USDOT Region V Regional University Transportation Center Final Report

Analysis of Travel Time Reliability on Indiana Interstates

By

Maria Martchouk
Graduate Student in Civil Engineering
Purdue University
mmartcho@purdue.edu

and

Fred Mannering, Principal Investigator
Charles Pankow Professor of Civil Engineering
Purdue University
flm@purdue.edu

DISCLAIMER

USDOT Region V Regional University Transportation Center Final Report

TECHNICAL SUMMARY

Analysis of travel time reliability on Indiana interstates

Introduction

Travel-time reliability is a key performance measure in any transportation system. It is a measure of quality of travel time experienced by transportation system users and reflects the efficiency of the transportation system to serve citizens, businesses and visitors. Travel-time reliability (the variability in travel times on the same route at the same time from one day to the next) is critical to travelers, shippers, receivers and carriers for trip decisions and on-time arrivals at destinations. Thus, understanding travel time reliability is important to travelers in order for them to plan their trips effectively as well as shippers for them to plan and select routes appropriately.

The first objective of this study is to formulate a methodology to obtain travel time data using Bluetooth technology on a freeway segment and collect travel time data. Bluetooth technology enables to collect real travel time data with a high sampling rate of up to 10% of the traffic flow. It also eliminates that need to use complex and often inaccurate algorithms use to calculate travel time from point speed data. Another benefit of this technology is it's the fact that it is relatively inexpensive to implement; every station where travel time is desirable needs to be equipped with a processing unit, power source and Bluetooth dongle.

The second objective is to observe daily and inter-daily variations as well as those due to poor weather conditions and estimate econometric models to predict travel time and variability. Within the context of this study, travel times were collected for two sections of freeway that experience heavy congestion during the peak hours. These travel times were then used to estimate three econometric models that predict travel time as a function of traffic flow parameters including speed and volume. The first model, which is linear regression with lagged dependent variable terms, aims to predict individual travel times during all times of day. The second model, a survival model, seeks to evaluate the probability of the trip lasting any specified length of time. In addition, it can predict the probability of exiting the freeway segment given that the vehicle has been traversing the segment up to that time. The third model seeks to describe travel time and the variability of travel time using the seemingly unrelated regression equations.

Findings

The study demonstrates how actual travel time data can be collected and processed using Bluetooth technology along two segments of a busy Interstate 69 in Indianapolis. Travel times were

collected for a period of two weeks and analyzed. Using these data, the inter-vehicle, inter-period and inter-day variability is evaluated. From the collected data, it is clear that at any point during the day, the travel time may vary by up to 100% of the average travel time. We speculate that this is due to individual driver behavior and different traffic speeds in adjacent lanes. This is a significant variation that needs to be taken into account when providing travel information to travelers. The inter-period variability is also pronounced in the collected data. The average travel time is observed to triple during the peak hours (as one might expect) and the standard deviation of individual travel times also increased during peak periods. And, despite a common conclusion that travel times during the same day of the week are fairly consistent, a comparison between two of the same days of the week indicates that it varies significantly. This variability is particularly pronounced during the peak hours.

The study also considers variability due to adverse weather. Travel time data collected during a snow event are compared to data obtained during adequate weather conditions. The analysis shows that both travel time and standard deviation of travel time are significantly different under adverse and normal weather conditions. In addition, the effect of the snow event can last for many hours thereafter due to the roadway not being cleared. This presents another difficulty in modeling the traffic following a snow event because the conditions on the road may depend on the response time of snow clearing personnel.

The study also demonstrates the application of three statistical models of the travel-time data. An autoregressive individual vehicle travel time model is estimated to determine factors that may generate variability. This model includes data obtained from Remote Traffic Microwave Sensors, including speed and volume as well as time of day information. It is found that average segment travel time, lagged travel time, volume and peak hour indicator all positively affect individual vehicle travel time. Meanwhile the night time parameter has a negative impact on travel time.

Next, duration models of travel time were estimated to provide insight as to when the variance of travel times become a problem. Duration models are useful in predicting how the probability of a car's duration of time on a roadway segment changes over time. Three duration models are estimated: all hours, peak hour and night time models. These models show that the point where the conditional probability of travel times becoming longer occurs roughly at the onset of level-of-service F conditions.

The third model predicts average travel time (for 15-minute intervals) as well as the inherent variability in this travel time over the 15-minute interval. This model can be used to predict traffic conditions in real time and convey the information to the roadway users via Variable Message Signs and web utilities.

Recommendations

Overall the study provides a framework for collecting and analyzing travel time data using Bluetooth technology and supplementing it with corresponding remote traffic microwave sensors data. In addition, the three statistical modeling alternatives used to establish a relationship between travel times and known traffic parameters show that there is great potential to be able to predict, with reasonable accuracy, mean travel times and travel time variability. Applying the methods and techniques demonstrated in this study to other freeway roadway segments in Indiana and throughout the US should provide even more useful insights into this important problem. Finally, it is recommended

that the third model (seemingly unrelated regressions model) be used as a basis (immediately) for providing real-time traffic information to travelers in the Indianapolis area.

Contacts

For more information:

Fred Mannering
Principal Investigator
Civil Engineering
Purdue University
flm@purdue.edu

NEXTRANS Center
Purdue University - Discovery Park
2700 Kent B-100
West Lafayette, IN 47906

nextrans@purdue.edu
(765) 496-9729
(765) 807-3123 Fax

www.purdue.edu/dp/nextrans

USDOT Region V Regional University Transportation Center Final Report

Analysis of Travel Time Reliability on Indiana Interstates

By

Maria Martchouk
Graduate Student in Civil Engineering
Purdue University
mmartcho@purdue.edu

and

Fred Mannering, Principal Investigator
Charles Pankow Professor of Civil Engineering
Purdue University
flm@purdue.edu

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Travel time and travel time reliability are important performance measures for assessing traffic condition and extent of congestion on a roadway. The study describes the use of Bluetooth technology to collect real travel time data on a freeway and applies it to obtain two weeks of data on Interstate-69 in Indianapolis. An autoregressive model, estimated based on the collected data, is then proposed to predict individual vehicle travel times on a freeway segment. This model includes speed, volume, time of day indicators, and previous vehicle travel times as independent variables. In addition to the autoregressive formulation, a duration model is estimated based on the same travel time data. The duration model of travel time provided insights into how one could predict the probability of a car's duration of time on a roadway segment changed over time. Interestingly, the three duration models estimated (all hours, peak hour and night time models) showed that the point where the conditional probability of travel times becoming longer occurs roughly at the onset of level-of-service F conditions. Finally, a seemingly unrelated regression equation model to predict travel time and travel-time variability is estimated. This model predicts 15-minute interval travel times and standard deviation of travel time based on speed, volume and time of day indicators. The model has a very good statistical fit and thus can be used in the field to compute real-time travel time using data available from remote traffic microwave sensors.

CHAPTER 1. INTRODUCTION

Congestion, delay and unreliability are terms that most associate with present-day travel. As the annual time spent in traffic jams rose from 16 hours per person in 1982 to 42 hours per person in 2003 for the largest 85 urban areas, the impact on cost of travel has been tremendous (FHWA, 2005). Both individuals and companies suffer from economic losses due to lost time as well as additional vehicle operating and detour costs. For this reason, travel time has been a critical measure used to evaluate conditions on the road and assess improvement projects (Mackie, 2001). Despite it being a popular statistic, until now travel time has been difficult to measure. For economic analyses the travel time is most commonly deduced from free-flow travel time. Other ways in which travel time is calculated is from floating car studies or by using speed obtained from loop detectors.

While travel time has always been the most important measure associated with effectiveness of the transportation system, recently a single measure of travel time has become insufficient. Travel times can vary during the day or between days. During peak hours, the travel time can increase significantly as compared to other time periods. In addition, heavily traveled corridors at or near capacity suffer from large travel time variation due to even small perturbations in traffic. An average travel time cannot capture these changes and a reliability measure is necessary.

Reliability of travel time has often been used interchangeably with variability and can refer to changes in travel time during the course of a day or changes from day to day. Seven sources of variability in travel times have been identified: incidents, work zones, weather, demand fluctuations, special events, traffic control devices and inadequate base capacity (FHWA, 2005).

Understanding travel time reliability is important to travelers in order for them to plan their trips effectively as well as shippers for them to plan and select routes

appropriately. Furthermore, reliability needs to be taken into account during project assessment in order to capture all the benefits of an improvement project.

The first objective of this study is to formulate a methodology to obtain travel time data using Bluetooth technology on a freeway segment and collect travel time data. Bluetooth technology enables to collect real travel time data with a high sampling rate of up to 10% of the traffic flow. It also eliminates that need to use complex and often inaccurate algorithms use to calculate travel time from point speed data. Another benefit of this technology is it's the fact that it is relatively inexpensive to implement; every station where travel time is desirable needs to be equipped with a processing unit, power source and Bluetooth dongle.

The second objective is to observe daily and inter-daily variations as well as those due to poor weather conditions and estimate econometric models to predict travel time and variability. Within the context of this study, travel times were collected for two sections of freeway that experience heavy congestion during the peak hours. These travel times were then used to estimate three econometric models that predict travel time as a function of traffic flow parameters including speed and volume. The first model, which is linear regression with lagged dependent variable terms, aims to predict individual travel times during all times of day. The second model, a survival model, evaluates the probability of the trip lasting any specified length of time. In addition, it predicts the probability of exiting the segment given that the vehicle has been traversing it up to that time. The third model describes travel time and standard deviation of travel time using the seemingly unrelated regression equations. This third model can be readily implemented to provide expected travel times as well as expected travel time deviations.

CHAPTER 2. LITERATURE REVIEW

For the past few decades much of the transportation research has been dedicated to development of performance measures that would enable congestions evaluation. Most agencies have focused on volume to capacity ratios and level of service measurements to describe the travel conditions (Highway Capacity Manual, 2000). These measurements compare the facility capacity to its utilization on a localized scale. On a larger spatial scale, hours of delay and vehicle miles traveled are used to assess the conditions. While effective at assessing the road's performance in comparison to the rest of the network, these measures do not easily translate well to the user experience. Recently a trend has emerged to develop performance indicators that express congestion and mobility in terms of something the system users can understand and appreciate. Two such measures include travel time and reliability (FHWA, 2005).

Travel time is a main factor in drivers' route choice decisions and its value has been studied for many decades. In more recent years, researchers hypothesized that the value of travel time models may have been omitting important reliability considerations and a model that includes both reliability as well as travel time should be used. Value of reliability has been explored both using the stated preference and revealed preference approach. Abdel-Aty et al. (1995) conducted a phone survey and formulated a binary logit model to determine the importance of reliability in route choice decision. He found that travel time reliability was equally important as travel time. Liu et al. (2004) used the revealed preference approach to estimate the value of travel time and reliability on SR 91. They estimated a mixed logit model by finding coefficients of travel cost, travel time and reliability that minimized the differences between the predicted and observed traffic flows. The authors concluded that the value of travel time was $12.81/hour, while the value of reliability was $20.63/hour. From these and other studies it is clear that reliability is important to the roadway users.

While there's pervasive evidence that both travel time and reliability are important considerations in travel decisions, these parameters are not easy to measure and predict.

2.1   <u>Methods of Estimating Travel Time</u>

Travel time evaluation methods consist of direct measurements and estimations. The most popular method to measure travel time directly has been the floating car technique (Robertson, 1994). This method uses a probe vehicle traveling with the traffic flow. Travel time and location are recorded along the route. A similar technique uses a GPS device to record the vehicle trajectory along with time stamps. This method provides more frequent sampling along the route and avoids human error. It also allows for viewing of the results using a GIS interface (Quiroga & Bullock, 1998). However, both methods only provide travel time for the probe vehicle, making it difficult to collect a large dataset for various routes and times of day.

Travel time estimation methods vary from simple estimation of average free flow travel time based on posted speed to complex algorithms based on combination of real-time 30-second dual loop detector data and historical data. Some of the most widely used estimation methods are based on dual-loop detector data. One method to evaluate the highway segment's travel time is based on an upstream and downstream speed (also referred to as trajectory method). The simplest way to estimate travel time over a segment is by dividing the segment distance by the average travel speed at the upstream and downstream detectors. This method is particularly appealing due to its simplicity and is still frequently used for real-time prediction (Tufte & Kothuri, 2007). Alternatively, instead of assuming that the vehicle traverses half the segment at the downstream speed and half the segment at the upstream speed, Cortes et al. (2002) proposed using a linear combination of speeds at upstream detector when the vehicle enters the segment and speed at downstream detector when the vehicle exits the segment. Initially the two parameters are set to 0.5 and are then calibrated for the given network. While this method allows one to capture network specific traffic effects, it requires travel time data for calibration. Another trajectory method used to estimate the travel time was proposed by Lint and is called piecewise linear speed based (PCSB) (Lint & Zijpp, 2007). Unlike the previous method discussed, the travel speeds over the segment are not assumed to be constant but rather have a linear shape. This allows the speeds to be continuous as section ends and models drivers adjusting their speed gradually in response to slower speeds ahead of them. Dual loop detector data can also be used to estimate the probability density function of segment travel times (Fan & Chen, 2009). The authors treat vehicle occurrence at the upstream and downstream station as time series and

calculate the auto-correlations and cross-correlations between the two time series. The probability density function of travel times is then derived using the Fredholm integral equation.

Another technology that measures travel time directly is the Automated Vehicle Identification (AVI) system consisting of an in-vehicle transponder and a roadside unit that receives the signal (Turner, 1996). This technology is most often integrated with tolling infrastructure. While this method provides good real-time travel time data it is expensive to implement. Furthermore, there are privacy concerns because transponders can be traced back to the vehicle.

2.2   Travel Time Reliability

Travel time reliability has been a subject of many studies. It is commonly accepted to refer to the level of consistency in transportation service for a mode, trip, route or corridor for a time period (Lomax et al., 2003). In general travel time variability has been classified into recurrent and non-recurrent, where recurrent variability is a result of insufficient capacity, while the non-recurrent variability is caused by transient events. Sources of non-recurrent congestion include accidents, inclement weather, construction and special events. Separating the causes for travel time variability is important in assessing the benefits of improvement projects and co-coordinating efforts to improve reliability (Bremmer et al., 2004).

Instead of classification by its source, variability can be categorized by its time frame. Bates et al. (1989) and Small et al. (1999) discussed variability as being inter-day, inter-period and inter-vehicle. Inter-day or day-to-day variability is caused by unexpected events such as construction or inclement weather. Inter-period or daily variation generally refers to the changes in travel time due to peak hour congestion. Meanwhile inter-vehicle variability is a result of individual driver behavior including lane changes and speed.

The existence of inter-period and inter-day variability has been long accepted. These are believed to arise due to variations in traffic resulting in delays the facility is at or near capacity. However, Mannering et al. (1990) empirically studied that the inter-vehicle variability that exists, which contradicts the assumption often made that everyone travels at the same speed within a given traffic flow. They hypothesized that drivers selected their speeds according to the equation,

$$SPEED_{kj} = \sigma ESP_{kj} + [1\text{-}q_{kj}/c_j]\, \delta SES_k + v_{kj} \qquad\qquad \text{Eq. 2.1}$$

where $SPEED_{kj}$ is the average speed of traveler k on route j, ESP_{kj} is the expected speed of traveler k on route j, q_{kj}/c_j is the volume to capacity ratio of route j, SES_k is a vector of socioeconomic characteristics of traveler j, v_{kj} is the disturbance term, σ and δ are estimable parameters.

Upon estimation the coefficient of δ was found to be significant, implying that the individual's speed varies significantly from the expected traffic speed, even at higher volumes.

2.3 Travel Time Reliability Measures

Regardless of the source or the type of variability, most common measures include various indices suggested by the FHWA (2005). Indices of reliability are commonly divided into statistical, buffer measures and tardy trip indicators. Statistical methods include travel time window and percent variation shown in equations 2.1 and 2.2. Both measures focus on estimating standard deviation of travel times and comparing it to the average travel time.

$$\textbf{Travel Window} = \textbf{Average Travel Time} \pm \textbf{Standard Deviation} \qquad \text{Eq. 2.2}$$

$$\text{Percent Variation} = \frac{\text{Standard Deviation}}{\text{Mean}} \, \text{x100\%}$$

$$\text{Eq. 2.3}$$

These statistical measures are effective in communicating the extent of unreliability to professionals, however, may not be meaningful to users because it is difficult for individuals to apply the concept of standard deviation to their individual travel time. They are also unable to capture variation due to different events separately thus providing a very general measure of reliability for the roadway.

The second category of methods is buffer measures, of which the most common is the buffer index. Buffer Index (BI) is a measure of trip reliability that expresses the amount of extra buffer time needed to be on time for 95 percent of the trips. This measure allows the traveler to estimate the extra percent of travel time that the trip may take due to varying congestion level.

$$\text{Buffer Index} = \left[\frac{\begin{array}{cc} \text{95th percentile} & \text{Average} \\ \text{Travel Time} & \text{Travel Time} \end{array}}{\begin{array}{c} \text{Average} \\ \text{Travel Time} \end{array}} \right] \text{x100\%}$$

Eq. 2.4

In addition to the buffer index, the planning time index can also be used, which estimates the extent by which the free-flow travel time will be exceeded.

$$\text{Planning Time Index} = \frac{\begin{array}{c} \text{95th percentile} \\ \text{Travel Time} \end{array}}{\begin{array}{c} \text{Travel Time Based} \\ \text{on Free - Flow Speed} \end{array}}$$

Eq. 2.5

The buffer index and planning time index are measures that most users can relate to because when planning a trip one would like to arrive on time in a vast majority of situations. The 95th percentile travel time ensures the user is only late 1 out of every 20 trips. The buffer measures can be used to calculate a single value of reliability for the road segment or different values that depend on time of day and day of the week. Bremmer et al. (2004) calculated the 95th percentile travel time for 12 commuter routes in Puget Sound by time of departure and provided a web-based tool for users to retrieve this information.

Tardy trip indicators, which include percent of unreliable trips and misery index are the third way to evaluate the variability in the travel time. Percent of unreliable trips is simply evaluated as the percent of trips with higher than acceptable travel times. The misery index is calculated as the average travel time subtracted from travel time from the top 20% of trips divided by average travel time.

% onTime=Percent Trip Times<[1.1*Mean Time]   Eq. 2.6

$$\text{Misery Index} = \frac{\begin{array}{cc} \text{Average travel time for} & \text{Average Travel Time} \\ \text{the longest 20\% of trips} & \end{array}}{\text{Average Travel Time}}$$

Eq. 2.7

In addition to the statistical methods of estimating travel time reliability, Elefteriadou (2006) proposed econometric modeling. She developed linear regression models to estimate average travel time for scenarios with different combinations of weather, accidents, congestion and work zones. Furthermore, the researchers determined probability of a reliable trip under various definitions of reliability. This study was unable to provide models for all scenarios due to the lack of data and did not clarify how the travel times for modeling were obtained.

2.4 Summary

All statistical methods of estimating travel time reliability rely on the knowledge of travel time. As discussed travel time measurement methods have been limited due to small sample as with floating car technique or require extensive infrastructure investment in the case of AVI. Meanwhile indirect methods of obtaining travel time rely only on travel speed, without the ability to take into account volume or additional congestion effects. The current study uses a new method of collecting travel times that is not cost prohibitive and allows one to capture up to 10% of the actual travel times of the traffic stream.

Travel time reliability is most commonly evaluated using statistical measures, buffer indices or tardy trip indicators. Statistical measures are useful in evaluating and comparing the performance of a roadway, however, may not be easy for users to understand. Buffer measures and tardy trip indicators can help users to plan their trips better; however, these can only capture historical trends and do not take into account current travel conditions. Econometric modeling enables to predict travel times based on current travel conditions. The set of econometric models proposed by Elefteriadou to predict travel time based on weather, incident, congestion and work zones are limiting because not all combinations of factors were available for model estimation. In addition, they were not estimated from actual travel time data but rather from calculated data. The econometric models formulated in this study are based on real travel time data. As will be shown, there is a single model instead of many different ones that can be applied in various circumstances.

CHAPTER 3. DATA

In order to evaluate travel time reliability, there is a need for accurate travel time data. As discussed in the literature review, in the past travel time was deduced from loop detector data, historical trends or floating car runs. This study, however, employed new Bluetooth technology to collect real travel time data from a freeway corridor in Indianapolis. In addition to travel time data, speed and volume data were obtained from Remote Traffic Microwave Sensor (RTMS) detectors located on the same segment of the freeway.

3.1 Travel Time Data

Collecting travel time data using Bluetooth technology is a new concept. This method was developed and first utilized by the Indiana Department of Transportation in collaboration with Professor Darcy Bullock (Wasson, 2008). It relies on capturing travelers in possession of Bluetooth-enabled devices that broadcast unique identifiers known as Media Access Control (MAC) addresses. MAC addresses are recorded at various stations and if the same MAC address is recorded at an upstream station and a downstream station, the travel time can be obtained by subtracting the time stamp at the downstream station from the upstream one. The study conducted by Wasson (2008) observed travel times on two corridors in Indianapolis, IN. The first corridor included a 5.8-mile section of freeway, while the second was a combination of arterial and freeway roads. This study paved the way for collecting real travel time data with an unprecedented sample rate.

The device that records MAC address and time stamp data can be a portable case or a laptop. The case includes a processing unit, battery and a Bluetooth dongle. The

cases deployed in this study included an AIRcable Server XR, 12 V car battery and power converter. In addition, two of the three deployed cases included a Zoom 4310BF USB Wireless Bluetooth Dongle and an extension tube. The extension arm allows to elevate the Bluetooth receiver in order to increase the number of recorded MAC addresses. The third case included the AIRcable Host XR receiver and did not require the extension arm because it was located on an overpass.

For the purposes of the current study, a segment of I-69 was selected between 96th Street and 116th Street. This corridor was selected because it is a heavily traveled roadway equipped with RTMS detectors that provide speed, volume and occupancy data. The first selected segment extends from Mile Marker (MM) 2.9 to Mile Marker 3.7. The second segment runs from Mile Marker 3.7 to Mile Marker 5.0.

The selected corridor, located in Northeast Indianapolis, experiences AADT of 95,690. Trucks with more than 2 axles comprise 20% of the vehicles. The first segment has four lanes traveling in both the Northbound and Southbound directions. The second segment has three lanes in both directions and includes an off-ramp in the Northbound direction.

The three locations where the Bluetooth cases were placed are shown in Figure 3.1.

After the cases were deployed, the batteries were replaced every 60 hours. Once the necessary data were collected, they were downloaded from the Server XR and unnecessary information was deleted. The resulting tables from the three detectors, which included MAC addresses and corresponding time stamps, were imported into MS Access. Since the Bluetooth receiver detects MAC addresses many times per second each vehicle may have multiple consecutive entries in the file. In the next step, it is necessary to eliminate the repeat detection entries. In the simplest algorithm, only the first detection time each day would be retained and thus there would only be one entry per vehicle per day. If the vehicle passes the station more than once a day the second time would be omitted. In this study many vehicles are expected to use the route to travel from home to work and from work home. Thus, the day is split into two periods, from 3 am until 3 pm and 3 pm until 3 am and the algorithm is created that retains the first time the vehicle is

detected at each station during each of these periods. In the last step, the MAC addresses at different detector stations are matched and travel times are calculated.

Figure 3.1 Bluetooth Detector Locations (MM= Mile Marker)

Because one of the case locations was on 116th street overpass, some of the recorded vehicles may have exited the freeway and were recorded on an overpass. These vehicles' travel times would be longer and are not representative of travel times on the freeway segment. In addition, some vehicles may have stopped on the shoulder or were traveling slow due to vehicle needing repair. Lastly, the vehicle may have been recorded at the upstream station, missed at the downstream station and then recorded at the downstream station later in the day when it was traveling in the opposite direction. All the aforementioned vehicle travel times needed to be removed prior to modeling because they are not representative of the travel time on the segment. A moving average of travel

times of preceding 5 vehicles and following 5 vehicles was determined for each vehicle. Then, travel times exceeding twice the moving average were excluded from the dataset.

For the purposes of this study, travel times were collected at these locations starting at 15:00 on Monday, November 10, 2009 through Friday, November 14, 2009 and Monday, November 17 through Thursday, November 19. The plot of individual vehicle travel times for road segment 1 located between Mile Marker 2.9 and Mile Marker 3.7 is shown in Figure 3.2 and Figure 3.3. Figure 3.4 and Figure 3.5 show the travel times for segment 2 located between Mile Marker 3.7 and Mile Marker 5.0. The Northbound travel times are plotted as positive values, while the southbound travel times are plotted as negative values.

Figures 3.2 through 3.5 show a relatively constant travel time for both segments during off-peak hours, indicating that the roadway experiences free-flow conditions the majority of the day. For the first segment the free flow travel time is 0.8 minutes in both Northbound and Southbound directions. For the second segment, the travel time is 1.1 minutes for both directions. During the pm peak, the congestion in the Northbound lanes results in tripling of the free-flow travel time. Meanwhile, the am peak affects travelers in the Southbound direction more than doubling their free flow travel time.
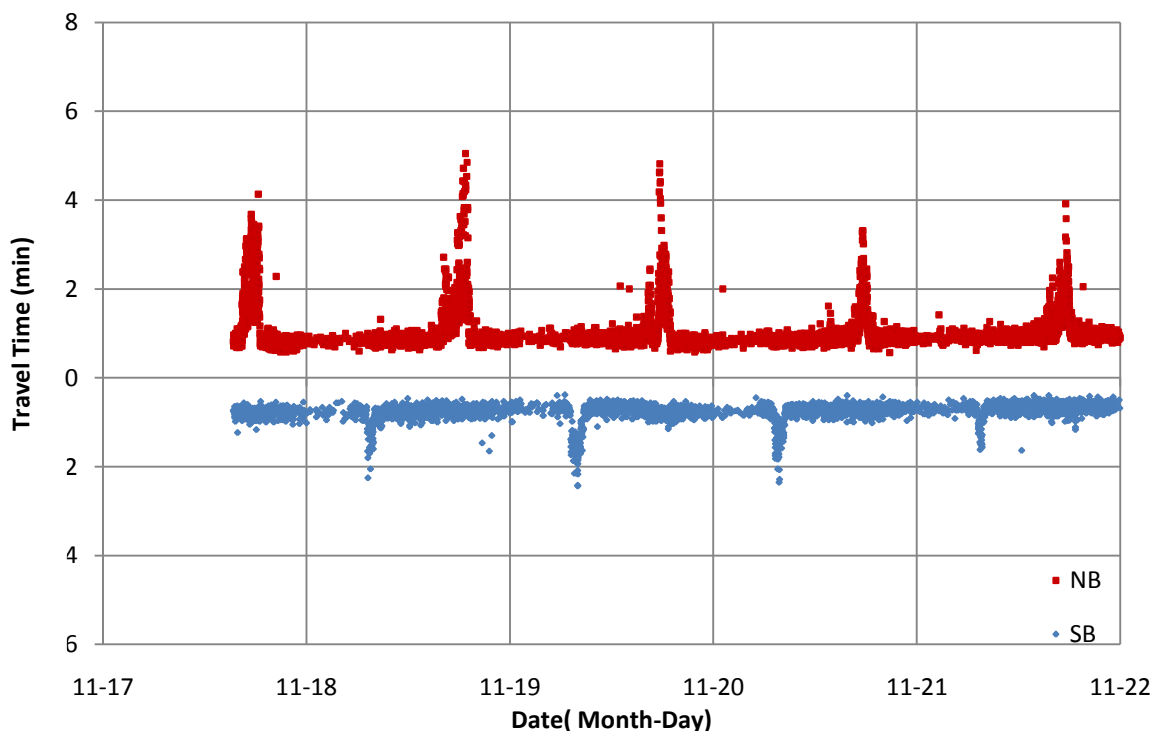
Figure 3.2- Travel Time between Mile Marker 2.9 and Mile Marker 3.7 from 11/17 through 11/22 (NB=Northbound, SB=Southbound)
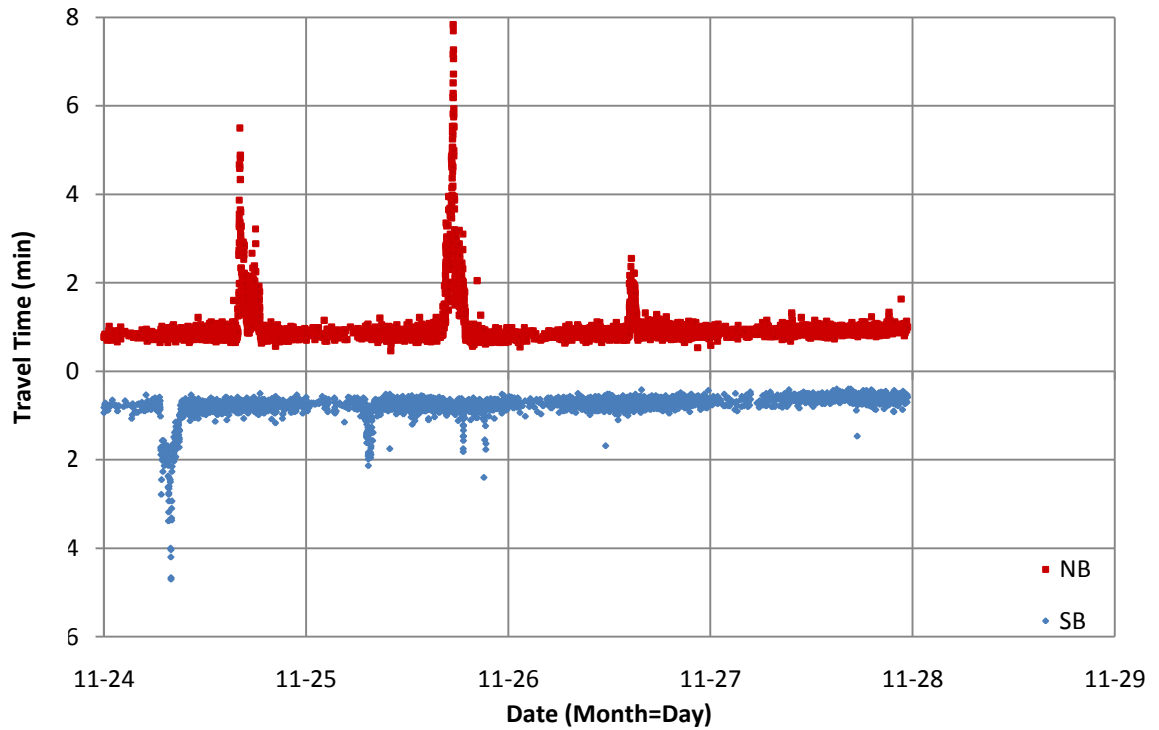
Figure 3.3 Travel Time between Mile Marker 2.9 and Mile Marker 3.7 from 11/24 through 11/28 (NB=Northbound, SB=Southbound)
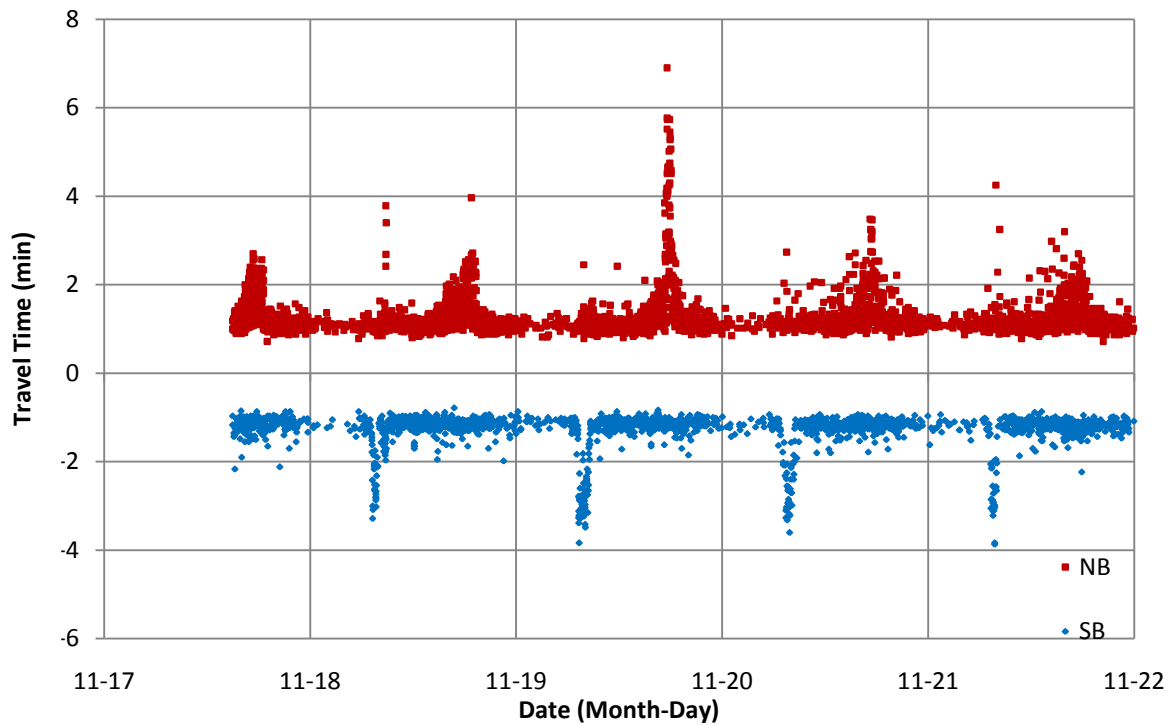
Figure 3.4 Travel Time between Mile Marker 3.7 and Mile Marker 5.0 from 11/17 through 11/22 (NB=Northbound, SB=Southbound)
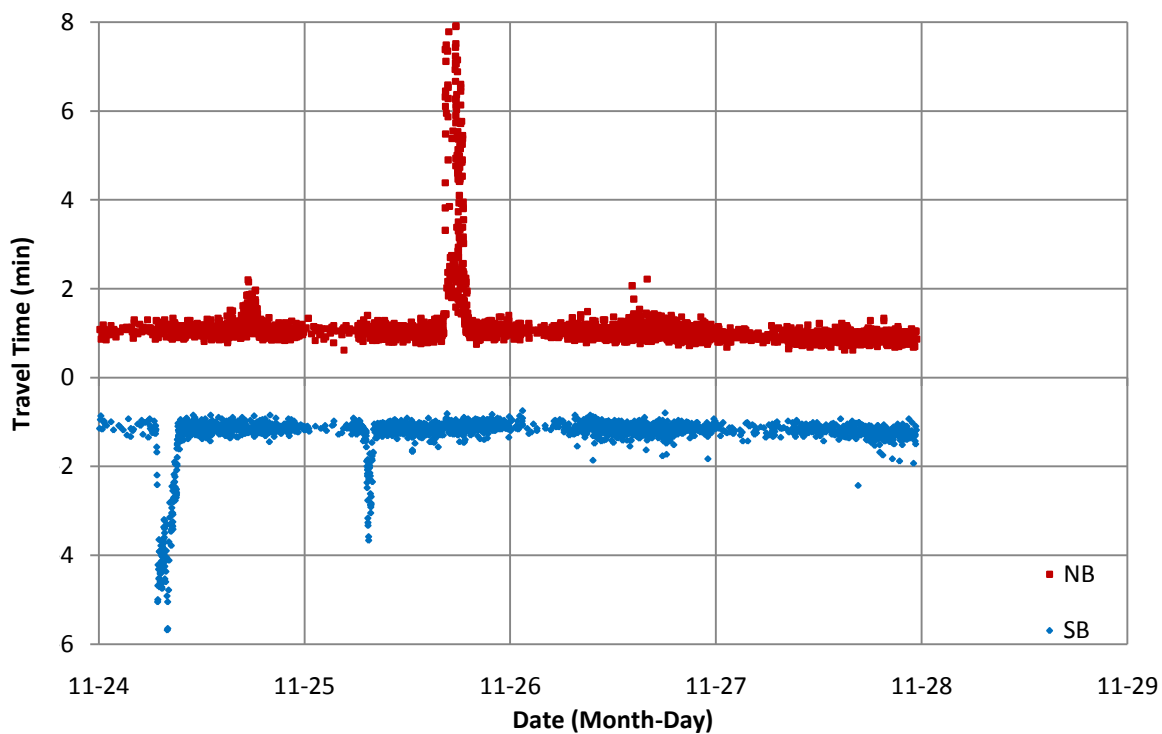
Figure 3.5 Travel Time between Mile Marker 3.7 and Mile Marker 5.0 from 11/24 through 11/28 (NB=Northbound, SB=Southbound)

3.2 Traffic Data

In order to evaluate the relationship between travel time and traffic characteristics, 30-second volume and speed data were obtained from the RTMS detectors located along the same segment of I-69. The map with locations of detectors is shown in Figure 3.6.

Figure 3.6 Locations of RTMS detectors with ID numbers

Because speeds and volumes were only available as point measures instead of segment measures such as travel times, it was necessary to calculate average speeds and volumes for each segment between bluetooth detectors.

In order to calculate the average speed over each segment, distances between detectors and bluetooth cases were determined as shown in Figure 3.7.
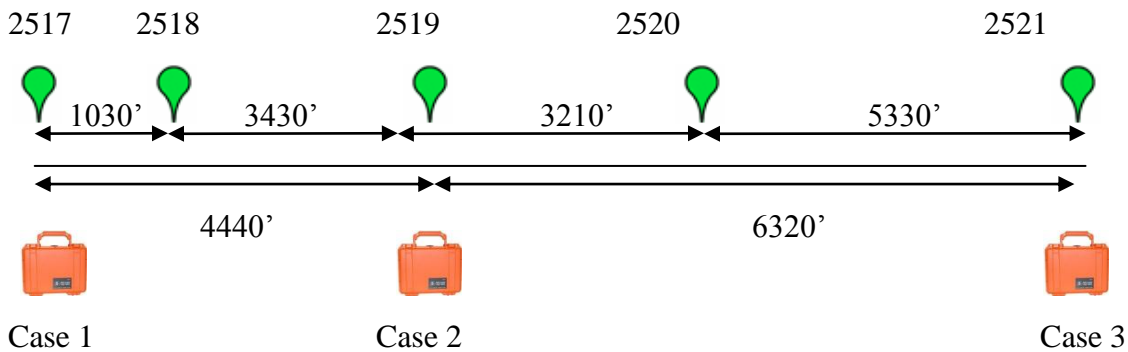
Figure 3.7 Details of Bluetooth Case Locations

The first step in calculating the average segment speed for each vehicle was to determine the 30-second speed at the upstream loop detector at the time that the vehicle departed the corresponding station. For example, the speed at detector 2517 was recorded at the time that the vehicle passed by case 1. Next, the time when the vehicle would reach the next loop detector was determined assuming constant travel rate and the 30-second speed from this loop detector was recorded. This procedure was repeated until the vehicle reached the end of the segment.

In order to determine average segment speed an assumption was made regarding the vehicle speed profile between two point speed estimates. It was assumed that the vehicle travels at the speed recorded at the upstream detector until it is halfway to the next detector at which point in time the vehicle travels at the speed of the downstream detector. This is shown in Figure 3.8.
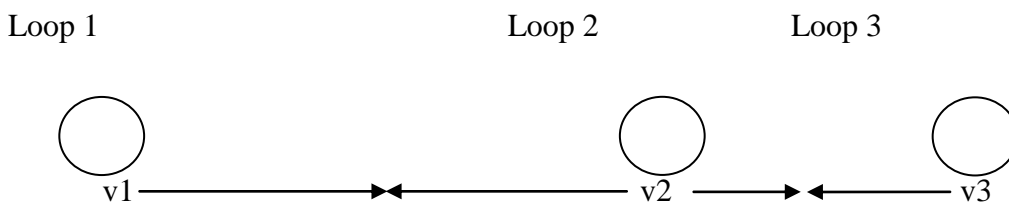
Figure 3.8 Vehicle Speed Profile

Average volumes were calculated in a similar fashion. The assumption this time was that the vehicle would be encountering the volume recorded at upstream station until

it reached the halfway point and then the volume would change to the downstream station.

The resulting database contained vehicle travel time over a segment, date and time of departure, average segment speed, and average segment volume. A plot of travel times, speeds and volumes for Wednesday November 19 is shown in Figure 3.9..

It is evident that individual vehicle travel time has a strong correlation with average travel speed; when the speed decreases the travel time rises. During off-peak hours both speed and travel time are at their free flow values. Meanwhile, during peak hours the speed drops to less than half of the free flow speed and the travel time rises up to 4 times the free flow value. One period when the travel time and speed measurements are not in agreement is at night. While the observed average speed is lower during night time than the free flow speed, the travel time remains close to its free flow value. This is explained by the fact that during night time there are fewer vehicles that are contributing to the average segment speed. Therefore, a single vehicle traveling at a significantly lower speed than the rest of the traffic stream can reduce the average speed. However, the travel time that is measures is individual vehicle travel time. Thus, if this specific vehicle is not recorded by the bluetooth detectors, there will not be any observations with a low travel time value.

Unlike the correlation between speed and travel time, there's a weaker relationship between travel time and volume. During night time the volumes are dramatically lower than during the day, however, the travel times are not different other than in the case of peak hour. During the peak hour, the volumes increase while the travel time decreases. However, the peak volumes and travel times occur at different times of the day. It was observed that the highest volumes on the freeway occur between 15:00 and 18:00, while the highest travel times and lowest speeds are observed between 16:00 and 19:00. This is explained by the traffic flow theory. The volumes are at their highest at the beginning of the congestion onset, when the speed has not decreased significantly. As the traffic enters the congested state, the speeds drop and fewer vehicles are able to traverse the segment resulting in lower volumes. Thus, in the forthcoming statistical models, delayed volume will be used.
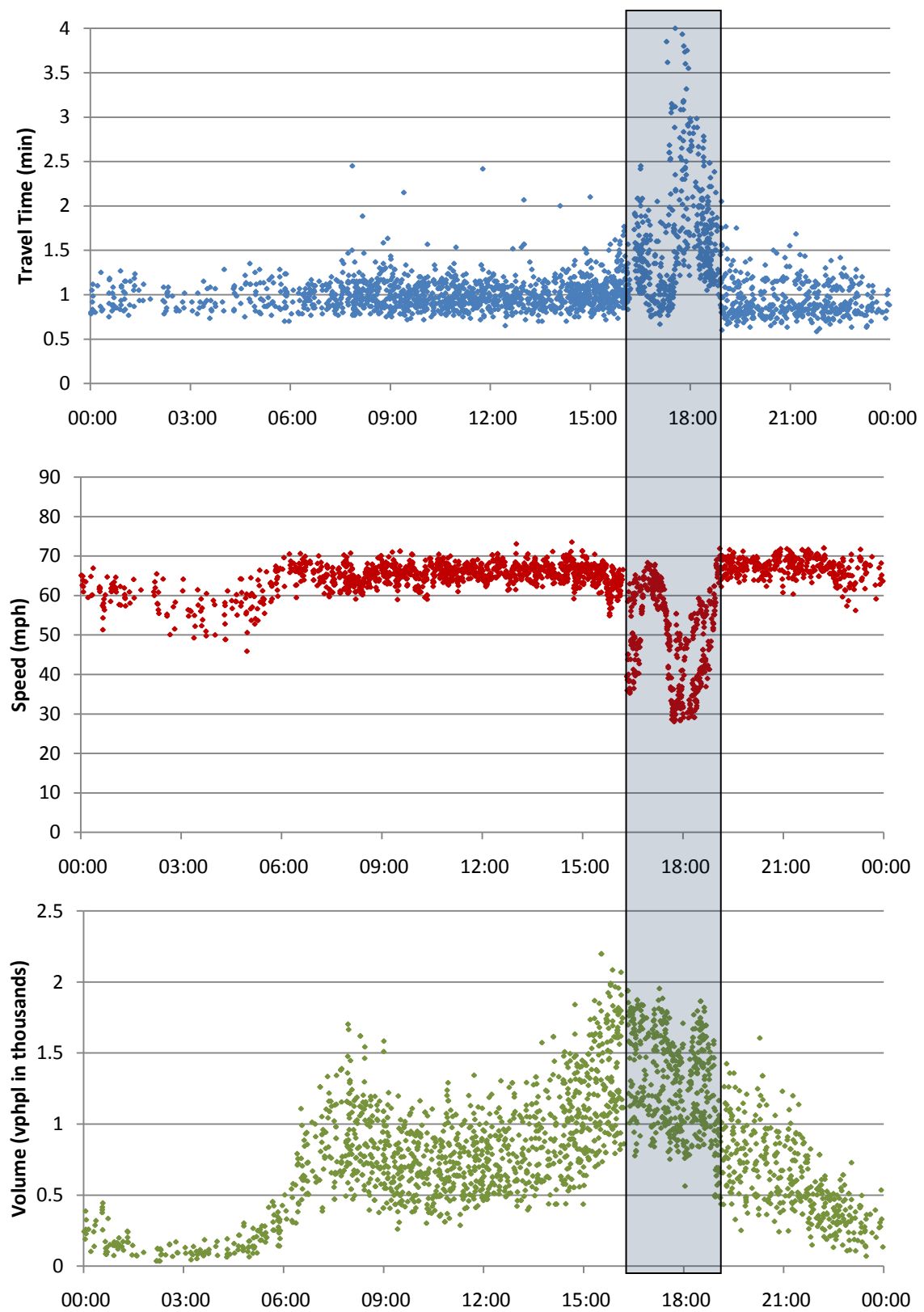
Figure 3.9 Plot of Travel Times, Speeds and Volumes

CHAPTER 4. Travel Time Variability

As mentioned in Chapter 2, travel time variability arises due to both recurrent and non-recurrent sources. Non-recurrent sources, include incidents, weather and special events. Recurrent congestion, meanwhile, is a result of insufficient capacity (Bremmer, 2004). The following analysis focuses on evaluating recurrent congestion on an inter-day, inter-period and inter-vehicle scales and non-recurrent variability due to inclement weather.

4.1 Recurrent Variability

4.1.1 Inter-vehicle

Previous travel time modeling attempts have typically predicted a single value of travel time for the given traffic conditions. Travel time was believed to vary in response to changes in volume or speed of the traffic stream. Thus vehicles traveling within a continuous traffic stream would have the same travel times. Mannering and his colleagues proposed that drivers could choose to travel at different speeds within the same traffic stream (Mannering, 1990). The data collected in this study support this conclusion as it is evident that vehicles traveling during the same short time period can have substantially different travel times.

Figure 4.1 shows that the variation in travel time for segment 1 in northbound direction is close to 20% of the travel time during non-congested conditions and can rise up to 100% for peak hours. This is an important finding because it indicates that it is insufficient to predict only the average travel time. This discovery can also have a profound effect on models of driver behavior because it is clear that lane selection and lane changes can greatly affect an individual's travel time.
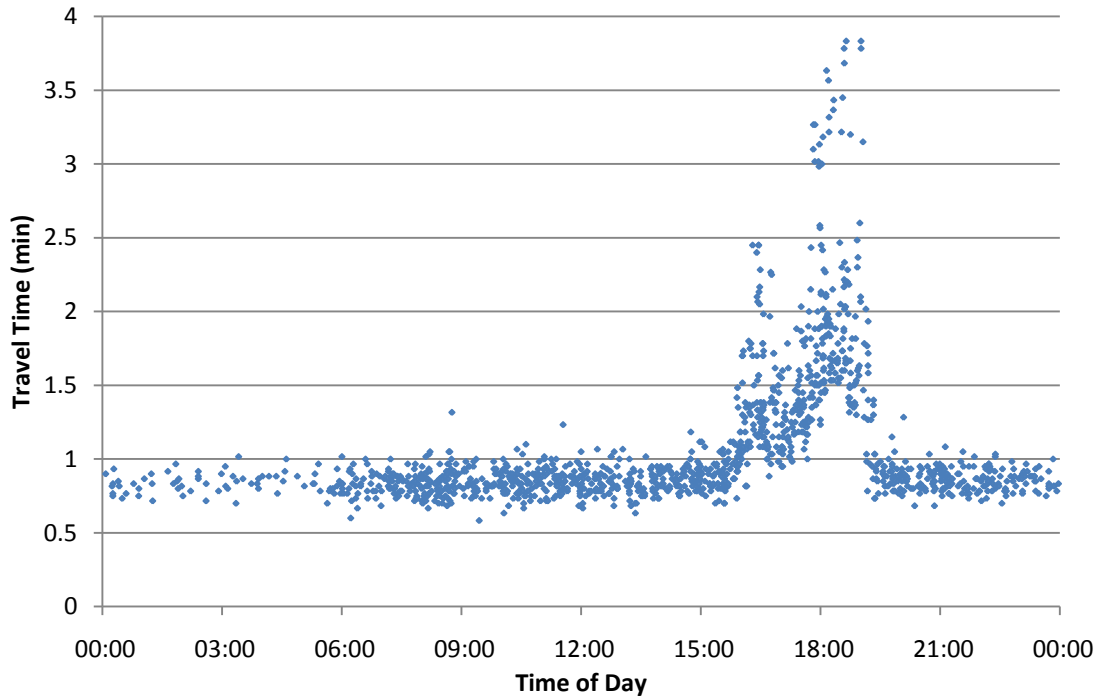
Figure 4.1 Travel Time Variation on Segment 1 in Northbound Direction

4.1.2   Inter-period

In addition to inherent travel time variation during the hour, travel time varies between different periods. As evident from Figure 4.1, most of the day, the road segment experiences light flow and corresponding free flow travel times. However, during afternoon peak hours, the travel time was observed to triple due to congestion. Unlike the inherent variation in travel times, changes in travel time during the peak hours have been previously evaluated and attributed to variations in volume.

4.1.3   Inter-day

In addition to hourly and daily variations in travel time, it may be of interest to compare travel times for two different days. Other studies suggest that travel times may vary for different days of the week due to changes in vehicle volume, however, same days of the week should have the similar travel times in the absence of unexpected events. In this study, data from two of the same weekdays were compared. The comparison was drawn between travel times and standard deviations of travel time

collected on Monday, November 17 starting at 15:00 through Tuesday November 18 at 15:00 and Monday November 25 starting at 15:00 through Tuesday November 26 at 15:00. The average travel time and standard deviation of the travel time were determined for every hour during the aforementioned days. These are shown in Table 4.1. In order to compare whether the mean travel times are different between the two days, the t-test can be performed. The t-test, which tests the null hypothesis that the two means are the same is described below,

$$t = \frac{\mu_1 - \mu_2}{s_p \sqrt{(\frac{1}{n_1} + \frac{1}{n_2})}}$$ Eq. 4.1

where, μ_1 is the average hourly travel time during the first day, μ_2 is the average hourly travel time during the second day, n_1 is the number of observations during the hour on the first day, n_2 is the number of observations during the hour on the second day, s_p^2 is the pooled variance of the samples defined as:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$ Eq. 4.2

The results of the average travel time comparison are described in Table 4.1. The table includes the mean travel time for each hour for the two days that are being examined, the value of the t-statistic and corresponding p-value. P-value indicates the probability that the mean travel times are the same. Thus a small p-value indicates that the average travel times during the specified time period are significantly different.

The table indicates that for Northbound traffic the largest difference in travel times occurs during the pm peak hours from 16:00 through 19:00. In addition, there are five other periods during the day when the average travel times are different at least at 90% confidence level. For Southbound traffic the average travel times are different during the am peak hour occurring from 8:00 until 9:00 and the pm peak hour from 17:00 through 18:00. These differences occur because the traffic flow is unpredictable in congested conditions. The average travel times are also different for four other periods

during the day. These results suggest that it is incorrect to assume that the travel times are
identical for same days of the week, in particular during the peak hours.

Table 4.1 Average Hourly Travel Times for November 17-18 and November 25-26

	Northbound				Southbound			
Hour	11/17 TT[1]	11/24 TT	T-stat	p-value	11/17 TT	11/24 TT	T-stat	p-value
15:00	0.843	0.859	0.806	0.421	0.796	0.771	1.849	0.067*
16:00	1.426	2.174	8.350	3.7E-15***	0.773	0.768	0.378	0.706
17:00	2.019	1.512	7.891	7.7E-14***	0.812	0.752	4.153	5.8E-05***
18:00	1.690	1.258	4.784	3.4E-06***	0.791	0.782	0.541	0.590
19:00	0.846	0.839	0.383	0.703	0.790	0.773	0.865	0.389
20:00	0.802	0.826	1.465	0.146	0.759	0.761	0.086	0.932
21:00	0.783	0.814	2.129	0.036**	0.782	0.786	0.193	0.847
22:00	0.766	0.837	3.705	3.9E-04***	0.779	0.759	0.908	0.367
23:00	0.814	0.834	1.111	0.272	0.801	0.742	2.229	0.031**
0:00	0.812	0.842	0.714	0.488	0.821	0.750	1.321	0.206
1:00	0.855	0.852	0.088	0.931	0.767	0.688	2.197	0.048**
2:00	0.800	0.858	0.838	0.417	0.795	0.752	1.193	0.247
3:00	0.837	0.832	0.142	0.889	0.662	0.774	1.530	0.154
4:00	0.913	0.833	1.800	0.099*	0.829	0.813	0.227	0.824
5:00	0.810	0.921	2.678	0.025**	0.712	0.701	0.359	0.722
6:00	0.836	0.844	0.322	0.750	0.752	0.744	0.437	0.663
7:00	0.837	0.857	0.906	0.369	1.235	1.277	0.740	0.460
8:00	0.852	0.844	0.318	0.751	0.879	0.769	5.126	1.3E-06***
9:00	0.855	0.856	1.005	0.318	0.760	0.733	1.161	0.248
10:00	0.853	0.819	1.833	0.070*	0.760	0.752	0.426	0.671
11:00	0.858	0.842	1.089	0.279	0.756	0.756	0.029	0.977
12:00	0.845	0.852	0.468	0.641	0.765	0.765	0.011	0.991
13:00	0.834	0.859	1.558	0.121	0.731	0.752	1.439	0.152
14:00	0.877	0.864	0.956	0.340	0.727	0.756	1.834	0.069*

TT[1] Travel Time

* Average values of travel time are significantly different at 90% confidence level

** Average values of travel time are significantly different at 95% confidence level

*** Average values of travel time are significantly different at 99% confidence level

In addition to variation in mean hourly travel time, it is important to assess the changes in the standard deviation of travel time. Comparison of the standard deviation of travel time between the two days is possible using the F-test. The F-test tests the null hypothesis that the two standard deviations are the same. It is described in equation 4.3.

$$F = \frac{s_1^2}{s_2^2}$$ Eq. 4.3

where s_1^2 is the sample standard deviation of hourly travel time during the first day, s_2^2 is the sample standard deviation of hourly travel time during the first day.

The results of the comparison including the standard deviations for the two days, F-test values and p-values are shown in Table 4.2.

Similarly to the average travel time, the largest differences in the standard deviation of travel time occurred during the afternoon peak hours for the Northbound traffic and the morning peak hours for the Southbound traffic. Once again this is explained by more unstable traffic flow during busier times of the day. In addition, there were several other periods when the standard deviations differed significantly.

Overall, the analysis indicates that the average travel times and standard deviations on the same day of the week can vary significantly. This difference is most pronounced during the peak hours. Thus, it may be incorrect to group the travel times by day of the week as is often done in studies.

Table 4.2 Hourly Standard Deviations of Travel Time for November 17-18 and November 25-26

	Northbound				Southbound			
Hour	11/17 TT[1]	11/24 TT	F value	p-value	11/17 TT	11/24 TT	F value	p-value
15:00	0.094	0.131	1.946	0.003***	0.071	0.073	1.049	0.439
16:00	0.553	0.871	2.482	1.4E-07***	0.083	0.075	1.242	0.169
17:00	0.635	0.376	2.843	1.5E-09***	0.087	0.081	1.142	0.292
18:00	0.767	0.446	2.954	8.7E-08***	0.093	0.085	1.205	0.231
19:00	0.109	0.092	1.384	0.103	0.090	0.093	1.059	0.417
20:00	0.090	0.086	1.106	0.354	0.101	0.118	1.361	0.187
21:00	0.077	0.073	1.110	0.361	0.089	0.094	1.112	0.369
22:00	0.087	0.084	1.087	0.397	0.096	0.085	1.273	0.259
23:00	0.060	0.070	1.369	0.223	0.101	0.072	1.969	0.064*
0:00	0.064	0.093	2.118	0.190	0.134	0.068	0.366	0.910
1:00	0.043	0.085	3.938	0.055*	0.071	0.073	0.791	0.608
2:00	0.073	0.145	3.980	0.098*	0.081	0.050	1.274	0.355
3:00	0.068	0.091	1.830	0.202	0.113	0.119	1.224	0.480
4:00	0.051	0.090	3.168	0.141	0.145	0.071	0.800	0.626
5:00	0.069	0.060	1.338	0.438	0.107	0.084	1.475	0.202
6:00	0.088	0.071	1.529	0.214	0.083	0.548	0.773	0.790
7:00	0.068	0.094	1.895	0.061*	0.300	0.777	1.406	0.085*
8:00	0.118	0.098	1.449	0.120	0.132	0.485	2.123	0.003***
9:00	0.069	0.070	1.492	0.121	0.070	0.118	5.338	8.3E-10***
10:00	0.090	0.099	1.206	0.257	0.096	0.102	0.623	0.949
11:00	0.082	0.064	1.634	0.054*	0.116	0.104	1.875	0.012**
12:00	0.085	0.074	1.294	0.171	0.101	0.087	0.643	0.966
13:00	0.088	0.097	1.223	0.211	0.095	0.085	0.812	0.821
14:00	0.079	0.091	1.320	0.105	0.094	0.095	1.031	0.450

TT[1] Travel Time

* Average values of travel time are significantly different at 90% confidence level

** Average values of travel time are significantly different at 95% confidence level

*** Average values of travel time are significantly different at 99% confidence level

4.2　　　　　Non-Recurrent Variability due to Adverse Weather

　　　　One of the main components of non-recurrent variability is adverse weather. In this chapter the weather impact was evaluated by collecting travel time data during a snowstorm and comparing them to a dataset collected during adequate weather conditions.

　　　　Travel time data were collected starting at 15:00 on January 27 through 11:59 on January 28. It was continuously snowing from 21:00 on January 26 until 6:00 on January 28 and thus by the time data collection commenced there was substantial snow cover. For comparison purposes travel time data were collected during sunny conditions starting on Tuesday November 18 through November 19. The resulting travel times are shown in Figure 4.2.
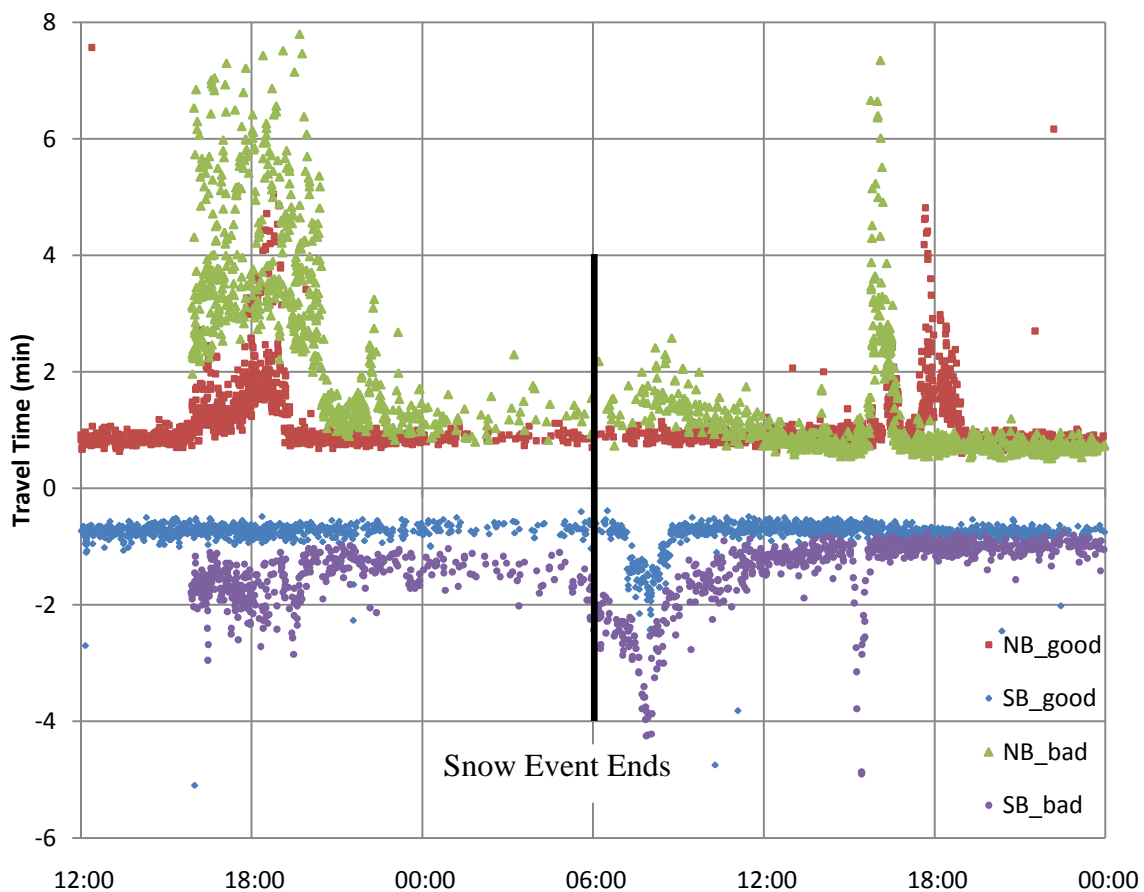

Figure 4.2- Travel Time for Snow Conditions and Sunny Day

From the graph it is evident that there is an increase in travel time which continues throughout the snowstorm event. The travel times appear to return to normal six hours following the cessation of snow. The travel time increase is further exacerbated during the pm peak hour in the Northbound direction, when the already high travel times appear to triple. In order to assess how much the travel times increase due to snowstorm activity, hourly averages of travel times were calculated and shown in Figure 4.3 and Figure 4.4. This change in travel time due to weather is clearly important to consider when assessing reliability.

In order to establish whether the average travel time is indeed different during the snow event as compared to regular conditions, t-test was performed on each pair of means. The results, shown in Table 4.3, indicate that for Northbound traffic the means are different at 90% significance level for all but two time periods (at 0:30 and 11:30). Meanwhile for Southbound traffic the means are significantly different at 95% confidence level for all time periods.
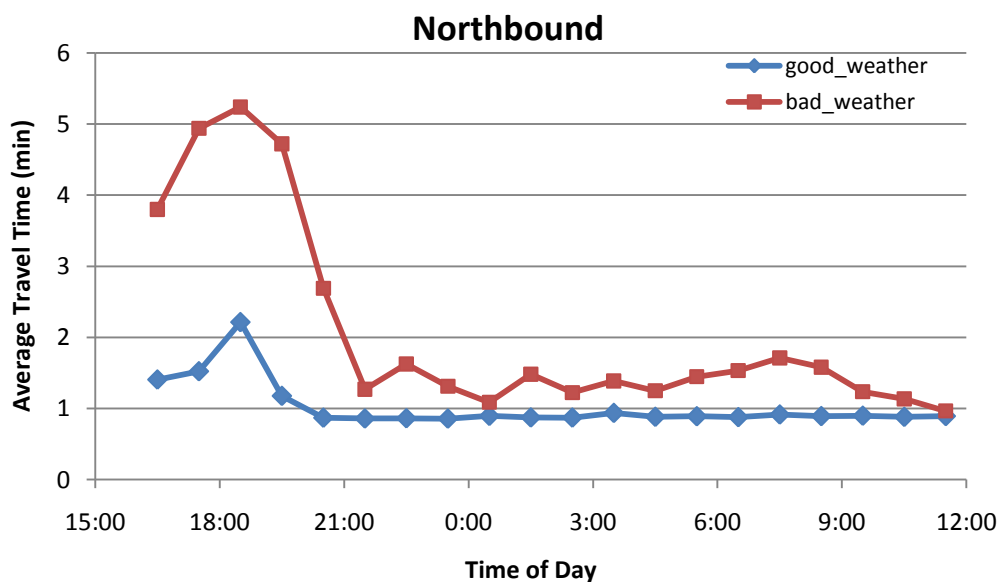

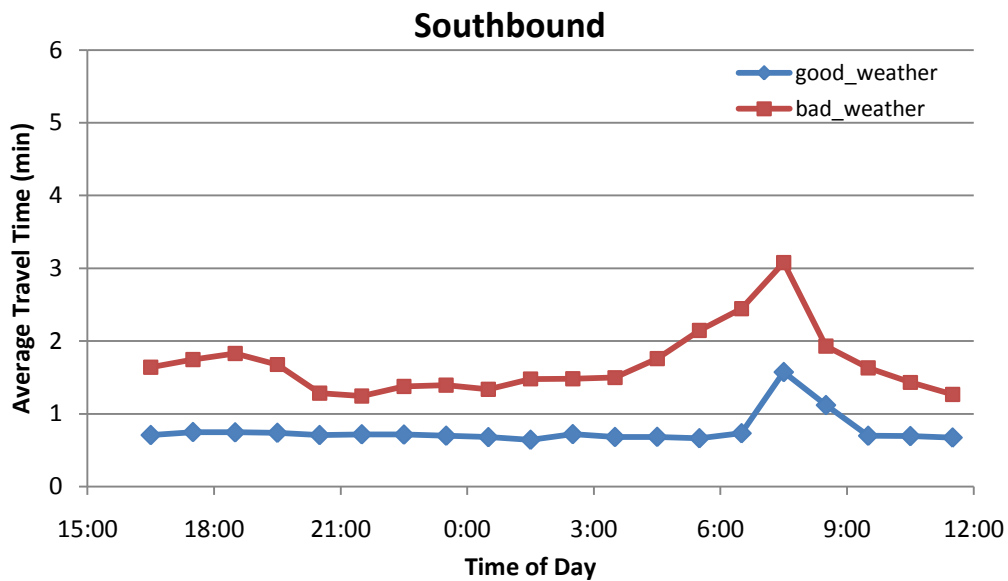Figure 4.3 Northbound Average Travel Time

Figure 4.4 Southbound Average Travel Time

Table 4.3 Comparison of Average Travel Time in Snow and Sunny Weather Conditions

		Northbound	Southbound
Hour Staring	Hour Ending	t-stat (p-value)	t-stat (p-value)
16:00	17:00	2.42 (0.008)	3.90 (0.000)
17:00	18:00	2.28 (0.012)	4.38 (0.000)
18:00	19:00	1.60 (0.056)	4.72 (0.000)
19:00	20:00	1.61 (0.054)	3.13 (0.001)
20:00	21:00	1.84 (0.034)	4.17 (0.000)
21:00	22:00	2.49 (0.007)	3.64 (0.000)
22:00	23:00	1.86 (0.033)	3.70 (0.000)
23:00	0:00	1.74 (0.044)	3.61 (0.000)
0:00	1:00	1.12 (0.136)	5.28 (0.000)
1:00	2:00	2.13 (0.024)	4.58 (0.000)
2:00	3:00	2.17 (0.026)	4.95 (0.000)
3:00	4:00	1.78 (0.047)	4.42 (0.000)
4:00	5:00	1.79 (0.049)	4.75 (0.000)
5:00	6:00	2.55 (0.01)	6.34 (0.000)
6:00	7:00	3.45 (0.002)	10.04 (0.000)
7:00	8:00	3.16 (0.002)	3.28 (0.001)
8:00	9:00	2.64 (0.005)	2.20 (0.015)
9:00	10:00	2.10 (0.02)	6.56 (0.000)
10:00	11:00	1.55 (0.063)	3.67 (0.000)
11:00	12:00	0.60 (0.274)	4.64 (0.000)

In addition to an increase in the average travel time, the variance in travel times during the same time period also rises. The variance of travel times is shown in Figure 4.5 and Figure 4.6.
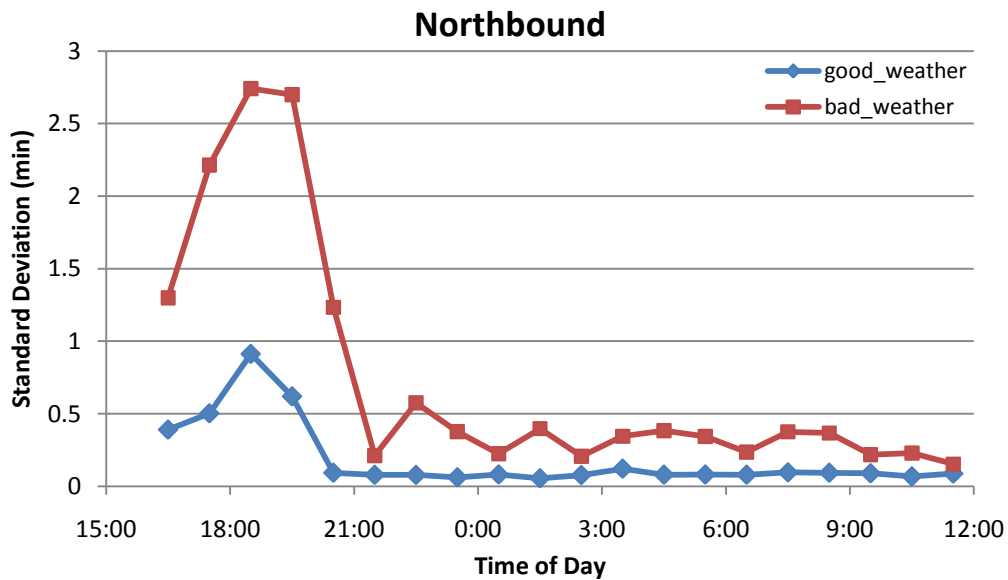

Figure 4.5 Variability in Northbound Travel Time for Snow Conditions and Regular Day
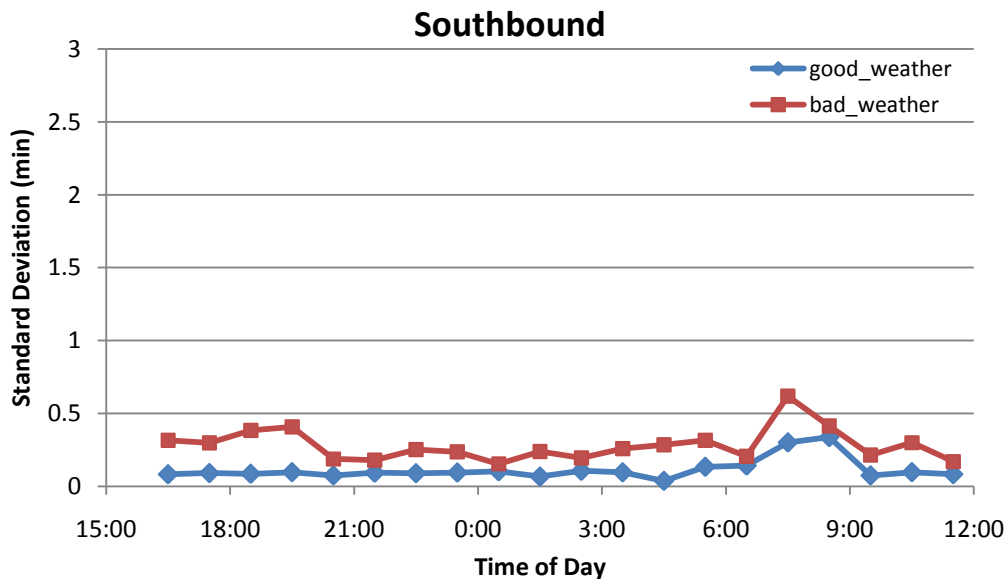

Figure 4.6 Variability in Southbound Travel Time for Snow Conditions and Regular Day

The graphs indicate that adverse weather results in increased variability in travel time. In order to assess whether the standard deviation of travel time during snow conditions is significantly different from the standard deviation of travel time under regular weather conditions, F-tests were performed for every observation.

Table 4.4 Comparison of Standard Deviations in Snow and Sunny Weather Conditions

		Northbound	Southbound
Hour Staring	Hour Ending	t-stat (p-value)	t-stat (p-value)
16:00	17:00	11.08 (0.000)	14.57 (0.000)
17:00	18:00	19.46 (0.000)	10.73 (0.000)
18:00	19:00	9.03 (0.000)	19.86 (0.000)
19:00	20:00	18.99 (0.000)	17.73 (0.000)
20:00	21:00	176.08 (0.000)	6.27 (0.000)
21:00	22:00	7.23 (0.000)	3.68 (0.000)
22:00	23:00	54.78 (0.000)	7.90 (0.000)
23:00	0:00	37.82 (0.000)	6.30 (0.000)
0:00	1:00	7.64 (0.000)	2.24 (0.061)
1:00	2:00	52.76 (0.000)	12.72 (0.000)
2:00	3:00	7.19 (0.016)	3.35 (0.081)
3:00	4:00	8.13 (0.004)	7.26 (0.013)
4:00	5:00	23.16 (0.013)	58.13 (0.000)
5:00	6:00	18.11 (0.000)	5.56 (0.000)
6:00	7:00	8.86 (0.002)	2.08 (0.031)
7:00	8:00	15.44 (0.000)	4.21 (0.000)
8:00	9:00	15.60 (0.000)	1.51 (0.095)
9:00	10:00	6.01 (0.000)	8.28 (0.000)
10:00	11:00	11.41 (0.000)	9.55 (0.000)
11:00	12:00	3.07 (0.000)	4.15 (0.000)

In conclusion, adverse weather conditions such as a snow event, result in higher travel time and standard deviation of travel time. This effect may last hours after the event depending on the severity and resulting road conditions. From the analysis, it is evident that adverse weather impact is difficult to quantify and incorporate in the models.

CHAPTER 5. AUTOREGRESSIVE MODEL

This chapter focuses on modeling travel times obtained from two Northbound sections of I-69, a segment from Mile Marker 2.9 to Mile Marker 3.7 and segment from Mile Marker 3.7 to Mile Marker 5.0. As mentioned in the previous chapter, travel times were available from Monday, November 17 through Friday November 22 and from Monday, November 25 through Wednesday, November 26. Loop detector data, including speeds and volumes was available for all the aforementioned days except Monday, November 25. Therefore, travel time data from this day were omitted from the dataset. Furthermore, on Tuesday November 26 there was an accident during the pm peak hour, thus causing abnormally high travel times that were also omitted from the modeling process. The final dataset consisted of 13234 individual travel time observations complemented by the traffic data.

The goal of this chapter is to develop a model that predicts travel time in terms of traffic parameters obtained from RTMS detectors and thereby evaluates the impact of changes in these parameters. Literature review indicated that parameters that may be important in predicting travel time include average segment speed, distance, and volume. In addition, peak hour parameter was incorporated to capture additional impact of congestion.

5.1 Model Parameters

Speed can be considered the single most important predictor of travel time. Ideally, if one could determine the exact speed profile of a vehicle traveling over a segment of the road, travel time would be known. However, only point speed estimates are generally available from loop detectors. Furthermore, speeds are not available for

individual vehicles but rather 30-second averages. In this study the 30-second point speeds were combined as explained in the previous section to yield an average segment speed, which was the incorporated into the model.

Distance is another critical variable, since an increase in the length of the segment will obviously increase the vehicle travel time.

Speed and distance variables are important in predicting travel times. These variables can be included in the model as linear terms. However, from a theoretical standpoint, vehicle travel time should be proportional to distance divided by speed and thus a term representing the average travel time on the segment was included instead of separate linear terms.

Despite incorporating speed into the model, volume can have an additional impact on the flow. Volume at the time of travel as well as volume one hour before the trip is included in the modeling process.

Finally, rush-hour parameter can have many different definitions because it is difficult to estimate its onset. In this study, rush hour was estimated to begin at 17:00 when most of the observed days vehicle speed dropped to half the free flow speed. The peak hour would cease at 18:30 pm, when vehicle speeds rose above the same threshold. This indicator variable is incorporated into the model in addition to the traffic data because it is possible that during the rush hour there is a breakdown in the traffic flow and average speed and volume data will be unable to capture the extent of increase in travel time.

5.2 Model Selection

Because travel time is a continuous variable, the first modeling formulation used was a linear regression with travel time as the dependent variable. Independent variables that were included in the model were average travel time, volume and rush hour indicator. Linear regression analysis yielded reasonable parameter estimates and model fit. However, a check on the Durbin-Watson statistic, which was significantly different from two, indicated serial correlation between observations. Serial correlation violates the

assumption that the error terms are independent between different observations and may result in parameter estimates being biased and inconsistent (Washington, Karlaftis, & Mannering, 2003). In this study serial correlation is not unexpected since time-series data are being analyzed. One approach that allows to remove the correlation is regression incorporating lagged dependent variable terms. The first-order autoregressive model form is described as,

$$Y_i = \beta X_i + \gamma Y_{i-1} + e_i \qquad\qquad \text{Eq. 5.1}$$

where β is a vector of estimated independent variable, X_i is the independent variable, γ is the coefficient of the lagged dependent variable, e_i is the error term.

Once this model is estimated, the Durbin Watson statistic is once again evaluated and additional lagged terms are added to meet the regression assumption.

5.3 Model Estimation

An autoregressive model was initially estimated with a single lagged dependent variable term. However, this attempt did not resolve serial correlation and more lagged terms were added, one at a time, until the Durbin Watson test indicated absence of serial correlation in the model. The final number of lagged variables was seven. It was also observed that the coefficients of second, third and fourth variables had a similar value as well as the coefficients of the fifth, sixth and seventh lagged terms. These were combined and represented as the average of second through fourth lags and the average of fifth through seventh lagged terms.

In addition to the lags of the dependent variable, average segment travel time and segment volume terms were found to be significant. The final model is shown in equation 5.2, and estimation results are detailed in Table 5.1.

$$Y = -0.0732 + 0.00425 * \text{Travel Time} + 0.00680 * \text{Volume} + 0.1733 * \text{Laged TT} +$$
$$0.3464 * \text{TT Lag4through6} + 0.1667 * \text{TTLag5through7} + 0.0462 * \text{Rush Hour}$$
$$\text{Eq. 5.2}$$

Table 5.1 Autoregressive Model of Travel Time

Parameter	Description	Value	Standard Error	T-statistic	p-value
Intercept		-0.0732	0.0135	-5.4	<0.0001
Travel Time	Average segment travel time, estimated as distance divided by speed (in minutes)	0.374	0.019448	19.23	<0.0001
Volume	Volume one hour before the vehicle enters the segment (veh/hr/lane)	0.006803	0.000719	9.46	<0.0001
Lagged Travel Time	Travel time of preceding vehicle (in minutes)	0.1733	0.008735	19.85	<0.0001
Travel Time Lag 2 through 4	Average travel time of preceding vehicle 2 through 4 (in minutes)	0.3464	0.0132	26.2	<0.0001
Travel Time Lag 5 through 7	Average travel time of preceding vehicle 5 through 7 (in minutes)	0.1667	0.0126	13.2	<0.0001
Rush Hour	Indicator variable equal to 1 if the vehicle is traveling between 17:00 and 18:30	0.0462	0.08742	5.28	<0.0001

Number of observations: 13,234
R^2: 0.62

5.4 Model Evaluation

Model fit was evaluated by calculating R squared statistic and by observing plots of actual and predicted values. R-squared statistic was determined to be 0.62. Plots of actual travel times and those predicted by the model for the two segments are shown in Figure 5.1 and Figure 5.2.
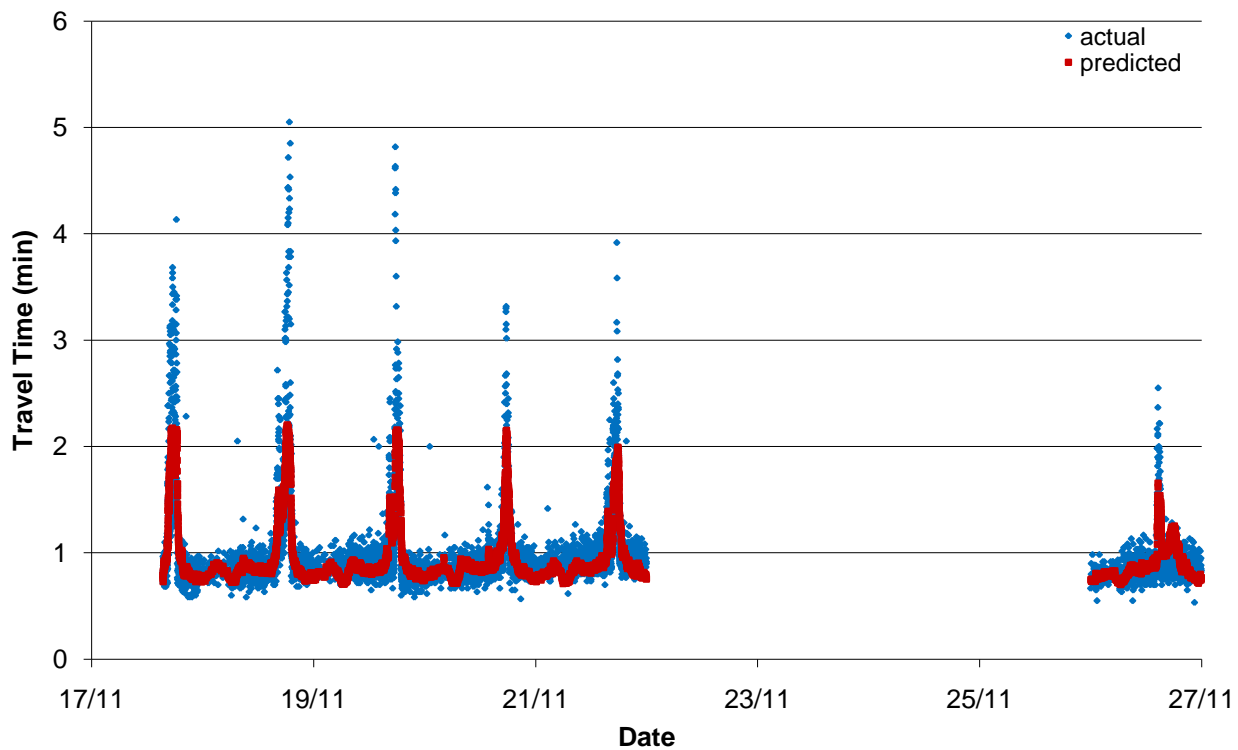
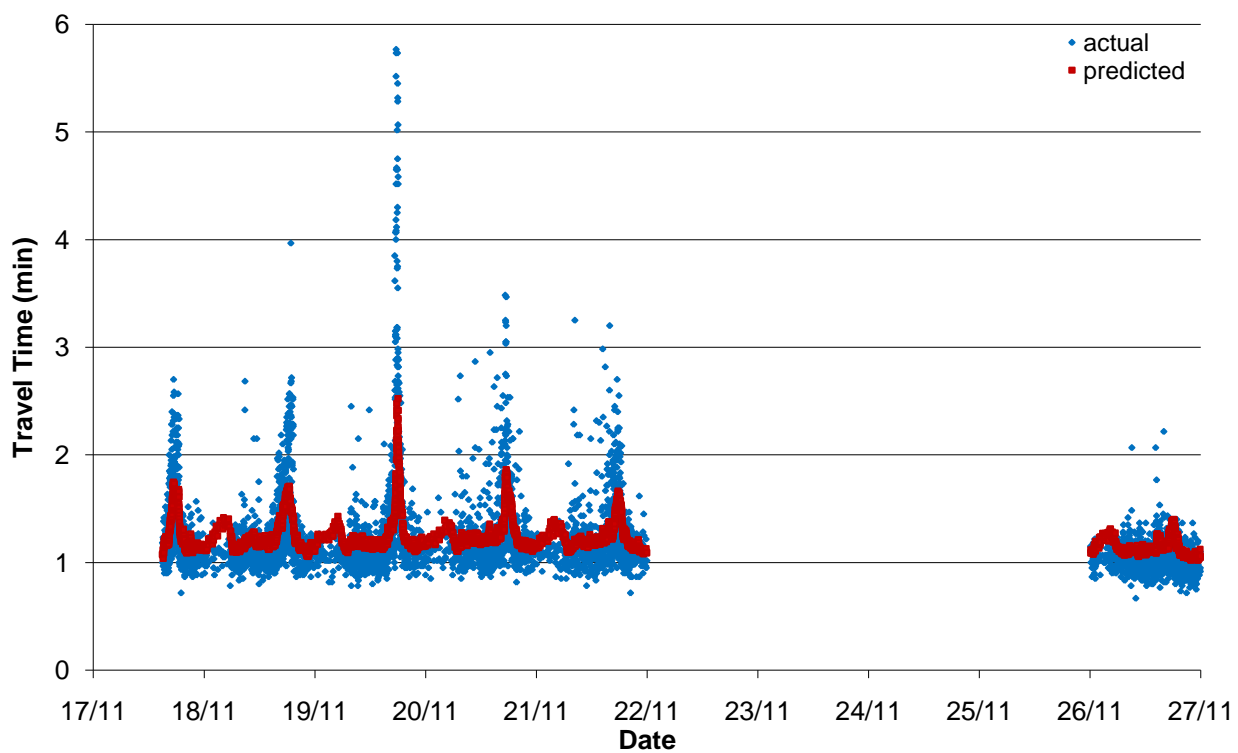Figure 5.1 Estimated and Actual Travel Time on Road Segment 1

Figure 5.2 Estimated and Actual Travel Time on Road Segment 2

The model appears to capture the trends in travel time during the day, showing an increase in travel times during the peak hours and a fairly constant travel time during the day. However, the travel times predicted during the night appear to be higher than the actual ones. This occurrence can be explained by observing the average travel speeds at night as shown in Figure 5.3.

There is an evident drop in vehicle speeds resulting in the model overpredicting the travel time. The drop in average speed can be explained by some drivers choosing to travel slower at night and thus decreasing the average segment speed. Since individual travel times are being modeled, if the slow drivers are not captured, the observed travel times may not decrease at night.
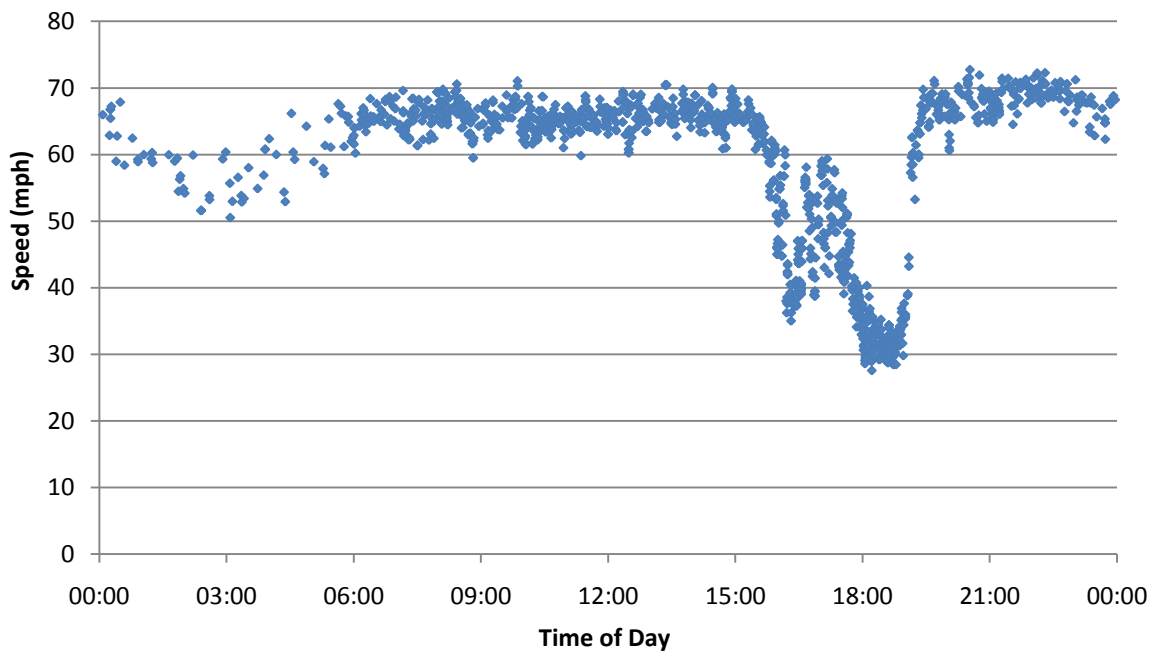

Figure 5.3 Daily Speed Profile on Segment 1

From the figures it also appears that the model underpredicts the highest travel times during the peak hours. However, from Figure 5.4 it is clear that there is a lot of

variance in individual vehicle travel times during the peak hour and the model is effective at predicting the average travel time value.
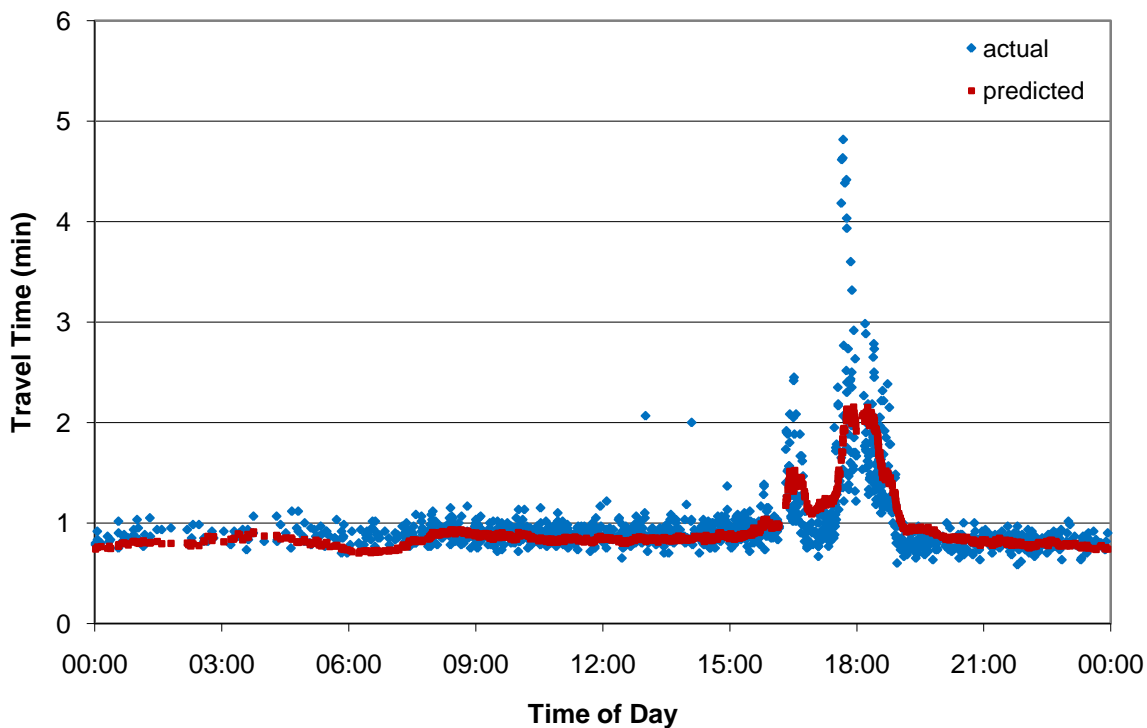

Figure 5.4 Travel Time on Segment 1

Next, it is important to assess the improvement that an autoregressive model yields over conventional ways of predicting travel times from travel speeds. Plotting travel times estimated as distance divided by the average travel speed is shown in Figure 5.5 and Figure 5.6.
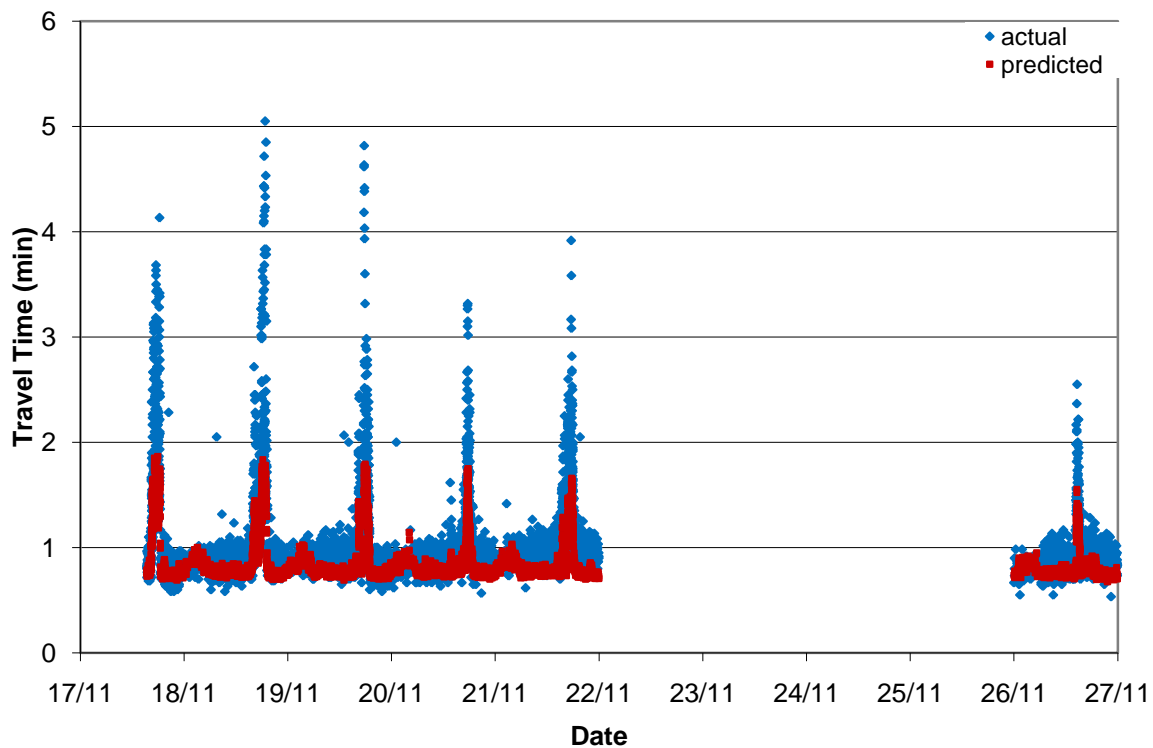
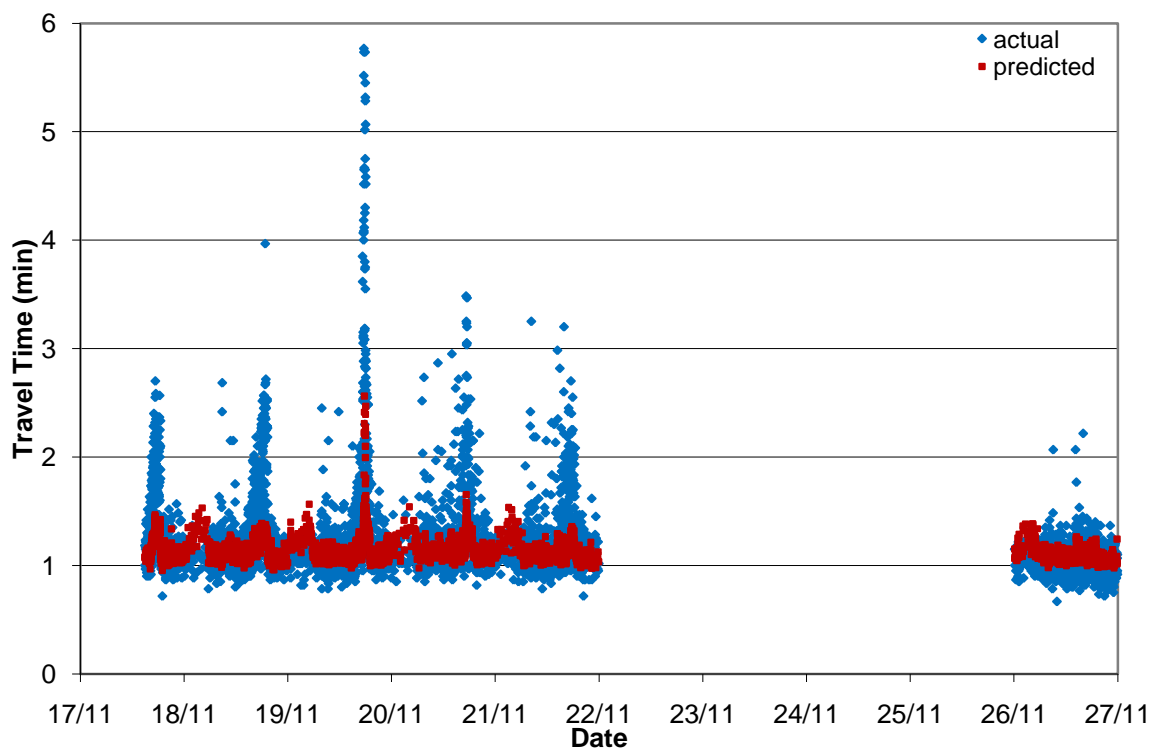Figure 5.5 Travel Times Estimated from Speed for Segment 1

Figure 5.6 Travel Times Estimated from Speed for Segment 2

For segment 1 the speed model fails to predict the extent of the increase in travel times during the peak hour and underpredicts travel times during the day. For the second segment the model performance is even less acceptable as random variations in travel times during the day exceed the variations during the peak hour. An objective comparison of the models can be performed using the root mean square error which corresponds to the distance between the predicted and actual estimate. It is described in equation 5.3.

$$RMSE = \sqrt{E\left[(Yi - Y)^2\right]}$$

Eq. 5.3

The autoregressive model yields a RMSE of 0.30, while the RMSE of the model only based on the segment speeds and distances is 0.36.

Another way to evaluate the goodness of fit of the model is to plot the predicted travel times values versus the actual ones. The scatter plots for two road segments are shown in Figure 5.7 and Figure 5.8.
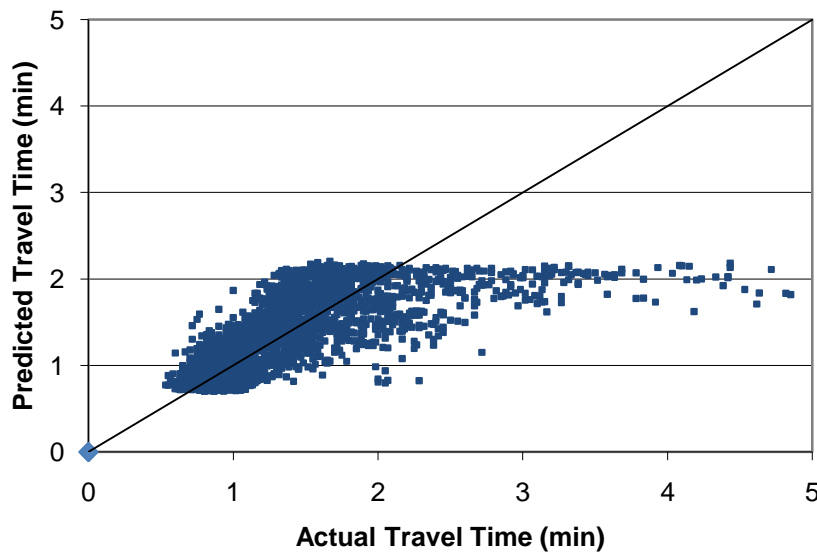

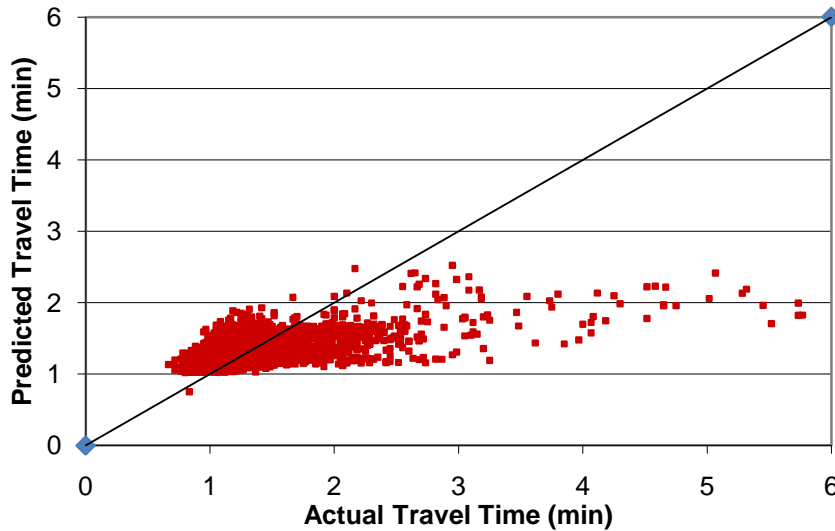Figure 5.7 Scatter Plot of Predicted vs. Actual Travel Times for Segment 1

Figure 5.8 Scatter Plot of Predicted vs. Actual Travel Times for Segment 2

It is evident that model performance is good for low travel times however appears to deteriorate for higher travel times. This may be attributed to a higher spread in travel times during the peak hours as explained previously. The model cannot incorporate individual driver behavior which results in both very high and low travel times.

5.5 Discussion

The estimated autoregressive model performs well compared to the simple model based only on speed and segment distance. As mentioned the predicted travel times deviate from the actual travel times during the night and peak hours. At night this is due to lower average travel speed on the freeway, which may not affect the individual drivers captured by the bluetooth detectors. During peak hours, inter-vehicle variability is especially high as discussed in chapter 4 and the model is only able to predict the average travel time value.

The model, estimated in this study can be used to replace the models used for real-time travel time prediction. As indicated in the previous section this model performs better than the one based only on the traffic speed, which is most commonly used in the

field. At the same it time it retains the computational simplicity and is easy for practitioners to understand.

CHAPTER 6. DURATION MODEL

Data used for the duration model are the same as those used for autoregressive model. The dataset includes 13234 individual vehicle travel time observations collected between November 17 and November 26. These data are complemented with 30-second volume and speed from loop detectors.

Parameters used in the duration model are analogous to the autoregressive model and include segment speed, volume, and segment length as well as time of day indicators.

In addition to predicting the actual travel time for a vehicle based on the current traffic conditions, one may be interested in how much longer it will take to complete the trip given that a motorist has been traveling for a specified amount of time. In order to answer this question, a duration model can be estimated. The duration model allows exploring the probability of completing the trip during the next time period given the elapsed travel time and estimating the percentage of vehicles exiting the segment at any trip duration.

6.1 Modeling Approach

Duration models can have non-parametric, semi-parametric or fully parametric form (Washington, Karlaftis, & Mannering, 2003). Non-parametric models do not make an assumption regarding the distribution of the duration times or the functional form of the effect of covariates on the hazard function. In contrast, fully parametric duration models assume the distribution of the duration times is either logistic, Weibull, exponential or Gamma and covariates influence the base hazard function as $\exp(\beta X)$.

Development of a duration model begins with specifying the cumulative distribution function $F(t)$

$$F(t) = Pr(T < t) \qquad \text{Eq. 6.1}$$

where Pr refers to probability of the vehicle's travel time, T, being less than a specified time t. The hazard function, h(t), which corresponds to the probability that the vehicle's travel time will end between t and t+dt given that it has lasted until t, is,

$$h(t) = f(t)/[1 - F(t)] \qquad \text{Eq. 6.2}$$

where f(t) is the derivative of the cumulative distribution function F(t)

The survival function S(t), which provides the probability that the trip duration is greater than or equal to some specified time t is,

$$S(t) = Pr(T \geq t) \qquad \text{Eq. 6.3}$$

The hazard ($h(t)$), density ($f(t)$), cumulative distribution ($F(t)$) and survivor functions ($S(t)$) are graphically illustrated in Figure 6.1.
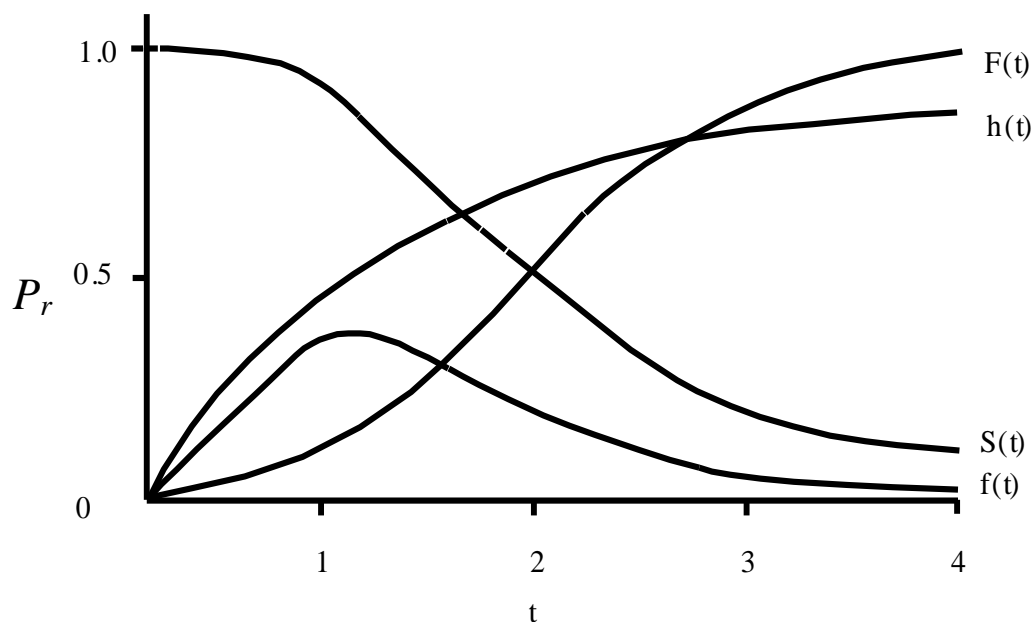
Figure 6.1 Illustration of $h(t)$, $f(t)$, $F(t)$ and $S(t)$ functions (Source: Washington et al., 2003)

The hazard function is important to determine how the probability of exiting the segment changes as the vehicle spends more time traversing it. The shape of the hazard function is critical in this assessment and can take on the shape of the curves shown in Figure 6.2. The first hazard function, $h_1(t)$, is monotonically decreasing implying that as

more time elapses since the vehicle entered the segment the probability of the trip ending decreases. The second hazard function, $h_2(t)$, increases initially and then proceeds to decrease. In this case, the vehicle would be more likely to exit the segment as more time elapses until the inflection point. Then, the probability of the vehicle exiting the segment would decrease as travel time increased. The third hazard function, $h_3(t)$, is monotonically increasing corresponding to the trip being more likely to end as travel time increases. Lastly, the fourth hazard function, $h_4(t)$, is constant, implying that the probability of exiting the segment does not change with the amount of time the vehicle spends traversing it.
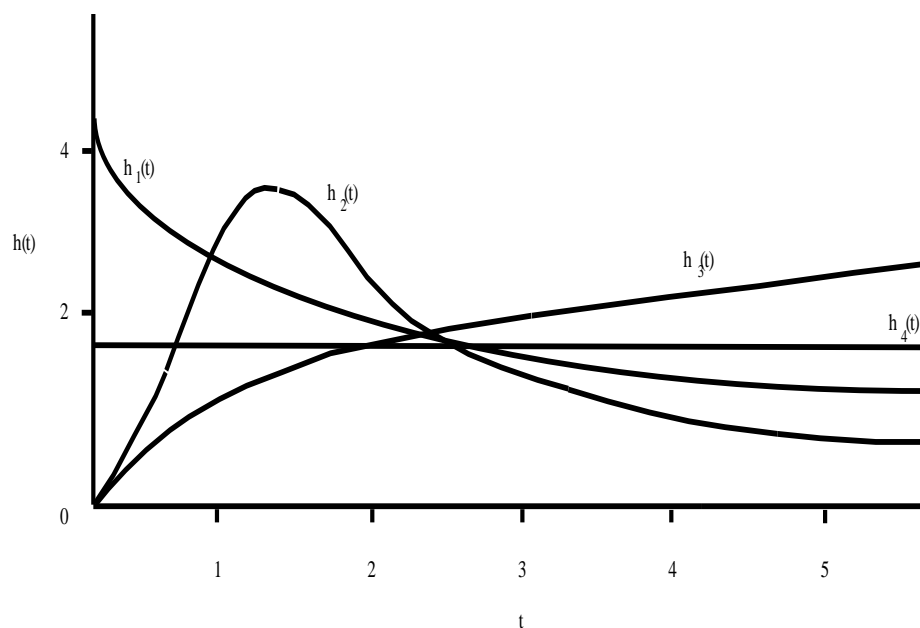
Figure 6.2 Hazard Functions (Source: Washington et al., 2003)

Because an assumption can be made regarding the distribution of travel times, a fully parametric duration model formulation was selected. Next, a distribution needs to be selected. The most common distributions include exponential, Weibull, log logistic and gamma. Exponential distribution restricts the hazard function to a constant value and thus the probability of trip ending at any point in time does not depend on how long the vehicle has been traveling. Weibull model assumes monotonically increasing or

decreasing hazard function. Log logistic distribution function allows for nonmonotonic hazard function and may have an inflection point. For the purposes of this study, log logistic function was selected due to being the least restrictive. Later in the chapter this selection will be tested to verify this assumption. The density function for this distribution is,

$$f(t) = \frac{\lambda P(\lambda t)^{P-1}}{[1+(\lambda t)^{P}]^{2}}$$

Eq. 6.4

And hazard function is:

$$h(t) = \frac{\lambda P(\lambda t)^{P-1}}{[1+(\lambda t)^{P}]}$$

Eq. 6.5

where P and λ are estimable parameters

Equation 6.5 indicates that if P<1, the hazard function is monotonically decreasing as the time spent traversing the segment increases. Meanwhile if P>1, the function is increasing to an inflection point and then decreases.

6.2 Model Estimation

First, a model was estimated that included all the observations. In addition to independent variables described earlier, lagged dependent variables were included. Unlike linear regression, duration models do not contain an assumption regarding the independence of error terms. However, the lagged variables are included because they help improve the model fit. Modeling results are shown in Figure 6.3.

Table 6.1 Overall Duration Model

Parameter	Description	Value	Standard Error	T-statistic	p-value
Intercept		-0.17803	0.020197	-8.815	<0.000
Speed	Average 30-second segment travel speed (in mph)	-0.01083	0.000247	-43.942	<0.000
Volume	Average 30-second volume one hour prior to the observation (in veh/hr/lane)	0.008933	0.000412	21.675	<0.000
Distance	Segment length (in miles)	0.0000995	0.00000195	50.980	<0.000
Lagged Travel Time	Travel time of the previous vehicle (in minutes)	0.157626	0.006124	25.739	<0.000
Second Lag of Travel Time	Travel time of the second last vehicle (in minutes)	0.12562	0.005863	21.426	<0.000
λ		0.947	0.012		
P		11.86	0.087		
Number of observations		13234			
Log likelihood		5997			

The model yields reasonable signs and coefficient values. As anticipated higher speed results in lower travel time, while higher volume, distance and preceding vehicle travel time all lead to increased travel time. All of the coefficients in the model are significant as indicated by their high t-statistics.

Next, hazard functions, which represent the rate of exiting the road segment at any point in time, are plotted. Since the two road segments are of different lengths and thus have different travel times, hazard functions for each of the segments are shown in separate plots (Figure 6.3 and Figure 6.4). In both cases the function starts at a value of minimum travel time for that segment. For segment 1 this time is 34 seconds

corresponding to average travel speed of 89 miles per hour. Segment 2 had the minimum travel time of 43 seconds corresponding to speed of 100 miles per hour. Since there are very few vehicles traveling at such speed, the probability of reaching the end of the road segment in the short travel time is low. As the travel time increases so does the rate of vehicles completing their trip within that time. The hazard function for the first segment reaches its peak at 56 seconds (average speed of 55mph), while for the second segment the highest rate occurs at 71 seconds (average speed of 60 mph).

It is hypothesized that the inflection point corresponds to the onset of congestion. Initially the probability of exiting the segment increases as the person spends more time traversing it, however, in congested condition despite the increase in the time that the traveler has already spent on the road, the probability of reaching the end of the segment may decline. For segment 1, which has a free flow speed of 67 mph, the highest rate of exiting the segment occurs at 55 mph. This is similar to the transition into Level of Service F according to the Highway Capacity Manual, which occurs at 53 mph. The level of service diagram is shown in Figure 6.5.
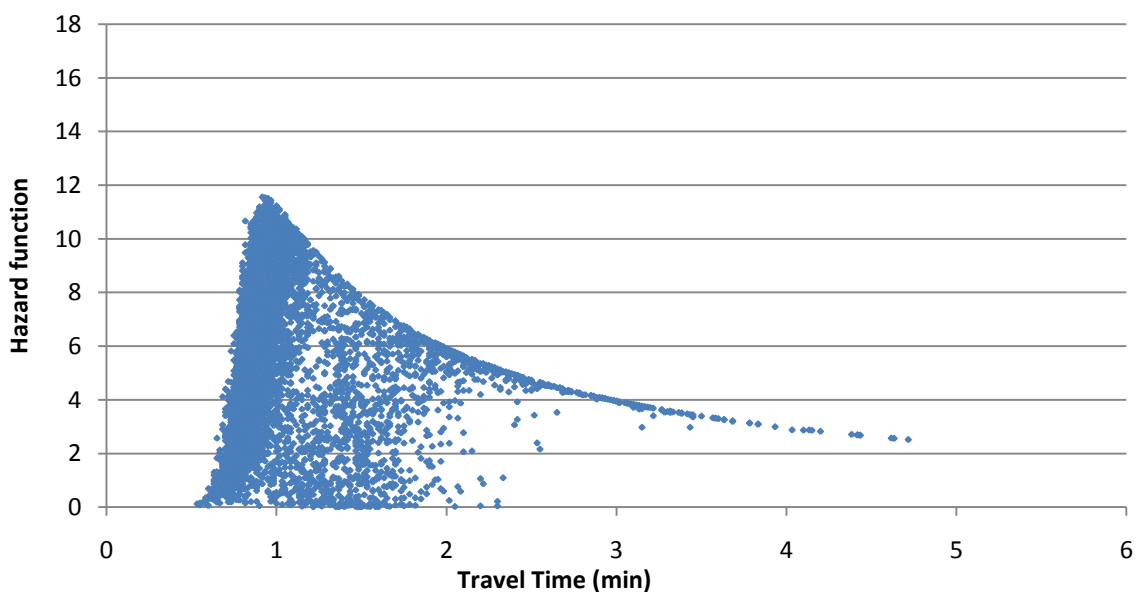

Figure 6.3 Hazard Function for Overall Duration Model for Segment 1

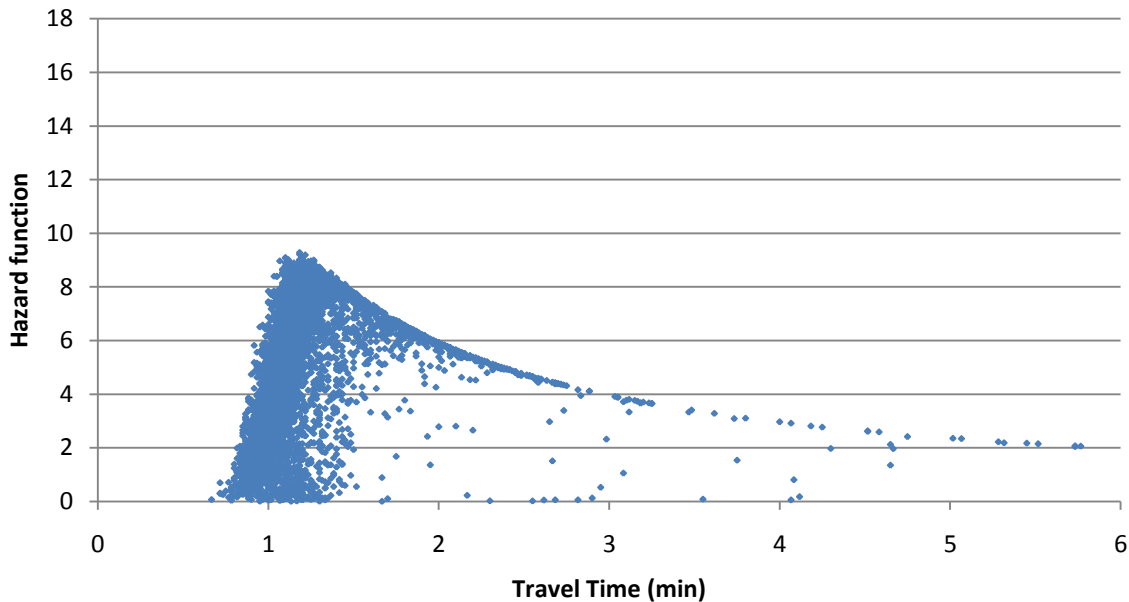Figure 6.4 Hazard Function for Overall Duration Model for Segment 2
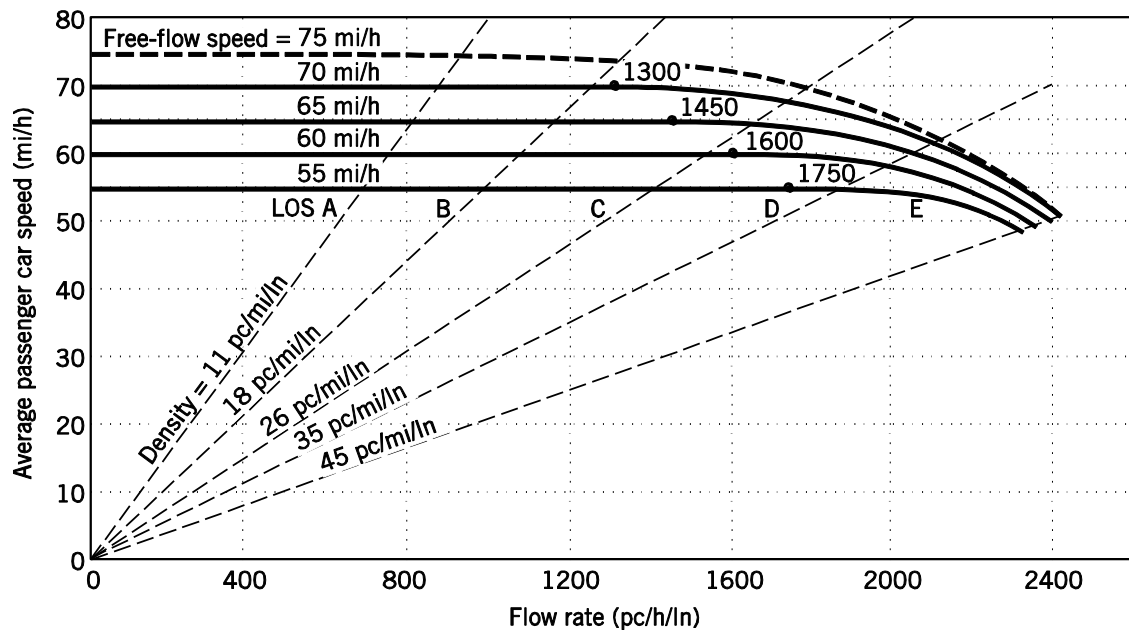

Figure 6.5 Speed Flow Curves and Level of Service Criteria (Source: HCM 2000)

In addition to observing how the probability of exiting the road segment varies with time, it may be of interest to observe the probability of traveling longer than the specified travel time. This probability can be observed from survival functions shown in

Figure 6.6 and Figure 6.7. Similarly to the hazard functions there is a spread in the values of the survival function for any travel time. This is caused by the deviation of independent parameters from the average value thus shifting the function. Most of the points are clustered along the curve that has the lowest probabilities of survival corresponding to the non-congested traffic conditions. In peak hour conditions, the survival function shifts to the right, increasing the probability of traveling longer. From the survival functions, it also evident that the travel times on the second segment are more consistent as the range of probabilities of traveling the specified length of time or longer is smaller.
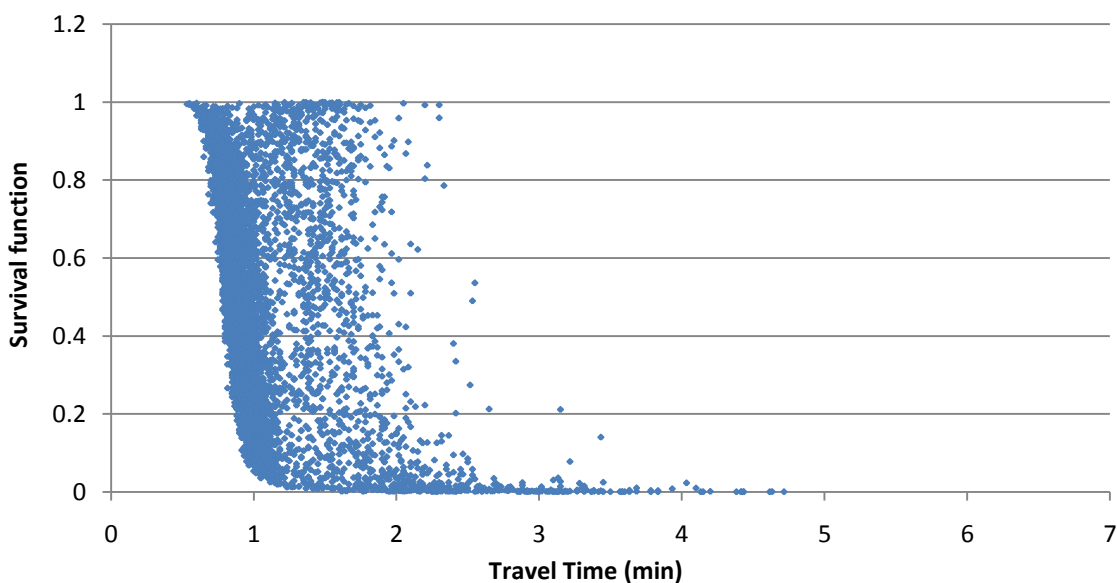

Figure 6.6 Survival Function for Overall Duration Model for Segment 1
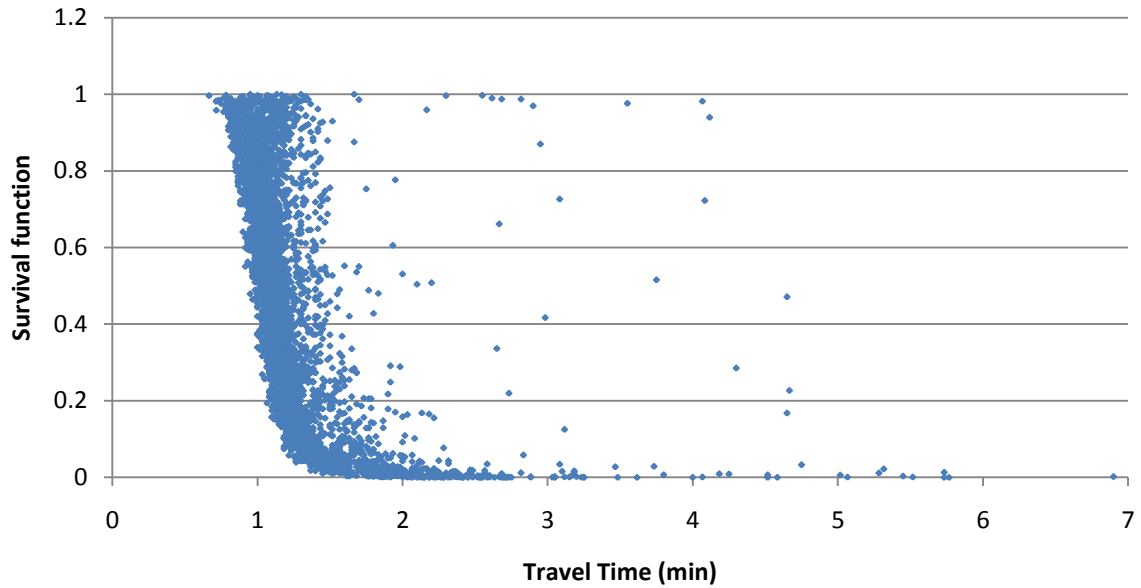
Figure 6.7 Survival Function for Overall Duration Model for Segment 2

6.2.1   Night Time Model

In addition to estimating a duration model for all observations, a separate model was developed for night time. Night time was assumed to begin at 22:00 and continue through 6:00. This time of the day corresponds to non-congested conditions and we expect to see a monotonically increasing hazard function because the longer the vehicle traveled the more likely it would be to exit the segment. Autoregressive models discussed in the previous section also indicated that the explanatory variables affected travel time differently during the day and at night. Thus the model also aimed to compare the parameter coefficients between day and night. The resulting model for night time is described in Table 6.2.

Table 6.2 Night Time Duration Model

Parameter	Description	Value	Standard Error	T-statistic	p-value
Intercept		-0.69783	0.040182	-17.367	<0.0000
Speed	Average 30-second segment travel speed (in mph)	-0.00219	0.000571	-3.830	0.0001
Volume	Average 30-second volume one hour prior to the observation (in veh/hr/lane)	0.008615	0.001503	5.730	<0.0000
Distance	Segment length (in miles)	.00008174	.00000472	17.334	<0.0000
Lagged Travel Time	Travel time of the previous vehicle (in minutes)	0.196927	0.024232	8.127	<0.0000
Second Lag of Travel Time	Travel time of the second last vehicle (in minutes)	0.149118	0.024157	6.173	<0.0000
λ		1.06	0.0027		
P		16.97	0.345		
Number of observations		1679			
Log likelihood		1403			

The estimated model produced intuitive parameter signs and coefficient magnitudes. One difference between the overall model and the night time model and is the impact of speed on travel time. Average segment speed appears to have a much smaller effect on travel time during night time. Meanwhile the lagged dependent variable has a greater impact on travel time during night time suggesting that drivers may be choosing to travel at similar speeds as nearby vehicles in free flow conditions.

As mentioned previously, it was hypothesized that congested conditions cause the hazard function to start decreasing after some point. Thus during the night in the absence of delays caused by accidents or road work, the hazard function should be monotonically increasing. Figure 6.8 and Figure 6.9 show the hazard function versus the travel time for trips between 22:00 and 6:00. For both segments, the inflection point exists, however, the rate of exiting the segment remains high. Also, very few points lie beyond the inflection point suggesting that there are just a few drivers traveling slower.

Night time conditions are characterized by absence of congestion, it is therefore anticipated that the survival function will decline rapidly with increasing travel time. Furthermore, due to relatively uniform vector of parameters, the survival functions should have little deviation from the base survival function. The resulting survival functions are shown in Figure 6.10 and Figure 6.11.
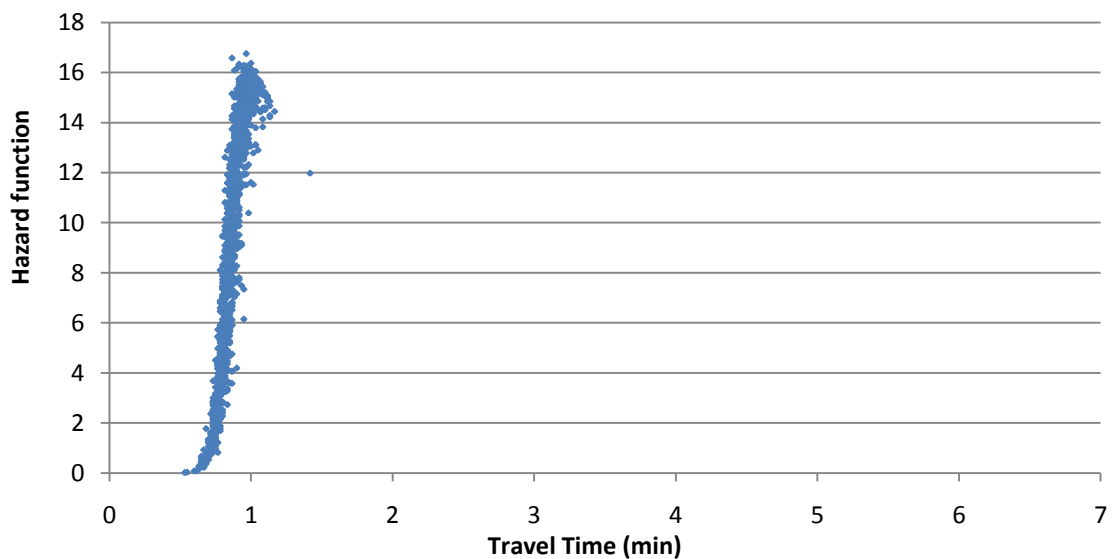

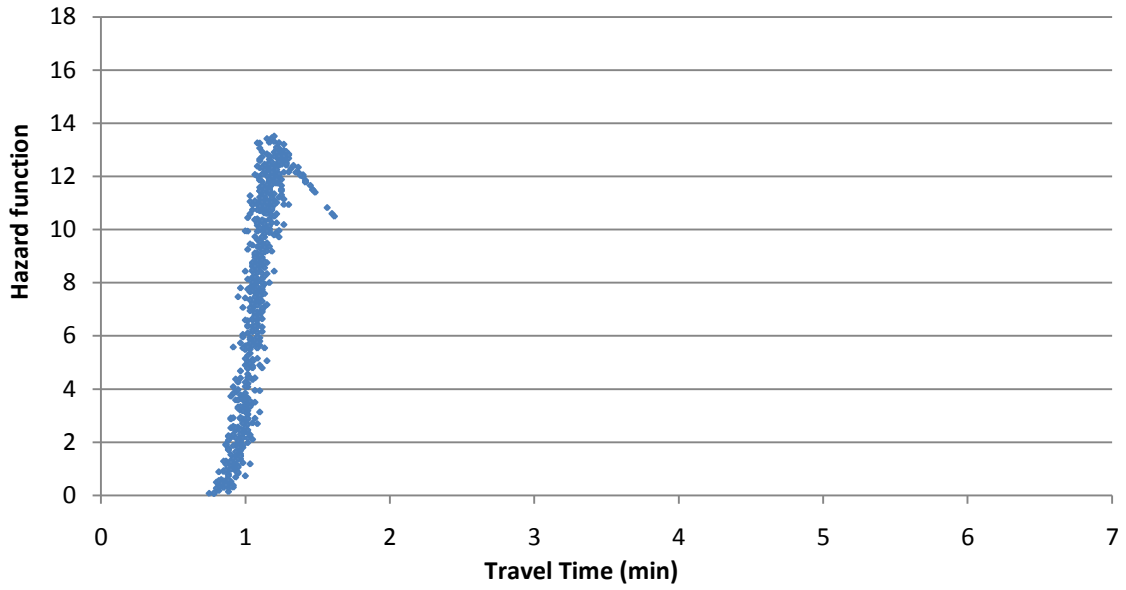Figure 6.8 Hazard Function for Night Time Duration Model for Segment 1

Figure 6.9 Hazard Function for Night Time Duration Model for Segment 2
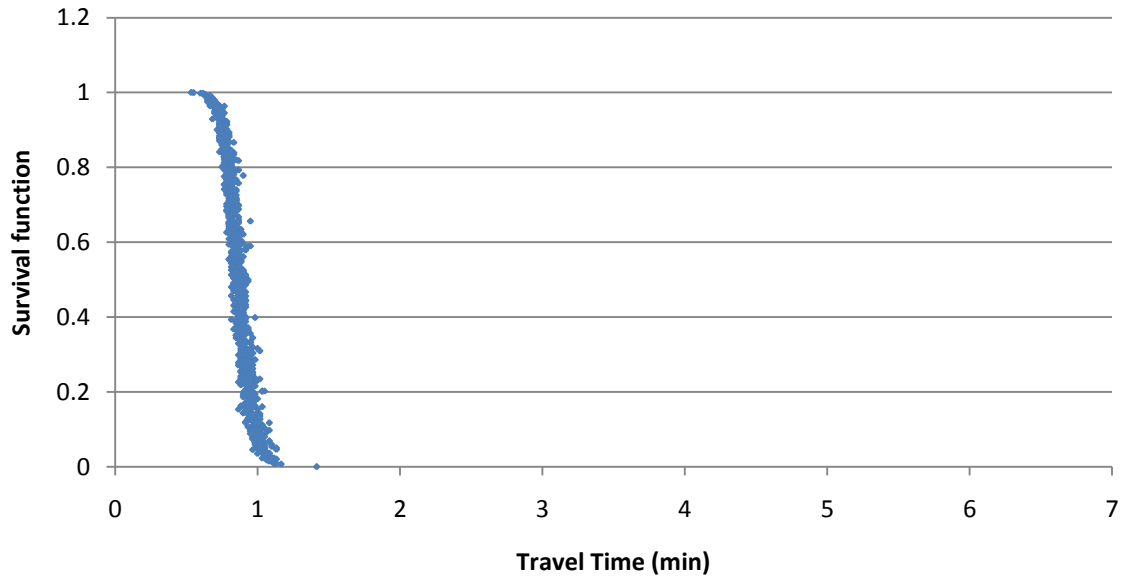
Figure 6.10 Survival Function for Night Time Duration Model for Segment 1
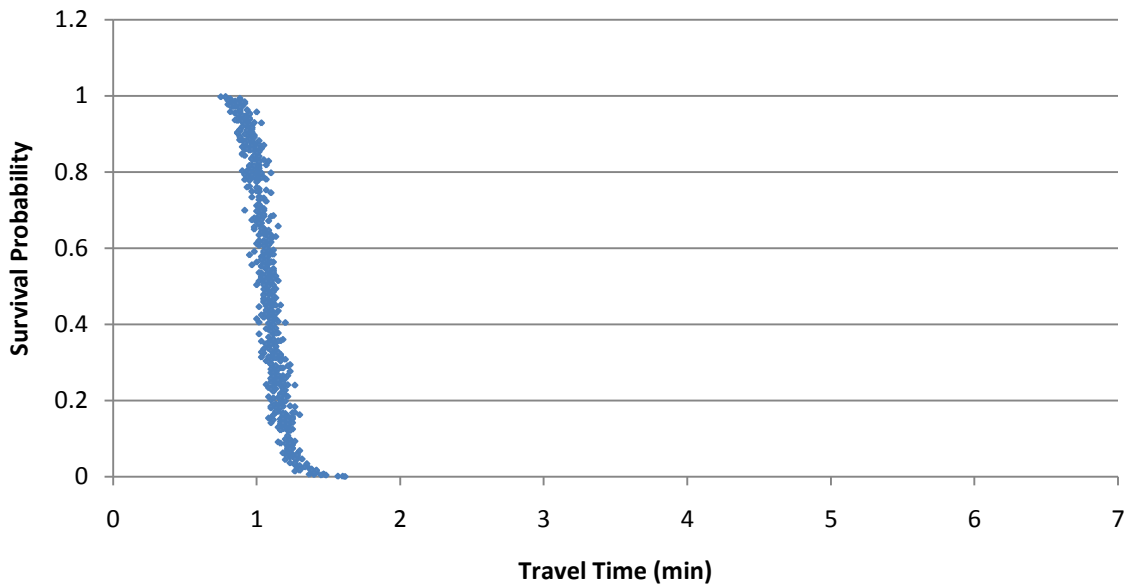
Figure 6.11 Survival Function for Night Time Duration Model for Segment 2

6.2.2 Peak Hour Model

The third model that was estimated is for peak hour conditions, lasting from 17:00 until 18:30. Once again the model was estimated to study the difference in how the explanatory variables affect the travel time during the congested traffic flow regime and changes in the hazard function. The resulting coefficient parameters and statistics are described in Table 6.3.

Similarly to the previous model estimates, all the parameter signs are and magnitudes are consistent with expectations. It appears that speed has a greater effect on the travel time during the peak hour than the overall model, while the preceding vehicle travel times have a smaller effect.

Table 6.3 Peak Hour Duration Model

Parameter	Description	Value	Standard Error	T-stat	p-value
Intercept		-0.19846	0.117726	-1.686	0.0918
Speed	Average 30-second segment travel speed (in mph)	-0.01662	0.000908	-18.303	<0.0000
Volume	Average 30-second volume one hour prior to the observation (in veh/hr/lane)	0.016688	0.003742	4.46	<0.0000
Distance	Segment length (in miles)	0.000175	.000016172	10.802	<0.0000
Lagged Travel Time	Travel time of the previous vehicle (in minutes)	0.113889	0.014068	8.095	<0.0000
Second Lag of Travel Time	Travel time of the second last vehicle (in minutes)	0.085658	0.013954	6.139	<0.0000
λ		0.645	0.045		
P		7.18	0.17		
Number of observations		1251			
Log likelihood		-30.7			

The hazard functions estimated for the vehicles during the pm peak hour, shown in Figure 6.12 and Figure 6.13, behave similarly to the hazard functions that include all the observations, initially increasing and then dropping. However, unlike the overall hazard function, it shows a significantly larger spread of rates and a lower maximum hazard rate.
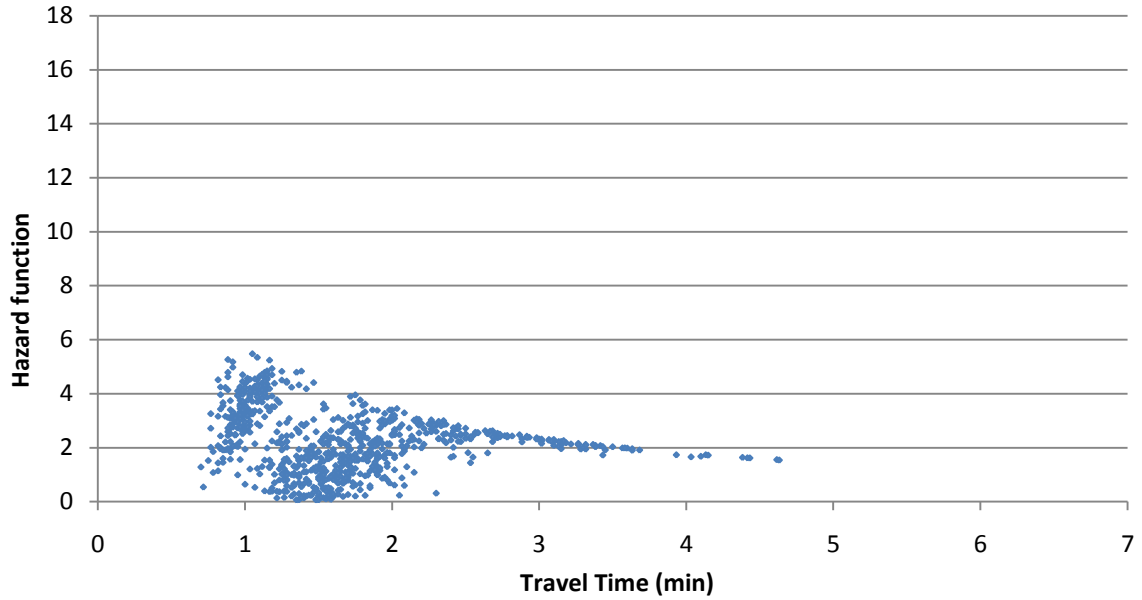
Figure 6.12 Hazard Function for Peak Hour Duration Model for Segment 1
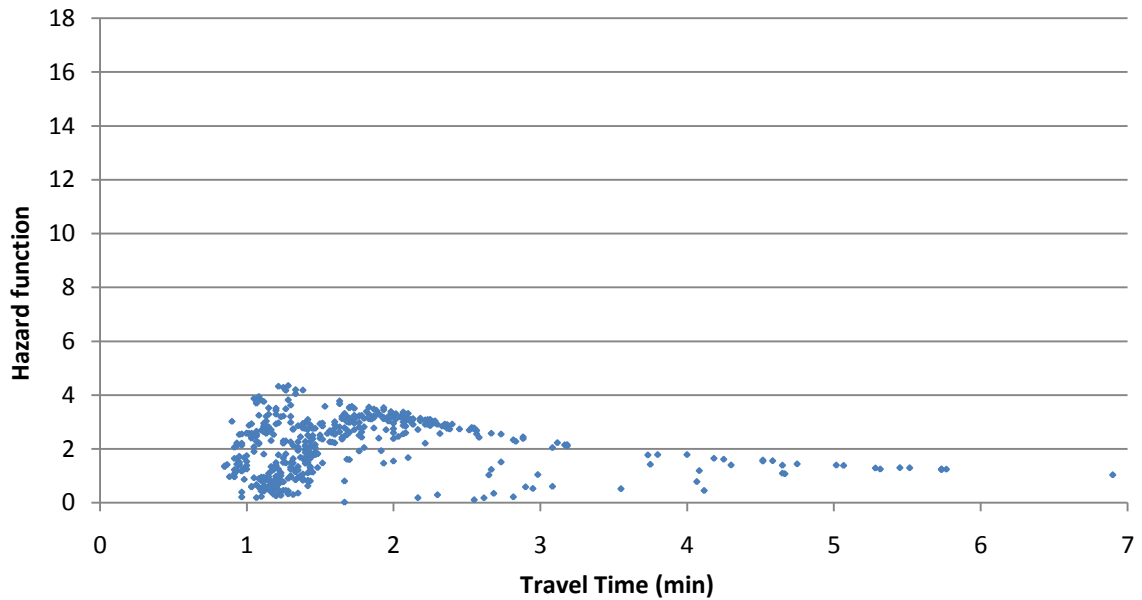
Figure 6.13 Hazard Function for Peak Hour Duration Model for Segment 2

The survival functions for peak hour model are expected to be shifted to the right, indicating lower probabilities of short travel times. Also, since there is a lot of variability

in speeds and lagged travel time variables, survival functions should have a larger spread. The resulting functions are shown in Figure 6.14 and Figure 6.15.
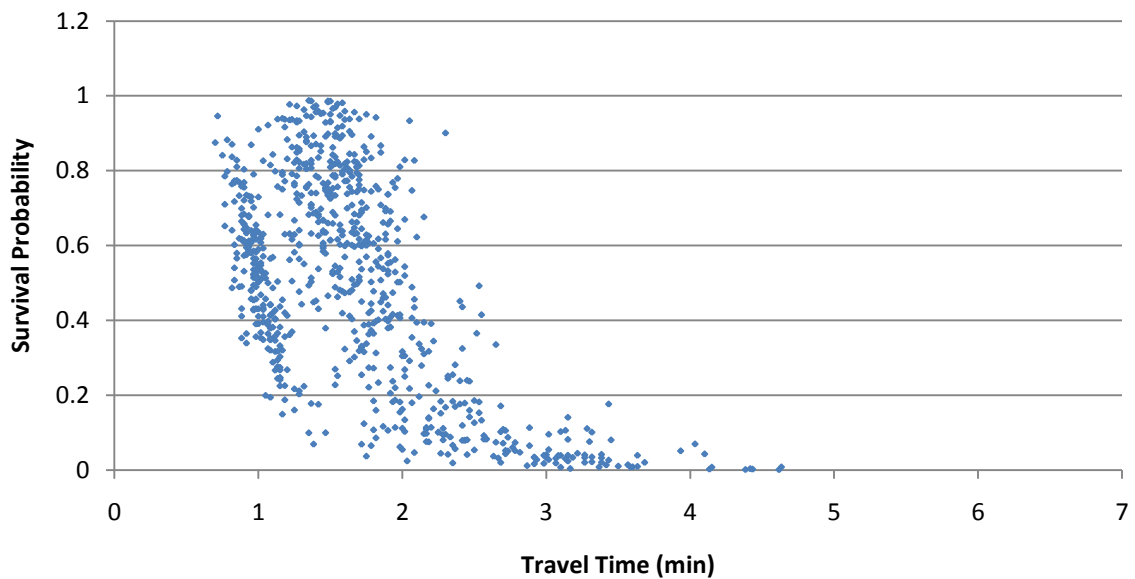

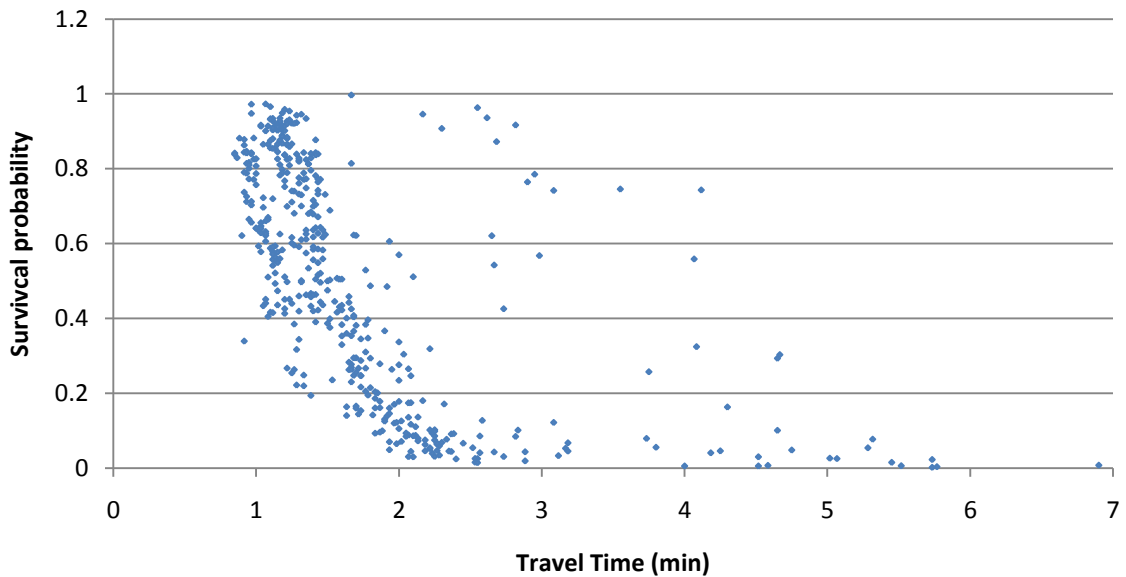Figure 6.14 Survival Function for Rush Hour Duration Model for Segment 1

Figure 6.15 Survival Function for Rush Hour Duration Model for Segment 2

6.3 Model Evaluation

Since the estimated model is non linear, R-squared cannot be used to evaluate the model fit. Instead the log likelihood of the converged model is compared to one with no coefficients. The converged log likelihood of the overall model is 5997, while the one with only the constant term is -4115. For the night time model the converged log likelihood is 1403; with only the constant it is 742. The converged rush hour model yields a log likelihood of -30.7, while the one with only the constant is -586. All models show a significant improvement over models with no coefficients.

Table 6.4 Comparison of Log Likelihood of Converged and Constant Only Models

Model	Log Likelihood Constant Only	Log Likelihood Converged
Overall	-4115	5997
Peak Hour	-586	-30.7
Night Time	742	1403

6.4 Discussion

The model indicates that the travel time is negatively affected by a decrease in average speed, an increase in volume one hour prior to observation period, as well as an increase in previous vehicle travel times. In addition, from the hazard functions it observed that during uncongested conditions, such as night time, the probability of reaching the end of the road segment increases with the travel time spent traversing the segment. However, in congested conditions the probability initially rises and then drops. This model enables to find the critical point after which the probability or exiting the segment drops corresponding to the onset of congestion. In addition, the survival function indicates how the probability of having any travel time varies depending on the time of day.

CHAPTER 7. TRAVEL TIME AND VARIANCE MODELING

As discussed before, travel time of vehicles traversing the same segment of the road at the same time can vary significantly. The standard deviation of travel time during a 15-minute interval ranges from 5% to 15% of the average travel time for uncongested conditions. During the peak hours, the standard deviation rises up to 50% of the average travel time. The inherent travel time variability within the traffic stream may be due to lane switching or traveling slower than the rest of the traffic stream. In the past, research has focused on identifying the increases in travel time due to the seven sources of unreliability discussed in the literature review section. The inherent variability in traffic flow has been overlooked until now. However, it is important to predict the variance as well as the travel time in order to provide users and decision makers with a better understanding of the traffic conditions. This chapter aims to predict the average travel time and standard deviation for 15-minute increments. This interval was selected because it provides sufficient data points for estimation while enabling to capture the short term perturbation in traffic flow.

7.1 Data

The modeling effort in this chapter uses the same travel time, speed and volume data as the previous sections. Data were collected from November 17, 2008 through November 22 and November 26. Unlike the individual vehicle travel time autoregressive model described in section 4, the proposed model aims to predict aggregate travel times and standard deviations of travel time. To prepare the dataset, average travel time and standard deviation of travel time was calculated for each time period. In addition, average speed and volume data were established for each of the 15-minute periods. The total

number of observations was 965. Other parameters that were included in the dataset were peak hour indicator and night time indicator. Lastly, since the travel time data were obtained from two segments of different lengths, a distance parameter was included.

7.2 Modeling Appoach

While it is possible to formulate two independent linear regression models for this prediction, this may not be the best modeling strategy. One of the standard linear regression assumptions states that the model has all of the information relating to the regression equation and variables, thereby making the estimated regression coefficients unbiased and efficient (Washington et al., 2003). However, if some information is not taken into account, the properties relating to the unbiasedness and efficiency of estimated coefficients cannot be determined. The omitted information may include the correlation between the error terms in the equations for average travel time and standard deviation. For example in the case of a bottleneck the travel time is likely to increase due to reduced travel speeds. The standard deviation of travel time would also be expected to rise because the traffic regime becomes unstable resulting in more variation. The Seemingly Unrelated Regression Equations (SURE) model can address correlation between the unobserved effects on travel time and the standard deviation. The SURE equation system is:

$TT=\beta_1 X+\varepsilon_1$

$STDEV=\beta_2 X+\varepsilon_2$ 						Eq. 7.1

where β_1 and β_2 are vectors of estimated parameters, X is a matrix of independent variables, ε_1 and ε_2 are error terms

In this case TT represents the average travel time during the 15 minute interval while STDEV corresponds to the standard deviation of travel time during the same time period. TT and STDEV variables do not have a direct effect on each other, however, since they are estimated during the same time period there may be unobserved parameters affecting both of them. The SURE formulation considers the correlation of disturbances

ε_1 and ε_2, making the estimates efficient and unbiased. Ordinary least squares parameters are estimated as,

$$\hat{\beta} = \left(X^T X \right)^{-1} X^T Y$$

<div align="right">Eq. 7.2</div>

where $\hat{\beta}$ is vector of parameters,

X is a matrix of data, X^T is the transpose of X, and Y is response vector

Seemingly unrelated equations are estimated using generalized least squares, which incorporates the correlation between error terms as follows,

$$\hat{\beta} = \left(X^T \Omega^{-1} X \right)^{-1} X^T \Omega^{-1} Y$$

<div align="right">Eq. 7.3</div>

where Ω is the correlation among equation error terms

7.3 Model Estimation

The resulting seemingly unrelated equations are:

$$TT = 1.58 + 0.00003 * \text{Volume} - 0.02 * \text{Speed} - 0.089 * \text{Night} + 0.222 * \text{Peak}$$
$$+ 0.695 * \text{Distance}$$

$$STDEV = 0.512 - 0.008 * \text{Speed} - 0.063 * \text{Night} + 0.125 * \text{Peak}$$
$$+ 0.166 * \text{Distance}$$

<div align="right">Eq. 7.4</div>

The statistics for the travel time and standard deviation models are shown in Table 7.1 and Table 7.2, respectively. Parameters that were insignificant at the 95% level were omitted from the models.

Table 7.1 SURE Model of Travel Time

Parameter	Explanation	Value	Error	T-stat	p-value
Constant	Intercept	1.584347	0.073171	21.653	0
Volume	15-minute average volume one hour before (vphpl)	.0000287	.0000093	3.09	0.002
Speed	15-minute average speed during the previous time period (mph)	-0.01967	0.001031	-19.087	0
Night Time	Indicator variable equal to 1 between 21:00 and 6:00	-0.10147	0.013477	-7.529	0
Peak Hour	Indicator variable equal to 1 between 17:00 and 18:30	0.327115	0.025205	12.978	0
Distance	Length of the segment (miles)	0.693571	0.032082	21.618	0
Number of Observations		965			
Adjusted R-squared		0.63			

Table 7.2 SURE Model of Standard Deviation

Parameter	Explanation	Value	Error	T-stat	p-value
Constant	Intercept	0.535054	0.04364	12.261	0
Speed	15-minute average speed during the previous time period (mph)	-0.00873	0.000627	-13.916	0
Night	Indicator variable equal to 1 between 21:00 and 6:00	-0.07212	0.007718	-9.344	0
Peak	Indicator variable equal to 1 between 17:00 and 18:30	0.169313	0.015102	11.211	0
Distance	Length of the segment (miles)	0.166265	0.019129	8.692	0
Number of Observations		965			
Adjusted R-squared		0.46			

The model estimates indicate that higher speed results in lower travel time because travel time is inversely proportional to speed. Higher segment volume one hour

prior to the observed travel time results in higher travel time, which is a result of heavier traffic. The reason that the volume one hour prior to the observation is considered is because once the congestion has set in and the speed has dropped, number of vehicles traversing the segment decreases. Thus, the highest volumes are observed prior to the most congested condition. At night, the travel time is also lower due to fewer vehicles being on the road. Meanwhile during the peak hours, which last from 17:00 until 18:30, the travel time is higher due to more congested conditions. Lastly, longer segment leads to a higher travel time, which is intuitive because travel time is proportional to distance. In the models of standard deviation of travel time, the parameters are the same as the travel time with the exception of the segment volume. Analogously to the travel time, higher speed and night time lead to a lower deviation, while peak hour factor increases the standard deviation. Distance between the detectors also has a positive effect on the standard deviation because the difference in travel time between a slow and fast driver on a longer segment is greater.

7.4 Model Evaluation

In order to evaluate the model fit, first the Chi-squared statistic is calculated. This statistic compares the estimated model to the one with no coefficients; it is equal to:

$$\chi^2 = 2[LL(\beta) - LL(0)] \qquad\qquad \text{Eq. 7.5}$$

Chi-squared values for the travel time and standard deviation models are 931 and 595, respectively. These values indicate that the estimated models are significantly different from the no coefficients models. Another way to assess the fit of the seemingly unrelated equation model is through the use of the R-squared value. The adjusted R-squared statistic is 0.63 for the travel time model and 0.46 for the standard deviation model. In addition to evaluating the model fit, it is important to ensure that all the parameters included in the models are relevant. As shown in Tables 7.2 and 7.3, the t-statistics for all parameters are above 1.96, making them significant at the 95% confidence level.

Another way to assess the effectiveness of the model is by visual assessment of predicted and actual travel times and standard deviations. Figure 7.1 and Figure 7.2 display actual and predicted average travel times for segments 1 and 2 on Wednesday November 20. Figure 7.3 and Figure 7.4 show the standard deviations on the two segments on November 20.
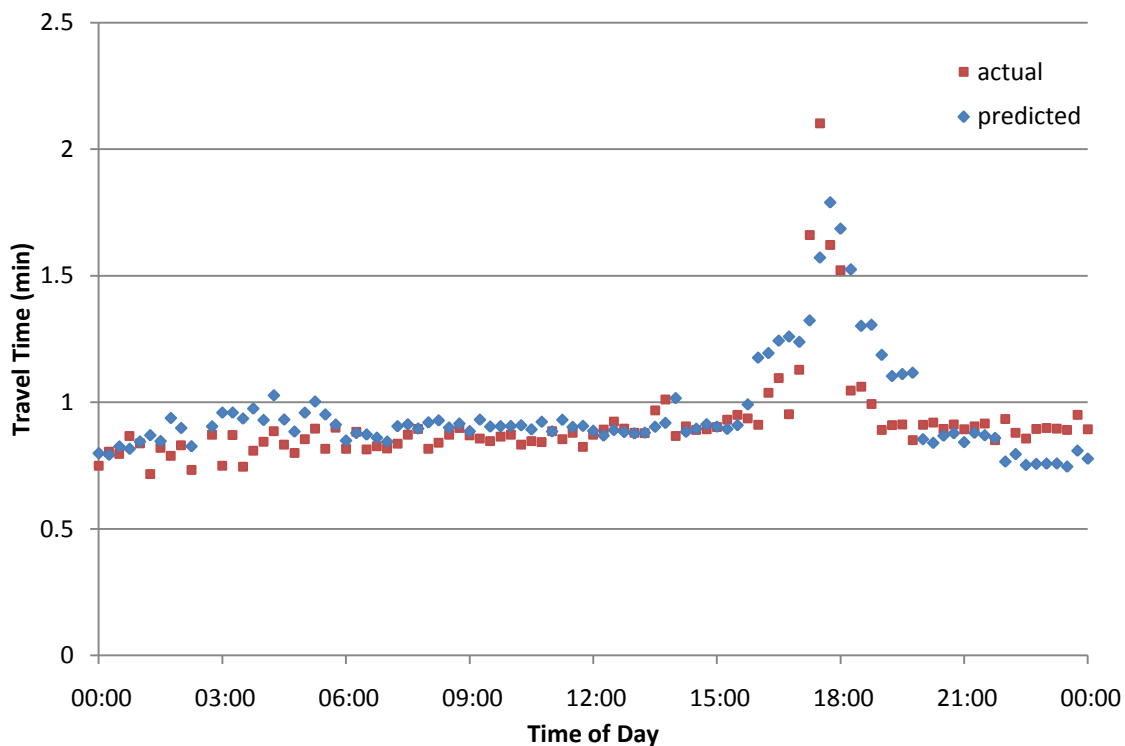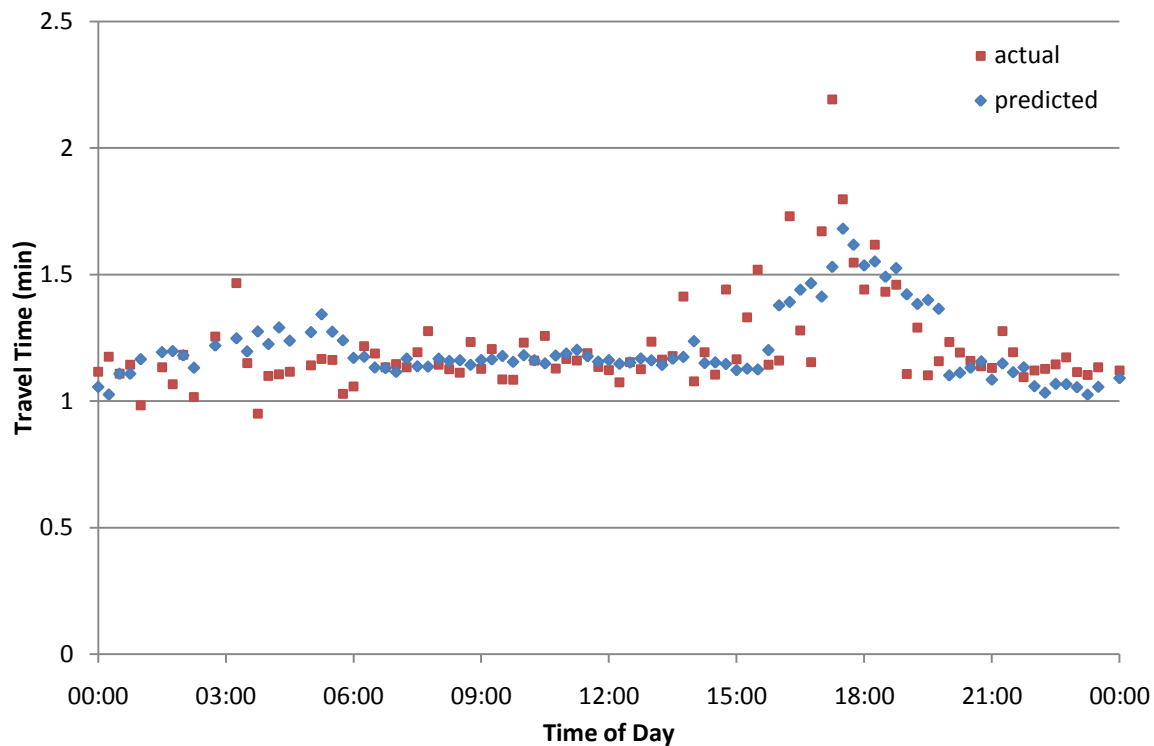

Figure 7.1 Segment 1 Travel Time on November 20

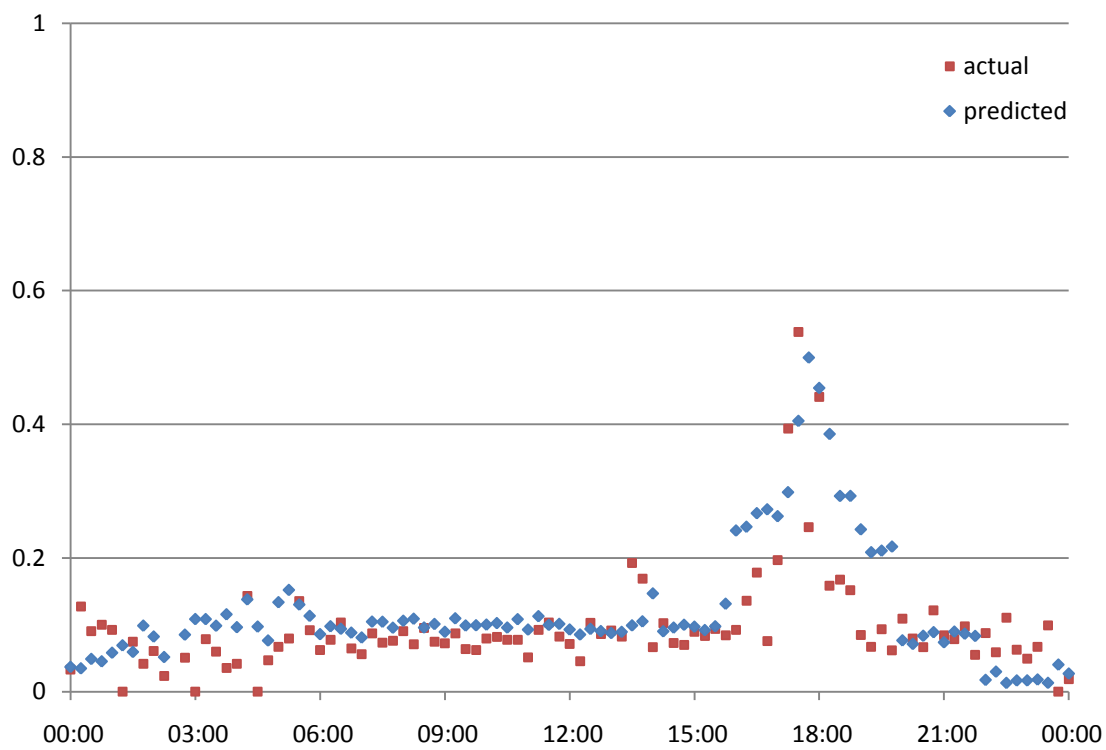Figure 7.2 Segment 2 Travel Time on November 20

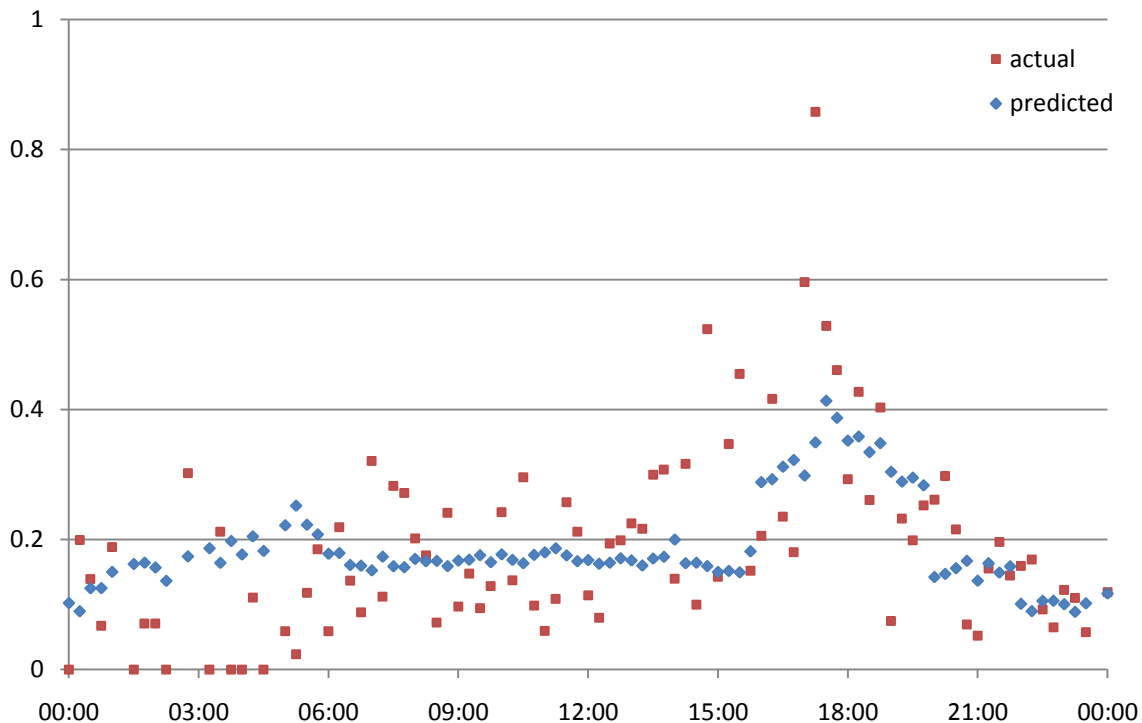Figure 7.3 Segment 1 Standard Deviation of Travel Time

Figure 7.4 Segment 2 Standard Deviation of Travel Time

It is evident from the graphs that the estimated model is effective at predicting the travel time and standard deviation on segment 1. The model also yields good results for the average travel time on segment 2, however, the standard deviation prediction is not as accurate. One reason for this inaccuracy is the presence of an exit ramp on the second segment which experiences extensive traffic backlogs. As a result there's spillover into the right lane which makes the traffic flow very inconsistent and difficult to predict.

Overall, the seemingly unrelated equations for travel time and standard deviation of travel time yield a good fit and ensure statistical validity of the estimates.

CHAPTER 8. SUMMARY AND CONCLUSIONS

The subject of travel time reliability has gained popularity in recent years. In general, sources of travel time variability can be grouped into recurrent, caused by insufficient capacity, and non-recurrent ones including incidents, weather and special events. Most commonly used methods to describe travel time reliability include statistical indices, buffer measures and tardy trip indicators. These measures are effective at conveying historical trends to the users, however, cannot capture the current traffic conditions. Moreover, they do not include inherent variability due to vehicles traveling at different speeds.

The study describes how actual travel time data can be collected and processed using Bluetooth technology along two segments of a busy Interstate 69 in Indianapolis. Travel times were collected for a period of two weeks and analyzed. Using these data, the inter-vehicle, inter-period and inter-day variability is evaluated. From the collected data, it is clear that at any point during the day, the travel time may vary by up to 100% of the average travel time. We speculate that this is due to individual driver behavior and different traffic speeds in adjacent lanes. This is a significant variation that needs to be taken into account when providing travel information to travelers. The inter-period variability is also pronounced in the collected data. The average travel time is observed to triple during the peak hours (as one might expect) and the standard deviation of individual travel times also increased during peak periods. And, despite a common conclusion that travel times during the same day of the week are fairly consistent, a comparison between two of the same days of the week indicates that it varies significantly. This variability is particularly pronounced during the peak hours.

The study also considers variability due to adverse weather. Travel time data collected during a snow event are compared to data obtained during adequate weather

conditions. The analysis shows that both travel time and standard deviation of travel time are significantly different under adverse and normal weather conditions. In addition, the effect of the snow event can last for many hours thereafter due to the roadway not being cleared. This presents another difficulty in modeling the traffic following a snow event because the conditions on the road may depend on the response time of snow clearing personnel.

The study considers three statistical models of the travel-time data. An autoregressive individual vehicle travel time model is estimated to determine factors that may generate variability. This model includes data obtained from Remote Traffic Microwave Sensors, including speed and volume as well as time of day information. It is found that average segment travel time, lagged travel time, volume and peak hour indicator all positively affect individual vehicle travel time. Meanwhile the night time parameter has a negative impact on travel time.

Next, duration models of travel time were estimated to provide insight as to when the variance of travel times become a problem. Duration models are useful in predicting how the probability of a car's duration of time on a roadway segment changes over time. Three duration models are estimated: all hours, peak hour and night time models. These models show that the point where the conditional probability of travel times becoming longer occurs roughly at the onset of level-of-service F conditions.

The third model predicts average travel time (for 15-minute intervals) as well as the inherent variability in this travel time over the 15-minute interval. This model can be used to predict traffic conditions in real time and convey the information to the roadway users via Variable Message Signs and web utilities.

Overall the study provides a framework for collecting and analyzing travel time data using Bluetooth technology and supplementing it with corresponding remote traffic microwave sensors data. In addition, the three statistical modeling alternatives used to establish a relationship between travel times and known traffic parameters show that there is great potential to be able to predict, with reasonable accuracy, mean travel times and travel time variability.

Applying the methods and techniques demonstrated in this study to other freeway roadway segments in Indiana and the US should provide even more useful insights into this important problem.

REFERENCES

Abdel-Aty, M., Kitamura, R., & Jovanis, P. (1995). Investigating effect of travel time variability on route choice using. *Transportation Research Record 1493*, 159-169.

Bates, J., Dix, M., & May, A. (1989). Travel time variability and its effect on time of day choice for the journey to work. *Proc. Seminar C, PTRC*, 293-311.

Bremmer, D., Cotton, K. C., Cotey, D., Prestrud, C., & Westby, G. (2004). Measuring congestion: Learning from operational data. *Transportation Research Record 1895*, 188-196.

Chang, J. S. (2009). Assessing travel time reliability in transport appraisal. *Journal of Transportation Geography*, In Press.

Chen, C., Petty, K., Skabardonis, A., Varaiya, P., & Jia, Z. (2001). Measurement system: Mining loop detector data. *Transportation Research Record 1748*, 96-102.

Cortes, C. E., Lavanya, R., Oh, J., & Jayakrishnan, R. (2002). General-purpose methodology for estimating link travel time with multiple-point eetection of traffic. *Transportation Research Record 1802*, 181-189.

Elefteriadou, L. (2006). A framework for defining and estimating travel time reliability. *85th Annual Transportation Research Board Conference*. Washington, DC.

Federal Highway Administration. (2005). *Traffic Congestion and Reliability: Trends and advanced strategies for congestion mitigation.* Cambridge Systematics Inc. and Texas Transportation Institute.

Highway Capacity Manual (2000). *Highway Capacity Manual.* Washington, D.C.: National Research Council.

Lint, J. W., & Zijpp, N. J. (2003). Improving a travel-time estimation algorithm by using dual loop detectors. *Transportation Research Record 1855*, 41-48.

Liu, H. X., Recker, W., & Chen, A. (2004). Uncovering the contribution of travel time reliability to dynamic route choice using real-time loop data. *Transportation Research Part A*, 38(6), 435–453.

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

Mackie, P. J., S. Jara-Diaz, S., & and Fowkes, A. S. (2001). The value of travel time savings in evaluation. *Transportation Research Part E*, 37 (2-3), 91-106.

Mannering, F. L., Abu-Eisheh, S. A., & Arnadottir, A. T. (1990). Dynamic Traffic Equilibrium with Discrete/Continuous Econometric Models. *Transportation Science*, 24 (2), 105-116.

Quiroga, C. A., & Bullock, D. (1998). Travel time studies with global positioning and geographic information systems: an integrated methodology. *Transportation Research Part C*, 6(2), 101-127.

Robertson, H. D. (1994). *Travel time and delay studies.* Washington, D.C.: Manual of Transportation Engineering Studies 4th ed., Institute of Transportation Engineers, Washington, DC..

Skabardonis, V. P. (2001). Measurement system: Mining loop detector data. *Transportation Research Record 1748*, 96-102.

Small, K. A., Noland, X. C., & Lewis, D. (1999). *NCHRP Report 431: Valuation of Travel Time Savings and Predictability in Congested Conditions for Highway User-Cost Estimation.* Washington, D.C.: National Research Council.

Tufte, K., & Kothuri, S. (2007). *Assessment and Refinement of Real-Time Travel Time Algorithms for use in Practice.* Oregon: Oregon Department of Transportation.

Turner, S. M. (1996). Advanced techniques for travel time data collection. *Transportation Research Record 1551*, 51-58.

Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2003). *Statistical and Econometric Methods for Transportation Data Analysis.* Chapman&Hall/CRC.

Wasson, J., Sturdevant, J., & Bullock, D. (2008). Real-time travel time estimates using media access control address matching. *Institute of Transportation Engineers Journal*, July, 20-23.