Spring 5-8-2015

Measuring Reliability in Dynamic and Stochastic Transportation Networks

Zifeng Wu
University of Nebraska-Lincoln, wuzi620@gmail.com

Follow this and additional works at: http://digitalcommons.unl.edu/civilengdiss

Part of the Transportation Engineering Commons

MEASURING RELIABILITY IN DYNAMIC AND STOCHASTIC

TRANSPORTATION NETWORKS

By

Zifeng Wu

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Civil Engineering

(Transportation System Engineering)

Under the Supervision of Professor Laurence R. Rilett

Lincoln, Nebraska

March, 2015

MEASURING RELIABILITY IN DYNAMIC AND STOCHASTIC TRANSPORTATION NETWORKS

Zifeng Wu, Ph.D.

University of Nebraska, 2015

Advisor: Laurence R. Rilett

As the traffic demand levels continue to grow in cities, more and more transportation systems experience instability during recurrent and non-recurrent congestion periods. Therefore, reliability has taken on increasing emphases in performance evaluation for transportation agencies, and in performance communication between agencies and the public. Existing reliability-related studies in transportation engineering focus on the long-term reliability of day-to-day travel time variations. This dissertation expands the reliability research literature with studies on the short-term reliability which is valuable for both real-time management and real-time traffic information systems.

This dissertation proposes a level of service reliability metric for system evaluation. Instead of using an average measurement, the confidence interval of a point estimate of the performance measurement of interest is incorporated to evaluate the reliability of each level of service for traffic systems. Bootstrap methods are applied to generate confidence intervals.

A reliability interval based on the travel time standard deviation is defined to describe short-term travel time variability for drivers' information. This dissertation investigates both estimation and prediction methodologies for the mean and reliability interval of travel time, using a five km arterial corridor consisting of three links as a test bed. Regarding the estimation methods, the first-order and second-order approximation

methods show superiority compared with the naïve sum method, which is widely applied to freeway corridors in practice. In terms of the prediction methodologies, the nonlinear autoregressive with exogenous inputs (NARX) neural network is shown to be effective to generate accurate reliability intervals in both the overall condition and the unexpected incident condition.

Finally, the proposed reliability metrics and estimation methodologies are applied on a bimodal traffic network with highway-railway at-grade crossings in Lincoln to evaluate the impact of train traffic on the roadway travel time reliability.

DEDICATION

This dissertation is dedicated to my parents whose constant love and support I could not do without, and to my younger brother whose encouragement always unintentionally reminds me of who I wanted to be when I was at his age. I love you all very much.

ACKNOWLEDGEMENTS

I extend my sincere gratitude to Dr. Justice Appiah, Dr. Bhaven Naik, and all my fellow students in the Nebraska Transportation Center. In addition, special appreciation is sent to Anita Fang, Dan Xu, and Shuihui Pan whose kindness, companionship, and encouragement carried me through those hardest moments.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

By definition, the transportation system is stochastic and dynamic because both demand and supply: 1) are not constant within a given time period, and 2) change as a function of time. With respect to time, transportation parameters can change both within the day and from day-to-day. Consequently, the costs to the operators, users, and non-users are also stochastic and dynamic. These costs could include travel time, monetary costs, emission costs, safety, etc. Historically, the costs have been represented by point estimates such as average travel time on a link. However, these point estimates do not describe the system performance comprehensively, because they lose information on the stochastic and dynamic nature of the system. Intuitively, these stochastic and dynamic costs are best represented by a probability distribution function that is a function of time. In the literature, the term reliability is typically used when describing metrics that seek to capture not only the measure of central tendency (e.g., mean, median, and mode), but also the measure of dispersion (e.g., range and variance). To date, a number of reliability metrics have been developed to address this issue.

In transportation network theory, the demand is represented by the desire to travel. It is typically represented by a production-attraction matrix (planning) or an origin-destination matrix (operations). These matrices attempt to capture the propensity of travel between two locations (e.g., zones or nodes). The supply is often considered to be the infrastructure controlled by the transportation agency. This could include physical infrastructure such as roads, number of lanes, intersections etc., as well as operational strategies such as traffic signal timings. Supply is often represented as the capacity of the

particular link, node, or network. Transportation models, including micro-simulation

models and the four-step planning model (Meyer and Miller 2001), attempt to model the

interaction between the demand (e.g., number of vehicles going from origin i to

destination j) and the supply (the physical network). The output from these models

includes various cost metrics (e.g., travel time, delay) on the nodes, links, and systems.

Historically, the output results of transportation models have been used to

measure reliability performance for coarse aggregate temporal levels at one hour or one

period (e.g., peak hours). Currently, more and more traffic measurements (e.g., travel

time and speed) are becoming available at a fine-grained aggregation level (e.g., 5 minute

interval) or even a discrete level (e.g., individual vehicle), as communication technology

advances. This requires a new way to take advantage of the available dataset to generate

more representative and effective reliability metrics for both traffic system planners and

system users.

1.1 Background

Several research categories for investigating uncertainties in the traffic system

have been developed, including reliability, vulnerability, and robustness (Taylor and

D'Este 2003, 2004; Ukkusuri 2005; Wakabayashi and Iida 1992; Yin et al. 2009). These

categories may be defined by the source of the uncertainty as shown in Table 1.1. Robust

design focuses on optimizing improvement schemes or traffic network design while

taking into account long-term demand uncertainty. The goal is to have the resulting

network insensitive to realizations of uncertain demand levels. The long-term demand

uncertainty could be attributed to external interruptions such as unexpected developments

of the socioeconomic system or prediction errors in travel demand modeling (Zhao and Kockelman 2002). For example, Ukkusuri (2005) proposed a robust network design model under uncertain demands and employed a multi-objective, evolutionary algorithm to identify very good solutions for the non-convex and non-differentiable optimization problem. Yin et al. (2009) studied robust improvement schemes using three optimization models: sensitivity-based, scenario-based, and min-max models.

Table 1.1 Various uncertainty studies in traffic networks

Research category	Source of uncertainty
Robust design	Long-term changes in demand
Vulnerability analysis	Long-term changes in supply (e.g., link failure due to earthquakes)
Reliability analysis	Short-term changes in demand (e.g., peak hour) Short-term changes in supply (e.g., capacity degradation due to incidents)

Vulnerability analysis identifies critical locations in large-scale, sparse, regional, or national transport infrastructure systems and evaluates the ability of transportation networks to withstand severe external disruptions, such as earthquakes or terrorism, which could lead to the complete or potential loss of a subset of network. Taylor and D'Este (2004) defined vulnerability by using the notion of accessibility, stating, "a network node is vulnerable if loss (or substantial degradation) of a small number of links significantly diminishes the accessibility of the node as measured by a standard index of accessibility and a network link is critical if loss (or substantial degradation) of the link significantly diminishes the accessibility of the network or of particular nodes, as measured by a standard index of accessibility". In his work, a definition of "significantly

diminish" was not provided. There are various indices of accessibility that can be considered. Taylor and D'Este (2003) applied generalized travel cost and the Hansen integral accessibility index to analyze the vulnerability of the Australian road network.

Reliability analysis concentrates on congested urban road networks, and the probability that a network will deliver a pre-defined level of performance given short-term uncertainties. The sources of uncertainties considered in the travel time reliability analysis in the Highway Capacity Manual (2010) include: 1) recurring variations in demand; 2) special events that produce temporary, intense traffic demands; and 3) severe weather, incidents, and work zones that reduce capacity.

Research focusing on reliability first began to gain momentum in the early 1990s. There have been various definitions of reliability. One of the most widely accepted definitions was given by Wakabayashi and Iida, which defined reliability as, "the probability of a device performing its purpose adequately for the period of time intended under the operating conditions encountered" (Wakabayashi and Iida 1992; Nicholson et al. 2003). The reliability in traffic system engineering has been defined in a number of ways such as connectivity reliability, travel time reliability, and capacity reliability (Chen et al. 2002; Ching and Hsu 2007; Tu 2008).

Based on previous research, this dissertation proposes a generic methodology for reliability analysis. This proposed approach is tested on both discrete (e.g., fine-aggregate) measurements from the empirical transportation system and from a well-calibrated, micro-simulation model replicating the mechanism of a real traffic network. Of particular note is that the approach will be scalable, meaning it can be used for various parameters (e.g., travel time, speed, and delay), various network levels (e.g., intersection,

corridor, and network), and various analysis periods (e.g., within a day, month, and year). The proposed methodology will: 1) benefit traffic planners by providing a realistic reliability evaluation methodology for stochastic and dynamic properties of traffic systems (i.e., intersections, arterials, OD routes), and 2) benefit individual travelers by providing time-dependent reliability indicators in an easy-to-understand format – a reliability interval of arrival time.

1.2 Problem Statement

A review of reliability analysis in transportation engineering reveals that previous studies on traffic operations centered on defining travel time reliability at the coarse temporal level (e.g., peak hour) using data generated from equilibrium-based methods or collected as averages in the real world (Chen et al. 2002; Haitham and Emam 2006). The fine temporal level (e.g., 15-min interval) is generally used in real-time applications, such as predicting average travel time for the next interval based on real-time and historic data collected by a traffic monitoring system. This dissertation will develop a new methodology to evaluate traffic system reliability based on new reliability metrics using either disaggregated or aggregated data (e.g., at individual vehicle level or average estimates for short-term intervals). This method can be used to measure both day-to-day variability and within-day variability, depending on the input data. The proposed metric indicating day-to-day variability, referred to as long-term reliability, is useful for traffic system planners who are concerned with the general performance of the traffic system. The metric based on within-day variability, referred to as short-term reliability, can be provided to traffic system users through a real-time traffic information system. The short-

term reliability metrics are also useful for traffic system operators to monitor real-time traffic situations and apply on-line management strategies. Methodologies will also be developed to estimate and predict the values of these metrics under various scenarios. The reliability metrics and reliability evaluation methodologies presented in this dissertation can be used for various traffic systems ranging from an intersection to a network between OD pairs. The problems to be addressed in this dissertation include three needs as discussed in the following sections.

1.2.1 Need a Generic Reliability Metric for Reliability Analysis

The concepts of transportation system reliability have been widely studied using data from equilibrium-based models for the entire period of interest (e.g., morning peak hour), under explicitly assumed distributions of demand and capacity (Chen et al. 2002; Du and Nicholson 1997). These studies considered the stochastic property of traffic parameters as static within the period of interest. Travel time reliability is the most commonly studied reliability metric, measured by buffer index, on-time probability, and/or a statistical range. However, traffic planners are also interested in different traffic parameters (e.g., travel time, safety, queue at intersection, etc.) for various design objectives. Furthermore, the stochastic property of traffic parameters such as average travel time is also dynamic in real traffic systems. To reflect the system reliability to system planners in a more comprehensive way, a generic reliability metric for system evaluation is needed. The metric needs to reflect the dynamic randomness of traffic parameters within the entire period of interest by accounting for short-term traffic fluctuations.

In contrast to the previous probability-based reliability metrics, this dissertation will propose an interval-oriented reliability metric. To reflect the uncertainty level in traffic systems, it is more representative to use not only point estimates of traffic parameters, but also the associated confidence intervals. Another advantage to using an interval-based metric is its capability to be correlated with level of service thresholds to yield the level of service reliability that can result in more representative system evaluation results.

1.2.2 Need a Short-Term Travel Time Reliability Metric

Most of the current real-time traffic information systems provide point estimates of key traffic parameters. For example, the Houston TranStar real-time traffic map provides average speed information. Some dynamic message signs at freeways are used to provide en-route travel time information in the format of a range around average travel time (Meehan and Rupert 2004; Oregon Department of Transportation 2005). The range, however, is not related to a statistic concept such as the confidence or prediction interval. For example, the Department of Transportation in Oregon "decided to display the travel time message in two-minute ranges during most times and in up to four-minute ranges during periods of heavy congestion." Previous research about travel time prediction has focused on the point estimates of freeway travel time (Park and Rilett 1998, 1999; Eisele 2001; van Lint et al. 2002; Xia 2006; Naik 2010; Fei et al. 2011). In addition, most in-vehicle en-route guidance systems provide a dynamic routing policy based only on the shortest expected travel time between the intermediate point and the destination node (Miller-Hooks 2001; Yang and Miller-Hooks 2004; Bell 2009). The limitation of the

average-based optimization is that it does not reflect the stochastic properties of real traffic systems. Thus, a new metric is needed to indicate the travel time reliability for different route options. In this dissertation, the reliability interval of arrival time is proposed as a short-term reliability metric, which is defined based on the mean and variance of route travel times.

1.2.3 Need to Identify Efficient Methodologies to Estimate and Predict the Proposed Reliability Metrics

A confidence interval allows the user to estimate the uncertainty of the point estimate. For example, the confidence interval of the mean travel time indicates the likely location of the population mean travel time and indicates the uncertainty of the population mean. The confidence interval is used in this dissertation as a measure of long-term reliability performance for traffic planners and managers. The estimation methods of the confidence interval of traffic parameters needs to be identified for the long-term reliability metric. Confidence intervals can be derived based on parameter mean and variance which are relatively straightforward for an individual parameter with some known distribution (e.g., normal distribution). Non-parametric techniques can be used to estimate confidence intervals for parameters with unknown distribution (Spiegelman 2010). When it comes to a compound parameter, such as route travel time, for which there is not enough of a sample to calculate the confidence interval directly, special methods are needed to consider the correlation among link travel times to yield accurate mean and variance estimations for a route or corridor. This dissertation will compare several existing estimation methods to find the most efficient one.

In terms of real-time traffic reliability information, individual drivers are also interested in the reliability of their own travel time, in addition to the mean travel time estimate for all travelers. In this dissertation, the reliability interval of arrival time is defined as an interval that can include arrival time of a driver departing at a given time with a certain level of confidence. Several prediction methods for the reliability interval are compared and validated through one empirical example.

1.3 Research Objectives

The objective of this dissertation is to develop generalized reliability measures and associated estimation and prediction methodologies to evaluate the long-term and shot-term reliability of traffic systems for traffic agencies and road users. The traffic system could be a single intersection, a corridor, an OD route, or a network. The long-term reliability is investigated in terms of level of service (LOS) reliability to give the overall LOS evaluation while accounting for day-to-day variability. The short-term reliability interval can be provided as real-time traffic information to assist drivers to make better decisions, which is enabled by the high penetration rate of Bluetooth technology and/or micro-simulation tool. The focus of this dissertation will be on the following aspects: 1) to propose a generic reliability metric for reliability evaluation for traffic agencies; 2) to test the efficiency of various methods in estimating corridor travel time mean and variance; and 3) to develop an efficient prediction model for arrival time reliability intervals under both regular conditions and unexpected congestions.

1.4 Structure of the Dissertation

This dissertation consists of six chapters. Chapter 1 introduces basic background information on reliability studies in traffic engineering, states the problems in need of consideration, and outlines the structure for this dissertation. Chapter 2 reviews the previous research in several associated topics, including existing reliability definitions and metrics, travel time prediction and estimation methodologies, en-route traffic information provision, and so on. This dissertation is based on these previous studies and further improves them. Chapter 3 presents a reliability metric for system evaluation and its estimation methodology, together with several examples to illustrate its application using simulated or real traffic data. Two simulation examples are presented to show the capability of the proposed metric to evaluate system reliability given demand and capacity variations in the simulated networks, respectively. The bootstrap algorithm is applied to estimate the reliability metric for individual traffic systems (e.g., an intersection, single link). Furthermore, methods to accumulate link-based statistics (e.g. mean and variance of travel time) into a corridor- or route-based statistic, considering correlation and dependency, are presented and validated using real traffic data. Chapter 4 turns to the reliability indicator for travelers – the corridor travel time reliability interval – and its prediction methodology. Different prediction models are compared using real traffic data to provide guidance for practitioners to choose the "best" model for a given objective. Chapter 5 applies the metrics and methodologies proposed in Chapter 3 and Chapter 4 on a simulated bimodal traffic network in Lincoln, Nebraska. This simulation model is well calibrated and validated. This chapter applies the proposed reliability metrics to quantify the impact of train traffic on the road way travel time reliability on

both route and OD levels. The expected results for travelers and traffic managers describe the system reliability in different perspectives, and have the potential to give operational benefits in real traffic management. The dissertation ends with a summary of findings and future directions presented in Chapter 6.

CHAPTER 2   LITERATURE REVIEW

2.1 Reliability for Traffic System Planners

2.1.1 Reliability for Traffic System Evaluation

Increasingly, reliability has become an important component of the performance evaluation of transportation systems for both system managers and system users. Over the past 20 years, a number of researchers have examined ways to study the reliability of the system. While the Kobe earthquake of 1995 promoted the interest of connectivity reliability, the current increased congestion and demand urges the diversity of reliability indicators such as travel time reliability and capacity reliability.

Connectivity reliability, also known as terminal reliability, was defined by Ching and Hsu (2007) as the probability that there is at least one route connecting the specific OD pair, while links and nodes are subjected to random failure events with known probability in real-world lifeline networks. Alternatively, connectivity reliability may also be defined as the probability that there exists at least one path with a certain traffic service level within a given time period. The certain traffic service level could be a simple physical connection as studied by Wakabayashi and Iida in 1992. Connectivity reliability does not reflect the capacity constraints of links, and is more useful to measure vulnerability of individual components and the network as a whole. One example is to evaluate the degree of performance satisfaction under extreme situations, such as natural disasters and terrorism attacks, which would lead to complete loss of links and nodes.

In daily operations, reliability measurements such as travel time reliability and capacity reliability have been used to assess performance under recurrent congestions, without a complete loss of network components. Du and Nicholson (1997a, b)

distinguished two kinds of travel time reliability according to the source of uncertainty: daily traffic variations and events that lead to link degradations. Iida (1999) assumed that path travel time, as a sum of link travel times which are normally and independently distributed, is also normally distributed, and the path travel time reliability is defined as, "the probability that a trip will arrive at its destination within a given period." Cassir (2001) defines the acceptable level of travel time as the travel time in normal conditions, plus a safety margin. Similarly, Chen et al. (2002) defines travel time reliability as the probability that a trip between a given OD pair can be made successful within a special interval of time. Haitham and Emam (2006) defines link travel time reliability as the probability that the expected travel time at degraded capacity is less than the link free flow travel time plus an acceptable tolerance; demands were assumed to be a normal distribution. Reliability engineering functions based on failure rate were used to estimate the travel time reliability (Haitham and Emam 2006). Tu (2008) summarized four approaches to define travel time reliability: 1) statistical range methods – considering travel time windows in the form of expected travel time, plus or minus a factor times the variance. This travel time window is a confidence interval for the mean travel time; 2) buffer time – the extra percentage travel time due to the travel time variability on a trip that a traveler should take into account in order to arrive on time; 3) tardy trip method – representing the travel time unreliability using the amount of trips that result in late arrivals, and focusing on the length of the delay of the worst trips; 4) probability-based methods – the probability that a trip between a given OD pair can be made successfully within a specified interval of time.

While travel time reliability is easy to understand by travelers in a traffic network, capacity reliability makes more sense from the viewpoint of network planners. Chen et al. (1999) defined capacity reliability as the probability that a network successfully accommodates a given level of travel demand; the capacity is the network reserve capacity (maximum network capacity) that was represented as a multiplier of an existing OD matrix. Furthermore, Chen et al. (2002) extended the research of capacity reliability by considering arc capacities as subject to random variations. Haitham and Emam (2006) modeled OD demand in normal distribution to find out link capacity reliability, which is the probability that the demand x_i is less than the mean link capacity (which is also distributed normally), plus certain acceptable additional flow. Du and Nicholson (1997b) proposed flow decrement reliability to measure reliability by the likelihood that the reduction in flow (as a result of supply-demand interactions) is not less than a threshold, for both OD pairs and the network. Heydeckera et al. (2007) proposed travel demand satisfaction reliability, which is defined as the probability that the road network can accommodate a given latent travel demand. Latent travel demand is estimated using elastic travel demand functions.

The studies above are mainly about developing reliability indicators to evaluate a traffic network. Some of them present sensitivity analysis of reliability indicators as a way to identify the critical components of a network for improvement. For example, Chen et al. (2002) performed a sensitivity analysis by computing derivations related to reliability measurements to identify critical arcs.

Instead of using data generated through demand and supply assumption, recently, there are also reliability analyses based on real traffic data, mainly on travel time data.

Barkley et al. (2012) presents a methodology to determine the optimal number of states

for travel time data. Their methodology also presents a process for distinguishing the

impact on the travel time state of different sources, of non-recurrent congestion.

2.1.2 Reliability for Traffic System Design

In addition to system evaluation, reliability indicators can also be incorporated in

the objective function or constraints of the models for designing traffic networks. Chen et

al. (2010) reviewed the transportation network design problems using bi-level models in

detail:

1) The Mean-variance model optimizes both the expected parameters and the

variance of parameters (e.g., travel time) (Chen et al. 2003; Sumalee et al. 2009;

Yin et al. 2009). Satish Ukkusuri (2005) addressed the discrete network design

problem, under long-term demand uncertainties, using a mean-variance model

with total system travel time as its performance measurement. Final decisions are

represented by a dummy variable to show whether or not to construct a new link.

2) The probability model approach uses probability as its optimization objective

(Chootinan et al. 2005; Chen et al. 2006; Sumalee et al. 2006; Chen et al. 2008).

For example, Sumalee (2006) proposed a stochastic network model with an

objective to maximize the network total travel time reliability, defined as the

probability that the network total travel time is less than a threshold under demand

uncertainties.

3) The chance-constrained model, first developed by Charnes and Cooper (1959),

models stochastic decision systems under the assumption that the "chance

constraint" will be held at least α times. The variable α is a percentage decided by researchers as a confidence level for the system's ability to meet the chance constraint. To design a reliable transportation network, the chance constraint is always related to a reliability definition in probability form (Lo and Tung 2000, 2003; Waller and Ziliaskopoulos 2001; Chen and Yang 2004). For example, Lo and Tung (2003) applied a link capacity chance constraint in their network capacity model to determine the maximum flow that a given network can carry. The chance constraint is that the probability of the traffic flow on link a, exceeding link a's capacity, needs to be less than α_a.

4) The Min-Max model offers a very conservative solution because such a model optimizes the worst-case performance.

5) The alpha reliable model applies the conception of value-at-risk in finance to determine an optimal capacity expansion scheme (Chen et al. 2007). The "alpha" is a user-specified confidence level to guarantee that the probability of total travel time in the network is less than total travel time budget (TTTB); the probability is also the objective of the model to minimize. Total travel time, in this instance, is a random variable due to the uncertainties from design variables (i.e., capacity expansion scheme) and the demand. A higher alpha level indicates a risk-averse design.

6) The scenario-based model is one of the most commonly used ways to deal with demand and capacity uncertainties. Yin et al. (2009) treated various demand levels with different probabilities as separate scenarios.

2.2 Reliability for Traffic System Users

2.2.1 Prediction Models in Traffic Analysis

Traffic parameter measurements are typically time series data. A fundamental property of time series data is its dependency among adjacent measurements. Analyzing the data generating mechanism and forecasting future outcomes can help traffic engineers manage a traffic system more efficiently. A number of macroscopic traffic parameters such as traffic flows, speeds, and occupancies are basic inputs for congestion management. For example, traffic flow forecasting will support the development of proactive traffic control strategies in advanced traffic management systems (ATMSs) and the evaluation of these real-time traffic control strategies. The microscopic parameter, individual travel time, is of great importance in advanced traveler information systems (ATISs). A report by Cheslow et al. (1992) about the architecture of intelligent transportation systems indicated that the ability to make continuous predictions of traffic flows and link travel times for several minutes into the future based on real-time traffic data is a major requirement for providing dynamic traffic control and guidance.

This dissertation aims to predict travel time together with its reliability information. The parameter of interest in this dissertation is travel time. According to the input data, prediction models can be separated as direct and indirect models. Indirect travel time prediction models derive predicted travel times based on other quantities such as flows and/or speeds (Van Lint 2005). Indirect models are applied due to the limitation of available travel time observations. In contrast, direct models predict travel times using previous travel time observations. Generally, prediction methodologies developed for

other parameters (e.g., volume) can also be applied to travel time. Thus, literatures of prediction models for other traffic parameters are also included in this part.

Generally, the models developed for prediction purposes in traffic engineering can be categorized as statistical models and artificial intelligence models. Statistical models include the historical average model, time-series models, dynamic linear models, Kalman filter, and non-parametric models. Artificial models include various neural network models.

Historic Average Model

A historic average method predicts future travel time for each time interval as the average value of all the past travel times at the corresponding time interval. This method is easy to calculate, and the prediction can be refined continuously by updating the historical average with new available observations added. This model, however, depends heavily on the repeatable nature of the traffic flow and cannot reflect sudden changes such as incidents.

Time-series Models

Commonly used time series models include the moving average (MA) model, exponential smoothing filter, auto-regression (AR) model, and auto-regressive integrated moving average (ARIMA) model. For a time-series dataset, adjacent values are usually related to one another. The MA(1) model represents the relationship between adjacent values in a time-series dataset by a process that a random error (e_{t-1}) at the previous time interval (*t-1*), plus a random error (e_t) at the current time interval (*t*), drives the series to yield the output for the mean centered series (Y_t) at current time interval (*t*), where (Y_t) is the difference between observation (y_t) and the mean of the series (μ), and (e_t) is the

white noise error term (McCleary and Hay 1980). The MA model assumes a stationary time-series dataset. The exponential smoothing filter is a direct method to forecast time series with no trend, assuming that the average level changes slowly over time. Exponential smoothing resembles the moving average method, but it gives higher weight to the most recent observation (Bowerman and O'Connell 1979).

In an auto-regression model, the predicted value of the time series is regressed, or expressed as a function of previous observation in the time series. For example, a pilot study by the Northern Region Operation (NRO) of the Virginia Department of Transportation (VDOT) applied an AR(2) model that assumes the predicted travel time at time t takes a linear, weighted form of the observed travel times in the previous two time intervals (Fei et al. 2011). To analyze the data-generating mechanism with both the auto-regression and moving average processes, an ARMA model is proposed to integrate the two models together with the assumption of the stationary characteristic. To relieve the constraints of this assumption, an ARIMA model predicts future data points based on the trends and variations from the previous data points by modeling the autocorrelation in a time series mathematically (Washington 2011). An ARIMA model is identified by (p, d, q), where p is the auto-regressive component, d is the integrated component, and q is the moving average component. The process to develop an ARIMA model consists of model identification, parameter estimation, and validation. Based on the ARIMA modelling philosophy, a seasonal autoregressive integrated moving average (SARIMA) model can account for the seasonality in data. In addition, explanatory variables can be incorporated into an ARIMAX or SARIMA(X) model. Cools et al. (2009) applied ARIMAX and SARIMA(X) models to daily traffic count data to study the impact of holidays on the

variability in daily traffic counts. Xia (2006) studied dynamic travel time prediction based on available data from single-loop detectors and incident reports using a SARIMA model with an embedded adaptive predictor. This method involves multiple-step ahead predictions for flow rate and occupancy in real time. The embedded predictor modifies prediction error based on updated traffic data for every five-minute interval.

One limitation of these linear based time series models (AR, MA, ARMA, and ARIMA) is their inability to deal with large variations and nonlinearity. This usually happens in transportation systems under recurrent and non-recurrent congestions when the system fails to generate continuous and stationary data series. In addition, there is a major difference between the typical time series prediction problems and the task of predicting travel times throughout a day. Most time series prediction models assume that the exogenous factors acting upon the dynamical system either remain constant or can be measured in the model, if they are time-varying (Amani et al. 2011). In a travel time prediction problem, however, the main exogenous factors – the demand and supply on a transportation network – varies widely throughout the day. These factors are typically difficult to quantify due to unpredictable reasons such as work zones, weather, and incidents.

Dynamic Linear Model

ARMA models can be usefully regarded in terms of dynamic linear models, but dynamic linear models can provide flexible framework in treating a non-stationary time series (Petris et al. 2007). A more general class of state-space models can be applied to analyze non-Gaussian and non-linear dynamic systems. Fei et al. (2011) applied a Bayesian dynamic linear model to predict short-term freeway travel time. To estimate

dynamic linear models, the Bayesian approach has both methodological and computational advantages. Bayesian forecasting is a learning process that revises sequentially the state of a priori knowledge based on newly available information. In a good forecasting model, there should be a routine way of learning during phases when predictions and decisions appear adequate, and an exceptional way when they seem unsatisfactory (West and Harrison 1997).

Kalman Filter

The Kalman filter is essentially a set of mathematical equations which can estimate the state of a dynamic process recursively while minimizing the mean of the squared error. To predict travel time, the process can be written into a state-space model by equations 2.1 and 2.2. The model was originally designed in a space tracking setting, where the state equation defines the motion equation for the position or state of a spacecraft, with the location represented as x_t, and y_t, which reflects the information observed such as velocity and azimuth (Shumway and Stoffer 2011).

$$\text{State equation: } x_t = \Phi x_{t-1} + w_t \tag{2.1}$$

$$\text{Observation equation: } y_t = A_t x_t + v_t \tag{2.2}$$

where:

x_t = the state at time interval *t*,

y_t = the observation at time interval *t*,

Φ = the state transition parameter,

A_t = the observation matrix, and

w_t , v_t = the white noises.

For travel time prediction, the state equation is an order-one, auto-regression model, determining the rule for the generation of the average travel time x_t (i.e., state of the traffic system) at time interval *t* from the past state x_{t-1}. The state transition parameter is represented by Φ. The variable y_t is the noisy observation(s) under the current system state, and A_t is the observation matrix. Both Φ and A_t can be defined in various versions as required in specific modeling problems. For example, they could be time-dependent or time-independent. The variables w_t and v_t are assumed to be white noises. The problem is to produce estimators for the underlying, unobserved x_t, given the data $Y_s = \{y_1, \dots, y_s\}$. The Kalman filter provides the solution to this problem with the advantage that it specifies how to update the filter from x_{t-1}^{t-1} to x_t^t once a new observation y_t is obtained, without having to reprocess the entire data set $y_1, \dots, y_t$ (Shumway and Stoffer 2011). The on-going Kalman filter is composed of predictor equations 2.3 and 2.4 and correction equations 2.5-2.7. Before the observation in time interval *t* becomes available, the average travel time at interval *t* is predicted based on previous observations $y_1, \dots, y_{t-1}$ by equation 2.3. The error covariance P_t^{t-1} is estimated by equation 2.4. The predictor equations yield a priori estimates for the next time interval. After the observation in time interval *t* becomes available, the current state at *t* is calculated by equation 2.6, accounting for both the a prior estimate and the inconsistency between the predicted observation and the actual observation. This logic estimates the current state (i.e., average travel time) adaptively by adjusting the a priori estimate with the error of prediction discounted by K_t, the Kalman gain. If the prediction (i.e., $A_t x_t^{t-1}$)

is lower than the actual observation y_t, the a posterior estimate of state x_t^t will be adjusted lower than the a priori estimate x_t^{t-1}, and vice versa. This feedback control will enable a prediction to account for the information from recent observations, which is very important in transition periods with unstable traffic conditions. During such a period, recent observations count for more information than the original trend to predict the system state.

$$x_t^{t-1} = \Phi x_{t-1}^{t-1} \tag{2.3}$$

$$P_t^{t-1} = \Phi P_{t-1}^{t-1} \Phi' + Q \tag{2.4}$$

$$K_t = P_t^{t-1} A_t' [A_t P_t^{t-1} A_t' + R]^{-1} \tag{2.5}$$

$$x_t^t = x_t^{t-1} + K_t(y_t - A_t x_t^{t-1}) \tag{2.6}$$

$$P_t^t = [I - K_t A_t] P_t^{t-1} \tag{2.7}$$

where:

x_t^{t-} , P_t^{t-1} = a priori estimate of state and covariance for time interval *t*,

x_{t-1}^{t-1} , P_{t-1}^{t-1} = state and covariance estimates from time interval *t-1*,

Φ = transition matrix,

A_t = matrix relating the state at previous time interval to the current time interval,

K_t = the Kalman gain for time interval *t*,

Q, R = noise covariance matrices,

x_t^t, P_t^t = state and covariance estimates at time interval *t*, and

y_t = the observation in time internal *t*.

The first application of the Kalman filter in traffic engineering was for traffic volume prediction (Okutani and Stephanedes 1984). The model based on the most recent prediction error and inputs from multiple links demonstrated a high degree of accuracy. Yang et al. (2004) presented a recursive algorithm based on the Kalman filter to dynamically predict short-term traffic volume. Kuchipudi and Chien (2003) developed a hybrid model based on the Kalman filter algorithm for dynamic travel time prediction. Xie et al. (2007) improved the performance of the Kalman Filter on volume prediction by adding discrete wavelet decomposition analysis to divide the original data into several approximate and detailed data. In this way, the noise in the original data was removed and the prediction accuracy was increased, compared to the direct Kalman filter model measured by mean absolute percentage errors and root mean square errors.

The standard Kalman filter (equations 2.3-2.7) was designed to estimate the state of a discrete-time controlled process that is governed by a linear stochastic difference equation. The Extended Kalman Filter (EKF), linearizing the current mean and covariance, can be used to address a non-linear process (Welch and Bishop 1995). One constraint of applying the Kalman filter directly in route travel time prediction is the delay of real-time observations. The filter has to wait until the trip is completed to receive input of new observations.

Nearest Neighbor Model

This non-parametric method is based on the hidden relationship between a large historical database and the current system state. After locating the current system state as a past time neighborhood with similar status, the states of the past systems in this

neighborhood are used to estimate the current state. This method aims for a satisfactory, rather than an optimal result. However, Karlsson and Yakowits (1987) stated that the nearest neighbor approach will result in an asymptotically optimal forecaster, meaning that for an input state vector containing *m* values, the nearest neighbor will asymptotically be at least as good as any *m*th order parametric model. The state vector for travel time prediction can include the travel times, traffic volumes, occupancies, and speeds in previous time intervals. The general methodology of a K nearest neighbor (KNN) model consists of five steps: 1) Build a historical database including traffic patterns such as free-flow, recurrent congestion, and non-recurrent congestion. 2) Define the neighborhood: Two basic approaches can be used to define the neighborhood. The kernal neighborhood has a fixed radius while the nearest-neighbor algorithm has a fixed sample size. 3) Calculate the distance through the absolute value distance or the Euclidean distance. 4) Find K for the nearest neighbor method. 5) Define the prediction function based on the average of the neighborhoods or the weighted average.

The KNN model has been applied to predict traffic volume and travel time (Davis and Nihan 1991; Bajwa et al. 2004; Robinson and Polak 2005). Kim et al. (2005) improved the KNN model by adding a traffic flow pattern recognition technique that uses the signs of changes in the past sequences of traffic volume to overcome the memory-less property of the KNN model. This improvement made by considering the sign changes, however, only reflects the qualitative trend. Furthermore, Qiao et al. (2012) developed a modified KNN model with trend adjustment (KNN-T model) so that the traffic trend effects can be included into the model for short-term travel time prediction. The trend adjustment in this study considered travel time trends both qualitatively and

quantitatively in terms of the signs of changes and the magnitudes of changes in travel times, respectively. For the case study using Bluetooth data on Route I-66 in Virginia, the KNN-T model outperformed the historical average method, ARIMA, Kalman filter, and KNN model for both all-day and peak-hour periods, evaluated by absolute percentage errors (MAPE).

Neural Network

Neural networks (NN) are popular in transportation problems mainly because of their ability to perform self-learning, work with multi-dimensional data, deal with a non-linear problem regarding a flexible model structure, and generate good prediction results (Karlaftis and Vlahogianni 2011). The topology generally used for prediction is a basic and fully connected back propagation multilayer perceptron (MLP) structure, consisting of one input layer, one hidden layer, and one output layer. The hidden layer is used to capture nonlinearity. Neural networks essentially train the connection weights as the hidden neurons learn to recognize different features of the total input space. The training is performed iteratively until the squared error between the computed and the desired output over all the training patterns is minimized (Washington et al. 2011). The trained neural network is capable to predict output values for future inputs. This topology was applied in various transportation prediction problems (Clark et al. 1993; Smith and Demetsky 1994; Park and Rilett 1998; Naik 2010).

This conventional NN structure using arbitrary squashing functions was shown to be theoretically able to approximate any measurable function from a finite-dimensional space to another finite-dimensional space, and to any desired degree of accuracy as long as the hidden layer has sufficient hidden neurons (Hornik et al. 1989). In practice,

however, this NN structure encounters the greatest difficulty in approximating functions with the input features that are not linearly separable. Two techniques were proposed to solve this problem associated with the conventional NN: 1) input feature transformation or a spectral basis neural network (Park et al. 1999), and 2) input-partitioning or a modular neural network is used to approximate with a combination of relatively simpler functions instead of one complex function (Park and Rilett 1998; Kisgyorgy and Rilett 2002).

Some researchers integrated two or more methods to generate hybrid models to pursue higher accuracy and efficiency. One example is to combine some clustering techniques with either time-series analysis or neural network models for integrated prediction. Yin et al. (2002) applied a two-module fuzzy-neural model to predict the traffic flows for an urban street network. A fuzzy approach was used to classify the input data into clusters. For each cluster, a conventional neural network approach was used to model the input-output relationship. Zheng et al. (2006) presented a Bayesian combined neural network approach for short-term freeway traffic flow prediction. In these studies, an appropriate method was applied to classify traffic flow patterns first. Then, a neural network model suitable for each traffic pattern was selected for modeling and prediction. It is important to note that a predictor trained for certain patterns (e.g., peak period traffic) will generate deficient prediction when the traffic pattern changes. Therefore, hybrid models were generally found to be more efficient than a singular predictor.

A feed-forward neural network, where the information flows in a unidirectional way from the input layer to the output layer, is not sensitive to previous processes, making it suitable for recognizing spatially and temporally separated patterns. Recurrent

neural networks can deal with the spatiotemporal dynamics through one or more feedback loops to feed predicted output signals or new output observations back to the input neurons. Van Lint et al. presented a recurrent neural network topology derived from a state-space formulation to predict freeway travel times.

Summary

In this dissertation, the requirement for a nonlinear predictor that can work under both stable and unstable state scenarios led to the application of one class of recurrent neural networks – the nonlinear autoregressive with exogenous inputs Model (NARX) neural network – to predict both mean and variance of corridor travel times on an urban arterial road.

2.2.2 Travel Time Prediction for Freeways

Advanced Traffic Information Systems (ATIS) aim to provide traffic information to help users make better pre-trip and en-route decisions, and receive reliable service out of the system. Real-time travel time information is an essential part for ATIS. Most researchers concentrated on the prediction of point estimates such as average travel time and percentiles. For example, Fei et al. (2011) presented a Bayesian dynamic linear model for short-term travel time prediction on a freeway stretch. The predicted travel time was considered as the sum of the median of historical travel times, time-varying random variations in travel time, and a model evolution error. Some studies also explored methods for travel time reliability prediction. Naik (2010) applied a neural network model to forecast the mean travel time on freeway sections, and bootstrap methods to estimate the standard error of the mean which could be used as a reliability measurement.

Based on the loess non-parametric statistical technique, Eisele (2001) outlined a procedure to estimate the travel time mean and variance from ITS data sources on freeways. Li and Rose (2011) applied a neural network model to predict the travel time range that was determined by the 10th and 90th percentiles in real time for the next ten-minute interval, and up to one hour ahead depending on variables indicating time of day, day of week, rain fall, and the travel time in last ten minutes.

2.2.3 Travel Time Prediction for Arterials

In comparison to freeway travel time, there is limited research with regard to arterial travel times and route travel time. This is because transportation agencies, in general, are more likely to instrument freeways than arterials because they carry more traffic and are not as extensive. In addition, estimation and forecasting are more difficult on arterials because of the complicated interactions among vehicles at intersections. However, with new data sources coming online, including GPS, cell phone, RFID, and Bluetooth, it becomes possible to analyze travel time on an arterial corridor and even on an OD route.

Due to sparse arterial data, research has focused on analytical methods for measuring intersection delay. Lin et al. (2004) decomposed the total delay on an arterial into link delay and intersection delay, and predicted arterial travel time based on the addition of link free flow time and the expected delay time at all the intersections along the arterial. Due to the metering effects of intersection signals, they assumed that the link travel time in midblock is not sensitive to the link flow that remains at medium or high levels without violating the capacity level. Thus, the problem of arterial travel time

prediction was reduced to estimate delay at intersections, which was done through the Webster delay formula and a calibrated transition matrix. The matrix represented the relationship of delay status at all intersections along the arterial. The advantage of this method is that there is no need for detailed calibration, which is difficult in real-time application. The method, however, is based on the existing delay formula that may not be qualified for oversaturated situations. Liu et al. (2006) proposed a hybrid model for predicting urban arterial travel time on the basis of so-called state-space neural networks (SSNNs) and an extended Kalman filter (EKF). The extended Kalman filter was incorporated to avoid laborious and sometimes impossible training and retraining sessions for real-time application. The improved SSNN used movement-separated traffic volumes collected by inductive loop detectors as its input, and predicted arterial travel time as its output. The three neurons in the hidden layer corresponded to the traffic conditions in terms of delays at the three intersections in the arterial for evaluation. This way of defining neurons reflects the relationship in physical traffic systems and avoids treating the neural network as a black-box. The SSNN was trained using observations of arterial travel time detected by two license plate cameras, and the performance of the model was compared with other SSNN trained by Levenberg-Marquardt (SSNNLM) and Kalman Filter (KF) methods. The results demonstrated the advantages of the hybrid model in terms of effectiveness and robustness for predicting arterial travel times.

2.2.4 Travel Time Prediction in Route Guidance Research

The research in route guidance application was generally placed into two classes: the time-adaptive least expected time (LET) hyperpath problem and the a priori LET path

problem. The time-dependent LET research aims to produce a set of path strategies that provide the traveler with the best next direction at each intermediate location depending on the actual arrival time at that location, while the a priori LET path research yields a unique LET path before starting the trip (Opasanon and Miller-Hooks 2006).

In real transportation network, link travel times are both dynamic and stochastic. Dynamic link travel times require the routing strategy to be based on the forecasts of the immediate future of the traffic. Stochastic link travel times require the routing strategy to account for the uncertainty of link travel time. Hall (1986) first investigated the shortest path problem in a dynamic and stochastic transportation network, and revealed that the standard shortest path algorithm may fail to find the expected shortest path in this case. Given the uncertainty about link travel times and other attributes, routing algorithms are required to find out all the paths that may be optimal, termed collectively as a hyperpath, to improve travel time reliability (Bell 2009). Given the dynamic nature of link travel times, routing algorithms are required to be adaptive. Many studies have focused on finding the adaptive hyperpaths in stochastic and time-varying transportation networks (Fu 2001; Miller-Hooks 2001; Yang and Miller-Hooks 2004; Kim et al. 2005; Fu et al. 2006; Gao and Chabini 2006; Opasanon and Miller-Hooks 2006; Ardakani and Sun 2012).

Other studies incorporated reliability explicitly into their routing algorithms through indicators such as travel time variance and reliability indices (Park 1998; Fu 2001; Fu et al. 2006; Kaparias et al. 2008; Kaparias and Bell 2009, 2010; Ardakani and Sun 2012). Among them, Park (1998) presented a heuristic two-stage strategy to identify multiple reasonable routes based on which "near-optimal path" was selected. Link travel

time reliability – in the form of forecasting errors – and link travel time variance were incorporated into multi-criteria objectives to improve the arrival time reliability in route searching. Kaparias et al. (2008) proposed a lateness reliability index and an earliness reliability index to indicate how much later and earlier than the expected arrival time the actual arrival may occur. A modified time-dependent A* algorithm, considering the lateness and earliness reliability indices, was tested by Kaparias and Bell (2009) by conducting experimental drives in the London Congestion Charging Zone and further illustrated in Kaparias and Bell (2010).

Fu and Rilett (1998) were the first to explicitly estimate a route's mean travel time and variance based on available link information in dynamic and stochastic transportation networks, within the context of route guidance application. This is a heuristic method to identify the a priori LET or the LET from any intermediate point to the destination point. In this dissertation, instead of developing an algorithm to find out the LET for drivers, the route information in the form of the mean route travel time and the arrival time standard deviation, will be calculated for different route choices through the heuristic methods in Fu and Rilett (1998). Drivers can choose the route based on the provided information. This solution is beneficial for commute drivers because: 1) commute drivers generally have several route choices in their mind and prefer to know the information on the known routes rather than be guided to some unfamiliar streets; and 2) drivers have the ability to decide the weights of efficiency (i.e., mean route travel time) and reliability (i.e., arrival time standard deviation), which may vary a lot due to different trip purposes.

2.3 Statistical Methods for Traffic Reliability Analysis

2.3.1 Non-Parametric Method for Confidence Interval Estimation

Due to data availability, traffic parameter statistics (e.g., mean and median) used for traffic system evaluation are generally estimated from samples rather than the whole population. The sample-based estimates, however, might not be exactly equal to the true population parameters, resulting in uncertainties for performance evaluation. Standard error is one indicator of such uncertainty. Naik (2010) applied ordinary bootstrap, block bootstrap, and gap bootstrap to estimate the uncertainty of the travel time prediction model. In this dissertation, the confidence interval of traffic parameter estimates will be used to evaluate the uncertainty of traffic system performance. In this section, various bootstrap methods for interval estimation are reviewed.

2.3.1.1 Standard Error Based Confidence Interval

Assuming that the estimator ($\hat{\theta}$) of the true parameter (θ) follows a normal distribution, the $(1 - 2\alpha)$ confidence interval can be approximated as $\hat{\theta} \pm z^{1-\alpha} \cdot \widehat{se}$, where $\hat{\theta}$ is the point estimate of θ and $\widehat{se}$ is the estimated standard error. When the sample size (*n*) is not large enough to make the assumption of normal distribution hold, $\hat{\theta} \pm t_{n-1}^{1-\alpha} \cdot \widehat{se}$ can generate efficient average estimates. These two methods are named as the standard confidence interval and the Student's t interval. They yield equal-tail intervals that are unable to represent the distribution skewness or other errors when $\hat{\theta}$ represents other statistics (e.g., median) instead of the mean.

To relieve the constraints of normal theory assumption and account for unequal tail, the bootstrap-*t* interval was proposed to estimate the distribution $\hat{t}$ directly from the data instead of making the assumption of normal or *t* distribution. The resulting interval is in the form of $[\hat{\theta} - \hat{t}^{1-\alpha} \cdot \widehat{se}, \hat{\theta} - \hat{t}^{\alpha} \cdot \widehat{se}]$. It is important to note that $\hat{t}^{1-\alpha}$ is not equal to $\hat{t}^{\alpha}$ in regards to skewness.

To apply this method, an efficient way to estimate the standard error estimator is necessary for the dataset with dependent structure. It is well-established that the standard error of the sample mean could be estimated using $\sqrt{s^2/n}$, where $s^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2 / (n-1)$. However, there is no such equation for most statistical estimators (e.g., median). In these instances, the bootstrap estimate of standard error first proposed by Efron in 1979, can be used. It is illustrated using the statistical estimator median as an example. The basic bootstrap algorithm starts with generating a large number of independent bootstrap samples: x^{*1}, x^{*2},..., x^{*B}, each of size n. The number of samples (B), generally ranges from 50 to 200 for standard deviation estimation. Bootstrap median replicates $s(x^{*1})$, $s(x^{*2})$,..., $s(x^{*B})$ can be calculated for each sample. The standard deviation of these replicates is the standard error estimator of the median $s(x)$, as shown in equation 2.8.

$$\widehat{se}_{boot} = \{\sum_{b=1}^{B}[s(x^{*b}) - s(\cdot)]^2/(B-1)\}^{1/2} \tag{2.8a}$$

$$s(\cdot) = \sum_{b=1}^{B} s(x^{*b})/B \tag{2.8b}$$

where:

$\widehat{se}_{boot}$ = the estimated standard error of median using bootstrap-*t* method,

B = the size of bootstrap sample, and

$s(x^{*b})$ = *b*th bootstrap median replicate.

Different from the standard intervals which are symmetric around zero, the asymmetric intervals resulting from bootstrap-*t* percentiles represent an improvement in coverage. It is particularly applicable to location statistics like the sample mean, median, and other percentiles, but is not trustworthy for more general problems such as setting a confidence interval for a correlation coefficient. An overall assessment of the three standard-error based confidence intervals are quoted from Efron and Tibshirani (1993):

> "The increase in accuracy of estimation for Bootstrap-*t* approximation is at the price of generality. The standard confidence interval applies to all samples, and all sample sizes; the student-t table applies to all samples of a fixed size n; the bootstrap-t table applies only to the given sample."

2.3.1.2 Percentile Based Confidence Interval

Although the bootstrap-*t* method can theoretically account for skewness and yield good theoretical coverage probabilities, it can yield somewhat erratic results in practice. Improved methods use percentiles instead of the standard error of bootstrapped estimates to identify the confidence limits.

If bootstrap distribution of $\theta^* = s(x^*)$ is roughly normal, then the standard normal and percentile intervals will nearly agree. The bootstrap distribution can be regarded as a normal distribution if sample size n approaches infinite, according to the

central limit theorem. However, this might not hold for small samples in which case the percentile interval is superior to the standard normal interval. Also, a percentile interval has transformation-respecting and range-preserving property. By range-preserving property, a percentile interval always falls within the allowable range of its estimator. Although percentile intervals are less erratic in practice compared to bootstrap-t intervals, they have less satisfactory coverage properties.

Given independent bootstrapped samples $x^{*1}, x^{*2}, \ldots, x^{*B}$, each of size n, bootstrap replicates $\hat{\theta}^*(b) = s(x^{*b}), b = 1, 2, \ldots, B$. Denote $\hat{\theta}_B^{*(\alpha)}$ as the 100αth empirical percentile (i.e., the value in the ordered list of the B replications of $\hat{\theta}_B^*$). The $(1 - 2\alpha)$ percentile interval would be $[\hat{\theta}_B^{*(\alpha)}, \hat{\theta}_B^{*(1-\alpha)}]$. It needs more bootstrap samples (B) for percentile estimation than for standard error estimation. Variable B should be greater than 500 or 1000 to make the variability of percentile estimators acceptably low.

2.3.1.3 Bias-Corrected and Accelerated (BCa) Interval

The BCa interval is an improved version of the percentile method in both theory and practice. Given enough sample size, the resulting interval would closely match exact confidence intervals in special situations, where the statistically exact interval is accessible through statistical theory, and give dependably accurate coverage probabilities in all situations. In addition, the BCa method is also transformation-respecting. Integrating the performance on accuracy and flexibility, the BCa method is recommended for general use by Efron and Tibshirani (1993).

The end points of the BCa interval is modified by acceleration ($\hat{a}$) and bias-correction ($\hat{z}_0$). The BCa interval of intended coverage $(1 - 2\alpha)$ is given by equation 2.9.

The notation $\Phi(\cdot)$ is the standard, normal cumulative distribution function and $z^{(\alpha)}$ is the 100αth percentile point of a standard normal distribution. For example, $\Phi(1.645) = 0.95$ and $z^{(0.95)} = 1.645$. It can be assessed from equation 2.9 that if $\hat{a}$ and $\hat{z}_0$ are zero, the interval is equal to the percentile interval. Non-zero $\hat{a}$ and $\hat{z}_0$ correct deficiencies of the previous standard and percentile methods.

$$\text{BCa: } (\hat{\theta}_B^{*(\alpha_1)}, \hat{\theta}_B^{*(\alpha_2)}) \tag{2.9a}$$

$$\alpha_1 = \Phi(\hat{z}_0 + \frac{\hat{z}_0 + z^{(\alpha)}}{1 - \hat{a}(\hat{z}_0 + z^{(\alpha)})}) \tag{2.9b}$$

$$\alpha_2 = \Phi(\hat{z}_0 + \frac{\hat{z}_0 + z^{(1-\alpha)}}{1 - \hat{a}(\hat{z}_0 + z^{(1-\alpha)})}) \tag{2.9c}$$

where:

$\hat{a}$ = acceleration,

$\hat{z}_0$ = bias-correction,

$z^{(\alpha)}$ = the 100α th percentile point of a standard normal distribution.

2.3.1.4 Modified Bootstrap

When the dataset is not composed of independent observations, the standard bootstrap method is not enough to get independent bootstrap samples, and modified bootstrap (e.g., block bootstrap) is needed. Specifically for the traffic dataset with dependent observations within one day, Lahiri et al. (2012) applied gap bootstrap to generate consistent and asymptotically unbiased estimates of standard error for a massive dataset with certain dependent structure.

2.3.1.5 Summary

This dissertation compares the coverage of the standard error based confidence interval and the BCa confidence interval. The BCa method is selected to calculate the confidence interval of individual traffic parameters.

2.3.2 Measurement of Effectiveness (MOE)

The estimated and predicted results need to be compared with real observations to measure the effectiveness of estimation and prediction methodologies. Commonly used MOEs include mean absolute error, mean absolute percentage error, and root mean squared error. The equations for calculation are as follows:

Mean absolute error (MAE): $MAE = \frac{1}{n} \sum_{t=1}^{n} |X(t) - \hat{X}(t)|$ (2.10)

Mean absolute percentage error (MAPE): $MAPE = \frac{1}{n} \sum_{t=1}^{n} |\frac{X(t) - \hat{X}(t)}{X(t)}|$ (2.11)

Root mean squared error (RMSE): $RMSE = \sqrt{\frac{\sum_{t=1}^{n}(X(t) - \hat{X}(t))^2}{n}}$ (2.12)

where:

  $X(t)$ = the real observation for time interval *t*,

  $\hat{X}(t)$ = the estimated or predicted value for the time interval *t*, and

   n   = the number of time intervals for analysis.

In addition, the CDF plot of absolute percentage error (APE) is also used in this dissertation to compare the distribution of estimation and prediction errors from various methods.

When there is a need to compare two sets of traffic volumes so that simulation models can be evaluated based on real traffic observations, the GEH statistic, first proposed by Geoffrey E. Havers, provides a way to incorporate both relative and absolute errors. One example is the fitness evaluation of calibration results for a simulated transportation network. The evaluation generally involves several link volume observations as the benchmarks to calibrate the simulated link volumes. The calibration algorithm aims to minimize the difference between the observed link volumes and the simulated link volumes. A difference of 500 vehicles on a link with a high volume may provide a better fit than a difference of 500 on a lightly trafficked link. The GEH statistic, in the form of the chi-squared statistic, can take the variation in volumetric differences into account (Train 2003). GEH is calculated by the mathematical formula as shown in equation 2.13.

$$GEH = \sqrt{\frac{2(V - v)^2}{V + v}} \qquad (2.13)$$

where:

V = the simulated traffic volume, and

v = the observed traffic volume.

2.3.3 Tests for Comparing Means and Distributions

Tests for comparing means and distributions are used to study the variation of traffic parameters in adjacent or close short-time intervals. A *t*-test can be used to check if the means of two independent and identically distributed samples differ from each other significantly. For example, a *t*-test can be used to compare the mean of travel times in two adjacent 5-minute intervals, within one day. A *t*-test can also be used to determine if two samples are significantly different from each other under the assumption that the test statistic follows a normal distribution. The Kolmogorov–Smirnov test, as an alternative to compare two sample distributions, does not require such an assumption. For samples consisting of paired observations of similar units, a paired *t*-test or a Wilcoxon signed test can compare the means.

2.3.4 Test for Normality

The Lilliefors test, testing the null hypothesis that the sample data came from a distribution in a normal family, is implemented by three steps (Lilliefors 1967). First, the population mean and variance are estimated based on the sample data. Next, as in the Kolmogorov-Smirnov test, the test statistic is identified as the maximum discrepancy between the empirical distribution function of sample data, and the cumulative distribution function of the normal distribution with the estimated mean and estimated variance from the first step. Lastly, if the maximum discrepancy is large enough to be statistically significant, the null hypothesis will be rejected.

CHAPTER 3   DEVELOPMENT OF RELIABILITY METRIC AND ESTIMATION METHODOLOGY FOR TRAFFIC MANAGEMENT

To quantitatively evaluate the quality of service in transportation systems, transportation professionals often use six levels of service (LOS A to F) that are defined based on a point estimate (i.e., average) of traffic performance measurements such as density. For example, LOS A on a freeway reflects the best operating conditions, while LOS F represents a hard-to-predict, stop-and-go condition (Highway Capacity Manual 2010). The six-level LOS definition easily communicates roadway performance to nontechnical decision makers. However, using the LOS based on average values might hide short-term variations in a traffic stream within an analysis period. Two traffic networks with the same average performance measurement could provide different service in terms of reliability. In this chapter, the concept of reliability, with respect to traffic network performance, will be defined based on the confidence intervals of performance measurements. The confidence interval is used to take into account the underlying uncertainties in the dynamic and stochastic traffic system. The proposed metric combines the LOS concept and reliability theory to yield a more comprehensive evaluation measurement. The concept of reliability developed in this dissertation can be applied to various quantitative traffic parameters and across varying time and space ranges to reflect short-term fluctuations in traffic systems.

3.1 Necessity of Performance Evaluation Considering Short-Term Fluctuations

Real travel time data recorded on one link of I-495 on Tuesday, Wednesday, and Thursday from 04/01/2012 to 09/30/2012, are used to demonstrate the existence of short-term travel time fluctuations, quantitatively. The 2.57 km link, between the Connecticut Avenue Interchange (I/C) and the Georgia Avenue I/C, is shown in figure 3.1. The link of interest is marked by "Start" and "End". The travel time information has been aggregated into five minute intervals.
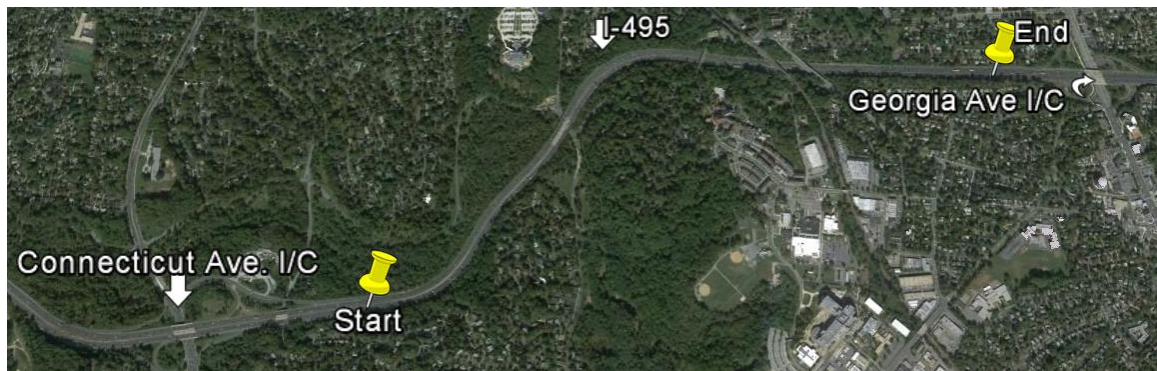
Figure 3.1 Geographical location of the real traffic time data

The paired t-test and the Wilcoxon test are used to test the differences between the first 5-minute interval in a hour (i.e., 16:00-16:05) and the five following 5-minute intervals within the same half hour (i.e., 16:05-16:10, 16:10-16:15, 16:15-16:20, 16:20-16:25, 16:25-16:30).

The null hypothesis H_0 of the paired t-test is that the mean difference d between the paired mean travel times for two different 5-minute intervals is zero, as shown in equation 3.1a.

$$H_0: d = \mu_1 - \mu_2 = 0 \tag{3.1a}$$

$$H_a: d \neq 0 \tag{3.1b}$$

where:

d = the difference in the mean values of the two groups being studied,

μ_1 = the mean of the observed travel times for the [16:00-16:05] interval, and

μ_2 = the mean of the observed travel times for the other 5-minute interval being

studied.

The Wilcoxon singed-rank test examines if the differences between each pair

come from a distribution with a median of zero (Dowdy et al. 2005). In the first step, the

absolute differences are ranked from the smallest to largest. The signs of differences (i.e.,

+ or -) are attached to their respective ranks to obtain signed ranks that are averaged to

obtain the mean of signed ranks ($\bar{r}$). The null hypothesis is that the expected value of $\bar{r}$ is

zero, as shown in equation 3.2a. If the null hypothesis is true, it implies that the

differences between the members of the two 5-minute intervals are just random and the

travel times in the two intervals are from the same distribution, as judged by the median

difference.

$$H_0: E(\bar{r}) = 0 \tag{3.2a}$$

$$H_a: E(\bar{r}) \neq 0 \tag{3.2b}$$

where:

r = the rank of absolute difference between the two groups being studied, and

$\bar{r}$ = the mean value of signed ranks (+/- *r*).

Table 3.1 shows the p-value of the paired t-test and Wilcoxon test for different 5-minute intervals within the half hour from 16:00 to 16:30. A p-value less than 0.05 indicates a rejection of the null hypothesis at the 95 percent confidence level. The paired t-tests indicate that the mean travel times of the 5-minute interval from 16:00 to 16:05 and the four intervals after 16:10 are significantly different at the 95 percent confidence level. The Wilcoxon tests indicate that the travel times of the 5-minute interval from 16:00 to 16:05 and the three intervals after 16:15 are from different distributions, judged by median difference, at the 95 percent confidence level.

Table 3.1 Results of the paired *t*-test and Wilcoxon signed test

[16:00-16:05] vs.	[16:05-16:10]	[16:10-16:15]	[16:15-16:20]	[16:20-16:25]	[16:25-16:30]
Paired *t*-test p-value	0.07	**0.04**	**0.03**	**0.02**	**0.00**
Wilcoxon signed test p-value	0.08	0.07	**0.03**	**0.05**	**0.00**

Note: 1). Sample sizes are 76
      2). Bold values indcates a statistically significant difference at 5 percent significance level
      3). All the 5-minute intervals are compared to the 16:00-16:05 interval

These results demonstrate the existence of short-term travel time fluctuations within a half hour period. For this specific example, the statistically significant change in travel time means occurred at 16:10. However, this is not fixed for all traffic systems. The traditional LOS approach uses the average estimate for the peak 15 minute interval to evaluate traffic network performance for the whole peak hour. By definition, it is not able to reflect the short-term fluctuations or dynamic property of system performance. Note

that short-term fluctuations may not make a significant change in system performance for traffic systems with stable traffic flow that operates under system capacity. However, such short-term fluctuation could lead to appreciable variation in system performance under unstable traffic situations, such as the signal pre-emption control strategy at highway-railway at grade crossings. In those situations, it is hypothesized that considering such short-term variability when quantifying the performance evaluation of traffic networks can yield more comprehensive, reliability-based evaluation results and can help communicate more accurate system performance to system users and policy makers. The reliability metric presented in this dissertation incorporates the confidence intervals of traffic measurements for short-term intervals into performance LOS evaluation, which can address the aforementioned issues regarding the system evaluation for unstable traffic systems.

3.2 Definition of a Generic Reliability Metric for Evaluation

In this dissertation, the reliability of the transportation system is defined as the ability of the system to adapt to internal changes while maintaining a satisfactory system performance. Internal changes include variations in demand and/or capacity under prevailing conditions. The generic reliability metric, therefore, is calculated as a percentage of the number of times when a dynamic and stochastic transportation system can provide satisfactory service given uncertainties in transportation demand and/or transportation supply. To analyze the system reliability, the variation of system performance due to interval changes needs to be represented efficiently, and the level of satisfactory performance variation needs to be specified. In equation 3.3, which is used to

calculate the reliability metric, C is the interval to represent system performance variation, and I is the interval to represent satisfactory performance variation. In this chapter, the confidence interval of performance measurements is used as C because, statistically, a confidence interval represents the uncertainty of population statistic estimates. Traffic planners can also define their own interval to represent the system performance variation of specific interest. The definition of "satisfactory system performance" needs to be determined by system planners through I. A common metric for system performance is the level of service (LOS) that relates a letter "grade" to quantitative traffic parameters (e.g., density, delay). Using the thresholds of LOS i as evaluation thresholds for I, the generic reliability metric can yield the reliability of LOS i for the system in evaluation. System planners can also determine their own evaluation intervals for a special evaluation objective.

The reliability metric for a given traffic parameter x (i.e., $R_{\hat{x}}$) represents the probability that for the whole period of interest (T), the confidence intervals (C) of parameter statistic $\hat{x}$, for all the short-time intervals (dt) in T, are included by the satisfactory evaluation interval (I) corresponding to the performance measurement $\hat{x}$.

$$R_{\hat{x}} = \frac{\int_0^T \varphi \cdot [min(C_U^{\hat{x}}, I_U^{\hat{x}}) - \max(C_L^{\hat{x}}, I_L^{\hat{x}})]dt}{\int_0^T (C_U^{\hat{x}} - C_L^{\hat{x}})dt} \quad \forall \ \hat{x} \tag{3.3a}$$

$$\varphi = \begin{cases} 0 & C_U^{\hat{x}} \leq I_L^{\hat{x}} \ or \ C_L^{\hat{x}} \geq I_U^{\hat{x}} \\ 1 & otherwise \end{cases} \tag{3.3b}$$

where:

x = the traffic parameter of interest (e.g., travel time), x̂ is the estimate of a statistic

of x (e.g., average travel time),

$R_{\hat{x}}$ = the reliability metric for x̂,

dt = the short-time interval,

T = the analysis period,

$C_U^{\hat{x}}$ = the time-dependent function of the upper bound of the confidence interval for $\hat{x}$

at the 95 percent confidence level,

$C_L^{\hat{x}}$ = the time-dependent function of the lower bound of the confidence interval for $\hat{x}$

at the 95 percent confidence level,

$I_U{}^{\hat{x}}$ = the upper threshold of the evaluation interval I for x̂,

$I_L{}^{\hat{x}}$ = the lower threshold of the evaluation interval I for x̂, and

φ = the inclusive factor determined by the relationship between C and I. If the

confidence interval area between $C_U^{\hat{x}}$ and $C_L^{\hat{x}}$ are not included by I, $\varphi = 0$. This

includes the cases where C is below $I_L{}^{\hat{x}}$, and where C is above $I_U{}^{\hat{x}}$, as shown in

figures 3.2a and 3.2b. If at least part of the confidence interval area is included by

I, $\varphi = 1$. Figures 3.2c-3.2f show different possibilities when C intersects with I,

and the gray areas indicate the portion of area C is included by area I. In figure

3.2c, area I is greater than area C, and all of C is included by I. In figure 3.2d,

area C is greater than area I and includes the entire I. Therefore, area I is equal to

the proportion of area C that is within the thresholds $I_U{}^{\hat{x}}$ and $I_L{}^{\hat{x}}$. In figure 3.2e,

the proportion of C within the thresholds of I is equal to $(I_U{}^{\hat{x}} - C_L{}^{\hat{x}}) \cdot dt$. In

figure 3.2f, the proportion of C within the thresholds of I is equal to $(C_U{}^{\hat{x}} - I_L{}^{\hat{x}}) \cdot dt$.
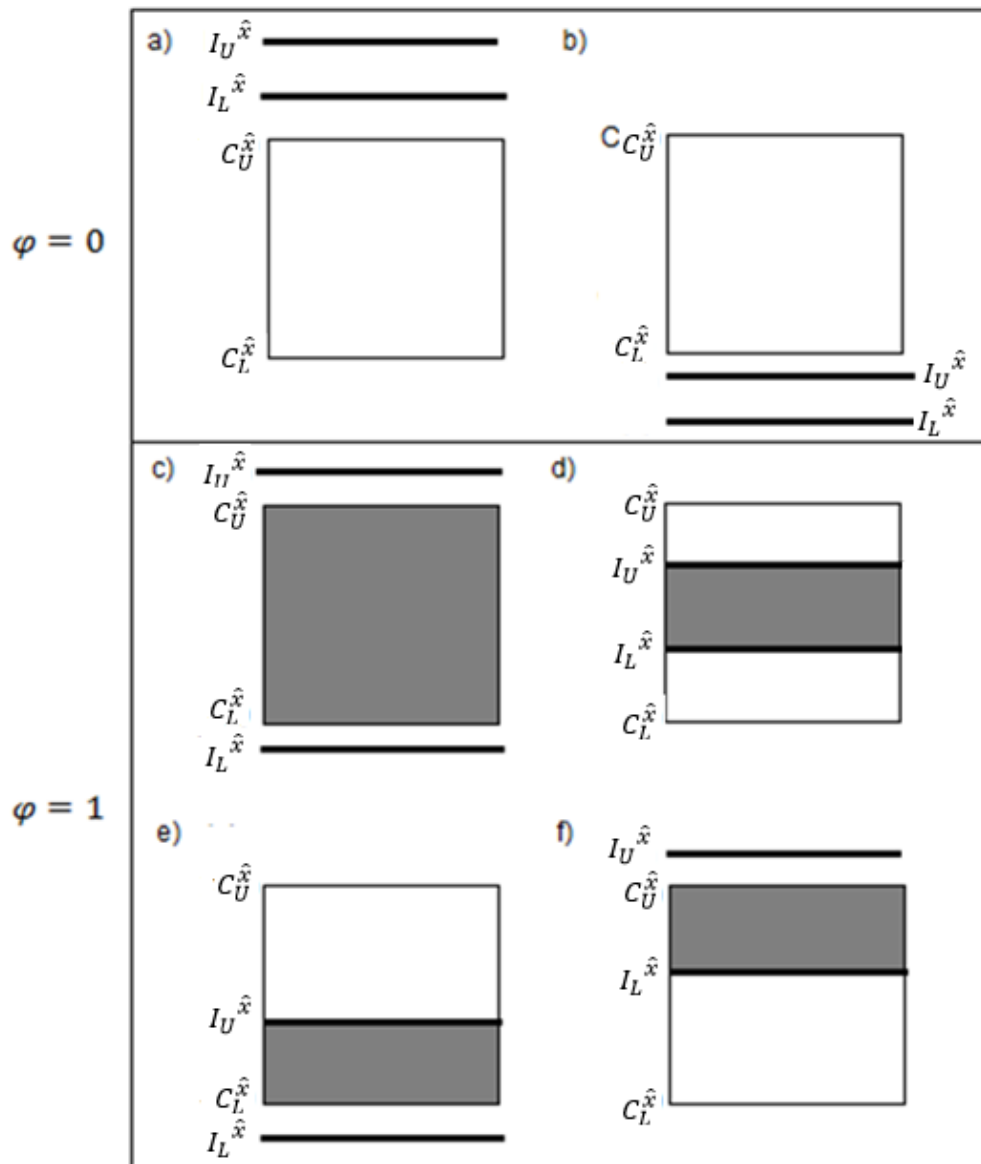
Figure 3.2 Values of φ and the corresponding relationships between confidence interval (C) and evaluation interval (I)

The variable x in equation 3.3a is objective-specific. For example, x might be: 1) the delay time or queue length for evaluating timing plans at an intersection, 2) the arterial travel time for evaluating the arterial coordination, or 3) the speed for freeway

evaluation. For the evaluation of the service level between a given OD pair, *x* could be the route travel time. In addition, researchers can choose which statistic (e.g., the estimated mean) to use for a given traffic parameter (e.g., travel time). For example, if the traffic parameter of interest is not normally distributed, then the median may be chosen because it limits the impacts of outliers. In studies related to speed limits, the 85th percentile may be appealing. This metric could be used as a quantitative standard for traffic agencies to evaluate the reliability-oriented performance of current and planned traffic systems. For example, a before-and-after comparison of the generic reliability metrics, in terms of the delay time and the number of conflicts, could be used to compare alternative signal timing strategies to identify which strategy is more reliable.

The evaluation interval *I* needs to be decided by traffic agencies and could vary by application and evaluation objective. In this dissertation, the LOS thresholds provided in the Highway Capacity Manual (HCM) are applied to analyze real traffic data collected for test beds in the field. In addition, the simulation examples for illustration purposes are given an assumed evaluation interval based on engineering judgment. In contrast to point estimates of traffic parameters, the metric shown in equation 3.3 considers variability due to short-term fluctuations within the period of interest. For example, consider the problem of evaluating the level of service at an intersection. The result would be a probability (e.g., *p*) where the confidence intervals of the average delay at the intersection fall into the LOS thresholds (e.g., B) defined in the HCM. In other words, the reliability metric would indicate that the intersection can provide the level of service B with a probability of *p*, over the time period of interest.

3.2.1 The Generic Reliability Metric and Demand Variation

In this section, the generic reliability metric is applied to evaluate the route travel time on a test bed during the afternoon peak period. The traffic demand for the afternoon peak period (D1) is 1,500 vehicles per hour. For comparison purposes, a demand level (D2) that is 15 percent higher than the current peak demand level is also analyzed. Figure 3.3 shows the test route that has an origin at the intersection of Vine Street and North Antelope Valley Parkway, and a destination at the intersection of N.16th Street and S Street. The route is 800 meters in length and consists of two signalized intersections: Vine Street and North16th Street, and Vine Street and North 17th Street. The two intersections are located 134 meters apart. Drivers may need to stop multiple times to get through the test route during the afternoon peak. The network was modeled in VISSIM, a microscopic traffic simulation tool. The simulation was run five times. Each run used a different random seed number. The simulation time was 70 minutes for each run. The first 15 minutes allowed the simulation to reach a steady state, and the last 55 minutes were used to collect data. The traffic parameter of interest is travel time (i.e., x = travel time; $\hat{x}$ = mean travel time). It is assumed that: 1) the confidence interval at the 95 percent confidence level lies within 1.96 standard deviations of the mean; and 2) assuming that 85 percent of the drivers regard their route travel times as acceptable, the upper threshold of evaluation interval I is the 85th percentile of the simulated travel times under current peak hour demand (D1). That is, the evaluation interval I for acceptable service is [0, 710] seconds.

Figures 3.4a and 3.4b show the resulting five-minute means and confidence intervals for two demand levels (i.e., D1 and D2), respectively. The green line in figure

3.4 shows the five-minute means of route travel time as a function of the elapsed simulation time. The blue and red lines show the upper and lower bounds of the confidence intervals (C) of the means as a function of the elapsed simulation time. The area between the upper bound and lower bound of the confidence intervals (i.e., blue and red lines in figure 3.4) is referred to as the confidence interval area. This confidence interval area is a numerical measurement of the variability in mean route travel time over a given analysis time period. The upper threshold of the user-specified evaluation interval (I) is shown as a black solid line in figure 3.4. For both demand levels, the mean route travel time, which is shown as a green line, is below the upper threshold. That is, using the mean route travel time for evaluation will result in the conclusion that an increase of 15 percent traffic demand will not change the system performance. In contrast, the 5-minute confidence interval (C) areas between the blue and red lines increases as demand increases. The impact of this change on performance evaluation can be quantitatively captured using the reliability metric proposed in this dissertation.

In this example, the reliability metric $R_{\hat{x}}$ is the probability that 5-minute confidence intervals of mean travel time $\hat{x}$ within the analysis period are included in the evaluation interval (I). The percentage of the confidence interval area within I of the overall confidence interval area is calculated as $R_{\hat{x}}$. The calculation procedure for current peak hour demand (i.e., D1=1,500 vehicle per hour) is summarized in table 3.2 as an illustration. It was found that: $R_{\hat{x}}$ is 0.73 for the level 1 demand (i.e., 1,500 vehicles per hour), while $R_{\hat{x}}$ is 0.69 for level 2 demand (i.e., 1,730 vehicles per hour). For this example, an increase of 15 percent in the traffic demand will degrade the reliability of providing acceptable service by 6 percent.

This example illustrates the ability of the proposed reliability metric to reflect the detailed change on system performance due to demand variation, using simulation data, and the assumed evaluation interval and confidence interval for mean travel time. For real applications, the evaluation thresholds can be based on the Highway Capacity Manual or engineering judgment. The confidence intervals need to be estimated using collected traffic data.
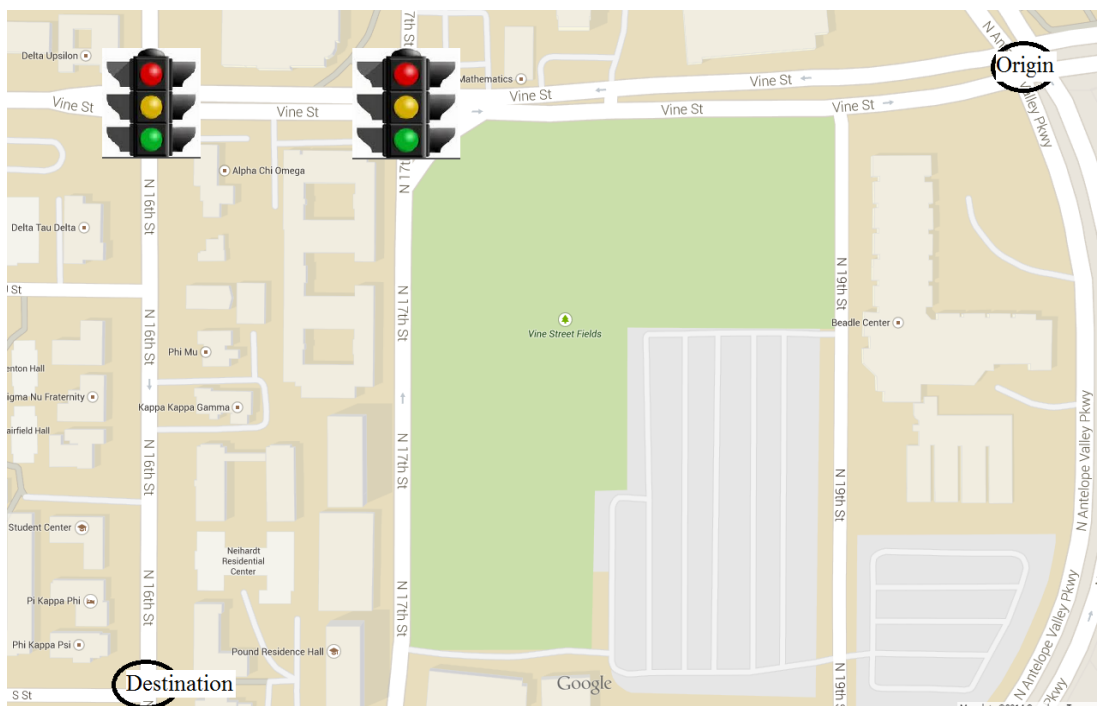
Figure 3.3 OD pair for example demonstration
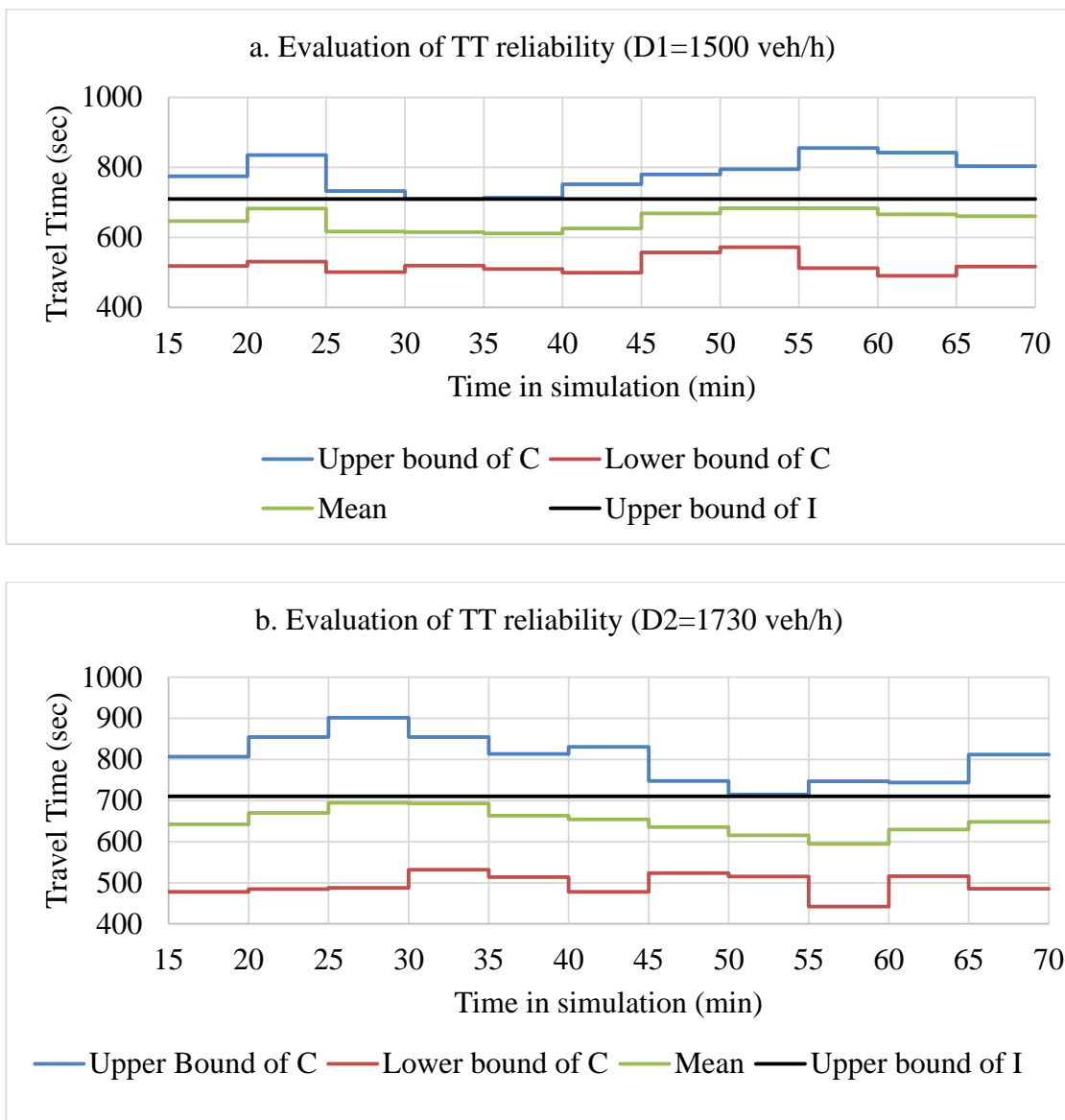
Figure 3.4 Travel time (TT) reliability evaluation for two demand levels

Table 3.2 Calculation of $R_{\hat{x}}$ for demand level D1 scenario

5-minute interval from t_i in simulation (seconds)	15	20	25	30	35	40	45	50	55	60	65
Mean TT (seconds)	646	683	617	615	612	626	668	683	684	666	660
$C_U^{\hat{x}}$ (seconds)	774	835	733	710	714	752	780	795	855	842	804
$C_L^{\hat{x}}$ (seconds)	518	530	501	520	510	499	557	572	513	490	517
φ	1	1	1	1	1	1	1	1	1	1	1
$min\left(C_U^{\hat{x}}, I_U^{\hat{x}}\right)$	710	710	710	710	710	710	710	710	710	710	710
$max\left(C_L^{\hat{x}}, I_L^{\hat{x}}\right)$	518	530	501	520	510	499	557	572	513	490	517
$min\left(C_U^{\hat{x}}, I_U^{\hat{x}}\right) - max\left(C_L^{\hat{x}}, I_L^{\hat{x}}\right)$	192	180	209	190	200	211	153	138	197	220	193
$C_U^{\hat{x}} - C_L^{\hat{x}}$	256	305	232	190	204	253	223	223	342	352	287

Note: 1) D1=1500 vehicle per hour

2) $I_U^{\hat{x}} = 710$ seconds. $I_L^{\hat{x}} = 0$ second; $\hat{x}$ = Mean Travel Time (TT)

3) $R_{\hat{x}} = \dfrac{\int_0^T \varphi \cdot [min\left(C_U^{\hat{x}}, I_U^{\hat{x}}\right) - max(C_L^{\hat{x}}, I_L^{\hat{x}})]dt}{\int_0^T (C_U^{\hat{x}} - C_L^{\hat{x}})dt} = 0.73$

3.2.2 The Generic Reliability Metric and Capacity Variation

A simple network was built in VISSIM to study the ability of the generic reliability metric to reflect the impact of capacity variations. As shown in figure 3.5, the simulation network consists of three links, each defined by its own physical length and speed limit. The upper threshold of the evaluation interval *I* for each link is defined as the travel time for a vehicle traveling at the speed limit. The lower threshold of *I* is 0 seconds. The length, speed limit, and evaluation threshold of each link are shown in table 3.3. The reliability metric used for this example will be based on the mean link travel time. This simulation example includes four assumptions: 1) for each ten-minute interval, the 95 percent confidence interval lies within 1.96 standard deviations of the mean link travel time; 2) the evaluation interval *I* represents the acceptable performance, that is, the satisfactory system is defined according to whether vehicles can travel at the speed limit or faster; 3) the OD demand is 2,500 vehicles per hour; and 4) the OD network is treated as a parallel system with three independent links, and the OD reliability is calculated through equation 3.4.

$$\mathbf{R}_{\bar{t}} = 1 - \prod_{j=1}^{J} \left[1 - R_{\bar{t}_j} \right] \tag{3.4}$$

where:

$\mathbf{R}_{\bar{t}}$ = the OD network reliability metric defined by the mean travel time $\bar{t}$.

$R_{\bar{t}_i}$ = the travel time reliability on link *i*, and

J = the number of independent links, *J*=3 in this example.
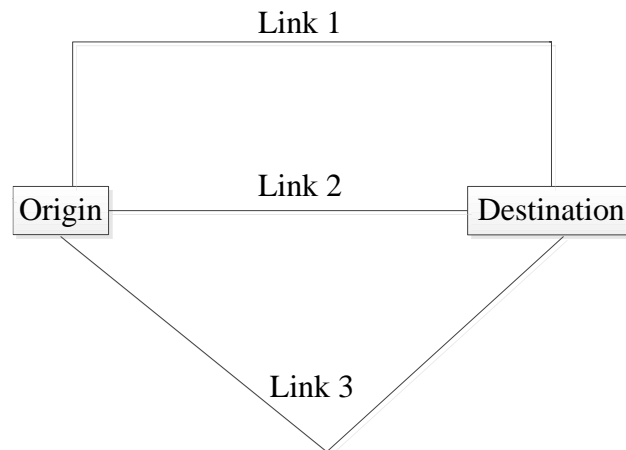
Figure 3.5 The simulation network

Based on ten simulation runs with different random seeds, the reliability metrics

of the three links and the OD network are calculated for two scenarios. Scenario 1 is a

network with all three links, as shown in figure 3.5. Scenario 2 is the same, with the

exception that link 2 is completely blocked. It can be seen that the loss of link 2 reduced

the link-level reliabilities for links 1 and 3 by 15 percent and 25 percent, respectively.

This means for scenario 1, links 1 and 3 can provide satisfactory performances 68 percent

and 90 percent of the time. If link 2 is blocked completely, the reliability metrics of links

1 and 3 can only provide satisfactory performances 58 percent and 68 percent,

respectively. On the OD level, the OD network reliability is reduced from 99 percent to

86 percent because of the loss of link 2.

Table 3.3 Reliability metric demonstration for capacity degradation scenarios

Reliability Metrics	Satisfactory Travel Time Threshold (s)	Length (m)	Speed Limit (km/h)	Reliability Metric Results		
				Complete Network	Broken Link 2	Degradation
Link 1 ($R_{\overline{t_1}}$)	36	500	50	68%	58%	15%
Link 2 ($R_{\overline{t_2}}$)	25	350	50	54%	-	-
Link 3 ($R_{\overline{t_3}}$)	35	390	40	90%	68%	25%
OD Network: $\mathbf{R}_{\overline{t}}=1-(1-R_{\overline{t_1}})(1-R_{\overline{t_2}})(1-R_{\overline{t_3}})$				99%	86%	12%

Note: Scenario 1: Links 1, 2, and 3 are all open.

   Scenario 2: Links 1 and 3 are open. Link 2 is closed.

3.3 Confidence Interval Estimation Methodology

The estimation methodology will be presented separately for a single link network (e.g., a basic freeway section) verses a multi-link traffic system (e.g., an urban arterial corridor composed of three links).

3.3.1 Evaluation of a Single Link Network: Freeway

3.3.1.1 Selection of Bootstrap Methods

The bootstrap methods, which can be used to estimate confidence intervals, may be classified as standard error (SE) based methods and percentile based methods (e.g., BCa method). As reviewed in section *2.3.1*, the main drawback of the SE-based methodology is its assumption of normality. Violation of this assumption will result in an unequal tail interval, which makes the estimates based on standard errors misleading. This limitation will be illustrated in this section using the I-495 data sample during the 5-

minute interval from 18:55 to 19:00 on Tuesday, Wednesday, and Thursday, from April to September in 2012. Both methods are used to generate confidence intervals at a 95 percent confidence level.

The confidence intervals of the mean speed are used to illustrate the results of the SE-based and BCa bootstrap methods. The histograms of 2,000 bootstrap mean replicates generated by the SE-based and BCa bootstrap methods are shown in the top and bottom plots in figure 3.6, respectively. A Lilliefors test validated that these estimated means of bootstrap samples are normally distributed. The red lines in the plots represent the confidence intervals of mean speeds in the unit of mph. The BCa and the SE-based methods yield very similar results: [43.0, 47.2] mph and [43.0, 47.1] mph, respectively. Both intervals include 95 percent of the mean replicates.

In contrast, the median speed is used as an example of the estimators that are not normally distributed. The histograms of 2,000 bootstrap median replicates generated by the two methods are shown in the top and bottom plots in figure 3.7, respectively. A Lilliefors test rejected the null hypothesis that the 2,000 bootstrap median replicates were from a normal distribution. The red lines in the top and bottom plots represent confidence intervals of median speeds in the unit of mph. The resulting confidence intervals using the BCa and the SE-based methods are [40.8, 52.0] mph and [36.5, 48.8] mph, respectively. As shown in the top plot in figure 3.7, the BCa method is able to work with non-normally distributed replicates with unequal tails as evidenced by the fact that its confidence interval includes 95 percent of the replicates. In contrast, the SE-based confidence interval in the bottom plot in figure 3.7 only includes 85 percent of all of the replicates. This dissertation aims to develop a generic framework to estimate the

reliability metric for system evaluation where some of the traffic measurement statistics (e.g., median speed) might not be normally distributed. Thus, the BCa method is selected for estimating the proposed generic reliability-based metric due to its capability to generate accurate results regardless of the estimator's distribution.
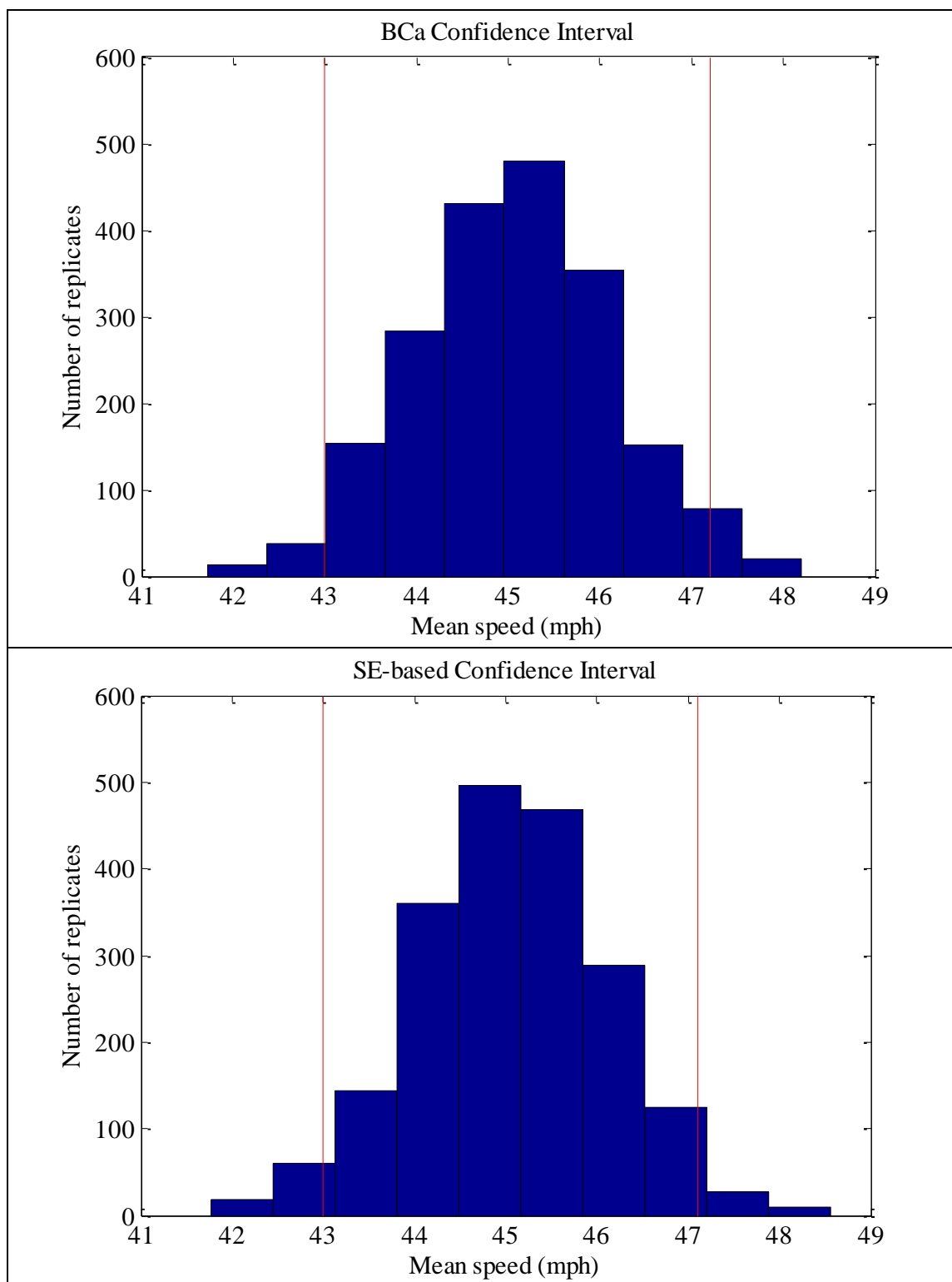
Figure 3.6 Histogram of 2,000 Bootstrap replicates of mean speed with two types of confidence intervals
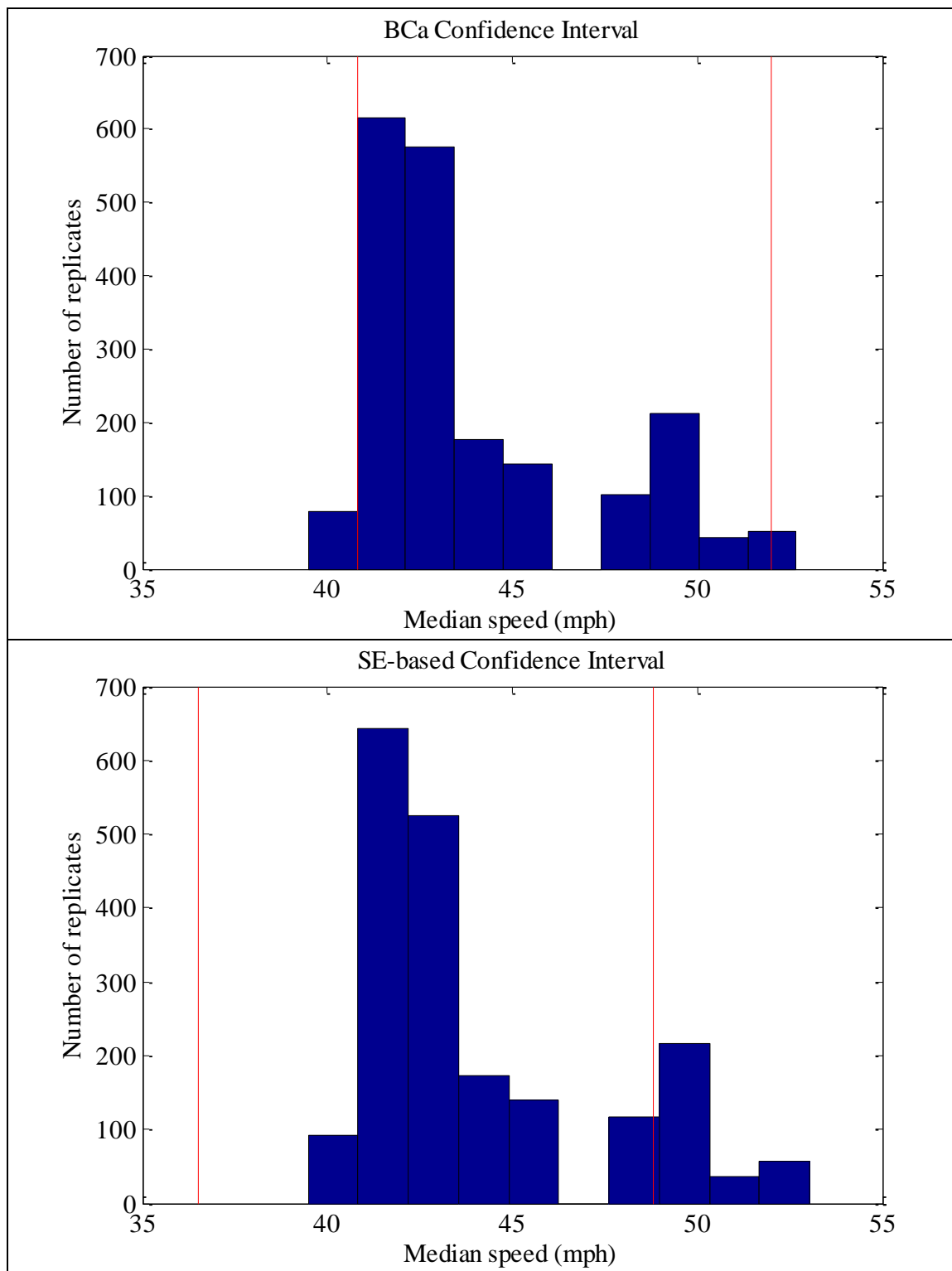
Figure 3.7 Histogram of 2,000 Bootstrap replicates of median speed with two types of confidence intervals

3.3.1.2 Evaluation of Level of Service (LOS)

Table 3.4 lists the HCM density threshold values for the LOS criteria for basic freeway segments. Using the density data collected on I-495 at 129 meters west of Stone-brook Drive, the LOS reliability metrics are calculated as shown in table 3.5. The density confidence intervals generated using the BCa method as a function of the time of day, together with the probabilities of the levels of service, are shown in figure 3.8.

Table 3.5 compares the traditional LOS evaluation based on mean estimates with the reliability-based LOS evaluation on an hour-by-hour level. Based on the traditional evaluation method, the facility provides LOS D for all three hours. The reliability-based metric indicates that there is a 3 percent chance that a vehicle will experience LOS E during the hour from 16:00 to 17:00. In contrast, the LOS was 100 percent LOS D from 17:00 to 18:00, and was 82 percent LOS D and 18 percent LOS C from 18:00 to 19:00. If travelers were aware that the later periods had higher reliability and they had flexibility in their departure time, they might adjust their departure time to avoid the worst hour of traffic.

Table 3.4 Level of service criteria for basic freeway segments

Level of Service	Density Range for Basic Freeway Sections (pc/mi/ln)
A	≤ 11
B	$> 11 \leq 18$
C	$> 18 \leq 26$
D	$> 26 \leq 35$
E	$> 35 \leq 45$
F	Demand Exceeds Capacity >45

Source: Roess, R et al., 2010

Table 3.5 Reliability-based LOS analysis results

Time Period	Mean Density (pc/mi/ln)	HCM LOS	Reliability-based LOS		
			C	D	E
16:00-17:00	33	D	0%	97%	3%
17:00-18:00	32	D	0%	100%	0%
18:00-19:00	28	D	18%	82%	0%
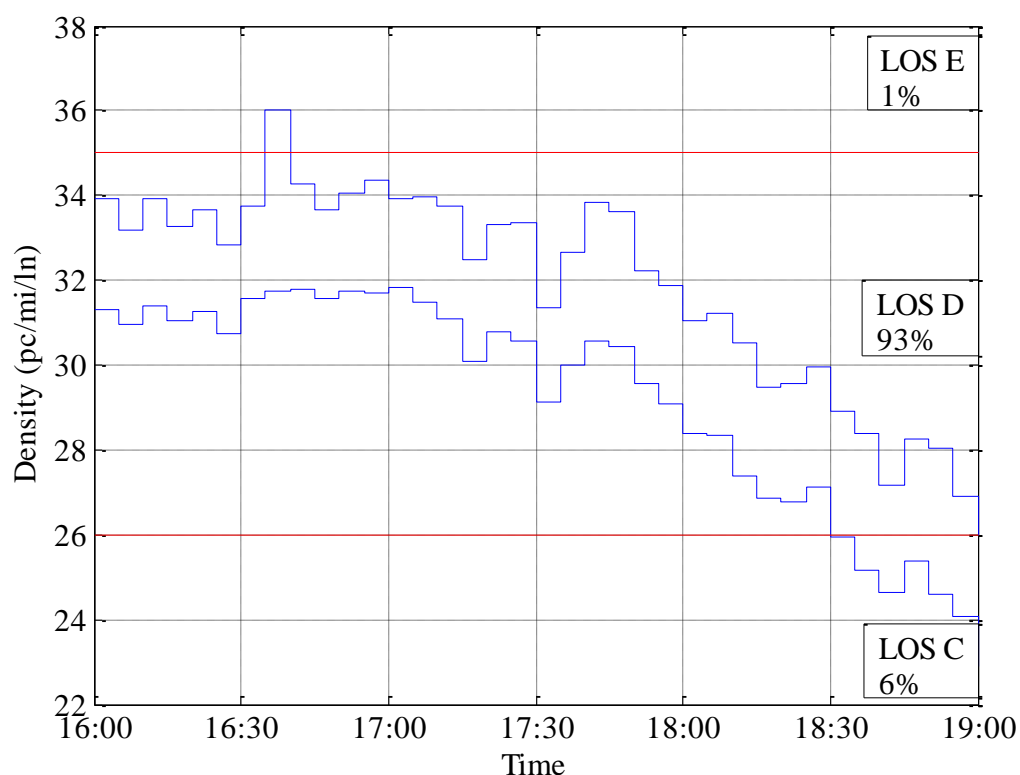
Figure 3.8 Five-minute confidence intervals of mean density

3.3.2 Evaluation of an Arterial Corridor

The level of service on an arterial corridor is generally related to the average travel speed on it. HCM uses an analytical method to calculate the delays on an urban street and to derive the average travel speed accordingly. In this section, the empirical

corridor travel time average and variance will be estimated based on link travel times collected on an arterial corridor. Given the estimated average corridor travel times, average travel speed is calculated, and subsequently, the reliability-based LOS is evaluated based on the confidence intervals of average travel speed.

3.3.2.1 Estimation Methodology

The most direct way to estimate average corridor travel time is to calculate it from direct measurements of probe vehicles traveling the corridor. In general, as the corridor distance increases, the number of vehicles traversing the corridor decreases. In addition, the time lag between the start of the vehicle journey and the time at which the travel time is reported will increase. In this situation, it might be advantageous to estimate the corridor time based on link measurements. The simplest way to do so is to add the link travel time means and variances, referred to as the naïve method by Eisele (2001). This naïve method can yield erroneous results if its assumption of independence between link travel times is violated. In dynamic traffic systems, the travel time on a roadway link is very likely to be dependent on the travel time on the preceding link, which will impact the arrival time of the current link, and in turn impact the travel time on the current link. For example, vehicles arriving in peak hours will experience higher travel times than those arriving in non-peak periods. That is, although the travel times on individual links can be regarded as statistically independent at a particular point in time, the travel times on adjacent links are often correlated. This correlation is referred to as "time-of-day correlation" (Fu and Rilett 1998). This section will introduce different methods used in section 3.3.2.2 to estimate the corridor travel time average and variance.

Approximation methods for mean and variance estimation. The link travel times in traffic networks are assumed as random variables with probability distributions dependent on the time of day. That is, the link travel times are modeled as a continuous-time stochastic process denoted as $\{X_a(t), t \in T\}$, where $X_a(t)$ is the travel time for vehicles entering the link a at time t, and T is a continuous parameter set. For each time instance t, the first-order probability density function (PDF) of $X_a(t)$ is denoted as $f_{X_a}(x_a, t)$. In this study, T is the time range of interest, (i.e., peak hours) although it could be $[0, \infty)$ theoretically.

Travel times on individual links at a particular point in time are assumed to be statistically independent (Fu and Rilett 1998). Because the individual link travel time is modeled as a time-dependent process, its dependence on the preceding link is accounted explicitly through the arrival time at the current link, which is related to the travel times on the preceding link. This correlation is referred to as time-of-day correlation and is modeled directly within the individual link's stochastic process. This way of addressing correlation avoided the more comprehensive data analysis incurred by considering the correlation between individual link travel times at a particular moment in time.

As shown in figure 3.9, the route example extracted from Fu and Rilett (1998) is used to illustrate the approximation formulas. After a vehicle starts off from the origin node, its arrival time at the next node in a given route depends on the link travel time. Additionally, its arrival time at a destination depends on all the link travel times and the departure time from the origin nodes. This process can be represented by a recursive equation (see equation 3.5). Let the random variable Y_i denote the arrival time at node i.

The departure time at the origin node 1, Y_1, is assumed to be deterministic and a known a priori. The PDF of Y_i is represented by $f_{Y_i}(y_i)$. The problem of estimating route travel time variability is, therefore, to estimate $f_{Y_i}(y_i)(i = N)$ given a departure time at node 1 and travel times on individual links. However, it is complex and computationally infeasible to estimate the $f_{Y_i}(y_i)$ when the links' travel times are both dynamic and stochastic (Fu and Rilett 1998). The first- and second-order approximation methods can approximate the route travel time mean and variance based on the first two moments of the link travel time PDF.
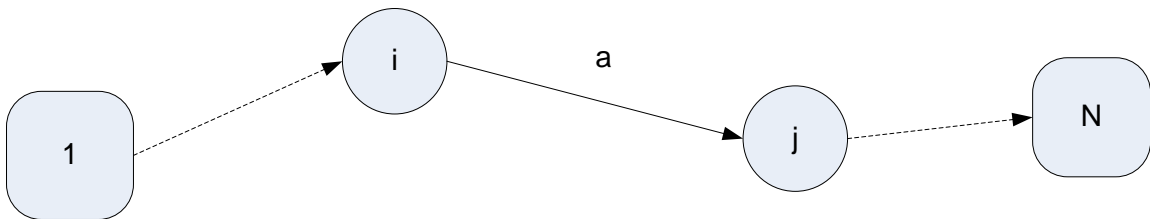
Source: Fig.1, Fu and Rilett (1998).

Figure 3.9 A route *p* from origin node 1 to destination node *N*

$$Y_j = Y_i + Z_a \tag{3.5}$$

where:

Y_j = the arrival time at link *j,*

Y_i = the arrival time at link *i,* and

Z_a = the travel time on link *a*.

The random variable Z_a is the travel time on link a under a given experiment. The parameters of the experiment include both the route and the departure time. The distribution of Z_a is conditional to the specific experiment, and it is different from the link travel time represented solely by a function of the time entering the link. The probability distribution of Z_a under a given time (i.e.,$Y_i = y_i$) will be the same as the probability distribution of $X_a(y_i)$, as shown in equation 3.6, where P is the link set for route p. In addition, $Z_a|Y_i = y_i$ also has the same mean and variance as the random variable $X_a(y_i)$, as shown in equations 3.7-3.8, where $\mu_{X_a}(y_i)$ and $v_{X_a}(y_i)$ need to be estimated through statistical methods.

$$P\{Z_a < x|Y_i = y_i\} = P\{X_a(y_i) < x\}, x \in P \tag{3.6}$$

$$E[Z_a|Y_i = y_i] = \mu_{X_a}(y_i) \tag{3.7}$$

$$VAR[Z_a|Y_i = y_i] = v_{X_a}(y_i) \tag{3.8}$$

where:

$\mu_{X_a}(y_i)$ = the distribution of the travel time mean given an arrival time, and

$v_{X_a}(y_i)$ = the distribution of the travel time variance given an arrival time.

To estimate the mean arrival time of node j, the recursive equation 3.5 can be further transformed into equation 3.9, where the second term is defined in equation 3.10. The PDF of arrival time $f_{Y_i}(y_i)$, however, is unavailable in the traffic database. What is available are the historical sample means and sample variances, or the forecasted means and variances for discrete periods throughout the day. To determine the second term in

equation 3.9 without $f_{Y_i}(y_i)$, the first- and second-order approximation methods expand $\mu_{X_a}(t)$ as a Taylor series around the point $t = E[Y_i]$, as shown in equation 3.11.

$$E[Y_j] = E[Y_i] + E[Z_a] = E[Y_i] + E[E[Z_a|Y_i] = E[Y_i] + E[\mu_{X_a}(Y_i)] \tag{3.9}$$

$$E[\mu_{X_a}(Y_i)] = \int_0^{+\infty} \mu_{X_a}(y_i) f_{Y_i}(y_i) dy_i \tag{3.10}$$

$$\mu_{X_a}(t) = \mu_{X_a}(E[Y_i]) + \mu'_{X_a} \cdot (t - E[Y_i]) + \frac{1}{2!}\mu''_{X_a} \cdot (t - E[Y_i])^2 + \cdots \tag{3.11}$$

$$E[\mu_{X_a}(Y_i)] \cong \int_0^{+\infty} \{\mu_{X_a}(E[Y_i]) + \mu'_{X_a} \cdot (y_i - E[Y_i])\} f_{Y_i}(y_i) dy_i$$
$$= \mu_{X_a}(E[Y_i]) \int_0^{+\infty} f_{Y_i}(y_i) dy_i + 0 = \mu_{X_a}(E[Y_i]) \tag{3.12}$$

As displayed in equation 3.12, the first-order approximation method estimates $E[\mu_{X_a}(Y_i)]$ by truncating the Taylor series of $\mu_{X_a}(t)$ at the linear term, assuming that the second and higher order derivatives are equal to zero. The first-order approximation model of the recursive formula is shown in equation 3.13.

$$E[Y_j] \cong E[Y_i] + \mu_{X_a}(E[Y_i]) \tag{3.13}$$

Furthermore, assuming the third and higher order derivatives of the Taylor series are zero, the second-order approximation model of the mean arrival time can be obtained, as shown in equation 3.14.

$$E[Y_j] \cong E[Y_i] + \mu_{X_a}(E[Y_i]) + \frac{1}{2!}\mu''_{X_a}(E[Y_i]) \cdot Var\,[Y_i] \tag{3.14}$$

In this dissertation, the link travel time function $\mu_{X_a}(t)$ is approximated in the

form of a second-order polynomial, as shown in figure 3.10.
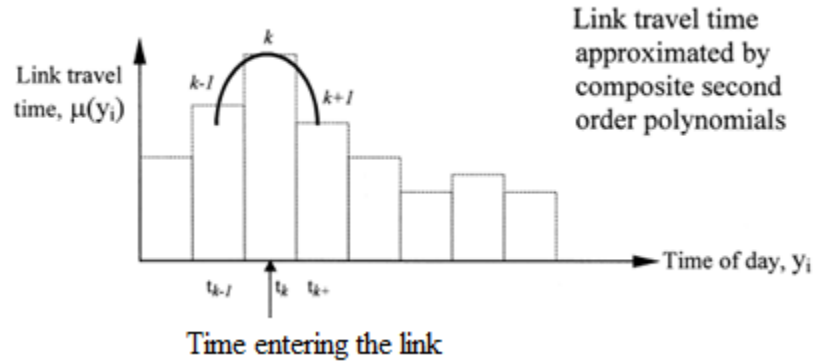
Source: Fig.5(d), Fu and Rilett (1998).

Figure 3.10 Time-dependent second-order link travel time functions

As shown in equation 3.14, the variance of arrival time at node i, $Var\,[Y_i]$, is

required to apply the second-order approximation method. Based on the recursive

equation 3.5, the variance of arrival time at node j, $Var\,[Y_j]$ can be written as:

$$Var\,[Y_j] = Var\,[Y_i] + Var[Z_a] + 2COV(Y_i, Z_a) \tag{3.15}$$

The last two terms of equation 3.15 can be transformed through equations 3.16

and 3.17 (Ross 1989).

$$Var[Z_a] = E[Var[Z_a|Y_i]] + Var[E[Z_a|Y_i]]$$
$$= E[Var[X_a(Y_i)]] + Var[E[X_a(Y_i)]$$
$$= E[v_{X_a}(Y_i)] + Var[\mu_{X_a}(Y_i)] \tag{3.16}$$

$$COV(Y_i, Z_a) = E[Y_i \cdot Z_a] - E[Y_i]E[Z_a]$$
$$= E[E[Y_i \cdot Z_a|Y_i]] - E[Y_i][E[Z_a|Y_i]]$$
$$= E[Y_i \cdot E[Z_a|Y_i]] - E[Y_i] E[\mu_{X_a}(Y_i)] \tag{3.17}$$
$$= E[Y_i \cdot \mu_{X_a}(Y_i)] - E[Y_i] E[\mu_{X_a}(Y_i)]$$

Thus, the variance of the arrival time at node j can be calculated by equation 3.18.

$$Var[Y_j] = Var[Y_i] + E[v_{X_a}(Y_i)] + Var[\mu_{X_a}(Y_i)] - 2E[Y_i \cdot \mu_{X_a}(Y_i)]$$
$$- 2E[Y_i] E[\mu_{X_a}(Y_i)] \tag{3.18}$$

Again, it is mathematically impractical to derive the functions $v_{X_a}(Y_i)$ and $\mu_{X_a}(Y_i)$. The first- and second-order approximation models for the variance of the arrival time at node j ($Var[Y_j]$) can be derived by replacing functions with truncated Taylor series expansions about point $E[Y_i]$.

By assuming that the second and higher derivatives of $v_{X_a}(Y_i)$ and $\mu_{X_a}(Y_i)$ are equal to zero, the first-order approximation method is shown in equation 3.19.

$$Var\left[Y_j\right] \cong A \cdot Var\left[Y_i\right] + \nu_{X_a}(E[Y_i]) \tag{3.19a}$$

$$A \cong \{1 + \mu'_{X_a}(E[Y_i])\}^2 \tag{3.19b}$$

By assuming that the third and higher derivatives of $\nu_{X_a}(Y_i)$ and $\mu_{X_a}(Y_i)$ are equal to zero, meaning the arrival time is symmetric and neither platykuric nor leptokuric, a simplified second-order approximation model for normally distributed arrival time is shown by equation 3.20.

$$Var\left[Y_j\right] \cong (A + B) \cdot Var\left[Y_i\right] + \nu_{X_a}(E[Y_i]) \tag{3.20a}$$

$$B \cong \frac{1}{2}\{\nu''_{X_a}(E[Y_i]) + \mu''^2_{X_a}(E[Y_i]) \cdot Var\left[Y_i\right]\} \tag{3.20b}$$

Similar to the mean estimation, second-order polynomial functions are used to approximate $\nu_{X_a}(Y_i)$ on link *a* for first- and second-order approximation methods.

Naïve sum method for mean and variance method. As a baseline to compare the efficiency of the first- and second-order approximation methods, the mean and variance of arrival time will also be estimated using the naïve method. With the independence assumption, the naïve method uses equations 3.21 and 3.22.

$$E\left[Y_j\right] = E[Y_i] + E[Z_a] \tag{3.21}$$

$$Var\left[Y_j\right] = Var\left[Y_i\right] + Var[Z_a] + 2COV(Y_i, Z_a) \cong Var\left[Y_i\right] + Var[Z_a] \tag{3.22}$$

Covariance-based method for variance estimation. The arrival travel time variance can also be estimated through equation 3.15.

$$Var\ [Y_j] = \ Var\ [Y_i] + Var[Z_a] + 2COV(Y_i, Z_a) \tag{3.15}$$

Cumulative sum method for mean estimation. The expected arrival time could be estimated by equation 3.23, derived from 3.10, where $\mu_{X_a}(.)$ is estimated by a step function, as shown in figure 3.11.

$$E[Y_j] \cong E[Y_i] + \mu_{X_a}(E[y_i]) = E[Y_i] + E[X_a|y_i] \tag{3.23}$$
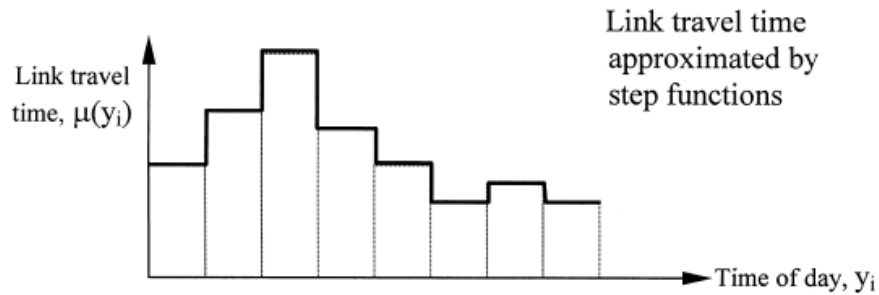
Source: Fig.5(b), Fu and Rilett (1998)

Figure 3.11 Time-dependent step link travel time functions

3.3.2.2 Estimation Methods Comparison

The test bed used to study the accuracies of different estimation methods is a three-link corridor on Westheimer Road, an east-west arterial in Houston, Texas. The corridor starts at Wilcrest Drive and ends at Eldridge Parkway. As shown in figure 3.12, the yellow pins indicate the locations of MAC readers that collect Bluetooth information. Once the location and time stamp information is obtained, travel time can be readily calculated.

Figure 3.12 The three-link corridor test bed

The corridor travel time data was collected for afternoon peak hours (16:00-19:00) from January to July in 2011. For each 15-minute interval, the true mean and variance of the arrival time (AT) at the Eldridge Parkway are calculated using the observed arrival travel times in the dataset. The performances of various estimation methods discussed in section 3.3.2.1 are evaluated by absolute percentage errors (APEs), which are calculated by equation 3.24.

$$APE = \frac{|Estimate\ Value - Observed\ Value|}{Observed\ Value} \times 100\% \qquad (3.24)$$

Table 3.6 summarizes the mean and median of APEs for mean estimation and variance estimation, respectively. For mean estimation, the naïve sum method results in a mean APE of 2.6 percent. The second-order approximation method yielded a mean APE of 1.9 percent and improved the estimation accuracy by 27 percent compared to the naïve sum method.

Because variance estimation could be significantly biased for intervals that have very few observations, the median APE is used to evaluate the overall performance of estimation methodologies. Also shown in table 3.6, the second-order approximation method improved the accuracy of variance estimation by 15 percent compared to the naïve sum method.

Table 3.6 Performance of various estimation methods for corridor travel time metrics

Estimation Methods for Mean Corridor Travel Time	Naïve Sum	Cumulative Sum	First-order Approximation	Second-order Approximation
Mean APE of Mean Estimation (%)	2.6	2.5	2.0	1.9
Improvement (%)	Baseline	4	23	27
Estimation Methods for Arrival Time Variance	Naïve Sum	COV-based	First-order Approximation	Second-order Approximation
Median APE of Variance Estimation (%)	7.2	6.8	6.1	6.1
Improvement (%)	Baseline	6	15	15

Figure 3.13 shows the cumulative density function plots of APEs for mean and variance estimations. Using this figure, traffic agencies can decide which method to use

according to their desired level of accuracy for their application. Assuming the traffic manager aims to estimate the corridor travel time mean with an APE of less than 5 percent, the approximation method can yield an acceptable estimation 93 percent of the time, while the naïve sum method yields an acceptable estimate 85 percent of the time. Assuming the traffic manager prefers a variance estimation with an APE less than 20 percent, the probability of obtaining satisfactory results are 90 percent and 82 percent by using the approximation method and the naïve sum method, respectively.
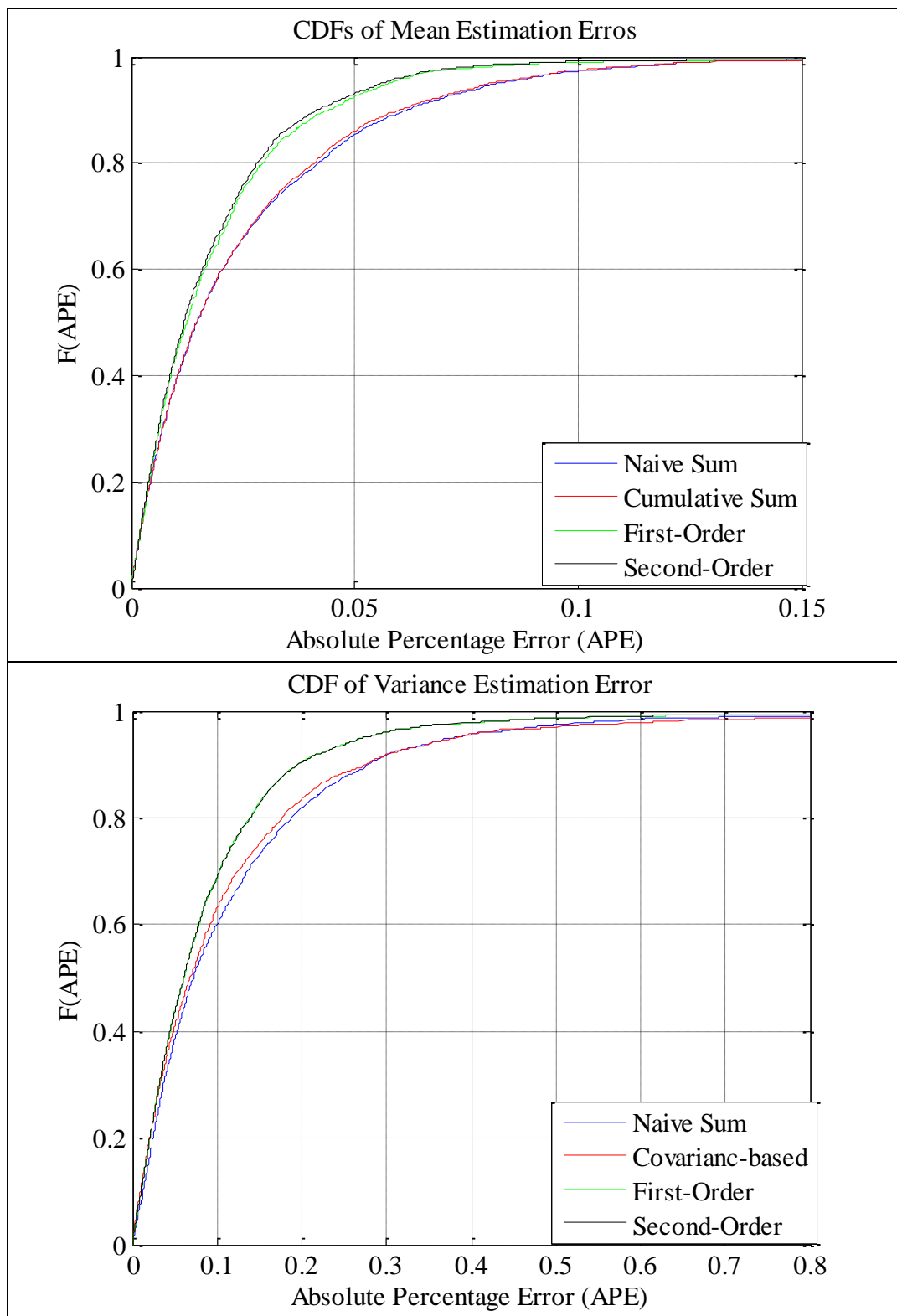
Figure 3.13 CDFs of estimation APEs

3.3.2.3 Evaluation of Level of Service

The estimation results of the second-order approximation method are used in this section to continue with the corridor evaluation. The estimated mean arrival times are converted to corridor travel times and then to average travel speeds. The confidence intervals of average speed are generated using the BCa bootstrapping method. The test corridor is defined as a Class II facility, and the level of service evaluation thresholds use the HCM 2000 standard shown in figure 3.14. The evaluation results of this corridor, based on seven months of data collection, are plotted in figure 3.15. It may be seen that the LOS is either B or C during the analysis period. The lower LOS (e.g., C) typically starts around 16:45 and ends around 18:45. During the whole analysis period from 16:00 to 19:00, the reliability of this corridor is 22 percent and 78 percent for LOS B and C, respectively.

Urban Street Class	I	II	III	IV
Range of free-flow speeds (FFS)	55 to 45 mi/h	45 to 35 mi/h	35 to 30 mi/h	35 to 25 mi/h
Typical FFS	50 mi/h	40 mi/h	35 mi/h	30 mi/h
LOS	Average Travel Speed (mi/h)			
A	> 42	> 35	> 30	> 25
B	> 34-42	> 28-35	> 24-30	> 19-25
C	> 27-34	> 22-28	> 18-24	> 13-19
D	> 21-27	> 17-22	> 14-18	> 9-13
E	> 16-21	> 13-17	> 10-14	> 7-9
F	≤ 16	≤ 13	≤ 10	≤ 7

Source: HCM 2000, Exhibit 15-2

Figure 3.14 HCM 2000 urban street LOS criteria
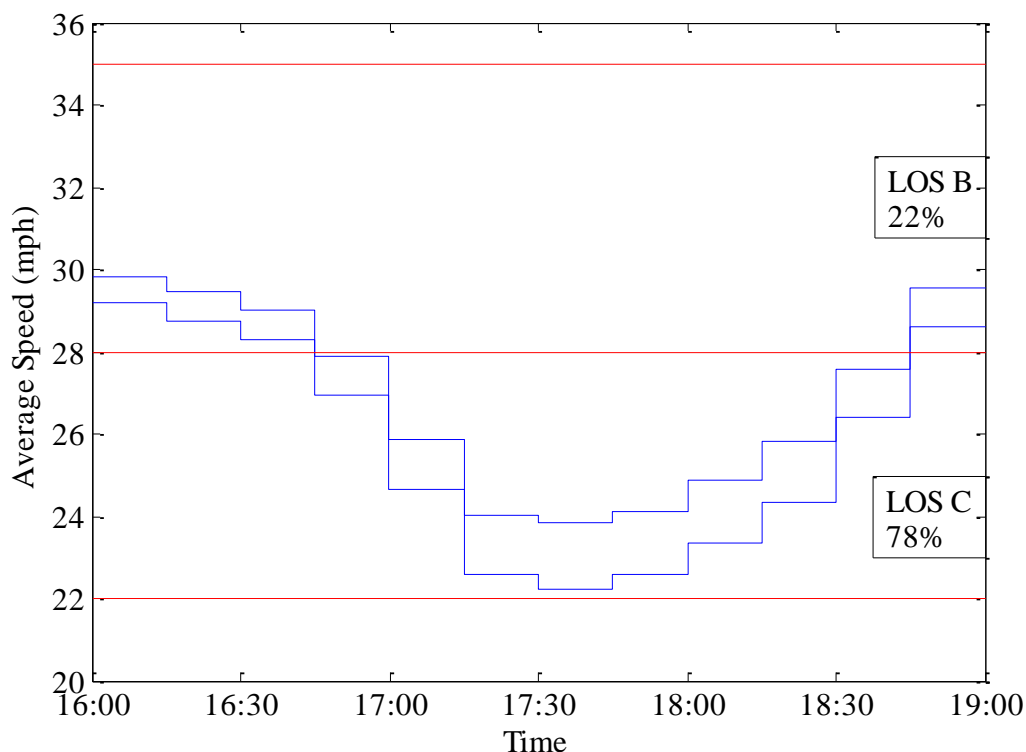
Figure 3.15 15-minute confidence intervals of mean travel speed

3.4 Concluding Remarks

This chapter presented the new reliability metric for the evaluation of traffic system performance. The metric is able to evaluate traffic system reliability in terms of a probability of obtaining a certain level of service within the period of interest. Time-dependent confidence intervals are incorporated to take into account the time-dependent fluctuations in traffic flow.

The travel time data for one I-495 segment were used to study the variability of travel time distributions within a half hour. Based on the six-month data, it was shown using a paired t-test that the mean travel time for the 5-minute interval from 16:00 to

16:05 was different from the mean for the 5-minute interval between 16:10 and 16:15 at a 95 percent confidence level. These results indicate that system performance needs to be monitored and evaluated at a scale finer than hourly if the user wishes for more detailed information on what is happening on the network.

To capture the fluctuations in performance, each short-time interval was analyzed separately using a new reliability-based level of service metric. The BCa bootstrap method was applied to obtain a confidence interval based on percentiles instead of standard error estimates. This approach was chosen so that confidence intervals could be obtained even for those datasets that are not normally distributed. The new metric was applied for a segment of I-495, and the traffic parameter of interest was density. Compared to the traditional LOS evaluation scheme based on an average density, the reliability-based results can provide more detail about how the system performs within a given period of interest. This would be helpful when the performance evaluation results are used to compare competing designs or management strategies.

In terms of applications on corridors that consist of multiple links, it is often difficult to obtain enough measurements of corridor travel times. This problem is particularly notable when the corridor under evaluation is very long. Thus, corridor travel time information such as mean and variance need to be first estimated from link travel times. A naïve way to do this is to sum the means and variances of travel times on all the links together by assuming that link travel times are independent. To relieve this unrealistic assumption, the first- and second-order approximation methods estimate the mean and variance of corridor travel time taking into account the correlations among

links. Bluetooth travel time data on a five-kilometer corridor of Westheimer Road in Houston, Texas, were used to compare the efficiency of the approximation methods with simple statistic methods. This is the first attempt to validate these approximation methods using empirical corridor travel times obtained using Bluetooth technology. The results justified the advantage of the two approximation methods for this three-link corridor. It is reasonable to conclude that for applications on longer corridors/paths, the accumulation of errors in the naïve method due to neglecting link-level correlations will be greater, and the advantage of the approximation methods will be even more apparent. It is also hypothesized that this approach would be better than the naïve method for real time application, which will be analyzed in chapter 4.

CHAPTER 4 DEVELOPMENT OF RELIABILITY INDICATOR AND PREDICTION

METHODOLOGY FOR TRAVELERS

4.1 Necessity of a Reliability Indicator for Traveler Information Systems

Most transportation systems are dynamic and stochastic in that travel times vary across space and time. This mainly arises because of: 1) recurring congestion such as the rush hour period, 2) operational treatments for unexpected disruptions (e.g., traffic signal preemption for emergency vehicles and highway railway at-grade crossings), and 3) traffic control devices and different roadway characteristics. For example, urban streets with traffic signals and conflicting cross street traffic introduce more variability in travel times than freeways and access-controlled highways, all else being equal. In traffic systems experiencing congestion, the provision of travel time reliability information is gaining importance among researchers, traffic operators, and drivers (Haitham and Emam 2006; Tu 2008; Barkley et al. 2012; Carrion and Levinson 2013).

Most current applications of Advanced Traveler Information Systems (ATIS) provide average estimates of the traffic parameters of interest through dynamic message signs and/or websites. Table 4.1 lists examples of dynamic message signs in terms of travel time information for freeway and arterial facilities. Examples (1) and (2) in table 4.1 display only instantaneous average travel times. In example (3) in table 4.1, a two-minute travel time interval is typically displayed during non-congested periods. In heavy congestion periods, an interval up to four minutes is applied to reflect a higher degree of uncertainty in expected travel times for travelers (Oregon Department of Transportation

2005). The use of time intervals was based on the experience of the traffic system operators.

Figure 4.1 is a snapshot of the online Houston Transtar Traffic Map where various colors indicate different traffic conditions in terms of average speeds calculated from travel times. The Houston Transtar uses Bluetooth technology to collect travel time information and uses the average of the link travel time measurements during the last 5 minutes as the current travel time information. The travel time information for a freeway corridor consisting of multiple links is based on the summation of the mean link travel times during the last 5 minutes (personal communition, TxDOT 2013). This summation approach assumes that link travel times are independent and that changes in route travel time over the period of interest are minimal. Intuitively, these assumptions are not reasonable for traffic systems with unstable traffic conditions.

The travel time at a route level is subject to variability due to the traffic control devices and different traffic conditions along the route. The level of variability increases as the length of the route inceases. Drivers who are interested in the route travel time can only get limited information from the average travel time estimate. In this case, a range within which they can expect individual arrival times with a certain level of confidence can provide them with a better idea of traffic information on the arterial corriodrs and/or the route. In fact, it has been found that travel time reliability is an important criterion in route choices and trip planning (Abdel-Aty et al. 1995; Carrion and Levinson 2013). This chapter represents travel time reliability with a reliability interval of arrival time. In contrast to the confidence interval, which by definition includes the population mean with

a certain statistical confidence, the reliability interval is defined to include a individual observation with certain statistical confidence.

This chapter will study several prediction models for the short-term corridor travel time (CTT) mean, standard deviation (SD), and associated reliability interval. The methodology can also be extended for longer and more diverse paths.

Table 4.1 Dynamic message signs (DMS) of travel time information

(1) DMS of freeway travel time in New Jersey.	
(2) Demonstration of DMS of arterial travel time on W Sand Lake Rd in Orlando by Post Oak Traffic Systems, Inc.	
(3) DMS of freeway travel time interval in Oregon.	

2:20pm 5/24/2013
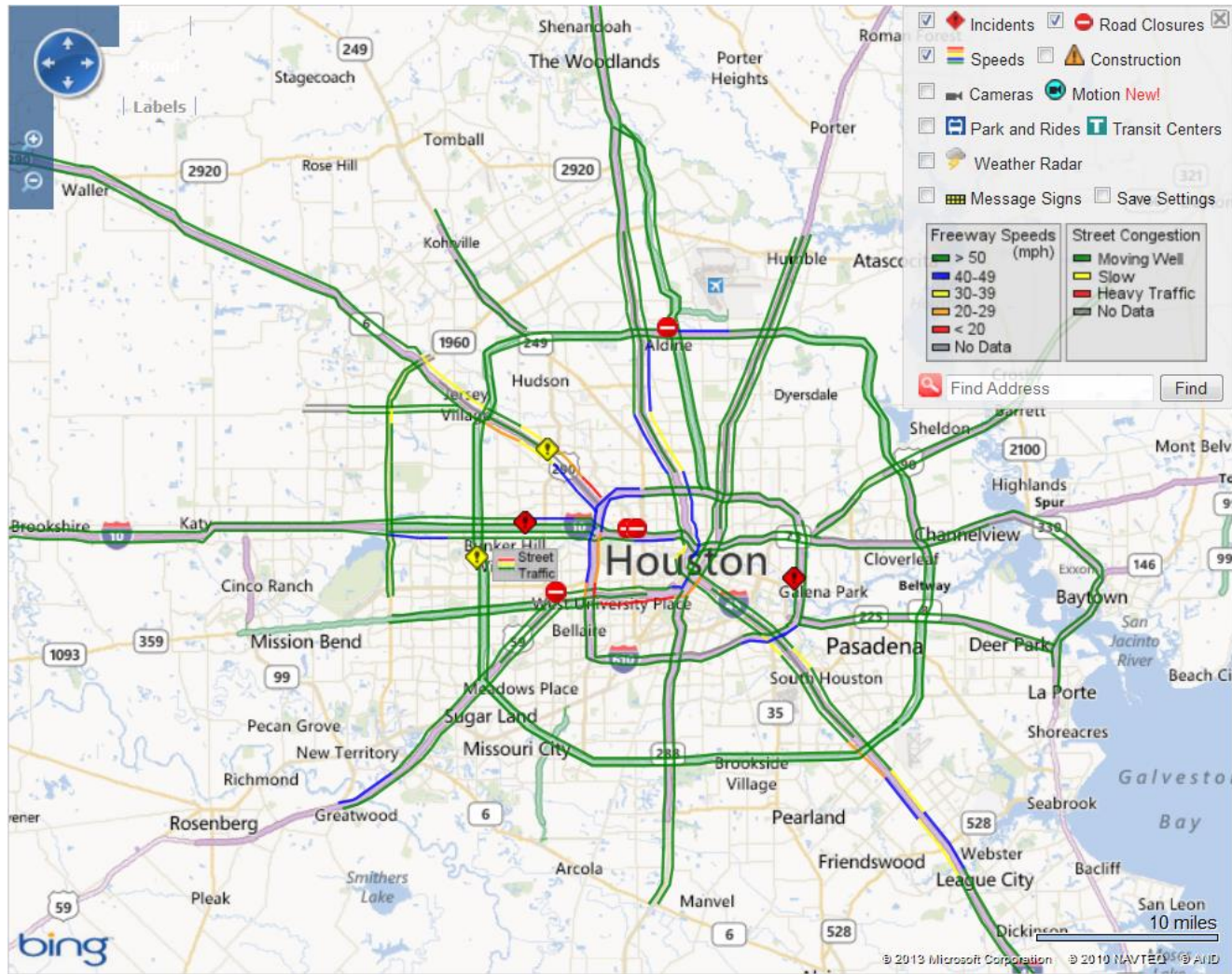
Figure 4.1 A snapshot of the Houston TranStar Traffic Map

4.2 Definition of Reliability Interval (RI)

The reliability interval of travel time R_{ijn}^{α} indicates the travel time variability between origin i and destination j, with the confidence level α for drivers departing during the n-th time interval. The reliability interval is predicted based on current and/or historic travel time measurements. Variable α is the percentage of drivers who depart in n-th time interval can expect to finish their trips within a RI. The reliability interval of travel time is defined as the expected travel time bounded by k^{α} times the predicted standard deviation of travel time, as shown in equation 4.1.

The reliability interval of the arrival time for the k-th driver between i and j, departing at d_k time, depends on his/her departure time d_k and the associated travel time reliability interval for the n-th time interval, as indicated by equation 4.2.

$$R_{ijn}^{\alpha} = \bar{x}_{ijn} \pm k^{\alpha} \cdot s_{ijn} \quad \forall i = 1 \ldots I, \forall j = 1 \ldots J, \forall n = 1 \ldots N \qquad (4.1)$$

$$A_{ijkn}^{\alpha} = d_k + R_{ijn}^{\alpha} \quad \forall i = 1 \ldots I, \forall j = 1 \ldots J, \forall n = 1 \ldots N, \forall k = 1 \ldots K \qquad (4.2)$$

where:

R_{ijn}^{α} = the predicted reliability interval of travel time between nodes i and j, for departure time within the n-th 15-minute interval, at the confidence level of α. The roadway between i and j could be a link, a corridor, or a path. In this case study, it is a corridor with three links.

k^{α} = the coverage parameter. This parameter is user defined and controls the spread of the interval.

$\bar{x}_{ijn}$ = the predicted mean travel time from i to j for vehicles departing during n-th time interval t_n.

s_{ijn} = the predicted standard deviation of travel time from i to j for vehicles departing during n-th time interval.

I, J = the number of origins and destinations, respectively.

N = the total number of 15-minute intervals within the analysis period. For a PM peak period from 4:00pm to 7:00pm, N is equal to 12.

A^{α}_{ijkn} = the predicted reliability interval of the arrival time for k-th driver between i and j, departing from i at d_k, which is in the n-th 15-minute interval.

K = the number of drivers in the dataset who departed in the n-th time interval.

α = the confidence level, which is the percentage of drivers who depart during n-th time interval can expect to arrive at j within A^{α}_{ijk}.

One key step to implementing the reliability interval in practice is to choose an appropriate value of k^{α}. Because the reliability interval is defined by the predicted sample mean and standard deviation, the coverage rate of a RI corresponding to a given k^{α} depends on the efficiency of prediction models. The coverage rate is defined as the percentage of the actual arrival times that are within the predicted reliability interval. Therefore, the value of k^{α} needs to be decided by traffic agencies through a preliminary study based on an analysis of historic travel time datasets and a pre-selected prediction model. Intuitively, the value will have to be updated on a regular basis given new available information. Because the instantaneous travel time prediction method (ITT) is

applied for freeway corridor mean travel time estimation, the k^α used in this dissertation

is set to the value with which the predicted reliability intervals, using ITT, includes actual

arrival times at the percentage of α or higher. Note that the coverage rate for a given k^α

can be used as an evaluation metric for different prediction methods. For example, the

higher the coverage rate of the predicted reliability intervals, the better the prediction

model is, all else being equal.

In addition to the coverage rate, the average interval range is the other indicator of

the efficiency of different prediction models. In general, there is a tradeoff between the

coverage rate and the interval range of a given reliability interval. While the coverage

rate indicates the ability of a prediction model to generate accurate reliability intervals for

drivers, the interval range represents the usability of the predicted reliability intervals. For

example, in a case study where the average corridor travel time is 442 seconds, a

reliability interval ranging from 1 to 1,000 seconds can provide a very high coverage rate,

but it would prove useless for drivers. Therefore, although a high coverage rate is desired,

drivers would also like to have reliability intervals as compact as possible. Both the

coverage rate and the average interval range need to be considered when identifying the

"best" prediction model. This issue will be discussed in the example problem later in this

chapter.

4.3 Case Study

An arterial corridor is used to demonstrate the calculation and prediction of arrival

time reliability intervals defined by equation 4.2. The test bed is a three-link corridor

along Westheimer Road in Houston, Texas. Houston TranStar and its partners have implemented an anonymous wireless address matching (AWAM) system to measure travel times along arterial roadways. The AWAM system is able to detect vehicles with Bluetooth networking devices such as cellular phones, mobile GPS systems, etc. The unique electronic address of each enabled Bluetooth device, known as a MAC address, can be detected by the roadside AWAM readers as the device passes the reader station. The AWAM readers then transmit the time and location of the detected device back to the AWAM host processing system. Subsequently, the travel time readings are matched and the individual link travel times are derived. These are used to calculate the average travel time and speed information which are then provided to drivers in real time (Puckett and Vickich 2010). Figure 4.2 is a snapshot of the on-line Houston TranStar Traffic map showing the three-link corridor in this case study, which is monitored by the AWAM system. Figure 4.3 shows the traffic information available from the website including the road name, cross street 1, cross street 2, distance, average travel time, and average speed.

In this section, the exploration of the Bluetooth travel time data will be expanded by predicting an arrival time reliability interval of the corridor trip for the next 15-minute interval. The three-link corridor, as marked by the four red circles in figure 4.2, is used as a test bed to compare efficiencies of various prediction methodologies. Each circle is a data collection station, and the links between adjacent circles are approximately 1.6 km in length, as shown in figure 4.2. The developer stated that this distance would provide an acceptable number of vehicle observations (Puckett and Vickich 2010). Each link includes multiple signalized intersections. The detailed information for each link is

summarized in table 4.2. Note that the number of traffic signals on each link were

obtained from Google map (2014).

Table 4.2 Link information of the corridor test bed in study

Cross street 1 (from)	Wilcrest	Kirkwood	Dairy Ashford
Cross street 2 (to)	Kirkwood	Dairy Ashford	Eldridge
Distance in km (mile)	1.6 (1.0)	1.6 (1.0)	1.8 (1.1)
Number of signals	2	2	3
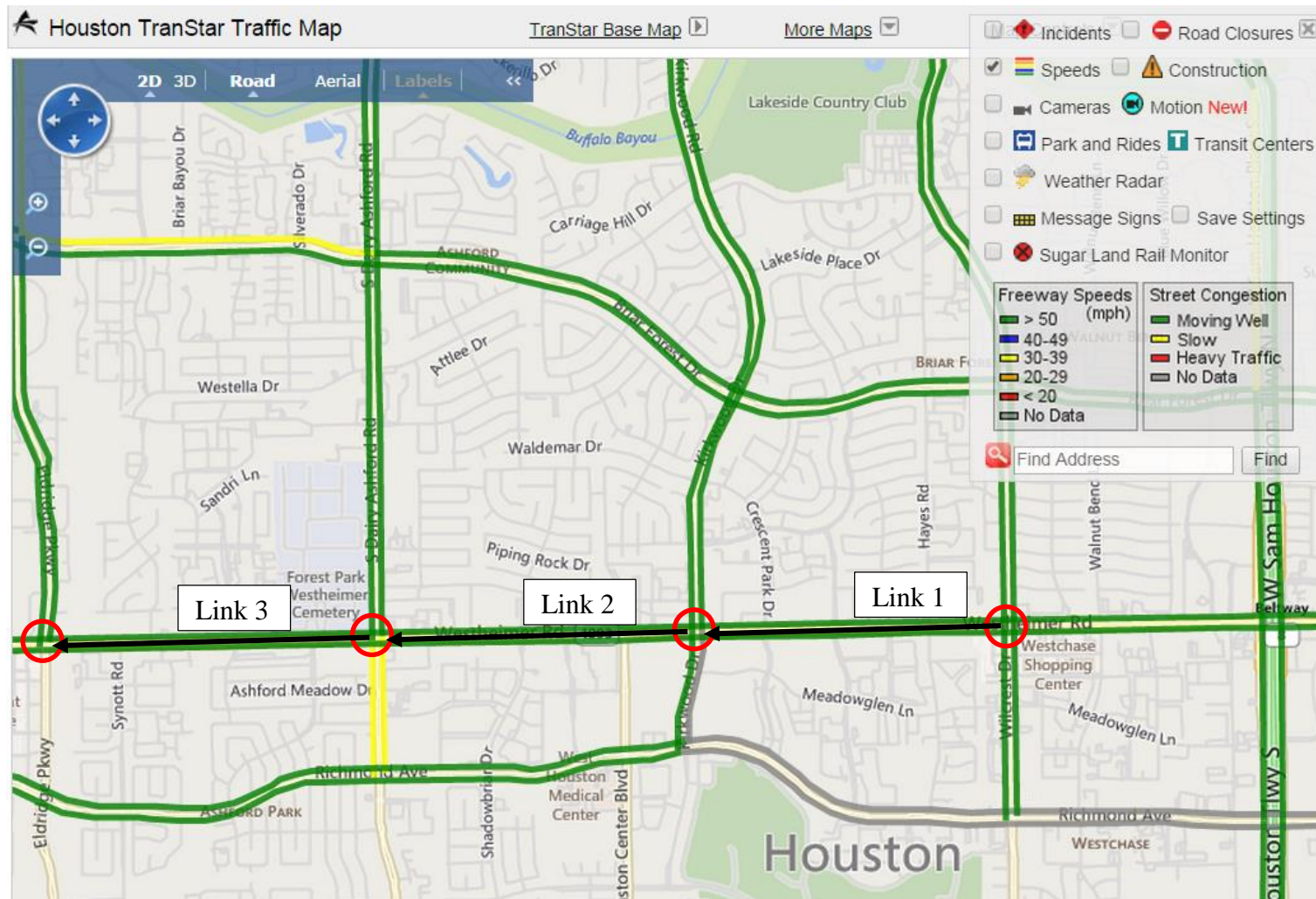
Figure 4.2 The test bed of a three-link arterial corridor

Link 1

Roadway Speed and Travel Time	
Westheimer Westbound from Wilcrest to Kirkwood	
Distance	1 miles
Travel Time	1 minute 41 seconds
Average Speed	36 MPH

Link 2

Roadway Speed and Travel Time	
Westheimer Westbound from Kirkwood to Dairy Ashford	
Distance	1 miles
Travel Time	1 minute 57 seconds
Average Speed	31 MPH

Link 3

Roadway Speed and Travel Time	
Westheimer Westbound from Dairy Ashford to Eldridge	
Distance	1.1 miles
Travel Time	1 minute 41 seconds
Average Speed	39 MPH

Figure 4.3 An example of real-time traffic information for each link in the test bed

4.4 Travel Time Data

4.4.1 Data Collection Technology

There are various individual vehicle collection methods including GPS-equipped probe vehicles, the automatic license plate recognition system, radio-frequency identification (RFID) tool tag technology, and Bluetooth technology. Their advantages and disadvantages are summarized in table 4.3 (Turner et al. 1998; Vo 2011). Bluetooth has proven to be a viable technology in terms of its relatively large penetration rate of travel time data samples for a given roadway segment. One research project showed that 5 to 7 percent of vehicles in a traffic stream were equipped with Bluetooth enabled devices (Tarnoff et al. 2009). This relatively large market penetration rate of Bluetooth data makes it applicable to calculate travel time standard deviations for short time intervals, which is necessary to evaluate short-term travel time variability using reliability intervals.

In addition, the Bluetooth measurement system is very cost-effective because the traffic agency does not have to spend any money equipping vehicles. It has been shown that the cost of this technology is one to two orders of magnitude below the costs for traditional toll tag reader equipment (Puckett and Vickich 2010). The cost advantage enables travel time to be collected on more roadway links. It also makes estimating arrival travel time reliability easier and potentially more accurate for relatively long trips.

Table 4.3 A comparison of various travel time collection technologies

GPS-Equipped Probe Vehicles	Pros: 1) Reduction in staff requirement compared to manual method 2) Relatively portable, reliable, and accurate 3) Generates automatic geo-coding of detailed speed data
	Cons: 1) Losing signals due to tall buildings, tunnels, etc. 2) Building the based map using a geographic information system (GIS) to use the incoming data 3) Limited sample size and privacy issues
Automatic License Plate Matching	Pros: 1) Decrease in data reduction time 2) Large sample size
	Cons: 1) Constrained by lighting conditions 2) Technologically intensive and unstandardized vendors
RFID	Pros: 1) Cost-effective compared to loop and video detection method 2) Takes advantage of preexisting RFID infrastructure 3) Larger sample size compared to probe vehicle method
	Cons: 1) Constrained by market penetration of RFID toll tags 2) Limited portable applications and infeasible to be implemented in arterials 3) Privacy issues
Bluetooth	Pros: 1) Low cost, standardized, non-proprietary equipment and protocols 2) Easy, non-intrusive field installation and maintenance with portable applications 3) Large penetration of data samples 4) Real-time summary calculations 5) Complete ownership of data by operating agency 6) No privacy issues
	Cons: Outliers from Bluetooth devices of non-vehicles

(Vo 2011)

4.4.2 Bluetooth Travel Time Data for the Case Study

The base datasets used in this study were collected from January 1st to July 31st in

2011 for the arterial corridor on westbound Westheimer Road, from Wilcrest Drive to

Eldridge Parkway. The link-based travel times were generated by matching MAC

readings between readers located at the beginning and end points of each link. The corridor-based travel times were generated by matching the readings between the readers at the beginning and end points of the corridor (e.g., between Wilcrest Drive and Eldridge Parkway).

4.4.2.1 Link Travel Time Distribution

Travel times on arterial links with traffic signals are expected to form a multiple-modal distribution because some vehicles will be stopped at one or more traffic signals while other vehicles will progress through the signals unimpeded. Using the travel time data from 16:00 to 19:00 on March 7, 2011 for the link from Wilcrest to Kirkwood, the distribution parameters were quantitatively investigated through the expectation-maximization (EM) algorithm, implemented by Mclust. Mclust is an R package developed by the University of Washington for model-based density estimation (Fraley and Raftery 2002; Barkley et al. 2012). The best number of modes and related parameters (i.e., mean and variance of each mode) are determined through a penalty on the number of model parameters and Bayesian Information Criterion (BIC), which is the value of the maximized log-likelihood.

The results for link Wilcrest-Kirkwood are summarized in figure 4.4. As expected, the histogram of the travel time data from 16:00 to 19:00 on March 7, 2011, shown in figure 4.4 (a), is bi-modal in nature. Figure 4.4 (b) shows the fitted distribution through the EM algorithm. It may be seen that the PDF has two components. The first component is a normal distribution with a mean of 110 seconds and a variance of 147

seconds^2. The second component is a normal distribution with a mean of 188 seconds and a variance of 767 seconds^2.

Figure 4.4 (c) and (d) shows the CDF and the quantile-quantile (Q-Q) plot, respectively, which are used for diagnostic purposes. The CDF plot in figure 4.4 (c) compares the estimated CDF (e.g., black curve) with the empirical distribution function (e.g., green dashed curve). It may be seen that the CDF of the empirical data and the estimated CDF match each other closely. The highest difference between empirical and estimated CDFs is less than 7 percent. The Q-Q plot in figure 4.4 (d) is a graphical technique to determine if the dataset of the estimated distribution and the observed dataset come from populations with a common distribution. A Q-Q plot shows the quantiles of the first dataset in relation to the quantiles of the second dataset. A quantile is defined as the percentage of points below the given value. For example, the 30 percent quantile is the point at which 30 percent of the data fall below the value, and 70 percent fall above. The 45-degree reference line is also plotted in the Q-Q plot. If the two sets come from a population with the same distribution, the points should fall approximately along this reference line. For this example, figure 4.4 (d) indicates that the estimated distribution of the first component models the empirical data very closely because the Q-Q plot lies on the 45-degree reference line. In contrast, while the second cluster doesn't model the empirical data as closely, it may be considered as an adequate approximation. It can be seen that the observed travel time dataset follows a bimodal distribution.
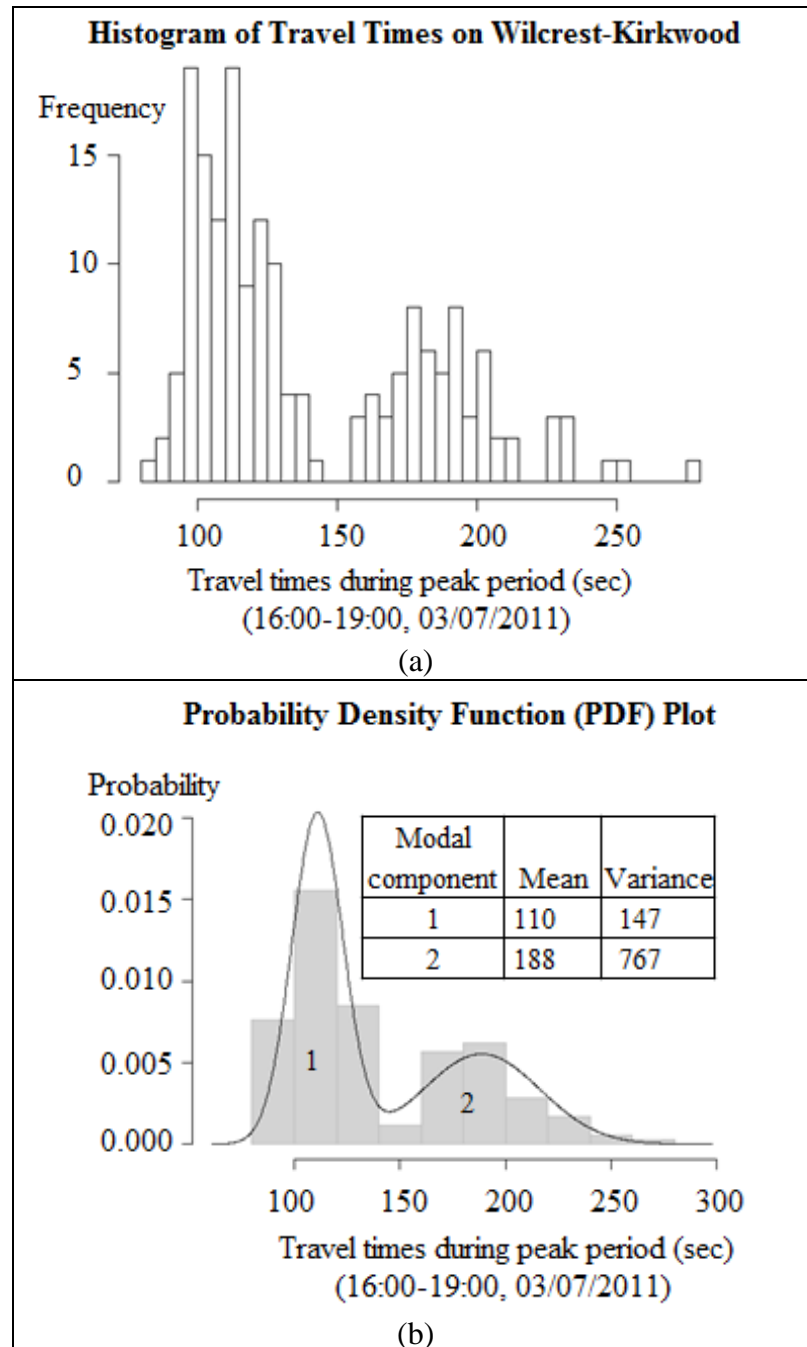
Figure 4.4 Results of model-based density estimation for travel time data
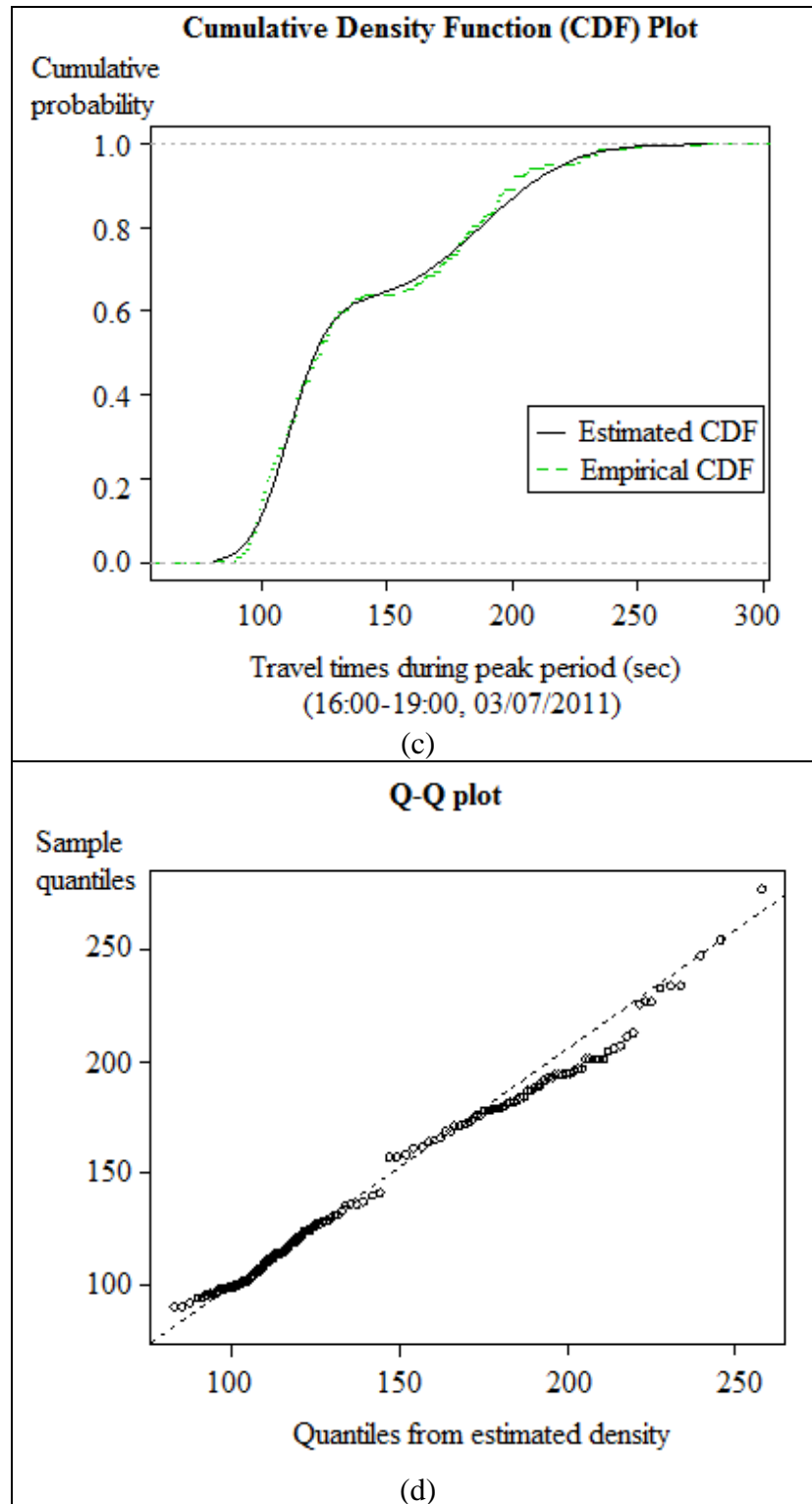
Figure 4.4 Results of model-based density estimation for travel time data (cont.)

4.4.2.2 Link-Based and Corridor-Based Dataset

The raw Bluetooth dataset from the AWAM systems in Houston was reduced into the format shown in table 4.4. Specifically, the raw data was transformed into link specific information. Each line includes the MAC address of the vehicle, the locations of the readers (cross streets 1 and 2), the two timestamps when the vehicle entered and exited each link, the associated estimated trip time in seconds (i.e., the time difference between the two timestamps), and the speed in mph. If a given vehicle has more than one Bluetooth-activated device being detected, it generates multiple data lines with exactly the same records for all the columns in table 4.4 except the MAC address. In this situation, information from only one device is kept so that the MAC address column can be used as the vehicle ID to distinguish different vehicles.

Subsequently, the three link based data sets were combined into one corridor based dataset. The corridor-specific dataset was derived by finding the sum of link-specific travel times for the same vehicle ID traveling through the three-link corridor. By definition, the number of observations in the corridor-specific dataset will be equal to or smaller than that of each of the three link-specific databases. Table 4.5 shows some examples of the corridor-specific dataset records including the MAC address of each vehicle, the timestamps entering and leaving all three links, the corresponding link travel times in seconds, and the corridor travel time in seconds. Note that for a given vehicle the timestamp for entering a downstream link should be the same as the timestamp for departing the upstream link for a single vehicle. There were some observations where the two timestamps were not identical, as shown in line 15 in table 4.5. It is hypothesized that

this is because the vehicle waited at an intersection and was recorded with multiple timestamps at the same reader station. Different timestamps, however, were selected to match link travel times for different links. To eliminate any inconsistency, the time a given vehicle left a link was set equal to the time the vehicle entered the following link. For example, the observation in line 15 in table 4.5 is adjusted, as shown in line 16.

Table 4.4 Examples of corridor-based travel time data

Vehicle ID	Cross street 1 (from)	Cross street 2 (to)	Enter Time	Exit Time	Travel Time (sec)	Speed (mph)
00:24:83:50:C1:48	Westheimer_Wilcrest	Westheimer_Kirkwood	2011/1/1 16:05:08	2011/1/1 16:07:56	168	21
00:12:1C:51:68:3D	Westheimer_DairyAshford	Westheimer_Kirkwood	2011/1/1 16:05:42	2011/1/1 16:07:36	114	32
10:1D:C0:96:67:40	Westheimer_Kirkwood	Westheimer_DairyAshford	2011/1/1 16:06:10	2011/1/1 16:07:56	106	34
00:26:5D:E1:08:52	Westheimer_Wilcrest	Westheimer_Kirkwood	2011/1/1 16:06:33	2011/1/1 16:07:57	84	43
00:18:C5:97:53:E0	Westheimer_Wilcrest	Westheimer_Kirkwood	2011/1/1 16:06:23	2011/1/1 16:08:00	97	37
00:12:1C:F5:C3:03	Westheimer_Wilcrest	Westheimer_Kirkwood	2011/1/1 16:06:51	2011/1/1 16:08:11	80	45
6C:9B:02:21:C1:29	Westheimer_DairyAshford	Westheimer_Eldridge	2011/1/1 16:07:35	2011/1/1 16:08:56	81	49
00:15:D3:7F:DF:14	Westheimer_DairyAshford	Westheimer_Eldridge	2011/1/1 16:07:40	2011/1/1 16:08:55	75	53
C0:38:F9:5A:1A:F0	Westheimer_DairyAshford	Westheimer_Kirkwood	2011/1/1 16:06:18	2011/1/1 16:09:28	191	19
00:21:FE:76:34:4B	Westheimer_DairyAshford	Westheimer_Kirkwood	2011/1/1 16:07:04	2011/1/1 16:09:55	171	21
10:1D:C0:96:67:40	Westheimer_DairyAshford	Westheimer_Eldridge	2011/1/1 16:07:56	2011/1/1 16:09:33	97	41
00:26:5D:E1:08:52	Westheimer_Kirkwood	Westheimer_DairyAshford	2011/1/1 16:07:57	2011/1/1 16:09:58	121	30
00:05:4F:49:9A:D7	Westheimer_DairyAshford	Westheimer_Kirkwood	2011/1/1 16:07:58	2011/1/1 16:09:46	108	33
D4:E8:B2:34:4E:6A	Westheimer_DairyAshford	Westheimer_Eldridge	2011/1/1 16:08:05	2011/1/1 16:09:33	88	45

Table 4.5 Examples of corridor-based travel time data (2011/1/1)

Line No.	Vehicle ID	Link 1 Wilcrest-Kirkwood			Link 2 Kirkwood-Dairy Ashford			Link 3 Dairy Ashford-Eldridge			Corridor
		Enter Time	Travel Time (sec)	Exit Time	Enter Time	Travel Time (sec)	Exit Time	Enter Time	Travel Time (sec)	Exit Time	Travel Time (sec)
1	00:24:90:C8:A7:04	15:46:43	80	15:48:03	15:48:03	106	15:49:49	15:49:49	82	15:51:11	268
2	FC:A1:3E:B8:BF:FB	15:55:46	86	15:57:12	15:57:12	111	15:59:03	15:59:03	93	16:00:36	290
3	10:1D:C0:96:67:40	16:04:43	87	16:06:10	16:06:10	106	16:07:56	16:07:56	97	16:09:33	290
4	00:26:5D:E1:08:52	16:06:33	84	16:07:57	16:07:57	121	16:09:58	16:09:58	86	16:11:24	291
5	C0:38:F9:47:97:1B	16:08:38	102	16:10:20	16:10:20	119	16:12:19	16:12:19	103	16:14:02	324
6	00:24:91:1A:86:90	16:09:29	175	16:12:24	16:12:24	122	16:14:26	16:14:26	93	16:15:59	390
7	00:10:18:E8:F1:D0	16:13:21	79	16:14:40	16:14:40	123	16:16:43	16:16:43	82	16:18:05	284
8	00:23:39:8C:CA:14	16:19:55	89	16:21:24	16:21:24	129	16:23:33	16:23:33	85	16:24:57	302
9	00:25:66:86:0E:B8	16:22:44	84	16:24:08	16:24:08	106	16:25:54	16:25:54	95	16:27:29	285
10	5C:59:48:70:3A:33	16:24:20	100	16:26:00	16:26:00	127	16:28:07	16:28:07	105	16:29:52	332
11	00:24:83:60:7D:12	16:29:12	82	16:30:34	16:30:34	120	16:32:34	16:32:34	88	16:34:02	290
12	44:4E:1A:72:D4:62	16:29:33	73	16:30:46	16:30:46	102	16:32:28	16:32:28	93	16:34:01	268
13	00:26:5F:D1:66:44	16:31:47	82	16:33:09	16:33:09	104	16:34:53	16:34:53	93	16:36:27	280
14	00:22:A9:2C:E2:73	16:32:20	155	16:34:55	16:34:55	126	16:37:01	16:37:01	91	16:38:31	371
15	E8:E5:D6:76:D0:2F	17:19:07	153	17:20:38	17:21:40	156	17:24:16	17:24:16	96	17:25:52	405
				↓							
16	E8:E5:D6:76:D0:2F	17:19:07	153	17:21:40	17:21:40	156	17:24:16	17:24:16	96	17:25:52	405

The link travel times in the link-based dataset include left-turn, right-turn, and through movements, while the link travel times in the corridor-based dataset, by definition, only include through movements. The only exception is that the vehicles travel along the corridor until the last link, where it is impossible to know their movements on the last link. Figure 4.5 shows the CDF plots of link travel times in the corridor-based dataset (LTT1) and in the link-based dataset (LTT2) for all three corridor links. In the figure, the blue solid lines represent the link travel times in the corridor-based dataset, which is specific for through movements. The red lines represent the link-based dataset including both turning and through movements. A Kolmogorov-Smirnov (K-S) test was conducted to test the consistency between the distribtions of the link travel times in the two different datasets for all three links. The null hypothesis and alternative hypothesis are:

$$H_0: LTT1 \text{ and } LTT2 \text{ are from the same continuous distribution.}$$

$$H_a: LTT1 \text{ and } LTT2 \text{ are from different continuous distributions.}$$

where:

$LTT1$ = link travel time observations in the corridor-based dataset,

$LTT2$ = link travel time observations in the link-based dataset.

The results of the two-sample Kolmogorov-Smirnov tests rejected the null hypothesis that the two samples are from the same continuous distribution at the 5 percent significance level for all of the three links. It can be seen from the figure for the Wilcrest-Kirkwood link that the link travel times in the corridor-based dataset are, on average, shorter than those in the link-based dataset. In contrast, the relationship is not

straightforward for the Kirkwood-DairyAshord and DiaryAshford-Eldridge links. For short travel times of less than 200 seconds, the corridor-based link travel times are lower than the corresponding link-based travel times. This is as expected because the latter includes left-turns and right-turns. The situation reversed after 200 seconds, but the greatest differences are only 3.5 percent and 7.4 percent for the corresponding corridor-based link travel times for the Kirkwood-DairyAshford and DairyAshford-Eldridge links, respectively. It is hypothesized for congested periods that the straight movements are affected more by congestion than the left-turn and right-turn movements for the DairyAshford-Eldridge link. It is important to be aware of the difference between the two kinds of datasets when developing prediction methodologies. Using link-based datasets as input to develop prediction models for mean corridor travel times can introduce additional bias at the input level.
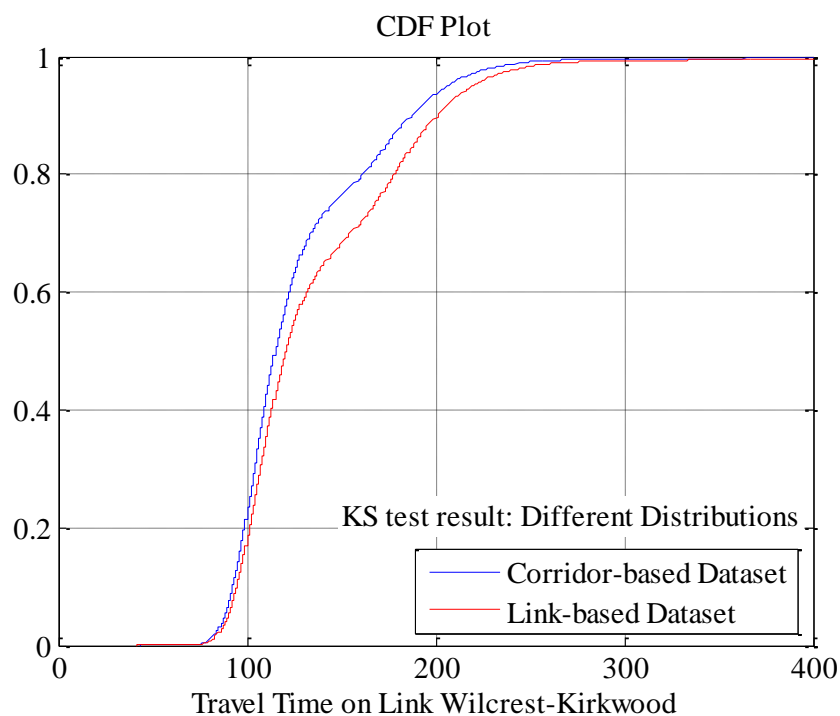
Figure 4.5 CDF plots of link travel times in corridor-based and link-based datasets
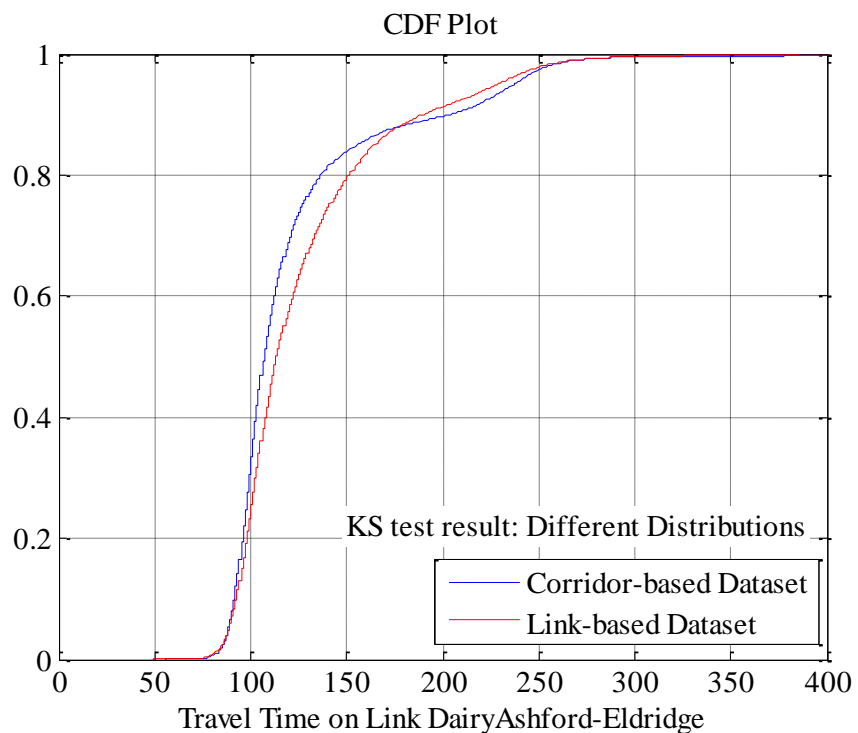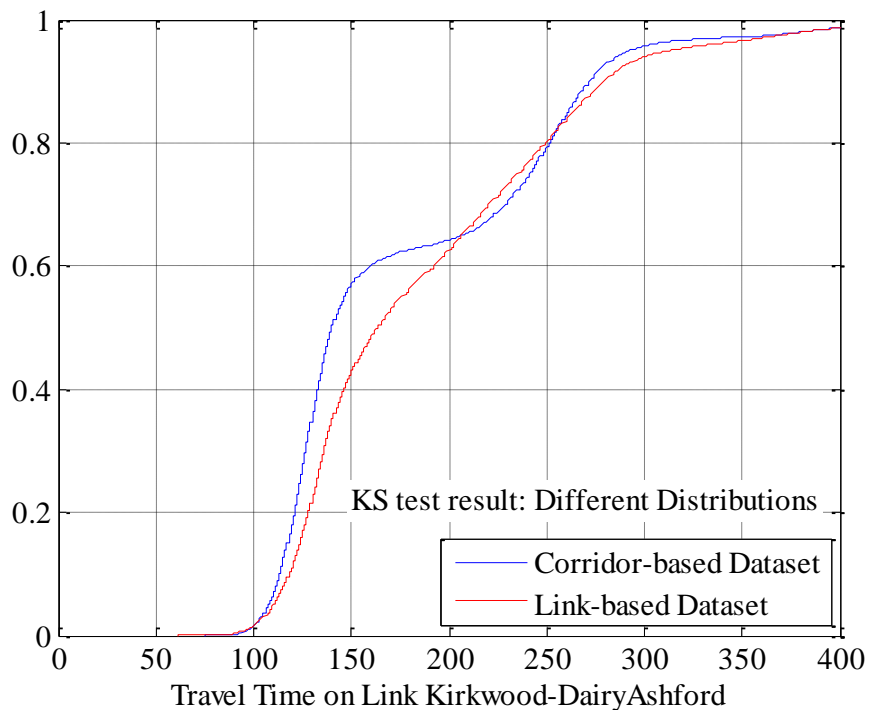
CDF Plot

Figure 4.5 CDF plots of link travel times in corridor-based and link-based datasets

(cont.)

However, the method to predict mean corridor travel time using corridor-based travel time observations also has its own constraints. First, the sample size of corridor-based datasets are always equal to or smaller than the corresponding link-based datasets. This is because not all of the vehicles that are identified at the link level travel down the entire corridor. In this case study, for example, the number of corridor-based travel time observations is 29,115, while the sizes of the three link-based travel time datasets are 65,728; 57,327; and 67,223, respectively. Secondly, the link travel times for a given vehicle, by definition, are shorter than the corridor travel times. Because travel time observations will not become available until the vehicle passes the last beacon, prediction models using corridor travel time data as an input have to use data that are not as recent as those using link travel time data as an input. For example, the average corridor travel time and the highest average link travel time in this study are 442 seconds and 182 seconds, respectively. Therefore, the corridor-based prediciton model will be using, at best, observations entering the corridor approximately 442 seconds before the current time. In contrast, the link-based prediction model can use observations that first entered the corridor approximately 182 seconds before the current time. This constraint of corridor travel time data is particularly true for applications on long OD routes (e.g., one hour) where it could be impractical to collect enough real-time observations to study the OD arrival time reliability. Because the exact effect of these issues is unknown, this dissertation will compare different combinations of model structures and data input formats for predicting the mean corridor travel time and the associated arrival time reliability.

4.4.3 Data Reduction and Outlier Identification

Outliers need to be identified and removed before conducting traffic analyses and/or developing prediction models. This is particularly true for the urban arterial corridor studied in this research, because the Bluetooth travel time data could include observations of non-vehicle travel times and trip chain travel times. Non-vehicle travel times in a multi-modal urban arterial could be from bicycles and pedestrians (O'Neil et al. 2006). A previous study showed that pedestrian traffic accounted for seven percent of Bluetooth observations (O'Neil et al. 2006). Trip chain travel times occur when the two matched observations are not from a single trip, resulting in an unreasonably long travel time. The links in this study are on the Westheimer Road arterial that is home to various attractions including shopping malls, schools, and restaurants. As such, the probability of drivers making trip chains around this area is high. Consider the example shown in figure 4.6. A vehicle driving west on Westheimer Road is detected by the first sensor at Kirkwood. The driver subsequently stops at Susie's Cakes to get some food for 15 minutes. The driver returns to Westheimer Road, proceeds westward, and is then detected at DairyAshford. The matching algorithm would generate a link travel time 15 minutes longer than the travel time between Kirkwood and DairyAshford without the stop, because the algorithm could not account for the intermediate stop without additional information. The goal of outlier identification would be to identify these measurements and remove them from the data.

Figure 4.6 One example of trip chain travel times

The existing methods for outlier identification include the smoothed histogram-based method, the Host Software filter algorithm, and the moving standard deviation method (Haghani et al. 2010; Puckett and Vickich 2010; Quayle et al. 2010). Boxel et al. (2011) proposed a statistical method based on the observations' standard residual in a robust Greenshields model. Each of these methods has its own advantages and limitations, as discussed below.

4.4.3.1 Smoothed Histogram-Based Method

Haghani et al. (2010) uses the smoothed histogram-based method to identify "abnormal" travel times for freeways. One example of an abnormal travel time was the detection of a specific vehicle at two consecutive Bluetooth sensors occurring on different days. They were matched in the travel time algorithm to generate unreasonably large travel time observations. In this method, the histogram of observed travel times is first generated. A moving average is then calculated to estimate the travel time distribution by "smoothing" the frequencies in the histogram. The algorithm identifies the peak in the moving averages and then searches for the first category on either side that violates the expected "down trend". The categories that violated the trend are treated as outliers. Thus, it requires the data to be a single modal distribution (Kim et al. 2012).

However, the travel time data in this case study is not a single modal distribution, as shown in section 4.4.2.2. Thus, the smooth histogram method is not a good choice for identifying outliers in the arterial travel time dataset used in this study.

4.4.3.2 Moving Standard Deviation (SD) Method

Equations 4.3a and 4.3b illustrate the calculation process of the moving standard deviation method. The threshold T is used to discriminate outliers (e.g., those unreasonably slow trips) and is calculated by adding $\pm\alpha$ local standard deviations to the mean. The local standard deviation σ is calculated by equation 4.3b, and this is referred to as the central mode. In essence, the local standard deviation for a given observation x is calculated by the sample of size $(u+1)$ centering on x.

$$T = \mu \pm \alpha \cdot \sigma \tag{4.3a}$$

$$\sigma = \sqrt{\frac{1}{u+1} \sum_{i=x-u/2}^{x+u/2} (p_i - \mu)^2} \tag{4.3b}$$

where:

T = the threshold of the moving SD method used to discriminate outliers;

μ = the mean of the neighborhood sample;

σ = the local standard deviation of the neighborhood sample;

α = the range parameter, which is based on experience, and often one or two are used (Quayle et al. 2010), $\alpha=2$ in this dissertation;

$(u + 1)$ = a user-set neighborhood sample size to base standard deviation

           comparisons on, u=30 in this dissertation;

x = the location of the current detection being assessed; and

p_i = the travel time value for detection i.

This method assumes that travel times are independent and normally distributed, which is violated for the urban arterial corridor used in this dissertation. As discussed in 4.4.2.1, the corridor travel time follows bi-modal distributions instead of one single normal distribution. In addition, the travel time observations might not be independent. The performance of the moving SD method depends on parameters such as neighborhood size $(u + 1)$ and the range parameter (α). That is, it might be unreasonable to assume the parameter values on different links during various time periods are all equal.

4.4.3.3 Gap Method

This method is based on the observation that there is often a large gap between non-outlier and outlier travel times. Kim et al. (2012) assumed the critical gap length to be equal to 0.5 times the median travel time. The basic host software system method developed by the Texas Transportation Institute (TTI) defines outliers as the travel time that differs from the current average for the roadway link by more than a certain percentage (e.g., 25 percent) (Puckett and Vickich 2010). The limitation of the gap-based approach is that it is slow to adapt to the rapidly-changing travel times during the transition period (Kim et al. 2012). Puckett and Vickich (2010) found that the algorithm worked well for freeway segments but was insufficient for signalized arterial segments

because of the wide variance in arterial travel times. The host software addressed this issue by allowing users to specify the algorithm and parameters that would yield the most effective results (Puckett and Vickich 2010).

4.4.3.4 A Statistical Method Based on a Robust Greenshields Model

Boxel et al. (2011) presented this statistical methodology for real-time deployments and tested its performance on both Interstate highways and urban arterial corridors. Outliers are identified by a data point's standard residual in a robust Greenshields model. A set of valid speed data points should have standardized residuals that are standardized normal. Outliers, however, don't have such a convenient distribution. The results validated the effectiveness of the model at identifying outliers from Bluetooth data. To compare the performance of the method for interstate and arterial corridor datasets, the Shapiro-Wilk (S-W) test was used to validate the normality of the standard residuals. They found the performance was better for Interstate highways than for urban arterials because the p value of the S-W test is more significant for the interstate dataset cleaned by this method compared to the cleaned arterial corridor dataset.

4.4.3.5 Robust Local Regression Based Method

The robust local regression model has been shown to be useful for identifying outliers (Cleverland 1979). The robust locally weighted regression scatterplot smoothing (RLOESS) model is a nonparametric method able to yield a robust regression surface out of scatter observations with outliers, and thus, those outliers have a limited impact on the

resulting threshold. An added advantage of this approach is that it does not require

assumptions about the data distribution. As part of this dissertation, an extended outlier

identification method was developed based on the RLOESS model. This new method

identifies outliers based on the residual of a data point in the RLOESS model by

following these steps:

1. Decide the user-set neighborhood sample size u to fit a local RLOESS model.

    Data points outside the span u around x_i (i.e., the current data point being

    processed) have no influence on the fit.

2. Compute the regression weights for each data point in the span u using

    equation 4.4.

$$w_i = (1 - |\frac{x - x_i}{d(x)}|^3)^3 \tag{4.4}$$

    where:

    w_i = the regression weight of the ith data point,

    x = the detection time of the current data point being assessed,

    x_i = the neighbor of x within the user-set span, and

    $d(x)$ = the distance along the abscissa from x to the most distant data

    point within the span.

3. A weighted, linear least-squares regression is performed on the raw travel

    time data. The smoothed values for the average travel time at each moment

    can be calculated from the regression model.

4. Calculate the residuals r_i from the smoothed value and the real observation.

5. Compute the robust weights for each data point in the span using equation 4.5.

    The robust weights help to remove the impact of outliers on the local

regression model. If the robust weight is 0, the associated data point is excluded from the smooth calculation.

$$w_i = \begin{cases} (1 - (r_i/5MAD)^2)^2, & |r_i| < 5MAD \\ 0, & |r_i| \geq 5MAD \end{cases} \qquad (4.5)$$

where:

w_i = the robust weight of the *i*th data point,

MAD = the median absolute deviation of the residuals, and

r_i = the residual of the *i*th data point.

6. Smooth the data again using the robust weights. The final smoothed value is calculated using both the local regression weight and the robust weight.

7. Repeat the previous step for a total of A iterations to get the final RLOESS regression results and final residual R_i.

8. If the resulting robust weight of x is less than a user-set threshold B, x is identified as an outlier.

In this algorithm, parameters u, A, and B are defined based on engineering judgment. In this dissertation u is set to 0.5 times the sample size of travel time measurements during the peak period in the day being examined. The sample size varies from 300 to 400 per day. Parameters A and B were set to 3 and 0.05, respectively.

Figures 4.7-4.9 compared the outliers screened by this RLOESS based method and the moving standard deviation (SD) method for the three links in this study. The PM peak period (15:45-19:00) on January 3, 2011 (Monday), is used for illustration purposes. Potential identification errors are circled by a red circle. It can be seen that the travel time data on link Wilcrest-Kirkwood, shown in figure 4.7, does not have an identifiable PM

peak on the selected day. That is, there is no rapid changes within the dataset. It can be seen from figure 4.7 (b), two potential outliers were not identified by the moving SD method. Because the local mean and SD used to distinguish outliers are sensitive to outliers, the new method based on RLOESS is able to classify the two points correctly, as shown in the upper plot in figure 4.7 (a).

The link between Kirkwood and Dairy Ashford experienced a slight increase of travel time from 5:00 pm to 6:45 pm. In figure 4.8 (b), the five observations identified as outliers by the RLOESS based method, but not by the moving SD method, are circled.

Compared to the Kirkwood-DairyAshford link, the link between Diary Ashford Road and Eldridge Road experienced a shorter peak period with more rapid travel time changes at the shoulders of the peak period. As shown in figure 4.9, both methods resulted in some unreasonable identifications as shown in the red circles. The ROLESS based method yielded a more conservative result because it identified more outliers for the period from 5:45 pm to 6:45 pm.

It is important to note that no matter which method is used for outlier identification, engineering judgment is always necessary to fine-tune the algorithm parameters for the best application performance. The RLOESS based method with a single parameter setting is able to generate consistent results for different links and different hours.
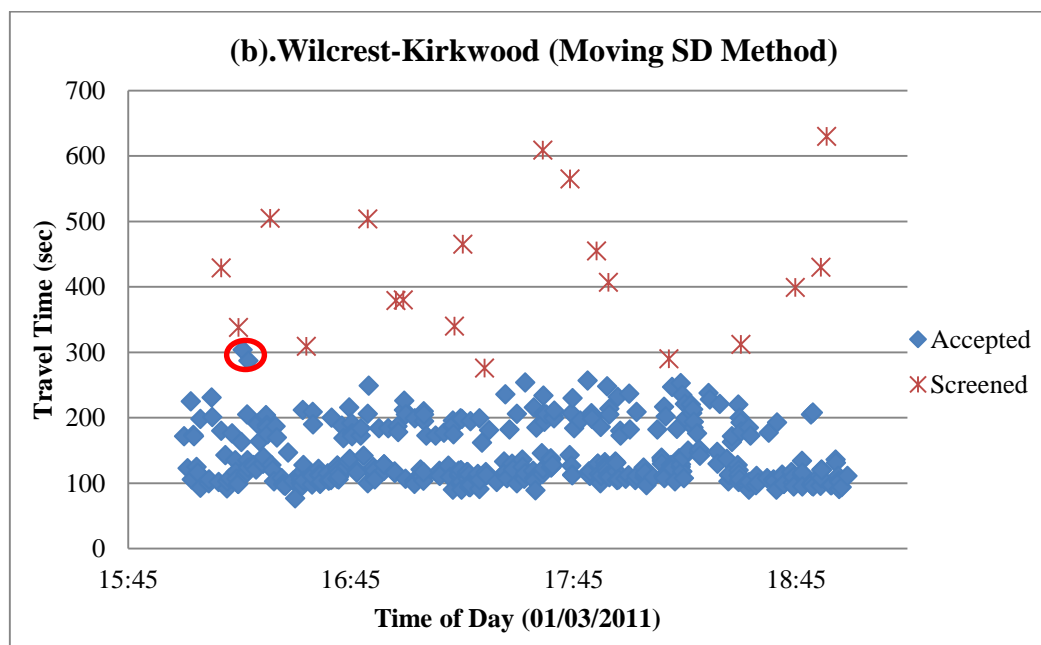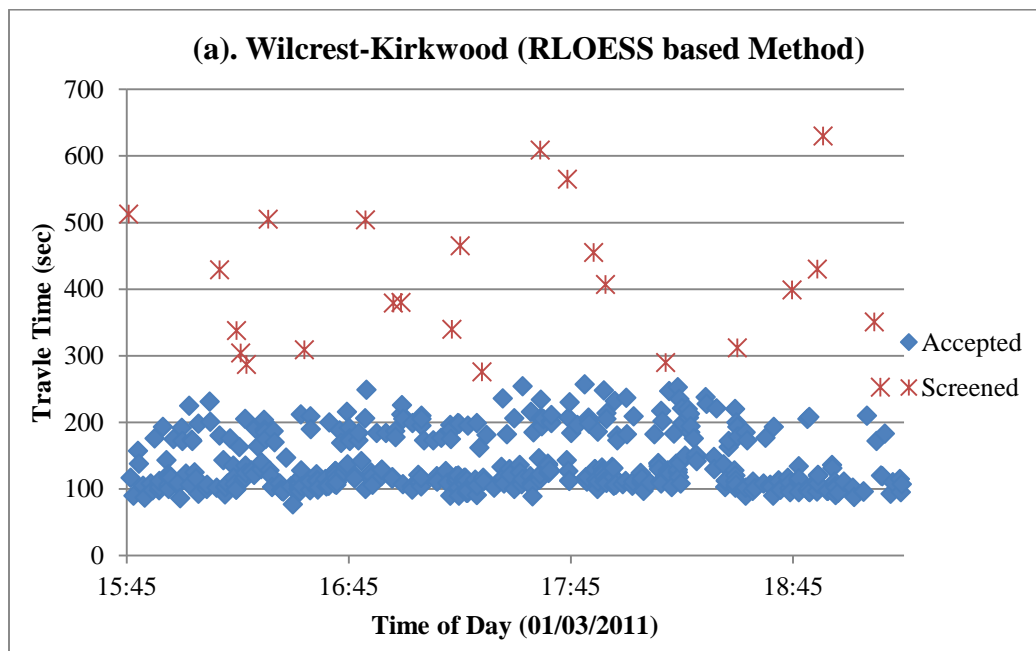
Figure 4.7 Outlier identification for Wilcrest-Kirkwood link
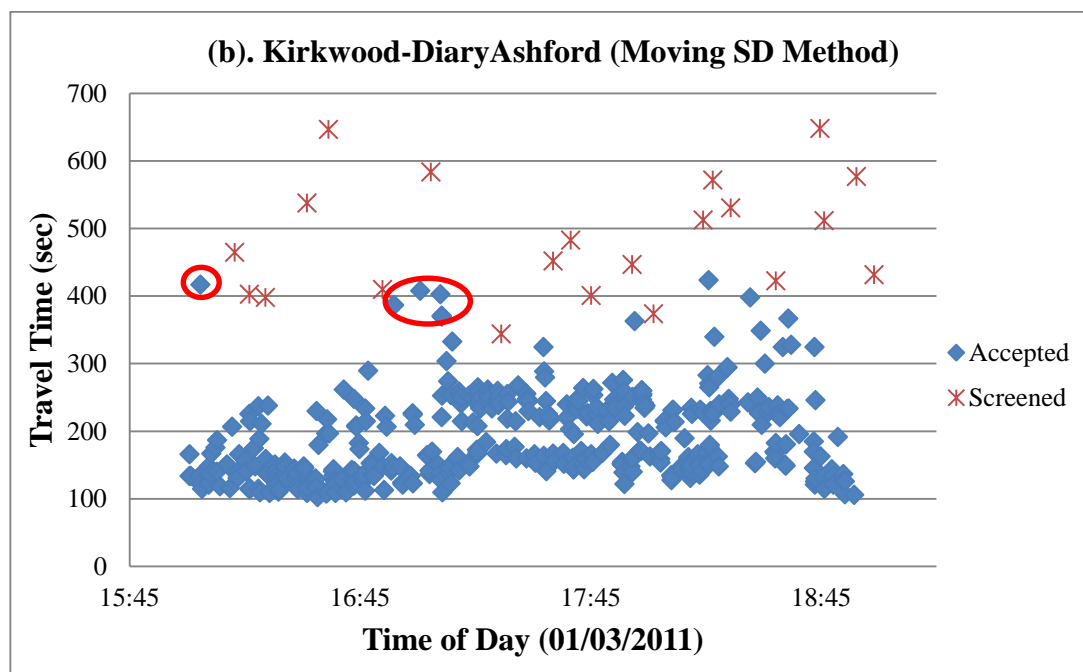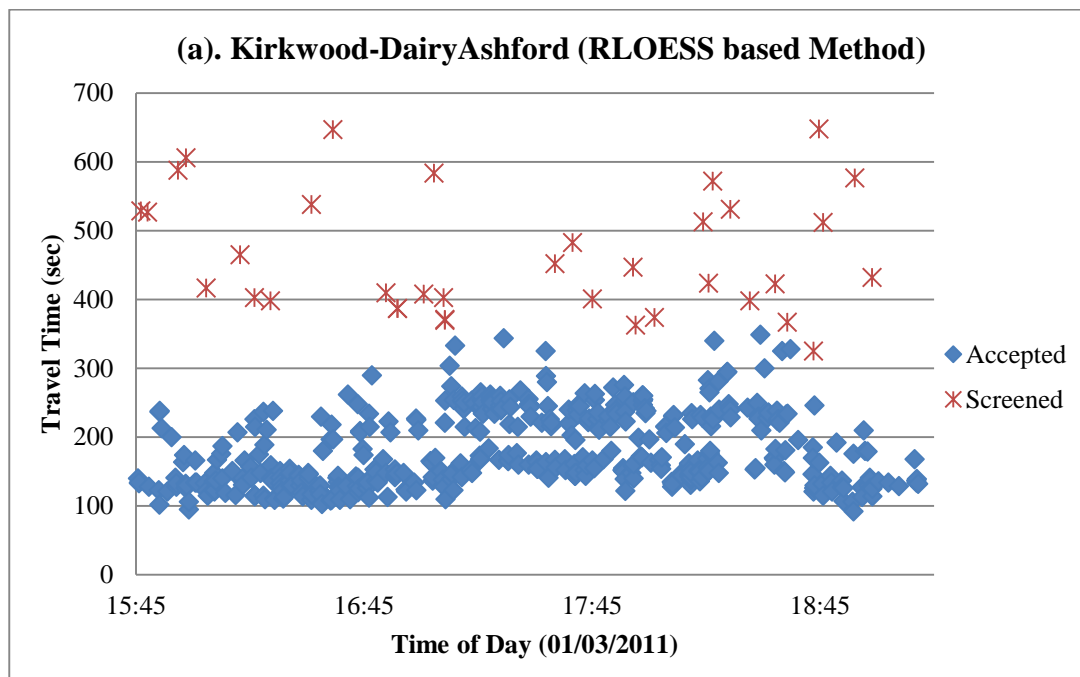
Figure 4.8 Outlier identification for Kirkwood-DairyAshford link
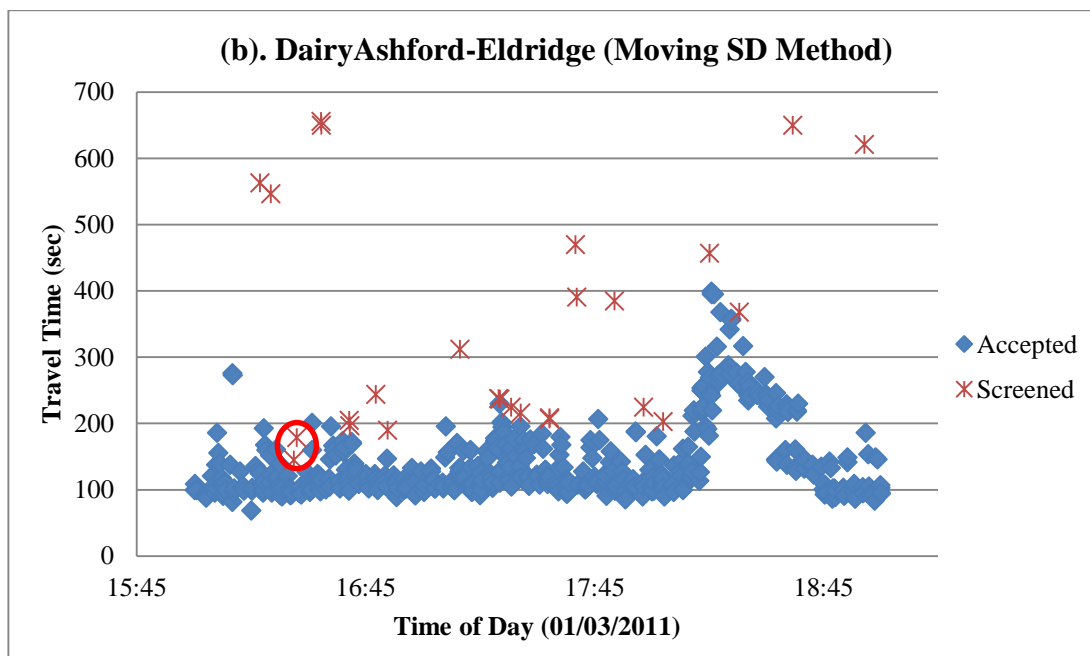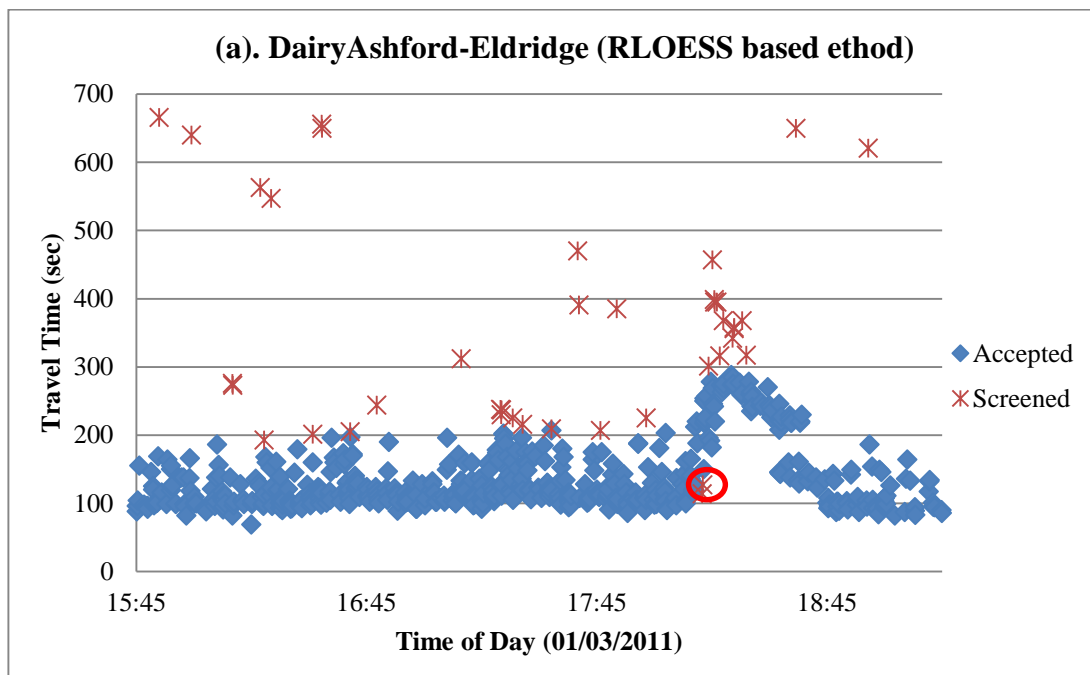
Figure 4.9 Outlier identification for DairyAshford-Eldridge link

4.5 Prediction Methodology and Results

The objective of this dissertation is to compare different combinations of model structures and data input formats to predict the reliability interval of arrival times. Every model studied in this chapter was estimated using link-based and corridor-based travel time data, respectively. Overall, 13 model formulations are tested using various combinations of the two kinds of datasets (e.g., link and corridor travel times) and four different modelling methods (e.g., neural network), as listed in table 4.6.

Table 4.6 The list of studied prediction models

Model No.	Model Structure
1	Corridor-based historical average travel time model (HTT)
2	Corridor-based instantaneous travel time model (ITT)
3	Corridor-based nonlinear autoregressive with exogenous inputs (NARX) neural network
4	Link-based HTT with naïve sum method for corridor prediction
5	Link-based HTT with accumulative sum method for corridor prediction
6	Link-based HTT with first-order approximation algorithm for corridor prediction
7	Link-based HTT with second-order approximation algorithm for corridor prediction
8	Link-based ITT with naïve sum method for corridor prediction
9	Link-based ITT with accumulative sum method for corridor prediction
10	Link-based ITT with first-order approximation algorithm for corridor prediction
11	Link-based ITT with second-order approximation algorithm for corridor prediction
12	Link-based feedforward neural network
13	Link-based NARX neural network

The link-based and corridor-based travel time datasets, discussed in section 4.4, were used to calculate link and corridor travel time means and variances for each 15-minute interval. For those models estimated using the link-based dataset, the corridor

travel time in the next 15-minute interval $T(i+1)$ is calculated based on the previous travel times on all corridor links. For those models estimated using the corridor-based dataset, $T(i+1)$ is predicted based on the previous corridor travel times.

The results will identify which model formulation works best for a specific objective. The efficiency of the mean corridor travel time prediction models is measured by the mean absolute percentage error. The prediction of the arrival time reliability intervals is evaluated through the coverage rates and average ranges of the predicted reliability interval of arrival time for the overall analysis period and an example incident period. The overall analysis period includes all the 15-minute intervals during the PM peak (i.e., 16:00-19:00) in June and July 2011. The example incident period is from 17:30 to 19:00 on July 21, 2011. The travel time data from January to May in 2011 are used to train the neural network models and to initiate the historical average model for the first weekday in June 2011. The dataset for June and July in 2011 is used to compare the performance of the trained neural networks, historical average models, and instantaneous travel time models. Although this case study uses Bluetooth travel time data, the methodology developed in this dissertation can be extended to other data collection methods that generate disaggregate measurements (e.g., individual travel time) such as automatic vehicle identification, GPS, and so on.

4.5.1 Prediction Models

4.5.1.1 Simple Prediction Models

Travel times at a future moment in time depend on the future demand for the network, the future capacity of the network, and the current state of the network. These

factors suggest two simple prediction methods that can be used as a baseline to compare with more advanced prediction models (Nikovski et al. 2005).

4.5.1.1.1 Historical Average Travel Time Method (HTT)

The historical average method predicts that the average and standard deviation (SD) of travel time for a given interval is simply the long-term average of averages and SDs over all the previous days for that specific interval (Park 1998, Bovy and Thijs 2000, Naik 2010). This prediction would be accurate if future demand and capacity were equal to their historic averages, and the current and past states had no influence on future states (Nikovski et al. 2005). In this case study, the long-term averages are calculated over all the weekdays in the five months before the day to be predicted.

1). For the corridor travel time dataset, the historical average method is shown in equations 4.6a and 4.6b.

$$\widehat{\overline{T}}_{d,p}^{c,H} = \frac{1}{R} \sum_{r=1}^{R} \overline{T}_{r,p}^{c} \quad \forall d = 1, \dots, D; p = 1, \dots, P \tag{4.6a}$$

$$\hat{s}_{d,p}^{c,H} = \frac{1}{R} \sum_{r=1}^{R} s_{r,p}^{c} \quad \forall d = 1, \dots, D; p = 1, \dots, P \tag{4.6b}$$

where:

$\widehat{\overline{T}}_{d,p}^{c,H}$ = the predicted average corridor travel time for day d amd period p. The superscript c indicates that it is based on the corridor travel time dataset, and the superscript H indicates that it is calculated through the historical

average method. The d represents the weekdays, in order, starting on June 1, 2011, and the p represents a particular 15-minute time period.

$\bar{T}_{r,p}^c$ = the average corridor travel time in period p and day r, using the corridor-based dataset.

$\hat{s}_{d,p}^{c,H}$ = the predicted standard deviation of corridor travel time for day d and period p, using the historical average method.

$s_{r,p}^c$ = the corridor standard deviation of travel time in period p and day r, calculated from the corridor-based dataset.

D = the number of days for which corridor travel time means and SDs are being calculated. For this example, there are D=43 days to estimate, and these correspond to the 43 weekdays for June and July of 2011.

R = the number of days used to calculate the $\hat{\bar{T}}_{d,p}^{c,H}$ and $\hat{s}_{d,p}^{c,H}$. In this example, R equals the 105 weekdays immediately preceding the current day d.

P = the number of 15-minute periods being analyzed. For this example, there are twelve 15-minute periods starting at 16:00 and ending at 19:00.

2). For the link travel time dataset, historical average method is shown in equations 4.7a, b, and c.

$$\hat{\bar{T}}_{d,p}^{l,H} = \frac{1}{R} \sum_{r=1}^{R} \bar{T}_{d,r}^l \quad \forall l = 1, \dots, L; \; \forall d = 1, \dots, D; p = 1, \dots, P \qquad (4.7a)$$

$$(\widehat{s^2})_{d,p}^{l,H} = \frac{1}{R} \sum_{r=1}^{R} (s^2)_{d,r}^l \quad \forall l = 1, \dots, L; \; \forall d = 1, \dots, D; p = 1, \dots, P \qquad (4.7b)$$

$$\left[\widehat{\overline{\text{T}}}_{d,p}^{c,H}, \hat{s}_{d,p}^{c,H}\right] = F\left(\widehat{\overline{\text{T}}}_{d,p}^{l,H}, (\widehat{s^2})_{d,p}^{l,H}\right) \quad l = 1, \ldots, L; \; \forall d = 1, \ldots, D; p = 1, \ldots, P \qquad (4.7c)$$

where:

$\widehat{\overline{\text{T}}}_{d,p}^{l,H}$ = the predicted average link travel time on link l for day d and period p. The superscript l indicates that it is a link travel time and the superscript H indicates that it is based on the historical average method. The d represents the weekdays, in order, starting on June 1, 2011, and the p represents a particular 15 minute time period.

$\overline{T}_{d,r}^{l}$ = the average travel time on link l in period p and day r.

$(\widehat{s^2})_{d,p}^{l,H}$ = the predicted link travel time variance for day d and period p. The superscript l indicates that it is a link-level estimate, and the superscript H indicates that it is based on the historical average method.

$(s^2)_{d,r}^{l}$ = the travel time variance on link l in period p and day r.

R = the number of days used to calculate $\widehat{\overline{\text{T}}}_{d,p}^{l,H}$ and $(\widehat{s^2})_{d,p}^{l,H}$. In this example, R equals the 105, indicating the 105 weekdays immediately preceding the current day d.

$F(\cdot)$ = a function used to estimate the predicted average and SD of corridor travel time based on the link travel time predictions on all the component links l. The total number of the component links on the corridor is L. In this example, L is 3.

The four estimation methods for estimating $F(\cdot)$, discussed in section 3.3.2, are examined in this chapter. They are the naive sum method, the cumulative sum method,

the first-order approximation method, and the second-order approximation method. The naive sum method is based on the assumption of independence between links. If this assumption is true, then the corridor travel time mean and variance at the 15-minute interval p are simply the sum of the link travel time means and variances at the same interval p, respectively. The cumulative method also assumes independence between links but considers the arrival time at each link. The corridor travel time mean and variance at the 15-minute interval p are the sum of the link travel time means and variances at the corresponding time interval within which the vehicle arrives at that link. The first- and second-order approximation methods use second-order functions to fit time-dependent link travel time means and variances for each link. Complete details of these later two methods may be found in section 3.3.2.

4.5.1.1.2 Instantaneous Travel Time Method (ITT)

The ITT method is fairly straightforward in that the predicted travel time mean and SD in the next interval ($i+1$) is equal to the travel time mean and SD in the previous time interval, respectively. This prediction would be accurate if future demand and capacity were equal to the current period, which is true when the network is in equilibrium (Park 1998, Nikovski et al. 2005, Naik 2010).

1). ITT method using the corridor travel time dataset

The mean and SD of corridor travel time for the 15-minute interval p in day d are equal to the corridor travel time average and SD for the previous 15-minute interval $p-1$ in the same day, respectively. Equations 4.8a and 4.8b illustrate the ITT method used to

calculate the corridor travel time mean and SD, respectively, using the corridor-based travel time dataset.

$$\hat{\bar{T}}_{d,p}^{c,I} = \bar{T}_{d,p-1}^{c} \quad \forall d = 1, \dots, D \ \ \forall p = 1, \dots, P \tag{4.8a}$$

$$\hat{s}_{d,p}^{c,I} = s_{d,p-1}^{c} \quad \forall d = 1, \dots, D \ \ \forall p = 1, \dots, P \tag{4.8b}$$

where:

$\hat{\bar{T}}_{d,p}^{c,I}$ = the predicted average corridor travel time for day d and period p. The superscript c indicates that it is based on the corridor travel time dataset, and the superscript I indicates that it is calculated by the instantaneous travel time or real time method. The d represents the weekdays, in order, starting on June 1, 2011, and the p represents a particular 15-minute time period.

$\bar{T}_{d,p-1}^{c}$ = the measured average corridor travel time for day d and period p-1. The superscript c indicates that it is based on the corridor travel time dataset. The d represents the weekdays, in order, starting on June 1, 2011, and the p-1 represents the 15-minute time period immediately preceding the current time period p. Note that this parameter only includes those vehicles who drove the entire corridor and arrived at the end node at some point during time period p-1.

$\hat{s}_{d,p}^{c,I}$ = the predicted standard deviation of the corridor travel times for day d and period p, using the ITT method.

$s_{d,p-1}^{c}$ = the measured corridor travel time SD for day d and period p-1.

D = the number of days for which corridor travel time predictions are being calculated. For this example, there are D=43 days to predict, and these correspond to the 43 weekdays for June and July of 2011.

P = the number of 15-minute periods being analyzed. For this example, there are twelve 15-minute period starting at 16:00 and ending at 19:00. *P* is equal to 12.

2). ITT method using the link travel time dataset

The mean and variance of link travel time for the *p*-th 15-minute interval in the day *d* are equal to the link travel time average and variance for the (*p-1*)-th 15-minute interval in the same day, respectively. Equations 4.9a and 4.9b are used to calculate the link travel time mean and variance predictions, respectively, for the ITT method using the link-based travel time dataset. Equation 4.9c estimates the corridor travel time mean and SD for the *p*-th 15-minute interval in the *d*-th day using link-level predictions.

$$\widehat{\bar{\mathrm{T}}}_{d,p}^{l,I} = \bar{T}_{d,p-1}^{l} \quad \forall d = 1, \dots, D \quad \forall p = 1, \dots, P \tag{4.9a}$$

$$(\widehat{s^2})_{d,p}^{l,I} = (s^2)_{d,p-1}^{l} \quad \forall d = 1, \dots, D \quad \forall p = 1, \dots, P \tag{4.9b}$$

$$\left[\widehat{\bar{\mathrm{T}}}_{d,p}^{c,I}, \hat{s}_{d,p}^{c,I}\right] = F\left(\widehat{\bar{\mathrm{T}}}_{d,p}^{l,I}, (\widehat{s^2})_{d,p}^{l,I}\right) \quad l = 1, \dots L; \forall d = 1, \dots, D \quad \forall p = 1, \dots, P \tag{4.9c}$$

where:

$\widehat{\bar{\mathrm{T}}}_{d,p}^{l,I}$ = the predicted average travel time on link *l* for day *d* and period *p*. The superscript *l* indicates that it is based on the travel time dataset of link *l*, and the superscript *I* indicates that it is calculated through the

instantaneous travel time method. The *d* represents the weekdays, in order, starting on June 1, 2011, and the *p* represents a particular 15-minute time period.

$\bar{T}_{d,p-1}^{l}$ = the measured link travel time average of link *l* for day *d* and period *p-1*. The *d* represents the weekdays, in order, starting on June 1, 2011. The *p-1* represents the 15-minute time period immediately preceding the time period *p*. Note that this parameter only includes those vehicles who drove the link *l* and arrived at the end node at some point during time period *p-1*.

$(\widehat{s^2})_{d,p}^{l,I}$ = the predicted variance of travel time on link *l* for day *d* and period *p*.

$(s^2)_{d,p-1}^{l}$ = the measured link travel time variance for link *l* for day *d* and period *p-1*.

$F(\cdot)$ = a function for converting link-level predictions to corridor travel time predictions. *L* is 3 in this study. The naive sum method, cumulative sum method, first-order and second-order approximation methods are applied to build this function.

4.5.1.2 Neural Network

In reality, the underlying assumptions for the HTT and ITT methods described above do not always hold. Both future demand and capacity are stochastic and may not be equal to the historic averages or the states at the previous time interval. Plus, the future state is closely related to the past and current states. For example, during the shoulders of the rush hour period, traffic states can change rapidly from one interval to the next, and the current travel times might not accurately predict travel time in the next time interval.

It is hypothesized that a dynamic prediction model for the prediction of short-term travel time is needed to provide accurate and reliable results. In this dissertation, dynamic neural network models are developed, and their performance is compared with the two simple methods.

Various neural network (NN) structures have been applied for short-term travel time prediction, including feedforward neural networks, modular neural networks, and spectral basis neural netowrks (Naik 2010; Park and Rilett 1998, 1999; Clark et al. 1993; Smith and Demetsky 1994; Kisgyorgy and Rilett 2002; Park et al. 1999). Dynamic neural networks (DNN) have been studied as well. Based on the literature review, only two sub-classes of DNN state-space neural networks (SSNNs) and time-delayed state-space neural networks (TDSSNNs) have been used for travel time prediction (Liu et al. 2006; Van Lint et al. 2005; Shen and Huang 2011; Zeng and Zhang 2013). The sub-class of DNN selected for this study is the nonlinear autoregressive with exogenous inputs (NARX) model, which has a short-memory to account for current and previous exogenous inputs $x(n), x(n-1), \ldots x(n-q+1)$ and for delayed values of the estimated outputs $\hat{y}(n), \hat{y}(n-1), \ldots, \hat{y}(n-q+1)$, as shown in figure 4.10 (Haykin 1999). As a dynamically recurrent neural model, NARX feeds the true or estimated ouputs back to the input layer when they become available. This is different from the SSNN structure which feeds the estimates of the network internal states back to the input layer (Zeng and Zhang 2013). Using the true travel time observations as feedback enables the model to respond to unexpected traffic incidents quickly. This is different from a feedforward neural network which models the information flows in a unidirectional way. A feedforward neural network without this type of feedback loop is not sensitive to newly available travel time information. Using

the link-based dataset as an illustration, this case study also compares the performance of the NARX model to the feedforward neural network with the same input.
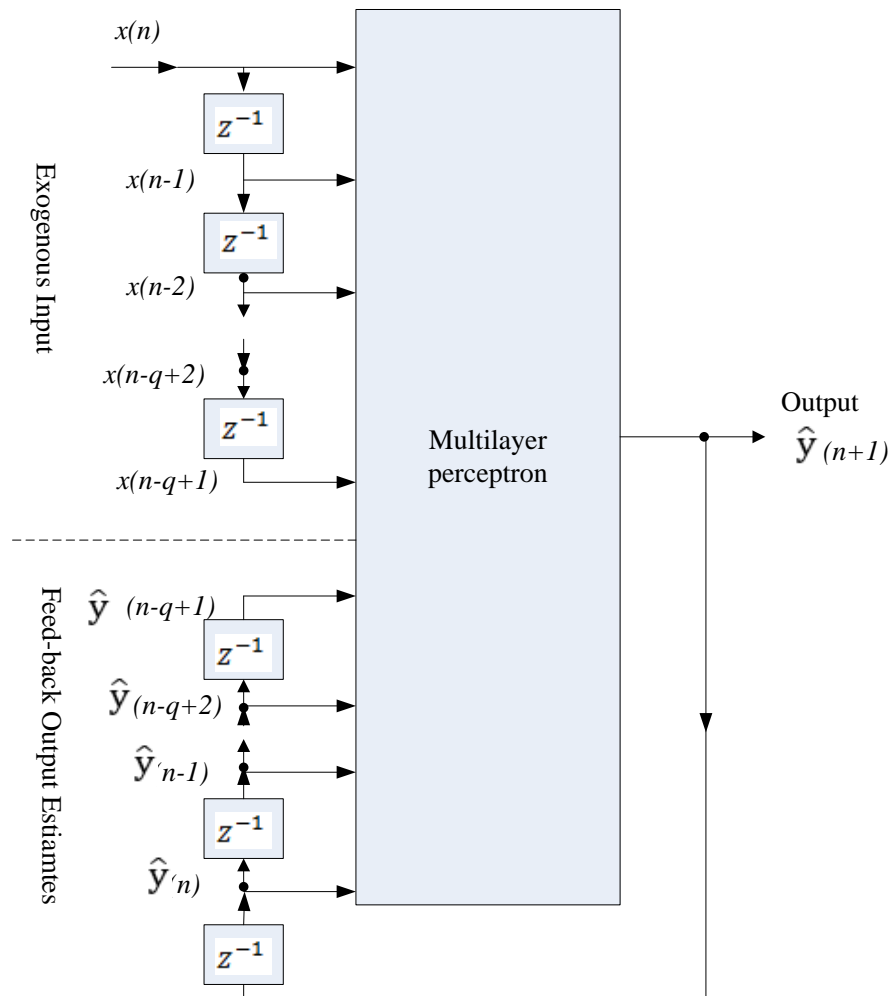
Figure 4.10 Nonlinear autoregressive with exogenous inputs (NARX) model

The dynamic behavior of the NARX model with estimated outputs as feedbacks, as shown in figure 4.10, can be described by equation 4.10.

$$\hat{y}(n+1) = F\big(\hat{y}(n), \dots \hat{y}(n-q+1), x(n), \dots, x(n-q+1)\big) \tag{4.10}$$

where:

$x(i)$ = the value of the input vector for the *i*th time interval, *i=n-q+1,…, n,*

$\hat{y}(i)$ = the value of the model output for the *i*th time interval, *i=n-q+1,…, n+1,*

q = the size of delay-line memories for the input and the recurrent output, and

F   = a nonlinear function of its arguments.

4.5.1.2.1 Corridor-Based NARX Model

The HTT and ITT methods discussed in section 4.5.1.1 do not require calibration because of their simplicity. In contrast, the NARX model needs calibration to identify the number of units for tapped-delay-line memories (*q*) and the number of neurons in the hidden layer (*n*). The selection of *q* and *n* are based on manually comparing the resulted mean square errors between the NARX network outputs and the target outputs, and choosing the model with the lowest mean square error. The number of hidden layers examined ranged from 7 to 12, and the number of units for tapped-delay-line memories ranged from 1 to 5. After comparing results from NARX with different numbers of delays and hidden neurons, the optimal values for *q* and *n* to predict the mean and standard deviation of corridor travel times are listed in table 4.7. Note that M-CTT stands for the mean corridor travel time, and SD-CTT stands for the standard deviation of corridor travel time.

Table 4.7 Structures of designed corridor-based NARX-1 models

Output $y(t)$	Input $x(t)$	Number of units for tapped-delay-line memories (q)	Number of neurons in the hidden layer (n)
M-CTT at interval (t)	M-CTT at interval $(t-1)$	4	10
SD-CTT at interval (t)	SD-CTT at interval $(t-1)$	1	10
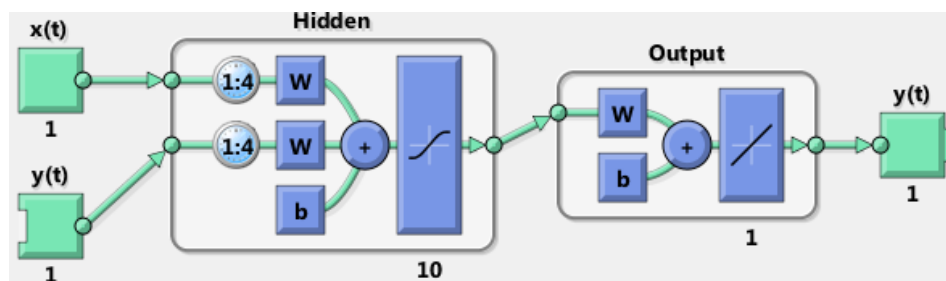
Figure 4.11 NARX neural network for the mean RTT prediction
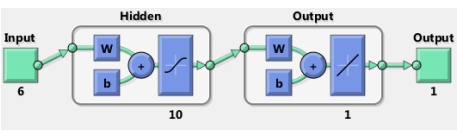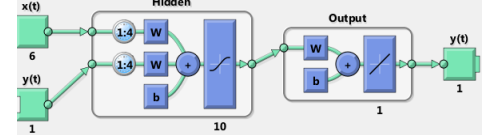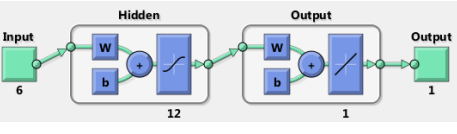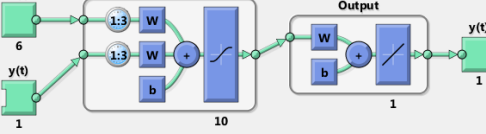
Figure 4.11 gives an illustration of the NARX structure for M-CTT prediction. The NARX model has 10 hidden neural neurons and 4 units of delay-line memories.

4.5.1.2.2 Link-Based Neural Network Model

The neural network model presented in this section predicts the corridor travel time mean and standard deviation using the link-based dataset as its input. The model structure implicitly accounts for correlation between links. The number of hidden layers examined ranged from 7 to 12, and the number of units for tapped-delay-line memories ranged from 1 to 5. Based on the lowest mean square errors between the network outputs

and the target outputs, the optimal values for q and n to predict the mean and standard deviation of corridor travel times are listed in table 4.8. The overall model structures for mean and standard deviation prediction are summarized in table 4.8. Both the feedforward and NARX models used the same input data and output format.

Table 4.8 Structures of designed neural networks

Output	Feedforward Network	NARX Network
Mean CTT	6 inputs: travel times and standard deviations on the three links	
STD CTT	6 inputs: travel times and standard deviations on the three links	

4.5.2 Prediction Results of Models Using Corridor-Based Data

The prediction results of the historical average travel time (HTT) model, the instantaneous travel time (ITT) model, and the NARX model using corridor-based travel time data are summarized in table 4.9. The results are separated into two parts. The first part shows the overall model performance for the PM peaks from 16:00 to 19:00 during June and July 2011. The second one shows the performance of the model during an example incident period between 17:30 and 19:00 on July 21, 2011. The percentage of arrival times that are within the predicted reliability interval is listed in the RI coverage rate cell. In terms of the overall performance of simple methods, the historical average

method performs better than the ITT method because it has: 1) a lower overall MAPE for the M-CTT prediction (e.g., 12 percent versus 17 percent), 2) a higher coverage rate (e.g., 98 percent versus 95 percent), and 3) a narrower reliability interval on average (e.g., 404 seconds versus 518 seconds). It is hypothesized that this occurred because: 1) the ITT method was negatively impacted by delayed corridor travel time observations, and 2) the historical traffic patterns of the corridor travel time are relatively steady for the period and days studied.

In general, the NARX neural network yielded the lowest overall MAPEs for both the mean and standard deviation predictions. Compared to the historical average method, the NARX-1 model has a 6 percent and 17 percent lower MAPE when predicting the M-CTT and SD-CTT, respectively.

It may be seen that during the non-recurring congestion period, the historical average and ITT approaches perform rather poorly as the coverage rates are only 23 percent and 19 percent, respectively. In contrast, the coverage rate of the NARX-1 model is 85 percent. The average range of the NARX model is only 86 seconds wider than those that resulted from the historical average method. Considering the 266 percent improvement in the coverage rate during the example incident period, it was hypothesized that an increase of 86 seconds in the RI interval range for a 5-kilometer corridor would be acceptable for drivers.

Table 4.9 Prediction efficiency of corridor-based models

1) Overall Performance				
Model	MAPE of M-CTT	MAPE of SD-CTT	RI Coverage Rate	RI Average Range (sec)
HTT	12%	29%	98%	404
ITT	18%	28%	95%	518
NARX	6%	12%	99%	490
2) Example Incident Period Performance				
Model	MAPE of M-CTT	MAPE of SD-CTT	RI Coverage Rate	RI Average Range (sec)
HTT	39%	34%	23%	404
ITT	34%	30%	19%	546
NARX	12%	49%	85%	489

Note: 1) All the 15-minute intervals from 16:00 to19:00 on weekdays of June and July 2011
      2) 15-minute intervals from 17:30 to 19:00 on July 21st, 2011

4.5.3 Prediction Results of Models Using Link-Based Data

4.5.3.1 Simple Methods

The first step of the prediction process is to predict the link travel time mean and

variance for the next 15-minute time interval for each link in the corridor. Once the

predicted link travel time mean and standard deviations are obtained, the corridor-level

prediction in the next 15-minute time interval will be calculated using the naïve sum

method, the cumulative sum method, the first-order, and the second-order approximation

methods.

The prediction results of the four estimation methods using link-level predictions

generated by the simple methods (i.e., HTT and ITT) are summarized in table 4.10. With

respect to the historical average prediction method, the results of the four corridor

estimation methods were essentially the same. For all four methods, the MAPEs for the

M-CTT and SD-CTT predictions were 14 percent and 22 percent, respectively. The

MAPEs for the example incident period were 28 percent and 24 percent, except for the naïve sum method.

For the ITT prediction method, the first- and second-order approximation methods performed the best for both M-CTT and RI predictions. However, it should be noted that the differences between the MAPEs of the first-order approximation and the second-order approximation are less than 1 percent. The approximation methods were found to be effective for the incident period when the approximation methods provided a higher coverage rate (e.g., 99 percent) with a lower average interval range (e.g., 520 seconds), as compared with the cumulative sum method that provided a coverage rate of 90 percent and an average interval range of 533 seconds.

Table 4.10 Prediction efficiency of link-based trivial models

Historical Average Travel Time Model				
1) Overall Performance				
Estimation Methods	MAPE of M-CTT	MAPE of SD-CTT	RI Coverage Rate	RI Average Range (sec)
Naïve Sum	14%	22%	98%	526
Cumulative Sum	14%	22%	98%	529
1st-order Approximation	14%	22%	98%	526
2nd-order Approximation	14%	22%	98%	526
2) Example Incident Period Performance				
Estimation methods	MAPE of M-CTT	MAPE of SD-CTT	RI Coverage Rate	RI Average Range (sec)
Naïve Sum	28%	27%	30%	524
Cumulative Sum	28%	24%	30%	514
1st-order Approximation	28%	24%	24%	509
2nd-order Approximation	28%	24%	24%	509
Instantaneous Travel Time Model				
1) Overall Performance				
Estimation methods	MAPE of M-CTT	MAPE of SD-CTT	RI Coverage Rate	RI Average Range (sec)
Naïve Sum	13%	46%	99%	525
Cumulative Sum	12%	48%	99%	531
1st-order Approximation	11%	45%	99%	527
2nd-order Approximation	11%	45%	99%	527
2) Example Incident				
Estimation methods	MAPE of M-CTT	MAPE of SD-CTT	RI Coverage Rate	RI Average Range (sec)
Naïve Sum	22%	82%	78%	537
Cumulative Sum	18%	66%	90%	533
1st-order Approximation	12%	59%	99%	520
2nd-order Approximation	12%	59%	99%	520

Note: 1) All the 15-minute intervals from 16:00 to19:00 on weekdays of June and July 2011
2) 15-minute intervals from 17:30 to 19:00 on July 21st, 2011

4.5.3.2 Neural Network Model

　　　Table 4.11 shows the results of the two neural network models using link-based travel time data. Compared to the feedforward network, the NARX-6 network MAPEs of M-CTT and SD-CTT were 3 percent lower and 1 percent lower for the overall condition, respectively. For the example incident period, the NARX-6 model had MAPEs of M-CTT and SD-CTT that were lower by 9 percent and 28 percent, respectively, as compared to the feedforward model. In addition, the NARX-6 model improved the coverage rate by 32 percent with a slight increase (e.g., 26 percent) on the average reliability interval range. It is hypothesized that these improvements are due to the NARX structure that feeds back real-time observations as part of its inputs, and that the resulting short-memory enables it to quickly adapt to unstable traffic conditions. This hypothesis, however, needs to be validated using additional empirical real data from unstable traffic conditions.

Table 4.11 Prediction efficiency of link-based NN models

1) Overall Performance				
Model	MAPE of M-CTT	MAPE of SD-CTT	RI Coverage Rate	RI Average Range (sec)
NARX	8%	20%	99%	492
Feedforward NN	11%	21%	99%	486
2) Example Incident Period Performance				
Model	MAPE of M-CTT	MAPE of SD-CTT	RI Coverage Rate	RI Average Range (sec)
NARX	11%	22%	85%	467
Feedforward NN	20%	50%	64%	395

Note: 1) All the 15-minute intervals from 16:00 to19:00 on weekdays of June and July, 2011
　　　2) 15-minute intervals from 17:30 to 19:00 on July 21st 2011

4.5.4 Overall Comparison and Conclusion

4.5.4.1 Mean Corridor Travel Time (M-CTT) Prediction

Based on the prediction results in tables 4.9, 4.10, and 4.11, the corridor-based NARX-1 model yielded the lowest overall MAPE of M-CTT predictions (6 percent), followed by the link-based NARX-6 model (8 percent). The link-based NARX-6 model generated the lowest MAPE of M-CTT for the example incident period (11 percent), followed by the corridor-based NARX-1 model (12 percent). These numbers show that, for situations without enough corridor-based travel time measurements, the NARX structure is a promising approach for developing a prediction model based on link-based travel time data with comparable efficiency to the corridor-based models.

Figures 4.12 and 4.13 compare the cumulative distribution function plots of the absolute percentage errors (APEs) of all the simple prediction models. Figure 4.12 compares the link-based HTT method using four link aggregation methods, with the corridor-based HTT methods. It was found that there was not much practical difference between the four link aggregation methods. This result is probably because the errors of link-level predictions using the historic average method are higher compared to the errors resulting from the estimation methods. The corridor-based historical average method outperformed all the other methods. In terms of the 90th percentile of APEs, the corridor-based historical average method improved the prediction results by 12 percent compared to the link-based simple methods.

Figure 4.13 compares the link-based ITT methods, using four estimation methods, with the corridor-based ITT method. The best model in this group is the link-based ITT prediction model with the second-order approximation algorithm, which has 90th

percentile APEs that are 12 percent and 38 percent lower compared with the link-based ITT methods using naïve sum or cumulative sum methods, and the corridor-based ITT method, respectively.
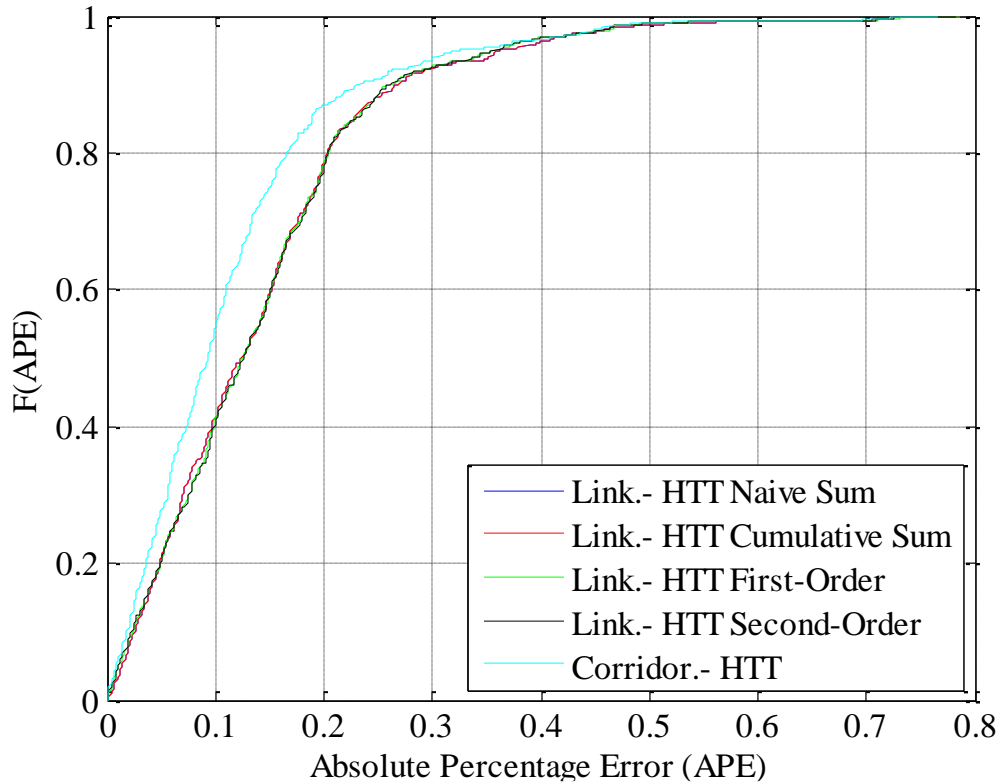
Figure 4.12 Simple methods comparison (corridor-based vs. link-based **HTT** methods)
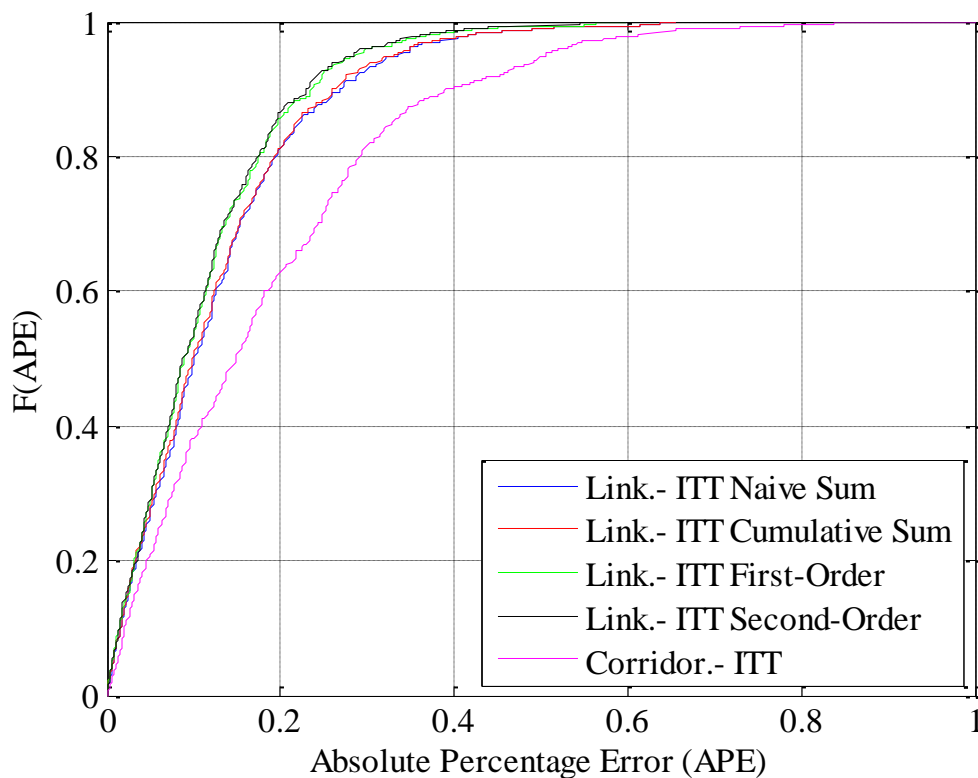
Figure 4.13 Simple methods comparison (corridor-based vs. link-based **ITT** methods)

Figure 4.14 compares the two "best" methods in the simple method category from figures 4.12 and 4.13 – the corridor-based HTT model and the link-based ITT model with the second-order approximation algorithm – to the neural network models. It can be seen that the feedforward neural network performed almost the same as the two "best" simple methods despite the extra requirements in structure design and data training. The NARX networks were able to reduce the 90th percentile of APE by half compared to the other three methods.
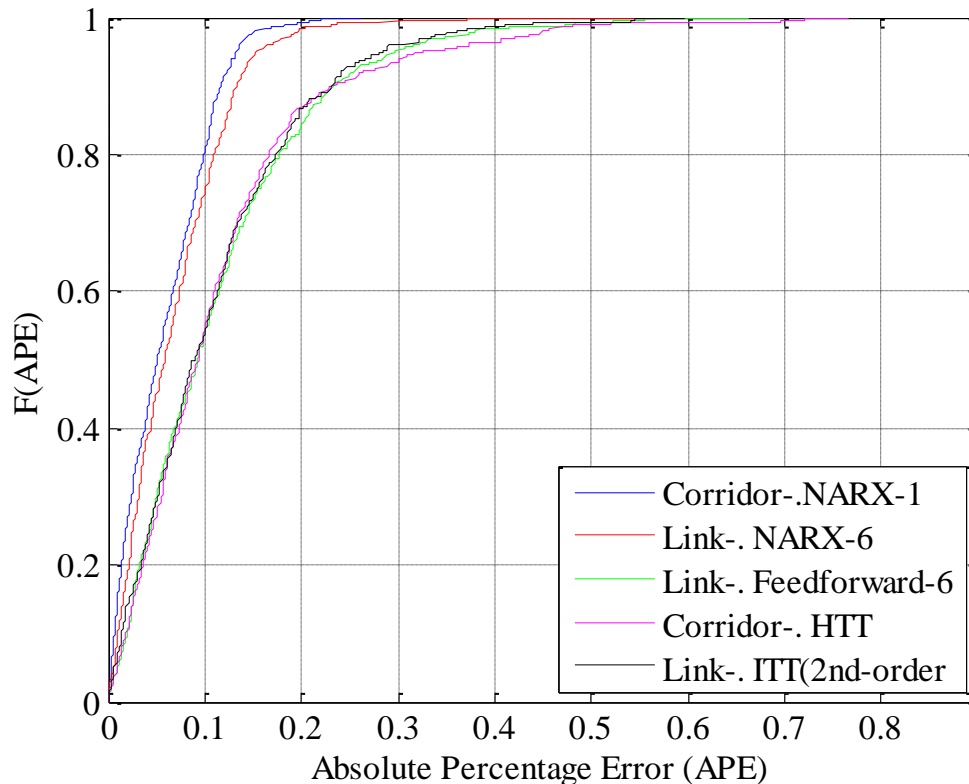
Figure 4.14 Outstanding simple methods vs. Neural network methods

4.5.4.2 Reliability Interval (RI) Prediction

As shown in table 4.9, the reliability interval of arrival time, based on one standard deviation, was able to effectively represent the corridor travel time in this case study because the corridor-based ITT model resulted in RIs with an overall coverage of 95 percent. A desirable prediction model needs to predict RIs with a high coverage rate and small range. Figures 4.15-4.18 compare the cumulative distribution function plots of the RI ranges of all the prediction models. In each figure, the coverage rates during the incident situation are shown in the legend.

Figure 4.15 shows that the CDFs of the interval ranges resulting from neural networks are located at the left of the CDFs corresponding to the historical methods. That indicates that neural network models are superior compared to the HTT method, given link-based travel time data as the input.

Figure 4.16 compares the link-based neural network models with link-based ITT prediction models. Although the link-based ITT methods can generate RIs with coverage rates comparable to the neural network models, as shown in the legend, their predicted RIs are generally wider, based on the location of corresponding CDF curves.

Figure 4.17 compares all the corridor-based prediction models, showing that the corridor-based NARX-1 model resulted in the most compact distribution of RIs and the highest coverage rate. More than 90 percent of the reliability intervals predicted by the NARX-1 model are within the range from 480 seconds to 520 seconds.

Figure 4.18 further compares all three neural network models. Overall, the distribution of RIs generated by the corridor-based NARX-1 network is the most compact. Given the slightly wider RIs of the link-based NARX-6 model, the coverage rate is 21 percent higher than that of the link-based feedforward neural network under the incident period. Therefore, the link-based NARX approach is still valuable for this particular application.
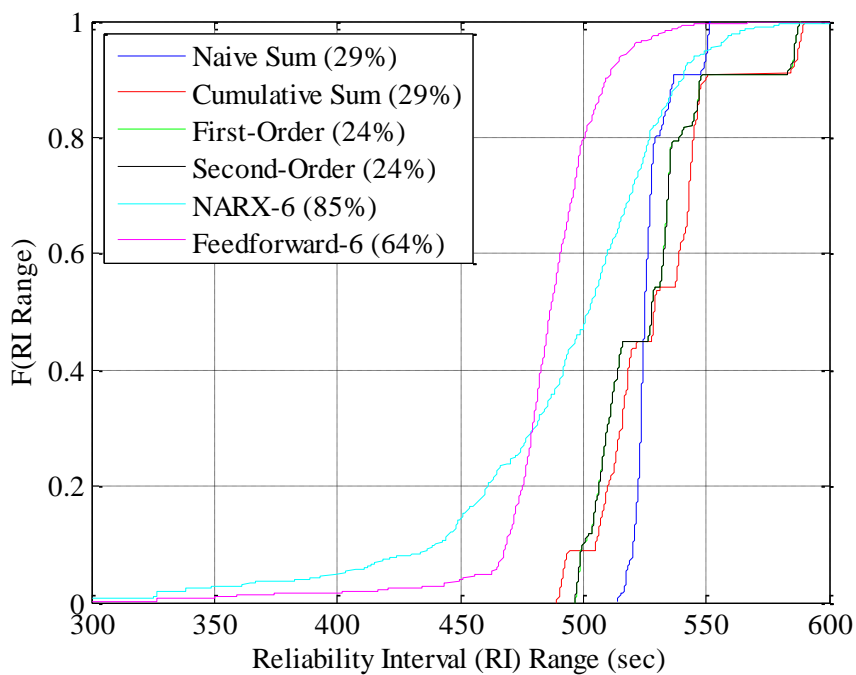
Figure 4.15 Link-based **HTT** methods vs. link-based NN methods
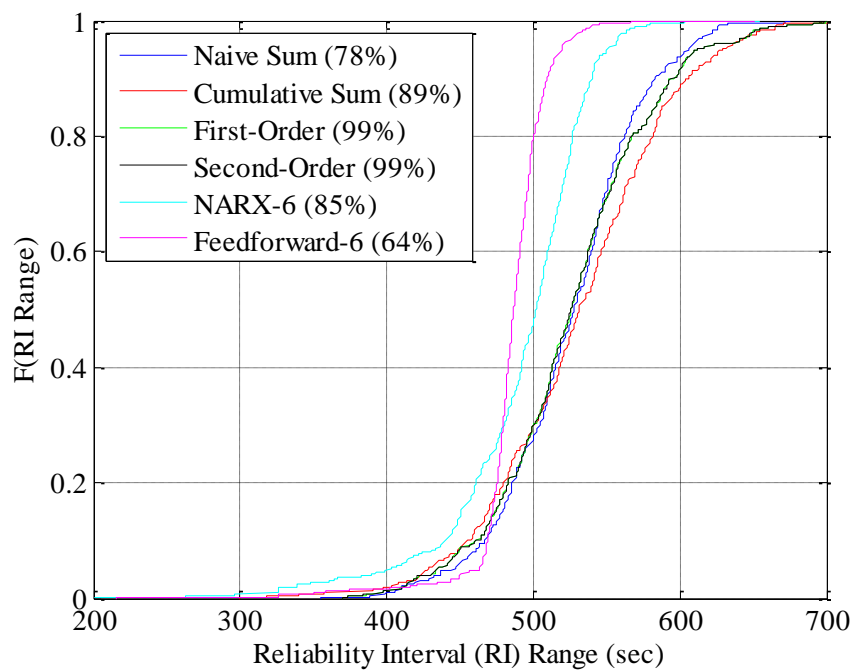
Figure 4.16 Link-based **ITT** methods vs. link-based NN methods
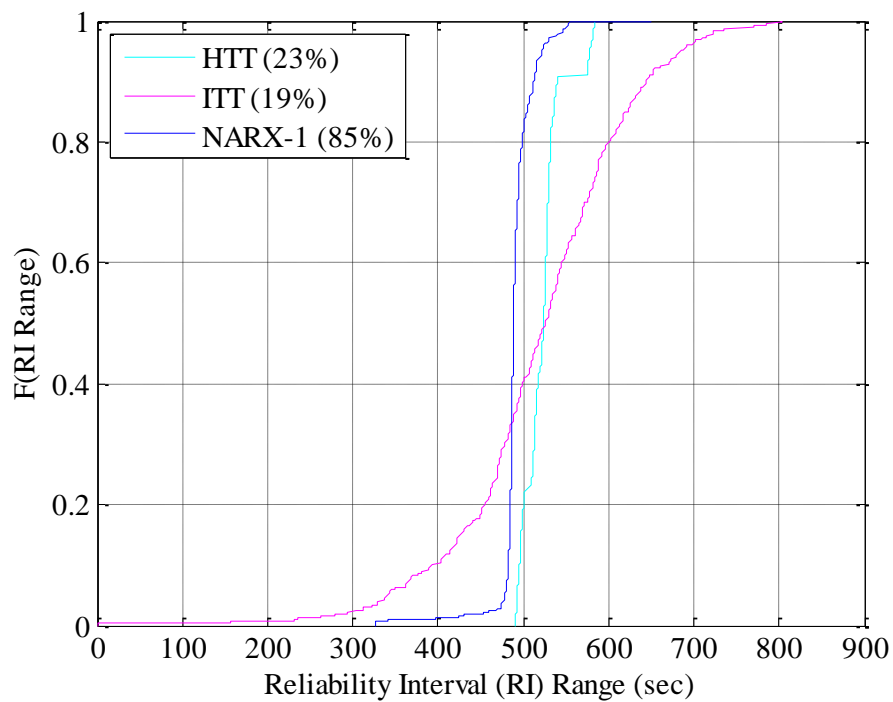
Figure 4.17 Corridor-based prediction methods
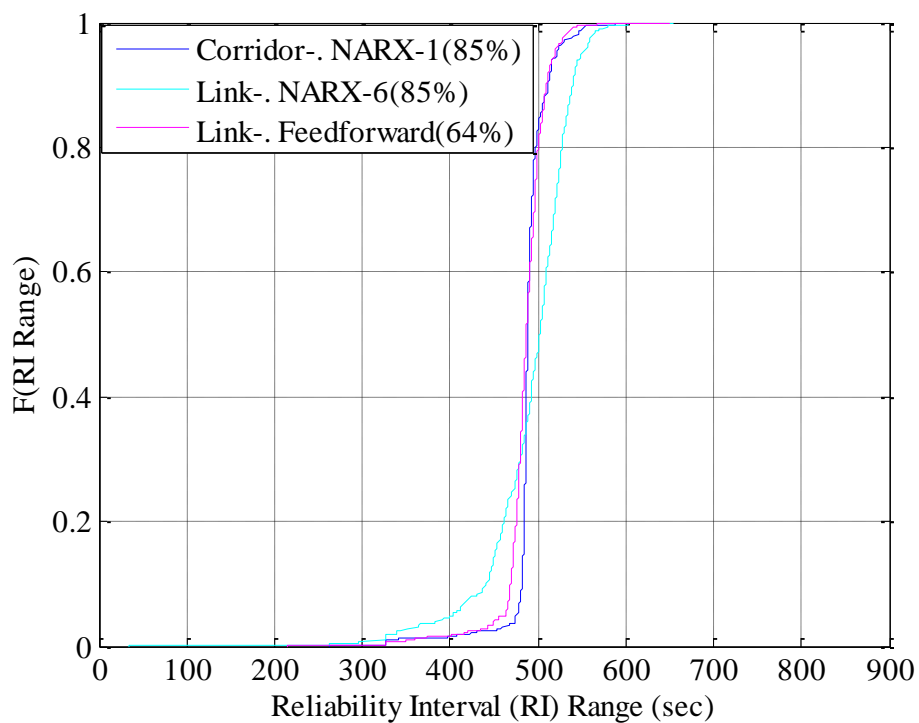
Figure 4.18 Neural network models

In practice, traffic agencies are interested in finding a solution which can maximize the coverage rate and minimize the interval ranges that are represented by the average interval range in this study. This type of decision making has two conflicting objectives. For the research results in this dissertation, there does not exist a single solution that optimizes both objectives simultaneously. In this case, a single or set of Pareto optimal solutions can be identified using figures 4.19 and 4.20. A solution is called Pareto optimal if one of the objective functions cannot be improved without degrading the other objective values. Without additional subjective preference information, all Pareto optimal solutions are considered equally good. Given an objective emphasized by traffic agencies (e.g., to maximize the reliability interval coverage during traffic incident periods or general peak periods), traffic engineers can choose the optimal application-specific model(s) according to figures 4.19 and 4.20.

Figure 4.19 plots the coverage rates and average ranges of various models under overall conditions. Overall, the link-based ITT prediction model with the approximation estimation method, the corridor-based historical average prediction model, the corridor-based NARX-1 model, and the link-based feedforward neural network model are all Pareto optimal solutions, each with its own advantage in either of the two objective functions. Agencies can make a final choice based on the data availability, computation capacity, and technician ability.

Similarly, if the information accuracy in the unexpected incident period is a particular concern for model choices, the Pareto optimal solutions for incident periods can be used to assist decision-making. Figure 4.20 plots the coverage rates and average ranges of the studied models under the selected incident period in this study. It can be

seen that all three of the Pareto optimal solutions are link-based models. These models are the link-based NARX-6 model, the link-based feedforward neural network model, and the ITT method with approximation estimation. However, this plot needs to be improved by including more incidents period data before it can be used in real applications.
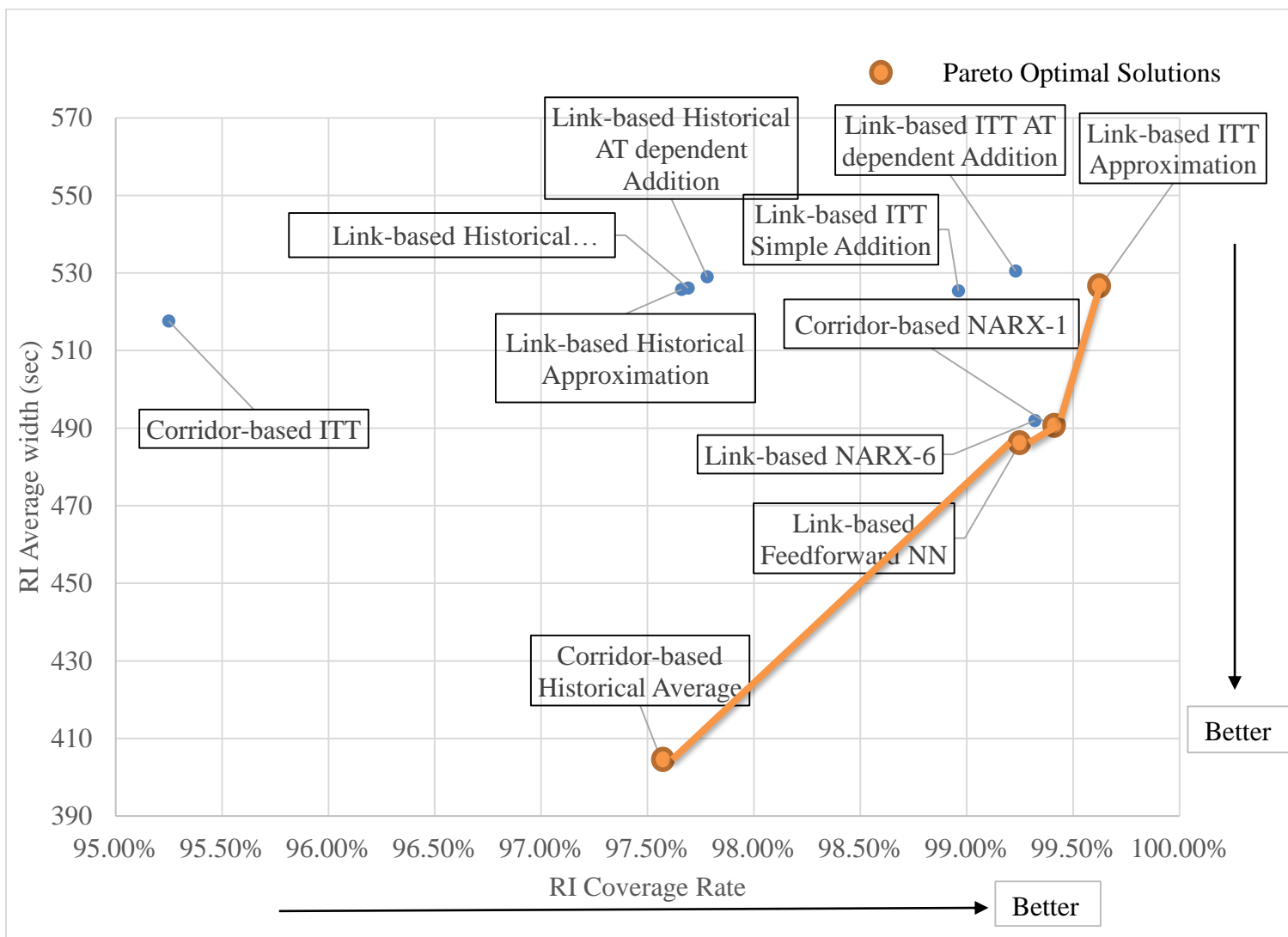
Figure 4.19 Pareto optimal solutions for the overall traffic condition
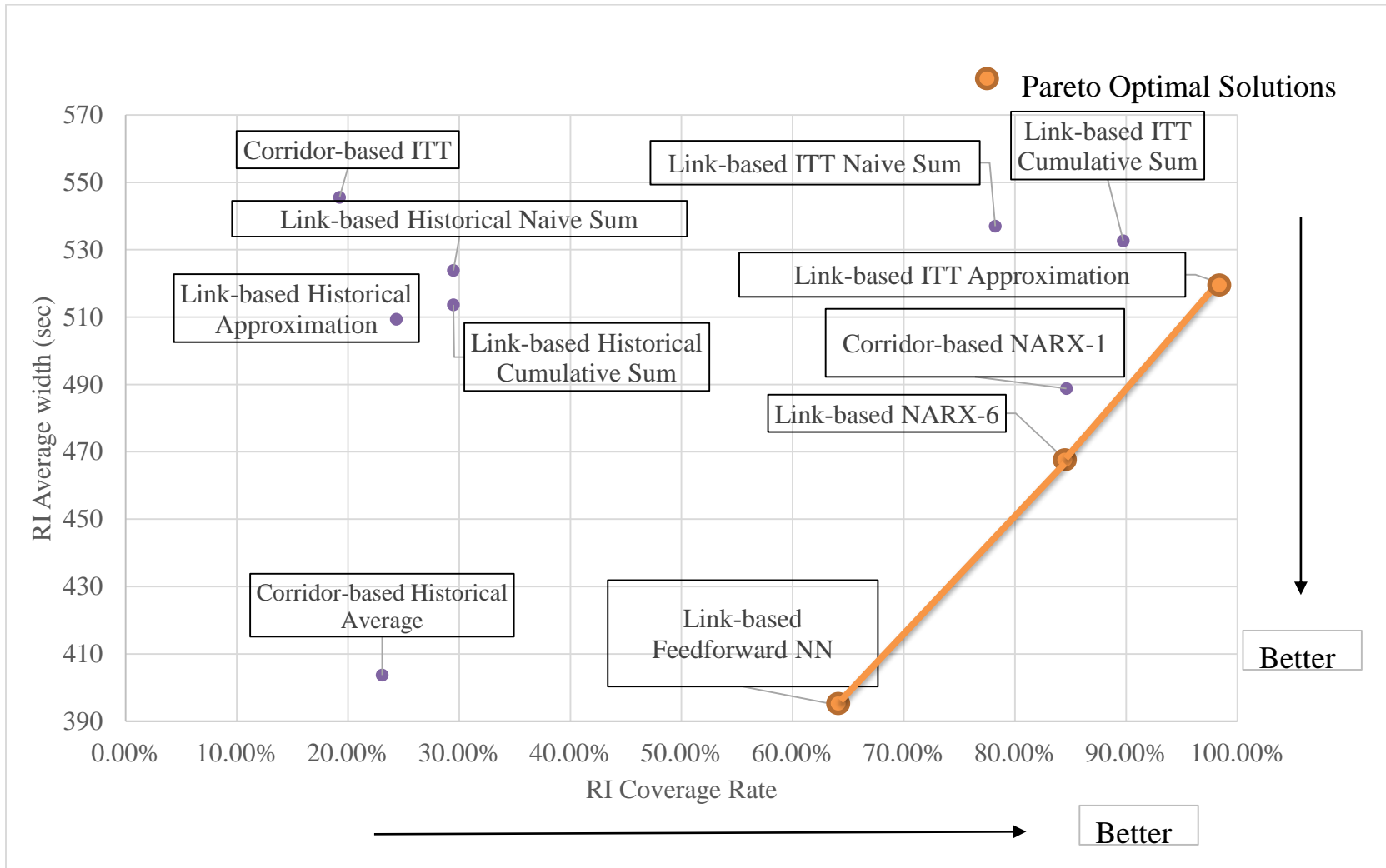
Figure 4.20 Pareto optimal solutions for the unexpected incident condition

4.5.4.3 Summary

After the comprehensive comparisons of various prediction methodologies, the NARX network models are recommended for both the corridor-based and link-based travel time data sets for predicting the mean corridor travel time and the reliability interval of the arrival time on a 15-minute level. For the applications of dynamic traffic information prediction, the NARX neural network, as a subclass of dynamically driven recurrent networks, can take advantage of the global feedback. This enables the recurrent network to acquire estimated or observed outputs and efficiently make predictions for nonlinear and adaptive situations. The travel time data from a real arterial corridor demonstrated the advantage and efficiency of the NARX neural network model in predicting both mean and reliability information.

For those simple models, using the approximation algorithm can help to improve the prediction performance on the corridor level while link-based data are used as input. Table 4.12 shows the aggregated results for the best HTT and ITT corridor prediction models as a function of input type (e.g., corridor travel times or link travel times). For the corridor travel time input, the historical average model had better overall performance than the ITT model. For the link travel time data set, the ITT resulted in lower MAPE in terms of corridor travel time mean prediction. During the example incident period, however, the link-based ITT model with the first- or second-order approximation algorithm yielded corridor travel time mean predictions with the lowest MAPE (e.g., 12 percent) and the arrival time reliability interval with the highest coverage rate (e.g., 99 percent). The high coverage rate, however, comes with a higher average interval range compared with other methods.

Table 4.12 Comparison of simple models

Overall Performance	Corridor Data Input		Link Data Input	
	HTT	ITT	HTT[1]	ITT[2]
MAPE of M-CTT	12%	18%	14%	11%
MAPE of SD-CTT	29%	28%	22%	45%
RI Coverage Rate	98%	95%	98%	99%
RI Average Range	404	518	526	527
Example Incident Period Performance	Corridor Data Input		Link Data Input	
	HTT	ITT	HTT[3]	ITT[3]
MAPE of M-CTT	39%	34%	28%	12%
MAPE of SD-CTT	34%	30%	24%	59%
RI Coverage Rate	23%	19%	24%	99%
RI Average Range	404	546	509	520

Note: 1) The four aggregation methods resulted in the same results for the HTT model.
2) The best of four aggregation methods for the ITT model is the first/second-order approximation algorithm.
3) The best of four aggregation methods for both HTT and ITT models is the first/second-order approximation algorithm.

CHAPTER 5   APPLICATIONS OF THE RELIABILITY METRICS IN A BI-MODAL

TRANSPORTATION NETWORK

In chapter 3 and 4, reliability metrics have been developed for network evaluation and traveler information systems. This chapter will apply those reliability metrics on the OD level to exemplify the potential benefits for both driver and system operator perspectives. To illustrate the concept, the impact of railway traffic on roadway traffic at highway railway at-grade crossings (HRGCs) will be analyzed using a simulation model from Lincoln, NE.

Traditionally, the Nebraska Department of Roads (NDOR) identifies potential locations for new grade separation structures based on exposure factor, crash costs, the elimination of vehicular delay, and other appropriate factors (NDOR 2011). For example, the minimum exposure factor of 50,000 for a single HRGC is required before a HRGC is considered for viaduct construction (NDOR 2011). The exposure factor is calculated as the product of the number of vehicles and the number of trains at a given HRGC for a day. The reliability metric for LOS proposed in chapter 3 enables traffic agencies to evaluate the HRGCs close to signalized intersections from the perspective of LOS reliability. The arrival time reliability interval that was presented in chapter 4 is applied to study the impact of train traffic on roadway travel time reliability. In addition, the chapter also provides an example of evaluating network service for a given OD pair based on the reliability intervals of multiple routes.

The data for this study is collected from a calibrated micro-simulation model. Micro-simulation models can capture characteristics of real transportation systems

including emergent properties such as capacity and congestion (Nagel and Rasmussen 1994). The output from the model includes both aggregated (e.g., average link travel time) and disaggregate data (e.g., link travel time for individual vehicles). In most micro-simulation models, vehicle speed and location data at the end of each simulation step (e.g., 1 sec.) can be output. Aggregated statistics such as the maximum, mean, and standard deviation for traffic parameters (e.g., speed) can be recorded for pre-set simulation intervals at specific points or links in the network. In addition, simulation models can examine scenarios under varying railway demands. All these elements make micro-simulation models a useful tool for studying the reliability of transportation networks, such as the test bed with HRGCs used in this dissertation. The existence of train traffic and HRGCs generates short-term interference to roadway traffic. Applying reliability-based metrics to monitor and manage such bimodal transportation networks has the potential to improve the efficiency of decision-making for both traffic engineers and drivers.

5.1 Simulation Model Setups

5.1.1 Benefits of Using Simulation Models for Reliability Analysis

In this dissertation, the micro-simulation software VISSIM is selected as the simulation tool. There are three benefits to using a well-calibrated micro-simulation model for the reliability-based evaluation of traffic systems.

1) The ability to record and output performance measures for individual vehicles are necessary for considering short-term traffic fluctuations. The availability of real data is becoming more prevalent with the deployment of probe-based travel time collection,

and the archiving systems for roadway segments have been implemented by many private sector companies (Winick 2012). Currently, probe-based data is mainly collected on freeways, on major and some minor urban arterials, and on some non-urban highways for business. For the roadways without real travel time data, simulation models can provide simulated travel times in exchange. The methodology using simulation data developed in this dissertation can readily be generalized to real-time systems once more empirical data become widely available.

2) Explicit assumptions inherent in aggregated models are not required, such as requiring capacity to be input rather than treated as an emergent characteristic. It is hypothesized that this results in a more realistic model for this application.

3) Traffic designers can test various proposed strategies and designs before implementation. The results of the reliability evaluation of various alternatives could assist the decision-making process. This saves time and money by allowing designers to exclude poor-performance options. As will be shown in this chapter, various train volumes, speeds, and lengths can be simulated easily.

5.1.2 A Bimodal Simulation Model

The bimodal transportation network used in this study is bounded by Cornhusker Highway, Holdrege Street, 27th Street, and 48th Street, as illustrated in figure 1.8. This is a 2.4 km by 3.2 km urban transportation network that includes three HRGCs. Corresponding to the physical map in figure 1.8, figure 5.1 is the simulation model of the network in the VISSIM environment. The blue lines represent a 3.2 km section of the Burlington Northern Santa Fe (BNSF) railroad tracks, while the dark-gray lines represent

the roadway transportation network. The three red circles represent the locations of the HRGCs, and the blue rectangle represents a grade separated roadway-railway crossing. The 16 numbered points (e.g., in brackets) shown in figure 5.1 indicate the locations where traffic volume data were collected by the City of Lincoln. This data will be used to calibrate the model.

The test network was modeled using dynamic traffic assignment in order to better capture the interaction between supply and demand. In a dynamic assignment, the travel time is a function of volume, and traffic demand will spread out over routes based on the route travel times.

In this dissertation, the traffic demand is added to the network at 5 percent increments. At each iteration, the demand is loaded to the path with the shortest travel time. After each iteration, the travel times are updated and the process repeats. After 20 iterations (e.g., 20×5%=100%), the network is fully loaded.
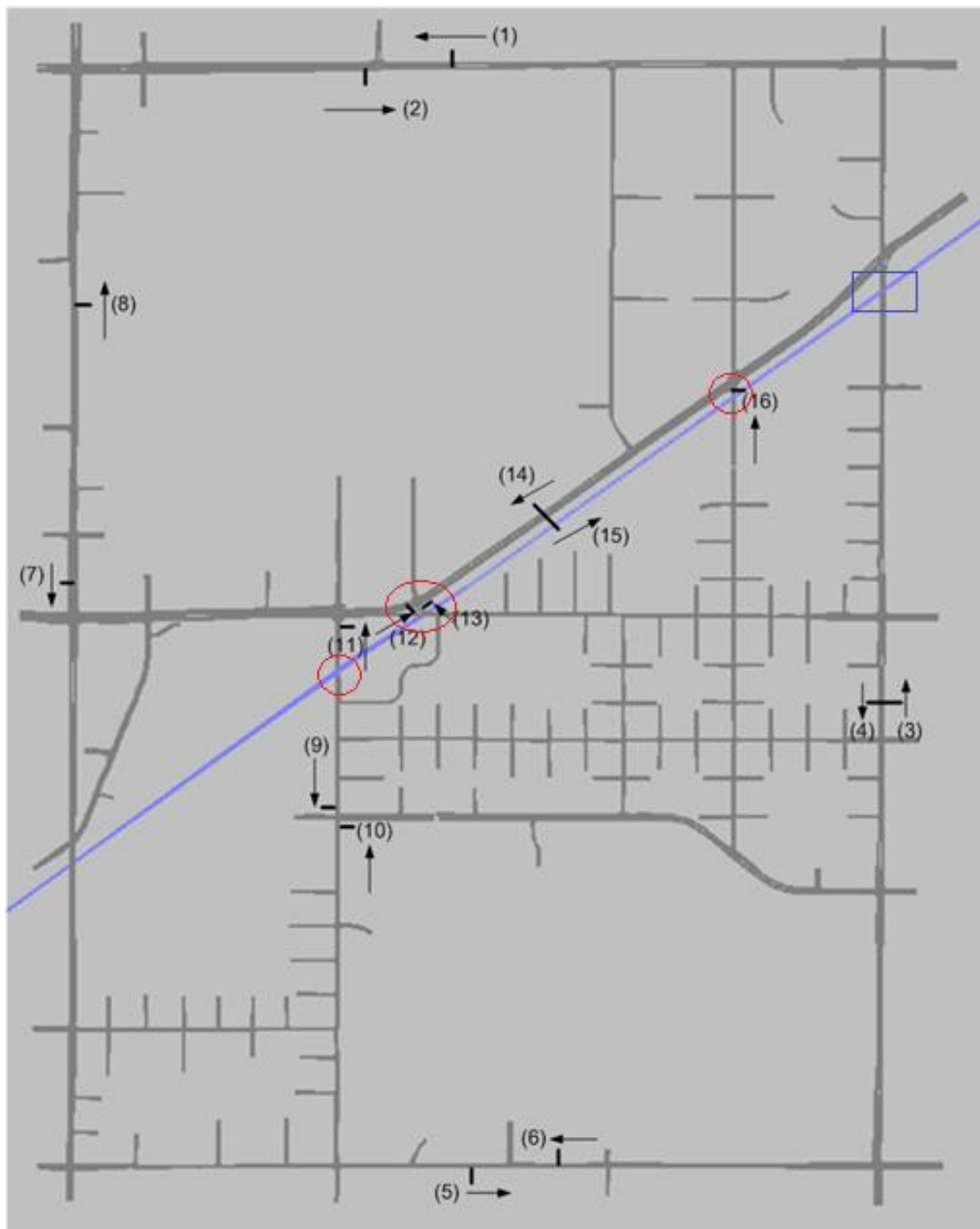
Figure 5.1 The simulation model in VISSIM of the bimodal transportation network

5.1.3 Railway Traffic Modelling

The three signalized intersections located at points (11), (13), and (16) in figure 5.1 are all within 200 meters of a HRGC. Consequently, the traffic signal timings follow standard train preemption logic when a train is present. The goal is to operate the traffic signal to allow vehicles to clear the HRGC before a train arrives. The train preemption logic of these three intersections were programmed using VISVAP in the VISSIM environment.

VISSIM requires users to first define vehicle classes before these classes can be used in a simulation. In this bimodal simulation model, a vehicle class named "train" is used to represent trains. The class has both length and speed attributes. Table 5.1 shows the train length distribution input in VISSIM. The share was identified based on an empirical train length distribution as shown in figure 5.2. During the simulation, VISSIM will generate trains with a length randomly drawn from the distribution shown in table 5.1. Note that VISSIM 5.40 does not allow direct input of train lengths. Instead, it has built-in train components (e.g., locomotives and train cars with given lengths). The train lengths are generated internally to VISSIM based on the input number of train components. It may be seen in figure 5.2 that the simulated train length distribution closely follows the observed train length distribution.

Table 5.2 shows the empirical speed distribution collected by Doppler radar located at the intersection of the Salt Creek Roadway and North Antelope Valley Parkway (Chen 2015). In contrast to vehicle length, VISSIM allows for direct input of speed distributions. Figure 5.3 displays the empirical speed distribution that is also used in the VISSIM simulation model.

Table 5.1 Simulated train length distribution

Share	Train Length (m)	Number of Locomotives	Number of Train Cars
0.1	998.6	1	37
0.1	1737.8	1	65
0.1	2213.0	1	83
0.5	2661.8	1	100
0.2	2978.6	1	112

Note: 1) The default locomotive length is 21.8 meters in VISSIM.
  2) The default length of a train car is 26.4 meters in VISSIM.
  3) Train Length = Locomotive Length + Number of Train Cars * Car Length.
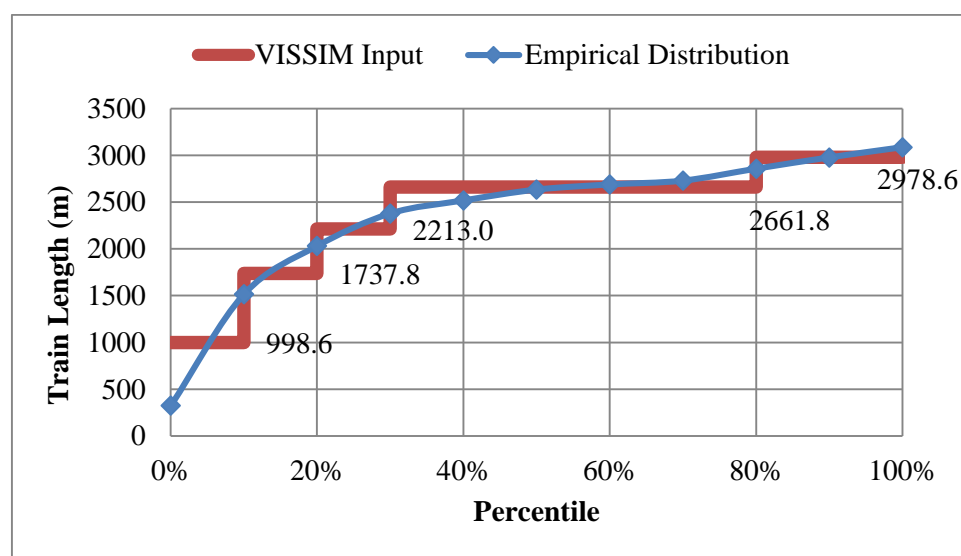
Figure 5.2 Empirical and simulated train length distributions

Table 5.2 Train speed distribution

Train Speed (km/h)	Cumulative Percentile
8.1	0%
22.7	10%
26.5	20%
29.7	30%
32.5	40%
34.8	50%
35.5	60%
39.5	70%
41.4	80%
45.9	90%
64.2	100%
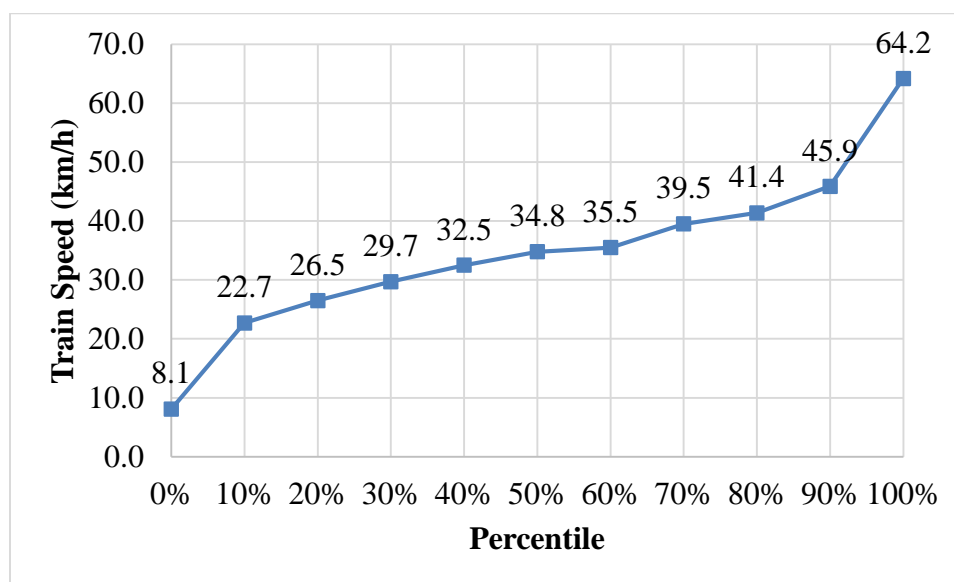
Figure 5.3 Simulated train speed cumulative distribution input into VISSIM

<u>5.2 Calibration and Validation</u>

5.2.1 Calibration Parameters

There are 24 parameters related to car following and lane changing logic in VISSIM. Based on previous calibration research on this network by Appiah in 2009, the eight parameters shown in table 5.3 will be used in the calibration process. Seven of them are VISSIM parameters. The last one, demand proportion, is incorporated to account for possible demand increases on this network after previous research. In addition, the calibration effort will directly account for the effect of trains. Note that trains were not analyzed in previous research. These eight parameters are described below and the default values are shown in table 5.3 (PTV 2011).

Waiting time before diffusion defines the maximum length of time a vehicle can wait at emergency stop positions for an acceptable gap to change lanes so that it can stay on its route. The default value is 60 seconds.

Minimum headway defines the maximum distance to the leading vehicle that must be available for a lane change in standstill condition. The default value is 0.5 meters. As this value decreases, the road capacity increases.

The *number of observed vehicles* determines the ability of vehicles in the network to predict the movement of other vehicles, and to react accordingly. The default value for urban driving behavior is 4. As this value increases, the run time of simulation also increases.

Maximum look-ahead distance defines the maximum distance allowed for drivers to "look ahead" and to react accordingly. Note that any event occurring outside this range will be ignored by the driver. The default value is 250 meters.

Average standstill distance, the additive and multiplicative part of the desired safety distance are model parameters for the Wiedemann 74 car following model that is suitable for urban traffic conditions. The distance *d* between two vehicles in meters is calculated by equation 5.1 in VISSIM.

$$d = ax + (bx_{add} + bx_{mult} \cdot z) \cdot \sqrt{v} \qquad (5.1)$$

where:

d $\quad$ = the average standstill distance,

ax $\quad$ = the average standstill distance that defines the average desired

$\qquad$ distance between stopped cars,

bx_{add} $\quad$ = the additive part of the desired safety distance,

bx_{mult} $\quad$ = the multiplicative part of the desired safety distance,

z $\quad$ = a value in the range [0, 1] normally distributed around 0.5 with a

$\qquad$ standard deviation of 0.15, and

v $\quad$ = the vehicle speed in m/s.

Table 5.3 Default values of model parameters

Calibration Parameter	Default Value
Waiting time before diffusion (sec)	60
Minimum headway (m)	0.5
Number of observed vehicles	4 - urban 2 - others
Max look-ahead distance (m)	250
Average standstill distance (m)	2
Additive part of safety distance	2
Multiplicative part of safety distance	3
Demand proportion	1.00

5.2.2 Calibration Algorithm

A genetic algorithm (GA) toolbox, first developed by the University of Sheffield (Yu et al. 2003), and subsequently revised by Wojtal (2012), is used to calibrate the simulation model. Figure 5.4 is a flowchart of the genetic algorithm. A randomly generated set of feasible chromosomes is used as the first generation to start the GA procedure. In this study, the generation set includes 30 chromosomes. Each chromosome corresponds to a solution of simulation parameters. The process is repeated until a pre-specified maximum number of iterations are completed. The maximum number of iterations used in this dissertation is 40. The best solution X* is the output of the calibrated parameters. Each step will be discussed with more details in the following sections.
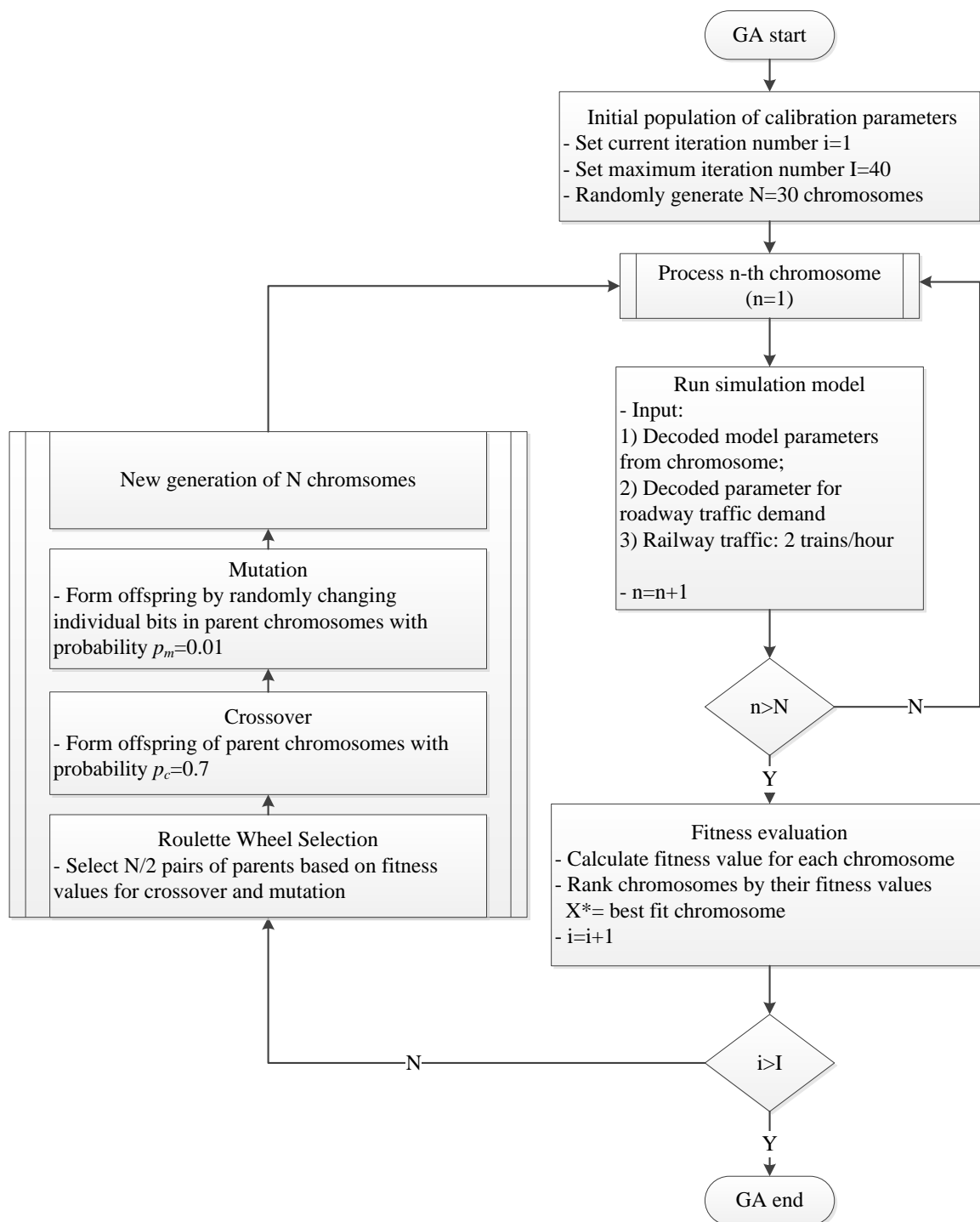
Figure 5.4 Genetic algorithm flowchart

5.2.2.1 Parameter Coding

Typical GA procedures require potential solutions to be encoded as a binary bit string that is called a chromosome. Each parameter to calibrate is encoded as a number of genes in the form of a chromosome. The number of genes needed for parameter x_i is calculated by equation 5.2.

$$n_i = Log(\frac{\max(x_i) - \min(x_i)}{a_i} + 1) \qquad (5.2)$$

where:

x_i = the *i*-th parameter to calibrate (i=1, … 8),

n_i = the number of genes necessary to encode parameter x_i,

$\max(x_i), \min(x_i)$ = the upper and lower bounds marking the range of the potential

            parameter values, and

a_i = the increment to change parameter values in GA procedures,

            which is calculated by equation 5.3.

$$a_i = \frac{\max(x_i) - \min(x_i)}{(2^{n_i} - 1)} \qquad (5.3)$$

The optimal values of a_i and n_i are determined by a trial and error procedure. The final results are summarized in table 5.4. Table 5.4 also lists all the upper and lower bounds used in this dissertation. They need to be determined based on the default values and engineering judgment. These bounds were based on previous studies (Appiah 2009).

Table 5.4 Encoding of calibration parameters

Parameters to calibrate (x_i)	Max(x_i)	Min(x_i)	Increment (a_i)	Number of genes (n_i)
1.Waiting time before diffusion (sec)	90	30	4	4
2.Minimum headway (m)	1	0.1	0.06	4
3.Number of observed vehicles	4	1	1	2
4.Max look-ahead distance (m)	300	200	6.67	4
5.Average standstill distance (m)	5	0.5	0.145	5
6.Additive part of safety distance	10	1	0.6	4
7.Multiplicative part of safety distance	10	1	0.6	4
8.Demand proportion	1.50	1.0	0.016	5

Once the GA algorithm is done, the chromosome solutions are translated into simulation parameters by equation 5.4. Table 5.5 shows one example chromosome solution in binary and decimal formats. The conversion is based on equation 5.4,

$$x_i = \min(x_i) + A \cdot \frac{\max(x_i) - \min(x_i)}{(2^{b_k} - 1)} \tag{5.4}$$

where:

b_k = the length of binary bit string corresponding to parameter x_i, and

A   = the value of binary bit string to base 10 of parameter x_i.

Table 5.5 Example of converting binary chromosome to simulation parameters

Parameters to calibrate (x_i)	Chromosome part (Binary format)	Simulation parameter (Decimal format)
1.Waiting time before diffusion (sec)	1110	60
2.Minimum headway (m)	0001	0.16
3.Number of observed vehicles	01	2
4.Max look-ahead distance (m)	1010	266.7
5.Average standstill distance (m)	00001	0.645
6.Additive part of safety distance	0010	2.2
7.Multiplicative part of safety distance	0010	2.2
8.Demand proportion	00011	1.048

Note: The chromosome solution is: 1110000101101000010010001000011.

5.2.2.2 Fitness Evaluation

A number of functions can be used to evaluate the "fitness" of a given chromosome (i.e., parameter vector). In this dissertation, the fitness values are calculated through the Geoffrey E. Havers (GEH) statistic. The model is calibrated to the traffic volumes measured at the 16 locations shown in figure 5.1. The objective of the calibration is to find the "best" set of values of the eight parameters that replicate the observed volumes. The parameter solutions with lower GEH values provide a more accurate simulation with respect to the traffic volumes. GEH is selected as the fitness objective due to its self-scaling feature. That is, a single acceptance threshold based on the GEH statistic can be used over a fairly wide range of traffic volumes (Appiah 2009). For example, consider two roads. The first has an observed volume of 100 veh/h and a simulated volume of 200 veh/h. The second has an observed volume of 1,100 veh/h and a simulated volume of 1,200 veh/h. If the mean absolute percentage error is used to evaluate the calibration error, it will be very difficult for researchers to select a single

percentage as the threshold to evaluate both roads. The GEH statistic, in a form of the

chi-squared statistic that incorporates both relative and absolute errors, reduces this

problem (Appiah 2009).

In this dissertation, the fitness value of each chromosome is calculated as an

averaged GEH value at all the 16 data collection points after the simulation model is run.

The GEH statistic at each point is calculated by equation 5.5.

$$GEH_i = \sqrt{\frac{2(V_i - v_i)^2}{(V_i + v_i)}} \qquad i = 1, \dots N \qquad (5.5)$$

where:

V_i = the simulated traffic volume at the data collection point i,

v_i = the observed traffic volume at the data collection point i, and

N   = the number of locations with observed traffic volume, N=16.

According to a technical report from the Oregon Department of Transportation

(DOT), the goodness of fit can be evaluated using the following rules (Oregon DOT

2006). If the GEH<5, the solution is a good fit to the problem. If 5<GEH<10, the solution

needs further investigation. If the GEH>10, the solution is a poor fit.

5.2.2.3 Genetic Operators

The next generation of chromosomes is generated from the current generation

through three genetic operators: selection, mutation, and crossover. The fitness values of

the 30 chromosomes in the current generation are calculated, and each chromosome is

ranked based on its fitness value. The selection operator selects the parents for the next generation based on the rank of each chromosome. The better the chromosomes are, the higher their chance of selection. The crossover operator exchanges the genes of two parents to create new offspring with a probability of 0.7. The mutation operator changes the bits of new offspring with the probability of 0.01. Complete details of the GA operators used in this dissertation may be found in Cao and Wu, 1999.

5.2.3 Calibration Results

After 40 iterations, each with 30 chromosomes, the best GEH obtained was 5.27. Although this value is slightly higher than the recommended value of 5, it is regarded as acceptable considering the large size of the network in this study. The calibrated parameter values are shown in table 5.6.

Table 5.6 Calibrated values of simulation parameters

Calibrate parameter	Default value	Calibrated value
Waiting time before diffusion (sec)	60	42
Minimum headway (m)	0.5	0.34
Number of observed vehicles	4	3
Max look-ahead distance (m)	250	266.7
Average standstill distance (m)	2	0.5
Additive part of safety distance	2	5.8
Multiplicative part of safety distance	3	8.8
Demand proportion	1.00	1.02

Figure 5.5 is a scatter plot of the simulated hourly traffic volumes and the observed traffic volumes for both the calibrated and un-calibrated cases. The plot shows that most of the calibrated scatter points are close to the 45-degree line, which indicates

that the calibrated parameters provide acceptable simulation results compared to the real traffic observations. The coefficient of correlation is 0.70, indicating a relatively strong linear relationship between the observed and simulated volume counts after calibration. In contrast, the un-calibrated scatter points generally have higher simulated volumes than the calibrated ones. Overall, the coefficient of correlation is 0.55 for the un-calibrated scenario.
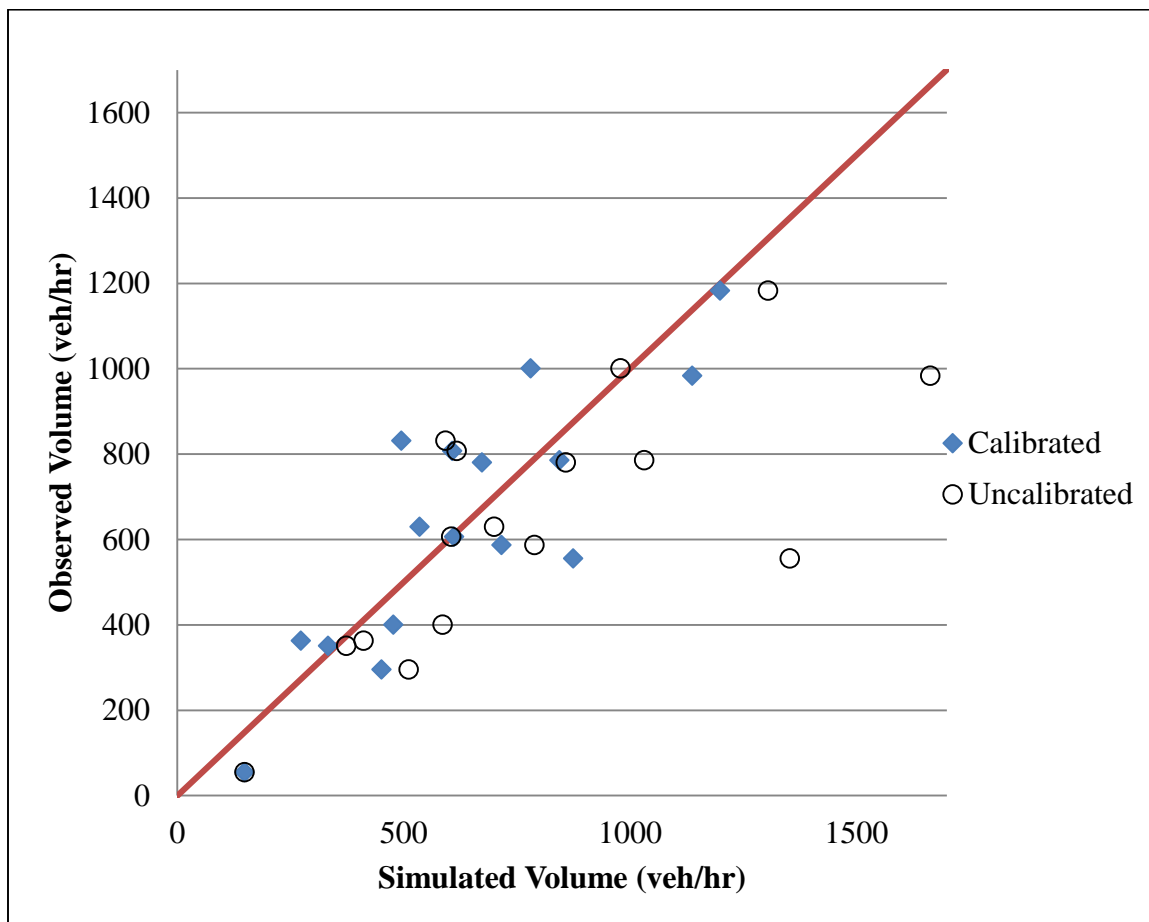
Figure 5.5 Simulated volume versus observed volume in field

5.2.4 Validation Results

The calibration algorithm adjusts the default parameters in VISSIM to make the

simulation model produce reasonable traffic volumes at the 16 locations shown in figure

5.1. It is important to validate the calibrated simulation model using simulation output

different from the output used for calibration.

The validation process is based on the turning ratios at the six intersections shown

in figure 5.6. The blue lines in figure 5.6 represent the railroad tracks in this test network.

Intersections II, III, and IV are the HRGC-related intersections. The mean absolute

percentage errors (MAPE) between the simulated and observed turning ratios are calculated for each movement and weighted by traffic volumes to yield the weighted MAPE for each intersection, as summarized in table 5.7. Table 5.7 lists the weighted MAPEs for both the before- and after-calibration cases. It can be seen that the calibration improves the weighted MAPE of turning ratios for intersection I, III, IV, and V, by 10 percent, 7 percent, 10 percent, and 12 percent, respectively. After calibration, the weighted MAPE for intersection II becomes 7 percent higher than the before-calibration network. On average, the calibration improves the weighted MAPE of turning ratios by 6 percent. Based on these results, it was decided that the calibrated VISSIM model could be used for reliability analyses.

Table 5.7 Weighted MAPE of turning ratio

No.	Intersection Name	Weighted MAPE (Before Calibration)	Weighted MAPE (After Calibration)	Improvement
I	29th &Cornhusker	7%	6.%	10%
II	33rd &Cornhusker	14%	15%	-7%
III	35th &Cornhusker	27%	25%	7%
IV	44th &Cornhusker	21%	19%	10%
V	33rd &Huntington	25%	22%	12%
Average		19%	17%	6%
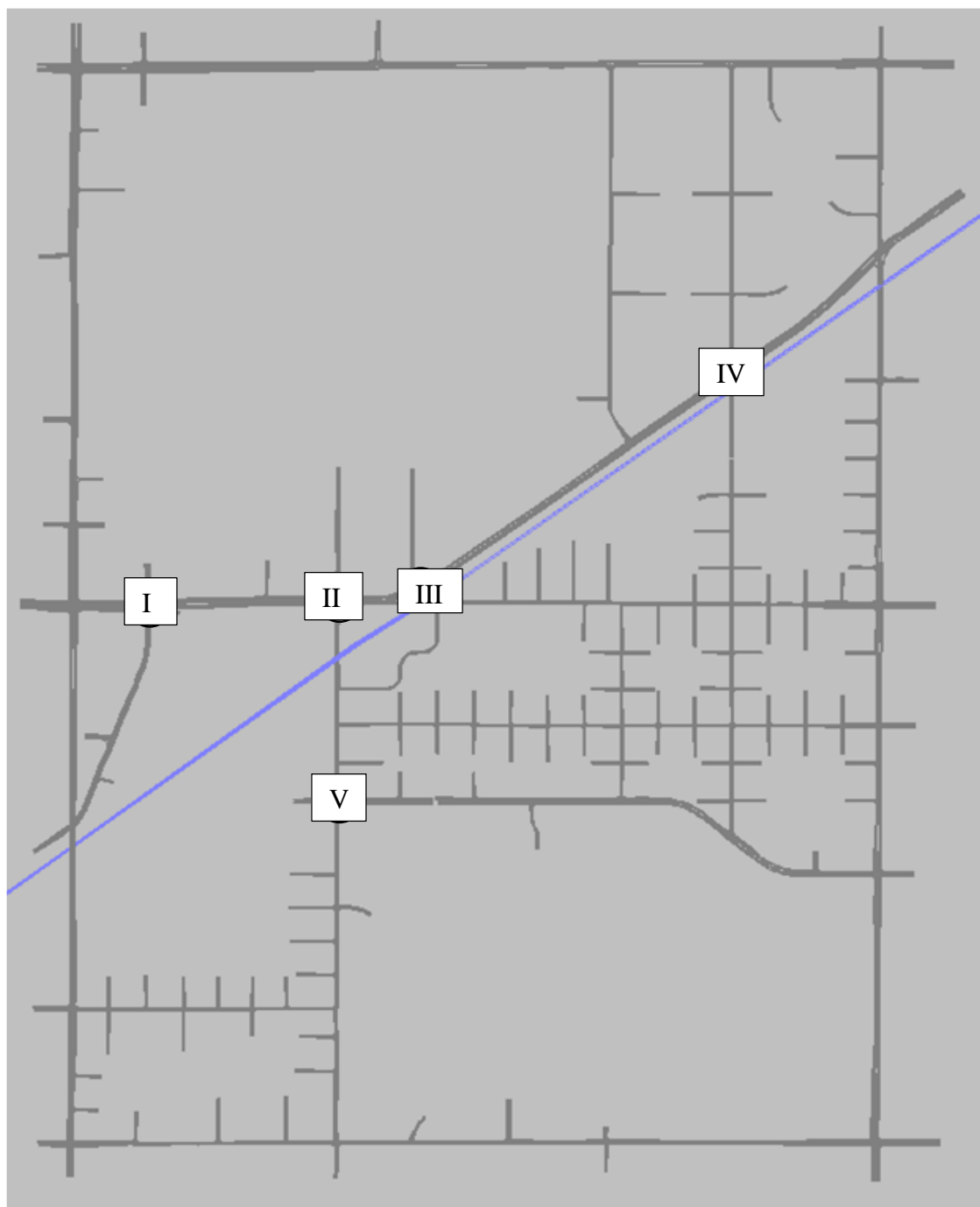
Figure 5.6 Six intersections used in the validation study

5.3 Reliability-Based LOS Evaluation of HRGC Related Intersection

5.3.1 Intersection Evaluation

    Table 5.8 compares the intersection evaluation method from the highway capacity

manual (HCM) and the simulation-based method.

Table 5.8 Comparison of different analysis methods

Methods	HCM	Simulation
Measure of Effectiveness	Delay	Delay, Stops, Queues
Best Applications	Operations, Signal Timing	Unusual Situations, Closely Spaced Intersections
Secondary Applications	Planning, Impact Studies, Roadway design	Operations, Signal Timing, Planning, Impact Studies, Roadway Design

Source: Husch, D. and J. Albeck (2003)

    The HCM presents an analytical method to calculate the control delay for each

movement in an intersection. The HCM LOS thresholds, in terms of average delay, can

be used to determine the operational LOS when details on intersection flow,

signalization, and geometrics are known. According to the HCM method, the intersection

LOS is directly related to the average control delay per vehicle. This metric is calculated

by aggregating the estimated average control delay of each approach. The aggregation is

based on volume, as shown in equations 5.6 and 5.7 (HCM 2000).

$$d_I = \frac{\sum_{a=1}^{A} d_a V_a}{\sum_{a=1}^{A} V_a} \tag{5.6}$$

where:

d_I = the delay per vehicle for an intersection I (s/veh),

A = the total number of approaches at intersection I,

d_a = the delay for approach a that is aggregated from the delays for the lane

groups in approach a by equation 5.7, and

V_a = the adjusted flow for approach a (veh/h).

$$d_a = \frac{\sum_{i=1}^{I} d_i V_i}{\sum_{i=1}^{I} V_i} \tag{5.7}$$

where:

d_i = the delay per vehicle for lane group i on approach a (s/veh),

I = the number of lane groups on approach a, and

V_i = the volume for lane group i on approach a.

The output from the HCM method includes average control delays for each lane group, approach, and the intersection as a whole, along with the corresponding LOS.

As summarized in table 5.8, the best application for simulation models are for unusual situations such as closely spaced intersections (Husch and Albeck 2003). This is because the simulation software can provide case-specific data collection points to record a variety of performance measurements such as delay, queue, and the number of stops. The target of this study is a signal intersection located very close to a highway-railway at-grade crossing. This scenario requires a special setup in VISSIM to collect the delay

resulting from both the intersection signal timing and the HRGC. Therefore, a simulation model built in VISSIM is selected to evaluate the operational performance at this intersection. This study will evaluate the reliability-based LOS based on simulation data. The HCM thresholds shown in table 5.9 will be used to calculate the LOS reliability on a specific lane group, and this analysis will consider the impact of trains in the network.

Table 5.9 LOS criteria for signalized intersection sin HCM2000

LOS	Control Delay per Vehicle (sec/veh)
A	≤10
B	>10-20
C	>20-35
D	>35-55
E	>55-80
F	>80

5.3.2 HRGC Related Intersections

Table 5.10 summarizes the basic information for the three HRGC related intersections in the test bed. At each signalized intersection, the inbound north bound (NB) lanes intersect with the railway. Consequently, there is a danger of vehicle queuing on the tracks when the NB signal is red. In addition, vehicles queuing up from the HRGC will experience delay due to the train traffic when there is a train present in the corridor. First, the volume to capacity (v/c) ratios for the lane groups on the NB approach are calculated without considering train traffic. The results for critical lane groups are summarized in table 5.10. The v/c ratios for critical lane-group *i* were calculated by equation 5.8. The relevant information used to calculate the v/c ratios was provided by the City of Lincoln.

The NB approach of the 44th and Cornhusker intersection only has one lane. The v/c ratio is 0.32, indicating that the approach can provide sufficient capacity for current traffic flows. The NB approach of the 33rd and Cornhusker intersection has two lanes – one is for left-turn movement and the other is for through and right-turn movements. The traffic on this approach consumes 84 percent of the overall capacity. The NB approach of the 35th and Cornhusker intersection also has two lanes – one exclusively for left-turn movements and the other for through and right-turn vehicles. The westbound (WB) Adams Street has one lane. When it turns into the NB 35th Street, after the HRGC, it becomes two lanes, as shown in figure 5.7. The left-turn bay on the NB 35th St is less than 28 meters in length. Based on empirical observations at the site, the section of road can accommodate approximately three vehicles. Given the limited lane capacity, the volume for the NB left-turn movement is 340 vehicles per hour (City of Lincoln 2008). Therefore, the v/c ratio for the left-turn lane on the NB of 35th Street approaches 1.05 during peak hours, indicating that the lane operates at saturation flow rate during the peak hour. When a train is present, it is hypothesized that the service level for 35th Street NB and Adams Street WB will be further degraded due to the extra delay that results from waiting for the train to clear the HRGC. Therefore, this lane group is selected to implement the reliability-based LOS evaluation.
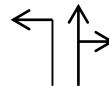
The physical configuration of the 35th and Cornhusker intersection is shown in figure 5.7, with the north bound left-turn (NBL) phase marked by a solid arrow. The control delay for each vehicle from point A to point B is output from VISSIM and is used to calculate the confidence interval of control delay for the left-turn movement. Point A is set 150 meters east of the HRGC because this encompasses the maximum queue length

observation at the site. Point B is located 60 meters west of the signalized intersection. The delay time measurements output by VISSIM is the difference between the travel time from A to B under the simulated scenario and the idealized case (e.g., no signal control, no train traffic, etc.).

As shown in table 5.9, control delay is used by current HCM to define the thresholds of different levels of service. It includes the initial deceleration delay, queue move-up time, stopped delay, and final acceleration delay. The purpose of setting a relatively long section for A-B is to ensure that all of the three delay components are measured.

Table 5.10 HRGC related intersections

HRGC Intersections	33rd & Cornhusker	35th & Cornhusker	44th & Cornhusker
NB Lane Configurations	←↑↑	←↑↳	←↑↳
V/C Ratio of NB Critical Lane Group	0.84 (NB)	1.05 (NBL)*	0.32 (NB)
Distance from the Stop Bar to HRGC (m)	170	30	28
Volume NBL (veh/h)	166	340	24
Volume NBT(veh/h)	51	15	21
Volume NBR(veh/h)	79	8	10
Green Time (second)	30	20	14

* Selected lane group for reliability analysis.

$$X_i = \frac{v_i}{s_i(\frac{g_i}{C})} \tag{5.8}$$

where:

X_i = the v/c ratio,

v_i = the observed volume of the critical lane group I,

$\frac{g_i}{C}$ = actuated green to cycle ratio, and

s_i = saturation flow rates.
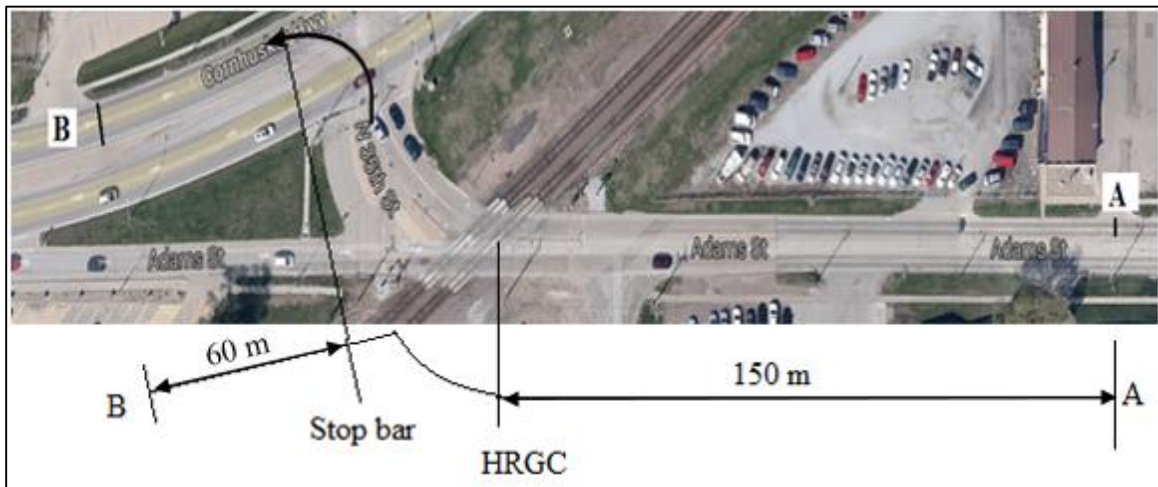
Figure 5.7 The intersection movement for analysis

5.3.3 Reliability-Based LOS of the Left-Turn Movement

Three scenarios were simulated for this reliability analysis: 1) no train traffic; 2) one train traveling at 60 km/h that occupies the HRGC from 1,185 seconds to 1,342 seconds during the simulation; and 3) one train traveling at 25 km/h that occupies the HRGC from 1,127 seconds to 1,542 seconds during the simulation. These three scenarios were selected to illustrate the impacts of trains with different speeds on the LOS

reliability of the northbound left-turn movement at intersection of 35th Street and Cornhusker Highway.

The time-dependent confidence intervals of the average delay for the left-turners for scenarios 1, 2, and 3 are shown in figure 5.8, 5.9, and 5.10, respectively. These confidence intervals are calculated using the BCa bootstrap method. In these figures, the x-axis represents the time since the simulation start time in unit of seconds, and the y-axis represents the delay time for the left-turners at the test bed. The first 600 seconds are used as the warm-up period, and the results from this period are not included in the figures. Therefore, the x-axis ranges from 600 seconds to 3,600 seconds. The horizontal blue lines are the upper and bottom boundaries of confidence intervals of the average delay for each 5-minute interval. The confidence interval of a given 5-minute interval is calculated for the vehicles that passed the point B in figure 5.7 within that 5-minute interval. The horizontal red lines are the level of service thresholds. In figure 5.9 and 5.10, the vertical black dashed lines indicate the period of train presence at the HRGC.

Figure 5.8 shows the control delay of left-turning vehicles when there is no train present at the HRGC. Figure 5.9 is the control delay for the scenario with a 60 km/h train present in the HRGC for 157 seconds. Note that the train occupies the HRGC during the 5-minute interval from 1,200 seconds to 1,500 seconds. For the no-train scenario, the confidence interval of average delay is [68, 78] seconds for this 5-minute interval (e.g., 1,200 seconds to 1,500 seconds). In contrast, the confidence interval is [136, 159] seconds for the 60km/h train scenario. In summary, the range of the confidence interval for average delay increases from 10 seconds to 23 seconds during the train event, indicating greater uncertainty in delay. Not surprisingly, the average delay during the 5-

minute interval when a train is present approximately doubles as compared to the no-train scenario.

Figure 5.10 shows the control delay for the 25 km/h train scenario. The train arrives at the HRGC at 1,127 seconds in the simulation and departs at 1,542 seconds. The train occupies the HRGC for 415 seconds. The confidence interval of average delay for the time period (e.g., 1,500 seconds to 1,800 seconds) is [339, 384] seconds. For the no-train scenario, the confidence interval of average delay is [73, 85] seconds during the 5-minute interval from 1,500 seconds to 1,800 seconds. In other words, the 25 km/h train increases the interval range from 12 seconds to 45 seconds, and results in an approximately five fold increase in control delay.

The graphs also show that the train traffic increases both travel time and travel time variation for approximately 20 minutes for scenario 2 and approximately 30 minutes for scenario 3. It may be seen that the longer the train occupies the HRGC, the greater the impact on roadway traffic delay.
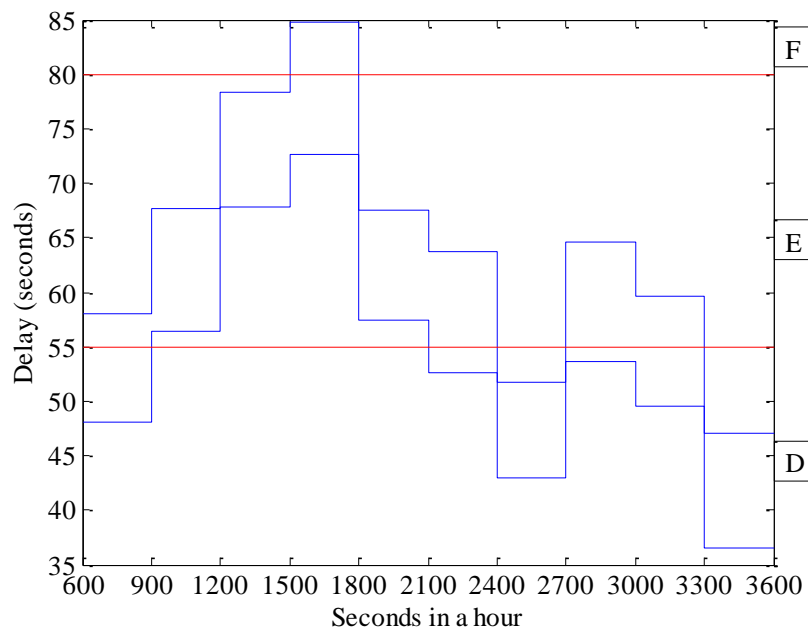
Figure 5.8 Time-dependent confidence intervals of average delay for the no-train scenario
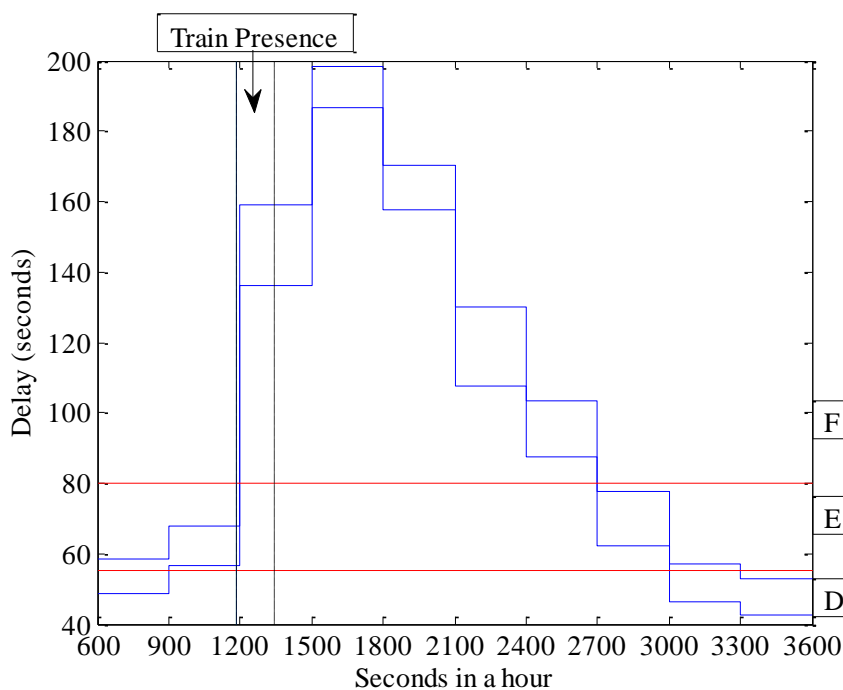
Figure 5.9 Time-dependent confidence intervals of average delay for scenario 2
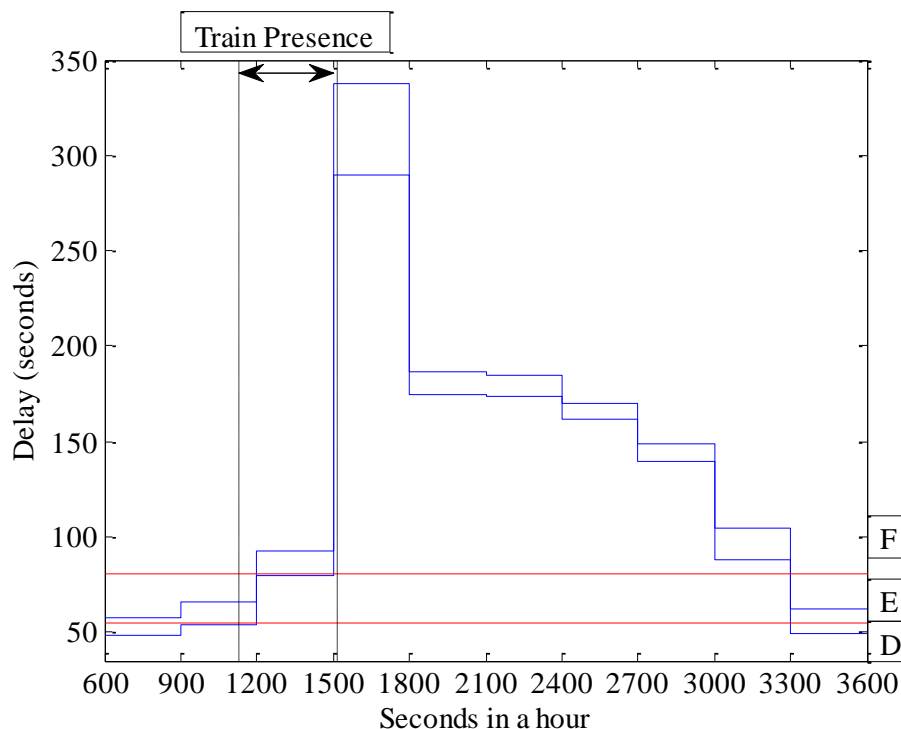
Figure 5.10 Time-dependent confidence intervals of average delay for scenario 3

As defined in chapter 3 (e.g., equation 3.3), reliability of a certain LOS is the probability that the average delay confidence intervals can be included in the thresholds of that LOS. The results of the reliability-based LOS analysis for the three scenarios are summarized in table 5.11.

Table 5.11 LOS reliability of the left-turn movement in this study

LOS	Scenario 1: No train	Scenario 2: One train at 60 km/h	Scenario 3: One train at 25 km/h	Thresholds based on average delay (sec)
F	0.05	0.60	0.79	>80
E	0.61	0.23	0.16	>55-80
D	0.34	0.17	0.05	>35-55

In scenario 1, where there is no train traffic, the probability of LOS E is 61 percent. In scenario 2, where a 60 km/h train occupies the HRGC for 157 seconds, the reliability of LOS E decreases to 23 percent. At the same time, the reliability of LOS F is increased from the 5 percent for the no-train scenario to 60 percent. In scenario 3, a train at 25 km/h occupies the intersection for 415 seconds. The reliability of left-turning vehicles at the studied approach experiencing LOS D was lowered from the 61 percent for the no-train scenario to 16 percent. This is 30 percent lower than the scenario 2 with a 60 km/h train. The LOS E reliability was increased to 79 percent, which is a 32 percent increase as compared to scenario 2.

These results indicated the complex impact that train traffic at a given HRGC has on roadway traffic. The longer the train occupies the HRGC, the lower the reliability of LOS D and E and the higher the reliability of LOS F. The concept of LOS reliability can reflect the different system performances for trains traveling at different speeds. For example, as compared to scenario 1 (i.e., no-train scenario), the LOS E reliability was reduced by 62 percent in scenario 2 (e.g., a train traveling at 60 km/h), while for scenario 3 (i.e., a train traveling at 25 km/h) the LOS E reliability was reduced by 74 percent. In contrast, as shown in table 5.12, the traditional LOS based on average delay indicates that both scenarios 2 and 3 had LOS F. The traditional approach is unable to distinguish the impacts from different train speeds. The reliability metric proposed in this dissertation provides traffic engineers a more comprehensive representation of the impact magnitude due to train traffic events at the intersections near HRGCs.

Table 5.12 LOS of the left-turn movement in this study

	Scenario 1: No train	Scenario 2: One train at 60 km/h	Scenario 3: One train at 25 km/h
Average Delay (sec)	60	106	156
HCM LOS	E	F	F

5.4 OD Based Reliability Information

Train traffic in a bimodal transportation network generally increases delay to

vehicle traffic at HRGCs as the vehicles wait for the train to clear the HRGC. Drivers

who use paths that include HRGCs will be particularly impacted. The arrival time

reliability concept developed in this dissertation can be used to measure this effect. If

drivers can be informed, in real time, regarding the reliability intervals of their route

travel time when there is a train present in the corridor, they would have a better idea

about their arrival time reliability, and could choose routes with better arrival time

reliability. It is hypothesized that this will reduce total network delay and increase overall

network performance.

5.4.1 Study Area

One OD pair in the test network is selected to demonstrate this application. The

origin is the University of Nebraska-Lincoln's east campus, and the destination is the

UPS customer service center, as shown in figure 5.11. Four potential routes are selected,

and they are representative of the options available to a driver. They are shown in

different colors in figure 5.11. It can be seen in figure 5.11 that drivers who choose

Route-2 have the opportunity to change to Route-1 before making their first right-turn.

Similarly, drivers who choose Route-3 have the opportunity to change to Route-4 before

making their second left-turn. Therefore, in this analysis, Route-1 and Route-4 are considered as the alternative routes for Route-2 and Route-3, respectively, when a train is present in the corridor. Relevant information for each route is shown in table 5.13. The free flow travel time is the travel time at the link speed limit.

Table 5.13 Basic information of the four routes to study

Route No.	Physical length (km)	Speed limit section		Number of signals	Free flow travel time (min)	Utilize HRGC
		Speed limit km/h (mph)	Section length (km)			
Route-1	3.96	56 (35 mph)	1.45	6	5	No
		64 (40 mph)	1.79			
		72 (45 mph)	0.72			
Route-2	2.29	56 (35 mph)	2.29	3	3	Yes
Route-3	4.67	56 (35 mph)	2.98	6	8	Yes
		40 (25 mph)	1.53			
		72 (45 mph)	0.16			
Route-4	6.16	56 (35 mph)	1.26	9	10	No
		40 (25 mph)	2.88			
		72 (45 mph)	2.03			
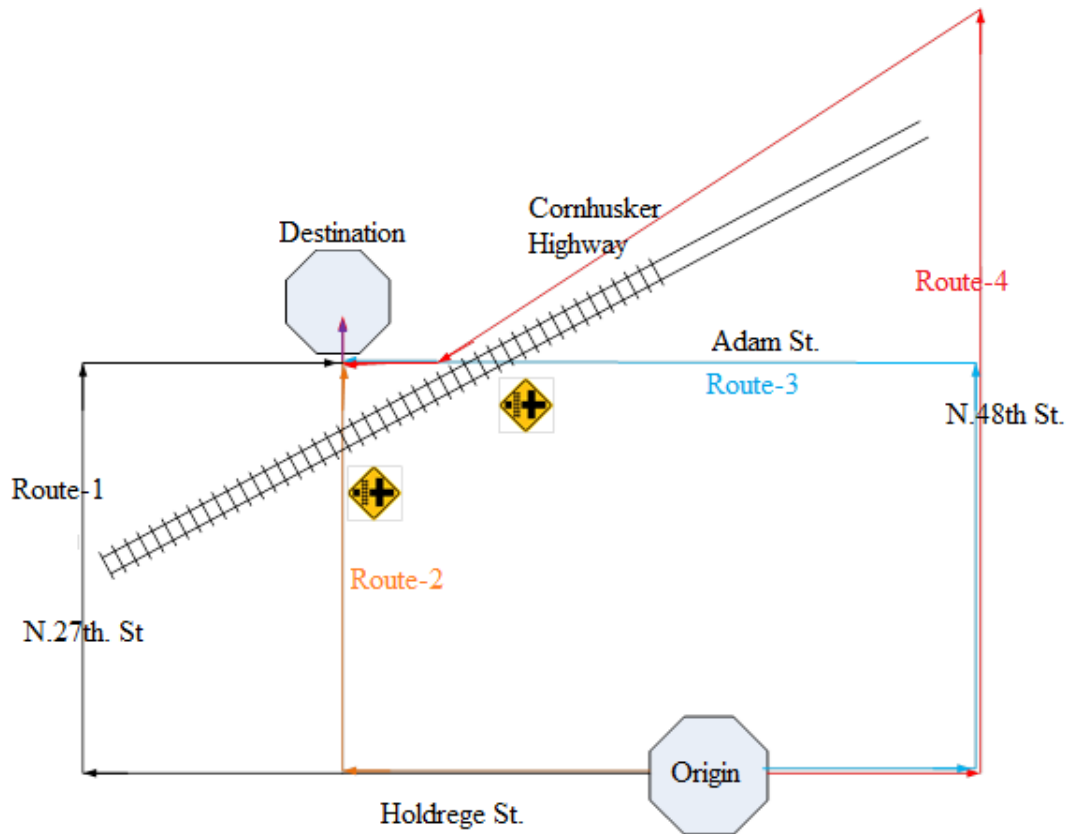
Figure 5.11 Four routes selected to study the OD pair from Campus (Origin) to UPS (Destination)

The train volume is set to 1 train per hour, traveling at 34 km/h, which is approximately the average train speed in the corridor, as shown in table 5.2. The network is simulated for one hour and the travel times are collected as illustrated in section 5.4.2. In section 5.4.3, the arrival travel time reliability for each route will be estimated using the approach developed in section 3.3.2.1. The impact of the train arrival on the OD-level network performance will be estimated using the methodology developed in section 5.4.4.

5.4.2 Simulated Travel Time Collection

The simulation model is used to generate travel time data to estimate travel time reliability intervals on the route and OD level. A specific way to collect travel time data needs to be defined before further research steps are taken. Conventionally, a road network is structured similar to picture (1) in figure 5.12, using a single node representing an intersection as shown in picture (2). Roadway links are the segments connecting a pair of adjacent nodes. Roadway travel time is defined as the travel time between the centers of the pair of nodes. Route travel time, however, is movement-specific. This requires a way to reflect movement-specific delay at intersections in the route. To do so, the travel time at a node is collected for each turning movement to include the movement-specific delay at the intersection. A movement-specific turning node is decomposed as three sub-links according to specific movements within the intersection as illustrated in picture (3) in figure 5.12. Travel times for C-l, C-t, and C-r are collected for the left-turning, going-through, and right-turning movements at the intersection, respectively. The data collection for route travel time estimation in this dissertation is based on movement-specific turning nodes, as shown in picture (4) in figure 5.12. The tool of travel time sections available in the simulation software VISSIM for data collection is set up at the locations necessary for collecting link and movement-specific node travel times.
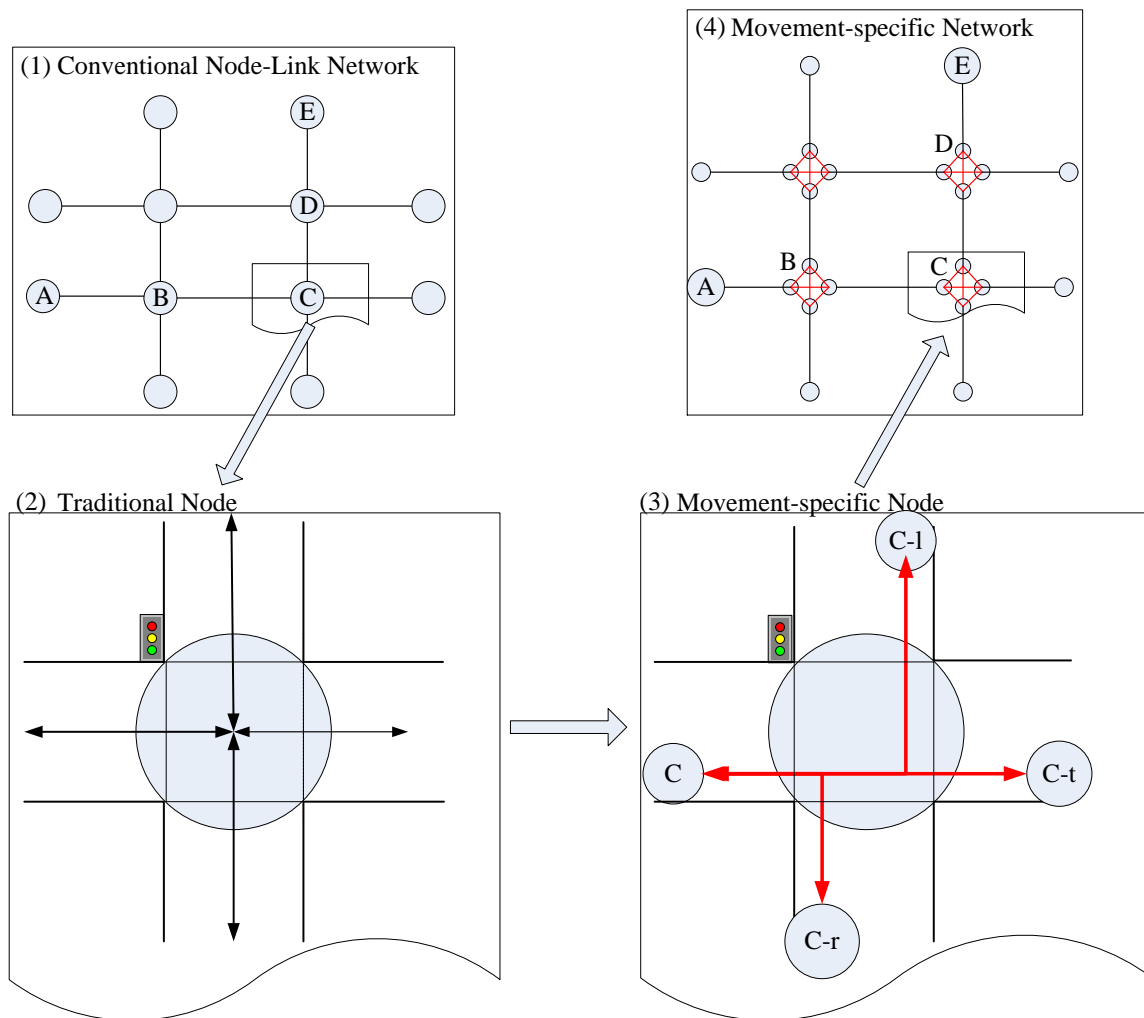
Figure 5.12 Node coding to study route travel time reliability

5.4.3 Reliability Information for Individual Drivers

As defined in equation 4.1, the reliability interval of route travel time is an interval centered on the average route travel time and bounded by plus and minus k standard deviations. In this study, k is set to 1 because the 4 routes are composed of urban arterials and the traffic conditions along these routes are similar to the test bed in chapter

4. In real applications where actual route travel times are available, k can be determined based on empirical experiments.

Figures 5.13 (a), (b), (c), and (d) show the time-dependent reliability intervals under the one-train scenario for Routes-1, 2, 3, and 4, respectively. The x-axis is the simulation time in seconds. The y-axis is the route travel time in minutes. The green line bounded with stars indicates the presence of the train at the HRGC. Note that only figures 5.13 (b) and (c) show the green lines because they are the two routes that cross the HRGC. The averages and standard deviations of route travel time used to construct the reliability interval bands are estimated by the first-order approximation method. Each reliability interval of route travel time is calculated for the trips that depart within the corresponding 5-minute interval shown on the x-axis. For the purpose of comparison, Figures 5.14 (a), (b), (c), and (d) show the time-dependent reliability intervals under the no-train presence scenario for Route-1, 2, 3, and 4, respectively. These plots provide straightforward information about the traffic situations on the four routes from 600 seconds to 3000 seconds. It would be beneficial to include this type of information in an on-line traveler information system so that drivers can plan their trips beforehand.

The means and standard deviations of route travel times for both the one-train scenario and the no-train scenario are summarized in table 5.14. As shown in table 5.13, Route-2 is shorter in physical distance than its alternative Route-1, and Route-3 is shorter than its alternative Route-4. However, table 5.14 reveals that the presence of a train on Route-2 and Route-3 generates enough delay to make their alternatives more attractive to drivers. As an example, consider the 5-minute interval from 600 to 900 seconds. The average route travel times of Route-1 and Route-2 in the no-train scenario are 6.2 and 5.6

minutes, respectively. In the one-train scenario, the average Route-2 travel time increases to 9.7 minutes, which is longer than Route-1. In addition, the standard deviation of Route-2 in the one-train scenario is 4.5 minutes while that of Route-1 is only 1.9 minutes.

Despite having to cross a HRGC, Route-2 is less impacted than Route-3 for the one-train scenario. The underlined numbers in table 5.14 indicate the four 5-minute intervals where Route-3 is negatively impacted, and the one 5-minute interval where Route-2 is negatively impacted. These results indicate that even with the presence of only one train, it could be difficult to evaluate the traffic conditions for each route by drivers. Providing time-dependent reliability interval information can assist drivers in route decision making.

Another application is to provide en-route information in the format similar to table 5.15 for the trip A-B, which can assist drivers at point A in choosing the most efficient route to get to point B. This kind of information can also be used in dynamic message sign systems to provide drivers with a better estimation of their arrival times along a particular route or corridor.
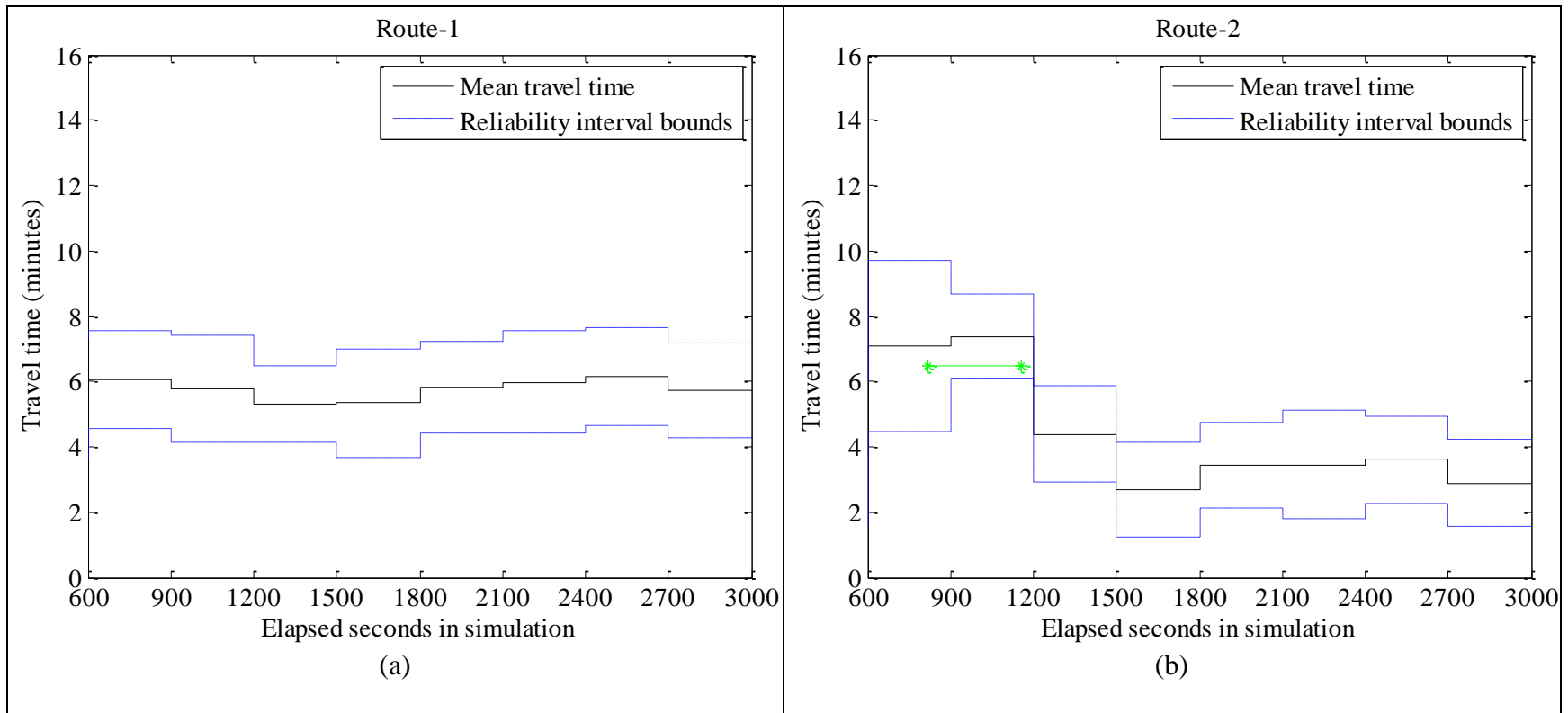
Figure 5.13 Time-dependent reliability intervals for route travel time under the one-train scenario
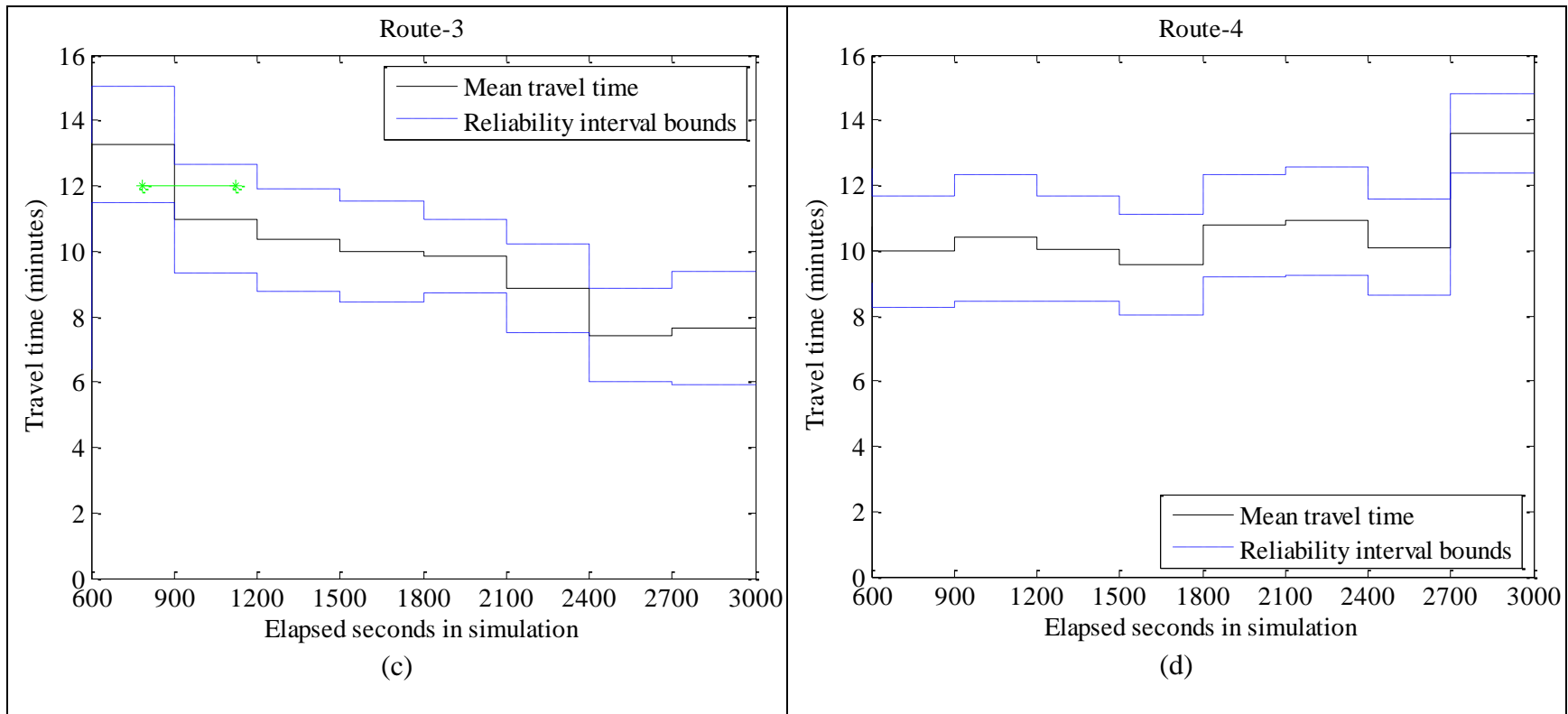
Figure 5.13 Time-dependent reliability intervals for route travel time under the one-train scenario (cont.)
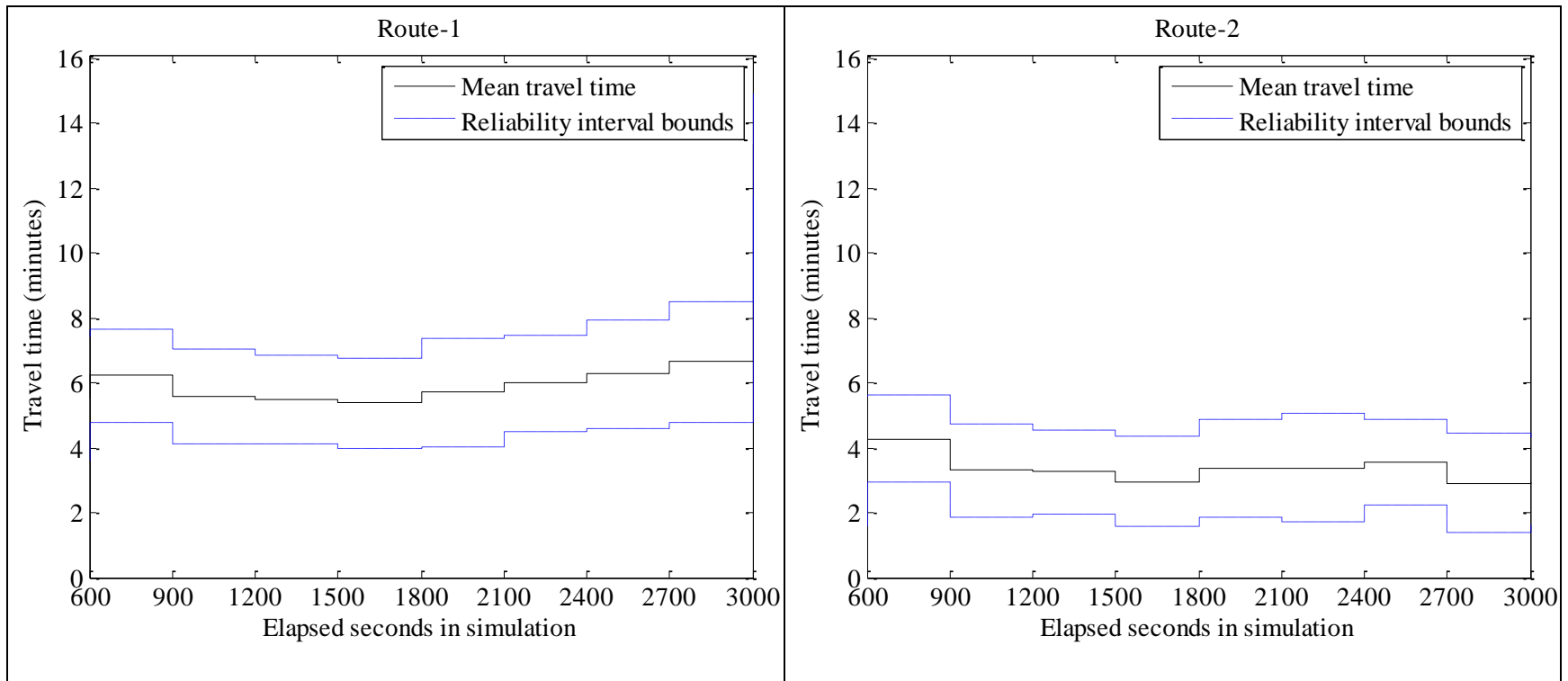
Figure 5.14 Time-dependent reliability intervals for route travel time under the no-train scenario
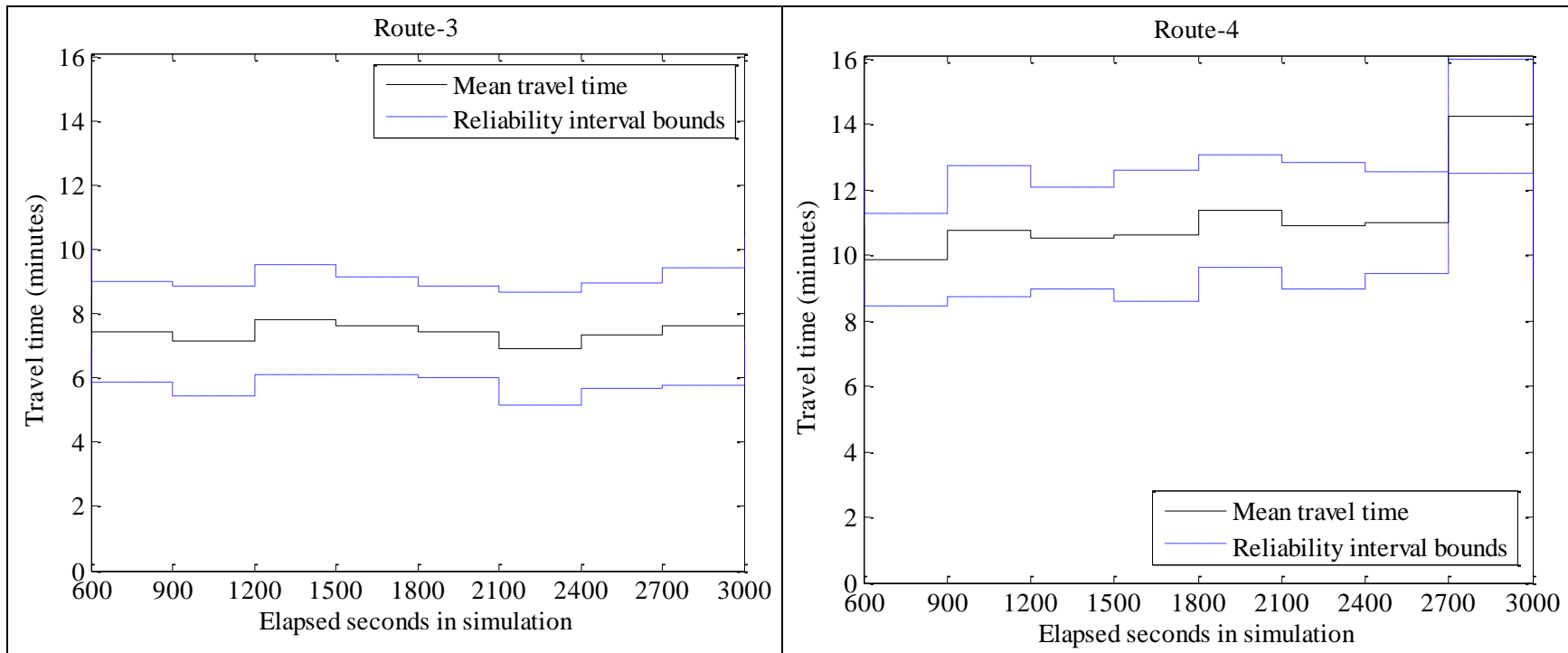
Figure 5.14 Time-dependent reliability intervals for route travel time under the no-train scenario (cont.)

Table 5.14 Means and standard deviations of route travel times (min.)

Route	Scenario	Statistic	5-minute Interval for departure								Average
			600-900	900-1200	1200-1500	1500-1800	1800-2100	2100-2400	2400-2700	2700-3000	
Route-1	no-train	Mean	6.2	5.6	5.5	5.4	5.7	6.0	6.3	6.6	6.7
		STD	1.4	1.5	1.4	1.4	1.7	1.5	1.7	1.9	1.5
Route-2	one-train	Mean	9.7 *	8.7 *	5.9	4.1	4.8	5.1	5.0	4.2	5.8
		STD	4.5 *	6.1 *	2.9	1.2	2.1	1.8	2.3	1.6	2.7
	no-train	Mean	5.6 *	4.8 *	4.5	4.3	4.9	5.0	4.9	4.4	4.8
		STD	2.9 *	1.9 *	2.0	1.6	1.8	1.7	2.2	1.4	1.9
Route-3	one-train	Mean	13.3 *	11.0 *	10.3	10.0	9.9	8.9	7.4	7.6	9.7
		STD	1.8 *	1.6 *	1.6	1.5	1.1	1.4	1.4	1.7	1.5
	no-train	Mean	7.4 *	7.1 *	7.8	7.6	7.4	6.9	7.3	7.6	7.6
		STD	1.6 *	1.7 *	1.7	1.5	1.4	1.7	1.6	1.8	1.6
Route-4	no-train	Mean	9.9	10.7	10.5	10.6	11.4	10.9	11.0	14.2	11.2
		STD	1.4	2.0	1.5	2.0	1.7	1.9	1.5	1.8	1.8

Note:

1) Routes-1 and 4 are minimally impacted by train traffic. Therefore, only no-train scenarios are listed in this table.

2) Routes-2 and 3 shown in bold are the two routes impacted by HRGCs.

3) Cells with (*) indicate the intervals impacted by train presences at the HRGCs.

4) Cells with underlined values indicate the impacted periods after a train has left. During these impacted periods, the difference between the means for one-train and no-train scenarios are higher than one minute.

Table 5.15 A-B travel time reliability interval information

8:10-8:15	Route-1	Route-2	Route-3	Route-4
Average (min)	6	7	13	10
+/- (min)	1	3	2	2

5.4.4 Reliability Evaluation for System Operators

System managers may be interested in quantifying the impact of train traffic on

the route performance and network performance for a given OD pair. This section

demonstrates how the metrics and techniques developed in this dissertation may be

helpful for this application.

5.4.4.1 Impact on Route Performance

It is assumed in this dissertation that system operators want to evaluate the

network in terms of the degree to which the network meets drivers' expectations. System

operators first need to define the evaluation thresholds that represent drivers'

expectations. For example, drivers may expect a longer travel time if they choose a route

with a longer distance and more signals. Therefore, the evaluation threshold, E_r, is

defined based on the ideal route travel time, as shown in equation 5.9. The ideal route

travel time t_i is based on the free flow travel time t_f and reasonable waiting time at

signals t_s, as shown in equation 5.10. It is assumed in this dissertation that drivers expect

to pass 75 percent of the signals without stopping and stop once at the other 25 percent of

the signals. Each stop at a signal is assumed to result in a 0.5 min delay. Note that this

method can be readily supplemented with empirical data with no loss in generality.

$$E_r = a * t_i \tag{5.9}$$

where:

E_r = the evaluation threshold of rth route.

a = the inflation factor to reflect driver expectation. It maybe postulated that in large cities such as Chicago and Los Angeles, this factor will be larger, and in smaller cities it may be smaller. In this dissertation, $a = 1.5$.

t_i = the ideal travel time of rth route, calculated by equation 5.10a.

$$t_i = t_f + t_d \tag{5.10a}$$

$$t_f = \frac{d_r}{v_r} \times 60 \tag{5.10b}$$

$$t_d = t_s \cdot n \cdot p_1 \tag{5.10c}$$

where:

t_f = the free flow travel time of rth route in minute,

d_r = the distance of rth route in mile,

v_r = the speed limit of rth route in mph,

t_d = the reasonable delay at signals,

t_s = the average delay of each stop at a signal ($t_s = 0.5$ min),

n = the number of signals of rth route, and

p_1 = the expected percentage of signals where a vehicle needs to stop ($p_1 =$ 25%).

Table 5.16 summarizes the evaluation results and route travel time reliability of Route-2 and Route-3. The route performance reliability is measured by the probability of the reliability intervals of route travel time included by the threshold. Consider Route-2 as an illustration. Figure 5.15 displays the time-dependent reliability intervals for Route-2 under the no-train scenario and the one-train scenario. The x-axis shows the elapsed seconds in simulation, and the y-axis shows the travel time in minutes. The red line is the evaluation threshold, and the green line indicates the presence of a train at the HRGC near the intersection of 35th Street and Cornhusker Highway. The reliability of Route-2 is lowered to 49 percent from 72 percent due to the presence of the train.

Table 5.16 Route information and reliability analysis under one-train scenario

Route No.	Evaluation threshold (min)	Reliability for 1-train scenario	Reliability for no-train scenario	Degradation
Route-2	4	0.49	0.72	32%
Route-3	9	0.23	0.95	76%
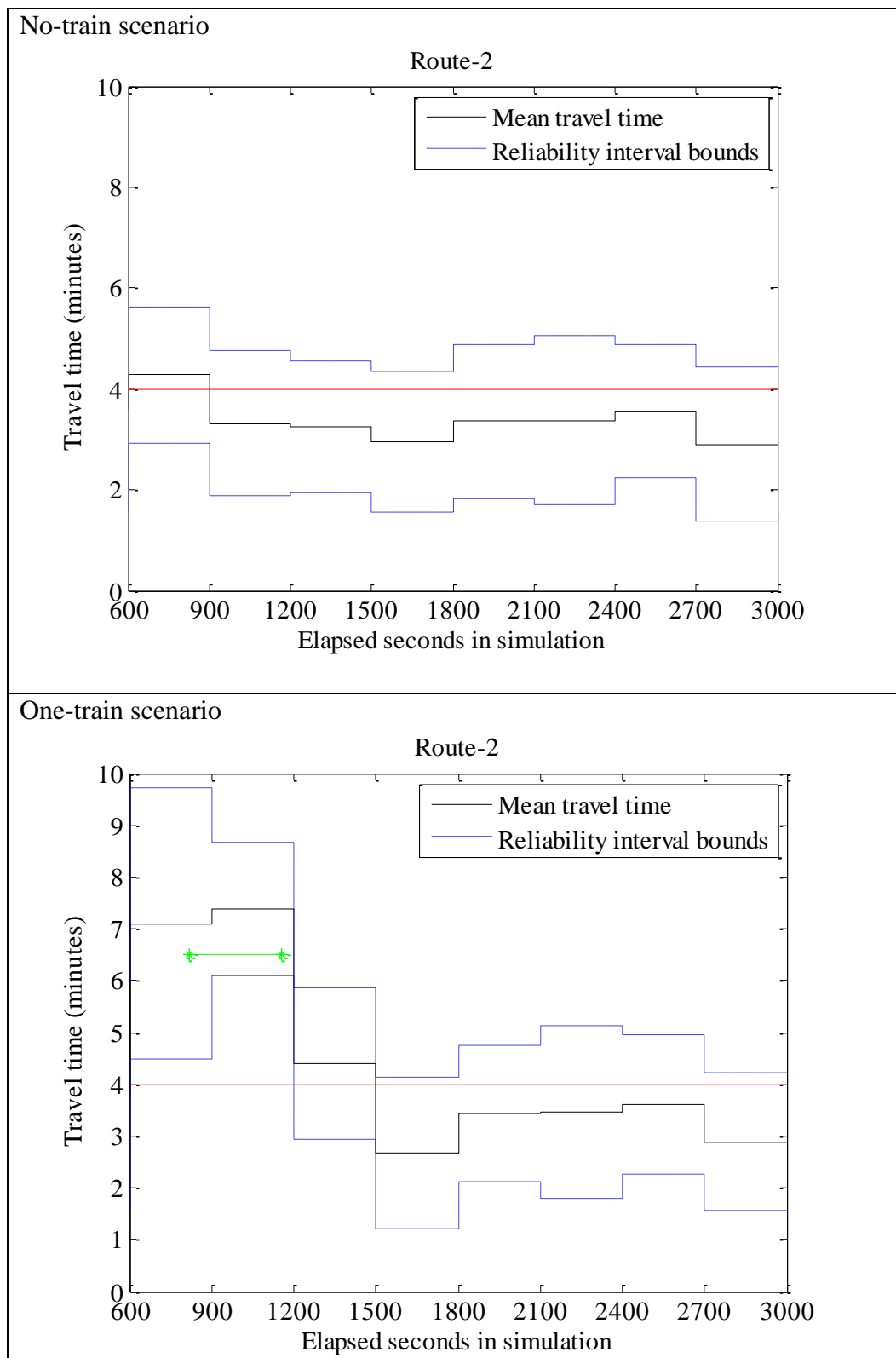
Figure 5.15 Time-dependent route travel time reliability intervals under with- and without-train scenarios

5.4.4.2 Impact on OD Level

The evaluation of OD service is based on the major routes serving a given OD pair. In the previous section, four routes were selected based on initial engineering judgment for the sample OD pair. In this section, the four routes are further separated into section components so that the route network can be analyzed as a complex system with parallel and series sub-systems, as shown in figures 5.16 and 5.17. The shared sections of different routes are defined as "components". For this example, there are three components. The independent sections of each route are defined as a "sub-route". There are four sub-routes for this example. All the "components" and "sub-routes" are illustrated in figure 5.16. The evaluation thresholds E_r for each component and sub-route are calculated based on equations 5.9 and 5.10. They are summarized together with the resulting reliability metrics for no-train and one-train scenarios in table 5.17.
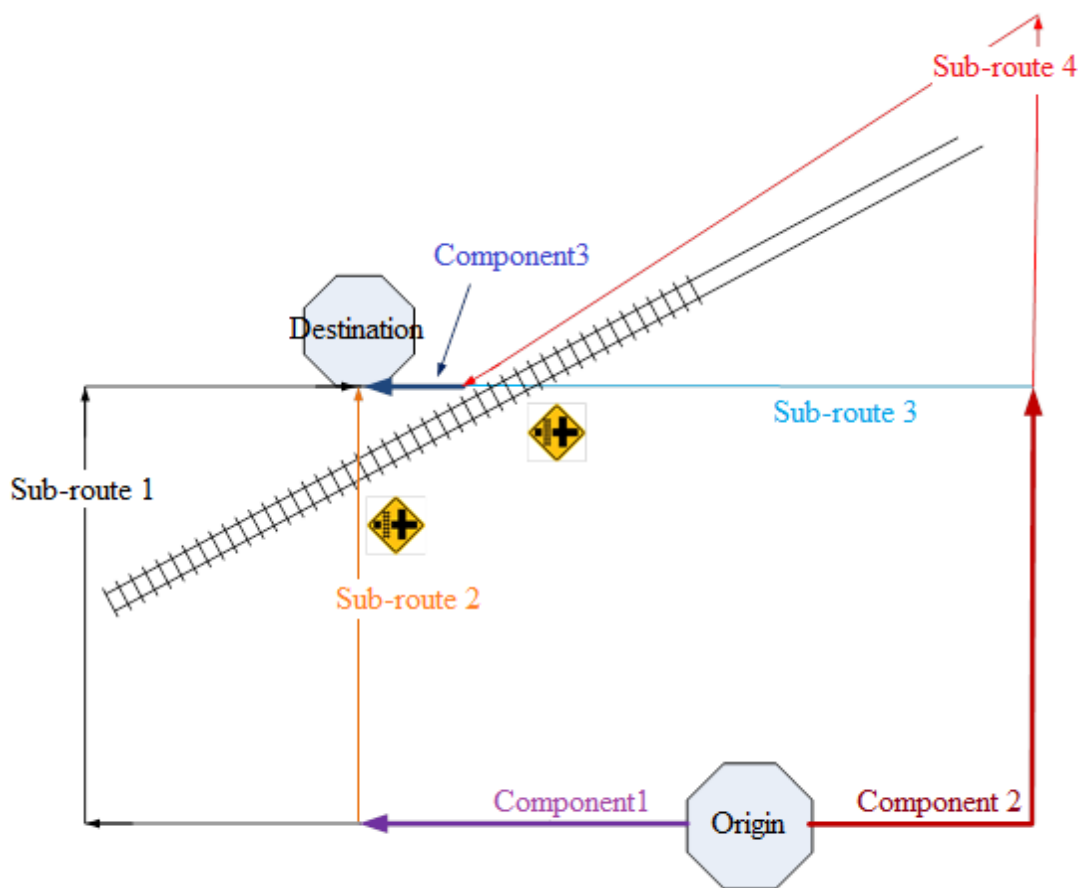
Figure 5.16 Section components for the route network connecting the OD pair
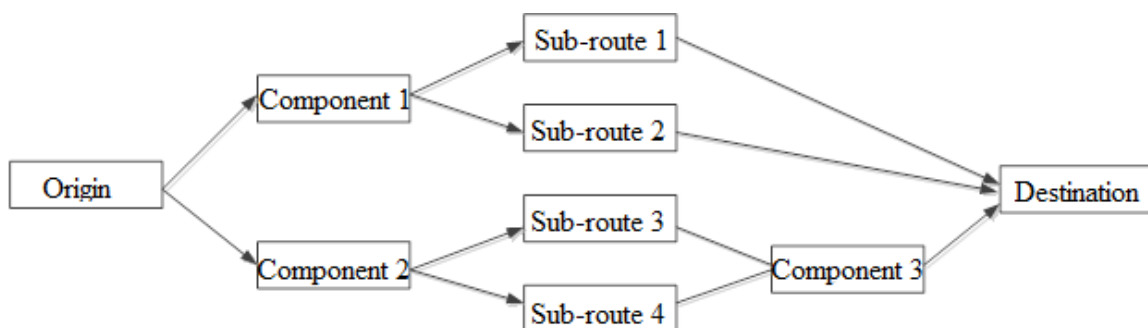
Figure 5.17 The complex system for the OD network

Table 5.17 Reliability of each components in the OD network

Components	Distance (km)	t_f (min)	Number of signals	t_d (min)	t_i	E_r	Reliability 0-train	1-train
Component 1	0.45	0.48	1.00	0.13	0.61	0.91	0.92	0.93
Component 2	2.77	3.24	3.00	0.38	3.61	5.42	0.96	0.98
Component 3	0.16	0.14	1.00	0.13	0.26	0.40	0.71	0.71
Sub-route 1	3.49	3.33	5.00	0.63	3.95	5.93	0.71	0.78
Sub-route 2	1.81	1.95	3.00	0.38	2.32	3.48	0.76	0.40
Sub-route 3	1.71	1.84	3.00	0.38	2.21	3.32	0.56	0.14
Sub-route 4	3.20	3.56	6.00	0.75	4.31	6.46	0.53	0.63

Note: The results are in two decimal digits because some components are shorter than 1 km in length, and the corresponding travel times are less than one minute.

In reliability engineering, the reliability of a parallel system can be calculated by equation 5.11a, and the reliability of a series system can be calculated by equation 5.11b.

$$R(Parallel\ system) = 1 - \prod_{j=1}^{k}(1 - R_j) \qquad (5.11a)$$

$$R(Series\ system) = 1 - \prod_{j=1}^{k} R_j \qquad (5.11b)$$

where:

k   = the number of components/sub-routes in the system, and

R_j = the reliability of jth components/sub-routes connecting the OD, $j=1, …, k$.

The reliability calculation of the OD network shown in figure 5.17 is based on equation 5.11. The overall OD network reliability is 0.93 for the no-train scenario. In the one-train scenario, the overall OD reliability is decreased to 0.78, which is a reduction of 16 percent. These types of analyses enable a reliability evaluation at the network level. It quantifies the impact of train traffic on reliability for an OD pair of interest and would

allow benefits for increasing reliability to be considered in benefit-cost studies for infrastructure improvements such as viaduct construction.

5.5 Concluding Remarks

This chapter provided a case study of a local bi-modal transportation system in order to illustrate the benefits of the reliability metrics developed in this dissertation. The confidence interval based reliability metric for system evaluation was applied to evaluate the left-turn movement at one intersection located near a HRGC. A comparison between the reliability based LOS and the traditional LOS evaluation indicated the reliability metric has the ability to reflect a more detailed system performance when the system is impacted by train traffic.

The travel time reliability interval was calculated for 4 selected routes for a sample OD pair in the test network. The arrival time interval information can be provided via a traveler traffic information system in order to assist drivers in choosing routes based on estimated mean travel times and travel time reliability. In addition, the real-time interval information can be used in dynamic message signs to notify drivers of reliability issues when a train is present in the corridor.

The presence of train traffic increased the mean and standard deviation of travel times on the routes passing through a HRGC so that they become the less attractive choices. The duration of the impacted period depends on the train speed, traffic demand at the intersection, and the intersection operation condition. In this study, Route-3 has an impacted period about 15 minutes longer than Route-2. Given the evaluation thresholds defined in this dissertation, the reliability of Route-3 is degraded by 76 percent while the

reliability of Route-2 is degraded by 32 percent for the one-hour analysis period. This result can also assist system operators in identifying the priority of improvement projects.

Lastly, the travel time reliability intervals over a one-hour period were used to evaluate the overall network performance with respect to the test OD pair. This OD-based reliability analysis provides an example of how the reliability metrics and estimation methodologies developed in this dissertation can be used to evaluate the reliability issue on a network level.

CHAPTER 6   CONCLUSION AND RECOMMENDATIONS

This dissertation has proposed innovative metrics to evaluate reliability for transportation systems, and presented methodologies to implement these metrics under different application scenarios. Although the methodologies are tested using Bluetooth data and simulation data in this dissertation, their applications can be readily extended to other data collection methods that generate individual travel time observations, such as automatic vehicle identification and GPS. As these technologies advance, they will provide more reliable data with increased sample sizes to further improve the performance of the presented methodologies. This chapter summarizes the key findings in this dissertation.

6.1 An Innovative Metric for Reliability Analysis in Transportation Engineering

Previous studies defined travel time reliability "in terms of how travel times vary over time" (e.g., hour-to-hour, day-to-day). To represent the influences of random events such as bad weather and unexpected incidents, current reliability metrics are calculated based on travel time measurements over a substantial portion of time (e.g., 6 months of data without the influence of winter weather). This definition takes on a macroscopic level to study the long-term reliability of traffic systems.

This dissertation proposes a new generic reliability metric that enables a reliability-based performance evaluation to account for dynamic and stochastic properties in transportation networks. It can be tailored for different evaluation objectives by specifying the performance measurement and statistic for different system levels ranging from a node to a network, and for both short-term and long-term scales.

This dissertation applies this new metric for a level of service evaluation which resulted in the reliability of certain levels of service for transportation systems. To evaluate long-term reliability, the confidence interval of performance measurements observed within a six-month period is calculated to represent the system performance variation. In addition, the level of service thresholds defined in the HCM are used to represent the satisfactory system performance. The overall reliability is measured by the probability that the level of service thresholds are able to include the time-dependent confidence intervals within the period of interest. This new metric can be used as a complement of the current reliability metrics, and to improve the level of service evaluation to a reliability-based level. To evaluate short-term travel time reliability within the peak period of a single day, an application example for a bimodal transportation network including highway-railway at-grade crossings is discussed in this dissertation. The reliability metric is calculated based on the user-defined travel time reliability intervals and evaluation intervals.

6.2 An Innovative Metric for Real-Time Reliability Information

Current real-time advanced traffic information systems (ATIS) focus on the average estimates of traffic parameters, such as average travel speeds and average travel times, for a section of the roadway. Existing ATIS generally include limited travel time reliability information which has been getting more attention as the uncertainty in traffic systems increases due to continuously growing congestion. In practice, the buffer index is usually used as a reliability indicator to help drivers plan their upcoming trips. A buffer index is calculated as the difference between the 95th percentile and the average travel

time, divided by the average. Based on travel time observations during a relatively long period (e.g., one year), the buffer index represents the extra travel time in relation to the average travel time needed to accomplish a trip on time, 95 percent of the time. However, the buffer index provides very little information regarding short-term travel time variability.

This dissertation proposes a short-term reliability indicator, the time-dependent reliability interval of the arrival time, to provide drivers with real-time information representing travel time variability for the current day. Using the reliability interval at the 15-minute level as an example, the reliability interval can be predicted for departure time in the next 15-minute interval. The real-time prediction enables drivers to decide their best departure time and route choice based on more recent traffic information compared with the long-term buffer index. The predicted reliability interval of corridor travel time is computed as the predicted mean travel time within the next 15-minute interval, and bounded by one predicted standard deviation in this case study. Based on the reliability interval of corridor travel times, the time-dependent reliability interval of arrival time is a predicted interval that can include at least *x* percent of the arrival times for trips departing during the next 15-minute interval, on average. The value of *x* needs to be validated after an experimental study on a target segment or corridor. In the case study of this dissertation, *x* is 95 under the overall situations.

6.3 Investigation of an Outlier Identification Method to Obtain Reliable Bluetooth Data

Most of traffic data analyses start with data "cleaning", which includes outlier identification, outlier removal, and missing data estimation. For example, Bluetooth

technology may detect travel times of non-vehicle modes and trip chains, resulting as outliers in the final dataset. This is particularly true for arterial corridor applications because urban arterials also serve for pedestrians and bicyclists. This dissertation proposed a new outlier identification algorithm, and compared its performance with the moving standard deviation method that is commonly used in current applications. The proposed method evaluates data points based on their corresponding residuals in a robust local regression scatter smoothing model, which is able to fit a regression surface with no constraint on parametric distributions and limited influences from outliers. In contrast, the moving standard deviation method generates a threshold that is used to identify outliers by adding one or two local standard deviations to the mean. The disadvantage of the moving standard deviation method is the assumption of an independent and normal distribution for the travel time dataset. For the arterial travel time dataset in this case study, it was shown that this assumption does not hold.

The advantage of the proposed methods over the moving standard deviation method was validated for the arterial corridor in this case study. The residuals in the local regression models were more adaptive to changing means and standard deviations, and thus the proposed method resulted in more efficient thresholds under unstable traffic conditions.

6.4 Investigation of Corridor Travel Time Mean and Variance Estimates

To analyze the travel time variability on a corridor or path level, different estimation methodologies are investigated to estimate corridor travel time mean and variance using link-based travel time observations. This is because the number of direct

travel time observations generally decreases as the length of a corridor or path increases. Because of the setup of the Bluetooth-based anonymous wireless address matching (AWMA) system, both corridor travel times and travel times on its component links can be generated for the test bed in this dissertation. This enables comparing the accuracy of different methods using real observations.

For the estimation of the corridor travel time mean, the four methods evaluated in this study are: the naïve addition method, the cumulative addition method, the first-order approximation method, and the second-order approximation method. The naïve addition method assumes independence between link travel times, and estimates corridor travel time by summing up the mean travel times on the component links for the departure time interval. This assumption generally holds for freeway corridors where vehicles travel under stable traffic conditions. However, for arterial corridors and urban streets during the period when traffic conditions change rapidly, the travel time on a certain link is very likely related to the arrival time on that link, which in turn depends on the travel time on preceding links. The other three methods address such correlation between link travel times by modeling link travel time as a function of the time of day. The accumulative summation method uses the step function that calculates the mean corridor travel time as the sum of the mean link travel time on link i within the time interval, corresponding to the arrival time at link i. The first-order and second-order approximation methods approximate the mean corridor travel time based on the Taylor series around the expected arrival time truncated at the linear and second-order terms, respectively. In this dissertation, the functions to model the link travel times are second order polynomials. The average of all the absolute percentage errors for the 15-minute intervals during PM

peaks from January to July of 2011 was used as a performance indicator of the different estimation methods. The second-order approximation method improves the performance by 26 percent compared with the naïve addition method. It is important to note that there is not much of a difference between the cumulative addition method and the naïve addition method. This is because the corridor test bed in this case study is about 5 km in length, while the estimation methods are applied for 15-minute intervals. Therefore, approximately 90 percent of vehicles travelled through the whole corridor within the same 15-minute interval during which they departed. It is hypothesized that the advantage of the cumulative addition would become more distinguishable if the study was conducted on a longer corridor and/or at a finer interval level.

For the estimation of arrival time variance for the studied corridor, the four methods evaluated in this study are the naïve addition method, the covariance based method, and the first-order and second-order approximation methods. The approximation methods had a 16 percent and 12 percent improvement in the median of absolute percentage errors compared with the naïve addition and the covariance based method, respectively.

6.5 Investigation of Prediction Models for Real-time Travel Time Reliability Information

Reliable short-term traffic parameter prediction models are very important for the successful implementation of a variety of ATIS applications, such as dynamic message signs and en-route guidance. This dissertation investigated one subclass of the dynamic recurrent neural networks, the nonlinear autoregressive with exogenous inputs (NARX)

model, to predict the short-term information of a corridor's travel time mean and variability information in the form of a reliability interval.

Previous research revealed that direct corridor travel time observations could produce more accurate mean corridor travel time predictions than link-level travel time observations while using Kalman filter techniques (Chen and Chien 2001). This dissertation investigated NARX models using corridor-level and link-level travel time data and compared the performance with naïve prediction methods – the historical average method and the instantaneous travel time (ITT) method – which are widely applied in current applications. Model performances are evaluated by the mean absolute percentage errors of the corridor travel time mean predictions, and by the coverage rate and average range of predicted reliability intervals (RI) of the arrival time. The coverage rate is the percentage of drivers arriving at the end of the corridor within the predicted reliability intervals. An efficient reliability interval should have a high coverage rate and, at the same time, be as small as possible.

For the three-link corridor studied in this dissertation, the NARX outperformed both the historical average method and the ITT model. Overall, the corridor-based NARX model was able to generate reliability intervals to include 99 percent of the corridor travel times. This coverage rate decreased to 84 percent for a non-recurrent incident period. The link-based NARX model provided comparable prediction results to the corridor-based NARX model in terms of the coverage rate of predicted reliability intervals of the arrival time. Furthermore, for the example non-recurrent incident period, the resulting reliability intervals of the link-based NARX model were able to provide the same coverage rate, 84

percent, with an average interval range that is 22 seconds smaller than that of the corridor-based NARX model.

For the non-recurrent incident period, the link-based second-order ITT model was able to provide the highest RI coverage rate among all the tested models. This high RI coverage rate, however, came with an increase in the RI average range. The RI average range was 52 seconds wider than that of the link-based NARX model for the three-link corridor.

6.6 Investigation of the Impact of Train Arrivals on Travel Time Reliability

Reliability issues in transportation systems are concerned with travel time uncertainty. Generally, the sources of travel time uncertainty can be categorized into recurrent and non-recurrent. The systematic demand pattern, such as the regular rush hours within the course of a day, is recurrent. Examples of non-recurrent sources include traffic incidents, weather, and malfunctions of traffic control devices. In bi-modal transportation networks with highway-railway at-grade crossings (HRGCs), the signal preemption for train traffic will incur an additional uncertainty on waiting times at the intersection and travel times on the associated routes. Using the proposed reliability metrics, this dissertation investigated the impact of train arrivals on the level of service at the intersection near HRGCs, and the travel time reliability on alternative routes of one impacted OD.

The reliability-based level of service analysis reveals that the impact of train traffic depends on the train speed, the resulting dwelling time at the intersection, and the arrival time. The time-dependent confidence intervals of the delay time illustrate the start

and dissipation of the impact in a more straightforward and detailed fashion. This analysis can serve as a supplement to the exposure factor that is currently used to warrant state train tax and federal rail safety funding for crossing improvements, and is calculated as a multiplication of the number of cars and the number of trains at an intersection each day.

To inform drivers with the travel time reliability related to train traffic, the time-dependent reliability intervals of four routes connecting one OD pair are estimated for with- and without-train scenarios. The presence of one train per hour can degrade the performance of the routes passing HRGCs, and thus, change the optimal route for the selected OD. This example validates the necessity of providing real-time route travel time reliability information to drivers. In combination with train arrival detection, the simulation model can be used to estimate the reliability information in real time. The proposed reliability metric can play a significant role in improving the traffic operations on such bi-modal transportation networks by guiding drivers to more efficient routes when there is a train coming.

6.7 Recommendations for Future Research

There are several directions to further extend the concepts and methodologies regarding reliability analysis in transportation engineering. The test bed used to develop and evaluate the estimation and prediction methodologies on the corridor level is 5 km in length. There is a need to validate the methodologies for longer corridor and path levels. This will become feasible as data collection technologies advance.

Another direction is to determine a well-agreed k value for the reliability interval through more empirical studies. The variable k is the number of the standard deviation that forms the reliability interval of the arrival time. In this case study, it is represented as one. Overall, the reliability intervals of the standard deviation of one are able to include 99 percent of the detected corridor trip makers' arrival times through a certain combination of the prediction model and the input dataset format. In real applications, k will probably depend on local traffic conditions and prediction model performance. In addition, other indicators like the 95th percentile could also be used to provide real-time reliability information. Studies related to the efficiency of associated prediction models for these indicators are needed to investigate the usability of various indicators.

This dissertation applied the proposed metrics and methodologies in several scenarios to assist decision-making in transportation engineering, such as how to evaluate the impact of train traffic on different routes for a certain OD pair. However, there are many other problems that can be addressed using the metrics presented in this dissertation. For example, real-time travel time reliability information can be used to develop automatic incident identification systems, or to identify the start and the end of congestion so that immediate traffic management strategies can be applied in time.

GLOSSARY

capacity reliability: the probability that a network successfully accommodates a given level of travel demand (Chen et al. 1999).

CDF: the cumulative distribution function (CDF) describes the probability that a real-valued random variable X with a given probability distribution will be found to have a value less than or equal to x.

connectivity reliability: the probability that there is at least one route connecting the specific OD pair, while links and nodes are subjected to random failure events with known probability in real-world lifeline networks (Ching and Hsu 2007).

confidence interval: is an interval associated with a parameter which is assumed to be non-random but unknown. It is a type of interval estimate of a population parameter, indicating the reliability of an estimate. After a sample is taken, the population parameter is either in the interval made or not, there is no chance. The significance level of a confidence interval indicates the probability that the confidence range captures this true population parameter given a distribution of samples (Hyndman 2013).

demand: the number of vehicle-based trips made within a particular unit of time in the traffic system. It is also OD specific.

dynamic: the state of traffic system changes over time because of the interactions between traffic control and the stochastic OD demand.

emergent property: all the performance properties in a transportation system (or any other man-made systems) result from the interacting individuals in the system, through dynamics (Nagel and Rasmussen 1994).

estimation: to obtain the statistics of traffic parameters using an available dataset that is always a sample of the population.

flexibility: the ability of a system to adapt to external changes while maintaining a satisfactory system performance. External changes include uncontrolled conditions that affect the system such as long-term changes of demand influenced by economics and population, a shift in spatial traffic patterns, changes in the price of fuel, etc. (Morlok and Chang 2004).

highway-railway at-grade crossing (HRGC): is an intersection where a railway track crosses a road at the same grade level, as opposed to the railway line crossing over or under using a bridge or tunnel.

level of service reliability: the probability that for the whole analysis period, the short-term confidence intervals of the performance measurement of interest are able to be included by a specified interval corresponding to the stratification of the performance measurements in level of service analyses.

long-term reliability: focuses on day-to-day variability in system performance within a relatively long period (e.g., one year).

MAC address: the unique electronic address of each enabled Bluetooth device.

metrics: a set of measurements that quantifies results. In this dissertation, metrics are performance measurements that quantify the efficiency of traffic systems.

network emergence: may be defined as the temporal process by which the macroscopic properties of a system or network alter due to the microscopic changes of its constituent parts (Manley and Cheng 2010).

OD connectivity reliability: the probability that there is at least one route connecting the specific OD pair while links and nodes are subjected to random failure events (Ching 2008).

OD: Origin-Destination, the start and ending points of a trip.

PDF: a probability distribution function of a continuous random variable is a function that describes the relative likelihood for this random variable to take on a given value.

prediction interval: is an interval associated with a random variable yet to be observed with a specified probability of the random variable lying within the interval. The significance level of a prediction interval indicates the probability that the predicted range captures the next actual travel time observation. (Hyndman 2013).

prediction: to make known in advance the traffic parameters in the next time interval based on current and historical data.

q-q Plot: the quantile-quantile (q-q) plot is a graphical technique for determining if two data sets come from populations with a common distribution. A q-q plot is a plot of the quantiles of the first data set against the quantiles of the second data set. A q-q plot also includes a 45-degree reference line. If the two sets come from a population with the same distribution, the points should fall approximately along this reference line.

railway preemption of traffic signals: a type of signal control at the traffic signals located in close proximity to a railroad crossing that allows the normal operation of signal lights to be preempted to assist the passage of trains (Ogden 2007).

reliability: is the ability of a system to adapt to internal changes while maintaining a satisfactory system performance. Internal changes include the variations of demand and capacity under prevailing conditions.

reliability interval of travel time: an interval associated with the predicted mean and standard deviation of travel times, which is expected to include the travel times departing at the next time interval (e.g., 15-minute) with a certain coverage rate.

resilience: is the system's ability to return to its stable state after strong perturbations from failure, disaster, or attack (Ip and Wang 2009).

short-term reliability: focuses on within-day variability in system performance for a relatively short period less than 24 hours (e.g., one hour or fifteen minutes).

stochastic: The performance parameters (e.g., travel time, speed) of a traffic system have a state or distribution that is randomly determined and that can be analyzed statistically, but is unlikely to have a precise prediction.

supply: the capacity of the roadway to accommodate vehicles within a unit of time in the traffic system.

travel time reliability: the probability that a trip between a given OD pair can be made successfully within a specific interval of time (Chen et al. 2002).

vulnerability: The concept of vulnerability can be divided into two parts: the probability of a hazardous event and the consequences of the event in a certain place or a node (Jenelius 2006).

REFERENCES

1. Abdel-Aty, M.A., Kitamura, R., Jovanis, P.P. 1995. "Investigation Effect of Travel Time Variability on Route Choice using Repeated Measurement Stated Preference Data." *Transportation Research Record: Journal of Transportation Research Board*, no. 1493, Transportation Research Board of the National Academies, Washington, D.C.:39-45.

2. Amani, P., M. Kihl, and A. Robertsson. 2011. "NARX-Based Multi-Step Ahead Response Time Prediction for Database Servers." *Proceedings of Intelligent Systems Design and Applications (ISDA) 11th International Conference*, Cordoba, Spain: 813-818.

3. Appiah, J. 2009. *Quantifying Uncertainties in Synthetic Origin-Destination Trip Matrix Estimates*. PhD Dissertation. Department of Civil Engineering, University of Nebraska-Lincoln.

4. Ardakani, K. and L. Sun. 2012. "Decremental Algorithm for Adaptive Routing Incorporating Traveler Information." *Computers & Operations Research* 39, no. 12: 3,012-3,020.

5. Bajwa, S., E. Chung, and M. Kuwahara. 2004. "An Adaptive Travel Time Prediction Model Based on Pattern Patching." Presented at 11th Intelligent Transportation Systems World Congress, Nagoya, Japan.

6. Barkley, T., R. Hranac, and K. Petty. 2012. "Relating Travel Time Reliability and Nonrecurrent Congestion with Multistate Models." *Transportation Research Record: Journal of Transportation Research Board*, no. 2278, Transportation Research Board of the National Academies, Washington, D.C.:13-20.

7. Bell, M.G.H. 2009. "Hyperstar: A Multi-Path A-star Algorithm for Risk Averse Vehicle Navigation." *Transportation Research Part B*, 43, no.1: 97-107.

8. Berdica, K., Z. Andjic, and A.J. Nicholson. 2003. "Simulating Road Traffic Interruptions: Does It Matter What Model We Use?" *Proceedings of 1st International Symposium on Transportation Network Reliability (INSTR)*, Kyoto University, Kyoto, Japan: 353-368.

9. Bowerman, B. L. and R. T. O'Connell. 1979. *Time Series and Forecasting*. Duxbury Press, North Scituate, Massachusetts.

10. Boxel, D.V., W. H. Schneider IV, and C. Bakula. 2011. "Innovative Real-time Methodology for Detecting Travel Time Outliers on Interstate Highways and Urban Arterials." *Transportation Research Record: Journal of the Transportation Research Board,* no. 2256, Transportation Research Board of the National Academies, Washington, D.C.: 60-67.

11. Bovy P.H.L., and R. Thijs. Estimators of travel time road networks, new developments, evaluation results, and applications. Delft University Press, Delft, The Netherlands, 2000.

12. Cambridge Systematics Inc. 2012. *Travel Time Data Collection. Prepared for the TRAVEL Department of Transportation, District IV*: 1-4 and 1-5.

13. Cao Y., Wu Q. 1999. "Teaching Genetic Algorithm using Matlab", Int. J. Elect. Enging. Educ., Vol. 36: 139–153.

14. Carrion, C. and Levinson, D. 2013. "Valuation of Travel Time Reliability from a GPS-based Experimental Design", *Transportation Research Part C*, Volume 35: 305-323.

15. Cassir, C. 2001. *A Flow Model for the Analysis of Transport Network Reliability*. PhD dissertation, University of Newcastle, United Kingdom.

16. Charnes, A. and W. W. Cooper. 1959. "Chance-Constrained Programming." *Management Science*, 6, no. 1: 73-79.

17. Chen, A. and C. Yang. 2004. "Stochastic Transportation Network Design Problem with Spatial Equity Constraint." *Transportation Research Record: Journal of the Transportation Research Board*, no. 1882, Transportation research board of the National Academies, Washington, D.C.: 97-104.

18. Chen, A., H. Yang, H. K. Lo, and W. H. Tang. 1999. "A Capacity Related Reliability for Transportation Networks." *Journal of Advanced Transportation*, 33, no. 2: 183-200.

19. Chen, A., H. Yang, H. K. Lo, and W. H. Tang. 2002. "Capacity Reliability of a Road Network: an Assessment Methodology and Numerical Results." *Transportation Research*, 36, no. 3: 225-252.

20. Chen, A., H. Yang, H.K. Lo, and W. H. Tang. 1999. "A Capacity Related Reliability for Transportation Networks." *Journal of Advanced Transportation*, 33, no. 2: 183-200.

21. Chen, A., J. Kim, Z. Zhou, and P. Chootinan. 2007. "Alpha Reliable Network Design Problem." *Transportation Research Record: Journal of the Transportation Research Board*, no. 2029, Transportation Research Board of the National Academies, Washington, D.C.: 49-57.

22. Chen, A., K. Subprasom, and Z. Ji. 2003. "Mean-Variance Model for the Build-Operate-Transfer Scheme under Demand Uncertainty." *Transportation Research Record: Journal of the Transportation Research Board*, no. 1857,

Transportation Research Board of the National Academies, Washington, D.C.:
93-101.

23.   Chen, A., P. Chootinan, and S. C. Wong. 2006. "New Reserve Capacity Model of
        a Signal-Controlled Road Network." *Transportation Research Record: Journal
        of the Transportation Research Board*, no. 1964, Transportation Research Board
        of the National Academies, Washington, D.C.: 35-41.

24.   Chen, A., S. Ryu, C. Yang, and S. C. Wong. 2010. "Alpha Reliable Network Design with
        Multiple Objectives and Demand Uncertainty." In *TRB 89th Annual
        Meeting Compendium of Papers DVD*.

25.   Chen, A., Z. Zhou, P. Chootinan, and S. C. Wong. 2008. "A Bi-Objective
        Reliable Network Design Model." *Presented at the 87th annual meeting of
        the Transportation Research Board*, Washington, D.C.

26.   Chen, H., H. Yang, H. K. Lo, and W. H. Tang. 2002. "Capacity Reliability of a Road
        Network: an Assessment Methodology and Numerical Results." *Transportation
        Research Part B*, 36: 225-252.

27.   Chen, M. and S. I. J. Chien. 2001. "Dynamic Freeway Travel Time Prediction with Probe
        Vehicle Data." *Transportation Research Record: Journal of the   Transportation
        Research Board,* no. 1768, Transportation Research Board of the National
        Academies, Washington, D.C.: 157-161.

28.   Chen, Y. 2014. *A Methodology of Network-Wide Signal Optimization for
        Networks with Multiple Highway-Railway At-Grade Crossings*. PhD
        Dissertation. Department of Civil Engineering, University of Nebraska-
        Lincoln.

29.   Cheslow, M., S. G. Hatcher, and V. M. Patel. 1992. *An Initial Evaluation of
        Alternative Intelligent Vehicle Highway Systems Architecture*. Report No.
        92w0000063, Mitre Corporation, Federal Highway Administration.

30.   Ching, J. 2008. "An Efficient Method for Evaluating Origin-Destination Connectivity
        Reliability of Real-World Lifeline Networks." *Computer-Aided Civil and
        Infrastructure Engineering*, 22, no. 8:584-596.

31.   Ching, J. and W. C. Hsu. 2007. "An Efficient Method for Evaluating Origin-
        Destination Connectivity Reliability of Real-World Lifeline Networks."
        Computer-Aided Civil and Infrastructure Engineering, 22, no. 8: 584-596.

32.   Cho, D. J. 2002. *Three Papers on Measuring the Reliability and Flexibility of
        Transportation System Capacity*. PhD Dissertation in Systems Engineering,
        University of Pennsylvania.

33. Chootinan, P., A. Chen, and H. Yang. 2005. "A Bi-objective Traffic Counting Location Problem for Origin-Destination Trip Table Estimation." *Transportmetrica*, 1, no. 1: 65-80.

34. City of Lincoln. 2008. "SYNCHRO Traffic Signal Data File", Obtained in May, 2009.

35. Clark, S. D., M. S, Dougherty, and H. R. Kirby. 1993. "The Use of Neural Networks and Time Series Models for Short-term Traffic Forecasting: A Comparative Study." *Transportation Planning Methods: Proceedings of PTRC 21st Summer Annual Meeting*: 151-162.

36. Cleveland, William S. 1979. "Robust Locally Weighted Regression and Smoothing Scatterplots." *Journal of the American Statistical Association,* 74 (368): 829-836.

37. Cools, M., E. Moons, and G. Wets. 2009. "Investigating the Variability in Daily Traffic Counts Using ARIMAX And SARIMA(X) Models: Assessing the Impact of Holidays on Two Divergent Site Locations." *TRB 88th Annual Meeting Compendium of Papers DVD.* Transportation Research Board, Washington, D.C.

38. D'Este, G.M. and M.A.P Taylor. 2003. "Network Vulnerability: An Approach to Reliability Analysis at the Level of National Strategic Transport Networks." *Proceedings of 1st International Symposium on Transportation Network Reliability (INSTR),* Kyoto University, Kyoto, Japan: 23-44.

39. Davis, G. and M. Nihan. 1991. "Nonparametric Regression and Short-Term Freeway Traffic Forecasting." *Journal of Transportation Engineering*, 117, no. 2: 178-188.

40. Dowdy, S., S. Weardon, and D. Chilko. 2005. *Statistics for Research*, Third Edition. Wiley Series in Probability and Statistics.

41. Du, Z. P. and A. Nicholson. 1997a. "Degradable Transportation Systems: an Integrated Equilibrium Model." Transportation Research B, 31, no. 3: 209-223.

42. Du, Z. P. and A. Nicholson. 1997b. "Degradable Transportation Systems: Sensitivity and Reliability Analysis." Transportation Research B, 31, no. 3: 225-237.

43. Efron, B. 1979. "Bootstrap Methods: Another Look at the Jackknife." *The Annals of Statistics*, 7, no. 1:1-26.

44.   Efron, B. and R. J. Tibshirani. 1993. *An Introduction to the Bootstrap*. Chapman & Hall, NY. London.

45.   Eisele, W. L. 2001. *Estimating Travel Time Mean and Variance Using Intelligent Transportation Systems Data for Real-Time and Off-Line Transportation Applications.* PhD dissertation, Department of Civil Engineering, Texas A&M University.

46.   Fei, X., C. Lu, and K. Liu. 2011. "A Bayesian Dynamic Linear Model Approach for Real-Time Short-Term Freeway Travel Time Prediction." *Transportation Research Part C,* 19, no. 6: 1306-1318.

47.   Fraley, C. and A. E. Raftery. 2002. "Model-Based Clustering, Discriminant Analysis, and Density Estimation." *Journal of the American Statistical Association*, 97, no. 458: 611-631.

48.   Fu, L. 2001. "An Adaptive Routing Algorithm for In-Vehicle Route Guidance Systems with Real-Time Information." *Transportation Research Part B*, 35, no. 8: 749-765.

49.   Fu, L. and L. R. Rilett. 1998 "Expected Shortest Paths in Dynamic and Stochastic Traffic Networks." *Transportation Research Part B*, 32, no. 7: 499-516.

50.   Fu, L., D. Sun, and L. R. Rilett. 2006. "Heuristic Shortest Path Algorithms for Transportation Applications: State of the Art." *Computers & Operations Research,* 33, no. 11: 3324-3343.

51.   Gao, S. and I. Chabini. 2006. "Optimal Routing Policy Problems in Stochastic Time-Dependent Networks." *Transportation Research Part B*, 40, no.2: 93-122.

52.   Haghani, A., H. Masoud, and F. S. Kaveh, S. Young, and P. Tarnoff. 2010. "Freeway Travel Time Ground Truth Data Collection Using Bluetooth Sensors." *Presented on the 89th Annual Meeting of the Transportation Research Board*, Washington, DC.

53.   Haitham, A. and E. B. Emam. 2006. "New Methodology for Estimating Reliability in Transportation Networks with Degraded Link Capacities." *Journal of intelligent transportation systems*, 10, no. 3: 117-129.

54.   Hall, R. 1986. "The Fastest Path through a Network with Random Time- Dependent Travel Time." *Transportation Science,* 20, no. 3: 182-188.

55.   Haykin, S. 1999. *Neural Networks: A Comprehensive Foundation*, 2nd ed., Upper Saddle River, NJ: Prentice Hall.

56.   Heydeckera, B. G., W. H. K. Lamb, and N. Zhang. 2007. "Use of Travel Demand

Satisfaction to Assess Road Network Reliability." *Transportmetrica*, 3, no. 2: 139-171.

57.   *Highway Capacity Manual.* 2010. Transportation Research Board Publication, Washington, DC.

58.   Hornik, K. M., M. Stinchcombe, and H. White. 1989. "Multiplayer Feed-Forward Networks Are Universal Approximators." *Neural Networks*, 2, no. 5: 5359-366.

59.   Husch, D. and J. Albeck. 2003. *Intersection Capacity Utilization: Evaluation Procedures for Intersections and Interchanges*. Trafficware.

60.   Hyndman, R. 2013. "The Difference between Prediction Intervals and Confidence Intervals." http://robjhyndman.com/hyndsight/intervals/. Accessed on May 30, 2013.

61.   Iida, Y. 1999. "Basic Concepts and Further Direction of Road Network Reliability Analysis." *Journal of Advanced Transportation*, 33, no. 2: 125-134.

62.   Ip, W.H. and D. Wang. 2009. "Resilience Evaluation Approach of Transportation Networks." *Proceedings of International Joint Conference on Computational Sciences and Optimization*.

63.   Jenelius, E., T. Petersen, L. Mattsson. 2006. "Importance and Exposure in Road Network Vulnerability Analysis." *Transportation Research Part A*, 40, no. 7: 537-560.

64.   Jones, E. G. 2008. "The University of Nebraska-Lincoln Highway-Rail Grade Crossing Test Bed." *Proceedings of the 18th Annual International Railway Safety Conference*.

65.   Jones, E. G., A. H. Khattak, and L. R. Rilett. 2009. "Highway-Rail Grade Crossing Test Bed System at the University Nebraska-Lincoln." *Proceedings of Transportation Research Board 88th Annual Meeting*, Transportation Research Board of the National Academics, Washington, D.C.

66.   Juri, N. R., A. Unnikrishnan, and S. T. Waller. 2007. "Integrated Traffic Simulation-Statistical Analysis Framework for Online Prediction of Freeway Travel Time." *Transportation Research Record: Journal of the Transportation Research Board*, No. 2039, Transportation Research Board of the National Academies, Washington, D.C.: 24-31.

67.   Kaparias, I. and M. G. H. Bell. 2009. "Testing A Reliable In-Vehicle Navigation Algorithm in the Field." *Intelligent Transportation Systems*, 3, no. 3: 314-

324.

68.   Kaparias, I. and M. G. H. Bell. 2010. "A Reliability-Based Dynamic Re-Routing Algorithm for In-Vehicle Navigation." *Proceeding of 13st International IEEE Annual Conference on Intelligent Transportation Systems*, Madeira Island Portugal.

69.   Kaparias, I., M. G. H. Bell, and H. Belzner. 2008. "A New Measure of Travel Time Reliability for In-Vehicle Navigation Systems." *Journal of Intelligent Transportation Systems*, 12, no. 4: 202-211.

70.   Karlaftis, M. G., and E. I. Vlahogianni. 2011. "Statistical Methods versus Neural Networks in Transportation Research: Differences, Similarities and Some Insights." *Transportation Research Part C*, 19, no. 3: 387-399.

71.   Karlsson, M., and S. Yakowitz. 1987. "Rainfall-Runoff Forecasting Methods, Old and New." *Stochastic Hydrology and Hydraulics,* no. 1: 303-318.

72.   Kim, D. S., J. D. Porter, M. E. Magana, S. Park, and A. Saeedi. 2012. *Wireless Data Collection System for Travel Time Estimation and Traffic Performance Evaluation*. Final Report, SPR 737 OTREC-RR012-06: 40-45.

73.   Kim, S., M. E. Lewis, and C. C. White. 2005. "Reduction for Non-Stationary Stochastic Shortest Path Problems with Real-Time Traffic Information." *IEEE Transactions on Intelligent Transportation Systems*, 6, no. 3: 273-283.

74.   Kim, T., H. Kim, and D. Lovell. 2005. "Traffic Flow Forecasting: Overcoming Memory-Less Property in Nearest Neighbor Non-Parametric Regression." *Proceedings of the 8th International IEEE Conference on Intelligent Transportation Research Board*. Washington, D.C.

75.   Kisgyorgy, L. and L. R. Rilett. 2002. "Travel Time Prediction by Advanced Neural Network." *Periodica Polytechnica Ser. Civil Engineering*, 46, no. 1: 15-32.

76.   Kuchipudi, C.M., and S. I. J. Chien. 2003. "Development of a Hybrid Model for Dynamic Travel Time Prediction." *Transportation Research Record: Journal of the Transportation Research Board*, No. 1855, Transportation Research Board of the National Academies, Washington, D.C.: 22-31.

77.   Lahiri, S. N., C. Spiegelman, J. Appiah, and L. R. Rilett. 2012. "Gap Bootstrap for Massive Data Sets with an Application to Transportation Engineering." *The Annals of Applied Statistics*, 6, no. 4: 1552-1587.

78. Li, R. and G. Rose. 2011. "Incorporating Uncertainty into Short-Term Travel Time Predictions." *Transportation Research Part C,* 19, no. 6: 1006-1018.

79. Lilliefors, H. 1967. "On the Kolmogorov–Smirnov Test for Normality with Mean and Variance Unknown." *Journal of the American Statistical Association*, 62: 399-402.

80. Lin, W., A. Kulkarni, and P. Mirchandani. 2004. "Short-Term Arterial Travel Time Prediction for Advanced Traveler Information Systems." *Intelligent Transportation Systems*, 8, no. 3: 143-154.

81. Liu, H., H. Van Zuylen, H. Van Lint, and M. Salomons. 2006. "Predicting Urban Arterial Travel Time with State-Space Neural Networks and Kalman Filters." *Journal of Transportation Research Board*, no. 1968, Transportation Research Board of the National Academies, Washington, D.C.: 99-108.

82. Lo, H. and Y. K. Tung. 2000. "A Chance Constrained Network Capacity Model." *Reliability of* Transport Networks Edited by M. Bell and C. Cassir, Research Studies Press Limited, Chapter 11: 159-172.

83. Lo, H. and Y.K. Tung. 2003. "Network with Degradable Links: Capacity Analysis and Design." Transportation Research Part B, 37, no. 4: 345-363.

84. Makridakis, S., S.C. Wheelwright, and R.J. Hyndman. 1998. Forecasting Methods and Applications, 3rd Edition, Wiley.

85. Manley E. and T. Cheng. 2010. "Understanding Road Congestion as an Emergent Property of Traffic Networks." *Proceedings of the 14th World Multi-conference on Systemics, Cybernetics, and Informatics: WMSCI 2010*, International Institute of Informatics and Systemics (IIIS): Caracas, Venezuela: 109-114.

86. McCleary, R., and R. Hay. 1980. *Applied Time Series Analysis for the Social Sciences*. Beverly Hills: Sage: 20.

87. Meehan, B. and B. Rupert. 2004. Putting Travelers in the Know. *Public Roads*, 68, no. 3, Publication Number: FHWA-HRT-05-002. http://www.fhwa.dot.gov/publications/publicroads/04nov/06.cfm. Accessed on Mar 15, 2014.

88. Meyer, M. and J.E. Miller. 2001. *Urban Transportation Planning: a Decision-Oriented Approach*, 2nd Edition. McGraw-Hill Companies.

89. Miller-Hooks, E. 2001. "Adaptive Least-Expected Time Paths in Stochastic Time-Varying Transportation and Data Networks." *Networks*, 37, no. 1: 35-52.

90. Morlok, E. K. and D. J. Chang. 2004. "Measuring Capacity Flexibility of a Transportation System." *Transportation Research Part A*, 38, no. 6: 405-420.

91. Nagel, K. and S. Rasmussen. 1994. *Traffic on the Edge of Chaos*. Working Paper 94-06-032, Santa Fe, Santa Fe Institute: N.M.

92. Naik, B. 2010. *A Generalized Non-Parametric Approach for Uncertainty Evaluation in Travel Time Prediction Models*. PhD Dissertation. Department of Civil Engineering, University of Nebraska-Lincoln.

93. Nebraska Department of Roads, "*Administration of State and Federal Highway-Rail Grade Crossing Safety Projects*." Nebraska Administrative Code, July 2011, Title 415:47.

94. Nicholson, A., J. Schmocker, G. H. B. Michael, and Y. Iida. 2003. "Assessing Transport Reliability: Malevolence and User Knowledge." *Proceedings of 1st International Symposium on Transportation Network Reliability (INSTR)*, Kyoto University, Kyoto, Japan: 1-22.

95. Nikovski, D., N. Nishiuma, Y. Goto, and H. Kumazawa. 2005. "Univariate Short-Term Prediction of Road Travel Times." *Proceedings of the 8th International IEEE Conference on Intelligent Transportation Systems*, Vienna, Austria: 1074-1079.

96. O'Neill, E., V. Kostakos, T. Kindberg, A. F. gen. Shiek, A. Penn, D. S. Fraser, and T. Jones. 2006. "Instrumenting the City: Developing Methods for Observing and Understanding the Digital Cityscape." *Proceedings of 8th International Conference on Ubiquitous Computing*: 315-332.

97. Ogden, B. D. 2007. *Railroad-Highway Grade Crossing Handbook*, FHWA report: FHWA-SA-07-010.

98. Okutani, I. and Y. Stephanedes. 1984. "Dynamic Prediction of Traffic Volume through Kalman Filtering Theory." *Transportation Research Part B*, 18, no. 1: 1-11.

99. Opasanon, S. and E. Miller-Hooks. 2006. "Multicriteria Adaptive Paths in Stochastic, Tim-Varying Networks." *European Journal of Operational Research*, 173, no. 1: 72-91.

100. Oregon Department of Transportation. 2005 *Travel Time Messaging on Dynamic Message Signs – Portland, OR*, Oct 2005, Publication Number: FHWA-hop-05-048. Accessed on Feb 15, 2014.

101.   Oregon Department of Transportation. 2006. *VISSIM Calibration and Validation*. Technical Report.

102.   Park, D. 1998. *Multiple Path Based Vehicle Routing in Dynamic and Stochastic Transportation Networks*. PhD Dissertation. Department of Civil Engineering, Texas A&M University.

103.   Park, D. and L. R. Rilett. 1998. "Forecasting Multiple-Period Freeway Link Travel Times Using Modular Neural Networks." *Transportation Research Record,* no. 1617, pp. 163-170.

104.   Park, D. and L. R. Rilett. 1999. "Forecasting Freeway Link Travel Times with a Multilayer Feed-Forward Neural Network." *Computer-Aided Civil and Infrastructure Engineering,* 1, no. 5: 357-367.

105.   Park, D., L. R. Rilett, and G. Han. 1999. "Spectral Basis Neural Networks for Real-Time Travel Time Forecasting." *Journal of. Transportation Engineering*, 125, no. 6: 515–523.

106.   Petris, G., S. Petrone, and P. Campagnoli. 2007. *Dynamic Linear Models with R*. Springer.

107.   PTV Planung Transport Verkehr AG. 2011. *VISSIM 5.40-01 User Manual*.

108.   Puckett, D. D. and M. J. Vickich. 2010. *Bluetooth-Based Travel Time/Speed Measuring Systems Development*. UTCM Project #29-00-17 Final Report.

109.   Qiao, W., A. Haghani, and M. Hamedi. 2012. "A Non-Parametric Model for Short-Term Travel Time Prediction Using Bluetooth Data." *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 17, No. 2, pp. 165-175.

110.   Quayle, S., P. Koonce, D. DePencier, and D. Bullock. 2010. "Arterial Performance Measures Using MAC Readers: Portland Pilot   Study." *Transportation Research Record: Journal of the Transportation Research Board,* no. 2192, Transportation Research Board of the National Academies, Washington, D.C.: 185-193.

111.   Robinson, S. and J. Polak. 2005. "Modeling Urban Link Travel Time with Inductive Loop Detector Data by Using KNN Method." *Transportation Research Record: Journal of the Transportation Research Board*, no. 1935, Transportation Research Board of the National Academies, Washington, D.C.: 47-56.

112. Roess, R. P., E. S. Prassas, and W. R. McShane. 2010. *Traffic Engineering*, 4th Edition, Prentice Hall.

113. SAS Institute. *SAS/STAT 9.2 User's Guide Chapter 50: The LOESS Procedure*. Accessed on May 28th, 2013.

114. Shen, L. and M. Huang. 2011. "Assessing Dynamic Neural Networks for Travel Time Prediction." *Applied Informatics and Communication: Communications in Computer and Information Science*, 224: 196-477.

115. Shumway, R. H. and D. S. Stoffer. 2011. *Time Series Analysis and Its Applications with R Examples*, 3rd Edition. Springer.

116. Smith, B. and M. Demetsky. 1994. "Short-Term Traffic Flow Prediction Neural Network Approach." *Transportation Research Record: Journal of the Transportation Research Board,* no. 1453, Transportation Research Board of the National Academies, Washington, D.C.: 98-104.

117. Spiegelman, C. H., E. S. Park, and L. R. Rilett. 2010. *Transportation Statistics and Microsimulation*. CRC Press.

118. Sumalee, A., D. P. Watling, and S. Nakayama. 2006 "Reliable Network Design Problem: the Case with Uncertain Demand and Total Travel Time Reliability." *Transportation Research Record: Journal of the Transportation Research Board*, No. 1964, Transportation Research Board of the National Academies, Washington, D.C.: 81-90.

119. Sumalee, A., P. Luathep, W. H. K. Lam, and R.D. Connors. 2009. "Transport Network Capacity Evaluation and Design under Demand Uncertainty." *Transportation Research Record: Journal of the Transportation Research Board*, No. 2090, Transportation research board of the National Academies, Washington, D.C.: 470-479.

120. Tarnoff, P. J., J. S. Wasson, S. E. Young, N. Ganig, D. M. Bullock, and J. R. Sturdevant. 2009. "Continuing Evolution of Travel Time Data Information Collection and Processing." *Presented at the 88th Transportation Research Board Annual Meeting*, Washington, D.C.

121. Taylor, M. A. P. and G. M. D'Este. 2004 "Safeguarding Transport Networks: Assessment of Network Vulnerability and Development of Remedial Measures." *Australian Journal of Multidisciplinary Engineering, Special Edition on Engineering a Secure Australia*: 13-22.

122.   Train, K. E. 2003. *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge, UK.

123.   Transportation Research Board. 2013. *Analytical Procedures for Determining the Impacts of Reliability Mitigation Strategies*. SHRP2 Reliability Research Report S2-L03-RR-1: 2.

124.   Tu, H. 2008. *Monitoring Travel Time Reliability on Freeways*. PhD Dissertation. Delft University of Technology.

125.   Turner, S. M., W. L. Eisele, R. J. Benz, and D. J. Holdener. 1998. *Travel Time Data Collection Handbook*. Report No. FHWA-PL-98-035.

126.   Ukkusuri, S. V.S.K. 2005. *Accounting For Uncertainty, Robustness and Online Information in Transportation Network*. PhD Dissertation, Department of Civil Engineering, University of Texas at Austin.

127.   Van Lint, J. W. C., S. P. Hoogendoorn, and H. J. Van Zuvlen. 2002. "Freeway Travel Time Prediction with State-Space Neural Networks Modeling State-Space Dynamics with Recurrent Neural Networks." *Transportation Research Record: Journal of the Transportation Research Board*, no. 1811, Transportation Research Board of the National Academies, Washington, D.C.: 30-39.

128.   Van Lint, J. W. C., S. P. Hoogendoorn, and H.J. Van Zuylen. 2005. "Accurate Freeway Travel Time Prediction with State-Space Neural Networks under Missing Data." *Transportation Research Part C,* 13, no. 5-6: 347-369.

129.   Vo, T. 2011. *An Investigation of Bluetooth Technology for Measuring Travel Times on Arterial Roads: A Case Study on Spring Street*. School of Civil & Environmental Engineering, Georgia Institute of Technology.

130.   Wakabayashi H. and Y. Iida. 1992. Upper and Lower Bounds of Terminal Reliability of Road Networks: an Efficient Method with Boolean Algebra. *Journal of Natural Disaster Science*, 14, no. 1: 29-44.

131.   Waller, S. T., and A. K. Ziliaskopoulos. 2001. "Stochastic Dynamic Network Design Problem." *Transportation Research Record: Journal of the Transportation Research Board*, no. 1771, Transportation Research Board of the National Academies, Washington, D.C.: 106-113.

132.   Wardrop, J. G. 1952. "Some Theoretical Aspects of Road Traffic Research." *Proceedings of Institute of Civil Engineers: Engineering Divisions*, 1, no. 3: 325–362.

133.   Washington, S. P., M. G. Karlaftis, and F. L. Mannering. 2011. *Statistical and*

Econometric Methods for Transportation Data Analysis, 2nd Edition. CRC Press.

134.   Welch, G. and G. Bishop. 1995. *An Introduction to the Kalman Filter Course Notes TR 95-041*, Department of Computer Science, University of North Carolina at Chapel Hill: 24.

135.   West, M. and J. Harrison. 1997. *Bayesian Forecasting and Dynamic Models*, 2nd edition. Springe.

136.   Winick, M. R. 2012. "Reliability of Personal Travel Time and Transportation System." *Transportation Planning Council Winter E-Newsletter*, Institute of Transportation Engineering.

137.   Wojtal, M. R. 2012. *Development of a methodology for analyzing safety treatments at isolated signalized intersections.* PhD Dissertation. Department of Civil Engineering, University of Nebraska-Lincoln.

138.   Xia, J. 2006. *Dynamic Freeway Travel Time Prediction Using Single Loop Detector and Incident Data.* PhD Dissertation. Department of Civil Engineering, University of Kentucky.

139.   Xie, Y., Y. Zhang, and Z. Ye. 2007. "Short-term Traffic Volume Forecasting Using Kalman Filter with Discrete Wavelet Decomposition." *Computer-Aided Civil and Infrastructure Engineering*, 22, no. 5: 326-334.

140.   Yang, B. Y. and E. Miller-Hooks. 2004. "Adaptive Routing Considering Delays Due To Signal Operations." *Transportation Research Part B*, 38, no. 5: 385-413.

141.   Yang, F., Z. Yin, H.X. Liu, and B. Ran. 2004. "Online Recursive Algorithm for Short Term Traffic Prediction." *Transportation Research Record: Journal of the Transportation Research Board*, no. 1879, Transportation Research Board of the National Academies, Washington, D.C.: 1-8.

142.   Yin, H., S. C. Wong, J. Xu, and C. K. Wong. 2002. "Urban Traffic Flow Prediction Using a Fuzzy-Neural Approach." *Transportation Research Part C*, 10, no. 2: 85-98.

143.   Yin, Y., S. M. Madanat, and X. Lu. 2009. "Robust Improvement Schemes for Road Networks under Demand Uncertainty." European Journal of Operational Research, 198, no. 2: 470-479.

144.   Yu, L., X. Li, and W. Zhuo. 2003. Airport Related Traffic and Mobile Emission
         Implications. Report No. TxDOT 4317, Texas Department of Transportation.

145.   Zeng, X. and Y. Zhang. 2013. "Development of Recurrent Neural Network
         Considering Temporal-Spatial Input Dynamics for Freeway Travel Time
         Modeling." *Computer-Aided Civil and Infrastructure Engineering*, 28, no.
         5: 359-371.

146.   Zhao, Y. and K. Kockelman. 2002. "The Propagation of Uncertainty through Travel
         Demand Models: An Exploratory Analysis." *Annals of Regional Science*, 36, no.
         1: 145-163.

147.   Zheng, W., D. Lee, and Q. Shi. 2006. "Short-Term Freeway Traffic Flow
         Prediction: Bayesian Combined Neural Network Approach." *Journal of
         Transportation Engineering*, 132, no. 2: 114-121.