ABSTRACT

Title of Document: VEHICULAR TRAFFIC MODELLING, DATA ASSIMILATION, ESTIMATION AND SHORT TERM TRAVEL TIME PREDICTION.

Kaveh Farokhi Sadabadi, Doctor of Philosophy, 2014

Directed By: Professor Ali Haghani, Civil and Environmental Engineering Department

This dissertation deals with the problem of short term travel time prediction. Traffic dynamics models and traffic measurements are in particular the tools in approaching this problem. Effectively, a data-driven traffic modeling approach is adopted. Assimilating key traffic variables (flow, density, and speed) under standard continuum traffic flow models is fairly straight-forward. In current practice, travel time (space integral of pace or inverse of speed) is obtained through trajectory construction methods. However, the inverse problem of estimating speeds based on travel times is generally under-determined. In this dissertation, appropriate dynamic model and solution algorithms are proposed to jointly estimate speeds and travel times. This model essentially paves the way to assimilate travel time data with other traffic measurements. The proposed travel time prediction framework takes into account the fact that in reality neither traffic models nor measurements are flawless. Therefore, optimal state estimation methods to solve the resulting state-space model

in real-time are proposed. Alternative optimality criterion such as minimization of the variance of estimate errors and minimization of the maximum (minmax) estimate errors are considered. Practical considerations such as occurrence of missing data, delayed (out of order) arrival of measurements and their impact on solution quality are addressed. Proposed models and algorithms are tested on datasets provided under NGSIM project.

VEHICULAR TRAFFIC MODELLING, DATA ASSIMILATION, ESTIMATION AND SHORT TERM TRAVEL TIME PREDICTION

By

Kaveh Farokhi Sadabadi

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2014

Advisory Committee:
Professor Ali Haghani, Chair
Professor Gang-Len Chang
Associate Professor David J. Lovell
Assistant Professor Barton A. Forman
Professor Christopher C. Davis

Dedication

To my parents, Shahpur and Khadijeh.

Acknowledgements

I would like to express my deep appreciation and gratitude to my advisor, Dr. Ali Haghani, for his unwavering support and encouragements during my time at the University of Maryland. This dissertation would not have been completed without his vision and patience.

I would also like to thank my committee members for their time and constructive comments that contributed to the improvements in my research. I was fortunate to be in a position to take advantage of their intellectual heft, insights, and very thoughtful suggestions.

I appreciate all the help and support I received from my colleagues at the Center for Advanced Transportation Technology (CATT). I learned a lot from working with this great group of professionals. Stan, Michael, and Phil have been really amazing colleagues and friends. Especially, I would like to acknowledge Tom Jacobs for his support and for being the greatest boss I have ever had. Go Ravens!

My family has played an important role in my success throughout my life. I am grateful to my parents who instilled in me a strong love for knowledge and the desire for learning. I know they made many sacrifices to give me the education and support I needed. I am immensely indebted to my siblings Hilda, Kambiz, Nahid, and Kamyar for their love and emotional support.

My friends and colleagues at the UMD transportation program provided a cordial and pleasant working environment. I certainly enjoyed having thought-provoking conversations with many of them. I am thankful to my officemates over the years, Taehyeong, Mona, Hadi, Susmit, Kiavash, Keivan, and Ali with whom I shared so

many memorable moments. Certainly the help and encouragements I received from my good friends Masoud, Shahab, Behrang, Gulsa, Rahul, Sevgi, Gina, Nikola, Mercedeh, Yashar, Wenxin, Yanru, Rafael, and Sepideh made my journey much easier. I sincerely treasure our friendships and hope that we will build on that in the future.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction and Motivation

This dissertation deals with the problem of vehicular traffic modeling and assimilation of field measurements in presence of travel time data. In particular, novel travel time models are derived which makes the assimilation of field measured travel times and conventional traffic models straight-forward. The proposed travel time models make the joint estimation of speed and travel times possible. Using proposed models, travel times in both anticipative (predictive) and retrospective modes can be integrated into the estimation process. Applications of the proposed models in offline and real-time estimation processes are presented. Numerical experiments on real-world datasets are presented and discussed.

Traditionally, local traffic state is described by density, speed, and flow measures. However, travel time experienced by vehicles over a distance has both spatial and temporal aspects and in that sense is not a local measure of traffic state. In a forward estimation setting, travel times are estimated by post-processing given standard traffic state (speed) estimates. Proposed methods for travel time estimation in essence attempt to construct vehicle trajectories based on given speeds. Simply put, trajectory construction methods perform numerical integration of speed domain to achieve travel time estimates. In general, these methods are numerically expensive. Besides, they are only as good as their input speed data.

With increasing access to the field travel time measurements, it is desirable to incorporate travel times into the traffic estimation process. This requires development of appropriate framework and models to relate global travel time measures with local traffic states such as speed. Essentially, this is what this dissertation sets out to accomplish. The following are three major hypotheses that form the basis of this research.

- There is a relationship between travel times and local speeds; or more accurately, between local

travel time variations and speeds.

- Inherently, there is value in field measurements. Incorporating travel time measurements into traffic estimation will increase the accuracy of estimations.
- Better estimates lead to better predictions. Travel time can be predicted more accurately when currently realized travel times are taken into account.

The underlying concept pursued in this research is that in order to achieve a good prediction the following two ingredients: a good estimate of the current state of the system, and an accurate model to propagate the current state into the future are needed. Traffic dynamics models and traffic data are in particular the tools in approaching this problem. Effectively, the proposed solution to this problem is a data-driven traffic modeling approach.

Intentionally, the framework proposed in this dissertation can accept and make use of any conceivable traffic data source. There are two reasons for this decision. First, each data source has an additional value and can potentially improve both the estimation and prediction quality. Second, a framework that is general enough to include all different possibilities of data availability is preferred because in practice it is not clear how much of each traffic data type will be available and how that will change over time. This would make the proposed models and methods more robust against changes in data availability and therefore makes it transferable.

The remainder of this chapter is dedicated to addressing various questions about travel time, its value as traffic data, and its measurement and prediction. The answers to these questions serve as motivation for this research. Also, the proposed approach in this research and its contributions are summarily discussed. Finally, an overview of the other chapters included in this dissertation concludes this chapter.

1.1 Why is travel time important?

Travel time is a crucial performance measure of any transportation system. Travel time is an indicator of congestion and the associated delay incurred by the system users. It also indirectly serves as a measure of how much fuel is being wasted in the system as well as how much pollutant is being released into the environment. Therefore, it is a common practice to use travel time as proxy for travel cost, including externalities, in most traffic studies.

Travel time as a measure of congestion overall has been on the rise. The congestion epidemic is an ever-increasing problem throughout the developed and developing world. In the United States, in 2007 a typical peak traveler on average has experienced up to 51 hours of delay which indicates a nearly 20% increase compared to just a decade ago. In terms of the extra cost incurred by average peak travelers due to congestion, this amount of delay translates into $1,081 which is roughly a 60% increase over the congestion cost a decade ago (Research and Innovative Technology Administration 2010).

A once popular solution to congestion problem, building extra capacity to existing road infrastructure is becoming more difficult. From a sustainability point of view which advocates a balanced approach to economy, environment and social equity we are reminded that we should not feed the fire of congestion solely with the fuel of added capacity. In addition, soaring land prices and construction costs in most congested urban areas would make it difficult to financially justify such widening projects.

Figure 1 illustrates the role of different factors in creating congestion nationwide (Cambridge Systematics, Texas Transportation Institute 2004). It should be noted that recurrent congestion which is primarily due to demand levels surpassing available capacity comprises about 40 percent of congestions (bottlenecks) overall, while non-recurrent congestion due to traffic

incidents, bad weather and work zones are responsible for other 50 percent altogether. This in fact suggests that adding capacity alone is not going to solve the congestion problem.

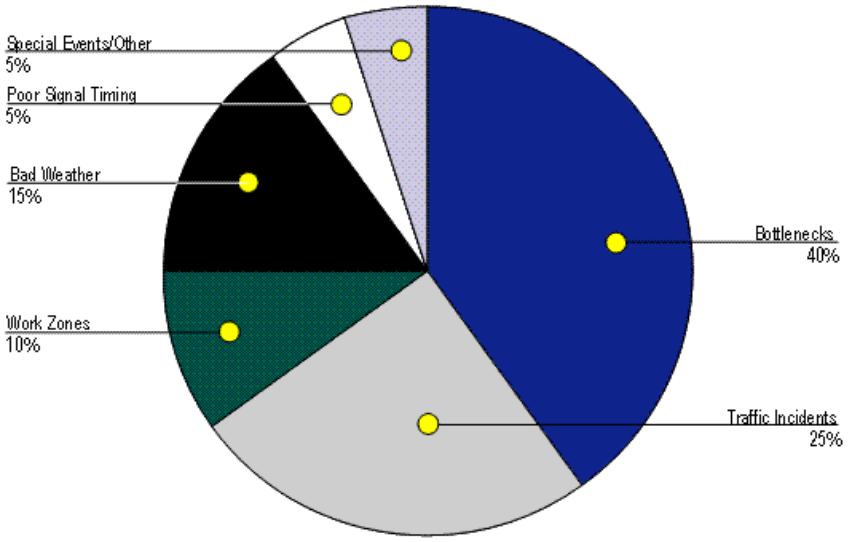


Figure 1. The national summary of the congestion sources (Cambridge Systematics, Texas Transportation Institute 2004).

It seems that a multipronged approach to congestion problem is more acceptable. The basic underlying notion in these solutions is to make a better use of existing capacities rather than adding new capacity to the system. To achieve this goal, road users' decisions have to be informed with accurate and reliable information regarding their planned trips and its associated costs. This entails provision of Advanced Traveler Information Systems (ATIS) off- and en-route. Likewise, system managers need to know what the traffic situation is and where it is heading in order to be able to make best decisions to optimize utilization of the existing facilities and to minimize the associated overall costs of system use. In this context, it is easy to recognize that travel time is the most essential variable that influences the decisions of both users and managers alike. For instance, at an operational level, it is commonly believed that road users consider travel time in their activity, destination, mode and route choice decisions. System managers, on the other hand, set traffic light timings, speed limits, ramp metering control parameters, and determine lane assignments, usage and pricing to minimize system wide travel

times.

1.2 How is travel time measured?

In the past it has been very difficult to measure travel time directly in the field. In order to measure travel time directly it is necessary to capture both spatial and temporal features of vehicle movements. Therefore measuring travel time is much more difficult than measuring cross sectional variables such as flow and time mean speed, or longitudinal variables such as density and space mean speed. In general, traditional methods of travel time measurement aside from being inaccurate for the most part are labor intensive, time consuming and therefore very expensive. This has been the root cause of limited travel time studies in practice in the past.

In the past two decades, however, advent of new technologies such as Global Positioning System (GPS) and cell-phone tracking has led to increased accuracy and more widespread travel time data availability in practice to the extent that today real-time travel time data is available on all major highways and urban facilities nationwide at an affordable price. Currently, traffic data industry is a multi-billion dollar industry and growing. Apart from public sector's interests in accurate travel time data for applications in performance measurement, traffic monitoring, management and traveler information systems, private sector has been the leading force in the traffic data market. Web-based mapping and travel guidance solutions pioneered by giant IT companies such as Yahoo! and Google, in-vehicle navigation systems marketed by TomTom and Garmin, and more recently by all major auto manufacturers such as General Motors, Ford, Lexus, and Toyota are only examples of this growing interest in traffic data and its value-added applications that are common-place today.

To make things even more interesting it should be noted that in recent years, travel time data has become much easier to obtain. For decades, classic Automatic Vehicle Location (AVL)

technologies such as probe and cellphone tracking complemented by a growing number of Automatic Vehicle re-Identification (AVI) technologies such as Automatic License Plate Readers (ALPR), and Radio Frequency Identification (RFID) have been part of traffic engineers' toolbox for travel time measurement. However, cost and privacy concerns associated with these technologies have limited these applications primarily to toll roads and certain corridors. The new wave of AVI technologies based on tagging consumer electronics carried inside vehicles in the traffic stream are more promising in this regard. Bluetooth as a short range wireless communication protocol has received widespread acceptance in consumer electronics devices such as cellphones, hands-free earpieces, navigation systems, laptops, gaming systems, cameras, etc. The Media Access Control (MAC) identity of each Bluetooth antenna is a unique hexadecimal number which potentially can serve as a proxy signature of the vehicle carrying that device in traffic stream. The Bluetooth MAC address is not tied to any database and therefore provides a high level of privacy protection for the public. Recent studies show that travel time samples in the order of three to five percent of hourly traffic volume can be obtained using Bluetooth sensors in the United States depending on the time of day and location of the highway (Haghani et al. 2010). Wireless traffic monitoring, such as Bluetooth and Wi-Fi, could potentially provide a low cost source of travel time measurements that can be spread throughout the road network without any serious requirements for periodic calibration or other maintenance concerns. Such systems are already implemented in some major metropolitan areas (Houston TranStar 2011).

1.3 How is travel time estimated?

For the most part of the 20th century, travel time has been indirectly estimated from other traffic variables such as volume or speed at a planning level. For instance, performance functions such

as well-known Bureau of Public Roads (BPR) type functions essentially establish a simple nonlinear relationship between hourly flow rates, nominal capacity and travel time in a highway. Later, Dynamic Traffic Assignment (DTA) and micro-simulation traffic models have been introduced in an attempt to take into account the time variability of traffic and its impacts on decisions made by system users. In these models travel time is implicitly obtained from the interaction between platoons of vehicles or individual vehicles, respectively. DTA uses mesoscopic traffic flow models while micro-simulations are typically based on car-following and lane changing models. (Transportation Network Modeling Committee 2010).

It should be noted that despite widespread use of these models in practice to formulate and evaluate various policies, designs and planned improvements still they are not accurate enough to closely reproduce any of operational level traffic measures including travel time.

1.4 Two types of travel time: retrospective (measured) or anticipative (predicted)

It is necessary to highlight the distinction between measured travel time data and what is anticipated to be experienced travel time in the future. Figure 2 illustrates this point where measured travel times are obtained when a vehicle is leaving the segment of interest, while the anticipated travel time is assigned to the time vehicle is entering the segment. Even though anticipated travel times are most useful for real time decisions it is currently common to use retrospectively measured travel times instead in all major applications.

There are several good reasons for such strong bias toward measured travel times in real world applications. First, travel time field measurements are more accurate than its predictions. Second, measured travel times maintain the causality conditions. The causality property suggests that reported travel time at any given time should reflect the existing traffic conditions that resulted in a vehicle taking so much time to traverse the segment of interest. The predicted travel times on

the other hand are based on hypothetical traffic conditions that are nonexistent yet but are anticipated to take place in the future times. Third, in practice travel time is reported as discrete values over consecutive time intervals. Obviously, the larger time intervals become the potentially larger variations in travel time within the time interval are masked by the smoothing effect of expected values or averaging operation. Therefore, at some granularity levels, the difference between measured and predicted travel times may not be of practical interest. However, over longer distances and shorter time intervals and in presence of congestion these differences may be very significant. Currently, as the best practice in freeway systems travel times are reported in real-time every minute over approximately one mile long segments.

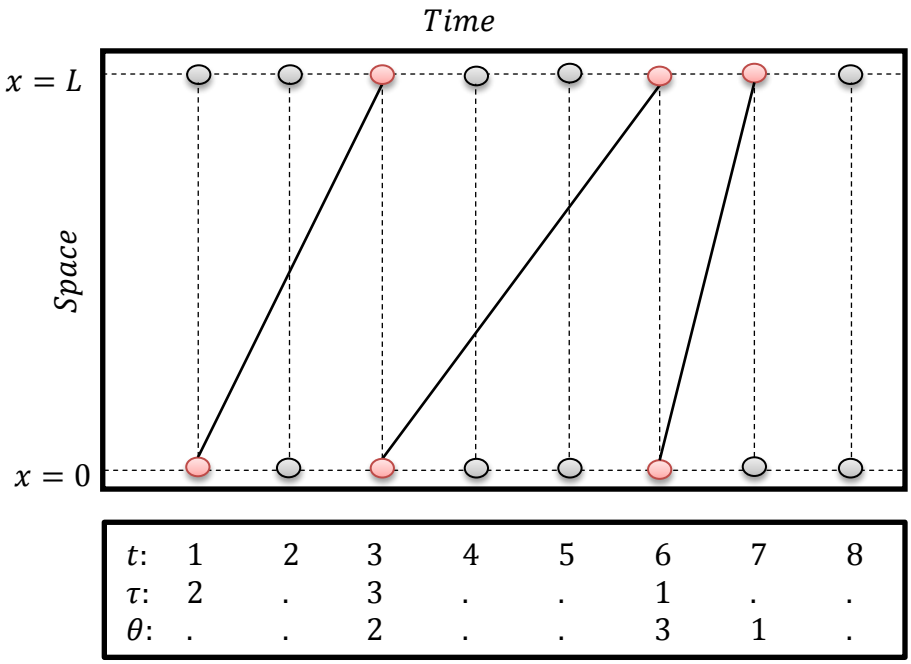


Figure 2. Observed versus anticipated travel times at a given time. (τ is the anticipated travel time of vehicle entering the segment at t, and θ is the observed travel time of vehicle leaving the segment at t)

Based on the above discussion and the existing trend to provide public and system managers with quality real-time travel time data, as well as major legislative mandates such as SAFETEA-LU and MAP-21, it is inevitable that in the near future travel time data in its measured retrospective and predicted anticipative forms is going to become a focal point of advanced

traffic operation and control applications as well as form the basis of informed decision making from both user and manager perspectives.

1.5 What are the sources of error in travel time data?

In general, any error in travel time estimation or prediction can be attributed to three different sources. First, any model is ultimately an attempt in representing the reality. Traffic models are primarily conservation laws expressed in the partial differential equation form. These models have originated from hydrodynamics which primarily deals with movements of continuous fluids. Application of these models to vehicular traffic which is essentially a discrete flow system with human drivers in it is a bit of stretch. However, in absence of better theories for vehicular traffic the conservation based models are the best available so far to model the evolution of traffic and to capture various phenomena known to be present in the traffic stream.

Second, input data are typically outcome of some physical measurement processes followed by transmission of data packets to a central processing unit. Accuracy, frequency and timeliness of the measurements greatly depend on the measurement and transmission technology used. Equipment calibration and maintenance have a great impact on data quality.

Third, estimation and prediction methodology plays an important role in getting the best results out of the model specification and its associated input data. The method should be optimal in some sense. Typically estimation methods seek to optimize one of several well-known objective functions representing least squares, maximum likelihood, and min-max.

1.6 What is accomplished in this research?

This dissertation is aimed at identifying and addressing all three error sources mentioned above in the context of travel time estimation and prediction. In this respect, it is necessary to develop a better understanding of vehicular travel times and their evolution in a typical highway system

according to existing vehicular traffic models. These models essentially conceptualize and define the evolutionary relationships of various traffic variables such as flow, speed and density. Even though travel time is theoretically related to speed, travel time has never been explicitly incorporated into traffic flow models. As mentioned before, this is partly due to the fact that calculating travel time based on speed leads to a line integral which is a function of both spatial and temporal movements of vehicles and as such depends on the trajectory of vehicles.

This dissertation proposes to model the relationship between travel time and speed as a partial differential equation. This equation is derived from first principles of kinematics. This modeling framework would relieve the need for trajectory data, but it requires that a continuity and derivability assumption be made on the travel time function. Note that these assumptions are not too strong in the case of travel time due to its definition as the line integral of speeds.

Both traffic and travel time models proposed in this dissertation are boundary value partial differential equations. Two major approaches exist to solve these equations numerically and efficiently. Finite Difference Methods (FDM) result in a piecewise constant approximation in the solution domain. This is a zero order approximation. Finite Element Method (FEM) on the other hand can be used to increase the order of approximation. This increased solution accuracy comes at the cost of solution efficiency. Therefore, a trade-off between approximation order and efficiency has to be considered. To keep solution methods scalable and to avoid unnecessary complications, in this dissertation FDM solution methods are pursued.

The state-space modeling framework is used to represent the traffic system dynamics in presence of different field measurements over time. The dynamics equations are non-linear in this setting. Therefore, optimal linear stochastic dynamic least square estimation method known as Discrete Kalman Filter (DKF) is not directly applicable to this case. Also, note that traffic dynamic

models are highly nonlinear; hence linearization does not provide accurate estimates of the system state. Therefore, conventional Extended Kalman Filter (EKF) cannot be used for this purpose. However, Ensemble Kalman Filter (EnKF) or Unscented Kalman Filter (UKF) which avoid linearization are applicable in this case. In particular, UKF method which repeatedly draws from the error distributions and sends them through the nonlinear equations to propagate the estimation mean and error covariance is adopted.

In this dissertation an alternative estimation method is also explored. Unscented H-infinity Filter (UHF, the min-max counterpart of UKF) attempts to minimize the maximum estimation error. This method virtually makes it possible to directly enforce an intelligently selected bound on the errors.

1.7 What are the contributions of this research?

Contributions of this research are briefly summarized in the following bullets:

- Deriving a partial differential equation that relates speed and travel times independent of vehicle trajectory,
- Deriving finite difference solutions of coupled first-order velocity-based traffic continuum models and travel time equation,
- Introducing a framework to systematically and explicitly assimilate travel time measurements in traffic estimation process,
- Extending an existing linear state-space model estimation method based on min-max paradigm to the case of highly non-linear joint traffic and travel time models,
- Proposing a delayed filter to explicitly take into account the delayed nature of anticipative travel time measurements with respect to the current time, and in general out-of-order arrival of traffic measurements,
- Sensitivity analysis of travel time estimates and predictions to the presence of various traffic data

sources, error types, and their respective magnitudes,

- Improving short-term travel time predictions

1.8 Preview of the rest of this document

The remainder of this dissertation is organized into six chapters. The following provide a brief description of the topics that are addressed in each chapter.

1.8.1 Chapter 2: Literature Review

First, a comprehensive review of the existing literature on the broad topic of travel time estimation and prediction is provided. In doing so, existing literature in this area is grouped into four main categories based on the type of measurements used. Specifically, methods based on Eulerian (spot speeds, counts), Lagrangian (probe trajectory), and integrated Lagrangian (vehicle re-identification) traffic data measurements are addressed. A brief review of state-space models used in optimal traffic estimation and prediction is provided. This review shows that state of the art in this area is far from perfect and there is substantial room for new and meaningful contributions.

1.8.2 Chapter 3: First-Order Continuum Traffic Flow Model

In this chapter, first a well-known first-order continuum traffic flow model is adopted to represent the dynamics of the system. An equivalent form of this model in terms of speed is derived. This model provides a theoretical framework to describe and analyze traffic processes on a variety of roadway facilities. A finite difference method for numerical solution of the velocity based equivalent of the first-order continuum traffic flow model is proposed.

1.8.3 Chapter 4: Travel Time Model

In this chapter, a first-order Partial Differential Equation (PDE) model relating local variations of travel time with local speeds is derived. At any point along the trip two travel times either with respect to the start or the end of the trip can be defined. The former definition amounts to a

retrospective view of travel time, while the latter leads to an anticipative or predictive definition of travel time. The proposed PDE model is capable of dealing with both travel time definitions with very slight variations required in the underlying model. In addition, some desirable properties of travel time such as stability and first-in first-out (FIFO) are briefly presented and their implications in terms of the proposed travel time models are discussed. Efficient finite difference approximations of the proposed PDE models are presented.

1.8.4 Chapter 5: Estimation Method

In this chapter, the joint traffic and travel time dynamics model is presented. The joint model is cast into a state-space form. All components of the proposed state-space model are specified. Optimal state estimation concept is introduced, and unscented methods to deal with nonlinear state propagation and conditioning steps are identified. An efficient method is introduced to enforce common-sense physical constraints on travel time states. Two optimal estimation methods capable of dealing with highly nonlinear state-space models are proposed. A delayed filtering approach is proposed to assimilate current travel time measurements as delayed predictive travel times.

1.8.5 Chapter 6: Numerical Experiments

NGSIM datasets are used to run different numerical experiments. Results of these experiments under different data assimilation and solution scenarios are presented. Estimation and prediction error measures are reported and computation times of different scenarios are investigated. Discussions on the impacts of estimation method, delayed filter, space-time aggregation level, speed-density relation, and traffic measurements used are provided.

1.8.6 Chapter 7: Conclusions and Future Directions for Research

This chapter outlines the contributions in this dissertation. It summarizes the high-level conclusions reached. And, it presents a set of directions to be pursued for future research.

Chapter 2: Literature Review

Travel time estimation and prediction problem can be classified in many different ways:

First, they can be grouped based on the facility type on which the problem needs to be solved. For instance, travel time estimation in facilities with interrupted flow should be treated differently from estimations performed on facilities with uninterrupted flows. While a lot of effort in the past has been spent to estimate travel times in the latter facility types (e.g. freeways), very few studies report on methods to estimate or predict travel times over the interrupted facilities (e.g. arterials).

Second, in many practical cases, proposed methods are limited to data readily available from existing traffic sensing technologies. This would include stationary sources such as inductive loop detectors and road side microwave radars. Vehicle re-identification data from license-plate or toll-tag readers can provide a sample of travel times. Finally, probes are capable of not only providing a travel time sample but they also will give insight to the evolutions of traffic conditions over space and time inside the segment under study. Methods to fuse data from different sources and to establish a hybrid estimate of travel time are gaining more popularity.

The third aspect that can be used to distinguish between different travel time estimation and prediction methodologies is the inductive (non-parametric) or deductive (parametric) nature of the proposed methods. In broad terms, inductive methods are data-driven and make extensive use of historic observations. Given a representative data set, inductive methods are shown to have a good performance in predicting travel times under recurrent traffic conditions. On the other hand, deductive methods take into account physical principles governing traffic operations and resulting interactions between different traffic parameters and various external factors affecting traffic. Therefore, deductive methods are capable to handle unforeseen traffic situations and are

equally useful in traffic control applications due to their normative nature as opposed to inductive models which have mere descriptive powers.

The fourth characteristic of reported models can be defined with regard to their adaptive or non-adaptive nature. In general, adaptive methods have more flexibility and are able to discern temporal changes in the traffic system under both recurrent and non-recurrent conditions. This is a very desirable feature since accurate travel time estimates are most needed when unforeseen conditions due to incidents, construction, inclement weather, and such arise.

Last, but not least, property of travel time models is the inclusion of a sound vehicular traffic model in the estimation process. Unfortunately, the majority of methods reported in the literature are solely based on generic statistical techniques and do not make any effort to take advantage of the existing knowledge on traffic flow theory.

Travel time estimation and prediction methods reported in the literature can be broadly classified into three groups according to their adopted methodology.

1. Conservation of flow
2. Kinematics
3. Statistical
4. Hybrid

2.1 Models Based on Eulerian Data

2.1.1 Conservation of Flow (Input-Output Curves)

The first group of methods for travel time estimation is based on the conservation of flow principle. Generally speaking, this principle states that vehicles entering a segment at upstream over some time along with the ones initially existing inside the segment are the ones that will leave the segment at the downstream during the same time or will remain in it at the end of the time period. This gives rise to the idea of obtaining travel times by comparing N-curves representing cumulative number of vehicles passing upstream (entering) and downstream

(exiting) of the segment. This idea was first presented by Newell (1993) in which cumulative number of vehicle arrivals at a sequence of locations on a highway are used to estimate travel times, flow variations and shockwave creation and propagation. Cassidy and Windover (1995) described a similar method for assessing the dynamics of freeway traffic. The methodology is more descriptive rather than normative (prescriptive). Figure 3 further illustrates the concept. In this figure, slope of the cumulative curves is equal to traffic flow (q); the vertical distance between two curves at each time represents the accumulation of vehicles on the segment (a) while the horizontal distance is equal to travel time (T) on the segment under study.

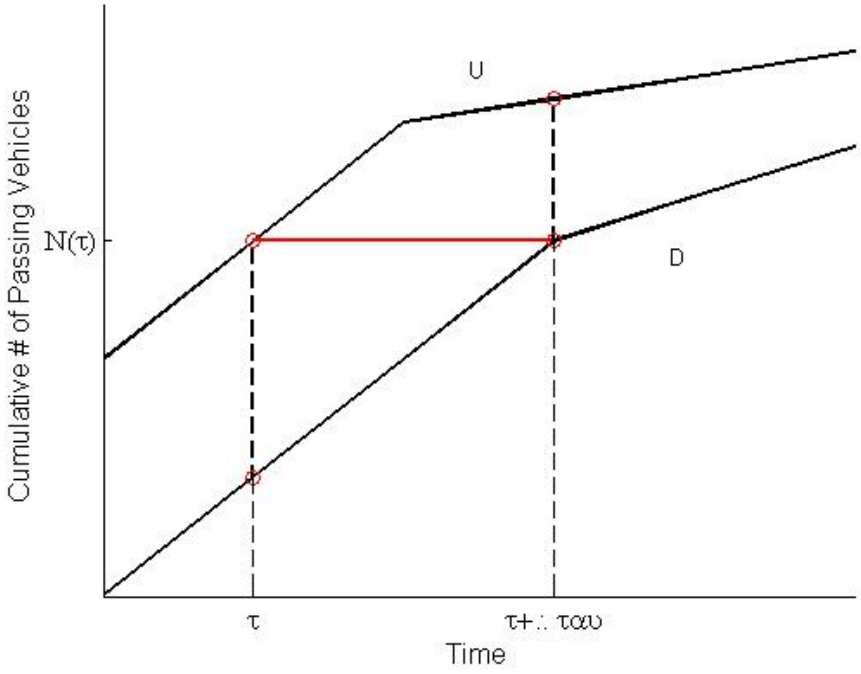


Figure 3. Cumulative input-output curves concept.

$$q(x,t) = \dot{N}(x,t) \tag{2.1}$$

$$N(U,t) = N(D, t + T(t)) \tag{2.2}$$

$$T(t) = N^{-1}(D, N(U,t)) - t \tag{2.3}$$

It may seem that counting the number of vehicles passing a point of the highway should be an

easy process. Inductive loop detectors and a variety of stationary sensors are used to accomplish this task, but it is well-known that counts obtained using these technologies are less than perfect. In fact, ILDs which are not calibrated properly are susceptible to a phenomenon called drifting in which passage of some vehicles are missed. Such technological deficiencies along with the necessary knowledge of the initial number of vehicles in the segment for these methods to work have been the main impediments in widespread use of these methods. To this we can add the fact that the concept is primarily suitable for segments with no access/egress points in the middle. Otherwise, number of vehicles entering and exiting in the middle of the segment should be taken into account.

Assuming that cumulative curves are continuous and smooth everywhere, Astraita (1996) took the derivative of both sides of equation (2.2) and derived the following relation between flow rates at upstream, downstream and travel time on the segment. It should be noticed that flow rates are easier to obtain and to work with than the cumulative number of vehicles.

$$q(D, t + T(t)) = \frac{q(U,t)}{1+T'(t)} \tag{2.4}$$

Carey et al. (2003) proposed a dynamic link travel time model based on the assumption that travel time is a non-decreasing function of the average surrounding flow experienced by a vehicle while traveling along the segment. They approximated this average flow as a linear combination of flow at the entrance and at the exit points of the segment as experienced by the vehicle.

$$T(t) = f(\beta q(U,t) + (1-\beta)q(D, t + T(t))) \tag{2.5}$$

After substituting for downstream flow rate using equation (2.4), they got the following model.

$$T(t) = f(\beta q(U,t) + (1-\beta)\frac{q(U,t)}{1+T'(t)}) \tag{2.6}$$

And, after inverting and rearranging they got the following first-order ordinary differential

equation.

$$T'(t) = -\frac{f^{-1}(T) - q(U,t)}{f^{-1}(T) - \beta q(U,t)} \quad (2.7)$$

Carey (2004) showed that this model has some desirable properties, such as causality, first in first out (FIFO) and similarity to the static model when flows are constant. Carey and Ge (2007) examined several discrete time approximation methods for numerical solution of their proposed model (2.5). These approximations are in fact simple forward and backward differencing methods that are widely used for solving differential equations with no closed form analytical solutions. They point out that simple approximate solutions may be violating FIFO property. Therefore, to keep the FIFO property in approximate solutions, regardless of the size of discrete time intervals, an alternate differencing method is suggested which applies the backward differencing method while moving forward in time. They concluded that this model can be equivalently solved as a simple optimization problem at each time interval. The optimization problem can be solved using simple line search algorithms such as golden section search.

Vanajakshi and Rillet (2006) proposed an adjustment algorithm based on generalized reduced gradient (GRG) method to fix problems associated with accuracy of inductive loop detector records. In essence, this method attempts to make smallest changes in the readings while maintaining the condition that cumulative flow at successive detector points should be smaller or equal to that amount at upstream points. Also, the constraint for allowing practically possible maximum number of vehicles on any road segment at any time is enforced under this methodology. These two conditions in fact hold up conservation of flow principle in the traffic stream. Vanajakshi et al. (2009) used these adjusted detector readings to improve on the travel time estimation method originally proposed by Nam and Drew (1996, 1998, and 1999). Vanajakshi et al. (2009) suggest that the congested flow model (Nam and Drew 1998) should be

used throughout, and that density to be estimated based on a source other than cumulative flows. They use the relationship between occupancy and density to estimate the latter. They report between 9 to 16 percent error in travel time estimates on a segment between two detector stations in their test case. This error increases up to 20 percent on a 2 mile corridor.

Waller et al (2007) adopted an ARIMA(3,1,2) to forecast inflows to the freeway segment under study, then they used a meso-simulation technique called cell transmission model (CTM) to simulate propagation and movements of vehicles inside the segment. Later, based on cumulative flow curves at the segment endpoints they were able to forecast travel time. On a 3 mile freeway segment, they reported 10 to 23 percent RMSE on travel times predicted 5 minutes ahead using this method when compared with travel times obtained from VISSIM micro-simulation.

2.1.2 Approximate Kinematic Models

The second group is comprised of kinematic models. Kinematics is a branch of mechanics which deals with motion without regard to forces or energies that may be exerted on the objects under study. The basic notion of kinematics is that point speed of a vehicle at any given time is equal to the derivative of its trajectory at that time. Therefore, we can derive the relation between distance traveled, speed and travel time in an integral form,

$$\dot{X}(t) = v(X(t), t) \tag{2.8}$$

$$X(T) = X(0) + \int_0^T v(X(s), s)ds \tag{2.9}$$

where,

$X(t)$ is the vehicle position at time t, and

$v(X(t), t)$ is the vehicle speed at time t.

The integral in equation (2.9) is difficult to estimate since in most cases the speed profile of a vehicle during its trip is not known. Instead, it is common to approximate this integral with point

speed measurements at multiple points along the segment over which travel time is to be estimated. Specifically, in highway applications, speeds at upstream and downstream of the segment are usually available.

$$X(T) \cong X(0) + \left[\frac{v(X(0),0)+v(X(T),T)}{2}\right]T \tag{2.10}$$

Therefore, travel time can be estimated as

$$T \cong \frac{2L}{v(X(0),0)+v(X(0)+L,T)} \tag{2.11}$$

where, L is distance traveled or length of the segment $[L = X(T) - X(0)]$.

Equation (2.11) essentially suggests an iterative method to estimate travel times which is called dynamic time slice method in the literature (Waller, et al. 2007). A further approximation of this formula would result in what is called instantaneous method in which downstream speed at the time vehicle enters the segment is used,

$$T \cong \frac{2L}{v(X(0),0)+v(X(0)+L,0)} \tag{2.12}$$

Figure 4 illustrates the times and locations for which speeds are available and are being used to predict travel times versus what speeds should be used. Obviously, these approximations only work under stable traffic conditions when there is not much change in vehicle speeds over space and time.

Lindveld et al. (2000) employed the harmonic mean of speeds to substitute the integral in equation (2.9)

$$X(T) \cong X(0) + \left[\frac{2}{\frac{1}{v(X(0),0)}+\frac{1}{v(X(T),T)}}\right]T \tag{2.13}$$

which results in the following estimate of travel time

$$T \cong \frac{L}{2}\left\{\frac{1}{v(X(0),0)} + \frac{1}{v(X(T),T)}\right\} \tag{2.14}$$

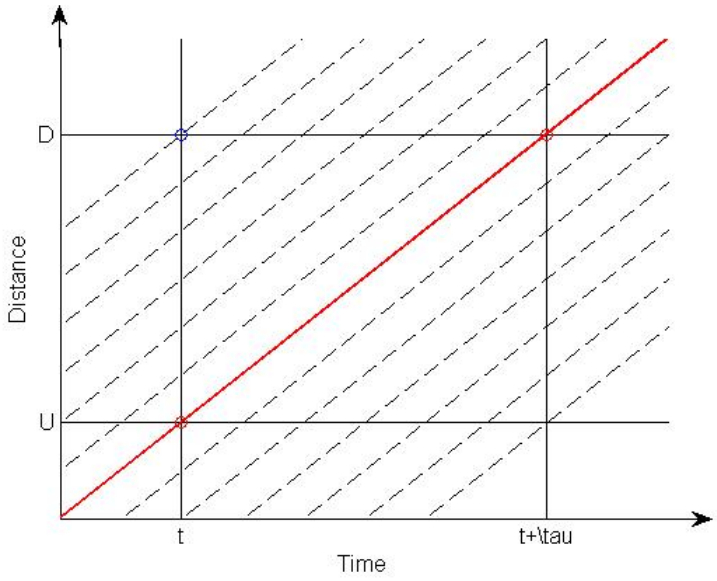


Figure 4. Illustration of typical speed-based travel time prediction concepts.

Further, they evaluated several kinematics based and flow correlation methods for travel time estimation and prediction in three different European sites (Amsterdam, Paris, and Padua-Venice). The input data for these methods generally comes from inductive loop detectors. Evaluation results show that these methods produce RMSEP in travel time estimation/prediction well above 10% under free flow conditions, while as congestion increases their performance rapidly deteriorates.

The kinematics methods are easy to use and provide inexpensive travel time estimation solutions which generally make use of existing sensing technologies and readily available data. However, they lose their accuracy as distances between consecutive sensing stations become large. Also, they are most accurate when traffic condition along the segment is stationary. As traffic conditions begin to change abruptly over time and/or space estimates from these methods become less reliable.

Various technologies are in use to measure vehicle speeds passing a given point on the highway. Inductive loop detectors are among the earliest sensors used for this purpose. In the single loop

setting the relationship between detector occupancy, volume, and vehicle length can be used to estimate spot speeds. For a single vehicle, relationship between time it has kept the detector in presence mode (t_i), detector length (l_d) , vehicle length (l_i) and its speed (v_i) is as follows,

$$t_i = \frac{l_d + l_i}{v_i} \tag{2.15}$$

However, one should keep in mind that data usually is not available at a single vehicle level; instead aggregate data (20~30 seconds) is typically provided by detectors. Therefore, occupancy of detector measured as fraction of time detector has been in presence mode in an interval is defined as below.

$$O = \frac{\sum_{i=1}^{q} t_i}{\Delta t} = \frac{1}{\Delta t} \sum_{i=1}^{q} \frac{l_d + l_i}{v_i} \tag{2.16}$$

where, Δt is the time interval, and q represents the number of detected vehicles in that same interval.

Kurkjian et al. (1980) used an approach based on the first-order method of moments to estimate spot speeds using a single inductive loop detector. They effectively set the summation in (2.16) equal to its average times number of vehicles resulting in the following

$$q . \frac{l_d + \bar{l_v}}{\bar{v}} = \sum_{i=1}^{q} \frac{l_d + l_i}{v_i} \tag{2.17}$$

where, $\bar{l_v}$, is the mean effective vehicle length, and $\bar{v}$ is the average speed during the interval.

Substituting (2.17) into (2.16) the following relationship between spot speed, flow and occupancy may be obtained.

$$\bar{v} = q . \frac{l_d + \bar{l_v}}{O . \Delta t} \tag{2.18}$$

It should be noted that in this setting average vehicle length is not directly measured. Normally, a constant average vehicle length is considered in the above equation. This is a biased estimator. Hall and Persaud (1989) proposed to adjust the estimator by multiplying a correction constant,

however they pointed out that the effect of the bias is not uniform and a constant adjustment factor is not sufficient.

Dailey (1999) applied the Taylor's expansion up to the first two moments of the space-mean speed measurements, resulting in a non-linear function of the population speed parameter. This function was then linearized and used as the observation equation of a state-space model which was then solved by Kalman filter for population speed parameter. Ye et al. (2006) pointed out that the expansion approach is not robust and greatly depends on the linearization, the choice of initial guess and/or changes in vehicular speed. Ye et al. (2006) and Bickel et al. (2007) also used Kalman filter method to estimate vehicular speeds. Hazelton (2004) performed Bayesian analysis and applied Markov Chain Monte Carlo (MCMC) approach based on the assumption that speed in consecutive intervals follow a random walk. This method simulates the posterior distribution of vehicle speeds with a great improvement on accuracy; however, this offline approach is not practical for online estimation. Li (2009a, b) proposed a non-Gaussian Kalman filter and a recursive method for online vehicular speed estimation

Ahmed and Cook (1977) proposed a Box-Jenkins type model for flow and occupancy time series obtained from inductive loop detectors. Their model is essentially an ARIMA(0, 1, 3) model. They compared the performance of this model with three different smoothing algorithms; namely, moving average, double exponential smoothing, and exponential smoothing with adaptive response (Trigg-Leach method). However, they did not report on any modeling effort based on either travel time or speed data.

In a double loop setting, ILDs are able to provide an accurate estimate of vehicle speed based on the passage time lapse and distance between two loops. D'Angelo et al. (1999) proposed a non-linear time series to predict point speeds at the location of dual loop detectors on freeway

segments in the short term. Then these point speeds are simply extended to an area from midpoint of the upstream segment to the midpoint of the downstream segment to evaluate travel times. Ishak and Al-Deek (2002) made a comprehensive analysis of this method. However, in their evaluation they did not use any ground-truth travel time or speeds. Instead, they compared predicted point speeds with observations from loop detectors. One of the major findings of Ishak and Al-Deek (2002) is that increasing rolling horizon (the duration of traffic evolution prior to current time used in predictions) would increase the relative error of travel time predictions. This is a counter-intuitive observation, since we expect a model should perform better when it uses more historical data as input. Additionally, they found that this method produces substantial errors under congested flow conditions. Relative errors of up to 30 percent are reported in less than 20mph range. In 20 to 50mph range errors are as high as 20 percent. Only, at free flow speeds higher than 50mph, relative errors are reported to be less than 5 percent.

Based on a simple shock wave analysis procedure and basic kinematic principle (8), Coifman (2002) proposed a method to build vehicle trajectories around the location of a dual loop detector placed in the middle or on either end of a basic freeway segment. These trajectories then can be used to estimate travel times on the freeway segment. Compared to the naïve travel time estimates such as (2.12) or (2.14), this method reduces the errors by almost half, but the average absolute error still remains at around 10 percent of the ground truth travel times. The accuracy of this method falls with increase in the length of the freeway segment under study. This method is based on the stationary assumption for traffic conditions all over the freeway segment and at all times. Therefore, under normal conditions where queues are formed and later dissipate, one detector depending on its location may not capture all the existing shock waves in the segment and time period of interest.

Sun et al., 2008 proposed a method based on interpolating point speeds read in three consecutive detector stations to estimate travel time on the segment between detectors. This method simply fits a quadratic speed trajectory on three point speeds at these detector stations. For any departure time at first station, it is not clear how one should determine the arrival times at two downstream stations. This is essential in building the speed trajectory. The reported test case exhibits errors of up to 55% in travel time estimates using this method.

2.1.2.1 Non-linear filtering

Treiber and Helbing (2002) proposed an adaptive smoothing method which is essentially a non-linear filter that transforms input stationary detector data into the smooth spatio-temporal functions. The non-linear filter is, in fact, an adaptive linear combination of two linear anisotropic low-pass filters each representing either free-flow or congested traffic status. Weight system in the linear filters is based on exponential functions with scaled relative space-time coordinates. The weight system in the upper combination level is a non-linear hyperbolic tangent function with bias toward congested traffic filter results. No quantitative measure for accuracy of travel time predictions using this method is given. However, visual evidence is given as to accuracy of estimations and predictions.

2.2 Models Based on Lagrangian Data

Lagrangian data is comprised of vehicle trajectory data obtained from tracing probe vehicles inside the road segment of interest. For this reason, this data type may also be called internal data. In this sense, full/partial vehicle trajectories and speed profiles are an example of such data. Full trajectory data is considered very rich since it basically provide a complete record of a vehicle movement and the speeds and travel time it has experienced. In general, trajectory data is both expensive and brings about a host of privacy issues. Therefore, in practice, internal data are still very rare even though GPS and cellphone tracking applications are becoming more popular

as traffic data sources. To address some of the privacy issues, some cellphone companies are using Virtual Trip Lines (VTL) concept to detect passage speed of a sample of vehicles at a set of pre-specified locations which amounts to Eulerian data similar to speed data collected at loop detectors. However, (near) real-time trajectory data as a bi-product of fleet management operations has become available in recent years in certain corridors. The latter provides a major source of Lagrangian traffic data at an affordable cost for travel time estimation for commercial purposes.

2.3 Models Based on Integrated Lagrangian Data

2.3.1 Automatic Vehicle Re-Identification (AVI)

Automatic license plate reader (ALPR), toll-tag readers and video processing systems capable of matching passing vehicles signatures between a pair of locations along the road are examples of these technologies. AVI data directly reflects realized travel times between two observation points, but at the same time it is more difficult to obtain compared to point measurements. Generally speaking, established traffic sensing technologies that are able to provide AVI data are both expensive and controversial in terms of exposing general public to privacy risks and therefore have found very limited geographical reach. As a result, earlier studies in this area tend to make use of widely available point sensors and to show that matching data from a pair of, for instance, loop detectors can result in accurate travel time estimates. In recent years, new emerging technologies have proved to be more effective in providing AVI data. Magnetic and Bluetooth matching sensors are examples of the new wave of AVI technologies.

Hoffman and Janko (1990) are the first who reported on implementing a travel time prediction system using AVI data. In their study, data was obtained from infra-red transmitter/receivers installed at over 230 signalized junctions and 10 locations on urban freeways in West Berlin. A small fleet of vehicles were equipped with the same infra-red capability as well as position

finding devices so that their passage time in front of static devices could be recorded. Their proposed prediction methodology mainly consists of forming a historic data set and estimation of average travel time for each time interval, then a correction factor in the form of ratio of the observed travel time in the last interval to that same interval's historic average is used to predict current interval's travel time. Unfortunately, no measure of accuracy for this method is reported.

Dailey (1993) proposed a signature matching method for travel time estimation which uses cross-correlations between 5 sec vehicle counts from upstream and downstream inductive loop detectors at relatively short distances (0.5 mile is used in the reported example). In this method no effort is made to evaluate speeds from occupancy and therefore there is no need to estimate average length of vehicles. The method chooses the lag associated with the maximum cross-correlation value as the mean travel time between two consecutive detectors. The minimum acceptable cross-correlation value is reported as 0.4 which is shown to correspond to the 15 percent occupancy level. It is postulated that with increase in the congestion level beyond this point, the rigidity in flow of traffic between two points diminishes. Therefore, the method is not suitable for congested situations. No effort to validate the results of this method against ground-truth data has been reported.

Coifman and Ergueta (2003) proposed an algorithm along with four separately designed filters to match signals between two consecutive dual loop detector stations on a single lane. The algorithm identifies a set of feasible upstream pulses for each downstream pulse; each pulse representing the passage of a vehicle. Then all vehicles in this set which have an estimated length range that includes that of the corresponding downstream vehicle are incorporated into a vehicle match matrix. Visual inspection of this matrix suggests that, under stable traffic conditions, correct matches should form a long vertical sequence of entries in the matrix. The method is

therefore based on finding the longest vertical sequence. To eliminate false positives, four tests are introduced; filter test, cone test, travel time test and multiple lane change test. Results of a reported test study on two separate 0.55 km freeway segments demonstrate the accuracy of the method to be extremely good in comparison with ground truth. A mere 1.45 percent average absolute percent error in travel time is reported in a case where there is no on/off ramp between two detector stations. However, in a setting where an off ramp exists on the studied segment no error measure is reported.

Coifman (2003) considered the case of a pair of double loop detectors located at two ends of a freeway lane. In order to detect the start of congestion, he suggested that outstanding vehicle length estimates from downstream station be compared with length estimates from upstream station within a time window reflecting free flow travel times on the segment. If in consecutive time intervals such matches are not found then it is suggested that traffic is in congested mode. However, if a match is found then it provides a travel time estimate. This method works best when larger number of trucks (or any longer vehicles) is present in the mix. In his numeric test, Coifman managed to match 5% of traffic using this method. In its basic case, this may not be very valuable information since free-flow travel time is more or less a known constant (small variation). Therefore, he extends this method to the congested case by considering different travel speed ranges which results in a rudimentary method for travel time estimation under any traffic condition using existing point sensor technology. It should be noted that quantity of matches and also quality of estimates will decrease as congestion increases because during congestion more vehicles change lanes.

2.3.1.1 Time Series Analysis/(Non)Linear Filtering

Generally, methods falling in this category are based on signal processing ideas. It is conceived that travel time observations when ordered on the basis of the sequence of time intervals at which

they have been measured provide a history of the evolution of a system.

$$\boldsymbol{Y} = [y_t] \tag{2.19}$$

Box, Jenkins, and Reinsel (1970) proposed statistical techniques to analyze time series. Auto Regressive Integrated Moving Average (ARIMA) models provide a standard modeling framework in a typical time series analysis. The $ARIMA(p, d, q)$ model is expressed as,

$$\left(1 - \sum_{i=1}^{p-d} \varphi_i B^i\right)(1 - B)^d y_t = \left(1 + \sum_{i=1}^{q} \theta_i B^i\right)\varepsilon_t \tag{2.20}$$

where,

p, is the order of auto-regressive terms,

d, is the number of sequential differencing needed to stationarize the time series

q, is the order of moving average terms,

φ, are the parameters of the auto-regressive part,

θ, are the parameters of the moving average part,

B, is the lag operator defined as $B^i(y_t) = y_{t-i}$, and

$[\varepsilon_t]$, are the error terms series assumed to be independent and identically distributed (i.i.d.) random normally distributed variables with mean equal to zero (white noise).

Dion and Rakha (2006) proposed a real-time adaptive exponential low-pass filtering algorithm for travel time estimation and prediction using very small sample AVI data (less than one percent of traffic volume) from toll-tag readers. They used toll-tag data from TransGuide system in San Antonio to demonstrate the method performance in predicting two minute time intervals. Aside from graphs, no other concrete measure of prediction accuracy is reported.

They assume that travel time is log-normal distributed. This algorithm uses a simple smoothing technique to forecast the future average and variance of travel time. The predicted average travel time is estimated according to the following equation,

$$ln(\hat{y}_{t+k}) = \begin{cases} \alpha.ln(y_t) + (1-\alpha).ln(\hat{y}_t) & ,k = 1 \\ \alpha.ln(y_t) + (1-\alpha).ln(\hat{y}_{t+1}) & ,k = 2 \\ \alpha.ln(\hat{y}_{t+k-2}) + (1-\alpha).ln(\hat{y}_{t+k-1}) & ,k > 2 \end{cases} \quad (2.21)$$

where,

y_t, is the observed travel time at time interval t,

$\hat{y}_t$, is the estimated travel time at time interval t,

α, is the smoothing factor to linearly combine log-normal of travel times at time interval t, and

k, is the number of time steps ahead for which prediction is being performed.

Based on the predicted travel time average and variance a range for valid travel time observations in the next time interval can be specified. Observations that fall outside this validity window are dismissed as outliers. Essentially, in this method, specification of the validity range is performed based on the following four factors:

- Expected average trip time and trip time variability in future time interval,
- Number of consecutive intervals without any readings since the last recorded trip time,
- Number of consecutive data points either above or below the validity range, and
- Variability in travel times within an analysis interval.

2.3.1.2 State-Space Models

Chen and Chien (2001) used probe vehicle travel times as measurements in a trivial Kalman filter to predict travel times on a freeway path segment. They use historic travel time data to estimate transition parameter,$\phi(t)$, in the system model.

$$x(t) = \phi(t-1).x(t-1) + w(t-1) \quad (2.22)$$

$$z(t) = x(t) + v(t) \quad (2.23)$$

where,

$$\phi(t) = \frac{\hat{x}_H(t+1)}{\hat{x}_H(t)} \quad (2.24)$$

and, $\hat{x}_H(t)$ is the historic travel time associated with time interval t.

The CORSIM simulations are the source of their probe travel time measurements. They report maximum relative errors of 5 percent in their travel time predictions when probe vehicles represent 1% of traffic. Their prediction accuracy does not improve proportionally by increasing probe vehicles to 3% of traffic though.

Barcelo et al. (2009) proposed a discrete Kalman filter (DKF) similar to Chen and Chien (2001) to estimate and predict travel times on a 40 km long freeway segment of AP-7 Motorway between Barcelona and the French border. However, the state transition function adopted in this work is set as the ratio of travel time estimates in the last two time intervals.

$$x(t) = \phi(t-1).x(t-1) + w(t-1) \tag{2.25}$$

$$z(t) = x(t) + v(t) \tag{2.26}$$

where,

$$\phi(t) = \frac{\hat{x}(t)}{\hat{x}(t-1)} \tag{2.27}$$

They used travel time measurements obtained from 6 Bluetooth vehicle re-identification sensors on each direction that were deployed anywhere from 4 to over 15 kilometers apart from each other. Raw travel time samples first have been filtered and aggregated in one minute time intervals. It is these one minute mean travel time estimates that are used in the DKF framework to predict travel times. Later, predictions are aggregated and reported in 5 minute time intervals. A very high correlation coefficient ($R^2 = 0.9863$) between the observed and predicted time series and a prediction MARE equal to 3.54% are reported. It should be noted that long distance and intercity nature of the data used to evaluate this method, to a large extent, would explain the high quality performance of this method in forecasting travel times. In this study, speeds below 70 km/h (45mph) are assumed to signal a congested condition which in itself reflects the high speed nature of operations on the facility under study.

2.4 Models Based on Eulerian and Integrated Lagrangian Data

In cases where both Eulerian data from two endpoints of the segment and travel time (integrated Lagrangian) observations between them are available then it is possible to investigate the relationship between the two data types. The effects of Eulerian data on travel time can be modeled and evaluated using Eulerian data as independent (descriptive) variables and travel time data as dependent variable. Essentially, in this setting, travel time can be modeled as an implicit/explicit function of the available Eulerian data.

$$y = f(\mathbf{x}) \tag{2.28}$$

When function $f(.)$ is not explicitly defined, inductive or statistical methods can be used to draw conclusions on the relationship between travel time and other Eulerian data. Non-parameteric models such as k-Nearest Neighbor (k-NN) are specifically of this type. On the other hand, when function $f(.)$ is assumed to take a linear form then linear regression models can be adopted to specify the relationship between travel time and the Eulerian data. However, in general, this relationship may be non-linear in nature. Therefore, general non-linear functions such as Artificial Neural Networks (ANN) may be used for this purpose.

Downside to these methods is that huge historic data sets are needed to calibrate the associated models. The results will highly depend on the extent of the historic data set and its representation of recurrent and non-recurrent traffic conditions. Moreover, these models tend to be site dependent, a property which limits the transferability of the estimated models.

2.4.1 Inductive/Statistical (Historic Data Based) Models

2.4.1.1 k-Nearest Neighbor Methods (k-NN)

These methods belong to the non-parametric category of travel time prediction methods. This implies that no assumption is necessary to be made on error distributions. Even though large historic data sets are necessary to make good predictions using k-NN, it is anticipated that over

time historic data set gets richer and therefore performance of the k-NN method in predicting travel times is expected to improve. In this method, given the input vector $\mathbf{x}$, the following inference on the prediction error magnitude is made.

$$\|\mathbf{x} - \mathbf{x}_k\| \leq \varepsilon_k \Longrightarrow \|y - y_k\| \leq \gamma_k \qquad k = 1, \dots, K \tag{2.29}$$

where,

$\mathbf{x}_k$, is the k-th nearest neighbor to input vector,

ε_k, is the measured distance between input vector $\mathbf{x}$ and its historic k-th nearest neighbor $\mathbf{x}_k$,

y_k, is the historic travel time associated with vector $\mathbf{x}_k$, and

γ_k, is the anticipated distance between predicted travel time y and its corresponding historic k-th nearest neighbor y_k.

Basically, equation (2.29) states that if input vector, $\mathbf{x}$, is close enough to its k-th nearest neighbor, $\mathbf{x}_k$, then its output, y, will be close enough to the historic output associated with the k-th nearest neighbor, y_k. Therefore, output y may be written as the sum of the k-th nearest neighbor's output, y_k, and (an unknown) function of the measured distance between input vectors, $g_k(\varepsilon_k)$.

$$y = y_k + g_k(\varepsilon_k) \qquad k = 1, \dots, K \tag{2.30}$$

The output, y, then can be approximated as a function of all K nearest neighbor outputs.

$$y \cong h(y_1, y_2, \dots, y_K) \tag{2.31}$$

Use of average function is a popular choice for function $h(.)$ in most circumstances.

$$y \cong \sum_{k=1}^{K} y_k / K \tag{2.32}$$

Handley et al. (1998) reported the first application of k-NN method to forecast travel times on a 25 mile southbound segment of I-5 in San Diego. The method takes into account weekday versus weekend, day of week, time of day, and the 30 second average traffic speeds reported from 116

loop detectors along this segment as four features based on which similarity between current conditions and historic observations are determined. In this application three nearest neighbors are selected and the average of their associated travel time is reported as predicted travel time for current time interval. This method resulted in a MARE of up to 20% during peak period and up to 7% during off-peak period.

Clark (2003) proposed a k-NN approach to forecast 10 minute time mean speeds from a set of loop detectors on the outer loop of London beltway M25. He used a set of four consecutive speed observations in the matching process to find 8 nearest neighbors in the historic database. The distance metric used in this study is the weighted sum of squares of distances between current and historic observations contributed from each parameter included in the analysis domain. Results show that speed forecasts will be best if only speeds are included in the process. A best MARE of 5% is reported for this method in predicting speeds 10 minutes ahead. However, based on the reported results, it seems that a naïve forecast (the current observation) will perform as well as the proposed method.

Robinson and Polak (2005) tested both isolated and combined effects the choice of distance metric, value of k, and local estimation measure will have on the performance of k-NN method in forecasting an urban arterial travel times. Observed travel time data in this study are obtained using a pair of license plate matching cameras installed at two ends of a one kilometer long segment in central London. They found that k-NN method is not too sensitive to the choice of distance metric, and that a robust local estimation method is preferable to other methods. Also, they found that the optimal value of k depends on the size of the historical database. In their case study, a k-NN method with k equal to 2160 using standardized Euclidean with variance as weights for distance measurement and a locally weighted scatter plot smoothing (LOWESS) as

estimation method was found to perform optimally. This method produced MARE equal to 18% in 15 minute travel time forecasts.

You and Kim (2000) reported on an early application of k-NN method on both freeway and arterial segments in Korea. The segments they studied, however, do not seem to reflect any serious congestion conditions. Similarly, Bajwa et al. (2005) reported on an application of k-NN method on data from five long freeway segments in Tokyo metropolitan area. They reported RMSEP more than 10% for their applications in congested segments.

2.4.1.2 Linear Regression Methods

When function in equation (2.28) is assumed to be of linear type, then it can be written as follows.

$$y = A\mathbf{x} \tag{2.33}$$

where, A is the coefficients vector and can be estimated using linear regression methods such as least squares.

Kwon et al. (2000) proposed a prediction method based on linear regression with stepwise variable selection. In their model historic travel time measurements are used as dependent variable against which flow and occupancy data from loop detectors are regressed as independent variables in a least squares error sense. They used data gathered on a 6.2 mile segment of I-880 south of Oakland, California for model evaluation. This data set includes measurements from double loop speed stations located at approximately one-third of a mile apart as well as probe travel time data (364 trips) from 20 weekday mornings. At 5 minute ahead, this linear regression method resulted in 9-15% MARE in travel time predictions. Obviously, this model is site specific and should be re-estimated for other segments using their corresponding data sets.

Zhang and Rice (2003) proposed a time varying coefficient (TVC) linear model to improve upon

a naïve predictor based on current speeds at two ends of a freeway segment. The method requires a large historic database to calibrate prediction model's coefficients. They reported on the method's performance on the north-bound direction of I-880 data set which was used by Kwon et al. (2000). While using historic dataset provides a slightly more than 10 percent MARE on travel time prediction, the TVC model has roughly 6% error in current travel time estimation and about 11% error at 30 minute forecasts.

Chakroborty and Kikuchi (2004) proposed a simple linear regression model to estimate auto travel times based on measured bus running travel times. The latter is obtained using GPS devices installed on buses and is equal to total bus travel time minus times bus spends stopped at stations along the segment. The method evaluations on five arterial segments in northern New Castle County, Delaware revealed that in the worst case 77% of predictions made were within 10% of floating car measurements. In this study, there were 28 to 30 travel time measurements made at each site.

Liu and Chang, 2006 reported on attempts to calibrate single linear regression models to account for increase in travel time due to accumulations on segments with constant and variable capacity drops at the downstream. Data obtained from CORSIM micro-simulation runs is used to estimate the models. Applying the method in practice is difficult since model calibration requires a large historic database and the count data to be used in the method are not accurately available.

2.4.1.3 Artificial Neural Network (ANN) Methods

ANN is a general non-linear function approximation system that is inspired by generic functions of biological neural networks. The idea behind ANN is that data processing happens at many simple data processing units called neurons. Typically, in an ANN these neurons are organized in layers in a feed-forward network. Associated with each link in the network is a weight that should be determined using a training procedure such as error back-propagation. Input to each

neuron is the weighted sum of outputs from neurons in the previous layer. Neurons act as a switch and depending on the input strength produce an output determined by an activation function. Identity, linear, binary step and sigmoid (S-shaped) functions such as logistic and hyperbolic tangent functions are among popular activation functions (Fausett, 1994). A feed forward ANN with L layers can be concisely represented as the following recursive equation.

$$y = f_L\left(\mathbf{w}'_{\mathbf{L-1},L} f_{L-1}\left(\dots f_1\left(\mathbf{w}'_{\mathbf{0},\mathbf{1}}\mathbf{x} + \boldsymbol{b}_{\mathbf{1}}\right) \dots + \boldsymbol{b}_{\boldsymbol{L-1}}\right) + \boldsymbol{b}_{\boldsymbol{L}}\right) \quad (2.34)$$

where,

$\mathbf{w}_{\boldsymbol{l},\boldsymbol{l+1}}$, is the specified weight matrix between neurons in consecutive layers l, and $l+1$,

$\boldsymbol{b}_{\boldsymbol{l}}$, is the bias vector in neurons of layer l, and

$f_l(.)$, is the vector of activation functions belonging to neurons of layer l.

Park and Rilett (1998) proposed a clustering and artificial neural network method to forecast travel times on an urban freeway. They use AVI travel time data on link segments from just over one mile to 5 mile long on eastbound US-290. This is part of the automatic tolling system TranStar in Houston, Texas. Application of this method resulted in 5 minute travel time forecasts with over 8% MARE. Errors nearly doubled in 25 minute forecasts when MARE reached 16%.

Rilett and Park (2001) report on applying a spectral basis neural network to directly forecast freeway corridor travel times. In this method an extra layer is added to the front of ANN which implements Fourier transform. The transformed basis functions then will be used in a series of hidden layers to build a forecast for corridor travel times. Again, performance of the method on a 12.8 km segment of eastbound US-290 in Houston is reported. MAPE in 5 minute forecast has been about 6% while this same measure for 25 minute ahead forecasts has been more than 15% which is not that different from their older results.

In recent years, ANN methods with transformed input data have become more common place.

Hamad et al. (2009) used Hilbert-Huang decomposition of the ILD speed signals as input to a speed predicting ANN. They tested this method on I-66 data. Their prediction MARE for 5 to 25 minute ahead during morning peak hour ranged from 6 to 10 percent.

2.4.2 Traffic Flow Theory Models

Traffic flow theory models can be categorized into two major groups. This classification is based on the level of detail at which a traffic stream is being modeled. Microscopic models track the movements of individual vehicles in traffic. These movements typically fall under two umbrella categories: car following and lane changing models. Microscopic models are computationally intensive and very difficult to calibrate and verify.

On the other hand, macroscopic traffic models deal with characteristics of a group of vehicles at an aggregate level. Variables such as flow and density are passage rate of vehicles at a cross section and their presence rate over a stretch of highway, respectively. Based on definition, these variables can be shown to be related through a third variable, namely space mean speed. This constitutive relationship along with an assumption on the form of dependence between speed and density leads to the so-called Fundamental Traffic Diagram (FTD).

Macroscopic models are in fact conservation laws expressed in the form of Partial Differential Equations (PDE). These models can be solved using exact methods such as method of characteristics. In real world applications, in general, it is difficult to obtain the exact solutions. Mesoscopic models approximate solutions to these conservation laws by breaking the solution domain into a series of smaller sub-domains. Finite Difference (FD) methods such as up-winding and Finite Element (FE) methods such as Galerkin are typically used to approximate the evolution of traffic variables over time and space.

2.4.2.1 Microscopic Simulation Models

Liu et al. (2006) reported on an online travel time prediction system customized for Ocean City,

Maryland. The system receives data from 10 stationary sensors sparsely located along 30 miles of US-50 and MD-90 between Salisbury and Ocean City, Maryland. In their study they used a calibrated micro-simulation model based on CORSIM software package to predict travel times in the system. Forecast traffic volumes at detector locations needed for micro-simulation module are determined from a historic database using a nearest neighbor method. No specific measure of accuracy regarding predicted travel times is reported.

2.4.2.2 Mesoscopic Simulation Models

Waller et al. (2007) adopted an ARIMA(3,1,2) to forecast inflows to the freeway segment under study, then they used a meso-simulation technique called cell transmission model (CTM) to simulate propagation and movements of vehicles inside the segment. Later, based on cumulative flow curves at the segment endpoints they were able to forecast travel time. On a 3 mile freeway segment, they reported 10 to 23 percent RMSE on travel times predicted 5 minutes ahead using this method when compared with travel times obtained from VISSIM micro-simulation.

2.4.2.3 Hybrid Models

Zou et al. (2007) proposed a method for travel time estimation over long freeway segments. Their method is an extension of Coifman (2002), which makes use of occupancy and speed data from stationary detectors located at either end of the segment. They first identify different recurrent traffic patterns based on a historic data set. Then for each pattern, they calibrate a parameterized model to estimate travel times over the segment assuming that speeds at each detector can be extended using a linear relationship to represent the average travel speeds on each half segment. Later, based on a piecewise exponential speed-occupancy relationship and the assumption that traffic conditions at detectors will propagate with a constant speed within the segment, an iterative method for trajectory approximation is proposed. Performance of this method is reported in comparison with travel times obtained from vehicle re-identification

conducted on two days' worth of video recordings at the endpoints of an over 10 mile long segment of I-70 between US-40 and I-695 east of Baltimore, Maryland. Results are counter-intuitive in the sense that the proposed method resulted in higher errors in free flow and heavy congestion conditions rather than in moderately congested periods. In free flow conditions errors up to 8.7% in travel time estimation are reported.

Yu et al. (2008) developed a hybrid model to predict travel times on a 7 mile long segment of US-50 leading to Ocean City, Maryland. They decomposed travel time to a trend and a variation component. A fuzzy weighted average of clusters in the historic data base is used to estimate the trend term, while a cluster-based artificial neural network calibrated again on historic data base is used to predict travel time variations. Performance of the proposed hybrid model is compared with results obtained from micro-simulation software CORSIM. An average error of 8.7% in predicted travel times throughout the day is reported.

2.5 State Space Models

State space models provide a systematic general framework to represent the dynamics of the system no matter how complicated the system under investigation is. Additionally, it allows for incorporation of various measurements that become available dynamically over time into the estimation process. In general, a state space model consists of a system and a measurement equation. The idea is, in some cases it is difficult to make direct observations of a system state, instead it might be easier to observe and measure its correlated variables. Then the problem is to dynamically obtain best estimates of the system state by observing the correlated variables' evolution over time.

In its simplest form both the system and measurement equations in a state space model are linear. The following is an example of a discrete-time linear state space model:

$$v_{n+1} = M_n v_n + w_n \quad (2.35)$$

$$y_n = H_n v_n + u_n \quad (2.36)$$

where,

v_n, is the $N \times 1$ column vector of state variables at time step n

y_n, is the $M \times 1$ column vector of measured variables at time step n

M_n, is the $N \times N$ square transition matrix representing dynamics of the system at time step n,

H_n, is the $M \times N$ state to measurement transition matrix of the system at time step n,

w_n, is the $N \times 1$ column vector of state dynamics errors at time step n, and

v_n, is the M$\times$ 1 column vector of measurement errors at time step n.

The best linear estimate of a state space model in the least square sense is obtained by Kalman filtering (Kalman, 1960 and 1961).

Nanthawichit et al., 2003 applied standard Kalman filter to the linearized approximation of a discretized version of the Payne's traffic model. The method primarily uses loop detector volume and speed measurements to estimate density and speed in a set of cells over time. Measurements from probe vehicles also may be included in this method in a very simplistic way. Average probe vehicle speed in each cell is regarded as measurement from an imaginary loop detector, if the cell in question does not include a loop detector. However, if a loop detector exists in the cell and we have two measurements from loop detector and probe vehicle in that cell at the same time interval, then the average of two measurements is used as measurement from the cell. Data generated by simulation software INTEGRATION has been used to evaluate methodology's accuracy. The suggested combined use of a traffic model along with stationary sensor and probe data in a Kalman Filter is shown to improve the travel time predictions by up to 36% compared to the autoregressive Kalman filter method proposed by Chen and Chien, 2001 which only uses

probe data.

Sun et al., 2004 proposed a Monte Carlo method based on a binary switching mode traffic model that only distinguishes between free-flow and congestion modes. In this method, first a fixed number of mode sample sequences with highest probability are identified, and then on each of these mode sample sequences a time varying Kalman filter is applied to estimate continuous traffic states (density). The a posteriori estimates of the continuous states are then computed as the weighted average of estimates from each Kalman filter. They use real stationary data from PeMS as well as simulation results from VISSIM to evaluate their method. They offer visual evidence that their proposed method is working well in estimating traffic mode; no quantitative measures are given though.

Chu et al, 2005 assumed traffic flow and density are homogeneous on a freeway segment which may even include multiple on/off ramps. Also, they assume that all entering and exiting boundary flows to and from the segment are measured by means of stationary traffic sensors such as loop detectors, and they receive travel time measurements from probe vehicles traversing the segment every once in a while. An adaptive Kalman filter is proposed in which density is adopted as state variable and travel time measurements are simply related to the average density on the segment through a time-varying coefficient. A simple method for estimation of noise statistics (mean and variance of errors in system and measurement equations) based on an earlier work is given. Data generated using PARAMIC simulation at 30 second intervals on a 0.82 mile freeway segment with one on- and one off- ramp is used to evaluate the proposed method. They reported 8% mean relative errors in travel time estimates under recurrent morning peak conditions with a 5% probe rate, while under non-recurrent conditions (10 minute long incident blocking the right lane of freeway) this error measure is increased to about 10%.

Wang and Papageorgiou, 2005 reported on using an extended Kalman filter to estimate density and speed on 500 meter freeway segments every 10 seconds. In their system equation a modified Payne-Witham model for dynamic speed estimation is used. Taylor series expansions are used to linearize the model equations at each point in time. Flow and speed measurements used in the estimation come from stationary traffic sensors at the boundaries of the freeway segment. Traffic data is generated using simulation based on the same traffic models as in the Kalman filter. In other words, Kalman filter is utilized o estimate traffic states given that we have full knowledge of actual traffic dynamics in the system. However, Kalman filter is used as a tool to identify the system state in presence of model and measurement noise. In their application, root mean square errors of the order 20% and 14% are reported for density and speed estimates, respectively. In the case of speed estimates, average absolute RMSE has been about 14 kilometers per hour (almost 9mph) on a 5 kilometer stretch of freeway. In a later work, Wang et al., 2007 used collected data from a 4.1 kilometer German highway to demonstrate the performance of their proposed methodology. In this case, no quantitative error measures for state estimates are reported.

Work et al., 2008 proposed an Ensemble Kalman filtering (EnKF) approach for highway traffic estimation in the presence of both stationary and probe vehicle data. They used a Velocity based Cell-Transmission Model (CTM-v) with a Greenshield's type fundamental diagram which makes it possible to work directly with measured speeds. They ran tests on a simulation model calibrated for I-880. This method resulted in 25% average relative error on speed estimates at 5% probe penetration rate.

Herrera and Bayen, 2009 used Cell-Transmission Model (CTM) with a triangular fundamental diagram to estimate density given a combination of boundary and probe measurements. Newton

relaxation and discrete Kalman filter are two methods that they used to estimate traffic conditions. They used data collected in Next Generation SIMulation (NGSIM) project on US highway 101 in California, as well as GPS probe data from Mobile Century data collection effort on interstate 880 in California.

Claudel and Bayen, 2008 proposed a method based on viability theory in optimal control to estimate a lower and upper bound for the number of vehicles that are initially present on the road segment under investigation based on both stationary and probe data. Then, using the conservation of vehicles principle this method is capable of estimating a range for travel time between the two end points of the segment. This method does not take into consideration presence of on and off ramps between the two end points. They tested their proposed method on US-101 dataset from NGSIM and I-880 from Mobile Century data collection effort. A mean relative error higher than 8% on travel time estimates is reported using this method.

Chen et al. (2011), and Sadek and Rakha (2012) proposed shock-fitting approach for a general flux function. Since, the PDE in general form is not conservative, Godunov upwind scheme cannot be directly applied. Instead, a Courant-Isaacson-Reece (CIR) scheme in regions where no shocks are present is adopted. In presence of shocks (discontinuities or jumps) the Rankine-Hugoniot jump condition should be still satisfied (LeVeque 1992). A shock-fitting procedure is introduced to adjust the CIR speeds at suspect points near a shock which in turn requires tracing individual characteristic wave speeds at all points and at all times. For this reason, while the method is theoretically appealing it cannot be readily applied in large-scale practical cases.

Mihaylova and Boel, 2004 used a Particle Filter (PF) to estimate traffic variables from a nonlinear state-space model. PF is essentially a Bayesian recursive estimation approach analogous to a Monte Carlo simulation. Therefore, PF is computationally expensive but is most

accurate in the case of nonlinear state-space models. Their numerical experiments on an undisclosed 0.5 kilometer long four lane freeway segment reported errors up to 10% in speed estimates.

Table 1 summarizes the distinctive features of relevant traffic state estimation studies reported in the literature which are particularly based on state-space models. Based on this table a few points should be noted. First, not many studies reported on the accuracy of travel time predictions. Second, as expected, time interval size plays a significant role in the accuracy of estimates. Third, travel time data has never been systematically incorporated into the estimation process. The only reported work that directly combines probe data as travel time with stationary data (Chu et al. 2005) does so through a simplifying assumption that travel time is an adaptive coefficient of density in the segment under study.

This dissertation builds on Work (2010) by incorporating travel time measurements into the estimation process. The CTM-v model coupled with a novel travel time model enable the direct assimilation of travel time measurements along with boundary speed measurements. Unscented approach is adopted to optimally filter the resulting nonlinear state-space model. Both Kalman and H-infinity type filters are considered. This setting allows for the joint speed and travel time estimations and predictions. Real world data from NGSIM project (US-101) is used to evaluate the performance of the proposed models and methods at very fine granularity levels (2 sec and 4 sec time intervals).

Table 1. Summary features of traffic speed/travel time estimation studies using state-space models.

Author(s)	Year	Traffic Model	Measurement(s)	Estimation Method	Data Source	Facility Type	Time Interval	Estimation Variable	Prediction Variable	Accuracy
Chen, Chien	2001		-Probe (travel time)	Auto-Regressive Kalman	CORSIM	Freeway (I-80 in NJ)	5 min	Travel Time	Travel Time (5 min)	MARE ~2% @ 1-3% probe
Treiber, Helbing	2002		-Stationary (flow, speed, density)	Adaptive Smoothing	Double ILD	Freeway (A8, A9, A5, Germany)	1 min	Density	Density (20 min)	Visual
Nanthawichit et al.	2003	Payne (linearized, and discretized)	-Stationary (flow, speed) -Probe (speed) -Combined	Kalman Filter	INTEGRATION	Freeway (Yokohane, Japan)	10 sec est., 3 min pred.	Speed	-Speed -Travel Time	MARE<3-26% MARE <4% @ 3% probe
Mihaylova, Boel	2004		-Stationary (flow, speed, density)	Particle Filter	METANET		10 sec	Flow Speed Density		
Sun et al.	2004	SMM	-Stationary (flow, density)	Mixture Kalman Filter	PeMS (ILD) VISSIM	Freeway (I-210 in CA)	2 sec	Speed		Visual
Chu et al.	2005	LWR (discretized)	-Stationary (flow, density) -Probe (travel time) -Combined	Adaptive Kalman Filter	PARAMICS	Freeway (I-405 in CA)	30 sec			MARE ~10% @ 5% probe
Wang, Papageorgiou	2005	Modified PW (linearized)	-Stationary (flow, speed)	Extended Kalman Filter	Kalman Filter		10 sec	Density Speed		MARE <19-21% MARE ~14%
Wang, Papageorgiou	2007	Modified PW (linearized)	-Stationary (flow, speed)	Extended Kalman Filter	ILD	Freeway (A92, Germany)	10 sec	Flow Speed Density		Visual
Herrera, Bayen	2008	CTM (triangular flux)	-Stationary (density) -Probe (position, speed)	-Newton Relaxation -Kalman Filter	NGSIM Mobile Century	Freeway (US-101, I-880 in CA)	1.2 sec 8 sec	Density		?
Claudel, Bayen	2008	Moskowitz HJ PDE	-Stationary (density) -Probe (position, speed)	LP	NGSIM Mobile Century PeMS	Freeway (US-101, I-880 in CA)		Travel Time		MARE >8% @ 5% probe
Work et al.	2008	CTM-v	-Stationary (speed) -Probe (position, speed)	Ensemble Kalman Filter	PARAMICS Mobile Century	Freeway (I-880 in CA)	2 sec	Speed		MARE 25% @ 5% probe
Barcelo et al.	2009		-Probe (travel time)	Auto-Regressive Kalman	Pilot project	Freeway (AP-7 in Spain)	5 min	Travel Time	Travel Time	MARE 3.5%

Chapter 3: First-Order Continuum Traffic Flow Model

In this chapter, first a well-known first-order continuum traffic flow model is adopted to represent the dynamics of the system. The proposed traffic flow model is a hyperbolic Partial Differential Equation (PDE) which specifies a conservation law. To be well-posed this PDE needs to be solved in presence of initial and/or boundary conditions. An equivalent form of this model in terms of speed is derived. This model provides a theoretical framework to describe and analyze traffic processes on a variety of roadway facilities. Currently, this model is used in an array of traffic operation and control applications worldwide. Therefore, it is essential to have both efficient and accurate solution methods for this model. A finite difference method for numerical solution of the velocity based equivalent of the first-order continuum traffic flow model is proposed. Two major speed-density relationships are introduced for use in the formulation and solution of the proposed traffic flow model.

3.1 Traffic Model (LWR-v)

Continuum traffic flow theory is a powerful tool to describe the evolution of macroscopic traffic parameters over time and space. This is in contrast to microscopic models of traffic flow which generally require meticulous handling of individual vehicles movements in the traffic stream. The most basic continuum traffic flow theory builds on two basic physical principles that is conservation of vehicles and the fundamental relationship between flow rate, density and speed. Conservation principle states that no vehicle is added or lost in traffic at any time other than the ones that enter or exit through the boundaries. This basic continuum theory was first proposed by Lighthill and Whitham (1955) and Richards (1956). Despite its simplicity, and therefore its inherent limitations (Daganzo, 1997), the so called kinematic wave theory or LWR model provides a good approximation to the dynamics of traffic flow which has proved to be useful for

most practical purposes.

Even though our ability to directly measure different traffic parameters has dramatically increased over the years, measurements are still widely different in terms of their accuracy and reliability. For instance, flow rate and density both can be obtained as a result of simple counting processes performed at a single point or at a pair of points, respectively. However, it is ironic that in practice, large inconsistencies between counts in consecutive stations (with no exit or entrance in between) exist. In the case of loop detectors this "drift" phenomenon is well known. In addition, at a macroscopic level, definition of density is a bit ambiguous in the sense that the length over which concentration of vehicles would affect a driver's behavior under normal conditions is not specified.

In contrast, speed measurements which theoretically are expected to be more difficult to obtain have proved to be a far more reliable source of traffic data. That is why, in this dissertation, a velocity-based equivalent of the LWR model is adopted to model traffic dynamics.

The rest of this section is organized as follows. First, the first-order continuum traffic flow model known as LWR model is briefly presented. Second, the finite difference based numerical solution to LWR model known as Cell Transmission Model (CTM) is presented. Two invertible speed-density relations are described. Finally, the derivation of the first-order speed based continuum traffic flow model based on LWR model is presented. Also, for the sake of completion, the speed based finite difference method equivalent to CTM presented by Work (2010) is summarized.

3.1.1 LWR Model

The first-order continuum traffic flow model proposed by Lighthill, Whitham and Richards (LWR) is considered in its differential from:

$$\rho_t(x,t) + q_x(x,t) = 0 \tag{3.1}$$

where,

$\rho(x,t)$ denotes traffic density at time t and at point x along the highway, and

$q(x,t)$ denotes traffic flow rate at time t and at point x along the highway.

This is the conservation law which essentially implies that no vehicle is born or lost along the highway. This Partial Differential Equation (PDE) is originally borrowed from hydrodynamics but has found widespread use in modeling vehicular traffic flow despite obvious differences between the two fields.

3.1.2 Cell Transmission Model

In practice, the LWR model presented in the previous section has to be approximated. In this section an approximation to LWR model proposed by Daganzo (1994) is briefly presented. Cell Transmission Model (CTM) is a finite difference PDE solution method which adopts a Godunov numerical scheme in order to approximate flow rates across cell boundaries.

Similar to any finite difference approach, in CTM both time and space dimensions are discretized. Time is divided into N time intervals $\{t_n | n = 0,1,\dots,N\}$ each of length $\Delta t = T/N$, and space is divided into M space cells $\{x_i | i = 0,1,\dots,M\}$ each of length $\Delta x = (b-a)/M$. To each space cell x_i at time interval t_n, an average density ρ_n^i is assigned. Then, density evolution at each space cell over time is given by,

$$\rho_{n+1}^i = \rho_n^i - \frac{\Delta t}{\Delta x}\left[G\left(\rho_n^i, \rho_n^{i+1}\right) - G\left(\rho_n^{i-1}, \rho_n^i\right)\right] \tag{3.2}$$

where the flow rate G between the two states ρ_1 and ρ_2 is defined as,

$$G(\rho_1,\rho_2) = \begin{cases} Q(\rho_2) & \text{if } \rho_c \le \rho_2 \le \rho_1 \\ Q(\rho_c) & \text{if } \rho_2 \le \rho_c \le \rho_1 \\ Q(\rho_1) & \text{if } \rho_2 \le \rho_1 \le \rho_c \\ \min(Q(\rho_1), Q(\rho_2)) & \text{if } \rho_1 \le \rho_2 \end{cases} \tag{3.3}$$

where ρ_c is the density corresponding with the maximum of the concave flux function $Q(\rho)$.

Figure 5 provides a graphical representation of the Godunov flux function in the ordered pair density (ρ_1, ρ_2) domain. Note that when upstream (sending) cell is congested $(\rho_c \leq \rho_1)$ and downstream (receiving) cell is not congested $(\rho_2 \leq \rho_c)$, then Godunov flux at the boundary of two cells is constant at the maximum possible flow (capacity). In other cases, Godunov flux is variable and equals either upstream or downstream flow rates. The line connecting (ρ_c, ρ_c) and $(0, \rho_{max})$ is where standing shockwaves form, and therefore the Godunov function is not differentiable.

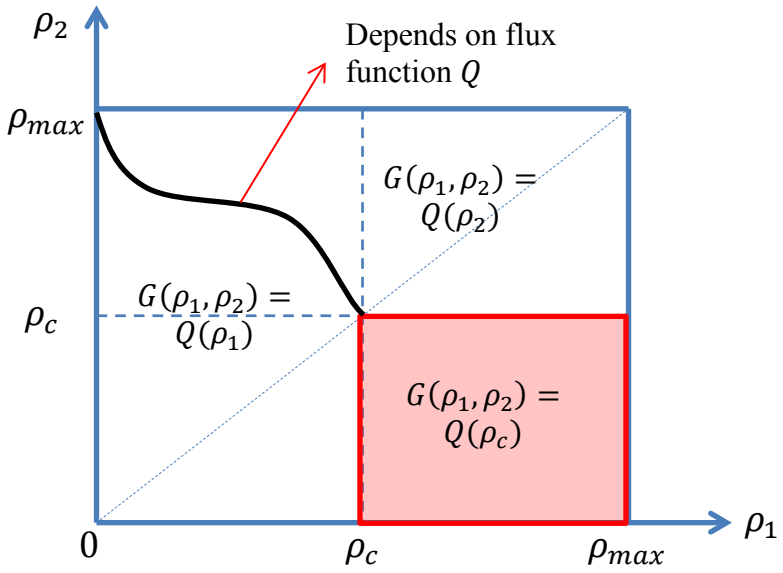


Figure 5. Godunov flux function representation in the density domain.

Bardos et al. (1979) have shown that nonlinear hyperbolic conservation laws of type (3.1) with strictly concave flux and initial condition (3.4) and the weak boundary conditions (3.5) and (3.6) in space $]a, b[\times]0, T[$ are well-posed.

$$\rho(x, 0) = \rho_0(x) \qquad , \forall x \in]a, b[\tag{3.4}$$

$$\begin{cases} \rho(a, t) = \rho_a(t) \text{ or} \\ Q'\big(\rho(a, t)\big) \leq 0 \text{ and } Q'\big(\rho_a(t)\big) \leq 0 \text{ or} \\ Q'\big(\rho(a, t)\big) \leq 0 \text{ and } Q'\big(\rho_a(t)\big) > 0 \text{ and } Q(\rho(a, t)) \leq Q\big(\rho_a(t)\big) \end{cases} \tag{3.5}$$

$$\begin{cases} \rho(b,t) = \rho_b(t) \text{ or} \\ Q'\big(\rho(b,t)\big) \geq 0 \text{ and } Q'\big(\rho_b(t)\big) \geq 0 \text{ or} \\ Q'\big(\rho(b,t)\big) \geq 0 \text{ and } Q'\big(\rho_b(t)\big) < 0 \text{ and } Q(\rho(b,t)) \leq Q\big(\rho_b(t)\big) \end{cases} \quad (3.6)$$

Note that above boundary conditions are in the weak form (designed to ensure that entropy solution to LWR PDE exists and is unique). The first line in each case is however, in the strong case where solution at boundary takes the value of the boundary condition. The second and third lines in each case attempt to maintain the condition that no outside waves enter the solution domain.

3.1.3 Speed-Density Relations

In vehicular traffic flow it is generally believed that speed and density are negatively correlated. Different empirical relationships between traffic speed and density are introduced over the years. (Del Castillo and Benitez, On the Functional Form of the Speed-Density Relationship-I: General Theory 1995). The first speed-density relationship proposed by Greenshields (1935) has a simple linear form:

$$v = V_{GS}(\rho) = v_f(1 - \rho/\rho_{max}) \quad (3.7)$$

where, v_f is the free flow speed, and ρ_{max} is the jam density of highway under prevailing conditions. Also, under stationary traffic conditions, by definition the following relationship between flow, speed and density holds:

$$q = \rho\, v \quad (3.8)$$

Substituting for speed from (3.3) in (3.4), the following quadratic relation between flow and density is obtained:

$$q = Q_{GS}(\rho) = v_f(\rho - \rho^2/\rho_{max}) \quad (3.9)$$

Recently, more sophisticated speed-density models are proposed. For instance, triangular fundamental diagram (Daganzo-Newell) is a two-regime model which assumes a constant free-

flow speed throughout the uncongested regime and a hyperbolic speed in the congested phase:

$$v = V_{DN}(\rho) = \begin{cases} v_{max} & ,\rho \le \rho_c \\ -w_f(1 - \rho_{max}/\rho) & ,\rho > \rho_c \end{cases} \tag{3.10}$$

where, w_f is a parameter indicating the wave speed. Note that to assure continuity of speed at critical density ρ_c the following should hold:

$$w_f = v_{max}(\rho_{max}/\rho_c - 1) \tag{3.11}$$

A close alternative to the triangular fundamental diagram (in fact a combination of Greenshields and triangular models) is the multi-regime hyperbolic-linear model. In the uncongested regime, this model maintains a linear relationship between speed and density, while in the congested regime it takes a hyperbolic form:

$$v = V_{HL}(\rho) = \begin{cases} v_{max}\left(1 - \frac{\rho}{\rho_{max}}\right) & ,\rho \le \rho_c \\ -w_f\left(1 - \frac{\rho_{max}}{\rho}\right) & ,\rho > \rho_c \end{cases} \tag{3.12}$$

And, the corresponding continuity condition is now given by the following expression:

$$w_f = v_{max}(\rho_c/\rho_{max}) \tag{3.13}$$

3.1.4 Velocity-Based Cell Transmission Model

Given an invertible speed-density relationship, at each discrete point (space cell i, and time step n) the CTM model can be rewritten in terms of the traffic speeds:

$$v_{n+1}^i = V\left(\rho_{n+1}^i\right) = V\left[V^{-1}\left(v_n^i\right) - \frac{\Delta t}{\Delta x}\left[\tilde{G}\left(v_n^i, v_n^{i+1}\right) - \tilde{G}\left(v_n^{i-1}, v_n^i\right)\right]\right] \tag{3.14}$$

where,

$$\rho = V^{-1}(v) \tag{3.15}$$

denotes the density as a function of speed, and

$\tilde{G}(v_1, v_2)$, is the speed-based Godunov type numerical flux function defined as

$$\tilde{G}(v_1, v_2) = \begin{cases} \tilde{Q}(v_2) & \text{if } v_1 \le v_2 \le v_c \\ \tilde{Q}(v_c) & \text{if } v_1 \le v_c \le v_2 \\ \tilde{Q}(v_1) & \text{if } v_c \le v_1 \le v_2 \\ \min\left(\tilde{Q}(v_1), \tilde{Q}(v_2)\right) & \text{if } v_1 \ge v_2 \end{cases} \tag{3.16}$$

where v_c is the speed corresponding with the maximum of the concave flux function $\tilde{Q}(v)$. Note that CTM-v model (3.14) is in general highly nonlinear in terms of the speed arguments.

Figure 6 provides a graphical representation of the Godunov flux function in the ordered pair speeds (v_1, v_2) domain. Note that when upstream (sending) cell is congested $(v_1 \le v_c)$ and downstream (receiving) cell is not congested $(v_c \le v_2)$, then Godunov flux at the boundary of two cells is constant at the maximum possible flow (capacity). In other cases, Godunov flux is variable and equals either upstream or downstream flow rates. The line connecting (v_c, v_c) and $(v_{max}, 0)$ is where standing shockwaves form, and therefore the Godunov function is not differentiable.

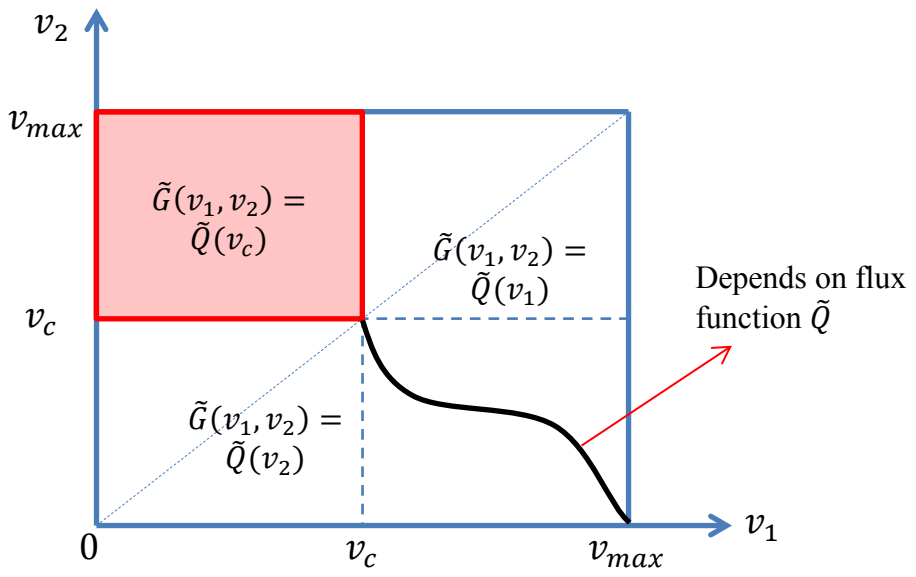


Figure 6. Godunov flux function representation in the speed domain.

Similar to the CTM case, the CTM-v model (3.14) with the initial condition (3.17) and the weak boundary conditions (3.18) and (3.19) in space $]a, b[\times]0, T[$ is well-posed.

$$v(x,0) = v_0(x) \qquad ,\forall x \in]a,b[\tag{3.17}$$

$$\begin{cases} v(a,t) = v_a(t) \text{ or} \\ \tilde{Q}'\big(v(a,t)\big) \leq 0 \text{ and } \tilde{Q}'\big(v_a(t)\big) \leq 0 \text{ or} \\ \tilde{Q}'\big(v(a,t)\big) \leq 0 \text{ and } \tilde{Q}'\big(v_a(t)\big) > 0 \text{ and } \tilde{Q}(v(a,t)) \leq \tilde{Q}\big(v_a(t)\big) \end{cases} \tag{3.18}$$

$$\begin{cases} v(b,t) = v_b(t) \text{ or} \\ \tilde{Q}'\big(v(b,t)\big) \geq 0 \text{ and } \tilde{Q}'\big(v_b(t)\big) \geq 0 \text{ or} \\ \tilde{Q}'\big(v(b,t)\big) \geq 0 \text{ and } \tilde{Q}'\big(v_b(t)\big) < 0 \text{ and } \tilde{Q}(v(b,t)) \leq \tilde{Q}\big(v_b(t)\big) \end{cases} \tag{3.19}$$

3.2 Summary

In this chapter a first-order and velocity-based continuum traffic flow model was introduced. Possible model and input data were discussed. Also, potential output from proposed modeling effort has been generally identified. Various features of desirable solution approaches to the proposed research problem in terms of prediction horizon, efficiency and accuracy have been generally specified.

Chapter 4: Travel Time Model

In this chapter, a first-order partial differential equation (PDE) model relating local variations of travel time with local speeds is presented. Travel time can be defined relative to either the start point or the end point of a trip. At any point along the trip, the former definition amounts to a retrospective (a posteriori) view of travel time, while the latter leads to an anticipative or predictive (a priori) definition of travel time. The proposed PDE model is capable of dealing with both travel time definitions with very slight variations required in the underlying model.

In addition, some desirable properties of travel time such as stability and first-in first-out (FIFO) are briefly presented and their implications in terms of the proposed travel time models are discussed. Efficient numerical solution schemes based on finite difference approximation of partial differentials are introduced. Also, appropriate boundary and initial value conditions required in solving the proposed PDE models are derived and presented. Finally, a discussion on the benefits and potential applications of proposed models is presented.

4.1 Preliminaries

A vehicle travel time can be defined as the line integral of inverse of vehicle speed along its trajectory. Based on this definition, it is clear that either the vehicle trajectory should be known, or the inverse problem of finding speeds based on travel time will be under-determined.

$$\tau = \int \frac{1}{\sqrt{1+v(X(s),s)^2}} ds \tag{4.1}$$

Given the speed field, constructing vehicle trajectory and travel time estimation is a direct problem. Theoretically, travel time estimation given speeds is a straight-forward process. However, in practice this process is very inefficient since it requires numerical approximations to the above line integral. Also, travel time estimates based on trajectory construction tend to have

poor quality since errors in speed estimates quickly add up in the process. Instances of this method resulted in errors up to 10 percent in travel time estimates over a half a mile segment.

Alternatively, in this dissertation, instead of the integral representation the focus is on the local travel time variations. In other words, differential equations relating travel time and local speeds seem to be more useful in this setting.

Similar to other traffic states, predictive (retrospective) travel time τ (θ) can be defined on the space-time domain as the minimum distance from (to) any given point (x, t) in the solution domain to (from) the downstream (upstream) boundary x_d (x_u). Figure 7 illustrates the travel time definitions at a given point (x, t) with respect to the upstream and downstream of a segment.

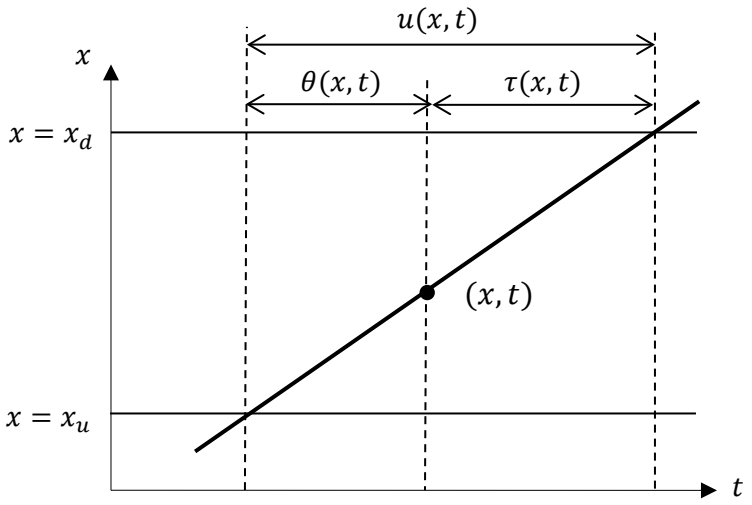


Figure 7. Travel time definitions.

Figure 8 illustrates the space-time domain with iso-distance contours representing the set of points from which travel time to the downstream boundary are the same. Also, this representation suggests that travel times observed at upstream (or any other point along the highway) is in fact a cross-section of various contours.

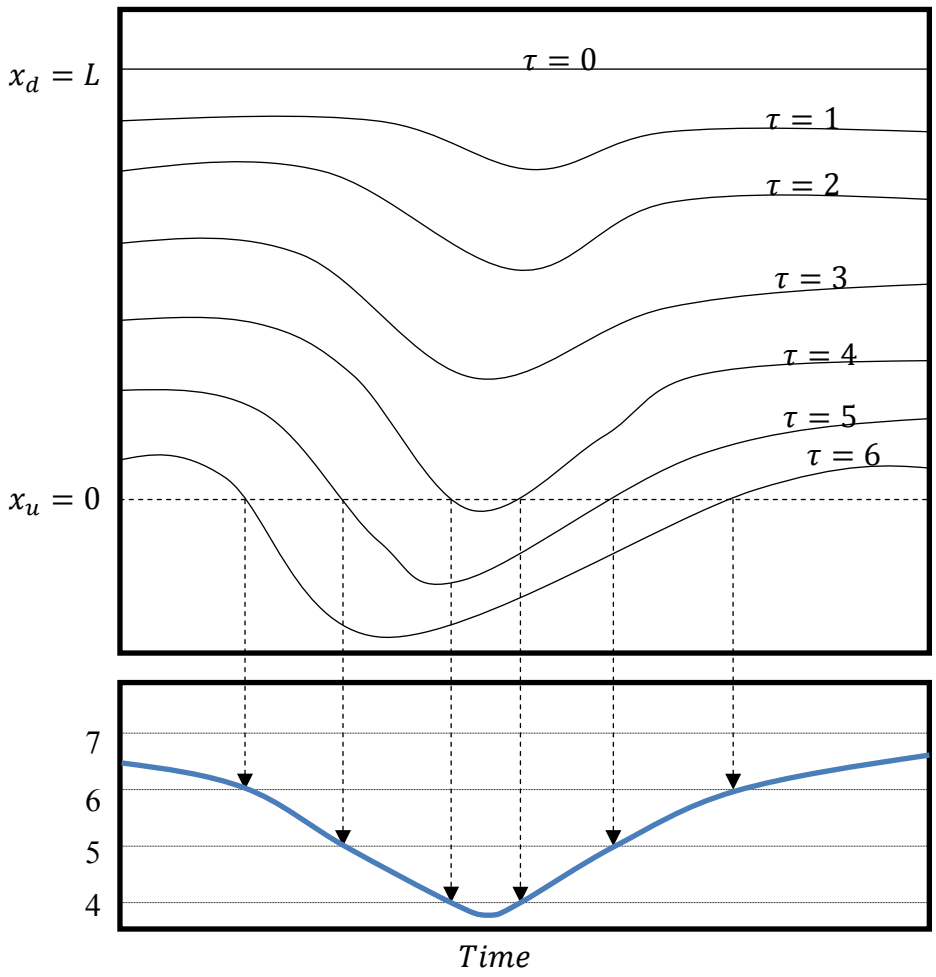


Figure 8. Concept of travel time as distance from downstream boundary in a wave propagation paradigm.

These definitions along with the assumption of smoothness (of underlying travel time function) result in partial differential equation representation of travel time variations in the space-time domain. Finite difference schemes may be used to numerically solve the proposed travel time model.

In the following sections derivation of first-order travel time PDE models are presented. Then, some desirable properties of travel time such as stability and first-in first-out (FIFO) are discussed. It is shown that proposed travel time models under certain circumstances uphold these properties. Finite difference schemes along with appropriate boundary conditions to efficiently

solve the proposed travel time models are presented.

4.2 First-Order Travel Time Model Derivation

Let $\tau(x,t)$ represent travel time from a point (x,t) in space x and time t coordinates to a given downstream point x_d. This definition specifies the so called a priori travel time since at point (x,t) travel time $\tau(x,t)$ has not yet realized. It should be noted that in what follows derivations and proposed solution schemes are based on this definition of travel time. However, it will be trivial to derive similar models in the case of a posteriori travel times.

Assuming smoothness and therefore existence of derivatives we can use Taylor's function expansion to obtain the travel time near a point (x,t) as

$$\tau(x+dx,t+dt) = \tau(x,t) + \tau_t dt + \tau_x dx + O((dt)^2) + O((dx)^2) \tag{4.2}$$

Figure 9 illustrates these definitions and the above relationship where the pair of points (x,t) and $(x+dx,t+dt)$ are located on a single vehicle trajectory. In this case, it is obvious that travel time to downstream at the second point has a simple relationship with the travel time at the first point, or more specifically the difference between the two travel times is equal to dt:

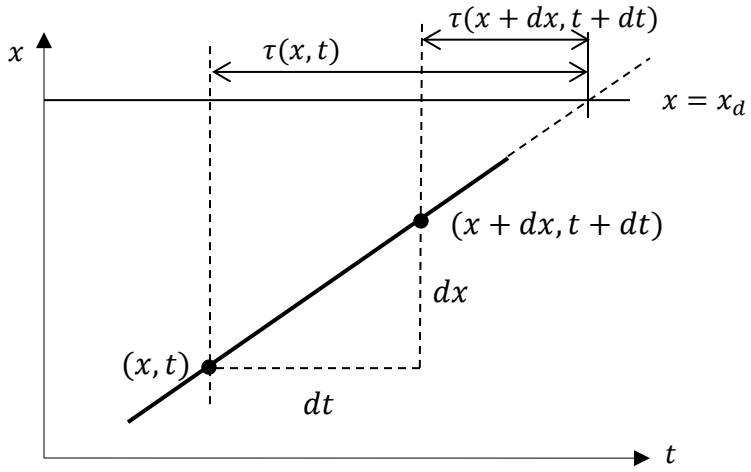


Figure 9. Schematic illustration of a vehicle trajectory in space-time domain.

$$\tau(x+dx,t+dt) = \tau(x,t) - dt \tag{4.3}$$

Substituting (4.2) into (4.1), $\tau(x,t)$ term cancels out on both sides, and then moving derivative terms to the right hand side, the following is obtained:

$$\tau_t dt + \tau_x dx = -dt + O((dt)^2) + O((dx)^2) \tag{4.4}$$

Now dividing both sides of (3) by dt, the following expression is obtained:

$$\tau_t + \tau_x \left(\frac{dx}{dt}\right) = -1 + O(dt) + O\left(\frac{(dx)^2}{dt}\right) \tag{4.5}$$

As dt goes to zero, in the limit $\frac{dx}{dt}$ goes to speed $v(x,t)$ which in vehicular traffic is typically assumed to be bounded from above by a known free flow speed v_f. Therefore, taking the limits from both sides of (4.4) the following first-order PDE equation is obtained:

$$\tau_t + v\tau_x = -1 \tag{4.6}$$

Note that equation (4.6) is hyperbolic with a non-zero right hand side. Hence, it is not conservative. Also, equation (4.6) suggests that gradient of a priori travel time (τ_x, τ_t) and the gradient of vehicle trajectory $(v, 1)$ are pointing to opposite directions. This means that as vehicle travels along its trajectory a priori travel time decreases.

Denoting travel time from an upstream point x_u to a point (x,t) in the solution domain by $\theta(x,t)$, the following first-order PDE equation for a posteriori travel times can be obtained:

$$\theta_t + v\theta_x = 1 \tag{4.7}$$

Note that equation (4.7) is also non-conservative. Also, according to (4.7) the gradient of a posteriori travel time (θ_x, θ_t) and the gradient of vehicle trajectory $(v, 1)$ are overlapping at all times. This implies that as vehicle travels along its trajectory a posteriori travel time increases.

Also, note that summing up models (4.6) and (4.7) and substituting total travel time between upstream and downstream $u = \theta + \tau$, the following conservative equation for total travel time of a vehicle along its trajectory can be obtained:

$$u_t + vu_x = 0 \tag{4.8}$$

Equation (4.8) suggests that the gradient of total travel time along the segment (u_x, u_t) and the gradient of vehicle trajectory $(v, 1)$ are perpendicular to each other at all times. In other words, total travel time of a vehicle stays constant along its trajectory. While an interesting result, and perhaps useful in certain cases, model form (4.8) is not as informative as model forms (4.6) and (4.7) as the latter provide more information about the vehicle travel time changes as it goes through the segment of interest.

4.3 Travel Time Model Properties

Note that travel time models (4.6) and (4.7) suggest a close relationship between local travel time variations and spot speeds exist. In other words, using these models travel time variations may be further characterized. In this section, proposed models behavior in terms of stability conditions, and first-in first-out properties of traffic are explored.

4.3.1 Stability

Under stable traffic flow condition, speeds over time and space do not change. In terms of travel times, stability in time direction suggests that travel time at a given point on the highway will be constant over time. In other words,

$$\tau_t = 0 \tag{4.9}$$

$$\tau_{tt} = 0 \tag{4.10}$$

In the space direction, however, stability has a different interpretation. Under stable conditions (constant speeds), at any given time, it takes a vehicle $\Delta t = \Delta x/v = n\Delta x$ additional time units to cover a distance Δx in front of it, where n is the speed inverse (pace). Given stable conditions, this time lapse is effectively the difference in travel time between two points along the space coordinate:

$$\Delta\tau = \tau(x + \Delta x, .) - \tau(x, .) = -\Delta t = -\Delta x/v = -n\Delta x \tag{4.11}$$

Therefore, in the limit as Δx goes to zero, the derivative of travel time with respect to space can be obtained:

$$\tau_x = -1/v = -n \tag{4.12}$$

$$\tau_{xx} = 0 \tag{4.13}$$

It is straight-forward to show that model (4.6) maintains the pair of first-order stability conditions given by equations (4.9) and (4.12). Similarly a set of stability conditions for a posteriori (retrospective) travel times can be derived:

$$\begin{cases} \theta_t = 0 \\ \theta_{tt} = 0 \\ \theta_x = 1/v = n \\ \theta_{xx} = 0 \end{cases} \tag{4.14}$$

4.3.2 First-In-First-Out (FIFO)

The FIFO condition implies no passing takes place in traffic. While this may not be a realistic assumption, existence of passing suggests that there has been a possibility for the front vehicle to go faster. For all practical purposes, sequence of vehicles in the entrance to and at the departure from the segment can be easily re-arranged in order to maintain the FIFO condition.

Let $D(x,t)$ denote departure time at the downstream point of a vehicle that entered the segment at a point x, and at time t. Clearly, $D(x,t)$ can be expressed in terms of the vehicle entrance time t and the vehicle anticipative travel time:

$$D(x,t) = t + \tau(x,t) \tag{4.15}$$

The FIFO condition asserts that if a pair of vehicles enter the segment at times t and $t + \Delta t$, then their departure times should follow the same order:

$$D(x,t) \leq D(x,t+\Delta t) \tag{4.16}$$

Substituting from (4.15) into both sides of (4.16) and simplifying the following can be obtained:

$$\tau(x,t) \leq \Delta t + \tau(x,t+\Delta t) \tag{4.17}$$

Rearranging (4.17) and taking the limit as the difference between the entrance times of two vehicles goes to zero, in order for FIFO condition to hold, the derivative of anticipative travel time with respect to time will have to be greater than or equal to negative one:

$$\tau_t = \lim_{\Delta t \to 0} \big(\tau(x, t + \Delta t) - \tau(x, t)\big)/\Delta t \geq -1 \quad (4.18)$$

Similarly, and denoting entrance time at the upstream point of a vehicle that departed the segment at point x and at time t by $A(x, t)$, the FIFO condition on a pair of vehicles departing the segment at times t and $t + \Delta t$ can be written as

$$A(x, t) \leq A(x, t + \Delta t) \quad (4.19)$$

Noting that entrance times can be expressed in terms of the departure time t and retrospective travel time

$$A(x, t) = t - \theta(x, t) \quad (4.20)$$

By substituting (4.20) into (4.19), then simplifying and taking the limit, the upper limit on retrospective travel time derivative with respect to time can be derived:

$$\theta_t = \lim_{\Delta t \to 0} \big(\theta(x, t + \Delta t) - \theta(x, t)\big)/\Delta t \leq 1 \quad (4.21)$$

Also, no passing condition of FIFO implies that anticipative travel time of a vehicle at a given point (x, t), $\tau(x, t)$, should be greater than or equal to the anticipative travel time of a vehicle which at the same time t is Δx distance units ahead of it, independent of front vehicles speed:

$$\tau(x, t) \geq \Delta x / v(x, t) + \tau(x + \Delta x, t) \quad (4.22)$$

Rearranging and taking the limit, it can be shown that in order to maintain FIFO condition, at any given point (x, t) the derivative of anticipative travel time with respect to space has to be smaller than or equal to the negative inverse of the local speed (pace):

$$\tau_x = \lim_{\Delta x \to 0} \big(\tau(x + \Delta x, t) - \tau(x, t)\big)/\Delta x \leq -1/v(x, t) = -n \quad (4.23)$$

$$\theta(x, t) + \Delta x / v(x, t) \leq \theta(x + \Delta x, t) \quad (4.24)$$

$$\theta_x = \lim_{\Delta x \to 0} \big(\theta(x+\Delta x, t) - \theta(x,t)\big)/\Delta x \geq 1/v(x,t) = n \tag{4.25}$$

Considering travel time models (4.6) and (4.7), it can be shown that if conditions (4.18) and (4.21) on the corresponding travel time derivatives with respect to time are satisfied, the proposed models will satisfy FIFO conditions.

4.4 Numerical Solution

Travel time models (4.6) and (4.7) can be solved numerically using a forward-time backward-space (FTBS) finite difference scheme. For this purpose we need to discretize the solution domain into cells. First, time duration T is divided into N time intervals $\{t_n | n = 0,1, \dots, N\}$ each of length $\Delta t = T/N$ and the highway length X is divided into M space cells $\{x_i | i = 0,1, \dots, M\}$ each of length $\Delta x = X/M$. To each space cell x_i at time interval t_n, a discrete average speed v_n^i and travel times τ_n^i (anticipative) and θ_n^i (retrospective) are assigned. Therefore, under smooth conditions the pair of travel time evolution equations at each space cell over time is given by,

$$\tau_{n+1}^i = \tau_n^i - \frac{\Delta t}{\Delta x} v_n^i \big(\tau_n^i - \tau_n^{i-1}\big) - \Delta t \tag{4.26}$$

$$\theta_{n+1}^i = \theta_n^i - \frac{\Delta t}{\Delta x} v_n^i \big(\theta_n^i - \theta_n^{i-1}\big) + \Delta t \tag{4.27}$$

4.5 Boundary Conditions

Based on definition, some boundary values are easy to determine. For instance, at the downstream point the value of anticipative travel time function is constantly equal to zero:

$$\tau(L,t) = 0 \tag{4.28}$$

Similarly, retrospective travel time function at downstream point is constantly equal to zero:

$$\theta(0,t) = 0 \tag{4.29}$$

Initial conditions are also assumed to be known at every point along the segment under consideration:

$$\tau(x,0) = F(x) \tag{4.30}$$

$$\theta(x,0) = G(x) \tag{4.31}$$

Travel time data provided by AVI technologies such as Bluetooth detection units can serve as additional boundary, initial or internal conditions. In our specific application such data sources are considered as boundary conditions since the pair of detectors are assumed to be placed at both ends of the segment of interest

$$\tau(0,t) = H(t) \tag{4.32}$$

$$\theta(L,t) = M(t) \tag{4.33}$$

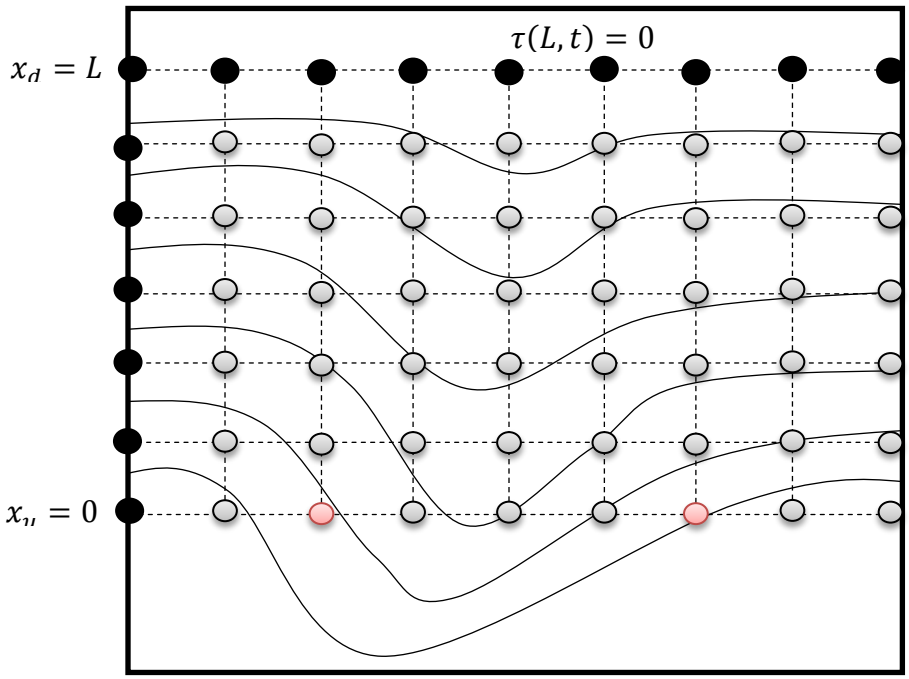


Figure 10. Space-time grid representation of the solution domain.

Figure 10 shows a typical grid in which initial and downstream boundary conditions on travel times are represented by dark nodes, while red nodes represent occasional travel time measurements at the upstream boundary.

According to travel time models (4.26) and (4.27), at the upstream cell $(i = 0)$, the travel time update equations will require an estimate of travel time in the ghost cell $(i = -1)$. The following

boundary conditions give an estimate of travel time in the ghost cell based on the upstream speed, upstream travel times, and the cell size:

$$\tau_n^{-1} = \tau_n^0 + \Delta x / v_n^0 \tag{4.34}$$

$$\theta_n^{-1} = \theta_n^0 - \Delta x / v_n^0 \tag{4.35}$$

4.6 Summary

A first-order travel time model was derived. This model relates the local speeds with local variations of travel time. The proposed travel time model has a non-conservative PDE form. Travel time can be defined relative to either the start point or the end point of a trip. At any point along the trip, the former definition amounts to a retrospective view of travel time, while the latter leads to an anticipative or predictive definition of travel time. With a slight modification the proposed PDE model is capable of dealing with both retrospective and anticipative travel times. Stability and first-in-first-out properties of the proposed travel time models were investigated and conditions under which these properties hold were identified. The proposed numerical solution for this model was obtained by substituting the relevant travel time partial derivatives with their corresponding first-order finite-difference approximations. Appropriate boundary and initial value conditions required in solving the proposed PDE models were derived and presented.

Chapter 5: Estimation Method

In this chapter, the joint traffic and travel time model is cast into a state-space modeling framework. Different components of the corresponding state-space model including state vector, input vector, measurements vector, dynamic model, and measurement equations are specified. The state-space model specification provides a framework to assimilate field measurements with the state dynamics models.

The proposed state-space model state estimation procedures are statistical tools that are optimally designed to achieve a pre-specified performance objective. Optimal state estimations are typically two-step state update processes. These methods are usually derived based on a set of assumptions about the nature of model and measurement errors. In their standard form, optimal state estimation methods are applicable to linear state-space models. Different methods are proposed to extend the optimal state estimation methods to the nonlinear case. In this dissertation, given the highly nonlinear nature of the proposed state-space model, unscented state and error covariance estimation methods are adopted.

Note that optimal state estimation methods are general procedures. Other than the modeled dynamics, some physical constraints may be known about the state of the system. For instance, in the proposed joint traffic and travel time model it is desirable that at all times estimated retrospective travel times at upstream and predictive travel times at downstream be equal to zero. A general Maximum Likelihood Estimation (MLE) method is adopted to enforce these perfect measurement constraints at each time step. This MLE method in fact performs a post-process of the general optimal state estimation results.

Retrospective travel time measurements at the current time are treated as delayed anticipative travel time measurements. Therefore, a delayed filtering method is proposed to condition the

current state estimates based on the additional information that is contained in these delayed measurements.

The rest of this chapter is organized as follows. First, the state-space model corresponding to the proposed joint traffic and travel time dynamics models is specified. Taking into account the nonlinearity of the proposed model, methods to optimally estimate the system state vector are identified. Then, a MLE method is presented to enforce the relevant perfect measurement constraints on the travel time estimates at the boundaries of the road segment. A pair of adaptive nonlinear filtering methods is presented which can handle the highly nonlinear form of the proposed dynamic models efficiently. A delayed filtering approach is proposed to assimilate the delayed upstream anticipative travel time measurements.

5.1 State Space Model

The most general representation of a discrete dynamical system can be given as:

$$\begin{cases} x_{n+1} = f(x_n, u_n, w_n) \\ y_n = h(x_n, v_n) \end{cases} \tag{5.1}$$

where,

x_n, is the system state vector at time interval n,

u_n, is the system input (parameters, boundary conditions) vector at time interval n,

w_n, is the system process error (noise) vector at time interval n,

y_n, is the system measurements vector at time interval n,

v_n, is the measurement error (noise) vector at time interval n,

f, is the multi-dimensional system dynamics function (model) , and

h, is the multi-dimensional measurement equation.

If $f(.)$ and $h(.)$ are explicit functions of time interval n then the system is time-varying.

Otherwise, the system is time invariant.

In this dissertation, traffic dynamics and the corresponding state estimation problem is cast as a state space model. In the following sub-sections different components of the state space model pertaining to the current traffic estimation problem are defined.

5.1.1 System state vector

State of the traffic on a given segment broken into M spatial cells at time interval n can be expressed by three $M \times 1$ sub-vectors components representing speeds, retrospective, and anticipative travel times in each cell. The system state vector is a $3M \times 1$ vector:

$$x_n = [V_n \quad \Theta_n \quad \mathrm{T}_n]^T \tag{5.2}$$

where,

$$V_n = [v_n^1 \quad \dots \quad v_n^M]^T \tag{5.3}$$

$$\Theta_n = [\theta_n^1 \quad \dots \quad \theta_n^M]^T \tag{5.4}$$

$$\mathrm{T}_n = [\tau_n^1 \quad \dots \quad \tau_n^M]^T \tag{5.5}$$

Equations (5.3)-(5.5) specify the speeds, retrospective and anticipative travel time state sub-vectors, respectively.

5.1.2 System input vector

Input to the model include the discretization (PDE solution stencil) parameters $(\Delta t, \Delta x)$, speed-density relationship parameters $\left(v_{max}, v_c, w_f\right)$, and set of traffic and travel time PDE solution boundary conditions $(v_n^0, v_n^{M+1}, \theta_n^0, \tau_n^0)$:

$$u_n = \left[\Delta t \quad \Delta x \quad v_{max} \quad v_c \quad w_f \quad v_n^0 \quad v_n^{M+1} \quad \theta_n^0 \quad \tau_n^0\right]^T \tag{5.6}$$

5.1.3 System measurement vector

In this dissertation measurements at the boundaries of the road segment are considered. Spot speed measurements from a pair of loop detectors installed at the two boundaries $\left(v_n^{1,m}, v_n^{M,m}\right)$,

and retrospective travel time measurement at the downstream $\left(\theta_n^{M,m}\right)$ are included in the measurement vector:

$$y_n = \left[v_n^{1,m} \quad v_n^{M,m} \quad \theta_n^{M,m}\right]^T \tag{5.7}$$

5.1.4 Dynamic model

System dynamic model is comprised of speed and travel time update models given in the earlier chapters. Model errors are assumed to be additive. Similar to the system state vector, the vector of dynamic models can be broken down into three distinct components:

$$f_n(.) = [f_{V,n}(.) \quad f_{\Theta,n}(.) \quad f_{\mathrm{T},n}(.)]^T + w_n \tag{5.8}$$

where, set of speed update models $f_{V,n}(.)$ is given by the following for every cell $\{i = 1, \dots, M\}$:

$$v_{n+1}^i = V\left[V^{-1}\left(v_n^i\right) - \frac{\Delta t}{\Delta x}\left[\tilde{G}\left(v_n^i, v_n^{i+1}\right) - \tilde{G}\left(v_n^{i-1}, v_n^i\right)\right]\right] \tag{5.9}$$

Similarly, set of travel time update models $f_{\Theta,n}(.)$ and $f_{\mathrm{T},n}(.)$ are given by the following set of equations for every cell $\{i = 1, \dots, M\}$:

$$\tau_{n+1}^i = \tau_n^i - \frac{\Delta t}{\Delta x} v_n^i\left(\tau_n^i - \tau_n^{i-1}\right) - \Delta t \tag{5.10}$$

$$\theta_{n+1}^i = \theta_n^i - \frac{\Delta t}{\Delta x} v_n^i\left(\theta_n^i - \theta_n^{i-1}\right) + \Delta t \tag{5.11}$$

And, w_n is the $3M \times 1$ model error vector at time interval n.

Note that all the system dynamic equations are nonlinear in system state variables. While speed-density relationship $V(.)$ and its inverse $V^{-1}(.)$ as well as Godunov flux function $\tilde{G}(.,.)$ introduced in Chapter 3 are the source of nonlinearity in the speed update equations (5.9), both travel time update equations (5.10) and (5.11) include multiplicative terms of speed and travel time variations over space.

5.1.5 Measurement equations

Measurement equations considered in this dissertation reflect direct measurement of some of the state variables. Error terms are assumed to be additive. Therefore, measurement equations have a linear form:

$$h(x_n, v_n) = Hx_n + v_n \tag{5.12}$$

where,

H, is a binary (0 or 1) matrix of the size $3 \times 3M$ with only one nonzero element in each row (columns 1, M, and $2M$ in rows 1, 2, and 3, respectively), and

v_n, is a 3×1 measurement error vector.

5.2 Optimal State Estimation

In presence of dynamic models and measurements, state of the system can be estimated by following a two-step approach. In the first step, given the initial state of the system, the dynamic model can be used to propagate (usually mean and covariance of) the state of the system into the next time step. In the second step, the a priori state estimates will be updated according to the new measurements. This can be regarded as the conditioning of the estimated system state based on the measurements.

Figure 11 summarizes the two principal steps involved in optimal state estimations. The variables used follow the same notation specified in the general definition of state-space models earlier in this chapter. Minus and plus signs indicate the a priori and a posteriori nature of state estimates, respectively.

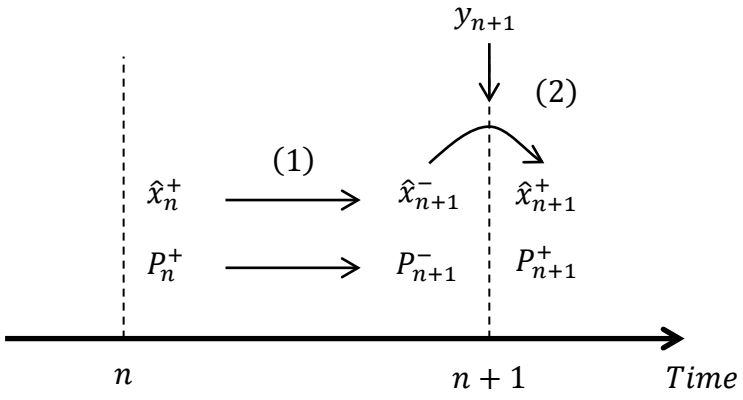


Figure 11 Optimal state estimation steps. (1-state and covariance propagation, 2-measurement update)

Usually state estimation is performed to achieve a desired performance. The Kalman type filters are designed to minimize the expected value of the state estimate error covariance. On the other hand, H-infinity filters are designed to minimize the worst-case estimation error.

Most filters used in optimal state estimation are designed with linear state-space models in mind. In order to apply the optimal state estimation methods to nonlinear cases two general approaches are identified. First, in the case of nonlinear models (especially the ones with a low degree of nonlinearity) the models can be linearized and then standard filters can be applied to the approximate linear models. The second group of methods involves approximating the state error distributions and performing the optimal state estimations on the set of points used in the approximation.

In this dissertation, to deal with highly nonlinear joint traffic and travel time model, unscented methods are used to approximate the mean and covariance of nonlinearly transformed random variables.

5.3 Unscented State and Covariance Propagation

Based on the insight that it is easier to approximate a probability distribution than it is to approximate a random nonlinear function or transformation, Julier et al. (2000) proposed an

unscented method to approximate the mean and covariance of a nonlinearly transformed random vector. This method forms the basis of the Unscented Kalman Filter (UKF) estimators which are presented in the next section. An unscented transformation takes advantage of the fact that a set of individual points (sigma points) in the state space can closely approximate the true distribution of a state vector, and nonlinear transformation on a single point can be easily performed.

The following equations propagate the mean $\hat{x}_n^+$ and covariance P_n^+ of the state estimates from one time interval n to the next time interval $n+1$. Note that to perform the time propagations, sigma points $\tilde{x}_n^{(s)}$ are added to the current (mean) state estimates $\hat{x}_n^+$ (note: in this dissertation state vector has $3M$ elements):

$$\hat{x}_n^{(s)} = \hat{x}_n^+ + \tilde{x}_n^{(s)} \quad s = 1, \dots, 2 \times 3M \tag{5.13}$$

where,

$$\tilde{x}_n^{(s)} = \left(\sqrt{(3M)P_n^+}\right)_s^T \quad s = 1, \dots, 3M \tag{5.14}$$

$$\tilde{x}_n^{(3M+s)} = -\left(\sqrt{(3M)P_n^+}\right)_s^T \quad s = 1, \dots, 3M \tag{5.15}$$

Note that $\sqrt{A}$ is the matrix square root of A such that $\left(\sqrt{A}\right)^T\left(\sqrt{A}\right) = A$, and subscript s indicates the s^{th} row of the matrix.

Once, the ensemble of state vectors (5.2)-(5.5) are drawn, the known nonlinear system equation $f(.)$ can be used to transform the sigma points into next time interval state estimate $\hat{x}_{n+1}^{(s)}$ vectors:

$$\hat{x}_{n+1}^{(s)} = f\left(\hat{x}_n^{(s)}, u_n\right) \quad s = 1, \dots, 2 \times 3M \tag{5.16}$$

Now, the $\hat{x}_{n+1}^{(s)}$ vectors can be combined to obtain the *a priori* state estimate (mean and error covariance) at time $n+1$:

$$\hat{x}_{n+1}^{-} = \frac{1}{2(3M)} \sum_{i=1}^{2(3M)} \hat{x}_{n+1}^{(s)} \tag{5.17}$$

$$P_{n+1}^{-} = \frac{1}{2(3M)} \sum_{i=1}^{2(3M)} \left[\left(\hat{x}_{n+1}^{(s)} - \hat{x}_{n+1}^{-} \right) \left(\hat{x}_{n+1}^{(s)} - \hat{x}_{n+1}^{-} \right)^{T} \right] + Q_n \tag{5.18}$$

5.4 Unscented Measurement Error and Covariance Estimation

Again, new sigma points $\tilde{x}_{n+1}^{(s)}$ based on the current best estimates of the mean and error covariance of the state are selected:

$$\hat{x}_{n+1}^{(s)} = \hat{x}_{n+1}^{-} + \tilde{x}_{n+1}^{(s)} \quad s = 1, \dots, 2 \times 3M \tag{5.19}$$

where,

$$\tilde{x}^{(s)} = \left(\sqrt{(3M) P_{n+1}^{-}} \right)_{s}^{T} \quad s = 1, \dots, 3M \tag{5.20}$$

$$\tilde{x}^{(3M+s)} = -\left(\sqrt{(3M) P_{n+1}^{-}} \right)_{s}^{T} \quad s = 1, \dots, 3M \tag{5.21}$$

Note that in general updating sigma points can be expensive. In case estimate accuracy is not heavily impacted, the sigma points that were obtained from the time update can be reused to save on computational burden. In this dissertation updated sigma points are used.

The known measurement equation $h(.)$ is used to transform the sigma points into measurement prediction vectors $\hat{y}_{n+1}^{(s)}$:

$$\hat{y}_{n+1}^{(s)} = H \hat{x}_{n+1}^{(s)} \tag{5.22}$$

Now, the $\hat{y}_{n+1}^{(s)}$ vectors can be combined to obtain the measurement predictions (mean and error covariance) at time $n + 1$:

$$\hat{y}_{n+1} = \frac{1}{2(3M)} \sum_{s=1}^{2(3M)} \hat{y}_{n+1}^{(s)} \tag{5.23}$$

$$P_{yy} = \frac{1}{2(3M)} \sum_{s=1}^{2(3M)} \left[\left(\hat{y}_{n+1}^{(s)} - \hat{y}_{n+1} \right) \left(\hat{y}_{n+1}^{(s)} - \hat{y}_{n+1} \right)^{T} \right] + R_{n+1} \tag{5.24}$$

Also, the cross covariance between $\hat{x}_k^{-}$ and $\hat{y}_k$ can be readily estimated based on the current estimates and their corresponding sigma points:

$$P_{xy} = \frac{1}{2(3M)} \sum_{s=1}^{2(3M)} \left[\left(\hat{x}_{n+1}^{(s)} - \hat{x}_{n+1}^{-} \right) \left(\hat{y}_{n+1}^{(s)} - \hat{y}_{n+1} \right)^T \right] \quad (5.25)$$

5.5 *Perfect Measurement and Constraint Enforcement*

In general it is possible that some constraints on the state variables are known to exist. Also, for some state variables perfect measurements might be available. These cases can be treated as a set of perfect measurements:

$$d_n = D_n x_n + \mathbf{0} \quad (5.26)$$

where,

d_n, is the $P \times 1$ vector of perfect measurements at time interval n,

D_n, is the $P \times 3M$ matrix of binary $(0,1)$ coefficients at time interval n, and

$\mathbf{0}$, is the $P \times 1$ vector of zero measurement errors.

In this dissertation, the perfect measurements at each time interval n include the retrospective travel time at upstream (θ_n^1), and the anticipative travel time at downstream (τ_n^M) which are both equal to zero. Therefore, the corresponding perfect measurement vector d_n is a 2×1 vector of zeroes, and the $2 \times 3M$ coefficient matrix D_n has only two non-zero elements (column $M + 1$ in row 1, and column $3M$ in row 2)

Conceptually it is possible to augment measurement equations (5.12) with perfect measurement equations (5.26) and use the augmented set of measurements in state estimation. However, this approach results in a singular error covariance matrix (R_n) which increases the possibility of numerical problems.

In this dissertation, an alternative approach to incorporate perfect measurements into estimation process is adopted. The approach is based on maximum likelihood estimation (MLE) of state variables subject to the set of perfect measurement constraints. Solving the Lagrangian of the optimization problem the following update equation for state variables are derived:

$$\tilde{x}_n = \hat{x}_n - P_n D_n^T (D_n P_n D_n^T)^{-1} (D_n \hat{x}_n - d_n) \quad (5.27)$$

Note that this update equation maintains that at a time interval n, the constrained state estimate $\tilde{x}_n$ is equal to the unconstrained state estimate $\hat{x}_n$, minus a correction term which depends on the current state estimate error covariance, and the current error with regards to the perfect measurements. Also note that this correction method is general and can be applied to both cases of a priori and a posteriori state estimates.

5.6 Unscented Kalman Filtering

In this dissertation Unscented Kalman Filter (UKF) is used to estimate the optimal state of the system. UKF is an extension of conventional Kalman Filter and is specifically designed to be applicable to highly nonlinear state-space models. In the rest of this section a complete description of UKF estimation process is presented.

Consider the discrete time nonlinear system state space model (5.1)-(5.12) with state vector x_n and measurement vector y_n at each time interval n, the generally nonlinear system $f(.)$ and measurement equations $h(.)$.

Also, mainly for simplification, make the following set of additional assumptions that system and measurement errors are white and not correlated:

$$E(w_k w_j^T) = Q_k \delta_{k-j}, \qquad \forall k, j \quad (5.28)$$

$$E(v_k v_j^T) = R_k \delta_{k-j}, \qquad \forall k, j \quad (5.29)$$

$$E(w_k v_j^T) = 0, \qquad \forall k, j \quad (5.30)$$

The UKF estimation method can now be described as the following procedure:

Initialization: As the first step the UKF can be initialized as follows:

$$\hat{x}_0^+ = E(x_0) \quad (5.31)$$

$$P_0^+ = E[(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^T] \quad (5.32)$$

Time Update: Perform time updates using equations (5.13)-(5.18) to obtain nonlinearly propagated mean ($\hat{x}_{n+1}^{-}$) and covariance of the system state (P_{n+1}^{-}) as *a priori* estimates.

Measurement Update: Now that time updates based on the system model are performed, it is time that measurements are taken into account to improve the a priori estimates. In general, a procedure similar to the nonlinear time update is adopted. Predicted measurement vector and cross-covariance matrix can be estimated using equations (5.19)-(5.25).

The measurement gain and state update equations are similar to the normal Kalman filter. The *a posteriori* state update equations (mean and error covariance) take into account improvement in state estimates as a result of new information in the measurement:

$$K_{n+1} = P_{xy}P_{yy}^{-1} \tag{5.33}$$

$$\hat{x}_{n+1}^{+} = \hat{x}_{n+1}^{-} + K_{n+1}[y_{n+1} - \hat{y}_{n+1}] \tag{5.34}$$

$$P_{n+1}^{+} = P_{n+1}^{-} - K_{n+1}P_{yy}K_{n+1}^{T} \tag{5.35}$$

Perfect Measurements Update: Use equations (5.26)-(5.27) to update a posteriori state estimates obtained from equation (5.34).

5.7 Unscented H-infinity Filtering

In most applications the nature and size of model and measurement errors are uncertain. Kalman filter presented in the previous section is based on a set of strong assumptions on model and measurement errors (that they are Gaussian and uncorrelated). These are strong assumptions on the nature of model and measurement error terms. In most practical scenarios model and measurement errors are not that well behaving. H-infinity filter is specifically designed for robust estimation. In fact, H-infinity filter does not make any assumptions on the nature of model and measurement errors.

Kalman filter is designed to minimize the expected value of the estimation error covariance. But, H-infinity filter is designed to minimize the worst-case estimation error. Denoting the squared L_2-norm of vector x weighted by matrix M as following:

$$\|x\|_M^2 = x^T M x \tag{5.36}$$

The objective function (to be minimized) in the case of H-infinity filter can be expressed as following:

$$J = \frac{\sum_{n=0}^{N-1} \|z_n - \hat{z}_n\|_{S_n}^2}{\|x_0 - \hat{x}_0\|_{P_0^{-1}}^2 + \sum_{n=0}^{N-1} \left(\|w_n\|_{Q_n^{-1}}^2 + \|v_n\|_{R_n^{-1}}^2 \right)} \tag{5.37}$$

where, P_0, Q_n, R_n, and S_n are symmetric positive definite matrices chosen based on the nature of specific problem. Note that despite obvious analogies these matrices are not referred to as error covariance matrices. This is to stress the fact that in H-infinity filter derivation no use of the statistical properties of model and measurement errors are made.

Also, note that cost function (5.37) provides a game theoretic interpretation of H-infinity filter design. In fact, the cost function (5.37) represents the ratio of weighted estimation error covariance and the sum of initial estimate and model and measurement normalized error covariances.

While minimization of (5.37) subject to (5.1)-(5.12) is not tractable, by choosing a performance bound (γ^2) and solving for a dynamic estimation strategy that satisfies the set threshold the H-infinity filter can be derived.

$$J = \frac{\sum_{n=0}^{N-1} \|z_n - \hat{z}_n\|_{S_n}^2}{\|x_0 - \hat{x}_0\|_{P_0^{-1}}^2 + \sum_{n=0}^{N-1} \left(\|w_n\|_{Q_n^{-1}}^2 + \|v_n\|_{R_n^{-1}}^2 \right)} < \gamma^2 \tag{5.38}$$

More details on derivation of H-infinity filter in the case of linear state-space models using Lagrange multiplier method to solve the resulting dynamic constrained optimization problem can be found elsewhere (Simon 2006). This dissertation extends the H-infinity filter to the nonlinear

case based on insights provided in the derivation of adaptive robust filters (Sayed 2003). In the rest of this section a complete description of UHF estimation process is provided.

Consider the discrete time nonlinear system state space model (5.1)-(5.12) with state vector x_n and measurement vector y_n at each time interval n, the generally nonlinear system $f(.)$ and measurement equations $h(.)$. The UHF method to optimally estimate a linear transform of state variables $z_n = L_n x_n$ (where matrix L_n is full-rank and user specified) can now be described as the following procedure:

Initialization: As the first step the UHF can be initialized as follows:

$$\hat{x}_0^+ = E(x_0) \tag{5.39}$$

$$P_0^+ = E[(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^T] \tag{5.40}$$

Time Update: Perform time updates using equations (5.13)-(5.18) to obtain nonlinearly propagated mean $(\hat{x}_{n+1}^-)$ and covariance of the system state (P_{n+1}^-) as *a priori* estimates.

Measurement Update: Now that time updates based on the system model are performed, it is time that measurements are taken into account to improve the *a priori* estimates.

The measurement gain and state update equations are similar to the normal H-infinity filter. The *a posteriori* state update equations (mean and error covariance) take into account improvement in state estimates as a result of new information in the measurement:

$$R_{0,n+1} = \begin{bmatrix} R & 0 \\ 0 & -\gamma^2 I \end{bmatrix} \tag{5.41}$$

$$R_{e,n+1} = R_{0,n+1} + \begin{bmatrix} H_{n+1} \\ L_{n+1} \end{bmatrix} P_{n+1}^- [H_{n+1}^T \quad L_{n+1}^T] \tag{5.42}$$

$$P_{n+1}^+ = P_{n+1}^- \left(I - [H_{n+1}^T \quad L_{n+1}^T] R_{e,n+1}^{-1} \begin{bmatrix} H_{n+1} \\ L_{n+1} \end{bmatrix} P_{n+1}^- \right) \tag{5.43}$$

$$K_{n+1} = P_{n+1}^- H_{n+1}^T (R + H_{n+1} P_{n+1}^- H_{n+1}^T)^{-1} \tag{5.44}$$

$$\hat{x}_{n+1}^+ = \hat{x}_{n+1}^- + K_{n+1}[y_{n+1} - \hat{y}_{n+1}] \tag{5.45}$$

This filter will be valid, if and only if, matrices $R_{0,n+1}$ and $R_{e,n+1}$ have the same inertia at every time steps. For this condition to be met, the two matrices must have exactly the same number of positive and negative eigenvalues. An investigation of the two matrices reveals that this condition will be met if parameter $\gamma^2 = 1/\theta$ meets the following condition:

$$\gamma^2 \geq \max\{\text{eig}\{L_{n+1}((P_{n+1}^-)^{-1} + H_{n+1}^T R^{-1} H_{n+1})^{-1} L_{n+1}^T\}\} \quad (5.46)$$

Note that $\max\{\text{eig}\{A\}\}$ denotes the maximum eigenvalue of the matrix A. Therefore by setting a parameter $(\alpha \geq 1)$ validity of the filter can be guaranteed at all times:

$$\gamma^2 = \alpha \max\{\text{eig}\{L_{n+1}[(P_{n+1}^-)^{-1} + H_{n+1}^T R^{-1} H_{n+1}]^{-1} L_{n+1}^T\}\} \quad (5.47)$$

Note that γ is a measure of filter performance. It is desirable for γ to be as small as possible. Hence, selection of parameter α has to be carried out with care. Optimal parameter α depends on the model and measurements. In this dissertation, based on trial and error and running multiple experiments with the proposed model, value of γ^2 is set equal to 500.

Perfect Measurements Update: Use equations (5.26)-(5.27) to update a posteriori state estimates obtained from equation (5.39).

5.8 Delayed Filtering

As was stated earlier, every travel time measurement can be regarded as either anticipative or retrospective depending on the adopted reference point. In real time, all current travel time measurements are retrospective in nature. However, in the proposed model formulation the same travel time measurement can be regarded as a delayed anticipative travel time measurement at the upstream of the segment. In order to assimilate this delayed measurement a delayed filter can be used. This delayed filter will run from the time interval to which the delayed anticipative travel time measurement belongs up to and including the current time interval. The proposed

delayed filter is expected to increase the accuracy of the current state estimates and therefore eventually to improve the state predictions.

Note that practicality of the delayed filter depends on the computation time of each state update, the magnitude of travel time measurements, and the size of time intervals used in model discretization. If time intervals are too short and travel times are comparatively too long, then it may be the case that delayed filter will take more time than allowed to catch up with the current time filter.

5.9 Summary

In this chapter the state-space model corresponding to the proposed joint traffic and travel time dynamics model was fully specified. Optimal state estimation methods in their most general two-step form were presented. Due to the highly nonlinear form of the proposed models a standard unscented approach to state and error covariance propagation was adopted. Similarly, in the second step a standard unscented approach was adopted to condition a priori system state estimates according to the estimated measurement error and covariance. A general MLE approach was identified to enforce perfect measurement constraints on the retrospective and anticipative travel time measurements at the upstream and the downstream boundaries of the segment, respectively. This method performs a post-process on optimally estimated states at each time step to ensure the above natural constraints hold at all times. Unscented Kalman filtering and H-infinity used in this dissertation for optimal state estimation were presented. Note that travel time measurements in current time represent retrospective travel times at the downstream. At the same time, these measurements represent delayed anticipative travel times at the upstream boundary of the segment. A delayed filtering approach was adopted to assimilate the delayed upstream anticipative travel time measurements into the proposed estimation process.

Chapter 6: Numerical Experiments

In this chapter a series of relevant numerical tests on traffic modeling, traffic data assimilation, traffic data estimation, and short-term prediction are presented. The focus will be on field measured travel times and assimilating them in the context of the proposed joint traffic and travel time dynamics model. Traffic datasets used in numerical experiments are fully described. Implementation details of the joint traffic and travel time model are specified. This includes specification of the adopted speed-density relations, and discretization levels at which solutions are obtained. Model errors are evaluated through an open-loop state estimation process in which perfect measurements (no added errors) are used as boundary values in solving the proposed finite-difference models. Also, based on literature and field experience, statistical properties of the measurement errors including magnitudes and correlations of the error terms are specified.

Different scenarios representing combinations of traffic datasets, speed-density relations, field measurements, space-time discretization schemes, and optimal state estimation methodologies are considered. Estimation results are presented for different scenarios when boundary speeds and travel time measurements are used in assimilation experiments either separately or together. Delayed filter impacts on state estimations are reported when boundary travel time measurements have been included in assimilations. Short-term state prediction results are presented under different scenarios. Computation times obtained by running scenarios on a desktop machine are reported.

6.1 Traffic Datasets

In this research standard traffic datasets prepared and made available under Next Generation SIMulation (NGSIM) project are used (FHWA 2006). NGSIM datasets provide a rich source of

accurate and detailed traffic data over a variety of facilities located in California and Georgia. These datasets contain high quality (resolution equal to one-tenth of the second) observations of the type, position, speed, and acceleration of every single vehicle that has been part of the traffic stream in the segment under study. In this dissertation one NGSIM dataset from California is selected to run numerical experiments. The US 101 dataset is fully described in the following section.

6.2 The US 101 Dataset

This is a 45 minute freeway dataset representing traffic flows on a segment of US Highway 101 (Hollywood Freeway) in the Universal City neighborhood of Los Angeles, California. The 45 minute dataset is broken into three smaller 15 minute datasets which display increasingly more congested traffic from 7:50AM to 8:35AM on June 15, 2005. The dataset represents vehicle trajectory data on a 2100 foot (640 meters) segment of southbound US 101.

Figure 12 illustrates the NGSIM study area on US 101 and its lane configuration. The segment consists of five main lanes throughout the section. The merge/weave section represented in the data includes an on-ramp off-ramp connected by an auxiliary lane. The dataset consists of detailed vehicle trajectory data, wide-area detector data and supporting data needed for behavioral algorithm research. In the rest of this chapter results of experiments on US 101 dataset are presented.

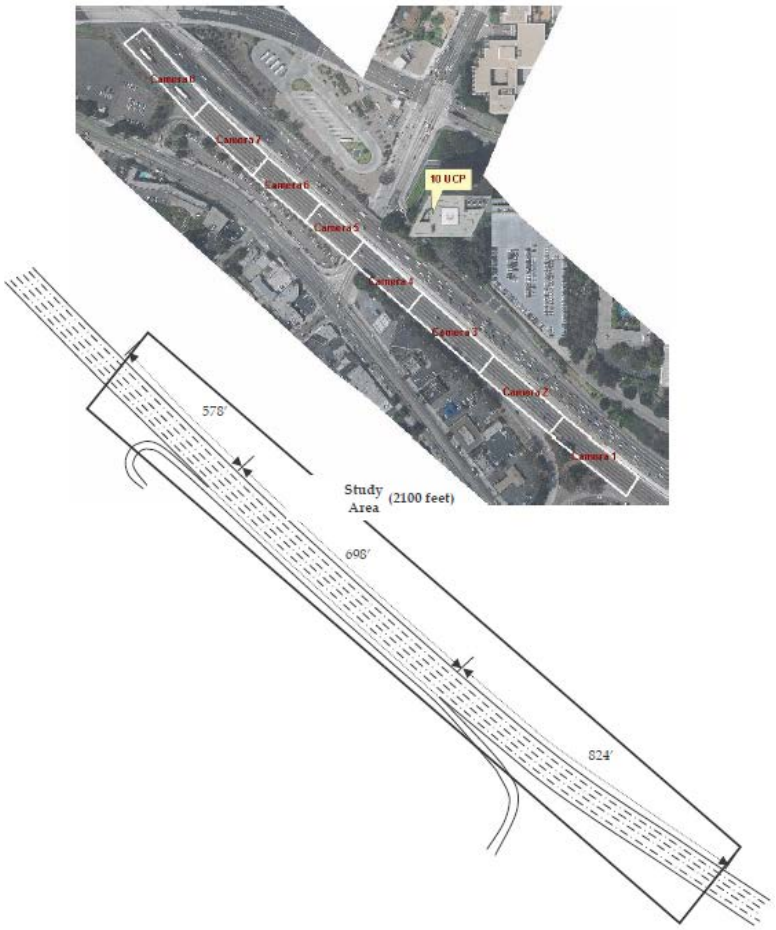


Figure 12 US 101 dataset study area schematic and camera coverage (Source: Cambridge Systematics, Inc. 2005).

As mentioned before, NGSIM raw data is available at the vehicle trajectory level. This raw dataset was processed to estimate speeds and travel times in each space cell and time interval. The processed results are used as ground-truth in this study. In the following sub-sections ground-truth speed and travel time observations during each 15 minute dataset is presented. In each case, ground-truth data is provided at two discretization levels:

- 210 feet by 2 seconds
- 420 feet by 4 seconds

6.2.1 Dataset 1: US 101 at 7:50AM-8:05AM

Figure 13 presents the speed ground-truth during the first 15 minute of US 101 dataset. The top graphs show the observed speeds at upstream and downstream boundaries as well as throughout the segment from 7:50AM to 8:05AM. The top set of graphs represents ground-truth speeds when space and time discretization is 210 foot and 2 seconds, respectively. The bottom graphs represent ground-truth speeds when space and time discretization is 420 foot and 4 seconds, respectively.

Note that while both discretization levels provide fine granularity, the first set is much more detailed. While multiple slow-downs are observed in terms of speed data, only one major shockwave is identified towards the end of this time period which was initiated at the downstream and propagated backward along the segment to the upstream. Also, as speeds hovered above 40 mph initially at the downstream for more than half the duration of this time the upstream speeds have been fluctuating as a series of minor shockwaves frequently arrived at the upstream. The initial point of these smaller shockwaves roughly corresponds with the end of auxiliary lane, therefore these mid-segment slowdowns may be attributed to the weaving operations at the area of on- and off-ramps.

Figure 14 and Figure 15 present the retrospective and anticipative travel time ground-truth during this 15 minute time period, respectively. Travel times substantially increase toward the end of this time period. Retrospective travel times at upstream and anticipative travel times at downstream are strictly equal to zero. Note that retrospective and anticipative travel times at the boundaries generally follow the same pattern observed at the opposite boundary. This is due to the definition and should be expected.

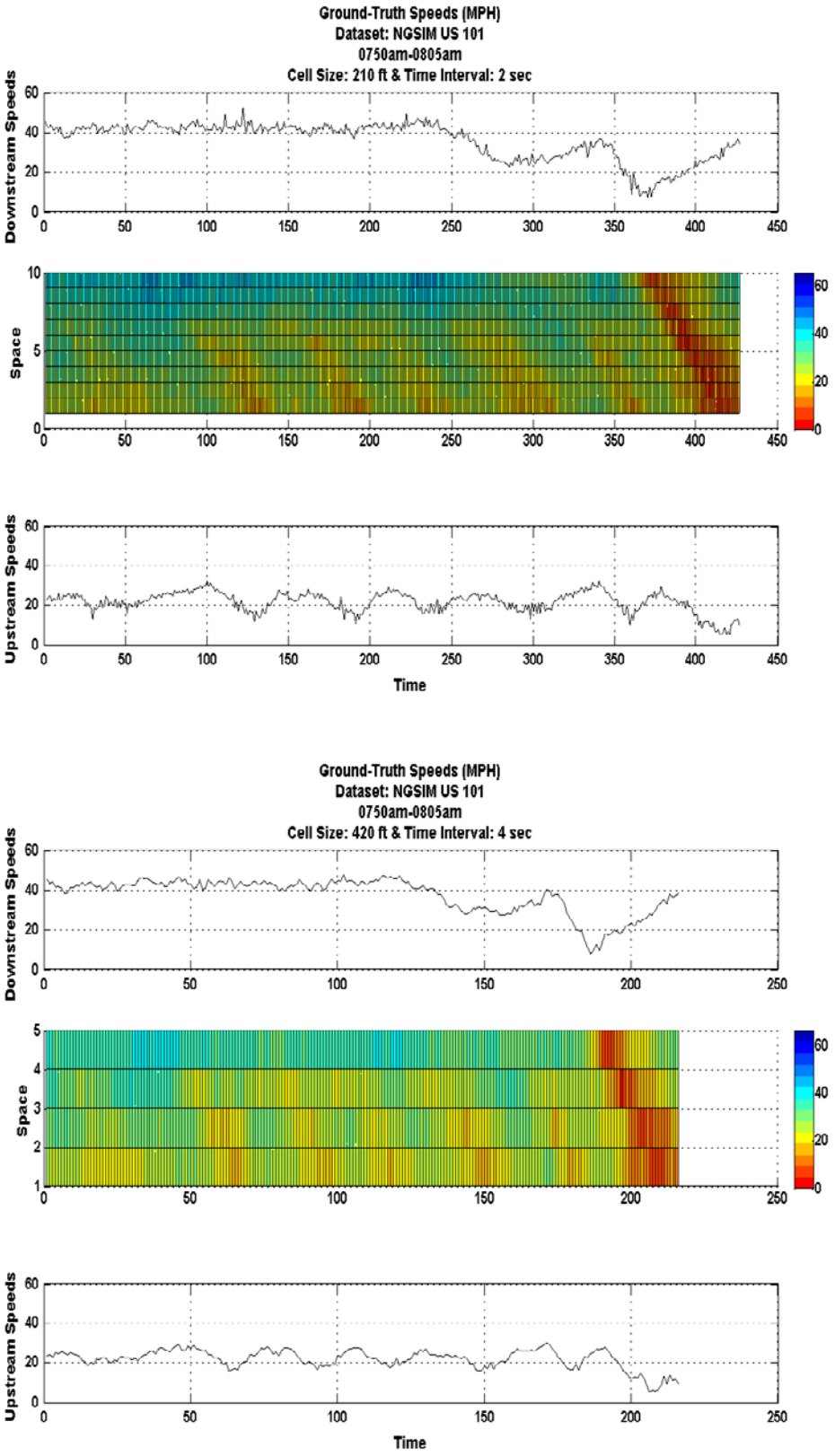


Figure 13. Speed ground truth US 101 at 7:50AM-8:05AM.

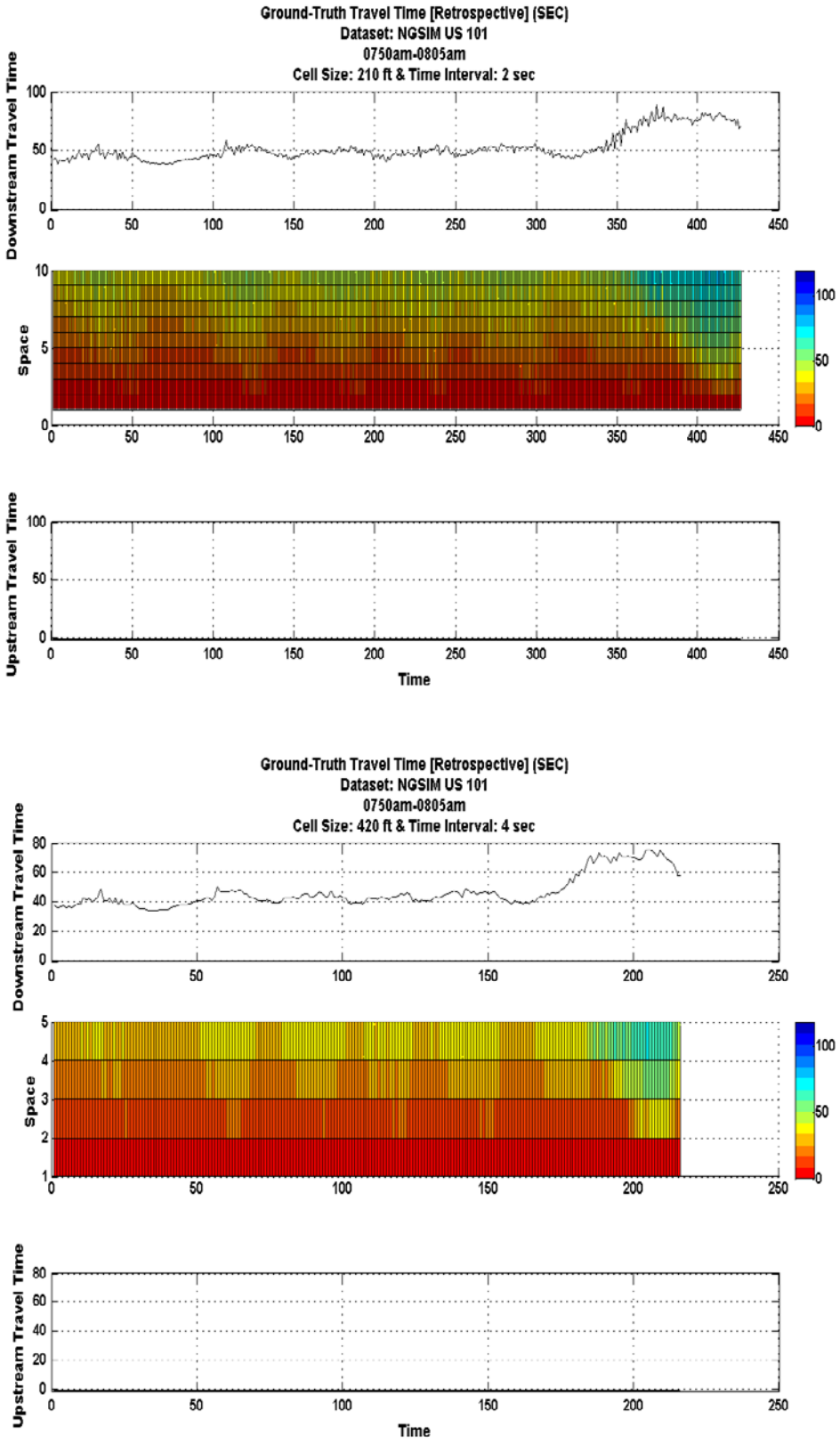


Figure 14. Retrospective travel time ground truth US 101 at 7:50AM-8:05AM.

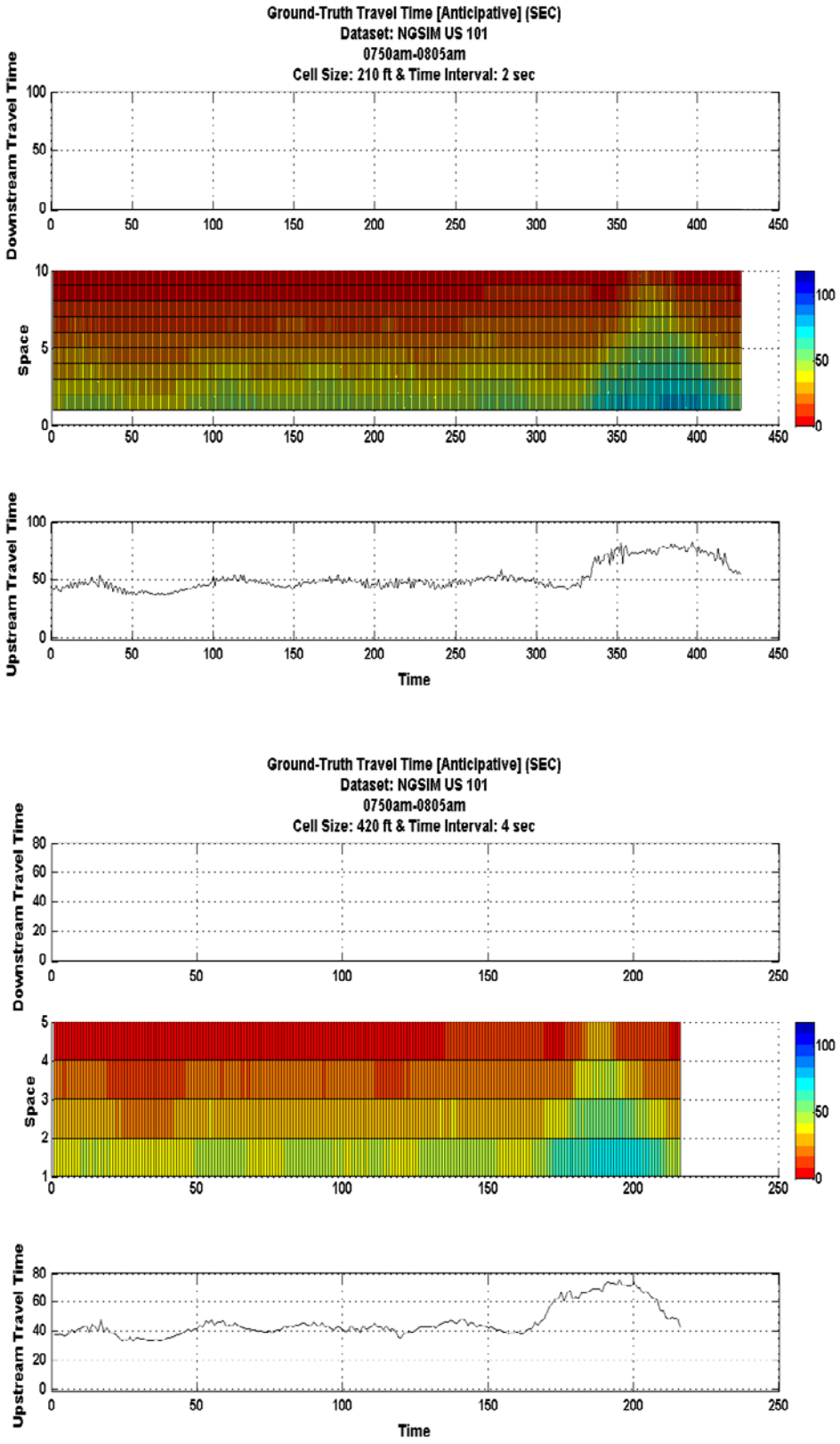


Figure 15. Anticipative travel time ground truth US 101 at 7:50AM-8:05AM.

6.2.2 Dataset 2: US 101 at 8:05AM-8:20AM

Figure 16 presents the speed ground-truth during the second 15 minute of US 101 dataset. The top graphs show the observed speeds at upstream and downstream boundaries as well as throughout the segment from 8:05AM to 8:20AM. The top set of graphs represents ground-truth speeds when space and time discretization is 210 foot and 2 seconds, respectively. The bottom graphs represent ground-truth speeds when space and time discretization is 420 foot and 4 seconds, respectively.

In this time period multiple slow-downs are observed in the segment in terms of speed data. At least three major shockwaves are identified during this time period which are initiated at the downstream and propagate backward along the segment to the upstream. Also, a number of smaller shockwaves are present which originated in the middle of the segment again at the end of weaving area. Boundary speeds are generally fluctuating more in the below 40 mph range which indicates stop and go traffic conditions are present.

Figure 17 and Figure 18 present the retrospective and anticipative travel time ground-truth during this 15 minute time period, respectively. Travel times go through two major cycles during which they double from about 50 seconds to almost 100 seconds. Again, retrospective travel times at upstream and anticipative travel times at downstream are strictly equal to zero. Note that retrospective and anticipative travel times at the boundaries generally follow the same pattern observed at the opposite boundary. This is due to the definition and should be expected.

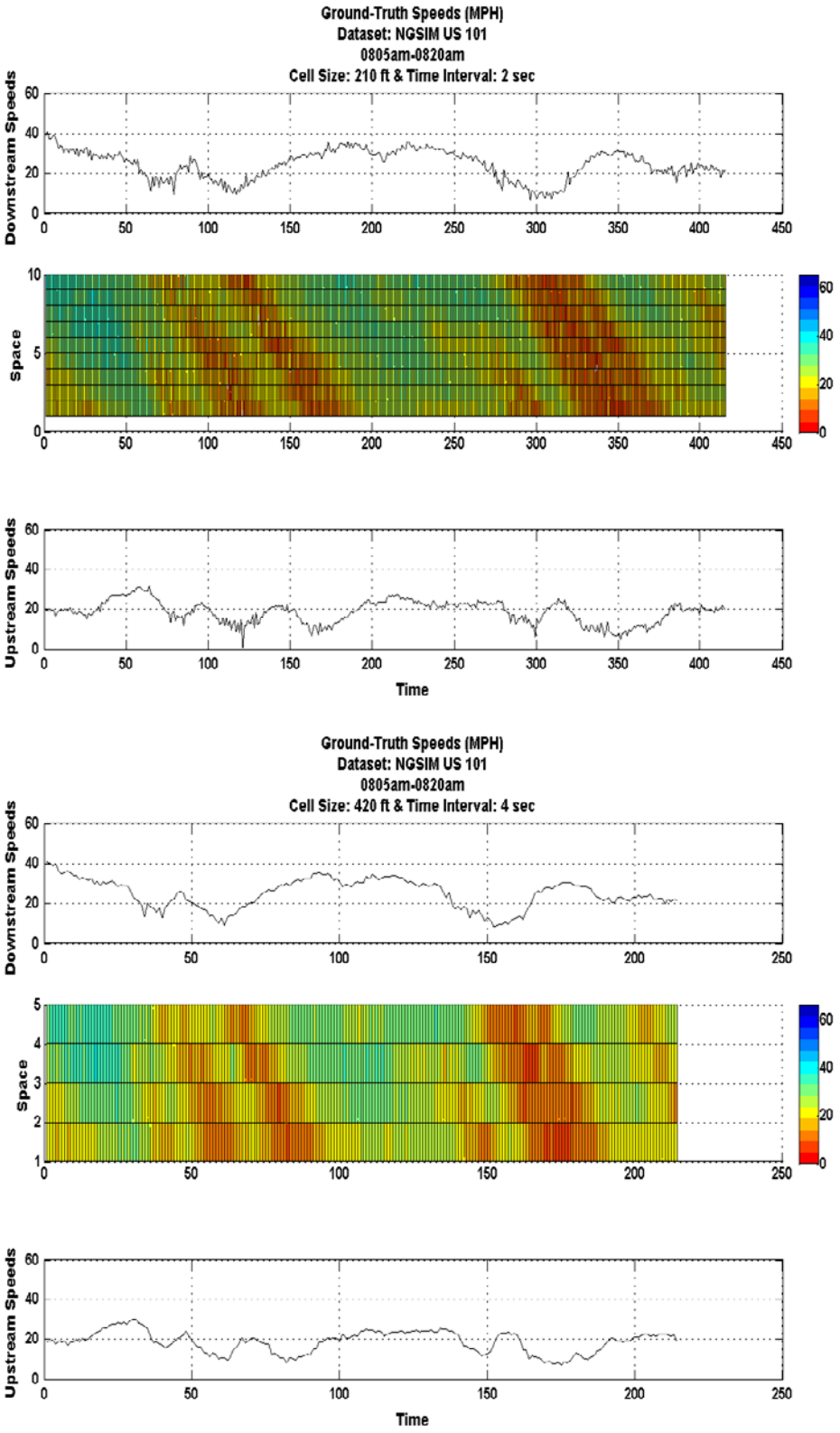


Figure 16. Speed ground truth US 101 at 8:05AM-8:20AM.

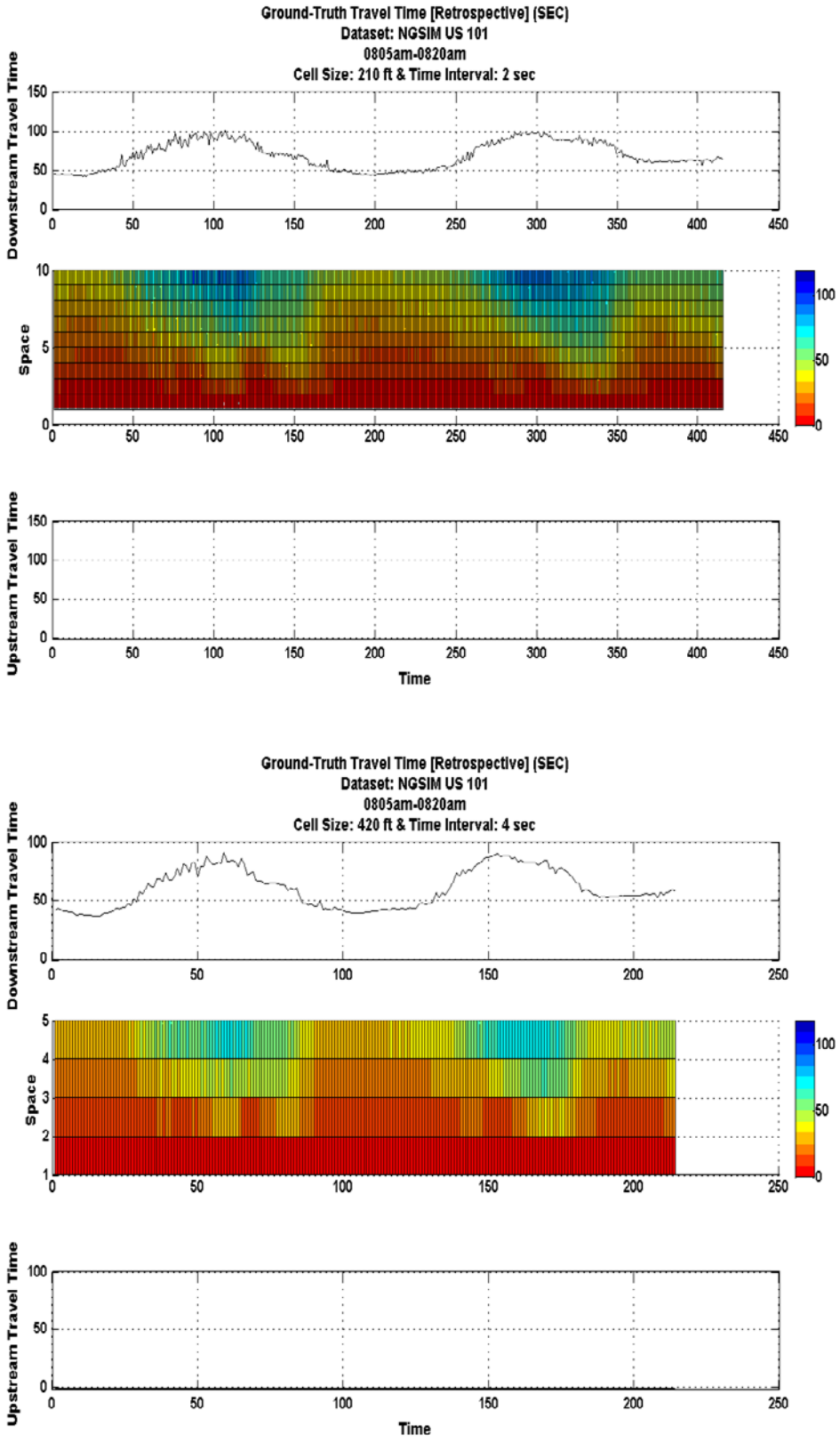


Figure 17. Retrospective travel time ground truth US 101 at 8:05AM-8:20AM.

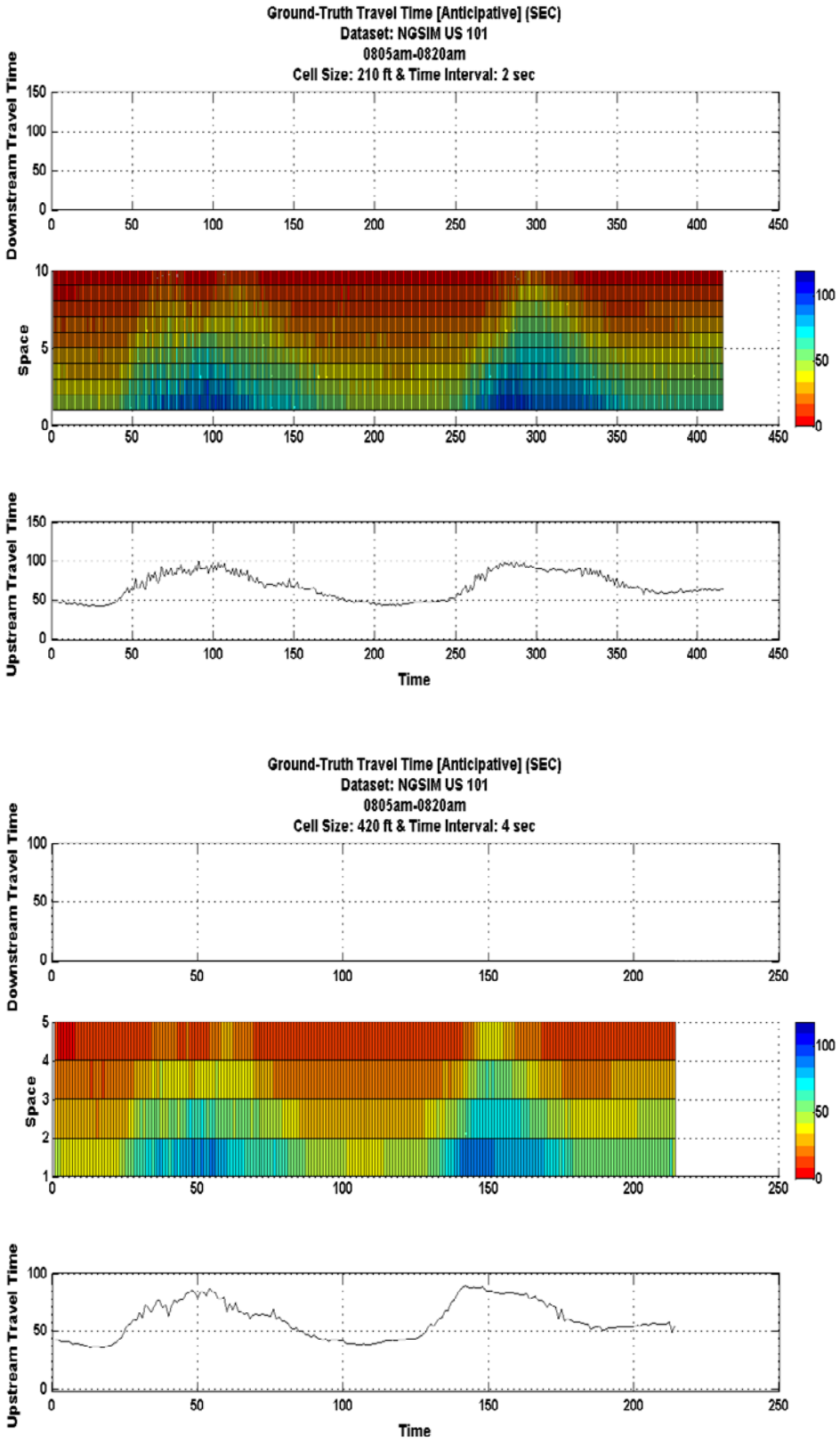


Figure 18. Anticipative travel time ground truth US 101 at 8:05AM-8:20AM.

6.2.3 Dataset 3: US 101 at 8:20AM-8:35AM

Figure 19 presents the speed ground-truth during the second 15 minute of US 101 dataset. The top graphs show the observed speeds at upstream and downstream boundaries as well as throughout the segment from 8:20AM to 8:35AM. The top set of graphs represents ground-truth speeds when space and time discretization is 210 foot and 2 seconds, respectively. The bottom graphs represent ground-truth speeds when space and time discretization is 420 foot and 4 seconds, respectively.

In terms of speed ground-truth, in this time period slow-downs are more frequently present and almost all of the observed shockwaves extend throughout the segment. At least four major shockwaves are identified during this time period which are initiated at the downstream and propagate backward along the segment to the upstream. It is conceivable that these shockwaves have masked or merged with smaller shockwaves which originate in the middle of the segment again due to the mid-section weaving area. Boundary speeds are generally fluctuating in the below 40 mph range at downstream and, for the most part, below 20 mph range at upstream which indicates heavily congested conditions are present.

Figure 20 and Figure 21 present the retrospective and anticipative travel time ground-truth during this 15 minute time period, respectively. Travel times go through two major cycles during this time period. However, the second cycle is more severe and longer-lived than the first cycle. Again, retrospective travel times at upstream and anticipative travel times at downstream are strictly equal to zero. Note that retrospective and anticipative travel times at the boundaries generally follow the same pattern observed at the opposite boundary. This is due to the definition and should be expected.

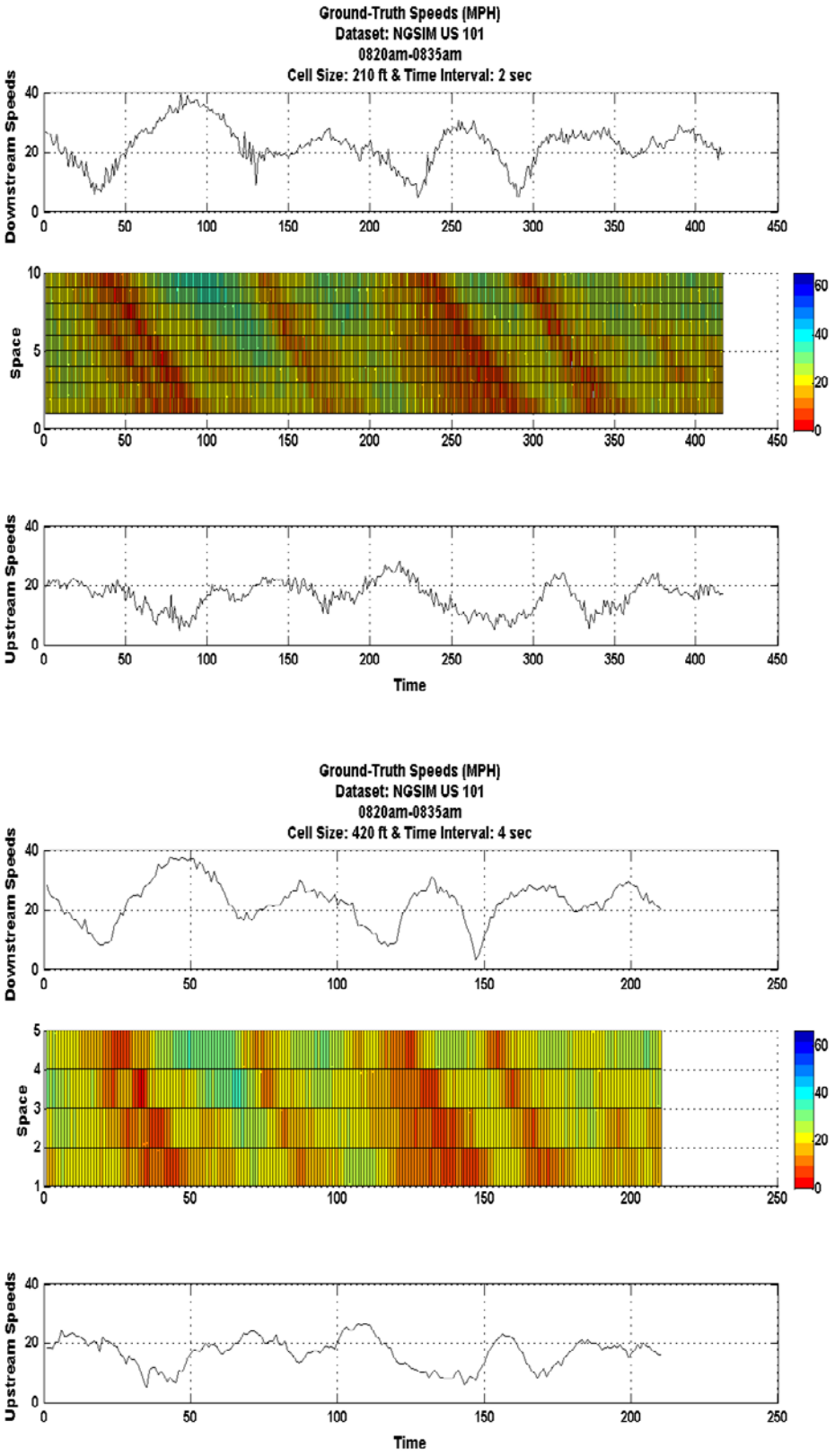


Figure 19. Speed ground truth US 101 at 8:20AM-8:35AM.

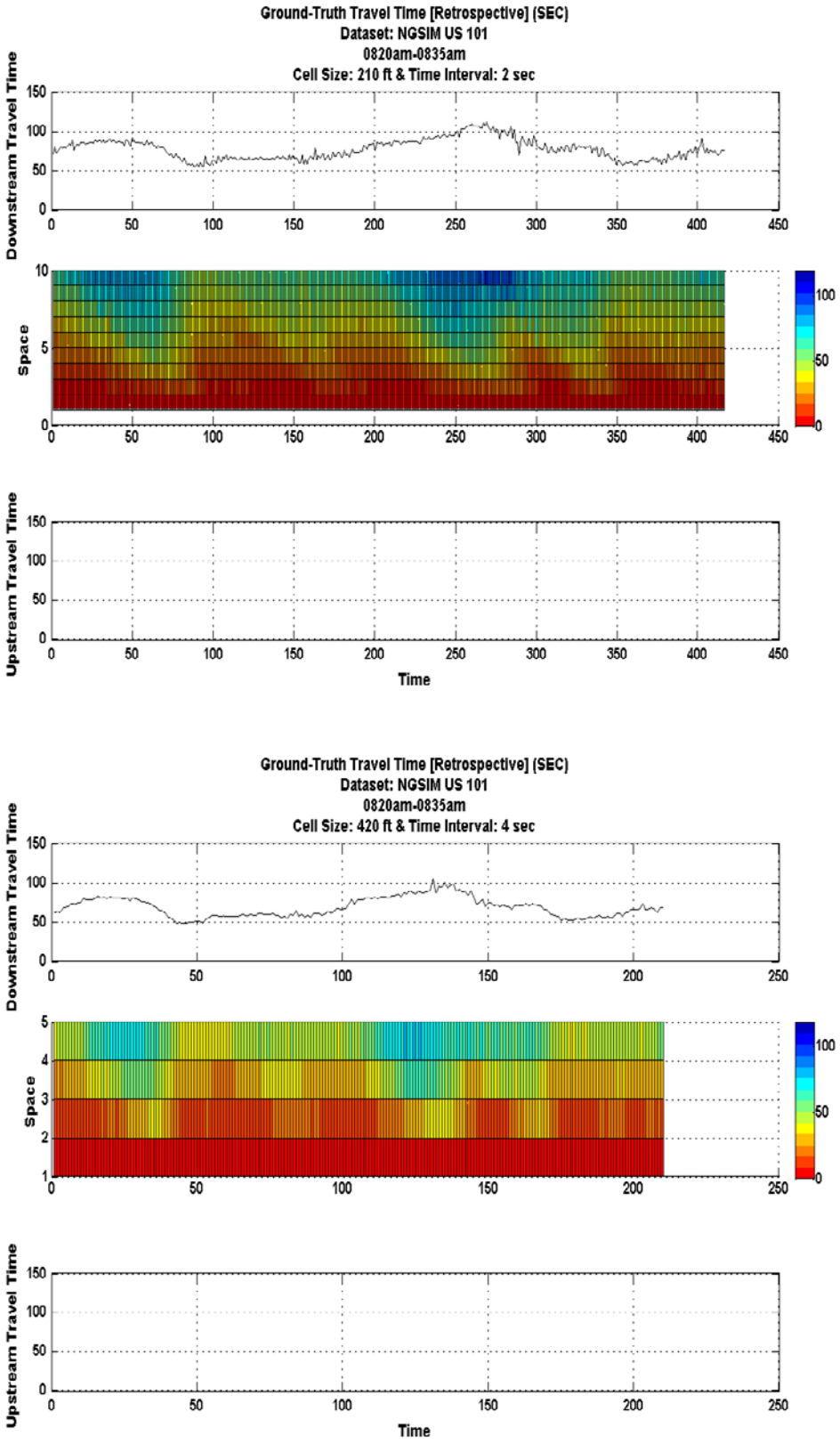


Figure 20. Retrospective travel time ground truth US 101 at 8:20AM-8:35AM.

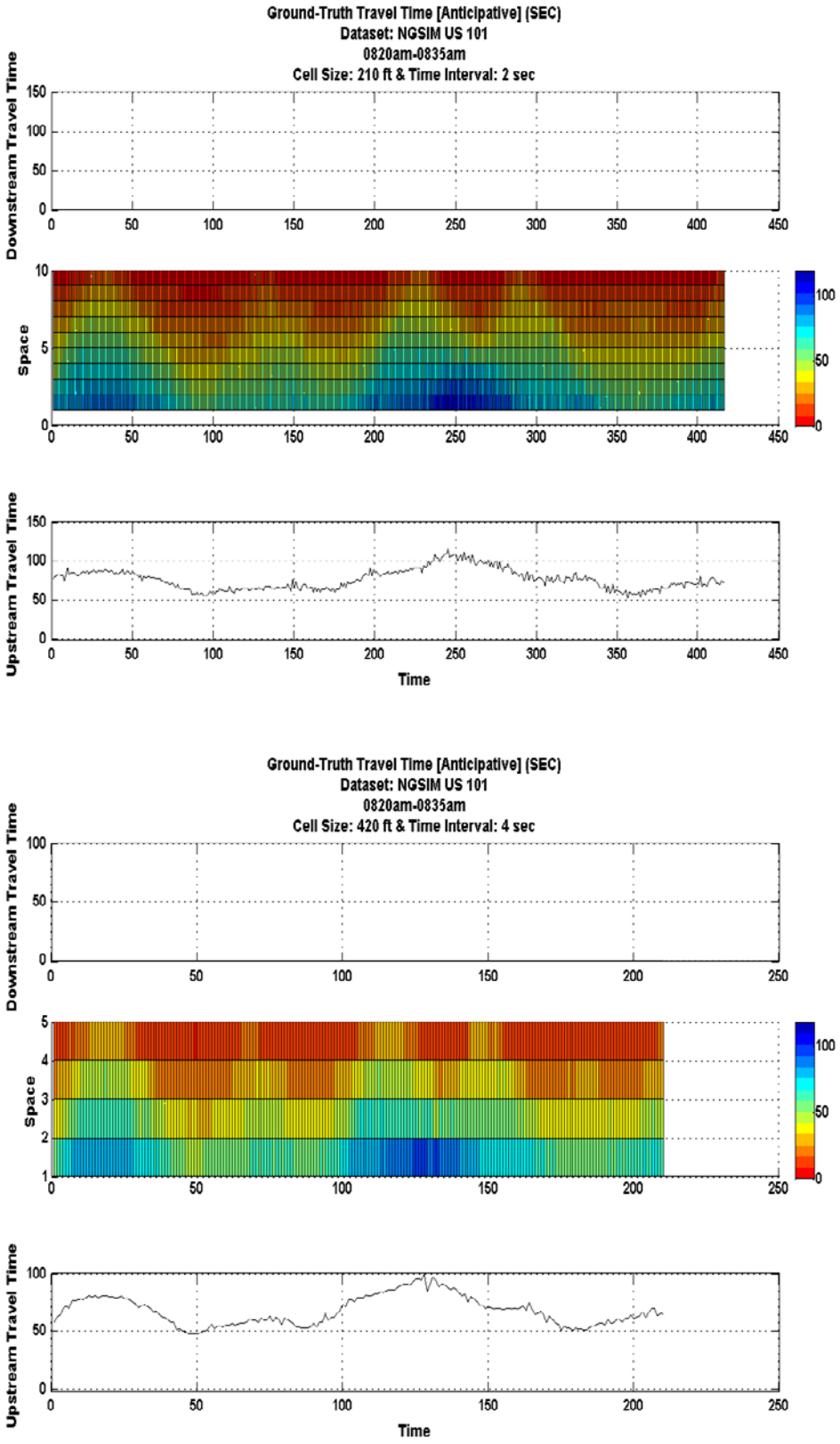


Figure 21. Anticipative travel time ground truth US 101 at 8:20AM-8:35AM.

6.3 *Traffic Modeling*

In this section the modeling effort as applied to the case of three aforementioned datasets is described. First, speed-density relationships used in the modeling and estimation process are presented. Then, using the naïve approach traffic and travel time model errors are estimated. Later, measurement errors in the case of speed and travel times are specified. The resulting model and measurement errors are used as input in the optimal estimation process.

6.3.1 Speed-density relationships (GS & HL)

Based on available data speed-density relationship parameters are estimated. Free-flow speed and jam density are estimated at 65mph and 200vpmpl, respectively. Hence, the Greenshields relationship can be specified as the following:

$$V_{GS}(\rho) = 65(1 - \rho/200) \quad , 0 \leq \rho \leq 200 \tag{6.1}$$

Also, the wave speed and critical density are estimated at 15mph and 45vpmpl, respectively. Hence, the hyperbolic-linear relationship can be specified as the following:

$$V_{HL}(\rho) = \begin{cases} 65\left(1 - \frac{\rho}{200}\right) & , \rho \leq 45 \\ -15\left(1 - \frac{200}{\rho}\right) & , \rho > 45 \end{cases} \tag{6.2}$$

Figure 22 illustrates the estimated speed-density relationships used in this dissertation. Greenshields relationship is shown on the left, while the hyperbolic-linear relationship is shown on the right. The top row of graphs represent the hypothesized relationships between speed and density, while the bottom row of graphs represent the resulting relationships between flow rate and density in each case.

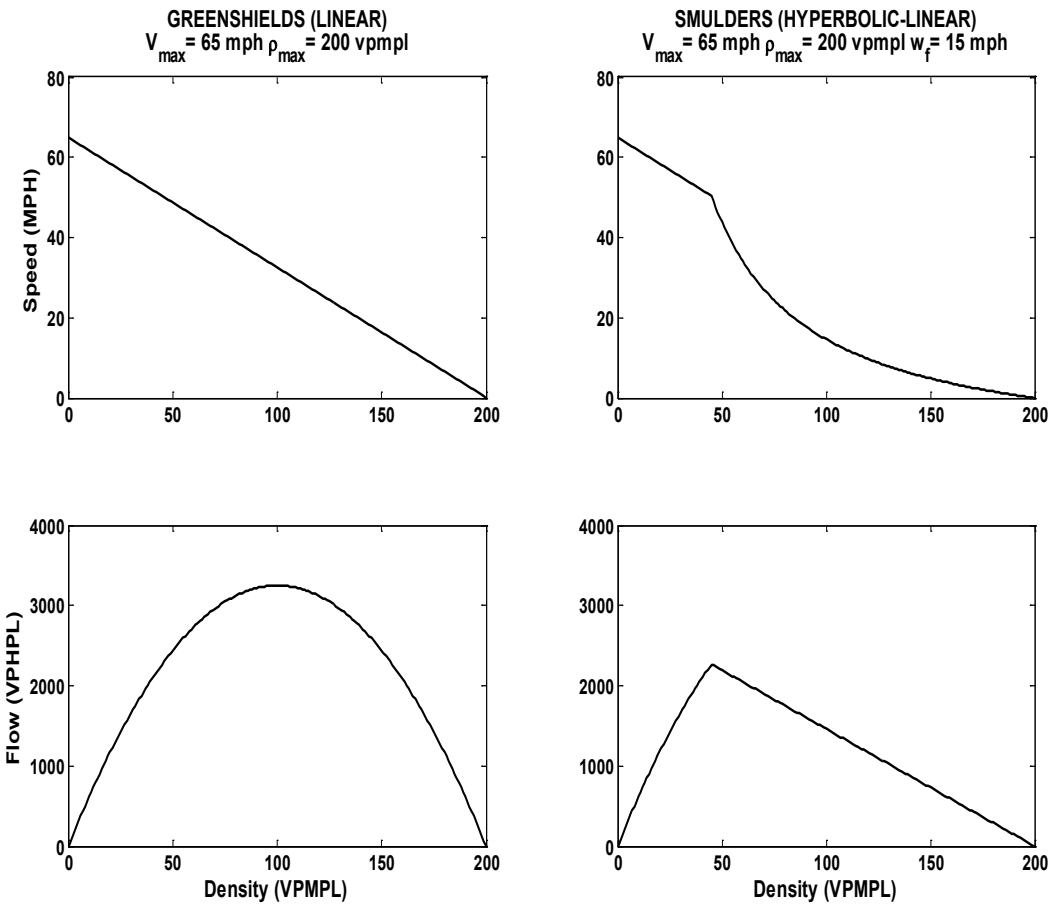


Figure 22. Speed-density relationships (Left: Greenshields, Right: Smulders, Top: Speed-Density, Bottom: Flow-Density).

6.3.2 Discretization Levels

Speed and travel time models are solved at two different discretization levels. At the first discretization level, space dimension is divided into 210ft (64m) cells, and the time update interval is chosen to be equal to two seconds. The second discretization level doubles both space and time dimensions. The cell sizes used in the second discretization level are 420ft (128m) long, and the time update is performed at every four seconds. It should be noted that in both cases the $\frac{\Delta x}{\Delta t}$ ratio is equal to 105 feet per second (71.6mph) which is slightly higher than the adopted free flow speed (65mph) satisfying the stability condition.

6.3.3 Modeling Errors

Models presented in Chapter 4 and ground-truth data are used to estimate model errors. The approach can be described as a naïve application of state-space model in which case exact measurements (ground-truth) are used as boundary values in solving the proposed finite difference traffic and travel time models. At every time interval model equations are used to advance the state variable estimates to the next time interval. In this case no attempt is made to improve the current estimate by taking into account the correlations between measurements and system states.

To estimate model errors, ground-truth data (no measurement error) is fed into the state-space model as initial and boundary conditions. Estimation is performed in the open loop sense; that is, no measurement update is performed to correct the model predictions. Estimation results are compared against the known ground-truth to evaluate model errors.

Table 2 summarizes the model error estimates obtained from this process. The top table reports the speed estimation errors based on the CTM-v model. The middle and bottom tables report the retrospective and anticipative travel time estimation errors using the THETA and TAU models, respectively. In these tables mean, standard deviation, and three quartile points of the error distributions in each time period are reported. Also, error measures are reported when either Greenshields or Smulders speed-density relation is used as well as when problems are solved at either of the two discretization levels.

Note that models are generally biased with error means substantially different from zero. The bias in the case of speeds is positive which indicate a tendency for CTM-v model to overestimate the speeds. This has naturally resulted in the underestimation of retrospective travel times. On the other hand, anticipative travel times are overestimated.

Travel time error measures indicate that travel time models are very sensitive to the speed quality. It is shown elsewhere (Sadabadi & Haghani, 2012) that travel time model errors are much lower when ground-truth speeds are used to estimate travel times.

Table 2. Error measures in naïve application of models to US 101 datasets.

CTM-v	Cell Size	Greenshields										Smulders									
		Error (mph)					Absolute Error (mph)					Error (mph)					Absolute Error (mph)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	2	6	-3	2	7	6	4	3	5	8	8	8	1	6	15	9	7	3	7	15
	420x4	2	6	-2	3	7	6	4	2	5	8	9	8	2	8	16	10	7	3	8	16
0805am-0820am	210x2	2	5	-2	1	5	4	3	1	3	6	2	4	-1	2	4	3	3	1	3	5
	420x4	2	5	-2	1	5	4	3	2	3	5	2	4	-1	2	5	4	3	1	3	5
0820am-0835am	210x2	2	5	-2	1	4	4	4	1	3	6	2	4	-1	2	4	3	3	1	3	5
	420x4	2	5	-1	2	4	4	3	2	3	5	2	4	-1	2	5	4	3	2	3	5

THETA	Cell Size	Greenshields										Smulders									
		Error (sec)					Absolute Error (sec)					Error (sec)					Absolute Error (sec)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	-9	6	-12	-8	-4	9	5	4	8	12	-12	7	-18	-12	-6	12	7	6	12	18
	420x4	-10	6	-13	-9	-5	10	6	5	9	13	-14	7	-20	-14	-8	14	7	8	14	20
0805am-0820am	210x2	-10	11	-15	-7	-2	11	10	3	7	15	-10	12	-16	-7	-2	12	11	3	8	16
	420x4	-10	9	-15	-8	-3	10	9	3	8	15	-11	10	-14	-8	-3	11	10	4	8	14
0820am-0835am	210x2	-12	9	-17	-11	-5	12	8	5	11	17	-12	9	-19	-10	-4	12	9	5	10	19
	420x4	-13	9	-19	-12	-6	13	9	6	12	19	-13	9	-20	-11	-6	13	9	6	11	20

TAU	Cell Size	Greenshields										Smulders									
		Error (sec)					Absolute Error (sec)					Error (sec)					Absolute Error (sec)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	8	5	4	8	12	8	5	4	8	12	12	7	6	12	17	12	6	6	12	17
	420x4	9	6	5	9	13	9	6	5	9	13	14	7	8	14	20	14	7	8	14	20
0805am-0820am	210x2	10	9	3	8	16	10	8	3	9	16	10	8	4	9	16	11	7	4	9	16
	420x4	10	9	3	9	16	11	8	4	9	16	11	8	5	10	17	11	8	5	10	17
0820am-0835am	210x2	11	10	3	9	16	12	9	4	10	16	12	8	6	11	16	12	8	6	11	16
	420x4	12	11	4	11	19	13	10	4	11	19	13	9	7	13	18	13	9	7	13	18

Figure 23 represents the naïve modeling error boxplots at each cell along the segment in the case of Dataset 1 with the finer discretization at 210ft by 2 sec. Spatial trend of speed estimate errors indicate that CTM-v produces overestimated speeds over the first half of the segment, while over the other half speeds are underestimated. Moving from upstream to downstream, THETA generally and increasingly underestimates the retrospective travel times, but towards the end of the segment the error distributions become stable. Note that TAU model generated a similar but completely opposite pattern by overestimating anticipative travel times with increasing error when moving in the direction of traffic along the segment.

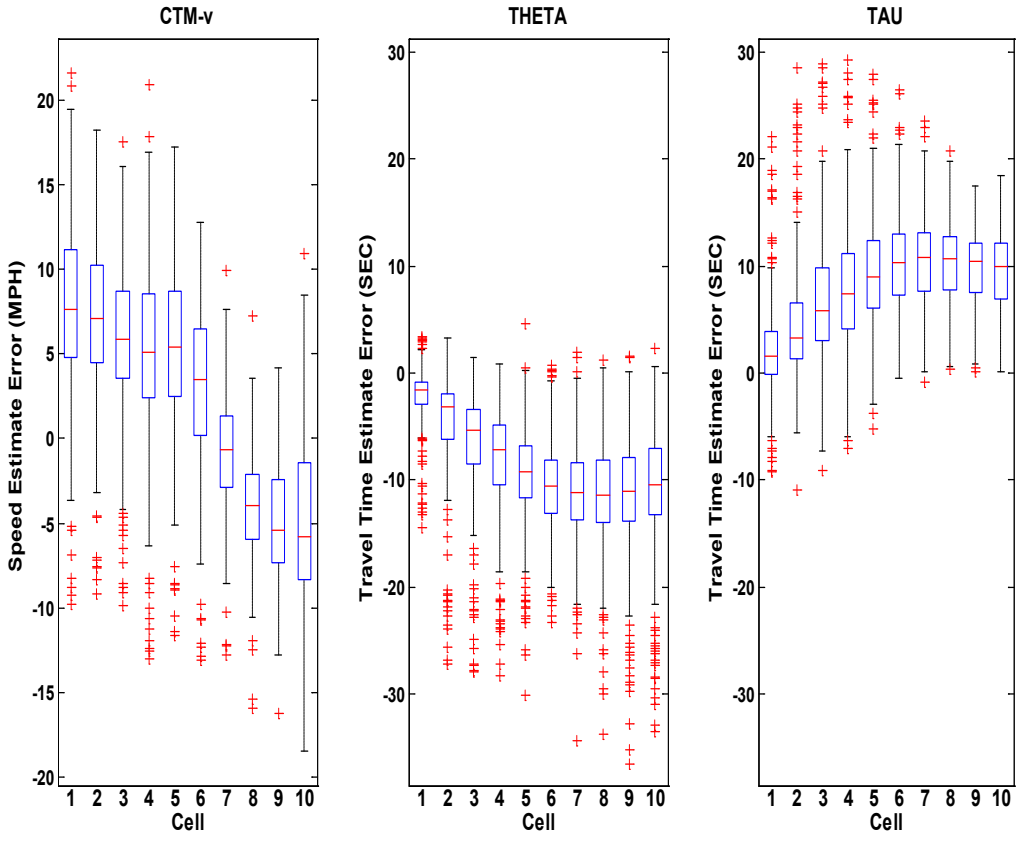


Figure 23. Error boxplots in naïve application of models to Dataset 1 (210ft x 2sec).

Figure 24 illustrates the temporal variations of model estimate errors. Clearly, in the mean error time series, strong temporal correlations between estimate errors in all models can be observed.

Again, note that for the most part, speed and anticipative travel times are overestimated while retrospective travel times are underestimated.

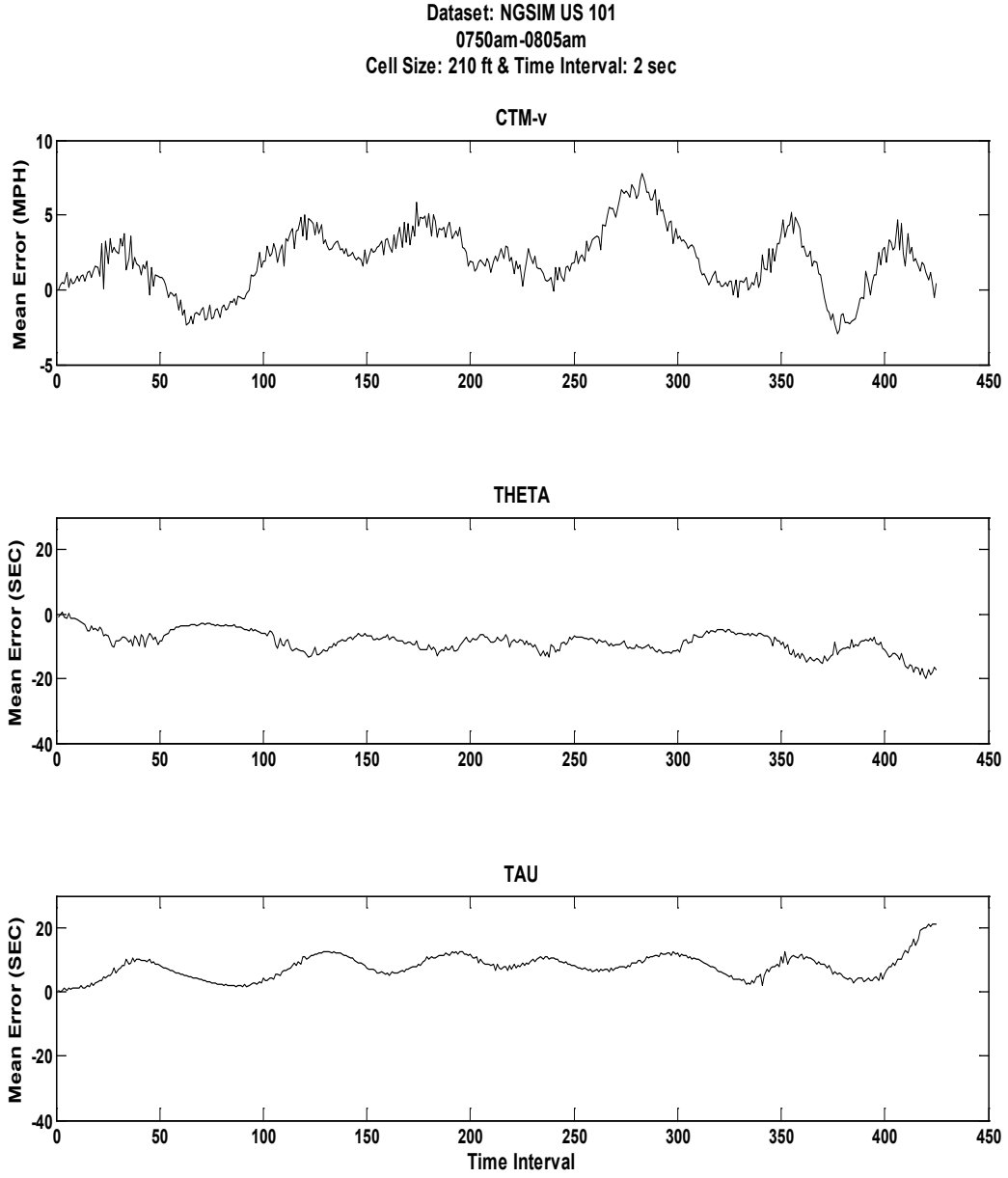


Figure 24. A typical temporal variation of mean model errors.

Taking advantage of error estimates obtained from this open-loop estimation process, model error covariance matrices for different lags are calculated. These covariance matrices are used in the optimal estimation process that will be presented in the next section.

6.3.4 Measurement errors

In reality, every measurement contains error. Hence, when measurements are taken into account, their accuracy levels should also be considered. In traffic applications, the measurement error reflects the accuracy of technology, aggregations and assumptions used in field data collections.

In this dissertation, for simplicity and clarity of analysis, measurements are assumed to incorporate an unbiased (zero mean) and uncorrelated Gaussian error (white noise). Nevertheless, the proposed models and estimation techniques are capable of handling other forms of measurement errors being non-Gaussian, biased, and or correlated of various forms (colored noise).

The white noise assumption also implies that measurement equipment is well-calibrated and properly maintained. Therefore, errors associated with measurements of the same state variable at different locations and at different times (as long as the same technology is used) can be simply specified by a diagonal covariance matrix.

In this dissertation, spot speed measurements are assumed to come from loop detectors or side-fire microwave radars located at the boundaries of the segment under investigation. Travel time measurements are assumed to come from a pair of automatic vehicle re-identification devices (such as Bluetooth traffic monitoring units, or license-plate reader cameras) also placed at the two boundaries of the segment.

Based on previous studies (Zwahlen, et al. 2005 & Ki and Baik 2006), it is assumed that speed measurement error has a 3mph standard deviation. Note that under congested conditions (effectively speeds less than 30mph) this amounts to Root Mean Squared Errors (RMSE) in speed measurements that are larger than 10% of the actual speeds.

Travel time measurement error using AVI technology such as Bluetooth detection is shown to be equal to half the size of reporting time interval used. Note that for obvious reasons, speed and

travel time measurement errors are not correlated with each other. Therefore, measurement error covariance function defined in section 5.1.5 can be specified as:

$$R_k = \begin{bmatrix} 3^2 & 0 & 0 \\ 0 & 3^2 & 0 \\ 0 & 0 & \left(\frac{\Delta t}{2}\right)^2 \end{bmatrix} \tag{6.3}$$

6.4 Estimation Results

Table 3 shows different aspects of numerical experiments performed in this study. In total 120 different experiments are performed. Error measures reported for each experiment is based on 10 instances generated by adding random errors to the ground-truth input measurements.

Table 3. Numerical experiments dimensions.

Item	Variations
Estimation method	Unscented Kalman Filter (UKF) Unscented H-infinity Filter (UHF)
Fundamental diagram	Greenshields (GS) Smulders (HL)
Discretization size	210 ft x 2 sec 420 ft x 4 sec
Datasets	0750am-0805am (D1) 0805am-0820am (D2) 0820am-0835am (D3)
Delay filter (when travel time measurements are used)	Yes No
Input measurement data	Speed Travel time (theta) Speed and travel time
Number of instances	10

Note that in the case of three US-101 datasets, the free-flow and maximum congested travel times between upstream and downstream of the segment are nearly 50 and 100 seconds, respectively. Therefore, at four second resolution about four to eight percent of travel time estimate errors between the two boundaries can be attributed to the discretization scheme used. Similarly, when solutions at two second resolution are considered the discretization errors will vary between two to four percent.

6.4.1 Boundary Measurements Used as Input: Speed

Table 4 shows the error in state estimates using UKF estimation method when only boundary speed measurements are used as input. As expected, these experiments resulted in very accurate speed estimates with maximum mean absolute error (MAE) obtained in all cases is only four miles per hour (6.4 km/hr). Travel time estimates on the other hand are more biased and mean absolute errors up to 12 seconds are reported during more congested periods. In general, while speed estimates using Smulders speed-density relationship are slightly worse than estimates using Greenshields relationship, travel time estimates based on the former relation are more accurate. Also, results indicate that increasing the discretization size (reducing resolution), in general, leads to smaller errors both on average and in the spread.

Table 4. Error measures in UKF estimation of state variables in US 101 datasets. (Speed inputs)

CTM-v	Cell Size	Greenshields										Smulders									
		Error (mph)					Absolute Error (mph)					Error (mph)					Absolute Error (mph)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	1	4	-1	1	3	3	3	1	2	4	3	4	0	2	5	4	4	1	3	6
	420x4	1	4	-1	0	2	2	3	1	1	3	1	4	-1	0	3	3	3	1	2	4
0805am-0820am	210x2	1	4	-1	0	3	3	3	1	2	4	0	3	-2	0	2	3	2	1	2	4
	420x4	1	4	-1	0	2	2	3	1	1	4	0	3	-1	0	1	2	2	1	1	3
0820am-0835am	210x2	1	4	-1	1	4	3	3	1	2	4	1	3	-1	0	3	3	2	1	2	4
	420x4	1	4	-1	0	2	2	3	1	1	3	0	3	-1	0	1	2	2	1	1	3

THETA	Cell Size	Greenshields										Smulders									
		Error (sec)					Absolute Error (sec)					Error (sec)					Absolute Error (sec)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	-1	6	-3	0	2	4	4	1	3	6	-2	5	-4	-1	1	4	4	1	3	5
	420x4	-1	5	-4	-1	1	3	4	1	2	5	-3	4	-5	-2	0	4	3	1	3	5
0805am-0820am	210x2	2	17	-6	0	7	11	13	2	6	15	-1	12	-6	0	5	8	9	2	5	13
	420x4	0	9	-5	0	3	6	6	1	4	9	-1	7	-5	0	2	5	5	1	3	7
0820am-0835am	210x2	5	18	-3	1	12	12	14	2	7	17	1	11	-5	0	6	8	8	2	5	12
	420x4	3	17	-4	1	9	12	13	2	6	18	-1	10	-6	0	3	7	7	1	4	10

TAU	Cell Size	Greenshields										Smulders									
		Error (sec)					Absolute Error (sec)					Error (sec)					Absolute Error (sec)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	-1	7	-3	0	3	4	5	1	3	6	-1	6	-3	0	2	4	5	1	3	6
	420x4	-2	6	-5	-2	1	4	4	1	3	6	-2	5	-4	-1	1	4	4	1	2	5
0805am-0820am	210x2	2	16	-5	1	9	11	11	3	7	17	0	12	-5	1	7	9	8	2	6	13
	420x4	-2	10	-5	0	4	7	7	2	5	10	-2	9	-5	0	3	6	6	1	4	8
0820am-0835am	210x2	7	17	-3	4	14	13	13	3	8	17	3	10	-3	2	9	8	7	3	6	12
	420x4	1	14	-7	1	7	10	10	3	7	15	-2	10	-7	-1	3	7	7	2	5	10

6.4.2 Boundary Measurements Used as Input: Travel Time

Table 5 shows the error in state estimates using UKF estimation method when only realized travel times measured at downstream (theta) are used as input. Results indicate that travel time estimates are generally improved; while speed estimates are generally deteriorated compared to when boundary speed measurements are used. Interestingly, the speeds are overestimated (positive bias) while retrospective travel times are generally underestimated (negative bias). The bias in anticipative travel time estimates varies between negative one and two seconds in all scenarios. While compared to Smulders, Greenshields relationship has led to slightly more accurate speed estimates in these cases, both relationships overall have resulted in very similar travel time estimates.

Table 5. Error measures in UKF estimation of state variables in US 101 datasets. (Travel time inputs, no delayed filter)

CTM-v	Cell Size	Greenshields										Smulders									
		Error (mph)					Absolute Error (mph)					Error (mph)					Absolute Error (mph)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	1	8	-3	2	6	6	5	2	5	9	2	7	-2	3	7	6	5	2	5	9
	420x4	-1	7	-5	0	4	5	4	2	5	8	0	6	-4	1	5	5	4	2	4	7
0805am-0820am	210x2	3	6	-1	3	7	5	4	2	4	8	6	6	1	5	10	6	5	2	5	10
	420x4	2	5	-2	1	5	4	3	2	4	6	3	5	0	3	7	5	4	2	4	7
0820am-0835am	210x2	4	6	0	4	8	6	4	2	5	8	6	6	2	6	10	7	5	3	6	10
	420x4	2	6	-2	1	5	5	4	2	4	7	4	6	0	4	7	5	4	2	4	7

THETA	Cell Size	Greenshields										Smulders									
		Error (sec)					Absolute Error (sec)					Error (sec)					Absolute Error (sec)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	-2	4	-4	-1	0	3	4	0	2	5	-2	4	-4	-1	0	3	4	0	2	5
	420x4	-2	4	-3	-1	0	2	3	0	1	4	-2	4	-3	-1	0	3	3	0	1	4
0805am-0820am	210x2	-4	7	-7	-1	0	5	6	0	2	7	-4	7	-7	-2	0	5	6	1	2	7
	420x4	-2	5	-3	0	0	3	5	0	1	3	-2	5	-3	0	0	3	5	0	1	4
0820am-0835am	210x2	-5	7	-9	-3	0	6	6	1	4	9	-5	7	-9	-3	0	6	6	1	4	9
	420x4	-3	6	-5	0	0	4	5	0	1	6	-3	6	-5	0	0	4	5	0	1	6

TAU	Cell Size	Greenshields										Smulders									
		Error (sec)					Absolute Error (sec)					Error (sec)					Absolute Error (sec)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	1	5	0	2	4	4	4	1	3	5	1	5	0	2	3	3	4	1	3	4
	420x4	2	5	0	2	4	3	4	0	3	5	2	5	0	2	4	4	4	0	3	5
0805am-0820am	210x2	-1	6	-3	0	3	4	5	1	3	6	-1	6	-3	0	3	4	5	1	3	6
	420x4	0	6	-1	0	2	4	4	0	2	6	0	6	-1	0	2	4	4	0	2	5
0820am-0835am	210x2	0	6	-3	0	5	5	4	1	4	7	1	6	-3	1	5	5	4	1	4	7
	420x4	0	6	-3	0	3	4	4	1	3	6	-1	6	-4	0	3	4	4	1	4	6

6.4.3 Boundary Measurements Used as Input: Speed and Travel Time

Table 6 shows the error in state estimates using UKF estimation method when boundary speeds as well as the realized travel times measurements at downstream (theta) are used as input. Results indicate that estimate errors increase with congestion, but they tend to decrease with increase in the discretization size. Speed estimate bias is between zero and two mph. Retrospective travel time estimate bias varies between minus three and minus one seconds, while its MAE varies between one and five seconds. Anticipative travel time estimate bias is between minus one and four seconds, while its corresponding MAE varies between two and seven seconds.

Table 6. Error measures in UKF estimation of state variables in US 101 datasets. (Speed and travel time inputs, no delayed filter)

CTM-v	Cell Size	Greenshields										Smulders									
		Error (mph)					Absolute Error (mph)					Error (mph)					Absolute Error (mph)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	1	3	-1	1	3	3	3	1	2	4	2	4	0	2	5	3	3	1	2	5
	420x4	1	3	-1	0	2	2	2	1	1	3	1	4	-1	0	3	3	3	1	1	4
0805am-0820am	210x2	1	4	-1	0	3	3	3	1	2	4	0	3	-2	0	2	3	2	1	2	4
	420x4	1	3	-1	0	2	2	3	1	1	4	0	3	-1	0	1	2	2	1	1	3
0820am-0835am	210x2	2	4	-1	1	4	3	3	1	2	4	1	4	-1	0	3	3	3	1	2	4
	420x4	1	4	-1	0	2	2	3	1	1	3	0	3	-1	0	1	2	2	1	1	3

THETA	Cell Size	Greenshields										Smulders									
		Error (sec)					Absolute Error (sec)					Error (sec)					Absolute Error (sec)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	-1	3	-2	0	1	2	3	0	1	3	-1	3	-3	0	0	2	3	0	1	3
	420x4	-1	2	-1	0	0	1	2	0	1	2	-1	2	-2	0	0	1	2	0	1	2
0805am-0820am	210x2	-3	7	-5	0	1	4	6	0	2	6	-3	6	-5	0	1	4	6	0	2	6
	420x4	-1	4	-2	0	1	2	3	0	1	3	-1	4	-1	0	1	2	3	0	1	3
0820am-0835am	210x2	-2	7	-5	0	1	5	5	0	3	7	-2	7	-5	0	1	5	5	0	3	7
	420x4	-1	5	-3	0	0	3	4	0	1	5	-1	5	-2	0	0	3	4	0	1	4

TAU	Cell Size	Greenshields										Smulders									
		Error (sec)					Absolute Error (sec)					Error (sec)					Absolute Error (sec)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	0	5	-1	0	2	3	4	0	1	3	0	5	-1	0	2	3	4	1	1	4
	420x4	0	4	-1	0	1	2	3	0	1	3	0	4	0	0	2	2	3	0	1	3
0805am-0820am	210x2	0	10	-3	0	5	7	7	1	4	10	0	8	-2	0	4	5	6	1	4	8
	420x4	-1	6	-2	0	2	4	5	0	2	5	-1	6	-2	0	2	4	5	0	2	5
0820am-0835am	210x2	4	9	0	3	8	7	7	2	5	9	1	7	-2	1	6	5	5	2	4	8
	420x4	0	5	-2	0	3	3	3	0	2	5	-2	5	-4	0	2	4	4	0	2	5

6.5 Delay Filter Impact

This section describes the impact of using delay filter on state variable estimation. Note that delay filter can only be applied when boundary retrospective travel time measurements are part of the input.

6.5.1 Boundary Measurements Used as Input: Travel Time

Table 7 shows the error in state estimates using UKF estimation method with delayed filter application when boundary travel time measurements at downstream (theta) are used as input. Results indicate that application of delay filter in this case has led to marginal improvements on the quality of both speeds and retrospective travel time estimates. However, using the delayed filter has led to mixed results in terms of the anticipative travel time estimates. By and large, when only travel time measurements are used as the input, using delayed filters has resulted in deterioration of the quality of anticipative travel time estimates.

Table 7. Error measures in UKF estimation of state variables in US 101 datasets. (Travel time inputs, delayed filter)

CTM-v	Cell Size	Greenshields										Smulders									
		Error (mph)					Absolute Error (mph)					Error (mph)					Absolute Error (mph)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	1	8	-3	2	6	6	5	2	5	9	2	7	-2	3	7	6	5	2	5	9
	420x4	-1	6	-5	0	3	5	4	2	4	7	0	6	-3	1	4	5	4	2	4	7
0805am-0820am	210x2	3	6	-1	3	7	5	4	2	4	7	5	6	1	5	9	6	5	2	5	9
	420x4	1	5	-2	1	5	4	3	2	4	6	3	5	0	3	7	5	3	2	4	7
0820am-0835am	210x2	4	6	0	4	8	6	4	2	5	8	6	6	2	6	9	7	5	3	6	9
	420x4	1	6	-3	1	4	4	4	2	4	7	3	5	-1	3	6	5	4	2	4	7

THETA	Cell Size	Greenshields										Smulders									
		Error (sec)					Absolute Error (sec)					Error (sec)					Absolute Error (sec)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	-2	4	-4	-1	0	3	4	0	2	5	-2	4	-4	-2	0	3	4	0	2	5
	420x4	-2	4	-3	-1	0	2	3	0	1	3	-2	4	-3	-1	0	2	3	0	1	3
0805am-0820am	210x2	-4	7	-7	-1	0	5	6	0	2	7	-4	7	-7	-1	0	5	6	0	2	7
	420x4	-2	5	-3	0	0	3	5	0	1	3	-2	5	-3	0	0	3	4	0	1	3
0820am-0835am	210x2	-5	7	-9	-3	0	6	6	1	3	9	-5	7	-9	-3	0	6	6	1	4	9
	420x4	-3	6	-5	0	0	4	5	0	1	5	-2	5	-5	0	0	4	5	0	1	5

TAU	Cell Size	Greenshields										Smulders									
		Error (sec)					Absolute Error (sec)					Error (sec)					Absolute Error (sec)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	1	6	0	2	4	4	4	1	3	5	1	5	0	2	3	4	4	1	3	5
	420x4	2	5	0	1	3	3	4	0	2	4	1	5	0	1	3	3	4	0	2	4
0805am-0820am	210x2	0	10	-3	0	5	7	7	1	4	11	-1	9	-3	0	5	6	7	1	4	10
	420x4	0	7	-1	0	3	5	6	0	2	7	0	7	-1	0	2	4	5	0	2	7
0820am-0835am	210x2	1	9	-4	1	7	7	6	2	5	10	1	9	-4	1	6	7	6	2	5	10
	420x4	0	8	-3	0	5	6	5	1	5	9	0	7	-3	0	5	5	5	1	4	8

6.5.2 Boundary Measurements Used as Input: Speed and Travel Time

Table 8 shows the error in state estimates using UKF estimation method with delayed filter application when boundary speeds as well as the realized travel times measurements at downstream (theta) are used as input. Results indicate that while speed and retrospective travel time estimate qualities for the most part have remained unchanged, applying the delayed filter has improved the quality of predictive travel times.

Table 8. Error measures in UKF estimation of state variables in US 101 datasets. (Speed and travel time inputs, delayed filter)

CTM-v	Cell Size	Greenshields										Smulders									
		Error (mph)					Absolute Error (mph)					Error (mph)					Absolute Error (mph)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	1	3	-1	1	3	3	3	1	2	4	3	4	0	2	5	3	3	1	2	5
	420x4	1	3	-1	0	2	2	2	1	1	3	1	4	-1	1	3	3	3	1	1	4
0805am-0820am	210x2	1	4	-1	0	3	3	3	1	2	4	0	3	-2	0	2	3	2	1	2	4
	420x4	0	3	-1	0	2	2	3	1	1	4	0	3	-1	0	1	2	2	0	1	3
0820am-0835am	210x2	1	4	-1	1	3	3	3	1	2	4	1	4	-1	0	3	3	3	1	2	4
	420x4	1	4	-1	0	2	2	3	1	1	3	0	3	-1	0	1	2	2	1	1	3

THETA	Cell Size	Greenshields										Smulders									
		Error (sec)					Absolute Error (sec)					Error (sec)					Absolute Error (sec)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	-1	4	-2	0	1	2	3	0	1	3	-1	3	-3	0	0	2	3	0	1	3
	420x4	-1	2	-1	0	0	1	2	0	1	2	-1	2	-2	0	0	1	2	0	1	2
0805am-0820am	210x2	-3	7	-5	0	1	4	6	0	2	6	-3	7	-5	0	1	4	6	0	2	6
	420x4	-1	4	-2	0	1	2	3	0	1	3	-1	4	-1	0	1	2	3	0	1	3
0820am-0835am	210x2	-2	7	-5	0	1	5	5	0	3	7	-3	7	-5	0	0	5	6	0	3	7
	420x4	-1	5	-3	0	0	3	4	0	1	5	-1	5	-2	0	1	3	4	0	1	4

TAU	Cell Size	Greenshields										Smulders									
		Error (sec)					Absolute Error (sec)					Error (sec)					Absolute Error (sec)				
Time Period		Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile	Mean	Std-Dev	25-ile	50-ile	75-ile
0750am-0805am	210x2	0	5	-1	0	1	3	4	0	1	3	0	5	-1	0	1	3	4	0	1	4
	420x4	0	4	-1	0	1	2	3	0	1	3	0	4	0	0	2	2	3	0	1	3
0805am-0820am	210x2	-2	10	-4	0	3	6	8	1	3	8	-2	9	-3	0	3	6	7	1	3	8
	420x4	-1	7	-2	0	2	4	5	0	2	5	-1	7	-2	0	2	4	5	0	2	6
0820am-0835am	210x2	-1	7	-4	0	3	5	5	1	4	7	0	7	-4	0	4	5	5	1	4	7
	420x4	-1	5	-4	0	1	4	4	0	2	5	-1	5	-3	0	2	4	4	0	3	6

When boundary speed measurements have been part of the input, invariably it has resulted in the best speed estimates. For the most part, speed estimates have been insensitive to the application of the delayed filter when retrospective travel times have also been used in the estimation. The best travel time estimates (both retrospective and anticipative) are obtained when both speeds and retrospective travel times are used as input. Application of delayed filters has had marginal impact on the quality of retrospective travel times. However, using the delayed filter has led to mixed results in terms of the anticipative travel time estimates. When only travel time measurements are used as the input, using delayed filters by and large has deteriorated the quality of anticipative travel time estimates. But, when speed and travel time measurements are used simultaneously as the input, using delayed filters has resulted in relatively similar or better anticipative travel time estimates especially during more congested time periods.

Also, note that in general errors become larger with increasing congestion, but they become smaller with increasing discretization cell sizes and aggregation time intervals.

Generally speaking, using Smulders relation generally did improve the results especially in the case of speed estimates. However, in the case of anticipative travel times the best estimates resulting from both relations with speed and travel time measurements as input and using the delayed filter ultimately are very similar.

Table 9. Maximum MAE measures in UKF estimation of state variables in US 101 datasets.

			Maximum MAE in Estimation		
			Speed	Retrospective Travel Time	Anticipative Travel Time
		Input Measurements	mph	sec	sec
Delay Filter	No	Speed	4/3	12/12	13/10
		Retro-Travel Time	7/5	6/4	5/4
		Speed & Retro-Travel Time	3/3	5/3	7/4
	Yes	Speed			
		Retro-Travel Time	7/5	6/4	7/6
		Speed & Retro-Travel Time	3/3	5/3	6/4

Table 9 summarizes the maximum MAE estimates in UKF estimation of three sets of state variables in the case of US 101 datasets. It is clear that when each state variable has been measured on the boundaries, on average its corresponding estimates have improved. When speed and retrospective travel times are both measured on the boundaries speed estimates with maximum three mph MAE at both discretization levels are obtained. Maximum MAE in travel time estimates are almost halved when discretization level has increased from two seconds to four seconds. Retrospective travel time estimates with maximum five and three second MAE are obtained in these cases. Anticipative travel time estimates have been slightly worse with maximum seven and four second MAE in the two discretization levels, respectively. Taking advantage of delayed filter did not have a clear impact on the maximum MAE of speed and travel time estimates.

Figure 25 exhibits the Mean Absolute Percentage Error (MAPE) of speed and travel time estimates when Greenshields relation and UKF estimation method is used. Graphs on the left and on the right in each case represent error measures when two and four second discretization schemes are adopted, respectively. The labels on the horizontal axis represent the index of 15 minute dataset for which the error estimates are being reported. The colored bars in each graph indicate the combination of input measurement and delayed filter scenarios used to estimate the three state variables. Note that MAPE reported in these graphs is an average over both space and time.

Figure 26 is similar to Figure 25 except that it exhibits the MAPE of speed and travel time estimates when Smulders relation is used. In both cases, note that in general errors become larger with increasing congestion, while they become smaller with increasing discretization cell sizes and aggregation time intervals.

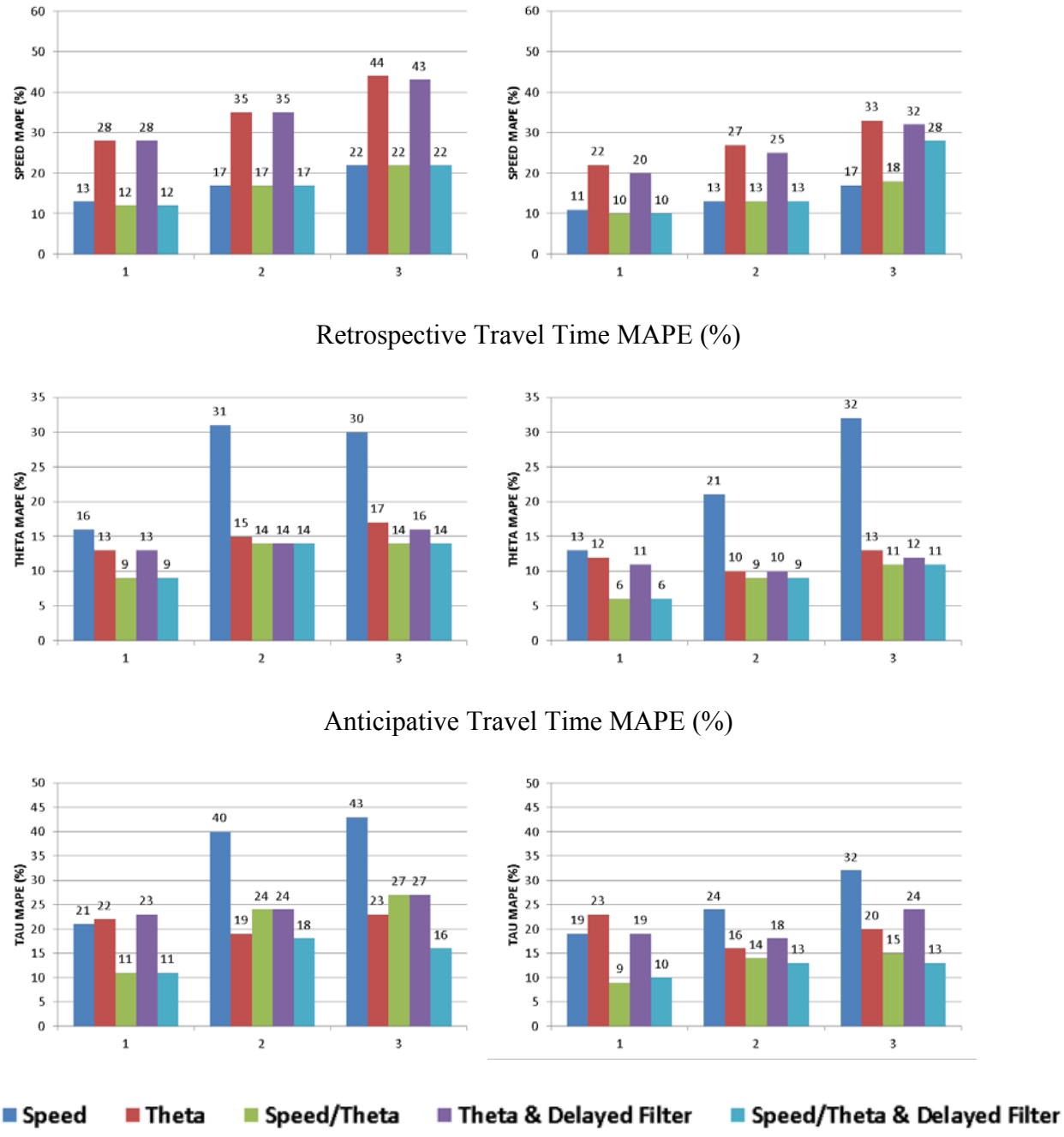


Figure 25. MAPE of estimates using Greenshields relation and UKF method. (Left: 210ftx2sec; Right: 420ftx4sec)

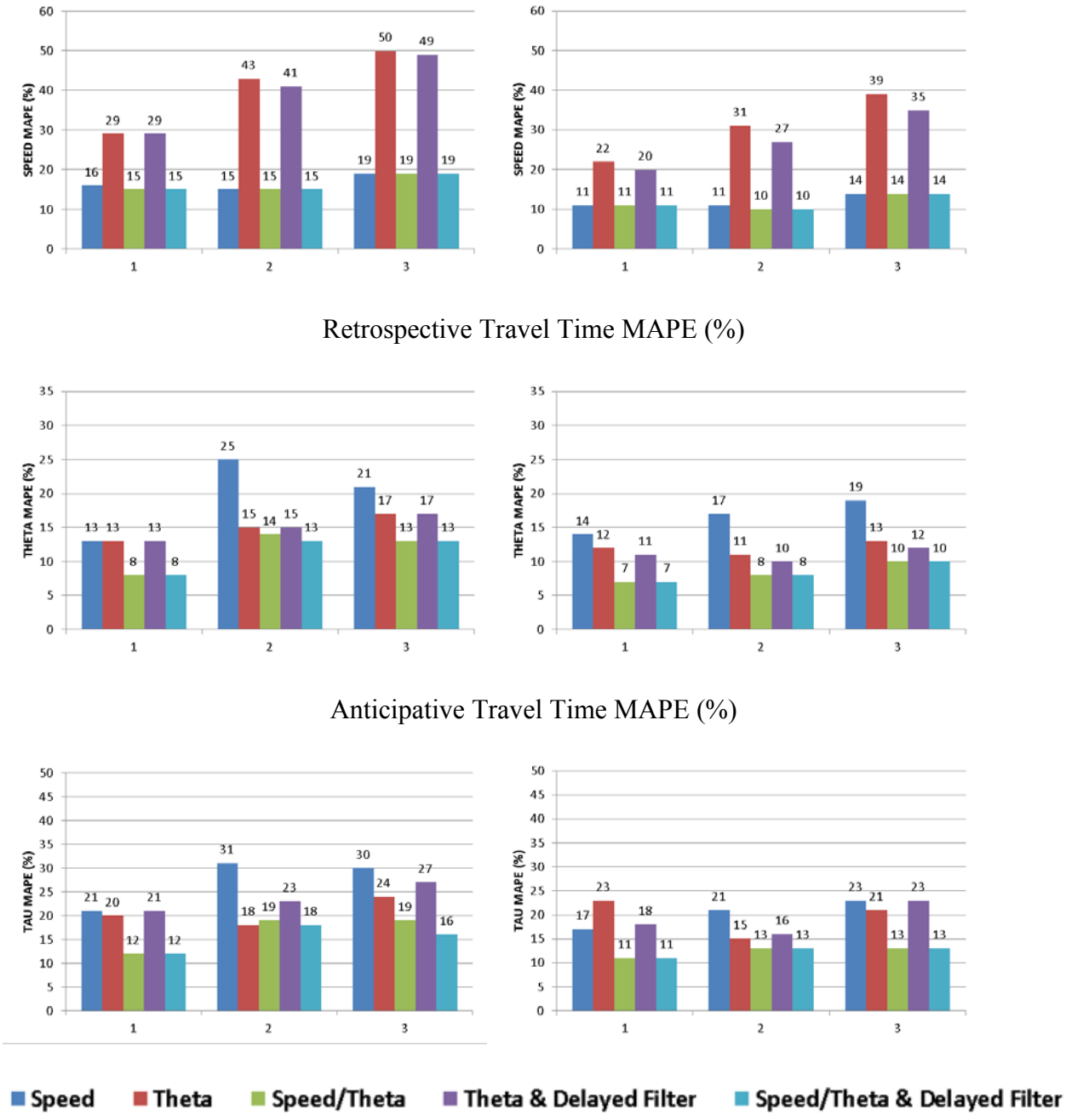


Figure 26. MAPE of estimates using Smulders relation and UKF method. (Left: 210ftx2sec; Right: 420ftx4sec)

Table 10. Overall MAPE estimates using UKF method in US 101 datasets.

Delayed Filter?	Input Data	Greenshields			Smulders		
		Speed	Theta	Tau	Speed	Theta	Tau
No	Speed	13/11 17/13 22/17	16/13 31/21 30/32	21/19 40/24 43/32	16/11 15/11 19/14	13/14 25/17 21/19	21/17 31/21 30/23
	Theta	28/22 35/27 44/33	13/12 15/10 17/13	22/23 19/16 23/20	29/22 43/31 50/39	13/12 15/11 17/13	20/23 18/15 24/21
	Both	12/10 17/13 22/18	9/6 14/9 14/11	11/9 24/14 27/15	15/11 15/10 19/14	8/7 14/8 13/10	12/11 19/13 19/13
Yes	Speed	/ / /	/ / /	/ / /	/ / /	/ / /	/ / /
	Theta	28/20 35/25 43/32	13/11 14/10 16/12	23/19 24/18 27/24	29/20 41/27 49/35	13/11 15/10 17/12	21/18 23/16 27/23
	Both	12/10 17/13 22/18	9/6 14/9 14/11	11/10 18/13 16/13	15/11 15/10 19/14	8/7 13/8 13/10	12/11 18/13 16/13

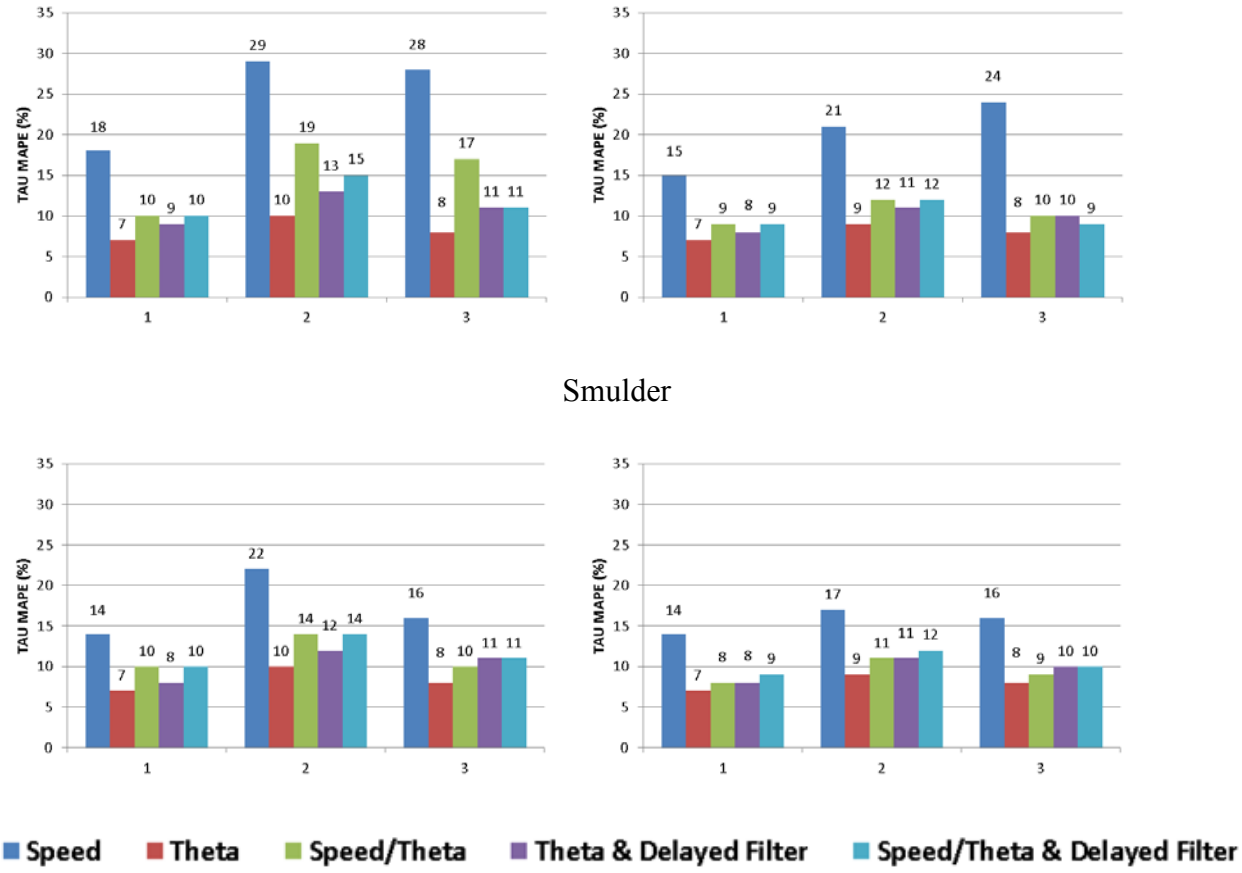
(210x2)/(420x4)

Table 10 summarizes the same overall MAPE measures as shown in Figure 25 and Figure 26 in a tabular format.

6.6 Prediction Results

Figure 27 shows the MAPE of predictive travel time estimates at the upstream of the segment under consideration using UKF estimation method. Upstream anticipative travel times are best estimated when only retrospective travel times are used as input measurements in the estimation. Invariably, using delayed filter in this case has slightly deteriorated results. However, when speed and retrospective travel times are both used as input measurements for the most part results have been insensitive to the use of delayed filter. Note that upstream predictive travel time estimates are significantly better than their corresponding overall estimates. This underlines the fact that estimates of travel time in internal cells (other than boundaries) exhibit worse than average errors.

Greenshields



Smulder

Speed Theta Speed/Theta Theta & Delayed Filter Speed/Theta & Delayed Filter

Figure 27. MAPE of upstream predictive travel time estimates using UKF method. (Left: 210ftx2sec; Right: 420ftx4sec)

Table 11 summarized the same upstream MAPE measures as shown in Figure 27 in a tabular format.

Table 11. Upstream anticipative travel time MAPE estimates using UKF method.

Delayed Filter?	Input Data	Greenshields			Smulders		
		Speed	Theta	Tau	Speed	Theta	Tau
No	Speed	/ / /	/ / /	18/15 29/21 28/24	/ / /	/ / /	14/14 22/17 16/16
	Theta	/ / /	/ / /	7/7 10/9 8/8	/ / /	/ / /	7/7 10/9 8/8
	Both	/ / /	/ / /	10/9 19/12 17/10	/ / /	/ / /	10/8 14/11 10/9
Yes	Speed						
	Theta	/ / /	/ / /	9/8 13/11 11/10	/ / /	/ / /	8/8 12/11 11/10
	Both	/ / /	/ / /	10/9 15/12 11/9	/ / /	/ / /	10/9 14/12 11/10

(210x2)/(420x4)

Figure 28 depicts the prediction speed error quartiles when prediction horizon has varied between zero (current time) and 60 seconds (one minute) ahead of the current time. The top, middle, and bottom row of graphs represent datasets with increasing congestion levels. Note that when speed and retrospective travel time measurement along with a delayed filter are used in estimation, resulting predictions errors exhibit highest level of stability. As can be seen in this case the prediction bias (median line) and the size of interquartile range is gradually but steadily increasing with congestion level. As prediction horizon increases, the errors become more skewed toward overestimation.

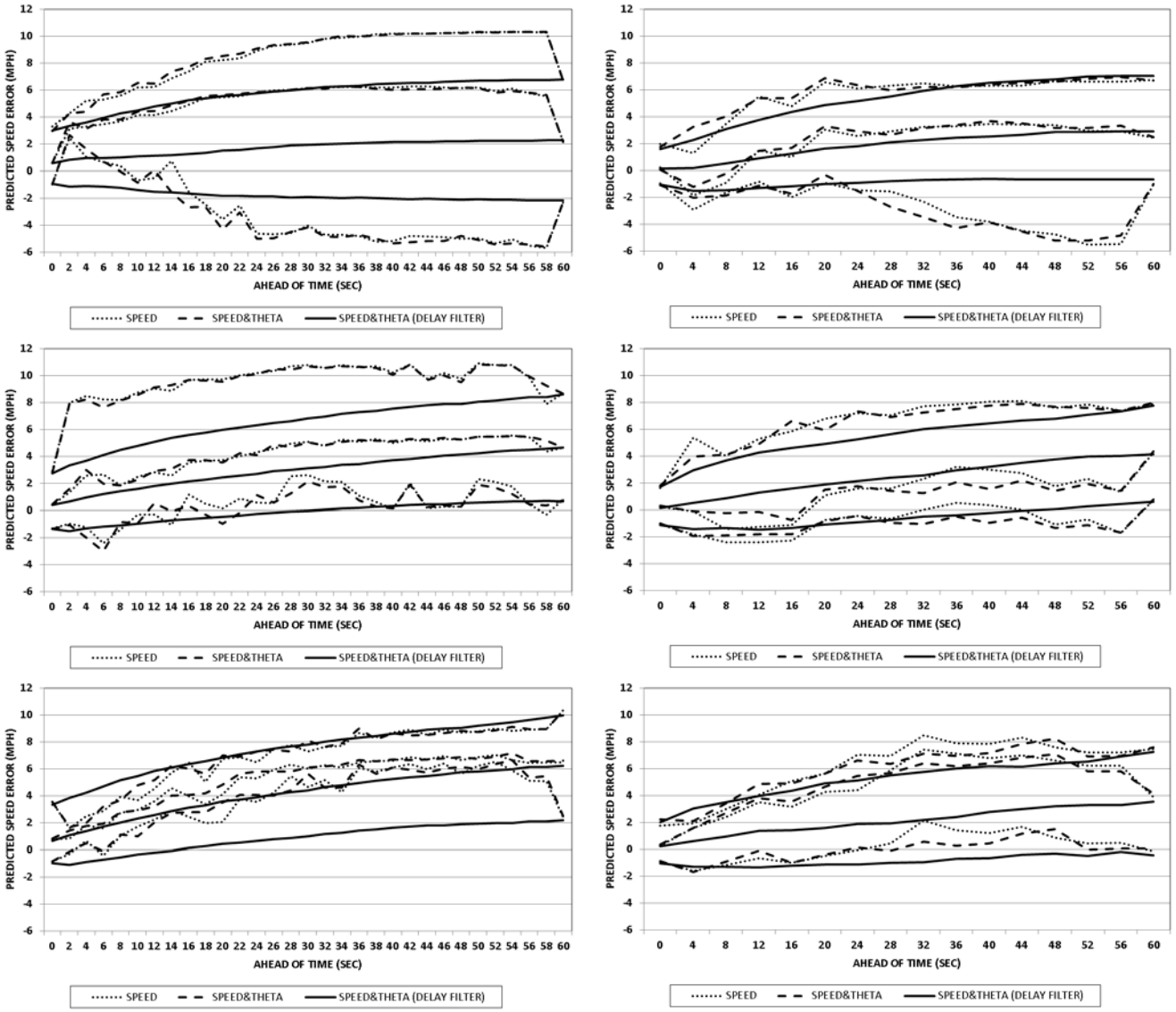


Figure 28. Prediction speed error quartiles. (top: D1, middle: D2, bottom: D3; left: 210x2, right 420x4)

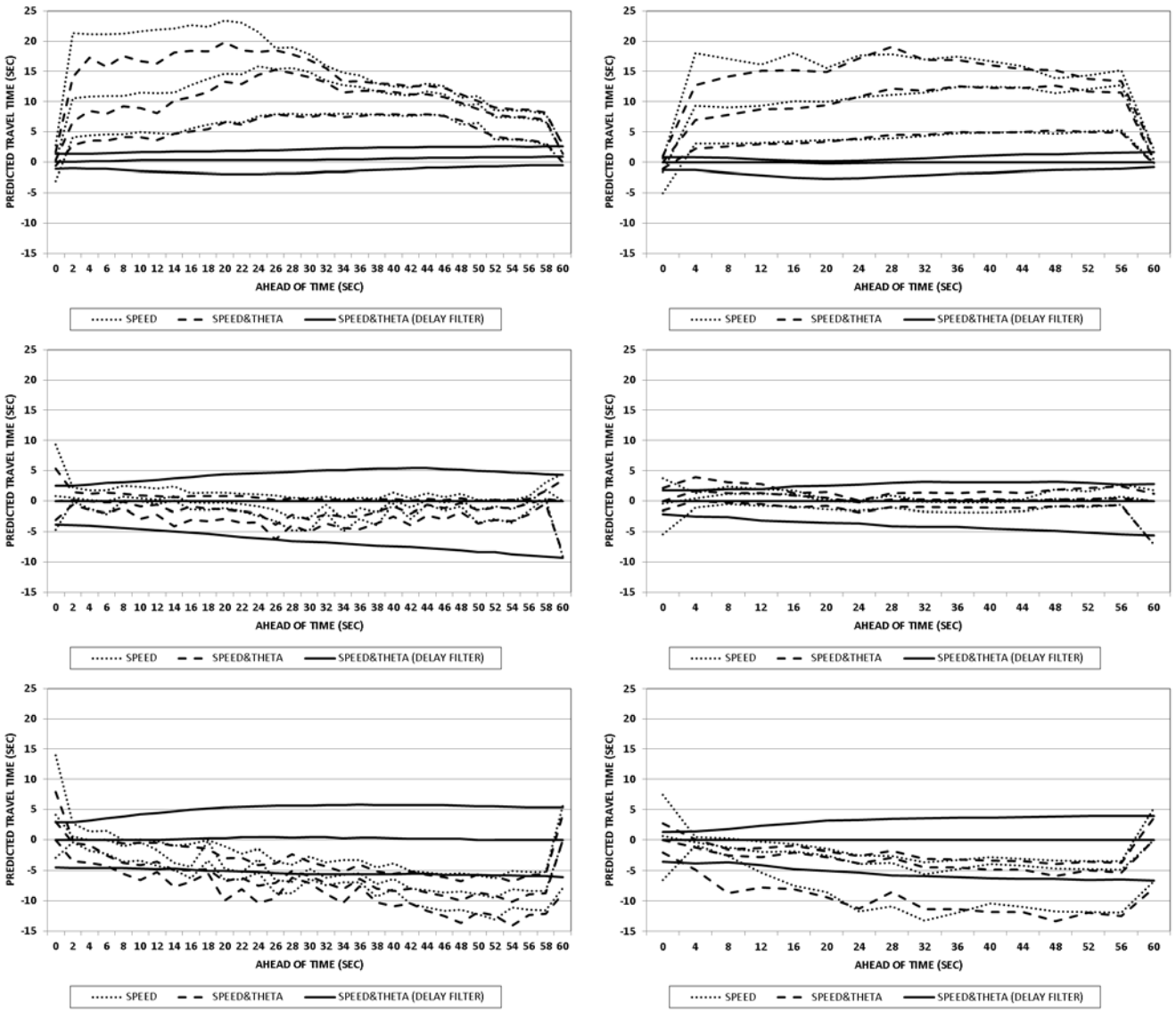


Figure 29. Prediction travel time error quartiles. (top: D1, middle: D2, bottom: D3; left: 210x2, right 420x4)

Figure 29 exhibits the prediction travel time error quartiles when prediction horizon extends up to one minute ahead of the current time. Note that in case of speed and travel time measurement input along with a delayed filter, prediction errors invariably have remained unbiased. Additionally, the interquartile range in these cases has quickly stabilized even if it initially started to grow with the prediction horizon.

6.7 Computation Time

The computation times reported in this section are obtained on a machine with Intel Core i7-860 processor (8M Cache, Quad-Core 2.80 GHz) and 4.00 GB RAM running 64 bit Microsoft Windows 7 Enterprise operating system. The algorithm is implemented and run in MATLAB R2012b.

Table 12. Average computation time for one time step estimation in US 101 dataset. (Speed and travel time inputs)

Delayed Filter Applied?			**No**		**Yes**	
Speed-Density Relation			**GS**	**HL**	**GS**	**HL**
Time Interval (sec)	**2**	**D1**	0.106	0.106	1.826	1.807
		D2	0.107	0.107	2.257	2.244
		D3	0.105	0.104	2.526	2.487
		Mean	**0.106**	**0.106**	**2.203**	**2.179**
	4	**D1**	0.028	0.028	0.242	0.229
		D2	0.028	0.028	0.305	0.290
		D3	0.027	0.028	0.324	0.301
		Mean	**0.028**	**0.028**	**0.290**	**0.273**

Doubling the update rate implies that the number of discretization cells in a given segment is practically reduced in half to meet the CFL condition. In the case of proposed model in this dissertation this leads to a 50 percent reduction in the size of state vector to be estimated. For instance, in the US 101 implementation scenarios when time update rate increases from two to four seconds, the state vector size is reduced from thirty to fifteen.

Table 12 summarizes the average computation times experienced for one time step estimation of state variables under different scenarios in US 101 dataset when speed and travel time measurements have been used as input. In all cases where delayed filter has been off, the average computation times at each step has been fraction of a second. Results indicate that in the current setting at very fine two second time update rate when delayed filter is used the proposed UKF estimation method may not be applied in real time. This is due to the fact that under this scenario the average time it takes to update the state vector estimates has been larger than two seconds. However, in all other cases the proposed UKF estimations can be performed in real time. Congestion level virtually has no effect on the computation times when delayed filter is not used, but in presence of delayed filters increasing congestion seems to slightly increase the computation times.

Also, note that at each time step the number of delayed filter updates on average is equal to the number of time steps that make for the average travel time on the segment. For instance, in two second time steps the forty second average travel time between the two ends of the segment can be covered in twenty time steps. This means to update the state vector in presence of a delayed filter at each time step roughly twenty simple updates should take place. Similarly, in the case of four second time steps, using a delayed filter on average is equivalent to ten single updates.

Interestingly, these ballpark estimates closely correlate with the ratio of average computation times spent in each pair of corresponding cases where delayed filters are turned on and off.

6.8 Unscented H-infinity Filter Results

In all scenarios tested the proposed UHF estimation method resulted in very similar error measures as the UKF estimation method. Although much care was spent to adjust the γ^2 parameter in order to obtain the desired constrained worst-case performance of UHF, it turned out that in the range of feasible γ^2 parameters the performance is not particularly different from UKF. This underlines the fact that the joint traffic and travel time model proposed in this dissertation is highly nonlinear. Note that this may have been exacerbated by very fine discretization levels adopted in numerical experiments reported in this dissertation.

6.9 Summary

In this chapter results of the numerical experiments performed with the proposed joint traffic and travel time model and estimation algorithms on NGSIM US 101 datasets were reported. Different scenarios in terms of estimation algorithm, speed-density relationships, discretization sizes, and the use of delayed filter were taken into account.

Numerical results indicated that in general estimation errors become larger with increasing congestion, while they become smaller with increasing discretization cell sizes and aggregation time intervals. Also, the best speed estimates were obtained when boundary speed measurements have been part of the input. Overall, speed estimates have been insensitive to the application of the delayed filter in presence of retrospective travel time measurements. Both speeds and retrospective travel times were needed to make better estimates of travel times (both retrospective and anticipative). Delayed filter has not been very effective in improving the estimates of retrospective travel times, but when speed and retrospective travel time

measurements were assimilated in presence of a delayed filter, anticipative travel times during congested periods have improved. Finally, using Smulders speed-density relation generally did improve the estimation results especially in the case of speed estimates.

Interestingly, upstream anticipative travel times were best estimated when only retrospective travel times were used as input measurements in the estimation. Using the delayed filter in this case did not improve the results. A combination of speed and retrospective travel time measurements along with the application of a delayed filter seemed to be most effective in containing the speed and travel time prediction error bias.

Chapter 7: Conclusions and Future Directions for Research

This chapter presents a concise summary of contribution made, conclusions reached, and lessons learned in this dissertation. Also future directions of research with the goal of improving upon the proposed models and methods in this dissertation are outlined. Different possibilities to apply the proposed travel time assimilation concepts to other real world cases are enumerated.

7.1 Contributions

The following is a brief outline of the main contributions of this dissertation:

- **Derivation of a first-order PDE relating speeds and travel times independent of vehicle trajectory.** In Chapter 4, a travel time model was derived. This is a first-order PDE which relates the local speeds and local variations in travel time. The proposed model can be applied to both retrospective and anticipative travel times with minimal change. Finite-difference approximate solutions for the travel time models were proposed. Stability and FIFO properties of the proposed travel time model were discussed.
- **Derivation of finite-difference solutions of coupled first-order velocity-based traffic continuum models and travel time equation.** In Chapter 5, the joint traffic and travel time dynamics model was specified. The traffic dynamics was modeled using first-order velocity-based cell transmission model (CTM-v) introduced in Chapter 3. Retrospective and anticipative travel time dynamics were modeled using first-order finite-difference travel time models proposed in Chapter 4 (named THETA and TAU, respectively).
- **Introduction of a framework to systematically and explicitly assimilate travel time measurements in traffic estimation process.** In Chapter 5, the proposed joint traffic and travel time model was cast as a state-space model whose state vector incorporate traffic speeds, retrospective, and anticipative travel times sub-vectors. The joint system model was shown to be highly nonlinear. The proposed framework takes into account additive model and measurement

errors. Also, the measurement model is capable of accepting boundary speed and travel time measurements.

- **Extension of an optimal state estimation method based on min-max paradigm to the case of highly non-linear joint traffic and travel time model.** In Chapter 5, the UHF estimation method for non-linear state-space models was presented. H-infinity type filters seek to minimize the maximum estimation errors in face of adversarial model and measurement errors. Unscented methods were adopted to nonlinearly propagate state estimate mean and covariance. Also, unscented methods were adopted in the conditioning step to adjust the a priori state estimates with information contained in the measurements. H-infinity filtering was proposed as an alternative to Kalman filtering in the conditioning step (measurement update).
- **Adoption of a delayed filter to explicitly take into account the delayed anticipative travel time measurements.** In Chapter 5, a delayed filter was proposed to assimilate any delayed anticipative travel time measurements. This is a brute-force approach but computational results showed that it works very well in all practical scenarios. In general, the proposed delay filter can be used to assimilate any delayed or out-of-order traffic measurements.
- **Sensitivity analysis of speed and travel time estimates/predictions to the presence of various traffic measurements, congestion levels, aggregation levels, speed-density relations, delayed filter, and optimal estimation methods.** In Chapter 6, results of numerical experiments conducted on various scenarios were reported. Scenarios were comprised of combinations of the following ingredients: three datasets with increasing level of congestion, availability of speed and travel time measurements to be used in the assimilations, two discretization schemes, two speed-density relations, presence or absence of the proposed delayed filter when travel time measurements were involved, and two optimal estimation methods applicable to nonlinear state-space models.
- **Improvement of traffic state estimates and short-term travel time predictions.** In Chapter 6,

results indicated that incorporation of field measured travel times into the proposed joint traffic and travel time model estimation reduced retrospective travel time estimates' maximum MAE from 12 seconds to 3 seconds at 4 second time intervals. Also, under the same conditions anticipative travel time estimates' maximum MAE was reduced from 10 seconds to 4 seconds. Note that these maximum MAE measures (3 and 4 seconds) were smaller than the discretization time used in their corresponding numerical experiments (4 seconds). The MAPE of upstream anticipative travel time (current time predictions) was reduced from 24% to 10% under congested conditions when retrospective travel time measurements along with speeds were used in the assimilations.

7.2 Conclusions

Numerical experiment results indicated that the proposed joint traffic and travel time model and estimation algorithms have been successfully used to seamlessly assimilate different field measured traffic data and in particular boundary travel time measurements into the traffic state estimation process. Speeds and travel times are most accurately estimated when speed and travel time measurements at the boundaries have been part of the input to the state estimation process, respectively. Increasing the discretization size (reduced granularity) resulted in more accurate estimates.

In general, considering the estimate biases speeds are overestimated while travel times are underestimated. The literature search showed that in comparable scenarios state-of-the-art using boundary along with five percent probe speed measurements, overall speed estimates MAPE has been 25% (Work, et al. 2008). Using boundary speed and travel time measurements resulted in speed and anticipative travel time estimate MAPEs varying in the 12%-22% and 11%-27% range, respectively. Using the proposed delayed filter in this case brought about an additional improvement in the anticipative travel time estimates where MAPEs vary in the 11%-18% range.

Smulders speed-density relation resulted in better estimates compared to the Greenshields relation. This was more evident in the case of speed estimates.

When anticipative travel times solely at the upstream of the segment are targeted, using current retrospective travel time measurements at the downstream of the segment in assimilation provide the best estimates. MAPE reported for the upstream anticipative travel times in this case varies in the 7%-10% range.

A combination of speed and retrospective travel time measurements along with the application of a delayed filter are most effective in containing the bias in multi-step ahead speed and travel time predictions. These scenarios resulted in unbiased multi-step ahead travel time predictions. At 2 second discretization, the interquartile range of one-minute-ahead speed and anticipative travel time prediction errors varied between -2mph to +10mph, and -10sec to +5sec, respectively. Increasing the discretization size to 4 seconds, the interquartile error range of these predictions was reduced to -1mph to +8mph, and -7sec to +4sec, respectively. Table 13 provides a summary of the results obtained in this dissertation in the same way as was presented earlier in Table 1 at the end of literature review chapter.

Table 13. Summary results of the proposed joint state-space model and estimation algorithms.

Traffic Model	Data	Estimation Method	Data Source	Facility Type	Time Interval	Estimation Variable	Prediction Variable	Accuracy
CTM-v	-Stationary (speed) -AVI (travel time)	-UKF -UHF	NGSIM	Freeway (US-101 CA)	2 sec 4 sec	-Speed -Travel Time	-Speed -Travel Time	MARE 12%-22% MARE 7%-10%

The proposed model and estimation algorithms can be extended to take into account any gradual changes in roadway and traffic conditions. For instance, if free flow speeds or minimum safe driving headways are impacted by changes in weather or lighting conditions, the corresponding parameters can be adaptively estimated in the proposed model with minor changes. However,

abrupt changes in underlying geometry such as lane closure due to incidents or other unexpected events with long-term impact on traffic need to be explicitly modeled. Having said that the proposed model can be used to identify cases where realities on the ground are substantially different from what is modeled. These discrepancies in a calibrated model could be indicative of events that have changed the underlying assumptions of the model. Obviously, in such cases all assumptions and parameters should be revisited and a new modeling and calibration process will be necessary.

This dissertation presents a modeling and estimation framework that can be used readily to integrate private and public sector data and to improve travel time estimations and predictions. Currently, travel time estimation and navigation products based on crowd-sourced AVL (probe) data are marketed by private sector. While probe data represents speed of traffic, its accuracy and coverage heavily depends on market penetration of probes. On the other hand, sparsely installed spot traffic detectors are mainly owned and run by public agencies which usually lack funding to routinely maintain and calibrate them. Spot detectors measure average local traffic conditions as it applies to all vehicles passing that location. Recently, private and public sector are beginning to appreciate the value in sharing their respective data. Some companies such as Waze (acquired by Google) are in the process of bridging that gap by offering public agencies access to crowd-sourced and anonymized probe data in return for access to public sector's road sensor data, as well as pre-planned construction and road closures.

Models and algorithms presented in this dissertation can be adopted by both private and public sector with minimum modifications. In the current practice, private sector's archival crowd-sourced traffic data is used to generate and update travel speed profiles for individual road segments and different day/time combinations. These profiles are used to validate current travel

speed measurements on different segments and to identify whether or not they resemble a specific existing pattern (past weather or incident events) in the historic data. Also, for the most part, any short- or long-term travel time predictions are based on these profiles. While it may not be practical to use continuum traffic flow models to represent traffic dynamics in a large network, historic speed profiles can be used to generate ad hoc models of traffic speed dynamics. These traffic models along with travel time models presented in this dissertation can form an ad hoc joint traffic and travel time model. Filtering techniques presented in this dissertation can be adopted to jointly estimate and predict traffic speeds and travel times on individual segments as well as on paths comprised of multiple segments.

7.3 Future Directions for Research

This line of research can be continued in at least five different directions. First, dynamic traffic and travel time models used can be improved. Second, proposed models and estimation techniques can be easily adjusted to take into account internal traffic measurements in addition to boundary measurements addressed in this dissertation. Third, other real world traffic datasets may be used to further verify the proposed model and estimation methods' performances in traffic estimation and short-term predictions. Fourth, for practical reasons, it is desirable for the proposed model and estimation methods to become capable of handling irregular shaped space-time discretization schemes. Fifth, different avenues for the use of proposed modeling and estimation framework in typical traffic control application can be explored. Each of these directions is further discussed in the following sections.

7.3.1 Model Improvements

7.3.1.1 Second-Order Traffic Model

The traffic model used in this dissertation is a first-order conservation law. The CTM-v model presented in Chapter 3 simply ensures vehicles are conserved over space and time. Besides,

under CTM-v model, speed-density relationship holds at all times. In other words, speeds are always at the equilibrium level with respect to the existing density. This practically results in abrupt and unrealistic changes in density and speed across shockwaves.

Second-order macroscopic traffic models improve upon the first-order model by taking into account the distinction between equilibrium and non-equilibrium speeds. In fact, the second-order traffic model while keeping the CTM-v model to describe equilibrium speed evolutions includes a second model to represent effects of relaxation, convection, and anticipation on non-equilibrium traffic speeds. The pair of models to describe equilibrium and non-equilibrium speed dynamics can be expressed as following:

$$v_{e,n+1}^{i} = V\left(\rho_{n+1}^{i}\right) = V\left[V^{-1}\left(v_n^i\right) - \frac{\Delta t}{\Delta x}\left[\tilde{G}\left(v_n^i, v_n^{i+1}\right) - \tilde{G}\left(v_n^{i-1}, v_n^i\right)\right]\right] \quad (7.1)$$

$$v_{n+1}^{i} = v_n^i + \frac{\Delta t}{\vartheta}\left[v_{e,n}^i - v_n^i\right] + \frac{\Delta t}{\Delta x} v_n^i\left[v_n^{i-1} - v_n^i\right] - \frac{\psi \Delta t}{\vartheta \Delta x}.\left[V^{-1}\left(v_n^{i+1}\right) - V^{-1}\left(v_n^i\right)\right] / \left[V^{-1}\left(v_n^i\right) + \kappa\right] \quad (7.2)$$

where, v_n^i and $v_{e,n}^i$ denotes the non-equilibrium and equilibrium traffic speeds at cell i during time interval n, respectively. Parameters ϑ, ψ, κ, should be defined appropriately. In equation (7.2) the second term represents relaxation between equilibrium and non-equilibrium speed states at each cell, the third term represents the impact of convection from the upstream cell, and finally the last term represents the impact of anticipation of downstream congestions on traffic speeds.

7.3.1.2 Second-Order Travel Time Model

Travel time model proposed in this dissertation is based on finite-difference approximation of first-order travel time partial derivatives with respect to space and time. To improve the proposed model accuracy, it is conceivable to use higher-accuracy finite-difference approximations. For instance, as was mentioned in chapter 4, travel time models (4.6) and (4.7) can be solved numerically using a forward-time backward-space (FTBS) finite difference scheme. Following

the same discretization scheme, travel time partials with respect to time and space can be approximated by the following:

$$\tau_t \cong \left(-\tau_{n+2}^i + 4\tau_{n+1}^i - 3\tau_n^i\right)/(2\Delta t) \tag{7.3}$$

$$\tau_x \cong \left(3\tau_n^i - 4\tau_n^{i-1} + \tau_n^{i-2}\right)/(2\Delta x) \tag{7.4}$$

Substituting these expressions in (4.6) will provide second-order accuracy to the proposed finite-difference anticipative travel time model. Similarly, the following pair of approximate partial derivatives with respect to time and space will provide second-order accuracy to the proposed finite-difference retrospective travel time model.

$$\theta_t \cong \left(-\theta_{n+2}^i + 4\theta_{n+1}^i - 3\theta_n^i\right)/(2\Delta t) \tag{7.5}$$

$$\theta_x \cong \left(3\theta_n^i - 4\theta_n^{i-1} + \theta_n^{i-2}\right)/(2\Delta x) \tag{7.6}$$

7.3.2 Internal Speed and Travel Time Measurements

In this dissertation, only speed and travel time measurements at the boundaries of the segment are used. It is conceivable that in presence of probes, relevant measurements from inside the segment can be obtained. Such internal measurements can be easily integrated into the proposed estimation and prediction process. The additional measurements potentially will increase the accuracy of state estimations and ensuing predictions.

7.3.3 Other Datasets (NGSIM & Mobile Century)

In this dissertation, applications of proposed estimation methodology on US 101 datasets are reported. It is possible to apply the proposed models and methods on other available datasets to examine their performance under different facility types, geometry and driver compositions. Under NGSIM project, detailed traffic data on three other facilities in California and Georgia are collected. I-80 dataset in San Francisco provides another rich opportunity to test the proposed methods on freeway type facilities. Also, two arterial datasets collected on Lankershim

Boulevard in Los Angeles, and Peachtree Street in Atlanta are made available through NGSIM project.

In addition, cell phone GPS data collected from 100 probe vehicles driving 6-10 mile loops continuously for 8 hours on freeway I-880 near Union City in the San Francisco Bay Area has become available (Herrera, Work, et al. 2010). The dataset named Mobile Century in addition to GPS logs of 77 mobile devices (at 3 second frequency) provides inductive loop detector data installed in the area, ground-truth travel time data on the northbound direction between Stevenson Blvd and Decoto Road and between Decoto Road and Winton Avenue.

7.3.4 Irregular Space/Time Discretization (Application to VPP Data)

Since 2008, private sector probe based space mean speed (SMS) data has become available to states along the I-95 Corridor on the eastern coast of the United States. Vehicle Probe Project (VPP) is an ongoing effort and has resulted in a comprehensive archive of speed data at one minute resolution on standard segments along the highways and arterials.

The segment definitions are based on industry developed Traffic Message Channel (TMC) standards overseen by Traveller Information Services Association (TISA). The TMC segments on freeway facilities are typically defined between consecutive on- and off-ramps. As a result, TMC segments are not uniformly sized.

Concurrent with VPP, automatic vehicle identification (AVI) travel time measurements using Bluetooth monitoring systems has been ongoing mainly for validation purposes. These efforts have resulted in a large collection of travel time datasets on some of the most congested highways in major urban areas in the continental United States.

This vast archival speed and travel time dataset can be used to further evaluate the proposed models and methods in this dissertation. However, in order to do that, it is necessary that methods be adjusted to take into account the non-uniform length of TMC segments and unequal

data resolutions considering the fact that speeds are reported every one minute while travel times are updated roughly every five seconds. Once, accuracy of the proposed methods in these conditions is established, it is possible to use them in real-time travel time prediction applications along select corridors with active user information services such as variable message signs (VMS) and other advisory systems.

7.3.5 Control Applications

In this dissertation, proposed models and estimation methods were primarily used in the estimation mode. This means speed and travel time measurements were fed to the model in order to make real-time estimates of state variables and to make better short-term predictions. This approach essentially treats the highway system as an open system. However, the proposed models can also be used to model important state variables when some control measures are present. Essentially, the proposed model and estimation method can be used to track and to evaluate the relevant performance measures (goals) of the closed traffic systems that are to be controlled.

For instance, in a freeway system, the proposed joint models and estimation method can be used to make use of travel time measurements in ramp metering (RM) and variable speed limit (VSL) control applications to prevent the congestion from developing in the susceptible regions of freeway traffic lanes.

In arterial systems, the major control devices used are traffic signals. The proposed models and estimation method can be used to integrate travel time measurements into adaptive signal re-timing decision makings locally, on a corridor level, and or on a regional basis.

7.4 Summary

In this chapter main contributions in this dissertation were briefly outlined. Conclusions reached based on experimental results were reiterated. Finally, five possible directions to continue this

line of research and to add to the traffic estimation and short-term prediction state-of-the-art were identified.

Chapter 8:Bibliography

Ahmed, M. S., and A. R. Cook. "Analysis of Freeway Traffic Time Series Data by using Box-Jenkins Techniques." *Transportation Research Record* (Journal of Transportatin Research Board of the National Research Council) 722 (1979): 1-9.

Ahmed, S. A., and A. R. Cook. "Application of Time Series Analysis Techniques to Freeway Incident Detection." *Transportation Research Record* (Journal of Transportation Research Board of the National Research Council) 841 (1982): 19-21.

Astraita, A. "A Continuous Time Link Model for Dynamic Network Loading Based on Travel Time Functions." *Proceedings of the 13th International Symposium of Transportation and Traffic Theory.* Oxford, UK: Elsevier, 1996. 79-102.

Bajwa, S. I., E. Chung, and M. Kuwahara. "Performance Evaluation of an Adaptive Travel Time Prediction Model." *Proceedings of the 8th International IEEE Conference on Intelligent Transportation Systems.* Vienna, Austria, 2005.

Barcelo, J., L. Montero, L. Marques, and C. Carmona. "Travel Time Forecasting and Dynamic Origin Destinaiton Estimation for Freeways Based on Bluetooth Traffic Monitoring." *Transportation Research Record: Journal of the Transportation Research Board* (Transportation Research Board of the National Academies), no. 2175 (2010): 19-27.

Bardos, C., A. Y. Leroux, and J. C. Nedelec. "First Order Quasilinear Equations with Boundary Conditions." *Communications in Partial Differential Equations* 4, no. 9 (1979): 1017-1034.

Beskos, D. E., and P. G. Michalopoulos. "An Application of the Finite Element Method in Traffic Signal Analysis." *Mechanics Research Communications* 11, no. 3 (1984): 185-189.

Beskos, D. E., P. G. Michalopoulos, and J. K. Lin. "Analysis of Traffic Flow by the Finite Element Method." *Applied Mathematical Modeling* 9 (October 1985): 358-364.

Bickel, P. J., C. Chen, J. Kwon, J. Rice, J. Van Zwet, and P. Varaiya. "Measuring Traffic." *Statistical Science* 22, no. 4 (2007): 581-597.

Box, G. E.P., G. M. Jenkins, and G. C. Reinsel. *Time Series Analysis: Forecasting and Control.* 4th. Hoboken, New Jersey: John Wiley, 2008.

Cambridge Systematics Inc. "NGSIM U.S. 101 Data Analysis: Summary Report." Federal Highway Administration, Washington, D.C., 2005.

Cambridge Systematics, Texas Transportation Institute. *Traffic Congestion and Reliability: Linking Solutions to Problems.* Federal Highway Administration, U.S. Department of Transportation, Washington, D.C.: Office of Operations, 2004.

Carey, M. "Link Travel Times I: Desirable Properties." *Networks and Spatial Economics* 4, no. 1 (2004): 257-268.

Carey, M., and Y. E. Ge. "Retaining Desirable Properties in Discretising a Travel Time Model." *Transportation Research Part B* 41, no. 1 (2007): 540-553.

Carey, M., Y. E. Ge, and M. Mc Cartney. "A Whole Link Travel Time Model with Desirable Properties." *Transportation Science* 37, no. 1 (February 2003): 83-96.

Carmi, A., P. Gurfil, and D. Kanevsky. "Methods for Sparse Signal Recovery Using Kalman Filtering with Embedded Pseudo-Measurement Norms and Quasi-Norms." *IEEE Transactions on Signal Processing* 58, no. 4 (April 2010): 2405-2409.

Cassidy, M. J., and J. R. Windover. "Methodology for Assessing Dynamics of Freeway Traffic Flow." *Transportation Research Record* (Transportation Research Board of the National Research Council) 1484 (1995): 73-79.

Chakroborty, P., and S. Kikuchi. "Using Bus Travel Time Data to Estimate Travel Time on Urban Corridors." *Transportation Research Record: Journal of the Trasnportation Research Board* (Transportation Research Board of the National Research Council), no. 1870 (2004): 18-25.

Chen, Hao, Hesham A. Rakha, and Shereef Sadek. "Real-Time Freeway Traffic State Prediction: A Particle Filter Approach." *14th International Conference on Intelligent Transportation Systems.* Washington, D.C., USA: IEEE, 2011. 626-631.

Chen, M., and S. I.J. Chien. "Dynamic Freeway Travel Time Prediction with Probe Vehicle Data." *Transportation Research Record* (Journal of the Transportation Research Board of the National Research Council) 1768 (2001): 157-161.

Chu, L., J.-S. Oh, and W. Recker. "Adaptive Kalman Filter Based Freeway Travel Time Estimation." *Presented at Transportation Research Board Annual Meeting.* Washington, D.C.: Transportation Research Board of the National Academies, 2005.

Clark, S. "Traffic Prediction Using Multivariate Nonparametric Regression." *ASCE Journal of Transportation Engineering* 129, no. 2 (March-April 2003): 161-168.

Claudel, C. G., A. Hofleitner, N. D. Mignerey, and A. M. Bayen. "Guaranteed Bounds on Highway Travel Times Using Probe and Fixed Data." *Presented in Transportation Research Board Annual Meeting.* Washington, D.C.: Transportation Research Board of the National Academies, 2009.

Claudel, C. G., and A. M. Bayen. "Guaranteed Bounds for Traffic Flow Parameters Estimation Using Mixed Lagrangian-Eulerian Sensing." *46th Annual Allerton Conference on Communication, Control, and Computing.* Allerton, Illinois, 2008.

Claudel, C. G., and A. M. Bayen. "Lax-Hopf Based Incorporation of Internal Boundary Conditions into Hamilton-Jacobi Equation. Part I: Theory." *IEEE Transactions on Automatic Control* 55, no. 5 (May 2010): 1142-1157.

Claudel, C. G., and A. M. Bayen. "Lax-Hopf Based Incorporation of Internal Boundary Conditions into Hamilton-Jacobi Equation. Part II: Computational Methods." *IEEE Transactions on Automatic Control* 55, no. 5 (May 2010): 1158-1174.

Coifman, B. "Estimating Travel Times and Vehicle Trajectories on Freeways using Dual Loop Detectors." *Transportation Research Part A* 36, no. 4 (2002): 351-364.

Coifman, B. "Identifying the Onset of congestion Rapidly with Existing Traffic Detectors." *Transportation Research Part A* 37 (2003): 277-291.

Coifman, B., and E. Ergueta. "Improved Vehicle Reidentification and Travel Time Measurement on Congested Freeways." *ASCE Journal of Transportation Engineering* 129 (2003): 475-483.

Daganzo, C. "A Variational Formulation of Kinematic Waves Basic Theory and Complex Boundary Conditions." *Transportation Research Part B* 39 (2005): 187-196.

Daganzo, C. "A Variational Formulation of Kinematic Waves: Solution Methods." *Transportation Research Part B* 39 (2005): 934-950.

Daganzo, C. F. *Fundamentals of Transportation and Traffic Operations.* Oxford: Pergamon, 1997.

Daganzo, C. F. "The Cell Transmission Model: A Dynamic Representation of Highway Traffic Consistent with the Hydrodynamic Theory." *Transportation Research Part B* 28, no. 4 (1994): 269-287.

Dailey, D. J. "A Statistical Algorithm for Estimating Speed from Single Loop Volume and Occupancy Measurements." *Transportation Research Part B* 33, no. 5 (1999): 313-322.

Dailey, D. J. "Travel Time Estimation Using Cross Correlation Techniques." *Transportation Research Part B* 27B, no. 2 (1993): 97-107.

D'Angelo, M. P., H. M. Al-Deek, and M. C. Wang. "Travel Time Prediction for Freeway Corridors." *Transportation Research Record* (Journal of Transportation Research Board of the National Research Council) 1676 (1999): 184-191.

Del Castillo, J. M., and F. G. Benitez. "On the Functional Form of the Speed-Density Relationship-I: General Theory." *Transportation Research Part B* 29B, no. 5 (1995): 373-389.

Del Castillo, J. M., and F. G. Benitez. "On the Functional Form of the Speed-Density Relationship-II: Empirical Investigation." *Transportation Research Part B* 29B, no. 5 (1995): 391-406.

Dion, F., and H. Rakha. "Estimating Dynamic Roadway Travel Times Using Automatic Vehicle Identification Data for Low Sampling Rates." *Transportation Research Part B* 40 (2006): 745-766.

Fausett, L. V. *Fundamentals of Neural Networks: Architecture, Algorithms, and Applications.* Englewood Cliffs, New Jersey: Prentice-Hall, 1994.

Federal Highway Administration, U.S. Department of Transportation. *Next Generation SIMulation Fact Sheet.* December 2006. http://www.fhwa.dot.gov/publications/research/operations/its/06135/index.cfm (accessed May 10, 2011).

Greenshields, B. "A Study of Traffic Capacity." *Highway Research Board*, 1935: 448-477.

Haghani, A., M. Hamedi, and K. F. Sadabadi. *I-95 Corridor Coalition Vehicle Probe Project: Validation of INRIX Data July-September 2008.* Research Report, Civil Engineering Department, University of Maryland at College Park, I-95 Corridor Coalition, 2009.

Haghani, A., M. Hamedi, K. F. Sadabadi, S. E. Young, and P. Tarnoff. "Data Collection of Freeway Travel Time Ground Truth with Bluetooth Sensors." *Transportation Research Record* (Journal of the Transportation Research Board of the National Academies), no. 2160 (2010): 60-68.

Hall, F. L., and B. N. Persaud. "Evaluation of Speed Estimates Made with Single Detector Data from Freeway Traffic Management Systems." *Transportation Research Record* (Journal of Transportation Research Board of the National Research Council) 1232 (1989): 9-16.

Hamad, K., M. T. Shourijeh, E. Lee, and A. Faghri. "Near Term Travel Speed Prediction Utilizing Hilbert Huang Transform." *Computer Aided Civil and Infrastructure Engineering* 24 (2009): 551-576.

Handley, S., P. Langley, and F. A. Rauscher. "Learning to Predict the Duration of an Automobile Trip." *Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining.* New York: AAAI Press, 1998.

Hazelton, M. L. "Estimating Vehicle Speed from Traffic Count and Occupancy Data." *Journal of Data Science* 2 (2004): 231-244.

Herrera, J. C., and A. M. Bayen. "Incorporation of Lagrangian Measurement in Freeway Traffic State Estimation." *Transportation Research Part B* 44 (2010): 460-481.

Herrera, J. C., D. B. Work, R. Herring, X. Ban, Q. Jacobson, and A. M. Bayen. "Evaluation of Traffic Data Obtained via GPS-Enabled Mobile Phones: The Mobile Century Field Experiment." *Transportation Research Part C* 18 (2010): 568-583.

Hoffman, G., and J. Janko. "Travel Times as a Basic Part of the LISB Guidance Strategy." *Proceedings of the Third International Conference on Road Traffic Control.* London, England: Institution of Electrical Engineers, 1990.

Houston TranStar. *Houston TranStar and Bluetooth Monitoring.* 2011. http://traffic.houstontranstar.org/bluetooth/transtar_bluetooth.html (accessed May 1, 2011).

Ishak, S., and H. Al-Deek. "Performance Evaluation of Short Term Time Series Traffic Prediction Model." *ASCE Jouranl of Transportation Engineering* 128, no. 6 (November-December 2002): 490-498.

Julier, S. J., and J. J. LaViola. "On Kalman Filtering with Nonlinear Equality Constraints." *IEEE Transactions on Signal Processing* 55, no. 6 (June 2007): 2774-2784.

Julier, S., J. Uhlmann, and H. Durrant-Whyte. "A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators." *IEEE Transactions on Automatic Control* 45, no. 3 (March 2000): 477-482.

Kalman, R. "A New Approach to Linear Filtering and Prediction Problems." *ASME Journal of Basic Engineering* 82 (March 1960): 35-45.

Kalman, R., and R. Bucy. "New Results in Linear Filtering and Prediction Theory." *ASME Journal of Basic Engineering* 83 (March 1961): 95-108.

Ki, Y.-K., and D.-K. Baik. "Model for Accurate Speed Measurement Using Double Loop Detectors." *IEEE Transactions on Vehicular Technology* 55, no. 4 (July 2006): 1094-1101.

Krukjian, A., S. Gershwin, P. Houpt, A. Willsky, and E. Chow. "Estimation of Roadway Traffic Density on Freeways using Presence Detector Data." *Transportation Science* 14 (1980): 232-261.

Kwon, J., B. Coifman, and P. Bickel. "Day to Day Travel Time Trends and Travel Time Prediction from Loop Detector Data." *Transportation Research Record: Journal of the Transportation Research Board* (Transportation Research Board of the National Academies), no. 1717 (2000): 120-129.

LeVeque, R. J. *Numerical Methods for Conservation Laws.* Basel: Birkhauser Verlag, 1992.

Li, B. "A Non-Gaussian Kalman Filter with Application to the Estimation of Vehicular Speed." *Technometrics* 51, no. 2 (May 2009): 167-172.

Li, B. "On the Recursive Estimation of Vehicular Speed using Data from a Single Inductance Loop Detector: A Bayesian Approach." *Transportation Research Part B* 43 (2009): 391-402.

Lighthill, M. J., and G. B. Whitham. "On Kinematic Waves: II. A Theory of Traffic Flow on Long Crowded Roads." *Proceedings of the Royal Society of London*, 1955: 317-345.

Lindveld, C. D. R., R. Thijs, P. H. L. Bovy, and N. J. Van der Zijpp. "Evaluation of Online Travel Time Estimators and Predictors." *Transportation Research Record* (Journal of Transportation Research Board of the National Research Council) 1719 (2000): 45-53.

Liu, R.-X., H. Li, and Z.-F. Wang. "The Discontinuous Finite Element Method for Red-and-Green Light Models for the Traffic Flow." *Mathematics and Computers in Simulation* 56 (2001): 55-67.

Liu, Y., and G.-L. Chang. "Estimation of Freeway Travel Time Based on Sparsely Distributed Detectors." *Paper Presented at the 9th International Conference on Applications of Advanced Technologies in Transportation Engineering.* Chicago, 2006.

Liu, Y., P.-W. Lin, X. Lai, G.-L. Chang, and A. Marquess. "Developments and Applications of a Simulation Based Online Travel Time Prediction System: Traveling to Ocean City, Maryland." *Transportation Research Record: Journal of Transportation Research Board* (Transportation Research Board of the National Academies), no. 1959 (2006): 92-104.

Lu, Y., S.C. Wong, M. Zhang, and C.-W. Shu. "The Entropy Solutions for the Lighthill-Whitham-Richards Traffic Flow Model with a Dsicontinuous Flow-Density Relationship." *Transportation Science* 43, no. 4 (November 2009): 511-530.

Lu, Y., S.C. Wong, M. Zhang, C.-W. Shu, and W. Chen. "Explicit Construction of Entropy Solutions for the Lighthill-Whitham-Richards Traffic Flow Model with a Piecewise Quadratic Flow-Density Relationship." *Transportation Research Part B* 42 (2008): 355-372.

Mehran, B., M. Kuwahara, and F. Naznin. "Implementing Kinematic Wave Theory to Reconstruct Vehicle Trajectories from Fixed and Probe Sensor Data." *Procedia Social and Behavioral Sciences: 19th International Symposium on Transportation and Traffic Theory* 17 (2011): 247-268.

Mihaylova, L., and R. Boel. "A Particle Filter for Freeway Traffic Estimation." *43rd IEEE Conference on Decision and Control.* Atlantis, Paradise Islands, Bahamas, 2004.

Nam, D. H., and D. R. Drew. "Analyzing Freeway Traffic under Congestion: Traffic Dynamics Approach." *ASCE Journal of Transportation Engineering* ?, no. ? (May-June 1998): 208-212.

Nam, D. H., and D. R. Drew. "Automatic Measurement of Traffic Variables for Intelligent Transportation Systems Applications." *Transportation Research Part B: Methodological* 33, no. ? (1999): 437-457.

Nam, D. H., and D. R. Drew. "Traffic Dynamics: Method for Estimating Freeway Travel Times in Real Time from Flow Measurements." *ASCE Journal of Transportation Engineering* ?, no. ? (May-June 1996): 185-191.

Nanthawichit, C., T. Nakatsuji, and H. Suzuki. "Application of Probe Vehicle Data for Real Time Traffic State Estimation and Short Term Travel Time Prediction on a Freeway." *Transportation Research Record: Journal of Transportation Research Board* (Transportation Research Board of the National Academies), no. 1855 (2003): 49-59.

Newell, G. F. "A Simplified Theory of Kinematic Waves in Highway Traffic, Part I: General Theory." *Transportation Research* 27B (1993): 281-287.

Ni, D., and H. Wang. "Trajectory Reconstruction for Travel Time Estimation." *Journal of Intelligent Transportation Systems* 12, no. 3 (2008): 113-125.

Okutani, I., D. E. Beskos, and P. G. Michalopoulos. "Finite Element Analysis of Freeway Dynamics." *Engineering Analysis* 3, no. 2 (1986): 85-92.

Park, D., and L. R. Rilett. "Forecasting Multiple Period Freeway Link Travel Times Using Modular Neural Networks." *Transportation Research Record: Journal of Transportation Research Board* (Transportation Research Board of the National Research Council), no. 1617 (1998): 163-170.

Rao, S. S. *Applied Numerical Methods for Engineers and Scientists.* Upper Saddle River, New Jersey: Prentice Hall, 2002.

Research and Innovative Technology Administration. *Pocket Guide to Transportation.* U.S. Department of Transportation. 2010. http://www.bts.gov/publications/pocket_guide_to_transportation/2010/ (accessed April 25, 2011).

Richards, P. I. "Shock Waves on the Highway." *Operations Research* 4 (1956): 42-51.

Rilett, L. R., and D. Park. "Direct Forecasting of Freeway Corridor Travel Times Using Spectral Basis Neural Networks." *Transportation Research Record: Journal of Transportation Research Board* (Transportatin Research Board of the National Research Council), no. 1752 (2001): 140-147.

Robinson, S., and J. W. Polak. "Modeling Urban Link Travel Time with Inductive Loop Detector Data by Using the k-NN Method." *Transportation Research Record: Journal of the Transportation Research Board* (Transportation Research Board of the National Academies), no. 1935 (2005): 47-56.

Sadabadi, K. F., and A. Haghani. "Real Time Solution of Velocity Based First Order Continuum Traffic Model Using Finite Element Method." *Transportation Research Record: Journal of the Transportation Research Board* (Transportation Research Board of the National Academies) xxxx (2011): xx-xx.

Sadek, Shereef A., and Hesham A. Rakha. "A New Speed Formulation Traffic Model for a General Flux Function." *Transportation Research Board Annual Meeting.* Washington, D.C.: National Academies, 2012. 1-14.

Sayed, A. H. *Fundamentals of Adaptive Filtering.* Hoboken, New Jersey: John Wiley & Sons, 2003.

Sethian, J. A. "Fast Marching Methods." *SIAM Review* 41, no. 2 (June 1999): 199-235.

Simon, D. *Optimal State Estimation: Kalman, H[infinity] and Nonlinear Approaches.* Hoboken, New Jersey: John Wiley & Sons, 2006.

Strikwerda, J. C. *Finite Difference Schemes and Partial Differential Equations.* Belmont, CA: Wadsworth & Brooks/Cole, 1989.

Sun, L., J. Yang, and H. Mahmassani. "Travel Time Estimation Based on Piecewise Truncated Quadratic Speed Trajectory." *Transportation Research Part A* 42, no. ? (2008): 173-186.

Sun, X., L. Munoz, and R. Horowitz. "Mixture Kalman Filter Based Highway Congestion Model and Vehicle Density Estimator and its Application." *Proceedings of the American Control Conference.* Boston, Massachusetts, 2004.

Transportation Network Modeling Committee. *A Primer for Dynamic Traffic Assignment.* Washington, D.C.: Transportation Research Board of the National Academies, 2010.

Treiber, M., and D. Helbing. "Reconstructing the Spatio Temporal Traffic Dynamics from Stationary Detector Data." *Cooper@tive Tr@nsport@tion Dyn@mics* 1, no. 3 (2002): 1-24.

Tsitsiklis, J. N. "Efficient Algorithms for Globally Optimal Trajectories." *IEEE Transactions on Automatic Control* 40, no. 9 (september 1995): 1528-1538.

Van Hinsbergen, C. P.IJ., and J. W.C. Van Lint. "Bayesian Combination of Travel Time Prediction Models." *Transportation Research Record: Journal of Transportation Research Board* (Transportation Research Board of the National Academies), no. 2064 (2008): 73-80.

Van Trier, J., and W. W. Symes. "Upwind Finite Difference Calculation of Travel Times." *Geophysics* 56, no. 6 (June 1991): 812-821.

Vanajakshi, L., and L. R. Rilett. "System Wide Data Quality Control of Inductance Loop Data Using Nonlinear Optimization." *ASCE Journal of Computing in Civil Engineering* 20, no. 3 (May-June 2006): 187-196.

Vanajakshi, L., B. M. Williams, and L. R. Rilett. "Improved Flow Based Travel Time Estimation Method from Point Detector Data for Freeways." *ASCE Journal of Transportation Engineering* 135, no. 1 (January 2009): 26-36.

Waller, S. T., Y. -C. Chiu, N. Ruiz-Juri, A. Unnikrishnan, and B. Bustillos. *Short Term Travel Time Prediction on Freeways in Conjunction with Detector Coverage Analysis.* University of Texas, Austin: Texas Department of Transportation and the Federal Highway Administration, 2007.

Wang, Y., and M. Papageorgiou. "Real Time Freeway Traffic State Estimation Based on Extended Kalman Filter: A General Approach." *Transportation Research Part B* 39 (2005): 141-167.

Wang, Y., M. Papageorgiou, and A. Messmer. "Real Time Freeway Traffic State Estimation Based on Extended Kalman Filter: A Case Study." *Transportation Science* 41, no. 2 (2007): 167-181.

Wong, G.C.K., and S.C. Wong. "A Wavelet-Galerkin Method for the Kinematic Wave Model of Traffic Flow." *Communications in Numerical Methods in Engineering* 16 (2000): 121-131.

Work, D. B., O.-P. Tossavainen, S. Blandin, A. M. Bayen, T. Iwuchukwu, and K. Tracton. "An Ensemble Kalman Filtering Approach to Highway Traffic Estimation Using GPS Enabled Mobile Devices." *Proceedings of the 47th IEEE Conference on Decision and Control.* Cancun, Mexico, 2008. 5062-5068.

Work, Daniel B. "Real-Time Estimation of Distributed Parameters Systems: Application to Traffic Monitoring." PhD Dissertation, Civil and Environmental Engineering, University of California, Berkeley, CA, 2010.

Work, Daniel B., Sebastian Blandin, Olli-Pekka Tossavainen, Benedetto Piccoli, and Alexandre M. Bayen. "A Traffic Model for Velocity Data Assimilation." *Applied Mathematics Research eXpress* 2010, no. 1 (2010): 1-35.

Ye, Z. R., Y. L. Zhang, and D. R. Middleton. "Unscented Kalman Filter Method for Speed Estimation using Single Loop Detector Data." *Transportation Research Record* (Journal of Transportation Research Board of the National Research Council) 1968 (2006): 117-125.

You, J., and T. J. Kim. "Development and Evaluation of a Hybrid Travel Time Forecasting Model." *Transportation Research Part C* 8 (2000): 231-256.

Yu, J., G.-L. Chang, H.W. Ho, and Y. Liu. "Variation Based Online Travel Time Prediction Using Clustered Neural Networks." *Paper Presented at the 11th International IEEE Conference on Intelligent Transportation System.* Beijing, China, 2008.

Zhang, X., and J. A. Rice. "Short Term Travel Time Prediction." *Transportation Research Part C* 11 (2003): 187-210.

Zou, N., J. Wang, G.-L. Chang, and J. Paracha. "A Hybrid Model for Reliable Travel Time Estimation on a Freeway with Sparsely Distributed Detectors." *Paper Presented at the ITS World Congress.* Beijing, China, 2007.

Zou, N., J. Wang, G.-L. Chang, and J. Paracha. "Application of Advanced Traffic Information Systems: Field Test of a Travel Time Prediction System with Widely Spaced Detectors."

Transportation Research Record: Journal of Transportation Research Board (Transportation Research Board of the National Academies), no. 2129 (2009): 62-72.
Zwahlen, H. T., A. Russ, E. Oner, and M. Parthasarathy. "Evaluation of Microwave Radar Trailers for Nonintrusive Traffic Measurement." *Transportation Research Record*, no. 1917 (2005): 127-140.