Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a, user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use" that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select "Pages from: first page # to: last page #" on the print dialog screen

ABSTRACT

OPTIMIZATION OF PATH BASED SENSOR SPACING ON A FREEWAY SEGMENT FOR TRAVEL TIME PREDICTION DURING INCIDENTS

by
Patricia Kathleen DiJoseph

Congestion on freeways is increasing and a key source of it is non-recurring incidents. Accurate vehicle travel time predictions are needed during these incidents in order for roadway users to make informed trip decisions. Path based sensors are becoming a leading technology in gathering real-time travel time data. The data is used to make travel time predictions that are then provided through various means, such as dynamic message signs, to roadway users. These types of sensor are located at stationary points along a roadway and collect individual vehicle travel time data from vehicles as they drive pass the sensors.

The accuracy of the predictions, in terms of representing future travel times, is dependent on many factors including the sensor spacing along the roadway, the duration and location of a traffic incident, and the uncongested and congested traffic speeds and traffic flows. Understanding the relationship between the travel time prediction accuracy and the different variables is necessary to optimize sensor spacing. In addition, because incidents occur at different times of the day, have varying durations, occur at different locations, and cause different capacity reductions depending on the severity of the incident, the sensor spacing cannot be based on one incident scenario. Instead, multiple incident scenarios, along with the probability of each occurring, needs to be taken into account.

Path based sensor spacing during incidents on a freeway segment is optimized in this dissertation. In addition, the marginal benefit of additional sensors is calculated. A mathematical model and a solution methodology are developed. The mathematical model applies macroscopic traffic principles and shock wave theory. It calculates the travel time prediction error by sensor spacing during an incident on a freeway segment.

The solution algorithm consists of four main steps. First, historical incident data for the roadway are gathered. Second, the mathematical model is applied to determine the average travel time prediction error by sensor spacing for each of the historical incidents. Third, the weighted average travel time prediction error by sensor spacing is calculated, which considers all the possible incidents and the frequency of each occurring. Fourth, the optimal spacing is chosen which minimizes the weighted average error.

The applicability of the model and solution methodology is demonstrated through a case study of a ten mile freeway segment in Northern New Jersey.

OPTIMIZATION OF PATH BASED SENSOR SPACING ON A FREEWAY SEGMENT FOR TRAVEL TIME PREDICTION DURING INCIDENTS

by
Patricia Kathleen DiJoseph

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Transportation

Department of Civil and Environmental Engineering

August 2013

.

APPROVAL PAGE

OPTIMIZATION OF PATH BASED SENSOR SPACING FOR TRAVEL TIME PREDICTION UNDER CONGESTION

Patricia Kathleen DiJoseph

Dr. Steven Chien, Dissertation Advisor — Date
Professor of Civil and Environmental Engineering, NJIT

Dr. Athanassios Bladikas, Committee Member — Date
Associate Professor of Mechanical and Industrial Engineering, NJIT

Dr. Janice Daniel, Committee Member — Date
Associate Professor of Civil and Environmental Engineering, NJIT

Dr. Lazar Spasovic, Committee Member — Date
Professor of Civil and Environmental Engineering, NJIT

Dr. Jian Yang, Committee Member — Date
Associate Professor of Management Science and Information Systems, Rutgers

BIOGRAPHICAL SKETCH

Author: Patricia Kathleen DiJoseph

Degree: Doctor of Philosophy

Date: January 2013

Undergraduate and Graduate Education:

- Doctor of Philosophy in Transportation,
  New Jersey Institute of Technology, Newark, New Jersey, 2013

- Master of Science in Transportation,
  New Jersey Institute of Technology, Newark, New Jersey, 2008

- Bachelor of Arts in Mathematics,
  Rutgers College, New Brunswick, New Jersey, 2005

Major: Transportation

Presentations and Publications:

DiJoseph, P. Optimization of the Number of Automatic Vehicle Identification Sensors for Accurate Travel Time Prediction, *Dana Knox Student Research Showcase*, New Jersey Institute of Technology, Newark, NJ, April 2013.

DiJoseph, P. and Chien, S., Optimizing Sustainable Feeder Bus Operation Considering Realistic Networks and Heterogeneous Demand, *Journal of Advanced Transportation*, Article first published online: May 2011, DOI: 10.1002/atr.174.

DiJoseph, P. and Chien, S., Optimal Service Planning for a Sustainable Transit System, *50th Annual Forum of the Transportation Research Forum*, Portland, OR, March 2009.

This dissertation is dedicated to the women who study Mathematics and Engineering and to those who give them the love, encouragement, and support to do so.

ACKNOWLEDGMENT

I thank my dissertation advisor, Dr. Steven I-Jy Chien, for his dedication, insight, and guidance. I also thank him for his confidence in me to achieve a PhD. I'd also like to thank my dissertation committee members: Dr. Athanassios Bladikas, Dr. Janice Daniel, Dr. Lazar Spasovic, and Dr. Jian Yang.

To my fellow graduate students, thank you for walking with me on this journey, Dr. Kitae Kim, Ms. Jiruttichut Leoviriyakit, Dr. Dejan Besenski, and Mr. Branislav Dimitrijevic.

Without the support of my family, I would never have pursued and completed a doctorate. My parents, Clare and John, my sisters Joanna, Theresa, and Lindsay, and my brother-in-law William, thank you so much for your love and encouragement.

I also appreciate support from the New Jersey Institute of Technology, the United States Department of Transportation, and the New Jersey Department of Transportation.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES
(Continued)

CHAPTER 1

INTRODUCTION

1.1 Background

Traffic congestion on urban freeways cost Americans over $48 billion in 2011. This included 2.2 billion hours of extra travel time and 1.16 billion gallons of wasted fuel. In addition, congestion diminished travel time reliability, thus drivers had to build buffer time into their trips to avoid being late. In fact, to ensure being on time for an important trip that would only take twenty minutes under free flow conditions, drivers had to allot an extra forty minutes of travel time (Schrank et al. 2012).

About half of all traffic congestion is caused by temporary disruptions that reduce the roadway capacity. These include incidents (25% of congestion), work zones (10% of congestion) and weather (15% of congestion) (Cambridge Systematics 2005). These non-recurring events dramatically reduce the reliability of the entire transportation system as they occur irregularly and are unexpected by drivers (FHWA 2011). The sources of congestion are shown in Figure 1.1.

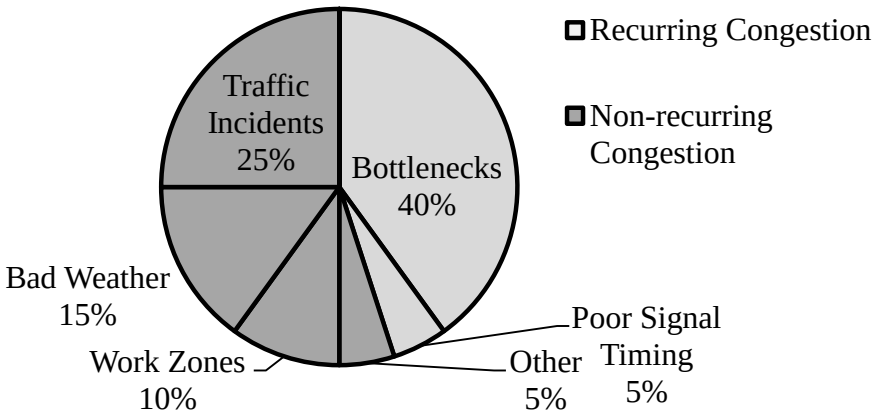


Figure 1.1 Sources of congestion.
Source: http://www.ops.fhwa.dot.gov/aboutus/opstory.htm

The Federal Highway Administration lists several strategies to reduce congestion and improve travel time reliability. These include the provision of traveler information. This can be accomplished through traveler information telephone services and travel time message signs. Also, they cite Section 1201 of the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU) which calls for the establishment of "a real-time system management information program to provide, in all States, the capability to monitor, in real-time, the traffic and travel conditions of the major highways of the United States and to share that information…" (FHWA 2011).

There has been an increase in the deployment of intelligent transportation systems that form the basis for real-time management of traffic on freeways including real-time traffic data collection technologies and dynamic message signs (DMS). The percentage of freeway miles with real-time data collection technologies increased 37% from year 2000 to year 2010. Also, the number of DMS deployed increased by 2,400 over the same time period. Figure 1.1 below illustrates these trends with the percentage of freeway miles on the left side y-axis and the number of DMS deployed on the right side y-axis (Gordon 2011).

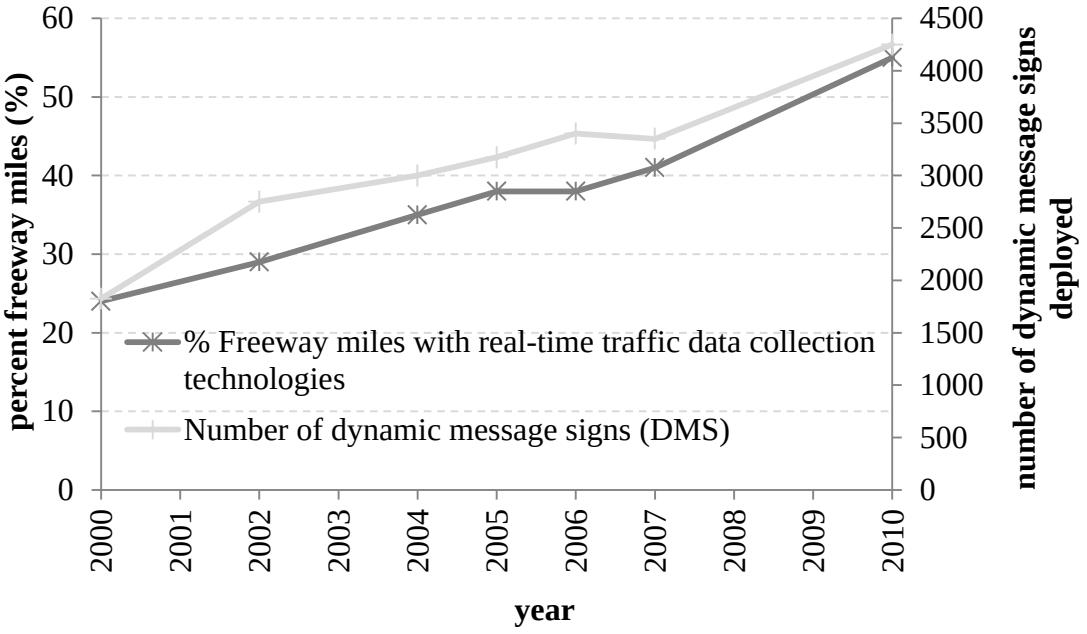


Figure 1.2 Freeway management deployment indicators.

Source: http://www.itsdeployment.its.dot.gov/FM.aspx

Data collection technologies can be categorized into two groups (Hagemann et al. 2010). The first group is point based sensors which historically have been used for highway data collection. These include loop detectors, microwave radar, and video image processing. These sensors detect anonymous vehicles' speeds which are then converted to travel time data by assuming the speed is constant for a freeway segment. Inductive loops have been the primary highway data collection devices since their inception in the 1960s. However, new technologies are now emerging and are becoming more widely used. Path based sensors were introduced in the mid-1990s as an alternative technology. These devices record when a probe vehicle reaches the sensor location and thus the probe vehicle's travel time can be calculated by comparing its time stamps at different locations. These include Bluetooth, automatic vehicle identification (AVI) sensors, toll tag readers, and GPS. Path based sensors' market share has been increasing; it went from

17% in 2004 to 37% in 2010 (RITA 2011). They are an improvement over point based sensors because they require less infrastructure investment and generate more direct travel time data (Hagemann et al. 2010). In addition, they are less expensive. To illustrate, the lifecycle cost for a 6 lane highway in 2012 dollars, including capital, maintenance, and operations costs, of an inductive loop detection system is $239,995 whereas it is $133,461 for the the path based sensor system TRANSMIT (Mouskos 1998).

As the use of path based sensor technologies continues to grow there is a need for understanding how to best achieve accurate travel time predictions from their data. This includes their optimal sensor spacing. As non-recurring congestion has a great negative impact on travel time reliability and incidents represent the largest portion of non-recurring congestion, it is very desirable to optimize path based sensor spacing under incident conditions.

1.2 Problem Statement

Traffic incidents represent the worst case scenario for vehicle travel time prediction because roadway conditions transition when the incident occurs and again when the incident is cleared (Bertini 2007). To assist roadway users trying to determine their travel times, path based sensors can be deployed on the roadway to collect real-time travel information during the incidents. The real-time travel information is then provided to the users as they approach the roadway, by various means such as dynamic message signs. The accuracy of this information in terms of representing future travel times is dependent not only on the sensor spacing but also on the roadway characteristics and incident characteristics. Therefore, understanding the relationship between the accuracy

and different variables is necessary to optimize the sensor spacing on a freeway segment. In addition, because incidents occur at different times of the day, have varying durations, occur at different locations, and cause different capacity reductions depending on the severity of the incident, the sensor spacing cannot be chosen just based on considering one incident scenario. Instead, multiple incident scenarios, and each of their related function of travel time prediction error by sensor spacing, need to be taken into account along with the probability of each occurring.

Although much research has been conducted to determine the optimal sensor spacing of point based sensors, the results cannot be applied to path based sensors because the sensors operate differently. One key difference is that point based sensors record speeds and path based sensors record travel times. Also, path based sensors work in pairs and point based sensors can produce results independent of each other. Furthermore, while point based sensors detect all vehicles on the roadway, path based sensors detect only those equipped with in-vehicle sensor technology such as Bluetooth devices or toll tags.

Research that has been conducted on the optimal spacing of path based sensors has not considered their performance during incidents. Instead, they focus on the probe sample size achieved, the origin-destination pairs covered by the sensors, or the assumed travel time distribution for the roadways.

It is desirable to develop an analytical model to optimize path based sensor spacing for a freeway segment such that the average travel time prediction error during incidents is minimized. Such a model can reliably evaluate various sensor spacing

considering different incident characteristics and can determine the optimal sensor spacing realizing that the type of incident that affects one freeway segment varies.

1.3 Objective and Work Scope

The objective of this research is to optimize path based sensor spacing on a freeway segment such that the average travel time prediction error per vehicle during traffic incidents is minimized. To formulate the optimal solution, different incident scenarios that occur on the freeway segment are taken into consideration, along with the probability of each scenario occurring.

To reach this objective a set of tasks has been compiled for this research project. These are as follows:

- Set the objective of the study.
- Conduct a literature review of previous studies regarding travel time prediction and traffic incidents.
- Model the travel time prediction error produced by path based sensors during transitions in the traffic state of a freeway segment. The transitions include one from an uncongested traffic state to a congested traffic state and vice versa.
- Develop an algorithm to optimize path based sensor spacing considering the range of possible incident locations, clearance times, and v/c ratios.
- Conduct a numerical analysis of both the model and algorithm. This includes comparing the modeled error with the real-world error, applying the algorithm to a case study, calculating the marginal benefit of additional sensors, and performing a sensitivity analysis.
- Summarize the findings from the research and suggest future applications of it.

Figure 1.3 below illustrates the work scope.

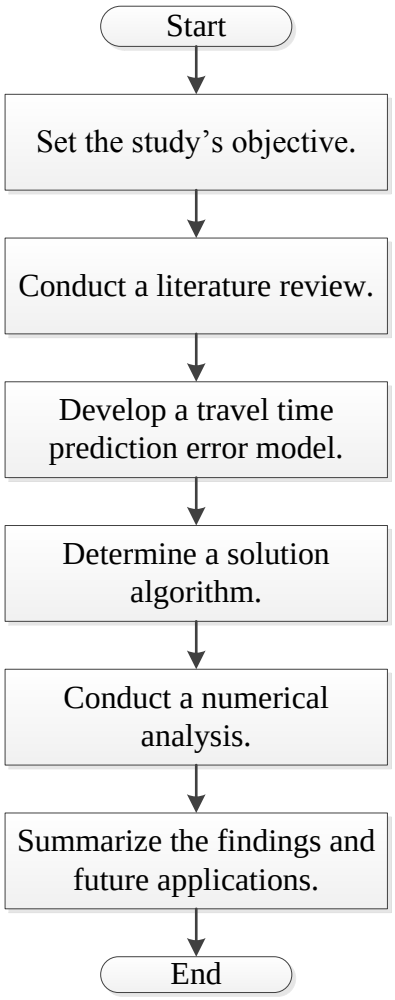


Figure 1.3 Work scope.

1.4 Methodology

In order to optimize sensor spacing for a freeway segment, a mathematical model and a solution algorithm are developed. First, the relationship between sensor spacing and travel time prediction error during an incident is modeled. Then, a step procedure to determine the optimal sensor spacing is determined.

The mathematical model applies macroscopic traffic principles and shock wave theory. The decision variables of the macroscopic model include the uncongested and

congested traffic flows and vehicle speeds and the incident's location and duration. Also included are the freeway segment length and sensor spacing. The outputs of the model are the relationship between vehicle travel time and vehicle departure time and the relationship between predicted travel time and vehicle departure time. By comparing these relationships, the travel time prediction error by departure time is calculated along with the average error for the incident study period.

The solution algorithm consists of four main steps. First, historical incident data is gathered including the frequency of each incident occurring. Second, the mathematical model is applied to determine the average travel time prediction error by sensor spacing for each of the historical incidents. Third, the weighted average error by sensor spacing is calculated, which considers all the possible incidents and the frequency of each occurring. Fourth, the optimal spacing is chosen which minimizes the weighted average error.

1.5 Dissertation Organization

This dissertation is organized into six chapters. Chapter 1 explains the motivation for this dissertation along with its objectives and work scope. Chapter 2 summarizes the efforts of previous studies related to various aspects of vehicle travel time prediction and optimal sensor spacing. Chapter 3 models the travel time prediction error by sensor spacing that occurs during the passage of shock waves on a freeway segment. Chapter 4 provides a methodology to optimize the sensor spacing. Chapter 5 numerically analyses the model developed in Chapter 3, and the methodology developed in Chapter 4, through a real-world scenario and a case study. Finally, Chapter 6 summarizes the finding from the dissertation and suggests future applications for it.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

The purpose of the literature review is to provide the necessary background information to develop the models and the methodology presented in this proposal. It is divided into seven sections. In the first section macroscopic traffic flow theory is discussed. The second section provides information on different types of sensors for travel time data collection. The third section outlines different types of travel time predictions and their benefits. The fourth section reviews methodologies that have been developed to optimize the spacing or allocation of sensors to freeway segments. The fifth section is on the relationship between probe sample size and its effect on travel time prediction. The sixth section summarizes incident duration prediction models. The last section provides guidelines on using traffic microscopic simulation software.

2.2 Macroscopic Traffic Flow Theory for Transitions in Traffic States

Traffic stream parameters fall into two broad categories: macroscopic and microscopic. Macroscopic parameters describe the traffic stream as a whole and there are three principle parameters. The first is volume, or rate of flow, and it is measured in vehicles per hour (vph). The second is speed and it is measured in miles per hour (mph). The third is density and it is measured in vehicles per mile (vpm). On the other hand, microscopic parameters describe the behavior of individual vehicles or pairs of vehicles within the traffic stream. Examples of microscopic parameters include the speed of individual vehicles (mph), headway (i.e., the time interval between successive vehicles as

they pass a point) (seconds), and spacing (i.e., the distance between successive vehicles in a traffic lane) (ft) (Roess et al. 2004).

The fundamental hypothesis of macroscopic traffic flow theory states that at any point on a road the flow is a function of the density. It also states that the speed equals the flow divided by the density. Lighthill and Whitham described the traffic flow function as a single curve (a parabola) in the flow versus density plane. Every point along the curve represents a possible traffic state the roadway can be in. The exact values of the function are particular to the roadway and can vary by time of day. However, a flow rate of zero vph always occurs at both a density of zero vph and at a very high density, called a jam density. The jam density occurs when there are so many vehicles on the roadway that motion stops. Between these two points, at the vertex of the curve, is the critical density where the maximum flow occurs. The traffic states that are represented on points of the curve that fall to the left of the critical density are classified as uncongested while traffic states that fall to the right of the critical density are classified as congested (Lighthill and Whitham 1955).

Newell (1993) expanded on the traffic flow function in Lighthill and Whitham (1955) by proposing a simplified, triangular shaped traffic flow function. In this research, the flow-density function was described as linearly increasing from the origin to the critical density and then linearly decreasing to the jam density. Windover and Cassidy (2001) empirically verified that this simplification of the function is reasonably accurate.

Shock wave analysis for transportation facilities was published by Richards (1955). It has been expanded upon in other research including Lighthill and Whitham

(1955) and Messer et al. (1973) who used shock wave analysis and the Greenshields' macroscopic traffic flow models to predict individual travel time and trajectories during a traffic incident. Also, Wirasinghe (1978) determined traffic delay from shock waves and May (1990) explained traffic flow fundamentals and shock waves.

According to the rules of macroscopic traffic dynamics, when the traffic state of a roadway changes, such as from one point on the flow-density function to another, a clear boundary is established in the time-space plane that demarks the old and new traffic states. This boundary is referred to as a shock wave. For example, when an incident occurs that blocks a lane, it reduces the available capacity. This causes a reduction in the traffic flow and an increase in the traffic density. The change in traffic state is not static. Instead, it changes over space and time. The reduced flow and increased density will not only occur at the incident location, but as time progresses will spread upstream of the incident location (May 1990).

A shock wave's characteristics include its speed, direction of travel, and whether its propagation results in an increase or decrease in the length of roadway that is congested. These characteristics can be visualized with a diagram of the flow-density function by a line drawn from the point on the flow-density function that represents the old traffic state to the point that represents the new traffic state. First, the slope of this line equals the shock wave speed, and therefore the speed equals the change in flow divided by the change in density. Second, the drawn line can have either a positive slope, negative slope, or a slope equal to zero. A positive slope means the wave will travel upstream, in the direction of traffic. A negative slope means the wave will travel downstream, in the direction against traffic. A slope equal to zero means the wave is

stationary. Third, if the traffic state changes from being uncongested to congested then the propagation of the shock wave results in an increase in the congested portion of the roadway. In contrast, if the traffic state changes from being congested to uncongested then the propagation of the shock wave results in a decrease in the congested portion of the roadway (May 1990).

A naming convention for the types of shock waves has been established in the literature. It is based on the waves' characteristics and is summarized in the table below.

Table 2.1 Shock Wave Naming Convention

		Wave's direction of travel	
		Downstream (with traffic)	**Upstream (against traffic)**
Change in the traffic state	**Uncongested to congested**	forward shock wave	backward shock wave
	Congested to uncongested	forward recovery wave	backward recovery wave

Source: (May 1990)

In addition, a frontal stationary shock wave indicates a wave that is at the downstream edge of a congested region and remains at the same position in space. A rear stationary shock wave indicates a wave that is at the upstream edge of a congested region and remains at the same position in space.

Nam and Drew (1998) analyzed shock wave methods as a macroscopic tool. The study was carried out by analyzing dynamic mechanisms of traffic under congestion based on the principle of conservation of vehicles. Study results indicated that the shock wave method was able to estimate both delay and travel time that satisfied the principle

of conservation of vehicles. However, it did so by requiring an assumption of constant flows and densities in traffic streams, which is not an adequate representation of reality.

Lee and Volpatti (2010) developed a methodology to calculate shock wave speed from roadway data (i.e., speed, volume, and occupancy) collected by loop detectors. The collected data was averaged on a 1-minute basis to eliminate large fluctuations of the values. It was then used to classify the traffic conditions as either uncongested or congested based on whether the roadway density (calculated by dividing the volume data by the speed data) was less than or greater than the critical density. Next, volume and density data for a 10 minute time period were plotted. Finally, the shock wave speed was predicted by measuring the slope of a line that best fit to the observed data points. The methodology was applied to data collected by loop detectors placed approximately .31 miles apart on the Gardiner Expressway in Toronto, Canada. It was found that a linear function could be fit to the data more objectively using least-square linear regression. Also, in some cases, although the volume and density did not consistently increase or decrease during the entire 10-minute study period, there were two distinct clusters of points that were assumed to be two different steady-state traffic states. The average volume and density were calculated for each cluster as a representative point and the slope of a line connecting these representative points was measured to determine the shock wave speed. There were some instances where no specific pattern of changes in volume and density was noticeable so it was concluded that a shock wave did not exist for these instances.

2.3 Traffic Sensors

Real time traffic data is collected with two categories of traffic surveillance technologies. The first is point based sensors which are used to measure the speeds of vehicles at the locations of the sensors along a roadway. Travel times are then calculated by assuming the speed is representative for the entire roadway segment. For segments with high variances of speed over their length, the error can be significant. However, a benefit is that these technologies do not require any specific vehicle to be sampled more than once and the sensors collect data from 100% of the vehicles. The second is path based sensors which record data from vehicles equipped with technology that transmits unique vehicle identification numbers such as E-ZPass tags, GPS units, and Bluetooth enabled devices. When one of these vehicles is detected by a path based sensor, the sensor records its unique identification number as well as the time at which the sensor detects it. By comparing data from an upstream sensor to a downstream sensor, the travel times of the vehicles are determined. Unlike the point based sensors, data is not collected from 100% of the vehicles on the roadway (Dowling Associates 1999).

A survey was conducted in August 2009 of state Department of Transportations' (DOTs) application of point and path based sensors on freeways and signalized arterials. A request to complete the survey was e-mailed to twenty-six members (state DOTs) of the Federal Highway Administration's Traffic Management Center Pool-Funded study. In addition, although not a member, the Ohio State DOT was included because it deploys point based sensors statewide. The survey had 16 respondents (59% response rate with 32% coverage of all states). The majority of the states surveyed, 9 out of 16, do not use path based sensors on freeways as of the time this survey was conducted. None of them

used path based sensors on signalized arterials. The survey also found that 9 out of 16 state DOTs space their point based sensors every half mile. However, some noted that a greater spacing of 1 to 2 miles may be applied for less congested areas. Also, 4 of the states use a spacing of less than a half mile. There is no consensus of point based sensor spacing on arterials (Chien and Spasovic 2010).

Haghani et al. (2010) evaluated blue tooth sensors, a type of path based sensors, for travel time estimation. Each Bluetooth enabled device has a Machine Access Control address. The devices continuously transmit their addresses for the purpose of identifying another device with which to communicate. A Bluetooth traffic monitoring system compares the time stamps of matching Bluetooth wireless networks addresses that are detected at successive stations. This data provides a measure of travel time and space mean speed. The maximum error in the estimates resulting from local inaccuracies in readings at each sensor is a function of the coverage radius of the sensors, the average traffic speed, error in travel time estimates, and the actual travel time between a pair of sensors. By comparing the accuracy of speed estimates obtained through the use of Bluetooth sensors and floating car data collected in Maryland and Northern Virginia, it was found that the maximum error in speed estimates will be less on longer segments. The research concluded that using Bluetooth sensors on freeway segments less than one mile long may deteriorate the quality of travel time estimates. Also, the sensors should not be deployed on highway segments that are parallel to arterials or that have both express and local lanes, as travel time from the roads will get mixed due to the signal radius. In addition, the average Bluetooth sensors sample between 2 to 3.4% of the vehicles in the traffic stream.

Mirchandani et al. (2009) described the use of toll tags as path based sensors. The use of automatic toll collection technologies to collect fees on toll roads is increasing. Data is collected from vehicles equipped with toll tags as they pass toll tag readers that are located either overhead or on the side of the road. The technologies collect data on the travel times from the point when the vehicle is first detected to the point it is next detected and can also collect volume.

Kim et al. (2011) and Liu et al. (2012) evaluated the accuracy of various travel time estimation technologies. These technologies included TRANSMIT readers, Bluetooth sensors, and INRIX. The TRANSMIT system uses transponder readers installed at the roadside that detect electronic toll-collection tags inside passing vehicles. Bluetooth sensors, also installed at the roadside, collect travel time of vehicles equipped with Bluetooth devices. The INRIX system collects speed information from trucks, delivery vans, and other fleet vehicles equipped with GPS devices as well as consumer cellular GPS-based devices. It integrates this data with fixed sensor data. Data collected by these three technologies was compared to data collected by probe vehicles. The probe vehicle data was assumed to be accurate because the drivers were instructed to travel at the average speed of the traffic stream by passing as many vehicles as passed them. Only two probe vehicles were dispatched per 15-minute study time interval. The study location was a 41-mile corridor of I-287 in central New Jersey and data was collected from 7:00 am to 10:00 am over several days. The study results indicated that the accuracy of the speed estimates by Bluetooth and TRANSMIT outperformed that with INRIX data. But, the accuracy of speeds reported by TRANSMIT was slightly worse under non-recurring congestion; and, link length did not affect the accuracy.

2.4 Travel Time Prediction and its Benefits

This section presents research on the benefits of travel time prediction. These benefits include travel time savings and improved on-time reliability. These findings are important because they can be used to convert travel time prediction error, which is in units of time (hr), into units of cost ($).

The objective of the study by Toledo and Beinhaker (2006) was to evaluate the potential travel time savings from Advanced Traveler Information Services (ATIS) which provide drivers with travel time and routing information. To begin their study, they classified different route guidance methods based on two characteristics. The first characteristic is the nature of the data and prediction method used to generate routing information. Based on the first characteristic, they classified route guidance methods into four levels of sophistication and computational requirements: static, historic, instantaneous, and predictive. For the first level, static, route guidance is based on static information about the network, such as distances, speed limits, and classification of road facility types. For the second level, historic, route guidance is based on travel times derived from a historical database that represents past traffic conditions. For the third level, instantaneous, route guidance is based on real-time traffic information collected in the field and the data analysis is done in real time. The information provides a snapshot of the current conditions in the network. For the fourth level, predictive, route guidance is based on predictions of future traffic conditions, which are made in real time. Both real time traffic data is required along with models which predict short-term future traffic conditions based on current conditions. The second characteristic to classify the different route guidance methods is the timing of when the route guidance information is provided

to drivers. It can either be provided pre trip, once at the beginning of the trip. Or, it can be provided en-route, the information is provided during the trip every time new information is available.

The different route guidance methods were tested with real world traffic data collected from inductive loop detectors on the Los Angeles highway network. The network is comprised of 51 nodes and 162 links. There are three important assumptions as part of the methodology to test the methods. First, the predictive method is assumed to provide the actual, shortest route travel time possible. Second, it was assumed that the routes chosen by drivers in response to travel time information did not affect traffic conditions due to the small market penetration of ATIS. Third, the collected traffic data is error free. (This assumption is relaxed in the sensitivity analysis.)

The results of Toledo and Beinhaker's study indicated that ATIS offer potential travel time savings of up to 14%, compared to drivers only having only static information. The results were also consistent with the hierarchy of the information levels. For example, considering both a.m. and p.m. peak trips, the travel time savings, compared to drivers only having only static information, for the historical information level, instantaneous provided pre-trip level, instantaneous provided en-route level, and predictive level were 1.3%, 5.5%, 9.1%, and 11.8%, respectively. Also, the savings were generally greater for the p.m. peak than the a.m. peak because the p.m. peak was more congested. Another finding was that the various ATIS levels significantly reduce travel time variability (in percentage terms compared to the static information). In fact, the impact on travel time variability was greater than that on the average travel time savings. As with travel times, the impact of the information increased with the sophistication of

the information provided and is larger for the p.m. peak period. As stated previously, the assumption of error free collected data was studied. To do this, random errors were introduced into the travel time information that is used to route drivers. The inaccuracy affected travel times in all information levels similarly and on average reduced travel time savings by 0.5%. Furthermore, the error in the information did not have a significant impact on variation of travel times.

Wunderlich et al. (2001) studied the benefits of incorporating ATIS use into a traveler's regular commuting pattern. The study was motivated by numerous previous field experiments and simulation studies which concluded that ATIS users experience little or no actual reduction in their in-vehicle travel time. It was shown that, even though over time the users realize only marginally reduced in-vehicle travel time, they do realize substantial time management benefits from improved on-time reliability, reduced early and late schedule delay (defined as time wasted by arriving too early and total accumulated lateness, respectively), and more trip predictability. Drivers that benefit from this include truckers who are delivering time sensitive materials and commuters who need to make daycare pickup requirements on their work-to-home trips. This information also reduces commuter stress. The methodology employed the Heuristic On-line Web-Linked Arrival Time Estimator process, developed by Mitretek, which utilizes the concept of a simulated yoked trial. A simulated yoked trial consists of two steps; in the first, traveler path and time of departure choices are established for a traveler who utilizes ATIS and one who does not. The user relies on the real-time estimates provided by the traveler information service and adapts his trip timing and route choice on a daily basis. The non-user is assumed to have estimates of average link travel times during the

commute period from prior experience in the network and selects the faster option as the habitual route and then budgets in additional time to account for expected day-to-day travel time variability to establish a habitual trip start time. In the second step, travel times and on-time performance for each traveler are reconstructed based on the trip timing and routes chosen in the first step. According to the authors, a critical conceptual difference between the simulated yoked study and previous ATIS field evaluations is that the paired trials are organized around the principle of destination and target time of arrival rather than on simultaneous release from trip origin. The use of target arrival times allows for the quantifiable reliability measures to be defined and tracked along with in-vehicle travel time. By applying their methodology to a real-life case study consisting of freeway, expressway, and major arterial facilities they drew the following conclusions: peak-period commuters who do not use ATIS were three to six times more likely to arrive late compared to counterparts who use it; cases where it clearly benefits the user outweighed cases where it clearly disadvantages the user by five to one; users in peak periods are more frequently on-time than conservative non-users, yet they experience only two-thirds as much early schedule delay as non-users. The late shock, the surprise of arriving late, is reduced by 81% through ATIS use.

The value of travel time delay was provided in the 2011 Urban Mobility Report by Schrank et al. (2011). The value was estimated to be $16 per hour of person travel and $88 per hour of truck time. Another useful statistic included was that the average number of persons in each vehicle during peak period travel is 1.25.

In conclusion, accurate travel time prediction benefits travelers. Primarily, it reduces hours spent in vehicles and improves on time reliability. When this time is converted into cost, it reduces cost also.

2.5 Methodologies for Locating Sensors

Previous research has presented different methodologies for determining the number and location of sensors. The research can be divided into two groups: ones that focus on point based sensors and ones that focus on path based sensors. Point based sensors record the speed of all vehicles that pass their location on a roadway while path based sensors record the time and identification numbers of vehicles that pass them that are equipped with detectable devices (i.e., GPS, toll tags, and Bluetooth). In this section of the literature review, research on point based sensors is presented first followed by research on path based sensors.

In order to optimize point based sensor density for a network with travel time predictions provided to drivers via a route guidance system, Chan and Lam (2002) proposed a bi-level programming model. The objective of the upper level problem was to maximize the precision of the travel time information and to minimize the construction and social costs of the speed detectors. Social costs included the installation and operations of the detectors. Many assumptions were made as part of the methodology. First, it was assumed that the predicted travel time error dispersion function was a function of three things: speed detector density, volume to capacity ratio, and the social cost of the speed detectors. Second, it was also assumed that the variance of the predicted travel time error is a nonlinear decreasing function of the speed detector density. Another assumption is that more expensive sensor technology is more accurate.

A final assumption was that as the volume to capacity ratio increased, the variation of individual actual travel time on a roadway increases. The methodology was demonstrated with an example network of the Tuen Mun road corridor in Hong Kong. It was found that as origin-destination (O-D) demand increased, the point based sensor detector density required increased because larger O-D demand resulted in larger segment travel time and larger measured travel time error variance.

Coifman (2002) presented a methodology for estimating link travel time with data collected by only one point based sensor. The sensor can be located at either the downstream or upstream end of the link. The collected data (vehicle headway and speed) was used to estimate a vehicle's trajectory at the sensor location. In addition, the methodology applied macroscopic traffic flow theory including a triangular flow density relationship and shock wave theory. It is assumed that the speeds of shock waves that occur during the study time period are known. The sensor data and shock wave speed information were used to estimate vehicles' trajectories on the link. A case study was presented of an 1800 ft. freeway link that experienced only free flow conditions during the study period. The results showed that the methodology can be applied to estimate travel time within 10% of the true travel time, on average. It was noted that the methodology should not be applied under recurring bottlenecks or incident conditions.

Liu et al. (2006) proposed rules and an iterative procedure for locating a limited number of point based sensors for a freeway corridor for which the recurring conditions are known. Travel time estimation performance was compared using different algorithms under various sensor location plans. The following rules for locating a limited number of sensors were proposed. First, if less than 50% of a segment is covered by the queue, then

divide the roadway into its free flow part and its congested part. Second, if more than 50% of the segment is covered by the queue, then the midpoint of the queue is the preferred sensor location. Third, for the part of the segment under free flow conditions, two end sensors will be enough to control estimation accuracy. Fourth, for segments completely covered with a queue, an even sensor spacing is preferable. Based on these rules, the study provided an iterative procedure for locating detectors. The rules were tested to determine the locations for ten sensors on a 20-mile long segment of the major commuting corridor of I-70 in Baltimore, Maryland. When the sensors were located according to the rules, the resulting average error was 68 seconds. In comparison, deploying the ten sensors at an even spacing resulted in an average error of 139 seconds and deploying sensors every 3000 ft resulted in an average error of 73 seconds.

Bartin et al. (2006) provided a methodology for determining the number and location of point based sensors for a given route with known traffic characteristics. The objective was to minimize travel time estimation error. The problem of determining the optimal number and location of sensors was described as a space discretization problem. The idea was to represent space-time trajectory in steps of a small interval instead of as a continuous space time trajectory. The space discretization problem was converted into a common clustering problem by assuming a uniform speed trajectory for each vehicle over space. The methodology was demonstrated with a hypothetical 12.5-mile freeway route modeled with PARAMICS, a microscopic simulation software. It was found that the relative gain in increasing the number of sensors decreases as the number of sensors increases and that after seven sensors there is no more gain.

Edara et al. (2008) developed a methodology based on a genetic algorithm. It identified the optimal locations of a finite set of point based sensors on a freeway corridor, which minimized predicted travel time error subject to available capital and maintenance budget constraints. Through the application of the methodology to case studies in Northern Virginia and Richmond, the following conclusions were drawn: given that traffic conditions change over time, detector placement will require periodic validation and possibly modification to ensure continued accuracy, detector density needs to be higher in congested areas, and detectors are required at merge areas near entrance ramps, especially when the acceleration lanes are short.

Ban et al. (2009) presented a methodology in which the problem of optimizing the locations of point based sensors is formulated as a dynamic programming (DP) model by first discretizing both time and space. In the methodology, it is assumed historic ground truth measurements are known, including the locations of recurring bottlenecks and the locations that experience free flow conditions. It is also assumed that some vehicle trajectories are available and travel times are calculated using aggregated sensor speeds. The proposed DP model and solution algorithm were illustrated with a case study. Vehicle trajectory data was collected from probe vehicles equipped with GPS-enabled cellular phones, obtained from the Mobile Century field experiment performed on 5.5 miles of I-880 in the San Francisco Bay Area, California. The probe vehicles achieved a 2% to 5% penetration rate on the highway throughout the day. The case studies showed that optimal sensor placement dictates multiple sensors in bottleneck areas and less in free flow areas, where one or two sensors are usually sufficient per free flow area. Also, the decrease in travel time estimation error is monotonic as the number of sensors

increases. Specifically, for an objective value of travel time error less than 80 seconds (which represents about 13% error since the travel time of the entire route is about ten minutes) eighteen sensors for the segment are required.

Bertini (2007) presented a methodology to optimize point based sensor spacing that applied first principles of traffic flow. Specifically, the methodology considered the travel time prediction error that resulted when a shock wave passed from the downstream end to the upstream end of a freeway segment. It was asserted that the passage of a shock wave represents the worst case scenario for travel time prediction and, therefore, the methodology can be interpreted as a form of robust decision analysis. In the model, sensors are evenly spaced and are placed with the midpoint method (i.e., the sensor is located at the middle of the segment). Also, vehicles receive their predicted travel times as they enter the segment, such as via dynamic message sign. There were two possible travel times provided: one that reflected a uniform, uncongested speed and one that reflected a uniform, congested speed. A change from one travel time prediction to the other was initiated when the shock wave passed the farthest downstream sensor. (In the case of a backward shock wave, the prediction changed from uncongested to congested time. In the case of a backward recovery wave, the opposite occurred.) Thus, shorter sensor spacing decreased the time from when a shock wave entered the segment until the time a change was made in the travel time prediction. The error index considered the magnitude of under and over prediction error in vehicle hours traveled (VHT). Under prediction was considered worse for drivers and thus had a larger weight than over prediction. In the case study, a hypothetical 1-mile freeway segment was assumed, along with the values for uncongested and congested traffic flow, speed, and shock wave

speeds. The sensor spacing was enumerated and the error was calculated for each of the spacings. During the passage of a backward shock wave, it was found that a shorter sensor spacing reduced the under prediction error but increased the over prediction error. This is because the shorter spacing decreased the amount of time until the shock wave was detected. Thus it decreased the amount of time there are uncongested predictions and it increased the amount of time there are congested predictions. In contrast, during the passage of a backward recovery wave, it was found that a shorter sensor spacing decreased the over prediction error but increased the under prediction error. This is because the shorter spacing decreased the amount of time until the shock wave was detected. Thus it decreased the amount of time there are congested predictions and it increased the amount of time there are uncongested predictions. When the results from both waves were taken into account and the errors were weighted, the optimal spacing was 0.33 mile. Future research listed included determining where to optimally place sensors, including in relation to known bottlenecks and incident locations, and the use of other travel time algorithms besides the midpoint method.

Bertini and Lovell (2009) expand upon the research of Bertini (2007). First, the same methodology was applied to optimize sensor spacing when a forward recovery wave passed from the upstream end to the downstream end of a segment. The results showed that to minimize travel time over prediction, shorter sensor spacing is better. However, to minimize under prediction, spacing equal to the segment length is optimal. In addition, the effect of locating sensors at the downstream end of a link, instead of at the midpoint, was studied. If a backward shock wave passes through the segment, there

is less travel time under prediction but a greater tendency toward over prediction with this location.

The methodologies presented in Bertini (2007) and Bertini and Lovell (2009) were improved in Feng et al. (2010). One way this was accomplished was that travel time predictions provided to vehicles were modeled as a function of the sensor density. Each time the shock wave reached each sensor on the segment, the travel time prediction was recalculated. With this new prediction methodology, the error results changed for the passage of a backward shock wave; both under and over prediction error decreased as sensor density increased. Second, the methodology for travel time estimation presented in Coifman (2002) was applied to calculate estimation error during the passage of a single, backward shock wave. The sensor spacing was varied. The freeway segments' traffic states (e.g., uncongested or congested) and the shock wave were modeled as in the previous research. The results indicated that if sensors are located at the downstream end of each link, only under estimation occurred. If sensors are located at the upstream end of each link, there is no estimation error. If sensors are located at the midpoint of each link, there is no overestimated error and underestimated error linearly decreases as sensor density increases. It was noted that these results are not applicable for travel time estimation error during the passage of a backward or forward recovery wave. Third, they presented a methodology to optimize sensor spacing so that the total cost of prediction/estimation error and sensor construction cost is minimized. Parameter values were assumed for the cost coefficients of under predicted/underestimated and over predicted/overestimated total travel time error ($/vehicle hour). It was noted that the values of the cost coefficients vary according to the diversity among trips at different

times of day, trip duration, and trip purpose. Also, a parameter value was assumed for the sensor cost per shock wave ($/sensor/shock wave). It was defined as the construction cost and maintenance cost of each sensor in its life cycle divided by number of shock waves over this time period. In addition, the total cost objective function was constrained by drivers' tolerance of relative travel time estimation error. The results indicated that the optimal sensor spacing is independent of link length. It only depends on two ratios: the cost coefficient of overestimated total travel time error to sensor cost per shock wave and the cost coefficient of under estimated total travel time error to the cost coefficient of overestimated total travel time error.

Danczyk and Liu (2011) developed a methodology to optimize the allocation of point sensors along a one-directional corridor for accurate performance monitoring purposes, subject to resource budgetary constraints. In the methodology, a freeway corridor is divided into a number of cells, where each cell is designated as a potential site to place a sensor, and the sensors are placed in the middle of the cell. To solve the nonlinear program for at which sites to place sensors, they transformed the problem into an equivalent mixed-integer linear model. Then, a customized Branch-and-Bound technique was proposed to solve the resource constrained shortest path problem. A case study was conducted on a 7.2-mile strip of I-94 in Minneapolis, Minnesota with real data captured during the PM peak period. Seven sensors were needed to achieve a relative travel time error of less than 0.5%. Also, the proposed model allocated sensors more accurately than a uniform-spacing configuration. A budgetary constraint was not considered in the case study.

More recently, researchers have been developing methodologies to determine optimal spacing of path based sensors. The following research summarizes many important findings.

Sherali et al. (2006) developed an algorithm for optimally locating path based sensors by maximizing the benefit that would accrue from measuring travel times on a transportation network, subject to budgetary constraints. The problem is formulated as a quadratic 0-1 optimization problem where the objective function parameters represent benefit factors that capture the relevance of measuring travel times as reflected by the demand and travel time variability along specified trips. It is assumed that the benefit factor of providing the travel times for each link is known. Also, in the methodology, it is assumed that the potential sensor locations are predetermined. The algorithm was applied to determine the sensor locations for a 20-mile freeway in San Antonio, Texas. It was found that the benefit of optimal sensor locations on a link had a positive relationship with the volume of demand and the variability of the link travel time. The results showed that the relationship between the number of sensors and the total benefit exhibits an S-curve behavior. It first increased at an increasing rate and then displayed diminishing marginal returns. For this study network, with just six optimally placed readers, the entire travel time variability in the network was nearly captured.

Mirchandani and He (2008) applied two objective functions for path based sensor location optimization on a network. The first objective was to locate a given number of sensors to maximize the total number of vehicle miles monitored. Therefore, routes with higher flows and longer distances were favored. The second objective was to maximize the utility of deploying sensors by placing them on routes with the greatest variance in

travel times and thus use the sensors to predict travel times. The variances for each link were assumed to be already known and also constant for the study period. It was also assumed that 10% of the network traffic flow acted as probe vehicles. A key difference in the two objectives is that only one sensor per route is needed for the first but two sensors per route are needed for the second so that the path based travel times can be found. A case study was provided to demonstrate the model for a network in Harris County, Texas which already had 27 tag readers installed on the sensor network. The travel time variance on each link on the network was estimated assuming it was proportional to the travel time on that link. The optimal locations for additional tag readers were solved for and the percent reduction in the total variance of the mean route travel times were calculated. With an additional five, ten, and fifteen tag readers, the variance was reduced by 60%, 72%, and 79%, respectively.

Mirchandani et al. (2009) expand upon Mirchandani and He (2008) by considering larger transportation networks for which optimization models are NP-hard and thus exact optimal solutions cannot be achieved in reasonable computation time. To be able to approximately solve the problems in a reasonable amount of time, the methodology developed a two-stage heuristic. In the first phase, a greedy construction heuristic was used to determine an initial solution. Then, in the second phase, an exchange procedure was developed that improved the value of the objective function. The models were applied to a portion of the Harris County, Texas network to evaluate the performance of the heuristic approach and to compare it to the non-heuristic approach. It was found that the difference between the optimal (found with CPLEX solver software) and the heuristic solution values were always small, showing that the heuristic was

efficient. In addition, it was also found that when the number of links in the network was small, CPLEX was faster but as the number increased the heuristic was faster. For example, for a scenario with 120 routes, the CPLEX computational time was 2.4 seconds while the heuristic computational time was 31 seconds. But, for a scenario with 300 routes, the CPLEX computational time was 14,360 seconds while the heuristic computational time was 988 seconds.

Mazare et al. (2012) optimized both point based and path based sensor spacing for travel time estimation. The accuracies of the estimates from the two types of sensors were compared. The work was a case study that used experimental probe data obtained from the Mobile Century field experiment. Study data was collected on one day on a stretch of the I-880 freeway. Traffic conditions on that day included periods of uncongested conditions, recurring congestion, and non-recurring congestion (a traffic accident). The experiment site was also covered with seventeen inductive loop detector stations. The data collected by the sensors was combined with a mathematical traffic model in a highway traffic estimation algorithm using a data assimilation technique called ensemble Kalman filtering. The results of the algorithm were compared against the true travel times experienced by drivers, obtained through license plate re-identification. A number of scenarios were created that considered various volumes of probe vehicles as well as different numbers of inductive loop detectors. Three limitations of the methodology were noted. First, due to privacy reasons, it was not possible to track probe vehicles and thus no continuous GPS records from probes were available for the estimation algorithm. Second, the estimation algorithm did not directly estimate travel times. Instead travel times are computed from the estimated velocity field assuming that

a vehicle travels at the mean speed reported in each cell (approximately 300m long). Third, the flow model required some historical flow information for calibration. The following four key results were found from the study. First, dynamic travel time estimates can be achieved with less than 10% error when using a flow model with data assimilation, by using either inductive loop detector data, probe data, or a mixture of both inductive loop detector data and probe data. Second, using data from more than eight inductive loop detector stations (average spacing 0.83 mi) did not improve the travel time estimation. The error remained constant between 6-13% depending on the time of day, regardless of the added loop detectors. Next, when sampling probe vehicles at a rate of 127.5 vph with more than 2.54 path based sensors per mile, increasing the number of probe measurements with the addition of more probe vehicles or additional sensors caused only small travel time accuracy improvement. Last, when combining loop detector data with probe vehicle data, better estimations were obtained, especially at low probe vehicle penetration rates.

2.6 Probe Sample Size

Path based sensors record data from probe vehicles. Probe vehicles are equipped with detectable devices such as GPS, toll tags, and Bluetooth that communicate with the path based sensors. The following section contains studies whose objective were to answer the question how many is enough probe vehicles.

Turner and Holdener (1995) studied the required minimum number of probe vehicles that are necessary to report real-time travel speeds and times for a desired statistical accuracy. They calculated the sample size with a regression equation where the dependent variables were the travel time coefficient of variation (mean travel

time/standard deviation), the desired confidence level, and the permitted relative error. The case study data came from three corridors near Houston on which path based sensors were already located at 1.8 to 5.1 mile spacing. The corridors consisted of freeways that have high occupancy vehicle (HOV) lanes. Data was collected during the peak congested periods on weekdays. The summary and analysis of the travel time data considered two separate time periods: 5 minute and 15 minute periods to represent the real life periods over which travel times from probe vehicles are averaged and then reported to drivers. The results showed that the HOV lanes have a lower average number of probe vehicles due to their low volume. Therefore, more probe vehicles are needed. In addition, it was found that for a 5 minute period with a 95% confidence level and a 10% relative error, the sample sizes range from 1 probe vehicle every 5 minutes for free-flow conditions (HOV lane segments) to 6 probe vehicles every 5 minutes for severely congested conditions. Also, sample sizes are slightly lower for a 90% confidence level and 10% relative error. It was concluded that the sample size results confirm the intuitive notion that congested freeways required greater number of probe vehicles than uncongested freeways. In addition, path based sensors were recommended to be installed every 2 to 3 miles on congested segments and every 3 to 4 miles on uncongested segments.

Srinivasan and Jovanis (1996) studied the number of probe vehicles required to estimate link travel times for a network during a peak period. Both travel time reliability and network coverage adequacy requirements were considered. A general heuristic algorithm was developed. It explicitly considered the time period for travel time estimation (e.g., 5, 10, or 15 minutes), the number of replications of travel time desired for each link during each measurement period, the proportion of links to be covered, and

the length of the peak period. The algorithm was implemented by using a simulation of the Sacramento, California network (170 mi^2) for the morning peak period. It was found that a substantial number of probe vehicles are required to estimate link travel times if all the classes of links, including minor arterials and local and collector streets, are to be covered adequately. However, the heavily traveled high-speed routes, such as the major arterials and freeways, require a much smaller fraction of probe vehicles. Specifically, the case study considered the probe sample size required for a 10-minute measurement period with an adequacy requirement of 80 percent of major arterials and freeway links reliably covered with a minimum of three probe vehicles in each link during each measurement period. It was found that about 3,500 probe vehicles were required for a 2-hour peak period. This number of probe vehicles constituted less than 5 percent of the total peak period volume.

Sen et al. (1997) studied the sample size of probes needed to obtain reliable link travel time estimates. The correlation of vehicles' travel times was studied. Data was analyzed from two signalized arterials that consisted of twelve links in a suburban area of Chicago. The data was collected by probe vehicles that represented 1-2% of the vehicle population. The research showed that most often, highly used links have high correlations between travel times of vehicles. Therefore, under situations of high congestion the sample sizes required would be smaller. In addition, under high-congestion levels, because volumes would be higher, even low deployment rates would usually achieve reasonable probe frequencies. There were two other important conclusions that were drawn. First, the variance of mean link travel time remains quite

far from zero no matter how many probes use the link. Second, after a certain number of probes per unit time, additional probes do not decrease the variance very much.

Chen and Chien (2000) developed a heuristic to determine the minimum number of probe vehicles needed for reliable link travel time estimation. The methodology took into account that vehicle travel time on some links is not normally distributed. A preliminary understanding of the impact of traffic volume on the minimum required number of probe vehicles was established. For the case study, 8 miles of I-80 in New Jersey were modeled with CORSIM, a microscopic simulation model software. Specifically, in the case study, demand was varied to 5 levels that had best-case V/C ratios of 0.26, 0.45, 0.53, 0.63, and 0.61 and corresponding worst-case V/C ratios of 0.41, 0.68, 0.82, 0.99, and 0.99. The minimum percentage of probe vehicles that corresponded to these levels were 7, 3, 3, 10, and 12, respectively, for the desired statistical accuracy (with 5 percent maximum relative error and 5 percent sample error limit) over a 5-minute time interval. It was noted that although the actual percentage number in this example is network specific, the general pattern of the relationship between traffic volume and minimum probe vehicle percentage is most likely generic.

Cheu et al. (2002) studied the relationship between probe vehicle sample size and travel time estimation accuracy for an arterial link. The methodology assumed a vehicle's position and time stamp were recorded at every intersection. A microscopic simulation tool, Version 2.0 of INTEGRATION model, was employed to microscopically model traffic flow and simulate probe vehicles on a network of major and secondary arterials in a town area of Singapore. Results showed that the improvement in the accuracy of link speed estimation diminishes when the probe vehicle population in the

network has reached 15%. Also, to achieve an absolute error in the estimated average link speed of less than 3.1 mi/hr at least 95% of the time, there should be at least 4% to 5% active probes in the total network traffic volume, or at least ten probes that have traversed a link.

Li et al. (2002) provided a methodology to determine the minimum probe vehicle sample size required for speed estimation. An equation was developed that is dependent on four variables. The first is the standard deviation of the collected data for the study segment. The second is the Z-value of the standard normal distribution of the data. It was noted this should be used instead of the t-value. Also, a user-selected allowable error in the estimate of the mean speed is included. Finally, a sample size adjustment factor for which the value is determined by the confidence interval is needed. The methodology required that data is collected with a minimum of 3 initial test runs. A case study was included for a 2-mile route with signals on IN-26 in Indiana. It was found that 5 to 10 probe vehicles can generate reliable results for travel time and delay studies. Also, the minimum sample size decreased as the length of the study route increased. This is because the dispersion of the travel speed decreased and the travel speed because more stable.

2.7 Incident Duration Prediction Models

The duration of an incident is comprised of three different phases – detection , response, and clearance phases – while some studies include a fourth phase – the recovery phase. The detection phase is the period of time between when the incident occurs and when it is detected by responders or traffic managers. The response phase is the period of time between when the incident is detected and when emergency responders arrive. The

clearance phase is between when responders arrive and when capacity is restored at the incident location. The final phase is between when capacity is restored and traffic flow on the freeway segment returns to normal conditions.

Incident duration has a positive relationship with congestion during an incident. In order to accurately predict travel times during an incident, reliable incident duration predictions are necessary. Much research has been conducted in order to produce models that predict the incident duration with available data during an incident. Incident and roadway characteristics that affect the duration have been studied although a consensus on which are most important has not been formed.

Garib et al. (1997) developed a regression model for predicting incident duration. Data from a 7.3-mile segment of I-880 in Oakland, CA was used for the model. It was collected during the morning and evening peak periods. Thus, the models' parameters were dependent on the input data. The incident duration model predicted 81% of the incident duration in a natural logarithmic format as a function of six independent variables. These variables were number of lanes affected, number of vehicles involved, whether trucks were involved in the incident, time of day, police response time, and weather conditions. No other variables tested either individually or jointly were found to be significant. Truck involvement, time of day, and weather conditions all only had values of zero or one. The truck involvement variable had a value of zero if none were involved and a value of one if they were involved. The time of day variable had a value of zero for the morning peak time and a value of one for the afternoon peak time. The weather conditions variable equaled zero if no rain and one otherwise. Note that snow would not be present in Oakland, CA. The values of the independent variables

coefficients were all positive except for the time of day and the weather condition coefficients. This indicated that the duration of an incident in expected to be less in rainy weather conditions than in dry weather conditions. Also, incident duration is less in the afternoon peak than in the morning peak. The authors suggest that traffic conditions are negatively correlated with incident duration because there is more urgency to clearing incidents during peak hours because of the congestion they cause. The model results also indicate that police response time is a highly significant factor in predicting the incident duration followed by weather condition, time of day, truck involvement, and finally the joint effect of the number of vehicles involved and the number of lanes affected. In addition, a lognormal distribution was used to describe incident duration where a shift to the left shows a larger proportion of short-duration incidents.

Wu et al. (1998) validated the incident duration estimation model of Wide Area Incident Management Support System (WAIMSS). The model predicts incident duration based on an estimation tree. The model was calibrated with incident data collected on freeways in Northern Virginia. It was found that the average error of the model was 14.2%. In addition, the data set strongly supported a non-normal duration distribution assumption. Although the incident duration had a shape similar to a lognormal distribution, it was rejected by several statistical significance tests. However, if the set of incidents is divided into subsets of incidents that have the same type and similar severity a normal distribution of duration was found. The global maximum for incident duration appeared to be around 30 minutes, which corresponded to the average duration for minor incidents, and a secondary local maximum appeared around 60 minutes, which corresponded to the average duration for moderate incidents. Also, it was found that

while most incidents last less than one hour, the spread of the incident duration is much wider than one hour. Furthermore, the authors suggested future studies that included the effect of congestion on incident duration. They found that congestion lengthens the incident duration mainly by affecting the travel time of response vehicles, and in some cases, affecting operations at the incident scene. In particular, response times to incidents on bridges are substantially extended, and efficiency of incident removal operations on bridges and ramps are reduced by limited space.

Smith and Smith (2001) tested three different models for predicting accident clearance time. They were a stochastic model, a nonparametric regression model, and a classification tree model. Data from 6,828 accidents on Virginia freeways was used to test the models. The independent variables considered for the forecasting models were as follows with their possible values shown in parentheses: time of day (peak or off-peak), day of the week (weekday or weekend), weather (normal or adverse), response agencies (EMS, Fire, etc), number of vehicles (one, two, or three or more), truck involvement (yes or no) and passenger bus involvement (yes or no). Statistical significance tests on the independent variables using analysis of variance (ANOVA) were performed. The ANOVA analysis showed that all of the independent variables are significant except for weather and FIRT (a response agency) response. The first forecasting model developed for the study was a stochastic model. This model attempts to describe the randomness of the events. One was to do so is through a probability density function which is an equation to describe continuous random variables for a specific distribution. The study concluded that the stochastic model could not be applied to forecasting future accident clearance times due to the lack of a probabilistic distribution that fit the clearance time

data. In particular, even though a histogram plot of the data showed a definite left-shifted tendency towards accidents with smaller clearance times, a chi-square test determined that neither a Weibull nor lognormal distribution adequately described the data. Even after dividing the accident clearance time data into different sets by accident severity (i.e., number of vehicles involved) and then trying to fit a probabilistic distribution to each set of data, a chi-square test again determined that neither a Weibull nor lognormal distribution adequately described the data. The same method was tried with dividing the data into different sets by time of day (i.e., peak period weekday, off-peak period weekday, and weekend) but still with the same results. The second forecasting model developed for the study was a nonparametric regression model. This model attempts to emulate a deterministic relationship between the accident characteristics and the clearance time. However, the results were not encouraging and the authors rejected this model. They found that the model had a very large average error that in most cases was larger than the model prediction value. For example, only 18% of the test accidents had a prediction error less than five minutes. The last forecasting model that was developed for the study was a classification tree model. For this type of model, the clearance time forecasts are assigned instead of mathematically calculated. The authors concluded that this model also cannot be applied as it was only correct in predicting accident clearance time 58% of the time. In summary, the authors found that none of the models were applicable. However, they stated this could be due to shortcomings in the accident characteristics available. For future studies, they recommended including the following accident information: number of lanes blocked due to the accident, the number of personal injuries, and the number of vehicles responding from each agency.

Ozbay and Noyan (2006) modeled incident clearance durations using Bayesian Networks (BN). BNs can be used to create dynamic incident duration estimation trees that can be extracted in the presence of a real incident for which data might only be partially available. The BN represents the associations of variables to the incident clearance time as a joint probability distribution over a set of random variables. The purpose of the BN approach is to draw inference, and then determine patterns and relationships among all the predictors as well as the target variable, which in this case is the incident clearance time. The variables or predictors applied were the type of incident, number of police vehicles, number of ambulances, number of fire engines, number of injuries, number of trucks involved, number of cars involved, total number of lanes, and type of roadway. To test the model, incident data from Northern Virginia was applied. The model was able to predict incident clearance times with 80% accuracy. In addition, it showed that most incidents last less than one hour.

Li and Cheng (2011) also applied Bayesian Networks for incident duration prediction. In particular, they developed a continuous probability distribution of duration model based on lateen Gaussian naïve Bayesian (LGNB) classifier, assuming duration fits a lognormal distribution. The purpose of LGNB is to classify incidents according to the fitted continuous probability distribution, but not the length of duration. They applied the model to data from 1,470 incidents in the Atlanta, Georgia metropolitan area. The following information was known for each incident: weekday or weekend, day or night, incident type, detection type, location type (freeway, ramp, intersection, arterial, other), lane type (lane, off road, shoulder, core area, none), number of lanes affected, fire presence, hazmat presence, injury count, fatality count, need police flag, need HERO

flag, need truck wreck flag, signal damage, guardrail damage, type and number of vehicles involved in incident, and duration (less than 30 minutes, 30 to 60 minutes, 60 to 90 minutes, 90 to 120 minutes, or over 120 minutes). Based on the incidents' characteristics, each incident was classified into one of four classes, with each class having its own probability distribution. A distribution fit test was performed on each distribution to determine that all four classes approximately fit the Gaussian distribution. A sensitivity analysis was performed to determine which incident variables affect the class (and duration probability distribution) to which an incident is assigned. It was found that only half the variables affect the class variables significantly. These include incident type, detection type, location type, lane type, signal damage, guardrail damage, and type and number of vehicles involved in incident. The type of incident was the major factor and the location type was also an important factor. The authors concluded that four classes of incidents classified by LGNB can be interpreted with different levels of severity and complexity. Note that time of day or weekday or weekend was not considered an important variable for determining incident duration.

Khattak et al. (2011) analyzed traffic incidents and presented an online tool (iMiT – Incident Management Integration Tool) that dynamically predicted incident durations, secondary incident occurrence, and associated incident delays. Incident duration prediction and secondary incident occurrence was based on statistical regression. An incident database covering 110 freeway miles of the Hampton Roads metropolitan area of Virginia was used. Analyses of the data showed that the average incident duration is fourteen minutes, 9.5% of incidents result in lane closures, 35% of incidents occur during the peak period, and 10% occur during bad weather. Ordinary least squares regression

models were developed which used the following variables to predict incident duration: time of day (either peak or off-peak), weather (bad or not), freeway corridor (categorical variable), average annual daily traffic (per 1000 vehicles), incident detection source, number of vehicles involved in incident, incident type (categorical variable), traffic lane closure (yes or no), emergency medical service response (yes or no), right shoulder affected (yes or no), ramp affected by incident (yes or no), and left shoulder affected by incident (yes or no). The key factors associated with longer incident durations were found to be crashes (as opposed to other types of incidents), freeway facility damage, more vehicles involved in incident, severe injuries, when incident affects the left shoulder or ramp, and longer lane closer times. A theoretically-based deterministic queuing model was applied to estimate delay. The main inputs were the incident severity which is directly related to incident reduced capacity, incident duration, traffic demand, and road geometry information such as number of lanes. The study empirically validated the incident duration model by comparing observed and predicted values. The Root Mean Square Error(RMSE) for predicted incident duration was 16.4% which was reasonable and indicated that the models could provide realistic predictions for most incidents. Also, it was found that the duration model did not predict extreme values very well because statistical models are based on capturing the central tendency in the data, rather than outliers.

A methodology to estimate incident impacts in real-time was developed by Hadi et al. (2011). A number of measures were used to quantify incident impacts including percentage lane blockage, incident duration, average incident delay, queue length, and the potential for secondary incidents. The percentage lane blockage has to be observed by

CCTV image operators or by incident responders. Models were developed to estimate incident durations, mobility impacts of incidents, and the potentials for secondary incidents based on various incident and traffic conditions. These models were developed based on archived incident and traffic detector data from the I-95 corridor in Broward Country, FL, which is managed by FDOT District 4. The incident data includes incident timestamps (detection, notification, responses, arrivals, and departures) and event details. For the models, the incident durations were predicted in two phases in real-time applications because incident information is generally known sequentially. During Phase 1, a simple approach is used based on historical data using the mean and 95% confidence interval of incidents of similar types to the current incident at the incident location. In Phase 2, more detailed information is known. Also, the model estimates incident response and lane clearance durations separately because the two different durations are influenced by different factors. Incident response duration is the time from the occurrence of the incident to the arrival of the first responder. It was predicted based on cross-classification considering the factors of night vs. day, weekdays vs. weekends, and the injury levels of the incidents. Lane clearance duration is the time between first responder arrival and the reopening of all travel lanes. It was predicted based on a model that utilizes the M5P tree algorithm. Overall, the M5P model shows that the significant factors in predicting lane clearance duration are the number of blocked lanes, number of responded service patrol vehicles, injury presence, number and type of vehicles involved, time of day (AM, PM, Midday, Night, Weekend), TMC verification and response time, incident type, number of lanes blocked, presence of CCTV cameras, and the presence of the Severe Incident Response Vehicle. To predict incident mobility impacts in real-time

including delays and queues, the authors provided the following methods that can be applied: queuing theory, shock wave analysis, or simulation analyses. The authors chose deterministic queuing analysis and left the other methods for future studies. To apply the queuing theory, the parameters required for the calculation include incident duration, traffic demands, and capacity with and without incidents. The authors note that traffic demands during incident conditions can be estimated based on the average historical values based on traffic detector data at the incident location and can be updated in real-time. The capacities during incident and non-incident conditions could be estimated based on the values presented in the Highway Capacity manual, although the capacity during incident conditions can also be estimated based on detector data.

Chien and Ozbay (2012) analyzed New Jersey incident data to determine New Jersey specific incident rates, response times, and clearance times. They found that crash response times are shorter than disablement incidents response times. Also, the response times varied by the type of incident management or detection. The average crash response time ranged by type of incident management from 8.4 minutes to 20 minutes and the average disablement incident response time ranged from 9.5 minutes to 25 minutes. The clearance times were affected by whether the incident occurred in the shoulder or lane and by the incident type. The average clearance time considering all incident types was 19 minutes for in-lane locations and 17 minutes for shoulder locations. For the study, the authors also conducted an analysis of the effect of time of day on response and clearance times. Regarding crash response time, there was no significant fluctuation with respect to time of day. For crash clearance times, there is no significant pattern for average duration or standard deviation values. In addition, from their

literature review, they concluded that, in general, the factors affecting total incident duration are open to discussion for their influence, especially on response time. Even for factors which are determined to affect response time, such as time of day, there is no general consensus about its real effect. They included a summary of previous studies' findings on factors that influence incident response and clearance times. These are incident/disablement type, severity, lane closure, number /type of vehicles involved, number/type of response agencies, and time of day, incident location (route, in-lane/shoulder), traffic conditions, seasonal /weekday variation, weather/environmental conditions, alcohol involvement, pavement operations, freeway damage/debris, response timing, and existence of insurance. In addition, they included a summary of previous studies' findings on average incident response times and clearance times. The ranges of times provided by the studies were from 7.5 to 33.0 minutes and 4.94 to 136.8 minutes, respectively, with the average times amongst the studies of 15 minutes and 30.75 minutes, respectively.

2.8 Traffic Microscopic Simulation

The Austroads Project NS1229 (Tudge et al. 2007) aimed to promote consistency in the application of micro simulation traffic models including AIMSUNG NG, Quadstone Paramics, and VISSIM. It summarized the values of parameters, such as the period of simulation and number of runs, currently employed by Austroads Project Group members; these values are recommended for use. For the period of simulation, one to two hours was recommended. It was noted that this period varies depending on demand data and network complexities. For the number of runs, five to six were recommended. Three additional recommendations were made in regards to the number of runs. First,

more runs are recommended for complex models. Second, a set of fixed random number seeds should be used for each scenario to ensure repeatability. Third, when analyzing the data from the runs, the median value should be used instead of the average and hence an odd number of runs (i.e.,e five, seven, or nine) make it easy to identify median values.

CHAPTER 3

MODEL OF TRAVEL TIME PREDICTION ERROR

Path based sensors provide travel time data in real time that is used for vehicle travel time prediction. Travel time predictions can be communicated to vehicles approaching a freeway segment via dynamic message signs (DMS). The accuracy of the predicted travel times directly affects vehicle occupants' ability to schedule their activities, plan their routes, and gain or maintain confidence in information technology systems. Therefore, it is essential that predicted travel times are as accurate as possible. Travel time predictions are most useful to drivers during transitions in the traffic state. An example of when a transition affects a freeway segment is when an accident occurs. The accidents result in a change from an uncongested traffic state to a congested traffic state. The change, which occurs over space and time, is represented with a shock wave. During a transition, vehicles' travel times vary by vehicle departure time as the segment becomes either more or less congested. Neither the posted speed limit or knowledge of the average, usual speed for the time of day (i.e., peak period) are applicable for predicting the travel time during a transition.

Transitions that affect a freeway segment include those which originate within the segment, such as at an incident or bottleneck location, and those that originate outside of the segment. For example, backward moving transitions (that move in the direction against vehicle travel) can occur downstream from (i.e., in front of) a segment and can eventually reach and pass through the segment. In addition, forward moving transitions (that move in the direction of vehicle travel) can occur upstream from (i.e., behind)

segment and can also eventually reach and pass through the segment. There are two types of transitions, recurring transitions, such as those that occur when the traffic demand changes from peak to off peak, and non-recurring transitions, which occur during an incident. Transitions are modeled with shock waves.

This dissertation develops a mathematical model that compares the travel times predicted for vehicles to traverse a freeway segment to the vehicles' actual travel times. The difference between the predicted and actual travel times is the travel time prediction error. The sum of the absolute value of the error for all vehicles in a study period is the total error, measured in vehicle hours of travel (VHT). Predicted travel times are calculated from path based sensor data where the sensors are located at a given spacing along a freeway segment. The error during a transition or transitions in the freeway segment's traffic state during a study period is considered. Six different study periods are included; three have a single transition in the study period and three have two transitions in the study period. The study period for the model starts at the departure time of the first vehicle affected by the transition or transitions. The study period ends at the departure time of the first vehicle unaffected by the transition or transitions.

There are four parts to the model. The first part classifies the study period into one of six Scenarios. The second part models vehicles' actual travel times on the freeway segment during the study period. The third part models vehicles' predicted travel times during the study period, which are the travel times expected by the vehicles' occupants. The fourth part models the travel time prediction error. Error indices include the total error and average error for the study period. The error is found by comparing the first

two models. The relationship amongst these four steps is illustrated in the flow chart below. The model inputs are explained in more detail in the assumptions section.

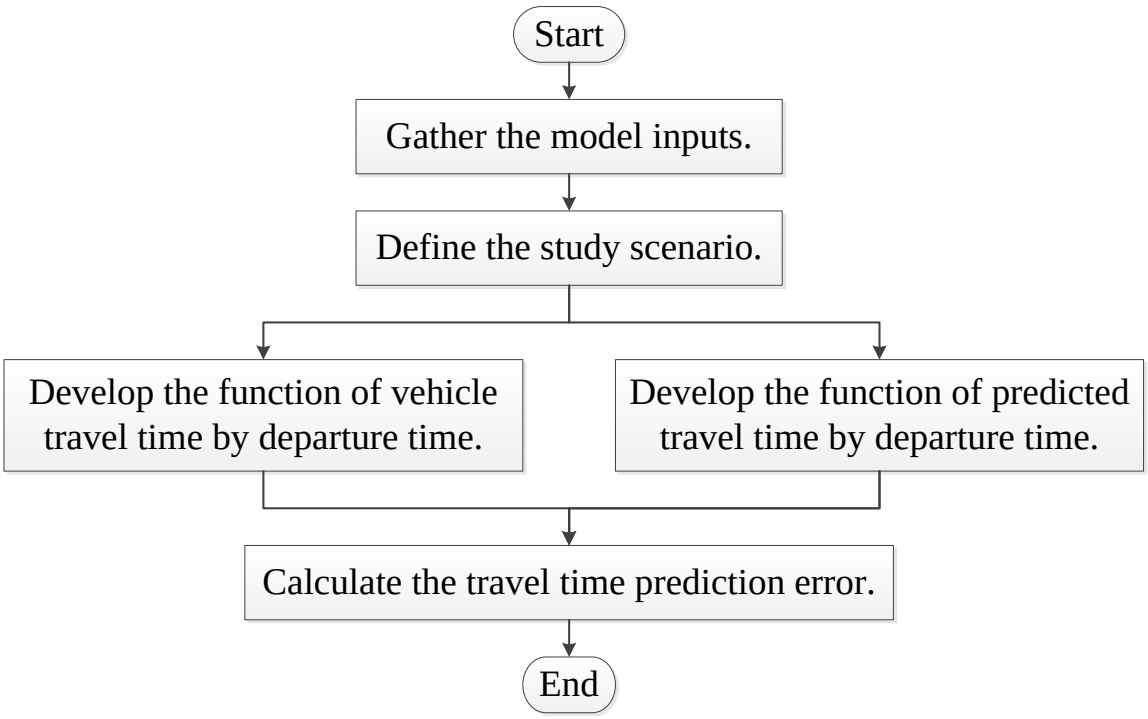


Figure 3.1 Overview of the travel time prediction error model.

Table 3.1 Variables Used in the Model and their Symbols and Descriptions

Sym-bol	**Description**	**Unit**
A_t	The actual travel time of a vehicle that enters the freeway segment at time t where $t = t_a$, t_b, $t_{a,s}$, $t_{a,r}$, $t_{b,s}$, $t_{b,r}$, etc.	hr
D	Freeway segment length	mi
E	The total travel time prediction error during a study period	VHT
e	The average travel time prediction error during a study period	hr
E_t	The travel time prediction error for a vehicle that departs at time t	hr
$f(t)$	Function of vehicle travel time by departure time	-
$g(t)$	Function of predicted travel time by departure time	-
$g_n(t)$	Function of predicted travel time by departure time for the link located between sensors n-1 and n	-
N	Number of sensors	sensor
n	Index of sensors	-
$p_{n,t}$	The predicted travel time from mile x_{n-1} to mile x_n at time t where $t = t_{a,n}$, $t_{b,n}$, $t_{a,s,n}$, $t_{a,r,n}$, $t_{b,s,n}$, $t_{b,r,n}$, etc.	hr
P_t	The predicted travel time for the freeway segment at time t	hr
q_c	Congested traffic flow	vph
q_u	Uncongested traffic flow	vph
S	Sensor spacing	mi
t	The departure time of a vehicle	hr
t_a	The departure time of the first vehicle to encounter the shock wave	hr
$t_{a,n}$	The time at which a link's downstream sensor detects the first vehicle that encountered the shock wave within the link	hr
$t_{a,r}$	The departure time of the first vehicle to encounter the backward recovery wave	hr
$t_{a,r,n}$	The time at which a link's downstream sensor detects the first vehicle that encountered the backward recovery wave within the link	hr
$t_{a,s}$	The departure time of the first vehicle to encounter the backward shock wave	hr
$t_{a,s,n}$	The time at which a link's downstream sensor detects the first vehicle that encountered the backward shock wave within the link	hr
t_b	The departure time of the last vehicle to encounter the shock wave	hr
$t_{b,n}$	The time at which a link's downstream sensor detects the last vehicle that encountered the shock wave within the link	hr
$t_{b,r}$	The departure time of the last vehicle to encounter the backward recovery wave	hr
$t_{b,r,n}$	The time at which a link's downstream sensor detects the last vehicle that encountered the backward recovery wave within the link	hr
$t_{b,s}$	The departure time of the last vehicle to encounter the backward shock wave	hr
$t_{b,s,n}$	The time at which a link's downstream sensor detects the last vehicle that encountered the backward shock wave within the link	hr

Table 3.1 Variables Used in the Model and their Symbols and Descriptions (continued)

Sym-bol	Description	Unit
t_F	The first departure time of the study period	hr
t_I	The time at which a backward shock wave and backward recovery wave intersect	hr
t_L	The last departure time of the study period	hr
t_w	The time at which a shock wave originates in the freeway segment	hr
$t_{w,r}$	The time at which a backward recovery wave originates in the freeway segment	hr
$t_{w,s}$	The time at which a backward shock wave originates in the freeway segment	hr
v_c	Congested vehicle speed	mph
v_u	Uncongested vehicle speed	mph
v_w	Shock wave speed	mph
$v_{w,r}$	Backward recovery wave speed	mph
$v_{w,s}$	Backward shock wave speed	mph
x_I	The location at which a backward shock wave and backward recovery wave intersect	mi
x_n	The location of sensor n	mi
x_w	The location at which a shock wave originates in the freeway segment	mi

3.1 System Assumptions

The proposed mathematical model is developed based on the fundamental macroscopic traffic flow theory. According to this theory, a freeway segment is in one of three states: an uncongested traffic state with average traffic speed v_u (mph) and flow q_u (vph), a congested traffic state with average traffic speed v_c (mph) and flow q_c (vph), or in a transition between the states. The traffic state can also transition more than once during the study period, from an uncongested state to a congested state then back to the uncongested state. An example of this is when an incident occurs (which causes a transition from the uncongested to congested state) and the subsequent clearance of the incident from the roadway (which causes a return to the uncongested state).

A transition is modeled as a shock wave where the shock wave travels at a constant velocity v_w (mph) through the freeway segment. The model considers that a transition can originate within the study segment or on another segment of the freeway.

To formulate the model, the following five assumptions are made:

1. The characteristics of the traffic state of a freeway segment during the study period are known, including the uncongested and congested speeds, flows, and densities. Also, the characteristics of the shock wave or waves that occur during the study period are known, including the type, direction, and velocity of a shock wave. The velocity of a shock wave equals the change in flow divided by the change in density. The traffic state, and thus the entering traffic flow, during the study period is either uncongested or congested.

2. All vehicles receive travel time predictions as they enter the freeway segment at mile zero, via DMS.

3. Vehicle travel time is modeled as a continuous function of vehicle departure time where departure time is defined as the time a vehicle enters the segment.

4. Sensors are modeled to report travel times continuously and instantaneously. Also, a 100% probe sample rate is assumed. Problems in the functioning of the sensors, such as detecting vehicles or transmitting data, are not considered. Thus,

vehicle predicted travel time is modeled as a continuous function of vehicle departure time.

5. The part of a freeway segment downstream, i.e., ahead of a transition location is unaffected by the transition and is in the uncongested traffic state.

The traffic state of a freeway segment undergoing a transition over time is determined by three characteristics of the transition. The first is whether the segment transitions from an uncongested traffic state to a congested traffic state or from a congested traffic state to an uncongested one. The second is at what location and at what time the transition begins to affect the segment. The third is at what location and at what time the transition terminates within the segment or exits the segment. This third characteristic is dependent on whether the change in the traffic state propagates upstream (i.e., in the direction against traffic) or downstream (i.e., in the direction with traffic) from the initial site of the transition and is also dependent on the speed at which the transition propagates through the segment. Based on this information, the traffic state of a freeway segment during a study period is determined. The boundary between the uncongested and congested traffic state on the freeway segment during the study period is represented with a shock wave.

3.2 Shock Wave Transition Scenarios

Six different transition scenarios are considered in this dissertation. Each transition is represented by a shock wave. In the first three scenarios, one transition affects the freeway segment during the study period. In contrast, in the next three scenarios, two transitions affect the freeway segment during the study period. The table below summarizes the characteristics of all the Scenarios.

Table 3.2 Characteristics of the Study Scenarios

	Scenarios					
	One	**Two**	**Three**	**Four**	**Five**	**Six**
Type of wave(s) in study period	backward shock	backward recovery	forward recovery	backward shock and backward recovery	backward shock and backward recovery	backward shock and backward recovery
Maximum number of waves a single vehicle can intersect	one	one	one	one	two	two
Location waves intersect	n/a*	n/a*	n/a*	Upstream from the segment	Upstream from the segment	Within the segment

* Only one wave is present in these scenarios

The characteristics of the transitions for Scenarios One, Two, and Three are discussed in detail in Section 3.2.2.1 and the characteristics of Scenarios Four, Five, and Six are discussed in detail afterward in Section 3.2.2.2.

3.2.1 Scenarios One, Two, and Three

In Scenarios One, Two, and Three, one transition, represented with a shock wave, affects the freeway segment during the study period. The characteristics of the transitions are described below.

- Scenario One
    - The freeway segment transitions from an uncongested traffic state to a congested traffic state during the study period.
    - The initial site of the transition is either downstream from the study segment (i.e., it originates in another freeway segment) or within the study segment.

- The length of segment under congestion increases over time in the direction against traffic (i.e., towards upstream).
- Therefore, the transition is represented with a backward shock wave.
- The transition terminates upstream from the segment (i.e it does not terminate within the segment).

The characteristics of Scenarios One are illustrated in Figure 3.2.

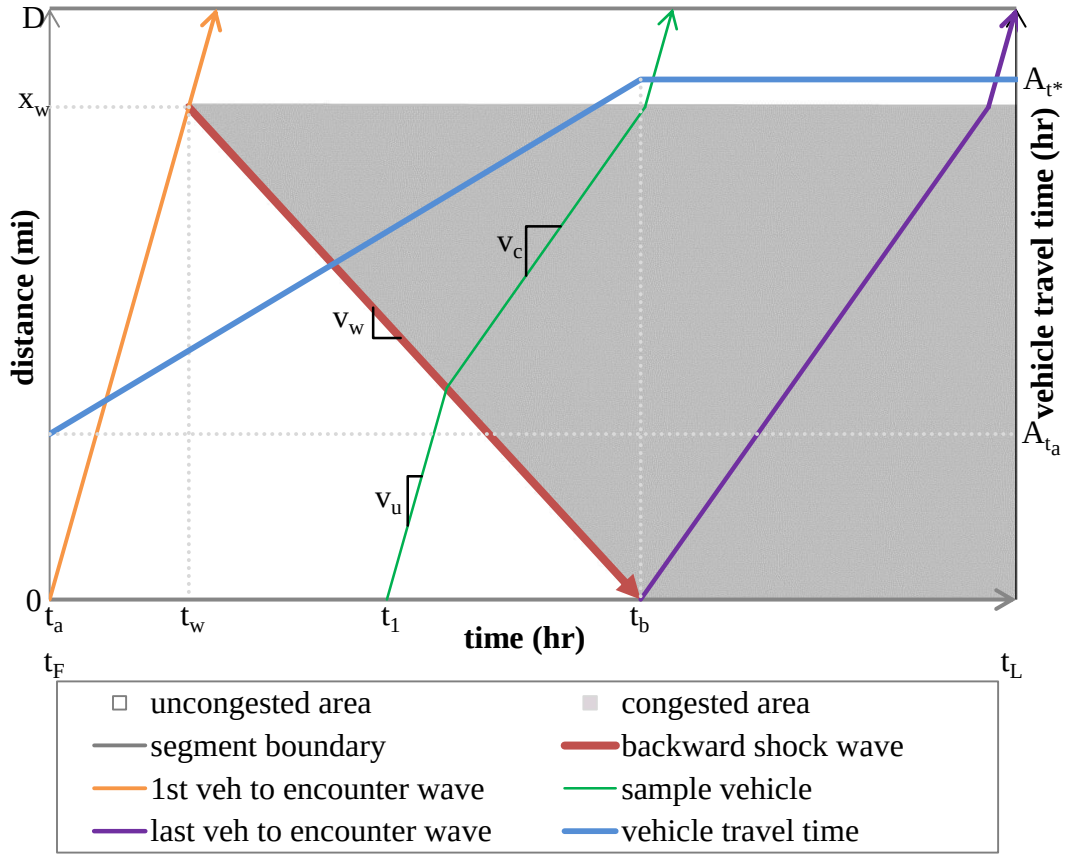


Figure 3.2 Freeway segment with a backward shock wave.
where $t_b \leq t^* \leq t_L$

In the figure, the x-axis represents time and the y-axis represents distance. The length of the segment equals D (mi). Mile zero is the upstream end of the freeway segment and mile *D* is the downstream end of the freeway segment. The trajectory of the backward shock wave is shown in red. The location at which the shock wave originates

within the freeway segment, such as an incident or bottleneck location, is denoted by x_w (mi). For study periods in which the wave initiates in a downstream segment past mile *D*, then x_w equals *D*. Also, the time at which the shock wave begins to affect the freeway segment is denoted by t_w (hr). The area in the time-space plane to the left of the backward shock wave is uncongested as it is shown with no shading. The area in the time-space plane to the right of the shock wave is congested and is represented by light gray shading. The figures also show the trajectories of vehicles over the freeway segment. Vehicles are modeled to enter at the upstream end of the freeway segment and exit at the downstream end of the segment as illustrated with the directional arrows. Vehicles' trajectories have two different slopes. The steeper one represents travel in the uncongested traffic state and the less steep one represents travel in the congested traffic state because a steeper trajectory means the vehicle travels farther in a shorter amount of time.

- Scenario Two
    - The freeway segment transitions from a congested traffic state to an uncongested traffic state.
    - The initial site of the transition is either downstream from the segment or within the segment.
    - The length of segment under congestion decreases over time in the direction against traffic (i.e., towards upstream).
    - Therefore, the transition is represented with a backward recovery wave.
    - The transition terminates upstream from the segment.

The characteristics of Scenarios Two are illustrated in Figure 3.3.

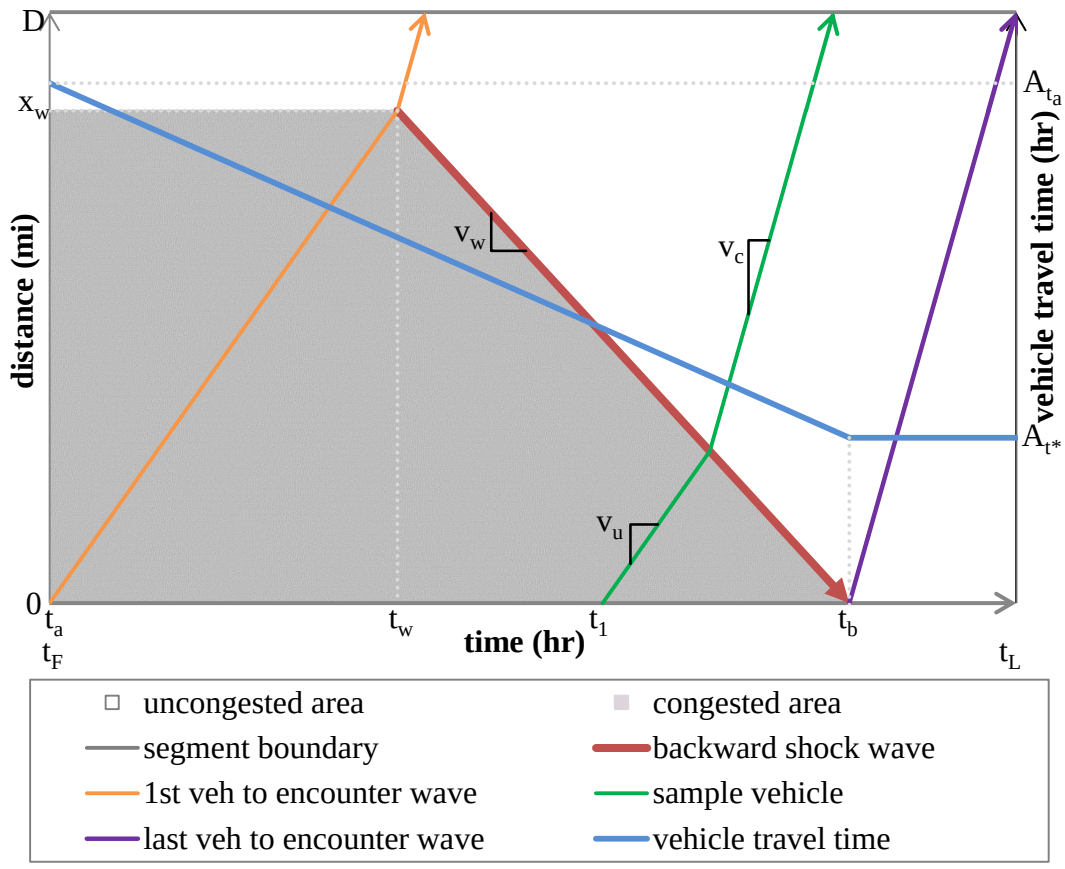


Figure 3.3 Freeway segment with a backward recovery wave.

where $t_b \leq t^* \leq t_L$

The description of this figure is the same as for Scenario One with two exceptions. First, in this figure, the trajectory of a backward recovery wave is shown in red. Second, the time-space plane to the left of the wave is congested and is represented with light gray shading and the time-space plane to its right is uncongested.

- Scenario Three
    - The freeway segment transitions from a congested traffic state to an uncongested traffic state.
    - The initial site of the transition is either upstream from the segment or within the segment.

- The length of segment under congestion decreases over time in the direction with traffic (i.e., towards mile *D*).
- Therefore, the transition is represented with a forward recovery wave.

- The transition terminates downstream from the segment.

The characteristics of Scenario Three are illustrated in Figure 3.4.

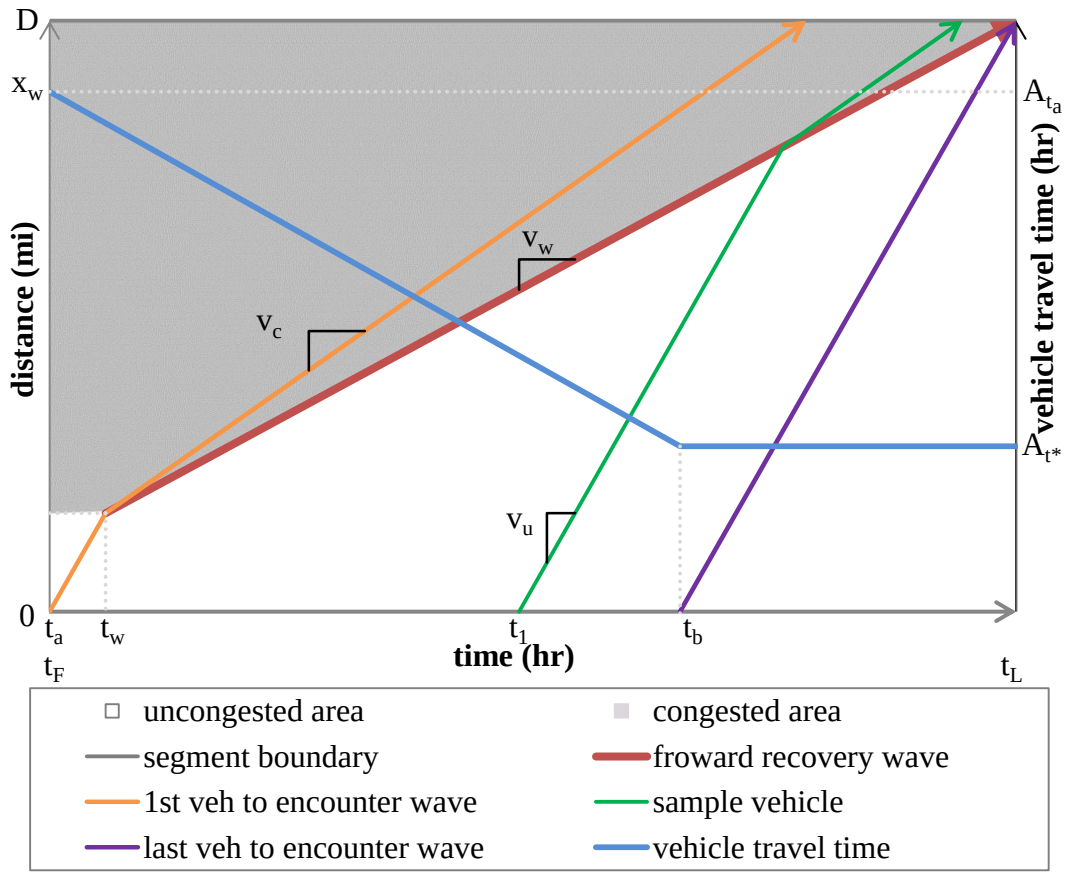


Figure 3.4 Freeway segment with a forward recovery wave.
where $t_b \leq t^* \leq t_L$

In the figure, the red line represents the trajectory of the forward recovery wave. The directional arrows show that the wave moves in the same direction as the vehicles on the segment. The time at which the shock wave originates in the freeway segment is denoted by t_w and the location at which the shock wave originates in the freeway segment

is denoted by x_w. If the wave initiates in an upstream segment prior to mile zero, then x_w equals zero. Similar to the backward recovery wave, the time-space plane to the left of the wave is congested and is represented with light gray shading and the time-space plane to its right is uncongested.

The differences in the characteristics of the transitions among the Scenarios are apparent in the figures. For example, in Scenario One, the length of the freeway segment under congestion increases over time. The congestion spreads from the initial site of the congestion towards the upstream end (i.e., mile zero) of the freeway segment. In Scenario Two, the length of the freeway segment under congestion decreases over time. The transition begins at a point toward the downstream end of the segment and continues towards the upstream end of the freeway segment. In Scenario Three, the length of the freeway segment under congestion also decreases over time. In contrast to Scenario Two, the transition begins at a point toward the upstream end of the segment and continues towards the downstream end (i.e., mile *D*) of the freeway segment.

Note that a single transition from an uncongested traffic state to a congested traffic state that travels upstream through the segment toward mile *D*, which is represented by a forward shock wave, is not included because it results in a period of no vehicle detection by path based sensors. However, the model can be modified in future studies to take this into account.

3.2.2 Scenarios Four, Five, and Six

The characteristics of Scenarios Four, Five, and Six are described next. In these Scenarios two transitions, each represented by its own shock wave, occur in the study period. The first is a transition from an uncongested traffic state to a congested traffic

state that moves upstream in the direction against the traffic flow. This transition is represented with a backward shock wave. The second is a transition from the congested traffic state back to the uncongested traffic state that also moves upstream through the segment. This transition is represented with a backward recovery wave. The backward shock wave and backward recovery wave are included within the same study period, instead of in two separate study periods like in Scenarios One and Two. This is done because the predicted travel times provided to vehicles that intersect the backward recovery wave are calculated based on data from vehicles which intersected the backward shock wave. An example of when the passage of a backward shock wave is followed by the passage of a backward recovery wave is during a traffic accident. A decrease in roadway capacity occurs at the accident location. This results in a backward shock wave. The subsequent clearance of the accident restores the capacity. This results in a backward recovery wave. The characteristics of Scenarios Four, Five, and Six are summarized below.

- Scenarios Four, Five, and Six
    - The freeway segment transitions from an uncongested traffic state to a congested traffic state and then returns to the uncongested state.
    - The transitions are represented with a backward shock wave and backward recovery wave.
    - The initial site of the transition is either downstream from the segment or within the segment.
    - The shock waves travel in the upstream direction.

The Scenarios are differentiated by considering two characteristics that affect the formation of the travel time prediction error model. If they are ignored, the model is not

applicable for studying all possible backward shock wave speeds, recovery wave speeds, and vehicle speeds on the freeway segment. The first characteristic is the maximum number of shock waves a vehicle can intersect while on the freeway segment. A vehicle may encounter none, one or both of the waves. The second characteristic is the location the waves intersect and thus terminate; the location can be either within the segment or upstream from the segment. If the waves intersect within the segment, then the part of the segment located upstream from the site of intersection remains in the uncongested traffic state for the entire study period. The characteristics for Scenarios Four, Five, and Six were provided previously in Table 3.2. In addition, the equations to calculate whether the shock waves are included within the same study period and the equations to determine if the study period is Scenario Four, Five, or Six are provided in detail in Section 3.4.2.

To illustrate Scenarios Four, Five, and Six, Figure 3.5 depicts a freeway corridor that experiences two transitions in its traffic state over time. The freeway corridor is comprised of three freeway segments all of which experience the transitions. The first transition is from an uncongested to a congested traffic state and it is represented with a backward shock wave. The second transition is from a congested to an uncongested traffic state and it is represented with a backward recovery wave. The trajectory of the backward shock wave is shown in red while the trajectory of the backward recovery wave is shown in black. The times at which the waves originate are denoted by $t_{w,s}$ and $t_{w,r}$, where the *s* and *r* subscripts stand for shock and recovery, respectively. The location the waves begin is denoted by x_w. A congested traffic state, which occurs in the time-space plane located between the trajectories of the shock waves, is represented with light gray

shading. The uncongested traffic state is represented with no shading. Vehicles are modeled to enter at the upstream end of the freeway segment and exit at the downstream end of the segment as illustrated with the directional arrows.

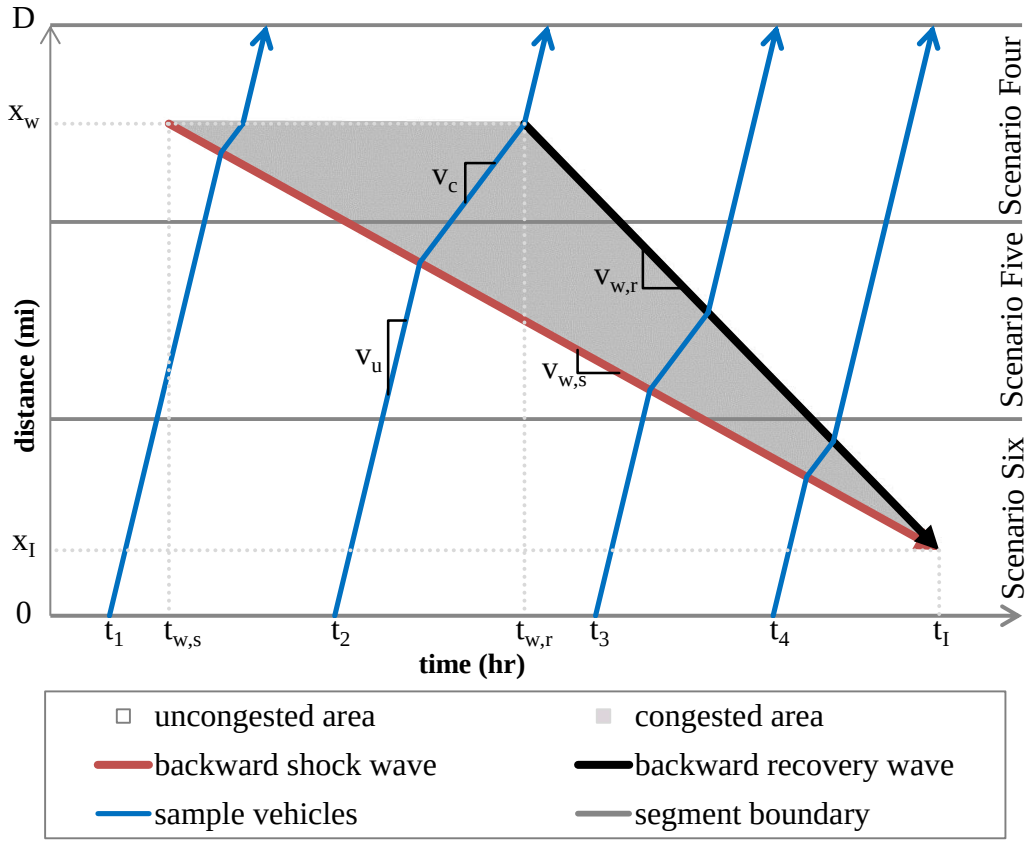


Figure 3.5 Freeway segments with backward shock and recovery waves.

The three freeway segments that comprise the corridor in Figure 3.5 experience either a Scenario Four, Five, or Six transition. The differences amongst the three Scenarios are described. First, the transition that affects the farthest downstream segment is classified as Scenario Four because vehicles intersect at most one shock wave within the segment. For example, the vehicle that departs at time t_1 only encounters the backward shock wave. In addition, the vehicle that departs at time t_2 is the first vehicle to encounter the backward recovery wave within the segment but it encounters the

backward shock wave in the previous, upstream segment. Another reason the transition is classified as Scenario Four is because the waves intersect upstream from the segment. Second, the transition in the middle segment is classified as Scenario Five. The change in classification from one freeway segment to the next occurs because over time, the waves' trajectories have less distance between them. Vehicles can encounter both shock waves within the middle segment, such as the vehicle that departs at time t_3. Note that not all vehicles encounter both shock waves within the segment, such as the vehicle that departs at time t_2. Another reason this is Scenario Five is because the waves intersect upstream from this segment also. Third, the transition is classified as Scenario Six in the farthest upstream segment. Vehicles can intersect both shock waves within the segment, such as the vehicle that departs at time t_4. Furthermore, the waves intersect within the segment. Therefore, the congestion also ends within this segment.

The characteristics of the different transition Scenarios have been provided. These include the number and type of shock waves that occur in the segment. Also, for Scenarios in which two transitions occur, the characteristics include whether the shock waves terminate within the segment and if vehicles intersect both of the shock waves within the segment. The categorization is necessary in order to model the relationships between actual travel time, predicted travel time, and departure time.

3.3 Vehicle Travel Time

The relationship between vehicle travel time and departure time over a study period is modeled. The relationship can be positive (travel times increase as departure times increase), negative (travel time decrease as departure times decrease), or constant (travel times are the same for all departure times). The type of relationship changes throughout

the study period. Specifically, it changes at the departure times of the first and last vehicles affected by the transition in the traffic state (i.e., the first and last vehicle to encounter the shock wave). Therefore, the relationship is determined using the following four steps for each Scenario. First, the departure times of the first and last vehicles to encounter the shock wave(s) are identified. Second, the travel times of these vehicles are calculated. Third, based on the calculations from the first two steps, the beginning and end of the study period are determined. Fourth, the information is used to develop equations that express the relationship between vehicle travel time and departure time for all other departure times within the study period. These steps are applied first to the Scenarios that consider one shock wave, Scenarios One, Two, and Three. The steps are then applied to the Scenarios that consider two shock waves, Scenarios Four, Five, and Six.

3.3.1 Scenarios One, Two, and Three

To begin, the departure times of the first and last vehicles to encounter the shock waves in Scenarios One, Two, and Three are determined. These times are denoted t_a and t_b, respectively. The departure times and trajectories of the first and last vehicles to encounter the shock wave are illustrated in Figures 3.2 through 3.4. The trajectory of the first vehicle that encounters the shock wave is illustrated in orange and the trajectory of the last vehicle that encounters the shock wave is illustrated in purple.

The first vehicle to encounter the shock wave does so at the instant and location the shock wave enters the segment (if the shock wave initializes outside of the segment) or at the instant and location the shock wave forms within the segment (if the shock wave initializes within the segment). This holds true because vehicle departure times are

assumed to be continuous. The departure time of the first vehicle to encounter a backward shock wave or a forward recovery wave is calculated with Eq. 3.1, and the departure time of the first vehicle to encounter a backward recovery wave is calculated with Eq. 3.2.

- Scenarios One and Three

$$t_a = t_w - \frac{x_w}{v_u} \tag{3.1}$$

- Scenario Two

$$t_a = t_w - \frac{x_w}{v_c} \tag{3.2}$$

As seen in Eqs. 3.1 and 3.2, the departure time is dependent on when and where the shock wave enters the segment and on the amount of time it takes a vehicle to reach this location.

Also, the last vehicle to encounter the shock wave does so at the instant and location the shock wave departs from the segment. The location is the upstream end of the segment if the shock wave is a backward moving wave or the location is the downstream end of the segment if the shock wave is a forward moving wave. The departure time of the last vehicle to encounter a backward shock wave or a backward recovery wave is calculated with Eq. 3.3 because both of these waves exit the segment at the upstream end. In the equation, subtraction is used instead of addition, because a backward moving wave has a negative velocity. The departure time of the last vehicle to encounter a forward recovery wave is calculated with Eq. 3.4 because this wave exits the segment at the downstream end.

- Scenarios One and Two

$$t_b = t_w - \frac{x_w}{v_w} \tag{3.3}$$

- Scenarios Three

$$t_b = t_w + \frac{D - x_w}{v_w} - \frac{D}{v_u} \tag{3.4}$$

Next, the travel times for the first and last vehicles to encounter the shock waves in Scenarios One, Two, and Three are calculated. The travel times are calculated with one of the following four equations. The first equation is used to calculate the travel time of a vehicle that travels the entire length of the segment at the uncongested speed. An example of such a vehicle is the first vehicle to encounter the backward shock wave in Figure 3.2. The travel time is calculated by dividing the segment length by the speed of the uncongested traffic state. The travel time is denoted A_t (hr) and is calculated with Eq. 3.5.

$$A_t = \frac{D}{v_u} \tag{3.5}$$

The second equation is used to calculate the travel time of a vehicle that travels the entire length of the segment at the congested speed. It is calculated by dividing the segment length by the speed of the congested traffic state as shown in Eq. 3.6.

$$A_t = \frac{D}{v_c} \tag{3.6}$$

The third equation is applied for a vehicle that travels from the beginning of the segment to the incident location at the uncongested speed and then continues to the end of the segment at the congested speed. This can occur when a shock wave originates within the segment. An example is the vehicle that departs at time t_a in Figure 3.4. The travel time is calculated with Eq. 3.7.

$$A_t = \frac{x_w}{v_u} + \frac{D - x_w}{v_c} \tag{3.7}$$

In contrast to the third equation, the last equation is applied to a vehicle that travels from the beginning of the segment to the incident location at the congested speed and then continues to the end of the segment at the uncongested speed. This can also occur when a shock wave originates within the segment. An example is the vehicle that departs at time t_b in Figure 3.2. The travel time is calculated with Eq. 3.8.

$$A_t = \frac{x_w}{v_c} + \frac{D - x_w}{v_u} \tag{3.8}$$

The specific application of the above equations for the first and last vehicles to encounter a shock wave in each Scenario is discussed. The equations used for Scenario One are discussed first. The travel time of the first vehicle to encounter the backward shock wave in Scenario One is calculated with Eq. 3.5 because the vehicle travels from the beginning to the end of the segment at the uncongested speed. The last vehicle to

encounter the wave travels at the congested speed from the beginning of the segment to the location the wave initiates. If the wave enters the segment at the downstream end, such that $x_w = D$, the travel time is calculated with Eq. 3.6 because the entire segment is in the congested traffic state once the wave passes through it. Otherwise, if the location is within the segment, such that $0 < x_w < D$ and as illustrated in Figure 3.1, the vehicle then travels at the uncongested speed from this location to the end of the segment. Therefore, its travel time is calculated with Eq. 3.8.

Next, the travel times of the first and last vehicles to encounter a backward recovery wave are discussed. The first vehicle to encounter the wave travels at the congested speed from the beginning of the segment to the location the wave initiates. If the location is downstream from the segment, the travel time is calculated with Eq. 3.6 because the entire segment is in the congested traffic state before the wave passes through it. Otherwise, if the location is within the segment such as illustrated in Figure 3.3, the vehicle then travels at the uncongested speed from this location to the end of the segment. Therefore, its travel time is calculated with Eq. 3.8. The travel time of the last vehicle to encounter the wave is calculated with Eq. 3.5 because the vehicle travels from the beginning to the end of the segment at the uncongested speed since the entire segment is in the uncongested traffic state once the wave passes through it.

Finally, the travel times of the first and last vehicles to encounter a forward recovery wave are discussed. If the wave initiates upstream from the segment, the travel time is calculated with Eq. 3.6 because the entire segment is in the congested traffic state before the wave passes through it. However, if the wave initiates within the segment such as illustrated in Figure 3.4, the first vehicle to encounter it travels at the uncongested

speed from the beginning of the segment to the location the wave initiates. It then travels at the congested speed from this location to the end of the segment. Therefore, its travel time is calculated with Eq. 3.7. The travel time of the last vehicle to encounter the wave is calculated with Eq. 3.5 because the vehicle travels from the beginning to the end of the segment at the uncongested speed since the entire segment is in the uncongested traffic state once the wave passes through it.

The next step is to determine the study periods for Scenarios One, Two, and Three. As defined previously, the study period starts at the departure time of the first vehicle affected by the shock wave and ends at the departure time of the last vehicle affected by the shock wave. The first and last departure times of the study period are denoted t_F and t_L, respectively. The study period therefore begins at t_a. Thus,

$$t_F = t_a \tag{3.9}$$

The equations to use for calculating t_a for each scenario were discussed previously and are summarized in Table 3.3. Vehicles are no longer affected by the shock wave once the last vehicle to encounter it exits the segment. Therefore, the study period ends when the vehicle that departs at t_b exits the segment. The value for t_L is calculated as follows

$$t_L = t_b + A_{t_b} \tag{3.10}$$

The equations to calculate t_b and A_{t_b} were discussed previously and are summarized in Tables 3.4 and 3.6, respectively.

Finally, the relationship between vehicle travel time and departure time is determined for the entire study period. In the previous steps, three unique departure times were determined, the time the study period begins which is the departure time of the first vehicle affected by the transition in the traffic state - t_a, the departure time of the last vehicle affected by the transition - t_b, and the time the study period ends - t_L. These three departure times divide the study period into two sub periods as follows:

1. The departure time of the first vehicle to encounter the backward shock wave (which is also the beginning of the study period) and the departure time of the last vehicle to encounter the backward shock wave.

2. The departure time of the last vehicle to encounter the backward shock wave and the time the study period ends.

Equations to calculate the travel times of vehicles within each of these two sub periods are developed.

Based on the departure times and travel times of the first and last vehicles to encounter a shock wave, an equation is developed to calculate the travel time of a vehicle that departs between t_a and t_b. This equation is developed based on two observations. First, the relationship between vehicle travel time and departure time over the study period is continuous. As stated in the assumptions, the function is continuous because vehicle departure times are assumed to be continuous. Second, the relationship between vehicle travel time and departure time is linear between t_a and t_b. Specifically, travel times increase at a constant rate from departure time t_a to time t_b in Scenario One. Also, travel times decrease at a constant rate from departure time t_a to time t_b in Scenarios Two and Three. The linearity is a result of two characteristics of the model. The first characteristic is that the traffic state changes at a constant rate. The second characteristic

is that vehicles' speeds are constant within each traffic state. Because the relationship is continuous and linear, a linear equation is used to describe a vehicle's travel time by its departure time. The rate of change in the equation (the slope) is found using the departure times and travel times of the first and last vehicles to encounter the shock wave. It equals the difference in their travel times divided by the difference in their departure times. The travel time of a vehicle that departs at time *t*, where *t* is between t_a and t_b, is found with Eq. 3.11.

$$A_t = \frac{A_{t_a} - A_{t_b}}{t_a - t_b}(t - t_a) + A_{t_a} \tag{3.11}$$

In the equation, A_{t_a} is the travel time of the first vehicle to encounter the shock wave and A_{t_b} is the travel time of the last vehicle to encounter the shock wave. The equations to determine the values of travel times and departure times were discussed previously and are summarized in Tables 3.3 through 3.6.

The travel times for vehicles that depart after time t_b until the end of the study period are also determined. These vehicles have the same travel time as the vehicle that departs at t_b because the traffic state is static after the shock wave exits the segment. The equations to calculate this travel time have been discussed previously.

Based on the above information, a function of vehicle travel time by departure time is produced for Scenarios One through Three. The function, denoted *f*, assigns to each departure time *t*, where $t_F \leq t \leq t_L$, an actual travel time. It is a piecewise defined function because it is defined by different formulas in different parts of its domain (i.e., $[t_F, t_L]$.

$$f(t) = \begin{cases} \frac{A_{t_a} - A_{t_b}}{t_a - t_b}(t - t_a) + A_{t_a}, & t_F \le t < t_b \\ A_{t_b} \quad , & t_b \le t \le t_L \end{cases} \tag{3.12}$$

Figures 3.2 through 3.4 display the function of vehicle travel time by departure time on the secondary y-axis with a blue line for Scenarios One, Two, and Three, respectively. For each Scenario, the function begins at the departure time of the first vehicle to encounter the shock wave, time $t_F = t_a$, and ends at the exit time of the last vehicle to encounter the shock wave, time t_L. For study periods in which the wave is a backward shock wave, the function linearly increases from t_a to t_b because the length of the segment that vehicles travel under congestion increases at a constant rate as the shock wave passes through the segment. On the other hand, if the wave is a backward recovery wave or forward recovery wave, the function linearly decreases from t_a to t_b since the length of the segment that vehicles travel under congestion decreases at a constant rate as the shock wave passes through the segment. Then, from the departure time of the last vehicle to encounter the wave until the end of the study period, the function is constant because the traffic state is no longer changing.

3.3.2 Scenarios Four, Five, and Six

In addition to addressing scenarios that include the passage of a single shock wave on a freeway segment, scenarios are developed in which the passage of a backward shock wave is followed by the passage of a backward recovery wave within the same study period. These include Scenarios Four, Five, and Six. Both waves begin at the same location, a site that has a decrease in capacity followed by a restoration of capacity, which can be caused by an incident. The decrease in capacity causes the backward shock

wave to form and the restoration of capacity causes the backward recovery wave to form. Over distance and time, the traffic state is uncongested to the left of the backward shock wave and to the right of the backward recovery wave but the traffic state is congested between the waves.

The waves are included within the same study period if at least one of the vehicles that intersects the backward shock wave travels on the freeway segment while at least one of the vehicles that intersects the backward recovery wave also does (i.e., the time at which the last vehicle to encounter the backward shock wave exits the segment is later than the departure time of the first vehicle to encounter the backward recovery wave). One can tell that a case study satisfies this condition if the inequality in Eq. 3.13 holds true.

$$t_{w,s} - \frac{x_w}{v_{w,s}} + \frac{2x_w}{v_c} + \frac{D - x_w}{v_u} \geq t_{w,r} \tag{3.13}$$

In the above equation, the times at which the backward shock wave and backward recovery wave enter the freeway segment are denoted by $t_{w,s}$ and $t_{w,r}$. The location at which they enter the freeway segment is denoted by x_w. The velocity of the backward shock wave and the velocity of the backward recovery wave are denoted by $v_{w,s}$ and $v_{w,r}$, respectively. However, if the inequality is not satisfied, the waves are separated into two study periods; the backward shock wave is classified as Scenario One and the backward recovery wave is classified as Scenario Two.

The three scenarios in which the passage of a backward shock wave is followed by the passage of a backward recovery wave, Scenarios Four, Five, and Six, are

differentiated. In Scenario Four, no vehicles intersect both shock waves. In Scenario Five and Six, at least one vehicle intersects more than one wave. Whether vehicles intersect at most one or both shock waves is dependent on the trajectories of the shock waves and on the congested vehicle speed. If the inequality in Eq. 3.14 is true, no vehicles intersect both waves and the study period is classified as Scenario Four.

$$\frac{x_w}{v_c} - \frac{x_w}{v_{w,s}} < t_{w,r} - t_{w,s} \tag{3.14}$$

In the above inequality, the left side represents the amount of time it takes for the backward shock wave to reach the beginning of the segment and a vehicle to pass through the entire congested area. The right side of the inequality represents the incident duration. Therefore, the inequality represents whether the backward shock wave reaches the beginning of the segment and a vehicle can pass through the entire congested area before the incident is cleared.

In both Scenario Five and Scenario Six the inequality in Eq. 3.11 is not satisfied. Therefore, a vehicle can intersect both shock waves while traversing the segment. However, Scenario Five and Scenario Six differ in the location at which the backward shock wave and the backward recovery wave intersect each other. In Scenario Five, the shock waves intersect upstream from the segment. However, in Scenario Six, the shock waves intersect within or at the upstream end of the segment. Whether or not shock waves intersect depends on the amount of time between their occurrences and on their speeds. The location at which two waves intersect, denoted x_I, is found with Eq. 3.15.

This equation is applicable if the speed of the backward recovery wave is faster than the speed of the backward shock wave.

$$x_I = \frac{v_{w,s}\left(-x_w + v_{w,r}t_{w,r}\right) - v_{w,r}\left(-x_w + v_{w,s}t_{w,s}\right)}{v_{w,r} - v_{w,s}} \tag{3.15}$$

If x_I is less than zero, then the waves intersect upstream from the segment and the study period is classified as Scenario Five. If x_I is greater than zero, then the waves intersect within the segment. If x_I equals zero, the waves intersect at the upstream end of the segment, which is mile zero. For either result, the study period is classified as Scenario Six. Scenario Six was illustrated in Figure 3.5. In the figure, the location of x_I is denoted on the y-axis. The flow chart below illustrates the process of determining the classification of the Scenarios for study periods with both a backward shock wave and backward recovery wave.

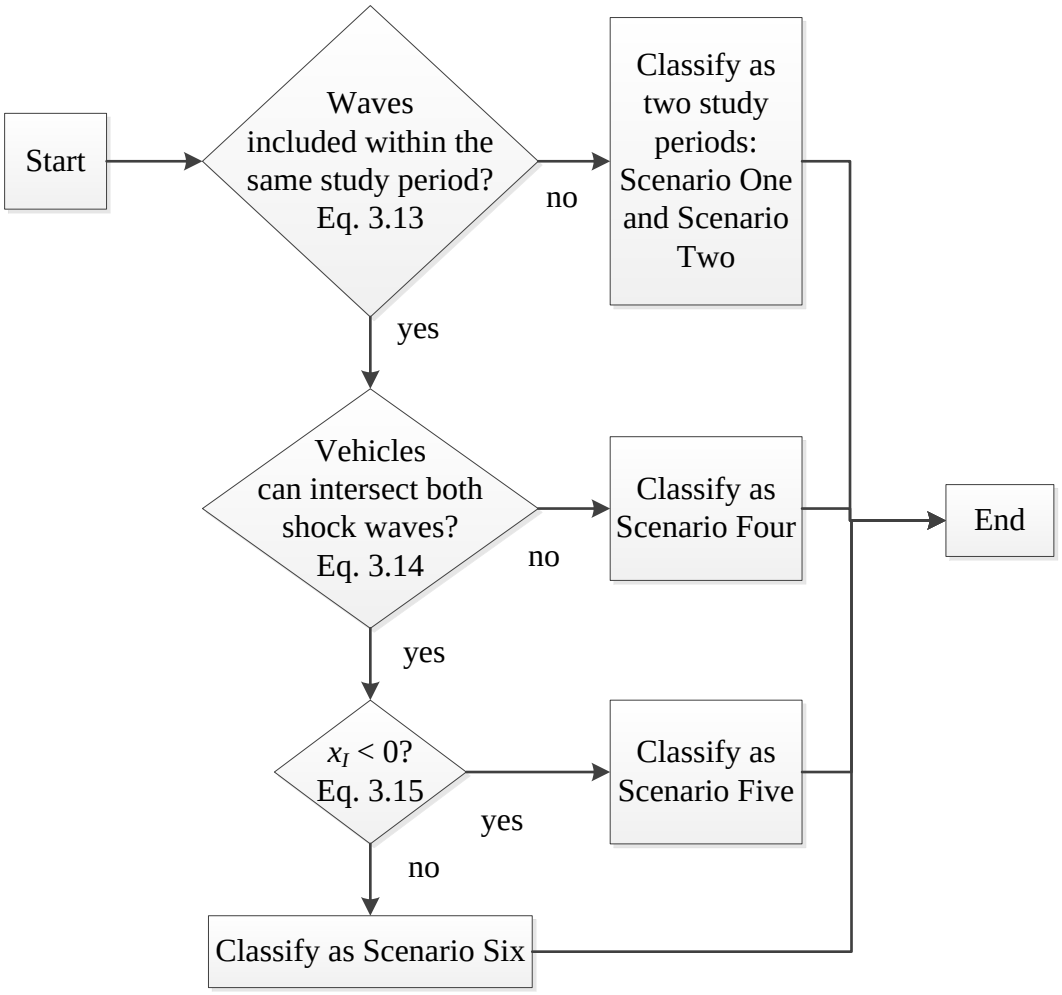


Figure 3.6 Flow chart of classification of Scenarios with two shock waves.

After distinguishing Scenarios Four, Five, and Six, the steps to determine vehicles' travel times are now applied. The first and second steps are discussed together. These steps are used to determine the departure times and travel times of the first and last vehicles to encounter the backward shock wave as well as the first and last vehicles to encounter the backward recovery wave. The departure times and trajectories of these vehicles are illustrated in Figures 3.7 through 3.9. The departure times of the first and last vehicles to encounter the backward shock wave are denoted $t_{a,s}$ and $t_{b,s}$. The trajectories of these vehicles are illustrated in orange and purple, respectively. The

departure times of the first and last vehicles to encounter the backward recovery wave are denoted $t_{a,r}$ and $t_{b,r}$. The trajectories of these vehicles are illustrated in green and pink, respectively.

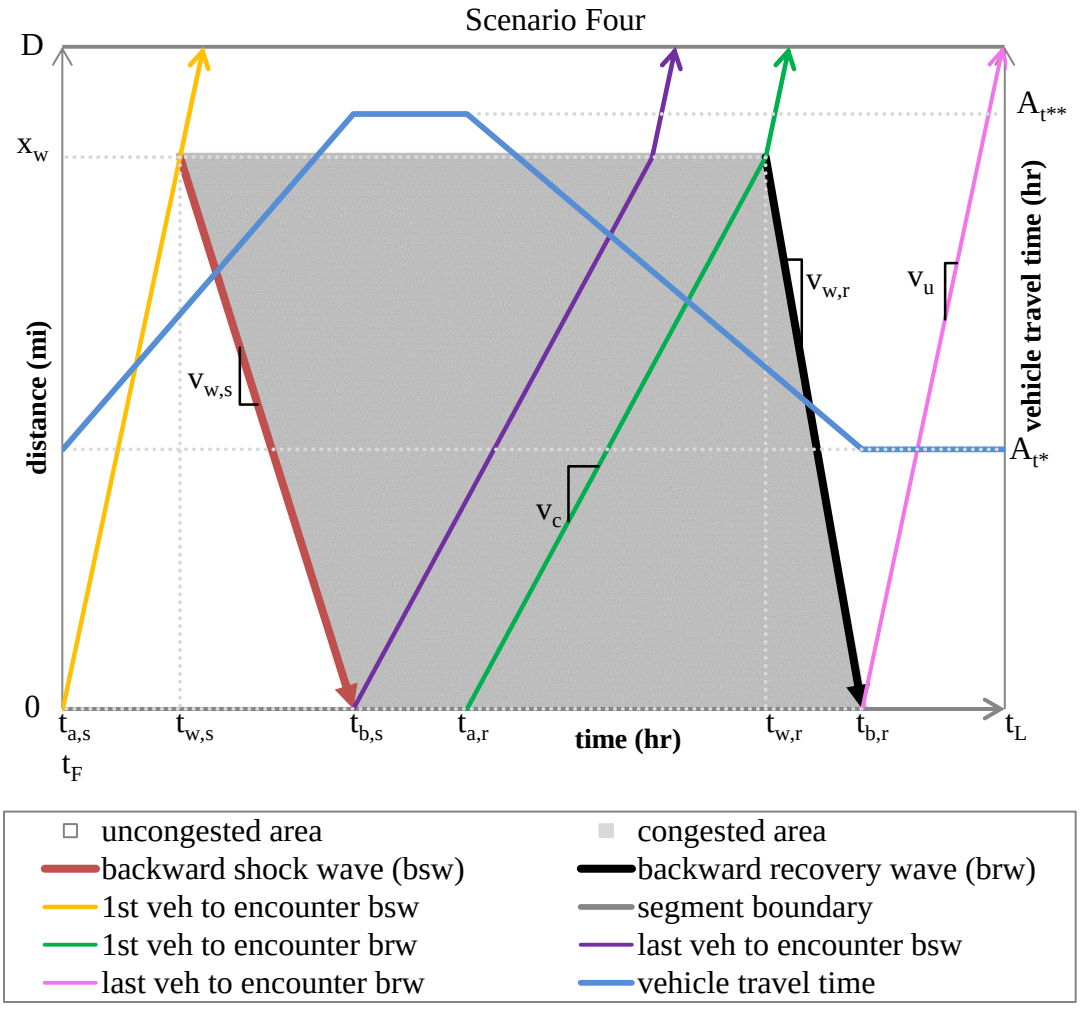


Figure 3.7 Freeway segment undergoing transition Scenario Four.
where $t^* = t_{a,s}$ and $t_{b,r} \leq t^* \leq t_L$, $t_{b,s} \leq t^{**} \leq t_{a,r}$

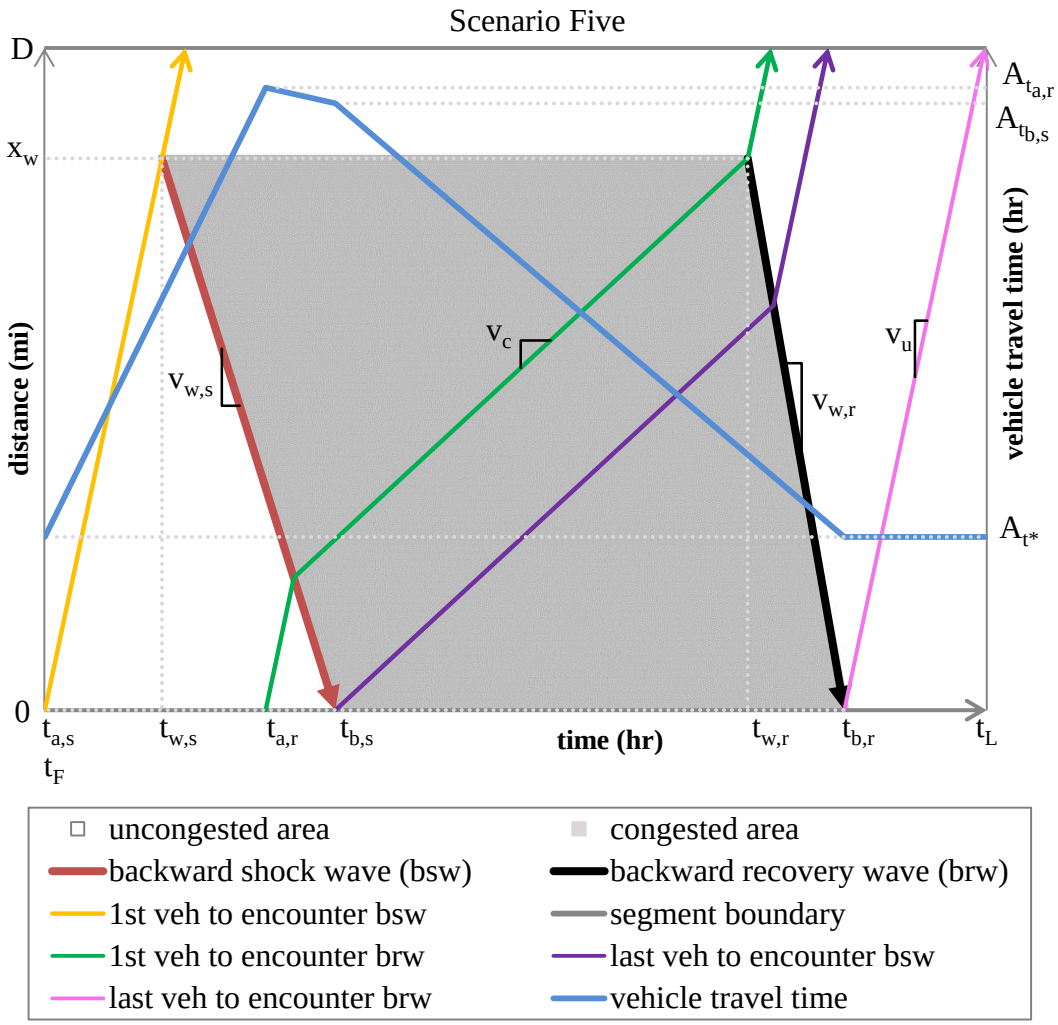


Figure 3.8 Freeway segment undergoing transition Scenario Five.
where $t^* = t_{a,s}$ and $t_{b,r} \leq t^* \leq t_L$

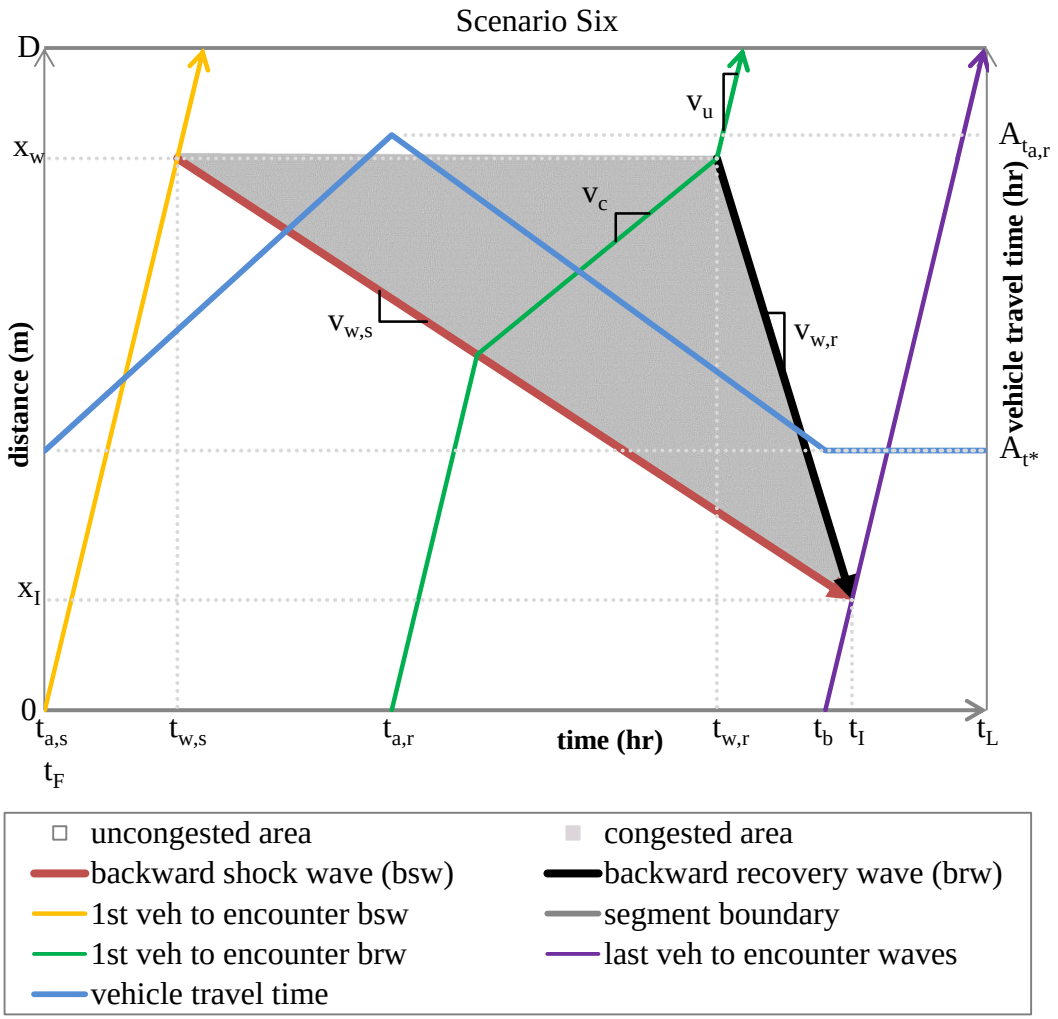


Figure 3.9 Freeway segment undergoing transition Scenario Six.
where $t^* = t_{a,s}$ and $t_b \leq t^* \leq t_L$

In Scenario Four, the first and last vehicles to encounter the shock waves do not encounter any other waves within the segment. This was also the case in Scenario One and Two, which considered the passage of a backward shock wave and backward recovery wave, respectively. Therefore, the travel times of the first and last vehicles to encounter the backward shock wave are calculated as explained for Scenario One. Also, the departure times and travel times of the first and last vehicles to encounter the backward recovery wave are calculated as explained for Scenario Two.

In Scenario Five, the first vehicle to encounter the backward shock wave and the last vehicle to encounter the backward recovery wave do not encounter any other waves within the segment. Therefore, the departure time and travel time of the first vehicle to encounter the backward shock wave is calculated as explained for Scenario One. Also, the departure time and travel time of the last vehicle to encounter the backward recovery wave is calculated as explained for Scenario Two. However, the last vehicle to encounter the backward shock wave encounters the backward recovery wave within the segment. While its departure time can be calculated as in Scenario One, its travel time cannot. Instead, its travel time, denoted $A_{t_{b,s}}$, is calculated with Eq. 3.16.

$$A_{t_{b,s}} = \left(\frac{-t_{b,s}v_c - x_w + t_{w,r}v_{w,r}}{v_{w,r} - v_c} - t_{b,s}\right)\left(1 - \frac{v_c}{v_u}\right) + \frac{D}{v_u} \tag{3.16}$$

In the equation, the first term on the right side represents the time at which the vehicle intercepts the backward recovery wave. The value inside the parenthesis therefore equals the travel time at the congested speed. The vehicle's departure time, $t_{b,s}$, is applied in the equation.

The first vehicle to encounter the backward recovery wave in Scenario Five encounters the backward shock wave within the segment. Its departure time is calculated with Eq. 3.17.

$$t_{a,r} = \left(\frac{-t_{w,r}v_c + t_{w,s}v_{w,s}}{v_{w,s} - v_c}\right)\left(1 - \frac{v_c}{v_u}\right) + \frac{v_c t_{w,r} - x_w}{v_u} \tag{3.17}$$

In the equation, the first term on the right side represents the time at which the vehicle intercepts the backward shock wave. The vehicle's travel time is found with Eq. 3.18.

$$A_{t_{a,r}} = t_{w,r} - t_{a,r} + \frac{D - x_w}{v_u} \tag{3.18}$$

In the equation, the first two terms on the right side represent the travel time from when the vehicle departs until it reaches the location the waves initiate. The second term on the right side represents the travel time from the location the waves initiate to the end of the segment.

In Scenario Six, the first vehicle to encounter the backward shock wave does not encounter any other waves within the segment. Therefore, its departure time and travel time are calculated as explained for Scenario One. The first vehicle to encounter the backward recovery wave encounters the backward shock wave within the segment. Therefore, its departure time and travel time are calculated as explained in Scenario Five for the first vehicle to encounter the backward recovery wave. Because the backward shock wave and backward recovery wave intersect within the segment, there is one vehicle that is the last to intersect both the backward shock wave and the backward recovery wave. To calculate its departure time, first the time at which the waves intersect is calculated. The time they intersect, denoted t_I, is found with Eq. 3.19

$$t_I = t_{w,s} - \frac{x_w - x_I}{v_{w,s}} \tag{3.19}$$

where x_I is found with Eq. 3.15. Then, the vehicle's departure time is calculated with Eq. 3.20.

$$t_b = t_I - \frac{x_I}{v_u} \tag{3.20}$$

This vehicle travels the entire segment at the uncongested speed because the traffic state is uncongested in the time-space plane outside of the boundaries determined by the shock waves' trajectories. The travel time of the vehicle is therefore calculated with Eq. 3.5.

The next step is to determine the study periods for Scenarios Four, Five, and Six. The study period begins at the departure time of the first vehicle affected by the backward shock wave. The departure time of this vehicle was defined previously to be $t_{a,s}$. Thus,

$$t_F = t_{a,s} \tag{3.21}$$

where the equations to calculate $t_{a,s}$ for each Scenario were provided previously. Vehicles are no longer affected by the shock waves once the last vehicle to encounter the backward recovery wave exits the segment. Therefore, for Scenarios Four and Five, the study period ends when the vehicle that departs at $t_{b,r}$ exits the segment. The value for t_L, the last departure time in the study period, is calculated as follows

$$t_L = t_{b,r} + A_{t_{b,r}} \tag{3.22}$$

where the equations to calculated $t_{b,r}$ and $A_{t_{b,r}}$ were provided previously. For Scenario Six, Eq. 3.10 is applied because the departure time of the last vehicle to encounter the shock wave in Scenario Six is denoted t_b.

Finally, the relationship between vehicle travel time and departure time is determined for the entire study period. The relationship is illustrated in Figures 3.7, 3.8, and 3.9 where the function of travel time by departure time is shown with a blue line on the secondary y-axis. Each Scenario is discussed separately.

The forth step is first applied to Scenario Four. The relationship between vehicle travel time and departure time is constant between the departure times corresponding to the following events:

1. The beginning of the study period (time t_F found with Eq. 3.21), which is also the departure time of the first vehicle to encounter the backward shock wave, and the departure time of the last vehicle to encounter the backward shock wave.
2. The departure time of the last vehicle to encounter the backward shock wave and the departure time of the first vehicle to encounter the backward recovery wave.
3. The departure time of the first vehicle to encounter the backward recovery wave and the departure time of the last vehicle to encounter the backward recovery wave.
4. The departure time of the last vehicle to encounter the backward recovery wave and the end of the study period (time t_L found with Eq. 3.22).

The first sub period is between the departure times of the first and last vehicles to encounter the backward shock wave, $t_{a,s}$ and $t_{b,s}$. During this sub period, travel times increase at a constant rate. The reasoning is the same as explained for Scenario One for the sub period between t_a and t_b. The travel time equation is derived in the same way as in Scenario One where the rate of change in the equation is found using the departure times and travel times of the first and last vehicles to encounter the backward shock wave. The

travel time of a vehicle that departs at time *t*, where *t* is between $t_{a,s}$ and $t_{b,s}$, is found with Eq. 3.23 which is a generalized form of Eq. 3.11.

$$A_t = \frac{A_x - A_y}{x - y}(t - x) + A_x \tag{3.23}$$

In the equation, $t_{a,s}$ is represented by *x* and $t_{b,s}$ is represented by *y*. Thus, A_x, equals the travel time for the vehicle that departs at $t_{a,s}$ and A_y equals the travel time for the vehicle that departs at $t_{b,s}$.

During the second sub period, between the departure times of the last vehicle to encounter the backward shock wave and the first vehicle to encounter the backward recovery wave, vehicle travel times are constant; the vehicles that depart during this time period do not encounter any waves. The travel times for these vehicles therefore equal the travel times of the vehicles that depart at $t_{b,s}$ and $t_{a,r}$.

Vehicles that depart during the third sub period encounter the backward recovery wave. Like in Scenario Two for departure times between t_a and t_b, the departure times decrease at a constant rate. The travel time equation for this sub period is derived the same way as in Scenario Two where the rate of change in the equation is found using the departure times and travel times of the first and last vehicles to encounter the backward recovery wave. The travel time of a vehicle that departs at time *t*, where *t* is between $t_{a,r}$ and $t_{b,r}$, is found with Eq. 3.23 where $t_{a,r}$ is represented by *x* and $t_{b,r}$ is represented by *y*. Thus, A_x equals the travel time of the first vehicle to encounter the backward recovery wave and A_y equals the travel time of the last vehicle to encounter the backward recovery wave.

The last sub period is from the departure time of the last vehicle to encounter the backward recovery wave until the end of the study period. Vehicles that depart during this time have the same travel time as the vehicle that departs at $t_{b,r}$ because the traffic state is static after the shock wave exits the segment.

Based on the above information, a function of vehicle travel time by departure time is produced for Scenario Four. The function, denoted *f*, assigns to each departure time *t*, where $t_F \leq t \leq t_L$, an actual travel time. Like the function for Scenarios One through Three, it is a piecewise defined function.

$$f(t) = \begin{cases} \frac{A_{t_{a,s}} - A_{t_{b,s}}}{t_{a,s} - t_{b,s}}\left(t - t_{a,s}\right) + A_{t_{a,s}}, & t_F \leq t < t_{b,s} \\ A_{t_{b,s}}, & t_{b,s} \leq t \leq t_{a,r} \\ \frac{A_{t_{a,r}} - A_{t_{b,r}}}{t_{a,r} - t_{b,r}}\left(t - t_{a,r}\right) + A_{t_{a,r}}, & t_{a,r} < t < t_{b,r} \\ A_{t_{b,r}}, & t_{b,r} \leq t \leq t_L \end{cases} \tag{3.24}$$

The function is illustrated in Figure 3.7. The function is depicted with a blue line on the secondary y-axis. As described by the piecewise function, the slope of the function changes at the departure times of the first and last vehicles to encounter the backward shock wave and the departure times of the first and last vehicles to encounter the backward recovery wave.

The forth step is next applied to Scenario Five. The relationship between vehicle travel time and departure time is constant between the departure times corresponding to the following events:

1. The beginning of the study period and the departure time of the first vehicle to encounter the backward recovery wave.

2. The departure time of the first vehicle to encounter the backward recovery wave and the departure time of the last vehicle to encounter the backward shock wave.

3. The departure time of the last vehicle to encounter the backward shock wave and the departure time of the last vehicle to encounter the backward recovery wave.

4. The departure time of the last vehicle to encounter the backward recovery wave and the end of the study period.

In Scenario Five, the first sub period is between the departures time of the first vehicles to encounter the backward shock wave and the backward recovery wave. During this period, the travel times linearly increase because the length of the segment under congestion linearly increases as the backward shock wave passes through the segment. The travel time of a vehicle that departs at time t, where t is between $t_{a,s}$ and $t_{a,r}$, is found with Eq. 3.23 where $t_{a,s}$ is represented by x and $t_{a,r}$ is represented by y.

The second sub period is between the departure time of the first vehicle to encounter the backward recovery wave and the last vehicle to encounter the backward shock wave. During this time period, the backward recovery wave passes through the segment and the length of the roadway in congested conditions linearly decreases. As a result, travel times during this period linearly decrease. The travel time of a vehicle that departs at time t, where t is between $t_{a,r}$ and $t_{b,s}$, is found with Eq. 3.23 where t_{ar} is represented by x and $t_{b,s}$ is represented by y.

The third sub period is between the departure time of the last vehicle to encounter the backward shock wave and the last vehicle to encounter the backward recovery wave. The travel times continue to decrease linearly, however at a faster rate than in the second sub period. This is because the length of the segment in congested conditions now is

only affected by the trajectory of the backward recovery wave and not the trajectory of the backward shock wave. The travel time of a vehicle that departs at time *t*, where *t* is between $t_{b,s}$ and $t_{b,r}$, is found with Eq. 3.23 where $t_{b,s}$ is represented by *x* and $t_{b,r}$ is represented by *y*.

The last sub period is from the departure time of the last vehicle to encounter the backward recovery wave until the end of the study period. Vehicles that depart during this time have the same travel time as the vehicle that departs at $t_{b,r}$ because the traffic state is static after the shock wave exits the segment.

Based on the above information, a function of vehicle travel time by departure time is produced for Scenario Five. The function, denoted *f*, assigns to each departure time *t*, where $t_F \leq t \leq t_L$, an actual travel time.

$$f(t) = \begin{cases} \frac{A_{t_{a,s}} - A_{t_{a,r}}}{t_{a,s} - t_{a,r}}\left(t - t_{a,s}\right) + A_{t_{a,s}}, & t_F \leq t < t_{a,r} \\ \frac{A_{t_{a,r}} - A_{t_{b,s}}}{t_{a,r} - t_{b,s}}\left(t - t_{a,r}\right) + A_{t_{a,r}}, & t_{a,r} \leq t < t_{b,s} \\ \frac{A_{t_{b,s}} - A_{t_{b,r}}}{t_{b,s} - t_{b,r}}\left(t - t_{b,s}\right) + A_{t_{b,s}}, & t_{b,s} \leq t < t_{b,r} \\ A_{t_{b,r}}, & t_{b,r} \leq t \leq t_L \end{cases} \tag{3.25}$$

The function is depicted in Figure 3.8 with a blue line on the secondary y-axis. As described by the piecewise function, the slope of the function changes at the departure times of the first and last vehicles to encounter the backward shock wave and the departure times of the first and last vehicles to encounter the backward recovery wave.

The maximum value is achieved at the departure time of the first vehicle to encounter the backward recovery wave while a local maximum is achieved at the departure time of the last vehicle to encounter the backward shock wave.

Last, the forth step is applied to Scenario Six. The relationship between vehicle travel time and departure time is constant between the departure times corresponding to the following events:

1. The beginning of the study period and the departure time of the first vehicle to encounter the backward recovery wave.
2. The departure times of the first vehicle to encounter the backward recovery wave and the last vehicle to encounter the shockwaves.
3. The departure time of the last vehicle to encounter the shock waves and the end of the study period.

In Scenario Six the first sub period is the same as for Scenario Five. Therefore, the equation developed in Scenario Five is applied. The second sub period is also the same for both Scenarios. The equations developed in Scenario Five are again applied with the exception that the variable t_b is used in the place of $t_{b,s}$. The last sub period is from the departure time of the last vehicle to encounter the waves until the end of the study period. Vehicles that depart during this time have the same travel time as the vehicle that departs at t_b because the traffic state is uncongested after the shock waves terminate.

Based on the above information, a function of vehicle travel time by departure time is produced for Scenario Six. The function, denoted *f*, assigns to each departure time *t*, where $t_F \leq t \leq t_L$, an actual travel time. Unlike the functions for Scenarios Four and Five, it only has three pieces because the waves intercept within the segment.

$$f(t) = \begin{cases} \frac{A_{t_{a,s}} - A_{t_{a,r}}}{t_{a,s} - t_{a,r}}\left(t - t_{a,s}\right) + A_{t_{a,s}}, & t_F \le t < t_{a,r} \\ \frac{A_{t_{a,r}} - A_{t_b}}{t_{a,r} - t_b}\left(t - t_{a,r}\right) + A_{t_{a,r}}, & t_{a,r} \le t < t_b \\ A_{t_b}, & t_b \le t \le t_L \end{cases} \quad (3.26)$$

The function is depicted in Figure 3.9 with a blue line on the secondary y-axis. The maximum value is achieved at the departure time of the first vehicle to encounter the backward recovery wave.

In summary, a model of vehicle travel time by departure time has been produced for Scenarios One through Six. The model was formed by first determining the departure times of the first and last vehicles to encounter the shock waves. Second, the travel times of these vehicles were calculated. Third, the study period was found to begin at the departure time of the first vehicle to encounter a shock wave and end at the time the last vehicle to encounter a shock wave exited the segment. Fourth, sub periods of the study period were determined for which the relationship (i.e., slope of the function) between vehicle travel time and departure time is constant and then the travel times for each sub period were calculated by applying the departure time and travel times of the first and last vehicles within each sub period.

Tables 3.3 through 3.7 summarize equations used to model vehicle travel time by departure time for each Scenario. Table 3.3 and Table 3.4 provide the equations to calculate the departure time of the first vehicle and the last vehicle to encounter each wave by Scenario. Table 3.5 and Table 3.6 provide the equations to calculate the travel

time of the first vehicle and the last vehicle to encounter each wave by Scenario. Table 3.7 provides the equations to determine the first and last departure time for each Scenario.

Table 3.3 Departure Time of the First Vehicle to Encounter a Wave by Scenario

		Scenario					
		1	2	3	4	5	6
Wave	Backward shock	3.1	n/a	n/a	3.1	3.1	3.1
	Backward recovery	n/a	3.2	n/a	3.2	3.17	3.17
	Forward recovery	n/a	n/a	3.1	n/a	n/a	n/a

n/a means non-applicable for the Scenario

Table 3.4 Departure Time of the Last Vehicle to Encounter a Wave by Scenario

		Scenario					
		1	2	3	4	5	6
Wave	Backward shock	3.3	n/a	n/a	3.3	3.3	3.19
	Backward recovery	n/a	3.3	n/a	3.3	3.3	3.19
	Forward recovery	n/a	n/a	3.4	n/a	n/a	n/a

n/a means non-applicable for the Scenario

Table 3.5 Travel Time of the First Vehicle to Encounter a Wave by Scenario

		Scenario					
		1	2	3	4	5	6
Wave	Backward shock	3.5	n/a	n/a	3.5	3.5	3.5
	Backward recovery	n/a	3.6 or 3.8	n/a	3.6 or 3.8	3.18	3.18
	Forward recovery	n/a	n/a	3.6 or 3.7	n/a	n/a	n/a

n/a means non-applicable for the Scenario

Table 3.6 Travel Time of the Last Vehicle to Encounter a Wave by Scenario

		Scenario					
		1	2	3	4	5	6
Wave	Backward shock	3.6 or 3.8	n/a	n/a	3.6 or 3.8	3.16	3.20
	Backward recovery	n/a	3.5	n/a	3.5	3.5	3.20
	Forward recovery	n/a	n/a	3.5	n/a	n/a	n/a

n/a means non-applicable for the Scenario

Table 3.7 First and Last Departure Times of a Study Period by Scenario

		Scenario					
		1	2	3	4	5	6
Wave	First vehicle	3.9	3.9	3.9	3.21	3.21	3.21
	Last vehicle	3.10	3.10	3.10	3.22	3.22	3.22

3.4 Predicted Vehicle Travel Time

In the previous section, 3.2.3, a model of the relationship between vehicle travel time and departure time was produced. Six different Scenarios were considered in which either one or two transitions in the traffic state of a freeway segment were considered. In this section, a model of the relationship between predicted travel time and departure time is developed considering the same Scenarios. First, the assumptions pertaining to modeling the predicted travel times are discussed. Then, the relationship between predicted travel times and departure time for each Scenario are modeled.

In order to formulate the model, system assumptions are made. The first assumption is that the sensors are evenly spaced along the freeway segment. The locations of path based sensors are as follows. Given a number of sensors, denoted *N*, one sensor is placed at the upstream end of the segment, which is mile zero, and one is placed at the downstream end of the segment, which is mile *D*. The remaining *N*-2 sensors are evenly spaced along the segment. The sensor spacing, denoted *S* (mi), is calculated with Eq. 3.27.

$$S = \frac{D}{N-1} \tag{3.27}$$

The location of sensor *n* (*n* = 1 to *N*) is mile x_n along the segment, where x_n is calculated with Eq. 3.28.

$$x_n = (n-1)S \tag{3.28}$$

where $1 \leq n \leq N$

The section of the freeway segment between two neighboring sensors, *n*-1 and *n*, is referred to as a link. The link is from location x_{n-1} to location x_n.

A freeway segment with *N* evenly spaced sensors can be seen in Figure 3.10. The sensors' locations are denoted on the y-axis. The locations are constant over time as shown with thick gray lines. The sensors divide the segment into *N*-1 links. The freeway segment is undergoing a transition from an uncongested to a congested traffic state that is represented with a backward shock wave. The backward shock wave originates in the link between sensor *n*-1 and sensor *n*. The study period is classified as Scenario One.

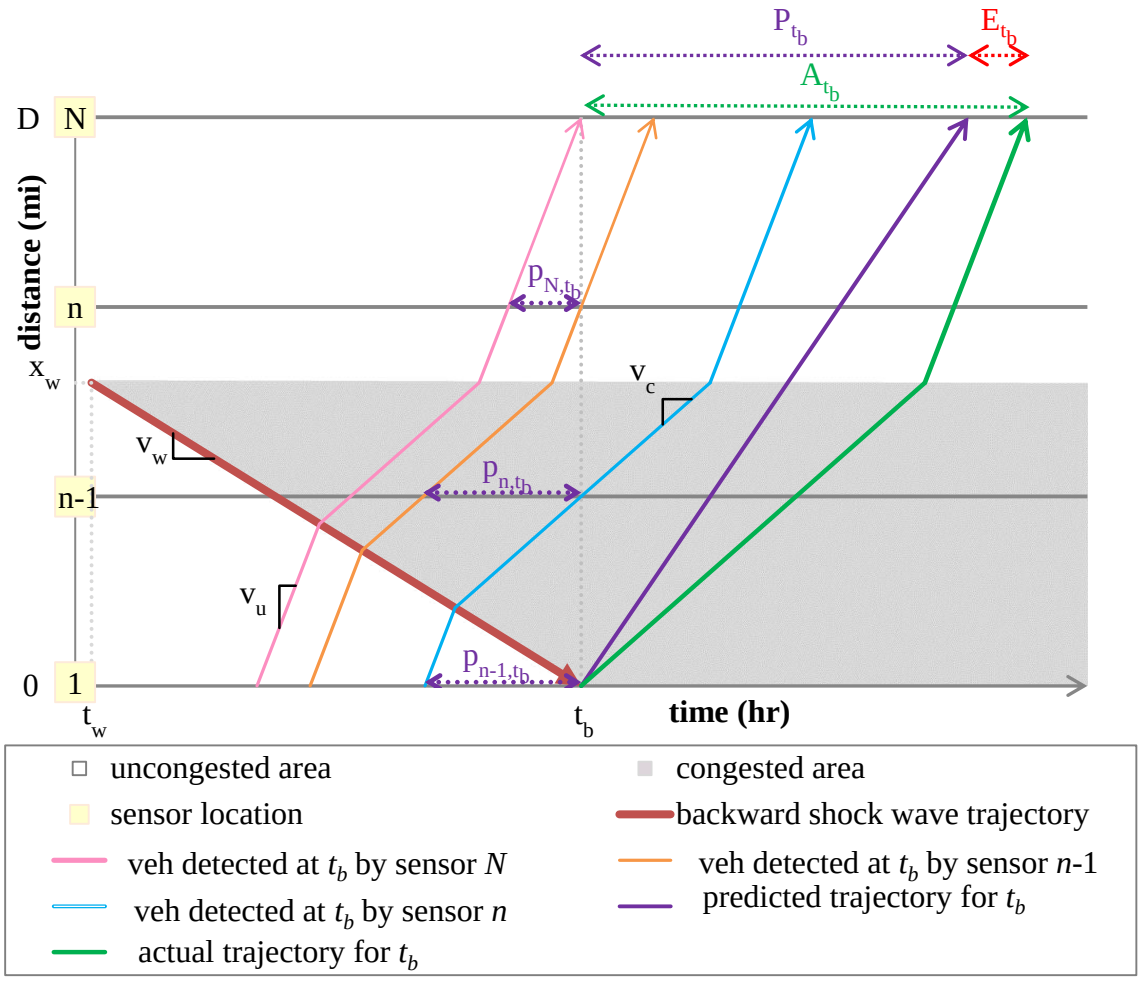


Figure 3.10 Freeway segment with backward shock wave and *N* sensors.

The data collected by two neighboring sensors, *n*-1 and *n*, is used to predict vehicles' travel times for the link defined by the two sensors. Each sensor collects data from passing vehicles equipped with traceable devices, such as Bluetooth devices. The data collected from a vehicle includes a unique identification number and the time at which it passes the sensor's location. By comparing the time at which a vehicle passes sensor *n-1* to the time at which it passes sensor *n*, the vehicle's travel time on a link is determined. The latest vehicle detected by sensor *n* provides the most recent travel time for mile x_{n-1} to mile x_n, and thus this travel time is used as the current predicted travel time for this link. The predicted travel time derived from a vehicle that is detected by sensor *n* at time *t* is denoted $p_{n,t}$.

The concept of predicted travel time is illustrated in Figure 3.10. In particular, the predicted travel time for the vehicle that departs at time t_b is demonstrated. The prediction is derived from the vehicles detected by the sensors at time t_b. The trajectory of the vehicles detected by sensors *N*, *n*, and *n*-1 at time t_b are depicted with a pink line, orange line, and blue line, respectively. The travel time from sensor *n* to sensor *N* of the vehicle whose trajectory is indicated with the pink line is denoted as p_{N,t_b} in the figure. It has this denotation because it is the predicted travel time derived at time t_b from the vehicle that is detected by sensor *N*. Likewise, the travel time from sensor *n-1* to sensor *n* of the vehicle whose trajectory is indicated with the orange line is denoted p_{n,t_b} and the travel time from sensor *1* to sensor *n-1* of the vehicle whose trajectory is indicated with the blue line is denoted p_{n-1,t_b}.

As it is assumed that predicted travel times are calculated instantaneously, the predicted travel time provided to a vehicle entering the segment at time *t* is derived from

the vehicles that are detected by the sensors at time *t*. As vehicles enter the freeway segment they receive a prediction of the amount of time it will take to travel from mile zero to mile D, the entire length of the segment. The prediction for a vehicle that departs at time *t* equals the sum of the most recent travel times calculated at time *t* by each pair of consecutive sensors (i.e., sensors *n*-1 and *n* for *n*=2 to *N*) as shown in Eq. 3.29.

$$P_t = \sum_{n=2}^{N} p_{n,t} \tag{3.29}$$

This concept is illustrated in Figure 3.10. The predicted trajectory for departure time t_b is shown with a purple line. (Note that the line is straight and therefore conveys a constant vehicle speed over the length of the segment. This reflects that vehicles' occupants assume the traffic state of the freeway segment is constant; the vehicles' occupants have no knowledge of where a transition in the traffic state occurs. Thus, the occupants assume the vehicle's speed will be constant for the entire length of the segment.) The predicted travel time of the vehicle, denoted P_{t_b} in the figure, equals the sum of the predictions from each link of the segment.

Predicted travel times can vary for the different links of a freeway segment during the same study period. The predicted travel times are affected by the number of waves that affect the link, whether or not the wave begins and/or ends with the link, and by the number of waves vehicle encounter while in the link. For example, if a link is located downstream from a backward moving wave's origin point or upstream from its termination point, then the predicted travel times for the link only reflect the uncongested

traffic state. Nine different link types are developed that consider these different characteristics. A freeway segment's links are categorized as fitting one of the nine link types to facilitate the calculation of the predicted travel times on each link. The table below summarizes links' characteristics for each link type. The characteristics are discussed in detail afterward.

Table 3.8 Characteristics of Link Types

Characteristics of the Link		Link Type								
		I	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IX**
Number of waves that affect the link		zero	one	one	two	two	two	two	two	two
Waves originate within the link		n/a	yes	no	yes	yes	yes	no	no	no
Waves terminate within the link		n/a	no	no	no	no	yes	no	no	yes
Vehicles can encounter both waves within the link		n/a	n/a	n/a	no	yes	yes	no	yes	yes
Scenarios During Which a Link Type Can Exist	One	yes	yes	yes	no	no	no	no	no	no
	Two	yes	yes	yes	no	no	no	no	no	no
	Three	yes	yes	yes	no	no	no	no	no	no
	Four	yes	no	no	yes	no	no	yes	no	no
	Five	yes	no	no	yes	yes	no	yes	yes	no
	Six	yes	no	no	yes	yes	yes	yes	yes	yes

n/a means non-applicable for the situation

The first link type, Type I, is not affected by a transition in the traffic state. This type of link may be present during any of the Scenarios. In Scenarios One, Two, Four, Five, and Six, a link receives this classification if the wave originates upstream of the freeway link (i.e., $x_w \leq x_{n-1}$). A link can also receive this classification in Scenario Six if the backward shock wave and backward recovery wave intersect downstream of the link (i.e., $x_n \leq x_I$ where x_I is found with Eq. 3.15). In Scenario Three, a link receives this classification if the wave originates downstream of the freeway section (i.e., $x_w \geq x_n$).

The next two link types only apply to links during Scenarios with one shock wave, i.e., Scenarios One, Two, and Three. Links are classified as Type II if the shock wave originates within it (i.e., $x_{n-1} < x_w < x_n$ for Scenarios One, Two, and Three). In contrast, links are classified as Type III if the shock wave passes through the entire link (i.e., $x_w \geq x_n$ for Scenarios One and Two and $x_w \leq x_{n-1}$ for Scenario Three).

The classifications of the links for a Scenario Three study period are illustrated in Figure 3.11. The freeway segment is undergoing a transition from a congested to an uncongested traffic state that is represented with a forward recovery wave. The freeway segment is comprised of *N*-1 links. The upstream link, located between sensors *1* and *n*, is classified as a Type II link because the forward recovery wave originates within this link. The downstream link, located between sensors *n* and *N*, is classified as a Type III link because the forward recovery wave passes through the entire link. The link types are indicated on the secondary y-axis of the figure.

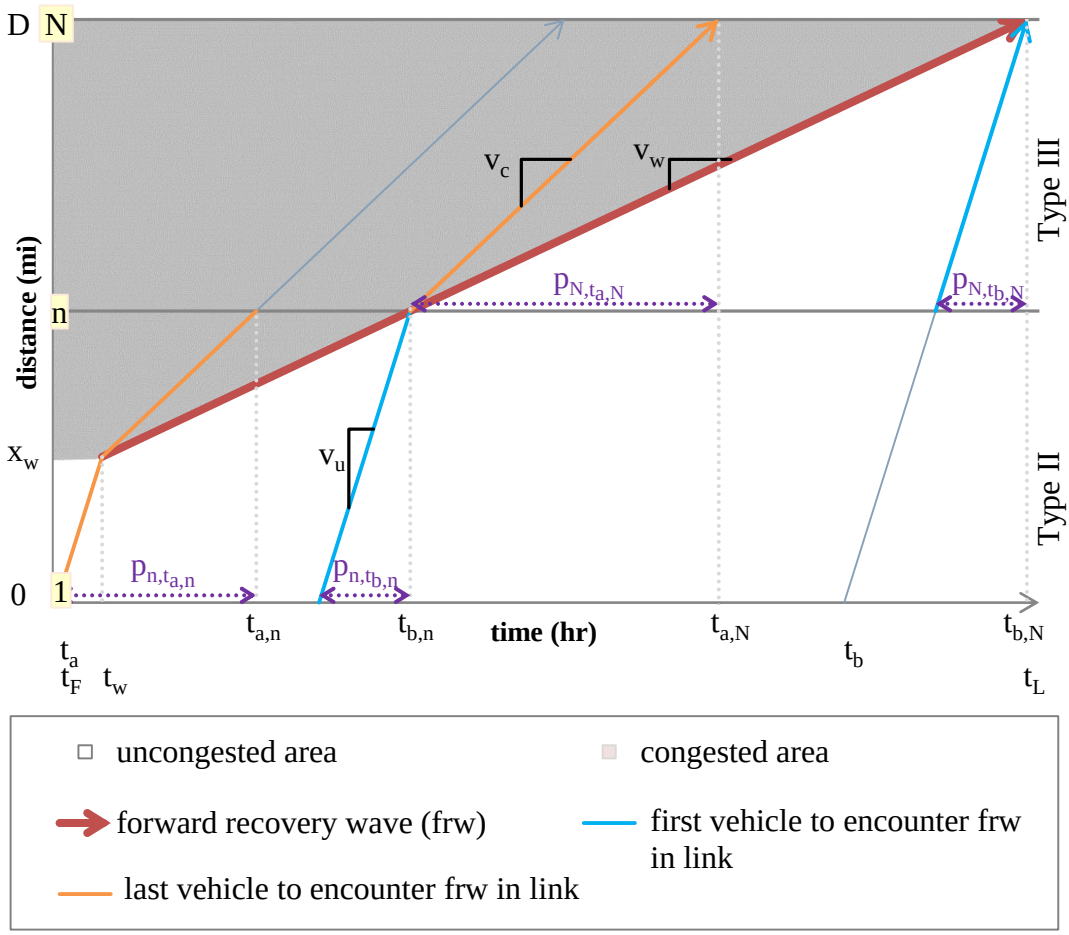


Figure 3.11 Freeway with links classified as Type II and Type III.

The last six link types only apply to links during Scenarios with two shock waves, i.e., Scenarios Four, Five, and Six. These categories consider if the waves originate within the link, terminate within the link, and/or whether vehicles can encounter both waves while traversing the link. Each one of the combinations of these characteristics is covered by one of the link types. The combinations take into account weather the waves end within a link, and then at least some of the vehicles encounter both waves while traversing the link. However, vehicles can encounter both waves even if the waves do not meet within a link.

The three characteristics considered for Categories IV through IX are provided in detail. First, the waves originate within a link if $x_{n-1} < x_w < x_n$. If this is true, the link is classified as Type IV, V, or VI. On the other hand, if the wave originates downstream from the link, the link is classified as Type VII, VIII, or IX. Second, the waves terminate within a link only during Scenario Six and if $x_{n-1} \leq x_I < x_n$. For this case, the link is classified as either Type VI or IX. Third, vehicles can encounter both waves while traversing the link only during Scenarios Five and Six and if the inequality in Eq. 3.30 holds true. If it is true, then the link is classified as Type V, VI, VIII, or IX. If it is not true, then the link is classified as Type IV or VII.

$$t_{w,s} - \frac{x_w - x_{n-1}}{v_{w,s}} + \frac{S}{v_c} > t_{w,r} - \frac{x_w - x_n}{v_{w,r}} \tag{3.30}$$

In the above inequality, the left side represents the amount of time it takes for the backward shock wave to reach the beginning of the link (mile x_{n-1}) and a vehicle to pass through the link while it is congested. The right side of the inequality represents the amount of time it takes the backward recovery wave to reach the link (mile x_n). Therefore, the inequality represents whether the backward shock wave can reach the beginning of the link and a vehicle can pass through the entire congested link before the backward recovery wave enters the link.

Figure 3.12 illustrates a freeway segment with *N* sensors and *N*-1 links. A backward shock wave and backward recovery wave originate within the downstream link, located between sensors *n* and *N*. Vehicles traveling within this link encounter at most one of the shock waves. Therefore, the link is classified as Type IV. The waves

pass through the entire upstream link located between sensors 1 and *n*. In addition, vehicles encounter at most one shock wave within the link. Therefore, the link is classified as Type VII. Although no vehicles encounter both waves within a single link, some vehicles do encounter the backward shock wave on the upstream link and then encounter the backward recovery wave on the downstream link. Therefore, in the figure a Scenario Five study period is depicted.

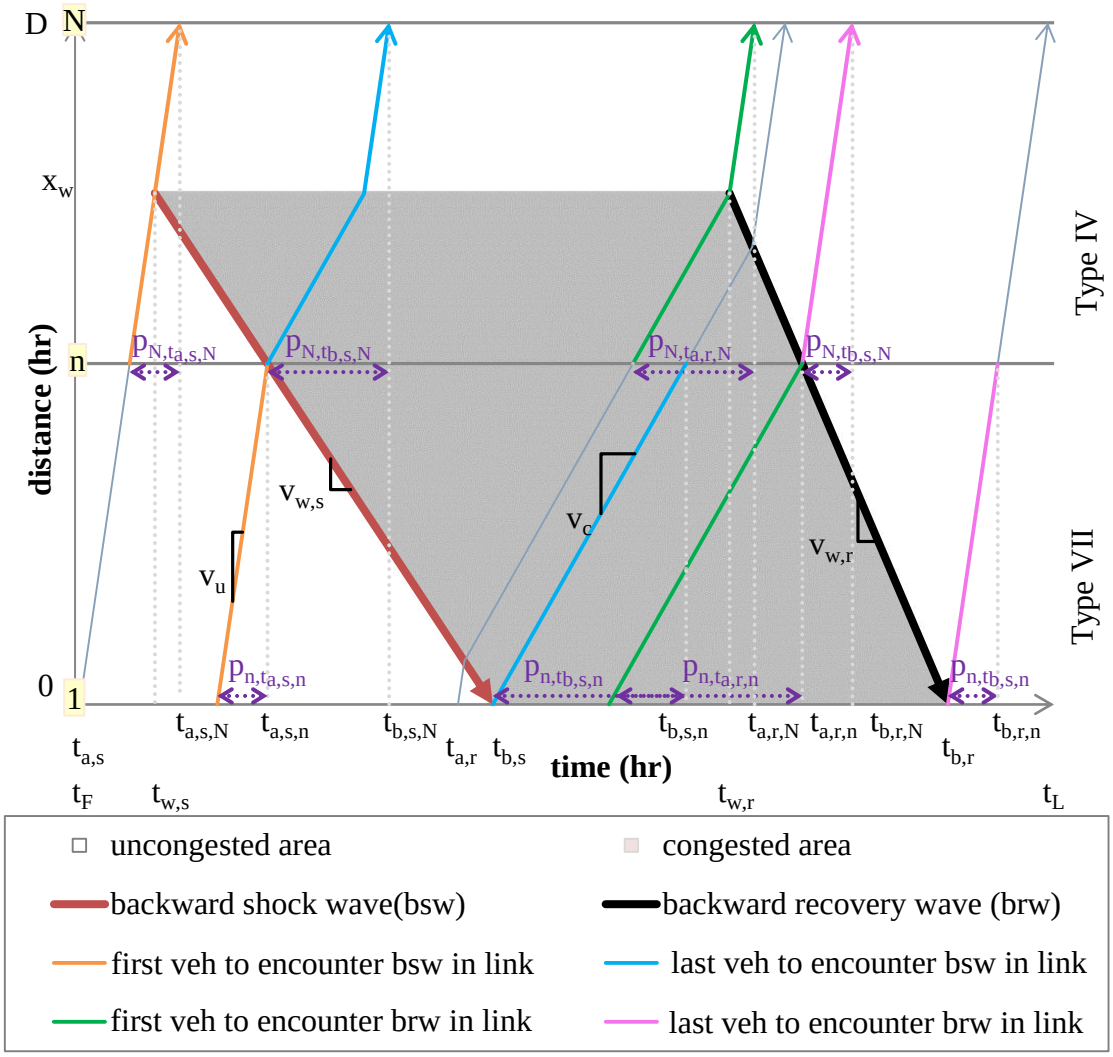


Figure 3.12 Freeway with links classified as Type IV and Type VII.

Based on the aforementioned assumptions, the relationship between predicted travel time and vehicle departure time is modeled for each Scenario. This relationship is developed by applying the following four steps which are listed below and illustrated in the flow chart in Figure 3.13.

1. Calculate the location of each of the *N* sensors.

2. Classify the link type located between each pair of consecutively located sensors.

3. For the link located between sensors *n*-1 and *n* (*n*=2 to *N*), calculate the function of predicted travel time by departure time, denoted $g_n(t)$.

4. Sum all the links' functions to determine the relationship between predicted travel time and departure time for the entire segment length (i.e., mile *0* to mile *D*), denoted $g(t)$.

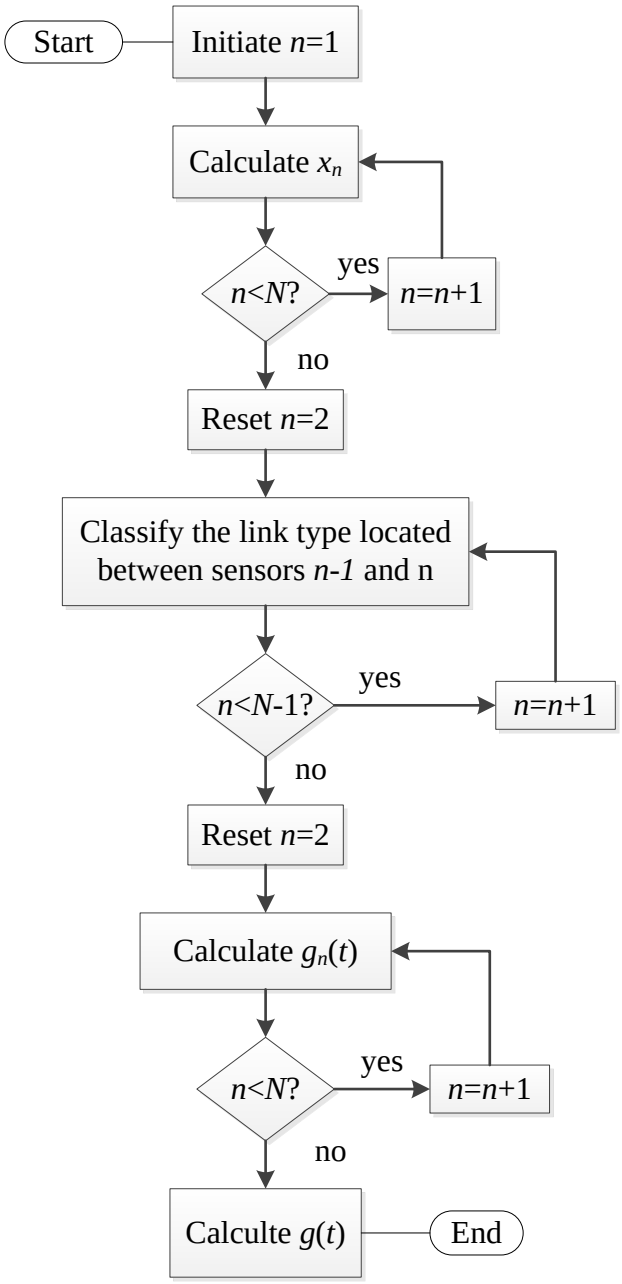


Figure 3.13 The steps to develop the function of predicted travel time by departure time.

The relationship between predicted travel time and departure time is constant for the duration of the study period for freeway links that are unaffected by the shock waves, such as Type I links. The traffic state of these links remains uncongested for the entire

study period. Therefore, the predicted travel time is calculated by dividing the link length, *S*, by the speed of the uncongested traffic state as shown in Eq. 3.31.

$$p_{n,t} = \frac{S}{v_u} \tag{3.31}$$

The function of predicted travel time by departure time for a link *n*, denoted $g_n(t)$, is therefore constant for Type I links and is represented with Eq 3.32 below.

$$g_n(t) = \frac{S}{v_u} \tag{3.32}$$

where $t_F \leq t \leq t_L$

For links affected by the shock waves, such as Type II-IX links, the relationship between predicted travel time and departure time changes over the study period. The predicted travel times are determined by applying the following three steps. First, the times at which the relationship changes are found. Second, the predicted travel times at these points are calculated. Third, this information is used to develop equations that express the relationship between predicted travel time and departure time for all other departure times within the study period. These steps are explained in detail as follows. The Scenarios that consider one shock wave, Scenarios One, Two, and Three are first considered. Then, the steps are applied to the Scenarios that consider two shock waves, Scenarios Four, Five, and Six.

3.4.1 Scenarios One, Two, and Three

In the first step to determine predicted travel times, the departure times at which the predicted travel times begin to increase or decrease are found along with the times they cease to increase or decrease. In Scenarios One through Three, these are the times at which a link's downstream sensor, i.e., sensor n, detects the first and last vehicles that encountered the shock wave within the link. These detection times are denoted $t_{a,n}$ and $t_{b,n}$, respectively. During the passage of a backward shock wave, $t_{a,n}$ represents when predicted travel times begin to increase and $t_{b,n}$ represents when predicted travel times cease to increase and instead become constant. In contrast, during the passage of a recovery wave, $t_{a,n}$ represents when predicted travel times begin to decrease and $t_{b,n}$ represents when predicted travel times cease to decrease and instead become constant.

The detection times and trajectories of the first and last vehicles to encounter a forward recovery wave within a link are illustrated in Figure 3.11 provided previously. Within each link, the trajectory of the first vehicle that encounters the recovery wave is illustrated in orange and the trajectory of the last vehicle that encounters the shock wave is illustrated in blue. The times the vehicles are detected by the downstream sensor of each link are also denoted. At these times, the relationship between predicted travel time and departure time changes.

The time at which a link's downstream sensor detects the first vehicle which encountered the shock wave within the link is calculated. Scenarios One, Two, and Three are considered. There are two considerations for the calculation. The first is determining where and when the vehicle encounters the shock wave within the link. The second is determining the vehicle's travel time from this location to mile x_n. For a Type II

link, the first vehicle to encounter the wave does so at the instant and location the wave forms within the link. For a Type III link, the first vehicle to encounter the wave does so the instant the wave enters the link; the wave enters the link at mile x_n if the wave is a backward shock wave or backward recovery wave, or it enters at mile x_{n-1} if the wave is a forward recovery wave. Therefore, the time the vehicle is detected by sensor n is calculated with the following equations, listed by Scenario and link type:

- Scenarios One and Two
    - Type II link

$$t_{a,n} = t_w + \frac{x_n - x_w}{v_u} \tag{3.33}$$

where $x_{n-1} < x_w < x_n$

    - Type III link

$$t_{a,n} = t_w - \frac{x_w - x_n}{v_w} \tag{3.34}$$

where $x_w \geq x_n$

- Scenario Three
    - Type II link

$$t_{a,n} = t_w + \frac{x_n - x_w}{v_c} \tag{3.35}$$

where $x_{n-1} < x_w < x_n$

    - Type III link

$$t_{a,n} = t_w + \frac{x_{n-1} - x_w}{v_w} + \frac{S}{v_c} \tag{3.36}$$

where $x_w \leq x_{n-1}$

As can be seen in Eqs. 3.33 and 3.35, for a link in which either a backward shock wave, backward recovery wave, or forward recovery wave initiates, the detection time is dependent on when and where the shock wave initiates within the link and on the amount of time it takes a vehicle to travel from this location to sensor *n*'s location at mile x_n. For example, in Figure 3.11, the first vehicle to encounter the forward recovery wave in the link between sensors 1 and *n* travels at the uncongested speed from mile 0 to mile x_w and then from mile x_w to mile x_n at the congested speed. In Eq. 3.34, for a link in which a backward moving wave initiates downstream from it, the detection time equals the time at which the backward moving wave reaches sensor *n*. In Eq. 3.36, *S* is found with Eq. 3.27. Also, in Eq. 3.36, for a link that has a forward recovery wave initiate upstream from it, the detection time equals the time at which the wave reaches sensor *n*-1 plus the amount of time it takes for a vehicle to travel from sensor *n*-1 to sensor *n* at the congested speed. For example, for the first vehicle to encounter the forward recovery wave in the link between sensors *n* and *N* in Figure 3.11, the vehicle encounters the wave at mile x_n and then travels at the congested speed until mile x_N.

Next, the time at which a link's downstream sensor detects the last vehicle which encountered the shock wave within the link is calculated. The last vehicle to encounter the shock wave does so at the instant and location the shock wave departs from the link. This location is mile x_{n-1} if the shock wave is a backward moving wave or this location is mile x_n if the shock wave is a forward moving wave. If the location is mile x_{n-1}, the then

vehicle's travel time from this location to mile x_n needs to be calculated. Therefore, the time the vehicle is detected by sensor *n* is calculated with one of the following equations, listed by Scenario and link type:

- Scenario One
    - Type II link

$$t_{b,n} = t_{a,n} - \frac{x_w - x_{n-1}}{v_w} + \frac{x_w - x_{n-1}}{v_c} \tag{3.37}$$

where $x_{n-1} < x_w < x_n$

    - Type III link

$$t_{b,n} = t_w - \frac{x_w - x_{n-1}}{v_w} + \frac{S}{v_c} \tag{3.38}$$

where $x_w \geq x_n$

- Scenario Two
    - Type II or III link

$$t_{b,n} = t_w - \frac{x_w - x_{n-1}}{v_w} + \frac{S}{v_u} \tag{3.39}$$

- Scenario Three
    - Type II or III link

$$t_{b,n} = t_w + \frac{x_n - x_w}{v_w} \tag{3.40}$$

The equations for a backward recovery wave, Eq. 3.37 and Eq. 3.38, take into account whether part of the link remains in uncongested conditions or if the entire link is in congested conditions after the wave passes through, respectively. The detection time by sensor *n* of the last vehicle to encounter a backward recovery wave for both Categories II and III is calculated with the same equation, Eq. 3.39, because in both instances the vehicle travels from mile x_{n-1} to mile x_n at the uncongested speed. Also, the detection time by sensor *n* of the last vehicle to encounter a forward recovery wave for both Categories II and III is calculated with the same equation, Eq. 3.40, because in both instances the vehicle is detected at the time the wave exits the link. For example, in Figure 3.11, both $t_{b,n}$ and $t_{b,N}$ occur when the forward recovery wave reaches the downstream end of each link.

The predicted travel times derived from the first and last vehicles to encounter the shock waves within a link are calculated. As a reminder, Scenarios One, Two, and Three are still being considered.

The predicted travel time derived from a vehicle that travels the entire link at the uncongested speed is first discussed. The entirety of a link is in uncongested conditions before a backward shock wave enters it and after a recovery wave departs from it. Therefore, the travel times of the following vehicles are calculated with this equation:

- The first vehicle to encounter the backward shock wave for a Type II or III link during Scenario One
- The last vehicle to encounter the backward recovery wave for a Type II or III link during Scenario Two
- The last vehicle to encounter the forward recovery wave for a Type II or III link during Scenario Three

The predicted travel time is calculated by dividing the link length, *S*, by the speed of the uncongested traffic state as in Eq. 3.31 provided previously. This equation is illustrated in Figure 3.11 with the predicted travel time derived from the vehicle detected at time $t_{b,N}$. The vehicle's travel time from mile x_n to mile x_N is denoted in the figure as $p_{N,t_{b,N}}$.

Next, the predicted travel time derived from a vehicle that travels the entire link at the congested speed is determined. The entirety of a link is in congested conditions after a backward shock wave exits it and before a recovery wave enters it. Therefore, the following vehicles' travel times are calculated with this equation:

- The last vehicle to encounter the backward shock wave for a Type III link during Scenario One
- The first vehicle to encounter the backward recovery wave for a Type III link during Scenario Two
- The first vehicle to encounter the forward recovery wave for a Type III link during Scenario Three

The predicted travel time is calculated by dividing the link length, *S*, by the speed of the congested traffic state as in Eq. 3.41.

$$p_{n,t} = \frac{S}{v_c} \tag{3.41}$$

This equation is also illustrated in Figure 3.11 with the predicted travel time derived from the vehicle detected at time $t_{a,N}$. The vehicle travels from mile x_n to mile x_N at the congested speed; the travel time from the one location to the other is denoted in the figure as $p_{N,t_{a,N}}$.

Third, travel time is predicted when a forward recovery wave originates within the link. This particular equation, Eq. 3.42, is used to calculate the predicted travel time derived from a vehicle that travels from the beginning of the link to the location the shock wave originates at the uncongested speed, and then continues to the end of the link at the congested speed. Therefore, the following vehicle's travel times is calculated with this equation:

- The first vehicle to encounter the forward recovery wave for a Type II link during Scenario Three

The predicted travel time is calculated with Eq. 3.42.

$$p_{n,t} = \frac{x_w - x_{n-1}}{v_u} + \frac{x_n - x_w}{v_c} \tag{3.42}$$

where $x_{n-1} < x_w < x_n$

The first term on the right hand side of the equations represents the travel time at the uncongested speed and the second term represents the travel time at the congested speed. In Figure 3.11, the predicted travel time derived from the first vehicle to encounter the forward recovery wave within the link between sensors 1 and *n*, denoted $p_{n,t_{a,n}}$, is calculated with this equation.

In contrast to Eq. 3.42, Eq. 3.43 is applicable when a backward shock wave or a backward recovery wave originates within the link. It is applied to calculate the predicted travel time derived from a vehicle that travels from the beginning of the link to the location the shock wave originates at the congested speed and then continues to the end of the link at the uncongested speed. Therefore, the following vehicles' travel times are calculated with this equation:

- The last vehicle to encounter the backward shock wave for a Type II link during Scenario One

- The first vehicle to encounter the backward recovery wave for a Type II link during Scenario Two

The predicted travel time is calculated with Eq. 3.43.

$$p_{n,t} = \frac{x_w - x_{n-1}}{v_c} + \frac{x_n - x_w}{v_u} \tag{3.43}$$

where $x_{n-1} < x_w < x_n$

The first term on the right hand side of the equations represents the travel time at the congested speed and the second term represents the travel time at the uncongested speed.

In the final step to develop the function of predicted travel time by departure time, the predicted travel times for the remaining departure times are determined. These departure times are grouped into three sets. The first set includes the departure times from the beginning of the study period until the detection time of the first vehicle to encounter the wave. The second set includes the departure times in between the detection times of the first and last vehicles to encounter the wave. The last set includes the departure times from the detection time of the last vehicle to encounter the wave until the end of the study period.

From the beginning of the study period of a Scenario, which is departure time t_F found with Eq. 3.9, until the detection time of the first vehicle to encounter the shock wave, which is time $t_{a,n}$, the predicted travel times are constant. The predicted travel times during this period are constant because the traffic state is static before the shock

wave enters the link. Therefore, the predicted travel times are calculated with the equation used to find the predicted travel time for departure time $t_{a,n}$.

Next, the relationship is constant from the detection time of the first vehicle to encounter the shock wave until the detection time of the last vehicle to encounter the shock wave, which is departure time $t_{b,n}$. Specifically, in Scenario One, predicted travel times linearly increase because, during this period, the backward shock wave passes through the link. Also, in Scenarios Two and Three, predicted travel times linearly decrease, because, during this period, the recovery wave passes through the link. A linear equation is developed to calculate the predicted travel time for link n for a vehicle that departs during this period. The independent variables of the equation include the departure times $t_{a,n}$ and $t_{b,n}$ and their associated predicted travel times $p_{n,t_{a,n}}$ and $p_{n,t_{b,n}}$. The predicted travel time on a link n for a vehicle that departs at time t, where t is between $t_{a,n}$ and $t_{b,n}$, is found with Eq. 3.44.

$$p_{n,t} = \frac{p_{n,t_{a,n}} - p_{n,t_{b,n}}}{t_{a,n} - t_{b,n}}\left(t - t_{a,n}\right) + p_{n,t_{a,n}} \tag{3.44}$$

where $t_{a,n} \leq t \leq t_{b,n}$

The equation is of the form $y = mx + b$ where the slope represents the change in predicted travel time over the change in departure time.

For the remainder of the study period, from the detection time of the last vehicle to encounter the shock wave until the end of the study period, time t_L found with Eq. 3.10, the predicted travel times are constant. This is because the traffic state is static after the shock wave exits the link. The predicted travel times during this period are the same

as the predicted travel time derived from the vehicle that is detected at time $t_{b,n}$. The equation to calculate $p_{n,t_{b,n}}$ has been discussed previously.

The predicted travel times by departure time are summarized with the following function shown with Eq. 3.45. The function, denoted g_n, assigns to each departure time t, where $t_F \leq t \leq t_L$, a predicted travel time for a link that is located between sensors n-1 and n and is classified as either Type II or III.

$$g_n(t) = \begin{cases} p_{n,t_{a,n}} & , \quad t_F \leq t \leq t_{a,n} \\ \dfrac{p_{n,t_{a,n}} - p_{n,t_{b,n}}}{t_{a,n} - t_{b,n}}\left(t - t_{a,n}\right) + p_{n,t_{a,n}}, & t_{a,n} < t < t_{b,n} \\ p_{n,t_{b,n}} & , \quad t_{b,n} \leq t \leq t_L \end{cases} \tag{3.45}$$

The relationship between predicted travel time and departure time for the case study with a forward recovery wave illustrated in Figure 3.11 is shown in Figure 3.14. For each link, the function is constant until the first vehicle to encounter the forward recovery wave is detected by the downstream sensor. This is time $t_{a,n}$ for the upstream link of the segment and time $t_{a,N}$ for the downstream link of the segment. Then, the function decreases as the predicted travel times are derived from vehicles that encounter the forward recovery wave within each link. The function is again constant when the first vehicle to travel the entire link at the uncongested speed is detected by the downstream sensor on the link. This is time $t_{b,n}$ for the upstream link of the segment and time $t_{b,N}$ for the downstream link of the segment. The maximum predicted travel time for the link between sensors n and N is greater than the maximum for the link between sensors 1 and n because part of the link between sensors 1 and n always remains uncongested.

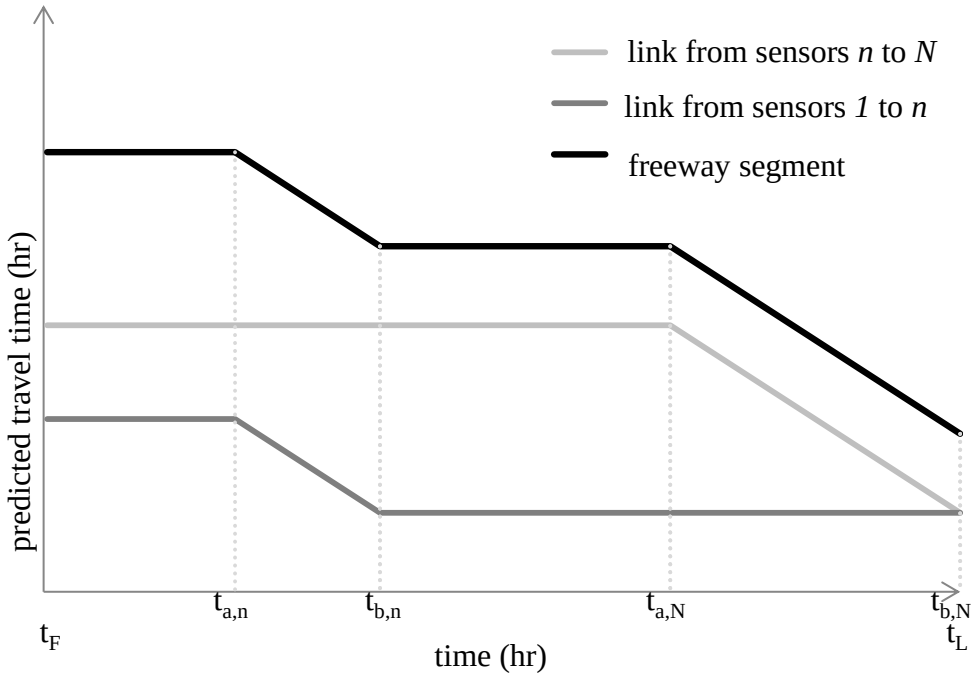


Figure 3.14 Predicted travel times for study period illustrated in Figure 3.11.

The predicted travel times for a link for the duration of a study period for Scenarios One, Two, and Three has been modeled. The detection times of the first and last vehicles to encounter the shock wave were calculated along with these vehicles' travel times on the link. From this information, the predicted travel times for the entire duration of the study period were found.

3.4.2 Scenarios Four, Five, and Six

Next, the relationship between predicted travel time and departure time is determined for Scenarios Four, Five and Six. There is both a backward shock wave and backward recovery wave in each of these Scenarios. In the first step, the departure times at which the predicted travel times begin to increase or decrease are found along with the times they cease to increase or decrease. The relationship between predicted travel time and

departure time changes when a link's downstream sensor, i.e., sensor *n*, detects the first and last vehicles that encountered the backward shock wave within the link. These times are denoted $t_{a,s,n}$ and $t_{b,s,n}$, respectively. In addition, the relationship changes when a link's downstream sensor detects the first and last vehicles that encounter the backward recovery wave within the link. These times are denoted $t_{a,r,n}$ and $t_{b,r,n}$, respectively.

The detection times and trajectories of the first and last vehicles to encounter a backward shock wave and backward recovery wave within a link are illustrated in Figure 3.12. The trajectory of the first vehicle that encounters the backward shock wave is illustrated in orange and the trajectory of the last vehicle that encounters the backward shock wave is illustrated in purple. In addition, the trajectory of the first vehicle that encounters the backward recovery wave is illustrated in green and the trajectory of the last vehicle that encounters the backward recovery wave is illustrated in pink.

The time at which a link's downstream sensor detects the first vehicle which encountered the backward shock wave within the link is calculated. For a link classified as Type IV, V, or VI, the first vehicle to encounter the backward shock wave does so at the instant and location the wave forms within the link, time $t_{w,s}$ and mile x_w. For a link classified as Type VII, VIII, or IX, the first vehicle to encounter the wave does so the instant the wave enters the link at mile x_n. The detection time by sensor *n* of the first vehicle to encounter a backward shock wave is calculated with one of the following equations:

- Scenarios Four, Five and Six
    - Type IV link or a Type VI link

$$t_{a,s,n} = t_{w,s} + \frac{x_n - x_w}{v_u} \tag{3.46}$$

where $x_{n-1} < x_w < x_n$

- Scenarios Four, Five, and Six
    - Type VII link or a Type VIII link

$$t_{a,s,n} = t_{w,s} - \frac{x_w - x_n}{v_{w,s}} \tag{3.47}$$

where $x_w \geq x_n$

As can be seen in Eq. 3.46, for links in which the backward shock wave initiates, the detection time is dependent on when and where the backward shock wave initiates within the link and on the amount of time it takes a vehicle to travel from this location to sensor *n*'s location at mile x_n. For example, in Figure 3.12, the first vehicle to encounter the backward shock wave in the link located between sensors *n* and *N* travels at the uncongested speed from mile x_w to mile *D*. Also, as can be seen in Eq. 3.47, for links in which the backward shock wave initiates downstream, the detection time equals the time at which the backward shock wave reaches sensor *n*. This is apparent in Figure 3.12 for the link located between sensors 1 and *n*.

Next, the time at which a link's downstream sensor detects the last vehicle which encountered the backward shock wave within the link is calculated. For Scenarios Four and Five, the vehicle encounters the wave at the instant and location the shock wave departs from the link. This location is mile x_{n-1}. For Scenario Six, the vehicle encounters the wave at the instant and location the shock wave terminates within the link. This

location is mile x_I. The detection time by sensor n of the last vehicle to encounter a backward shock wave is calculated with one of the following equations, provided by link type:

- Scenarios Four, Five, and Six
    - Type IV link

$$t_{b,s,n} = t_{w,s} - \frac{x_w - x_{n-1}}{v_{w,s}} + \frac{x_w - x_{n-1}}{v_c} + \frac{x_n - x_w}{v_u} \tag{3.48}$$

where $x_{n-1} < x_w < x_n$

    - Type V or VIII link

$$t_{b,s,n} = \frac{x_{n-1} - x_w + t_{w,r} v_{w,r}}{v_{w,r} - v_c}\left(1 - \frac{v_c}{v_u}\right) + \left(t_{w,s} - \frac{x_w - x_{n-1}}{v_{w,s}}\right)\left(\frac{v_c\left(v_{w,r} - v_u\right)}{v_u\left(v_{w,r} - v_c\right)}\right) + \frac{S}{v_u}$$

where $x_w > x_{n-1}$ (3.49)

    - Type VII link

$$t_{b,s,n} = t_{w,s} - \frac{x_w - x_{n-1}}{v_{w,s}} + \frac{S}{v_c} \tag{3.50}$$

where $x_w \geq x_n$

    - Type VI or IX link

$$t_{b,n} = t_I + \frac{x_n - x_I}{v_u} \tag{3.51}$$

where $x_{n-1} < x_w < x_n$

In Eq. 3.48, for links in which the backward shock wave initiates and no vehicles encounter both shock waves, the detection time equals the time at which the backward shock wave exits the link plus the travel time of the vehicle that enters the link at this time. As can be seen in Figure 3.12 for the link between sensors n and N, this vehicle travels at the congested speed from the location of sensor n-1 to the location the wave initiates and then at the uncongested speed for the remainder of the link. For links in which the vehicles encounter both waves but the waves do not terminate within the link, Eq. 3.49 takes into account that the vehicle from which the predicted travel time is derived encounters the backward recovery wave within the link. This vehicle travels at the congested speed from the location of sensor n-1 to the location it encounters the backward recovery wave and then travels at the uncongested speed for the remainder of the link. In Eq. 3.50, for links in which the backward shock wave initiates downstream and no vehicles encounter both shock waves, the detection time equals the time at which the backward shock wave exits the link plus the amount of time it takes a vehicle to travel from the location of sensor n-1 to the location of sensor n at the congested speed. For example, in Figure 3.12, the vehicle with detection time $t_{b,s,n}$ encounters the backward shock wave at mile zero and then travels at the congested speed until mile x_n. In Eq. 3.51, for links in which the waves terminate, the time and location the backward shock wave and backward recovery wave intersect, t_I and x_I, are applied to determine when the vehicle that is at mile x_I at time t_I then reaches sensor n. In this equation, the detection time is denoted $t_{b,n}$ to represent that one vehicle is the last to encounter both the backward shock wave and backward recovery wave.

Next, the time at which a link's downstream sensor detects the first vehicle which encountered the backward recovery wave within the link is calculated. For a link classified as Type IV, V, or VI, the first vehicle to encounter the wave does so at the instant and location the wave forms within the link. For a link classified as Type VII, VIII, or IX, the first vehicle to encounter the wave does so the instant the wave enters the link at mile x_n. The detection time by sensor n of the first vehicle to encounter a backward recovery wave during Scenarios Four, Five, or Six is calculated with one of the following equations:

- Scenarios Four, Five, and Six
    - Type IV link, V link, or a Type VI link

$$t_{a,r,n} = t_{w,r} + \frac{x_n - x_w}{v_u} \tag{3.52}$$

where $x_{n-1} < x_w < x_n$

    - Type VII link, Type VIII link, or a Type IX link

$$t_{a,r,n} = t_{w,r} - \frac{x_w - x_n}{v_{w,s}} \tag{3.53}$$

where $x_w \geq x_n$

As can be seen in Eq. 3.52, for links in which the backward recovery wave initiates, the detection time is dependent on when and where the backward recovery wave initiates within the link and on the amount of time it takes a vehicle to travel from this location to sensor n's location at mile x_n. For example, in Figure 3.12, the first vehicle to encounter

the backward recovery wave in the link located between sensors *n* and *N* travels at the uncongested speed from mile x_w to mile *D*. Also, as can be seen in Eq. 3.53, for links in which the backward recovery wave initiates downstream, the detection time equals the time at which the backward recovery wave reaches sensor *n*. This is illustrated in Figure 3.12. The time at which the first vehicle to encounter the backward recovery wave in the link between sensors one and *n* is detected, time $t_{a,r,n}$, occurs when the backward recovery wave reaches mile x_n.

Last, the time at which a link's downstream sensor detects the last vehicle which encountered the backward recovery wave within the link is calculated. The detection time of the last vehicle to encounter the backward recovery wave is calculated. For Scenarios Four and Five, the last vehicle to encounter the backward recovery wave does so at the instant and location the shock wave departs from the link. This location is mile x_{n-1}. The detection time by sensor *n* of the last vehicle to encounter the backward recovery wave is calculated with the following equation:

- Eq. 3.54 for a Type IV or VII link during Scenario Four or Five and a Type V or VIII link during Scenario Five

$$t_{b,r,n} = t_{w,r} - \frac{x_w - x_{n-1}}{v_{w,r}} + \frac{S}{v_u} \tag{3.54}$$

In Eq. 3.54, the detection time equals the time at which the backward recovery wave exits the link plus the amount of time it takes a vehicle to travel from the location of sensor *n*-1 to the location of sensor *n* at the uncongested speed. In Figure 3.12, one can see in both links that the last vehicle to encounter the backward recovery wave travels the entire link length at the uncongested speed. In Scenario Six, there is one vehicle that is the last to

encounter both the backward shock wave and backward recovery wave. This vehicle's detection time was previously calculated with Eq. 3.51.

The predicted travel times derived from the first and last vehicles to encounter the backward shock wave and backward recovery wave within the link are calculated. First, the predicted travel time derived from a vehicle that travels the entire link at the uncongested speed is discussed. A link is in uncongested conditions before a backward shock wave enters it and after a recovery wave departs from it. Therefore, the following vehicles' travel times are calculated with this equation:

- The first vehicle to encounter the backward shock wave for link Categories IV-IX in Scenarios Four, Five, and Six
- The last vehicle to encounter the backward recovery wave for link Categories IV, V VII, and VIII in Scenarios Four, Five, and Six
- The last vehicle to encounter both waves for link Categories VI and IX in Scenario Six

The predicted travel time is calculated by dividing the link length, *S*, by the speed of the uncongested traffic state as shown previously with Eq. 3.31. This equation is illustrated in Figure 3.12 with the predicted travel time derived from the vehicles detected at times $t_{a,s,N}$, $t_{a,s,n}$, $t_{b,r,N}$, and $t_{b,r,n}$. The vehicles' travel times on the links are denoted on the x-axis with dotted blue lines.

Next, the predicted travel time derived from a vehicle that travels the entire link at the congested speed is calculated. The shock waves must initiate outside of the link in order for the entire link to be in the congested state. In addition, a vehicle can travel the entire link at the congested speed only if it departs after the backward shock wave exits the link and the vehicle reaches mile x_n before the backward recovery wave enters the

link. These conditions are only possible for a link classified as Type VII. Therefore, the following vehicles' travel times are calculated with this equation:

- The last vehicle to encounter the backward shock wave for a Type VII link during Scenarios Four, Five, and Six
- The first vehicle to encounter the backward recovery wave for a Type VII link during Scenarios Four, Five, and Six

The predicted travel time is calculated by dividing the link length, *S*, by the speed of the congested traffic state as was shown previously with Eq. 3.41. This equation is illustrated in Figure 3.12 with the predicted travel time derived from the vehicles detected at times $t_{b,s,n}$, and $t_{a,r,n}$. Both of these vehicles travel between sensors 1 and *n* at the congested speed.

The next calculation is applicable when the shock waves initiate within the link. It determines the predicted travel time derived from a vehicle that travels from the beginning of the link to the location the shock wave originates at the congested speed and then continues to the end of the link at the uncongested speed. Therefore, the following vehicles' travel times are calculated with this equation:

- The last vehicle to encounter the backward shock wave for a Type IV link during Scenarios Four, Five, and Six.
- The first vehicle to encounter the backward recovery wave for a Type IV link during Scenarios Four, Five, and Six.

The predicted travel time is calculated as was shown previously with Eq. 3.43. In Figure 3.12, the predicted travel times $p_{N,t_{b,s,N}}$ and $p_{N,t_{a,r,N}}$ derived from the vehicles detected at times $t_{b,s,N}$ and $t_{a,r,N}$, respectively, are calculated as such.

The next equation, Eq. 3.55, is used to calculate the predicted travel time derived from a vehicle that departs at the time the backward shock wave exits the link and encounters the backward recovery within the link. It is therefore the predicted travel time derived from the following vehicle:

- The last vehicle to encounter the backward shock wave for a Type V or VIII link during Scenarios Four, Five, and Six.

The equation to calculate the predicted travel time, Eq. 3.55, employs the detection time of the vehicle, $t_{b,s,n}$, already found with Eq. 3.49.

$$p_{n,t_{b,s,n}} = t_{b,s,n} - \left(t_{w,s} - \frac{x_w - x_{n-1}}{v_{w,s}} \right) \tag{3.55}$$

where $x_w > x_{n-1}$

In Eq. 3.55, the predicted travel time is found by subtracting the vehicle's departure time into the link, which is the time the backward shock wave exits the link, from the time the vehicle exits the link, $t_{b,s,n}$.

Next, Eq. 3.56 is used to calculate the predicted travel time derived from a vehicle that is the first to encounter the backward recovery wave within a link classified as Type V or VI. The equation to calculate the predicted travel time of this vehicle needs to take into account that the waves initiate within the link and the vehicle encounters the backward shock wave within the link. The following vehicles' travel times are calculated with this equation:

- The first vehicle to encounter the backward recovery wave for a Type V or VI link during Scenarios Five and Six.

The equation to calculate the predicted travel time, Eq. 3.56, employs the detection time of the vehicle, $t_{a,r,n}$, already found with Eq. 3.52.

$$p_{n,t_{a,r,n}} = t_{a,r,n} - \left[\left(\frac{-t_{w,s}v_{w,s} + t_{w,r}v_c}{v_c - v_{w,s}}\right)\left(1 - \frac{v_c}{v_u}\right) - \frac{x_w - t_{w,r}v_c - x_{n-1}}{v_u}\right] \quad (3.56)$$

where $x_{n-1} < x_w < x_n$

In Eq. 3.56, the first term on the right hand side is the time the vehicle exits the link. The term within the brackets is the time the vehicle enters the link. This term takes into account that the vehicle encounters the backward shock wave within the link. The predicted travel time is therefore found by subtracting the vehicle's departure time into the link from its exit time from the link.

Whereas Eq. 3.56 is used to calculate the predicted travel time derived from a vehicle that is the first to encounter the backward recovery wave within a link classified as Type V or VI, Eq. 3.57 is used for a link classified as Type VIII or IX. Therefore, the equation needs to take into account that the waves initiate downstream from the link and the vehicle encounters the backward shock wave within the link. The following vehicles' travel times are calculated with this equation:

- The first vehicle to encounter the backward recovery wave for a Type VIII or IX link during Scenarios Five and Six.

The equation to calculate the predicted travel time, Eq. 3.57, employs the detection time of the first vehicles to encounter the backward shock wave and backward recovery wave within the link because they represent the times at which the shock waves enter the link. These variables $t_{a,s,n}$ and $t_{a,r,n}$, were previously calculated with Eqs. 3.47 and 3.53, respectively.

$$p_{n,t_{a,r,n}} = t_{a,r,n} - \left[\left(\frac{-t_{a,s,n}v_{w,s} + t_{a,r,n}v_c}{v_c - v_{w,s}}\right)\left(1 - \frac{v_c}{v_u}\right) - \left(\frac{x_n - t_{a,r,n}v_c - x_{n-1}}{v_u}\right)\right]$$

where $x_w \geq x_n$ (3.57)

The equation is similar to Eq. 3.56 in that the predicted travel time is found by subtracting the vehicle's departure time into the link from its exit time from the link.

Next, the last step to develop the function of predicted travel time by departure time is applied. The relationship between predicted travel time and departure time changes at the detection times by sensor *n* of the first and last vehicles to encounter the shock waves. Therefore, the function is constant in between these detection times. The equation that expresses the function is developed. The chronological order of the detection times of the first and last vehicles to encounter the shock waves are the same for link Categories IV and VII, for Categories V and VIII, and for Categories VI and IX and so the development of the function is discussed for each of these Condition groups.

The third step is first applied to link Categories IV and VII. The relationship between predicted travel time and departure time is constant between the departure times corresponding to the following events:

1. The beginning of the study period (time t_F found with Eq. 3.21) and the detection time of the first vehicle to encounter the backward shock wave.
2. The detection times of the first and last vehicles to encounter the backward shock wave.
3. The detection times of the last vehicle to encounter the backward shock wave and the first vehicle to encounter the backward recovery wave.
4. The detection times of the first and last vehicles to encounter the backward recovery wave.

5. The detection time of the last vehicle to encounter the backward recovery wave and the end of the study period (time t_L found with Eq. 3.22).

The predicted travel times between the first set of departure times, t_F and $t_{a,s,n}$, are constant because the traffic state is static before the backward shock wave enters the link. Therefore, the predicted travel times are calculated with the equation used to find the predicted travel time for departure time $t_{a,s,n}$, which is Eq. 3.46 in link Type IV and Eq. 3.47 in link Type VII.

Next, predicted travel times linearly increase between the second set of departure times, $t_{a,s,n}$ and $t_{b,s,n}$, as a result of the backward shock wave passing through the link. The predicted travel time on a link n for a vehicle that departs at time t, where t is between $t_{a,s,n}$ and $t_{b,s,n}$, is found with Eq. 3.58, which is a generalized form of Eq. 3.44.

$$p_{n,t} = \frac{p_{n,t_x} - p_{n,t_y}}{t_x - t_y}(t - t_x) + p_{n,t_x} \qquad (3.58)$$

where $t_{a,n} \leq t \leq t_{b,n}$

In the equation, $t_{a,s,n}$ is represented by t_x and $t_{b,s,n}$ is represented by t_y. Thus, p_{n,t_x} equals $p_{n,t_{a,s,n}}$, the predicted travel time derived from the first vehicle to encounter the backward shock wave and p_{n,t_y} equals $p_{n,t_{b,s,n}}$, the predicted travel time derived from the last vehicle to encounter the backward shock wave.

Between the third set of departure times, $t_{b,s,n}$ and $t_{a,r,n}$, the predicted travel times are constant. The predicted travel times during this period are constant because the traffic state is static after the backward shock wave exits the link but before the backward

recovery wave enters the link. Therefore, the predicted travel times are calculated with the equation used to find the predicted travel time for detection time $t_{b,s,n}$, which is Eq. 3.48 in link Type IV and Eq. 3.50 in link Type VII.

For the departure times that fall between $t_{a,r,n}$ and $t_{b,r,n}$, which are the detection times of the first and last vehicles to encounter the backward recovery wave and the fourth set of departure times, predicted travel times linearly decrease. An equation previously developed, Eq. 3.58, is applied to determine the predicted travel time on a link n for a vehicle that departs at time t, where t is between $t_{a,r,n}$ and $t_{b,r,n}$. In the equation, $t_{a,r,n}$ is represented by t_x and $t_{b,r,n}$ is represented by t_y. Thus, p_{n,t_x} equals $p_{n,t_{a,r,n}}$, the predicted travel time derived from the first vehicle to encounter the backward recovery wave and p_{n,t_y} equals $p_{n,t_{b,r,n}}$, the predicted travel time derived from the last vehicle to encounter the backward recovery wave.

For the remainder of the study period, from the detection time of the last vehicle to encounter the backward recovery wave until the end of the study period, the predicted travel times are constant. This is because the traffic state is static after the backward recovery wave exits the link. Therefore, the predicted travel times are calculated with the equation used to find the predicted travel time for departure time $t_{b,r,n}$, which is Eq. 3.31.

The predicted travel times for the study period are summarized with the below function. The function, denoted g_n, assigns to each departure time t, where $t_F \leq t \leq t_L$, a predicted travel time for a link that is located between sensors n-1 and n and is classified as either Type IV or VII.

$$g_n(t) = \begin{cases} p_{n,t_{a,s,n}} & , \quad t_F \le t \le t_{a,s,n} \\ \frac{p_{n,t_{a,s,n}} - p_{n,t_{b,s,n}}}{t_{a,s,n} - t_{b,s,n}}\left(t - t_{a,s,n}\right) + p_{n,t_{a,s,n}}, & t_{a,s,n} < t < t_{b,s,n} \\ p_{n,t_{b,s,n}} & , \quad t_{b,s,n} \le t \le t_{a,r,n} \\ \frac{p_{n,t_{a,r,n}} - p_{n,t_{b,r,n}}}{t_{a,r,n} - t_{b,r,n}}\left(t - t_{a,r,n}\right) + p_{n,t_{a,r,n}}, & t_{a,r,n} < t < t_{b,r,n} \\ p_{n,t_{b,r,n}} & , \quad t_{b,r,n} \le t \le t_L \end{cases} \tag{3.59}$$

The relationship between predicted travel time and departure time for the case study illustrated in Figure 3.12 with links classified as Type IV and VII is shown in Figure 3.15. There are three lines. The light gray line depicts the function for the link located between sensors 1 and n; this link is classified as Type VII. The dark gray line depicts the function for the link location between sensors n and N; this link is classified as Type IV. The thick black line is the sum of the first two lines; it is the predicted travel time vehicles are provided as they enter the segment. For each link, the function is constant until the detection time of the first vehicle to encounter the backward shock wave. Then, it increases between the detection times of the first and last vehicles to encounter the backward shock wave. Next, the function is again constant between the detection times of the last vehicle to encounter the backward shock wave and the first vehicle to encounter the backward recovery wave. After the detection time of the first vehicle to encounter the backward recovery wave, the function decreases until the detection time of the last vehicle to encounter the backward recovery wave. Finally, the function is constant until the end of the study period. The maximum predicted travel time for the

link between sensors 1 and *n* is greater than the maximum for the link between sensors n and *N* because part of the link between sensors *n* and *N* always remains uncongested.

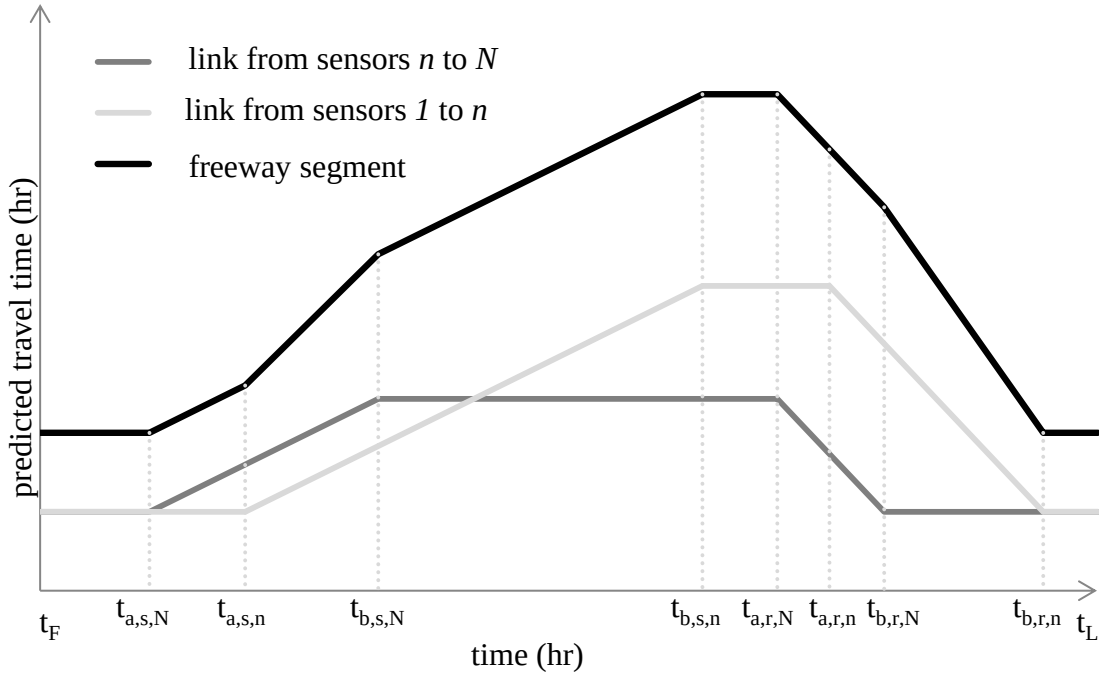


Figure 3.15 Predicted travel times for freeway segment shown in Figure 3.12.

The third step is next applied to link Categories V and VIII. The relationship between predicted travel time and departure time is constant between the departure times corresponding to the following events:

1. The beginning of the study period and the detection time of the first vehicle to encounter the backward shock wave.

2. The detection times of the first vehicle to encounter the backward shock wave and the first vehicle to encounter the backward recovery wave.

3. The detection times of the first vehicle to encounter the backward recovery wave and the last vehicle to encounter the backward shock wave.

4. The detection times of the last vehicle to encounter the backward shock wave and the last vehicle to encounter the backward recovery wave.

5. The detection time of the last vehicle to encounter the backward recovery wave and the end of the study period.

Because the first and fifth sets of departure times are the same as for Type IV and VII links, they will not be discussed again.

Between the second set of departure times, $t_{a,s,n}$ and $t_{a,r,n}$, predicted travel times linearly increase as a result of the backward shock wave passing through the link. An equation previously developed, Eq. 3.58, is applied to determine the predicted travel time on a link n for a vehicle that departs at time t, where t is between $t_{a,s,n}$ and $t_{a,r,n}$. In the equation, $t_{a,s,n}$ is represented by t_x and $t_{a,r,n}$ is represented by t_y. Thus, p_{n,t_x} equals $p_{n,t_{a,s,n}}$, the predicted travel time derived from the first vehicle to encounter the backward shock wave and p_{n,t_y} equals $p_{n,t_{a,r,n}}$, the predicted travel time derived from the first vehicle to encounter the backward recovery wave.

Between the third set of departure times, $t_{a,r,n}$ and $t_{b,s,n}$, predicted travel times linearly decrease as a result of the backward recovery wave entering the link. Equation 3.58 is again applied. It is used to determine the predicted travel time on a link n for a vehicle that departs at time t, where t is between $t_{a,r,n}$ and $t_{b,s,n}$. In the equation, $t_{a,r,n}$ is represented by t_x and $t_{b,s,n}$ is represented by t_y. Thus, p_{n,t_x} equals $p_{n,t_{a,r,n}}$, the predicted travel time derived from the first vehicle to encounter the backward recovery wave and p_{n,t_y} equals $p_{n,t_{b,s,n}}$, the predicted travel time derived from the last vehicle to encounter the backward shock wave.

Predicted travel times linearly decrease at a faster rate between the fourth set of departure times, $t_{b,s,n}$, and $t_{b,r,n}$, then between the third set of departure times. This is a result of only the backward recovery wave being present in the link. Once again Eq. 3.58

is applied to describe the relationship between predicted travel time and departure time. In the equation, $t_{b,s,n}$ is represented by t_x and $t_{b,r,n}$ is represented by t_y. Thus, p_{n,t_x} equals $p_{n,t_{b,s,n}}$, the predicted travel time derived from the last vehicle to encounter the backward shock wave and p_{n,t_y} equals $p_{n,t_{b,r,n}}$, the predicted travel time derived from the last vehicle to encounter the backward recovery wave.

The predicted travel times for the study period are summarized with the function below. The function, denoted g_n, assigns to each departure time t, where $t_F \leq t \leq t_L$, a predicted travel time for a link that is located between sensors n-1 and n and is classified as either Type V or VIII.

$$g_n(t) = \begin{cases} p_{n,t_{a,s,n}} & , \quad t_F \leq t \leq t_{a,s,n} \\ \dfrac{p_{n,t_{a,s,n}} - p_{n,t_{a,r,n}}}{t_{a,s,n} - t_{a,r,n}}\left(t - t_{a,s,n}\right) + p_{n,t_{a,s,n}}, & t_{a,s,n} < t < t_{a,r,n} \\ \dfrac{p_{n,t_{a,r,n}} - p_{n,t_{b,s,n}}}{t_{a,r,n} - t_{b,s,n}}\left(t - t_{a,r,n}\right) + p_{n,t_{a,r,n}}, & t_{a,r,n} \leq t \leq t_{b,s,n} \\ \dfrac{p_{n,t_{b,s,n}} - p_{n,t_{b,r,n}}}{t_{b,s,n} - t_{b,r,n}}\left(t - t_{bs,n}\right) + p_{n,t_{b,s,n}}, & t_{b,s,n} < t < t_{b,r,n} \\ p_{n,t_{b,r,n}} & , \quad t_{b,r,n} \leq t \leq t_L \end{cases} \tag{3.60}$$

Last, the third step is applied to link Categories VI and IX. For these links, one vehicle is the last to encounter both the backward shock wave and backward recovery wave. Therefore, there are only four groups of departure times. The relationship between predicted travel time and departure time is constant between the departure times corresponding to the following events:

1. The beginning of the study period and the detection time of the first vehicle to encounter the backward shock wave.

2. The detection times of the first vehicle to encounter the backward shock wave and the first vehicle to encounter the backward recovery wave.

3. The detection times of the first vehicle to encounter the backward recovery wave and the last vehicle to encounter the shock waves.

4. The detection time of the last vehicle to encounter the shock waves and the end of the study period.

Because the first set of departure times is the same as for Type IV and VII links and the second set of departure times is the same as for Type V and VIII links, they will not be discussed again.

Between the third set of departure times, $t_{ar,n}$ and $t_{b,n}$, predicted travel times linearly decrease as a result of the backward recovery wave entering the link. Equation 3.58 is again applied. It is used to determine the predicted travel time on a link n for a vehicle that departs at time t, where t is between $t_{ar,n}$ and $t_{b,n}$. In the equation, $t_{ar,n}$ is represented by t_x and $t_{b,n}$ is represented by t_y. Thus, p_{n,t_x} equals $p_{n,t_{a,r,n}}$ the predicted travel time derived from the first vehicle to encounter the backward recovery wave and p_{n,t_y} equals $p_{n,t_{b,n}}$, the predicted travel time derived from the last vehicle to encounter the waves.

For the remainder of the study period, from the detection time of the last vehicle to encounter the waves until the end of the study period, the predicted travel times are constant. This is because the traffic state is uncongested after the waves terminate. Therefore, the predicted travel times are calculated with the equation used to find the predicted travel time for departure time $t_{b,n}$, which is Eq. 3.31.

The predicted travel times for the study period are summarized with the function below labeled as Eq. 3.61. The function, denoted g_n, assigns to each departure time t, where $t_F \leq t \leq t_L$, a predicted travel time for a link that is located between sensors n-1 and n and is classified as either Type VI or IX. The function is comprised of four pieces.

$$g_n(t) = \begin{cases} p_{n,t_{a,s,n}} & , \quad t_F \leq t \leq t_{a,s,n} \\ \dfrac{p_{n,t_{a,s,n}} - p_{n,t_{a,r,n}}}{t_{a,s,n} - t_{a,r,n}}\left(t - t_{a,s,n}\right) + p_{n,t_{a,s,n}}, & t_{a,s,n} < t < t_{a,r,n} \\ \dfrac{p_{n,t_{a,r,n}} - p_{n,t_{b,n}}}{t_{a,r,n} - t_{b,n}}\left(t - t_{a,r,n}\right) + p_{n,t_{a,r,n}}, & t_{a,r,n} \leq t \leq t_{b,n} \\ p_{n,t_{b,n}} & , \quad t_{b,n} \leq t \leq t_L \end{cases} \tag{3.61}$$

The function of predicted travel time by departure time was modeled for each link of the freeway segment. The detection times of the first and last vehicles to encounter the shock waves were calculated along with these vehicles' travel times on the link. From this information, the predicted travel times for the entire duration of the study period were found for a link. For a freeway segment, a total of N-1 of these functions is produced. The functions are summed to determine the predicted travel time for the entire length of the segment, denoted $g(t)$, as shown in Eq. 3.62 below.

$$g(t) = \sum_{n=2}^{N} g_n(t) \tag{3.62}$$

This function expresses the predicted travel times provided to vehicles' occupants as their vehicle enters the freeway segment at mile zero.

3.5 Travel Time Prediction Error

In this section, the travel time prediction error for the study period of each Scenario is measured. Scenarios One through Six are discussed together. Prediction error is the difference between the actual travel time, modeled in section 3.2.3, and the predicted travel time, modeled in section 3.2.4. First, the error for a given departure time, *t*, is calculated. Then, the total error for a study period is determined. Finally, the average error for the study period is found. Both the total error and average error take into account the traffic flow during the study period.

The travel time prediction error for a given departure time is calculated. The difference in actual travel time and predicted travel time for departure time *t* is the travel time prediction error, denoted E_t (hr). It is calculated as shown in Eq. 3.63.

$$E_t = P_t - A_t \tag{3.63}$$

For example, in Figure 3.10, the predicted travel time for the vehicle that departs at time t_b is denoted P_{t_b} and is shown with a purple dotted line. Also, the actual travel time for the vehicle that departs at time t_b is denoted A_{t_b} and is shown with a dotted green line. The difference is these times, the prediction error, is illustrated with a dotted red line and is denoted E_{t_b}.

Prediction error can be either positive or negative. A positive error value represents over prediction and a negative value represents under prediction. Over prediction occurs when travel times decrease during a study period. Therefore, there is over prediction during Scenarios Two or Three. For example, Figure 3.16 below displays the prediction error for the study period first illustrated in Figure 3.11. The study period consisted of a freeway segment undergoing a transition from a congested to an uncongested state represented with a forward recovery wave. In the figure, the light gray line depicts vehicles' travel times, the thick black line depicts vehicles' predicted travel times, and the dark gray line depicts the prediction error throughout the study period. The thick black line is the same as that shown in Figure 3.14. Throughout the study period, the predicted travel time function is above the actual travel time function (i.e., the thick black line is higher than the light gray line) which indicates over prediction.

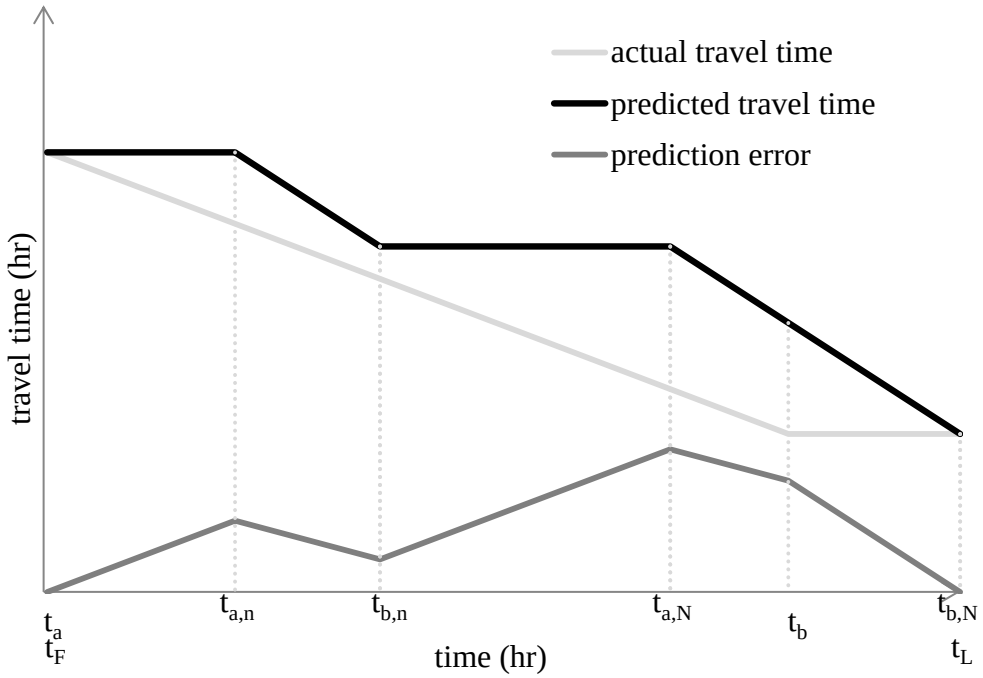


Figure 3.16 Prediction error for the study period illustrated in Figure 3.11.

Under prediction occurs when travel times increase during a study period. Therefore, there is under prediction during Scenario One. For study periods during which travel times initially increase and then later decrease, such as during Scenarios Four, Five, and Six, travel times are under predicted during the beginning of the study period but are over predicted later on in the study period. For example, Figure 3.17 below displays the prediction error for the study period first illustrated in Figure 3.12. The study period consisted of a freeway segment undergoing a transition from an uncongested to a congested state, represented with a backward shock wave, and then undergoing a transition to return to the uncongested state, represented with a backward recovery wave. In the figure, the light gray line depicts vehicles' travel times, the thick black line depicts vehicles' predicted travel times, and the dark gray line depicts the prediction error throughout the study period. The thick black line is the same as that shown in Figure 3.15. At the beginning of the study period, the predicted travel time function is lower than the actual travel time (i.e., the thick black line is below the light gray line) which indicates under prediction. However, between the departure time of the last vehicle to encounter the backward shock wave and the detection time by sensor *n* of this vehicle, the error function crosses the actual travel time function. The error function remains above the actual travel time function until time $t_{b,r,n}$. Thus, there is over prediction during this period. Then, between time $t_{b,r,n}$ and the end of the study period, time t_L, the functions are the same which indicates zero error.

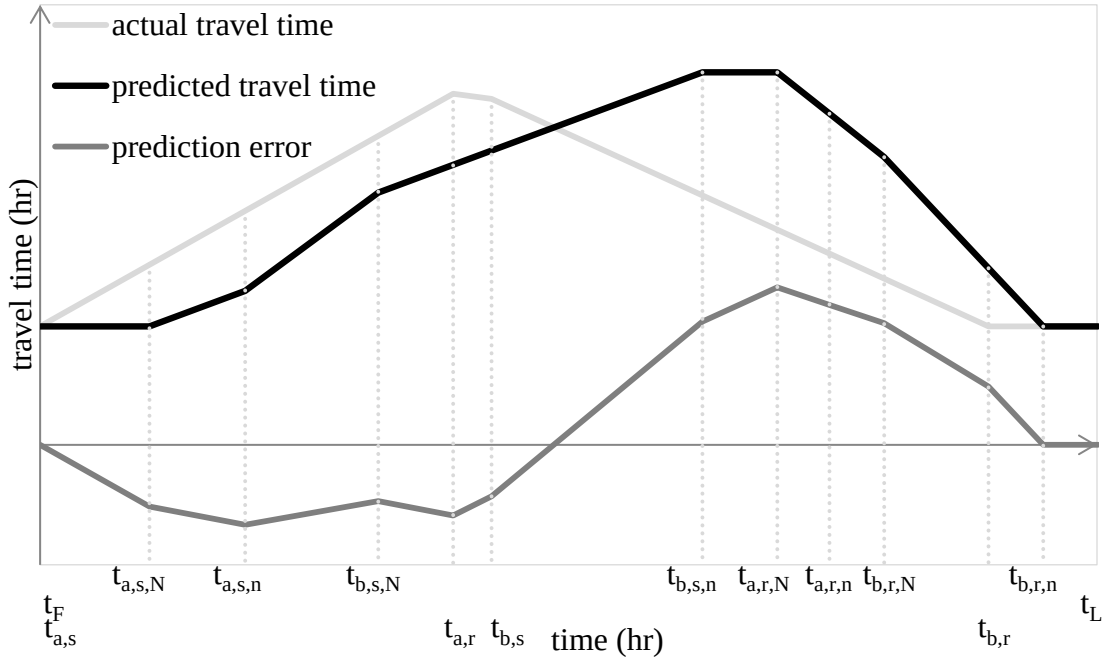


Figure 3.17 Prediction error for study period illustrated in Figure 3.12.

In order to calculate the total error and the average error of a study period, the traffic flow during the study period needs to be determined. The traffic flow for a given departure time reflects the traffic state at mile zero, the upstream end, of the freeway segment at that time. For example, if the beginning of the segment (i.e., mile zero) is uncongested, then the uncongested traffic flow is applied. In contrast, if the beginning of the segment is congested, then the congested traffic flow is applied. The traffic flow may change during a study period if a shock wave reaches mile zero of the segment. For example, in Scenario One depicted in Figure 3.2, an uncongested traffic flow is used for departure times before the backward shock wave reaches mile zero, from time t_F to time t_b, and a congested traffic flow is used after, from time t_b to time t_L. In contrast, in Scenario Two depicted in Figure 3.3, a congested traffic flow is used for departure times before the backward recovery wave reaches mile zero, from time t_F to time t_b, and an

uncongested traffic flow is used after, from time t_b to time t_L. Similarly, if a forward recovery wave originates upstream from the segment, then the traffic flow is congested from the beginning of the study period until the wave enters the segment, which occurs at time t_w, and is uncongested after the wave enters the segment until the end of the study period. However, if a forward recovery wave originates within the segment, such as in Figure 3.4, then the uncongested flow is used for the entire study period.

For study periods in which two waves reach mile zero, such as in Scenarios Four and Five shown in Figures 3.7 and 3.8, the traffic flow changes from uncongested to congested and then returns to uncongested. First, the traffic flow is uncongested from the beginning of the study period, time t_F, until the backward shock wave reaches mile zero at time $t_{b,s}$. Then, the traffic flow is congested from time $t_{b,s}$ until the time at which backward recovery wave reaches the upstream end, time $t_{b,r}$. Finally, the traffic flow is uncongested from time $t_{b,r}$ until the end of the study period, time t_L. However, if a shock wave never reaches mile zero during the study period, such as during Scenario Six shown in Figure 3.9, the traffic flow is assumed to be uncongested for the entire study period.

Next, the total error for a study period, denoted E, is determined. The functions of actual travel time by departure time, $f(t)$, and travel time by departure time, $g(t)$, are applied. These functions were found for each Scenario previously in sections 3.2.2 and 3.2.3. The absolute value of the area between the two functions represents the difference in the actual hours traveled by vehicles and the predicted hours traveled by vehicles for a study period. The area is calculated as shown in Eq. 3.64 below. The absolute value of the error is used so that the positive and negative errors of the study period do not cancel each other out.

$$|g(t) - f(t)| = \begin{cases} g(t) - f(t), & g(t) \geq f(t) \\ f(t) - g(t), & f(t) \geq g(t) \end{cases} \quad (3.64)$$

The area has a unit of (hours of travel – hour) because the travel times are formally called hours of travel and the study period is measured in hours. This area is then multiplied by the traffic flow (veh/hr) to determine the total error for the study period. The unit of total error is therefore vehicle hours of travel or VHT. The equations to calculate the total error are provided below for each Scenario:

- Scenario One

$$E = q_u \int_{t_F}^{t_b} |g(t) - f(t)| dt + q_c \int_{t_b}^{t_L} |g(t) - f(t)| dt \quad (3.65)$$

- Scenario Two

$$E = q_c \int_{t_F}^{t_b} |g(t) - f(t)| dt + q_u \int_{t_b}^{t_L} |g(t) - f(t)| dt \quad (3.66)$$

- Scenario Three if $x_w \leq x_{n-1}$

$$E = q_c \int_{t_F}^{t_a} |g(t) - f(t)| dt + q_u \int_{t_a}^{t_L} |g(t) - f(t)| dt \quad (3.67)$$

- Scenario Three if $x_w > x_{n-1}$ and Scenario Six

$$E = q_u \int_{t_F}^{t_L} |g(t) - f(t)| dt \quad (3.68)$$

- Scenario Four and Scenario Five

$$E = q_u \int_{t_F}^{t_{b,s}} |g(t) - f(t)| dt + q_c \int_{t_{b,s}}^{t_{b,r}} |g(t) - f(t)| dt + q_u \int_{t_{b,r}}^{t_L} |g(t) - f(t)| dt \tag{3.69}$$

As can be seen in Eq. 3.65, 3.66, 3.67, and 3.69, for study periods during which there are two traffic flows, both an uncongested and congested traffic flow, the function is separated into sub periods during which the flow is constant. Then, the area of each sub period is found and multiplied by its traffic flow. In addition, as can be seen in Eqs. 3.65 through 3.69, the total error has a positive value because the absolute value of the difference in the functions is used. The values for the independent variables in each equation are calculated according to the given Scenario.

Last, the average error during a study period, denoted *e*, is calculated. The average error has a unit of hours of travel, or more simply, hour. It equals the total error divided by the number of vehicles that depart during the study period. The equations to calculate the total error are provided below for each Scenario:

- Scenario One

$$e = \frac{E}{q_u(t_b - t_f) + q_c(t_L - t_b)} \tag{3.70}$$

- Scenario Two

$$e = \frac{E}{q_c(t_b - t_f) + q_u(t_L - t_b)} \tag{3.71}$$

- Scenario Three if $x_w \leq x_{n-1}$

$$e = \frac{E}{q_c(t_a - t_f) + q_u(t_L - t_a)} \tag{3.72}$$

- Scenario Three if $x_w > x_{n-1}$ and Scenario Six

$$e = \frac{E}{q_u(t_L - t_f)} \tag{3.73}$$

- Scenario Four and Scenario Five

$$e = \frac{E}{q_u(t_a - t_f + t_L - t_{b,r}) + q_c(t_{b,r} - t_{b,s})} \tag{3.74}$$

The values for the independent variables in each equation are calculated according to the given Scenario. The average error has a positive value because the terms used to calculate it are positive.

In conclusion, three error indices were provided. These are the error per departure time, total error per study period, and average error per study period. The study period includes all vehicles affected by the shock wave or waves. In order to calculate their values, both the actual travel times calculated in Section 3.2.2 and the predicted travel times calculated in Section 3.2.3 are applied. In addition, the equations and the independent variables used in the equations are Scenario specific.

3.6 Summary

In summary, this chapter presents a methodology to determine vehicles' travel times and predicted travel times on a freeway segment undergoing a transition in its traffic state. Six different study period scenarios are considered, including three with a single transition in the traffic state of a freeway segment, represented by the passage of a shock wave through the freeway segment, and three with two transitions in the traffic state, represented by the passage of two shock waves through the freeway segment. The inputs to the models are the segment's characteristics and the characteristics of the transitions, and the shock waves which represent them. In particular, the following segment's characteristics are considered:

- The segment length
- The number of sensors on the segment
- The uncongested traffic state with its associated speed and traffic flow

Also, the following incident characteristics are considered:

- The congested traffic state with its associated speed and traffic flow
- The speed of the shock waves
- The location at which the shock waves originate
- The amount of time between the shock waves, referred to as the incident duration

As a reminder, the uncongested traffic state is the condition that is present before the backward shock wave passes and after the backward recovery wave passes and the congested traffic state is the condition that is present after a backward shock wave passes or before a recovery wave passes.

From these two models and their associated inputs, the total and average travel time prediction error for each study period scenario is produced. The difference in the actual and predicted travel times is the error. In Chapter 4, these models are applied to optimize the number of sensors on a freeway segment such that the travel time prediction error can be minimized.

CHAPTER 4

METHODOLOGY

The objective function, weighted average travel time prediction error for a freeway segment under incident conditions, is formulated in this chapter. The travel time prediction error during an incident is decided by various factors such as the path based sensor spacing on the segment and the incident's duration, location, and volume to capacity (v/c) ratio. The v/c ratio represents the relationship between the average uncongested traffic flow during the incident and the maximum traffic flow possible at the incident location. It has a direct impact on the congested traffic flow and speed and on the backward shock wave speed and the backward recovery wave speed. The model developed in Chapter 3 is applied to calculate the average travel time prediction error during an incident.

The optimal solution is obtained by considering that the incidents that affect a freeway segment have different durations, v/c ratios, and locations and that the optimal solution for each incident may differ. For example, the optimal spacing for an incident that occurs during the off-peak period may not be the same as the optimal spacing for an incident that occurs during the peak period. The different incidents that can occur on the freeway segment are included in the objective function. This is achieved by determining the error by sensor spacing for each of the incidents and then also determining the probability of that incident occurring. The weighted average error by sensor spacing is then calculated. The objective function is minimized with an optimal spacing of sensors.

Chapter 4 is organized as follows. First, a set of system assumptions is made to formulate the research problem. Second, the model is formulated. Third, constraints on the optimal solution are developed. Fourth, solution methods are discussed. Last, the chapter is summarized.

Table 4.1 Variables Used in the Methodology and Their Symbols and Descriptions

Symbol	Description	Unit
B	Yearly sensor budget	$/year
b	Yearly cost per sensor	$/sensor/year
d	Minimum distance required between sensors	mi/sensor
□	Weighted average prediction error	min
i	Index of incident duration classes	-
j	Index of v/c ratio classes	-
k	Index of incident locations	-
L	Total number of incident locations	location
N^*	Optimal number of sensors	sensor
N_{max}	Maximum number of sensors	sensor
N_{min}	Minimum number of sensors	sensor
S^*	Optimal sensor spacing	mi
x_{w_k}	Incident location k	mi
Y	Total number of duration classes	range
y_i	incident duration that represents all durations within range i	hr
Z	Total number of v/c ratio classes	range
z_j	v/c ratio that represents all v/c ratios within range j	-

4.1 System Assumptions

The formulated model is designed to optimize the sensor spacing on a freeway segment such that the prediction error that occurs during an incident is minimized. The methodology is developed for incidents that cause a capacity reduction at the incident site, such as an accident blocking a lane, and then has a capacity restoration, such as when the accident is cleared. The reduction and restoration of the capacity each result in a transition in the traffic state of the freeway segment. The first transition is from an

uncongested traffic state to a congested traffic state and the second transition is from a congested traffic state back to the uncongested traffic state. A backward shock wave occurs as a result of the first transition and a backward recovery wave occurs as a result of the second transition, as discussed in Chapter 3. Both waves initiate at the incident location and travel in the upstream direction against traffic. Thus, study periods that satisfy the following scenarios described in Chapter 3 are considered: an incident with a backward shock wave and backward recovery wave represented by Scenarios One and Two, respectively, or by Scenarios Four, Five, or Six. The suitable scenario for the incident is determined by its duration and v/c ratio. Study periods that involve other types of waves, such as Scenario Three which contains a forward recovery wave, are not discussed in the methodology, but can be included with slight modifications.

To formulate the research problem, a system assumption is made about the availability of historical incident data. It is assumed that a database of historic incidents is available which includes the incidents' durations and v/c ratios. From the database, one can see that the previous incidents on the freeway segment had different durations and different v/c ratios. Also, it is assumed that the future incidents on the freeway segment will have different durations and v/c ratios and that these characteristics for the future incidents are the same as those in the past.

Therefore, the set of possible future incidents contains one incident per every combination of each possible duration, v/c ratio, and incident location. These assumptions are in addition to those presented previously in Chapter 3.

4.2 Model Formulation

The travel time prediction error for each of the possible future incidents that can occur on the freeway segment is calculated for the optimization model. As stated in the system assumptions, the set of possible future incidents contains one incident per every combination of each possible duration, v/c ratio, and incident location. A set of discrete incident locations are chosen. These can be the sites of past incidents or bottlenecks. Or, if the entire length of the segment is a potential site, then discrete points are chosen, such as every mile or half mile. The total number of incident locations is denoted L and the locations are indexed on k where x_{w_k} represents incident location k. In addition, to decrease the number of possible incidents to consider for the optimization model, a smaller set of incident durations are chosen to represent all the possible incident durations from the database and a smaller set of v/c ratios are chosen to represent all the possible v/c ratios from the database. To choose the durations to include in the smaller set, the following two steps are taken. First, a frequency distribution of the incident durations from the historical database is developed. The durations are grouped into Y mutually exclusive classes with equal class width (e.g, 0-1hr, 1-2hr, >3hr). An exception can be made to allow for unequal class intervals to avoid a large number of empty classes (e.g., the last class is for all durations greater than 3 hours). Second, a discrete duration from within each class (e.g. the average incident duration in the class) is chosen to represent all durations within the class. Therefore, if there are Y total incident duration classes, where the classes are indexed on i, let y_i be the incident duration that represents all the durations within class i.

Likewise, to choose the v/c ratios to include in the smaller set, the following two steps are taken. First, a frequency distribution of the v/c ratios from the historical database is developed. The durations are grouped into Z mutually exclusive classes with equal class width, allowing for the exception for an unequal class interval (e.g, 0-.5, .5-1, >1). Second, a discrete v/c ratio from within each class is chosen to represent all v/c ratios within the class. Therefore, if there are Z total v/c ratio classes, where the classes are indexed on j, let z_j be the v/c ratio that represents all the v/c ratios within class j.

It is acceptable to form the classes and choose one value to represent all the values within a range because incidents of similar durations or incidents of similar v/c ratios will have comparable relationships between the travel time prediction error and sensor spacing. The set of possible future incidents is thus reduced to one incident per every combination of a duration that represents a range (i.e., y_i) with a v/c ratio that represents a range (i.e. z_j) with an incident location (i.e. x_{w_k}). For example, if there are Y classes of durations, Z classes of v/c ratios, and L incident locations, the set of possible future incidents thus includes a total of Y, multiplied by Z, multiplied by L, incidents.

The probability of each incident occurring is dependent on the frequency of its duration and the frequency of its v/c ratio in the historic database and on the likelihood of an incident occurring at the location under consideration. The probability for incident duration y_i is denoted $P(y_i)$ and equals the frequency of duration class i. For instance, if thirty percent of historic incidents had durations within class i, then $P(y_i)$ equals 0.3. Similarly, the probability for v/c ratio z_j is denoted $P(z_j)$ and equals the frequency of v/c ratio class j. The probability for incident location x_{w_k} is denoted $P(x_{w_k})$ and equals the likelihood of an incident occurring there. If incident locations are continuously

distributed, than the probability for each location is equal. For the durations, v/c ratios, and locations, the sum of the probabilities for all incident conditions equals one, as illustrated in Eq. 4.1.

$$\sum_{i=1}^{Y} P(y_i) = \sum_{j=1}^{Z} P(z_j) = \sum_{k=1}^{L} P(x_{w_k}) = \sum_{i=1}^{Y}\sum_{j=1}^{Z} P(y_i)P(z_j)\,P(x_{w_k}) = 1 \qquad (4.1)$$

Therefore, the probability of an incident occurring with a duration of y_i, a v/c ratio of z_j, and at location x_{w_k} equals the product of $P(y_i)$, $P(z_i)$, and $P(x_{w_k})$.

The travel time prediction error for each incident is dependent on its duration, v/c ratio, location, and sensor spacing. This relationship is found with the model provided in Chapter 3. To optimize the spacing, the travel time prediction error for each incident needs to be taken into account. To do so, the weighted average error of all incidents is calculated. The weighted average error equals the sum of each incident's average prediction error multiplied by the probability of that incident occurring. The weighted average travel time prediction error for a freeway segment with N sensors, denoted ε_N (hr), can therefore be formulated as follows

$$\varepsilon_N = 60\sum_{i=1}^{Y}\sum_{j=1}^{Z}\sum_{k=1}^{L} P(y_i)P(z_j)P(x_{w_k})e_{i,j,k,N} \qquad \forall\, N \qquad (4.2)$$

where $e_{i,j,k,N}$ is the average error of an incident with the following characteristics: the incident is located at mile x_{w_k}, the incident duration (i.e., the elapsed time between $t_{w,r}$,

and $t_{w,r}$) equals y_i, and the uncongested and congested traffic states correspond to v/c ratio z_j where the shock wave speeds correspond to these traffic states' flows and speeds. Also, the freeway segment has N evenly spaced sensors. The model in Chapter 3 is applied to calculate the average error per incident. The error equation is multiplied by sixty to convert from the unit of hour to minutes.

4.3 Constraints

The constraints considered in this study include the maximum and minimum number of sensors per freeway segment, denoted N_{max} and N_{min}, respectively. For the first constraint, the maximum number of sensors can be restricted both by budgetary and technology limits. For a yearly budget of B ($) for purchase, installation, and maintenance costs, and a yearly cost per sensor of b ($), the maximum number of sensors due to budgetary limits is found with the top equation in Eq. 4.3. In addition, if the sensor technology requires that the sensors be a minimum of d miles apart to minimize interference between sensor signals, then the maximum number of sensors due to technology constraints is found with the bottom equation in Eq. 4.3.

$$N_{max} = \begin{cases} \frac{B}{b}^{-}, & \frac{B}{b} < \frac{D}{d} \\ \frac{D}{d}^{-}, & \frac{B}{b} \geq \frac{D}{d} \end{cases} \tag{4.3}$$

Note that in Eq. 4.3, the negative sign indicates to round down to the nearest integer to ensure that the constraint is not violated. Also, if only one of the maximum constraints is applicable, then only it should be considered. Next, the second constraint ensures that

path based travels times can be calculated for the freeway segment. Therefore, a minimum of two sensors per freeway segment is required. Thus, N_{min} = 2. The constraints can be expressed in terms of sensor spacing, instead of number of sensors, by calculating the equivalent spacing for the given number of sensors and segment length. Eq. 3.27 from Chapter 3 is applied.

4.4 Optimization Model

Based on the discussion in Sections 4.1 through 4.3, the studied model of the number of sensors that minimizes the weighted average total travel time prediction error on a freeway segment during incidents, subject to the budgetary and technical constraints, is formulated as follows:

Minimize

$$\varepsilon_N = 60 \sum_{i=1}^{Y} \sum_{j=1}^{Z} \sum_{k=1}^{L} P(y_i) P(z_j) P(x_{w_k}) e_{i,j,k,N} \qquad \forall N \tag{4.4}$$

Subject to

$$N_{min} \leq N \leq N_{max}$$

4.5 Solution Method

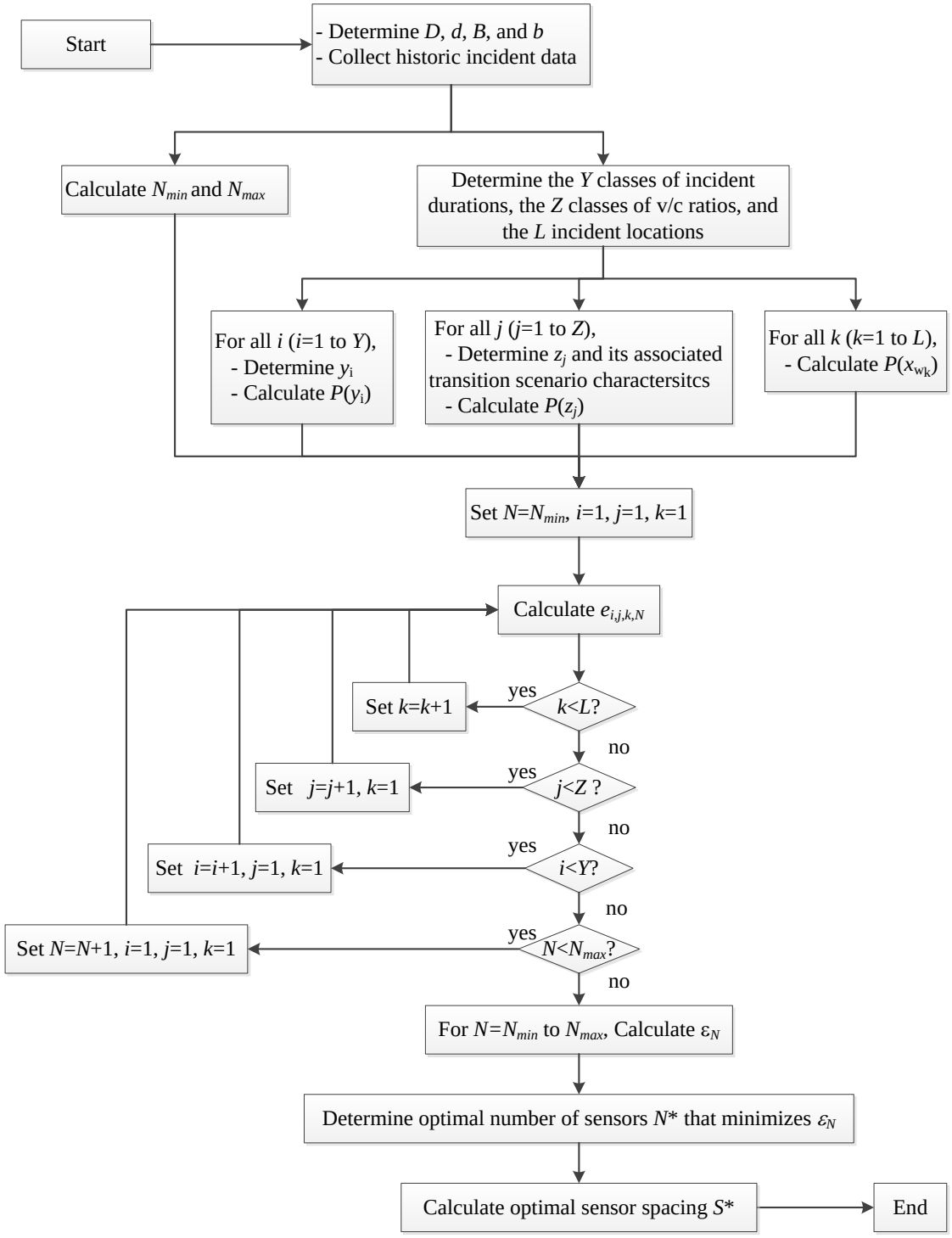
The objective of this study is to develop a model to optimize the sensor spacing on a freeway segment. The optimal number of sensors minimizes the weighted average travel time prediction error during incidents. Therefore, the decision variable is the sensor

spacing. The steps to determine the optimal spacing are described below and are shown in Figure 4.1.

1. Obtain model input data including the freeway segment length (D), the minimum amount of space needed between sensors (d) and their budget and costs (B and b). Also, collect historical incident data including the incidents' durations and v/c ratios.
2. From the data found in Step 1, calculate the constraints (N_{min} and N_{max}) and analyze the historic incident duration distribution and v/c ratio distribution to choose the Y incident duration classes and the Z v/c ratio classes. Also, determine the L incident locations.
3. For each incident duration class, determine a discrete duration (y_i) to represent it. Also, calculate $P(y_i)$, the probability of an incident with duration y_i occurring. For each v/c ratio class, determine a discrete ratio to represent it (z_j) and its associated traffic states' and shock waves' characteristics. Also, calculate $P(z_j)$, the probability of an incident with v/c ratio z_j occurring. For each incident location (x_{w_k}) calculate $P(x_{w_k})$, the probability of an incident occurring there.
4. Calculate the travel time prediction error for every possible future incident (i.e. every combination of y_i, z_j, and x_{w_k} for all i, j, and k) when there are N sensors on the segment. Iterate through each possible number of sensors, ($N_{min} \leq N \leq N_{max}$). The model developed in Chapter 3 is applied to calculate the error.
5. For each possible number of sensors, calculate the weighted average of the total prediction error with Eq. 4.2.

6. Choose the number of sensors which results in the smallest weighted average prediction error, denoted N^*.
7. Covert N^* to the optimal sensor spacing, denoted S^*, with Eq. 3.27.

More details for Steps 1, 3, and 4 are discussed in Sections 4.5.1, 4.5.2, and 4.5.3, respectively.



The notations are listed in Tables 3.1 and 4.1.

Figure 4.1 Solution algorithm for the optimal number of sensors.

4.5.1 Input Data

According to Step 1 of the Solution Method, input data is needed. The input data, which are the freeway and incident characteristics, can be obtained from multiple sources. One source is historical data that is found in incident logs or that was collected through sensors, the floating car technique, videos, etc. Another source is the Highway Capacity Manual as it can be referenced to calculate traffic flow.

4.5.2 Incident Case Studies

Step 3 instructs that for each v/c ratio range *j*, determine an incident case study whose v/c ratio falls within range *j*. Specifically, determine the incident's uncongested and congested vehicle speed and flows and the speeds of the backward shock wave and backward recovery wave as these are the inputs to the travel time prediction model. One way to determine this information is through the application of microscopic simulation software.

4.5.3 Computer Code for the Travel Time Prediction Model

For Step 4, the travel time prediction error is calculated with the model provided in Chapter 3. The model is therefore applied a number of times equal to the product of *L*, *N*, *Y*, and *Z*. In order to facilitate the calculations for average travel time prediction error, an application was developed in Java to automate the model's calculations and step procedures described in Chapter 3. The application calculates the average travel time prediction error for a study period with a backward shock wave and backward recovery wave. Inputs to the model include the segment characteristics, the different incident scenarios, durations, and locations to consider, and the different numbers of sensors to

test. The calculations for 300 incident scenarios, considering forty different possible numbers of sensors for each scenario, were computed in less than two seconds. The code is provided in Appendix A.

4.6 Summary

In this chapter, the objective weighted average travel time prediction error function was formulated. The error during the entirety of an incident was considered including the error from the passage of the backward shock wave and the error from the passage of the backward recovery wave. The waves are consequences of the capacity reduction and the subsequent capacity restoration at the incident site. The objective function is formulated so that it considers that incidents with different durations, v/c ratios, and locations affect the study freeway segment. The error per incident is weighted by the probability of an incident occurring with its duration and v/c ratio and at its location. To determine the error per incident, the model developed in Chapter 3 is applied.

The decision variable is the sensor spacing. The optimization model can be solved by applying the steps listed in Section 4.5. The methodology is demonstrated in Chapter 5 with a case study.

CHAPTER 5

NUMERICAL EXAMPLE

This chapter demonstrates the applicability of the travel time prediction model developed in Chapter 3 and the optimization algorithm presented in Chapter 4. In Section 5.1, the macroscopic principles and assumptions used in the mathematical model are evaluated by comparing the model's results to a microscopic travel time prediction model which represents a "real-world" scenario. In Section 5.2, a case study is presented. Finally, in Section 5.3, a sensitivity analysis was performed to determine how the class size used for both the incident duration frequency distribution and the v/c ratio frequency distribution affected the optimal sensor spacing.

5.1 Validity of the Mathematical Model

The purpose of this section is to explore the validity of the mathematical model presented in Chapter 3, especially its use of macroscopic principles and assumptions. The macroscopic traffic principles applied in the model included the use of an average vehicle speed and average traffic flow and the use of shock waves to represent transitions. The model's assumptions included continuous vehicle departure times and continuous sensor detection of vehicles. The model also assumed that predicted travel times are calculated instantaneously and are then continuously updated. The validity of the model was explored by comparing its results to the "real-world" travel time prediction error during a traffic incident. In the real-world scenario, microscopic travel time data was analyzed, such as individual vehicles' departure times, travel times, and predicted travel times.

Also, the predicted travel times were calculated and updated every minute instead of instantaneously. In comparison to the mathematical model, with this real-world scenario, the macroscopic nature and assumptions were relaxed.

The "real-world" travel time prediction error by sensor spacing was calculated with microscopic travel time data collected during an incident. In particular, Paramics microscopic software was utilized to simulate an incident. A ten mile, three lane segment in Northern New Jersey was developed with a capacity estimated to be 8,250 vph. The input demand was 4,950 vph. The simulation period was two hours and an incident was programmed to occur after fifteen minutes that blocked one lane (the outside lane) at the midpoint of the segment for twenty minutes. The simulation was repeated five different times, each with a different seed value that was randomly generated.

Individual vehicle data was collected during the simulation period. This was achieved by placing sensors on the segment at one twelfth mile spacing. Each sensor recorded the time at which each vehicle passed it along with the unique identification number of the vehicle. Paramics outputted the information from each of the sensors into its own file for a total of 120 separate comma separated value (CSV) files per run. The information was then manipulated and combined into one file. A macro, written in Visual Basic, was developed to semi-automate the conversion of the files so that the data can be analyzed. The macro converted the files from CSV to Excel (xlsx), changed the order of the columns so that the vehicle identification number is the first column, which is needed for the VLOOKUP function in Microsoft Excel, and changed the number format of the time stamps so that they could be analyzed. In their original state, time stamps from the second hour of a simulation period could not be distinguished from the

time stamps from the first hour of the simulation period if they were not first converted into the number format. (This macro is included in Appendix B.) By comparing vehicles' time stamps from one sensor to the next, individual vehicles' trajectories along the segment were calculated. This included each vehicle's departure time, actual travel time, and travel time between each pair of sensors. Based on the sensor spacing being tested, such as one fourth mile, one third mile, etc., the appropriate pairs of sensors were determined from which to compute predicted travel times. Predicted travel times were computed as one minute averages. The predicted travel time for vehicles departing each minute equaled the sum of the previous minute's average travel times between each sensor pair.

As a concrete example of the prediction calculation, suppose the sensor spacing was five miles on the ten mile segment. The prediction provided to vehicles departing from minute t to minute t+1 was calculated as follows. First, there was a set of vehicles detected by the sensor at mile five from minute t-1 to minute t. These vehicles' travel times from mile zero to mile five were calculated by comparing their time stamps at the mile zero and mile five sensors. Then, the average travel time of these vehicles on the link was calculated. Concurrently, during the same one minute time period, vehicles were detected by the sensor at mile ten. These vehicles' travel times from mile five to mile ten were calculated and the average travel time on that link during the time period was recorded. Next, the two average travel times, the one from mile zero to mile five and the one from mile five to mile ten, were summed to find the travel time for the segment. Finally, this travel time was used as the predicted travel time for minute t to minute t+1.

From the simulation data and its analysis, the relationship between actual travel time and vehicle departure time and the relationship between predicted travel time and departure time were determined for various sensor spacing. For each possible spacing, the two functions were then compared to calculate the travel time prediction error. This process was repeated with the output data from each of the five seeds' runs. The resulting average travel time prediction error by sensor spacing for vehicles affected by the incident is displayed in Figure 5.1. How to determine the length of a study period was previously discussed in Chapter 3.

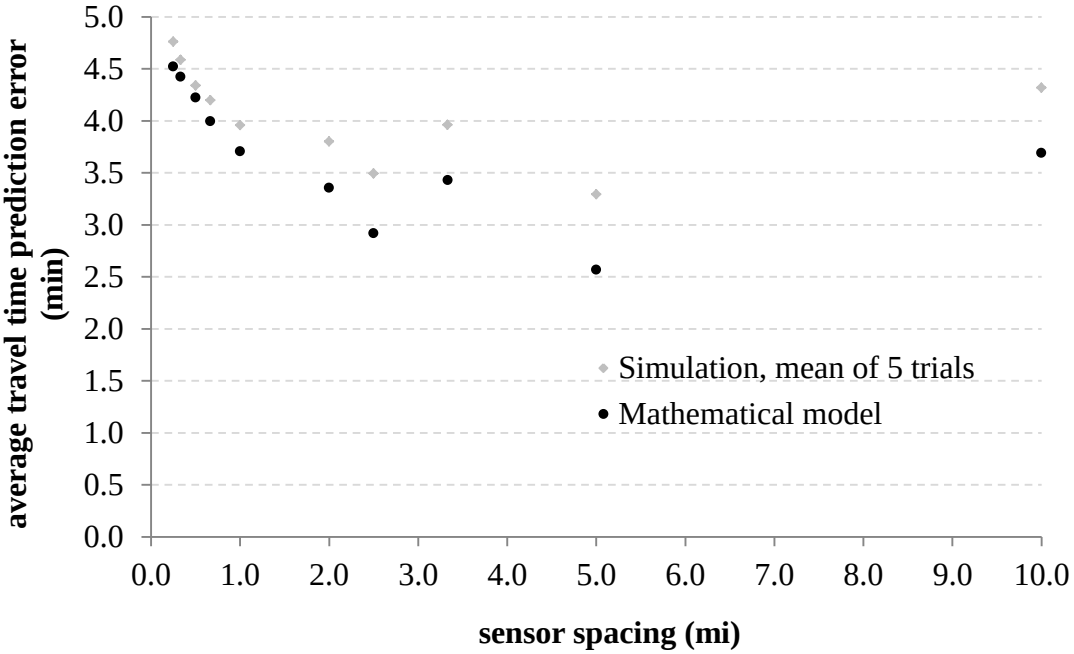


Figure 5.1 Travel time prediction error versus sensor spacing for the mathematical model and the micro simulation model.

One can see from Figure 5.1 that, in general, the average error decreased as the sensor spacing increased. As explained in more detail later in Section 5.2.4, this is a result of the short incident duration.

The travel time prediction error by sensor spacing for the real-world incident was also calculated using the mathematical model so that the two methods could be compared. The macroscopic input values for the model were determined by analyzing the Paramics output data which included the data mentioned above and another set of output files that contained each link's vehicle flow per minute. A link was a 0.25 mile section of the segment. The following incident characteristics were determined by analyzing the simulation output data: the average uncongested vehicle speed and average uncongested vehicle flow, the average congested vehicle speed and the average congested vehicle flow, the backward shock wave speed, the backward recovery wave speed, and the volume-to-capacity (v/c) ratio. Also, the volume "v" to capacity "c" ratio was calculated as follows. First, the volume equaled the uncongested traffic flow. Second, the incident capacity equaled the segment capacity (8,250 vph) multiplied by the percentage of freeway capacity that was available at the incident location when one of the three lanes was blocked, which is 49% (PB Farradyne 2000).

The values for the incident characteristics are summarized in Table 5.1. Note they are the average value from the results from the five different runs. In addition, the incident characteristics for three other Paramics input demands, 1,650 vph, 3,300 vph, and 6,600 vph, which were also run in the simulation with five different seed values each, are also provided. The negative shock wave speeds denote that the waves travel upstream, in the direction against the flow of traffic.

Table 5.1 Incident Characteristics Determined from Paramics Micro Simulation

Case Study Incidents							
Paramics Input	Calculated Outputs						
Demand (vph)	q_u (vph)	q_c (vph)	v_u (mph)	v_c (mph)	$v_{w,s}$ (mph)	$v_{w,r}$ (mph)	v/c ratio
1650	1664	1545	70.6	70.6	n/a*	n/a*	0.41
3300	3075	2396	66.8	7.9	-2.6	-17.0	0.76
4950	5121	2434	64.7	7.5	-8.0	-14.8	1.27
6600	6449	2646	62.9	7.0	-13.8	-14.1	1.60

* No shock waves were detected due to the low v/c ratio.

The mathematical model developed in Chapter 3 was run with the incident characteristics found for the incident that had a demand of 4,950 vph. As stated before, the incident duration was twenty minutes and the segment length was ten miles. Figure 5.1 displays the results from the micro simulation model and the mathematical model. One can see that the results are exciting for two reasons. First, the relationship between prediction error and sensor spacing follows the same shape for both models. Second, the average difference between the two models is only 23 seconds. The maximum error between the two models is forty four seconds and the minimum error is seven seconds. It is noted that the mathematical model under predicts the error and also that the mathematical model performs better for shorter sensor spacing than for longer sensor spacing. Based on these results, the author is confident in the mathematical model.

5.2 Case Study

A case study was developed to demonstrate the travel time prediction error model developed in Chapter 3 and the methodology presented in Chapter 4. In the case study, path based sensor spacing was optimized for a ten mile, three lane freeway segment in

Northern New Jersey. The optimal spacing minimized the weighted average travel time prediction error during incidents on the segment. The possible future incidents on the segment, along with the probability of each occurring, were determined by analyzing a database of historical incidents, as described in Section 5.2.1. In addition, characteristics of the traffic state of the freeway segment for different incident v/c ratios, such as the uncongested and congested vehicle speeds, the uncongested and congested traffic flows, and the shock wave speeds, were found through microscopic modeling, as described in Section 5.2.2. The average error per each incident was calculated with the model developed in Chapter 3 and the weighted average travel time prediction error was calculated with the methodology described in Chapter 4. The optimization results are provided in Section 5.2.3. The results were then analyzed in Section 5.2.4.

5.2.1 Historical Incident Database

From a historical incident database, the frequency distributions of incident durations and v/c ratios were determined. The historical incident database developed for the NJDOT report "Development of New Jersey Rates for the NJCMS Incident Delay Model" (Chien and Spasovic, 2012) was employed for this study. It included information on 57 incidents that occurred on three lane freeway segments in New Jersey from August 2008 to May 2009 that met the following requirements: the freeway segment was three lanes, the closure type was blockage of either the median, one, or two lanes, or the closure type was a shoulder disablement or shoulder accident, and the incident duration or v/c ratio was not an outlier. The following incident characteristics were included in the database:

- route
- roadway type
- number of lanes

- free flow speed
- hourly volume
- percentage of trucks in the traffic flow
- incident duration
- incident start time
- incident end time
- closure type
- number of lanes closed.

Note that the roadway capacity was not included and neither was the incident v/c ratio. However, both of these values were calculated from the provided information using current traffic engineering practices. The volume was calculated as the average hourly volume during the time period the incident occurred and the incident capacity equaled the roadway capacity multiplied by the percentage of freeway capacity available under incident conditions. This factor is shown in Table 5.2 below for different lane closure types.

Table 5.2 Percentage of Freeway Capacity Available Under Incident Conditions

Number of Lanes per Direction	Shoulder Disablement	Shoulder Accident	Lanes Blocked		
			One	Two	Three
2	0.95	0.81	0.35	0.00	N/A
3	0.99	0.83	0.49	0.17	0.00
4	0.99	0.85	0.58	0.25	0.13
5	0.99	0.87	0.65	0.40	0.20
6	0.99	0.89	0.71	0.50	0.25
7	0.99	0.91	0.75	0.57	0.36
8	0.99	0.93	0.78	0.63	0.41

Source: PB Farradyne. "Traffic Incident Management Handbook." Federal Highway Administration Office of Travel Management, 2000.

Note that median closure was not included in the table so the capacity reduction factor for shoulder disablement was used for it.

The durations of the incidents in the database are shown in Figure 5.2 below. One can see that the shape resembles a log normal curve which is consistent with the literature on the spread of incident durations.

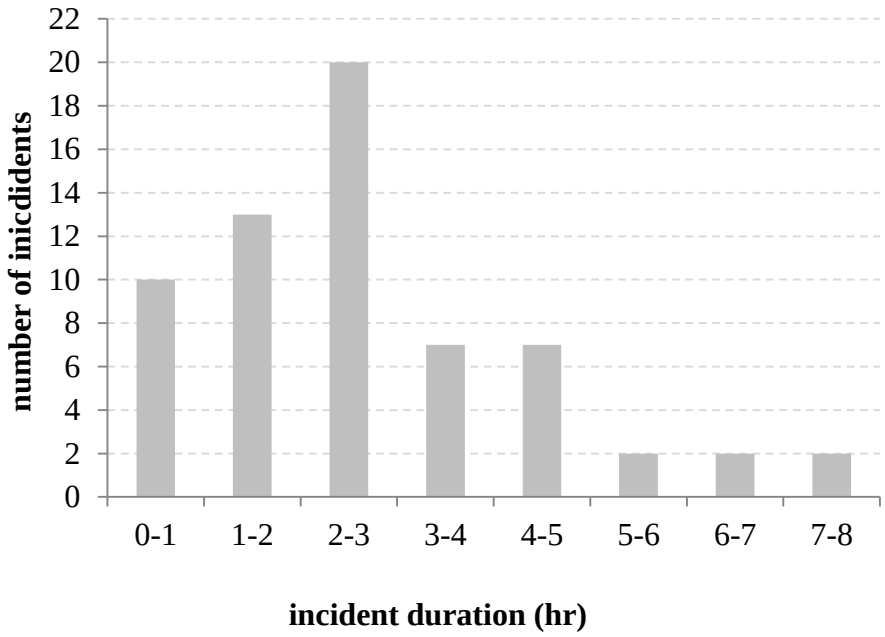


Figure 5.2 The incident durations on New Jersey's three lane freeways.

The frequency distribution of incident durations to use as part of the solution algorithm was determined with the following steps. First, incident durations greater than five hours were considered outliers and thus removed from the set leaving 57 incidents. Then, to determine the number of classes into which to divide the incidents into, the "2 to the k rule" was applied. Six classes were formed (two raised to the sixth power is greater than 57, the number of incidents in the database). Next, the class width was found by dividing the maximum incident duration (300 minutes) by the number of classes (six). Therefore, there were a total of six classes (Y=6) each with a width of fifty minutes as shown in Table 5.3. For each class, its average incident duration (i.e., the average duration of all the incidents included within the class) was chosen to represent it and thus

was the value of y_i. The value of y_i, along with the probability of it occurring, denoted $P(y_i)$, are shown in Table 5.3.

Table 5.3 Incident Duration Ranges

Range Index	Incident Durations within Range (min)	y_i (min)	$P(y_i)$ (%)
1	0 to 50	33	12.3
2	50 to 100	72	17.5
3	100 to 150	129	26.3
4	150 to 200	173	22.8
5	200 to 250	220	8.8
6	250 to 300	273	12.3

Also, the frequency distribution of the incident v/c ratios was determined and is shown in Figure 5.3 below.

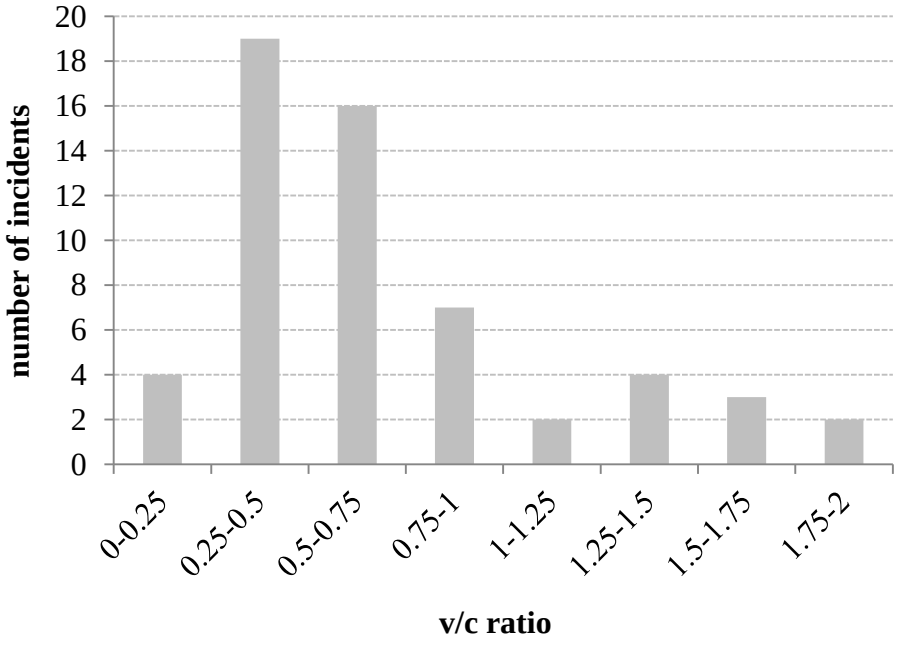


Figure 5.3 The v/c ratios on New Jersey's three lane freeways.

The maximum v/c ratio was 2.0. The six classes, as six was chosen as explained for the incident duration, were determined as follows. For v/c ratios less than 0.5, congestion is minimal and shock waves are not detectable. Therefore, v/c ratios less than 0.5 were grouped into one class. Also, v/c ratios between 1.5 and 2 were grouped into one class because not many incidents fall into this range. For the remaining v/c ratios between 0.5 and 1.5, four classes of width of 0.25 were made. Therefore, there were a total of six classes (Z=6) and they are shown in Table 5.4. For each class, its average v/c ratio was chosen to represent it and thus was the value of z_j. The value of z_j, along with the probability of it occurring, denoted $P(z_j)$, are shown in Table 5.4.

Table 5.4 Incident v/c Ratio Ranges

Range Index	Incident v/c Ratios within Range	z_j	$P(z_j)$ (%)
1	0.0 to 0.5	0.3	40.3
2	0.5 to 0.75	0.6	28.1
3	.75 to 1.0	0.9	12.3
4	1.0 to 1.25	1.0	3.5
5	1.25 to 1.5	1.3	7.0
6	1.5 to 2.0	1.8	8.8

In summary, the historical incident database was analyzed to determine the durations and v/c ratios of the future incidents that may occur on the study segment. Because there are six possible incident durations and six possible v/c ratios, there are a total of 36 different incident scenarios to include in the optimization model. Also, incidents were assumed to occur at all locations within the segment with an equal possibility. The following locations were chosen to represent the locations: every mile from mile one through mile ten (i.e., mile 1, 2, 3, etc.) for a total of ten incident locations.

Therefore, a total of 360 incidents were included in the case study, these are the 36 different incident scenarios each at every one of the ten possible incident locations.

5.2.2 Incident Characteristics by v/c Ratio

From the previous section it was found that there are six possible v/c ratios to include in the optimization model. However, based on the literature review and micro simulation testing, if the incident v/c ratio is less than 0.5 shock waves do not occur. Based on the principles of the model developed in Chapter 3, if there are no shock waves, there will be no travel time prediction error. Thus, for the incident scenarios with a v/c ratio less than 0.5, the value for the average travel time prediction error equals zero.

For the remaining five v/c ratio classes, the incident characteristics were determined by using Paramics micro simulation software. As v/c ratio is not an input into the software, instead four different demands scenarios were run in Paramics and their v/c ratios were then calculated. These incident scenarios were presented previously in Table 5.1. Because their v/c ratios were not exact matches for the v/c ratios from the incident database, the value for their characteristics were applied and interpolated to determine the incident characteristics for the case study's v/c ratios. Their values are provided in Table 5.5 below.

Table 5.5 Incident Characteristics by v/c Ratio

v/c ratio	q_u (vph)	q_c (vph)	v_u (mph)	v_c (mph)	$v_{w,s}$ (mph)	$v_{w,r}$ (mph)
0.3	n/a*	n/a*	n/a*	n/a*	n/a*	n/a*
0.6	3075	2396	66.8	7.9	-2.6	-17.0
0.9	3480	2404	66.4	7.8	-3.6	-16.6
1.0	4216	2417	65.6	7.7	-5.6	-15.8
1.3	5121	2434	64.7	7.5	-8.0	-14.8
1.8	6449	2646	62.9	7.0	-13.8	-14.1

* Prediction error was modeled to be zero for this v/c ratio and thus the incident characteristics were not calculated.

One can see from the data that there is a positive relationship between v/c ratio and uncongested traffic flow, congested traffic flow, and backward shock wave speed. Also, there is a negative relationship between v/c ratio and uncongested traffic speed, congested traffic speed, and backward recovery wave speed.

5.2.3 Optimization Results

The weighted average travel time prediction error was calculated for the case study. To do so, first, the function of average travel time prediction error by sensor spacing was calculated for the 360 incidents described in Section 5.2.2 by applying the model from Chapter 3. The number of sensors tested was two through 41, increased in increments of one (e.g., 2, 3, etc.) for a total of forty possible number of sensors to deploy on the segment. This translates to a maximum sensor spacing of ten miles or the length of the segment, and a minimum sensor spacing of a quarter mile. Then, the function of weighted average travel time by sensor spacing was calculated by applying the methodology from Chapter 4.

The function of travel time prediction error by sensor spacing for different incident scenarios are displayed in Figures 5.4, 5.5, and 5.6 below. Figure 5.4 displays the function when the incident duration is 33 minutes and the incident location is five miles, Figure 5.5 displays the function when the incident duration is 72 minutes and the incident location is five miles, and Figure 5.6 displays the function when the incident duration is 220 minutes and the incident location is also five miles. In each of the figures the function is displayed by v/c ratio. Note that the graphs reflect that for v/c ratios less than 0.5 the travel time prediction error is considered to be zero.

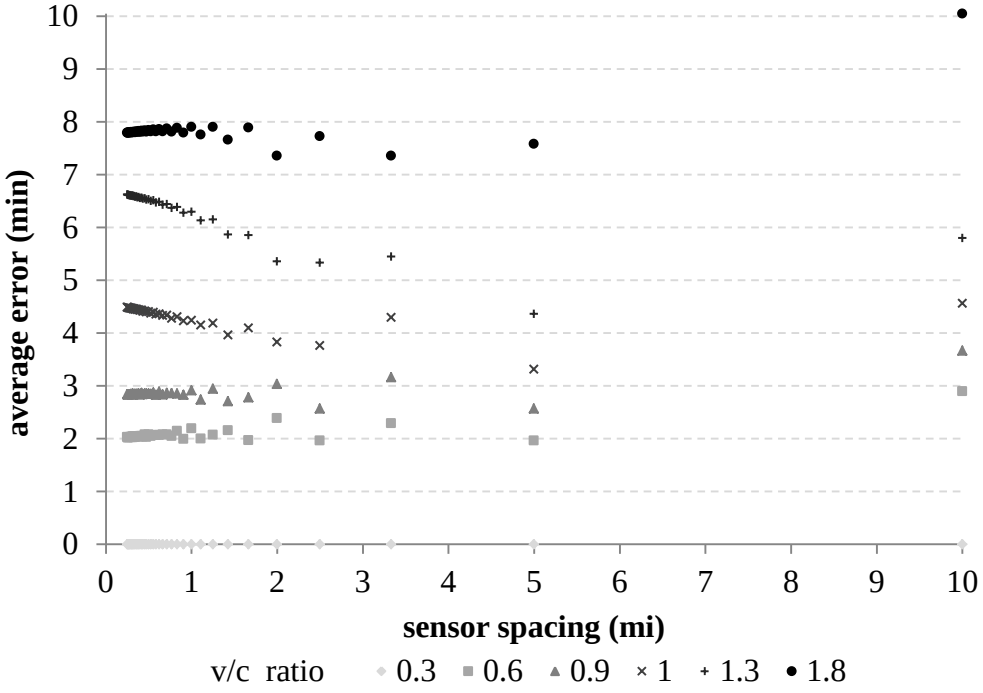


Figure 5.4 Travel time prediction error for a 33 minute incident located at 5 miles.

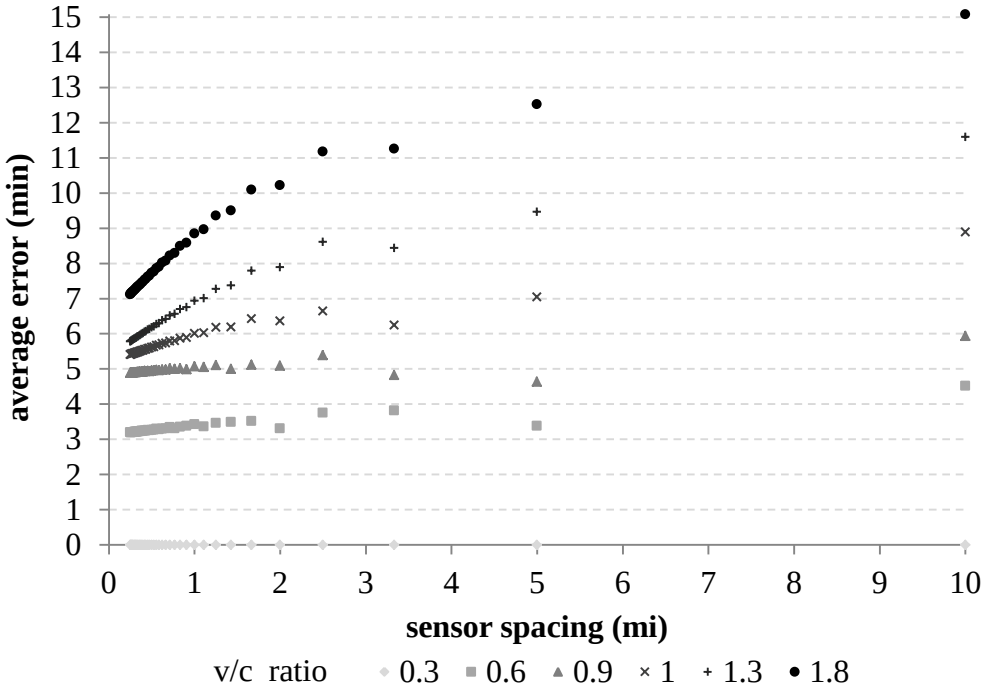


Figure 5.5 Travel time prediction error for a 72 minute incident located at 5 miles.

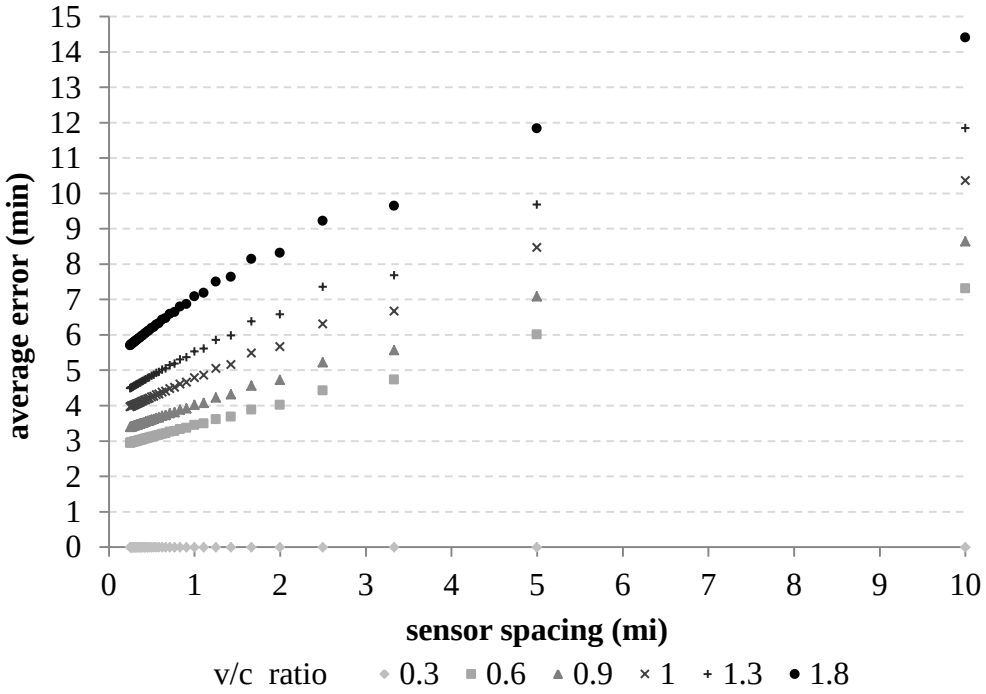


Figure 5.6 Travel time prediction error for a 220 minute incident located at 5 miles.

After determining the average travel time prediction error for each incident in the case study, the function of weighted average travel time prediction error by sensor spacing was calculated. For each sensor spacing, the weighted average error equaled the sum of the average travel time prediction error for each of the 360 incidents multiplied by the probability of that specific incident occurring. The function is shown in Figure 5.7 below.

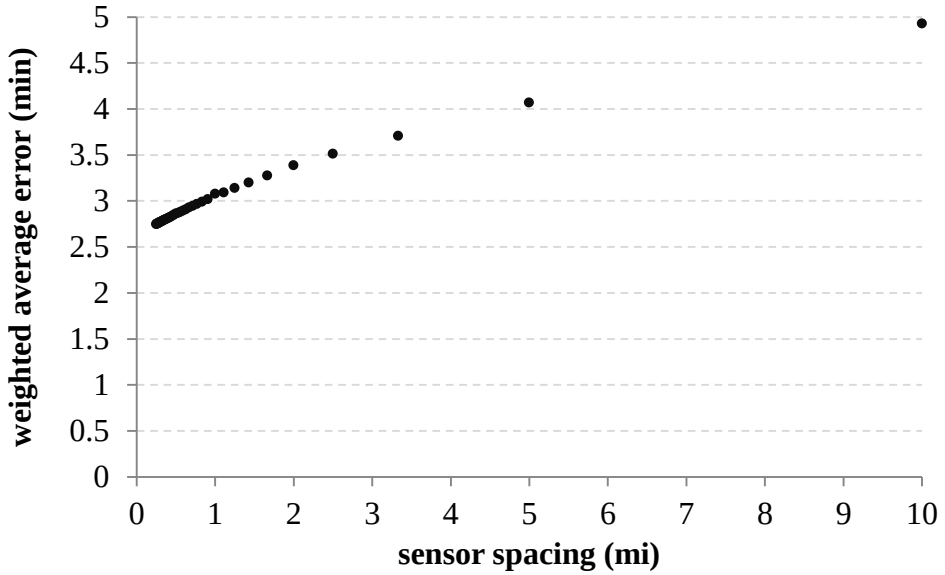


Figure 5.7 Weighted average travel time prediction error by sensor spacing.

From Figure 5.7, one can see that, overall, travel time predictions became more accurate as the sensor spacing decreased. However, the improvement in the error decreased as the spacing decreased. These results are also depicted through graphing the weighted error by the number of sensors, instead of by the sensor spacing, in Figure 5.8.

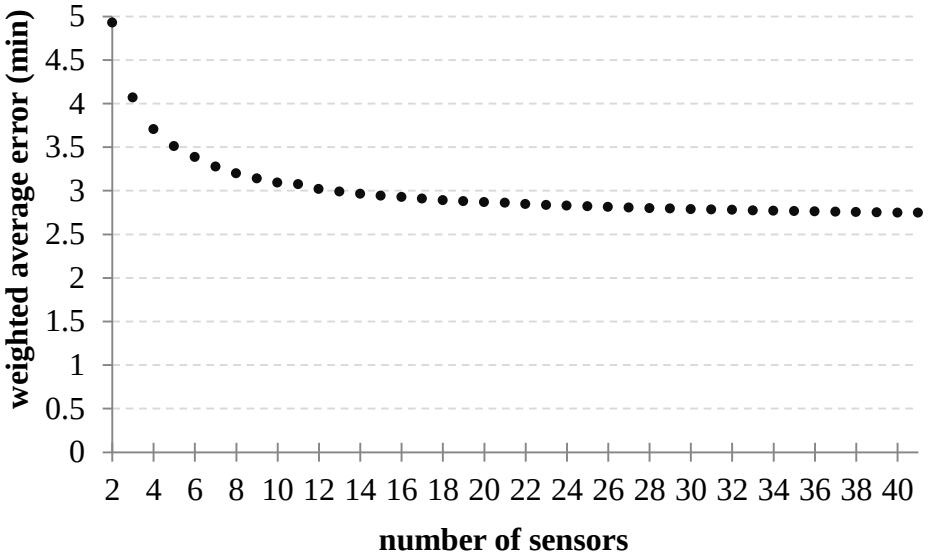
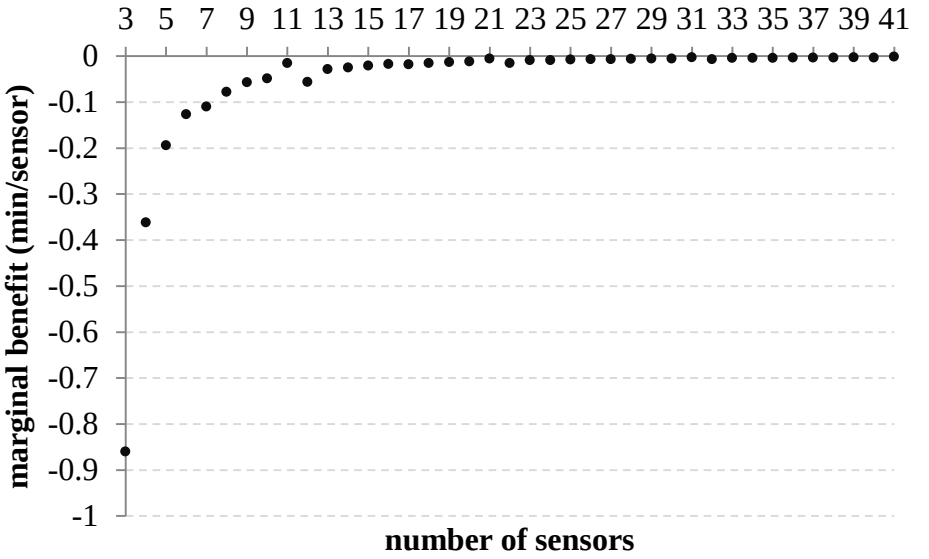


Figure 5.8 Weighted average travel time prediction error by number of sensors.

The marginal benefit of increasing the number of sensors was calculated. The marginal benefit equaled the change in the weighted average error that arose from an additional sensor on the freeway segment. It is graphed for number of sensors from 3 to 41 in Figure 5.9. The marginal benefit is negative because additional sensors decrease the weighted average error.



Figures 5.9 Marginal benefit by number of sensors.

Three observations were made from Figure 5.9. First, the marginal benefit is significant when comparing the average error for a small number of sensors. For example, the marginal benefit of increasing from two to three sensors is almost a minute improvement in the average travel time prediction error. Second, after eight sensors, which is equivalent to a senor spacing of 1.43 miles, the marginal benefit is less than 0.1 minutes and approaches zero with additional sensors. In fact, it is less than 0.01 minutes after 23 sensors, which is equivalent to a sensor spacing of 0.45 miles. Third, there seems to be a "bump" at eleven sensors which is equivalent to a sensor spacing of one mile. When there were less than eleven sensors, the spacing was greater than one mile. With eleven sensors, a sensor was located directly at every milepost so no matter which incident location was considered, there was a sensor located directly at the incident location. Also, when the number of sensors was greater than eleven, the spacing was less than one mile. In support of this reasoning, there are also "bumps" at 21 and 31 sensors, both of which also result in a sensor always being located directly at the incident location. In conclusion, increasing the number of sensors, which means decreasing the sensor spacing, improved the average error. However, the improvement in the error diminishes with each additional sensor.

5.2.4 Relationship between Incident Characteristics, Sensor Spacing, and Error

In the case study, the travel time prediction error varied not only by the sensor spacing but also by the incident duration, v/c ratio, and location. This is because both the v/c ratio and the incident duration affect the spread of congestion. With higher v/c ratios, the backward shock wave speed is faster and thus the congestion spreads more quickly. Also, with longer incident durations, the return to uncongested conditions occurs at a

later time. So, a small v/c ratio and short duration resulted in less congested conditions over space and time than a large v/c ratio and long duration. In addition, the incident location impacts how much of the segment is affected by congestion as well as the distance between the incident location and the nearest sensors. The following five observations were made regarding the relationship between travel time prediction error and sensor spacing, incident duration, v/c ratio, and location. Note they are general observations and further analysis will be partaken in future studies to verify the causes of the relationships.

First, the travel time prediction error varies by sensor spacing, incident duration, v/c ratio, and location. Therefore, when optimizing sensor spacing for a freeway segment, it is important to consider the range of possible incidents that may occur on the segment.

Second, holding the sensor spacing and incident duration constant, the average error increased as the v/c ratio increased. As the v/c ratio increased the backward shock wave speed significantly increased. Compared to a slower shock wave speed, a faster shock wave speed results in a faster rate of change between vehicle departure time and travel time. Thus, there is a greater difference between the predicted travel times provided to the departing vehicles and their actual travel times.

Third, holding the sensor spacing and v/c ratio constant, the average error increased as the incident duration increased. With short durations, the congestion does not affect much of the roadway and therefore the congested travel times are not much greater than the uncongested travel times. Note that this observation does not always hold true such as when comparing long durations against each other (e.g., 72 minutes

versus 220 minutes for a10 mile sensor spacing and a v/c ratio of 1.8). In these instances, the long duration results in a period of fully congested conditions and thus stable travel times which lead to accurate travel time predictions.

Fourth, the difference in the prediction error between applying different sensor spacing was greater when the duration and v/c ratio increased. For example, when the incident duration was 72 minutes and the v/c ratio was 0.6, the difference in the travel time prediction error between applying 0.25 mile sensor spacing and ten mile sensor spacing was about one minute. However, when the incident duration was 72 minutes and v/c ratio was 1.8, the difference was about eight minutes.

Fifth, decreasing the sensor spacing on a freeway segment did not always improve the accuracy of travel time predictions and in some instances even made the predictions less accurate. This was observed for all the v/c ratios when the incident duration was 33 minutes and for the smaller v/c ratios of 0.6 and 0.9 when the duration was 72 minutes. Appendix C provides an example to illustrate why this occurs.

5.3 Sensitivity Analysis

A sensitivity analysis was performed to determine how the class size used for both the incident duration frequency distribution and the v/c ratio frequency distribution affected the optimal sensor spacing. The incident characteristics found through micro simulation, described previously and shown in Table 5.1, were applied and a ten mile freeway segment with incident locations at each mile (i.e., mile 1, 2, 3, etc.) was modeled like in the case study.

Originally, in the case study, the class size was six. For the sensitivity analysis class sizes of two, three, four, and eight were also applied. For each of the possible class

sizes for the incident duration distribution, the characteristics for each of its classes (i = 1 to Y) are provided in Table 5.6 below. These characteristics are the range of the durations represented by each class, the average duration within each class (denoted y_i), and the probability of an incident within range i occurring (denoted $P(y_i)$). Also, for each of the possible class sizes for the v/c ratio distribution, the characteristics for each of its classes (j = 1 to Z) are provided in Table 5.7, including the range of the v/c ratios represented by each class, the average v/c ratio within each class (denoted z_j), and the probability of an incident within range j occurring (denoted $P(z_j)$).

Table 5.6 Incident Durations by Class Size

	Number of Incident Duration Classes (Y)											
	2			3			4			8		
i	range (min)	y_i (min)	$P(y_i)$ (%)	range (min)	y_i (min)	$P(y_i)$ (%)	range (min)	y_i (min)	$P(y_i)$ (%)	range (min)	y_i (min)	$P(y_i)$ (%)
1	0-150	90	56	0-100	56	30	0- 75	47	23	0- 38	19	5
2	150-300	210	44	100-200	149	49	75-150	120	33	38-75	55	18
3				200-300	251	21	150-225	184	30	75-113	95	12
4							225-300	267	14	113-150	135	21
5										150-188	169	19
6										188-225	211	11
7										225-263	249	7
8										263-300	286	7

Table 5.7 Incident v/c Ratios by Class Size

	Number of v/c Ratio Classes											
	2			3			4			8		
j	range	z_j	$P(z_j)$ (%)	range	z_j	$P(z_j)$ (%)	range	z_j	$P(z_j)$ (%)	range	z_j	$P(z_j)$ (%)
1	0- 1	0.5	81	0- .5	0.4	61	0- .5	0.3	40	0- .25	0.2	7
2	1- 2	1.5	19	.5- 1	0.9	28	.5- 1	0.7	40	.25- .5	0.4	33
3				1- 1.5	1.7	11	1- 1.5	1.2	11	.5- .75	0.6	28
4							1.5- 2	1.8	9	.75- 1	0.9	12
5										1- 1.25	1.0	4
6										1.25- 1.5	1.3	7
7										1.5- 1.75	1.6	5
8										1.75- 2	2.0	4

Changing the class size changed the model of the prediction error as shown in Figure 5.10 below. When the class size was small, such as for two or three classes, the results indicated that the average error by spacing was as little as 1.25 and 2 minutes, respectively. However, when the class size was increased, so that the range of durations and range of v/c ratios included within each class was smaller, the results indicated the average error was greater. This is because the class size dictated what incident characteristic values (i.e., the durations and v/c ratios to test) were included in the travel time prediction model. For example, when the class size was two, 81% of the incidents were categorized into the first class of v/c ratios, which ranged from zero to one, and the average v/c ratio for incidents within that class was 0.5. Therefore, the model assumed that 81% of incidents had zero travel time prediction error and thus the weighted average error reflected this. However, in reality, only 40% of incidents have v/c ratios less than 0.5.

Another observation is that the difference between four, six, and eight classes by sensor spacing was less than eight seconds. This indicates that continuously increasing the class size, which means increasing the number of times the prediction model from Chapter 3needs to be run, does not make a significant difference.

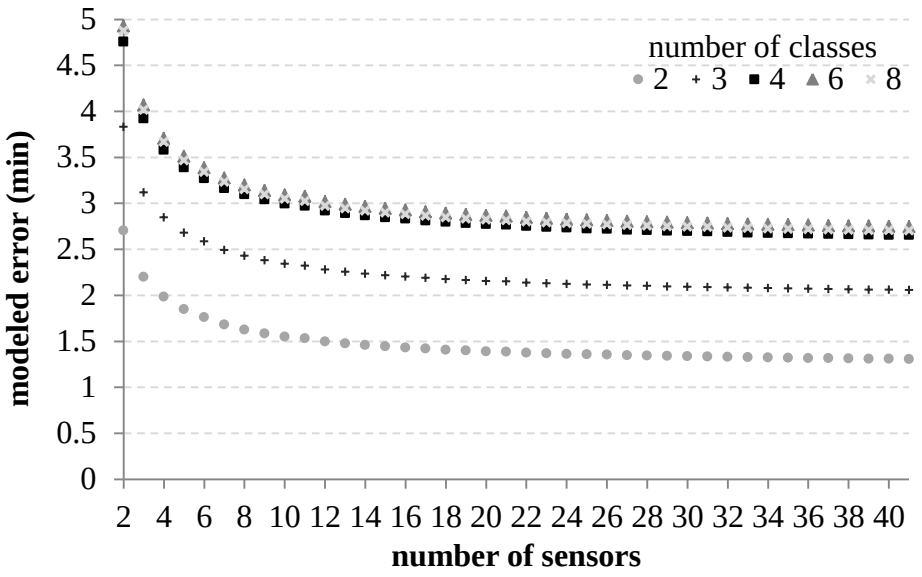


Figure 5.10 Modeled error by number of classes and number of sensors.

5.4 Summary

This chapter demonstrated the validity of the mathematical model developed in Chapter 3 by comparing it to a real-world incident scenario and it demonstrated the feasibility of applying the solution algorithm developed in Chapter 4 through a case study. Also, through the case study and sensitivity analyses the following important conclusions were drawn regarding optimal sensor spacing:

- The travel time prediction error by sensor spacing and the optimal sensor spacing are dependent on the incident duration, v/c ratio, and location.

- The distance between the incident location and the nearest sensor affects the travel time prediction error.

- Long sensor spacing is preferable when the spread of congestion over space and time is insignificant, such as for incidents with low v/c ratios and short durations.

- Short sensor spacing is preferable when the spread of congestion is large, such as for incidents with high v/c ratios and long durations.

- The marginal benefit of sensors decreases towards zero as the number of sensors increases.

- The class size chosen for the incident duration frequency distribution and the v/c ratio distribution affects the outputs obtained from the solution algorithm. Applying a class size of six for the case study, and thus adhering to the "2 to the *k*" rule was appropriate as a smaller class size would have underestimated the average prediction error and a larger class size would have produced similar results as to the class size of six but would have increased the number of incidents to include in the model.

CHAPTER 6

CONCLUSIONS AND FUTURE RESEARCH

This dissertation contributes to the field of Transportation and Intelligent Transportation Systems by providing the optimal path based sensor spacing on freeway segments. Congestion on freeways is increasing and a key source of it is non-recurring incidents. Accurate vehicle travel time predictions are needed during these incidents in order to aid roadway users and allow them to make informed trip decisions. Path based sensors are becoming a leading technology in gathering travel time data. These types of sensors are located along a roadway and collect individual vehicle travel time data by communicating with technology located inside the vehicles that drive pass the sensors. This dissertation provided the relationship between path based sensor spacing and travel time prediction error during all incidents that affect a freeway segment. This included the marginal benefit of additional sensors. From this contribution, Departments of Transportation and other policy makers can determine the optimal sensor spacing on their freeways so that the overall average travel time prediction error during incidents is minimized.

A model was provided to determine the travel time prediction error by path based sensor spacing during the passage of shock waves on a freeway segment. A shock wave occurs when the traffic state of a roadway transitions from being uncongested to congested or vice versa. The type of shock waves considered were a backward shock wave, backward recovery wave, and forward recovery wave; in addition, study periods that included both a backward shock wave and backward recovery wave were included. Model inputs included the uncongested and congested vehicle speeds and traffic flows,

the speed of the shock wave or waves, the location the shock waves originated, the amount of time between when shock waves occurred on the segment, the segment length, and the sensor spacing.

In addition, a methodology to optimize the sensor spacing was developed. An objective weighted average travel time prediction error function was formulated and considered that incidents with different durations, v/c ratios, and locations affect the study freeway segment. In the weighted average error function, the error per incident was weighted by the probability of an incident occurring with its duration and v/c ratio and at its location. To determine the error per incident, the mathematical model developed in Chapter 3 was applied.

A real world example of a freeway segment in northern New Jersey was introduced to demonstrate the applicability of the developed model and solution algorithm to optimize the studied problem. A historical incident database provided the distribution of incident durations and volume to capacity ratios. Also, the marginal benefit of increasing the number of sensors on the freeway segment was calculated. A sensitivity analysis was conducted for investigating the relationship between model parameters and the optimal sensor spacing.

6.1 Conclusions

The optimal path based sensor spacing for a freeway segment can be found by applying the presented methodology. The major findings and conclusions were as follows:

- The travel time prediction error varied by the sensor spacing, incident duration, v/c ratio, and location. Therefore, sensor spacing cannot be optimized by only considering one historic incident for the segment.

- Holding the sensor spacing and incident duration constant, the average error increased as the v/c ratio increased. As the v/c ratio increased the backward shock wave speed significantly increased. Compared to a slower shock wave speed, a faster shock wave speed meant the travel times by departure times were changing more rapidly. Thus, there is a greater difference between the predicted travel times provided to the departing vehicles and their actual travel times.

- Long sensor spacing is preferable when the spread of congestion over space and time is insignificant, such as for incidents with low v/c ratios and short durations. Short sensor spacing is preferable when the spread of congestion is large, such as for incidents with high v/c ratios and long durations.

- The marginal benefit decreases as the number of sensors increases. The benefit is significant when comparing the average error for a small number of sensors. However, the improvement in the error diminishes towards zero with each additional sensor. Therefore, increasing the number of sensors, which meant decreasing the sensor spacing, improved the average error. This was illustrated for the case study in Figure 5.9.

- The number of classes chosen for the incident duration frequency distribution and the v/c ratio distribution affects the outputs obtained from the solution algorithm. Too few classes can underestimate the average prediction error but too many classes can increase the amount of time to apply the methodology while not significantly improve the estimation. Therefore, the number of classes should be chosen carefully when applying the methodology.

6.2 Future Research

Future research areas related to the optimal sensor spacing problem are listed below:

- The model can be enhanced to consider the prove vehicle sample size on the freeway segment and also to include incident scenarios that have intermediate traffic demand fluxes and/or partial capacity restoration prior to full capacity restoration.

- The model can also be improved upon by including as an input variable the amount of time between when travel time predictions are updated for roadway users, such as on a DMS.

- The developed methodology can be improved by considering more than one origin-destination for which travel time predictions are provided to drivers. This can be achieved by modeling shock wave behavior across multiple freeway segments and also optimizing the sensor spacing considering the demand matrix.

- As the incident location affects the accuracy of the travel time predictions by sensor spacing, activating or deactivating some of the sensors on the segment during an incident can improve the travel time predictions. In addition, if an incident lasts for more than one time period, the optimal spacing may change. Guidelines can be developed to determine when and how to choose which sensors should and should not be active during an incident.

- The model, methodology, and results can contribute to and be extended, but not limited to, several applications of the USDOT's Intelligent Network Flow Optimization research (i.e., Queue Warning, Dynamic Speed Harmonization, and Cooperative Adaptive Cruise Control). The objectives of these projects are to maximize roadway system productivity, enhance roadway safety and capacity, and reduce overall fuel consumption (Mahmassani et al. 2012).

APPENDIX A

JAVA APPLICATION

In this Appendix, a code is provided. It is for an application that was developed in Java that automates the mathematical model's calculations and step procedures provided in Chapter 3. The application calculates the average travel time prediction error for a study period with a backward shock wave and backward recovery wave. Inputs to the model include the segment characteristics, the different incident scenarios, durations, and locations to consider, and the different numbers of sensors to test.

```
public class TTPEModelPDiJoseph {
	public static void main(String[] args)
	{
		//arrays runs many v/c ratios iteratively; this is designed
for 8 v/c ratios, change number accordingly
		//variables currentVC through current QC all go with one
v/c ratio
		//each array number goes for one v/c ratio, for ex, all
that have [0] go with 1 v/c ratio, all that have [1] go with another
v/c ratio, etc
		double[] currentVC = new double[8];  //veh congested speed
		double[] currentVU1 = new double[8];  //veh uncongested
speed
		double[] currentVWbs = new double[8];  //backward shock
wave speed
		double[] currentVWbr = new double[8]; //backward recovery
wave speed
		double[] currentQU = new double[8];  //veh uncongested
traffic flow
		double[] currentQC = new double[8];  //veh congested
traffic flow
		//end of arrays set up   and starting below is the first
v/c ratio info
		currentVC[0] = 7.90;
		currentVU1[0] = 66.80;
		currentVWbs[0] = -2.60;
		currentVWbr[0] = -17.00;
		currentQU[0] = 3075.00;
		currentQC[0] = 2396.00;
		//second v/c ratio info
		currentVC[1] = 7.80;
		currentVU1[1] = 66.40;
		currentVWbs[1] = -3.60;
		currentVWbr[1] = -16.60;
		currentQU[1] = 3480.00;
		currentQC[1] = 2404.00;
		//third v/c ratio info
		currentVC[2] = 7.70;
		currentVU1[2] = 65.60;
		currentVWbs[2] = -5.60;
		currentVWbr[2] = -15.80;
		currentQU[2] = 4216.00;
		currentQC[2] = 2417.00;
		//fourth v/c ratio info
		currentVC[3] = 7.50;
		currentVU1[3] = 64.70;
		currentVWbs[3] = -8.00;
		currentVWbr[3] = -14.80;
```

```
            currentQU[3] = 5121.00;
            currentQC[3] = 2434.00;
            // fifth v/c ratio info
            currentVC[4] = 7.00;
            currentVU1[4] = 62.90;
            currentVWbs[4] = -13.80;
            currentVWbr[4] = -14.10;
            currentQU[4] = 6449.00;
            currentQC[4] = 2646.00;
            // sixth v/c ratio info
            currentVC[5] = 0;
            currentVU1[5] = 0;
            currentVWbs[5] = 0;
            currentVWbr[5] = 0;
            currentQU[5] = 0;
            currentQC[5] = 0;
            // seventh v/c ratio
            currentVC[6] = 0;
            currentVU1[6] = 0;
            currentVWbs[6] = 0;
            currentVWbr[6] = 0;
            currentQU[6] = 0;
            currentQC[6] = 0;
            // eighth v/c ratio
            currentVC[7] = 0;
            currentVU1[7] = 0;
            currentVWbs[7] = 0;
            currentVWbr[7] = 0;
            currentQU[7] = 0;
            currentQC[7] = 0;  //end of different v/c ratios that will
be tested
            double[] currentDur = new double[8];  // runs many incident
durations iteratively// different durations that will be tested
            currentDur[0]=0.32;
            currentDur[1]=0.92;
            currentDur[2]=1.59;
            currentDur[3]=2.25;
            currentDur[4]=2.81;
            currentDur[5]=3.52;
            currentDur[6]=4.15;
            currentDur[7]=4.76;  // end of different durations that
will be tested
      // end of 8 classes */

            double[] currentXW = new double[10];  // runs many incident
locations iteratively// different locations that will be tested
            currentXW[0]=1.00;
            currentXW[1]=2.00;
            currentXW[2]=3.00;
            currentXW[3]=4.00;
            currentXW[4]=5.00;
            currentXW[5]=6.00;
            currentXW[6]=7.00;
            currentXW[7]=8.00;
            currentXW[8]=9.00;
```

```
            currentXW[9]=10.00;  // end of different incident locations
that will be tested
            int[] number = new int[40];  // runs many numbers of
sensors iteratively
            number[0]=2;
            number[1]=3;
            number[2]=4;
            number[3]=5;
            number[4]=6;
            number[5]=7;
            number[6]=8;
            number[7]=9;
            number[8]=10;
            number[9]=11;
            number[10]=12;
            number[11]=13;
            number[12]=14;
            number[13]=15;
            number[14]=16;
            number[15]=17;
            number[16]=18;
            number[17]=19;
            number[18]=20;
            number[19]=21;
            number[20]=22;
            number[21]=23;
            number[22]=24;
            number[23]=25;
            number[24]=26;
            number[25]=27;
            number[26]=28;
            number[27]=29;
            number[28]=30;
            number[29]=31;
            number[30]=32;
            number[31]=33;
            number[32]=34;
            number[33]=35;
            number[34]=36;
            number[35]=37;
            number[36]=38;
            number[37]=39;
            number[38]=40;
            number[39]=41;



            int p; // counter for number of sensors
            int pp; // counter for incident duration
            int v; //counter for v/c ratio
            int xx;  // counter for incident location
            for (v =0; v<=4; v++ ) // loop on v/c ratio
            {
                  System.out.println("V/C RATIO GROUP = "+v);
```

```
                  for (xx =0; xx<=9; xx++ ) // loop on incident
LOCATION
                  {
                        System.out.println("INCIDENT LOCATION =
"+currentXW[xx]);
                        for (pp =0; pp<=7; pp++ ) // loop on incident
duration
                        {
                              System.out.println("INCIDENT DURATION =
"+currentDur[pp]);
                              for (p =0; p<=39; p++ ) // loop on number
of sensors
                                    {
            double L=10.00;  // length of segment (mi)
            double vC=currentVC[v]; // average vehicle speed under
congested conditions (mi/hr)
            double vU1=currentVU1[v]; // average vehicle speed under
normal (uncongested conditions)  (mi/hr)
            double vU2=vU1; // must = vU2 otherwise, fix programming.
some eqs have already been corrected avg veh speed under normal
(uncongested conditions) after recovery wave (mi/hr)
            double vWbs=currentVWbs[v];  // shock wave speed for
backward shock wave; (mi/hr)
            double tWbs=1.00;  // time backward shock wave initiates in
the freeway segment (hrs after time 0)
            double xWbs=currentXW[xx];  // location waves initiate,
equals L if waves initiate downstream of the segment (mi)
            double vWbr=currentVWbr[v];  // shock wave speed for
backward recovery wave; (mi/hr)
            double incidentDuration=currentDur[pp]; // incident
duration in hrs
            double tWbr=tWbs+incidentDuration; // time backward
recovery wave initiates in the freeway segment (hrs after time 0)
            double qU = currentQU[v]; // uncongested traffic flow
(veh/hr)
            double qC = currentQC[v]; // congested traffic flow
(veh/hr)
            double VHT; // total travel time error in VHT
            double xWbr=xWbs;  // location of BRW at time tWbr (mi)
            int N;  // number of important vehicles affected by a shock
wave
            int actN; // number of important vehicles b/c of their
actual travel time
            int estN; // number of important vehicles b/c of their
estimated travel time
            int S = number[p]; // number of sensors
            double[] xS = new double[S];  // location of sensors
            int i,j,k,both;  // indexes, both is calculated further
down, equals 0 if no vehicles will encounter both shock waves, equals 1
if at least one does
            double s =L/(S-1); // sensor spacing
            double A, Abs, AtC, AdC, AtU1, C; // starting departure
time for vehs that hit both shock waves   and variables used for
intermediate calculations
            double tVbs; // time when veh hits bsw (hrs after 0)
```

```
            double xVbs; // location where veh hits bsw (miles into
segment)
            double tVbr; // time when veh hits brw (hrs after 0)
            double xVbr; // location where veh hits brw (miles into
segment)
            double tW;  // time waves intersect each other
            double xW;  // location at which waves intersect each other
            double xWin; // waves meet within segment if yes, value =1,
else = 0
            double[] eWin = new double[S-1];  // if waves meet within
segment
            double tDeptE; // used in calculating Ec[j]
            double xDeptE; // used in calculating Ec[j]
            double negArea = 0.00;  // used to find total amount of
negative error
            double posArea = 0.00;  // used to find total amt of pos
error, if waves meet within the segment then all q entering segment is
qu
            double startQc=tWbs+xWbs/-vWbs; // beginning of time period
when Q is qc  used in total error calcs
            double endQc=tWbr+xWbr/-vWbr;  // end of time period when Q
is qc
            double posAreaQu = 0.00; // to find total amt of pos error
with q = qu
            double posAreaQc = 0.00;  // to find total amt of pos error
with q = qc
            double negAreaQu = 0.00; // to find total amt of pos error
with q = qu
            double negAreaQc = 0.00;  // to find total amt of pos error
with q = qc
            double avgNeg;
            double avgPos;
            double avgNegQu;
            double avgNegQc;
            double avgPosQu;
            double avgPosQc;
            double avgErrAbs;
            double durAvgNeg;
            double durAvgPos;
            double durNegQu;
            double durNegQc;
            double durPosQu;
            double durPosQc;
            double vhtNeg;
            double vhtPos;
            double vhtNegQu;
            double vhtNegQc;
            double vhtPosQu;
            double vhtPosQc;
            double vhtAbs;
            double end;  //end of study period
            double pdNoError = 0;  // 1 if have a pd of 0 error
            double ptNoError = 0;  // 1 if just have 1 pt of 0 error
            double endPdNoError = 0.00;
            double noErrorAreaQu = 0.00;
```

```
            double noErrorAreaQc = 0.00;
            double durNoErrorQc=0.00;
            double durNoErrorQu=0.00;

            double noFakeIntercept=0;

            //notations for vehicles that first encounter a backward
shock and then backward recovery

            double tU1; // actual veh tt for a veh affected by a bsw,
tt till hits first wave // tf1
            double dU1;  // distance vehicle travels in uncongested 1
conditions (hr)
            double tC;   // amount of time vehicle travels in congested
conditions (hr)
            double dC;   // distance vehicle travels in congested
conditions (mi)
            double tU2;  // amt of time veh travels in uncongested 2
conditions (hr)
            double dU2;  // distance vehicle travels in uncongested 2
conditions (hr)

            double B = tWbs - xWbs/vU1;
            double D = tWbr + xWbr/(-vWbr);

            double[] eA = new double[S-1];
            double[] eB = new double[S-1];
            double[] eC = new double[S-1];
            double[] eD = new double[S-1];
            double[] eBoth = new double[S-1];  // veh hit both shock
waves together in segment
            double[] tUS = new double[S-1]; // amt of time a veh
travels in uncongested conditions from an incident location downstream
to sensor j, applies when eWin[j] = 2 or -1
            int count = 0;
            int counter = 0;
            int change=0;
            double intercept=0; // 0 value to initiate, value
calculated at the end
            int M=0;



            for (j=0; j<=S-1; j++)
            {
                  xS[j]=L-j*s;
            }

            // if bsw won't meet brw then need to change below
            tW = (xWbr-vWbr*tWbr-xWbs+vWbs*tWbs)/(vWbs-vWbr); // time
bsw meets brw so assumes they will meet :-)
            xW = vWbs*tW+xWbs-vWbs*tWbs;  // location bsw meets brw,
will be negative if outside segment :-)
```

```
      if (xW>=0)  // if the waves meet within the roadway segment do
the following calculations
      {
            xWin =1;  // denotes the waves meet within the roadway
segment
            actN=3; // number of important times to do act t.t.
calculations for: first to hit BSW, first to hit BRW, veh hits where
waves meet
      //    end = tW+(L-xW)/vU1; // time veh that hits where waves meet
reaches downstream end
            for (j=0; j<=S-2; j++)  // sensor j = 0 farthest downstream
            {

            if((xWbs<(S-1-j)*s)&&(xWbs>(S-2-j)*s)&&(xW<(S-1-
j)*s)&&(xW>=(S-2-j)*s))  // shock waves initiate in the spacing between
sensor j and sensor j+1

            // and shock waves meet within sensors j and j+1
            {
                  eWin[j]=-1;
                  tUS[j]=(xS[j]-xWbs)/vU1;
            }
            else if((xWbs<(S-1-j)*s)&&(xWbs>(S-2-j)*s))  // shock waves
initiate in the spacing between sensor j and sensor j+1
            {
                  eWin[j]=2;
                  tUS[j]=(xS[j]-xWbs)/vU1;

            }
            else if (xWbs<=(S-2-j)*s)  // shock waves initiate upstream
of the spacing between sensor j and sensor j+1
            {
                  eWin[j]=3;
                  count = count+1;
            }
            else if (xW>=(S-1-j)*s)  // shock waves never hit sensor j
bc they meet downstream of sensor j
            {
                  eWin[j]=3;
                  count = count+1;
            }
            else if ((xW<(S-1-j)*s)&&(xW>=(S-2-j)*s))  // shock wave
meeting pt within sensor j and j+1
            {
                  eWin[j]=1;
            }

            else  // shock waves meet below sensor j and j+1
            {
                  eWin[j]=0;
            }
                  }

            estN=4*(S-1-count)-1;   // number of important times to do
est. t.t. calculations for
```

```
	}

	else
	{
		xWin =0;  // denotes the waves meet not in the roadway
segment
		actN=4;   // number of important times to do act t.t.
calculations for

		for (j=0; j<=S-2; j++)  // sensor j = 0 farthest downstream
		{

		if((xWbs<(S-1-j)*s)&&(xWbs>(S-2-j)*s))  // shock waves
initiate in the spacing between sensor j and sensor j+1
		{
			eWin[j]=2;
			tUS[j]=(xS[j]-xWbs)/vU1;

		}
		else if (xWbs<=(S-2-j)*s)  // shock waves initiate upstream
of the spacing between sensor j and sensor j+1
		{
			eWin[j]=3;
			count = count+1;
		}
		else if (xW>=(S-1-j)*s)  // shock waves never hit sensor j
bc they meet downstream of sensor j   **
		{
			eWin[j]=3;
			count = count+1;

		}
		else  // shock waves meet below sensor j and j+1
		{
			eWin[j]=0;
		}
		}
		estN=4*(S-1-count);   // number of important times to do
est. t.t. calculations for

	}

		N = actN + estN;


			// notations for vehicles for estimation calculation
			double[][] Te= new double[N][S-1]; //  est travel
time for vehicles to calculate their actual travel time
			double[] tArr = new double[N];  // time vehicle i is
at sensor j
			double[] TTe = new double[N];  // total travel time
estimate for veh
			double[] error = new double[N];  // error for each
point
```

```
                  double[] relError = new double[N];  // relative error
for each point  (error)/act travel time
                  double errorAvg = 0;  // average error over the
incident duration
                  double[] xDept = new double[N]; // dept location for
vehs to calculate their actual travel time
                  double[] tDept = new double[N]; // dept location for
vehs to calculate their actual travel time
                  double[] Ta= new double[N]; //  actual travel time
for vehicles to calculate their actual travel time



            // dept time of first vehicle to encounter both waves
            if (tWbr-xWbr/vC>=tWbs+xWbs/(-vWbs))
            {
                  both=0;   // no vehicles will encounter both shock
waves
                  A = tWbs + xWbs/(-vWbs);
                  C = tWbr - xWbr/vC;

            }
            else
            {
                  both=1;
                  if (xWin<1)
                  {
                  C = tWbs + xWbs/(-vWbs);
                  }
                  else
                  {
                  C = tW - xW/vU2;
                  }
                  Abs=(xWbr-tWbr*vC-xWbs+vWbs*tWbs)/(vWbs-vC);
                  AtC=tWbr-Abs;
                  AdC=AtC*vC;
                  AtU1=(xWbr-AdC)/vU1;
                  A = tWbr-AtC-AtU1;
            }

            for (j=0; j<=S-2; j++)
            {
                  if (eWin[j]==0)
                  {
                        if(tWbr+(xWbr-(L-s*j))/(-vWbr)-
s/vC>=tWbs+(xWbs-(L-(j+1)*s))/(-vWbs))
                        {
                              eBoth[j]=0;
                        }
                        else
                        {
                              eBoth[j]=1;
                        }
                  }
                  if (eWin[j]==2)
```

```
                {
                      if(tWbr-(xWbr-(L-s*(j+1)))/vC>=tWbs+(xWbs-(L-
(j+1)*s))/(-vWbs))
                      {
                            eBoth[j]=0;
                      }
                      else
                      {
                            eBoth[j]=1;
                      }
                }
                if (eWin[j]==1||eWin[j]==-1)
                {
                      eBoth[j]=1;
                }
          }



          double[] U = new double[N];   //unsorted important veh dept
times
          U[0]=tWbs-xWbs/vU1;  // dept time of first veh to encounter
bsw
          tVbs=(xWbs-tWbr*vC-xWbs+vWbs*tWbs)/(vWbs-vC);  //
calculations for U[1]
          tC=tWbr-tVbs; // calculations for U[1]
          dC=tC*vC; // calculations for U[1]
          tU1=(xWbs-dC)/vU1; // calculations for U[1]

          if (xW>=0)   // waves meet in segment
          {
                U[1]=tWbr-(tC+tU1);
                U[2]=tW-xW/vU1;
                counter = 3;
          }

          else  // important dept  times (for actual) when shock
waves don't meet in segment
          {
          if (both==1)   // waves don't meet in segment but vehs
still hit both
          {
                U[1]=tWbr-(tC+tU1);
                U[2]=tWbs+xWbs/-vWbs;         // first veh to dept
that doesn't encounter bsw
                U[3]=tWbr+xWbr/-vWbr;         // first veh to dept
that doesn't enounter brw (or bsw)
                counter = 4;
          }
          else  // if both ==0 so no vehs encounter both shock waves
          {
                U[1]= tWbs+xWbs/-vWbs; // first veh in all congested
                U[2]=tWbr-xWbr/vC;            // last veh in all
congested
```

```
                  U[3]=tWbr+xWbr/-vWbr;         // first veh to dept
that doesn't enounter brw (or bsw)
                  counter = 4;
            }
            }
            for (j=0; j<=S-2; j++)
            {
                  if (eWin[j]==0)  // shock waves travel through, but
meet below, sensors j and j+1
                  {

                        eB[j]= tWbs + (xWbs-(L-j*s))/(-vWbs);
                        U[counter]=eB[j];
                        counter = counter + 1;

                        eD[j]= tWbr + (xWbr-(L-(j+1)*s))/(-vWbr) +
s/vU2;
                        U[counter]=eD[j];
                        counter = counter + 1;

                        if (eBoth[j]<1)  // vehs won't hit both waves
in segment
                        {
                              eA[j]= tWbs + (xWbs-(L-(j+1)*s))/(-
vWbs)+s/vC;
                              U[counter]=eA[j];
                              counter = counter + 1;

                              eC[j]= tWbr + (xWbr-(L-s*j))/(-vWbr);
                              U[counter]=eC[j];
                              counter = counter + 1;
                        }

                        else
                        {
                              eA[j]= tWbr + (xWbr-(L-s*j))/(-vWbr);
                              U[counter]=eA[j];
                              counter = counter + 1;

                              tDeptE=tWbs+(xWbs-(L-(j+1)*s))/(-vWbs);
                              xDeptE=L-(j+1)*s;
                              tC=(xWbr-tWbr*vWbr-xDeptE+tDeptE*vC)/(vC-
vWbr)-tDeptE;
                              dC= vC*tC;
                              dU2=s-dC;
                              tU2=dU2/vU2;
                              eC[j]=tDeptE+tC+tU2;
                              U[counter]=eC[j];
                              counter = counter + 1;
                        }

                  }

                  if (eWin[j]==1) // waves originate downstream of, but
meet between, sensors j and j+1
```

```
                  {
                        eB[j]= tWbs + (xWbs-(L-s*j))/(-vWbs);
                        U[counter]=eB[j];
                        counter = counter + 1;

                        eA[j]= tWbr + (xWbr-(L-s*j))/(-vWbr);
                        U[counter]=eA[j];
                        counter = counter + 1;

                        eC[j]=tW + (L-s*j-xW)/vU1;
                        U[counter]=eC[j];
                        counter = counter + 1;
                  }

                  if (eWin[j]==-1)  //waves originate, and meet,
between sensors j and j+1
                  {
                        eB[j]= tWbs + (L-s*j-xWbs)/vU1;
                        U[counter]=eB[j];
                        counter = counter + 1;

                        eA[j]= tWbr + (L-s*j-xWbr)/vU1;
                        U[counter]=eA[j];
                        counter = counter + 1;

                        eC[j]=tW + (L-s*j-xW)/vU1;
                        U[counter]=eC[j];
                        counter = counter + 1;
                  }

                  if (eWin[j]==2)  // shock waves initiate within
sensors j and j+1
                  {
                        eB[j]= tWbs + ((L-j*s)-xWbs)/(vU1);
                        U[counter]=eB[j];
                        counter = counter + 1;
                        eD[j]= tWbr + (xWbr-(L-(j+1)*s))/(-vWbr) +
s/vU2;
                        U[counter]=eD[j];
                        counter = counter + 1;

                        if (eBoth[j]<1)  // vehs won't hit both waves
in segment
                        {
                              eA[j]= tWbs + (xWbs-(L-(j+1)*s))/(-
vWbs)+(xWbs-(L-(j+1)*s))/vC + (L-j*s-xWbs)/vU1;
                              U[counter]=eA[j];
                              counter = counter + 1;

                              eC[j]= tWbr + ((L-s*j)-xWbr)/vU1;
                              U[counter]=eC[j];
                              counter = counter + 1;
                        }
                        else  // some vehs do hit both waves in segment
                        {
```

```
                              eA[j]= tWbr + ((L-s*j)-xWbr)/vU1;
                              U[counter]=eA[j];
                              counter = counter + 1;
                              tDeptE=tWbs+(xWbs-(L-(j+1)*s))/(-vWbs);
                              xDeptE=L-(j+1)*s;
                              tC=(xWbr-tWbr*vWbr-xDeptE+tDeptE*vC)/(vC-
vWbr)-tDeptE;
                              dC= vC*tC;
                              dU2=s-dC;
                              tU2=dU2/vU2;
                              eC[j]=tDeptE+tC+tU2;
                              U[counter]=eC[j];
                              counter = counter + 1;
                        }
                  }

            }

            for (i=0; i<= counter-2; i++)  // start of section to omit
duplicated U[]'s
            {
                  for (j=i+1; j<=counter-1; j++)
                  {
                        if (U[i]==U[j])
                        {
                              for (k=i; k<=counter-2;k++)
                              {
                                    U[k]=U[k+1];
                              }
                              counter=counter-1;
                              j=counter;
                              i=i-1;
                              N=N-1;
                        }
                  }
            }  // end of section to omit duplicated U[]'s



            for (i=0; i<=counter-1; i++)   //puts dept times in order
from smallest to largest
            {
                  tDept[i]=10000.00;
            }
            for (i=0; i<=counter-1; i++)
            {
                  for (j=0; j<=counter-1; j++)
                  {
                        if (U[j]<=tDept[i])
                        {
                              tDept[i]=U[j];
                        }

                  }
                  for (j=0; j<=counter-1; j++)
```

```
                {
                      if (U[j]==tDept[i])
                      {
                            U[j]=10000+j;
                      }
                }
          }   //ends puts dept times in order from smallest to
largest
          for (i=0; i<=counter-1; i++)
          {

//          System.out.println("tDp["+ tDept[i]); //  prints out
departure time for each important veh
//                System.out.println(tDept[i]);
          }

          end=tDept[counter-1];  // end of pd to study error?



          for (i=0; i<=N-1; i++)
          {
          tArr[i]=tDept[i];
          }



          //actual travel time calculations
          for (i=0; i<=N-1; i++)
          {
                // TT for vehs only traveling in vU1
                if (tDept[i]<=B)
                {
                      Ta[i]=L/vU1;
                      }

                //travel time for vehs encountering only a bsw
                else if (tDept[i]<=A)
                {
                      tU1=(xWbs-tWbs*vWbs-
xDept[i]+tDept[i]*vU1)/(vU1-vWbs)-tDept[i];
                      dU1= vU1*tU1;
                      dC=xWbs-dU1;
                      tC=dC/vC;
                      Ta[i]=tU1+tC+(L-xWbs)/vU1;
                }


                // travel time for vehs encountering bsw & brw
                else if((tDept[i]<C)&&(both > 0))
                {
                      tVbs=(xWbs-tWbs*vWbs-
xDept[i]+tDept[i]*vU1)/(vU1-vWbs);
                      tU1= tVbs-tDept[i];
                      dU1= vU1*tU1;
```

```
                        tVbr=(xWbr-tWbr*vWbr-xDept[i]-dU1+tVbs*vC)/(vC-
vWbr);
                        tC=tVbr-tVbs;
                        dC= vC*tC;
                        dU2=L-dU1-dC;
                        tU2= dU2/vU2;
                        Ta[i]=tU1+tC+tU2;
                        }

                  else if((tDept[i]<C)&&(both < 1))
                  {
                        Ta[i]=xWbs/vC+(L-xWbs)/vU1;
                        }

                  //travel time for vehs encountering a brw
                  else if (tDept[i]<D && xWin<1)  //correct symbols
                  {
                        tC=(xWbr-tWbr*vWbr-xDept[i]+tDept[i]*vC)/(vC-
vWbr)-tDept[i];
                        dC= vC*tC;
                        dU2=L-dC;
                        tU2=dU2/vU2;
                        Ta[i]=tC+tU2;
                        }


                  else
                  {
                        Ta[i]=L/vU1;
                              }
//          System.out.println("Ta["+i+"] = "+Ta[i]); // prints out
actual travel time for each important veh
            }// end of actual travel time calculations

            //begining of estimated travel time section
            for (i=0; i<=N-1; i++)
            {
                  for (j=0; j<=S-2; j++)
                  {
                        if (eWin[j]==3)
                        {
                              Te[i][j]=s/vU1;
                        }

                        else if (tArr[i]<=eB[j])
                        {
                              Te[i][j]=s/vU1;
                        }

                        else if (tArr[i]<=eA[j])  // hit BSW
                        {
                              if(eWin[j]==0||eWin[j]==1)  // "||"
should mean "or"
                              {
```

```
                              tVbs=(xS[j]-tArr[i]*vC-
xWbs+vWbs*tWbs)/(vWbs-vC);
                              tC=tArr[i]-tVbs;
                              dC=tC*vC;
                              tU1=(s-dC)/vU1;
                              Te[i][j]=tC+tU1;
                              }

                              else  // for eWin[j]== -1 or 2
                              {
                                    tVbs=(xWbs-(tArr[i]-tUS[j])*vC-
xWbs+vWbs*tWbs)/(vWbs-vC);
                                    tC=(tArr[i]-tUS[j])-tVbs;
                                    dC=tC*vC;
                                    tU1=(xWbs-dC-xS[j+1])/vU1;
                                    Te[i][j]=tC+tU1+tUS[j];
                              }
                        }

                        // travel time for vehs encountering bsw & brw
                        else if((tArr[i]<eC[j])&&(eBoth[j] > 0))
                        {
                              if(eWin[j]==0||eWin[j]==1)
                              {
                                    tVbr=(xWbr-tWbr*vWbr-
xS[j]+tArr[i]*vU2)/(vU2-vWbr);
                                    tU2=tArr[i]-tVbr;
                                    dU2=tU2*vU2;
                                    xVbr=xS[j]-dU2;
                                    tVbs=(xVbr-vC*tVbr-
xWbs+vWbs*tWbs)/(vWbs-vC);
                                    tC=tVbr-tVbs;
                                    dC= vC*tC;
                                    dU1=s-dU2-dC;
                                    tU1=dU1/vU1;
                        //          Te[i][j]=tU1+tC+tU2;
                                    Te[i][j]=tC+(s-dC)/vU1;
                              }
                              else
                              {
                                    tVbr=(xWbr-tWbr*vWbr-xWbr+(tArr[i]-
tUS[j])*vU2)/(vU2-vWbr);

                                    tU2=(tArr[i]-tUS[j])-tVbr;
                                    dU2=tU2*vU2;
                                    xVbr=xWbs-dU2;
                                    tVbs=(xVbr-vC*tVbr-
xWbs+vWbs*tWbs)/(vWbs-vC);
                                    tC=tVbr-tVbs;
                                    dC= vC*tC;
                                    dU1=xWbs-dU2-dC;  //here is wrong!!
                                    tU1=dU1/vU1;
                              //    Te[i][j]=tU1+tC+tU2+tUS[j];
                                    Te[i][j]=tC+(s-dC)/vU1;
                              }
```

```
                        }

                        // travel time for vehs traveling in all
congested
                        else if((tArr[i]<eC[j])&&(eBoth[j] < 1))
//wrong!!
                        {
                              if(eWin[j]==0||eWin[j]==1)
                              {
                                    Te[i][j]=s/vC;
                              }
                              else
                              {
                                    Te[i][j]=tUS[j]+(xWbs-xS[j+1])/vC;
                              }
                        }

                        //travel time for vehs encountering a brw
                        else if ((tArr[i]<eD[j]) &&
(eWin[j]==0||eWin[j]==2))
                        {
                              if(eWin[j]==0)
                              {
                                    tVbr=(xWbr-tWbr*vWbr-
xS[j]+tArr[i]*vU2)/(vU2-vWbr);
                                    tU2=tArr[i]-tVbr;
                                    dU2=tU2*vU2;
                                    tC=(s-dU2)/vC;
                                    Te[i][j]=tU2+tC;
                              }
                              else  // eWin[j]==2  here:)
                              {
                                    tVbr=(xWbr-tWbr*vWbr-xWbr+(tArr[i]-
tUS[j])*vU2)/(vU2-vWbr);
                                    tU2=(tArr[i]-tUS[j])-tVbr;
                                    dU2=tU2*vU2;
                                    tC=(xWbs-dU2-xS[j+1])/vC;
                                    Te[i][j]=tU2+tC+tUS[j];
                              }
                        }

                        else  // *:-)*for vehs that travel after waves
are done, shock wave theory suggests to use VU1  NEED TO FIX STILL
                        {
                              Te[i][j]=s/vU1;
                        }
                              TTe[i]=TTe[i]+Te[i][j]; // sums the
t.t.e. provided by all the sensors at time tArr[i]
                  }


            }
            //end of estimating travel time section
```

```
            for (i=0; i<=N-1; i++)
            {
//          System.out.println("TTe["+i+"]="+TTe[i]); // prints out
travel time estimates for each important veh
            }


            for (i=0; i<=counter-1; i++)
            {
                  error[i]=TTe[i]-Ta[i];           // use for reg error
      //          error[i]=(TTe[i]-Ta[i])/Ta[i];  //  use for relative
error
                  if((error[i]<=.000001) && (error[i]>=0))
                  {
                        error[i]=0.00;
                  }
                  else if ((error[i]>=-.000001) && (error[i]<=0))
                  {
                        error[i]=0.00;
                  }
                  else
                  {
                  }
//                System.out.println("error["+i+"] = "+error[i]); //
prints out travel time error for each important veh
            }

            int shape = 0;

            for (i=1; i<=counter-2; i++)
            {
                  if(Math.abs(error[i])==0.00)
                  {
                        intercept=tDept[i];
                  noFakeIntercept=noFakeIntercept+1;
                  shape=counter-1;
                  if(Math.abs(error[i+1])==0.00)
                  {
                        pdNoError = pdNoError+1;
                        for (j=i+1; j<=counter-2; j++)
                        {
                              if (Math.abs(error[j])==0.00)
                              {
                                    endPdNoError=tDept[j];
                              }
                              else
                              {
                                    j=counter-1;
                                    i=counter-1;
                              }
                        }
                  }
                  else
                  {
                        ptNoError=ptNoError+1;
```

```
                        i=counter-1;
                  }


                  }
                  else
                  {
                  }
            }
            if (noFakeIntercept<1) // so do no have any periods of 0
error
            {
                  for (i=0; i<=counter-2; i++)
                  {
                        if (error[i]*error[i+1]<0)
                        {
                              intercept = -error[i]/((error[i+1]-
error[i])/(tDept[i+1]-tDept[i]))+tDept[i]; // point where changes from
neg error to pos error
                        i=counter-1;
                  shape=counter;
                        }
                  }
            }
            else
            {

            }


            if(xWin==1)//  Qc never occurs because the waves intersect
within the segment
            {
                  startQc=end;
                  endQc=end;
            }
            double[] area = new double[shape];

            if (noFakeIntercept>0) //so have period of 0 error
            {
                        for (i=0; i<=counter -2; i++)
                  {
                        area[i]=.5*(tDept[i+1]-
tDept[i])*(error[i+1]+error[i]);

                        if
((tDept[i]>=intercept)&&(tDept[i+1]<=endPdNoError))
                        {
                              if(tDept[i+1]<=startQc)
                              {
                                    System.out.println("fix code, there
is a pd of no error Qu");
                              }
                              else
                              {
```

```
      noErrorAreaQc=noErrorAreaQc+Math.abs(area[i]);
                              }
                        }
                        else if(tDept[i+1]<=intercept)
                        {
                              if(tDept[i+1]<=startQc)
                              {

      negAreaQu=negAreaQu+Math.abs(area[i]);
                              }
                              else
                              {

      negAreaQc=negAreaQc+Math.abs(area[i]);
                              }
                        }

                        else if(tDept[i+1]<=startQc)  //so also greater
than intercept
                        {
                              posAreaQu=posAreaQu+area[i];
                        }
                        else if(tDept[i+1]<=endQc)
                        {
                              posAreaQc=posAreaQc+area[i];
                        }
                        else
                        {
                              posAreaQu=posAreaQu+area[i];
                        }
                  }
            }
            else
            {
            for (i=0; i<=counter -1; i++)
            {
                  if (error[i-M]*error[i+1-M]>=0)  // if don't cross
y=0 to go from one point to the next
                  {
                  area[i]=.5*(tDept[i+1-M]-tDept[i-M])*(error[i+1-
M]+error[i-M]);

                        if(tDept[i+1-M]<=intercept)
                        {
                              if(tDept[i+1-M]<=startQc)
                              {

      negAreaQu=negAreaQu+Math.abs(area[i]);
                              }
                              else
                              {

      negAreaQc=negAreaQc+Math.abs(area[i]);
                              }
```

```
                        }

                        else if(tDept[i+1-M]<=startQc)  //so also
greater than intercept
                        {
                              posAreaQu=posAreaQu+area[i];
                        }
                        else if(tDept[i+1-M]<=endQc)
                        {
                              posAreaQc=posAreaQc+area[i];
                        }
                        else
                        {
                              posAreaQu=posAreaQu+area[i];
                        }
                  }
                  else   // if do cross y=0 so then have to add in the
point (intercept,0)
                  {

                        area[i]=.5*error[i-M]*(intercept-tDept[i-M]);
                              if(tDept[i+1-M]<=startQc)
                              {

      negAreaQu=negAreaQu+Math.abs(area[i]);
                              }
                              else
                              {

      negAreaQc=negAreaQc+Math.abs(area[i]);
                              }
                        i=i+1;
                        area[i]=.5*error[i-M]*(tDept[i-M]-intercept);

                              if(tDept[i+1-M]<=startQc)  //so also
greater than intercept
                              {
                                    posAreaQu=posAreaQu+area[i];
                              }
                              else if(tDept[i+1-M]<=endQc)
                              {
                                    posAreaQc=posAreaQc+area[i];
                              }
                              else
                              {
                                    posAreaQu=posAreaQu+area[i];
                              }
                   M=M+1;  // so the index of the areas is okay with
the index of the points
                  }
            }
            }

            if (noFakeIntercept>0)  //so have period of 0 error
            {
```

```
                  if(startQc<=intercept)
                  {
                        durNegQu=startQc-B;
                        durNegQc=intercept-startQc;
                        if (pdNoError>0)
                        {
                              durNoErrorQc=endPdNoError-intercept;
                              durPosQc=endQc-endPdNoError;
                        }
                        else
                        {
                              durPosQc=endQc-(tWbr-xWbr/vC);
                        }
                        durPosQu=end-endQc;

                  }
                  else
                  {
                        durNegQu=intercept-B;
                        durNegQc=0;
                        durPosQc=endQc-(tWbr-xWbr/vC);
                        durPosQu=end-endQc;
                  }
            }
            else
            {
                  if(intercept<=startQc)
                  {
                        durNegQu=intercept-B;
                        durPosQu=startQc-intercept+(end-endQc);
                        durPosQc=endQc-startQc;
                        durNegQc=0;
                  }
                  else
                  {
                        durNegQu=startQc-B;
                        durNegQc=intercept-startQc;
                        durPosQc=endQc-intercept;
                        durPosQu=end-endQc;
      //                System.out.println("fake intercept and it
occurs after startQc");
      //                System.out.println("DDD");
                  }
            }
            if (s<L)  // adds pd of 0 error at end for sensor spacing
shorter than s=L
            {
                  if (xW<0) // waves meet outside of segment
                  {
                        durNoErrorQu = tWbr + xWbr/-vWbr + L/vU1 - end;
                  }
                  else // waves meet inside segment
                  {
                        durNoErrorQu = tW + (L-xW)/vU1 - end;
                  }
```

```
            }
            else
            {
            }  // do nothing


            double durNeg=durNegQu+durNegQc;
            double durPos=durPosQu+durPosQc;
            vhtNeg=negAreaQu*qU+negAreaQc*qC;
            vhtPos=posAreaQu*qU+posAreaQc*qC;
            double vhtNoError=qC*noErrorAreaQc;
            vhtAbs=vhtNeg+vhtPos+vhtNoError;
            avgNegQu=negAreaQu/durNegQu;
            avgNegQc=negAreaQc/durNegQc;
            avgNeg=vhtNeg/(durNegQu*qU+durNegQc*qC);
            avgPosQu=posAreaQu/durPosQu;
            avgPosQc=posAreaQc/durPosQc;
            double avgPosDenom=durPosQu*qU+durPosQc*qC;
            avgPos=vhtPos/avgPosDenom;
            if (xWin>=1)  // waves meet within segment, no period of Qc
            {
            avgErrAbs=vhtAbs/(qU*(tW+(L-xW)/vU1-tDept[0]));
            }
            else if ((tWbs-(xWbs/vWbs)+(xWbs/vC)+(L-xWbs)/vU1)<tWbr-
(xWbr/vC)) // have mid period of 0 error because duration makes it 2
separate study periods
            {
                  avgErrAbs=vhtAbs/(qU*((L/vU1)+tWbs-(xWbs/vWbs)-
tDept[0])+qC*(tWbr-(xWbr/vWbr)-(tWbs-(xWbs/vWbs)))-qC*(tWbr-(xWbr/vC)-
(tWbs-(xWbs/vWbs)+(xWbs/vC)+(L-xWbs)/vU1)));
            }
            else
            {
                  avgErrAbs=vhtAbs/(qU*((L/vU1)+tWbs-(xWbs/vWbs)-
tDept[0])+qC*(tWbr-(xWbr/vWbr)-(tWbs-(xWbs/vWbs))));
            }

      System.out.println(avgErrAbs);

}  // p closing bracket
            }//pp closing bracket
            } //xx closing bracket
      } // v closing bracket

      }// void main bracket
      }//opening bracket
```

APPENDIX B

MACRO FOR PARAMICS FILES

A macro, written in Visual Basic, was developed to semi-automate the conversion of Paramics output files so that the data can be analyzed. The output files contained individual vehicle data that was collected during the simulation period. This was achieved by placing sensors on the segment at one twelfth mile spacing. Each sensor recorded the time at which each vehicle passed it along with the unique identification number of the vehicle. Paramics outputted the information from each of the sensors into its own file for a total of 120 separate comma separated value (CSV) files per run. The macro converts the files from CSV to Excel (xlsx), changed the order of the columns so that the vehicle identification number is the first column, which is needed for the VLOOKUP function in Microsoft Excel, and changed the number format of the time stamps so that they could be analyzed. The information was then manipulated and combined into one file.

```
Sub ConvertCSVtoXLS()
    Dim strCSVFile As String
    Dim strXLSFile As String

    'Change Input and Output folders to relevant location
    Const strInputFolder As String = "C:\Temp\CSV\"
    Const strOutputFolder As String = "C:\Temp\XLS\"

    strCSVFile = Dir(strInputFolder & "*.csv")

    Do While strCSVFile <> ""
        strXLSFile  =  Left(strCSVFile,  InStrRev(strCSVFile,  "."))  &
"xlsx"
        Workbooks.Open strInputFolder & strCSVFile
        ' vehicleIdandTimeFormat Macro

'  Moves  the  Vehicle  ID  column  into  the  first  column  position  and
changes the time format to number
    Columns("D:D").Select
    Selection.Cut
    Columns("A:A").Select
    Selection.Insert Shift:=xlToRight
    Columns("B:B").Select
   Selection.NumberFormat = "0.00"

    ' end formatting columns

        ActiveWorkbook.SaveAs strOutputFolder & strXLSFile, _
        FileFormat:=xlOpenXMLWorkbook, CreateBackup:=False
        ActiveWorkbook.Close False
        strCSVFile = Dir
    Loop

End Sub
```

APPENDIX C

ILLUSTRATIVE EXAMPLE OF PREDICTION ERROR

The purpose of this section is to demonstrate that longer sensor spacing may result in less travel time prediction error than shorter spacing for some incidents. This is illustrated with an example. In particular, the travel time prediction error is examined for an incident with a v/c ratio of 0.76 and duration of .67 hours on a ten mile three lane freeway segment. The incident's characteristics were provided previously in Table 5.1 in Chapter 5. The function of travel time prediction error by sensor spacing for this incident is shown in Figure C.1.

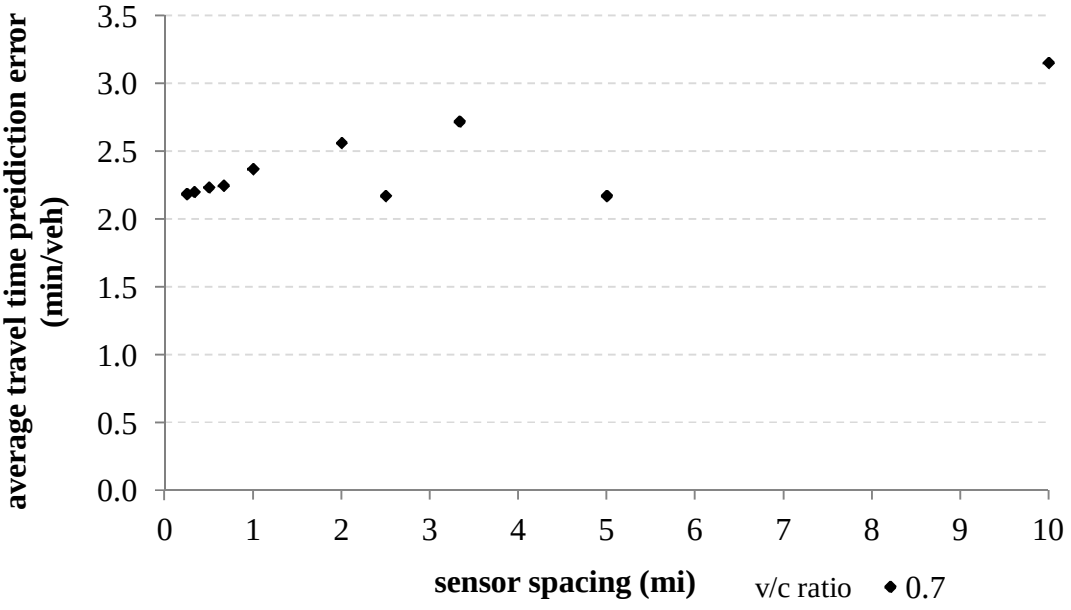


Figure C.1 Travel time prediction error for a 40-minute incident located at 5 miles.

As can be seen in Figure C.1, a sensor spacing of five miles (i.e., three sensors on the ten mile segment) resulted in less travel time prediction error than a sensor spacing of three and one third miles (i.e., four sensors on the ten mile segment). To study why the error was worse for the shorter spacing, three functions were calculated: actual travel time by departure time, predicted travel time by departure time for a five mile sensor spacing, and predicted travel time by departure time for a three and one third mile sensor spacing. These functions can be seen in Figure C.2 below.

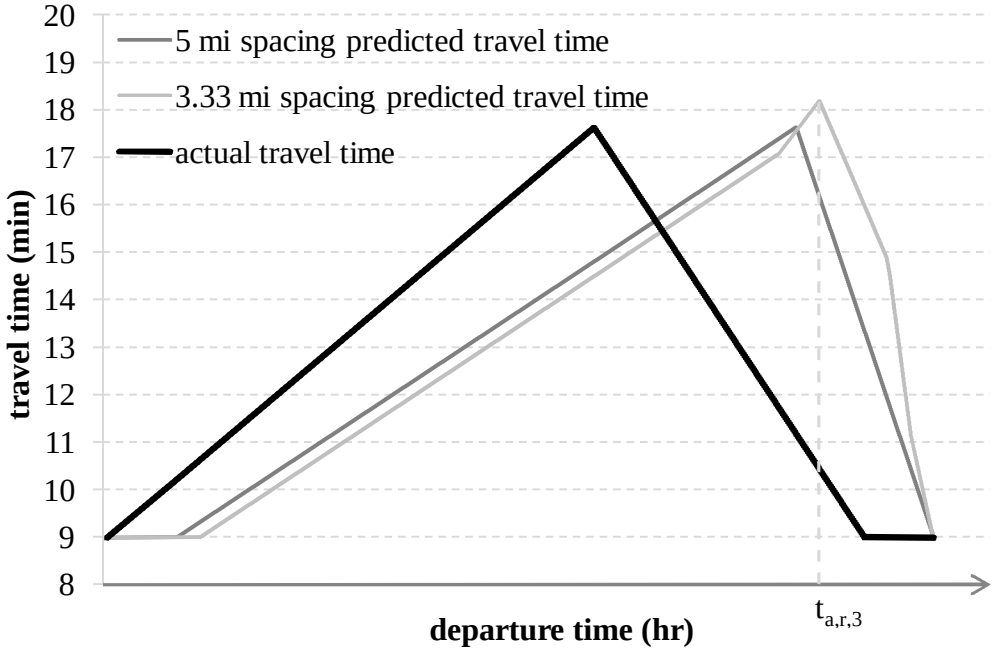


Figure C.2 Predicted travel time and actual travel time by departure time.

From the graph in Figure C.2 one can see that, for almost all of the departure times, the travel times predicted for the longer sensor spacing (i.e. five miles) are more accurate than the predicted travel times for the shorter sensor spacing. In order to understand this, the predicted travel time for a specific vehicle departure time is derived.

The departure time $t_{a,r,3}$ was chosen and it is denoted in the figure. This departure time corresponds to the detection time of the first vehicle (hence index *a*) that encountered the backward recovery wave (hence index *r*) by the third sensor (hence index 3) when the spacing was three and a third miles. This detection time, as it is related to this occurrence, can be seen in Figure C.4.

The departure time $t_{a,r,3}$ was chosen to be further analyzed for three reasons which are illustrated in Figure C.2. First, the travel time prediction error for this departure time was worse for the shorter sensor spacing than for the longer spacing. Second, this departure time had the largest predicted travel time for the sensor spacing of three and one third miles. Third, the predicted travel time was greater than the largest actual travel time of any vehicle.

The travel time prediction for the vehicle that departed at time $t_{a,r,3}$ was derived from the travel time of the vehicles detected at time $t_{a,r,3}$ by sensors $n = 2$ to N on the segment. Therefore, when there was five mile sensor spacing, and thus N equaled three, the predicted travel time was derived from the vehicles detected by the second and third sensors as illustrated in Figure C.3. Also, when there was three and a third mile sensor spacing, and thus N equaled four, the predicted travel time was derived from the vehicles detected by the second, third, and fourth sensors as illustrated in Figure C.4. The derivations of the predicted travel times are illustrated in Figures C.3 and Figure C.4 for the spacing of five miles and the spacing of three and a third mile, respectively. The sensors' locations are denoted on the y-axis and are shown to be constant over time with thin gray lines. The trajectories of the vehicles detected by the sensors at time $t_{a,r,3}$, from which the travel time predictions are calculated, are shown in blue. The portion of the

trajectory between sensors *n*-1 and *n* is highlighted as it is used in the prediction. The trajectory of the vehicle that departed at time $t_{a,r,3}$, and therefore received the travel time prediction, is shown with a green line.

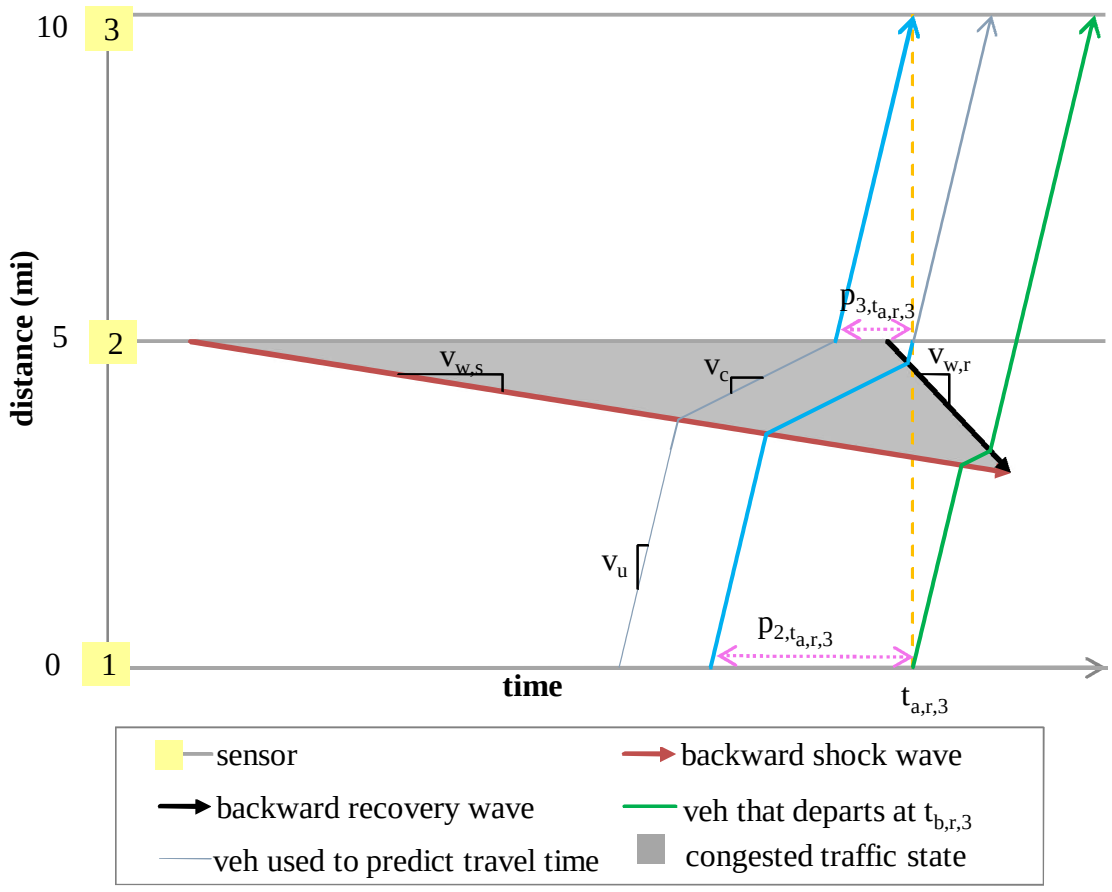


Figure C.3 Derivation of the predicted travel time for five mile spacing.

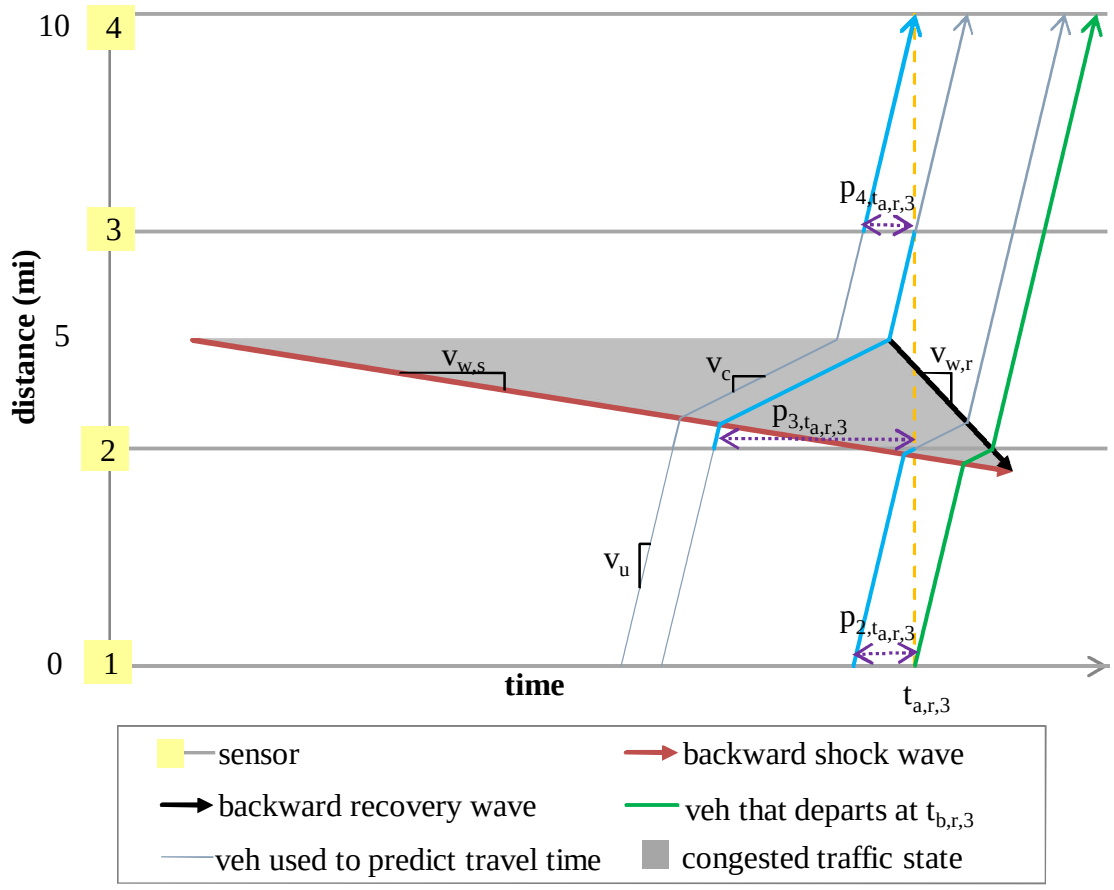


Figure C.4 Derivation of the predicted travel time for three and one third mile spacing.

From Figure C.4, it is apparent why the predicted travel time for departure time $t_{a,r,3}$ was greater than the longest actual travel time of any vehicle. As can be seen in the figure, the predicted travel time for the link between sensors two and three is derived from the vehicle which traveled the farthest distance possible at the congested speed. Also, the predicted travel time for the link between sensors one and two is derived from a vehicle which traveled part of the link at the congested speed. Therefore, the predicted travel time calculated from these two vehicles is greater than the longest actual travel time because it reflects a longer distance in congested conditions than any one vehicle actually drives through. In contrast, in Figure C.3, the predicted travel time is only derived from one vehicle that encountered the congested conditions and this vehicle did

not travel the farthest possible distance in the congested conditions. The prediction from the other vehicle, detected by sensor three, only reflected uncongested conditions.

The vehicle's actual travel time was compared to the predicted travel time from the five mile spacing and from the three and one third mile spacing. The comparison is shown in Figure C.5. In the figure, three vehicle trajectories are shown. The first, shown with a green line, is the actual trajectory of the vehicle that departs at time $t_{a,r,3}$. The two other trajectories, shown with a pink line and a purple line, are the predicted trajectories for departure time $t_{a,r,3}$. The pink line represents the trajectory predicted with the use of five mile sensor spacing. This corresponds to the prediction in Figure C.3. The purple line represents the trajectory predicted with the use of three and a third mile sensor spacing. This corresponds to the prediction in Figure C.4. The predicted trajectories are depicted with a constant speed to reflect that a DMS displays a predicted travel time for the entire length of the segment. The prediction is not broken down for different sections of the segment.

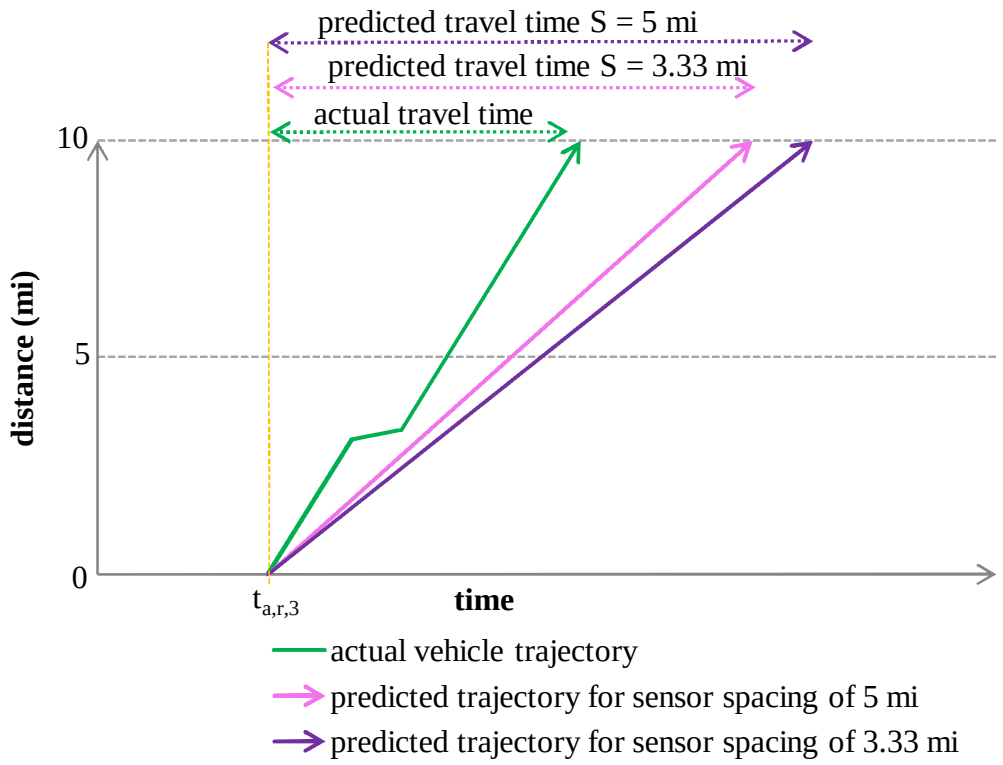


Figure C.5 Actual versus predicted vehicle trajectories.

Figure C.5 shows that the travel time was over predicted with both sensor spacing. The over prediction is a result of the vehicle departing at time $t_{a,r,3}$ traveling less distance in congested conditions than the vehicles used to predict its travel time. However, there is less of a difference between the predicted travel time and actual travel time for the five mile spacing than the three and a third mile spacing. This is because the longer sensor spacing results in less of the prediction trajectories to come from vehicles that traveled in congested conditions. In conclusion, this example demonstrated that longer sensor spacing can provide more accurate travel time predictions.

REFERENCES

Ban, X., Chu, L., Herring, R., Margulici, J.D. (2009) Optimal Sensor Placement for Both Traffic Control and Traveler Information Applications. *Proceedings of Transportation Research Board 88th Annual Meeting*, Washington, D.C.

Bartin, B., Ozbay, K., Iyigun, C. (2006) A Clustering Based Methodology for Determining the Optimal Roadway Configuration of Detectors for Travel Time Estimation. *IEEE Intelligent Transportation Systems Conference*, Toronto, Ontario, Canada, pp. 659-664.

Bertini, R.L. (2007) Toward Optimal Sensor Density for Improved Freeway Travel Time Estimation and Traveler Information. *Proceedings of IEEE Intelligent Transportation Systems Conference*, Seattle, WA.

Bertini, R.L., Lovell, D.J. (2009) Impacts of Sensor Spacing on Accurate Freeway Travel Time Estimation for Traveler Information. *Journal of Intelligent Transportation Systems*, Vol. 13, No. 2, pp. 97-110.

Chan, K.S., Lam, W.H.K. (2002) Optimal Speed Detector Density for the Network with Travel Time Information. *Transportation Research Part A: Policy and Practice*, Vol. 36, No. 3, pp. 203-223.

Chen, M., Chien, S.I.J. (2000) Determining the Number of Probe Vehicles for Freeway Travel Time Estimation by Microscopic Simulation. *Transportation Research Record*, No. 1719, pp. 61-68.

Cheu, R.L., Xie, C., Lee, D.H. (2002) Probe Vehicle Population and Sample Size for Arterial Speed Estimation. *Computer-Aided Civil and Infrastructure Engineering* Vol. 17, No. 1, pp. 53-60.

Chien, S.I.J., Spasovic, L. (2010) Determining the Location of Vehicle Sensors for Accurate Travel Time Estimation. *New Jersey Department of Transportation: Task Order 82*, New Jersey Institute of Technology ITS Resource Center, Newark, NJ.

Chien, S.I.J., Spasovic, L. (2012) Development of New Jersey Rates for the NJCMS Incident Delay Model. *New Jersey Department of Transportation*, New Jersey Institute of Technology ITS Resource Center, Newark, NJ.

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

Danczyk, A., Liu, H.X. (2011) A Mixed-Integer Linear Program for Optimizing Sensor Locations Along Freeway Corridors. *Transportation Research Part B: Methodological*, Vol. 45, No. 1, pp. 208-217.

Dowling Associates. (1999) Travel Time Data Collection Pilot Project. Final Report. *Metropolitan Transportation Commission*. Oakland, CA.

Edara, P. K., Guo, J., Smith, B. L., and McGhee, B. L. (2008) Optimal Placement of Point Detectors on Virginia's Freeways: Case Studies of Northern Virginia and Richmond. V*irginia Transportation Research Council: Report No. VTRC 08-CR3*, Charlottesville, VA.

Feng, W., Bigazzi, A.Y., Kothuri, S., Bertini, R.L. (2010) Freeway Sensor Spacing and Probe Vehicle Penetration: Impacts on Travel Time Prediction and Estimation Accuracy. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2178, pp. 67-78.

Garib, A., Radwan, A.E., Al-Deek, H. (1997) Estimating Magnitude And Duration of Incident Delays. *Journal of Transportation Engineering* Vol.123, No. 6, pp. 459-466.

Hadi, M., Zhan, C., Xiao, Y. (2012) Prediction and Visualization of Incident Impacts for Real-Time Traffic Management Center Operations. *Proceedings of ITS America 22nd Annual Meeting & Exposition*, National Harbor, MD.

Haghani, A., Hamedi, M., Sadabadi, K.F., Young, S., Tarnoff, P. (2010) Data Collection of Freeway Travel Time Ground Truth with Bluetooth Sensors. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2160, pp. 60-68.

Khattak, A.J., Wang, X., Zhang, H. (2011) iMiT: A Tool for Dynamically Predicting Incident Durations, Secondary Incident Occurrence, and Incident Delays. *Proceedings of Transportation Research Board 89th Annual Meeting*, TRB Paper: 11-0803, Washington, D.C.

Kim, K., Chien, S.I.J., Spasovic, L.N. (2011) Evaluation of Technologies for Freeway Travel Time Estimation: Case Study of I-287 in New Jersey. *Proceedings of Transportation Research Board 90th Annual Meeting*, Washington, D.C.

Lee, C., Volpatti, S. (2010) Effects of Shock Waves on Freeway Crash Likelihood. *The Open Transportation Journal*, Vol. 4, pp. 61-70.

Li, D., Cheng, L. (2011) Bayesian Network Classifiers for Incident Duration Prediction. *Proceedings of Transportation Research Board 90th Annual Meeting*, Washington, DC.

Li, S., Zhu, K., van Gelder, B.H.W., Nagle, J., Tuttle, C. (2002) Reconsideration of Sample Size Requirements for Field Traffic Data Collection with Global Positioning System Devices. *Transportation Research Record,* No. 1804, pp. 17-22.

Lighthill, M.J., Whitham, G.B. (1955) On Kinematic Waves. II. A Theory of Traffic Flow on Long Crowded Roads. *Porceedings of the Royal Society*, Vol. A, No. 229, pp. 317-345.

Liu, X., Chien, S.I.J., Kim, K. (2012) Evaluation of Floating Car Technologies for Travel Time Estimation. *Journal of Modern Transportation*, Vol. 20, No. 1, pp. 49-56.

Liu, Y., Lai, X., Chang, G.-L. (2006) Detector Placement Strategies for Freeway Travel Time Estimation. *IEEE Intelligent Transportation Systems Conference*, Toronto, Ontario, Canada, pp. 499-504.

Mahmassani, H., Rakha, H., Hubbard, E., Lukasik, D. (2012) Concept Development and Needs Identification for Intelligent Network Flow Optimization (INFLO) – Assessment of Relevant Prior and Ongoing Research. *United States Department of Transportation ITS Joint Program Office Research and Innovative Technology Administration (RITA)*, Report No. FHWA-JPO-13-011, Washington, D.C.

Owens, N., Armstrong, A., Sullivan, S., Mitchell, C., Newton, D., Brewster, R., and Trego, T. (2010) Traffic Incident Management Handbook. *Federal Highway Administration: Report No. FHWA-HOP-10-013*, Washington, D.C.

May, A. (1990) *Traffic Flow Fundamentals*. Prentice Hall, New Jersey.

Mazaré, P.-E., Tossavainen, O.-P., Bayen, A., Work, D.B. (2012) Trade-offs Between Inductive Loops and GPS Probe Vehicles for Travel Time Estimation: Mobile Century Case Study. *Proceedings of Transportation Research Board 91st Annual Meeting*, Washington, D.C.

Messer, C.J., Dudek, C.L., Friebele, J.D. (1973) Method for Predicting Travel Time and Other Operational Measures in Real-Time during Freeway Incident Conditions. *Highway Research Record*, Vol. 461, pp. 1-16.

Mirchandani, P.B., Gentili, M., He, Y. (2009) Location of Vehicle Identification Sensors to Monitor Travel-Time Performance. *IET Intelligent Transport Systems,* Vol. 3, No. 3, pp. 289-303.

Mirchandani, P.B., He, Y. (2008) Sensor Locations on a Network to Predict Travel Times. *Proceedings of Transportation Research Board 87th Annual Meeting*, Washington, D.C.

Nam, D.H., Drew, D.R. (1998) Analyzing Freeway Traffic Under Congestion: Traffic Dynamics Approach. *Journal of Transportation Engineering*, Vol. 124, No. 4, pp. 208-212.

Newell, G.F. (1993a) A Simplified Theory of Kinematic Waves in Highway Traffic, Part I: General Theory. *Transportation Research Part B: Methodological*, Vol. 27, No. 4, pp. 281-287.

Newell, G.F. (1993b) A Simplified Theory of Kinematic Waves in Highway Traffic, Part II: Queueing at Freeway Bottlenecks. *Transportation Research Part B: Methodological*, Vol. 27, No. 3, pp. 289-303.

Ozbay, K., Noyan, N. (2006) Estimation of Incident Clearance Times Using Bayesian Networks Approach. *Accident Analysis & Prevention*, Vol. 38, No. 3, pp. 542-555.

Qin, L., Smith, B.L. (2001) Characterization of Accident Capacity Reduction. *Center for Transportation Studies at the University of Virginia*, Charlottesville, VA.

Richards, P.I. (1956) Shock Waves on the Highway. *Operations Research*, Vol. 4, No. 1, pp. 42-51.

Roess, R.P., Prassas, E.S., McShane, W.R. (2004) Traffic Engineering, Third ed. *Pearson Eductiation, Inc.*, Upper Saddle River, New Jersey.

Schrank, D., Eisele, B., Lomax, T. (2012) 2012 Urban Mobility Report. *Texas A&M Transportation Institute*, College Station, TX.

Schrank, D., Lomax, T. Eisele, B. (2011) 2011 Urban Mobility Report. *Texas A&M Transportation Institute*, College Station, TX.

Sen, A., Thakuriah, P., Zhu, X.Q., Karr, A. (1997) Frequency of Probe Reports and Variance of Travel Time Estimates. *Journal of Transportation Engineering*, Vol. 123, No. 4, pp. 290-297.

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

Smith, K.W., Smith, B.L. (2001) Forecasting the Clearance Time of Freeway Accidents. *Center for Transportation Studies at the University of Virginia*, Charlottesville, VA.

Srinivasan, K.K., Jovanis, P.P. (1996) Determination of Number of Probe Vehicles Required for Reliable Travel Time Measurement in Urban Network. *Transportation Research Record,* No. 1537, pp. 15-22.

Toledo, T., Beinhaker, R. (2006) Evaluation of the Potential Benefits of Advanced Traveler Information Systems. *Journal of Intelligent Transportation Systems,* Vol. 10, No. 4, pp. 173-183.

Tudge, R., Luk, J. (2007) Microsimulation Standards Technical Note 2. *ARRB Group,* New South Wales, AU.

Turner, S.M., Holdener, D.J. (1996) Probe Vehicle Sample Sizes for Real-Time Information: The Houston Experience. *Proceedings of Proceedings of the Annual Meeting of ITS America 6th,* No. 1, pp. 287-295.

Windover, J.R., Cassidy, M.J. (2001) Some Observed Details of Freeway Traffic Evolution. *Transportation Research Part A: Policy and Practice,* Vol. 35, No. 10, pp. 881-894.

Wirasinghe, S.C. (1978) Determination of Traffic Delays from Shock-Wave Analysis. *Transportation Research /UK/,* pp. 343-348.

Wu, W., Kachroo, P., Ozbay, K. (1998) Validation of WAIMSS Incident Duration Estimation Model. *Faculty Publications (ECE)* Paper 129, pp. 3234-3239.

Wunderlich, K. E., Hardy, M. H., Larkin, J. J., Shah, V. P. (2001) On-Time Reliability Impacts of Advanced Travelere Information Services (ATIS): Washington, DC Case Study. *Mitretek Systems,* McLean, Virginia.