UC Berkeley
Research Reports

Title
Considering Risk-Taking Behavior in Travel Time Reliability

Permalink
https://escholarship.org/uc/item/59v844cr

Authors
Recker, Will
Chung, Younshik
Park, Jiyoung
et al.

Publication Date
2005

Considering Risk-Taking Behavior in Travel Time Reliability

Will Recker, Younshik Chung, Jiyoung Park, Lesley Wang, Anthony Chen, Zhaowang Ji, Henry Liu, Matthew Horrocks, Jun-Seok Oh

The contents of this report reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California. This report does not constitute a standard, specification, or regulation.

Final Report for Task Order 4110

January 2005

CALIFORNIA PARTNERS FOR ADVANCED TRANSIT AND HIGHWAYS

CONSIDERING RISK-TAKING BEHAVIOR IN TRAVEL TIME RELIABILITY

REPORT prepared for California PATH

By

Will Recker, Younshik Chung, Jiyoung Park, and Lesley Wang
Institute of Transportation Studies
University of California, Irvine
CA 92697-3600

Anthony Chen, Zhaowang Ji, Henry Liu, and Matthew Horrocks
Department of Civil and Environmental Engineering
Utah State University
Logan, UT 84322-4110

Jun-Seok Oh
Department of Civil and Construction Engineering
Western Michigan University
Kalamazoo, MI 49008-5316

October 2004

CONSIDERING RISK-TAKING BEHAVIOR IN TRAVEL TIME RELIABILITY

ABSTRACT

Travel time variability is increasingly being recognized as a major factor influencing travel decisions and, consequently, as an important performance measure in transportation management. In this research project, we examine a number of questions related to travel time variability: How should travel time variability be quantified at both the section level as well as at the route level?; How do travelers value travel time and its reliability?; How much does the travel time reliability contribute to travelers' route choices?; How much variation is there in travelers' preferences regarding the potential tradeoff between reliability and travel time itself?; How can travel time variability be incorporated into the route choice models for transportation planning purposes?; and, How can the effects of travel time reliability be incorporated in considering risk-taking behavior in route choice models? Answering these questions can help in the design and evaluation of transportation planning and managing strategies.

Keyword: Travel time variability, travel time reliability, mixed logit, route choice models, traffic assignment

EXECUTIVE SUMMARY

Recent empirical studies (Abdel-Aty *et al*., 1996; Small *et al*., 1999; Kazimi *et al*., 1999; Lam, 2000; Lam and Small, 2001; Ghosh, 2001; Bates *et al*., 2001; Noland and Polak, 2002) suggest that travelers are interested in not only travel time saving but also reduction of travel time variability. Variability in network travel times introduces uncertainty for travelers such that they do not know exactly when they will arrive at their respective destinations. Thus, it is considered as a risk (or an added cost) to a traveler making a trip. This risk may be manifest in a willingness to pay a premium (e.g., through use of toll roads) to avoid congestion and to achieve greater reliability in travel times.

Although travel time reliability ostensibly plays an important role in the traveler's route choice behavior, there has been little basic research directed toward an understanding of the effects of reliability on the route choice decision making of the traveler; many questions remain unanswered. How to quantify travel time variability in both section level as well as route level, how do travelers value travel time and its reliability, how much does the travel time reliability contribute to travelers' route choices, how much variation is there in travelers' preferences regarding the potential tradeoff between reliability and travel time itself, and how to incorporate travel time variability into the route choice model for transportation planning purposes, and how to estimating travel time reliability considering risk-taking behavior in route choice models? Answering these questions can help in the design and evaluation of transportation planning and managing strategies. The objective of this research is to develop methods to address the above questions. Specific research activities included in this report are:

- Development of a GIS database to facilitate travel time variability analysis.
- Development of computation procedures to calculate freeway travel time variability for both section level and route level.

4

- Development of a sound methodology to uncover the contribution of travel time reliability in route choice using real-time loop data from the California Advanced Transportation Management Systems (ATMS) Testbed at UC Irvine combined with real-world data from the State Route 91 Value Pricing Project in Orange County, California to estimate a mixed logit model that explicitly considers travelers' risk perceptions and preferences in making route choice decisions under an uncertain environment.
- Development of route choice models that account for the variations of travel time in the form of risk with perception errors and different behavioral preferences.
- Integration of risk-taking behavior into the estimation of travel time reliability.
- Conducting numerical experiments to examine the effects of incorporating travel time variability and risk-taking behavior into the route choice models and its impact on the estimation of travel time reliability under demand and supply variations.

The significance of understanding travel time reliability (or reduction of travel time variability) is important both for the users as well as for traffic managers. From the users' point of view, travel time variability reduction means more predictable travel times. Better predictable travel times imply improved scheduling matches since travel time can be predicted more precisely. For freight transportation, reduction of travel time variability can improve just-in-time inventory management. From the traffic managers' point of view, less travel time variability means better stability of the quality of service. That is, there are less speed variations along a path (i.e., less stop and go conditions). This directly contributes to less fuel consumption and emission since there are fewer vehicles undergoing acceleration/deceleration cycles. It should be noted that travel time variability reduction is a direct measure of pollution emission without the need to resort to proxy variables, which may introduce inaccuracy in calculating vehicle emission. In addition, reducing travel time (or speed) variations has the potential to increase traffic safety (or decrease accident risk). Results in this research project contribute to better planning, management, and evaluation of transportation networks.

Recommendations

The importance of travel time variability in route choice decisions that lead to aggregate demand on California's freeways has been documented by this study. However, the understanding of the complex effects of travel time variability is still in its infancy and applications to transportation management have not yet been developed. The first step for advancement to the application stage is to build a database for detailed analysis. As an example, a GIS-based database composed of the historic traffic data from freeways in Orange County, California, was developed under this project. The main purpose of such a database is to provide the analysis framework to analyze the potential impacts of travel time variability at both section-level as well as route-level. In day-to-day management, real-time traffic data used for traffic operations is often discarded without archiving, or archived in such a raw form that it presents a barrier to fusion with current field data. These historical archived traffic data can play an important role in real-time management by measuring freeway performance relative to travel time variability from a long-term perspective. We recommend that the type of database developed in this study be incorporated as standard practice in Caltrans TMCs as a foundation for better understanding travel time variability as a measure of transportation system reliability. The next step would be to identify the sources of travel time variability and to study how to remove the source of variability from the context of transportation system management.

The results of this study confirm that travel time reliability can have significant influence on traveler's route choice behavior and that it cannot be ignored in any model which purports to predict behavior or provide a basis for performance evaluation. In the study, we formulated traveler's route choice as a mixed-logit model, with the coefficients in the model representing individual traveler's preferences or tastes to travel time, reliability and cost, and applied the model to the California State Route 91 value-pricing project. Based on travelers' choice of whether or not to pay a congestion-based toll in order to use express lanes, we were able to estimate how travelers value travel time and travel-time reliability. Using a Monte Carlo simulation procedure to simulate risk perceptions and preferences in making route choice decisions under an uncertain environment, numerical

results were also presented to examine what the aggregate impact of changes in variability caused by demand and supply variations might have on network assignment and how travelers with different risk-taking behaviors respond to these changes. These models represent an initial foray into better understanding how individual travelers will react to the changing characteristics of the traffic patterns they experience in their daily commute. Although encouraging, the results are as yet not generalizable for use in traffic management. We recommend that, in conjunction with the formation of the databases noted above, additional investigation be pursued into development and calibration of a family of route choice/traffic assignment models that explicitly incorporate route travel time variability in the decisions of California motorists.

The framework provided by this research is expected to be useful in examining what the aggregate impact of changes in variability caused by demand and supply variations might have on network assignment and how travelers with different risk-taking behaviors respond to these changes. Although this research project has addressed many questions related to travel time variability, our understanding of travel time variability is still in its infancy and applications to real-world transportation management have not yet developed. Further studies are needed to better understand travel time variability and how to use such measures for transportation management.

Implementation Strategy

The obvious implementation strategy for "regularizing" the compilation of the travel time variability databases essential for the framework would be to incorporate these calculation procedures in PEMS. Presumably, this could be accomplished using the raw data input that already is part of PEMS, together with new route enumeration procedures that would identify likely paths between major decision points in the respective freeway networks. A series of calibrated route choice models could also be incorporated in PEMS that would access both the historical travel time variability matrices associated with the potential routes, together with the real-time data already incorporated in PEMS.

Contents

Figure List

Table List

Chapter One

1. INTRODUCTION

Travel time variability has recently become an important measure in transportation management and has begun being treated as a major factor influencing travel decisions. Recent empirical studies (Abdel-Aty *et al*., 1996; Small *et al*., 1999; Kazimi *et al*., 1999; Lam, 2000; Lam and Small, 2001; Ghosh, 2001; Bates *et al*., 2001; Noland and Polak, 2002) suggest that travelers are interested in not only travel time saving but also reduction of travel time variability. Variability in network travel times introduces uncertainty for travelers such that they do not know exactly when they will arrive at their respective destinations. Thus, it is considered as a risk (or an added cost) to a traveler making a trip. This risk may be manifest in a willingness to pay a premium (e.g., through use of toll roads) to avoid congestion and to achieve greater reliability in travel times.

Although travel time reliability ostensibly plays an important role in the traveler's route choice behavior, there has been little basic research directed toward an understanding of the effects of reliability on the route choice decision making of the traveler; many questions remain unanswered. How to quantify travel time variability in both section level as well as route level, how do travelers value travel time and its reliability, how much does the travel time reliability contribute to travelers' route choices, how much variation is there in travelers' preferences regarding the potential tradeoff between reliability and travel time itself, and how to incorporate travel time variability into the route choice model for transportation planning purposes, and how to estimating travel time reliability considering risk-taking behavior in route choice models? Answering these questions can help in the design and evaluation of transportation planning and managing strategies. The goal of this research project is to develop methods to address the above questions. Specifically, the objectives are to:

1. Develop a Geographical Information Systems (GIS) database to facilitate traffic analysis for freeways in Orange County, California.
2. Develop computation procedures to calculate freeway travel time variability for both section level and route level.
3. Develop a mixed logit model that explicitly considers travelers' risk perceptions and preferences when choosing routes based on expected travel time and travel time variability.
4. Estimate the mixed logit model using real-time loop data from the California Advanced Transportation Management Systems (ATMS) Testbed at UC Irvine combined with real-world data from the State Route (SR) 91 Value Pricing Project in Orange County, California.
5. Develop route choice models that account for the variations of travel time in the form of risk with perception errors and different behavioral preferences.
6. Develop traffic assignment procedures for solving the risk-taking route choice models.
7. Integrate risk-taking behavior into the estimation of travel time reliability.
8. Conduct numerical experiments to examine the effects of incorporating travel time variability and risk-taking behavior into the route choice models and its impact on the estimation of travel time reliability under demand and supply variations.

The significance of understanding travel time reliability (or reduction of travel time variability) is important both for the users as well as for traffic managers. From the users' point of view, travel time variability reduction means more predictable travel times. Better predictable travel times imply improved scheduling matches since travel time can be predicted more precisely. For freight transportation, reduction of travel time variability can improve just-in-time inventory management. From the traffic managers' point of view, less travel time variability means better stability of the quality of service. That is, there are less speed variations along a path (i.e., less stop and go conditions). This directly contributes to less fuel consumption and emission since there are fewer vehicles undergoing acceleration/deceleration cycles. It should be noted that travel time variability reduction is a direct measure of pollution emission without the need to resort to proxy variables, which may introduce inaccuracy in calculating vehicle emission. In addition, reducing travel time (or speed) variations has the potential to increase traffic safety (or

decrease accident risk). Results in this research project contribute to better planning, management, and evaluation of transportation networks.

The report is organized as follows. Chapter Two reviews the route choice models and its applications in surface transportation. Chapter Three describes the GIS database development and computation procedures for calculating freeway travel time variability. Chapter Four discusses the Mixed Logit model used to estimate the value of time and value of reliability using real-time loop data from the State Route (SR) 91 Value Pricing Project in Orange County, California. Chapter Five explains the implementation of the traffic assignment procedures for the route choice models described in Chapter Two. Chapter Six integrates the risk-taking route choice models into the travel time reliability evaluation procedure. Chapter Seven provides a summary of the findings in this research project.

Chapter Two

2. ROUTE CHOICE MODELS AND ITS SIGNIFICANCE IN SURFACE TRANSPORTATION APPLICATIONS

Because the travelers' selections of different paths will result in different levels of travel time reliability, travel time reliability analysis is closely related to the route choice problem. It is well known that the route choice problem is one of the most fundamental topics in transportation studies since route choice modeling is widely applied in transportation planning, network design, intelligent transportation system (ITS), etc. Figure 2.1 depicts the role of route choice modeling in applications of surface transportation. This framework consists of three basic applications: I. The Traffic Assignment Problem, II. Traffic Management and Control, and III. Traveler Information Systems. Brief descriptions about each application are given in the following.
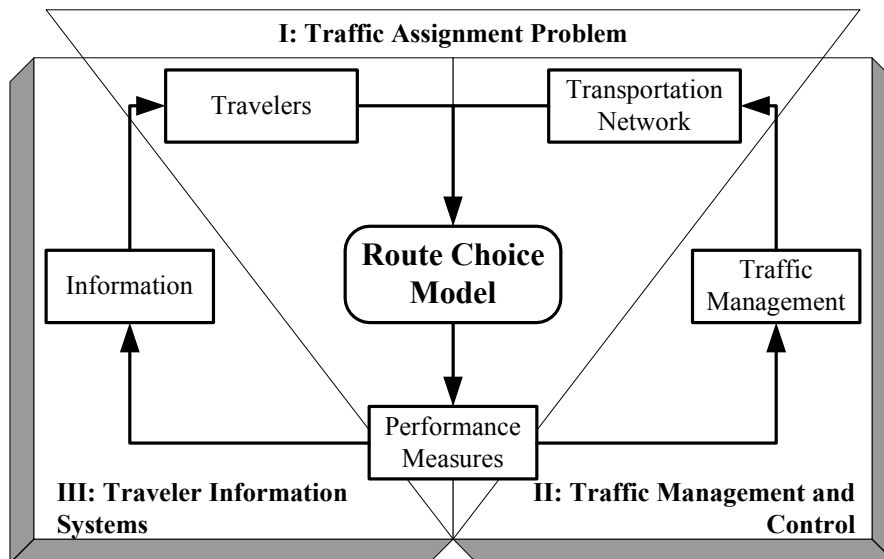
Figure 2.1: Route Choice Model Applications

I: Traffic Assignment Problem

The traffic assignment problem is the final step in the 4-step travel forecasting process (Ortuzar and Willumsen, 2001; Meyer and Miller, 2001). Given constant travel demand between origin-destination (O-D) pairs (i.e., travelers), and travel cost functions for each link of the network (i.e., transportation network), the traffic assignment problem is to determine the traffic flow pattern as well as network performance measures (e.g. total system travel time, vehicle miles of travel, vehicle hours of travel, fuel consumption and emission, etc.). The key component, in the traffic assignment problem, is the route choice model which represents individual route choice decision between various O-D pairs such that traffic flow patterns for the whole transportation network can be predicted. In route choice models, congestion can be explicitly considered through the travel cost functions and the interactions with route choice decisions of the travelers.

II: Traffic Management and Control

Traffic management and control is a two-way game (Bovy and Stern, 1990). That is, the traffic manager sets the traffic management and control strategies to regulate traffic based on the performance measures resulting from the aggregation of individual route choice decisions. It should be recognized that the traffic management and control strategies, set by the traffic manager, can only influence (not control) the route choice decisions of travelers. Thus, it is important to understand the factors that influence route choices so that better and more effective strategies can be developed to combat congestion. Several studies on the commonly used management and control strategies, used to account for travelers' route choices in a congested network, are listed in the following:

- Ramp metering control (Yang et al., 1994; Yang and Yagar, 1994).
- Signalized intersection control (Allsop, 1974; Allsop and Charlesworth, 1977; Fisk, 1984; Cantarella et al., 1991; Van Vuren and Van Vliet, 1992; Smith and Van Vuren, 1993; Yang and Yagar, 1995; Wong and Yang, 1997; Chiou, 1999; Wong and Yang, 1999).
- Road pricing (Ferrari, 1995; Yang and Lam, 1996; Yang and Bell, 1997; Yang and Huang, 1998).

III: Traveler Information Systems

Traveler information systems are considered as a promising technology to improve traffic conditions by helping travelers making more efficient travel choices. Unlike traffic management and control, the effectiveness of traveler information systems is determined by the quality of information disseminated to travelers, how travelers perceive the use of the information, and how they respond to such information (Adler and Blue, 1998). In the real world, information is rarely perfect. Its quality depends on many factors such as the surveillance system, the ability of the system to filter and project traffic information, the timeliness of its dissemination, and most importantly the interaction between information and route choice decisions. Many previous studies (e.g., van Vuren and Watling, 1991; Yang, 1998; Lo *et al*., 1999) assumed that the interaction between predicted information and route choice decisions is consistent with the Wardropian traffic flow patterns. This is a rather strong assumption because it implies that predicted information will be correct (or perfect) in the future and that travelers' route choice decisions comply with the *user equilibrium* principle. As illustrated in loop III of Figure 2.1, this circular relationship between predicted information and travelers' route choice decisions is difficult to sustain in practice given that predicted information is rarely perfect and travelers do not necessarily respond to information according to the *UE* principle (Emmerink, 1997).

From Figure 2.1, it can be clearly observed that route choice models are the core module for traffic assignment, traveler information system, and traffic control/management. Currently, the widely accepted route choice model (i.e., *user equilibrium* model) is based on strong assumptions that the network travel times are deterministic for a given flow pattern. All travelers are perfectly aware of the travel times on the network and always capable of identifying the minimum travel time route (Sheffi, 1985; Patriksson, 1994; Bell and Iida, 1997). To overcome the deficiencies of the *user equilibrium* model, some researchers have proposed random utility models (i.e., *stochastic user equilibrium* models) to relax the assumption of perfect knowledge of network travel times, thus allowing travelers to select routes based on their perceived travel times (Dagazon and Sheffi, 1977; Fisk, 1980). Due to variations in travelers' perceptions of travel times, travelers do not always choose the correct minimum travel time route. However, the assumption of network uncertainty remains unresolved.

Besides lacking the inclusion of network uncertainty into the route choice models, the criteria currently used (e.g. travel time, distance, or weighted travel time and distance) may not adequately reflect the behavior of travelers in congested networks, and more importantly their response to traffic management and information strategies (Dial, 1996, 1997; Blue *et al*., 1997, Park, 1998; Chen *et al*., 2000). Given that there is an increasing use of innovative information and communication technologies to manage traffic, it is important to understand how route choice decisions are made.

2.1   Route Choice Models

The objective of a route choice model, in the traffic assignment problem, is to describe how individual route choice decisions, represented by an origin-destination trip table interacts with the transportation network represented by a set of link performance functions so that the resulting network loads are accurately predicted (Sheffi, 1985). In the available literature, several route choice models have been proposed which differ in:

(1) Characterization of network travel times (i.e., deterministic or stochastic).

(2) Traveler's knowledge of network travel times (i.e., with or without perception error).

(3) Route choice behavior

     a. Criterion/criteria used in route choice decision process.

     b. Route cost structure (i.e., additive or nonadditive).

     c. Route choice preference (e.g., risk averse or risk prone).

Recently, Chen and Recker (2001) provided a classification scheme of the route choice models under the presence of congestion, using network uncertainty and perception error as follows:

Table 2.1: Classification of Route Choice Models

		Perception Error?	
		No	Yes
Network	No	DN-DUE	DN-SUE
Uncertainty?	Yes	SN-DUE	SN-SUE

Where

> DN = Deterministic Network
>
> SN = Stochastic Network
>
> DUE = Deterministic User Equilibrium
>
> SUE = Stochastic User Equilibrium

The above classification scheme results in four distinct route choice models. In each model, the following common assumptions are made:

- The travel time for every link on the network is assumed to be an increasing function of the flow of vehicles on the link.
- Each traveler makes a rational route choice decision based on minimizing some criteria related to average travel times or some disutility measure based on average travel times and their variances as follows:

$$U_k = V_k + \varepsilon_k,$$ (2.1)

where V_k is the systematic component of the disutility of route k; ε_k is the random error term of route k; and U_k is the total disutility of route k.

For comparison and discussion purposes, we define the following notation for the route choice models discussed in this chapter.

x_a	Traffic flow of link a
$\widetilde{C}_a$	Perceived travel cost of link a
c_a	Measured travel cost of link a
ε_a	Perception error of link a
$\widetilde{C}_k$	Perceived travel cost of path k
c_k^{rs}	Measured travel cost of path k between OD pair rs
ε_k^{rs}	Perception error of path k between OD pair rs
V_k	Systematic component of the utility of route k
ε_k	Random error term of route k
U_k	Total utility of route k
$\bar{t}_k$	Average travel time of route k
f_k^{rs}	Traffic flow of path k between OD pair rs
U_k^{rs}	Expected disutility of path k among OD pair rs
$f_k^{rs^*}$	Equilibrium traffic flow of path k among OD pair rs
f	Vector of path flow f_k^{rs}
$C_k^{rs}(\mathbf{f})$	Expected utility for traveling of route k between r and s
$DU(\widetilde{C}_k^{rs})$	Disutility value of path k between OD pair rs for given $\widetilde{C}_k^{rs}$
$\widetilde{C}_k^{rs}$	Perceived travel cost of path k between OD pair rs
P_k^{rs}	Probability of path k is taken between OD pair rs
δ_{ak}^{rs}	Path link index matrix, δ_{ak}^{rs} is equal to 1 for link a on path k and 0 otherwise
$E[\eta(T_k)]$	Expected utility of route k
$\eta(t_k)$	The utility function describing the risk-taking preference of travelers on route k
$f_{T_k}(t_k)$	The travel time probability density function (pdf) of route k

RS	OD pair Set
χ^{rs}	Path set of OD pair rs
q^{rs}	Traffic demand between OD pair rs
u_{rs}	Minimum travel cost between OD pair rs
$S_{rs}[\widetilde{C}^{rs}(X)]$	Satisfactory function of perceived path travel cost
$\widetilde{C}^{rs}(X)$	Vector of perceived path travel cost for given OD pair rs
$F^{f}(X)$	Column vector of $(U_{k}^{rs} - \pi^{rs}, \quad \forall rs \in RS, \quad \forall k \in \chi^{rs})$
$F^{\pi}(X)$	Column vector of $(\sum_{k \in \chi^{rs}} f_{k}^{rs} - q^{rs}, \quad \forall rs \in RS)$
π	Vector of the minimum expected disutility π^{rs}
η^{rs}	Expected disutility between OD pair rs
F	Network mapping function that map traffic demand to path flow
π^{rs}	The expected utility between OD pair r and s
θ	A positive parameter

2.2   The DN-DUE Model

The DN-DUE model ignores both network uncertainty and perception error. Essentially, this model assumes that travelers consider only the expected values of network travel times and they are perfectly aware of these expected travel times on the network. According to the utility function specified above, this means

$$V_{k} = -\theta\, E[T_{k}] = -\theta\, \bar{t}_{k} \qquad \text{and} \qquad \varepsilon_{k} = 0, \tag{2.2}$$

where $\bar{t}_{k}$ is the average travel time of route k, and θ is a positive parameter.

The criterion used for route choice is to minimize the expected value of route travel time, which is obtained by adding up the average travel times on all the links belonging to the route. The choices of routes made by all travelers result in a network flow allocation such that *all used routes between every origin-destination pair have equal average travel times and no unused*

route has a lower average travel time according to Wardrop's First Principle (1952). Since travel time variability is not considered in the route choice decision, all travelers in the DN-DUE model are implicitly assumed to be risk neutral. Depending on the route cost structure, formulations of the DN-DUE model include these four approaches: (i) mathematical programming (e.g., Beckmann et al., 1956; Sheffi, 1985), (ii) nonlinear complementarity problem (Aashtiani, 1979), (iii) variational inequality (e.g., Dafermos, 1980; Nagurney, 1999), and (iv) fixed point (Asmuth, 1978). All four approaches can be used to formulate the additive traffic equilibrium problem (i.e., route cost structure is simply the sum of the link costs on that route). The additive assumption allows one to express route cost in terms of the sum of link costs, and also enables the problem to be solved without the need to store routes despite route flow variables remain in the constraint set (Gabriel and Bernstein, 1997; Lo and Chen, 2000). The Frank-Wolfe algorithm is perhaps the best known algorithm that takes advantage of this additive assumption when solving large-scale transportation networks. Although routes are generated in each iteration of the column generation procedure, the algorithm does not store or make direct use of the routes. Instead, it operates directly in the link-flow space using only two vectors of link flows to perform the convex combinations step to update its solution. Solution approaches, that use link-based variables, are sometimes known as link-based algorithms. Link-based solution algorithms for the additive traffic equilibrium problem include the enhanced Frank-Wolfe (FW) algorithm (LeBlanc *et al*., 1975; Fukushima, 1984; Weintraub *et al*., 1985; Janson and Gorostiza, 1987; Lee and Nie, 2001), PARTAN algorithm (LeBlanc *et al*., 1985; Florian et al., 1987; Arezki and Van Vliet, 1990), restricted simplicial decomposition (RSD) algorithm (Hearn *et al*., 1985), and origin-based algorithm (Bar-Gera and Boyce, 2002). There are also the route-based algorithms that solve the same problem explicitly using route-flow variables which require storing the links of each individual route. Solutions resulting from a route-based algorithm provide both the aggregate link-flow solutions and the individual route-flow solutions that are not readily available from a link-based algorithm. Route-based solution algorithms for the additive traffic equilibrium problem include the OD-based Frank-Wolfe algorithm (Chen, 2001; Chen *et al*., 2002), disaggregate simplicial decomposition (DSD) algorithm (Larsson and Patriksson, 1992), and gradient projection (GP) algorithm (Bertsekas and Gafni, 1982; Jayakrishnan *et al*., 1994), and conjugate gradient projection algorithm (Lee *et al*., 2003). For a comprehensive review and

computational study of the route and link-based solution algorithms for the additive traffic equilibrium problem, refer to Chen *et al.* (2000, 2002) and Lee *et al.* (2002).

When the route cost structure is non-additive (i.e., route cost structure is not a simple sum of the link costs on that route), it is no longer feasible to solve the problem with just link-flow variables since there is no simple way of converting the non-additive route cost to equivalent link costs. Non-additive traffic equilibrium problems must be formulated and solved explicitly in the route-flow space. As adeptly discussed by Gabriel and Bernstein (1997), there are many situations in which the additive route cost structure is inadequate for addressing factors affecting a variety of transportation policies. Some of the examples include: (a) path-specific tolls and fares – most existing fares and tolls in the United States are not directly proportional to travel time or distance, (b) nonlinear valuation of travel time – small amounts of time are valued much less than larger amounts of time, and (c) emissions fees – emissions of hydrocarbons and carbon monoxides are a nonlinear function of travel times. Compared to the additive traffic equilibrium problem, there exist only a few solution algorithms for the non-additive traffic equilibrium problem. These include the nonsmooth equations/sequential quadratic programming (NE/SQP) method (Bernstein and Gabriel, 1997), gradient projection algorithm (Scott and Bernstein, 1998), gradient method with Armijo stepsize for solving Fischer's gap function (Lo and Chen, 2000), and self-adaptive projection and contraction method (Chen *et al.*, 2000, 2001).

For a more detailed overview of route choice models, solution algorithms, and applications of the DN-DUE model, refer to Bell and Iida (1997), Cascetta (2001), Patriksson (1994), and Sheffi (1985).

2.3   The DN-SUE Model

Due to the unrealistic assumption that all travelers have perfect knowledge of the network conditions, Daganzo and Sheffi (1977) extended the Wardrop's UE condition by introducing a perception error into the route choice process as follows:

$$V_k = -\theta\, E[T_k] = -\theta\, \bar{t}_k \qquad \text{and} \qquad \varepsilon_k \neq 0\,, \qquad\qquad (2.3)$$

In this model, each traveler is assumed to have some perceptions of the mean travel times on each link of the network, which include a random error term. Each traveler's route choice criterion is to minimize the *perceived* value of the route travel time, which can be obtained by adding up the *perceived* travel times on all the links belonging to the route. The choices of routes by all travelers result in a network flow allocation such that *no traveler can reduce his/her perceived travel time by unilaterally changing to another route* (Sheffi, 1985). This definition is an extension of the UE model, known as the *stochastic user equilibrium* model. Similar to the DN-DUE model, all travelers in DN-SUE model are risk neutral since only the mean travel times are considered in the route choice decision process.

Due to variations in travelers' perceptions of travel times, travelers do not always select the correct minimum travel time route. Route choice models proposed under this approach can have different specifications according to modeling assumptions of the random error term. The two most commonly used random error terms are Gumbel (Dial, 1971) and normal (Daganzo and Sheffi, 1977) distributions, which result in the logit and probit-based route choice models. Logit-based route choice models have a closed-form probability expression, an equivalent mathematical programming formulation (Fisk, 1980), and can be solved using both path enumeration techniques (Ben-Akiva et al., 1984; Cascetta *et al*., 1997, 2002) and column generation techniques (Bell *et al.*, 1993; Bell, 1994; Chen and Alfa, 1991; Damberg *et al*., 1996; Leurent, 1997; Maher, 1998). The drawbacks of the logit model are: (1) the inability to account for path overlapping (or correlation) among routes and (2) the inability to account for perception variance with respect to trips of different lengths. These two drawbacks stem from the logit's underlying assumptions that the random error terms are independently and identically distributed (IID) with the same, fixed variances (Sheffi, 1985). Probit-based route choice models, on the other hand, do not have such drawbacks since it handles path overlapping and identical variance problems by allowing covariance between the random error terms for pairs of routes. However, the probit model does not have a closed-form solution and it is computationally burdensome when the choice set contains more than a handful of routes. Due to the lack of a closed-form probability expression, solving the probit-based route choice model requires either Monte Carlo simulation (Sheffi and Powell, 1982) or Clark's approximation method (Maher and Hughes, 1997). Other specifications of the random error term include uniform distribution (Burrell,

1968), gamma distribution (Bovy, 1990), and lognormal distribution (Von Falkenshausen, 1976; Cantarella and Binetti, 1998).

Recently, there are renewed interests to improve the logit-based route choice model due to the importance of route choice models in Intelligent Transportation Systems (ITS) applications, particularly applications to advanced traveler information systems (ATIS). Several modifications or generalizations of the logit structure have been proposed to relax the IID assumptions. These extended logit models include the C-logit (Cascetta *et al*., 1996), path-size logit (Ben-Avika and Bierlaire, 1999; Ramming, 2002), cross-nested logit (Prashker and Bekhor, 1998; Vovsha, and Bekhor, 1998), paired combinatorial logit (Bekhor and Prashker, 1999; Gliebe et al., 1999; Prashker and Bekhor, 1998, 2000), generalized nested logit (Bekhor and Prashker, 2001), and logit kernel (Bekhor *et al*., 2002). Despite the recent advancements in the logit model and its adaptations to the route choice problem, all of the above models do not address the issues of travel time variability and choice behavior under uncertainty.

2.4   The SN-DUE Model

The two route choice models presented above assume that the network travel times are deterministic for a given flow pattern. In reality, path (link) travel times comply probability distributions for a given flow pattern and describes the variations of travel times on the network. Such variations could result from the differences in the mix of vehicle types on the network for the same flow rates, differences in driver reactions under various weather and driving conditions, differences in delays experienced by different vehicles at intersections, etc. Variability in travel times introduces uncertainty for travelers, such that they do not know with certainty when they will arrive at their destination. Thus, it is considered as a risk (or an added cost) to a traveler making a trip. Route choice decisions under network uncertainty often involve tradeoffs between expected travel time and travel time variability. This observation is supported by recent empirical studies (Abdel-Aty *et al*., 1996; Ghosh, 2001; Kazimi *et al*., 2000; Lam, 2000; Lam and Small, 2001; and Small *et al*., 1999) that found travelers are interested in not only travel time saving but also reduction of travel time variability. Abdel-Aty *et al*. (1996) found that travel time variability is one of the most important factors in route choice decisions. Specifically, about 54% of the

respondents in the survey indicated that travel time variability is either the most important or second most important reason for choosing their daily commute routes. Small *et al.* (1999) found that both passengers and freight carriers are strongly averse to the scheduling mismatches because they cannot predict precisely what their travel time will be. For this reason, they will pay a premium to avoid congestion and to achieve greater reliability in travel times. From the two value-pricing projects in Southern California in the United States, Ghosh (2001), Kazimi *et al.* (2000), Lam (2000), and Lam and Small (2001) consistently found that travelers are willing to pay a substantial amount to reduce variability in travel time. Suffice to say, travel time variability is an important factor for travelers when making their route choice decisions under risk or circumstances where they do not know, with certainty, the outcome of their decisions.

In the SN-DUE model, network uncertainty is explicitly considered, but perception error is ignored. Instead, link travel times are explicitly treated as random variables. For a given set of flows, there is a probability density function (pdf) associated with the route travel times. Because the travel time variability is included in this model, different travelers may respond to such variations differently depending on their risk-taking preferences. The risk in this case is the variability associated with network travel times.

$$V_k = E[\eta(T_k)] = -\theta \int \eta(t_k) f_{T_k}(t_k) dt_k \qquad \text{and} \qquad \varepsilon_k = 0, \tag{2.4}$$

where $E[\eta(T_k)]$ is the expected utility of route k; $\eta(t_k)$ is the utility function describing the risk-taking preference of the traveler on route k; and $f_{T_k}(t_k)$ is the pdf of route k.

To the best of our best knowledge, Soroush (1984) might be the first to give an analytical formulation for the SN-DUE model as a nonlinear complementarity problem (NCP) formulation. Let $F : R^n \mapsto R^n$ (i.e., F is a mapping to itself). The NCP is a system of equations and inequalities stated as:

Find x such that

$$x \geq 0, \qquad F(x) \geq 0, \qquad x^T \cdot F(x) = 0, \qquad\qquad (2.5)$$

where

$$x = \begin{pmatrix} f_k^{rs}, & \forall\, k,r,s \\ \pi^{rs}, & \forall\, r,s \end{pmatrix} \text{ and } F(x) = \begin{pmatrix} C_k^{rs}(\mathbf{f}) - \pi^{rs}, & \forall\, k,r,s \\ \sum\limits_k f_k^{rs} - q^{rs}, & \forall\, r,s \end{pmatrix}, \qquad (2.6)$$

$C_k^{rs}(\mathbf{f}) = E[\eta(T_k(\mathbf{f}))]$ is the expected utility for traveling on route k between r and s,

f_k^{rs} is the flow on route k between r and s,

$\mathbf{f}$ is a vector of route flows,

π^{rs} is the expected utility between r and s,

q^{rs} is the travel demand between r and s.

The above NCP is for a fixed travel demand case, but it is easily relaxed to become elastic by changing the travel demands as a function of the expected utilities (see Nagurney, 1999). Since travel time variability is explicitly considered in the SN-DUE model, different risk-taking preferences can be used to model how travelers respond to such variations in network travel times. Depending on the behavioral nature of travelers, they can be classified as risk averse, risk prone, or risk neutral. For instance, a risk averse traveler will trade off a reduction in travel time variability with some increases in expected travel time, whereas a risk prone traveler may choose a route with a greater variability so as to increase the possibility of a smaller travel time. A risk neutral traveler would choose a route based on only expected travel time without consideration of its variability. It should be noted that the risk neutral model in the SN-DUE model is essentially the same as the DN-DUE model in which each traveler makes his/her route choice based only on the mean travel times.

In the SN-DUE model, travelers are assumed to have perfect knowledge of the variable nature of network travel times. In general, this model may be suitable for peak-hour traffic where regular commuters have a good idea of the mean and variance of network travel times. Recently, Yin

and Ieda (2001) applied this NCP formulation with quadratic disutility functions to assess the performance reliability of road networks. A gap function was used to first convert the NCP into an unconstrained optimization problem, and then applied a route-based algorithm, proposed by Lo and Chen (2000), to solve this smooth, unconstrained reformulation. Uchida and Iida (1993) used the notion of effective travel time (i.e., mean travel time + safety margin) to model network uncertainty in the route choice model. The safety margin is defined as a function of travel time variability that serves as a measure of risk averseness in their risk-based, route choice model. An effective travel time model requires the specification of travel time variability as an input depicting the stochasticity of the network, which may be difficult to obtain in the field. Because of this, Bell and Cassir (2002) provided an alternative approach based on game theory, that circumvents the need to specify, *a priori*, the travel time distribution. In this game, the network users seek the best routes, subject to link failure probabilities, that are selected by the demons trying to cause a maximum damage to network users. However, it should be recognized that there are a number of restrictive assumptions underlying this game theory formulation of risk averse route choice model. Recently, Liu *et al*. (2004) proposed a Mixed Logit model to formulate the SN-DUE case and the newly collected loop data from SR 91 in California is used to calibrate the unknown coefficient in the model.

2.5 The SN-SUE Model

In the SN-SUE model, both the variability of network travel times as well as traveler perception errors are accounted for as follows:

$$V_k = E[\eta(T_k)] = -\theta \int \eta(t_k) f_{T_k}(t_k) dt_k \qquad \text{and} \qquad \varepsilon_k \neq 0. \qquad (2.7)$$

Mirchandani and Soroush (1987) were the first to propose this generalized traffic equilibrium model that incorporates both probabilistic travel times and variable perceptions in the route choice decision process. In the SN-SUE model, each traveler i is assumed to have a variable perception error $\xi_i \sim N(\mu_i, \theta_i)$ with $\mu_i \sim N(0, \tau)$ and $\theta_i \sim G(\alpha, \beta)$, respectively. In other words, the perception error for each traveler i is normally distributed with a variable mean μ_i

and a variable variance θ_i, where μ_i is assumed to be normally distributed over the population of travelers with zero mean and τ as the variance, and θ_i is assumed to be gamma distributed over the population of travelers with parameters α and β. This variable perception error allows each individual traveler to experience a different travel time for a given set of flows. This is different from the probit-based DN-SUE model in which the random error term only accounts for the randomness of the travelers' perceived travel times (i.e., $\xi_i \sim N(0,\tau)$) and treats the randomness of link travel times in the form of expected values.

Following Mirchandani and Soroush (1987), the perception error for traveler i for a link of unit travel time is assumed to be $\xi_{ia}\big|_{t_a=1} \sim N(\mu_i, \theta_i)$. Then, the conditional perception error for a link with some travel time is $\xi_{ia}\big|_{t_a=t_a} \sim N(\mu_i t_a, \theta_i t_a)$, and the conditional moment generating function[1] (MGF) for the distribution of perception error is

$$M_{\xi_{ia}|t_a=t_a}(s) = E\left[\exp\left(\xi_{ia|t_a=t_a}\right)\right] = \exp\left[t_a s\left(\mu_i + \frac{1}{2}\theta_i s\right)\right], \tag{2.8}$$

where s is a real number. Hence, the MGF of the perceived travel time for traveler i for link a can be written as

$$M_{\tilde{t}_{ia}}(s) = E\left[\exp\left(s\tilde{t}_{ia}\right)\right]. \tag{2.9}$$

Using the definition that the perceived travel time is the sum of the expected link travel time and the individual perception error, the MGF of the perceived travel time for traveler i for link a becomes

---

[1] Moment generating function is a convenient way to derive expressions for the means and standard deviations of functions of random variables (Ang and Tang, 1990).

$$
\begin{aligned}
M_{\tilde{t}_{ia}}(s) &= E\left[\exp\left(s\left(t_a + \xi_{ia}\right)\right)\right] \\
&= E\left[\exp\left(st_a\right)E\left(\exp\left(\xi_{ia|t_a=t_a}\right)\right)\right]. \\
&= E\left[\exp\left(st_a\right)M_{\xi_{ia|t_a=t_a}}(s)\right]
\end{aligned}
\tag{2.10}
$$

Substituting $M_{\xi_{ia|t_a=t_a}}(s)$ into equation (2.10), we obtain

$$
\begin{aligned}
M_{\tilde{t}_{ia}}(s) &= E\left[\exp\left(st_a\left(1+\mu_i+\frac{1}{2}\theta_i s\right)\right)\right] \\
&= M_{t_a}\left[s\left(1+\mu_i+\frac{1}{2}\theta_i s\right)\right]
\end{aligned}
\tag{2.11}
$$

Thus, the MGF of the perceived travel time is specified in terms of the MGF of the actual travel time and the variable parameters of the traveler i perception error distribution. Using the perceived travel time derived from the MGF and the expected disutility for traveling route k between r and s from the systematic component of the utility function, the perceived expected utility can be expressed as follows:

$$
E\left[\tilde{\eta}\left(\tilde{T}_k\right)\right] = -\theta \int \tilde{\eta}\left(\tilde{t}_k\right) f_{\tilde{T}_k}\left(\tilde{t}_k\right) d\tilde{t}_k,
\tag{2.12}
$$

where $E\left[\tilde{\eta}\left(\tilde{T}_k\right)\right]$ is the perceived expected utility of route k; $\tilde{\eta}\left(\tilde{t}_k\right)$ is the perceived utility function describing the risk-taking preference of the traveler on route k; and $f_{\tilde{T}_k}\left(\tilde{t}_k\right)$ is the perceived pdf of route k. The values of both $E\left[\tilde{\eta}\left(\tilde{T}_k\right)\right]$ and $\tilde{T}_k$ are unobserved and will be different from the values of $E\left[\eta\left(T_k\right)\right]$ and T_k. Because $E\left[\tilde{\eta}\left(\tilde{T}_k\right)\right]$ is a function of a variable perception error, where both μ_i and θ_i are randomly distributed across the population of travelers, the perceived expected utility will differ from traveler to traveler. Hence the route choice probability can be stated as follows:

$$
P_k^{rs} = \operatorname{Prob}\left(E\left[\tilde{\eta}\left(\tilde{T}_k^{rs}\right)\right] \geq E\left[\tilde{\eta}\left(\tilde{T}_l^{rs}\right)\right], \quad \forall l \neq k \in K_{rs}\right) \quad \forall k,r,s,
\tag{2.13}
$$

where P_k^{rs} is the probability of choosing route k between r and s, $\widetilde{T}_k$ and $\widetilde{T}_l$ are the perceived travel time of route k and route l between r and s; and K_{rs} is the set of routes between r and s.

Given the travel demands between all origin-destination (OD) pairs (q_{rs}) and the route choice probabilities (P_k^{rs}), the route flow assignment is simply

$$f_k^{rs} = P_k^{rs} q_{rs}, \quad \forall\, k,r,s. \tag{2.14}$$

The link flow on each link can then be calculated as

$$x_a = \sum_r \sum_s \sum_k f_k^{rs}\, \delta_{ka}^{rs}, \quad \forall\, a, \tag{2.15}$$

where δ_{ka}^{rs} is one if link a is on route k between r and s, and zero otherwise.

Following the work of Daganzo (1983) and Cantarella (1997), the SN-SUE model can be formulated as a fixed-point problem.

$$\mathbf{f}^* = \mathbf{P}\!\left(E\!\left[\tilde{\eta}\!\left(\widetilde{T}(\mathbf{f}^*) \right) \right] \right) \mathbf{q}, \tag{2.16}$$

where $\mathbf{f}^*$ is a vector of equilibrium path flows; $\mathbf{P}$ is matrix of route choice probabilities; and $\mathbf{q}$ is a vector of OD travel demands. Existence of solutions of the fixed-point formulation can be analyzed through Brouwer's theorem (see pages 653-658 in Cascetta (2001)) for the theoretical properties and solution algorithms for the fixed point formulation). Since the feasible set is non-empty (for any given positive travel demands with at least one route serving between each origin-destination pair), compact, and convex, and the expected utility function is continuous, the fixed-point formulation will admit at least one solution.

Similar to the SN-DUE model, travelers in the SN-SUE model can be either risk averse, risk prone, or risk neutral based on the assumptions about the behavioral preference of the travelers.

Risk neutral travelers in the SN-SUE model are the same as the DN-SUE model since all travelers make their route choice decisions solely based on the mean network travel times. A critical difference is that the SN-SUE travelers use the *perceived* expected utility (as opposed to expected utility in the SN-DUE model) as the route choice criteria. Hence, *equilibrium is achieved when no traveler can reduce his/her perceived expected disutility by unilaterally changing to another route*.

Recently, there has been an increasing interest in this class of route choice models. Tatineni *et al.* (1997) conducted experiments on a large-scale network in the Chicago Area to compare the traffic loadings of the DN-DUE, DN-SUE, and SN-SUE models; Boyce *et al.* (1999) and Liu *et al.* (2002) extended the SN-SUE model to a dynamic setting; and Chen *et al.* (2000), Chen and Recker (2001), and Chen *et al.* (2002) applied all four route choice models to evaluate network capacity reliability and travel time reliability of road networks.

Chapter Three

3. GIS DATABASE DEVELOPMENT AND CALCULATION OF FREEWAY TRAVEL TIME VARIABILITY

3.1 Introduction

Travel time variability has recently become an important measure in transportation system management and has begun being treated as a major factor influencing travel decisions. Despite its increasing importance, the travel time variability has not been widely quantified in practice, let alone used in real applications. In this chapter, a real case study is conducted to describe how the travel time variability is measured in the existing freeway system of Orange County, California.

Bates *et al*. (1987) and Small *et al*. (1999) have described three distinct factors on travel time variability. These factors include inter-day variability, inter-period variability, and inter-vehicle variability. Sources of such seasonal or day-to-day variation are regarded as results from demand fluctuations, accidents, road construction, and weather conditions. Personal driving styles and the behavior of traffic control devices along a route are reasons for causing travel variability. Abdel-Aty *et al*. (1996) have found that travel time variability is the most important factor for choosing the route from a survey. Perceiving that changes in departure times are also consequence of changes in congestion, Small *et al*. (1999) have attempted to fit an econometric model that treats scheduling considerations using preference survey data, and calibrated the value of reliability in terms of reduction in travel time variation. Lam and Small (2001) have also evaluated the value of travel time and reliability using the difference between the 90 percentile and median of travel time. Although there have been empirical and theoretical findings addressing the importance of travel time variability, less attention has been paid to estimating or calculating the travel time reliability. Recently there have been some efforts on quantifying such travel time variability

(Levinson and Zhang, 2001; Richardson, 2003; Lomax 2003). Travel time variability is treated as a way of measuring travel reliability (Lomax, 2003; Chen *et al*., 2003).

This case study incorporates existing freeway traffic surveillance systems to provide systematic measures of travel time variability. While real-time traffic data are a useful source for traffic operations, their long-term analysis provides precious information in identifying non-recurrent traffic conditions from temporal and spatial traffic variations. The main purpose of this case study is to develop a database to help understand traffic variability of point measures and route-level measures by incorporating existing traffic surveillance systems. This case study develops an integrated GIS database for traffic analysis in Orange County, California, and reports temporal and spatial traffic variability as measures of travel reliability.

3.2   Building Database

3.2.1   Overview

The main source of data for calculation of travel time variability is the Testbed's direct link to Caltrans District 12 TMC that provides real-time data links from area freeways directly to dedicated Testbed research laboratories located at University of California, Irvine (UCI). Currently, both historical freeway data (all Caltrans District 12 systems) and real-time freeway data are available in the Testbed lab PTL network via a CORBA interface. In this study, travel time variability is analyzed during a year period from March 2001 to February 2002, and the original lane-by-lane thirty-second data are aggregated into five-minute interval by loop detector stations.

Figure 3.1 shows the overall procedure of the travel time variability computation using the existing loop detector data from the UCI ATMS Testbed. This procedure consists of three major parts.

- Part 1: Analysis of Point/Section Measures
- Part 2: Analysis of Route-based Measures
- Part 3: Development of GIS (Geographical Information Systems) Database

23

The first part is to analyze traffic measures based on the existing loop detector stations. After calculating speeds at each point, section travel times are calculated for each section defined by the location of loop detectors. The focus of this part is to investigate travel time variability at the section level.

The second part is to calculate route-based travel time between freeway ramps. This part calculates route travel times for all feasible routes enumerated based on section travel times on the corresponding routes. An interest is to investigate how the variability of section travel times affects the route travel time. From drivers' points of view, route travel time and its variability are of more concern than those in sections.

The third part is to incorporate the section and route travel time measures into a GIS database to visualize and manage travel time and variability measures. Although the GIS database is used to visualize all information at the initial stage, the tasks in the first and second part are conducted within the GIS database.
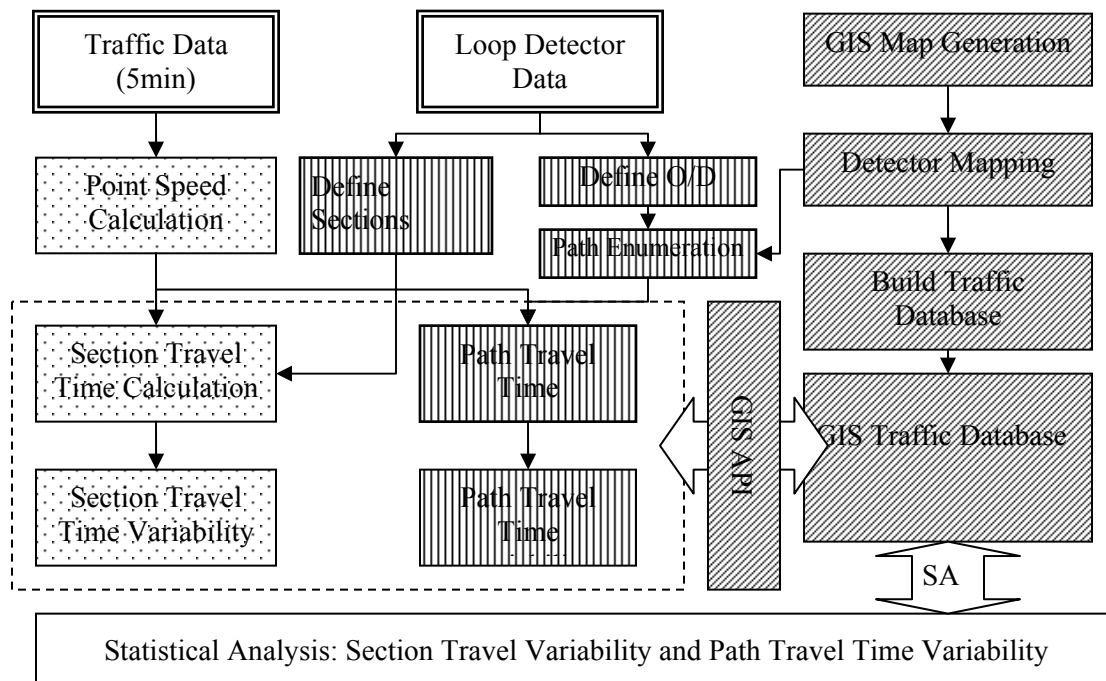
Figure 3.1: Overall Procedure for Travel Time Variability

24

Tasks in this chapter can be broken down into six parts:

Task 1: Data aggregation into five-minute interval

Task 2: Point speed calculation

Task 3: Define freeway sections, origins, destinations, and enumerate feasible paths

Task 4: Calculation of section travel times and variability

Task 5: Calculation of route travel times and variability

Task 6: GIS database construction

Task 7: Data analysis

The core part of this study is to analyze various travel time variability measures. Standard deviations are defined as measures of travel time variability. More specifically the normalized standard deviations are evaluated as representative measures of section travel times and route travel times. Travel time variability can be caused either by travel demand changes and/or by traffic incidents. In this study, travel time variability is viewed from three different angles.

- Day-to-day variability
- Within-day variability
- Spatial variability

The first variability is the day-to-day variation that makes difficult the prediction of travel time by travelers or system operators. The reliable transportation infrastructure will provide smaller day-to-day variation. The second variability is within-day variation that displays how much the traffic pattern fluctuates during a day or a given period of day. As a similar variability as a route level, spatial variation is defined as variability that drivers are experiencing while driving along a route.

3.2.2 Traffic Data and Data Availability

As described above, this study uses one-year historical traffic data from March 2001 to February 2002. Freeway traffic data for all of Orange County are available from the UCI Testbed research laboratory. The original data are available in the following format:

- vds id and timestamp
- lane and loop count
- volume, occupancy, and status

The current UCI Testbed provides traffic counts and occupancy for each lane at each detector station, every thirty seconds. In this study, lane-by-lane traffic data were aggregated into each detector station with five-minute intervals, in order to obtain valid traffic data from each point. A simple method was applied in aggregating traffic measures. All lane-by-lane data available during a five-minute interval is used, and the average was taken as a representative measure. Figure 3.2 shows loop detector stations in the study area.
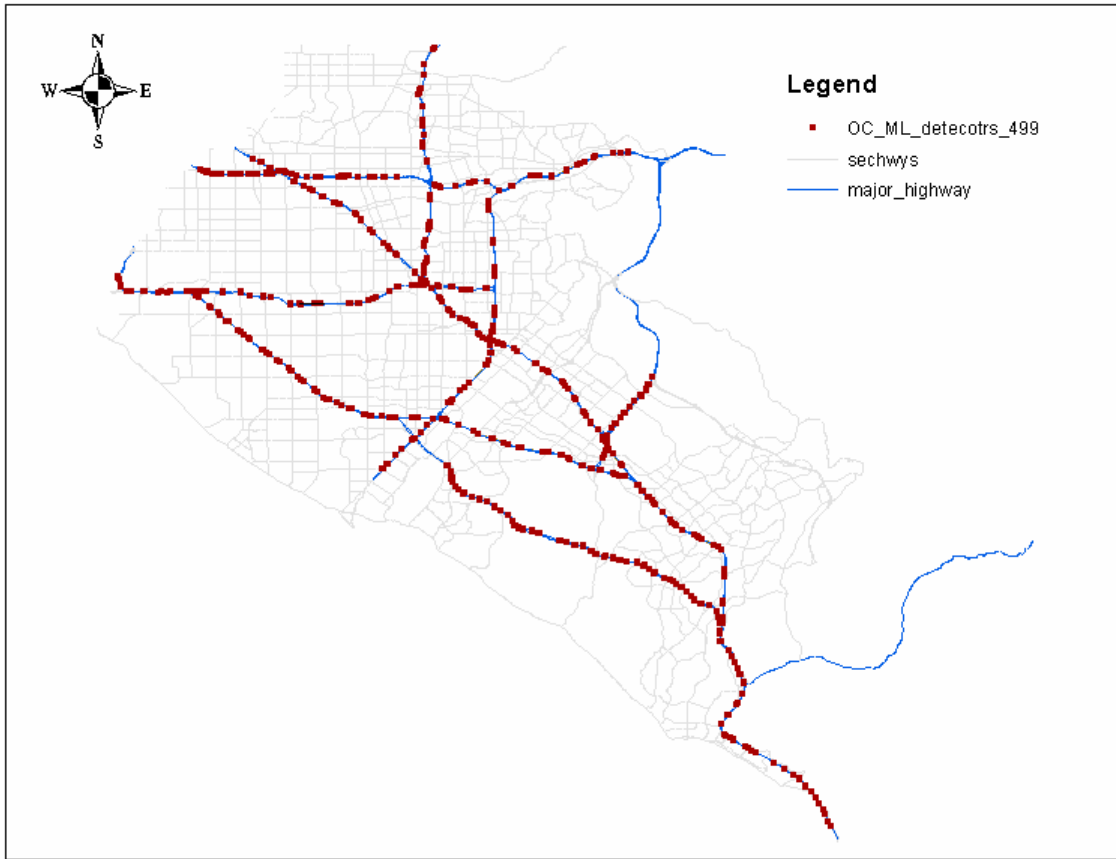
Figure 3.2: Loop Detector Stations in Study Area

Loop detectors in Orange County are classified into mainlines, on-ramps, off-ramps, freeway-to-freeway connectors, etc., and each detector station includes lane-by-lane traffic data. Since our main interests are investigating travel time on mainline freeways, this study only considered traffic data from the main-line loop detector stations. In the study area, there are 499 mainline loop detector stations. However, many loop detector data were missing during the study period. Table 3.1 shows the number of loop detector stations by available data rate. Due to freeway construction and communication problems, a quarter of detectors were not functioning, and only a 61.3% of loop stations were able to provide more than 50% of traffic data.

Table 3.1: Loop Detector Data Availability

% available data	Number of detector stations	%	Cumulative number of detector stations	%
95 - 100%	0	0.0%	0	0.0%
90 - 95%	0	0.0%	0	0.0%
85 - 90%	1	0.2%	1	0.2%
80 - 85%	18	3.6%	19	3.8%
75 - 80%	76	15.2%	95	19.0%
70 - 75%	68	13.6%	163	32.7%
65 - 70%	61	12.2%	224	44.9%
60 - 65%	41	8.2%	265	53.1%
55 - 60%	24	4.8%	289	57.9%
50 - 55%	17	3.4%	306	61.3%
45 - 50%	8	1.6%	314	62.9%
40 - 45%	9	1.8%	323	64.7%
35 - 40%	2	0.4%	325	65.1%
30 - 35%	1	0.2%	326	65.3%
25 - 30%	3	0.6%	329	65.9%
20 - 25%	9	1.8%	338	67.7%
15 - 20%	3	0.6%	341	68.3%
10 - 15%	7	1.4%	348	69.7%
5 - 10%	4	0.8%	352	70.5%
0 - 5%	23	4.6%	375	75.2%
0%	124	24.8%	499	100.0%

3.2.3   Reliability of Traffic Data Estimates by Data Availability

Since the aggregated, five-minute traffic data are based on the available samples during the five-minute interval, the quality of traffic data depends upon the data availability or the sample rate. In order to determine which loop detectors to include in the database, with respect to the data availability, we investigated potential errors by the sample rate using the Monte Carlo simulation technique. The experiment was conducted with the data on I-405, where reliable data were provided during the period. In this experiment, we randomly chose a given percentage of samples and estimated the five-minute aggregate measures. The averaged aggregated measures from the sample were compared with the actual average value (true mean) with a sample of 100%. After repeating the experiment for 30 times, we compared MAPE (Mean Absolute Percentage Error) as in equation below, as an overall performance measure. In this experiment, we used three traffic measures, volume, occupancy, and speed. However, speeds used in this

experiment were estimated from volume and occupancy assuming a g-factor because speeds were not available in single loop detectors as in Orange County.

$$MAPE = \frac{1}{N}\left(\sum_{n=1}^{N}\frac{|x_n - \bar{x}|}{\bar{x}}\right) \times 100$$

(3.1)

where

N = the total number of samples,

x_n = nth sample,

$\bar{x}$ = true mean traffic measure.

Figure 3.3 and Figure 3.4 show two sample cases of simulation experiment results. The first figure shows changes of measures with various sample rates during the experiment period (4:00 – 10:00 AM). The estimation errors decrease when approaching the peak time periods and may have been due to an increase in samples during the peak hour. In order to see the impact of congestion level on the estimates, we also compared the estimation errors (MAPE) with respect to the level of occupancy, as shown in the second figure. In most cases, the estimation errors were lowest when the traffic occupancy levels were within 10 – 20%. Another comparison, shown in Figure 3.3 (d) and Figure 3.4 (d), showed the estimated speeds with the true mean speeds. Figures 3.3 and 3.4 show that estimates with 10% sample may result in large range of errors while the estimates with more than 50% of data provide reasonable estimates with small errors.

(a) Volume estimation error by sample rate
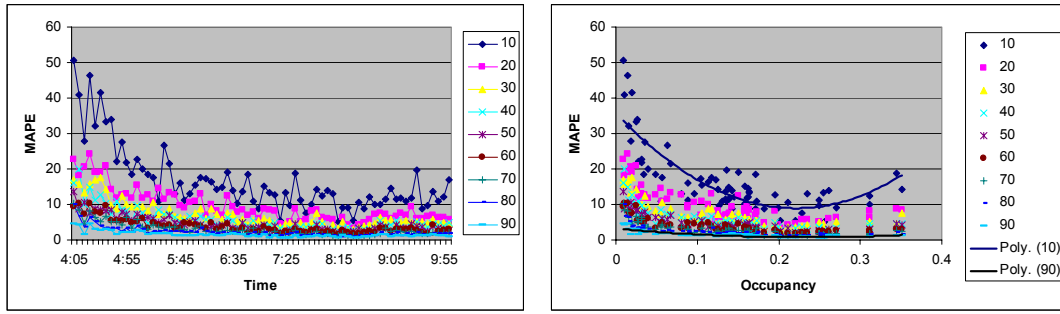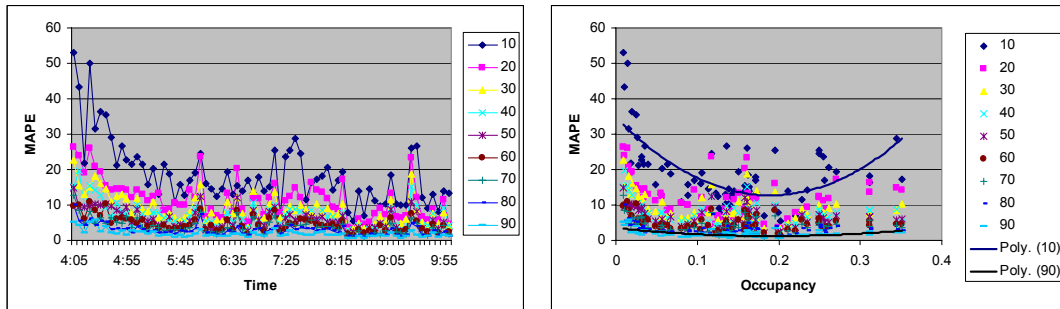
(b) Occupancy estimation error by sample rate

(c) Speed estimation error by sample rate

(d) Comparison of speed estimate with true mean

10% sample   30% sample   50% sample
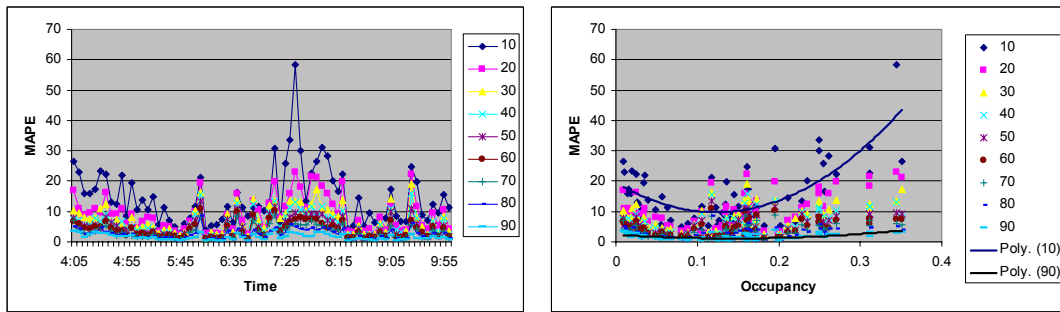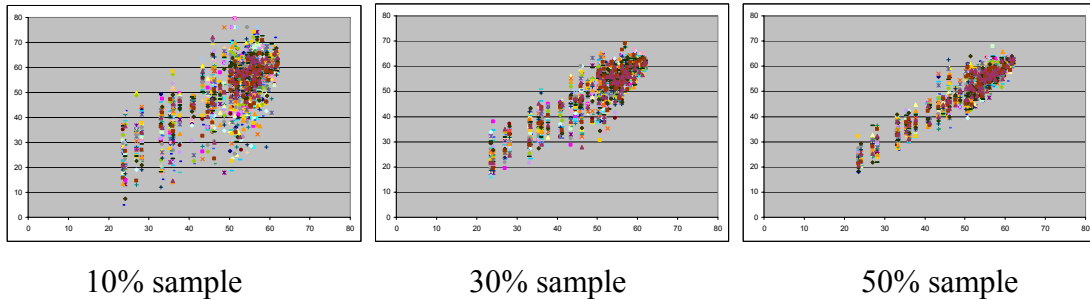
Figure 3.3: Quality of Estimates by Sample Rates (Detector on I-405, postmile 3.04)

(a) Volume estimation error by sample rate
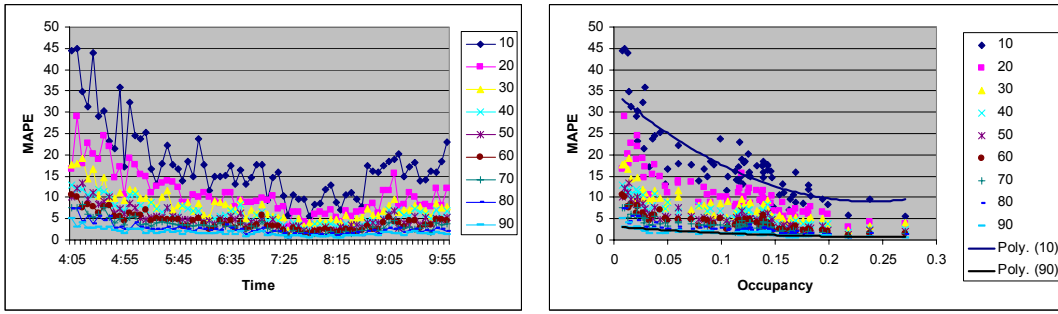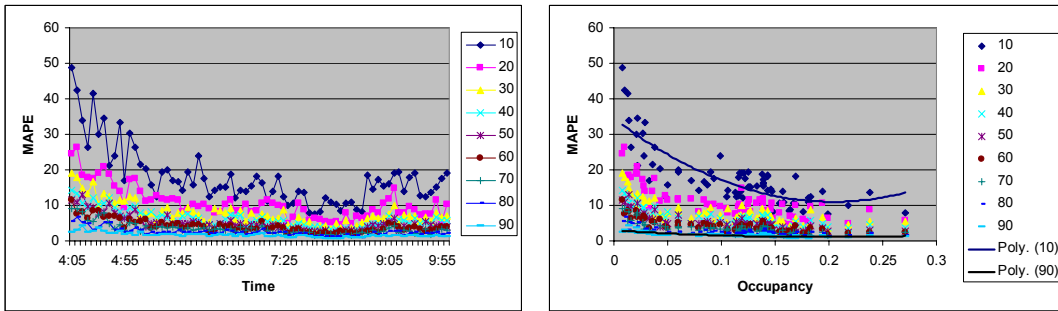
(b) Occupancy estimation error by sample rate

(c) Speed estimation error by sample rate

(d) Comparison of speed estimate with true mean

10% sample   30% sample   50% sample
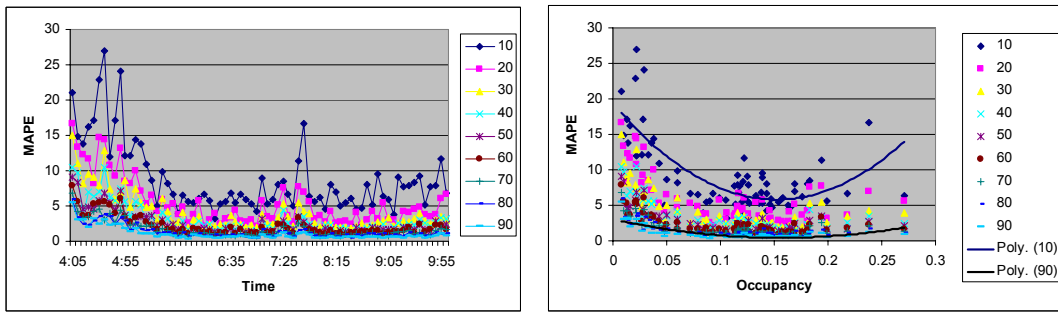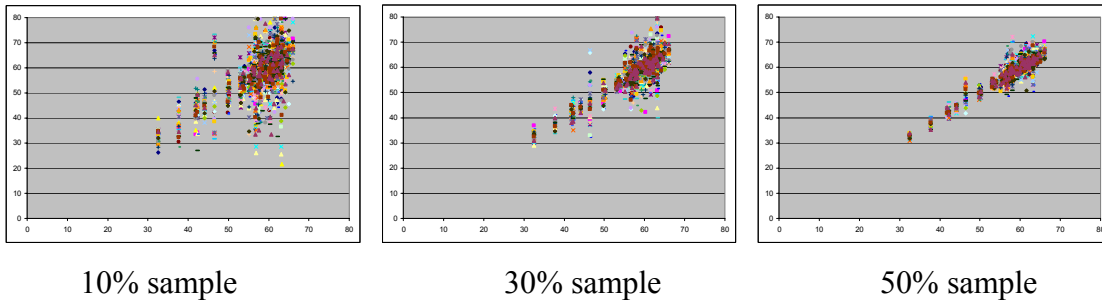
Figure 3.4: Quality of Estimates by Sample Rates (Detector on I-405, Postmile 3.86)

31

Overall errors, in terms of MAPE, are compared in Figure 3.5 with respect to the data availability. Certainly, as the sample rate increases, the error decreases. As shown in the figure, the simulation experiment shows that traffic data with a sample of 40% will be able to provide estimates with less than 5% of MAPE for all three traffic measures (volume, occupancy, and speed). Data with a 20% of samples are expected to provide estimates with less than 10% of MAPE. These errors are contributed by speed variation of lanes and the temporal variability of traffic measures during a five-minute period. In this study, we chose the sample rate of 20% as a minimum data requirement, which may provide estimates with less than 10% error.
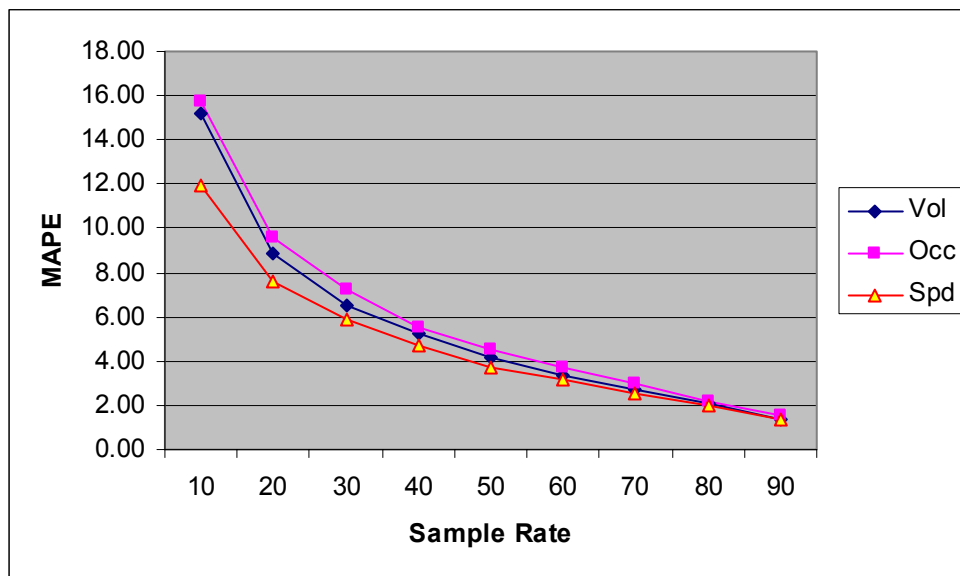
Figure 3.5: MAPE of Aggregated Measures by Sample Rates

Out of 499 mainline loop detector stations, we included 338 loop detector stations into our analysis after eliminating 161 loop detectors that include less than 20% data points during the analysis period (March 2001 – February 2002). Missing data at the loop detector stations for further analysis were interpolated using the yearly mean value of the corresponding five-minute period. The completed dataset includes traffic volume and occupancy for every five-minute period during 52 weeks.

3.2.4 Estimation of Speed and Section Travel Time

Loop detectors in Orange County are single loop detectors that provide only traffic volumes and occupancies, requiring the estimation of the travel speeds by assuming a so-called "g-factor," the summation of the average vehicle length, and the effective detection length. The average speed, $\bar{u}$, can be calculated as:

$$\bar{u} = \frac{g \times q}{5280 \times o},$$

(3.2)

where

q = flow rate (veh/h)

o = occupancy

The most critical value in calculating the speed is the assumed g-factor. In fact, applying appropriate g-factors is a key in estimating speeds from single loop detector data. This study calculates a g-factor representing each hour for each loop detector station as follows:

Step 1: Calculate initial g-factor by assuming the free-flow speed (75 mph) when the occupancy is lower than 0.06.

Step 2: Find an average g-factor representing each hour interval based on the identified initial g-factors over the 53 weeks.

Step 3: Apply smoothing parameter (0.9) for the next time period

$$\hat{g}(t+1) = p \cdot \hat{g}(t) + (1-p) \cdot g(t+1),$$

(3.3)

where

$\hat{g}(t+1)$ = g-factor for time step t+1,

$g(t+1)$ = initial average g-factor for time step t,

p = smoothing parameter (0.9).

By applying the above procedure, representative g-factors for each hour of day are calculated for all 388 loop detector stations. These g-factors are basic information for speed calculations in this study. Based on these g-factors, speeds are calculated every five-minute period, for all 52 weeks.

33

In this study, a freeway section corresponds to a mainline loop detector station. Within a section, a corresponding detector is located in the middle of the section as depicted in Figure 3.6. Based on the sections and corresponding estimated speeds, travel times for each section are calculated for every five-minute interval during the one-year analysis period.
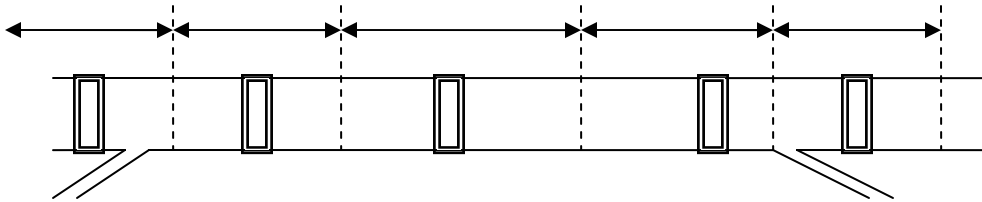
Figure 3.6: Section Definition and the Corresponding Detector Location

3.2.5 Route Travel Time Estimation

3.2.5.1 *Build a Network for Travel Time Study*

To find the feasible paths in the test area, the test network is constructed based on the loop detector position. As described in Figure 3.6, a node is set to be a median point between two adjacent loop detectors and a link is defined by two nodes. Based on the on-ramp and off-ramp locations, zones are defined as starting and ending points of trips.

For the route travel time analysis, we defined a smaller size network in the Irvine area, where alternative routes can be found and more reliable data are available. As a result, the total number of nodes in the network is 281 and of links is 345. Although there are 48 possible zones, in the smaller study network, we defined 10 representative zones for route travel time calculation as shown in Table 3.2 and Figure 3.7.

Table 3.2: Zones in the Network

Zone ID	Freeway	Entrance/Exit Street
1	I-5	Junipero Serra
2	I-5	Oso park
3	I-5	Tustin Ranch
4	I-5	4th street
5	SR-55	Fair
6	SR-55	Edinger
7	SR-55	4th street
8	I-405	Culver
9	I-405	Harbor
10	SR-73	New port coast
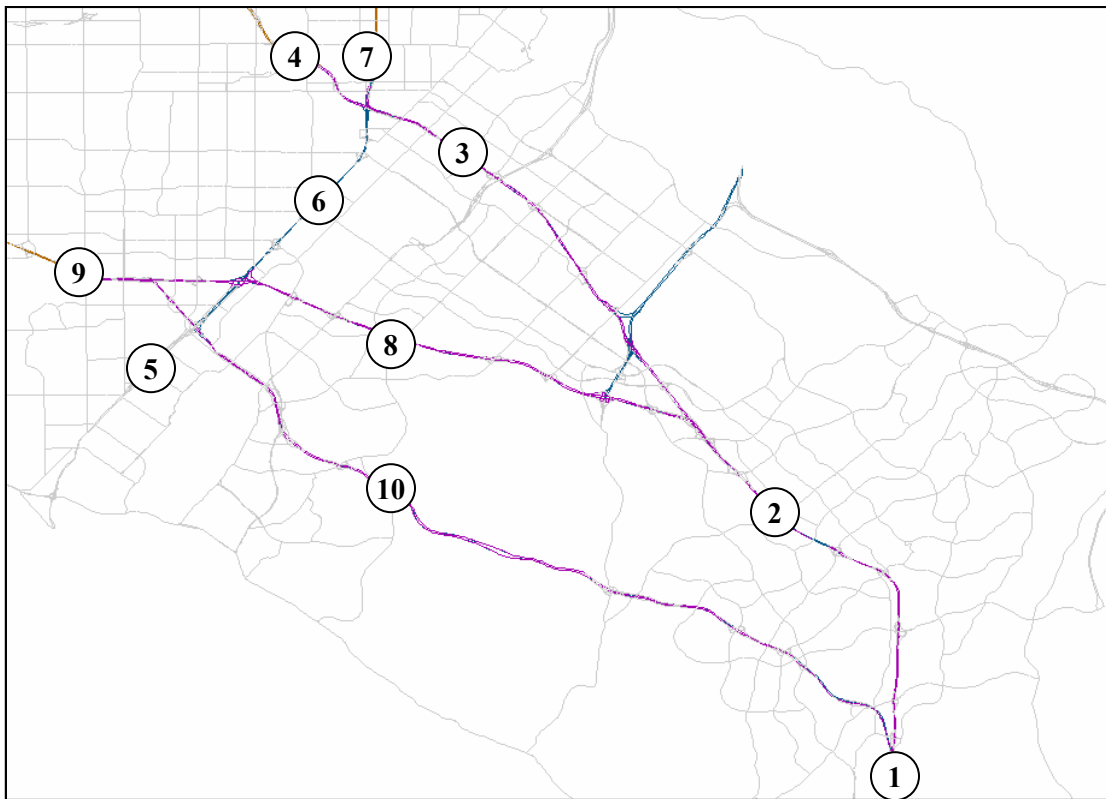
Figure 3.7: Selected Zones in the Test Network

3.2.5.2 *Enumeration of Feasible Routes*

For the purposes of this study, feasible routes between an origin and a destination, we determined using the all-feasible-path enumeration algorithm developed by Kim *et al*. (2004). The algorithm is based on conventional tree building algorithms. While general tree building algorithms only store the shortest paths or specific paths from an origin node to all other nodes, the all-path algorithm stores all paths from an origin to all other nodes. Similarly to conventional tree building algorithms, the all-path algorithm also branches spans to all nodes. The algorithm is divided into two sub-algorithms: 1) spanning branches and 2) sorting and storing paths. Since this all-path algorithm searches all feasible paths, the number of paths drastically increases as the network size increases. In this study, the number of feasible paths is constrained by the relative path length compared to the shortest route. In other word, the algorithm searches all paths that are shorter than α times the shortest path. Here, α is specified by users. This algorithm is more flexible and general than the conventional k-shortest path algorithm in that the number of paths for each OD pair is not fixed but determined depending on the network topology and the α value. The detail algorithm is described as follows.

Notation

BR_b^n : The nth link on a branch b,

BN_b^n : The nth node on a branch b,

BC_b^n : The branch cost of the branch b until link n,

$Closed\_BR_b^n$: If BR_b^n is closed, 1, otherwise, 0,

$NC(m)$: The travel cost from origin to node m,

$Link(m)_i^{start}$: The ith link emanating from node m,

$C(Link(m)_i^{start})$: The link cost of $Link(m)_i^{start}$,

$I[Link(m)_i^{start}]$: The head node of $Link(m)_i^{start}$,

$J[Link(m)_i^{start}]$: The tail node of $Link(m)_i^{start}$,

$Path[o][J[Link(m)_i^{start}]]^n$ = nth link on a path from origin to $J[Link(m)_i^{start}]$,

α : A parameter for bounding a branch.

Step 1. Initialization

Set $BR_b^0 = 0$, $BC_b^0 = 0$,

Set n=1,

$NC(m) = \infty \quad \forall\, m$,

m=o,

Step 2. Spanning branches

If $BC_b^n + C(Link(m)_1^{start}) \le \alpha \cdot NC(J[Link(m)_1^{start}])$, then set $BR_b^1 = Link(m)_1^{srart}$.

If $J[Link(m)_1^{start}]$ is already visited by BR_b^n, $Link(m)_1^{srart}$ does not be added to

BR_b^n,

$Closed\_BR_b^n = 1$,

Otherwise, $Closed\_BR_b^n = 1$,

If there is another link emanating from the origin, b=b+1,

Step 3. Storing paths

$Path[o][J[Link(m)_i^{start}]]^n = BR_b^n$ for all branches b,

Step 4. Stopping criteria

For all branches, if $Closed\_BR_b^n = 1$, then stop,

Otherwise, n = n+1, go to Step 2.

The path enumeration algorithm above identifies numerous alternative routes for a given OD pair. In this study, the paths were bounded up to 30% longer than the shortest path by applying 1.3 as the parameter for bounding a branch (α value). By applying the above algorithm, a total of 170 paths were identified. These paths are included for further path travel time analysis.

3.2.5.3 *Path Travel Time Calculation*

With the five-minute interval section travel time data, the time-dependent path travel times are calculated. These path travel times are also calculated every five-minute departure interval at the origin. In other word, travel times for a path connecting an origin and a destination is calculated by departure time at the origin, based on the time-dependent section travel times. Figure 3.8 shows an example calculation of the travel time of path k with four consecutive sections. The travel time at a section is selected based on the time when a driver would drive after traveling the previous sections. The path travel time departing at time t is obtained by summing the corresponding section travel times.

Path k

Link1   Link2   Link3   Link4

0. t_0: Departure
   Link1 travel time: τ_1 at t_0
   τ_1

1. $t_0+\tau_1$: $t_0+l\Delta t \leq t_0+\tau_1 < t_0+(l+1)\Delta t$
   Link2 travel time: τ_2 at $t_0+l\Delta t$
   τ_2

2. $t_0+\tau_1+\tau_2$: $t_0+m\Delta t \leq t_0+\tau_1+\tau_2 < t_0+(m+1)\Delta t$
   Link3 travel time: τ_3 at $t_0+m\Delta t$
   τ_3

3. $t_0+\tau_1+\tau_2+\tau_3$: $t_0+n\Delta t \leq t_0+\tau_1+\tau_2+\tau_3 < t_0+(n+1)\Delta t$
   Link4 travel time: τ_4 at $t_0+n\Delta t$
   τ_4

   Path travel time $TT_k = \tau_1+\tau_2+\tau_3+\tau_4$
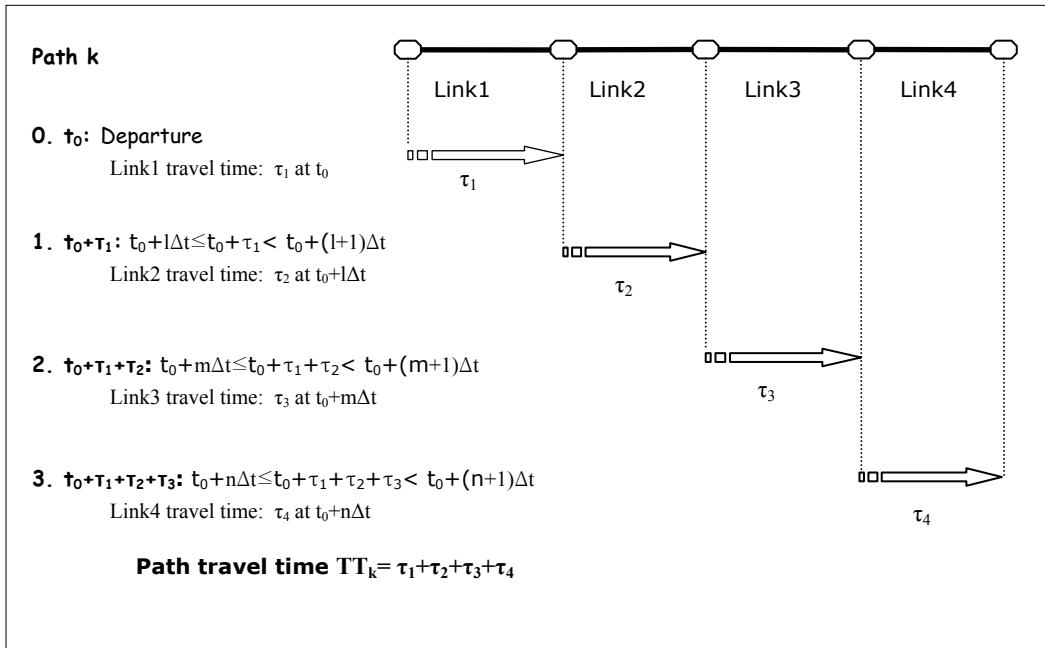
Figure 3.8: Calculation of Path Travel Time

3.2.6 GIS Database Construction

3.2.6.1 *Overall Procedure for GIS DB*

The objective of this part is to construct a base framework for utilizing Geographic Information Systems (GIS) and base database to analyze the travel time variability. There are many commercially available GIS software packages, such as ArcView GIS, ArcInfo, ArcGIS, MapInfo, TransCAD, and GeoMedia. Since UCI has unlimited access to the ArcView GIS and ArcGIS, which are the most popular packages for GIS users, and their segmentation and referencing tool make it useful for the analysis of transportation data, this research has utilized the ArcGIS produced by the Environmental Systems Research Institute (ESRI), Inc.

Traffic data used in this project is collected from loop detectors in Orange County area. Since the one-year data set is based on five-minute intervals and all 499 detectors, manipulated in the project, the size of the data is over 50,000,000 records. Because of this, database management systems (DBMSs) such as Oracle, MS-SQL server, Informix, Sybase, and MySQL are necessary to efficiently manage the huge database. MySQL, which is the world's most popular open source database server, was selected for this project.

After preparing two computer packages, GIS and database, a basemap is required for displaying traffic road network, traffic facilities, and features such as loop detector location, accident location, etc. Although Topologically Integrated Geographic Encoding Referencing (TIGER) files are freely available as a basemap, the TIGER files are not accurate enough for this project. Therefore, a map produced by Navigation Technology Corp. (NavTech) was used in this project. The NavTech map included several data layers including road network discriminated by the functional class and attribute data, and included more accurate information than other digital maps. However, the NavTech map was still impertinent to directly use it for the project, because each highway layer should be separated from the highway layer for using "Linear Referencing".

In this project, the detector location and accident location are referred by highway postmile, which are declared by Caltrans, but the NavTech map cannot identify the highway postmile. Since most traffic related data sets are collected by different agencies, or collected for different purposes, a common referencing system is required for integrating all data sets. As mentioned,

39

the detector location and accident location, as the reference events, are referred on the NavTech map using "Linear Referencing Technique" in ArcGIS. To efficiently query and analyze the traffic data from MySQL, it is very important to design the database structure. Other analysis would be processed with spatial analysis of GIS, database functionalities, and visual programming. Figure 3.9 shows the process for this part.
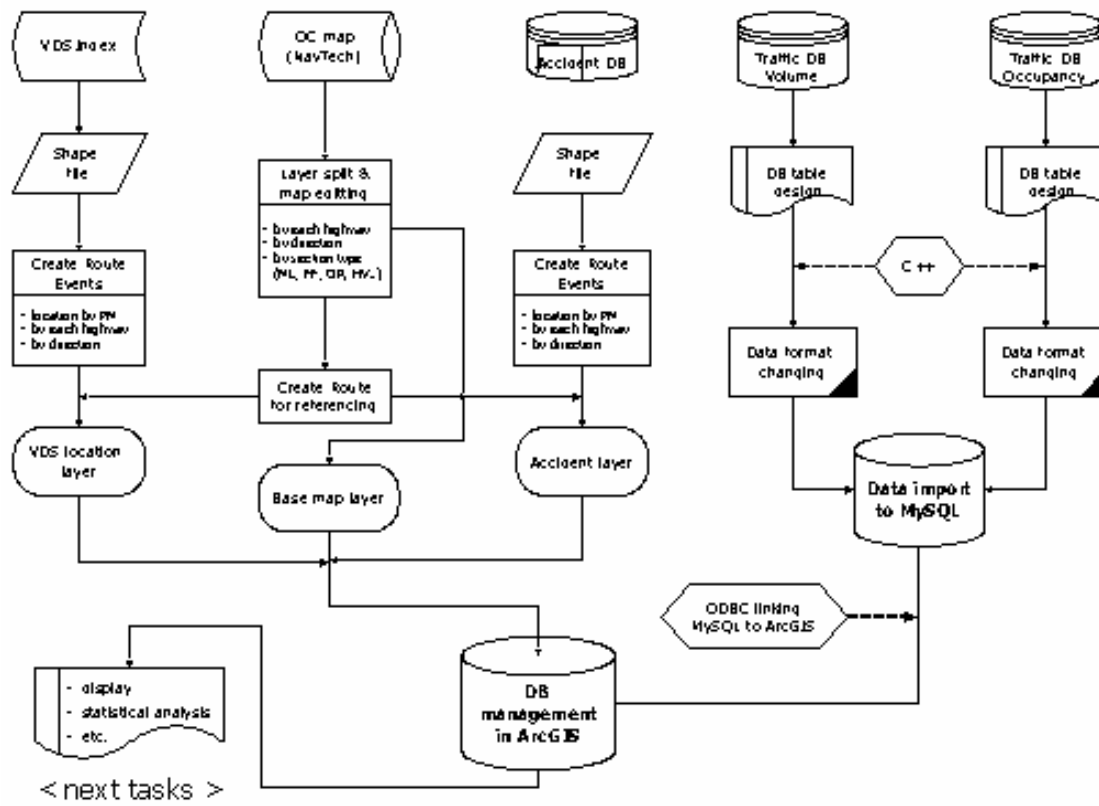
Figure 3.9: Task Process for GIS Database

4.2.6.2 *Linear Referencing for Map Editing*

To match events on the GIS map, coordinate data is required. Likewise, x-y coordinates are required to represent loop detector locations and accident locations on the GIS map. Although the loop detector location data have x-y coordinates, as described in Table 3.3, locations do not exactly match with the highway links in the GIS map. Fortunately, both the loop detector data and the accident data have the postmile attribute in their records and all point events related to them can be positioned on the GIS map using Linear Referencing System in ArcGIS. Instead of x-y coordinates, linear features such as highways, streets, railroads, rivers, and telephone lines can be recorded by using a single relative position (ESRI, 2001). That is, all event features can be identified by a known measurement, such as postmiles on the linear feature. For example, 1204198 VDS in Table 3.2 identifies the position on 0.65 miles away from the starting point of the I-5. The postmile attribute rather than x-y coordinates is used in identifying the location of VDS. Table 3.4 represents types of loop detectors.

Table 3.3: Data Format for Loop Detector Data

VDS_ID	Class	FWY Name	Dir	Postmile	Latitude	Longitude	Location
1204198	ML	5	N	0.65	33.405577	-117.59843	S. LUIS REY
1204209	FR	5	N	1.24	33.41139	-117.60107	MAGDALENA
1204211	ML	5	N	1.26	33.41196	-117.6014	MAGDALENA
1204204	OR	5	N	1.26	33.413536	-117.60217	MAGDALENA
1204228	FR	5	N	1.83	33.416473	-117.60365	EL CAMINO REAL
1204230	ML	5	N	1.83	33.418488	-117.60458	EL CAMINO REAL
1204222	OR	5	N	1.84	33.421011	-117.60672	EL CAMINO REAL
1204242	FR	5	N	2.45	33.425546	-117.6093	PRESIDIO
1204244	ML	5	N	2.47	33.427031	-117.61036	PRESIDIO

Table 3.4: Description of Loop Detector Types

VDS type	Description	VDS type	Description
CD	Collector/distributor	HF	HOV off ramp to arterial
CH	HOV collector/distributor	HI	HOV mainline Ingress
FC	Off collector/distributor	HN	HOV on ramp to arterial
FF	Fwy/Fwy connector	HV	Mainline HOV
FH	HOV off collector/distributor	ML	Mainline
FR	Off ramp	NH	HOV on collector/distributor
HB	HOV fwy/fwy connector	OC	On collector/distributor
HE	HOV mainline egress	OR	On ramp

Since all detector locations and accident locations are based on each highway postmile, all routes for linear referencing should be composed by each highway (e.g., I-5, I-405, SR-55, SR-22, etc.). Therefore, it is necessary to separate highway layers by highway name. Also, each highway layer should be separated by the direction because all highway layers have the direction such as north bound, south bound, east bound and west bound. An additional separation process is required for loop detectors, since loop detectors are discriminated by the detector type or location such as mainline, on-ramp, off-ramp, freeway to freeway, HOV, etc. Figure 3.10 shows an example of the route expansion by highway name and Figure 3.11 is an example of adding a ramp and a collector.

| Before adding links of CA-55 | After adding links of CA-55 |
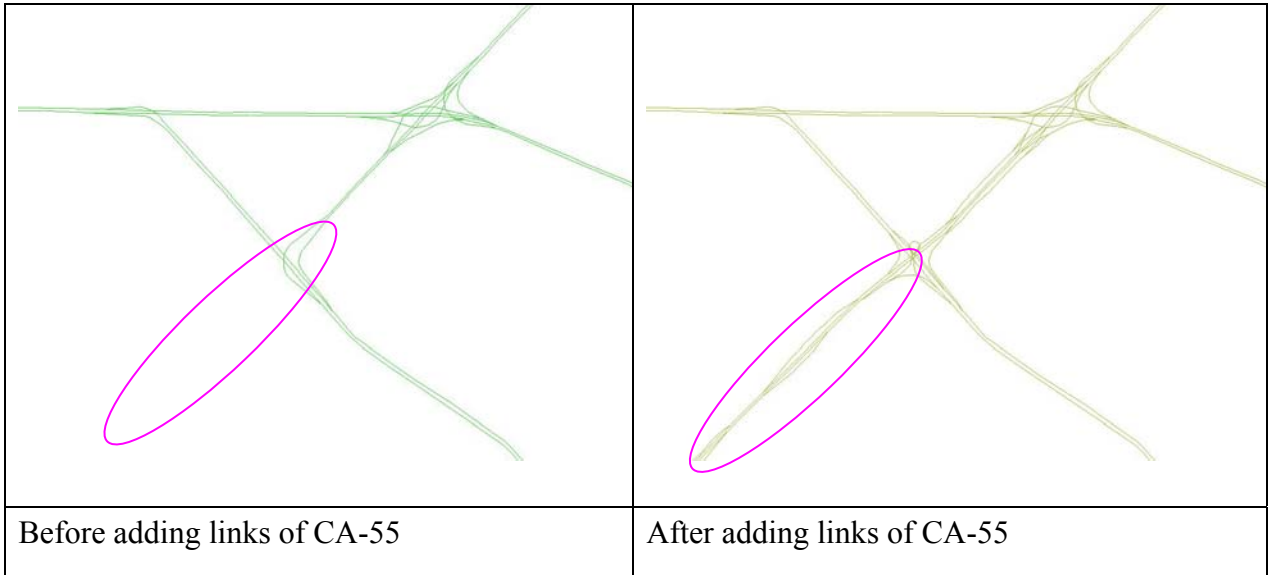
Figure 3.10: An Example of the Map after Adding Links

| Before adding ramp and collector | After adding ramp and collector |
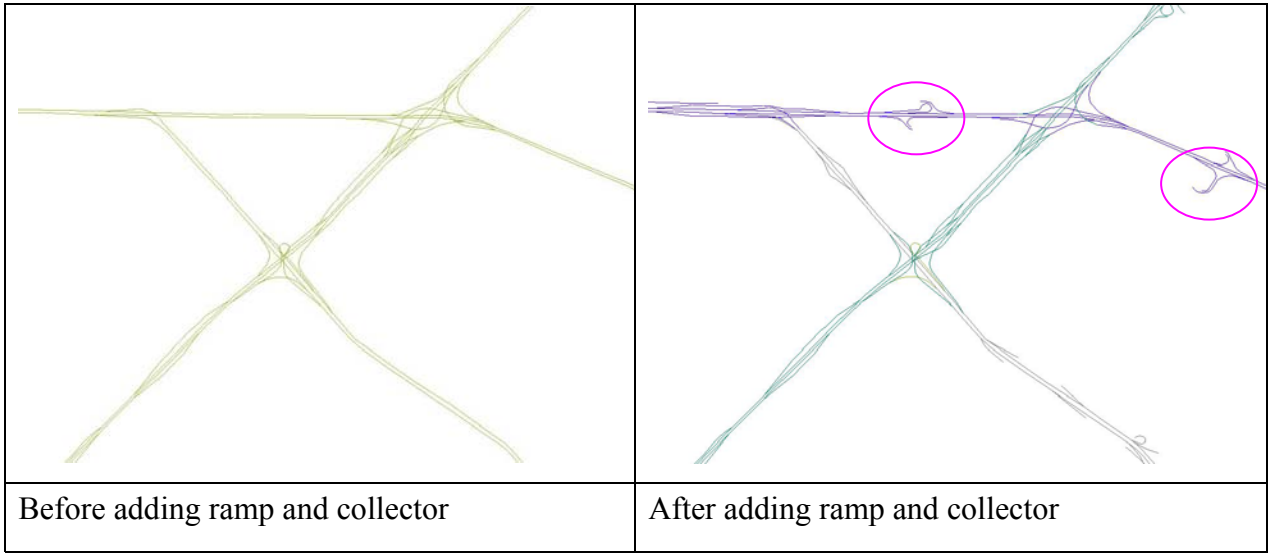
Figure 3.11: An Example of the Map after Adding Ramps and Collector Links

Routes are defined in GIS by using linear referencing. A route consists of one or more line features sequentially connected. Since each route has an identifier stored in a field and an associated measurement system, attributes are assigned to a route. Then, all events along a linear feature can be identified by this measurement system. Using the "Create Route Wizard" tool of the linear referencing in Arc Toolbox of ArcGIS, routes are generated. Both the detector layer and the accident data layer should be separated similar to a route layer.

4.2.6.3 *Projection of Detector Location and Accident Location on GIS Map*

Based on route layers and separated event layers for detector location, detector locations on the highway layer are projected using the "Route Events Geoprocessing Wizard" tool of ArcMap. This operation derives point events from point features by locating the features along a route reference. Figure 3.12 depicts some of the mainline detector locations, projected on the GIS map using linear referencing. The example shows how detectors are properly located in GIS map.

| Before matching ML detector location | After matching ML detector location |
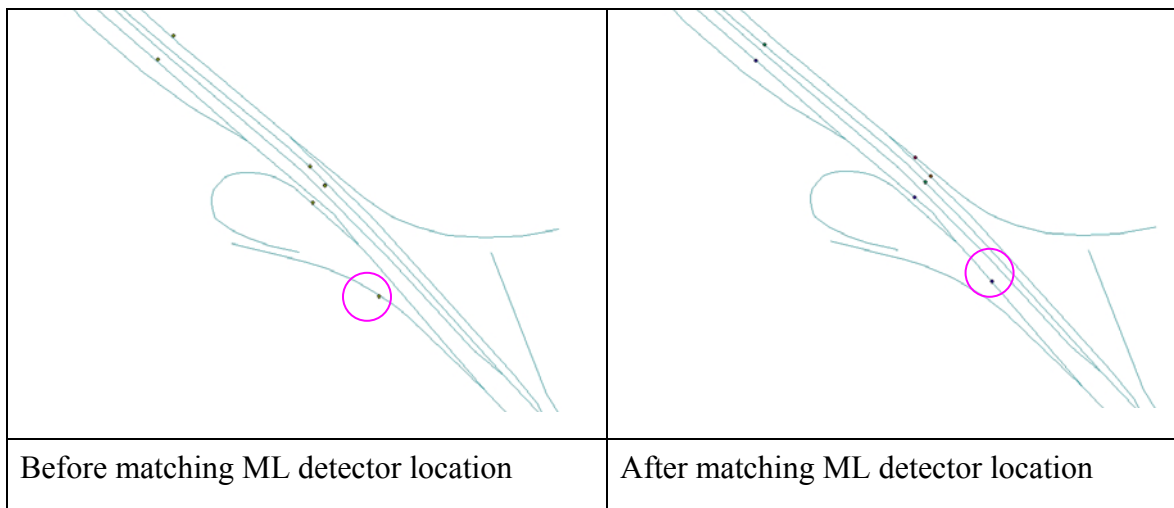
Figure 3.12: An Example of the Projection of ML Detector on GIS Map

Likewise, since there is no longitude and latitude data in the accident dataset, the linear referencing method is required to project the accident locations on the GIS map. Once creating a route feature class from a line feature class using Create Routes Wizard of Arc Toolbox in ArcGIS, accident locations along the routes can be identified. Figure 3.13 depicts the revised location of accidents and detectors by using linear referencing.
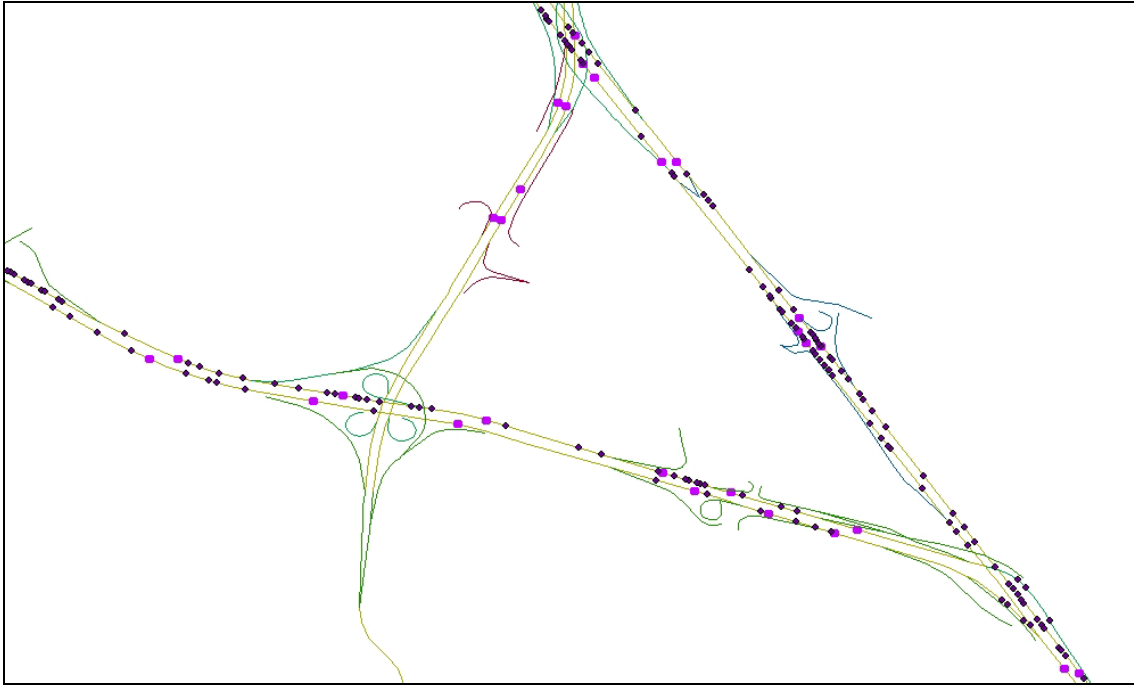
Figure 3.13: Location of Accident and Detector

3.2.7   Database Management and Analysis

3.2.7.1 *MySQL for Database Management*

Traffic volume and occupancy data are added to the database. Since ArcGIS or MS-Access has a limitation to handle and build the big database, MySQL version 5.0.0 was employed to build and handle all base datasets. Then, the MySQL database is linked to ArcCatalog of ArcGIS using ODBC (Open Database Connectivity) driver. Table 3.5 represents the merged traffic data table. In the table, VDS_ID and time_code are set as the primary key, and others are set as the Index key.

Table 3.5: Structure of the Database in MySQL

ID	VDS_ID	Time_code	Weekday	Volume	Occupancy	Speed
1	1204198	2001-03-01 00:00:00	Wed	370	0.009	N/A
2	1204198	2001-03-01 00:05:00	Wed	256	0.01	N/A
3	1204198	2001-03-01 00:10:00	Wed	244	0.01	N/A
⋮	⋮	⋮	⋮	⋮	⋮	⋮
5245479	1209438	2002-02-28 23:50:00	Wed	-9	-9	.
52454880	1209438	2002-02-28 23:55:00	Wed	-9	-9	.

3.2.7.2 *Linking SAS for Statistical Analysis*

Although the basic statistics can be obtained from the database, using SQL (Structured Query Language), commercial statistical analysis software can make more powerful analysis than using a DBMS alone. ODBC (Open Database Connectivity) driver enables other programs to access the database for further statistical analysis. ODBC driver is usually used when independent or simultaneous access to different data sources is required. Recent DBMSs support ODBC that provides a way for client programs to access a variety of databases or data sources. MySQL have also released a 32-bit ODBC driver, known as MySQL ODBC 3.51 Driver. In this study, the database, built using MySQL, is connected to ArcGIS and SAS for effective data management with GIS capability and further statistical analysis. Figure 3.14 shows a database connection to SAS using ODBC.
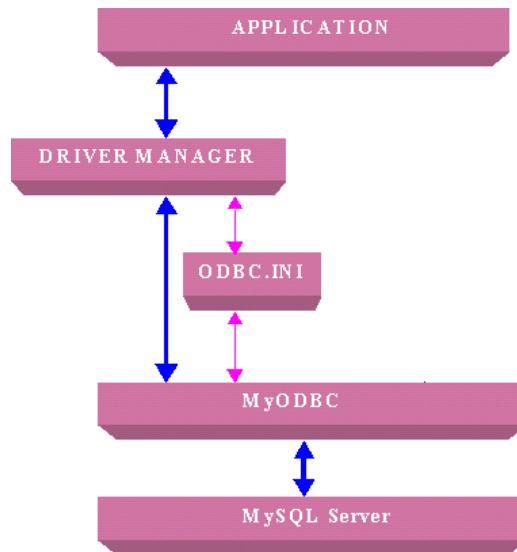
Figure 3.14: ODBC Connection Architecture (source: www.mysql.com)
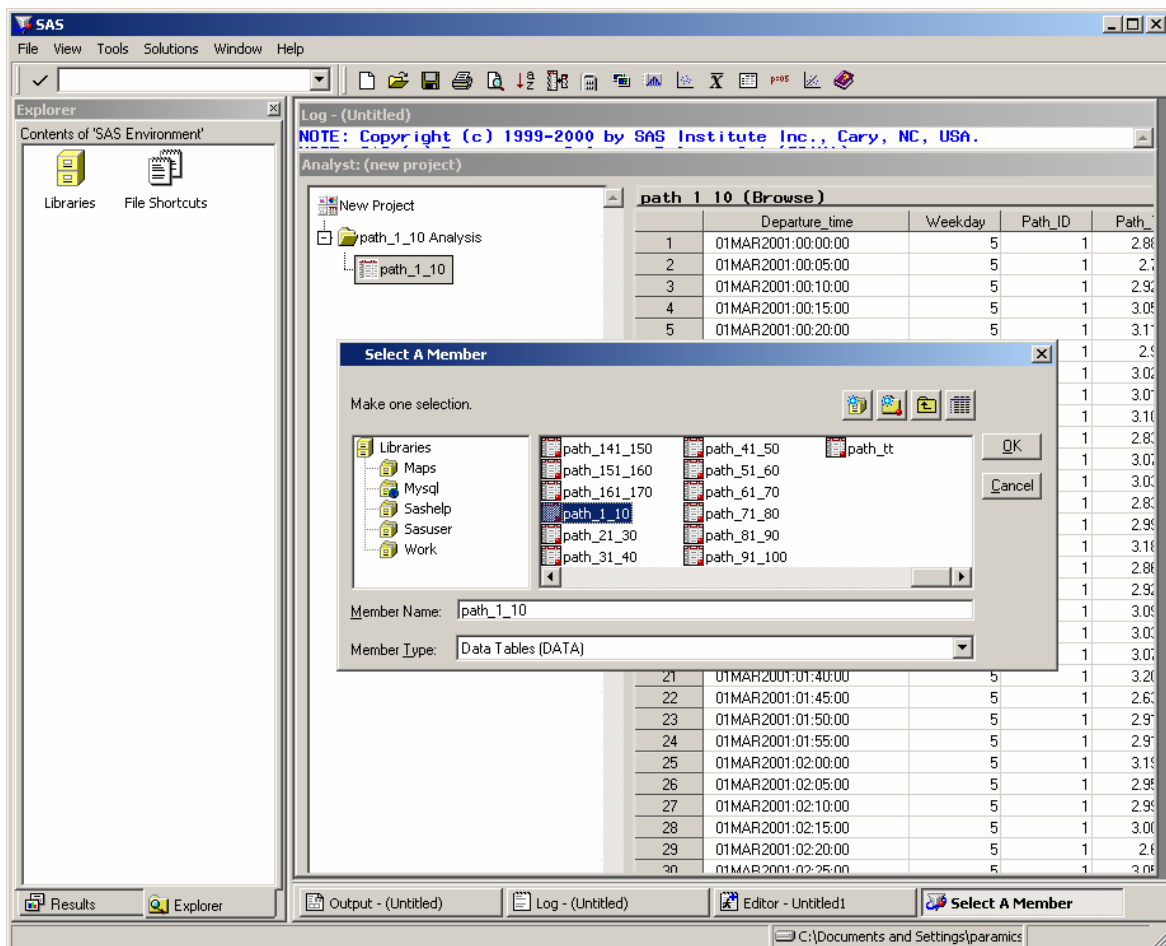
Figure 3.15: Connecting MySQL Database to SAS

3.3 Analysis of Section Travel Time Variability

3.3.1 Measures of Section Travel Time Variability

Section travel time variability is firstly measured using our traffic database. While many existing studies have investigated traffic variables in terms of traffic volume, occupancy, speed, and/or travel time itself, this study focuses on analyzing the variability of section travel times. In this study, the variability is measured from two different aspects. The first aspect is the variability of day-to-day travel time, and the second one is within-day variability. Table 3.6 shows the difference between day-to-day travel time variability and within-day variability.

Table 3.6: Measuring Travel Time Variability for Each Section

Time (t)	Day (d)						Day-to-day Variability
	3/1/01	3/2/01	……	D	……	2/28/02	
00:00-00:05	$\tau_{l,1,1}$	$\tau_{l,2,1}$	……	$\tau_{l,d,1}$	……	$\tau_{l,365,1}$	$DV_{l,1}$
00:05-00:10	$\tau_{l,1,2}$	$\tau_{l,2,2}$	……	$\tau_{l,d,2}$	……	$\tau_{l,365,2}$	$DV_{l,2}$
00:10-00:15	$\tau_{l,1,3}$	$\tau_{l,2,3}$	……	$\tau_{l,d,3}$	……	$\tau_{l,365,3}$	$DV_{l,3}$
⋮	⋮	⋮	……	⋮	……	⋮	⋮
t	⋮	⋮	……	⋮	……	⋮	⋮
⋮	⋮	⋮	……	⋮	……	⋮	⋮
23:45-23:50	$\tau_{l,1,187}$	$\tau_{l,2,187}$	……	$\tau_{l,d,187}$	……	$\tau_{l,365,187}$	$DV_{l,187}$
23:50-24:00	$\tau_{l,1,188}$	$\tau_{l,2,188}$	……	$\tau_{l,d,188}$	……	$\tau_{l,365,188}$	$DV_{l,188}$
Within-day Variability	$WV_{l,1}$	$WV_{l,2}$		$WV_{l,d}$		$WV_{l,365}$	

The day-to-day variability is measured to investigate how the travel time for a given time of day is fluctuating day-to-day. The higher variability of travel time implies that travel times are hard to predict from both travelers and system operator's point of view, which may lead to travelers' earlier departure for them to avoid the late arrival at destination. A standard deviation of data usually represents the level of data fluctuation. In this study, standard deviation and normalized standard deviation are used as measures of variability. The normalized standard deviation is calculated in order for the measures to be comparable, regardless the difference in length or travel time. Accordingly, three basic statistics, mean, standard deviation, and normalized standard deviation, are calculated, and the day-to-day travel time variability is represented by the (normalized) standard deviation. The time interval t in this study is five-minutes, and the analysis

period can be all weekdays during a one-year period or a specified period. A special interest is whether the travel time variability differs by day or month.

$$\bar{\tau}_{lt} = \frac{\sum\limits_{d \in D} \tau_{ldt}}{N_D},$$ (3.4)

$$\sigma_{lt} = \sqrt{\frac{\sum\limits_{d \in D} \left(\tau_{ldt} - \bar{\tau}_{lt}\right)^2}{N_D}},$$ (3.5)

$$V_{lt} = \frac{\sigma_{lt}}{\bar{T}_{lt}},$$ (3.6)

where,

$\bar{\tau}_{lt}$ = Average travel time of section l for time interval t during the period of D

σ_{lt} = Standard deviation of travel time of section l for time interval t during the period of D

V_{lt} = Normalized standard deviation of travel time of section l for time interval t during the period of D

τ_{ldt} = travel time of section l for time interval t on day d

N_D = the total number of days during the period of D

The other travel time variability is measured during a given period, within a day, which we call within-day travel time variability. Such variability can be measured during an hour period or during a day period. The within-day variability explains how much the traffic pattern fluctuates during a given period within a day. Certainly, the measures will vary by the period applied. When a day is applied, within-day variability can be a representative value for a day; however, such a daily fluctuation may not be a meaningful measure from a traveler's standpoint. Rather, a traveler may be concerned about the possible fluctuation during the period when he/she travels. In such circumstance, the within-day travel time variability, during a time of day, may be more meaningful. In this study, the within-day variability during an hour period and during a day period are compared since there is no consensus on the analysis interval.

$$\bar{\tau}_{ldp} = \frac{\sum\limits_{t \in P} \tau_{ldt}}{N_P}, \tag{3.7}$$

$$\sigma_{ldp} = \sqrt{\frac{\sum\limits_{t \in P}\left(\tau_{ldt} - \bar{\tau}_{ldp}\right)^2}{N_P}}, \tag{3.8}$$

$$V_{ldp} = \frac{\sigma_{ldp}}{\bar{T}_{ldp}}, \tag{3.9}$$

where

$\bar{T}_{ldp}$ = Average travel time of section l during a given time period P within a day d

σ_{ldp} = Standard deviation of travel time of section l during a given period P within a day d

V_{ldp} = Normalized standard deviation of travel time of section l during a given period P within a day d

T_{ldt} = travel time of section l at time t on day d

N_P = the total number of time intervals during a given period P

3.3.2 Day-to-Day Travel Time Variability

3.3.2.1 *Measures of Day-to-Day Travel Time Variability*

The measures of day-to-day travel time variability are calculated every five-minute interval for a day by loop detector station based on all weekdays (a total of 251 weekdays except holidays during the year). This statistic is to show how a travel time during a given five-minute interval is different from other days' travel time during the same five-minute period. Accordingly, each section of freeway has 288 measures of day-to-day travel time variability for each five-minute interval.

In order to compare the difference between sections, the each time interval's variability measures measured over a year period are averaged for a given period (N_T), 288 five-minute intervals for a day, or 12 five-minute intervals for an hour, as shown equations below.

$$\bar{T}_l = \frac{\sum\limits_{t \in T} \bar{T}_{lt}}{N_T}, \ \sigma_l = \frac{\sum\limits_{t \in T} \sigma_{lt}}{N_T}, \ \text{and} \ V_l = \frac{\sum\limits_{t \in T} V_{lt}}{N_T} \qquad (3.10)$$

Table 3.7 compares daily average and morning peak hour (7:00 – 8:00 AM) day-to-day travel time variability by loop detector stations. In the table, travel time variability of a representative section for each directional freeway is compared as an example. The average speeds were calculated from the average travel time for comparison purpose. The average normalized standard deviations (N-STD) are comparable to each other as the same magnitude, since they were normalized by dividing by the average travel time. In this example, SR-22 eastbound shows the largest travel time variation in daily average while the I-405 southbound shows the largest during morning peak hour. However, note that these measures do not represent overall freeway's travel time variability since such variability measures vary by location. When being compared between daily average and AM peak average, the sample average travel time during AM peak is 50% higher than daily average, the average travel time variability, STD or N-STD, during AM peak is almost double or more than the whole day average. This means that both travel time and its variability are larger during morning peak hour.

Table 3.7: Day-to-Day Travel Time Variability by Loop Detector Stations

Freeway	vds_id	Daily average				7:00 - 8:00			
		Travel time (sec)	Speed (mph)	STD	N-STD	Travel time (sec)	Speed (mph)	STD	N-STD
5N	1204532	35.89	69.22	5.13	0.1216	62.76	39.58	28.43	0.4442
5S	1205182	25.80	57.91	4.26	0.1485	45.93	32.53	12.47	0.2738
405N	1201510	51.62	61.03	10.02	0.1667	47.67	66.08	6.01	0.1256
405S	1201805	37.11	57.72	8.16	0.1585	95.70	22.38	50.56	0.5211
22E	1202785	52.77	36.84	13.39	0.2295	72.88	26.67	21.48	0.2959
22W	1202840	28.36	70.44	3.69	0.1255	29.82	67.00	3.86	0.1293
55N	1210172	22.41	65.06	3.74	0.1325	29.23	49.89	7.60	0.2300
55S	1203284	33.76	59.71	6.01	0.1392	99.10	20.34	33.65	0.3426
57N	1202263	92.53	55.44	15.45	0.1305	85.13	60.26	14.95	0.1662
57S	1202422	64.95	64.57	9.30	0.1166	139.28	30.11	50.64	0.3623
91E	1208226	51.13	53.86	11.96	0.1682	40.35	68.25	2.36	0.0586
91W	1208260	16.75	63.95	2.83	0.1607	24.56	43.60	5.78	0.2359
Average		42.76	59.65	7.83	0.1498	64.37	43.89	19.82	0.2655

4.3.2.2 *Time-of-Day Travel Time and Day-to-Day Variability*

Travel time of a section changes over time, as does its travel time variability. Figure 3.17 shows changes of the normalized STD as a measure of travel time variability. Each section of freeway shows different patterns. The patterns of travel time variability look similar to the levels of traffic congestion, as they show higher variability during AM or PM peak hours. In many cases, each direction of freeway shows the opposite pattern. For example, while the I-405 southbound shows very high variation during AM peak hours, the northbound shows higher variation during PM peak hours. SR-57 and SR-91 also show very clear distinction between directions.

Figure 3.18 Figure 3.18 shows changes in average travel times during time of day, with the interval of plus/minus one standard deviation ($\pm \sigma$). Assuming that the travel time pattern follows a normal distribution, the travel time for any five-minute period during any weekday will fall within the range with probability of 68.27% (the shaded area in Figure 3.16). The known probabilistic distribution can be used as a way of identifying abnormal traffic conditions. Presumably, abnormally short or long travel times are results of traffic incidents or drastic demand changes by an event.
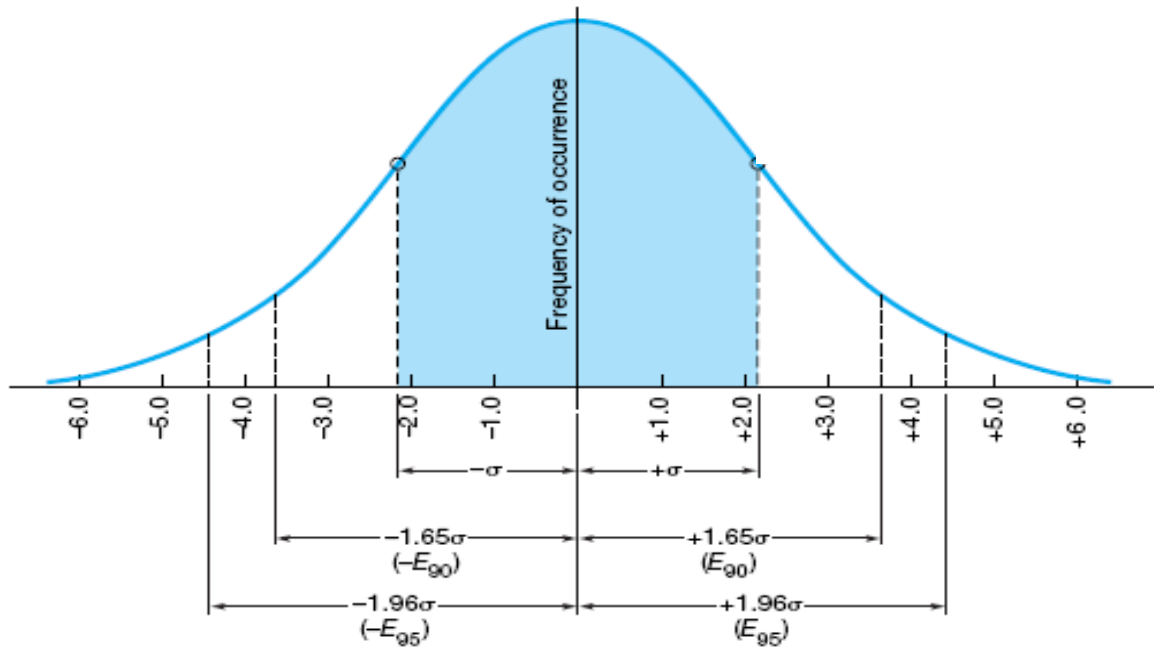
Figure 3.16: Normal Distribution
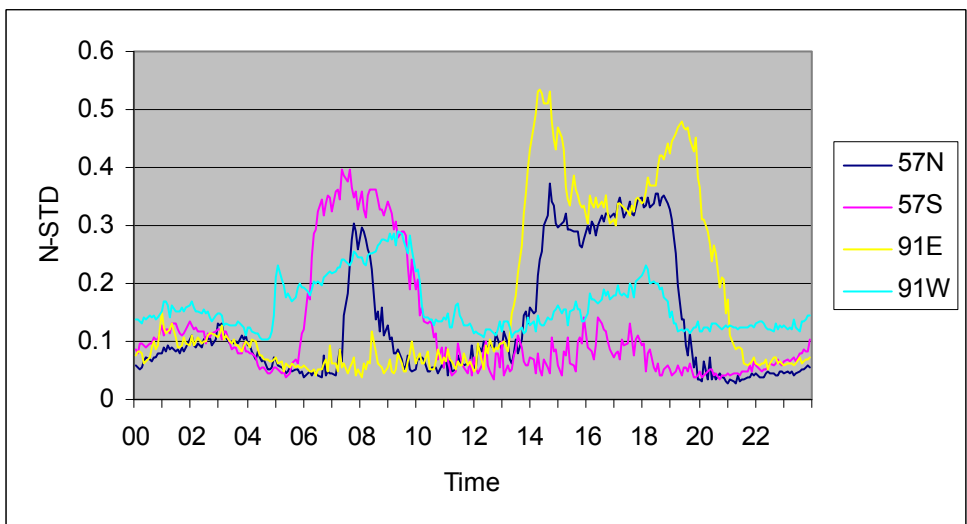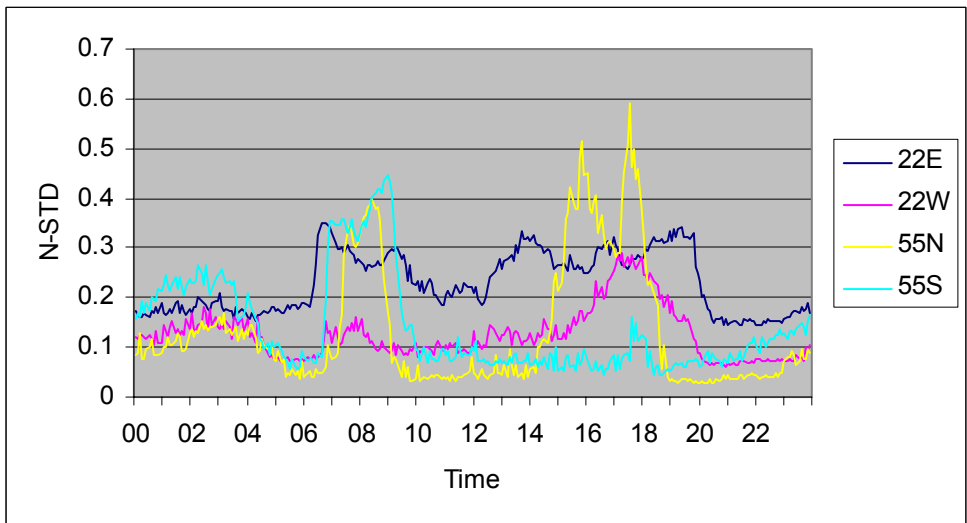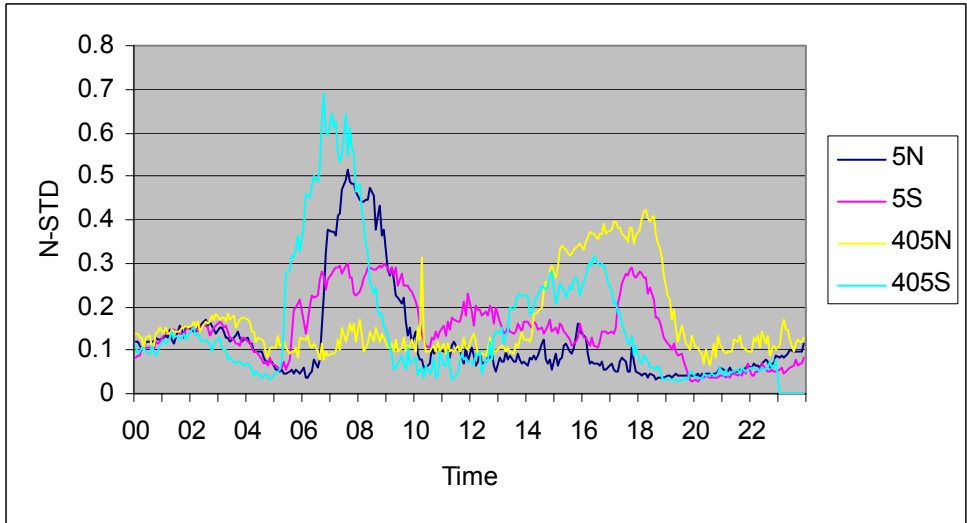
Figure 3.17: Day-to-Day Travel Time Variability

55
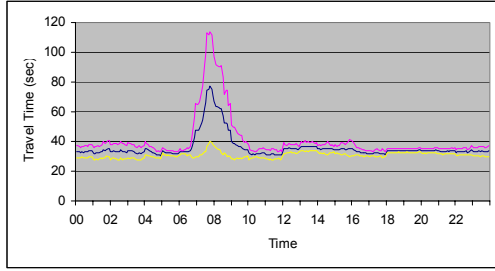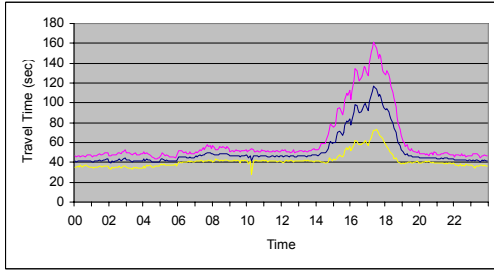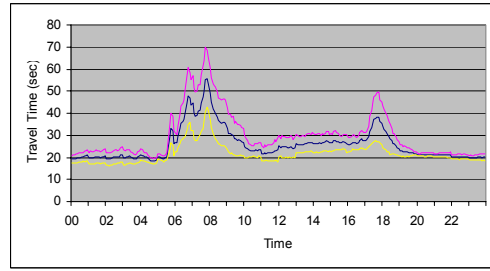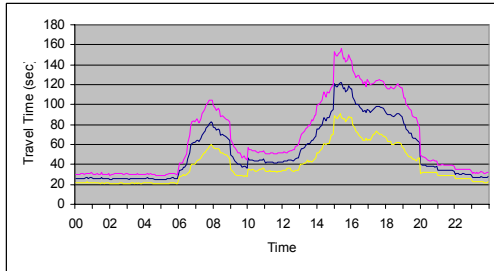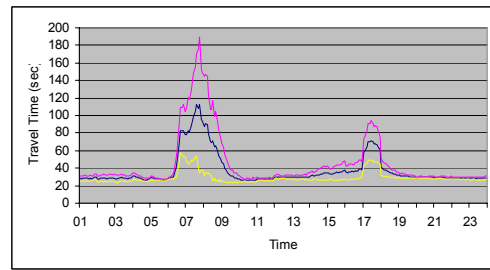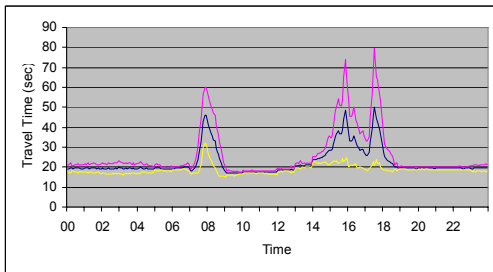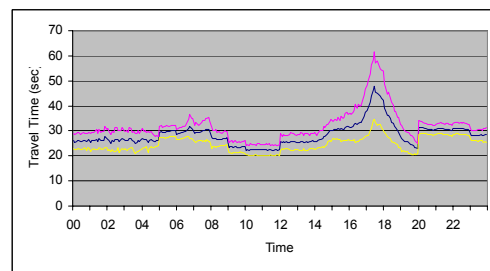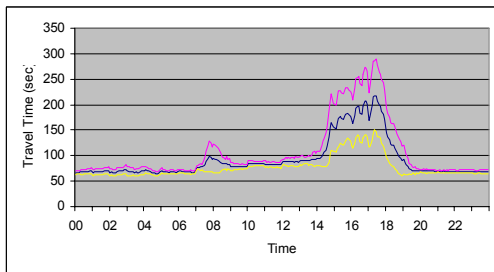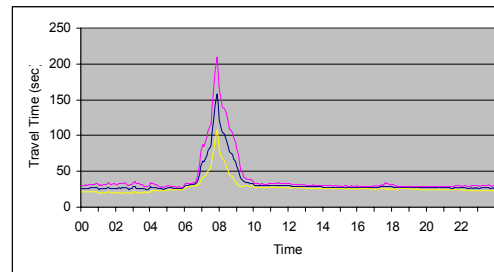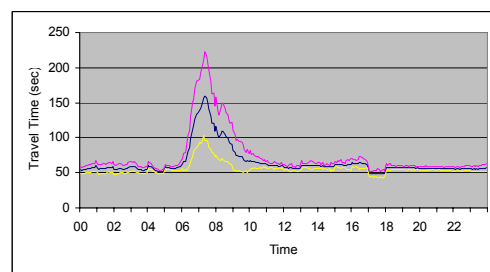
Figure 3.18: Day-to-Day Average Travel Time and Range of Variability
Note: Middle line shows mean travel time; upper and lower lines show ± one standard deviation

56

4.3.2.3 *Relationship Between Travel Time and Day-to-Day Variability*

In general, the standard deviation of travel time is known to be highly correlated with travel time as well as the day-to-day travel time variability. Table 3.8 shows the correlations between day-to-day average travel times and their standard deviations. These correlations were calculated based on 288 data points representing each five-minute intervals. The correlations between travel times and standard deviations are mostly greater than 0.9, and the average correlation for the sample locations is 0.945. The correlations between travel times and the normalized STDs are not as high as those between travel times and STD, but still show fairly high values (0.772 on average).

Table 3.8: Correlation Between Travel Time and Its Day-to-Day Variability

Correlation	vds_id	Correlation	
		Between travel times and their STD	Between travel times and their N-STD
5N	1204532	0.972002	0.894451
5S	1205182	0.938999	0.775226
405N	1201510	0.977827	0.885063
405S	1201805	0.957295	0.867532
22E	1202785	0.977886	0.757644
22W	1202840	0.786967	0.589239
55N	1210172	0.961554	0.899053
55S	1203284	0.964463	0.674607
57N	1202263	0.964427	0.824285
57S	1202422	0.974965	0.872792
91E	1208226	0.951549	0.695800
91W	1208260	0.907648	0.530446
Average		0.944632	0.772178

Note: based on 288 five-minute data

3.3.2.4 *Does Travel Time Variability Differ by Month or by Weekday?*

We have measured day-to-day travel time variability based on all weekday during a one-year period. In this section, we investigate how such day-to-day variability possibly differs by month or by weekday. With the same sections of freeways, firstly, average travel times and their N-STDs are compared. As shown in Figure 3.19, monthly average travel times are similar over the year period. However, the monthly average N-STDs are a little lower during the summer months.

In addition, the variability measures are a bit higher during December compared to other months. This may have been due to many seasonal activities during December.
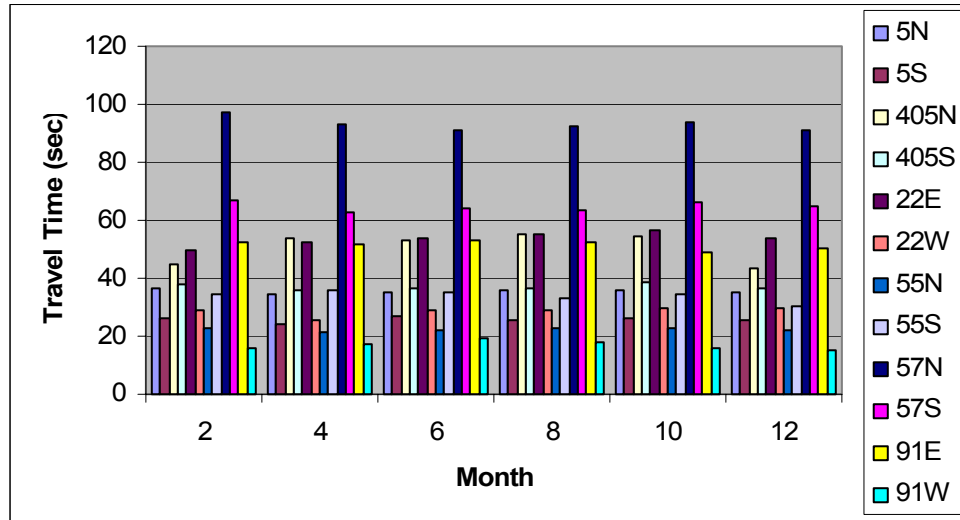
(a) Average Travel Time by Month
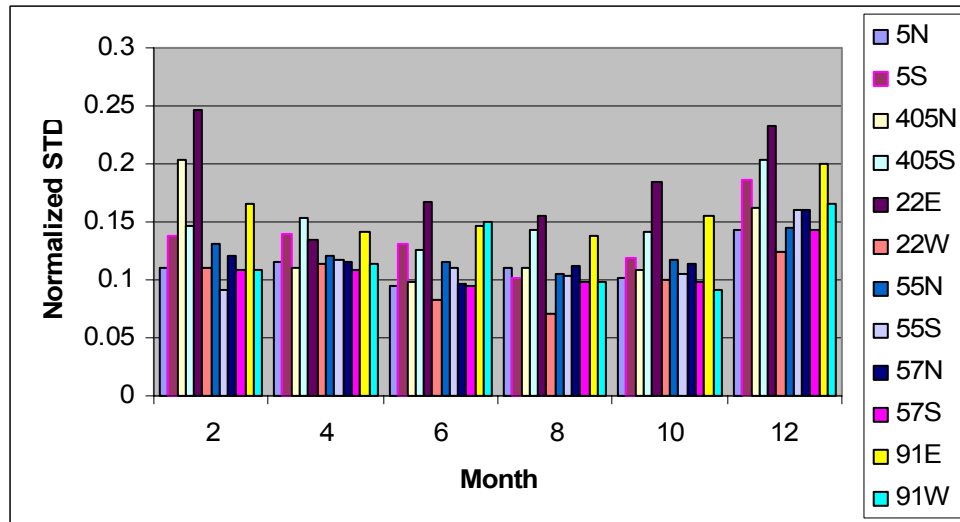
(b) Average Travel Time Variability by Month

Figure 3.19: Travel Times and Their Day-to-Day Variability by Month

Figure 3.20 compares the day-to-day travel time variability during weekdays. In this case, both travel times and their variability do not show significant difference overall. However, SR-91

eastbound shows a slightly longer travel time and higher variation on Friday. This is interpreted as characteristics of SR-91 that attracts heavy traffic during weekend travel.
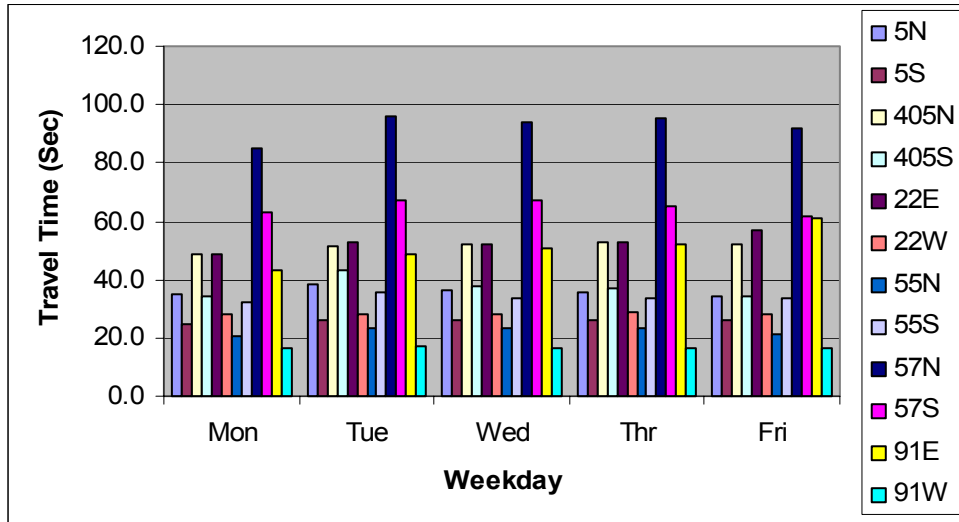
(a) Average Travel Time by Weekday
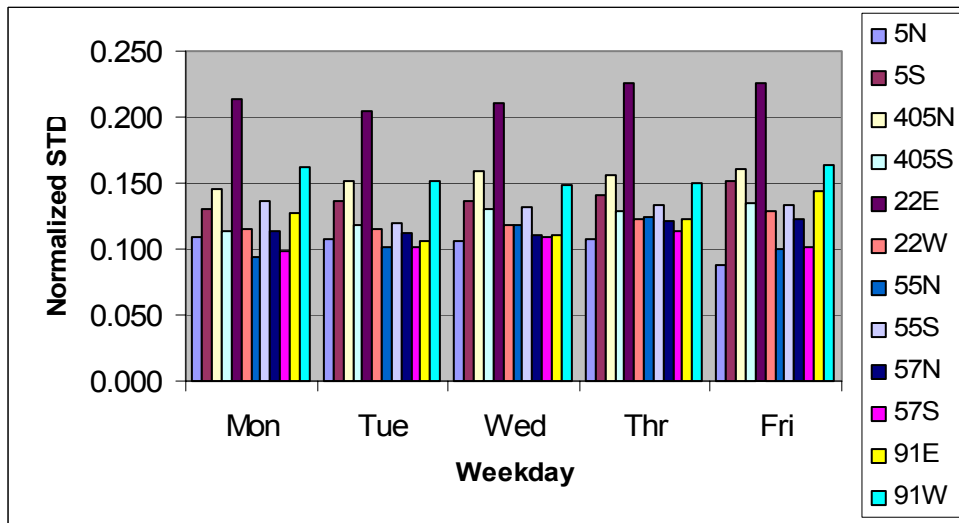
(b) Average Travel Time Variability by Weekday

Figure 3.20: Travel Times and Their Day-to-Day Variability by Weekday

3.3.3 Within-Day Travel Time Variability

3.3.3.1 *Measures of Within-Day Travel Time Variability*

The measures of within-day travel time variability are calculated for an hour interval or for a day with five-minute travel time data. This statistic shows how travel time fluctuates during a given period. In order to compare the difference between sections, daily within-day variability and hourly variability are calculated for each day and averaged over a year period.

As shown in Figure 3.21, each freeway section shows different within-day travel time variability by each time of day. Table 3.9 provides average measures of daily within-day travel time variability and hourly variability during the AM peak, and compares measures of within-day variability with those of day-to-day variability. Of course, the daily measures of within-day variability are greater than those during the AM peak. On average, measures of daily variability are 2.3 times bigger than those of the AM peak's. Measures of daily within-day variability are greater in most sample locations when compared with day-to-day variability, but measures of day-to-day variability are greater than those of within-day during the peak hour. That is, based on hourly measures, day-to-day shows larger variability.

Table 3.9: Within-day Travel Time Variability

Freeway	vds_id	Daily average variability		Variability during 7 - 8	
		Day-to-day	Within-day	Day-to-day	Within-day
5N	1204532	0.1216	0.2798	0.4442	0.2556
5S	1205182	0.1485	0.3332	0.2738	0.2072
405N	1201510	0.1667	0.3783	0.1256	0.0502
405S	1201805	0.1585	0.5408	0.5211	0.2470
22E	1202785	0.2295	0.5843	0.2959	0.1842
22W	1202840	0.1255	0.2008	0.1293	0.0508
55N	1210172	0.1325	0.3662	0.2300	0.3790
55S	1203284	0.1392	0.6357	0.3426	0.3901
57N	1202263	0.1305	0.4642	0.1662	0.1180
57S	1202422	0.1166	0.3666	0.3623	0.2520
91E	1208226	0.1682	0.6880	0.0586	0.0302
91W	1208260	0.1607	0.3637	0.2359	0.1194
Average		0.1498	0.4335	0.2655	0.1903
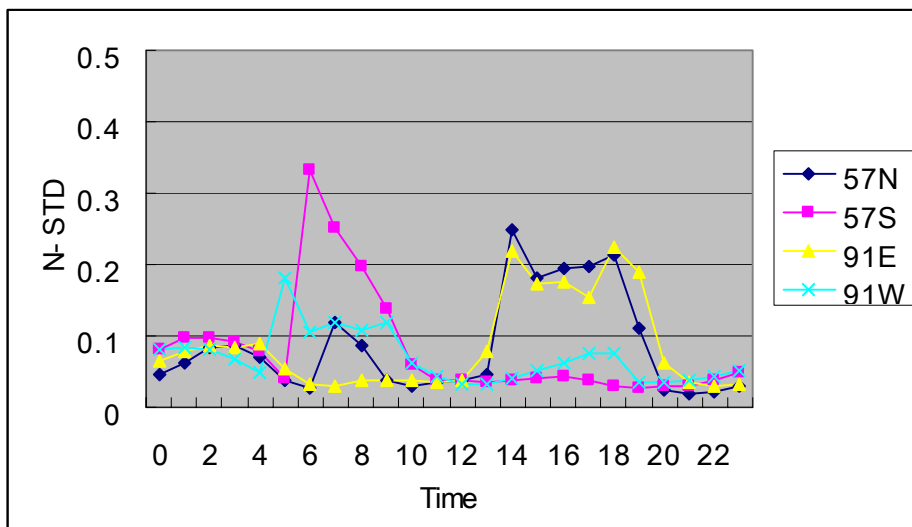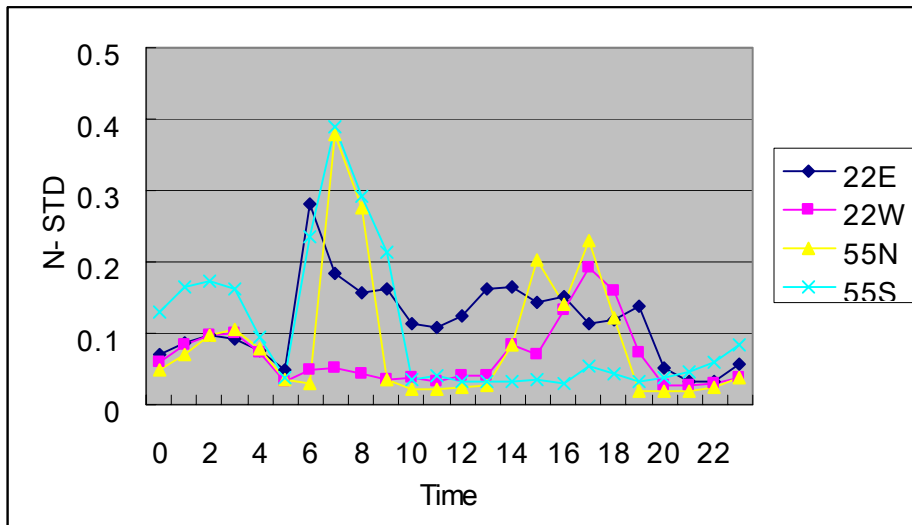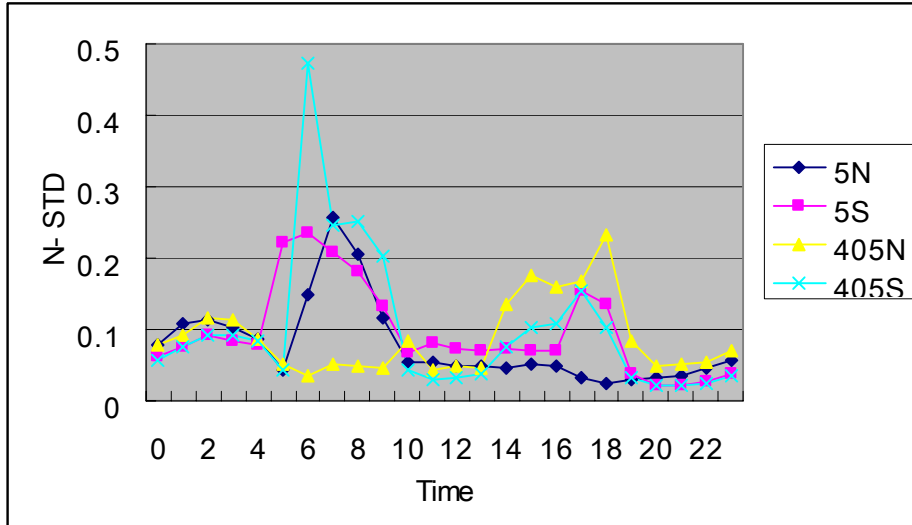
Figure 3.21: Within-Day Travel Time Variability
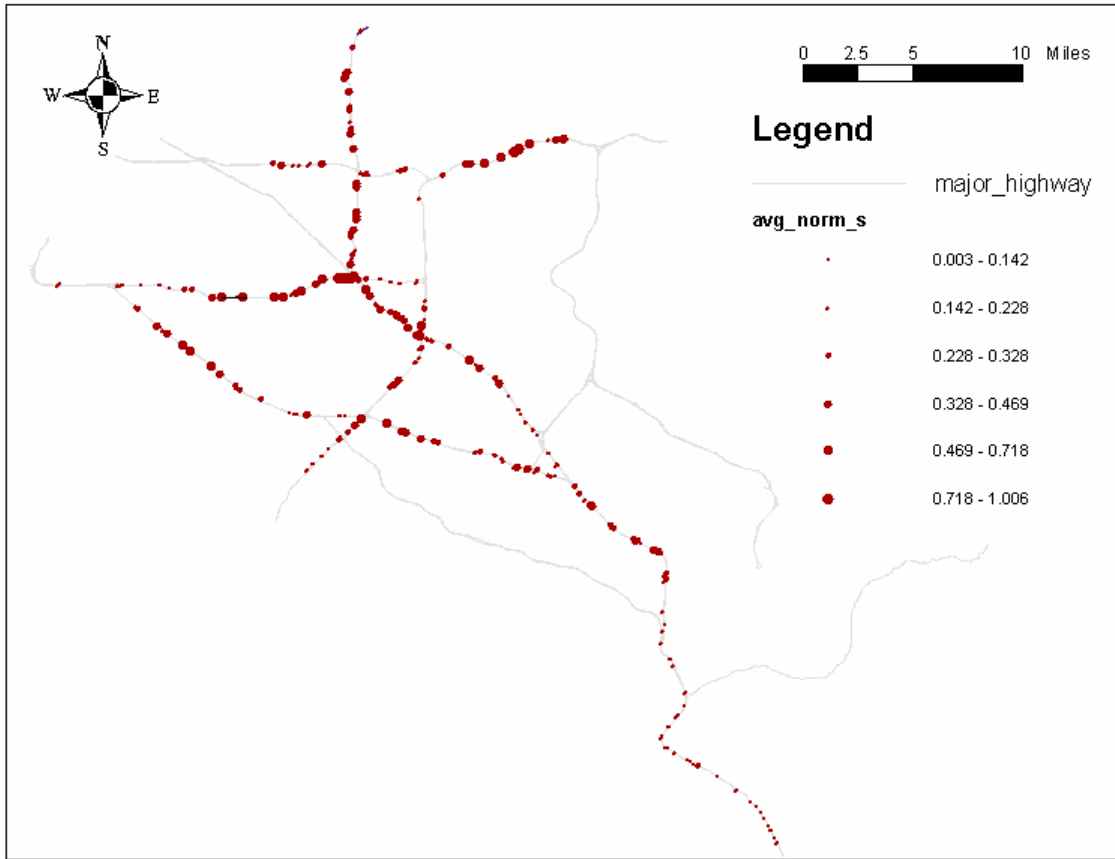
Figure 3.22: Measures of Within-Day Variability in GIS Map

Figure 3.22 shows measures of travel time variability in the GIS map. In the map, the larger dots indicate higher travel time variability. Some portions of I-5, SR-22, and SR-91 show larger travel time variability. Figure 3.23 compares the relationship between day-to-day variability and within-day variability based on one-hour variability measures in normalized standard deviation. Although the relationship varies by freeway section, the overall correlation coefficient for all sample locations is 0.8078. Figure 3.24 shows both day-to-day and within-day travel time variability as one standard deviation range from mean travel time. As addressed before, day-to-day variability shows larger band than with-day variability.
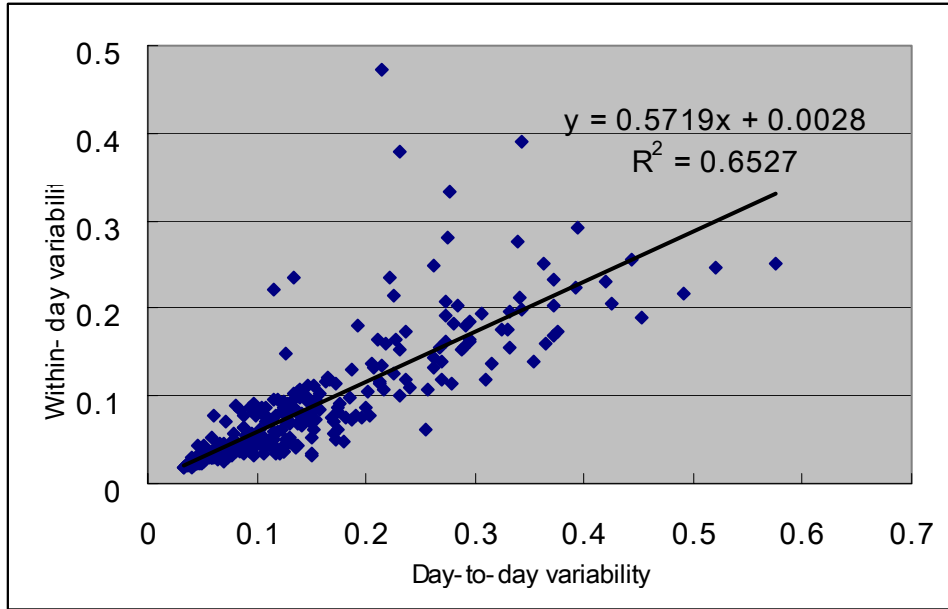
Figure 3.23: Relationship Between Within-Day Variability and Day-to-Day Variability
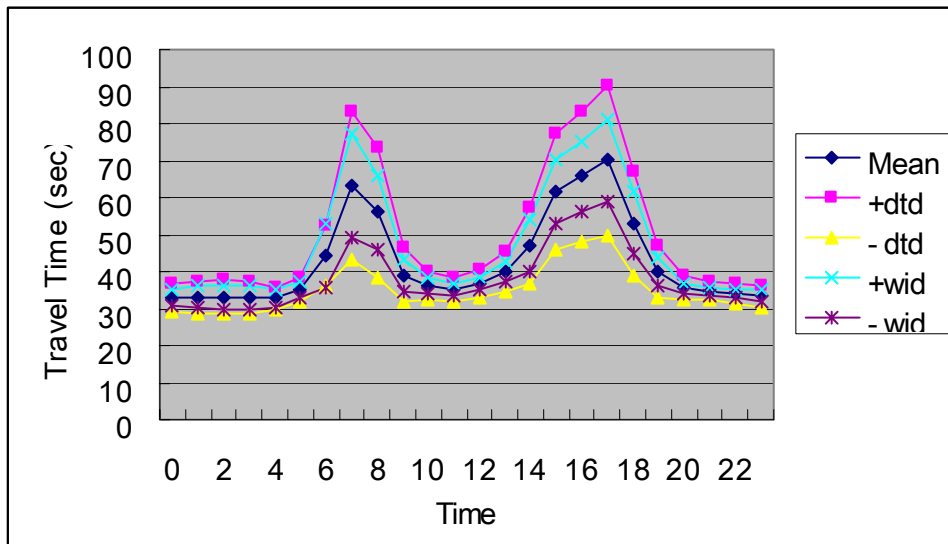
Figure 3.24: Mean Travel Time and Its Variability
Note: one standard deviation range

3.4 Analysis of Route Travel Time Variability

3.4.1 Measures of Route Travel Time Variability

In the previous section, we have investigated on measures of travel time variability within a point or a section of freeway level. However, from the travelers' point of view, the travel time and its variability along their travel route are more important than those at a point or a section. This section deals with such travel time and variability measures in the route level and compares these measures between possible alternatives.

Route travel time variability is measured in three aspects, day-to-day and within-day, as in the section travel time variability and spatial. In this study, the spatial variability is defined as variation of speeds along a route. Such spatial variability addresses the level of smoothness of traffic along a route, and potentially higher spatial variation may lead to unsafe traffic maneuvers. In that regards, such spatial variability could be a measure for drivers to consider their route decision. These three measures of route travel time variability are computed as follow:

Day-to-day variability

$$
\bar{\tau}_{kt} = \frac{\sum_{d \in D} \tau_{kdt}}{N_D}, \quad \sigma^{DTD}{}_{kt} = \sqrt{\frac{\sum_{d \in D} \left(\tau_{kdt} - \bar{\tau}_{kt}\right)^2}{N_D}}, \quad \sigma^{DTD}{}_{k} = \frac{\sum_{d \in T} \sigma^{DTD}{}_{kt}}{N_T}, \tag{3.11}
$$

Within-day variability

$$
\bar{\tau}_{kdP} = \frac{\sum_{t \in P} \tau_{kdt}}{N_P}, \quad \sigma^{WID}{}_{kdP} = \sqrt{\frac{\sum_{t \in P} \left(\tau_{kdt} - \bar{\tau}_{kd}\right)^2}{N_P}}, \quad \sigma^{WID}{}_{kP} = \frac{\sum_{d \in D} \sigma^{WID}{}_{kd}}{N_D}, \tag{3.12}
$$

Spatial variability

$$
\sigma^{S}{}_{kdt} = \sqrt{\frac{\sum_{s \in S_k} \left(u_{dt's} - \bar{u}_{dt}\right)^2}{N_{S_k}}}, \quad \sigma^{S}{}_{k} = \frac{\sum_{d \in D} \sum_{t \in T} \sigma^{S}{}_{kdt}}{N_D \cdot N_T}, \tag{3.13}
$$

where

k = path

d = day

t = departure time

t' = time at the link for the trip departing at time t at the origin

s = section

N_D = the number of days

N_T = the number of time intervals

N_P = the number of time intervals during a given period P

N_{S_k} = the number of sections on path k

τ_{kdt} = Travel time of path k at time t on day d

$\bar{\tau}_{kt}$ = Day-to-day average travel time of route k at time t

$\bar{\tau}_{kdP}$ = Within-day average travel time of route k during period p on day d

σ^{DTD}_{kt} = Standard deviation of day-to-day route travel time on path k at time t

σ^{DTD}_{k} = Average standard deviation of day-to-day route travel time on path k

σ^{WID}_{kdP} = Standard deviation of within-day route travel time on path k during period p on
        day d

σ^{WID}_{kP} = Average standard deviation of within-day route travel time on path k during
        period p

u_{dts} = Speed of section s at time t on day d

$\bar{u}_{dt}$ = Spatial average speed for the trip departing at time t on day d

σ^{S}_{kdt} = Standard deviation of speeds on path k departing at time t on day d

σ^{S}_{k} = Average standard deviation of speeds on path k

3.4.2   Route Travel Time and Variability among Alternative Routes

As described in section 3.2.5.1, we chose 10 zones and enumerated 170 feasible paths using our path enumeration algorithm. Table 3.10 shows path travel times and variability for some interesting origin and destination pairs, that include multiple feasible paths. The percentage of

the fastest among other alternative routes is also included in the table to compare with mean travel time and travel variability measures. The percentage of the fastest is based on all five-minute intervals over the whole one-year period, and the fastest path found from our database is not necessary the most common path by drivers. In most cases, the measures of day-to-day route travel time variability during peak hour is greater than those during the morning peak hour. This information will be precious when coupled with the actual driver's route choice data. Analyzing drivers' route choice behavior, associated with measures of route travel time variability, is of interest. Such variability measures will also be an important factor in travel time information systems especially when much uncertainty exists.

Table 3.10: Route Travel Time and Variability Between Alternatives

Origin	Dest.	Path	Mean travel time (min)	Daily Travel Variability			Variability 7:00 – 8:00 AM		% of the fastest
				Day-to-day	With-in day	Spatial	Day-to-day	With-in day	
1	7	1	20.35	1.4576	1.9635	16.525	3.0921	1.3112	0.8
		2	19.93	1.4330	1.9439	14.940	3.1343	1.3587	25.4
		3	18.75	0.5814	1.6368	9.514	0.7958	0.9604	73.8
2	5	1	15.35	0.8557	1.9484	13.944	2.3266	1.4634	22.8
		2	15.92	0.8413	1.8559	13.343	2.0893	1.0943	0.0
		3	13.97	0.1829	0.3444	9.909	0.2382	0.3533	77.2
2	9	1	15.51	0.9341	2.1919	12.477	2.3741	1.4972	32.9
		2	16.01	0.9170	2.1117	11.483	2.1381	1.1216	0.0
		3	15.84	0.9231	2.1731	17.676	2.1925	1.1800	0.4
		4	17.65	0.9646	2.2426	13.657	2.3813	1.4587	0.0
		5	14.38	0.2565	0.4329	13.727	0.2563	0.3752	66.6
3	10	1	10.97	0.9000	1.3482	28.742	0.6787	0.7837	71.0
		2	11.03	0.8441	1.3547	17.142	0.6347	0.7114	29.1
6	1	1	17.11	0.7946	1.3377	10.896	0.5580	0.4310	32.2
		2	17.86	0.8466	1.9678	12.025	0.5060	0.3501	11.1
		3	17.06	0.7659	1.2282	14.029	0.5798	0.4227	56.7
6	2	1	14.67	0.7218	2.0851	12.132	0.4492	0.2942	34.1
		2	14.33	0.6717	1.4115	10.221	0.5094	0.3778	58.3
		3	14.57	0.6491	1.2214	12.868	0.5303	0.3719	7.6
9	2	1	14.72	0.7759	2.0250	11.160	0.4943	0.2615	79.2
		2	16.58	0.8357	2.2202	11.122	0.5036	0.3008	0.0
		3	18.06	0.8471	2.1987	12.528	0.5152	0.2898	0.0
		4	14.92	0.3779	0.5072	14.595	0.6326	0.3971	20.8

3.4.3 Changes in Route Travel Time and Variability

This section reports an example of travel time and its variability to show details of route travel time analysis. The example route runs from the south to the north along I-5 and I-405.
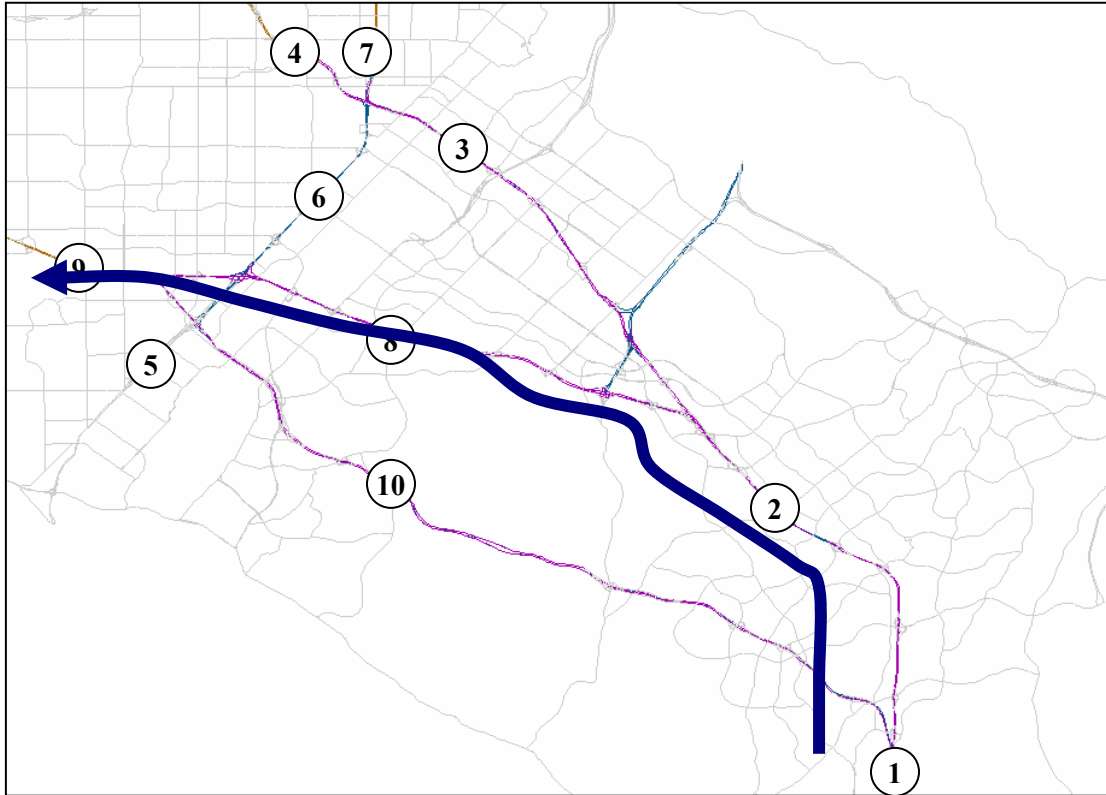
Figure 3.25: Test Route

3.4.3.1 *Changes in Travel Time and Travel Time Distribution by Departure Time*

First, we investigate changes in both travel time and its distribution by departure time. As shown in Figure 3.26, the route travel time changes in respect to time. The overall pattern is similar to changes in section travel time as reported in the previous section. Figure 3.27 shows travel time distributions during different time periods based on travel times during five-minute over the whole one-year period. While the distribution for all day appears as Gamma distribution in shape, each period shows different shapes with different ranges (deviation). For instance, during the morning peak (7 − 8 AM), the range of travel times is wider than other periods.
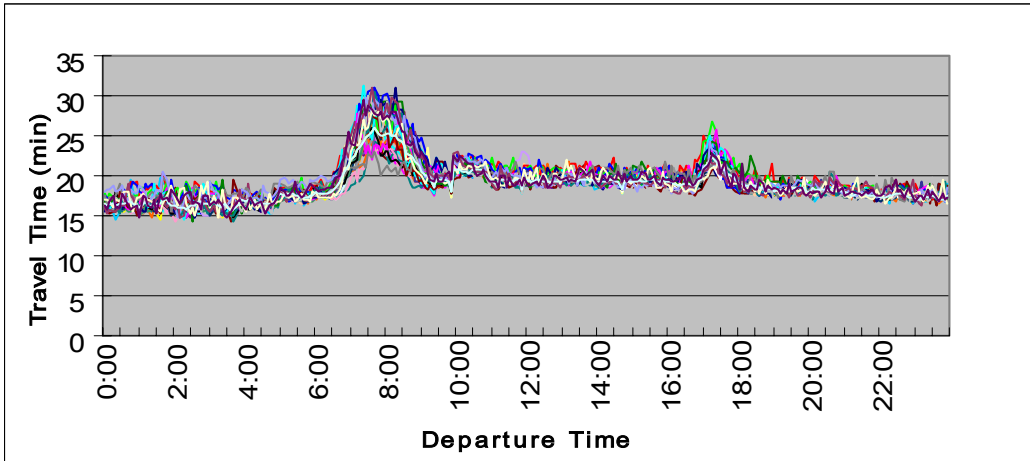
Figure 3.26: Route Travel Time by Departure Time
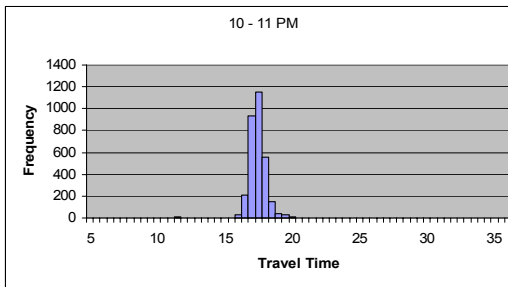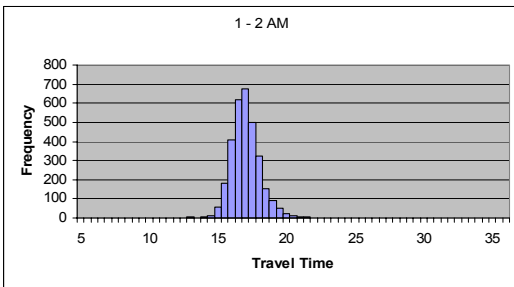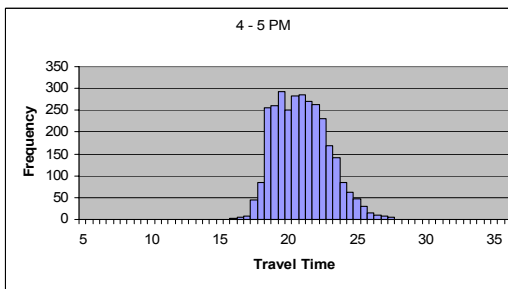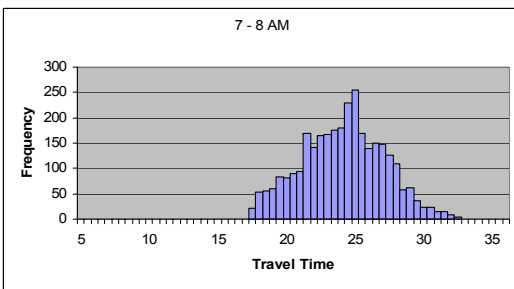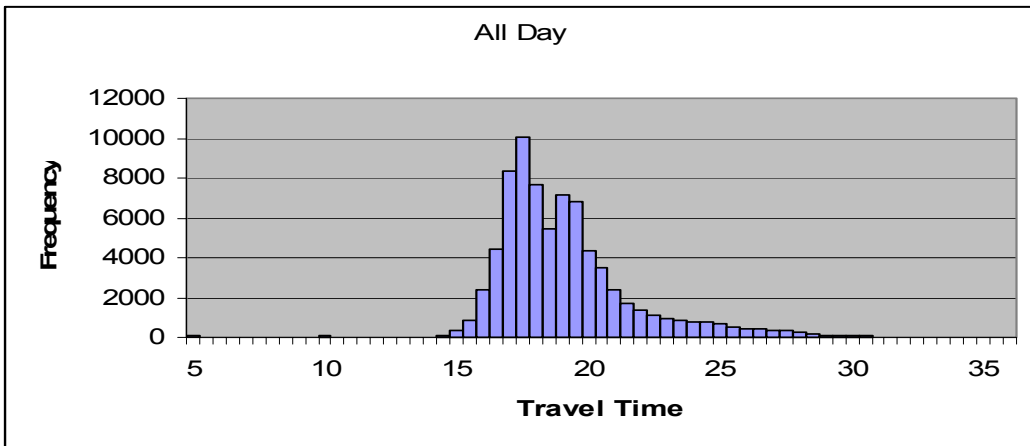
Figure 3.27: Travel Time Distribution by Departure Time Period

68

3.4.3.2 *Day-to-Day Travel Time Variability*

In fact, travel time distributions in the previous section explain characteristics of each route travel time. These distributions are mixture of both day-to-day and within-day variability. In this section, measures of day-to-day travel time variability are reported. Figure 3.28 shows the range of route travel times with ± one standard deviation of each five-minute interval, and Figure 3.29 shows changes in day-to-day variability by departure time. As in the section travel time cases, the route travel time variability is also highly correlated with the travel time itself. Figure 3.30 shows their linear relationship. The correlation coefficient between travel time and its standard deviation (variability) turns was 0.84.
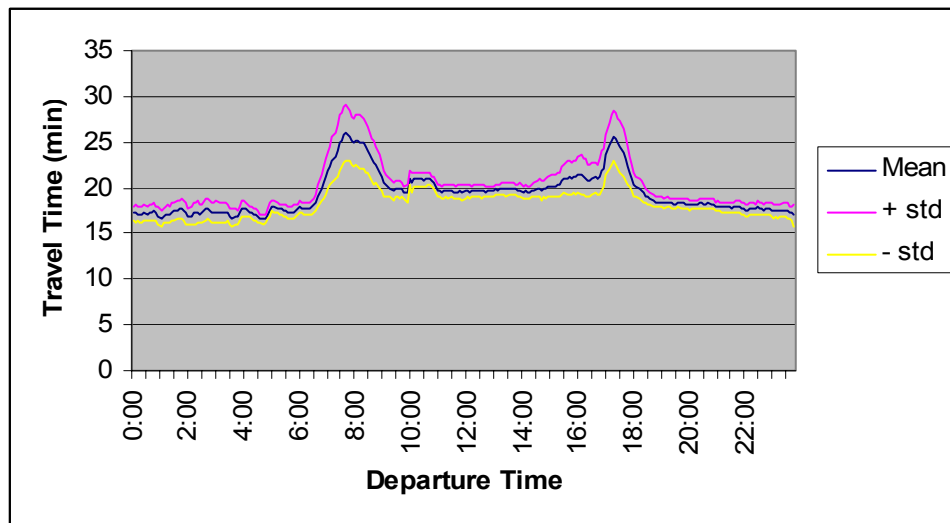
Figure 3.28: Range of Day-to-Day Travel Time by Departure Time
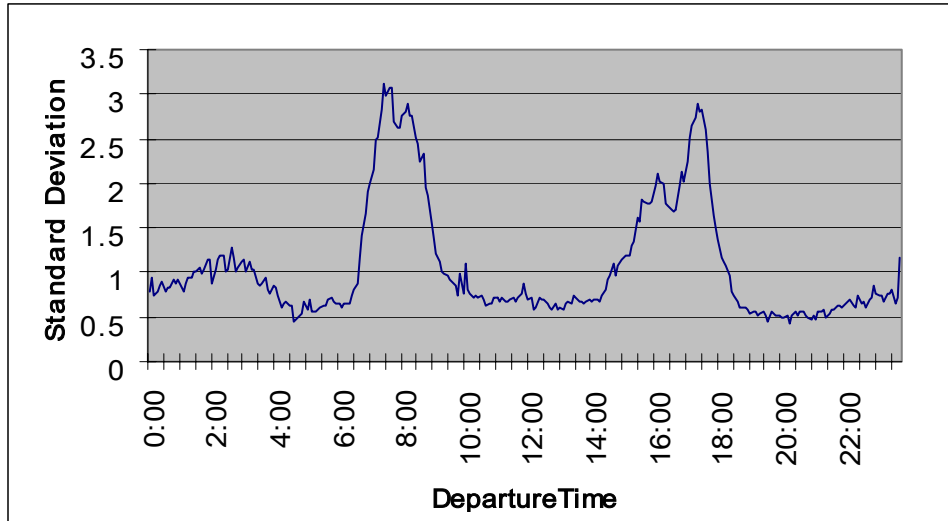
Figure 3.29: Day-to-Day Variability of Travel Time by Departure Time
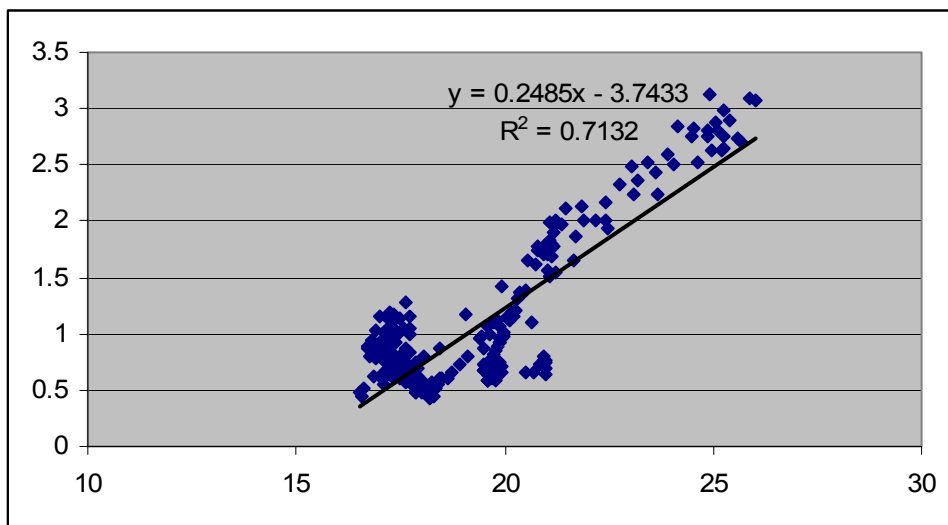
Figure 3.30: Relationship Between Travel Time and Day-to-Day Travel Time Variability

3.4.3.3 *Within-Day Variability of Route Travel Time*

Within-day variability can be measured either by day or by hour. In this section, a one-hour period is applied to compare the within-day variability with the day-to-day variability. In Figure 3.31, both variability measures are based on one-hour period. Although both measures are also highly correlated with each other, as shown in Figure 3.32, the measures of day-to-day variability are greater than those of within-day.
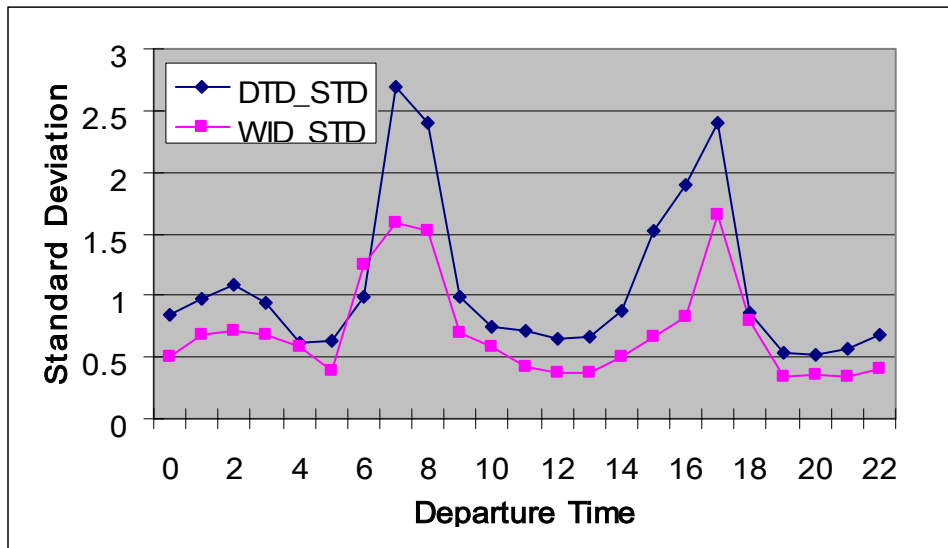
Figure 3.31: Route Travel Time Variability in Both Day-to-Day and Within-Day
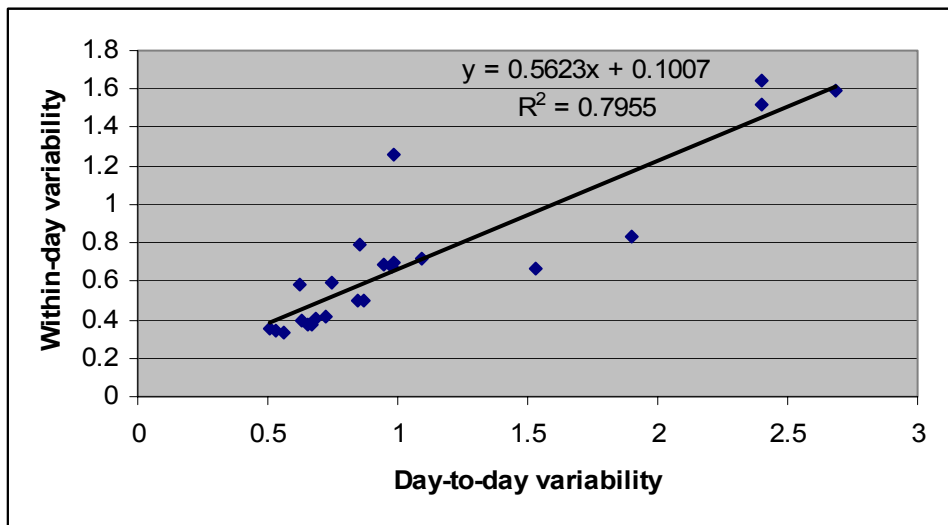
Figure 3.32: Relationship Between Day-to-Day Variability and Within-Day Variability

Table 3.11 shows the relationships among travel time, day-to-day variability, and within-day variability. As shown in Figure 3.32, day-to-day variability is highly correlated with both within-day variability, and travel time. The correlation between travel time and within-day variability is relatively lower than the others.

Table 3.11: Correlation Among Travel Time, Day-to-Day Variability, and Within-Day Variability

	Travel time	Day-to-day variability	Within-day variability
Travel Time	1.00	0.8458	0.7404
Day-to-day variability	-	1	0.8919
Within-day variability	-	-	1

Note: based on one-hour average measures

3.5   Concluding Remarks

Recently the importance of travel time variability has been widely perceived by researchers and practitioners. However, the understanding of travel time variability is still in its infancy and applications to transportation management have not yet been developed. The first step for advancement to the application stage is to build a database for detailed analysis.

This chapter has developed a GIS-based database composed of the historic traffic data from freeways in Orange County, California. The main purpose of the database is to provide the analysis framework to analyze the travel time variability in both section-level as well as route-level. In many cases, traffic data is used for traffic operations without storing the data. The archived traffic data can play an important role as a way of measuring freeway performance. Such historic data enables us to analyze such travel time variability from a long-term perspective. The database developed in this chapter is expected to help researchers and practitioners understand travel time variability as a measure of transportation system reliability by providing the analysis framework. The next step is to identify the sources of travel time variability and to study how to remove the source of variability from the context of transportation system management.

In this chapter, three measures of travel time variability were identified, such as day-to-day variability, within-day variability, and spatial variability. These measures characterize each section of freeway and/or each route between an origin and a destination. While the section-level variability is based on traffic data, the route-level travel variability is more concerned with the traveler's point of view and the affects of variability on their route choice. Further studies are needed to better understand the travel time variability and how to use such measures for transportation management.

Chapter Four

4. MIXED LOGIT ROUTE CHOICE MODEL

Travel time reliability has generally been surmised to be an important attribute of transportation systems. In this chapter, we study the contribution of travel time reliability in travelers' route choice decisions. Traveler's route choice is formulated as a mixed-logit model, with the coefficients in the model representing individual traveler's preferences or tastes towards travel time, reliability and cost. Unlike the traditional approach involving the use of traveler surveys to estimate model coefficients and thereby uncover the contribution of travel time reliability, we instead apply the methodology to real-time loop detector data, and use genetic algorithm to identify the parameter set that results in the best match between the aggregated results from traveler's route choice model and the observed time-dependent traffic volume data from loop detectors. Based on freeway loop data from California State Route 91, we find that the estimated median value of travel-time reliability is significantly higher than that of travel-time, and that the estimated median value of degree of risk aversion indicates that travelers value a reduction in travel time variability more highly than a corresponding reduction in the travel time for that journey. Moreover, travelers' attitudes towards congestion are not homogeneous; substantial heterogeneity exists in travelers' preferences of travel time and reliability. Our results validate results from previous studies involving the California State Route 91 value-pricing project that were based on traditional traveler surveys and demonstrate the applicability of the approach in travelers' behavioral studies.

4.1 Introduction

It is accepted that a wide range of factors influences the route choice of individual travelers. In addition to such factors as perceived travel time, monetary cost, comfort, and safety, the

reliability of travel time has been generally conceded to be an important factor, particularly for trips, such as journey-to-work, where time constraints (e.g. arrival time) may impose significant penalties on an individual. Reliability, by its nature, implies something about the certainty or stability of travel time of any particular trip under repetition. As such, reliability is closely associated with the statistical concept of variability. Variability could result from the differences in the mix of vehicle types on the network for the same flow rates, differences in driver reactions under various weather and driving conditions, differences in delays experienced by different vehicles at intersections, and such random incidents as vehicle breakdown and signal failure, etc. Variability in network travel times introduces uncertainty for travelers in that they do not know with certainty when they will arrive at their respective destinations. This risk (or added cost) to a traveler making a trip may be manifest in a willingness to pay a premium (e.g., through use of toll roads or HOV lane) to avoid congestion and to achieve greater reliability in travel times.

Although travel time reliability ostensibly plays an important role in the traveler's route choice behavior, there has been little empirical work directed to an understanding of the effects of reliability on the route choice decision making of the traveler; many questions remain unanswered. How do travelers value travel time and its reliability, how much does the travel time reliability contribute to travelers' route choices, and how much variation is there in travelers' preferences regarding the potential tradeoff between reliability and travel time itself? Answering these questions can help in the design and evaluation of transportation planning and operation strategies, but requires that this attribute be accounted for explicitly in the modeling of travelers' choice.

The above questions could be studied either through direct or indirect methods (Jackson and Jucker, 1981). The direct method involves posing a series of questions to a sample of travelers in a certain population, usually either as a revealed preference (RP) survey, in which the actual behavioral response to the traffic condition is reported, or as a stated preference (SP) survey, in which a traveler's behavior in hypothetical scenarios is reported. The indirect method involves inferring the answers to these behavioral questions from observed data describing the flows on alternative routes connecting an origin-destination pair.

Most previous research, that has attempted to address issues of reliability using direct methods, has analyzed RP data and/or SP data. Abdel-Aty *et al.* (1996) conducted a study to investigate the effect of travel time variability on route choice using repeated measurement SP data. Their results indicated the significance of both the degree of travel time variation and traffic information on route choice. Bates *et al.* (2001) provided a comprehensive overview of the theory underlying the valuation of reliability, and discussed the empirical issues in data collection. Because of the difficulties of finding real choice situations with sufficient variation, to allow statistically reliable estimates to be obtained for RP data, they acknowledged the value of SP data, and applied SP data in the study of valuation of reliability in passenger rail services. Although SP is usually *de facto* the only realistic possibility for data collection, Lam and Small (2001) leveraged the opportunity of a road pricing project and measured values of travel time and reliability from 1998 RP data on actual behavior of commuters on State Route 91 (SR91) in Orange County, California. Recently, Small *et al.* (2002) continued their previous studies by combining both RP and SP data on SR 91 to empirically identify the varied nature of traveler preference for travel time and reliability. They found that highway users exhibit substantial heterogeneity in their valuation of travel time and reliability.

However, both RP and SP data have drawbacks. Although a greater level of detail data could be obtained from the survey from the individual traveler, and these data may lead to a higher degree of accuracy for the estimation, data collection for the survey of large sample size and its following analysis are time-consuming and expensive. For SP data, it is not sufficient to say that the response to hypothetical situations really reflects traveler's behavioral choice to actual situations. For instance, because people tend to overstate the time delays they actually experienced, it is common that they respond more to a given actual time saving than to a hypothetical time saving of the same amount. Therefore, the estimated value of travel time may be lower than the actual. Although there is a general econometric tradition for favoring RP data, there are often serious problems in achieving the level of detail in data that is ideally required. In the study of travel time reliability, as noted by Bates (2001), it is virtually impossible to find RP situations where there is sufficient perceived variation to allow statistically reliable estimates -- notable exceptions to this are the studies conducted by Lam and Small (2001) and Small *et al.* (2002) with the California State Route 91 value-pricing project.

Alternatively, advancements in traffic surveillance and monitoring technologies, including real-time data from inductive loop detectors, can provide valuable aggregated information that ostensibly resulted from the disaggregated individual travel route choices. Instead of surveying motorists on their choices in the direct method, we propose an indirect method to answer some behavioral questions, in such a way, that the estimated choice probability resulting from a route choice model matches the revealed probability exhibited in the real-time loop detector data. To the best of our knowledge, there are but very few research studies that investigate behavioral issues using the indirect method, and this chapter aims to fill this gap.

In this chapter, traveler's route choice is formulated as a mixed-logit model (also known as random coefficient logit and logit kernel), which generalizes the standard logit model by allowing the coefficient associated with each observed variable to vary randomly across individuals (McFadden and Train, 2000; Bhat, 2001; and Bhat and Castelar, 2002). The coefficients in the model represent individual traveler's preferences or tastes toward travel time, reliability and cost. To find the distribution of the model coefficients and thereby uncover the contribution of travel time reliability in the dynamic route choice, we use Genetic Algorithm (GA) to identify the parameter set that results in the best match between the aggregated results from traveler's route choice model and the observed time-dependent traffic volume data from loop detectors.

We apply the proposed approach to measure value of time (VOT), value of reliability (VOR), and degree of risk aversion (DORA) simultaneously using data on actual travel behavior drawn from a real pricing context. A recent value pricing project on a major commuting highway, State Route 91 (SR91) in Orange County, California, gives travelers the option to travel free on the regular lanes or to pay a time-varying price for express travel on toll lanes situated along the median of the highway. Based on their respective choices of whether or not to pay a toll for the congestion-free travel, we observe the outcome from the choice probability between the two parallel routes in the form of loop detector data on thirty-second averages of count and occupancy. We find that travel time reliability plays an important role in traveler's decision making for route choice. Moreover, the results indicate that travelers value travel-time reliability substantially higher than they do travel-time savings. Our results validate results from such

previous studies as Lam and Small (2001) and Small *et al*. (2002), and demonstrate the applicability of our approach in travelers' behavioral studies.

4.2 Route Choice Structure and Formulation

Travelers' route choice among the available options will reflect their perception of the costs and benefits associated with each option. If the costs or benefits are perceived to be uncertain, the choice will be influenced by the travelers' attitude to that uncertainty. We incorporate the stochasticity of route travel time as a measure of the risk associated with the selection of specific routes. On the basis of the perceived distribution of network travel times, travelers are assumed to behave differently when considering routes for which their perceived travel times have a probabilistic component. Some are risk averse, choosing routes with longer expected travel times but smaller variations. Others, the risk takers, may choose routes with shorter expected travel times but greater variations in travel time reliability.

In modeling this behavior, we assume that the individual traveler has a subjective probability distribution of travel time for each available route. Additionally, we assume that there exists an objective distribution of travel time based on actual measurement over a suitably defined time period. This stochastic travel time reflects the intrinsic fluctuations in the transportation network, as noted, due to particular weather conditions, unpredictable lane closures, traffic accidents, etc. It is not necessarily the case that the subjective and objective probability distributions over route travel time are identical. They may differ, and when this is the case, the existence of perception errors is witnessed). However, for travelers in peak-hour commuting trips, it seems reasonable to assume that these two distributions are the same, i.e., drivers have no misperceptions of either travel time or travel time variation. This is the case for the model considered here; we leave the incorporation of traveler's perception errors for future studies.

We assume that travelers consider travel time, travel-time reliability (i.e., risk), and out-of-pocket monetary cost (such as toll) in their choice of route. Moreover, we assume that they value any tradeoff between travel time and travel-time reliability differently depending on individual tastes, and that such tastes are distributed in the population in a manner that covers a spectrum in

terms of the degree of risk aversion. We further assume that drivers are rational and are maximizing some utility measure, and suggest a common disutility functional form, but with coefficients (disutility weights) that reflect individual's preference. Traveler's perceived disutility is a function of the route travel time, travel time variability, monetary cost, and the individual traveler's attitude toward these three variables at each time. Specifically, we assume that the traveler is faced with a choice among P_{rs} alternative freeway routes between a freeway on-ramp origin r and a corresponding freeway off-ramp s. In this formulation, we assume that surface street travel is irrelevant to the choice and that the freeway origin and destination are fixed. There exists a choice set of route alternatives that is known to the individual, consisting of R different routes. Under these assumptions, the disutility to traveler n of travel, commencing at time t, along path p linking on-ramp origin r and off-ramp destination s is specified as:

$$U_{np}(t) = \beta_n' x_{np}(t) + \varepsilon_{np}, \tag{4.1}$$

where $x_{np}(t)$ is a vector of observed variables (including alternative specific constants), β_n is a corresponding coefficient vector that may vary over individuals but does not vary across alternatives or time, and ε_{np} is an unobserved extreme value random term that captures the idiosyncratic effect of all omitted variables that are not individual specific. ε_{np} is assumed to be identically and independently distributed across all choice occasions and independent of $x_{np}(t)$ and β_n.

Since we assume that the traveler's subjective distribution is identical to the objective distribution of route travel time, we omit n from the subscripts of $x_{np}(t)$. Therefore, $x_p(t)$ is a vector of observed variables based on the actual measurements from the field. Specifically,

$$x_p(t) = [T_p(t), R_p(t), C_p(t)]', \tag{4.2}$$

where $T_p(t)$ measures route travel time, $R_p(t)$ measures travel-time reliability, and $C_p(t)$ is the monetary toll cost. Therefore, the value of travel time (VOT), value of travel-time reliability (VOR), and the degree of risk aversion (DORA) are defined as:

$$VOT_n = \frac{\partial U_{np}(t)/\partial T_p(t)}{\partial U_{np}(t)/\partial C_p(t)} = \frac{\beta^T}{\beta^C};$$ (4.3)

$$VOR_n = \frac{\partial U_{np}(t)/\partial R_p(t)}{\partial U_{np}(t)/\partial C_p(t)} = \frac{\beta^R}{\beta^C};$$ (4.4)

$$DORA_n = \frac{\partial U_{np}(t)/\partial R_p(t)}{\partial U_{np}(t)/\partial T_p(t)} = \frac{\beta^R}{\beta^T}.,$$ (4.5)

where $\beta_n = [\beta_n^T, \beta_n^R, \beta_n^C]'$, a vector of coefficients reflecting individual n's particular tastes toward travel time, reliability, and monetary cost. As the notation indicates, the models we consider are specified so that VOT, VOR, and DORA depend on the individual traveler n but not on the choice instant t. $DORA_n$ reflects the degree of risk aversion, i.e., the extent to which travel time variability is undesirable to traveler n. The larger the value of DORA, the higher the perceived cost of uncertainty, and the more risk averse the traveler.

Preference heterogeneity is introduced by assuming that the coefficients β_n are random variables. These coefficients are assumed to vary over travelers based on individual characteristics in the population with density $f(\boldsymbol{\beta})$. This density is a function of parameters $\boldsymbol{\Theta}$ that represent, for example, the mean and the covariance of the β's in the population. This specification, known as "mixed" logit or "random coefficients" logit, is identical to the standard logit except that $\boldsymbol{\beta}$ varies over decision makers rather than being fixed. As such, the probability that traveler n will select path p, conditioned on $\boldsymbol{\beta}_n$, is given by:

$$L_{np}(\boldsymbol{\beta}_n; t) = \frac{e^{\boldsymbol{\beta}' \mathbf{x}_{np}(t)}}{\sum_{\forall j \in P} e^{\boldsymbol{\beta}' \mathbf{x}_{nj}(t)}}$$ (4.6)

The unconditional probability is the integral of $L_{np}(\boldsymbol{\beta}_n; t)$ over the distribution of all possible values of $\boldsymbol{\beta}_n$, i.e.,

$$P_{np}(t) = \int \frac{e^{\boldsymbol{\beta}' \mathbf{x}_{np}(t)}}{\sum\limits_{\forall j \in P_{rs}} e^{\boldsymbol{\beta}' \mathbf{x}_{nj}(t)}} \cdot f(\boldsymbol{\beta}) d\boldsymbol{\beta} \qquad (4.7)$$

Equation (4.7) is the general form for the so-called Mixed Logit probability. Train (2002) provides an excellent overview of the properties of such models and procedures for their estimation. Choice probabilities can be estimated using Monte-Carlo simulation to integrate the computationally difficult parts of the preference distribution. Assuming that the parameters of $f(\boldsymbol{\beta})$ are $\boldsymbol{\Theta}$, the unbiased estimator $\breve{P}_{np}$ for P_{np} can be obtained as follows:

1. Select trial values for $\boldsymbol{\Theta}$.

2. Draw Q values of $\boldsymbol{\beta}$ from $f(\boldsymbol{\beta}|\boldsymbol{\Theta})$; label the qth such value $\boldsymbol{\beta}^q$

3. Calculate $L_{np}(\boldsymbol{\beta}^q; t)$; $q = 1, 2, \ldots, Q$

4. Compute average simulated probability as:

$$\breve{P}_{np}(t) = \frac{1}{Q} \sum_{q=1}^{Q} L_{np}(\boldsymbol{\beta}^q; t) \qquad (4.8)$$

An issue of terminology arises with Mixed Logit models (Train, 2002). There are two sets of parameters in a Mixed Logit model. First, we have the parameters $\boldsymbol{\beta}$, which enter the logit formula. These parameters have density $f(\boldsymbol{\beta})$. The second set of parameters describes this density. For example, if we assume that $\boldsymbol{\beta}$ is normally distributed with mean b and covariance W, then b and W are parameters that describe the density $f(\boldsymbol{\beta})$.

In this section, we assume that the parameters $\boldsymbol{\beta}_n$ have independent normal distributions. If we further place the traveler in one of the M groups defined by their access to information (the assumption being that more risk adverse travelers may be more inclined to search out travel-time information), we would specify $\beta \sim N(b_m, W_m); m \in M$. With the assumption that the random parameters β^T, β^V, β^C for time, reliability and cost have normal distributions, i.e., $\beta^T \sim N(b^T, W^T)$, $\beta^V \sim N(b^V, W^V)$, $\beta^C \sim N(b^C, W^C)$, the parameter sets that need to be

estimated are $\{b^T, W^T, b^V, W^V, b^C, W^C\} \subset \Theta$. The estimation of Θ will define VOT, VOR, and DORA, and thereby uncover the relative roles of travel time and travel time reliability in route choice, as well as identify the distribution of the population along the risk dimension. We let Ω, called the parameter space, denote the set of all possible values that parameters Θ could assume. The estimation of Θ is to search Ω and find the best Θ satisfying certain criteria. Traditionally, these parameters in the Mixed Logit model are estimated based on RP and/or SP data by simulated maximum likelihood estimation (Small *et al*, 2002). However, since our observed data are aggregated responses in the form of loop counts rather than the individual responses of each traveler, we use a genetic algorithm to identify the parameter set that results in the best match between the aggregated results from traveler's route choice model and the observed time-dependent traffic volume data from loop detectors. We provide details of this procedure in the next section.

4.3   Estimation Procedure and Solution

Traditionally, the parameters in the mixed logit model are estimated based on RP and/or SP data by simulated maximum likelihood estimation (SMLE), as described in Train (2002). However, since our observed data are aggregated responses in the form of loop counts rather than the individual responses of each traveler, the essential concept of our estimation procedure is to find the set of parameter values that results in the best match between the aggregated results from traveler's route choice model and the observed time-dependent traffic volume data from the loop detectors. Therefore, the problem considered here is a minimization program of the difference between the volume data generated from the mixed logit route choice model (we assume the dynamic origin-destination matrix is given) and the observed loop counts. Because the route choice probability in the minimization program does not admit a closed form, as shown in Equation (4.7), gradient-based optimization methods require expensive computational effort to calculate the derivatives numerically and often result in finding a local optimal solution. We therefore adopt genetic algorithm to solve the minimization program.

4.3.1 Estimation

Consider freeway trips originating from origin O located at on ramp r during time interval $t_o - \Delta t_o$. The total number of such trips is evident from the thirty-second loop counts Q at r as:

$$\bar{Q}_r(t_o) = \sum_{j=0}^{\Delta t_o/30} Q_r(t_o - j).$$

(4.9)

Since the dynamic O-D matrix $O(t) \to D(t)$ is presumed to be known, the total number of trips originating at $O = r$, and bound for destination $D = s$, during time interval $t_o - \Delta t_o$ is given as:

$$\bar{Q}_{rs}(t_o) = \frac{\sum_{j=0}^{\Delta t_o/30} O_{rs}(t_o - j)}{\sum_{\forall k \in S} \sum_{j=0}^{\Delta t_o/30} O_{rk}(t_o - j)} \cdot \bar{Q}_r(t_o) \; ; \; S = \{k \in \text{ set of all off ramps}\}.$$

(4.10)

From Equation (4.7), the expected number of these trips to use any path p is given by:

$$\bar{Q}_{rs}^p(t_o) = P_{rs}^p(t_o) \cdot \bar{Q}_{rs}(t_o),$$

(4.11)

where

$$P_{rs}^p(t_o) = \int \frac{e^{\boldsymbol{\beta}' \mathbf{x}_p(t_o)}}{\sum_{\forall j \in P_{rs}} e^{\boldsymbol{\beta}' \mathbf{x}_j(t_o)}} \cdot f(\boldsymbol{\beta}) d\boldsymbol{\beta}.$$

(4.12)

For any loop station i on the path $p = \{r, 1, 2, \ldots i, \ldots s\}$, we denote the travel time from origin r to loop station i for trips starting at time t_0 as $t_i^*(t_o)$. Then the expected time at which the flow contribution from $\bar{Q}_{rs}^p(t_o)$ first will be counted (i.e., show up in the loop count poll) is $t_i^*(t_o)$; the expected end of the contribution from $\bar{Q}_{rs}^p(t_o)$ will occur at $t_i^*(t_o + \Delta t_o)$, or $\Delta t_o/30$ polls later. For purposes of identification, we expand the notation on $t_i^*(t_o)$ to include reference to path p from r to s, i.e.,

$$t_i^*(t_o) \to t_i^*(t_o; p_{rs}),$$

(4.13)

where p_{rs} denotes path p from on-ramp r to off-ramp s. Consider the observed loop count at some station i on the path p_{rs} over the time interval $t - \Delta t_o$, i.e.,

$$\hat{Q}_i(t) = \sum_{j=0}^{\Delta t_o / 30} Q_i(t - j). \tag{4.14}$$

Let

$$H(t_i^*(t_o, p_{rs}), \Delta t_o) = \begin{cases} 1, \ t_i^*(t_o, p_{rs}) - \Delta t_o \le t \le t_i^*(t_o, p_{rs}) \\ 0, \ \text{otherwise} \end{cases}. \tag{4.15}$$

Then, an estimate of $\hat{Q}_i(t)$ is given by:

$$\hat{\hat{Q}}_i(t) = \sum_{\substack{\forall t_o < t \\ i \in p_{rs}}} \sum_{\forall p_{rs}} H(t_i^*(t_o, p_{rs}), \Delta t_o) \cdot \bar{Q}_{rs}^p(t_o) \tag{4.16}$$

or

$$\hat{\hat{Q}}_i(t) = \sum_{\substack{\forall t_o < t \\ i \in p_{rs}}} \sum_{\forall p_{rs}} H(t_i^*(t_o, p_{rs}), \Delta t_o) \cdot P_{rs}^p(t_o) \cdot \bar{Q}_{rs}(t_o). \tag{4.17}$$

Alternatively, using the estimate for $P_{rs}^p(t_o)$ given by Equation (4.7),

$$\hat{\hat{Q}}_i(t) = \sum_{\substack{\forall t_o < t \\ i \in p_{rs}}} \sum_{\forall p_{rs}} H(t_i^*(t_o, p_{rs}), \Delta t_o) \cdot \breve{P}_{rs}^p(t_o) \cdot \bar{Q}_{rs}(t_o). \tag{4.18}$$

$\hat{\hat{Q}}_i(t)$ is a function only of known values and the unknown parameters of $f(\beta|\Theta)$. A standard approach to selecting the values of Θ is to minimize the mean square error (MSE) between the estimate $\hat{\hat{Q}}_i(t)$ and its true (observed) value $\hat{Q}_i(t)$ over some specified time period $t_1 \le t \le t_2$, i.e.,

$$\min_{\Theta} (MSE) = \int_{t_1}^{t_2} \sum_{\forall i} [\hat{Q}_i(t) - \hat{\hat{Q}}_i(t)]^2 dt \tag{4.19}$$

or, for distinct 30-second counting intervals,

$$\min_{\Theta} (MSE) = \sum_{t=1}^{t=\frac{t_2-t_1}{30}} \sum_{\forall i} [\hat{Q}_i(t) - \hat{\hat{Q}}_i(t)]^2 . \tag{4.20}$$

4.3.2 Solution Method

Our solution method to Equation (4.20) is based on genetic algorithms (GAs). GAs are heuristic search algorithms that attempt to search the solution space in a "smart" manner on the basis of natural selection and natural genetics. When using GAs to solve an optimization problem, each solution is encoded in a string (called chromosome), which is the concatenation of sub-strings corresponding to the set of decision variables. The entire population of such strings (solutions) is called a generation. Operators such as selection, crossover, and mutation are applied to parent chromosomes to create child chromosomes. The performance of each chromosome is evaluated by a fitness function, which corresponds to the objective function of the optimization program. Chromosomes that have high fitness values have high opportunities to reproduce, by cross-breeding with other chromosomes in the population. Detailed discussions of general GAs are available in Goldberg (1989).

The convergence of GAs has been proven by Holland (1975). Although not guaranteed to find the optimal solution, GAs often are successful in finding a solution with high fitness. GAs are also considered robust because at any time step of a search (or generation), GA progresses towards the optimal solution from a population of points, instead of starting the search at a single point, which increases the likelihood that the global, rather than a local, optimum will be found (Gen and Cheng, 2000). The population-based search procedure, together with stochastic operators used in GAs for reproducing child chromosomes in the next generation, are essential concepts for GAs to locate better solutions for complex and noisy objective functions than do such conventional techniques as gradient-based search methods.

The GA procedure used in the estimation of parameter set Θ that satisfies the condition defined by Equation (4.20) is summarized in Figure 4.1. The first step in the estimation procedure is to generate a number of random individuals as the initial population, each carrying a chromosome that represents a feasible solution. From the initial population, each chromosome is first decoded into the actual parameter values and fed into the mixed-logit dynamic route choice model described in Section 4.2. Since the O-D matrices are given, traffic assignment based on the mixed-logit route choice model can be preformed. The fitness function, in which the objective function of Equation (4.20) is embedded, is then evaluated using the estimated volumes from the route choice model and the observed volumes from the field data. If the stopping criterion is not met, a set of GA operators (including selection, crossover and mutation) are applied to the chromosomes in the current population to produce offspring. The reproduction cycle including decoding of chromosomes and fitness evaluation is repeated until the stopping criterion is met or the predetermined number of generations is reached. The major components in this genetic-algorithm-based estimation procedure, including encoding and decoding genetic chromosomes, evaluating the fitness of each chromosome, and reproducing child chromosomes by selection, crossover and mutation, are described in the following.

GA (*fit_threshold, max_generation, p, r, m*)

   fit_threshold: a threshold specifying the termination criteria.
   max_generation: maximal generations will be performed in the procedure.
   p: the number of chromosomes (solutions) to be included in the population.
   r: the fraction of population to be replaced by crossover at each generation (step).
   m: the mutation rate.

1. Initialize population G: Generate p chromosomes at random.

2. Decode each chromosome in the population G.

3. Evaluate: Compute fitness for each chromosome in the population G.

4. While (fitness of best chromosome is less than fit_threshold or

       the number of generations is less than max_generation) do

   Create a new generation of chromosomes, G_{new}:

- Select: Probabilistically select *(1-r)*p* members from the current population G, and add to the new generation G_{new}.

- Crossover: Probabilistically select *(r*p)/2* pairs of chromosomes from the current population G. For each pair, produce two offspring by applying the crossover operators. Add all offspring to the new generation G_{new}.

- Mutate: Choose *m* percent of the new generation G_{new} with uniform probability. Apply the mutation operator.

- Update: $G = G_{new.}$

- Evaluate: Compute fitness for each chromosome in the new generation.

- Find the highest fitness chromosome.

5. Decode the chromosome with the best fitness and obtain the best solution to this problem.

Figure 4.1: GA-based Estimation Procedure

4.3.2.1 *Encoding and Decoding Chromosomes*

To apply GA to a given problem, a suitable encoding scheme for the parameter set must first be determined. Of the various encoding methods that have been proposed, the most popular representation structures are binary vector and floating vector (Gen and Cheng, 2000). Because of its ease of implementation, a binary encoding method was applied to represent the parameter set in this research.

In binary encoding, each decision variable is represented by a binary substring, and these substrings are concatenated to form a longer string, i.e., the chromosome, which represents the set of decision variables. The length of the substring is determined by the range of the decision variable and the level of precision. Let x be the real-valued decision variable and let it have a domain $[x_{min}, x_{max}]$. The length of the binary string used to represent x can be obtained from:

$$L = INT(\log_2(\frac{x_{max} - x_{min}}{D} + 1) + 1),$$ (4.21)

where L is the length of binary string, D is the desired precision of variable x, and *INT* is the truncate operation to convert a real number into an integer. Then the decoding from a binary string to a real variable x is computed as:

$$x = x_{min} + (x_{max} - x_{min}) * \frac{A}{2^L - 1},$$ (4.22)

where A is the value of binary string base 10.

To identify x_{min}, x_{max} and D, a preliminary analysis and understanding of variable x is required. Specifically, if variable x is very sensitive, the desired precision needs to be higher. As a result, a longer string would be used in the representation scheme for variable x.

In our study, the parameter sets that need to be estimated are $\{b^T, W^T, b^V, W^V, b^C, W^C\} \subset \Theta$, which contain 6 decision variables. Each parameter is represented by 6 binary bits, so the

chromosome which represents the parameter set will have 36 binary bits. An example of genetic representation is shown in Figure 4.2.

100010	001001	010101	100001	000001	111000
b^T	W^T	b^V	W^V	b^C	W^C

Figure 4.2: An Example of Genetic Representation of 6 Decision Variables with 36 Bits

4.3.2.2 *Fitness Function*

A fitness function is required to measure the "goodness" of each chromosome. The fitness function used here is the linear scaling of the objective function in the minimization program of Equation (4.20), which represents how well the estimate volume $\hat{\hat{Q}}_i(t)$ and its true (observed) value $\hat{Q}_i(t)$ match, as shown in Equation (21):

$$f_k = a \sum_{t=1}^{t=\frac{t_2-t_1}{30}} \sum_{\forall i} [\hat{Q}_i(t) - \hat{\hat{Q}}_i(t)]^2 + b, \tag{4.23}$$

where f_k is the fitness value of the kth chromosome and a and b are the scaling factors. Linear scaling is introduced to avoid two significant difficulties in the fitness proportionate selection process: premature convergence termination at early generations, and stalling at late generations (Goldberg, 1989). Parameters a and b are selected so that the average fitness is mapped to itself and the best fitness is increased by a designed multiple of the average fitness.

4.3.2.3 *Selection Process*

Selection is an operation through which chromosomes are picked for reproduction with a probability proportional to their fitness. In this study, a combination of fitness proportionate selection and elitism strategy is adopted for the reproduction process. In the fitness proportionate selection, also called roulette wheel selection, the probability that a chromosome will be selected is given by the ratio of its fitness to the fitness of the entire population, as shown in the following:

$$q_k = \frac{f_k}{\displaystyle\sum_{m=1}^{pop\_size} f_m},$$

(4.24)

where q_k is the probability of selecting chromosome k to produce offspring, f_k is the fitness value of the kth chromosome in the current generation and *pop_size* is the population size.

The elitism strategy keeps a certain number of the top chromosomes that have the highest fitness values and propagates to the next generation. This procedure ensures that the best solution in the next generation is not worse than the one in the current generation.

4.3.2.4 *Crossover Operation*

The crossover operator produces two new offspring from two parent strings, by copying selected bits from each parent. The bit at position i in each offspring is copied from the bit at position i in one of the two parents. The choice of which parent contributes the bit for position i is determined by an additional string called the crossover mask. With the crossover mask, crossover operation can be performed at single-point or two-point, as illustrated in Figure 4.3. In our study, both single-point crossover and two-point crossover are used.

	Parent Strings	Crossover Mask	Offspring
Single-point crossover			
	1110100100	1111100000	1110101010
	0000101010		0000100100
Two-point crossover			
	1110100100	0011110000	1100100100
	0000101010		0010101010
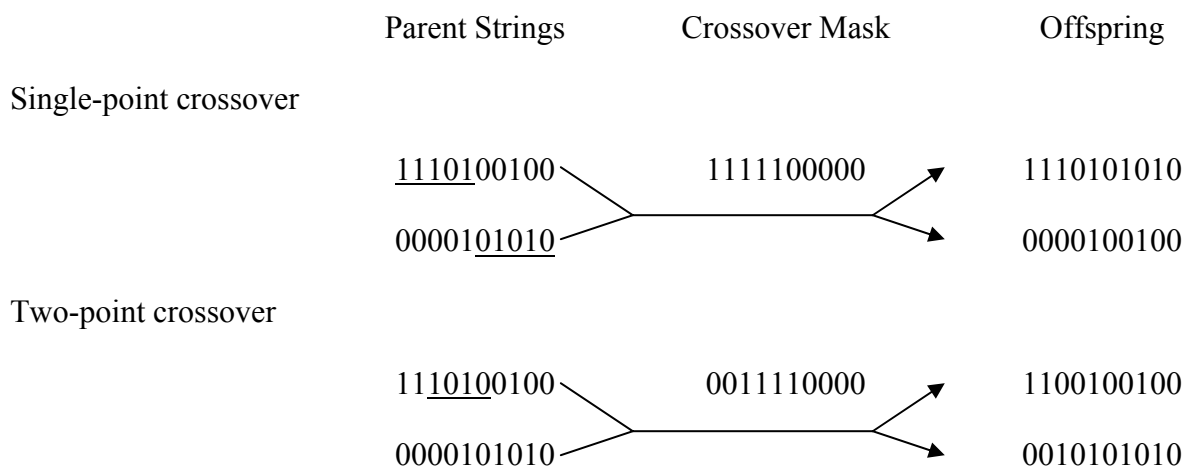
Figure 4.3: Illustration of Crossover Operation (ten bits string is used as an example)

90

Mutation is a process to overcome the local optimum problem. Each bit of a selected string is allowed to mutate according to a predetermined mutation probability, thereby reducing the likelihood that the search process will get stuck in a local optimum.

4.4 Empirical Data Collection and Processing

4.4.1 Study Site

We apply the proposed method to newly collected data, concerning route choice in the California State Route 91 value-pricing project. The SR 91 toll lanes, located between the SR 91/SR 55 junction in Anaheim, CA and the Orange/Riverside County Line, are the world's first fully-automated privately operated toll lanes (Sullivan, 2000). The express lanes extend approximately about ten miles along the former median of the Riverside Freeway (SR 91), connecting rapidly growing residential areas in Riverside and San Bernardino Counties to job centers in Orange and Los Angeles Counties to the west. SR 91 in eastern Orange County includes four regular freeway lanes (91F) and two express lanes (91X) in each direction. Motorists who wish to use express lanes must register and carry identifying electronic transponders (the so-called *FasTrak*) to pay a toll that varies hourly according to a preset schedule. Tolls in the express lanes vary hour-by-hour to control demand and maintain free flow traffic, in contrast to often congested traffic conditions in the adjacent free lanes. Tolls on westbound traffic during morning commute hours ranged from $1.65 (at 4-5 a.m.) to $3.30 (at 7-8 a.m., Monday-Thursday). Within the SR 91 corridor, the Eastern Toll Road (ETR) competes with the 91X for trips to Irvine and vicinity. However, since the 91X has no entrance or exit between its starting and ending points, ETR users must use the highly congested 91F for access.

We regard the SR 91 toll road portion as a two-route network. One route is the 91X, and the other is 91F, both 10 miles in length. This gives motorists the option to travel free on regular roads or to pay a time-varying price for congestion-free express travel on a limited part of their journey. Because of the toll pricing structure, observation is that traffic consistently moves at a

free-flow speed of approximately 75mph, even during peak hours on this facility. Consequently, we assume that travel time on 91X is deterministic (reliable) and equal to 8 minutes, corresponding to a speed of 75 mph. However, the free lanes are often congested during the morning peak hours (5-9 a.m.), and travel time on the 91F is rather stochastic and unreliable, presenting a relatively "clean" real-world experimental environment to study the relative contributions of travel time and travel-time reliability in the route choice decision process.

4.4.2 Travel Time Reliability Data Collection and Processing

Obtaining accurate measures of travel conditions, especially the appropriate measurement of travel time reliability, is a formidable task. We use actual field measurements (floating cars) of travel time on 91F taken at different times during morning peak period. Our data was obtained from the study of Small *et al*. (2002). The data consists of peak period travel time on 91F for 11 days: first on October 28, 1999, and then on July 10-14 and September 18-22, 2000. Data were collected from 4:00am to 10am on each day, and include a total of 210 observations of travel time along the 10-mile stretch of 91F at different times of day encompassing the morning peak period. Interested readers may refer to their paper for more details on the travel time data collection and processing techniques.

In order to construct measures of travel time and its reliability, we consider both the central tendency and the dispersion of the travel time distribution. Measures of central tendency include the mean and the median, and measures of dispersion include the standard deviation, the inter-quartile difference such as the 90th-50th or 80th-50th, ratio of standard deviation to mean, and percent of observations that exceed the mean by some specific threshold, etc. The nature of these measures is that they are positive, monotonically increasing functions of variability. We assume that motorists, especially commuters in the morning peak hours, are concerned with the probability of significant delay, and are likely to pay particular attention to the upper tail of the distribution of travel times. Among the candidate measures that capture this effect, we arbitrarily use the difference between the upper quartile and the median. To make our results comparable to Small *et al*. (2002), we use the same measures of central tendency and dispersion, i.e., median and the 80th-50th percentile differences.

Figure 4.4 shows the raw field observations of travel time savings (i.e., the difference between the 91F and 91X travel times over the 10-mile stretch). The non-parametric estimates of mean, median, and 80th percentile are calculated and displayed. Median time savings reach a peak of 5.6 minutes around 7:15 a.m. Figure 4.5 shows the median travel time savings and the 80th-50th percentile differences. The latter reaches a peak around 8:10 a.m. Correlations between these two measures are insignificant.
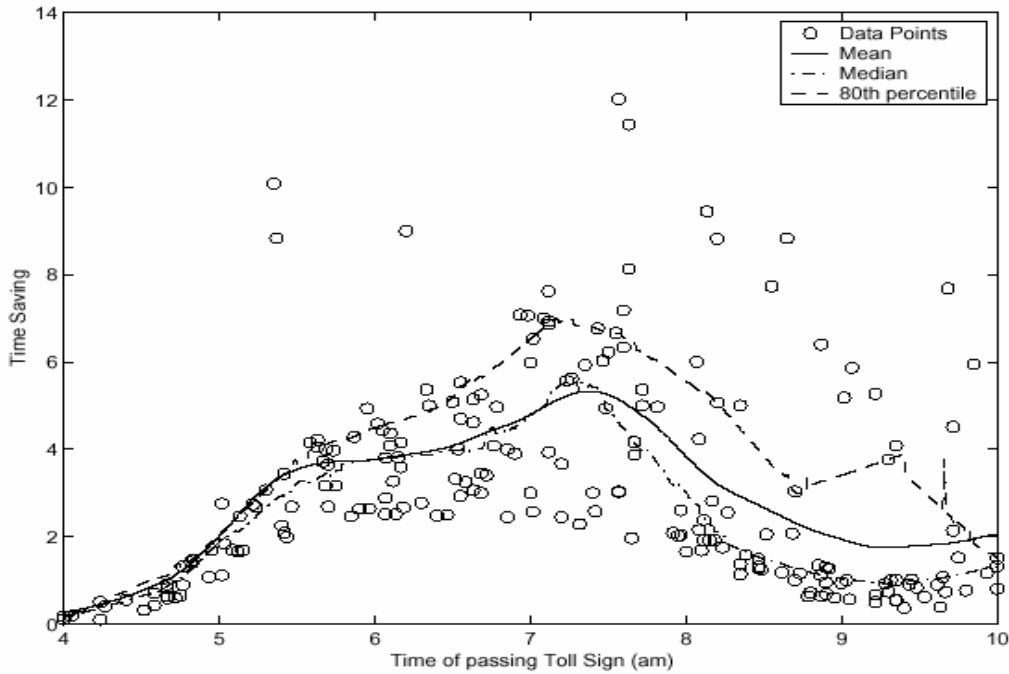
Figure 4.4: Travel Time Saving (From the study of Small *et al*., 2002)
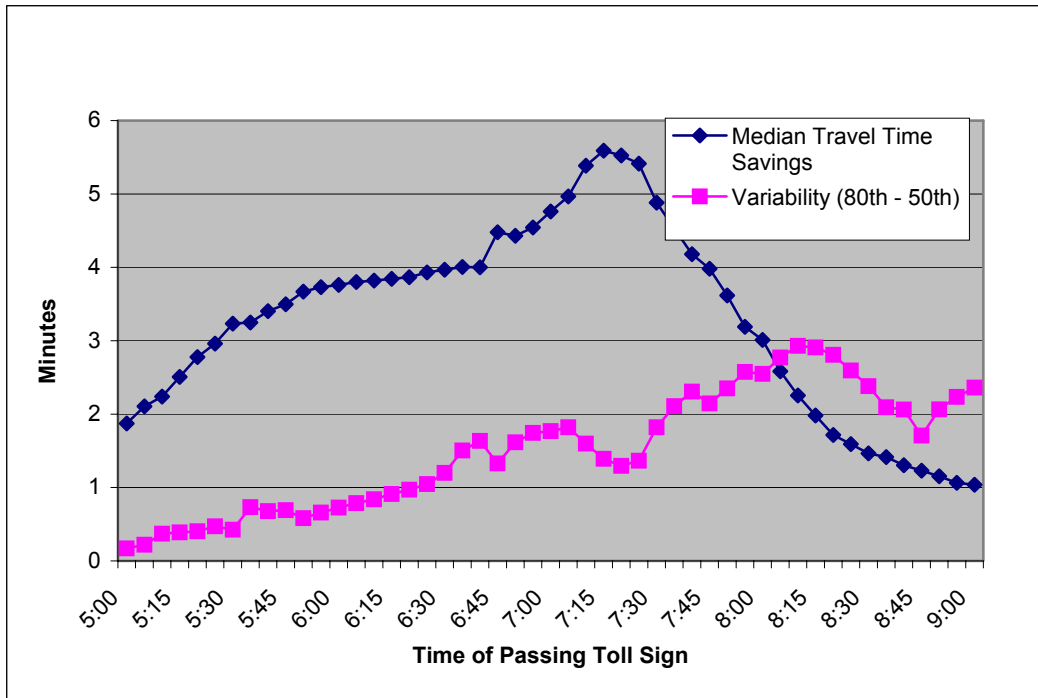
Figure 4.5: Measurements of Travel Time and Its Variability

4.4.3 Volume Data Collection and Processing

Along the stretch of SR 91, under consideration, loop detector stations are spaced at a distance of one every mile. Each loop detector station includes 6 loops covering all lanes of 91F and 91X. Volume data were collected using 30-second loop detector data and aggregated into five-minute intervals. The data consists of volumes on 91X, 91F, and ETR for 30 weekdays from September 17 to November 16, 2001. Since our study is concerned with the traveler's choice probability between 91X and 91F, the volume data of ETR was subtracted from 91F because ETR users have no option but to use 91F. Figure 4.6 shows the traffic flow on both 91X and 91F. As shown in Figure 4.7, the percentage of travelers taking 91X reaches a peak at around 8:00 a.m.
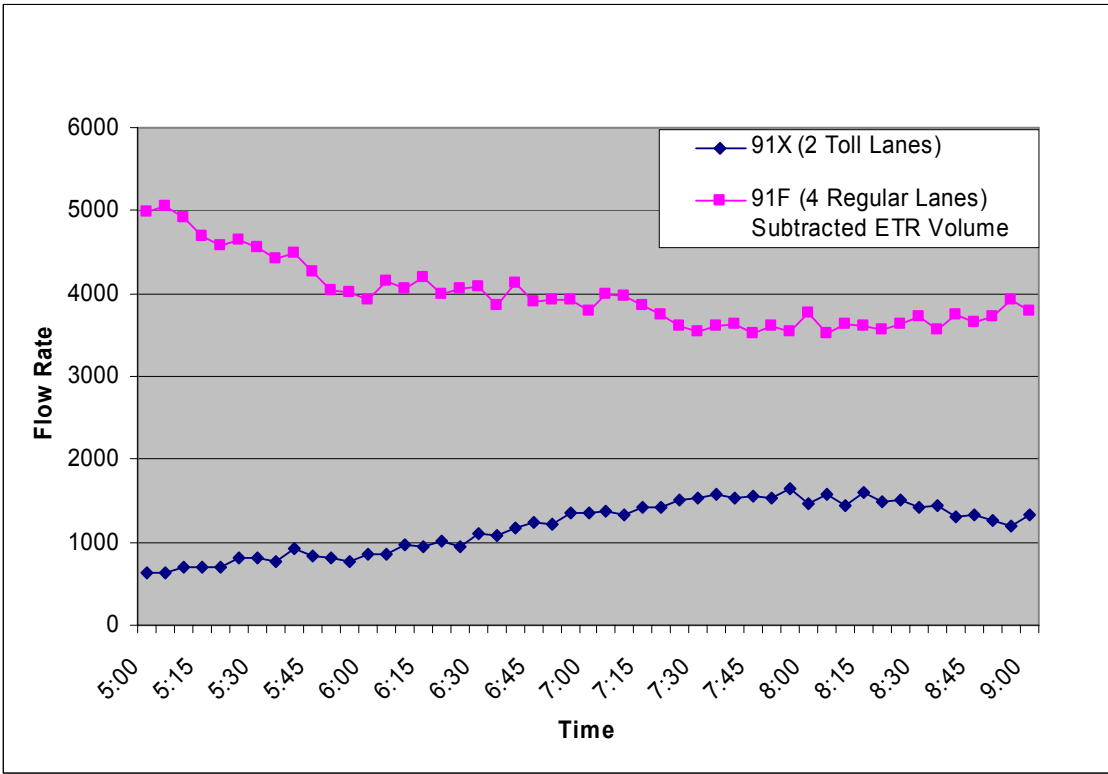
Figure 4.6: Traffic Flow on SR91 (September 18, 2001)
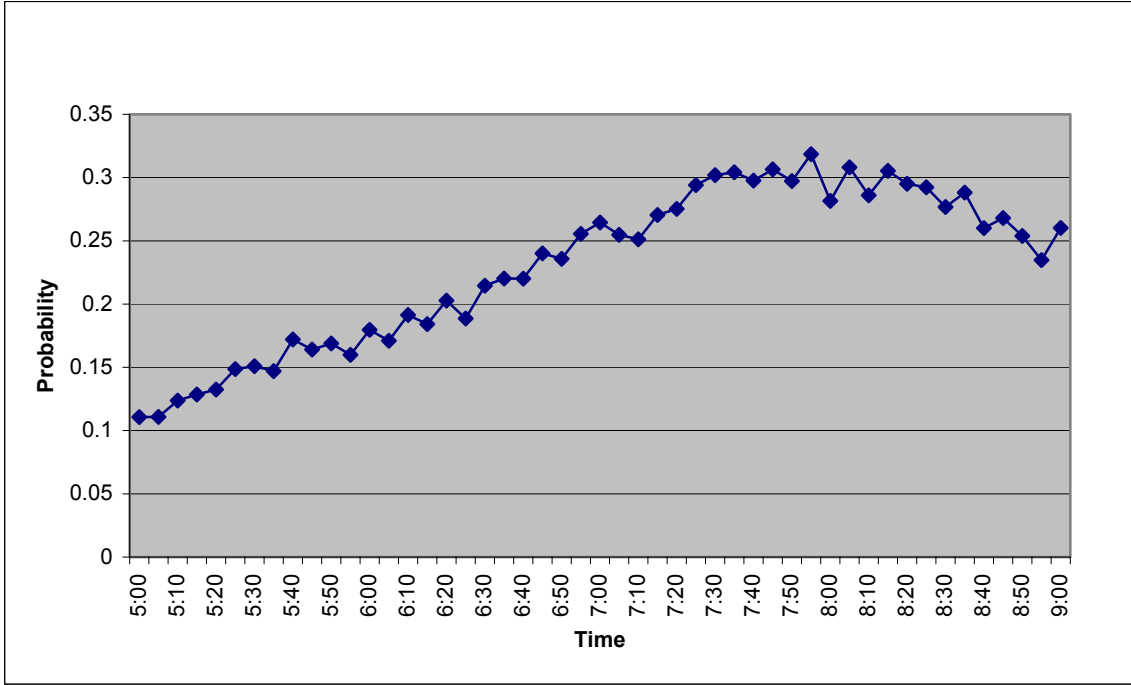
Figure 4.7: Percentage of Travelers Taking Toll Lanes from Loop Detector Data
(September 18, 2001)

4.5   Results and Analysis

After data collection and processing, we applied the GA algorithm to identify the parameter set that produced the best match to the volume data revealed from the loop detectors. The parameters used in the GA algorithms are shown in Table 4.1. To evaluate each chromosome, the choice probability estimated from mixed-logit was calculated using 2,000 random draws from a normal distribution of the components of $\boldsymbol{\beta}$ for the Monte-Carlo simulations. The convergence of the GA is shown in Figure 4.8, which displays the decreasing MAER values with the number of generations.

To estimate the parameters Θ with certain confidence level, we performed 30 GA runs with the volume data from 30 days, with each GA run corresponding to the volume data from one particular day. Figure 4.9 shows the identified values for parameters Θ with 30 GA runs (The standard deviation *Sd* is given instead of variance *W*). The statistical estimates of parameters Θ from these 30 GA runs are shown in Table 4.2.

Table 4.1: Control Parameters for the GA

Population size	20
Maximal Number of Generations	50
Chromosome Length Per Parameter	6
Crossover probability	0.6
Mutation probability	0.033
Elitism flag	1
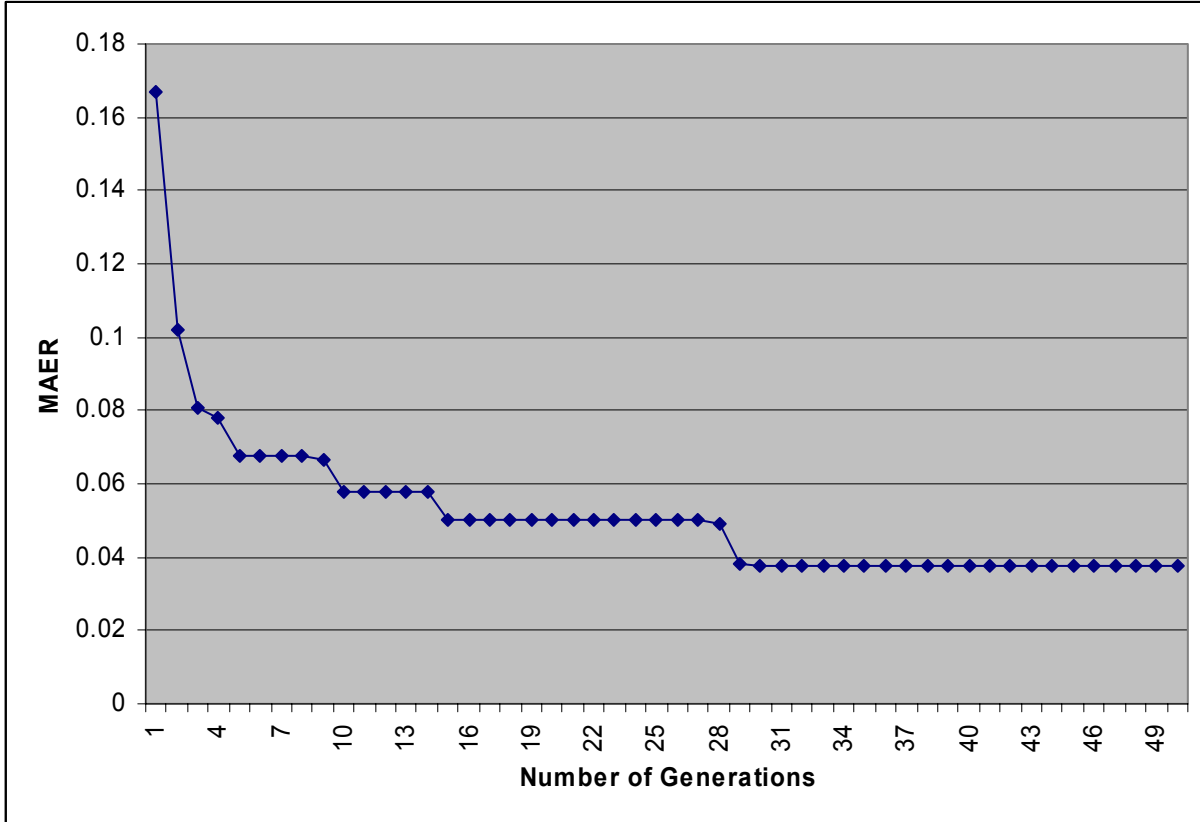
Figure 4.8: GA Convergence Curve with Best Fitness Values

Table 4.2: Estimated Parameter Values

Parameters	Range	Estimated Value (Median / [5%-ile, 95%-ile])
b^T	[10, 40]	15.39 / [11.05, 20.34]
Sd^T	[0, 10]	5.28 / [0.77, 9.07]
b^V	[10, 40]	25.66 / [20.08, 29.38]
Sd^V	[0, 10]	6.14 / [2.54, 9.38]
b^C	[1, 2]	1.26 / [1.0, 1.81]
Sd^C	[0, 1]	0.55 / [0.04, 0.94]
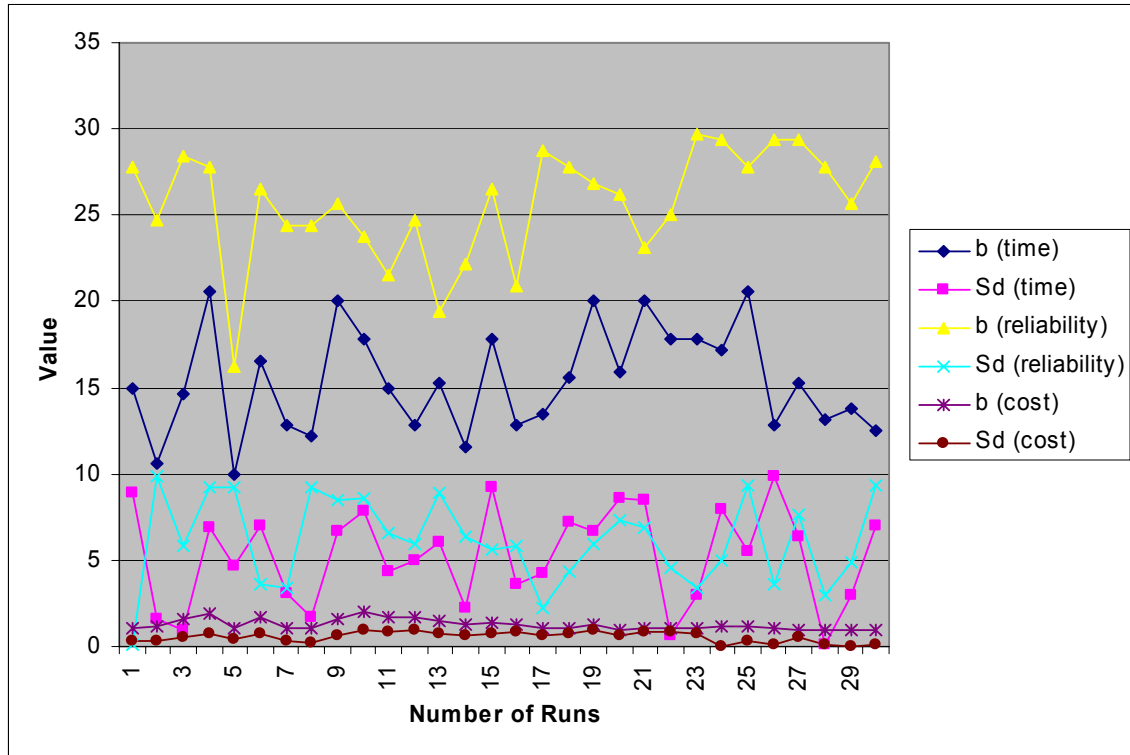
Figure 4.9: Best Estimated Parameters with 30 GA Runs

From the estimated statistical distributions of the parameter estimates, travelers' implied VOT, VOR and DORA, and the extent of their heterogeneity can be determined by Monte-Carlo simulations. We performed 2,000 random draws from normal distributions of **β** ($\beta^T \sim N(b^T, W^T)$, $\beta^V \sim N(b^V, W^V)$, $\beta^C \sim N(b^C, W^C)$), and calculated VOT, VOR and DORA using Equation (4.3 - 4.5). Percentile values, including 25%-ile, 50%-ile (median), and 75%-ile, were then obtained from the 2,000 values of VOT, VOR and DORA. Travelers' heterogeneity is measured as the inter-quartile difference, i.e., the difference between the 75th and 25th percentile values, because it is unaffected by high upper-tail values occasionally found in the calculation of ratios. We repeated this process for every parameter set $\{b^T, W^T, b^V, W^V, b^C, W^C\}$ identified by the 30 GA runs. The estimates of the median and heterogeneity of VOT, VOR, and DORA are shown in Table 4.3. In Table 4.3, we note that the confidence interval represents uncertainty due to statistical error, not heterogeneity. A positive

98

5th percentile value means the quantity is significantly greater than zero according to a conventional one-sided hypothesis test at a 5 percent significance level.

Table 4.3: Estimated Values of Time and Reliability

	Median Estimate	90% Confidence Interval [5%-ile, 95%-ile]
Value of Time ($/hour)		
Median	12.81	[8.66, 16.08]
Heterogeneity (75th-25th)	8.72	[4.16, 16.73]
Value of Reliability ($/hour)		
Median	20.63	[12.81, 28.47]
Heterogeneity (75th-25th)	13.06	[5.62, 23.44]
Degree of Risk Aversion (DORA)		
Median	1.73	[1.26, 2.13]
Heterogeneity (75th-25th)	0.97	[0.41, 1.75]

As shown in Table 4.3, the median value of time is $12.81, and the median value of reliability is $20.63. Since median time savings in our data peaks at 5.6 minutes in the rush hour and unreliability peaks at 3 minutes, the average commuter would pay $1.20 to realize time savings and pay $1.03 to avoid this possibility of unanticipated delay. In other words, travelers with the median VOT and VOR would save $2.23 from travel time and its reliability if they use 91X, but they need to pay $3.30 for the toll. So less than half travelers will choose to use the express lanes, and this is confirmed from the loop detector data, as shown in Figure 4.7. Regarding travelers' heterogeneity towards travel time and its reliability, both measures of heterogeneity in the cases of VOT and VOR are more than 60% of their median values, indicating that commuters exhibit a wide distribution of preferences for speed and reliability. By recognizing the heterogeneity in travelers' preference and offering choices that caters to their preferences, road pricing policies can increase transportation efficiency.

Similar results for the median values of VOT and VOR are found in the study of Small *et al*. (2002). In their study, they estimated VOT and VOR using the combination of RP and SP data. The median value of VOT estimated from RP data is $20.20, and $9.46 from SP data. Our results, in this regard, falls between these two values. In terms of VOR, the estimated median value from their study using RP data is $19.56, which is very similar with our result. Their

estimate of VOR from SP data is not comparable to ours since different measures of un-reliability are used. Our results validate the analysis from their study and demonstrate the applicability of an approach based on conventional loop data in the study of travelers' behavior.

The median value of DORA is 1.73, i.e., the disutility caused by certain amount of travel time unreliability is 1.73 times more than that caused by travel time of the same amount. For instance, assuming the commuting alternatives are a 20-minute commuting with essentially no possibility of significant delays and a commuting alternative that normally takes 10 minutes but has a variability of about 6 minutes. If the traveler has a DORA equal to 1.73, he/she will be almost indifferent between these two choices. A traveler with a DORA greater than 1.73 is more risk averse and will choose the first alternative. In other words, travelers with DORA greater than 1.0 value more highly a reduction in variability than a comparable reduction in the mean travel time. These travelers are willing to go out of their way to decrease the possibility of a delay, either because they dislike the risk of being delayed or dislike the discomfort normally associated with a delay, such as stop and go traffic. From the operator's point of view, this finding implies that traffic management strategies aimed at reducing travel time variability, such as incident management, deserve serious attention.

4.6 Summary

It is generally accepted that travel time reliability can have significant influence on traveler's route choice behavior and that it cannot be ignored in any model which purports to predict behavior or provide a basis for evaluation. With respect to the valuation of reliability, such direct methods as revealed preference surveys and/or state preference surveys have been used extensively in previous studies. In this chapter, we proposed an indirect method to study the contribution of travel time reliability in traveler's route choice behavior. We formulated traveler's route choice as a mixed-logit model, with the coefficients in the model representing individual traveler's preferences or tastes to travel time, reliability and cost. Unlike the traditional approach to estimate these coefficients with RP and/or SP data by simulated maximum likelihood estimation, we adopt genetic algorithm to identify the coefficients that enable the flows resulting from route choice model to best match the time-dependent traffic volume data obtained from loop detectors. Such an approach eliminates both the cost and biases inherent to RP and SP survey techniques.

We applied the proposed method to newly collected data concerning route choice in the California State Route 91 value-pricing project. Based on travelers' choice of whether or not to pay a congestion-based toll in order to use express lanes, we are able to estimate how travelers value travel time and travel-time reliability. We find that the estimated median value of travel-time reliability is substantially greater than that of travel-time, and the median value of degree of risk aversion is greater than 1, indicating that travelers value more highly a reduction in variability than in the mean travel time saving for that journey. Moreover, travelers' attitude towards congestion is not homogeneous; in fact, substantial heterogeneity exists in travelers' preference of travel time and reliability. The results of our study yield important insights into commuters' route choice in general and the tradeoffs among travel time, reliability, and monetary cost. Our results validate the analysis from some previous studies and demonstrate the applicability of the approach in the study of travelers' behavior.

Chapter 5

5. IMPLEMENTATION OF THE ROUTE CHOICE MODELS

In this chapter, implementation issues of the route choice models described in Chapter Two are discussed. Detailed traffic assignment procedures are provided for readers interested in implementing the route choice models. In addition, numerical results are provided to examine the effects of route choice models on network flow allocations.

5.1 Implementation Issues

All four route choice models described in Chapter Two can be implemented using the Frank-Wolfe (FW) based algorithm. The main differences, between the route choice models, are the criterion used to determine the search direction and the one-dimensional search schemes used to determine the step size. Table 5.1 briefly summarizes the criterion and line search scheme used in the FW based algorithm to implement the four route choice models.

Table 5.1: Implementation of the Route Choice Models

Route Choice Model	Search Direction	Step Size
DN-DUE	Minimize expected travel time	Bisection
DN-SUE	Minimize expected *perceived* travel time	1/n
SN-DUE	Minimize expected disutility	1/n
SN-SUE	Minimize expected *perceived* disutility	1/n

Before describing the implementation details, we discuss the issues involved in implementing the four route choice models. These issues include: (a) the estimation of link travel times and

variance, (b) the assumptions with regard to the perception error distributions, (c) the estimation of travel disutility functions for describing different risk-taking behavior, (d) the estimation of link disutilities, and (e) the choice of stochastic loading used in the FW-based traffic assignment procedure.

5.1.1 Estimation of Link Travel Times and Variance

All four route choice models account for congestion effects using a travel time function which is modeled as an increasing function of flow of vehicles on the link. Each traveler makes a rational route choice decision based on minimizing some criteria related to average travel times or some disutility function based on average travel times and their variances. Link travel times and variances are assumed to be directly dependent on the flow on the link. The link travel time function used in the route choice models is the standard Bureau of Public Road (BPR) function given below:

$$t_a = t_a^f \left(1 + 0.15 \left(\frac{v_a}{c_a} \right)^4 \right), \tag{5.1}$$

where v_a, t_a^f, and c_a are the flow, free-flow travel time, and capacity on link a. Depending on the availability of data to estimate travel time variability, these links are assumed either to have deterministic travel times that vary only with their flows or to follow certain distribution that characterizes the variation associated with the time moving along the link.

5.1.2 Assumptions on Perception Error

Due to the unrealistic assumption that all travelers have perfect knowledge of the network conditions, the SUE models relax the perfect knowledge assumption by introducing a perception error into the route choice process to allow for variations in travelers' perceptions of network conditions. For the DN-SUE model, each traveler i is assumed to have some perceptions of the expected travel time for each link a which include a random error term. In this research, we use the probit model to account for perception error in the DN-SUE model. The randomness of the ith traveler's perceived travel time on link a is assumed to follow a normal distribution as follows:

$$\xi_a^i \sim N\big(0, \tau_a\big), \tag{5.2}$$

where the mean of the perception error is assumed to equal zero and the variance of the perception error is assumed to equal τ_a. Sheffi (1985) suggests relating the variance of the perception error to the free-flow travel time by setting

$$\tau_a = \beta\, t_a^f, \tag{5.3}$$

where β is a proportionality constant (i.e., the variance of the perceived travel time over a road segment of unit travel time at free flow. For the SN-SUE model, each traveler i is assumed to have a variable perception error

$$\xi_a^i \sim N\big(\mu_a^i, \theta_a^i\big), \tag{5.4}$$

where $\mu_a^i \sim N\big(0, \tau_a\big)$ and $\theta_a^i \sim G\big(\alpha_a, \beta_a\big)$ (Mirchandani and Soroush, 1987). The parameter of μ_a^i is assumed to be normally distributed over the population of travelers with an expected value equal to zero and a variance that is proportionally to the free-flow travel time on link a. The parameter θ_a^i is assumed to gamma distributed over the population with parameters α_a and β_a. This variable perception error allows each individual traveler to experience a different travel time for a given set of flows. This is different from the probit-based DN-SUE model, which only accounts for the randomness of the travelers' perceived travel times (i.e., $\xi_i \sim N\big(0, \tau\big)$) and treats the randomness of link travel times in the form of expected values.

5.1.3   Estimation of Travel Disutility Functions

Travel time variability introduces uncertainty for travelers such that they do not know exactly when they will arrive at the destination. Thus, it is considered as a risk (or an added cost) to a traveler making a trip. To model the travelers' attitudes towards risk, disutility functions are needed to represent different risk-taking behaviors. Depending on the behavioral nature of

travelers, they can be classified as risk neutral, risk averse, or risk prone. Since a risk neutral traveler would choose a route based on only expected travel time without consideration of risk (or travel time variability), we simply model the risk neutral travelers using either the DN-DUE model or the DN-SUE model depending on whether perception error is considered or not.

For the SN-DUE and SN-SUE models, following the study by Tatineni *et al.* (1997), we also use the exponential functional form to describe risk averse and risk prone behaviors. Suppose that the travelers' disutility (without perception error) has the following forms:

Risk Averse: $U(T) = a_1 [\exp(a_2 T) - 1]$,   (5.5)

Risk Prone: $U(T) = b_1 [1 - \exp(-b_2 T)]$,   (5.6)

where T is the travel time in minutes for a given path, a_1, a_2, b_1, and b_2 are positive parameters to be estimated. The parameters a_2 (b_2) are a measure of risk aversion (proneness). As a_2 (b_2) gets larger, the traveler becomes more averse (seeking) to risk.

To estimate the parameters of the disutility function given in Equations (5.5) and (5.6), we need to make assumptions on how travelers value route travel time according to their risk-taking behavior. For illustration, the estimation of parameters for the risk averse case are explained. Suppose that all travelers have a disutility of 1 unit for a route that is 5 minutes, and they are indifferent between a route that has a travel time of 3.33 minutes for certain and a route that has an equal chance of having a travel time that is either close to zero minute or is equal to 5 minutes (i.e., an average of 2.5 minutes on this route). Travelers who took the longer travel time route (3.33 minutes) would pay a risk premium of 0.83 minute (i.e., 3.33 − 2.5) to avoid uncertainty. This is a typical behavior for risk averse travelers. Given these two assumptions, two equations, to solve for the values of the two parameters a_1 and a_2, are established as follows:

$$U(5) = a_1 [\exp(a_2\ 5) - 1] = 1.0 \qquad\qquad (5.7)$$

$$U(3.33) = a_1 [\exp(a_2\ 3.33) - 1] = 0.5 \qquad\qquad (5.8)$$

Solving the above equations for a_1 and a_2 give the following disutility function for a risk averse traveler:

$$U(T) = 0.309 \ [\exp(0.289 \ T) - 1] \qquad\qquad (5.9)$$

Using similar assumptions for the risk prone travelers, the disutility function for a risk prone traveler is:

$$U(T) = 1.309 \ [1 - \exp(-0.289 \ T)] \qquad\qquad (5.10)$$

For the risk neutral behavior, we can use the DN-SUE model, which uses a linear disutility function, by setting the disutility equal to the expect travel time. The shapes for these different risk taking behaviors are provided in Figure 5.1. For details of the derivation of the parameters for the disutility function can be found in Tatineni (1996) and Soroush (1984).
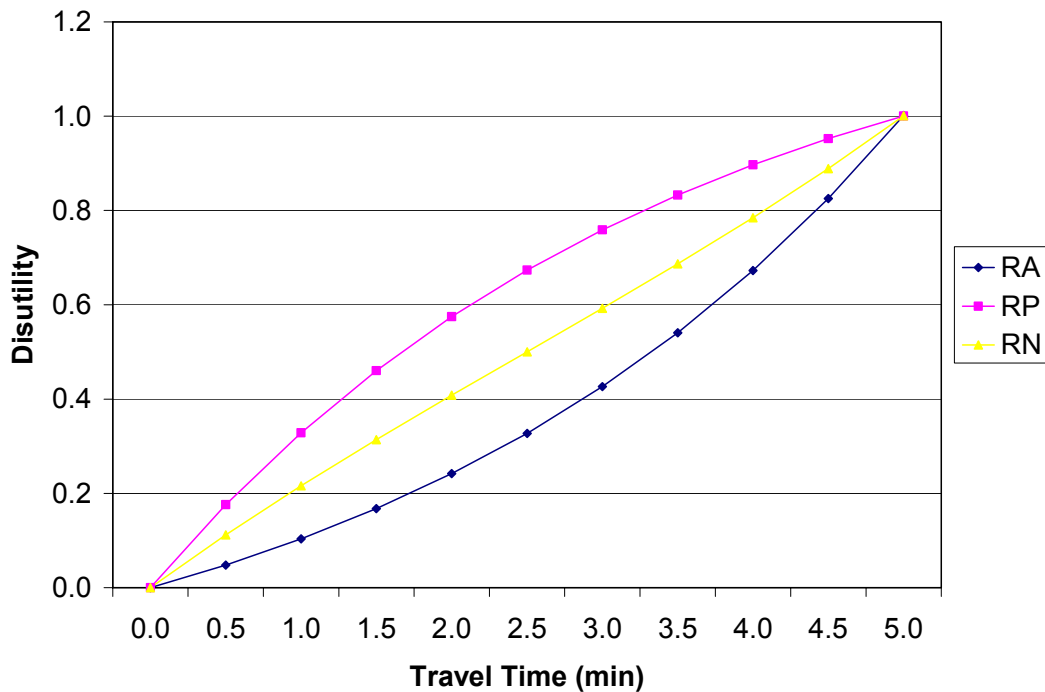
Figure 5.1: Disutility Functions for the Risk-Taking Route Choice Models

5.1.4 Estimation of Link Disutilities

Recall from Section 2.5, the *perceived* link travel times are obtained using the moment generating function (MGF) to link the variable perception error of the traveler and the distribution of link travel times. The *perceived* link travel time for individual i is given in equation (2.11) and is repeated here for convenience.

$$M_{\tilde{t}_{ia}}(s) = M_{t_a}\left[s\left(1 + \mu_a^i + \frac{1}{2}\theta_a^i s\right)\right],$$

(5.11)

where s is a real number, μ_a^i and θ_a^i are the parameters of the perception error of individual i conditioned on link a. Using equation 5.11, the first and second order moments can be derived easily taking the first and second derivatives with respect to s and evaluate it at $s = 0$ as follows:

$$E\left(\tilde{t}_{ia}\right) = \bar{t}_a\left(1 + \mu_a^i\right),$$

(5.12)

$$E\left(\tilde{t}_{ia}^2\right) = \left(\sigma_a^2 + \bar{t}_a^2\right)\left(1 + \mu_a^i\right) + \theta_a^i\,\bar{t}_a,$$

(5.13)

where $\bar{t}_a$ and σ_a^2 are the mean and variance of travel time on link a.

To estimate the *perceived* link disutilities, we incorporate the *perceived* link travel times into the travel disutility functions for each risk-taking behavior. Suppose that the traveler's *perceived* disutility function, $\tilde{U}(\tilde{T})$, is exponential with respect to $\tilde{T}$ and is given by

$$\tilde{U}(\tilde{T}) = \text{sgn}(b)\exp(b\tilde{T}),$$

(5.14)

where

$$\text{sgn}(b) = \begin{cases} 1 & if\ b > 0 \\ -1 & if\ b < 0 \end{cases}.$$

(5.15)

$\widetilde{T}$ is the *perceived* travel time for a given path. We note that when b is positive, then the disutility function is convex and travelers are "*constant risk averse*", that means the risk averseness does not change when travel time varies; however, when b is negative, then the disutility function is concave and travelers are "*constant risk prone*", that means the risk proneness does not change when travel time varies. The expected *perceived* disutility of path k is

$$
\begin{aligned}
\widetilde{U}_k &= E[\,\text{sgn}(b)\exp(b\widetilde{T}_k)\,] \\
&= \text{sgn}(b)E[\,\exp(b\widetilde{T}_k)\,]. \\
&= \text{sgn}(b)M_{\widetilde{T}_a}(b)
\end{aligned}
\tag{5.16}
$$

Since link travel times are assumed to be statistically independent, we can rewrite the above formula as below,

$$
\widetilde{U}_k = \text{sgn}(b)\prod_{a\in k} M_{\widetilde{T}_a}(b).
\tag{5.17}
$$

Substitute equation (5.11), we get

$$
\widetilde{U}_k = \text{sgn}(b)\prod_{a\in k} M_{T_a}[b(1+\mu_a^i + \frac{b\theta_a^i}{2}],
\tag{5.18}
$$

where μ_a^i and θ_a^i are sampled from $N(0,\tau)$ and $G(\alpha,\beta)$. The expected *perceived* disutility of path k is therefore:

$$
\begin{aligned}
E\left[\widetilde{U}_k\right] &= E\left[\text{sgn}(b)\prod_{a\in k} M_{T_a}[\,b(1+\mu_a^i + \frac{b\theta_a^i}{2}]\,\right] \\
&= E[\,\text{sgn}(b)\,]E\left[\prod_{a\in k} M_{T_a}[b(1+\mu_a^i + \frac{b\theta_a^i}{2}]\right]
\end{aligned}.
\tag{5.19}
$$

Using the parameters estimated in equation (5.9), the *perceived* link disutilities for the risk averse behavior can be derived as follows:

$$\tilde{U}_a = 0.02406 \cdot (\bar{t}_a + \bar{t}_a \cdot \mu_a^i + \frac{(0.02406 \cdot \theta_a^i \cdot \bar{t}_a)}{2}) +$$

$$\sigma_a^2 \cdot \left[0.02406 \left(1 + \mu_a^i + \frac{(0.02406 \cdot \theta_a^i)}{2}\right) \right]^2 \left[\frac{1}{2} + \frac{(0.02406 \cdot \gamma)}{3} \left(1 + \mu_a^i + \frac{(0.02406 \cdot \theta_a^i)}{2}\right) \right] \qquad (5.20)$$

When the travelers have no perception error, the *perceived* distribution of link travel times is equal to the actual distribution of link travel times (i.e., $\mu_a^i = 0$ and $\theta_a^i = 0$). The link disutilities are just a function of the behavioral type (e.g., risk averse behavior) and the distribution of travel times:

$$U_a = 0.02406\, \bar{t}_a + (0.02406)^2 \sigma_a^2 \left[\frac{1}{2} + \frac{(0.02406 \cdot \gamma)}{3} \right]. \qquad (5.21)$$

Using the parameters estimated in equation (5.10), the *perceived* link disutilities and link disutilities for the risk prone can also be similarly derived as follows:

$$\tilde{U}_a = 0.02406(\bar{t}_a + \bar{t}_a \mu_a^i + \frac{(-0.02406 \cdot \theta_a^i \bar{t}_a)}{2}) -$$

$$\sigma_a^2 \left[-0.02406 \left(1 + \mu_a^i + \frac{(-0.02406\theta_a^i)}{2}\right) \right]^2 \left[\frac{1}{2} + \frac{(-0.02406\gamma)}{3} \left(1 + \mu_a^i + \frac{(-0.02406\theta_a^i)}{2}\right) \right] \qquad (5.22)$$

and

$$U_a = 0.02406\, \bar{t}_a - (0.02406)^2 \sigma_a^2 \left[\frac{1}{2} + \frac{(-0.02406\gamma)}{3} \right] \qquad (5.23)$$

The details of the derivation can be found in Tatineni (1996) and Soroush (1984).

5.1.5 Choice of Stochastic Loading

For the probit-based DN-SUE model, there exists no closed-form solution for the route choice probability. A Monte Carlo simulation sampling technique is performed to construct a stochastic loading (Sheffi, 1985). At each iteration of the solution algorithm, the OD flows have to be

loaded on the minimum *perceived* travel time routes between each OD pair. To find the minimum *perceived* travel time routes, at every iteration, several sets of perception errors are sampled and the routes with the lowest *perceived* travel times are found for each sample. Instead of a single all-or-nothing loading, the OD flows are then averaged among every route between the origin and the destination. The stochastic flows estimated thus are averaged between successive iterations of the solution algorithm based on a pre-specified averaging rule until terminating criterion is satisfied. This Method of Successive Averages (MSA), as it is generally known, is used to solve the probit-based DN-SUE model (Sheffi and Powell, 1982).

An advantage of using the exponential functional form to represent risk-taking behavior is that the disutility associated with a route can be estimated by summing the link disutilities on a given route (see Section 5.1.4). This allows the classical Dijkstra-type shortest path algorithm to be used in the stochastic loading step to find the minimum expected disutility route in the SN-DUE model and the minimum expected *perceived* disutility route in the SN-SUE model.

5.2   Traffic Assignment Procedures

The implementation of the DN-DUE and the DN-SUE models is based on the FW algorithm and the method of successive averages (MSA) algorithm, respectively. The details of these two algorithms are well described in the excellent text by Sheffi (1985): pages 119-120 for the FW algorithm and page 327 for the MSA algorithm. For the SN-DUE and the SN-SUE models that account for network uncertainty and risk taking behavior, both can also be solved using the MSA algorithm with a different criterion used to perform the stochastic network loading procedure to determine the search direction. Here we provide the detailed steps for implementing these four route choice models.

5.2.1   Implementation of the DN-DUE Model

The criterion used in the DN-DUE model is to minimize the expected value of route travel time, which can be obtained by adding up the average travel times of all the links belonging to the route. The overall flowchart of the traffic assignment procedure for solving the DN-DUE model is given in Figure 5.2 and summarizes below:

Figure 5.2: Flowchart for Solving the DN-DUE Model

Step 0: *Initialization*. Generate an initial path for each OD pair.

0.1   Set $x_a(0) = 0$, $t_a = t_a[x_a(0)]$, $\forall a$ and $K_{rs}(0) = \varnothing$.

0.2   Set iteration counter $n = 1$.

0.3   Solve the shortest path problem: $\bar{k}_{rs}(n)$, $K_{rs}(n) = \bar{k}_{rs}(n) \cup K_{rs}(n-1)$, $\forall r, s$.

0.4   Perform AON assignment: $f^{rs}_{\bar{k}_{rs}(n)} = q_{rs}$, $\forall r, s$.

0.5   Assign path flows to links: $x_a(n) = \sum\limits_{r \in R} \sum\limits_{s \in S} \sum\limits_{k \in K_{rs}(n)} f^{rs}_k(n) \delta^{rs}_{ka}$, $\forall a$.

Step 1: *Increment iteration counter.* $n = n + 1$.

Step 2: *Update link travel time.* $t_a(n) = t_a[x_a(n-1)], \forall\, a$.

Step 3: *Direction Finding.* Perform all-or-nothing assignment.

    3.1   Solve the shortest path problem: $\bar{k}_{rs}(n), \ \forall\, r, s$.

    3.2   Augment path $\bar{k}_{rs}(n)$ to the path set $K_{rs}(n-1)$ if it is not already in the set:

        If $\bar{k}_{rs}(n) \notin K_{rs}(n-1)$, then $K_{rs}(n) = \bar{k}_{rs}(n) \cup K_{rs}(n-1), \forall\, r, s$; otherwise, tag the

        shortest path among the paths in $K_{rs}(n-1)$ as $\bar{k}_{rs}(n)$ and set

        $K_{rs}(n) = K_{rs}(n-1), \forall\, r, s$.

    3.3   Perform all-or-nothing assignment: $h^{rs}_{\bar{k}_{rs}(n)} = q_{rs}, \forall\, r, s$.

    3.4   Determine auxiliary link flows: $y_a(n) = \sum_{r \in R} \sum_{s \in S} \sum_{\bar{k}_{rs}(n)} h^{rs}_{\bar{k}_{rs}(n)} \delta^{rs}_{\bar{k}_{rs}(n)a}, \forall a$.

Step 4: *Line Search.* Find the value of α_n that solves:

$$\underset{0 \le \alpha \le 1}{Min} \ \sum_a \int_0^{x_a(n-1) + \alpha_n[y_a(n) - x_a(n-1)]} t_a(w) dw$$

Step 5: *Move.* Update path and link flows.

    5.1   Set path flows: $f_k^{rs}(n) = f_k^{rs}(n-1) - \alpha_n[h^{rs}_{\bar{k}_{rs}(n)} - f_k^{rs}(n-1)], \forall k \in K_{rs}(n), r, s$.

    5.2   Assign path flows to links: $x_a(n) = \sum_{r \in R} \sum_{s \in S} \sum_{k \in K_{rs}(n)} f_k^{rs}(n) \delta^{rs}_{ka}, \ \forall\, a$

Step 6: *Convergence Test.* Terminate the algorithm if it satisfies the stopping criterion (the

    current solution, $x_a(n)$, is the set of equilibrium link flows); otherwise, go to Step 1.

5.2.2 Implementation of the DN-SUE Model

The criterion used in the DN-SUE model is to minimize the expected *perceived* value of route travel time, which can be obtained by adding up the *perceived* travel times of all the links belonging to the route. The overall flowchart of the traffic assignment procedure for solving the DN-SUE model is given in Figure 5.3 and summarizes below:
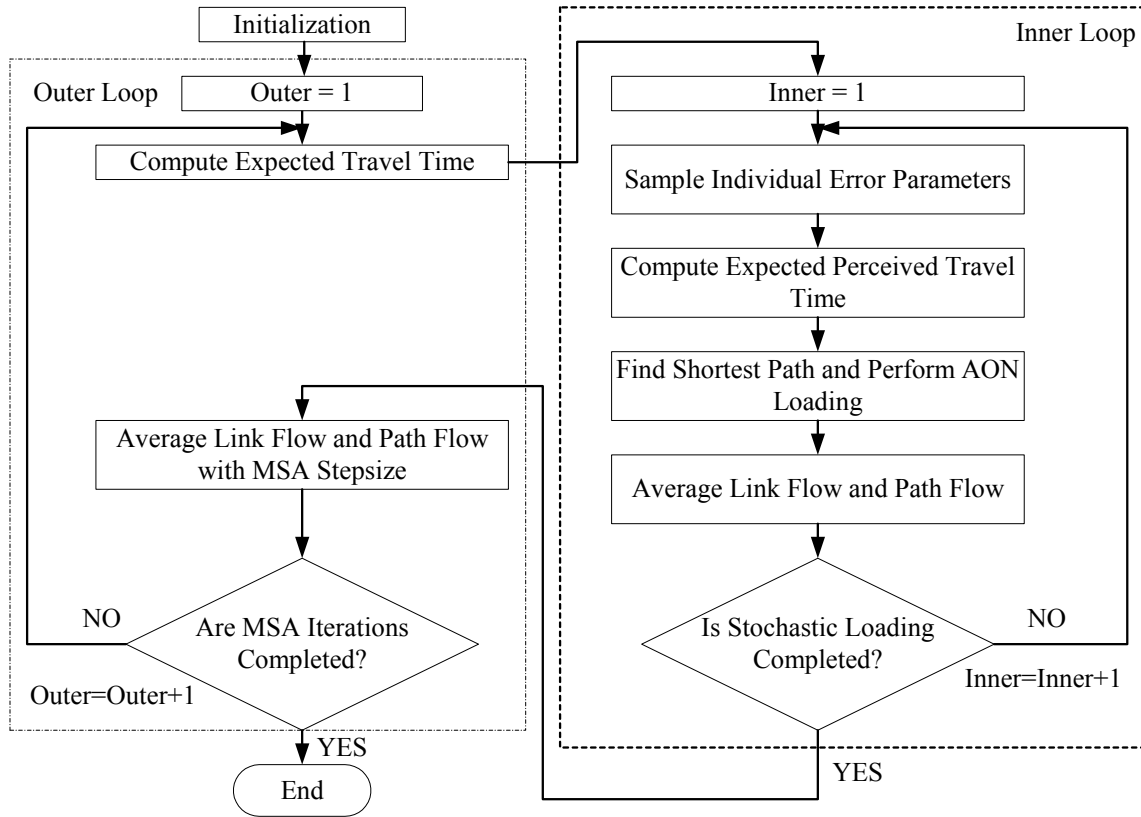
Figure 5.3: Flowchart for Solving the DN-SUE Model

Step 0: *Initialization*. Generate an initial path for each OD pair.

0.1   Set $x_a(0) = 0$, $t_a = t_a[x_a(0)]$, $\forall\ a$ and $K_{rs}(0) = \varnothing$.

0.2   Set iteration counter $n = 1$.

0.3   Solve the shortest path problem: $\bar{k}_{rs}(n)$, $K_{rs}(n) = \bar{k}_{rs}(n) \cup K_{rs}(n-1)$, $\forall\ r, s$.

0.4   Perform AON assignment: $f^{rs}_{\bar{k}_{rs}(n)} = q_{rs}$, $\forall\ r, s$.

0.5   Assign path flows to links: $x_a(n) = \sum\limits_{r \in R} \sum\limits_{s \in S} \sum\limits_{k \in K_{rs}(n)} f_k^{rs}(n)\delta_{ka}^{rs}$, $\forall\, a$.

Step 1: *Increment iteration counter.* $n = n + 1$.

Step 2: *Update link travel time.* $t_a(n) = t_a[x_a(n-1)]$, $\forall\, a$.

Step 3: *Direction Finding.* Perform a stochastic assignment based on *perceived* travel times.

3.1   Set temporary path set $P_{rs}(0) = \varnothing$, $\forall\, r, s$, set $y_a(0) = 0$, $\forall\, a$ and

       $h_k^{rs}(0) = 0$, $\forall\, k, r, s$, and set inner iteration counter $m = 1$.

3.2   Draw perception errors from a Normal distribution: $\varepsilon_a^m \sim N(0, \sigma_a^2)$, $\forall\, a$.

3.3   Compute *perceived* link travel times: $\tilde{t}_a^m = t_a(n) + \varepsilon_a^m$, $\forall\, a$.

3.4   Solve the shortest path problem based on *perceived* travel times: $\bar{k}_{rs}(m)$, $\forall\, r, s$.

3.5   Augment path $\bar{k}_{rs}(m)$ to the path set $P_{rs}(m-1)$ if it is not already in the set:

       If $\bar{k}_{rs}(m) \notin P_{rs}(m-1)$, then $P_{rs}(m) = \bar{k}_{rs}(m) \cup P_{rs}(m-1)$, $\forall\, r, s$; otherwise, tag the

       shortest path among the paths in $P_{rs}(m-1)$ as $\bar{k}_{rs}(m)$ and set

       $P_{rs}(m) = P_{rs}(m-1)$, $\forall\, r, s$.

3.6   Perform all-or-nothing assignment: $g_{\bar{k}_{rs}(n)}^{rs} = q_{rs}$, $\forall\, r, s$.

3.7   Determine temporary link flows: $u_a(m) = \sum\limits_{r \in R} \sum\limits_{s \in S} \sum\limits_{\bar{k}_{rs}(m)} g_{\bar{k}_{rs}(m)}^{rs} \delta_{\bar{k}_{rs}(m)a}^{rs}$, $\forall a$.

3.8   Average auxiliary link flows: $y_a(m) = [(m-1)y_a(m-1) + u_a(m)]/m$, $\forall\, a$

3.9   Average auxiliary path flows:

$h_k^{rs}(m) = [(m-1)h_k^{rs}(m-1) + g_{\bar{k}_{rs}(m)}^{rs}(m)]/m$, $\forall\, k \in P_{rs}(m), r, s$

3.10   If a specified criterion (i.e., $m \geq M$, where M is the maximum number of inner

       iterations) is met, then step and go to Step 3.11. Otherwise, set counter m=m+1

       and go to Step 3.2.

3.11 Update path set: If $P_{rs}(m) \notin K_{rs}(n-1)$, then $K_{rs}(n) = P_{rs}(m) \cup K_{rs}(n-1), \forall r, s$

and tag all the shortest paths found in the inner iteration as $\bar{k}_{rs}(n)$; otherwise, tag the shortest paths among the paths in $K_{rs}(n-1)$ as $\bar{k}_{rs}(n)$ and set

$$K_{rs}(n) = K_{rs}(n-1), \forall r, s .$$

Step 4: *Line Search.* Set $\alpha_n = 1/n$

Step 5: *Move.* Update path and link flows.

5.1 Set path flows: $f_k^{rs}(n) = f_k^{rs}(n-1) - \alpha_n [h_{\bar{k}_{rs}(n)}^{rs} - f_k^{rs}(n-1)], \forall k \in K_{rs}(n), r, s.$

5.2 Assign path flows to links: $x_a(n) = \sum_{r \in R} \sum_{s \in S} \sum_{k \in K_{rs}(n)} f_k^{rs}(n) \delta_{ka}^{rs}, \quad \forall a$

Step 6: *Convergence Test.* Terminate the algorithm if it satisfies the stopping criterion (the current solution, $x_a(n)$, is the set of equilibrium link flows); otherwise, go to Step 1.

5.2.3 Implementation of the SN-DUE Model

The criterion used in the SN-DUE model is to minimize the expected value of route travel disutility, which can be obtained by adding up the disutilities of all the links belonging to the route. The link disutilities are a function of average travel time, travel time variability, and the disutility function describing travelers' risk-taking behavior. The overall flowchart of the traffic assignment procedure for solving the SN-DUE model is given in Figure 5.4 and summarizes below:

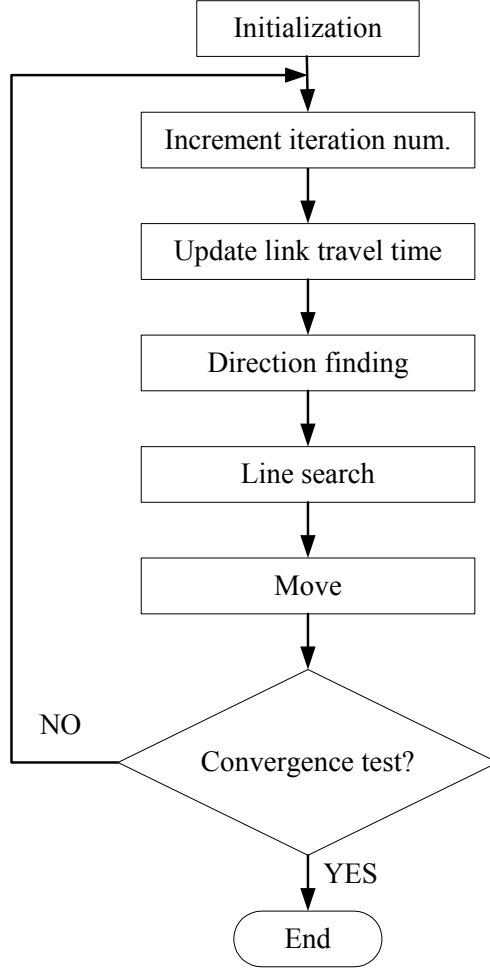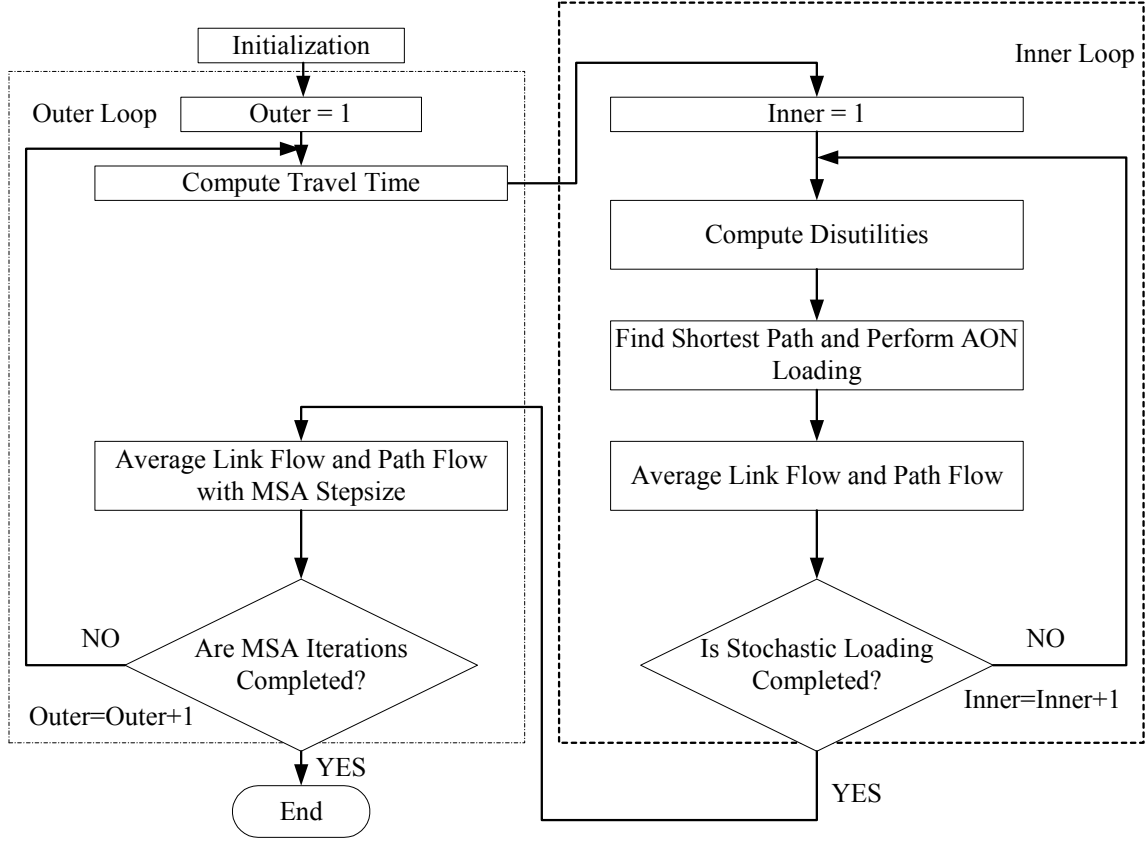**Figure 5.4**: Flowchart for Solving the SN-DUE Model

Step 0: *Initialization.* Generate an initial path for each OD pair.

0.1   Set $x_a(0) = 0$, $t_a = t_a[x_a(0)]$, $\forall\, a$ and $K_{rs}(0) = \varnothing$.

0.2   Set iteration counter $n = 1$.

0.3   Solve the shortest path problem: $\bar{k}_{rs}(n)$, $K_{rs}(n) = \bar{k}_{rs}(n) \cup K_{rs}(n-1)$, $\forall\, r, s$.

0.4   Perform AON assignment: $f^{rs}_{\bar{k}_{rs}(n)} = q_{rs}$, $\forall\, r, s$.

0.5   Assign path flows to links: $x_a(n) = \sum_{r \in R} \sum_{s \in S} \sum_{k \in K_{rs}(n)} f^{rs}_k(n) \delta^{rs}_{ka}$, $\quad \forall\, a$.

Step 1: *Increment iteration counter.* $n = n + 1$.

Step 2: *Update link travel time.* $t_a(n) = t_a[x_a(n-1)], \forall a$.

Step 3: *Direction Finding.* Perform a stochastic assignment based on the link disutilities.

    3.1   Set temporary path set $P_{rs}(0) = \varnothing, \forall r,s$, set $y_a(0) = 0, \forall a$ and

            $h_k^{rs}(0) = 0, \forall k,r,s$, and set inner iteration counter $m = 1$.

    3.2   Compute link disutilities using equation (4.21) and equation (4.23): $U_a^m, \forall a$.

    3.3   Solve the shortest path problem based on link utilities: $\bar{k}_{rs}(m), \forall r,s$.

    3.4   Augment path $\bar{k}_{rs}(m)$ to the path set $P_{rs}(m-1)$ if it is not already in the set:

            If $\bar{k}_{rs}(m) \notin P_{rs}(m-1)$, then $P_{rs}(m) = \bar{k}_{rs}(m) \cup P_{rs}(m-1), \forall r,s$; otherwise, tag the

            shortest path among the paths in $P_{rs}(m-1)$ as $\bar{k}_{rs}(m)$ and set

            $P_{rs}(m) = P_{rs}(m-1), \forall r,s$.

    3.5   Perform all-or-nothing assignment: $g_{\bar{k}_{rs}(n)}^{rs} = q_{rs}, \forall r,s$.

    3.6   Determine temporary link flows: $u_a(m) = \sum\limits_{r \in R} \sum\limits_{s \in S} \sum\limits_{\bar{k}_{rs}(m)} g_{\bar{k}_{rs}(m)}^{rs} \delta_{\bar{k}_{rs}(m)a}^{rs}, \forall a$.

    3.7   Average auxiliary link flows: $y_a(m) = [(m-1)y_a(m-1) + u_a(m)]/m, \forall a$.

    3.8   Average auxiliary path flows:

   $h_k^{rs}(m) = [(m-1)h_k^{rs}(m-1) + g_{\bar{k}_{rs}(m)}^{rs}(m)]/m, \forall k \in P_{rs}(m), r, s$

    3.9   If a specified criterion (i.e., $m \geq M$, where M is the maximum number of inner

            iterations) is met, then step and go to Step 3.10. Otherwise, set counter m=m+1

            and go to Step 3.2.

    3.10   Update path set: If $P_{rs}(m) \notin K_{rs}(n-1)$, then $K_{rs}(n) = P_{rs}(m) \cup K_{rs}(n-1), \forall r,s$,

            and tag all the shortest paths found in the inner iteration as $\bar{k}_{rs}(n)$; otherwise, tag

            the shortest paths among the paths in $K_{rs}(n-1)$ as $\bar{k}_{rs}(n)$ and set

            $K_{rs}(n) = K_{rs}(n-1), \forall r,s$.

Step 4: *Line Search.* Set $\alpha_n = 1/n$

Step 5: *Move.* Update path and link flows.

    5.1   Set path flows: $f_k^{rs}(n) = f_k^{rs}(n-1) - \alpha_n[h_{\bar{k}_{rs}(n)}^{rs} - f_k^{rs}(n-1)], \forall k \in K_{rs}(n), r, s.$

    5.2   Assign path flows to links: $x_a(n) = \sum_{r \in R} \sum_{s \in S} \sum_{k \in K_{rs}(n)} f_k^{rs}(n)\delta_{ka}^{rs}, \quad \forall\, a$

Step 6: *Convergence Test.* Terminate the algorithm if it satisfies the stopping criterion (the current solution, $x_a(n)$, is the set of equilibrium link flows); otherwise, go to Step 1.

5.2.4   Implementation of the SN-SUE Model

The criterion used in the SN-SUE model is to minimize the expected *perceived* value of route travel disutility, which can be obtained by adding up the *perceived* disutilities of all the links belonging to the route. The link *perceived* disutilities are a function of average travel time, travel time variability, assumption of perception error distribution, and the disutility function describing travelers' risk-taking behavior. The overall flowchart of the traffic assignment procedure for solving the SN-SUE model is given in Figure 5.5 and summarizes below:
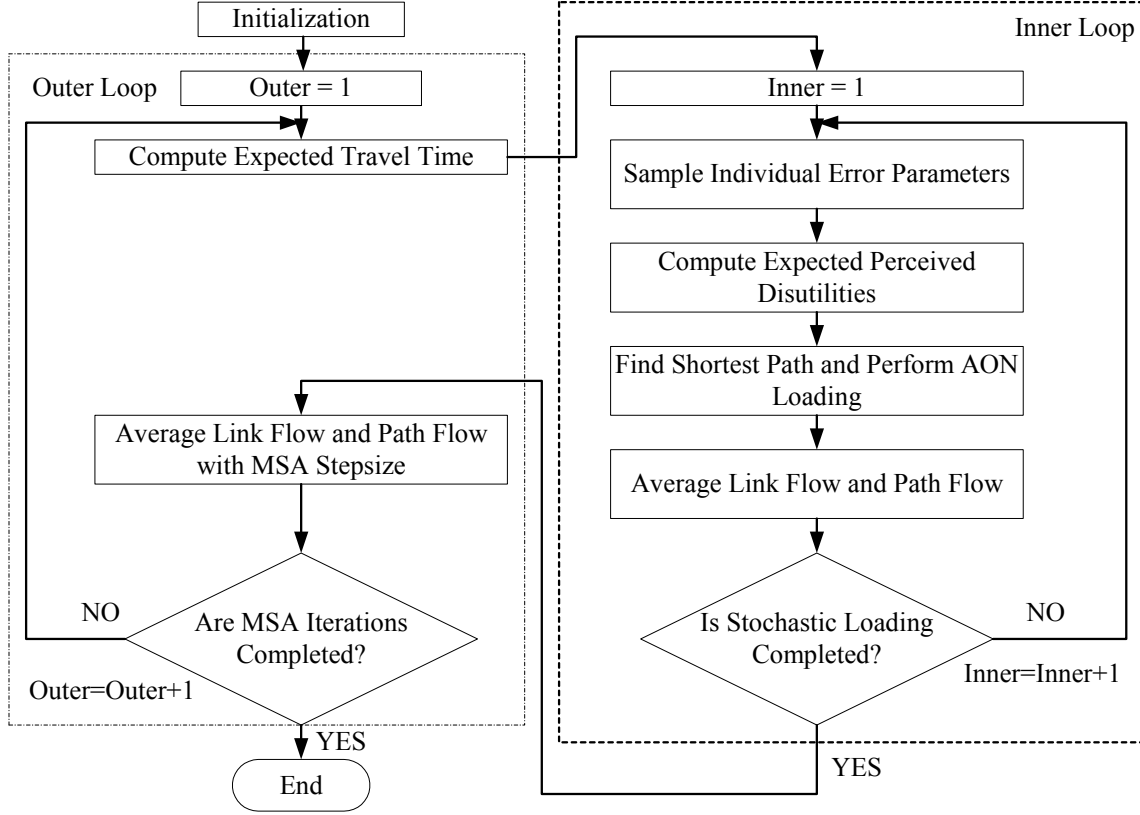
Figure 5.5: Flowchart for Solving the SN-SUE Model

Step 0: *Initialization*. Generate an initial path for each OD pair.

    0.1   Set $x_a(0) = 0$, $t_a = t_a[x_a(0)]$, $\forall\, a$ and $K_{rs}(0) = \varnothing$.

    0.2   Set iteration counter $n = 1$.

    0.3   Solve the shortest path problem: $\bar{k}_{rs}(n)$, $K_{rs}(n) = \bar{k}_{rs}(n) \cup K_{rs}(n-1)$, $\forall\, r, s$.

    0.4   Perform AON assignment: $f^{rs}_{\bar{k}_{rs}(n)} = q_{rs}$, $\forall\, r, s$.

    0.5   Assign path flows to links: $x_a(n) = \sum\limits_{r \in R} \sum\limits_{s \in S} \sum\limits_{k \in K_{rs}(n)} f^{rs}_k(n)\delta^{rs}_{ka}$, $\forall\, a$.

Step 1: *Increment iteration counter*. $n = n + 1$.

Step 2: *Update link travel time.* $t_a(n) = t_a[x_a(n-1)], \forall a$.

Step 3: *Direction Finding.* Perform a stochastic assignment based on the *perceived* link disutilities.

3.1 Set temporary path set $P_{rs}(0) = \varnothing, \forall r,s$, set $y_a(0) = 0, \forall a$ and

$h_k^{rs}(0) = 0, \forall k,r,s$, and set inner iteration counter $m = 1$.

3.2 Draw perception errors: $\varepsilon_a^m \sim N(\mu_a^i, \theta_a^i), \forall a$.

3.3 Compute *perceived* link disutilities using equation (4.20) and equation (4.22):

$\tilde{U}_a^m, \forall a$.

3.4 Solve the shortest path problem based on *perceived* link disutilities: $\bar{k}_{rs}(m), \forall r,s$.

3.5 Augment path $\bar{k}_{rs}(m)$ to the path set $P_{rs}(m-1)$ if it is not already in the set:

If $\bar{k}_{rs}(m) \notin P_{rs}(m-1)$, then $P_{rs}(m) = \bar{k}_{rs}(m) \cup P_{rs}(m-1), \forall r,s$; otherwise, tag the

shortest path among the paths in $P_{rs}(m-1)$ as $\bar{k}_{rs}(m)$ and set

$P_{rs}(m) = P_{rs}(m-1), \forall r,s$.

3.6 Perform all-or-nothing assignment: $g_{\bar{k}_{rs}(n)}^{rs} = q_{rs}, \forall r,s$.

3.7 Determine temporary link flows: $u_a(m) = \sum_{r \in R} \sum_{s \in S} \sum_{\bar{k}_{rs}(m)} g_{\bar{k}_{rs}(m)}^{rs} \delta_{\bar{k}_{rs}(m)a}^{rs}, \forall a$.

3.8 Average auxiliary link flows: $y_a(m) = [(m-1)y_a(m-1) + u_a(m)]/m, \forall a$

3.9 Average auxiliary path flows:

$h_k^{rs}(m) = [(m-1)h_k^{rs}(m-1) + g_{\bar{k}_{rs}(m)}^{rs}(m)]/m, \forall k \in P_{rs}(m), r,s$

3.10 If a specified criterion (i.e., $m \geq M$, where M is the maximum number of inner iterations) is met, then step and go to Step 3.11. Otherwise, set counter m=m+1 and go to Step 3.2.

3.11 Update path set: If $P_{rs}(m) \notin K_{rs}(n-1)$, then $K_{rs}(n) = P_{rs}(m) \cup K_{rs}(n-1), \forall r,s$, and tag all the shortest paths found in the inner iteration as $\bar{k}_{rs}(n)$; otherwise, tag the shortest paths among the paths in $K_{rs}(n-1)$ as $\bar{k}_{rs}(n)$ and set

$K_{rs}(n) = K_{rs}(n-1), \forall r,s$.

Step 4: *Line Search*. Set $\alpha_n = 1/n$

Step 5: *Move*. Update path and link flows.

    5.1   Set path flows: $f_k^{rs}(n) = f_k^{rs}(n-1) - \alpha_n [h_{\bar{k}_{rs}(n)}^{rs} - f_k^{rs}(n-1)], \forall k \in K_{rs}(n), r, s.$

    5.2   Assign path flows to links: $x_a(n) = \sum\limits_{r \in R} \sum\limits_{s \in S} \sum\limits_{k \in K_{rs}(n)} f_k^{rs}(n) \delta_{ka}^{rs}, \quad \forall a$

Step 6: *Convergence Test*. Terminate the algorithm if it satisfies the stopping criterion (the current solution, $x_a(n)$, is the set of equilibrium link flows); otherwise, go to Step 1.

5.3 Numerical Experiments

In this section, we present some numerical results using the four route choice models discussed above. The numerical experiment is constructed to show the effects of using different criteria in the route choice models on the network assignment results, particularly focusing how travel time variability on certain links affect network flow allocations. Figure 5.6 depicts the network topology of the Sioux Falls network. In this network, there are 24 nodes, 76 links, and 182 OD pairs. The OD table is provided in Table 5.2, and the network characteristics (link free-flow travel time and link capacity) are provided in Table 5.3.

5.3.1 Experimental Setup

(1)   Link Travel Time and Travel Time Variability

Sixteen links (highlighted in Figure 5.6) are selected to have a travel time variability (i.e., standard deviation of travel time) that is equal to 70% of its free-flow travel times. All other links are assumed to have no travel time variability. The link cost function is of Bureau of Public Road (BPR) type.

(2)   Perception Error

For the DN-SUE model, each traveler i is assumed to have a fixed perception error $\xi_i \sim N(0, \tau) = N(0, 0.01)$ sampled from a normal distribution (Sheffi, 1985). For the SN-SUE model, each traveler i is assumed to have a variable perception error $\xi_i \sim N(\mu_i, \theta_i)$ where $\mu_i \sim N(0, \tau) = N(0, 0.15)$ is sampled from a normal distribution, and $\theta_i \sim G(\alpha, \beta) = G(0.01, 0.5)$ is sampled from a Gamma distribution (Mirchandani and Soroush, 1987).
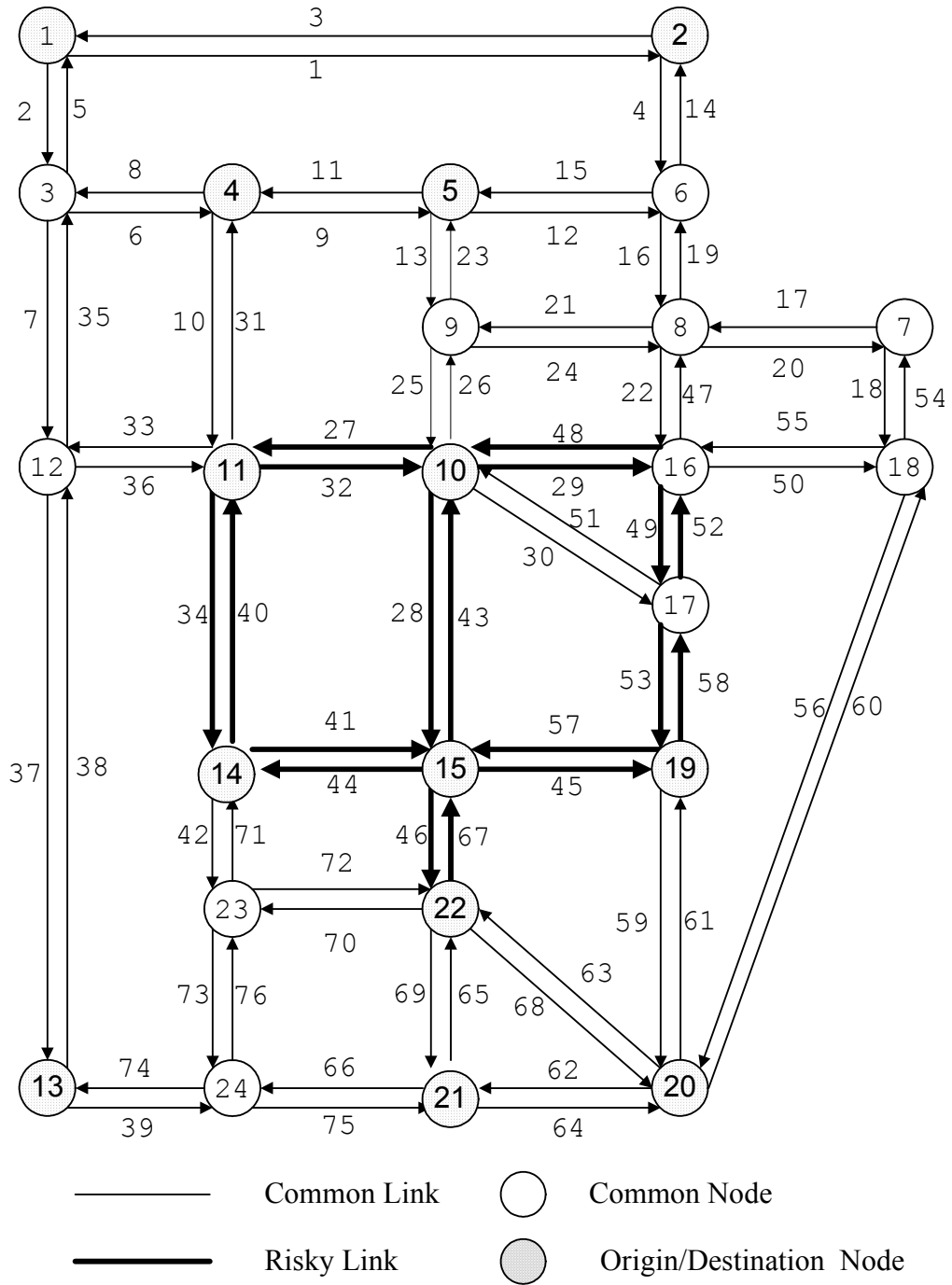
Figure 5.6: The Sioux Falls Network

Table 5.2: Origin-Destination Trip Table (x10^3 veh/hr)

	1	2	4	5	10	11	13	14	15	19	20	21	22	24
1		1.20	1.20	1.20	0.98	1.00	1.14	0.90	0.86	0.82	0.54	0.54	0.70	0.67
2	1.20		1.14	1.18	1.00	1.02	0.82	0.86	0.85	1.18	0.54	0.62	0.61	0.54
4	1.20	1.14		1.20	0.98	0.97	0.86	0.82	0.76	0.73	0.56	0.58	0.54	0.73
5	1.20	1.18	1.2		1.03	0.88	0.83	0.80	0.74	0.66	0.73	0.74	0.85	0.54
10	0.98	1.00	0.98	1.03		1.21	0.82	0.90	1.20	1.06	0.86	0.82	0.88	0.54
11	1.00	1.02	0.97	0.88	1.21		0.85	1.20	1.01	0.86	0.67	0.55	1.00	0.95
13	1.14	0.82	0.86	0.83	0.82	0.85		0.79	0.78	0.62	0.54	0.56	0.61	1.20
14	0.90	0.86	0.82	0.80	0.90	1.20	0.79		1.20	1.03	0.86	0.79	0.82	1.03
15	0.86	0.85	0.76	0.74	1.20	1.01	0.78	1.20		1.20	1.15	1.04	1.20	0.83
19	0.82	1.18	0.73	0.66	1.06	0.86	0.62	1.03	1.20		1.20	1.01	1.00	0.73
20	0.54	0.54	0.56	0.73	0.86	0.67	0.54	0.89	1.15	1.20		1.20	1.20	0.55
21	0.54	0.62	0.58	0.74	0.82	0.55	0.56	0.79	1.04	1.01	1.20		1.20	1.20
22	0.70	0.61	0.54	0.85	0.88	1.00	0.61	0.82	1.20	1.00	1.20	1.20		1.03
24	0.67	0.54	0.73	0.54	0.54	0.95	1.20	1.03	0.83	0.73	0.55	1.20	1.03	

Table 5.3: Link Characteristics of the Sioux Falls Network

Link No.	Capacity (1,000 veh per hour)	Length (mile)	Free-flow travel time (hour)	Link No.	Capacity (1,000 veh per hour)	Length (mile)	Free-flow travel time (hour)
1	25.90	3.6	0.06	39	5.09	2.4	0.04
2	23.40	2.4	0.04	40	4.88	2.4	0.04
3	25.90	3.6	0.06	41	5.13	3.0	0.05
4	4.96	3.0	0.05	42	4.92	2.4	0.04
5	23.40	2.4	0.04	43	13.51	3.6	0.06
6	17.11	2.4	0.04	44	5.13	3.0	0.05
7	23.40	2.4	0.04	45	15.65	2.4	0.04
8	17.11	2.4	0.04	46	10.32	2.4	0.04
9	17.78	1.2	0.02	47	5.05	3.0	0.05
10	4.91	3.6	0.06	48	5.13	3.0	0.05
11	17.78	1.2	0.02	49	5.23	1.2	0.02
12	4.95	2.4	0.04	50	19.68	1.8	0.03
13	10.00	3.0	0.05	51	4.99	4.8	0.08
14	4.96	3.0	0.05	52	5.23	1.2	0.02
15	4.95	2.4	0.04	53	4.82	1.2	0.02
16	4.90	1.2	0.02	54	23.40	1.2	0.02
17	7.84	1.8	0.03	55	19.68	1.8	0.03
18	23.40	1.2	0.02	56	23.40	2.4	0.04
19	4.90	1.2	0.02	57	15.65	2.4	0.04
20	7.84	1.8	0.03	58	4.82	1.2	0.02
21	5.05	6.0	0.10	59	5.00	2.4	0.04
22	5.05	3.0	0.05	60	23.40	2.4	0.04
23	10.00	3.0	0.05	61	5.00	2.4	0.04
24	5.05	6.0	0.10	62	5.06	3.6	0.06
25	13.92	1.8	0.03	63	5.08	3.0	0.05
26	13.92	1.8	0.03	64	5.06	3.6	0.06
27	10.00	3.0	0.05	65	5.23	1.2	0.02
28	13.51	3.6	0.06	66	4.89	1.8	0.03
29	5.13	3.0	0.05	67	10.32	2.4	0.04
30	4.99	4.8	0.08	68	5.08	3.0	0.05
31	4.91	3.6	0.06	69	5.23	1.2	0.02
32	10.00	3.0	0.05	70	5.00	2.4	0.04
33	4.91	3.6	0.06	71	4.92	2.4	0.04
34	4.88	2.4	0.04	72	5.00	2.4	0.04
35	23.40	2.4	0.04	73	5.08	1.2	0.02
36	4.91	3.6	0.06	74	5.09	2.4	0.04
37	25.90	1.8	0.03	75	4.89	1.8	0.03
38	25.90	1.8	0.03	76	5.08	1.2	0.02

(3)   Disutility Function

To model the travelers' attitudes towards risk, an exponential disutility function is used to represent the risk averse behavior of the travelers.

$$\tilde{\eta}\left(\tilde{T}_k\right)= a_1 \left(e^{a_2 \tilde{T}_k} - 1\right), \qquad (5.24)$$

where $\tilde{T}_k$ is the *perceived* travel time distribution on route k; a_1, and a_2 are positive parameters to be estimated. The parameters a_2 is a measure of risk aversion. As a_2 gets larger, the traveler becomes more averse to risk. With appropriate boundary conditions, the final form of the disutility function used in the implementation is:

$$\tilde{\eta}\left(\tilde{T}_k\right)= 0.310 \left(e^{7.22 \tilde{T}_k} - 1\right). \qquad (5.25)$$

Both SN-DUE and SN-SUE models used the above *perceived* disutility function as the criterion in the route choice decision. The only difference between these two models is that SN-DUE is without perception error (i.e., $\xi_i \sim N(\mu_i, \theta_i) = N(0,0)$) and SN-SUE is with perception error (i.e., $\xi_i \sim N(\mu_i, \theta_i)$ where the parameters are set as above).

5.3.2   Comparison of Different Route Choice Models

Table 5.4 provides some aggregate measures (total travel time, number of used routes, and number of routes per OD pair) that can be used to compare the assignment results from the four route choice models. In terms of total travel time, the DN-DUE model has the lowest value, followed by the DN-SUE, SN-SUE, and SN-DUE models. Since the DN-DUE and DN-SUE models are based on the assumption that travelers choose routes to minimize expected (*perceived*) travel times, it can be expected to have lower total travel time when compared to the two SN route choice models. That is, the risk averse travelers in the SN-DUE and SN-SUE models are willing to choose routes that are longer to avoid the sixteen risky links containing travel time variability. This observation can be verified by comparing the link-flow patterns provided in Table 5.5. Indeed, the flow allocations to these sixteen risky links are much less than those allocated by the two risk neutral models (DN-DUE and DN-SUE). Between the DN-DUE

126

and DN-SUE models, the DN-SUE model has a slightly higher total travel time. This is probably due to the perception error. However, it should be recognized that the DN-SUE model could have lower total travel time than the DN-DUE model in certain situations. As for the SN-DUE and SN-SUE models, the perception error in the SN-SUE model may overshadow the risk averse behavior and cause some travelers to use the risky links that result in a slightly lower total travel time than those in the SN-DUE model which assumes to have perfect knowledge about the variable nature of network travel times. In terms of the number of routes and number of routes per OD pair, the DN-DUE model is the lowest (1.12 routes per OD or 204 routes in total) and highest for the SN-SUE model (6.68 routes per OD or 1215 routes in total). This set of results can be explained using arguments similar to above.

Table 5.4: Aggregate Measures from the Four Route Choice Models

	DN-DUE	DN-SUE	SN-DUE	SN-SUE
Total travel time (Thousand vehicles hour)	21.88	24.60	47.89	41.38
Number of used paths*	204	944	600	1215
Number of paths per OD pair	1.12	5.19	3.30	6.68

* Only the path with path flow greater than 0.01 (Thousand vehicles) are counted

Table 5.5 provides the link-flow results on the sixteen links containing travel time variability for all four route choice models. As expected and mentioned above, traffic flows assigned to these risky links are much lower for the SN-DUE and SN-SUE models compared to those in the DN-DUE and DN-SUE models. This is because the risk averse travelers in the SN-DUE and SN-DUE models choose routes to avoid these risky links. Between the DN-DUE and DN-SUE models (without and with perception error), no particular trend can be observed within the subset of network links. As for the SN-DUE and SN-SUE models, it appears that the flow allocations by the SN-DUE model are similar those assigned by the SN-SUE model in this set of risky links. In general, we would expect some distinct differences between the DUE and SUE models, but these differences might be overshadowed by the congestion level as illustrated in Sheffi and Powell (1981). As the congestion level increases, the differences between the DUE and SUE models become more and more similar. This leads us to examine the effects of the congestion level.

Table 5.5: Link-Flow Results from the Four Route Choice Models (thousand vehicles)

Link ID	DN-DUE	DN-SUE	SN-DUE	SN-SUE
27	5.012	8.597	0.158	1.222
28	10.879	10.710	0.229	3.337
29	3.395	3.405	0.095	1.371
32	5.008	8.250	0.161	1.409
34	7.039	7.957	3.532	4.911
40	7.034	7.983	3.550	4.741
41	3.940	6.130	0.294	1.849
43	10.884	10.558	0.243	2.978
44	3.952	6.091	0.300	2.079
45	6.634	8.116	0.306	2.515
48	3.395	3.609	0.095	1.050
48	6.020	4.418	0.227	1.314
52	6.021	4.328	0.229	1.002
53	6.020	5.570	3.065	3.263
57	6.652	8.105	0.307	2.412
58	6.021	5.659	3.094	3.285

5.3.3 Comparison of Route Choice Models under Different Demand Levels

To examine the effect of congestion, we vary the demand by uniformly multiplying the trip table in Table 5.2 by a demand control factor of 0.6, 0.8, 1.2 and 1.4. The original trip table has a control demand factor of 1.0. To assess the effect of congestion in flow allocations of the four route choice models, we use the link flows obtained from the DN-DUE model as the benchmark, and calculate the deviations for each of the other three route choice models as follows:

$$Dev = \sum_a \frac{\left| x_a^{DN\_DUE} - x_a^{Model} \right|}{x_a^{DN\_DUE}} \cdot 100\%, \tag{5.26}$$

where *Dev* is the link-flow deviations from the DN-DUE model; *Model* here refers to one of the DN_SUE, SN_DUE and SN_SUE models; and *a* refers to all links in the network or just the sixteen links containing travel time variability.

Table 5.6 provides the link-flow deviations of the three route choice models for different demand control factors. For each route choice model, two sets of link-flow deviation results are provided: one for all links in the network and the other for the sixteen risky links. As the demand control factor increases, the deviations get smaller for both sets of links (risky links in the DN-SUE model is exceptional because this set of links may not reflect the effects of congestion). This suggests that congestion does play a role in the flow allocation of the route choice models. Intuitively, this result is expected since all route choice models are assumed to be flow-dependent based on the BRP function. As the level of congestion increases, there are less alternative routes for those with no perception error. Hence, the DN-DUE and DN-SUE models would have similar flow allocations. This result agrees with the study by Sheffi and Powell (1981). As for the SN-DUE and SN-SUE models, the criterion used for route choice is different (i.e., both average travel time and travel time variability) and the route choice behavior is also different (i.e., risk averse as opposed to risk neutral); thus, the flow allocations would also be different from the DN-SUE model. However, the flow differences for all links between the SN-DUE and SN-SUE models are minor as the demand control factor increases.

Table 5.6: Link-Flow Deviation Comparison as a Function of Demand Control Factor

Demand Control Factor	DN-SUE		SN-DUE		SN-SUE	
	All Links	Risky Links	All Links	Risky Links	All Links	Risky Links
0.6	17.9	20.6	80.7	96.3	65.8	75.6
0.8	13.9	21.4	73.3	91.9	61.6	68.0
1.0	12.5	21.2	66.8	83.8	59.0	60.4
1.2	11.7	18.4	60.2	78.0	56.0	53.0
1.4	11.4	14.7	55.3	75.9	50.7	47.1

5.3.4 Summary

In this chapter, we described the implementation of traffic assignment procedures for solving the four route choice models described in Chapter two. In addition, numerical experiments were conducted to examine the effects of using four route choice models on network flow allocations. The results showed that there are significant differences on the flow allocations between the risk neutral travelers (in the DN-DUE and DN-SUE models) and the risk averse travelers (in the SN-DUE and SN-SUE models) when there are travel time variability in the network. This is because of the criterion (expected disutilities rather than expected travel times) and the risk averse behavior used in the route choice decision process. Considering several recent empirical studies on travel behaviors, they all suggested that travel time variability is an important factor for travelers when making their route choice decisions. Thus, it is necessary to incorporate travel time variability into the route choice model to enhance the predictive capability of the route choice model. In addition, it is important to systematically examine the behavioral issues related to perceptions of travel time variability and risk preferences and how these issues affect individual route choice decisions and collectively on network assignment results.

Chapter Six

6.   Travel Time Reliability

6.1   Introduction

The potential sources of disruption to transportation networks are numerous, ranging from one extreme of major exceptional events (e.g., natural or man-made disasters) to the other extreme of minor regular events (e.g., daily recurrent congestion). The scale, impact, frequency, and predictability of these abnormal events also vary enormously. Disruptions to the transportation network can have a major impact not only on economic productivity but also on our daily lives. The negative impacts of transportation disruptions can be pervasive. Perhaps, the best example to illustrate the vulnerability of transportation networks is the air traffic network after the 9/11 Event. It completely halted all commercial air traffic in the United States and the impacts on travel disruptions and economic losses were spread worldwide.

In fact, the transportation system has been identified as the most important lifeline (Nicholson and Du , 1997) in the event of a disaster (e.g., earthquakes, floods, tornadoes, hurricanes, landslides, and others), Disasters of these kinds in recent years have exposed the vulnerability of lifeline systems and the need to mitigate the risk consequent to failure of these systems. In particular, recent events such as the 1989 Loma Prieta, 1994 Northridge, California, the 1995 Kobe, Japan, and the recent earthquakes in Turkey, Taiwan, Mexico, and India have provided compelling evidence that transportation systems are of paramount importance to restoring normalcy. Because the restoration of other lifelines (e.g., water supply, electrical power, sewer, communication, and others) depends strongly on the ability to move people and equipment to damaged sites, an unreliable transportation system would therefore hinder the restoration process and increase not only economic loss but also fatalities that are difficult to quantify.

Moreover, when planning a reliable transportation system, one should consider not only disasters, but also everyday disturbances, such as traffic congestion arising from daily traffic peaking or irregular traffic accidents. The reliability of the transportation system reflects the quality of service it would normally provide under different disruptions. In the 21st century, Iida (1999) argued that because of increased value of time, travelers would desire a more stable transportation system with less travel time uncertainty so that they can be confident of arriving at their destination on schedule. In addition, the value of time has increased in recent years due to increased economic activity and improvements in the quality of life. Road networks are now required to have high degree of reliability to ensure drivers smooth travel under normal traffic flow fluctuations and to avoid serious unexpected delays due to disruptions within the network. These new requirements have generated an urgent need for better understanding of road network reliability under recurrent and non-recurrent congestion.

6.2   Review of Transportation Reliability Measures

Reliability is generally defined as the probability that the system of interest has the ability to perform an intended function or goal (Ang and Tang, 1990.). Traditionally, transportation network reliability studies were primarily concerned with two problems: *connectivity reliability* and *travel time reliability* (Bell and Ida, 1997). *Connectivity reliability* is concerned with the probability that network nodes are connected (Iida, and Wakabayashi, 1989). *Travel time reliability* is concerned with the probability that a trip between a given OD pair can be made successfully within a given time interval or a specified level-of-service (Asakura and Kashiwadanu, 1991; Asakura, 1998; Bell *et al*, 1999). Recently, the reliability of transportation networks has become an increasingly important issue due to its critical status as the most important lifeline in the restoration process following the occurrence of a disaster (Nicholson and Du, 1997). It has attracted many researchers to develop various indicators to assess the reliability of a transportation network. Here we summarize these reliability indicators in Table 6.1 and briefly describe them below.

Connectivity reliability is concerned with the probability that network nodes are connected. A special case of connectivity reliability is the terminal reliability which concerns the existence of a

132

path between a specific O-D pair (Iida, and Wakabayashi, 1989). For each node pair, the network is considered successful if at least one path is operational. A path consists of a set of roadways or links which are characterized by zero-one variables denoting the state of each link (operating or failed). Capacity constraints on the links are not accounted for when determining connectivity reliability. This type of connectivity reliability analysis may be suitable for abnormal situations, such as earthquakes, but there is an inherent deficiency in the sense that it only allows for two operating states: operating at full capacity or complete failure with zero capacity. The binary state approach limits the application to everyday situations where arcs are operating in-between these two extremes. Therefore, the reliability and risk assessment results obtained through this approach may be misleading for normal conditions.

Travel time reliability is concerned with the probability that a trip between a given O-D pair can be made successfully within a given time interval and a specified level-of-service (Asakura and Kashiwadanu, 1991; Bell *et al.*, 1999). This measure is useful when evaluating network performance under normal daily flow variations. Bell *et al.* (1999) proposed a sensitivity analysis based procedure to estimate the variance of travel time arising from daily demand fluctuations. Asakura (1998) extended the travel time reliability to consider capacity degradation due to deteriorated roads. He defined travel time reliability as a function of the ratio of travel times under the degraded and non-degraded states. This definition of reliability can be used as a criterion to define the level of service that should be maintained despite the deterioration certain links in the network. Chen and Recker (2001) further examined the effects of considering risk-taking behavior in the calculation travel time reliability.

Travel demand reduction is concerned with the probability that decrement rate of OD flow is less than a given intolerable value under a degradable network. Nicholson and Du (1997) provide two definitions to evaluate the reduction of travel demand: OD sub-system and system reliabilities. The decrement rate of OD flow is defined as the ratio of the travel demand reduction due to degradation of network over the travel demand of a non-degradable network. The flow decrement rate can vary between zero (i.e., no degradation) and unity (i.e., degradation is so severe that the travel demand is zero). A system surplus was also suggested as a performance measure to assess the socioeconomic impacts of the system degradation.

Travel demand satisfaction reliability (DSR) is concerned with the probability that the network can accommodate a given travel demand satisfaction ratio. This ratio is defined as the equilibrium travel demand (i.e., travel demand that can be satisfied by using the transportation network) over the latent travel demand (i.e., total travel demand that intends to use the transportation network). The latent travel demand is the sum of the equilibrium travel demand or the satisfied travel demand and non-satisfied travel demand stemmed from the latent demand events (LDE) such as horseracing or weekend events. The key feature of this reliability measure is that it attempts to distinguish the difference between recurrent travel demand and latent travel demand under a degradable transportation network.

Encountered Reliability is concerned with the probability that a trip can be made successfully without encountering link degradation on the least (expected) cost path. Level of information to the users is important in the encountered reliability since users will often try to avoid degraded links and links which may be degraded. In addition, different users may behave differently. Risk averse users are concerned with avoiding disruptions that they are willing to travel longer, while risk neutral users will still travel on their preferred routes based on expected cost considerations regardless of the probability of encountering disruptions along their preferred routes.

Capacity reliability is concerned with the probability that the network capacity can accommodate a certain volume of traffic demand at a required service level (Chen *et al*., 1999, 2000, 2002). Link capacities for a road network can change from time to time due to various reasons such as the blockage of one or more lanes due to traffic accidents, and are considered as random variables. The joint distribution of random link capacities can be experimentally obtained or theoretically specified. Capacity reliability explicitly considers the uncertainties associated with link capacities by treating roadway capacities as continuous quantities subject to routine degradation due to physical and operational factors. Readers may note that when the roadway capacities are assumed to take only discrete binary values (zero for total failure and one for operating at ideal capacity), then capacity reliability includes connectivity reliability as a special case. This measure addresses the issue of adequate capacity planning for a highway network in order to accommodate the growing passenger traffic demand. It has the potential of being useful

at the system level, in planning roadway capacity expansion projects, and planning the timing and location of various road improvement projects.

Table 6.1: Principal Characteristics of Definitions of Road Network Reliability
(Modified from Lam and Zhang, 2000)

Reliability Index	Uncertainty	Performance Indicator	Probability Definition	Reliability Aspect for	
				Users	Planners
Connectivity (Iida and Wakayabashi, 1989)	Disruption of road links	$\theta_c=1$ if connect and $\theta_c=0$ if disconnect	Connected and disconnected network	Minimal Usefulness	Good Usefulness
Travel time (threshold based) (Asakura and Ksahiwadani, 1991)	Fluctuation of daily traffic flow	Specified travel time θ_t	Travel time less than a specified value	Good Usefulness	Good Usefulness
Travel time (level-of-service base) (Asakura, 1998)	Degradable link capacity	Specified network service θ_L	Service level less than a specified value	Minimal Usefulness	Good Usefulness
Travel demand reduction (Nicholson and Du, 1997)	Degradable link capacity	Intolerable decrement rate of OD flow θ_o	Decrement rate less than a specified value	Minimal Usefulness	Good Usefulness
Travel demand satisfaction (Lam and Zhang, 2000)	Degradable link capacity	Demand ratio	Travel demand satisfaction ratio greater than a specified value	Good Usefulness	Minimal Usefulness
Encountered reliability (Bell and Schmocker, 2002)	Disruption or degradation of road links	Least costs	Not encountering a link degradation	Good Usefulness	Minimal Usefulness
Network capacity (Chen *et al.*, 1999, 2000, 2002)	Degradable link capacity	Required demand level θ_D	Network reserve capacity greater than a specified value	Minimal Usefulness	Good Usefulness

6.3 Travel Time Reliability Evaluation Procedure

In this research, we are primarily interested in the effects of route choice models on estimating travel time reliability under demand and supply variations. Travel time reliability is concerned with the probability that a trip between a given OD pair can be made successfully within a given time interval and at a specified level-of-service (Asukura and Kashiwadani, 1991; Bell *et al*., 1999). Under this reliability measure, there are two potential measures - path travel time reliability and OD travel time reliability - that are of interest to the travelers and traffic managers (Bell *et al*., 1999). Path travel times are computed by summing up the link travel time on a given route. For the OD travel times, they are computed as a weighted average of the path travel times, where the weights are the path flows.

- Path travel time reliability is defined as the probability that the travel time of a given path is within an acceptable threshold.
- OD travel time reliability is defined as the probability that the weighted average travel time of a given OD pair is within an acceptable threshold.

Travelers are more concerned about the path travel time reliability, because it directly affects their route choice decisions. OD travel time reliability measures all relevant paths used by travelers to define an aggregate measure for the level-of-service between a given OD pair. Thus, traffic managers or planners can use the OD travel time reliability as a proxy to evaluate the performance of a given OD pair (i.e., OD sub-system reliability in terms of travel time). Similar to the system and OD sub-system reliability measure defined by Nicholson and Du (1997), the OD travel time reliability can also be extended to the system level to measure the performance of the whole network.

Travel Time Reliability Estimation Procedure

The travel time reliability evaluation procedure presented below is based upon a Monte Carlo simulation framework developed by Chen and Recker (2001). Link capacities (C_a) and OD demands (Q_{rs}) are treated as random variables. Assuming that the random variables C_a and Q_{rs} follow a known probability distribution, a random variate generator is used to generate the values of C_a for each link and Q_{rs} for each OD pair that preserve the provided distribution properties (e.g., mean, variance, correlation, etc.). For each set of link capacities and OD demands generated, four different route choice models discussed above are used to estimate the OD and path travel time reliabilities.

The travel time reliability procedure is described as follows:

Step 0:   Set sample number $k := 1$.

Step 1:   Generate a vector of OD demands $\mathbf{Q}_k = (\ldots, Q_{rs}, \ldots)$ and/or a vector of arc capacities $\mathbf{C}_k = (\ldots, C_a, \ldots)$.

Step 2:   Perform DN-DUE, DN-SUE, SN-DUE, and SN-SUE route choice models with same the OD demand vector $\mathbf{Q}_k$ and arc capacity vector $\mathbf{C}_k$.

Step 3:   Collect statistics from Step 2 to compute travel time reliability.

Step 4:   If sample number k is less than the required sample size k_{max}, then increment sample number $k := k + 1$ and go to Step1. Otherwise, terminate the procedure.

6.4   Numerical Experiments

The travel time reliability evaluation procedure is demonstrated using a small network depicted in Figure 6.1. This network consists of 5 nodes, 7 links, and 2 OD pairs. The base demand for OD pairs (1, 4) and (1, 5) are 15.0 and 18.8, respectively.
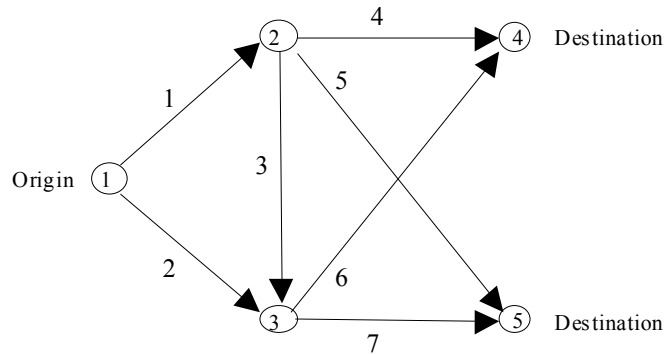
Figure 6.1: Test Network

The free-flow travel times, for each link on the network, as well as, the ideal capacity values, are shown in Table 6.2. The link travel time function, used in the route choice models, is the standard Bureau of Public Road (BPR) function given below:

$$t_a = t_a^f \left(1 + 0.15 \left(\frac{v_a}{c_a} \right)^4 \right),$$

(6.1)

where v_a, t_a^f, and c_a are the flow, free-flow travel time, and capacity on link a.

Table 6.2: Link Free-flow Travel Times and Ideal Link Capacities

Link #	Free-flow travel time (min)	Capacity (veh/min)	Travel time variability (min)
1	4.0	25.0	4.0
2	5.2	25.0	0.52
3	1.0	15.0	0.1
4	5.0	15.0	0.5
5	5.0	15.0	0.5
6	4.0	15.0	0.4
7	4.0	15.0	0.4

6.4.1 Experimental Setup

In order to model the risk-taking route choice models (i.e., SN-DUE and SN-SUE), we assume that all links in the network have some variability associated with their travel times. For link 1, the variability (e.g., standard deviation[2]) of the travel time is assumed to be 100 percent of the link's free-flow travel time. For all other links, the travel times are assumed to vary up to 10 percent of the link's free-flow travel time.

Travel time variability introduces uncertainty for travelers such that they do not know exactly when they will arrive at the destination. Thus, it is considered as a risk (or an added cost) to a traveler making a trip. To model the travelers' attitudes towards risk, disutility functions are needed to represent different risk-taking behaviors. Following the study by Tatineni *et al*. (1997), we also use the exponential functional form to describe risk averse and risk prone behaviors. For the risk neutral behavior, we simply use the DN-SUE model, which minimizes the perceived expect travel time and ignores travel time variability. Suppose that the travelers' disutility has the following forms:

Risk Averse: $U(T) = a_1 [\exp(a_2 T) - 1]$,   (6.2)

Risk Prone: $U(T) = b_1 [1 - \exp(-b_2 T)]$,   (6.3)

---

[2] Other measure such as the difference between the 95th percentile travel time and median travel time (Lam and Small, 2001) can also be used here.

where T is the random travel time in minutes for a given route, a_1, a_2, b_1, and b_2 are positive parameters to be estimated. The parameters a_2 (b_2) are a measure of risk aversion (proneness). As a_2 (b_2) gets larger, the traveler becomes more averse (seeking) to risk. To estimate the parameters of the disutility function given in Equations (6.2) and (6.3), we need to make assumptions on how travelers value route travel time according to their risk-taking behavior. For illustration, the estimation of parameters for the risk averse case is explained. Suppose that all travelers have a disutility of 1 unit for a route that is 5 minutes, and they are indifferent between a route that has a travel time of 3.33 minutes for certain and a route that has an equal chance of having a travel time that is either close to zero minute or is equal to 5 minutes (i.e., an average of 2.5 minutes on this route). Travelers who took the longer travel time route (3.33 minutes) would pay a risk premium of 0.83 minute (i.e., 3.33 − 2.5) to avoid uncertainty. This is a typical behavior for risk averse travelers. Given these two assumptions, two equations, to solve for the values of the two parameters a_1 and a_2, are established as follows:

$$U(5) = a_1 \left[\exp(a_2\, 5) - 1 \right] = 1.0 \tag{6.4}$$

$$U(3.33) = a_1 \left[\exp(a_2\, 3.33) - 1 \right] = 0.5 \tag{6.5}$$

Solving the above equations for a_1 and a_2 give the following disutility function for a risk averse traveler:

$$\text{Risk Averse:} \quad U(T) = 0.309 \left[\exp(0.289\, T) - 1 \right] \tag{6.6}$$

Using similar assumptions for the risk prone travelers, the disutility function for a risk prone traveler is:

$$\text{Risk Prone:} \quad U(T) = 1.309 \left[1 - \exp(-0.289\, T) \right] \tag{6.7}$$

The details of the estimation of the parameters are described in Section 5.1.2 and can also be found in Tatineni (1996). An advantage of using the exponential functional form to represent risk-taking behavior is that the disutility associated with a route can be estimated by summing

the link disutilities on a given route (see Section 5.1.4 on the estimation of link disutilities using the moment generating function). This allows the classical Dijkstra-type shortest path algorithm to be used in finding the minimum expected disutility route in the SN-DUE model and the minimum expected *perceived* disutility route in the SN-SUE model (see Section 5.1.5 on the choice of stochastic loading).

To simulate the uncertainty of the traffic demands, a normal distribution is used. Three levels of mean demand are specified for each OD pair by varying the base OD demand by ± 25% (see Table 6.3). For each level of mean OD demand, three levels of standard deviation (σ) of the OD demand is specified to reflect the variation of the OD demand as follows: (1) $\sigma = 1.5(q_{rs}/3)$ for high variation, (2) $\sigma = 1.0(q_{rs}/3)$ for medium variation, and (3) $\sigma = 0.5(q_{rs}/3)$ for low variation. These three variations of traffic demands can be interpreted as the relative accuracy of each level of mean OD demand (Mette and Bent, 1996). Given that both the mean and standard deviation of OD demands are defined for a total of nine combinations (3 means x 3 standard deviations), random samples for each OD pair can be generated according to the standard normal distribution as follows:

$$Q_{rs} = q_{rs} \pm Z\sigma, \tag{6.8}$$

where:

Q_{rs} = random demand between OD pair (*r,s*),

Z = random variable generated from N(0,1),

σ = standard deviation of the OD demand.

Table 6.3: Three Levels of Mean Demands between Each OD Pair

Origin	Destination	Low mean demand	Medium mean demand (base case)	High mean demand
1	4	N(11.25, σ)	N(15.00, σ)	N(18.75, σ)
1	5	N(14.10, σ)	N(18.80, σ)	N(23.50, σ)

σ is the standard deviation of the OD demand which reflects the variation of the OD demand.

Road capacity is not deterministic since it is influenced by a wide variety of physical and operational characteristics (e.g. roadway type, geometric configuration, mixture of vehicle types, weather conditions, etc.). These characteristics often degrade the capacity of a road network. To simulate the uncertainty of link capacities, let $\mathbf{C} = \mathbf{C}_0 - \varepsilon$, where $\mathbf{C} = \{\dots, C_a, \dots\}$ is a vector of random link capacities, $\mathbf{C}_0 = \{\dots, C_{a0}, \dots\}$ is a vector of ideal link capacities, and ε is a vector of random variables representing link capacity degradation. The random link capacity degradation vector (ε) ranges from 0 (i.e., operating at ideal capacity) to $\mathbf{C}_0$ (i.e., complete failure with zero capacity). For simplicity, a uniform distribution is used to generate the random link capacity degradation. The mean link capacity degradation is 25% of the ideal link capacity with three levels of standard deviation. To ensure high accuracy, 10,000 samples are used to generate the numerical results of the simulation experiment for the case study. As shown in Figure 6.2, the simulated distributions of OD demand for the medium mean demand case are approximately normal with predefined mean and standard deviations. The slight skewness in the high level of standard deviation is due to truncation error of negative OD demands. Similarly, the simulated link capacities also replicate the theoretical values of the predefined uniform distribution, as indicated in Table 6.4. Therefore, the number of samples drawn is sufficient to study the travel time reliability measures.

Three scenarios are constructed to test how different route choice models respond to uncertain environment caused by demand and supply variations and its effects on estimating travel time reliability. Scenario I examines the effects of route choice models on estimating travel time reliability caused by daily fluctuations of demand only; scenario II considers both demand and supply variations; and scenario III examines the effects of risk-sensitive travelers on the estimation of path travel time reliability.
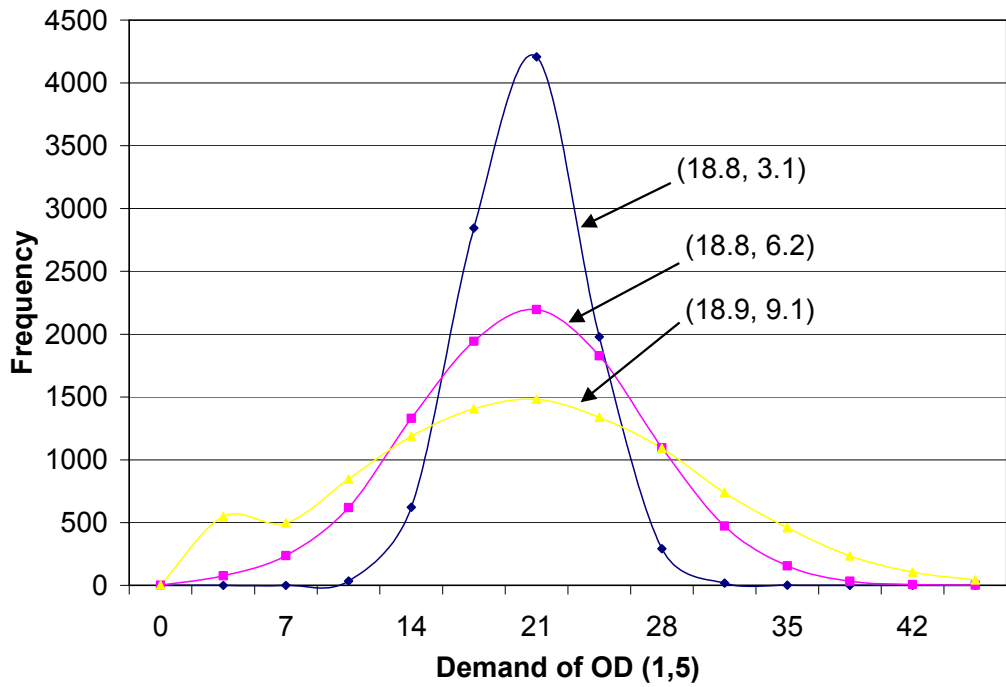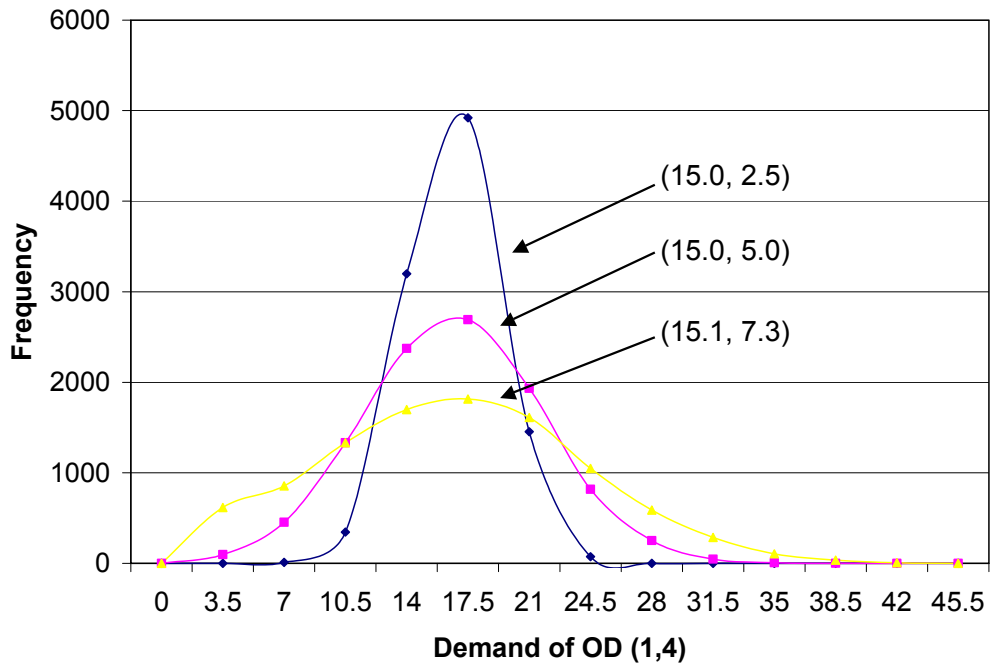
Figure 6.2: Estimated OD Demand Distributions Using 10,000 Samples

Table 6.4: Comparisons of Theoretical and Estimated Link Capacity

Link Number	Low Standard Deviation			
	Theoretical		Estimated	
	Mean	Std. Dev.	Mean	Std. Dev.
1	18.75	0.9021	18.7457	0.9050
2	18.75	0.9021	18.7470	0.9005
3	11.25	0.5413	11.2553	0.5464
4	11.25	0.5413	11.2569	0.5421
5	11.25	0.5413	11.2457	0.5414
6	11.25	0.5413	11.2454	0.5452
7	11.25	0.5413	11.2456	0.5412
	Medium Standard Deviation			
1	18.75	1.8042	18.7413	1.8099
2	18.75	1.8042	18.7440	1.8009
3	11.25	1.0825	11.2607	1.0928
4	11.25	1.0825	11.2638	1.0842
5	11.25	1.0825	11.2415	1.0829
6	11.25	1.0825	11.2408	1.0905
7	11.25	1.0825	11.2411	1.0823
	High Standard Deviation			
1	18.75	2.7063	18.7370	2.7148
2	18.75	2.7063	18.7409	2.7013
3	11.25	1.6238	11.2660	1.6393
4	11.25	1.6238	11.2707	1.6263
5	11.25	1.6238	11.2372	1.6243
6	11.25	1.6238	11.2362	1.6357
7	11.25	1.6238	11.2367	1.6235

6.4.2 Scenario I

Scenario I considers only demand variations. Link capacities are assumed to be deterministic and operating at the ideal capacities.

6.4.2.1 *OD Travel Time Comparison for Different Route Choice Models*

Mean and standard deviation (shown in parenthesis) of OD travel times for 9 combinations of demand variations (3 means x 3 standard deviations) and for four route choice models are shown in Table 6.5. For the SN-DUE and SN-SUE models, risk averse behavior is used for comparison with the DN-DUE and DN-SUE models. Except the DN-DUE model where all used paths of a given OD pair have the minimum travel time equal to the OD travel time, the OD travel times for the other three models are computed as a weighted average of the travel times on the different paths, where the weights are the path flows. The general trend for all the models is that as the mean demand increases, OD travels times for both OD pairs (1,4) and (1,5) also increase. This is also true for the case of a constant demand and only the standard deviation of OD demand is increased. In terms of the route choice models, the DN-DUE model has the lowest expected OD travel times and the SN-DUE model has the highest expected OD travel times. This is because the DN-DUE model is based on the assumption that travelers choose routes that minimize travel times, while the SN-DUE model with risk averse behavior without perception error tries to avoid paths that contain the high variability link and takes longer paths that result in higher OD travel times. For the two SUE models that contain a perception error, their OD travel times are in-between those given by the DN-DUE and SN-DUE models. This result can be explained as follows. Because of the perception error, travelers in the DN-SUE model cannot always perceive exactly the minimum travel time path and results in taking paths that are slightly longer with higher expected OD travel times. Similarly, travelers in the SN-SUE model with risk averse behavior and perception error cannot always identify paths that have the least travel time variability, and sometimes choose paths that contain some variability but lower travel times. This evidently results in lower OD travel times than those found by the SN-DUE model.

Table 6.5: OD Travel Time Comparisons for Different Route Choice Models under Demand Variations

Mean/Std. Dev.	OD (1,4)					
	DN-DUE			DN-SUE		
	LSD	MSD	HSD	LSD	MSD	HSD
LM	9.2 (0.0)	9.2 (0.1)	9.3 (0.2)	9.2 (0.0)	9.2 (0.1)	9.3 (0.2)
MM	9.3 (0.1)	9.4 (0.2)	9.5 (0.5)	9.4 (0.1)	9.4 (0.3)	9.6 (0.5)
HM	9.6 (0.2)	9.8 (0.6)	10.1 (1.2)	9.7 (0.3)	9.9 (0.6)	10.2 (1.2)
Mean/Std. Dev.	SN-DUE			SN-SUE		
	LSD	MSD	HSD	LSD	MSD	HSD
LM	9.5 (0.0)	9.5 (0.1)	9.5 (0.2)	9.2 (0.0)	9.2 (0.1)	9.3 (0.3)
MM	9.5 (0.1)	9.6 (0.4)	9.8 (0.7)	9.4 (0.2)	9.6 (0.5)	9.7 (0.7)
HM	10.2 (0.6)	10.2 (0.8)	10.4 (1.2)	10.1 (0.5)	10.1 (0.7)	10.3 (1.2)
Mean/Std. Dev.	OD (1,5)					
	DN-DUE			DN-SUE		
	LSD	MSD	HSD	LSD	MSD	HSD
LM	9.2 (0.0)	9.3 (0.1)	9.3 (0.2)	9.2 (0.1)	9.2 (0.2)	9.3 (0.8)
MM	9.4 (0.1)	9.5 (0.4)	9.7 (0.8)	9.5 (0.2)	9.6 (0.4)	9.8 (1.0)
HM	9.8 (0.3)	10.0 (0.9)	10.5 (1.9)	9.9 (0.4)	10.1 (0.9)	10.6 (2.0)
Mean/Std. Dev.	SN-DUE			SN-SUE		
	LSD	MSD	HSD	LSD	MSD	HSD
LM	9.4 (0.0)	9.4 (0.1)	9.5 (0.3)	9.2 (0.1)	9.3 (0.2)	9.4 (0.4)
MM	9.5 (0.2)	9.8 (0.6)	10.0 (1.0)	9.5 (0.3)	9.7 (0.6)	9.9 (0.9)
HM	10.3 (0.7)	10.5 (1.0)	10.9 (1.9)	10.3 (0.6)	10.3 (1.0)	10.7 (1.9)

LM=Low Mean, MM=Medium Mean, HM=High Mean, LSD=Low Standard Deviation, MSD=Medium Standard Deviation, HSD=High Standard Deviation

6.4.2.2 *Effects of Route Choice Models on the Estimation of OD Travel Time Reliability*

In Figure 6.3, we show the OD travel time reliability curves for both OD pairs (1,4) and (1,5) for four route choice models using the medium demand with medium standard deviation (i.e., the MM, MSD cell in Table 6.5). On the X-axis, we plot the travel time threshold value, which can be used as an indicator for level of service (a smaller value indicates a higher level of service). The Y-axis is the OD travel time reliability, which is defined as the probability that the weighted travel time for a given OD pair is within an acceptable level of service. This measure, as mentioned above, is an aggregate measure of all relevant used paths serving an OD pair. In both OD pairs, the OD travel time reliability under the DN-DUE model is always higher compared to the other three models. Since travelers in the DN-DUE model assume to have perfect information and use travel time as the sole criterion for route choice, it is reasonable to have such results. However, for the DN-SUE model, travelers are assumed to make route choices with imperfect information, hence the OD travel time may be slightly sub-optimal resulting in lower OD travel time reliability. For the SN-DUE and SN-SUE models which assume risk averse behavior, route choices are based on minimizing disutilities and perceived disutilities, respectively, which are based on expected link travel time and travel time variability leading to higher OD travel times or lower OD travel time reliabilities. Thus, it would seem reasonable that the OD travel time reliability is lower when route choices are assumed to be based on imperfect information and even more so when they are based on minimizing disutilities that are a function of factors other than link travel times.
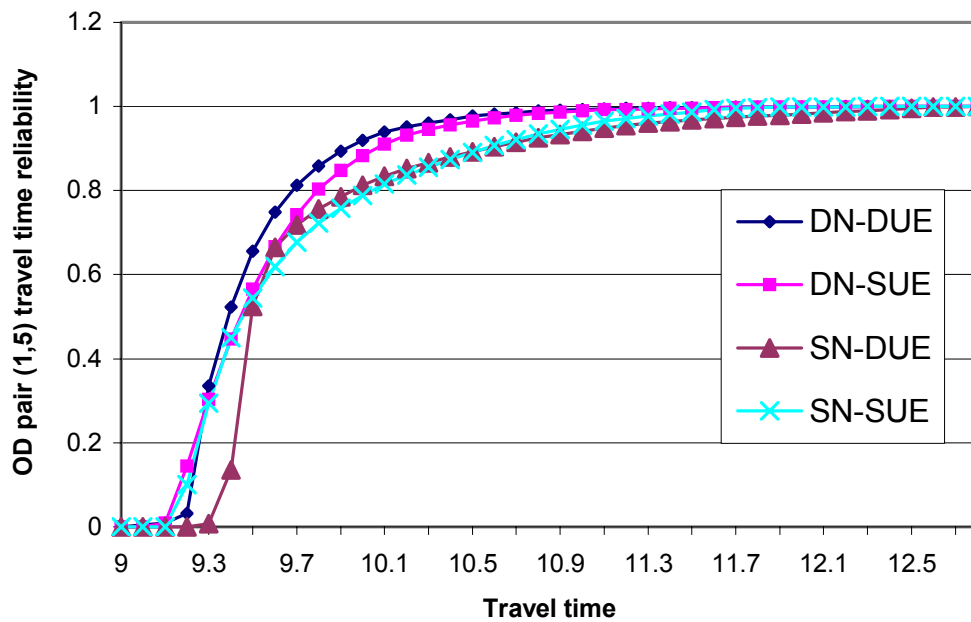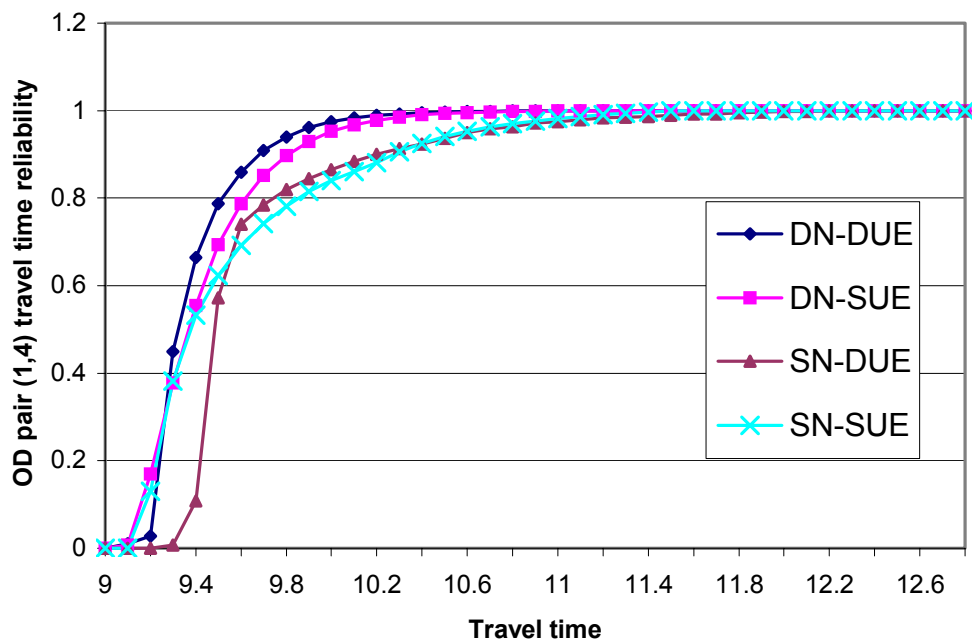
Figure 6.3: Effects of Route Choice Models on the Estimation of OD Travel Time Reliability under Demand Variations

6.4.2.3 *Effects of Route Choice Models on the Estimation of Path Travel Time Reliability*

For brevity, we show only the results of path 1 (link sequence 1->4) and path 2 (link sequence 2->6), which are two major alternate paths for OD pair (1,4) using the medium mean demand and medium standard deviation of demand (i.e., cell (MM,MSD) in Table 6.5) as an example. Figure 6.4 shows the effect of different route choice behaviors on estimating path travel time reliabilities. Similar to the OD travel time reliability curves, the X-axis is the travel time threshold value, and the Y-axis is the path travel time reliability which is defined as the probability that the travel time of a given path is within an acceptable threshold value. For the DN-DUE and DN-SUE models, travelers are assumed to make their route choices solely based on average travel times. No consideration is given to travel time variability. The travel time reliability on path 1 and path 2 are approximately equal. For the SN-DUE and SN-SUE models that consider both average travel time and travel time variability, less travelers are allocated to path 1 (1->4) because it contains link 1 which assumes to have a high variability in travel time, while more travelers are allocated to path 2 (2->6) that has less variability in travel time. Because more travelers use path 2, travel time of path 2 is higher than that of path 1. This evidently decreases the travel time reliability of path 2 for the risk averse travelers in the two SN models. The differences in travel time reliability between the route choice models implemented without and with perception error are minor between the DN-DUE and DN-SUE models, but are more visible between the SN-DUE and SN-SUE models. This result seems to suggest that perception error may have an effect on the risk taking behavior. Similar results (not shown here) are also observed for path 4 (1->5) and path 5 (2->6), which are the two major paths for OD pair (1,5).
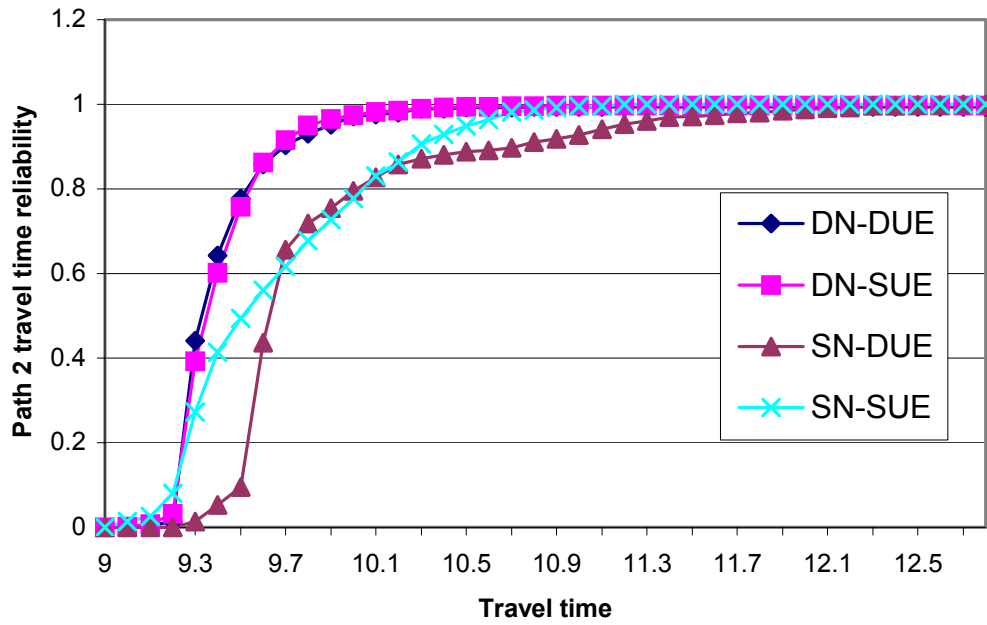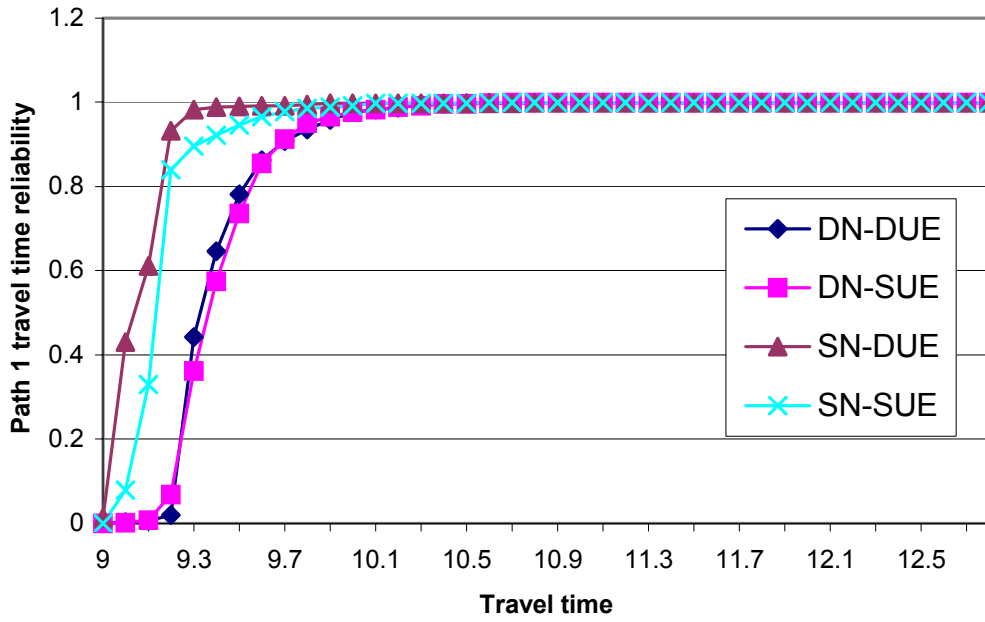
Figure 6.4: Effects of Route Choice Models on the Estimation of Path Travel Time Reliability under Demand Variations

6.4.3 Scenario II

Scenario II considers both demand supply variations. Link capacities are assumed to degrade at a mean of 25% of the ideal capacities with three levels of standard deviation. Using the medium demand case with three levels of standard deviation (MM-LSD, MM-MSD, and MM-HSD) and mixing them with the capacity variations (LSD, MSD, and HSD), there are 9 combinations (3 demand variations x 3 supply variations).

6.4.3.1 *OD Travel Time Comparison for Different Route Choice Models*

Similar to scenario I, the mean and standard deviation (shown in parenthesis) of OD travel times are shown in Table 6.6. Again, we use the risk averse behavior for the two route choice models that consider network uncertainty. The results are similar to those presented in scenario I. In general, the average OD travel times and standard deviation of OD travel times for both OD pairs and for all route choice models increase as the variations caused by demand and supply increase. In terms of the route choice models, the DN-DUE model has the lowest expected OD travel times, followed by the DN-SUE and SN-SUE models, and the SN-DUE model has the highest expected OD travel times. The same reasons explained in Scenario I also apply here.

Table 6.6: OD Travel Time Comparisons for Different Route Choice Models under Both Demand and Supply Variations

Demand/Supply	OD (1,4)					
	DN-DUE			DN-SUE		
	LSD	MSD	HSD	LSD	MSD	HSD
MM-LSD	9.8 (0.3)	9.8 (0.4)	9.8 (0.4)	9.9 (0.3)	9.9 (0.4)	10.0 (0.5)
MM-MSD	10.0 (0.8)	10.0 (0.8)	10.0 (0.9)	10.1 (0.8)	10.1 (0.9)	10.2 (1.0)
MM-HSD	10.4 (1.6)	10.4 (1.7)	10.5 (1.9)	10.5 (1.6)	10.5 (1.7)	10.6 (1.9)
Demand/Supply	SN-DUE			SN-SUE		
	LSD	MSD	HSD	LSD	MSD	HSD
MM-LSD	10.3 (0.6)	10.3 (0.6)	10.4 (0.7)	10.3 (0.5)	10.3 (0.5)	10.3 (0.5)
MM-MSD	10.4 (0.9)	10.4 (0.9)	10.5 (1.0)	10.3 (0.8)	10.3 (0.9)	10.4 (1.0)
MM-HSD	10.7 (1.5)	10.8 (1.6)	10.8 (1.8)	10.6 (1.5)	10.7 (1.7)	10.7 (1.9)
Demand/Supply	OD (1,5)					
	DN-DUE			DN-SUE		
	LSD	MSD	HSD	LSD	MSD	HSD
MM-LSD	10.0 (0.5)	10.0 (0.5)	10.1 (0.6)	10.1 (0.4)	10.2 (0.5)	10.2 (0.5)
MM-MSD	10.3 (1.2)	10.3 (1.3)	10.4 (1.4)	10.4 (1.2)	10.5 (1.3)	10.5 (1.4)
MM-HSD	10.9 (2.5)	11.0 (2.7)	11.1 (3.0)	11.0 (2.6)	11.1 (2.7)	11.3 (3.0)
Demand/Supply	SN-DUE			SN-SUE		
	LSD	MSD	HSD	LSD	MSD	HSD
MM-LSD	10.7 (0.8)	10.7 (0.8)	10.7 (0.9)	10.5 (0.6)	10.5 (0.6)	10.5 (0.6)
MM-MSD	10.8 (1.2)	10.8 (1.3)	10.9 (1.4)	10.6 (1.2)	10.7 (1.2)	10.7 (1.4)
MM-HSD	11.3 (2.5)	11.3 (2.6)	11.4 (2.9)	11.1 (2.5)	11.2 (2.6)	11.3 (2.9)

MM-LSD=Medium Mean-Low Standard Deviation, MM-MSD=Medium Mean-Medium Standard Deviation, MM-HSD=Medium Mean-High Standard Deviation

6.4.3.2 *Effects of Route Choice Models on the Estimation of OD Travel Time Reliability*

Similar to scenario I, we show the OD travel time reliability curves in Figure 6.5 using the medium mean and medium standard deviation of demand and medium standard deviation of capacity case (i.e., cell (MM-MSD,MSD)) in Table 6.5. Again, the OD travel time reliability under the DN-DUE model is the highest, followed the DN-SUE model, and the two SN models have the lowest OD travel time reliability. This is to be expected since the two DN models directly minimize travel times, albeit that the DN-SUE model is allocating flows based on imperfect information, and the two SN models that assume risk averse behavior are minimizing disutilities which are not solely based on travel times. In general, the OD travel time reliabilities for both OD pairs are significantly lower than those in scenario I. Using OD (1,4) as an example, the OD travel time reliability at a travel time threshold value of 10.4 for the DN-DUE model is 1.0 under scenario I, but it drops to 0.79 under scenario II when supply variation is also considered. Hence, it seems that the uncertainty due to supply variation can significantly contribute to the decrease of travel time reliability.
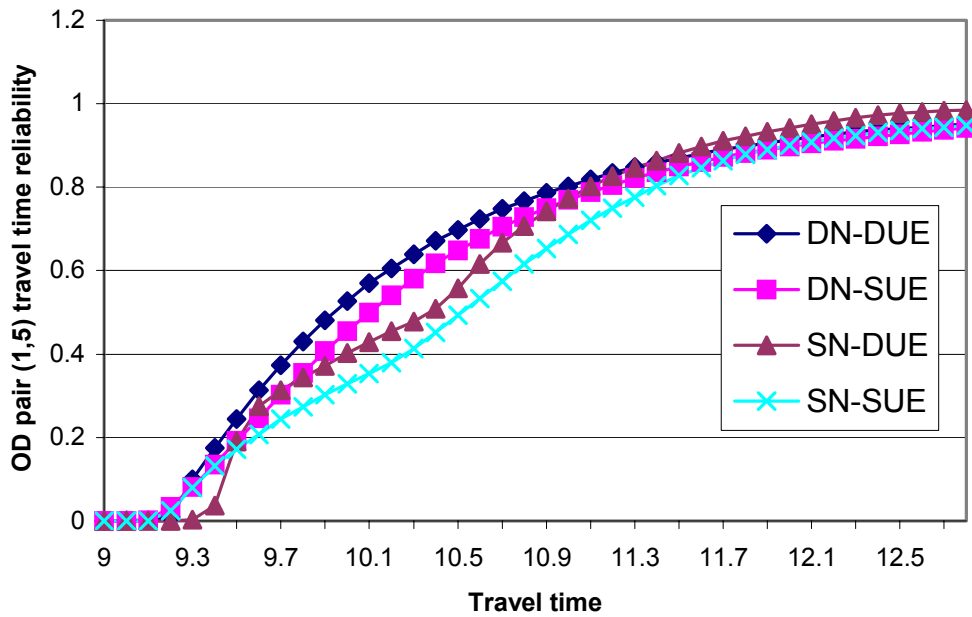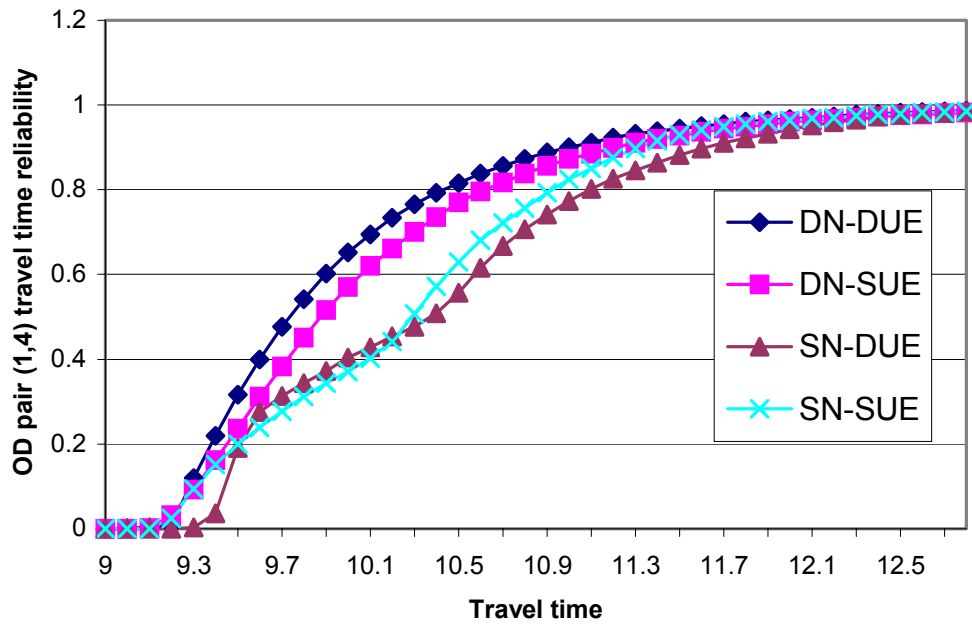
Figure 6.5: Effects of Route Choice Models on the Estimation of OD Travel Time Reliability under Both Demand and Supply Variations

6.4.3.3 *Effects of Route Choice Models on the Estimation of Path Travel Time Reliability*

Again for brevity, we show only the results of path 4 (1->5) and path 5 (2->7), which are two major alternate paths for OD pair (1,5) using cell (MM-MSD,MSD) in Table 6.6. Figure 6.6 plots the path travel time reliability curves for all four route choice models on path 4 and path 5. Despite the paths shown in Figure 6.6 are for a different OD pair, the results are generally agreed with those presented in Figure 6.4 of scenario I. That is, the two SN models that assume risk averse behavior allocate less flow to path 4 to avoid the high travel time variability link 1, which is part of path 4. Since path 5 is the only alternative path for OD pair (1,5) that does not contain link 1, the two SN models allocate more flow to path 5, which results in higher path travel time and lower path travel time reliability. When both demand and supply variations are considered, the path travel time reliability curves are in general lower than those in scenario I that considers only demand variations. Using the DN-DUE model with a travel time threshold value of 10.5 minutes as an example, the travel time reliability of path 5 is 0.963 for scenario I (not shown here), but drops to 0.702 for scenario II.
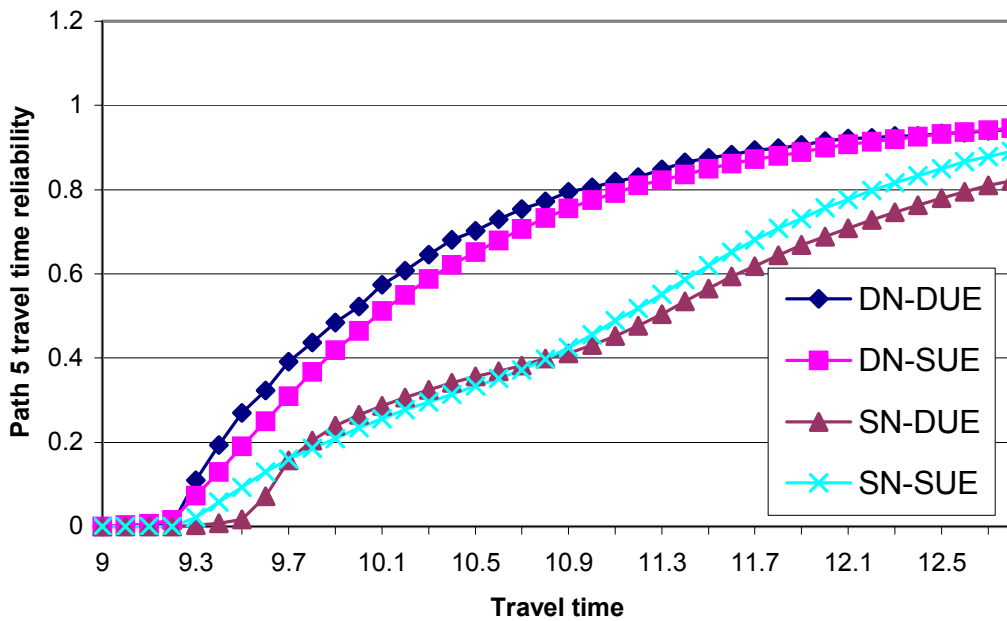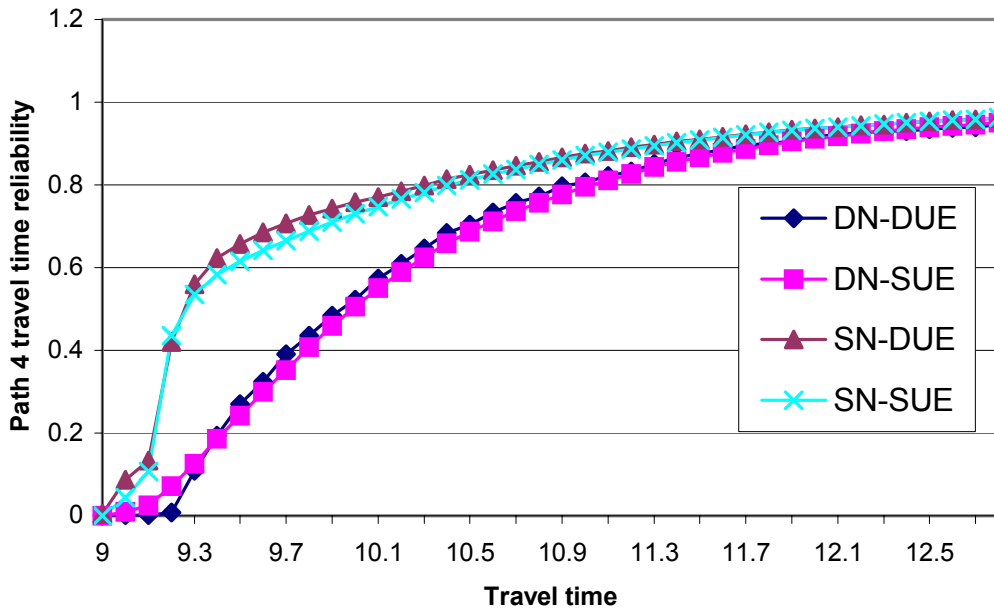
Figure 6.6: Effects of Route Choice Models on the Estimation of Path Travel Time Reliability under Both Demand and Supply Variations

157

6.4.4 Comparison Between Scenario I and Scenario II

Comparing the results of scenario II with those in scenario I, it is found that the both the mean and standard deviation of OD travel times are higher. The higher OD travel times are due to the 25% capacity reduction on all links in the network, while the larger standard deviation of OD travel times is due the additional variation coming from the supply side. Thus, it seems necessary to account for both demand and supply variations when assessing the performance of a road network.

6.4.4.1 *Link Flow Allocation Comparisons*

In Table 6.7, we compare the link flow allocations between the two scenarios, using the results from the medium demand and medium standard deviation of demand case (i.e., cell (MM,MSD)) in Table 6.5 and from the medium mean and medium standard deviation of demand and medium standard deviation of capacity (i.e., cell (MM-MSD,MSD)) in Table 6.6. For brevity, we focus the comparison on link 1 and link 2 which are the only two links emanating from origin 1 (see Figure 6.1). Link 1 has a lower free-flow travel time but higher variability associated with the travel time, while link 2 has a higher free-flow travel time but lower variability. Hence, one would expect that the DN-DUE model would allocate more flow to link 1, the SN-DUE model assuming risk averse behavior would allocate less flow, and the DN-SUE and SN-SUE models that contain perception errors in their route choice models would allocate flow in-between these two limits. This expectation appears to hold true for scenario I. Because of the 25% capacity degradation in scenario II, the flow allocation between link 1 and link 2 in the DN-DUE model is slightly more evenly distributed compared to those in scenario I. This flow allocation is optimal in terms of travel time minimization, as reflected in Table 6.5 where the expected OD travel times are lowest among the four route choice models. Because of the perception error in the DN-SUE model, the flow allocation over the network is slightly sub-optimal in terms of higher expected OD travel times (also see Table 6.8 for the sub-optimal flow allocation to path 3 and path 6) despite that a larger amount of flow is allocated to link 1. For the SN-DUE and SN-SUE models, flow allocations over the network are similar for both scenarios.

Table 6.7: Link Flow Comparisons Between Scenario I and Scenario II

Link #	Scenario I				Scenario II			
	DN-DUE	DN-SUE	SN-DUE	SN-SUE	DN-DUE	DN-SUE	SN-DUE	SN-SUE
1	20.89 (2.37)	20.59 (4.29)	12.12 (4.68)	14.64 (2.94)	18.80 (3.51)	19.70 (4.04)	13.70 (5.08)	14.85 (4.84)
2	12.96 (5.78)	13.26 (3.74)	21.72 (3.97)	19.20 (5.58)	15.04 (4.95)	14.14 (4.14)	20.14 (3.92)	19.00 (4.30)
3	4.17 (1.38)	7.51 (1.01)	0.03 (0.02)	3.17 (0.79)	2.38 (1.49)	5.98 (0.85)	0.18 (0.62)	2.21 (1.08)
4	7.41 (2.26)	5.60 (2.08)	5.02 (3.01)	4.75 (2.08)	7.25 (2.48)	5.82 (2.38)	5.82 (2.83)	5.26 (2.71)
5	9.30 (2.83)	7.47 (2.91)	7.07 (3.51)	6.73 (2.85)	9.18 (3.07)	7.90 (3.19)	7.71 (3.49)	7.38 (3.48)
6	7.60 (2.75)	9.41 (2.93)	10.00 (2.35)	10.27 (3.07)	7.77 (2.63)	9.19 (2.70)	9.19 (2.60)	9.76 (2.54)
7	9.53 (3.51)	11.36 (3.36)	11.72 (3.15)	12.10 (3.51)	9.65 (3.30)	10.93 (3.14)	11.12 (3.18)	11.45 (3.01)

6.4.4.2 *Path Flow Allocation Comparisons*

For completeness, we also compare the path flow patterns of both scenarios in Table 6.8. Similar to the link flow allocation comparison, we use the results from the medium mean and medium standard deviation of demand in scenario I (i.e., cell (MM,MSD)) in Table 6.5 and from the medium mean and medium standard deviation of demand and medium standard deviation of capacity in scenario II (i.e., cell (MM-MSD,MSD)) in Table 6.6 as an example for the path flow allocation comparison. There are 6 paths in total serving the 2 OD pairs. Paths 1 (1->4), 2 (2->6), and 3 (1->3->6) are for OD pair (1,4), and paths 4 (1->5), 5 (2->6), and 6 (1->3->7) are for OD pair (1,5). Based on the assumptions of the route choice model, it is apparent that each route choice model allocates a different amount of flow to each path. For example, the DN-DUE model seems to allocate more flow to path 1 and path 4, which contain link 1 (lower free-flow travel time but higher travel time variability), while the two SN models that assume risk averse behavior allocate more flow to path 2 and path 5 to avoid link 1. It is also interesting to observe that the SN-DUE model, which assumes to have perfect knowledge of the variable nature of network travel times, allocate significantly less flow to path 3 and path 6 compared to those allocated by the other three route choice models. This is because both path 3 and path 6 contain link 1 (the high variability link) and have higher average travel times. The differences in flow

159

allocation to the paths in the network between the route choice models implemented without and with perception error are also evident from the results presented in Table 6.8. Thus, it would appear that the different assumptions in modeling route choice behavior could have an effect on the path flow patterns as well as the link flow patterns.

Table 6.8: Path Flow Comparisons Between Scenario I and Scenario II

Path #	Scenario I				Scenario II			
	DN-DUE	DN-SUE	SN-DUE	SN-SUE	DN-DUE	DN-SUE	SN-DUE	SN-SUE
1	7.41 (2.24)	5.60 (2.90)	5.02 (3.01)	4.75 (2.08)	7.25 (2.49)	5.82 (2.38)	5.82 (2.83)	5.26 (2.71)
2	6.01 (3.06)	5.97 (2.75)	10.01 (2.36)	8.69 (3.03)	7.05 (2.66)	6.39 (2.27)	9.17 (2.55)	8.67 (2.59)
3	1.61 (1.39)	3.48 (2.44)	0.01 (0.22)	1.58 (0.42)	0.72 (1.12)	2.81 (0.71)	0.02 (0.63)	1.10 (0.50)
4	9.30 (2.76)	7.48 (3.49)	7.07 (3.51)	6.73 (2.85)	9.18 (3.10)	7.88 (3.17)	7.71 (3.49)	7.38 (3.49)
5	6.96 (3.05)	7.33 (3.74)	11.70 (3.15)	10.51 (3.72)	7.99 (3.15)	7.74 (2.76)	10.97 (3.08)	10.33 (3.05)
6	2.57 (0.94)	4.06 (2.63)	0.02 (0.20)	1.59 (0.56)	1.66 (0.97)	3.19 (0.72)	0.15 (0.88)	1.13 (0.69)

6.4.5 Scenario III

Scenario III examines the effect of parameter a_2 in the risk averse disutility function on the estimation of path travel time reliability. As mentioned above, the value of a_2 measures the degrees of risk aversion to network uncertainty. Larger values of a_2 imply that travelers are more averse to risk and willing to pay a higher premium to avoid uncertainty. Using the same procedure above, we estimate different values for a_2 by assuming that all travelers have a disutility of 1 for a route that is 1, 3, 5, 7, or 9 minutes. The shapes of these disutility functions are illustrated in Figure 6.7.
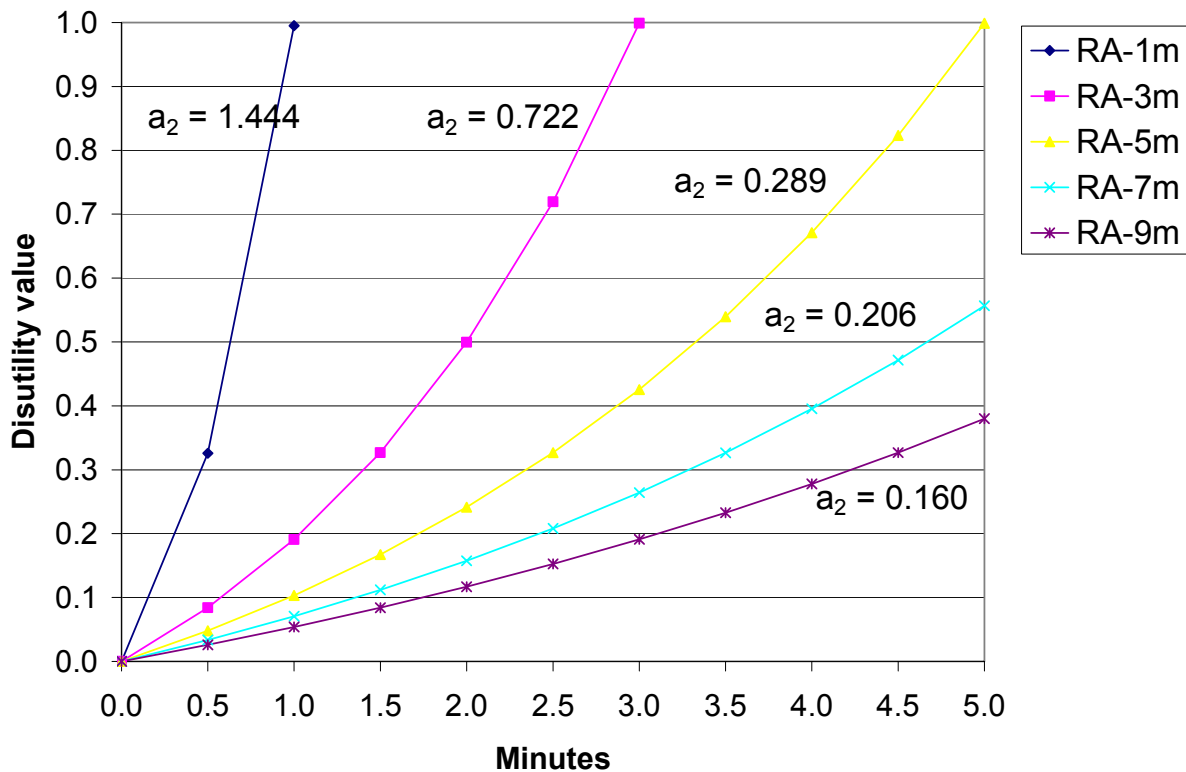

Figure 6.7: Risk Averse Disutility Function with Different a_2 Values

For brevity, we show only the results of path 1 (1->4) and path 2 (2->6), which are two alternate paths for OD pair (1,4) using the risk averse case of the SN-SUE model under the medium mean demand and medium standard deviation of demand (i.e., cell (MM,MSD) in Table 6.5) as an example. Based on the degrees of risk aversion to network uncertainty, Figure 6.8 shows that as

the value of a_2 increases (or 9m to 1m), the SN-SUE model assigns lesser and lesser flow to path 1, which contains the high travel time variability link (i.e., link 1 which has a variability of 100 percent of the link's free-flow travel time). It is apparent that these risk averse travelers are willing to pay a higher premium by taking a longer travel time route as shown in the travel time distribution on path 2 in Figure 6.9. Because more travelers use path 2 (i.e., path flow distribution on path 2 for the 1m is shifted to the right in Figure 6.8), travel times on path 2 is higher than that on path 1 as indicated in Figure 6.9. This evidently decreases the travel time reliability for path 2, particularly for the highest value of a_2 (or 1m) in Figure 6.10. For a given value of travel time threshold (say 10.9 min.), the travel time reliability is 1.00 for path 1 and 0.67 for path 2.
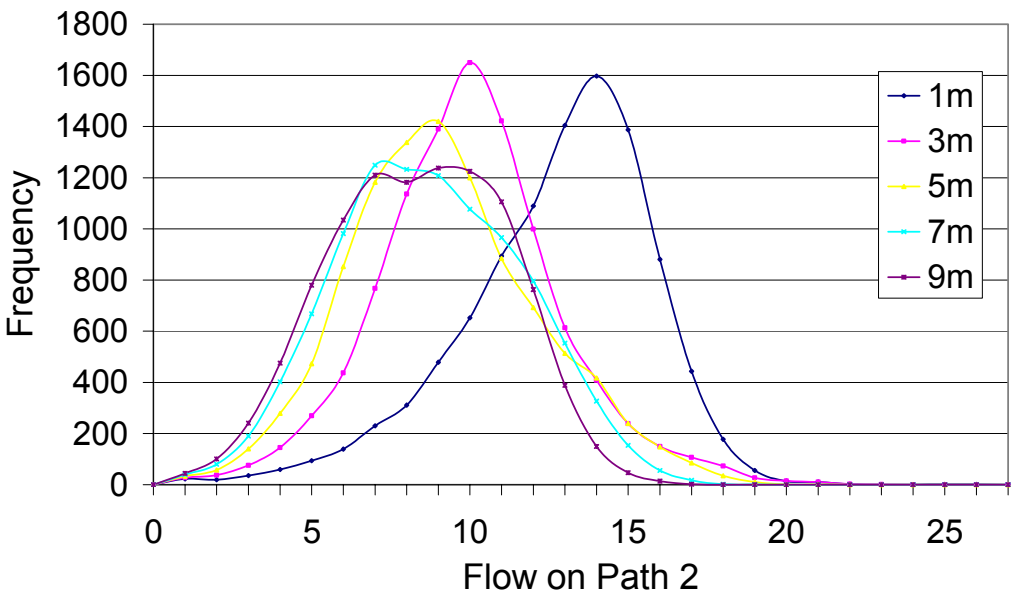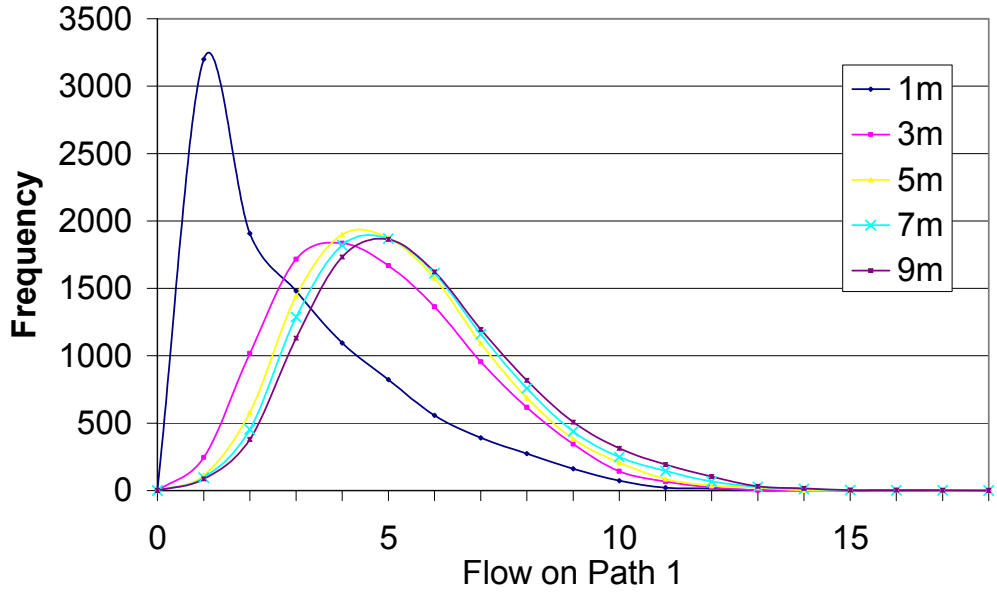
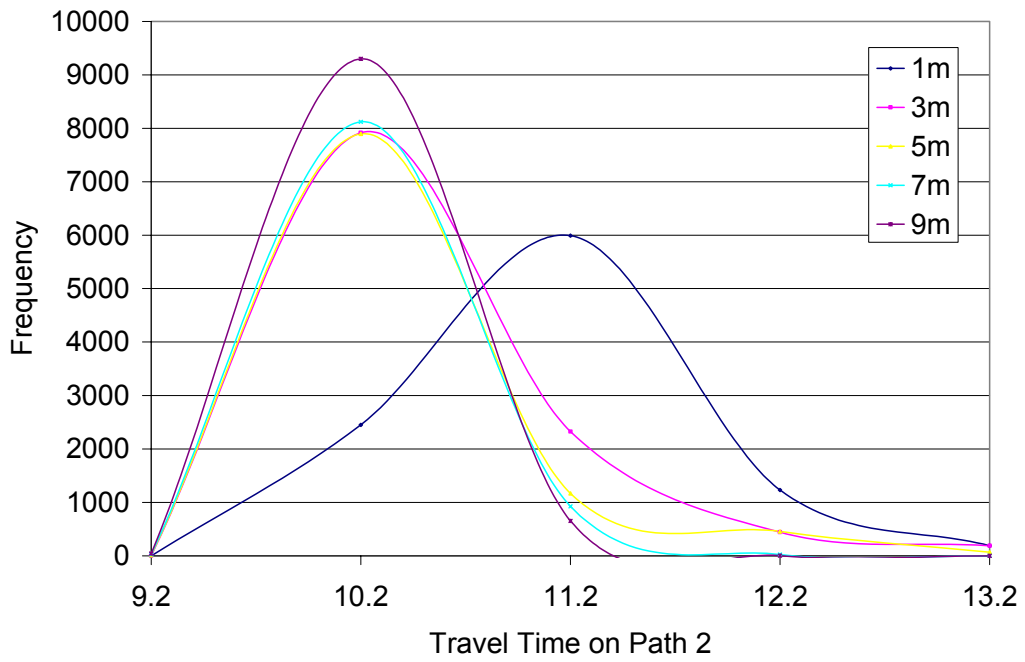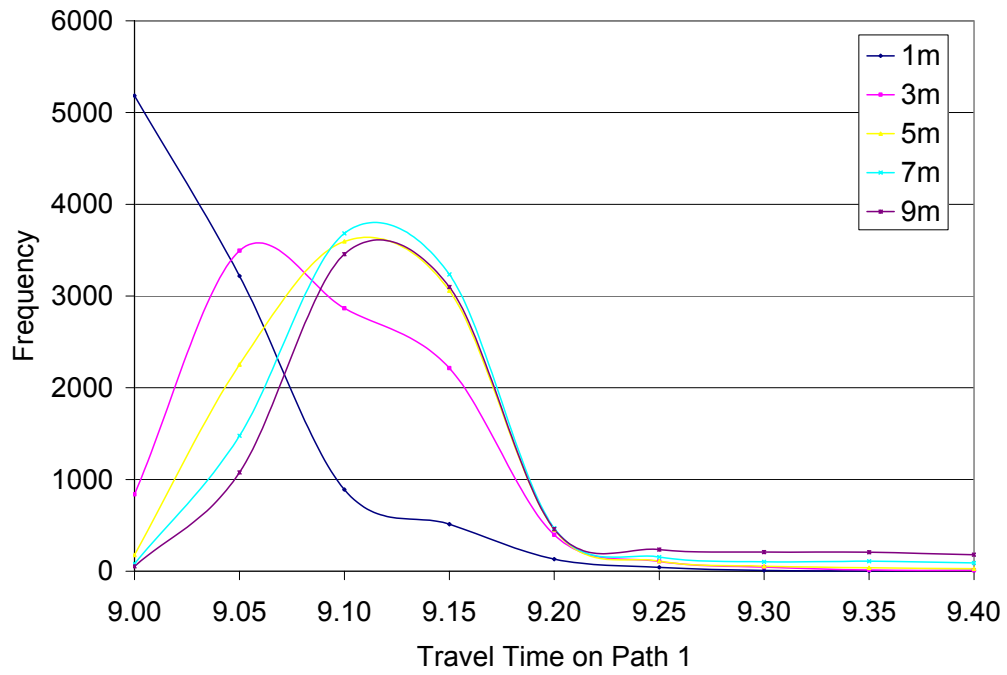**Figure 6.8**: Path Flow Distribution for Different a_2 Values

Figure 6.9: Path Travel Time Distribution for Different a_2 Values
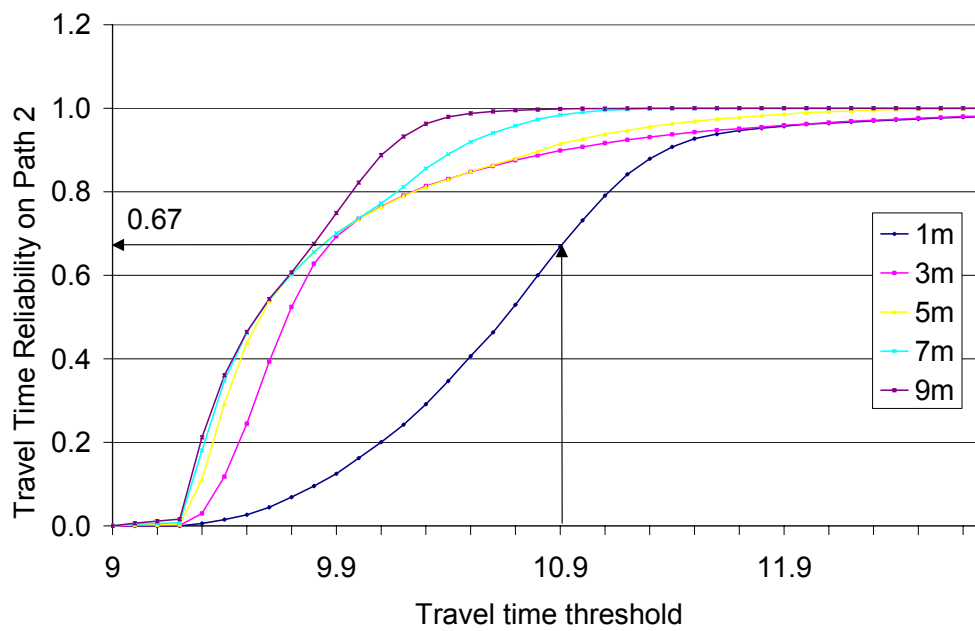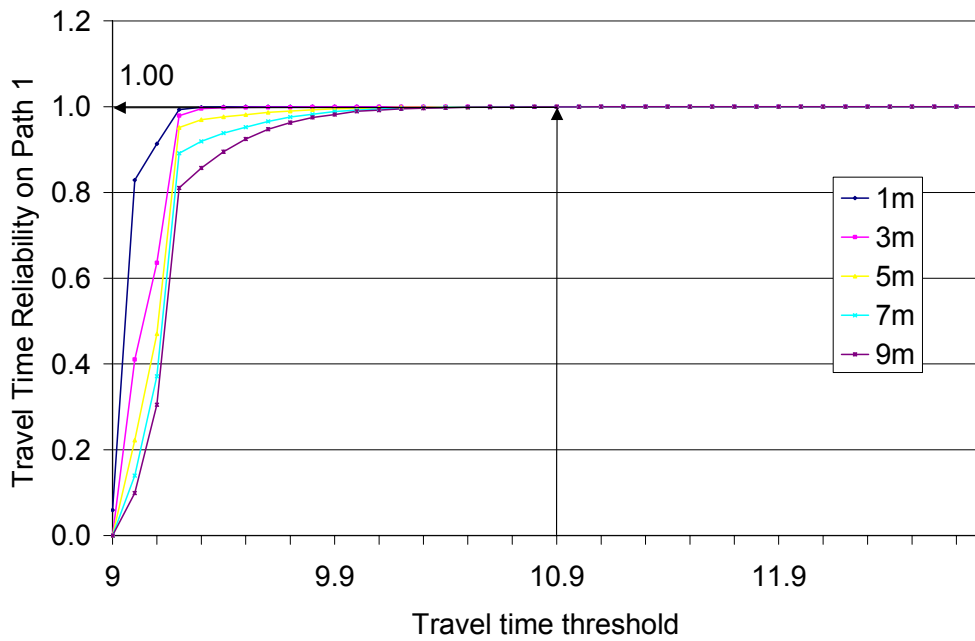
Figure 6.10: Path Travel Time Reliability for Different a_2 Values

6.5   Summary

In this chapter, we have presented a review of transportation reliability measures and focused on travel time reliability. A Monte Carlo simulation procedure was used to simulate risk perceptions and preferences in making route choice decisions under an uncertain environment. Numerical results were also presented to examine what the aggregate impact of changes in variability caused by demand and supply variations might have on network assignment and how travelers with different risk-taking behaviors respond to these changes.

In general, the results from the four route choice models presented in Table 2.1 of Chapter 2 may have a significant impact on the estimation of travel time reliability measures, and it is important to examine how the different assumptions made in the route choice models affect the calculation of the travel time reliability measures under an uncertain environment. Based on the results of the three scenarios constructed for the numerical experiment, the following conclusions can be drawn:

- As the mean and standard deviation of OD demand increase, the expected and standard deviation of OD travel time also increase.
- The weighted OD travel time is lowest for the DN-DUE model, followed by the DN-SUE and SN-SUE models, and highest for the SN-DUE model.
- OD travel time reliability is lower when route choices are assumed to be based on imperfect information and even more so when they are based on minimizing disutilities that are a function of factors other than travel time.
- Additional variation from capacity degradation can significantly increase not only the OD travel time but also the standard deviation of OD travel time.
- Perception error may contribute to sub-optimal path and link flow allocations since the network assignment is based on imperfect information.
- Travelers who possess risk averse behavior are likely to pay a premium to avoid uncertainty.
- As the degree of risk aversion to network uncertainty increases, travel time also increases and results in lower travel time reliability.

The numerical results, albeit based on a small network, reveal the significance of the route choice model on the estimation of travel time reliability measures. The different assumptions used in modeling route choice behavior could result in different estimates of travel time reliability. However, we note that the conclusions presented above are only relevant to the particular network and parameters used in this study.

The SN-SUE model used to estimate travel time reliability is for a single-class, risk-taking route choice model based on a pre-specified disutility function to model a specific risk preference (e.g. risk averse) for all travelers in the population. In reality, different travelers with the same characteristics facing the same choice situation may respond differently according to various degrees of risk aversion to travel time uncertainty. This is important in studying route choice behavior at the individual level, and how these individual responses to network uncertainty collectively affect the transportation systems performance. One way to extend the SN-SUE model to a multi-class route choice model is to segment the population into groups (e.g., 10% risk prone, 30 % risk neutral, and 60 % risk averse). This can be accomplished by using the diagonalization algorithm (Sheffi, 1985). However, it is difficult to observe how travelers make tradeoff decisions between expected travel time and travel time variability in actual choice situations. Another possible way to account for different responses to travel time variability across individuals is to use the random-coefficient (or mixed) logit model described in Chapter 4 to account for the degrees of risk of aversion by capturing the taste variation across individuals resulting in differences in their responses to route-specific attributes (e.g., expected travel time and travel time variability) or in differences in their intrinsic preferences for route alternatives. For the risk-taking route choice model, the random-coefficients logit model captures individuals' risk perceptions and preferences with randomly distributed coefficients on the value of time and value of travel time variability, resulting in a continuous multi-class network equilibrium model to describe the risk-taking route choices of heterogeneous travelers. The goal is to enhance the predictive capability of the route choice model and to expand the behavioral framework of route choice modeling.

Chapter Seven

7.   CONCLUDING REMARKS

Recently the importance of travel time variability has been perceived by many researchers and practitioners. However, there has been limited basic research directed toward understanding the effects of travel time variability. This research project has attempted to develop methods to address a number of questions relating to travel time variability. Specifically, this research project has developed:

- A GIS database and computation procedures to analyze freeway travel time variability.
- A mixed logit model to uncover the contribution of travel time reliability.
- A set of route choice models that accounts for the variations of travel time in the form of risk, perception errors, and behavioral preferences.
- A set of traffic assignment procedures for the above route choice models.
- A travel time reliability evaluation procedure.

The main purpose of the GIS database is to provide an analysis framework to analyze travel time variability for both section level and route level of the freeways in Orange County, California. The framework is expected to help researchers and practitioners understand different types of travel time variability: day-to-day variability, within-day variability, and spatial variability. The mixed logit model developed in this research project provides an indirect method to study the contribution of travel time reliability in traveler's route choice behavior. Unlike the traditional approach that estimates the values of time and reliability using reveal preference (RP) and/or stated preference (SP) data using the simulated maximum likelihood estimation method, this research developed a genetic algorithm procedure to identify these parameter values using the real-time traffic volume data obtained loop detectors. Such an approach eliminates both the cost

and biases inherent in RP and SP survey techniques. The route choice models described in Chapter Two are the core component of many surface transportation applications. Despite its importance, the state-of-the-practice route choice models do not account for travel time uncertainty in the models. The purpose of this chapter is to provide an up-to-date review of the route choice models, particularly emphasizing on the route choice models that accounts for the variations of travel time in the form of risk, perception errors, and behavioral preferences. Traffic assignment procedures were implemented for four route choice models. Numerical experiments were carried out to examine the effects of route choice models on network assignment results. The results showed that there are significant differences on the flow allocations between the risk neutral travelers in the DN-DUE and DN-SUE models and the risk averse travelers in the SN-DUE and SN-SUE models. The reason is because of the criterion and risk-taking behavior used in the route choice decision process. By incorporating travel time variability into the route choice models, it enhances the predictive capability of the route choice models and could potentially lead to better means of reducing traffic congestion, wasteful travel, and loss of productivity, and at the same time, improving network capacity utilization and travel time reliability. The travel time reliability evaluation procedure provides an analysis framework to estimate travel time reliability that explicitly considers the effects of route choice models under demand and supply variations. The framework is expected to be useful in examining what the aggregate impact of changes in variability caused by demand and supply variations might have on network assignment and how travelers with different risk-taking behaviors respond to these changes. Although this research project has addressed many questions related to travel time variability, our understanding on travel time variability is still in its infancy and applications to real-world transportation management have not yet developed. Further studies are needed to better understand travel time variability and how to use such measures for transportation management.

8. REFERENCE

Aashtiani, H. (1979) The multi-modal traffic assignment problem. Ph.D. Dissertation, Operations Research Center. MIT Press, Cambridge, Massachusetts, USA.

Abdel-Aty, M., Kitamura, R., and Jovanis, P. (1996) Investigating Effect of Travel Time Variability on Route Choice Using Repeated Measurement Stated Preference Data. *Transportation Research Record* 1493, 39-45.

Adler, J.L. and Blue, V.J. (1998) Toward the design of intelligent traveler information systems. *Transportation Research* 6C, 157-172.

Arezki, Y. and Van Vliet, D. (1990) A full analytical implementation of the PARTAN/Frank-Wolfe algorithm for equilibrium assignment. *Transportation Science*, 24, 58-62.

Asmuth, R. (1978) Traffic network equilibria. Ph.D. Dissertation, Department of Operations Research. Stanford University, Stanford, California, USA.

Allosp, R. E. (1974) Some possibility for using traffic control to influence trip distribution and route choice. *Proceedings of Sixteen International Symposium on Transportation and Traffic Theory*, 345-374.

Allsop, R. E. and Charlesworth, J. A. (1977) Traffic in a signal-controlled road network: an example of different signal timings inducing different routes. *Traffic Engineering and Control* 18, 262-264.

Ang, A.H.S. and Tang, W.H. (1990) *Probability concepts in engineering planning and design, vol.1 basic principle; vol. II - decision, risk and reliability*. John Wiley & Sons.

Asakura, Y. (1998) Reliability Measures of an Origin and Destination Pair in a Deteriorated Road Network with Variable Flow. In: *Transportation Networks: Recent Methodological Advances. Selected Proceedings of the 4th EURO Transportation Meeting (Ed. M.G.H. Bell)*, Pergamon Press: Oxford, pp. 273-287.

Asakura, Y. and Ksahiwadani, M., (1991) Road Network Reliability Caused by Daily Fluctuation of Traffic Flow. *Proceedings of the 19th PTRC Summer Annual Meeting*, Brighton, pp. 73-84.

Bar-Gera, H. and Boyce, D.E. (2002) Origin-based network assignment. *In Advances in Transportation Network Modelling (*edited by M. Patriksson and M. Labbe*)*, Kluwer Academic Publishers.

Bate, J., Dix, M., and May, A. (1989) Travel time variability and its effect on time of day choice for the journey to work, Transportation Planning Methods, *Proceedings of Seminar CPTRC*, 1989, pp 293-311.

Bates, J., Polak, J., Jones, P., and Cook, A., (2001) *The Valuation of Reliability for Personal Travel, Transportation Research* 37E, 191-229.

Beckmann, M. J., C.B. McGuire, and C.B. Winsten (1956) Studies in the Economics of Transportation. Yale University Press, New Haven, Connecticut, USA.

Bekhor, S. and Prashker, J.N. (1999) Formulations of extended logit stochastic user equilibrium assignments. *Proceedings of the 14th International Symposium on Transportation and Traffic Theory*, Jerusalem, Israel, 351-372.

Bekhor, S. and Prashker, J.N. (2001) A stochastic user equilibrium formulation for the generalized nested logit model. Paper presented at *the 80th Annual Meeting of the Transportation Research Board*, Washington, D.C., USA.

Bekhor, S., Ben-Akiva, M.E., and Ramming, M.S. (2002) Adaptation of logit kernel to route choice situation. Paper presented at *the 81th Annual Meeting of the Transportation Research Board*, Washington, D.C., USA.

Bell, M.G.H. (1994) A simple alternative to Dial's logit assignment algorithm. *Transportation Research* 29B, 369-386.

Bell, M.G.H. and Cassir, C. (2002) Risk-averse user equilibrium traffic assignment: an application of game theory. *Transportation Research* 36B, 671-681.

Bell, M.G.H., Cassir, C., Iida, Y., and Lam, W.H.K. (1999) A Sensitivity based Approach to Reliability Assessment, *Proceedings of the 14th International Symposium on Transportation and Traffic Theory*, Jerusalem, Israel, 283-300.

Bell, M.G.H. and Iida, Y. (1997) *Transportation Network Analysis*. John Wiley & Sons, New York.

Bell, M.G.H., Lam,W.H.K., Ploss, G., and Inaudi, D. (1993) Stochastic user equilibrium assignment and iterative balancing. *Proceedings of the 12th International Symposium on Transportation and Traffic Theory*, University of California, Berkeley.

Ben-Akiva, M., Bergman, M.J., Daly, A.J., and Ramaswamy, R. (1984) Modeling inter urban route choice behavior. *Proceedings of the 9th International Symposium on Transportation and Traffic Theory*, VNU Science Press.

Ben-Akiva, M. and Bierlaire, M. (1999) Discrete choice methods and their applications to short term travel decisions. *Handbook of Transportation Science*, R.W. Hall ed., Kluwer Publishers.

Bernstein, D. and Gabriel, S. (1997) Solving the nonadditive traffic equilibrium problem. In *Proceedings of the Network Optimization Conference*, edited by Pardalos, P. Hearn, D., and Hager, W., Springer.

Bertsekas, D. and Gafni, E. (1982) Projection methods for variational inequalities with application to the traffic assignment problem. *Mathematical Programming* 17, 139-159.

Bhat, C.R. (2001), *Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model*, Transportation Research Part B, 35, pp. 677-693.

Bhat, C.R., and Castelar, R. (2002), *A Unified Mixed Logit Framework for Modeling Revealed and Stated Preferences: Formulation and Application to Congestion Pricing Analysis in the San Francisco Bay Area*, Transportation Research Part B, 36, pp. 593-616.

Blue, V.J., List, G.L., and Adler, J.L. (1997) Real-time multiple objective path search for in-vehicle route guidance systems. *Transportation Research Record* 1588, 159-167.

Bovy, P.H.L. (1996) Stochastic traffic assignment technique enhancements for congested networks. *Transportation Modelling for Tomorrow*. Delft University Press, The Netherlands.

Bovy, P.H.L. and Stern, E. (1990) *Route Choice: Wayfinding in Transport Networks*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Boyce, D.E., Ran, B., and Li, I.Y. (1999) Considering travelers' risk-taking behavior in dynamic traffic assignment. *Transportation Networks: Recent Methodological Advances,* M.G.H. Bell (ed.), Elsevier, Oxford, 67-81.

Burrell, J.E. (1968) Multipath route assignment and its application to capacity restraint. *Proceedings of the Fourth International Symposium on Transportation and Traffic Theory,* Karlsruhe, Germany.

Cantarella, G.E. (1997) A general fixed-point approach to multimode multi-user equilibrium assignment with elastic demand. *Transportation Science* 5, 541-553.

Cantarella, G.E. and Binetti, M. (1998) Stochastic equilibrium traffic assignment with value-of-time distributed among users. *International Transactions in Operational Research* 31, 107-128.

Cantarella, G. E., Improta, G. and Sforza, A. (1991) Iterative procedure for equilibrium network traffic signal setting. *Transportation Research* 25A, 241-249.

Cascetta, E. (2001) *Transportation Systems Engineering: Theory and Methods*. Kluwer Academic Publishers.

Cascetta, E., Nuzzolo, A., Russo, F., and Vitetta, A. (1996) A modified logit route choice model overcoming path overlapping problems: specification and some calibration results for interurban networks. *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, Leon, France, 697-711.

Cascetta, E., Russo, F., and Vitetta, A. (1997) Stochastic user equilibrium assignment with explicit path enumeration: comparison of models and algorithms. *Preprint of the 8th IFAC Symposium on Transportation Systems*, Chania, Greece.

Cascetta, E., Russo, F., Viola, F., and Vitetta, A. (2002) A model of route perception in urban road networks. *Transportation Research* 36B, 577-592.

Chen, A. (2001) Effects of flow update strategies on implementation of the Frank-Wolfe algorithm for the traffic assignment problem, *Transportation Research Record* 1771, 132-139.

Chen, A., Jayakrishnan, R., and Tsai, W.K. (2002) Faster Frank-Wolfe traffic assignment with new flow update scheme. *Journal of Transportation Engineering* 128, 31-39.

Chen, A., Lee, D.-H., and Jayakrishnan, R. (2002) Computational study of state-of-the-art path-based traffic assignment algorithms, *Mathematics and Computers in Simulation* 59, 509-518.

Chen, A., Lee, D.-H., Nie, Y., and Jayakrishnan, R. (2000) Path and Link Traffic Assignment Algorithms: A Comprehensive Computational Study, paper presented at *the 6th International Conference on Applications of Advanced Technologies in Transportation Engineering*, Singapore.

Chen, A., Lo, H.K., and Yang, H. (2001) A *self-adaptive* projection and contraction algorithm for the traffic equilibrium problem with path-specific costs, *European Journal of Operations Research* 135(1), 27-41.

Chen, A., Oh, J., Park, D., and Recker, W. (2000) An Algorithm for the Elastic, Bi-Criteria Traffic Equilibrium Problem with Nonlinear Path Costs, paper presented at the 79th Transportation Research Board Annual Meeting, Washington, D.C.

Chen, A. and Recker, W. (2001) Considering risk taking behavior in travel time reliability. Paper presented at *the 80th Annual Meeting of the Transportation Research Board*, Washington, D.C., USA.

Chen, A., Tatineni, M., Lee, D. H., and Yang, H. (2000) Effect of Route Choice Models on Estimating Network Capacity Reliability, *Transportation Research Record* 1733, 63-70.

Chen, A., Yang, H., Lo, H.K., and Tang, W. (1999) A Capacity Related Reliability for Transportation Networks, *Journal of Advanced Transportation*, Vol. 33, No. 2, 183-200.

Chen, A., Yang, H., Lo, H.K., and Tang, W. (2002) Capacity Reliability of a Road Network: An Assessment Methodology and Numerical Results, *Transportation Research* 36B, 225-252.

Chen, C., Skabadonis, A., and Varaiya, P (2003) Travel-time reliability as a measure of service, *Transportation Research Record* 1855, pp 74-79.

172

Chen, M. and Alfa, A.S. (1991) Algorithms for solving Fisk's stochastic traffic assignment model. *Transportation Research* 25B, 405-412.

Cheu, R., Jin, X., Ng, K., and Ng, Y., (1998) Calibration of FRESIM for Singapore Expressway Using Genetic Algorithm, *ASCE Journal of Transportation Engineering*, Vol. 124, No. 6, 526-535.

Chiou, S-W (1999) Optimization of area traffic control for equilibrium network flows. *Transportation Science* 33, 279-289.

Dafermos, S. (1980) Traffic equilibrium and variational inequality. *Transportation Science* 14, 42-54.

Daganzo, C. (1983) Stochastic network equilibrium with multiple vehicle types and asymmetric, indefinite link cost Jacobians. *Transportation Science* 17, 282-300.

Daganzo, C. and Sheffi, Y. (1977) On stochastic models of traffic assignment. *Transportation Research* 14B, 243-255.

Daganzo, C. and Sheffi, Y. (1982) Unconstrained extremal formulation of some transportation equilibrium problems. *Transportation Science* 16, 332-360.

Damberg, O., Lundgren, J.T., and Patriksson, M. (1996) An algorithm for the stochastic user equilibrium problem. *Transportation Research* 30B, 115-131.

Dial, R. (1971) A probabilistic multipath traffic assignment model which obviates path enumeration. *Transportation Research* 5, 83-111.

Dial, R. (1996) Bicriterion traffic assignment: basic theory and elementary algorithms. *Transportation Science* 30, 93-111.

Dial, R. (1997) Bicriterion traffic assignment: efficient algorithms plus examples. *Transportation Research* 31B, 357-379.

Du, Z.P. and Nicholson, A.J. (1997) Degradable transportation systems: sensitivity and reliability analysis. *Transportation Research* 31B, No. 3, 225-237

Emmerink, R.H.M. (1997) *Information and Pricing in Road Transportation*. Springer.

ESRI (2001) Linear Referencing and Dynamic Segmentation in ArcGIS 8.1, ESRI technical paper, May, 2001.

Ferrari, P. (1995) Road pricing and network equilibrium. *Transportation Research* 29B, 357-372.

Ferrari, P. (2002) Road network toll pricing and social welfare. *Transportation Research* 36B, 471-483.

Fisk, C. (1980) Some developments in equilibrium traffic assignment. *Transportation Research* 14B, 243-255.

Fisk, C. S. (1984) Optimal signal controls on congested networks. *Proceedings of the Ninth International Symposium on Transportation and Traffic Theory,* VNU Science Press, Delft, The Netherlands, 197-216.

Florian, M., Guelat, J., and Spiess, H. (1987) An efficient implementation of the 'PARTAN' variant of the linear approximation method for the network equilibrium problem. *Networks*, **17**, 319-339.

Fukushima, M. (1984) A modified Frank-Wolfe algorithm for Solving the traffic assignment problem. *Transportation Research*, 18B, 169-177.

Gabriel, S. and Bernstein, D. (1997) The traffic equilibrium problem with nonadditive path costs. Transportation Science, Vol. 31, No. 4, 337-348.

Gaver, D.P., JR. (1968) Headstart Strategies for Combating Congestion. Transportation Science, **2**, 172-181.

Gen, M., and Cheng, R., (2000) *Genetic Algorithms and Engineering Optimization*, New York: Wiley.

Gliebe, J.P., Koppleman, F.S., and Ziliaskopoulos, A. (1999) Route choice using a paired combinatorial logit model. Paper presented at *the 78th Annual Meeting of the Transportation Research Board*, Washington, D.C., USA.

Ghosh, A. (2001) Heterogeneity in value of time: Revealed and stated preference estimates from the I-15 congestion pricing project. Paper presented at *the 80th Annual Meeting of the Transportation Research Board*, Washington, D.C., USA.

Hearn, D., Lawphongpanich, S., and Ventura, J.A. (1985) "Finiteness in Restricted Simplicial Decomposition," *Operations Research Letters*, 4, No. 3, 125-130.

Hicks, J.E., Boyce, D.E., and Sen, A. (1992) Static network equilibrium models and analyses for the design of dynamic route guidance systems, Technical Report, ADVANCE Project, University of Illinois, Chicago.

Holland, J. H. (1975), *Adaptation in Natural and Artificial Systems*, University of Michigan Press (reprinted in 1992 by MIT Press, Cambridge, MA).

Iida, Y., (1999) Basic concepts and future directions of road network reliability analysis. *Journal of Advanced Transportation*, Vol. 33, No. 2, 125-134.

Iida, Y., Wakayabashi, H., (1989) An approximation method of terminal reliability of road network using partial path and cut set. Proceedings of *the 5th World Conference*, vol. IV, Yokohama, Japan, pp. 367-380.

Jackson, W.B. and Jucker, J.V., (1981) *An Empirical Study of Travel Time Variability and Travel Choice Behavior*, *Transportation Science*, Vol. 16, No. 4, pp460-475.

Janson, B. and Gorostiza, C.Z. (1987) The problem of cyclic flows in traffic assignment. *Transportation Research* **21B**, 299-310.

Jayakrishnan, R., Tsai, W.K., Prashker, J.N., and Rajadhyaksha, S. (1994) A faster path-based algorithm for traffic assignment. *Transportation Research Record* 1443, 75-83.

Kazimi, C., Brownstone, D., Ghosh, A., Golob, T.F., and van Amelsfort, D. (2000) Willingness-to-pay to reduce commute time and its variance: Evidence from the San Diego I-15 congestion pricing project. Paper presented at *the 79th Annual Meeting of the Transportation Research Board*, Washington, D.C., USA.

Kim, H. Oh, J., and Jayakrishnan, R., (2004) A dynamic simulation model for urban taxi service, UCI-ITS Working paper.

Kim, K., and Rilett, L.R., (2001) Genetic-Algorithm-based Approach for Calibrating Microscopic Simulation Models, *IEEE Intelligent Transportation Systems Proceedings*, Oakland, California, pp.698-704.

Knight, T.E., (1974) An approach to the evaluation of changes in travel unreliability: A 'Safety Margin' hypothesis, *Transportation* 3, 393-408.

Lam, T. (2000) The effect of variability of travel time on route and time-of-day choice, Ph.D. Dissertation, Department of Economics, University of California, Irvine.

Lam, T. and Small, K. (2001) The value of time and reliability: measurement from a value pricing experiment. *Transportation Research* 37E, 231-251.

Lam, W.H.K., and Xu, G., (2000) *Calibration of Traffic Flow Simulator for Network Reliability Assessment*, Chapter 10 in the book "Reliability of Transport Networks", Edited by Bell, M., and Cassir C., Research Studies Press LTD.

Lam, W.H.K. and Zhang, N. (2000) A New Concept of Travel Demand Satisfaction Reliability for Assessing Road Network Performance, paper presented at Matsuyama Workshop on Transport Network Analysis, Matsuyama, Japan.

Larsson, T. and Patriksson, M. (1992) Simplicial decomposition with disaggregate representation for the traffic assignment problem. *Transportation Science*, 26, 4-17.

LeBlanc, L.J., Morlok, E.K., and Pierskalla, W.P. (1975) An efficient approach to solving the road network equilibrium traffic assignment problem. *Transportation Research*, **9**, 309-318.

LeBlanc, L.J., Helgason, R.V., and Boyce, D.E. (1985) Improved efficiency of the Frank-Wolfe algorithm for convex network problems. *Transportation Science*, 19, 445-462.

Lee, D.-H. and Nie, Y. (2001) Accelerating strategies and computational studies of the Frank-Wolfe algorithm for the traffic assignment problem. *Transportation Research Record* 1771, 97-105.

Lee, D.-H., Nie, Y., and Chen, A. (2003) A conjugate gradient projection algorithm for the traffic assignment problem, *Mathematical and Computer Modeling*, Vol. 37, 863-878.

Lee, D.-H., Nie, Y., Chen, A., and Leow, Y.C. (2002) Link and path based traffic assignment algorithms: a computational and statistical study, *Transportation Research Record* 1783, 80-88.

Leurent, F.M. (1997) Curbing the computational difficulty of the logit equilibrium assignment model. *Transportation Research* 31B, 315-326.

Levinson, D. and Zhang, L. (2001) Travel time variability after a shock: the case of the Twin Cities ramp meter shut off, Presentation paper at *First International Symposium on Transportation Network Reliability*, Kyoto, Japan, Jul. 30 – Aug. 1, 2001.

Liu, H, Ban, X., Ran, B., and Mirchandani, P. (2002) An analytical dynamic traffic assignment model with probabilistic travel times and travelers' perceptions. *Transportation Research Record* (in press).

Liu, H., Recker, W., Chen, A. (2004) Uncovering the Contribution of Travel Time Reliability to Dynamic Route Choice Using Real-time Loop Data, submitted to Transportation Research 38A, 435-453.

Lo, H. (2002) Trip Travel Time Reliability in Degradable Transport Networks. Transportation and Traffic Theory. Edited by M. Taylor, Elsevier Science, 541-560.

Lo, H.K. and Chen, A. (2000) Traffic equilibrium problem with route-specific costs: formulation and algorithms. *Transportation Research* 34B, 493-513.

Lo, H.K., Chen, A., and Yang, H. (1999) Travel time minimization in route guidance with elastic market penetration. *Transportation Research Record* 1667, 25-32.

Lo, H. and Tung, Y.K. (2003) Network with degradable Links: Capacity Analysis and Design. Transportation Research, 37B, 345-363.

Lomax, T., Turner, T., and Margoitta, R. (2003) Monitoring Urban Roadways in 2001: Examining Reliability and Mobility with Archived Data, FHWA-OP-03-141,Texas Transportation Institute and Cambridge Systematics, Inc.

Lomax, T., Schrank, D., Turner, T., and Margoitta, R. (2003) Selecting Travel Time Reliability Measures, Texas Transportation Institute and Cambridge Systematics, Inc.

McFadden, D., and Train, K., (2000) *Mixed MNL Models for Discrete Response, Journal of Applied Econometrics*, 15, pp447-470.

Maher, M. (1998) Algorithms for logit-based stochastic user equilibrium assignment. *Transportation Research* 32B, 539-549.

Maher, M. and Hughes, P. (1997) A probit based stochastic user equilibrium assignment problem. *Transportation Research* 31B, 341-355.

Mette, K.S. and F. Bent (1996) Accuracy of traffic forecasts and cost estimates on large transportation projects. *Transportation Research Record*, 1518, pp. 65-69.

Meyer, M.D. and Miller, E.J. (2001) *Urban Transportation Planning: A Decision-Oriented Approach,* Second Edition. McGraw Hill.

Mirchandani, P. and Soroush, H. (1987) Generalized traffic equilibrium with probabilistic travel times and perceptions. *Transportation Science* 21, 133-152.

Nagurney, A. (1999) *Network Economics: A Variational Inequality Approach.* Kluwer Academic Publishers.

Nicholson, A.J. and Du, Z.P. (1997) Degradable Transportation Systems: an Integrated Equilibrium Model. Transportation Research, vol. 31B, 209-224.

Ortuzar, J. de D., and Willumsen, L.G. (2001) *Modeling Transport*. Third Edition, John Wiley & Sons Ltd.

Park, D. (1998) Multiple Path Based Vehicle Routing in Dynamic and Stochastic Transportation Networks, Ph.D. Dissertation, Department of Civil Engineering, Texas A&M University.

Patriksson, M. (1994) *The Traffic Assignment Problem: Models and Methods*. VSP, Utrecht, The Netherlands.

Pells, SR (1987) The Evaluation of Reductions in Travel Time Variability. PhD Dissertation, Institute for Transport Studies, School of Economic Studies, University of Leeds

Prashker, J.N. and Bekhor, S. (1998) Investigation of stochastic network loading procedures. *Transportation Research Record* 1645, 94-102.

Prashker, J.N. and Bekhor, S. (2000) Congestion, stochastic and similarity effects in stochastic user equilibrium models. . *Transportation Research Record* 1733, 80-87.

Poylak, J.W., (1987) Travel Time Variability and Departure Time Choice: A Utility Theoretic Approach. Transport Studies Group, Discussion Paper 15. (London: Polytechnic of Central London).

Ramming, M.S. (2002) Network knowledge and route choice, Ph.D. Dissertation, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology.

Richardson, A. J., (2003) Travel time variability on an urban freeway, The Urban Transport Institute (TUTI) Report 25-2003.

Scott, K. and Bernstein, D. (1998) Solving a traffic equilibrium problem when paths are nonadditive. Paper presented at the 78th Annual Meeting of the Transportation Research Board, Washington, D.C., USA.

Senna, L.A.D.S., (1994) The Influence of Travel Time Variability on The Value of Time, Transportation, 21, 203-228.

Sheffi, Y. (1985) Urban transportation networks: Equilibrium analysis with mathematical programming methods, *Prentice Hall, England Cliffs, NJ.*

Sheffi, Y. and Powell, W. (1981) A comparison of stochastic and deterministic traffic assignment over congested networks. *Transportation Research* 15B (1), 53-64.

Sheffi, Y. and Powell, W. (1982) An algorithm for the equilibrium assignment problem with random link times. *Networks* 12 (2), 191-207.

Small, K.A., (1982) The Scheduling of Consumer Activities: Work Trips. American Economic Review, 72, 467-479.

Small, K.A., Noland, R., Chu, X., and Lewis, D. (1999) Valuation of Travel-Time Savings and Predictability in Congested Conditions for Highway User-Cost Estimation. NCHRP Report 431, Transportation Research Board, National Research Council.

Small, K. A., Winston, C., and Yan, J., (2002) *Uncovering the Distribution of Motorists' Preference for Travel Time and Reliability: Implications for Road Pricing*, Submitted to Economica.

Smith, M. J. and Van Vuren, T. (1993) Traffic equilibrium with responsive traffic control. *Transportation Science* 27, 118-132.

Soroush, H. (1984) Routes and flows through stochastic networks. Ph.D. Dissertation, Operations Research and Statistics Program. RPI, Troy, New York, USA.

Sullivan, E., (2000) *Continuation Study to Evaluate the Impacts of the SR91 Value-Priced Express Lanes*, Final Report Submitted to the Department of Transportation of the State of California, See http://gridlock.calpoly.edu/sr91/sr91.htm.

Tatineni, M. (1996) Solution properties of stochastic route choice models. Ph.D. Dissertation, University of Illinois at Chicago.

Tatineni, M., Boyce, D.E., and Mirchandani, P. (1997) Comparisons of deterministic and stochastic traffic loading models. *Transportation Research Record* 1607, 16-23.

Thomson, JM (1968) The Value of Traffic Management, Journal of Transport Economics and Policy, 1(1), 3-32.

Train, K., (2002) *Discrete Choice Methods with Simulation*, Cambridge University Press.

Uchida, T. and Iida, Y. (1993) Risk assignment: a new traffic assignment model considering risk of travel time variation. *Proceedings of the 12th International Symposium on Transportation and Traffic Theory*. Elservier, Amsterdam, 89-105.

Van Vuren, T. and Van Vliet, D. (1992) *Route Choice and Signal Control*. Athenaeum Press Ltd., Newcastle upon Tyne.

Van Vuren, T. and Watling, D. (1991) A multiple user class assignment model for route guidance. *Transportation Research Record*, 1306, 21-31.

Von Falkenhausen, H. (1976) Traffic assignment by a stochastic model, *Proceedings of the Fourth International Confeence on Operational Research*, 415-421.

Vovsha, P. and Bekhor, S. (1998) The link nested logit model: overcoming the route overlapping problem. *Transportation Research Record* 1645, 133-142.

Wakabayahsi, H. and Kameda, H. (1992) Network Performance of Highway Systems under Earthquake Effects: A Case Study of the 1989 Loma Prieta Earthquake. Proceedings U.S. - Japan Workshop on Earthquake Disaster Prevention for Lifeline Systems, 215-232, Public Works Research Institute, Japan.

Wardrop, J.G. (1952) Some theoretical aspects of road traffic research, Proceedings of the Institute of Civil Engineers, Part II, 1, 325-378.

Weintraub, A., Ortiz, C., and Gonzalez, J. (1985) Accelerating convergence of the Frank-Wolfe algorithm. *Transportation Research*, 19B, 113-122.

Wong S. C. and Yang, C. (1999) An iterative group-based signal optimization scheme for traffic equilibrium networks, *Journal of Advanced Transportation* 33, 201-217.

Wong S. C. and Yang, H. (1997) Reserve capacity of a signal-controlled road network, *Transportation Research* 31B, 397-402.

Yang, H. (1998) Multiple equilibrium behaviors and advanced traveler information systems with endogenous market penetration. *Transportation Research* 32B, 205-218.

Yang, H. and Bell, M. G. H. (1997) Traffic restraint, road pricing and network equilibrium. *Transportation Research* 31B, 303-314.

Yang, H. and Lam, W. H. K. (1996) Optimal Road Tolls under Conditions of Queuing and Congestion. *Transportation Research* 30A, 319-332.

Yang, H. and Huang, H.J. (1998) Principle of marginal-cost pricing: How does it work in a general network? *Transportation Research* 32A, 45-54.

Yang, H. and Yagar, S. (1994) Traffic assignment and traffic control in general freeway-arterial corridor systems. *Transportation Research* 28B, 463-486.

Yang, H. and Yagar, S. (1995) Traffic assignment and signal control in saturated road network, *Transportation Research* 29A, 125-139.

Yang, H., Yagar, S. Iida, Y. and Asakura, Y. (1994) An algorithm for inflow control problem on urban freeway networks with user-optimal flows, *Transportation Research* 28B, 123-139.

Yin, Y. and Ieda, H. (2001) Assessing performance reliability of road networks under nonrecurrent congestion. *Transportation Research Record* 1771, 148-155.