor trip (R_1) can also be calculated using the first column as input into Equation 6.6, yielding a percent reduction of 8%.

The third column contains the volume that each link experienced (in this case, $n_{Bi} + n_{Ai}$ is used instead of v_i) and the fourth column contains the length of each link (L_i). The volume-weighted average travel time savings per kilometre (MOE_2) can be calculated using the values of the second, third and fourth columns as input into Equation 6.4, yielding savings of 16 seconds per kilometre. The impact of the signal retiming (in terms of the volume-weighted impacts) as a fraction of the average before travel time per kilometre (R_2) can also be calculated using all four columns of data as input into Equation 6.7, yielding a percent reduction of 12%.

The fifth column contains the results of the t-test performed on the difference in the mean for each link. Two links were found to have differences in the mean travel times that were not found to be statistically significant with each indicated by a “0” in a cell highlighted yellow. The volume-weighted statistically significant average travel time savings per kilometre (MOE_3) can be calculated using the values of the second, third, fourth and fifth columns as input into Equation 6.5, yielding savings of 16 seconds per kilometre. The impact of the signal retiming (in terms of the volume-weighted statistically significant impacts) as a fraction of the average before travel time per kilometre (R_3) can also be calculated using all five columns of data as input into Equation 6.8, yielding a percent reduction of 12%.

The developed measures show that for the southbound direction during the morning peak period a driver will save an average of 16 seconds per kilometre on their trip, which represents a 12% average reduction in the mean travel time per kilometre (R_3). Conversely, much smaller savings are evident for the northbound direction. This is expected due to the fact that traffic patterns along this arterial corridor indicate heavy commuter traffic travelling southbound into the central business district of Toronto in the morning. Therefore, while the exact signal retiming measures implemented on this corridor are unknown, it is clear that these measures intended to facilitate the progression of traffic travelling southbound at the expense of traffic travelling northbound.

The volume-weighted average savings per kilometre and the volume-weighted statistically significant average savings per kilometre can also be seen to yield identical values for each direction. This shows that the majority of the calculated travel time savings at the link level are statistically significant, which supports the notion that the signal retiming measures successfully produced a statistically significant reduction in the travel time for each direction of this corridor. It interesting to note that for the northbound direction the value of the volume-weighted statistically significant average savings is greater than that of the volume-weighted average savings. This can be explained by the fact that one link was excluded (by way of the parameter γ_i being equal to zero) for which travel time losses were produced as a result of signal retiming. Consequently, the volume-weighted statistically significant average savings “lost” a negative value during the calculation process, resulting in an increase relative to the volume-weighted average savings.

According to the percent reduction in the mean (R_3), the mean was reduced by 12% for the southbound direction, the direction for which it appears coordination was desired. This percent reduction is not far off from the value of $\beta = 0.10$ examined in Chapter 5, which suggests that $\beta = 0.10$ is appropriate for being able to statistically confirm such a reduction in the mean travel time in the direction for which coordination is believed to be intended.

Period:	Morning Peak
Period Start:	7:00:00am
Period End:	10:00:00am

Southbound						Northbound				
Before Mean (s)	Mean Difference (s)	Volume	Length (km)	Mean Difference Significant?		Before Mean (s)	Mean Difference (s)	Volume	Length (km)	Mean Difference Significant?
					7 - Gordon Baker					
223	12	2052	2.5	1	|	227	-6	1885	2.5	0
					6 - Van Horne					
357	85	4206	2.0	1	|	211	-4	2920	2.0	1
					5 - Consumers					
376	56	2989	2.5	1	|	362	35	2314	2.5	1
					4 - Rowena					
251	-10	4441	2.4	1	|	274	-20	3554	2.4	1
					3 - Elvaston					
428	-11	998	4.0	0	|	418	22	1329	4.0	1
					2 - Crescent Town					
308	18	504	1.8	0	|	333	35	696	1.8	1
					1 - Kingston					

Southbound: Total Volume 15190

MOE 1 (s)	150.6	% Reduction 1	7.8
MOE 2 (s/km)	15.8	% Reduction 2	11.8
MOE 3 (s/km)	15.6	% Reduction 3	11.6

Northbound: Total Volume 12698

MOE 1 (s)	61.9	% Reduction 1	3.4
MOE 2 (s/km)	1.0	% Reduction 2	0.8
MOE 3 (s/km)	1.4	% Reduction 3	1.2

Figure 6.1: Measure of effectiveness diagram for the morning peak period for the field Bluetooth data

Figure 6.2 shows the measure of effectiveness diagram for the evening peak period for the field Bluetooth data. Greater average savings are evident for the northbound direction (20 seconds per kilometre, a 14% reduction in the mean) compared to the southbound direction (8 seconds per kilometre, a 6% reduction in the mean). Similar to commuter traffic travelling southbound in the morning, it is clear that implemented signal retiming measures intended to facilitate the movement of commuter traffic travelling northbound out of the city during the evening peak period at the expense of traffic travelling southbound.

Period:	Evening Peak
Period Start:	3:30:00pm
Period End:	6:30:00pm

Southbound						Northbound				
Before Mean (s)	Mean Difference (s)	Volume	Length (km)	Mean Difference Significant?		Before Mean (s)	Mean Difference (s)	Volume	Length (km)	Mean Difference Significant?
					7 - Gordon Baker					
256	-61	2881	2.5	1	|	219	-10	2224	2.5	1
					6 - Van Horne					
383	96	4229	2.0	1	|	288	21	3539	2.0	1
					5 - Consumers					
315	-4	2864	2.5	0	|	379	56	2235	2.5	1
					4 - Rowena					
254	-4	3516	2.4	0	|	380	87	4282	2.4	1
					3 - Elvaston					
500	34	1029	4.0	1	|	562	-29	1119	4.0	1
					2 - Crescent Town					
317	-52	603	1.8	1	|	435	150	616	1.8	1
					1 - Kingston					

Total Volume 15122

MOE 1 (s)	9.2	% Reduction 1	0.5
MOE 2 (s/km)	7.6	% Reduction 2	5.6
MOE 3 (s/km)	8.3	% Reduction 3	6.0

Total Volume 14015

MOE 1 (s)	276.2	% Reduction 1	12.2
MOE 2 (s/km)	19.8	% Reduction 2	13.7
MOE 3 (s/km)	19.8	% Reduction 3	13.7

Figure 6.2: Measure of effectiveness diagram for the evening peak period for the Bluetooth data with error incorporated

It must be noted that these final values should not solely be relied upon by a practitioner when making conclusions regarding the impacts of a signal retiming study. The entire measure of effectiveness diagram should be examined to provide context as to why the final measures of effectiveness produced the values that they did. For example, in Figure 6.2 the southbound direction results indicate that two links experienced statistically significant dis-benefits, thus resulting in lower overall savings. As a result, the features of each link that influence the final measures of effectiveness must be considered by a practitioner when presenting the results of a signal retiming study to the policymakers and the public.

6.4.2 Measures of Effectiveness for the Simulated Bluetooth Data

Figure 6.3 shows the measure of effectiveness diagram for the morning peak period for the simulated Bluetooth data with error incorporated. Statistically significant savings are evident for the northbound direction (3 seconds per kilometre, a 3% reduction in the mean) while no statistically significant travel time savings are evident for the southbound direction. This indicates that the signal retiming measures implemented in the simulation improved conditions for the northbound direction and facilitated the progression of peak morning traffic towards Highway 401 while traffic in the southbound direction did not receive significant improvements.

Period:	Morning Peak
Period Start:	7:30:00am
Period End:	10:30:00am

Southbound

Before Mean (s)	Mean Difference (s)	Volume	Length (km)	Mean Difference Significant?
221	1	153	1.9	0
170	6	255	2.4	0

3 - Eagle
|
2 - Bishop
|
1 - Todd

Northbound

Before Mean (s)	Mean Difference (s)	Volume	Length (km)	Mean Difference Significant?
173	12	235	1.9	1
228	7	255	2.4	0

Southbound Total Volume 408

MOE 1 (s)	6.5	% Reduction 1	1.7
MOE 2 (s/km)	1.6	% Reduction 2	1.9
MOE 3 (s/km)	0.0	% Reduction 3	0.0

Northbound Total Volume 490

MOE 1 (s)	18.2	% Reduction 1	4.5
MOE 2 (s/km)	4.4	% Reduction 2	4.7
MOE 3 (s/km)	2.9	% Reduction 3	3.1

Figure 6.3: Measure of effectiveness diagram for the morning peak period for the simulated Bluetooth data with error incorporated

Figure 6.4 shows the measure of effectiveness diagram for the peak evening period for the simulated Bluetooth data with error incorporated. Statistically significant savings are evident for the southbound direction (48 seconds per kilometre, a 25% reduction in the mean) while no statistically significant savings are evident for the northbound direction. This is expected due to the fact that traffic volumes during the evening peak period are much higher than those of the morning peak period; therefore, any improvements in the signal timings along this corridor will be most pronounced during this time period. This can be coupled with the fact that Cambridge can be characterised as a bedroom community in that commuter traffic travels southbound along Hespeler Road during the evening peak period as it arrives from the surrounding cities accessed via Highway 401 (this phenomenon was also observed in Chapter 3). It can therefore be understood that volumes for the southbound direction will be much greater than those of the northbound direction for the evening peak period and that any improvements in the signal timing plans along this corridor will impact the southbound evening scenario the most.

Period:	Evening Peak
Period Start:	3:00:00pm
Period End:	6:00:00pm

Southbound

Before Mean (s)	Mean Difference (s)	Volume	Length (km)	Mean Difference Significant?
546	139	309	1.9	1
251	64	379	2.4	1

3 - Eagle
|
2 - Bishop
|
1 - Todd

Northbound

Before Mean (s)	Mean Difference (s)	Volume	Length (km)	Mean Difference Significant?
206	11	195	1.9	0
230	7	167	2.4	0

Southbound Total Volume 688

MOE 1 (s)	202.8	% Reduction 1	25.4
MOE 2 (s/km)	47.5	% Reduction 2	25.4
MOE 3 (s/km)	47.5	% Reduction 3	25.4

Northbound Total Volume 362

MOE 1 (s)	17.9	% Reduction 1	4.1
MOE 2 (s/km)	4.5	% Reduction 2	4.3
MOE 3 (s/km)	0.0	% Reduction 3	0.0

Figure 6.4: Measure of effectiveness diagram for the evening peak period for the simulated Bluetooth data with error

For a comparison of the average percent reduction in the mean per corridor trip for all types of simulated data, Figure 6.5 compares the values of this measure of effectiveness for the simulated Bluetooth data with and without error, GPS data and true travel time data. The first measure of effectiveness was examined instead of the second or third measures of effectiveness so that the GPS probe vehicle travel time data could also be compared to the true travel time data. The measure of effectiveness diagrams for the Bluetooth travel time data without error and for the true travel time data are located in Appendix E.

It can be seen in this figure that the morning peak period results for the simulated GPS travel time data differ from the results for the true travel time data more than the Bluetooth travel time data do with or without error. This can be attributed to the number of observations being low enough such that the GPS data sampling rate is unable to capture the behaviour of the true travel time data. For the Bluetooth data, it can be seen that the results with and without error are similar to the true results, showing that the sampled Bluetooth data with or without error are representative of the truth. It can be concluded that the incorporation of error does not cause a noticeable deviation from the truth, indicating that the presence of “missed outliers” does not significantly impact the developed measure of effectiveness.

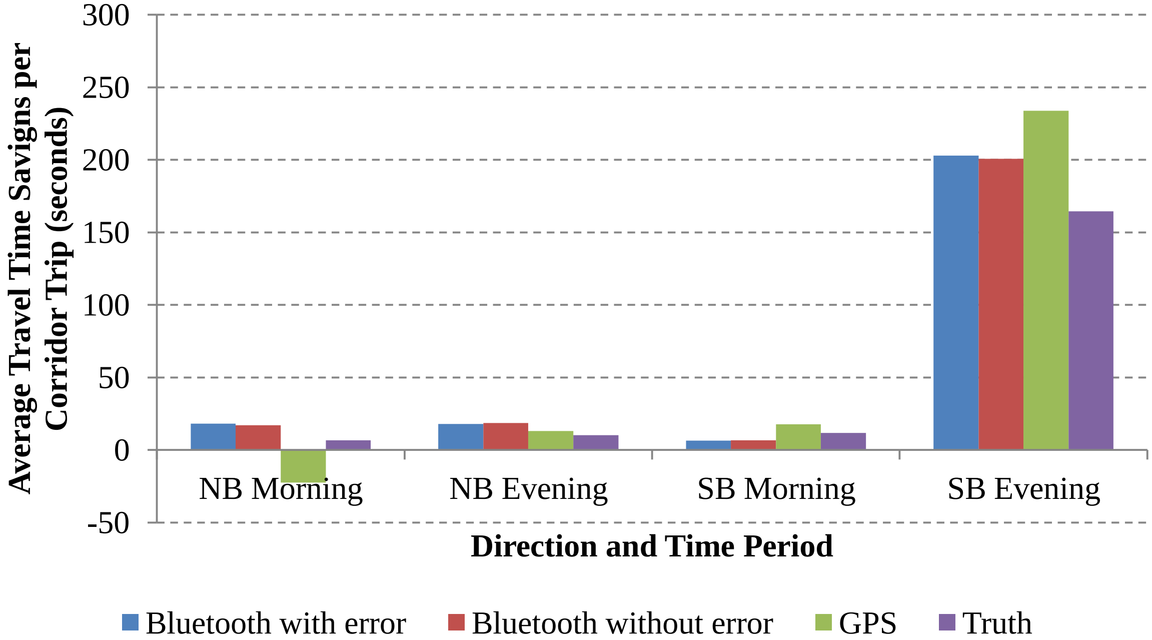

Figure 6.5: Comparison of the average travel time savings per corridor trip for the simulated Bluetooth, GPS and true travel time data

6.5 Summary

In this chapter, the development of a methodology to evaluate the implementation of signal retiming measures was described. These resulting measures of effectiveness were then applied to the field and simulated Bluetooth data. The following findings were concluded from this chapter:

- Measures of effectiveness were developed that incorporated the difference in the mean, the number of vehicles that experienced the impacts of signal retiming measures and the statistical significance of the difference in the mean. The measure that combines all three of these components is the recommended measure for evaluating signal retiming measures.
- A supplementary measure, the percent reduction in the mean, is a measure of effectiveness that was developed as an alternative to the absolute measures of effectiveness developed previously. This measure can also provide insight into possible values of β, the minimum percent reduction in the mean that a practitioner wants to statistically confirm when identifying the duration of the data collection period.

- Only the mean was able to be upgraded from the link level to the path level. It is difficult to justify a method of upgrading the standard deviation because the covariance is unknown.
- Using these measures of effectiveness to evaluate the field corridor using Bluetooth travel time data revealed savings in the average travel time during the peak commuting periods in the direction of traffic for which volumes were expected to be highest (i.e. southbound morning and northbound evening). Using this measure of effectiveness to evaluate the simulated corridor using Bluetooth travel time data revealed savings in the average travel time for during the peak commuting period in the direction of traffic for which volumes were expected to be highest (i.e. southbound evening).
- A comparison of the simulated Bluetooth data with and without error and the GPS data to the true travel time data was performed via the average travel time savings per corridor trip. The results for the GPS data were found to deviate from the results of the true travel time data. The low sample size of the GPS data explains this deviation. The Bluetooth data with and without error closely matched the true travel time data, showing that the sampled Bluetooth data with or without error is representative of the truth.

7 Influence of Outliers on Measures of Effectiveness

7.1 Introduction

This chapter describes the investigation into the influence that outliers have on the measures of effectiveness developed in Chapter 6. First, the methodology involved with testing the effect that outliers have on the measure of effectiveness is described. This methodology is then applied to the Bluetooth data collected in the field and the results for each combination of direction and time period are described. This chapter concludes with a summary of the major observations made regarding the sensitivity of the developed measure of effectiveness to the presence of outliers.

7.2 Methodology

The goal of this part of the investigation was to perform a basic sensitivity analysis on the travel time data with respect to outliers. In other words, this chapter intended to answer the question: "If a certain subset of the collected travel times is classified as a group of outliers and are excluded from the data, how will the resulting measures of effectiveness change?" Answering this question helps in determining whether or not outliers significantly impede the use of Bluetooth travel time data to evaluate a signal retiming study. This is an issue that a practitioner used to the absence of outliers in GPS probe vehicle travel time data is likely to be concerned about.

To answer this question, the exclusion of a certain top percentile travel times was explored. This strategy was more advantageous than using static boundaries to identify outliers, which comes with the assumption that traffic conditions are to remain relatively constant over time. Static boundaries would have been inaccurate because constant long-term traffic conditions do not exist in reality (and are a major reason why signal retiming studies are performed in the first place). The use of a percentile is also more advantageous than the use of heuristic measures, which become more laborious as the proportion of the vehicle population that is Bluetooth-enabled increases and traffic patterns change. While more complex outlier detection algorithms that more accurately reflect reality have been developed, the goal of this part of the investigation was simply to provide insight into the effect that excluding an explicitly defined collection of travel times has on the results of a signal retiming evaluation.

Of the previously developed measures of effectiveness, only the first (the average travel time savings per corridor trip) was examined in this chapter. The second and third measures of effectiveness (the volume-weighted average savings per kilometre and the volume-weighted statistically significant average savings per kilometre) were not examined so that only the basic effect of outliers on the developed measure of effectiveness could be measured.

With two directions and two time periods of data available, a total of four scenarios were examined. Top percentiles ranging from zero (no travel times excluded) to 99 (the top 99th percentile of travel times excluded) were designated for application to the field Bluetooth travel time data. Therefore, the top n^{th} percentile of the Bluetooth travel time data was excluded for a given link, direction, time period, and Before or After data set. For each percentile and for each combination of direction and time period, the sum of the link-level differences in the means was then calculated. For comparison, the results for the Traffax interquartile range algorithm are also presented and are referred to as the “included filter” results.

Since travel times that are outliers in reality (produced by a situation similar to when a vehicle leaves the corridor and then returns) do not exist in the simulated environment defined in Chapter 3, the simulated Bluetooth data were not examined in this part of the investigation.

Regression was used to model the relationship between the average travel time savings per corridor trip and the top percentile of outliers excluded for each combination of direction and time period. Whether or not the data were sensitive to outliers for each combination of direction and time period was assessed based on the statistical significance and magnitude of the y-intercept and slope coefficient of the model. R-squared values were not considered because identifying the strength of the correlation between the dependent and independent variables was not desired; instead, it was desired to identify the nature of the relationship between the dependent and independent variables through an assessment of the magnitudes and the significance of the model parameters.

It must be noted that in this part of the investigation weekend data were excluded and that travel times labelled as outliers by the included filter were not excluded.

It was expected that there would be three discernible regions for the average corridor travel time as a function of the top percentage of travel times excluded, each of which are illustrated in Figure 7.1 for a given link:

- For lower percentiles from zero (no data excluded) to P, a sharp decrease would be evident in the average corridor travel time. This would occur due to the removal of abnormally high travel times from the data (i.e. travel times that are likely to be outliers in reality).
- Moving towards higher outlier percentiles between P and Q, it was expected that in this region the average corridor travel time would experience a gradual decrease due to the removal of travel times that are unlikely to be outliers in reality. The margin between being labelled an outlier and being labelled valid would be very small, resulting in a gradual negative linear slope.
- From outlier percentile Q to 99 (i.e. all but 1% of the data excluded), travel times that incurred even minor amounts of delay during their trip would be excluded. A sharp decrease in the average corridor travel time is evident as the margin between being labelled an outlier and being labelled valid becomes very large. The Before and After curves would then converge upon the same value, the free speed travel time (i.e. the time taken to traverse the corridor without incurring any delay).

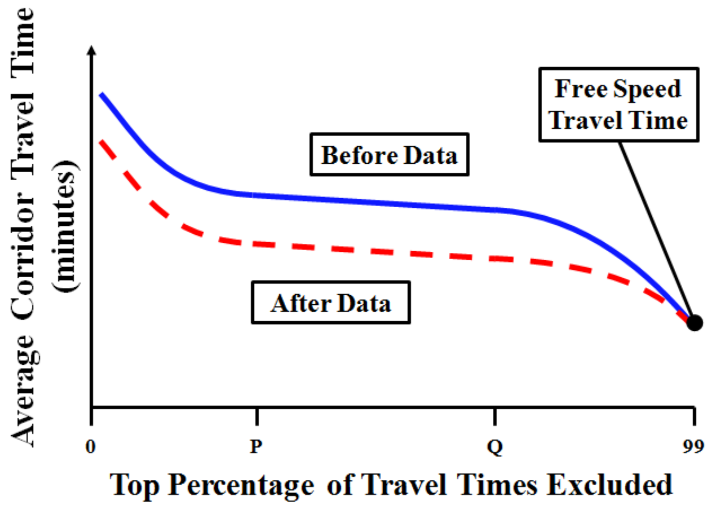

Figure 7.1: Expected relationship between the average corridor travel time and the top percentage of travel times excluded

7.3 Outlier Sensitivity Analysis of the Field Bluetooth Data

Figure 7.2 shows the average corridor travel time as a function of the top percentile of travel times excluded for the northbound direction and the morning peak period for the field Bluetooth data. The curves of the before and after data abide by the three regions that were expected: (1) a steep decrease in the average corridor travel time as the outlier percentile increases for lower percentiles (approximately 0-5%); (2) a shallow linear decrease in the average corridor travel time for most percentiles (approximately 5-95%); and (3) a sharp linear decrease in the average corridor travel time for the highest percentile (approximately 95-99%). Similar trends were observed for other directions and time periods.

Since it is clear from this figure that three distinct regions exist, regression analysis was performed only for part of the second region in which a gradual decrease in the average corridor travel time can be observed for the bulk of the tested outlier percentiles. Only part of this region was examined, specifically up to 30% because the slopes of the curves become relatively constant at an outlier percentile of 30%. This meant that regression analysis was only performed from an outlier percentile of 5% to an outlier percentile of 30%.

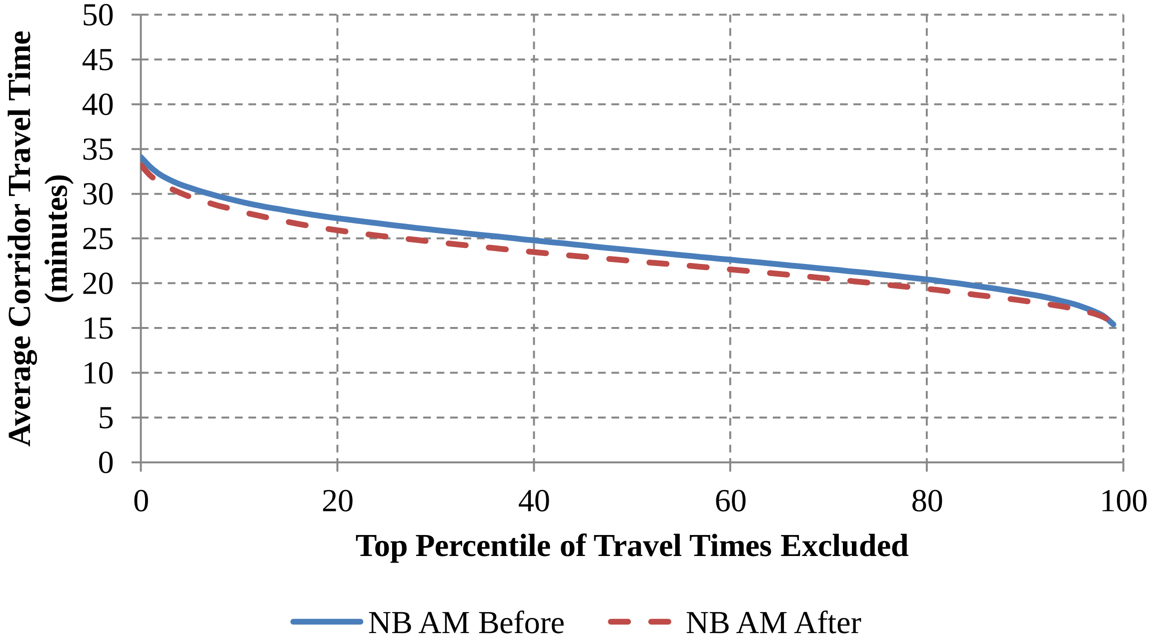

Figure 7.2: Average corridor travel time as a function of the top percentile of travel times excluded for the northbound direction and the morning peak period for the field Bluetooth data

Figure 7.3 shows the cumulative travel time frequency distribution for a given link, direction and time period for a series of top percentages of travel times excluded for the field Bluetooth data of the Before period. This figure illustrates the effect that excluding a certain percentile of the data has on the resulting travel time distribution. It can be seen that as the percentage of travel times excluded increases the travel time distribution shifts to the left, indicating a smaller frequency of longer travel times within the distribution.

The included filter can be seen to produce a cumulative travel time frequency distribution similar to when the top five percent of travel times are excluded.

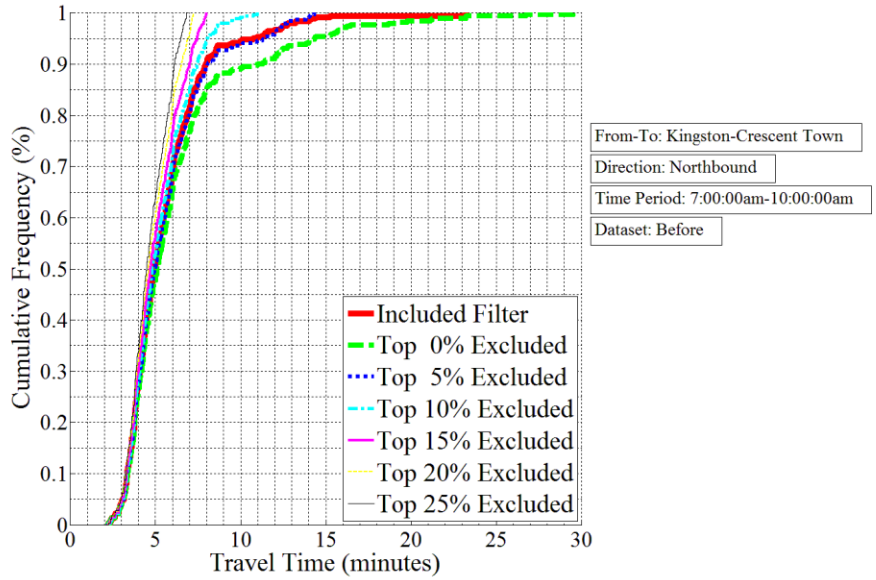

Figure 7.3: Sample cumulative travel time frequency distributions for a series of top percentages of travel times excluded for the field Bluetooth data of the Before period

Figure 7.4 shows the mean travel time as a function of the top percentage of travel times excluded for a given link, direction and time period for the field Bluetooth data. It can be seen that the mean decreases rapidly as the outlier percentile increases while for higher percentiles the decrease is less pronounced. This was expected as very high travel times are filtered out at the lower percentiles only the relatively similar travel times remain as the higher percentiles are reached. In other words, at the higher tested percentiles there are simply very few abnormally higher travel times remaining in the data, resulting in a smaller reduction in the mean as the top percentage of travel times increases.

The included filter can be seen to produce a mean travel time similar to when approximately the top five to seven percent of travel times are excluded.

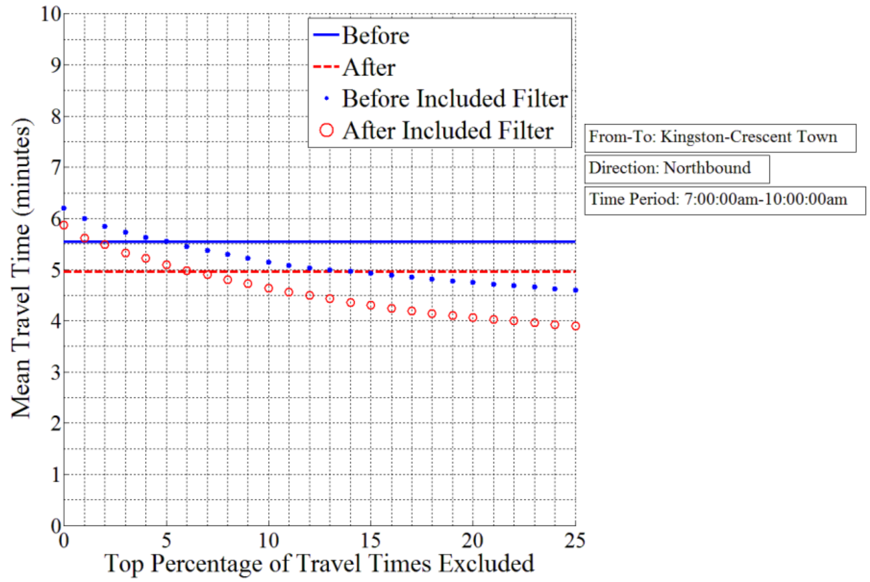

Figure 7.4: Mean travel time as a function of the top percentage of travel times excluded for a given link, direction and time period for the field Bluetooth data

Figure 7.5 shows the number of travel time observations as a function of the top percentage of travel times excluded for a given link, direction and time period for the field Bluetooth data. It can be seen that the number of observations decreases linearly as the top percentile of travel times excluded increases. This is expected as the exclusion of a top percentile for a continuous variable such as travel time will result in a fixed number of travel times being removed for each incremental increase in the top percentage of travel times excluded.

The included filter can be seen to produce a number of observations similar to when approximately the top five to ten percent of travel times are excluded.

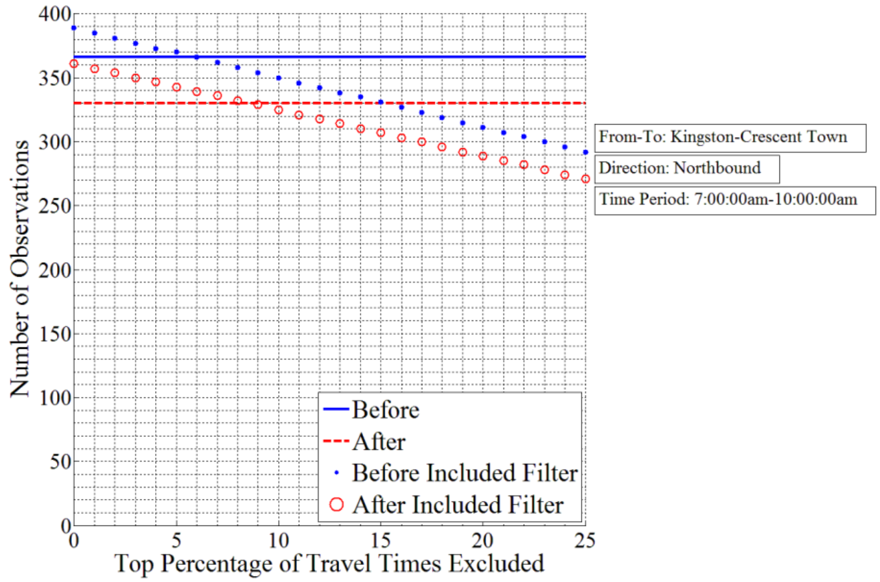

Figure 7.5: Number of observations as a function of the top percentage of travel times excluded for a given link, direction and time period for the field Bluetooth data

Figure 7.6 shows the travel times labelled as "valid" and as "outliers" as a function of time of day for an outlier percentile of 5% and for a given link, direction and time period for the field Bluetooth data. The cut-off between the data labelled "valid" and the data labelled "outliers" is clear and exists at approximately the 14-minute mark. This shows that the outlier percentile exclusion process is successfully excluding a selected top percentile of the data.

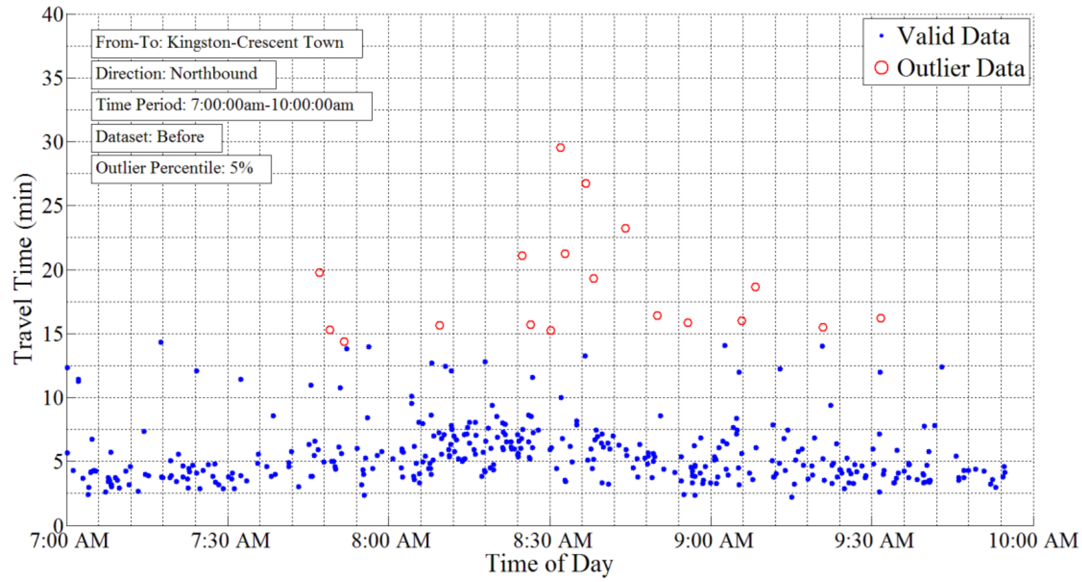

Figure 7.6: Travel time as a function of time of day for valid and outlier travel times for a given link, direction and time period for the Before period for the field Bluetooth data

Figure 7.7 shows the average savings per corridor trip as a function of the top percentage of travel times excluded for the morning peak period for the field Bluetooth data. The greater travel time savings observed for the southbound morning commute into the city are evident compared to the smaller savings for northbound morning traffic. The included filter can also be seen to produce a value similar to when the top five to ten percent of travel times are excluded for the northbound direction.

Negative values also appear at the highest percentiles. While unexpected, this is likely caused by several factors. At the highest outlier percentiles, the Before and After travel time data are very similar to each other because only travel times very close to the free speed are left. This means that the difference in the mean is unlikely to be a reliable measure. Considering that only the top percentile of travel times are being filtered for in this investigation, it is also possible that there also exist unrealistic travel times with very small magnitudes that are not being considered for exclusion. At the highest outlier percentiles, these types of travel times would greatly influence the results. It must also be remembered that the collected Bluetooth travel time data is a sample of the overall population. For the small sample sizes at the highest outlier percentiles, it is unlikely that the Bluetooth data is representative of the true free speed. Consequently, highly variable values of the mean can be expected.

The included filter produces a value much higher than any percentile value for the southbound direction (150 seconds compared to approximately 110 seconds). It can be seen that at an outlier percentile of zero (i.e. no data excluded) travel time savings of approximately 105 seconds are produced. This is a value more in line with the results of the other percentiles and shows that in this particular instance the included filter excluded a portion of travel times that in reality were not outliers. Since the examination of the nature of the included filter was beyond the scope of this investigation, the nature of this discrepancy was not reported on further.

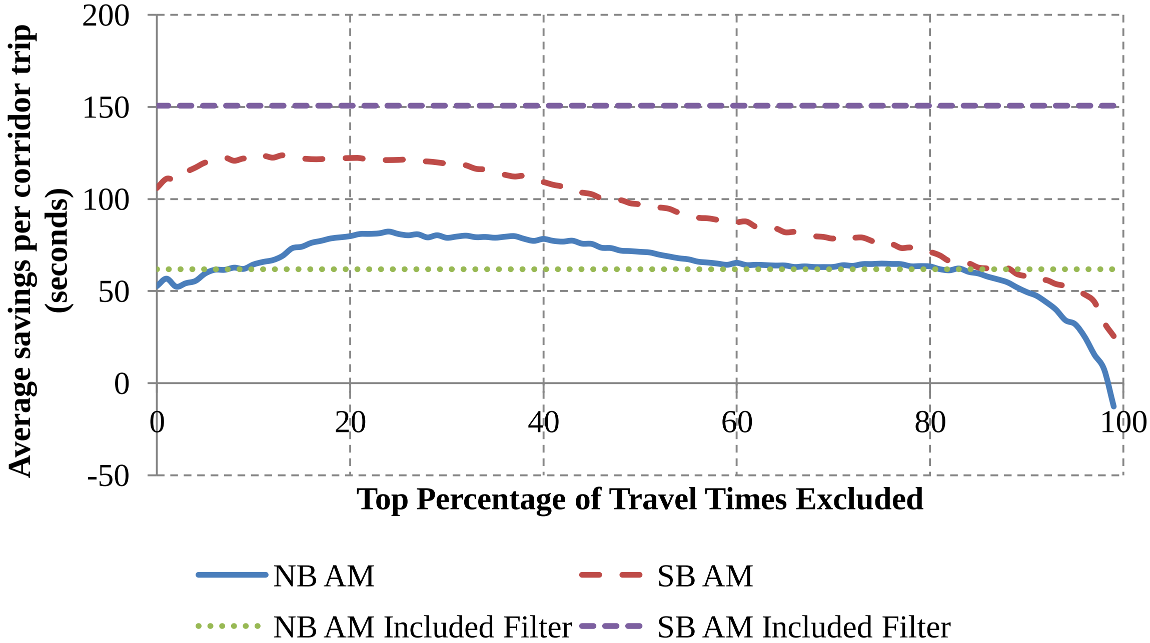

Figure 7.7: Sum of the differences in the mean as a function of the top percentage of travel times excluded for the morning peak period for the field Bluetooth data

Figure 7.8 shows the results of a regression analysis carried out for the 5-30% outlier percentile range for the evening peak period field Bluetooth data. The slope coefficient and the y-intercept was found to be statistically significant for the northbound direction while only the y-intercept was found to be statistically significant for the southbound direction. In comparison to the magnitudes of the y-intercepts, the slope coefficients are small, which indicates the relatively insignificant impact that each slope coefficient has on the measure of effectiveness as the top percentage of travel times changes. The relatively small magnitude (and lack of significance) of the slope shows that for the morning peak period outliers have a relatively insignificant influence on the average savings per corridor trip for the field Bluetooth data.

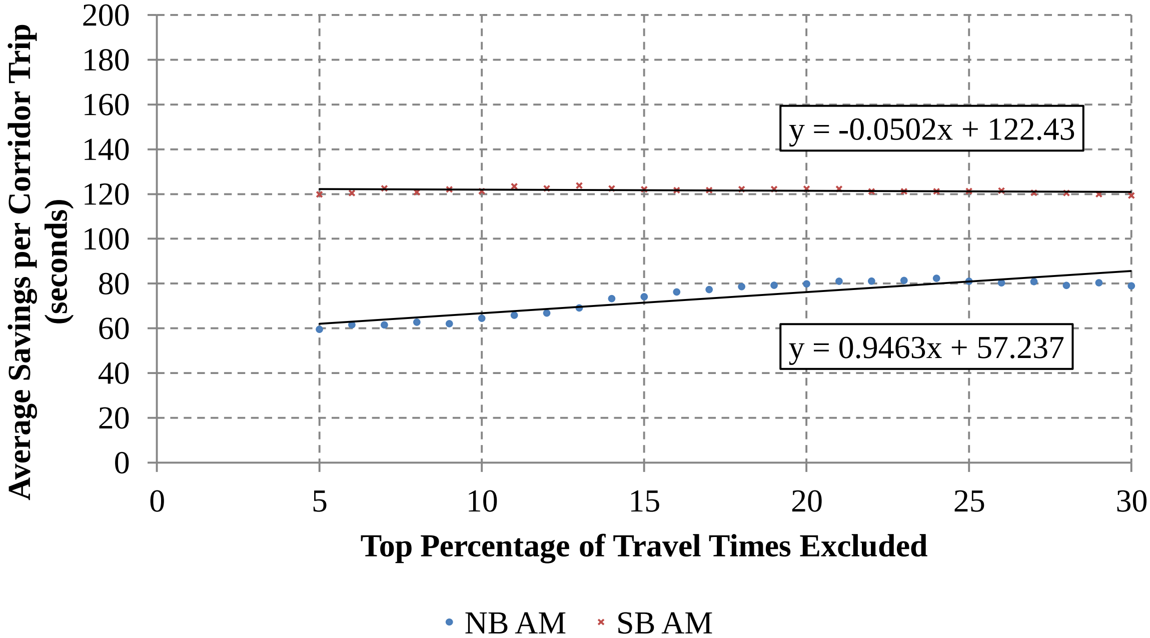

Figure 7.8: Regression analysis of the morning peak period field Bluetooth data as a function of the top percentage of travel times excluded

Figure 7.9 shows the average savings per corridor trip as a function of the top percentage of travel times excluded for the evening peak period for the field Bluetooth data. The greater travel time savings observed for the northbound evening commute out of the city are evident compared to the smaller savings for southbound evening traffic. The included filter can also be seen to produce values similar to when the top five percent of travel times are excluded for both directions.

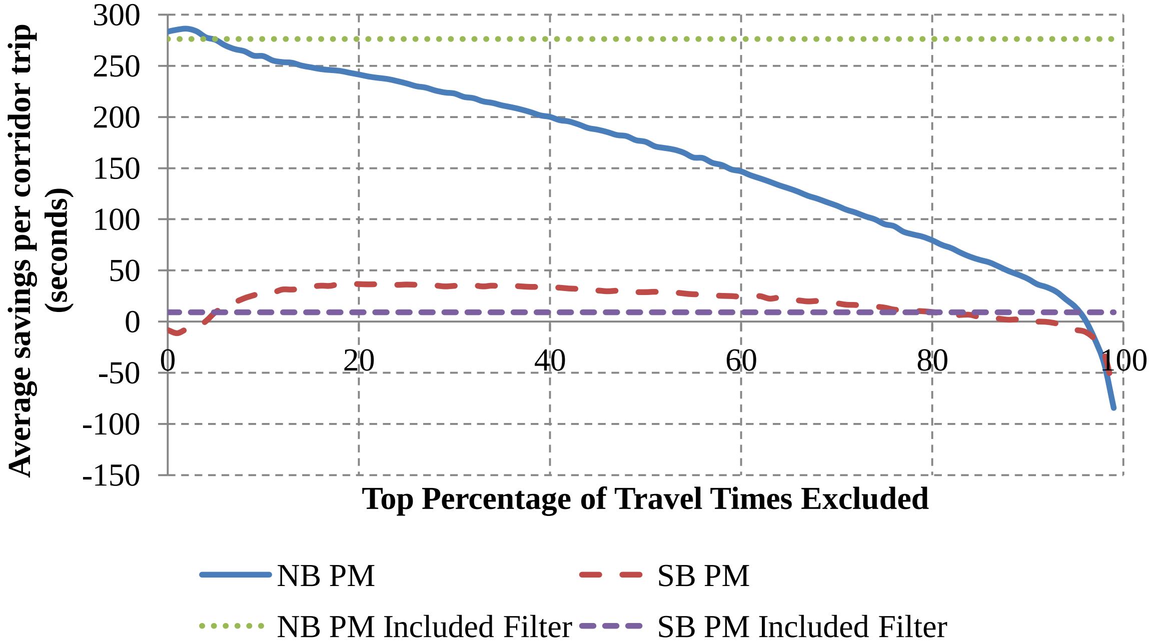

Figure 7.9: Sum of the differences in the mean as a function of the top percentage of travel times excluded for the evening peak period for the field Bluetooth data

Figure 7.10 shows the results of a regression analysis carried out for the 5-30% outlier percentile range for the evening peak period field Bluetooth data. The slope coefficient and the y-intercept were found to be statistically significant for both directions. In comparison to the magnitudes of the y-intercepts, the slope coefficients are again very small. As seen for the morning peak period, this shows that for the evening peak period outliers also have an insignificant influence on the average savings per corridor trip for the field Bluetooth data.

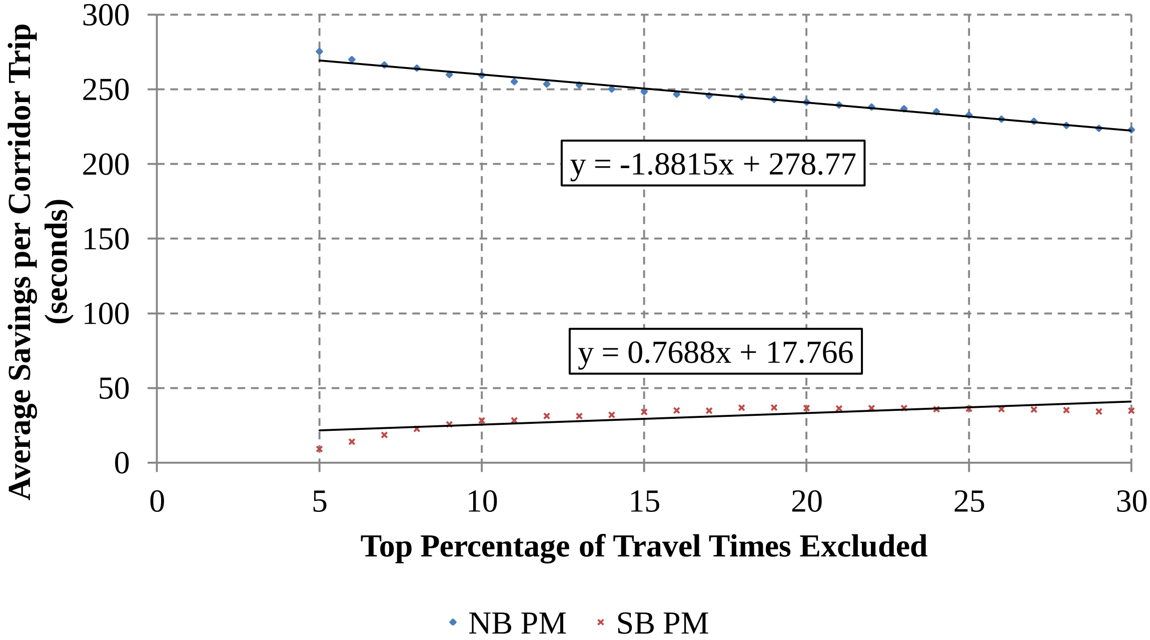

Figure 7.10: Regression analysis of the evening peak period field Bluetooth data as a function of the top percentage of travel times excluded

This analysis has shown using a basic percentile filtering method and regression that outliers do not have a significant influence on the developed measure of effectiveness, thus giving a practitioner the confidence to apply it in the evaluation of a signal retiming study despite the presence of outliers. This analysis also showed that the included filter operates within the five to ten percent range of the outlier percentile filtering method. In other words, the included filter produce values similar to when the outlier percentile filter is set to between five and ten percent. This suggests that the included filter has been designed to filter out the top five to ten percent of the travel time data.

It is important to state that these conclusions are contingent upon the definition of an outlier that was described in Chapter 1 (i.e. travel times are labelled as outliers based on travel time magnitude). Further investigation into what should define an outlier (i.e. time of day, direction, location, etc.) is essential to definitively concluding the negligible effect of outliers in a signal retiming study.

7.4 Summary

In this chapter, the influence that outliers had on the average savings per corridor trip was examined. This influence was examined to determine whether or not outliers significantly impede the use of Bluetooth travel time data to evaluate a signal retiming study. The following findings were concluded from this chapter:

- Outliers were found to have an insignificant impact on the average savings per corridor trip for the field Bluetooth data. This was concluded based on the relatively small magnitudes of the slope coefficient of the regression models used to examine the relationship between the average savings per corridor trip and the top percentage of travel times excluded.
- The Traffax interquartile range filter included in the field Bluetooth data (known as the "included filter") produces values in line with outlier percentiles generally between five and ten percent. This suggests that the included filter has been designed to filter out the top five to ten percent of the data.
- The testing of a basic outlier percentile filter on Bluetooth travel time data collected in the field environment revealed that outliers have a negligible influence on the developed measure of effectiveness. This gives a practitioner the confidence to apply the developed measure of effectiveness to Bluetooth data collected for the evaluation of a signal retiming study in spite of the presence of outliers.

8 Conclusions and Recommendations

8.1 Conclusions

The implementation of signal retiming measures has proven to be a cost-effective method of increasing the capacity of the existing network. The possibility of using Bluetooth technology in the evaluation of signal retiming studies presents further opportunities for improvements in cost savings and user benefits. However, the need for a practitioner-ready method to utilize these data for this purpose persists. This need formed the basis upon which this investigation was conducted.

This investigation explored Bluetooth travel time data collected in the field and in a simulated environment. The various characteristics of a signal retiming study itself were examined and measures of effectiveness to evaluate signal retiming studies were developed and tested. The main conclusions of this research are described in this chapter.

8.1.1 Comparison of Path-Level and Link-Level Bluetooth Travel Time Data

It was found to be difficult to statistically differentiate between the travel time data collected during the Before and After periods using the small number of travel times collected for Bluetooth-enabled vehicles that traversed the entire corridor. This made it infeasible to use the travel times of vehicles that traversed the entire corridor to evaluate signal retiming measures. Bluetooth travel time data collected at a much smaller scale was found to be a more feasible option due to producing a greater amount of travel times to work with. It was therefore recommended that for evaluating a signal retiming study a corridor should be subdivided into smaller links through the placement of multiple Bluetooth detectors.

8.1.2 Detector Location, Data Collection Period and Detector Spacing

Having established that Bluetooth travel time data can show significant differences in travel times between the Before and After periods of a signal retiming study, the parameters important to evaluating signal retiming measures were examined.

To ensure that sources of travel time variability are controlled for, detectors should be placed at midblock locations. If this commitment too great in terms of time and resources, minor street intersections can be used as detector locations instead. It would also be preferable to power the detectors using city infrastructure, provided the process to obtain permission is relatively easy. Using the field data, a regression model was developed to estimate the duration of the data collection period based on the characteristics of the corridor. Using the estimated of the data collection period, basic recommendations were made regarding the spacing of detectors.

8.1.3 Measures of Effectiveness

Measures of effectiveness were developed to evaluate the impacts of signal retiming measures on an arterial corridor. These measures of effectiveness incorporated the difference in the mean, the number of vehicles that experienced this difference in the mean, and the statistical significance of this difference in the mean. These measures provide a practitioner with an idea of the travel time savings or losses produced for the corridor, the number of vehicles that experienced these savings or losses on a per kilometre basis, and whether or not these savings or losses were statistically significant. A supplementary measure of effectiveness, the percent reduction in the mean, was also developed.

8.1.4 Influence of Outliers on Measures of Effectiveness

Evaluation of the field data indicated that the average travel time savings per corridor trip (i.e. MOE_1) was relatively insensitive to the presence or removal of an upper percentile of travel times, at least in the range of the 70^{th} to 95^{th} percentile travel times (i.e. the exclusion of the top 5% to 30% of travel times). These results suggest that the proposed measures of effectiveness are relatively insensitive to the performance of the outlier detection algorithm. Even a relatively naïve outlier detection algorithm such as an interquartile filter appear to be adequate to properly identify very long travel times (i.e. 95^{th} percentile and above) as outliers. Furthermore, these results suggest that the use of different outlier detection algorithms is unlikely to have a large impact on the measures of performance (assuming these outlier detection algorithms are reasonable accurate at identifying outlier).

8.2 Future Research

Possible subjects of future research that can help improve the findings of this investigation and the understanding of signal retiming evaluation include:

- Further investigation into the possible causes of missed Bluetooth detections discussed in Chapter 3. Comparing the field Bluetooth data to some true travel time data for various situations also collected in the field will confirm or refute the validity of each of these possible reasons.
- Since it was found for the field corridor that vehicles generally do not traverse the entire corridor, path-level and link-level travel time data should be examined for other corridors for which it is known that vehicles generally utilize the entirety of the corridor during the peak commuting periods. This will help to confirm or refute the notion that path-level data are unsuitable for evaluating signal retiming measures.
- Testing the model developed to estimate the duration of the data collection period on other field and simulated environments. This will validate the transferability of the model and can be used to develop a library of regression coefficients and values of β (the minimum percent reduction in the mean travel time that a practitioner wants to be able to statistically confirm) for a variety of corridor characteristics.
- Examining other variables that could improve the explanatory power of the model to estimate the data collection period. This could include the type of land use along the corridor or the characteristics of the roadway (i.e. the number of bus stops per direction, the presence or absence of a median, the presence or absence of bicycle lanes, etc.).
- Investigating possible methods of upgrading the standard deviation from the link level to the path level. While this was not possible based on the current understanding of the covariance associated with adjacent distributions of standard deviations, developing a viable method of doing this will aid in producing a concise measure of effectiveness that captures the influence of signal retiming measures on the variability of a vehicle's travel time.
- Examining the influence that more sophisticated outlier algorithms have on the developed measure of effectiveness. This will provide a more realistic idea of the influence of outliers in Bluetooth travel time data as outlier detection algorithms more sophisticated than percentile filtering have been developed; these algorithms were summarized in the literature review.
- Further investigation into what should define an outlier is essential to definitively concluding the negligible effect of outliers in a signal retiming study.

References

Araghi, B., Christensen, L., Krishnan, R., & Lahrmann, H. (2012). Application of Bluetooth Technology for Mode-Specific Travel Time Estimation on Arterial Roads: Potentials and Challenges. Annual Transport Conference at Aalborg University, Aalborg University, Aalborg, Denmark.

Ban, X., Chu, L., Herring, R., & Margulici, J. (2009). Optimal Sensor Olacement for Both Traffic Control and Traveler Information Applications. Compendium of Papers CD-ROM of 88th Transportation Research Annual Meeting, Washington, DC, USA.

Cambridge Systematics. (2012). Travel Time Data Collection White Paper. Florida Department of Transportation, District IV.

City of Toronto. Road Safety. Retrieved February 7, 2015, from http://www1.toronto.ca/wps/portal/contentonly?vgnextoid=747c4074781e1410VgnVCM10000071d60f89RCRD.

Day, C., Brennan, T., Hainen, A., Remias, S., & Bullock, D. (2012). Roadway System Assessment Using Bluetooth-Based Automatic Vehicle Identification Travel Time Data. (No. M5b—Handbook for Application of Bluetooth Traffic Monitoring on Arterials: Final Guidebook). West Lafayette, Indiana: Purdue University.

Day, C., Bullock, D., Li, H., Remias, S., Hainen, A., Freije, R., Stevens, A., Sturdevant, J., & Brennan, T. (2014). Performance Measures for Traffic Signal Systems: An Outcome-Oriented Approach.

Day, C., Haseman, R., Premachandra, H., Brennan, T., Wasson, J., Sturdevant, J., & Bullock, D. (2010). Evaluation of Arterial Signal Coordination. Transportation Research Record: Journal of the Transportation Research Board, 2192(1), 37-49.

Delcan & Lura Consulting. (2013). City of Toronto Congestion Management Plan 2014-2018. City of Toronto.

Edara, P., Guo, J., Smith, B., & McGhee, C. (2008). Optimal Placement of Point Detectors on Virginia's Freeways: Case Studies of Northern Virginia and Richmond. (No. VTRC 08-CR3). Virginia Department of Transportation.

Ernst, J., Day, C., Krogmeier, J., & Bullock, D. (2012). Probe Data Sampling Guidelines for Characterizing Arterial Travel Time. Transportation Research Record: Journal of the Transportation Research Board, 2135(1), 173-181.

Ernst, J., Krogmeier, J., & Bullock, D. (2012). Kullback-Leibler Comparison Framework for the Evaluation of Travel Time Distribution Estimates. 2012 15th International IEEE Conference Intelligent Transportation Systems (ITSC), Anchorage, Alaska, USA. 564-569.

Ernst, J., Krogmeier, J., & Bullock, D. (2014). Estimating Required Probe Vehicle Re-Identification Requirements for Characterizing Link Travel Times. Intelligent Transportation Systems Magazine, 6(1), 50-58.

Federal Highway Administration. (2015a). Arterial Management Program. Retrieved January 10, 2015, from http://www.ops.fhwa.dot.gov/arterial_mgmt.

Federal Highway Administration. (2015b). Flexibility in Highway Design, Chapter 3: Functional Classification. Retrieved January 10, 2015, from http://www.fhwa.dot.gov/environment/publications/flexibility/ch03.cfm.

Friesen, M., & McLeod, R. (2014). Bluetooth in Intelligent Transportation Systems: A Survey. International Journal of Intelligent Transportation Systems Research, 1-11.

Gettman, D., Fok, E., Curtis, E., Kacir, K., Ormand, D., Mayer, M., & Flanigan, E. (2013). Measures of Effectiveness and Validation Guidance for Adaptive Signal Control Technologies. (No. FHWA-HOP-13-031). Federal Highway Administration.

Google. (2015, January 9). V 7.1.2.2014. Cambridge, Ontario, Canada [map], 43° 23' 42.62" N, 80° 19' 39.85" W, Eye alt 8.66 km. Retrieved from http://www.earth.google.com.

Google, TerraMetrics, National Oceanic and Atmospheric Administration, & DigitalGlobe. (2015, January 9). V 7.1.2.2014. Toronto, Ontario, Canada [map], 43° 43' 49.30" N, 70° 09' 06.59" W, Eye alt 25.65 km. Retrieved from http://www.earth.google.com.

Google. (2014). Google maps. Retrieved February 9, 2015, from https://www.google.ca/maps/@43.4666274,-80.0802788,19z?hl=en.

Gordon, R. (2010). Traffic Signal Retiming Practices in the United States. (No. 409). Transportation Research Board.

Guiard, V., & Rasch, D. (2004). The Robustness of Two Sample Tests for Means. Psychology Science, 46(4), 549.

Haghani, A., & Hamedi, M. (2013). Application of Bluetooth Technology in Traffic Detection, Surveillance and Traffic Management. Journal of Intelligent Transportation Systems: Technology, Planning and Operations, 17(2), 107-109.

Haghani, A., Hamedi, M., Sadabadi, K., Young, S., & Tarnoff, P. (2010). Data Collection of Freeway Travel Time Ground Truth with Bluetooth Sensors. Transportation Research Record: Journal of the Transportation Research Board, 2160(1), 60-68.

HDR Corporation. (2008). Costs of Road Congestion in the Greater Toronto Area: Impact and Cost Benefit Analysis of the Metrolinx Draft Regional Transportation Plan. Greater Toronto Airports Authority.

Hu, Y. (2013). Short-Term Prediction of Freeway Travel Times using Data from Bluetooth Detectors. (Doctoral, University of Waterloo).

Imtech Traffic & Infra UK Ltd., Siemens Traffic Controls & TRL Limited. (2014). The development of SCOOT. Retrieved January 10, 2015, from http://www.scoot-utc.com/Development.php?menu=Overview.

Institute of Electrical and Electronics Engineers. (2005). IEEE standard 802.15.1.

Institute of Transportation Engineers. (2015). Traffic signal timing. Retrieved January 10, 2015, from http://www.ite.org/signal/index.asp.

Iteris. (2011). ITS Innovate Idea Project: Arterial Travel Time Monitoring System Using Bluetooth Technology. Minnesota Department of Transportation.

Jie, L., Van Zuylen, H., Chunhua, L., & Shoufeng, L. (2011). Monitoring Travel Times in an Urban Network Using Video, GPS and Bluetooth. Procedia-Social and Behavioral Sciences, 20, 630-637.

Khoei, A., Bhaskar, A., & Chung, E. (2013). Travel Time Prediction on Signalised Urban Arterials by Applying SARIMA Modelling on Bluetooth Data. 36th Australasian Transport Research Forum (ATRF) 2013.

KMJ Consulting. (2011). Using Bluetooth to Measure Travel Time Along Arterial Corridors. City of Philadelphia.

Koonce, P., Rodegerdts, L., Lee, K., Quayle, S., Beaird, S., Braud, C., & Urbanik, T. (2008). Traffic Signal Timing Manual. (No. FHWA-HOP-08-024).Federal Highway Administration.

Lee County Department of Transportation. (2008). Benefits Summary for this Signal Retiming Project. Retrieved January 10, 2015, from http://www.leesignaltiming.com/results.htm.

Li, R., Chai, H., & Tang, J. (2013). Empirical Study of Travel Time Estimation and Reliability. Mathematical Problems in Engineering, 2013.

Liu, X., Chien, S., & Kim, K. (2012). Evaluation of Floating Car Technologies for Travel Time Estimation. Journal of Modern Transportation, 20(1), 49-56.

Lomax, T., Schrank, D., Turner, S., & Margiotta, R. (2003). Selecting Travel Reliability Measures. Texas Transportation Institute.

Lund Research. (2013). Independent T-test using SPSS. Retrieved February 4, 2015 from https://statistics.laerd.com/spss-tutorials/independent-t-test-using-spss-statistics.php.

Moghaddam, S. (2014). Performing Short-Term Travel Time Prediction on Arterials. (Doctoral, University of Waterloo).

Moghaddam, S., & Hellinga, B. (2012). Quantifying Measurement Error in Arterial Travel Times measured by Bluetooth Detectors. Transportation Research Record: Journal of the Transportation Research Board, 2395(1), 111-122.

Moghaddam, S., Noroozi, R., & Hellinga, B. (2014). Real Time Prediction of Roadway Travel Times for Traveller Information and Proactive Control. OTC & CITE Joint Conference, Kitchener, Ontario, Canada.

Oh, S., & Choi, K. (2004). Optimal Detector Location for Estimating Link Travel Speed in Urban Arterial Roads. KSCE Journal of Civil Engineering, 8(3), 327-333.

Oppenlander, J. (1976). Sample Size Determination for Travel Time and Delay Studies. Traffic Engineering, 46(9).

Peng, C. (2009). Chapter 13: Analysis of Variance. Data Analysis Using SAS (pp. 363). SAGE Publications.

PTV AG. (2014). PTV Vissim 7 User Manual. (pp. 195). PTV AG, Karlsruhe, Germany.

Quayle, S., Koonce, P., DePencier, D., & Bullock, D. (2010). Arterial Performance Measures with Media Access Control Readers. Transportation Research Record: Journal of the Transportation Research Board, 2192(1), 185-193.

Quiroga, C., & Bullock, D. (1998). Determination of Sample Sizes for Travel Time Studies. ITE Journal, 68(8), 92-98.

Rakha, H., El-Shawarby, I., Arafeh, M., & Dion, F. (2006). Estimating Path Travel-Time Reliability. 2006 IEEE Intelligent Transportation Systems Conference, Toronto, Ontario, Canada. 236-241.

Roth, J. (2010). A Time Series Approach to Removing Outlying Data Points from Bluetooth Vehicle Speed Data. (Master's, University of Akron).

Sherali, H., Desai, J., & Rakha, H. (2006). A Discrete Optimization Approach for Locating Automatic Vehicle Identification Readers for the Provision of Roadway Travel Times. Transportation Research Part B: Methodological, 40(10), 857-871.

The Mathworks, Inc. (2015). Lognormal Distrubtion. Retrieved April 7, 2015, from http://www.mathworks.com/help/stats/lognormal-distribution.html.

Tindale-Oliver & Associates. (2008). Martin County Advanced Traffic Management Systems (ATMS) Assessment. (No. 181011-02.07). Martin County, Florida.

Turner, S., Eisele, W., Benz, R., & Holdener, D. (1998). Travel Time Data Collection Handbook. (No. FHWA-PL-98-035).

Van Boxel, D., Schneider, W., & Bakula, C. (2011). Innovative Real-Time Methodology for Detecting Travel Time Outliers on Interstate Highways and Urban Arterials. Transportation Research Record: Journal of the Transportation Research Board, 2256(1), 60-67.

Van Lint, J., & Van Zuylen, H. (2005). Monitoring and Predicting Freeway Travel Time Reliability: Using Width and Skew of Day-to-Day Travel Time Distribution. Transportation Research Record: Journal of the Transportation Research Board, 1917(1), 54-62.

Vo, T. (2011). An Investigation of Bluetooth Technology for Measuring Travel Times on Arterial Roads: A Case Study on Spring Street. (Georgia Institute of Technology).

Wakabayashi, H. (2010). Travel Time Reliability Indices for Highway Users and Operators. Fourth International Symposium on Transportation Network Reliability, University of Minnesota, Minneapolis, Minnesota, USA, 79-95.

Wang, J., Robinson, B., Shelby, S., Cox, K., & Townsend, W. (2010). Evaluation of ACS Lite Adaptive Control Using Sensys Arterial Travel Time Data. ITS America 20th Annual Meeting & Exposition.

Wang, W. (2014). Traffic Signal Coordination Study Along Overlea Boulevard. City of Toronto (unpublished).

Wasson, J., & Bullock, D. (2012). Database Architecture and Query Structures for Probe Data Processing.

Wasson, J., Sturdevant, J., & Bullock, D. (2008). Real-Time Travel Time Estimates Using Media Access Control Address Matching. ITE Journal, 78(6).

Young, S. (2008). Bluetooth Traffic Monitoring Technology: Concept of Operation & Deployment Guidelines. University of Maryland Center for Advanced Transportation Technology.

Appendix A Chapter 3 Link-Level Field GPS Travel Time Data

Table A.1: Mean travel times of the link-level field GPS travel time data

From	To	Direction	Time Period	GPS Probe Mean Travel Time (min)
1	2	NB	Morning	4.14
1	2	NB	Mid-day	3.89
1	2	NB	Evening	4.4
2	3	NB	Morning	6.56
2	3	NB	Mid-day	6.46
2	3	NB	Evening	6.55
3	4	NB	Morning	3.92
3	4	NB	Mid-day	3.48
3	4	NB	Evening	5
4	5	NB	Morning	4.91
4	5	NB	Mid-day	3.9
4	5	NB	Evening	5.88
5	6	NB	Morning	3.29
5	6	NB	Mid-day	3.75
5	6	NB	Evening	4.73
6	7	NB	Morning	2.96
6	7	NB	Mid-day	3.25
6	7	NB	Evening	3.52
7	6	SB	Morning	3.56
7	6	SB	Mid-day	3.08
7	6	SB	Evening	3.59
6	5	SB	Morning	6.43
6	5	SB	Mid-day	3.66
6	5	SB	Evening	5.35
5	4	SB	Morning	4.9
5	4	SB	Mid-day	3.85
5	4	SB	Evening	4.92
4	3	SB	Morning	3.84
4	3	SB	Mid-day	3.27
4	3	SB	Evening	3.77
3	2	SB	Morning	6.19
3	2	SB	Mid-day	6.57
3	2	SB	Evening	7.17
2	1	SB	Morning	4.12
2	1	SB	Mid-day	3.59

2	1	SB	Evening	4.66

Appendix B Chapter 3 Simulated Travel Time Data

Table B.1: Statistical test results for path-level simulated true travel time data

From	To	Direction	Time Period	Sample Size		Mean Travel Time (min)		Before Mean - After Mean (min)	Travel Time Stdev (min)		Before Stdev - After Stdev (min)	Stdevs Different?	Means Different?
				Before	After	Before	After		Before	After			
Todd	Eagle	NB	Morning	472	483	6.653	6.499	0.154	0.799	0.756	0.043	N	Y
Todd	Eagle	NB	Evening	306	312	7.465	7.251	0.214	0.894	0.872	0.023	N	Y
Eagle	Todd	SB	Morning	636	619	6.746	6.531	0.215	0.746	0.718	0.028	N	Y
Eagle	Todd	SB	Evening	828	872	13.358	10.518	2.841	5.837	3.032	2.805	Y	Y

Table B.2: Statistical test results for link-level simulated true travel time data

From	To	Direction	Time Period	Sample Size		Mean Travel Time (min)		Before Mean - After Mean (min)	Travel Time Stdev (min)		Before Stdev - After Stdev (min)	Stdevs Different?	Means Different?
				Before	After	Before	After		Before	After			
Todd	Bishop	NB	Morning	1276	1269	3.748	3.674	0.074	0.547	0.526	0.021	Y	N
Todd	Bishop	NB	Evening	851	859	3.871	3.794	0.077	0.622	0.595	0.027	Y	N
Bishop	Eagle	NB	Morning	1111	1134	2.748	2.713	0.035	0.560	0.559	0.002	N	N
Bishop	Eagle	NB	Evening	1000	1010	3.372	3.281	0.091	0.782	0.729	0.053	Y	Y
Eagle	Bishop	SB	Morning	1142	1138	3.782	3.670	0.112	0.706	0.660	0.046	Y	Y
Eagle	Bishop	SB	Evening	1282	1293	2.806	2.722	0.084	0.462	0.442	0.020	Y	Y
Bishop	Todd	SB	Morning	1997	2106	9.063	7.131	1.932	4.789	2.908	1.881	Y	N
Bishop	Todd	SB	Evening	1880	1871	3.971	3.160	0.811	1.528	0.690	0.839	Y	Y

Table B.3: Link-level simulated GPS travel time data

From	To	Direction	Time Period	Sample Size		Mean Travel Time (min)		Before Mean - After Mean (min)
				Before	After	Before	After	
Todd	Bishop	NB	Morning	5	5	3.747	3.443	0.303
Todd	Bishop	NB	Evening	5	5	4.047	4.283	-0.237
Bishop	Eagle	NB	Morning	5	5	2.210	2.887	-0.677
Bishop	Eagle	NB	Evening	5	5	3.693	3.240	0.453
Eagle	Bishop	SB	Morning	5	5	3.790	3.760	0.030
Eagle	Bishop	SB	Evening	5	5	12.027	10.857	1.170
Bishop	Todd	SB	Morning	5	5	2.990	2.723	0.267
Bishop	Todd	SB	Evening	5	5	5.287	2.560	2.727

Table B.4: Statistical test results for path-level simulated Bluetooth travel time data without error

From	To	Direction	Time Period	Sample Size		Mean Travel Time (min)		Before Mean - After Mean (min)	Travel Time Stdev (min)		Before Stdev - After Stdev (min)	Stdevs Different?	Means Different?
				Before	After	Before	After		Before	After			
Todd	Eagle	NB	Morning	47	58	6.649	6.469	0.179	0.803	0.800	0.003	N	N
Todd	Eagle	NB	Evening	37	29	7.306	6.993	0.313	0.923	0.661	0.262	N	N
Eagle	Todd	SB	Morning	41	42	6.695	6.436	0.259	0.628	0.677	-0.049	N	N
Eagle	Todd	SB	Evening	58	66	14.110	10.210	3.900	6.041	2.915	3.126	Y	Y

Table B.5: Statistical test results for link-level simulated Bluetooth travel time data without error

From	To	Direction	Time Period	Sample Size		Mean Travel Time (min)		Before Mean - After Mean (min)	Travel Time Stdev (min)		Before Stdev - After Stdev (min)	Stdevs Different?	Means Different?
				Before	After	Before	After		Before	After			
Todd	Bishop	NB	Morning	136	119	3.768	3.655	0.113	0.539	0.523	0.016	N	N
Todd	Bishop	NB	Evening	108	127	2.833	2.662	0.171	0.589	0.629	-0.040	N	Y
Bishop	Eagle	NB	Morning	83	84	3.823	3.699	0.124	0.539	0.533	0.006	N	N
Bishop	Eagle	NB	Evening	105	90	3.414	3.227	0.187	0.757	0.718	0.039	N	N
Eagle	Bishop	SB	Morning	77	76	3.672	3.643	0.028	0.707	0.630	0.077	N	N
Eagle	Bishop	SB	Evening	133	122	2.807	2.723	0.083	0.509	0.449	0.060	N	N
Bishop	Todd	SB	Morning	157	152	9.031	6.758	2.273	4.700	2.971	1.729	Y	Y
Bishop	Todd	SB	Evening	195	184	4.163	3.090	1.073	1.700	0.672	1.028	Y	Y

Appendix C Chapter 4 Link-Level Field Bluetooth Data

Table C.1: Sample sizes of the aggregated link-level field Bluetooth data

From	To	Direction	Time Period	Before Sample Size	After Sample Size
1	2	NB	Morning	367	330
1	2	NB	Evening	337	280
2	3	NB	Morning	683	658
2	3	NB	Evening	579	552
3	4	NB	Morning	1782	1785
3	4	NB	Evening	2162	2133
4	5	NB	Morning	1202	1120
4	5	NB	Evening	1139	1104
5	6	NB	Morning	1465	1459
5	6	NB	Evening	1636	1907
6	7	NB	Morning	911	976
6	7	NB	Evening	1027	1199
7	6	SB	Morning	985	1070
7	6	SB	Evening	1424	1460
6	5	SB	Morning	1931	2280
6	5	SB	Evening	2064	2170
5	4	SB	Morning	1398	1596
5	4	SB	Evening	1364	1505
4	3	SB	Morning	2121	2329
4	3	SB	Evening	1781	1744
3	2	SB	Morning	494	507
3	2	SB	Evening	556	476
2	1	SB	Morning	229	277
2	1	SB	Evening	308	297

Table C.2: Mean travel times of the link-level field Bluetooth data

From	To	Direction	Time Period	Before Mean (min)	After Mean (min)
1	2	NB	Morning	5.558	4.959
1	2	NB	Evening	7.259	4.746
2	3	NB	Morning	6.959	6.617
2	3	NB	Evening	9.364	9.820
3	4	NB	Morning	4.560	4.893
3	4	NB	Evening	6.337	4.882
4	5	NB	Morning	6.017	5.443
4	5	NB	Evening	6.303	5.391
5	6	NB	Morning	3.510	3.580
5	6	NB	Evening	4.795	4.439
6	7	NB	Morning	3.776	3.886
6	7	NB	Evening	3.643	3.805
7	6	SB	Morning	3.713	3.512
7	6	SB	Evening	4.259	5.268
6	5	SB	Morning	5.943	4.522
6	5	SB	Evening	6.376	4.768
5	4	SB	Morning	6.267	5.340
5	4	SB	Evening	5.249	5.322
4	3	SB	Morning	4.187	4.343
4	3	SB	Evening	4.240	4.287
3	2	SB	Morning	7.127	7.312
3	2	SB	Evening	8.320	7.772
2	1	SB	Morning	5.124	4.826
2	1	SB	Evening	5.273	6.147

Table C.3: Standard deviations of the travel times of the link-level field Bluetooth data

From	To	Direction	Time Period	Before Stdev (min)	After Stdev (min)
1	2	NB	Morning	2.569	2.781
1	2	NB	Evening	3.721	1.514
2	3	NB	Morning	1.460	1.714
2	3	NB	Evening	3.216	3.057
3	4	NB	Morning	1.565	1.921
3	4	NB	Evening	1.569	1.314
4	5	NB	Morning	2.284	2.179
4	5	NB	Evening	1.881	1.719
5	6	NB	Morning	0.910	0.874
5	6	NB	Evening	1.425	1.327
6	7	NB	Morning	1.204	1.313
6	7	NB	Evening	0.843	0.904
7	6	SB	Morning	0.982	1.000
7	6	SB	Evening	1.048	2.188
6	5	SB	Morning	2.255	1.357
6	5	SB	Evening	3.049	2.741
5	4	SB	Morning	2.444	1.770
5	4	SB	Evening	1.287	1.534
4	3	SB	Morning	0.914	1.066
4	3	SB	Evening	0.875	1.332
3	2	SB	Morning	2.212	2.590
3	2	SB	Evening	2.025	1.627
2	1	SB	Morning	2.427	2.368
2	1	SB	Evening	2.075	2.757

Table C.4: F-test and t-test results for the link-level field Bluetooth data

From	To	Direction	Time Period	Stdevs Different?	Means Different?
1	2	NB	Morning	N	Y
1	2	NB	Evening	Y	Y
2	3	NB	Morning	Y	Y
2	3	NB	Evening	N	Y
3	4	NB	Morning	Y	Y
3	4	NB	Evening	Y	Y
4	5	NB	Morning	N	Y
4	5	NB	Evening	Y	Y
5	6	NB	Morning	N	Y
5	6	NB	Evening	Y	Y
6	7	NB	Morning	Y	N
6	7	NB	Evening	Y	Y
7	6	SB	Morning	N	Y
7	6	SB	Evening	Y	Y
6	5	SB	Morning	Y	Y
6	5	SB	Evening	Y	Y
5	4	SB	Morning	Y	Y
5	4	SB	Evening	Y	N
4	3	SB	Morning	Y	Y
4	3	SB	Evening	Y	N
3	2	SB	Morning	Y	N
3	2	SB	Evening	Y	Y
2	1	SB	Morning	N	N
2	1	SB	Evening	Y	Y

Appendix D Chapter 5 Regression Analysis Raw Data

Table D.1: AADT map data for the field Bluetooth corridor

From	To	Link	24 Hour NB	24 Hour SB	AM Peak NB	AM Peak SB	PM Peak NB	PM Peak SB
Kingston	Gerrard	1	6624	4301	625	302	570	336
Gerrard	Danforth	1	10150	10856	753	729	787	903
Danforth	St Clair	1 and 2	15251	8144	737	514	1098	680
St Clair	Eglinton	2	14702	12359	977	748	1152	1016
Eglinton	Lawrence	2 and 3	17059	11639	1200	802	1323	891
Lawrence	Ellesmere	3 and 4	18710	17007	1265	1187	1436	1305
Ellesmere	401	4	18317	20200	1396	1522	1475	1654
401	Sheppard	4 and 5	17098	19065	1227	1229	1536	1978
Sheppard	Huntingwood	5	16089	12579	808	1297	1055	1304
Huntingwood	Finch	5 and 6	10978	14219	927	1137	969	1258
Finch	McNicoll	6	9363	10426	904	677	798	1211
McNicoll	Passmore	6 and beyond	10486	10386	726	736	1164	1236

Table D.2: Aggregated AADT data for the field Bluetooth corridor

Link	24 Hour NB	24 Hour SB	AM Peak NB	AM Peak SB	PM Peak NB	PM Peak SB
1	10675	7767	705	515	818	640
2	15671	10714	971	688	1191	862
3	17885	14323	1233	995	1380	1098
4	18042	18757	1296	1313	1482	1646
5	14722	15288	987	1221	1187	1513
6	10276	11677	852	850	977	1235

Table D.3: General data for the field Bluetooth corridor

Link	Length (km)	Signals	Unsignalized Intersections	Intersections	Posted Speed (km)	Minimum Lanes per Direction
1	1.8	7	8	15	60	2
2	4	9	29	38	60	2
3	2.4	6	8	14	60	2
4	2.5	8	7	15	60	2
5	2	6	3	9	50	1
6	2.5	6	4	10	50	1

Appendix E Chapter 6 Application of Measures of Effectiveness using Simulated Bluetooth Travel Time Data without Error and using Simulated True Travel Time Data

Period:	Morning Peak
Period Start:	7:30:00am
Period End:	10:30:00am

Southbound

Before Mean (s)	Mean Difference (s)	Volume	Length (km)	Mean Difference Significant?	
					3 - Eagle
220	2	153	1.9	0	|
					2 - Bishop
168	5	255	2.4	0	|
					1 - Todd

Total Volume 408

MOE 1 (s)	6.7	% Reduction 1	1.7
MOE 2 (s/km)	1.6	% Reduction 2	1.9
MOE 3 (s/km)	0.0	% Reduction 3	0.0

Northbound

Before Mean (s)	Mean Difference (s)	Volume	Length (km)	Mean Difference Significant?
170	10	235	1.9	1
226	7	255	2.4	0

Total Volume 490

MOE 1 (s)	17.1	% Reduction 1	4.3
MOE 2 (s/km)	4.1	% Reduction 2	4.4
MOE 3 (s/km)	2.6	% Reduction 3	2.8

Figure E.1: Measure of effectiveness diagram for the morning peak period for the simulated Bluetooth data without error

Period:	Evening Peak
Period Start:	3:00:00pm
Period End:	6:00:00pm

Southbound

Before Mean (s)	Mean Difference (s)	Volume	Length (km)	Mean Difference Significant?	
					3 - Eagle
542	136	309	1.9	1	|
					2 - Bishop
250	64	379	2.4	1	|
					1 - Todd

Total Volume 688

MOE 1 (s)	200.8	% Reduction 1	25.4
MOE 2 (s/km)	47.0	% Reduction 2	25.4
MOE 3 (s/km)	47.0	% Reduction 3	25.4

Northbound

Before Mean (s)	Mean Difference (s)	Volume	Length (km)	Mean Difference Significant?
205	11	195	1.9	0
229	7	167	2.4	0

Total Volume 362

MOE 1 (s)	18.6	% Reduction 1	4.3
MOE 2 (s/km)	4.6	% Reduction 2	4.5
MOE 3 (s/km)	0.0	% Reduction 3	0.0

Figure E.2: Measure of effectiveness diagram for the evening peak period for the simulated Bluetooth data without error

Period:	Morning Peak
Period Start:	7:30:00am
Period End:	10:30:00am

Southbound

Before Mean (s)	Mean Difference (s)	Volume	Length (km)	Mean Difference Significant?
227	7	2280	1.9	1
168	5	2575	2.4	1

3 - Eagle | 2 - Bishop | 1 - Todd

Northbound

Before Mean (s)	Mean Difference (s)	Volume	Length (km)	Mean Difference Significant?
165	2	2245	1.9	0
225	4	2545	2.4	1

Southbound Total Volume 4855

MOE 1 (s)	11.8	% Reduction 1	3.0
MOE 2 (s/km)	2.8	% Reduction 2	3.0
MOE 3 (s/km)	2.8	% Reduction 3	3.0

Northbound Total Volume 4790

MOE 1 (s)	6.6	% Reduction 1	1.7
MOE 2 (s/km)	1.5	% Reduction 2	1.7
MOE 3 (s/km)	1.0	% Reduction 3	1.1

Figure E.3: Measure of effectiveness diagram for the morning peak period for the simulated true travel time data

Period:	Evening Peak
Period Start:	3:00:00pm
Period End:	6:00:00pm

Southbound

Before Mean (s)	Mean Difference (s)	Volume	Length (km)	Mean Difference Significant?
544	116	4103	1.9	1
238	49	3751	2.4	1

3 - Eagle | 2 - Bishop | 1 - Todd

Northbound

Before Mean (s)	Mean Difference (s)	Volume	Length (km)	Mean Difference Significant?
202	5	2010	1.9	1
232	5	1710	2.4	1

Southbound Total Volume 7854

MOE 1 (s)	164.6	% Reduction 1	21.0
MOE 2 (s/km)	41.6	% Reduction 2	21.1
MOE 3 (s/km)	41.6	% Reduction 3	21.1

Northbound Total Volume 3720

MOE 1 (s)	10.1	% Reduction 1	2.3
MOE 2 (s/km)	2.4	% Reduction 2	2.4
MOE 3 (s/km)	2.4	% Reduction 3	2.4

Figure E.4: Measure of effectiveness diagram for the evening peak period for the simulated true travel time data