

UC Berkeley

Research Reports

Title

Development of a Path Flow Estimator for Deriving Steady-State and Time-Dependent Origin-Destination Trip Tables

Permalink

<https://escholarship.org/uc/item/8q85121w>

Authors

Chen, Anthony
Chootinan, Piya
Recker, Will
et al.

Publication Date

2004-09-01

CALIFORNIA PATH PROGRAM
INSTITUTE OF TRANSPORTATION STUDIES
UNIVERSITY OF CALIFORNIA, BERKELEY

Development of a Path Flow Estimator for Deriving Steady-State and Time-Dependent Origin-Destination Trip Tables

**Anthony Chen, Piya Chootinan,
Will Recker, H. Michael Zhang**

**California PATH Research Report
UCB-ITS-PRR-2004-29**

This work was performed as part of the California PATH Program of the University of California, in cooperation with the State of California Business, Transportation, and Housing Agency, Department of Transportation; and the United States Department of Transportation, Federal Highway Administration.

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 4135

September 2004

ISSN 1055-1425

CALIFORNIA PARTNERS FOR ADVANCED TRANSIT AND HIGHWAYS

Development of a Path Flow Estimator for Deriving Steady-State and Time-Dependent Origin-Destination Trip Tables

REPORT prepared for California PATH

By

Anthony Chen and Piya Chootinan
Department of Civil and Environmental Engineering
Utah State University
Logan, UT 84322-4110

Will Recker
Institute of Transportation Studies
University of California, Irvine
Irvine, CA 92697-3600

H. Michael Zhang
Department of Civil and Environmental Engineering
University of California, Davis
Davis, CA 95616

September 2004

Development of a Path Flow Estimator for Deriving Steady-State and Time-Dependent Origin-Destination Trip Tables

ABSTRACT

The origin-destination (O-D) trip table is a key input required for traffic assignment and simulation models utilized to analyze a wide variety of transportation applications. The main goal of this research is to develop an economical and quick method for estimating O-D trip tables from traffic counts. Path flow estimator (PFE), originally developed by Bell and Shield (1995), has been further developed to improve the reliability and efficiency of O-D trip table estimates. The research reported herein includes only the development of the steady-state O-D estimator. In this study, the original PFE model was carefully examined in several aspects to gain more insight for further improvements. Currently, the PFE has been successfully applied to estimate the steady-state O-D trip tables for the Irvine Testbed network in Orange County, California as well as some other real networks. The primary results demonstrate that PFE has the capability to correctly estimate the total and individual O-D demands when proper information is provided. They also indicate that the number and locations of traffic counts significantly influence the quality of O-D estimates as each observation contributes different amount and quality of information. The most difficult task observed thus far is the estimation of spatial pattern of O-D demands even when traffic counts were collected on all network links. These issues and the development of time-dependent PFE will be investigated in the second phase under Task Order 5502.

Keyword: Origin-destination Trip Table Estimation, Path Flow Estimator, Stochastic User Equilibrium

Development of a Path Flow Estimator for Deriving Steady-State and Time-Dependent Origin-Destination Trip Tables

EXECUTIVE SUMMARY

The origin-destination (O-D) trip table is a key input required for traffic assignment and simulation models utilized to analyze a wide variety of transportation applications. The traditional survey techniques, such as household survey, roadside interview, etc., are expensive, time consuming, and labor intensive; therefore, they are not considered appropriate due to financial and time constraints. The development of economical and quicker methods for estimating O-D trip tables using readily available traffic measurements is the main goal of this research (Task Order 4135). The work proposed under Task Order 4135 is to develop a path flow estimator (PFE) as a software tool for deriving both steady-state and time-dependent origin-destination trip tables for various transportation applications, from long-term transportation planning to short-term transportation management and information applications. The PFE, originally developed by Bell and Shield (1995), is a one-stage network observer that estimates path flows (hence O-D flows) and path travel times from traffic counts collected from a transportation network. The core component of the PFE is a logit path choice model (i.e., demand) that interacts with link cost functions (i.e., supply) to produce a stochastic user equilibrium traffic pattern. The PFE is a proven operational model that has been tested in a number of projects for both off-line and on-line applications in Europe.

Phase one of this project has primarily explored the PFE for estimating steady-state O-D trip tables and developed several promising algorithms. Research activities in the first phase include formulating and developing algorithms, based on Bell and Shield's original model, to solve the steady-state PFE, enhancing the PFE solution algorithms, adapting the PFE to accept different types of inputs (e.g., traffic counts with and without measurement errors, initial set of paths, initial trip table, etc.), examining the effect of different versions of PFE on O-D trip table estimates, developing strategies to handle operational problems in PFE, comparing with bi-level O-D trip table estimation techniques, validating the PFE with real world data from various sources (e.g., portion of the Orange County Transportation Analysis Model, OCTAM; traffic counts, movement counts, etc.), and quantifying accuracy, reliability, and quality of the O-D trip table estimates. In addition to the development of solution algorithms for PFE, a prototype of graphical user interface (GUI) has been developed to facilitate data communication and result display between users and the PFE. However, it is not fully functioning at this time. Some incomplete parts of the first phase research activities may carry on to the second phase (this is dependent on the availability of data from other projects).

Currently, the PFE has successfully been applied to estimate the steady-state O-D trip table for the Testbed Irvine network in Orange County, California. This network consists of three major freeways (I-5, I-405, and SR-133), and several arterials in the City of Irvine. In this application, the Irvine network and associated demand data were extracted from the OCTAM, which contains data for the whole county. The extracted network is composed of 163 nodes, 496 links connecting 39 traffic analysis zones (TAZ), 28 external stations, and 1,547 O-D pairs. Besides the Irvine network, the PFE has also been tested on several hypothetical and real networks in

order to study the properties of its estimation results. Using both simple and real networks, we have demonstrated that PFE has the capability to correctly estimate the total and individual O-D demands when proper information is provided. Based on the preliminary results, it has been determined that the selection of observed links (locations of traffic counts) plays an important role in the O-D estimation problem, as each observation contributes different amount and quality of information. Though the quality of estimates can be improved with the usage of more traffic counts, this may be impractical when the budget for data collection is scarce. From the small network, the results show that if there is at least one observation on each path, the total demand of the network can be correctly captured. In addition, higher observed traffic volumes appear to contribute more to the quality of O-D estimation. The most difficult task observed thus far is the estimation of the spatial patterns of O-D demands. Even when all network links are measured, the individual O-D demands may not be estimated correctly. Some of these issues will be further examined in the second phase under Task Order 5502. The focus of phase two is to develop a time-dependent version of the PFE for deriving time-dependent O-D trip table needed to support a wide variety of transportation applications in advanced transportation management and information systems (ATMIS).

TABLE OF CONTENT

	Page
EXECUTIVE SUMMARY	i
LIST OF TABLES.....	vi
LIST OF FIGURES	vii
1 OVERVIEW	1
1.1 Project Background	1
1.2 Project Status.....	2
1.3 Organization of Report	3
2 LITERATURE REVIEW.....	5
2.1 Steady-state Trip Table Estimation	5
2.1.1 Simple Network with no Route Choice	5
2.1.2 Network with no Congestion.....	6
2.1.3 Network with both Route Choice and Congestion	7
2.2 Time-dependent Trip Table Estimation.....	9
2.2.1 Intersection-Oriented Dynamic O-D Estimation.....	9
2.2.2 Network-Oriented Dynamic O-D Estimation	10
2.3 Performance Measures of O-D Trip Table Estimates	15
2.3.1 Statistical Measures of Estimation Error	15
2.3.2 Maximum possible relative error (MPRE).....	15
2.3.3 Total Demand Scale (TDS).....	16
3 PATH FLOW ESTIMATOR.....	18
3.1 Notations	18
3.2 Mathematical Formulation.....	19
3.3 Solution Procedure	20
3.3.1 Iterative Balancing Technique.....	20
3.3.2 Column Generation.....	22
3.4 Examining O-D Trip Table Estimated by PFE	22
3.4.1 Inconsistencies of Traffic Counts.....	22

3.4.2	Effects of Number and Locations of Traffic Counts	23
3.4.3	Effect of Measurement Errors of Traffic Counts	33
3.5	Derivation of Intersection Turning Movements by PFE	37
3.5.1	A Simple Isolated Intersection	38
3.5.2	A Signalized Arterial	41
3.5.3	Mixed Freeway and Surface Streets	45
4	SENSITIVITY ANALYSIS FOR pFE MODEL	49
4.1	Derivation of Sensitivity Expression.....	49
4.2	Determination of $\nabla_y \mathbf{H}^{-1}$ and Sensitivities of Path Flow Estimates.....	53
4.3	Determination of the Independence Constraints	56
4.4	Variance of Estimation Results	59
4.4.1	First Test Network	59
4.4.2	Second Test Network.....	61
5	DEVELOPMENT OF THE BI-LEVEL ESTIMATOR.....	67
5.1	Lower Level Problem: Route Choice Models.....	68
5.1.1	Deterministic User Equilibrium	70
5.1.2	Logit-Based Stochastic User Equilibrium.....	70
5.1.3	Path Based Algorithms for Solving Traffic Assignment Problem	73
5.1.4	Direction Finding (item 11).....	76
5.1.5	Line Search or Stepsize Selection (item 13)	78
5.2	Upper Level Problem: Statistical Inferences	79
5.2.1	Maximum Entropy Estimator.....	79
5.2.2	Generalized Least Square.....	79
5.3	Solution Techniques	80
5.3.1	Iterative Procedure.....	80
5.3.2	Genetic Algorithms.....	81
6	DATA COLLECTION.....	84
6.1	Network Data	84
6.2	O-D Demand Data From OCTA Model.....	85
6.3	Traffic Counts	86
6.4	Travel Time Patterns	86

7	DEVELOPMENT OF GRAPHICAL USER INTERFACE	89
7.1	Introduction.....	89
7.2	System Architecture	89
7.3	Implementing the PFE GUI (Prototype).....	90
7.3.1	Required Files and Folders.....	90
7.3.2	Ruining PFE	90
7.3.3	Displaying PFE results.....	91
7.4	Enhancement of PFE GUI	92
7.4.1	Network Editor	93
7.4.2	Data Conversion	94
7.4.3	Estimation Control.....	94
7.4.4	Diagnostic Tools.....	94
8	CONCLUSIONS AND FURTHER RESEARCHES	95
8.1	Findings and Conclusions	95
8.1.1	Software	95
8.1.2	Data Collection.....	95
8.2	Tasks to be Complete under the Continuation.....	95
9	REFERENCE.....	97

LIST OF TABLES

	Page
Table 1-1 Applications of PFE (Bell and Grosso, 1998).....	1
Table 2-1 Classification of O-D Trip Table Estimation Methods.....	8
Table 3-1 Link Characteristics of Hypothetic Network	25
Table 3-2 O-D Demand Estimates with Different Sets of Traffic Counts	25
Table 3-3 Path Set of Hypothetic Network.....	26
Table 3-4 Summary of Estimation Results for Irvine Network	29
Table 3-5 Estimation Results for Freeway O-D pairs of Irvine Network	30
Table 3-6 Effect of Error Bounds Selection on the Estimation Results	35
Table 3-7 Link Characteristics of an Isolated Intersection.....	39
Table 3-8 Intersection Turning Movement Delays (Turn Penalties)	40
Table 3-9 Estimation Results using Entry Flows only ($RMSE_{OD} = 1,314.04$)	40
Table 3-10 Estimation Results using Entry and Exit Flows ($RMSE_{OD} = 160.95$)	40
Table 3-11 O-D Trip Table Estimated using Observations on Entry and Exit Links	44
Table 3-12 Comparisons of Turning Movement Estimations ($RMSE = 150.187$)	47
Table 4-1 Estimated Path Flows	59
Table 4-2 Link Flows, Link Capacities, and Dual Variables.....	59
Table 4-3 Sensitivities of Path Flows to Link Observations and Link Capacities.....	61
Table 4-4 Estimated Path Flows	62
Table 4-5 Link Flows, Dual Variables, and Link Characteristics.....	62
Table 4-6 Dual Variables of Independent Constraints with the Maximal Observed Flows.....	63
Table 4-7 Sensitivities of Path Flow Estimates to Link Observations	64
Table 4-8 Estimated Path Flows with Simultaneously Perturbed Link Observations	65
Table 6-1 Node Numbering Systems between OCTAM and PFE.....	85
Table 6-2 List of Intersections with Traffic Counts.....	87

LIST OF FIGURES

	Page
Figure 1-1 Irvine network, Orange County, California.	3
Figure 3-1 Synthetic traffic counts for hypothetical network.....	24
Figure 3-2 Locations of different sets of traffic counts for Irvine network.....	28
Figure 3-3 Comparison of true and estimated O-D demands with different sets of traffic counts.	31
Figure 3-4 Comparison of true and estimated freeway O-D using different sets of traffic counts.	32
Figure 3-5 Comparison of observed link flows and estimated link flows at different selected error bounds.	36
Figure 3-6 Convention of link orientation.	37
Figure 3-7 Intersection layout, a) turning-movement counts, and b) entry and exit flows.....	39
Figure 3-8 Layout of Arterial Network, a) entry and exit flow observations, and b) turning movement counts.....	43
Figure 3-9 Estimated turning movements for arterial network ($RMSE_{turn} = 209.048$).	44
Figure 3-10 Comparison of derived and observed turning movements.	45
Figure 3-11 Layout of mixed freeway and surface streets, City of Santa Rosa, CA, USA.	46
Figure 3-12 Comparison between estimated and observed traffic counts.	47
Figure 4-1 Test network with 6 link and 4 O-D pairs.....	59
Figure 4-2 Test network with 7 links and 2 O-D pairs.	62
Figure 6-1 Irvine Network, Orange County, California.	84
Figure 6-2 Intersection map for Irvine network.	88
Figure 7-1 Location of PFE required folders and PFE executable file.....	91
Figure 7-2 Display PFE outputs dialog box.	92
Figure 7-3 Views of PFE outputs.....	93

1 OVERVIEW

1.1 Project Background

The origin-destination (O-D) trip table is a key input required for traffic assignment and simulation models utilized to analyze a wide variety of transportation applications, ranging from long term planning of transportation infrastructure to day-to-day operations of transportation management and information systems. Traditionally, the O-D trip tables are estimated through large-scale surveys such as household interviews, roadside interviews, license plate surveys, etc. These survey techniques are expensive, time consuming, and labor intensive. Moreover, if there is a rapid change in land-use development, the trip tables will soon become outdated. Therefore, the O-D survey-based methods are not considered appropriate due to financial and time constraints. The development of economical and quicker methods for estimating O-D trip tables using readily available traffic measurements is the main goal of this research. The work proposed under Task Order 4135 is to develop a path flow estimator (PFE) as a software tool for deriving both steady-state and time-dependent origin-destination trip tables for various transportation applications, from long range transportation planning to short-term transportation management and information applications.

The PFE, originally developed by Bell and Shield (1995), is a one-stage network observer that estimates path flows (hence O-D flows) and path travel times from traffic counts collected from a transportation network. The core component of the PFE is a logit path choice model (i.e., demand) that interacts with link cost functions (i.e., supply) to produce a stochastic user equilibrium traffic pattern. The PFE can be implemented for both off-line transportation planning and on-line traffic management applications. The PFE is a proven operational model that has been tested in a number of projects in Europe. The specific objectives of these projects, in regards to the PFE, are listed in Table 1-1.

Table 1-1 Applications of PFE (Bell and Grosso, 1998)

Projects	Objectives regarding the PFE
CLEOPATRA	Specification and validation of the PFE for various forms of urban traffic management and control system in Lyon, Stockholm, Toulouse and Turin
COSMOS	Specification and validation of the PFE in support of the MOTION traffic signal control system from Siemens and the MAKSIMOS rerouting optimization tool from the Technische Universitat Hamburg
PTV	Integration of the PFE into the VISION software from PTV GmbH for transportation planning applications
AIUTO	Extension of the PFE to multi-modal networks with specific reference to the City of York
EPSRC	Use of the PFE to study traffic signal control with specific reference to the City of York
TASTe	Use of the PFE, as part of a TDM analysis tool-box, to study Park-and-Ride in Barcelona
DETR	Use of the PFE to study traffic diversion due to motorway tolls

To accomplish the goal of this research, the PFE is tailored to meet the needs of Caltrans as well as developed further to improve the reliability and efficiency of O-D trip table estimates. This research effort consists of two phases for the development of the steady-state and time-dependent O-D trip table estimator for transportation applications. The first phase (nearing completion) is focused on the development of the PFE for deriving steady-state O-D trip tables. Research activities in the first phase include formulating and developing algorithms, based on Bell and Shield's original model, to solve the steady-state PFE, enhancing the PFE solution algorithms, adapting the PFE to accept different types of inputs (e.g., traffic counts with and without measurement errors, initial set of paths, initial trip table, etc.), examining the effect of different versions of PFE on O-D trip table estimates, developing strategies to handle operational problems in PFE, comparing with bi-level O-D trip table estimation techniques, validating the PFE with real world data from various sources (e.g., portion of the OCTAM model from OCTA, traffic counts, movement counts, etc.), and quantifying accuracy, reliability, and quality of the O-D trip table estimates. Some incomplete parts of the first phase research activities may carry on to the second phase (this is dependent on the availability of data from other projects).

1.2 Project Status

The specific objectives for phase one of Task Order 4135 are to:

1. Develop a path flow estimator for steady-state O-D trip tables,
2. Develop solution algorithms for PFE to accept different types of inputs,
3. Examine the effect of inputs on the estimation of O-D trip tables,
4. Compare with the bi-level O-D trip table estimator,
5. Collect data for a case study,
6. Quantify the accuracy, reliability, and quality of O-D trip table estimates,
7. Validate the steady-state PFE model,
8. Develop a graphical user interface for PFE.

At present, we have completed objectives 1 to 4 (part of the reason for this delay is that the contract was awarded several months later than the planned project start date). We are currently in the process of collecting various traffic data (e.g., freeway data from the Irvine Testbed area of Orange County, traffic counts from the surface streets, initial O-D trip table extracted from portion of the OCTAM model obtained from OCTA, cordon flows, etc.) to perform the analysis in objectives 6 and 7. For objective 8, a preliminary prototype that links the PFE to TransCAD has been developed. However, it is not fully functional at this time. Some tasks in objectives 5 to 8 will be continued in phase two.

Note that we have successfully implemented PFE to estimate the O-D trip table for the Testbed Irvine network in Orange County, California as depicted in Figure 1-1, below. This network consists of three major freeways (I-5, I-405, and SR-133), and several arterials in the City of Irvine. In the application, the Irvine network and associated demand data were extracted from the Orange County Transportation Analysis Model (OCTAM), which contains data for the whole county. The extracted network is composed of 163 nodes, 496 links connecting 39 traffic analysis zones (TAZ), 28 external stations, and 1,547 O-D pairs. The results of these initial tests

demonstrate that PFE has the capability to correctly estimate the total demand and individual O-D demands when proper information is provided.

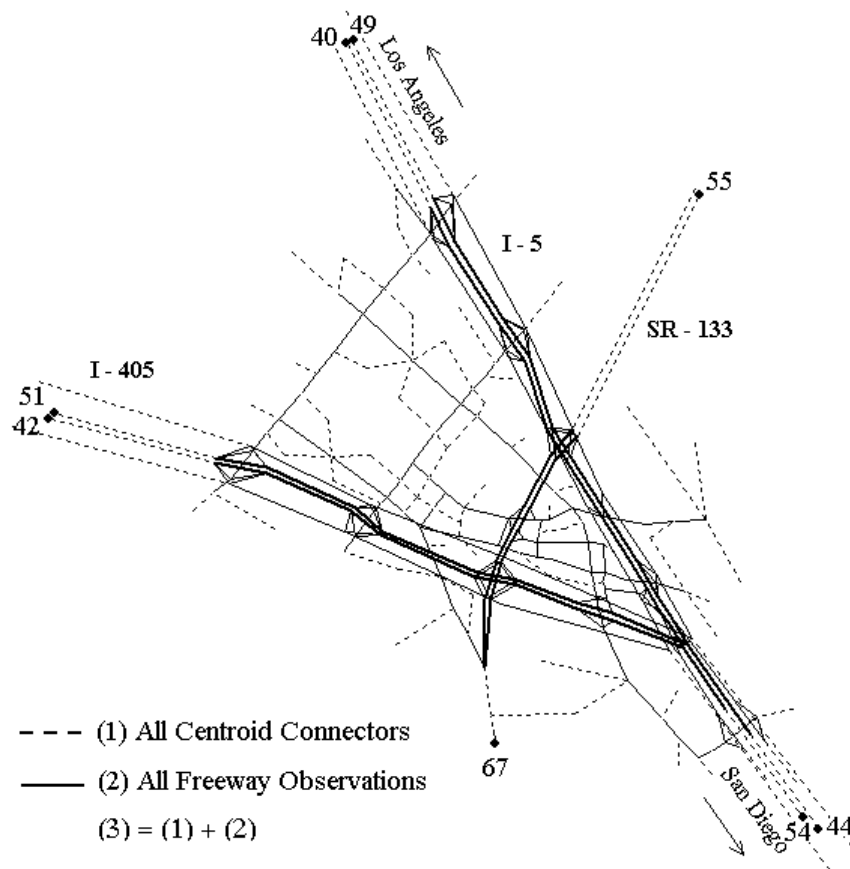


Figure 1-1 Irvine network, Orange County, California.

1.3 Organization of Report

The organization of this report is as follows. The reviews of the steady-state and time-dependent O-D estimation approaches are presented in Chapter 2. All relevant reviews including the indices of accuracy, reliability, and quality of O-D trip table estimates are also provided. Chapter 3 provides the mathematical formulation of the original PFE as well as its solution procedure. Chapter 4 examines the characteristics of O-D trip tables estimated by PFE in regards to the characteristics of input data (e.g., the number and locations of traffic counts, measurement errors, etc.). The findings from these experiments serve as the guidelines for enhancing the performance of PFE and the reliability of its estimation in the later parts. The sensitivity analysis of PFE is presented in Chapter 4. Chapter 5 details the development of the estimator based on the bi-level programming approach. The data used to conduct all analyses in this project are

described in Chapter 6. Chapter 7 provides the detailed development of the graphical user interface for PFE. The findings from research activities in the first phase of this project are addressed in the last chapter.

2 LITERATURE REVIEW

A common input required in many transportation models is an origin-destination (O-D) trip table. The O-D trip table depicts the spatial distribution of trips among traffic analysis zones (TAZ) in a transportation network. It is required for traffic assignment and utilized by many simulation models for analyzing a wide variety of transportation applications ranging from long-term strategic transportation planning to short-term transportation management on a daily basis. Traditionally, the O-D trip table is estimated through large-scale surveys such as household survey, roadside interviews, and license plate matching, etc. These survey-based techniques are expensive, time consuming, and labor intensive. In addition, if there is a rapid change in land-use patterns, the estimated trip table will soon become outdated. Therefore, the survey-based methods are often considered inappropriate due to financial and time constraints. Accordingly, the need to develop economical and quicker methods using readily available traffic counts for O-D trip table estimation has motivated researchers, during the past several decades, to devote significant efforts to this important topic.

Researches on the O-D trip table estimation can be broadly classified using two factors: network configuration and time horizon. The factors are further divided into sub-categories. For example, network configuration is separated into simple networks with no route choice (e.g., intersection, freeway segment, etc.), general networks with no congestion, and networks with both route choice and congestion. The time horizon is divided into two scales: steady-state and time-dependent cases. Using this classification scheme, we summarize the relevant research efforts on the O-D trip table estimation problem in Table 2-1.

2.1 Steady-state Trip Table Estimation

2.1.1 *Simple Network with no Route Choice*

The problem of estimating turning flows at an intersection or split ratios at a freeway segment from counts on the inflows and outflows is a special case of the O-D trip table estimation problem. Even for these simple networks, the O-D trip table estimation problem is under-specified (i.e., the number of links on which traffic flows are observed is generally less than the number of O-D pairs in the trip table). In other words, traffic counts alone are generally not sufficient to identify the true O-D trip table. Additional information, in the form of a prior trip table, is required to determine a unique solution (or a most likely solution judged by its similarity to the prior trip table). Various methods for estimating steady-state O-D trip tables (see Jeffreys and Norman, 1977; Mekky, 1979; Norman et al., 1979; Van Zuylen, 1979; Hauer et al., 1981; Bell, 1984a) were developed to increase the observability of the under-specified problem.

Jeffreys and Norman (1977) discussed some general properties (i.e., total inflow equals to total outflow, zero diagonal elements, etc.) required for the turning flow matrix to be realistic and feasible. They provided matrix manipulation schemes to generate additional feasible and realistic matrices from the initial feasible pattern. Later, most of studies in this category have been focused on the application of mathematical models for identifying the most probable

turning flow matrix of which the row and column sums respectively satisfy the known total inflow and outflow of an intersection. Mekky (1979) proposed a well-known log-linear model, which was derived as the solution to the constrained optimization problem (entropy maximization). Van Zuylen (1979) developed a similar model using the minimum information approach. Based on the work of Van Zuylen (1979), the most probable matrix is obtained by adding the minimum amount of information to the prior trip table. Prior turning probabilities can also be incorporated in these models. As reported by Van Zuylen (1979) and some follow-up studies (Hauer et al., 1981; Schaefer, 1988), turning probabilities (relative size of turning volumes) heavily affect the accuracy of turning flow estimates and have been considered as one of the critical inputs. This type of information is typically obtained based on either long-term or short-term historical traffic data. It has been reported that a higher degree of accuracy in turning flow estimates was obtained when turning probabilities were specified according to intersection and approach type (e.g., arterial or collector street) rather than the average value for each turning movement (Hauer et al., 1981). In addition, the solution based on historical data assumes that there is no substantial change in land-use or travel demand patterns. This assumption may not be applicable for long-range forecasting of turning movements (Schaefer, 1988).

2.1.2 Network with no Congestion

O-D trip table estimation methods, developed for networks with no congestion, basically assume that route choice proportions can be independently determined outside the estimation process. This independence assumption (also known as proportional assignment) is justified in the rather exceptional case where all links are measured and all link costs may therefore be estimated via link cost functions with the measured link flows (Bell and Iida, 1997). The methods for estimating steady-state O-D trip tables, using the assumption of proportional assignment, include information minimization (Van Zuylen, 1978), entropy maximization (Van Zuylen and Willumsen, 1980; Brenninger-Gothe et al., 1989; Lam and Lo, 1991), maximum likelihood (Spiess, 1987; Cascetta and Nguyen, 1988; Lo et al., 1996; Hazelton, 2000), generalized least square (Robillard, 1975; Cascetta, 1984; Hendrickson and Mcneil, 1985; Bell, 1991a), Bayesian inference (Maher, 1983), and shortest augmenting paths (Barbour and Fricker, 1996). Many of the statistical O-D estimators listed above are equivalent under certain assumptions.

Van Zuylen and Willumsen (1980) presented two closely related models for estimating the most likely trip table, which exactly reproduces traffic counts. One model is based on the principle of information minimization (Van Zuylen, 1978) and the other is based on the principle of entropy maximization. Traffic counts were assumed consistent (e.g., total inflow equals to total outflow at a node), and the internal consistency between traffic counts and assumed proportional assignment were required. The maximum likelihood model proposed by Spiess (1987), though different in principle, shares the same requirements. These requirements are rarely satisfied in real applications in which traffic counts are obtained from different sources with different levels of reliability. Later studies, on these two models, mostly relaxed the first assumption by allowing deviations between traffic counts and estimated link flows. The problem then becomes solving for the trip table that closely reproduces traffic counts when assigned onto the network according to the assumed route choice proportion. The relative beliefs in different sources of data (e.g., traffic counts, target trip table) are also allowed and incorporated into the estimation.

With the proportional assignment and assumption of inactive non-negativity constraint of O-D flows, the closed form solutions of these models were derived. Later, Bell (1991) explicitly considered the non-negativity constraints of O-D flows into the derivation of O-D trip table estimates.

2.1.3 Network with both Route Choice and Congestion

For general networks with both route choice and congestion, the proportional assignment assumption is no longer valid due to the dependency between route choice and O-D trip table (Bell and Iida, 1997). Because of this, it becomes necessary to incorporate a route choice model into the estimation process. A bi-level programming approach, by endogenously determining route choice proportion while estimating O-D trip table, is one of the possible solutions to ensure this interdependency (Fisk, 1988; Fisk, 1989; Yang et al., 1992; Maher et al., 2001). In the bi-level programming approach, the upper-level problem uses one of the statistical techniques mentioned earlier (e.g., generalized least squares) to select the most appropriate O-D trip table, whereas the lower-level problem endogenously determines route choice proportions (e.g., deterministic user equilibrium or stochastic user equilibrium) that are compatible with the estimation. Though the interdependency issue is resolved, the bi-level programming approach could pose a computational difficulty when estimating O-D trip tables for large-scale networks.

The analytical and computational difficulties of bi-level programming can be relaxed by making the following assumption. If a complete set of traffic counts constitutes an equilibrium flow pattern, the corresponding optimal trip table can then be described by an under-specified system of linear equations or a convex polytope (Yang et al., 1994). With this assumption, the bi-level program can be transformed into a single level optimization problem, which is equivalent to an under-specified system of linear equations with non-negative variables. Many researchers (Nguyen, 1977; 1984; Turnquist and Gur, 1979; Gur et al., 1980; LeBlanc and Farhangian, 1982; Sheffi and Barnhart, 1987; Fisk, 1988; Yang et al., 1994) took advantage of this assumption and formulated a single level equilibrium-based O-D trip table estimation problem. The solution to the equilibrium-based O-D trip table estimation problem, however, is not unique; there exist many possible O-D trip tables that satisfy the equilibrium link flow pattern. The multiplicity of solutions often requires a secondary optimization problem to overcome the non-uniqueness issue.

Instead of estimating O-D flows directly, Sherali et al. (1994) reformulated the O-D trip table estimation problem as a linear program to estimate path flows (hence O-D flows). This approach requires that all links are measured and assumes a deterministic user equilibrium (DUE) flow pattern. As noted above, the path flows are not uniquely defined under the DUE assignment and the requirement of the measurement of all links makes this approach less practical. Bell and Shield (1995) and Bell et al. (1997) further extended this approach to a non-linear path flow estimator (PFE) based on a stochastic user equilibrium (SUE) assumption; this gives unique flows and does not require all links to be measured.

Table 2-1 Classification of O-D Trip Table Estimation Methods

Network Configuration	Time Horizon	
	Steady State	Time Dependent
Simple networks with no route choice (e.g., intersection, freeway, etc.)	Jeffreys and Norman (1977) Mekky (1979) Norman et al. (1979) Van Zuylen (1979) Hauer et al. (1981) Bell (1984a) Schaeffer (1988)	Cremer and Keller (1981, 1984, 1987) Nihan and Davis (1987,1989) Keller and Ploss (1987) Bell (1991b) Nihan and Hamed (1992) Davis (1993) Nanne et al. (1994) Change and Wu (1994) Wu and Change (1996) Sherali et al. (1997) Lan and Davis (1999)
Networks with no congestion (i.e., proportional assignment)	Robillard (1975) Van Zuylen and Willumsen (1980) Bell (1983, 1991a) Maher (1983) Cascetta (1984) Mcneil and Hendrickson (1985) Spiess (1987) Cascetta and Nguyen (1988) Lam and Lo (1991) Barbour and Fricker (1994) Bierlaire and Toint (1995) Liu and Fricker (1996) Lo et al. (1996) Hazelton (2000)	Willumsen (1984) Okutani (1987) Ashok and Ben-Akiva (1993, 2000) Cascetta et al. (1993)
Networks with both route choice and congestion	Nguyen (1977, 1984) Turnquist and Gur (1979) Gur et al. (1980) LeBlanc and Farhangian (1982) Sheffi and Barnhart (1987) Fisk (1988, 1989) Yang et al. (1992, 1994, 2000) Sherali et al. (1994) Yang (1995) Bell and Shield (1995) Bell et al. (1997) Maher et al. (2001)	Janson and Southworth (1992) Bell et al. (1996, 1998) Sherali and Park (2001)

2.2 Time-dependent Trip Table Estimation

O-D matrix estimation models were first developed in the steady-state context in which the “average” O-D demands are estimated using the average traffic counts collected during a relatively long period. The time-varying nature of link flows and O-D demands are ignored in the static models. To overcome the limitation of static models, different dynamic O-D estimation models have been proposed. The existing dynamic approaches can be categorized into two classes: intersection-oriented and network-oriented¹. Intersection-oriented methods, which are often applied to handle isolated intersections or singular freeway segments, assume that the complete information for all the origin (entry) and destination (exit) counts are known.

2.2.1 Intersection-Oriented Dynamic O-D Estimation

The idea of “dynamic” O-D estimation originated from a series of pioneering works by Cremer and Keller (1981, 1984, 1987), where time-dependent traffic counts were first employed to transform the under-determined static model to an over-determined dynamic model. The goal of this type of dynamic model is to identify the time-varying split parameters (or turning proportions) that represent O-D demands at intersections or along small freeway segments. In the Cremer-Keller intersection model, the sequence of O-D flows and exit flows are assumed to be dependent on the time-varying patterns of entry flows through a linear relationship. Different algorithms have been proposed to solve this intersection-oriented dynamic model. Among the competing methods are the non-recursive approach, such as the cross-correlation method (Cremer and Keller, 1987), the constrained ordinary least squares method (Cremer and Keller, 1987, Sherali et al., 1997), the iterative maximum likelihood technique (Nihan and Davis, 1989), fixed point method (Nihan and Hamed, 1992), and the recursive approach, such as the method of recursive estimation (Cremer and Keller, 1987), recursive least squares (RLS) (Nihan and Davis, 1987), and the Kalman filtering procedure (Cremer and Keller, 1987, Nihan and Davis, 1987). Nihan and Davis (1987) showed that all recursive methods relate to a family of recursive prediction error (RPE) techniques and later Nihan and Davis (1989) compared RPE with the iterative maximum likelihood method.

Bell (1991) made the first attempt to extend the Cremer-Nihan dynamic O-D estimation model to general networks, in which two approaches are suggested to permit the distribution of travel times to span a number of different intervals. The first approach, which is suggested for single intersections and small networks, assumes that the fastest vehicles between any O-D pair reach the exit within one time interval, and that travel times at each exit follow geometrical distributions. Each exit flow y_j at time interval t is represented by the convex combination (weighed according to the so-called platoon dispersion factor) of y_j at $t-1$ and the sum of O-D flows headed toward the exit j at t . In the second approach, the vehicles are permitted to stay on the network for a pre-specified number of intervals, and y_j is the summation of O-D flows headed toward the exit j embarked in all these pre-specified time intervals. Both approaches adopt the constrained weighted least square method to solve sequentially for each exit. Note that Bell’s

¹ Ashok (2000) termed these as “closed networks” and “open network” respectively.

model does not really capture the time-varying nature of O-D split fractions since they are assumed to be constant over time. Thus, its essential idea is making use of time-dependent counts to iteratively refine the previous estimate of the split fractions.

Chang and Wu (1994) generalized Bell's model to consider various travel times consumed by vehicles between each O-D pair. This is more representative of a realistic environment since all vehicles do not have the same travel time. The decision parameters include not only the time-dependent O-D split parameters but also the dynamic assignment parameters. The latter set of parameters, related to the time-dependent link travel time, is introduced to define the proportion of O-D flows during the previous time interval, that arrive at any given exit during some future interval. The introduction of dynamic assignment parameters significantly increases the number of state variables to be estimated. Thus, a simplification is made by assuming that all vehicles that reach exit j during time t are distributed within intervals $t - n_{ij}^t$, where n_{ij}^t denotes the number of time intervals that the travel time between O-D pair i, j at time t will cover (If n_{ij}^t is not integer, another interval is added to remove rounding errors). To estimate dynamic travel time (hence generate proper assignment factors), the authors proposed a simple two-step method based on the simple speed-density-volume relationship. This further requires mainline traffic counts to be available. An extended Kalman filtering procedure is used to solve the nonlinear dynamic system but the satisfaction of constraints on the estimated variables is not guaranteed. The major limitations of Chang and Wu's approach are twofold. First, the travel time computation seems to be too simplistic for accuracy on congested networks. Second, the model requires not only all incoming flows, but also all mainline traffic counts to be known.

Wu and Chang (1996) further revised their earlier model to include constraints established through screenlines, which are the hypothetical cuts that intersect with a set of links and divide the network into two parts. The basic idea of the modified model comes from the assumption that most vehicles from origin i , arriving at the screenline during interval t , should embark during some intervals readily known from travel time estimation. Consequently, the fundamental measurement equation relates the screenline flows to O-D flows and time lag factors. Since one can choose as many as possible screenlines for a given urban network, sufficient constraints can be obtained for a more reliable estimation. Later, Chang and Tao (1996) extended the concept of screenlines to cordon lines (these are intended for more general networks). The measurement equation used in this model is similar to that in Chang and Wu (1994), but the dynamic assignment parameter is assumed known. Again, a Kalman filtering is used to estimate the state during each time interval and, as shown by the limited results, up to 20% improvements, over the basic scenario, have been observed when using cordon lines. The model of Chang and Tao (1996) seems problematic since it does not address how the assignment matrix might be obtained; a question that cannot be ignored in general network estimation models (which the proposed model claimed to be). Constructing cordon lines arbitrarily poses another serious feasibility issue, namely, how the entry and exit flows of all cordon lines should be observed.

2.2.2 Network-Oriented Dynamic O-D Estimation

Another group of studies, which we refer to as the network-oriented dynamic O-D estimation, focuses on extending the concept of steady-state O-D estimation by assuming the knowledge of a

dynamic assignment matrix and/or historical dynamic O-D trip tables. A dynamic assignment matrix describes the contribution of O-D demands to traffic observations at any space-time point. Mathematically it can be written as:

$$\sum_r \sum_s \sum_t q_{rs}^t p_{rs,h}^{a,t} = x_{ah}^* \quad \forall r, s, t, h, \quad (2.1)$$

where $p_{rs,h}^{a,t}$ denotes the proportion of the demand q_{rs}^t occupying link a during time interval h , and x_{ah}^* is the observation made over time h on link a .

The first known work, along this line, is contributed by Willumsen (1984), who proposed to extend the entropy maximization technique in order to handle the time-dependent traffic counts, thus ending up with a dynamic over-determined formulation. Willumsen suggested using a traffic simulation model (e.g., CONTRAM) to determine the dynamic assignment matrix. Based on Willumsen's pioneering work, four approaches have been proposed to tackle the network-oriented dynamic O-D estimation problem. The following information presents these models in detail.

Okutani's Model

To our knowledge, Okutani (1987) is the first researcher who applied the Kalman filtering technique in the network-oriented dynamic O-D estimation problem. Thus, his model can be viewed as the first approach suitable for on-line application on general networks. In this approach, the measurement equation is defined as the same fundamental equation (2.1) but a vector of random errors is introduced. The transition equation given in Okutani's model is autoregressive, taking the following form:

$$q_{rs}^{t+1} = \sum_{l=t-p}^t \sum_{r'} \sum_{s'} f_{rs}^{r's'}(t, l) q_{r's'}^l + e_{rs}^t, \quad (2.2)$$

where p is the number of lagged O-D flows assumed to affect the O-D flows in interval $t+1$ and $f_{rs}^{r's'}(t, l)$ reflects the effects of $q_{r's'}^l$ on q_{rs}^{t+1} . e_{rs}^t is a random error term. Okutani applies a standard linear Kalman filter to obtain optimal estimates for O-D flows in each time interval. Although it is of great interest for the potential on-line application, Okutani's original paper provided no explicit way to compute dynamic assignment matrices. Furthermore, as Ashok et al. (1993) argued, the use of an autoregressive transition equation (2.2) makes it impossible to "capture the complex structure of activities that result in the spatial and temporal pattern of trip making".

Cascetta et al. (1993) Model

Cascetta et al. (1993) suggested an optimization framework for estimating dynamic O-D flows, which can be viewed as an extension of classic static O-D estimators. Equation (2.1) is employed to describe the basic measurement relation and a target O-D matrix is assumed to yield a unique solution. Cascetta et al.'s model takes the following form:

$$\min d1(\mathbf{q}, \bar{\mathbf{q}}) + d2(\mathbf{x}, \mathbf{x}^*), \text{ Subject to: } \mathbf{q} \geq 0, \quad (2.3)$$

where $\bar{\mathbf{q}}$ is the target O-D flow, $\mathbf{x}$ and $\mathbf{x}^*$ represent the estimated and observed link volumes, respectively. Different distance functions will lead to estimators $\hat{\mathbf{q}}$ (the optimal solution of equation 2.3) with different statistical properties. For example, a log-likelihood function and generalized least squares (GLS) function will respectively lead to a maximum likelihood estimator and GLS estimator. Two different estimators are presented to address different applications. A simultaneous estimator designed for off-line application infers, in one step, the entire set of time-dependent O-D flows by using link traffic counts from all time intervals. On the other hand, a sequential estimator derives the O-D flows for a given time interval by using both previous O-D estimates and the current and previous traffic counts. The sequential estimator is suitable for on-line application because its computation time is relatively modest and it is able to make use of a priori estimated O-D matrices.

The authors also showed that the dynamic assignment matrix could be determined by a two-step stochastic route choice model. In the first step, fractions of time-dependent paths are given by a discrete choice model. Second, path flows are mapped onto links using a dynamic network loading (DNL) procedure, hence producing a stochastic dynamic assignment matrix. Cascetta et al's model assumes that the time-dependent link travel times through the whole network, at every time interval, are observable.

Ashok and Ben-Akiva (1993) Model

Ashok and Ben-Akiva (1993) proposed an improved version of Okutani's model. Instead of using the Okutani's autoregressive specification for O-D flows, Ashok and Ben-Akiva introduced the notion of deviation of O-D flows from target O-D flows and reformulated the transition equation as follows:

$$q_{rs}^{t+1} - \bar{q}_{rs}^{t+1} = \sum_{l=t-p}^t \sum_{r'} \sum_{s'} f_{rs}^{r's'}(t, l) (q_{r's'}^l - \bar{q}_{r's'}^l) + e_{rs}^t, \quad (2.4)$$

Obviously, by imposing the target O-D flows, the estimate is expected to avoid the risk of losing structural information about trip patterns. The determination of the assignment matrix in the Ashok and Ben-Akiva's model follows the ideas from Cascetta et al. (1993), namely, either an entirely observable network or a workable DTA model has to be available.

In a subsequent work, Ashok (1996a) indicated that the assignment matrix itself is an estimate whose random errors should not be ignored. This is because an assignment matrix is obtained from random variables such as link travel times and path fractions. It was shown that the estimator would become biased and inconsistent if this issue is not correctly addressed. In order to remedy the problem, a revised procedure is presented to take the stochastic assignment matrix into account. Treating all assignment factors as decision variables is not acceptable because it brings about an unbearable computational load. The author thus gives a simplification in which only $t+p$ state equations are added. To determine the state augmentation required to transform

this stochastic formulation to the standard transition and measurement equations, the method has to estimate each state variable as many times as the size of the lag, when applying a Kalman filter procedure. Most recently, Ashok and Ben-Akiva (2000) suggested an alternative approach defining the state-vector by deviation of departure rates from each origin and the shares headed to each destination. Except the different form of transition equations, the approach has a similar framework as those proposed previously by Ashok and Ben-Akiva (1993).

Ashok and Ben-Akiva (1993, 1996, 2000) reported encouraging results using various case studies and field tests, and indicated their method turned out to be robust. However, Sherali et al. (2001) discovered that, when attempting to implement Ashok-Ben-Akiva approach for comparison purposes, the matrix, need to be inverted when updating the Kalman gain matrix, is always singular. As Sherali et al. (2001) pointed out that the failure might be a result of a violation of some assumptions for applying Kalman filters. This observation called for special attention in checking data requirements and verifying standing assumptions when implementing Kalman filters in practice.

Sherali et al. (2001) Model

Mainly designed for off-line application, Sherali et al's model is a path-based formulation, which describes the dynamic assignment map with an equation similar to (2.1) but replacing O-D flows with path flows. Using path flows rather than O-D flows as decision variables gives rise to significant modeling merits, though at expense of the increase of decision variables. In this model, the dynamic estimation problem is formulated as a constrained least squares (CLS) problem which seeks to determine a set of time-dependent shortest path flows that reproduce the observed link counts as closely as possible. Specifically, Sherali et al's CLS estimation model reads:

$$\min \frac{1}{2} \sum_{t=1}^T \sum_{a \in A} \left[\sum_k \sum_t f_{rs}^{kt} d_{rsh}^{a,kt} - \bar{x}_{ah}^* \right]^2 + m \sum_k \sum_t c_{rs}^{kt} f_{rs}^{kt}, \quad (2.5)$$

$$\text{Subject to: } f_{rs}^{kt} \geq 0, \forall, k \in K, t, r, s. \quad (2.6)$$

Where K denotes all paths existing between all O-D pairs, f_{rs}^{kt} denote flow on path k departing at time t , and c_{rs}^{kt} the corresponding path travel cost. The second term in the objective function, a weighted total system cost, not only helps to produce a unique estimate but also “guides the solution toward more likely efficient paths.” A path generation algorithm is devised to solve the optimization problem iteratively. The algorithm begins with solving a restricted master problem based on an initial choice of a set of O-D paths, then attempts to augment the master problem with paths from time-dependent shortest path search upon a time-space expanded network. The method will terminate and claim an optimum if no new time-dependent shortest paths can be found.

Sherali et al's paper itself gives no explicit description on the construction of assignment matrices, although the authors imply that Cascetta et al's DNL method (1993) might be applicable in their model. Furthermore, this model suggested capturing network dynamics by

time-dependent link delays, yet an explicit functional form for link delays has not been described, and it is not clear how the delays can be updated according to dynamic path solutions at each iteration. Other limitations of this model include the requirement for a complete observation (which is typically not available in reality), and its inability to model the random errors and to deliver the statistical properties of estimates.

Bell's time-dependent path flow estimator (TDPFE)

Bell's time-dependent path flow estimator (1997) is distinct from the dynamic O-D estimation models reviewed so far, in the sense that it is not a strict dynamic approach. In TDPFE, trips starting in one period will always be completed within the same period unless inadequate road supplies prevent doing so. That is to say, the propagation of traffic flow and spatial and temporal evolution of congestion are simply ignored. It is believed that congestion phenomena can be captured by carrying over queued vehicles (who fail to make their ways to destinations in the current period) from one period to subsequent periods. To some extent, this approach is valid because the underlying notion of traffic congestion is in some forms of queue, and by explicitly introducing queues and associated delay, the temporary overloading of roads can be modeled.

TDPFE partitions links into two sets, the measured and unmeasured link set. The link-path incidence matrix is accordingly divided into set A for measured links and set B for unmeasured links. In TDPFE, the following nonlinear mathematical program is solved sequentially for each time period t .

$$\min \frac{1}{q} \mathbf{f}_t' (\ln \mathbf{f}_t - 1) + \mathbf{c}_t' \mathbf{f}_t + 0.5 \mathbf{u}_t' S^{-1} \mathbf{u}_t, \quad (2.7)$$

Subject to:

$$A \mathbf{f}_t = \bar{\mathbf{x}}_t, \quad (2.8)$$

$$B \mathbf{f}_t + \mathbf{u}_{t-1} \leq \mathbf{s} + \mathbf{u}_t, \quad (2.9)$$

where q is the dispersion parameter for stochastic route choice, $\mathbf{c}_t$ is a vector of path cost in period t , $\mathbf{u}_t$ is a vector equilibrium link queue in period t , and $\mathbf{s}$ is a vector of link capacity. According to the objective function, TDPFE seeks to maximize the path entropy by spreading the trips across the paths. This is done to concentrate the trips on least cost paths, and to discourage the queuing. Constraint (2.8) and (2.9) guarantees that the optimal path flow pattern would replicate observed link flows while satisfying link capacity constraints. It should be noted that link flows at each period could exceed corresponding capacities because queue is allowed to appear on links. Further, the existence of $\mathbf{u}_{t-1}$ in constraint (2.9) brings up the queue accumulated in previous periods into the current period, thus taking the dynamic congestion effects into account.

2.3 Performance Measures of O-D Trip Table Estimates

2.3.1 Statistical Measures of Estimation Error

In the literature, statistical measures are often used to quantify the quality of O-D estimates. Mean absolute error (MAE), root mean square error (RMSE), and relative error (RE) are examples of the statistical measures. They indicate the closeness between the observed (true) and estimated values, which could be link flows or O-D flows (if known). However, they may not be sufficient to quantify the quality of O-D estimates since the true O-D trip table is often unknown in practice. These measurements are given as follows.

Mean absolute error (MAE):

$$\text{MAE} = \left(\sum_{rs} |q_{rs} - \hat{q}_{rs}| / \sum_{rs} \hat{q}_{rs} \right) \times 100\%. \quad (2.10)$$

Root mean square error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{|\text{RS}|} \sum_{rs} (q_{rs} - \hat{q}_{rs})^2}. \quad (2.11)$$

Relative error (RE):

$$\text{RE} = \sqrt{\frac{1}{|\text{RS}|} \sum_{rs} \left(\frac{q_{rs} - \hat{q}_{rs}}{\hat{q}_{rs}} \right)^2} \times 100\%. \quad (2.12)$$

Total demand deviation (TDD):

$$\text{TDD} = \frac{\sum_{rs} q_{rs} - \sum_{rs} \hat{q}_{rs}}{\sum_{rs} \hat{q}_{rs}} \times 100\%. \quad (2.13)$$

where q_{rs} and $\hat{q}_{rs}$ are the estimated and true O-D trip table respectively. $|\text{RS}|$ is the number of O-D pairs in the network. It is clear that higher quality estimation of O-D trip tables is indicated by the smaller values of these measurements.

2.3.2 Maximum possible relative error (MPRE)

As the true O-D trip table is usually not available, the exact estimation errors, defined by one of the measurement errors presented earlier, is never known in practice. Yang et al. (1991) proposed the maximum possible relative error to assess the reliability of the estimated O-D trip table. To determine MPRE, it is assumed that the link use proportion is accurate and the traffic

counts are error free. Both the estimated and true O-D trip tables share one common property, which is the ability to reproduce traffic counts when assigned according to the assumed link use proportion. As mentioned by Yang et al. (1991), MPRE represents the upper bound (worst case) of the relative error of a particular O-D trip table estimate from the true but unknown O-D trip table. In other words, this measure assumes that the estimated O-D trip table happens to deviate the most from the true O-D trip table.

Let l_{rs} be the relative error between the estimated and true O-D flows of O-D pair rs .

$$l_{rs} = (\hat{q}_{rs} - q_{rs}) / q_{rs}, \quad (2.14)$$

The MPRE can be obtained by first solving the following optimization problem.

$$\text{Max} \sum_{rs} l_{rs}^2, \quad (2.15)$$

subject to:

$$\sum_{rs} \sum_k l_{rs} \cdot q_{rs} \cdot p_k^{rs} \cdot d_{ka}^{rs} = 0, \quad \forall a \in M, \quad (2.16)$$

$$l_{rs} \geq -1, \quad \forall rs \in RS. \quad (2.17)$$

Equation (2.16) presents the fact that both trip tables can exactly reproduce traffic counts. Equation (2.17) indicates the lower limit of the relative deviation, which can be verified from equation (2.14). The MPRE is then given by:

$$\text{MPRE} = \sqrt{(\sum_{rs} l_{rs}^2) / |RS|}. \quad (2.18)$$

2.3.3 Total Demand Scale (TDS)

Recently, Bierlaire (2002) proposed the use of total demand scale (TDS) to quantify the intrinsic under-determinate nature of the O-D estimation problem (e.g., the number of O-D pairs to be estimated is often much more than the number of traffic counts). The TDS measure, which is independent of the O-D estimation method, quantifies the quality of O-D estimates based on route choice proportions, network topology, and traffic counts. It can be computed by solving two linear programs.

$$\text{TDS} = f_{\max} - f_{\min}, \quad (2.19)$$

$$\text{where:} \quad f_{\max} = \text{Max}_{q_{rs}} \sum_{rs} q_{rs}, \text{ and } f_{\min} = \text{Min}_{q_{rs}} \sum_{rs} q_{rs}, \quad (2.20)$$

$$\text{subject to: } \sum_{rs} \sum_k q_{rs} p_k^{rs} d_{ka}^{rs} = \hat{v}_a, \quad \forall a \in M, \quad (2.21)$$

$$q_{rs} \geq 0, \quad \forall rs \in RS, \quad (2.22)$$

where p_k^{rs} is the proportion of travelers using path k between O-D pair rs , and $\hat{v}_a$ is the measured link flow. Using the TDS measure, the quality of O-D estimates using PFE can be quantified by the following three possibilities:

1. If the TDS is zero, it indicates that this set of observations has correctly captured the total demand of the network.
2. If the TDS is greater than zero, it indicates that, with this set of observations, the total demand of the network cannot be captured precisely but within a possible range.
3. If the TDS approaches infinity, it indicates that the demand of at least one O-D pairs is not captured by this set of traffic counts.

Possibility 1 is the most desirable since the only concern left is the determination of spatial pattern of the O-D estimates. Possibility 2 requires additional work for dealing with both uncertainty of total demand and spatial O-D distribution. Possibility 3 is the most problematic, and may often arise in practice. Under the third case, the unobserved O-D pairs are revealed by the maximization problem (e.g., $q_{rs} \Rightarrow \infty$) and additional observations are needed in order to improve the quality of O-D estimation.

3 PATH FLOW ESTIMATOR

Path Flow Estimator (PFE) was originally developed by Bell and Shield (1995) as a one-stage network observer. It is able to estimate path flows and path travel times from traffic counts obtained from detectors and other types of detection devices in transportation networks. This one-stage approach avoids the analytical and computational difficulties of the bi-level programming formulation. The core component of PFE is a logit-based path choice model, which interacts with link cost functions to produce a stochastic user equilibrium (SUE) traffic pattern. The interaction between travel times and route choices is modeled in similar ways to that of supply and demand interaction in the market place. The “market clearing” price and the quantity consumed are equivalent to the set of flows and the set of travel times equilibrated to a SUE traffic pattern (Bell and Iida, 1997). The theoretical advantage of PFE lies in the assumption of SUE, which allows the selection of non-equal travel time paths due to imperfect knowledge of network travel times. As a result, PFE gives unique path flows and does not require traffic volumes on all links to be measured.

3.1 Notation

The following symbols are used in this paper.

d_{ka}^{rs}	=	Path-link indicator: 1 if link a is on path k between origin r and destination s , and 0 otherwise
e_a	=	Measurement error, $[0, 1]$, for observation on link a
e_{rs}	=	Measurement error, $[0,1]$, for a priori O-D flow from origin r to destination s
q	=	Dispersion parameter
A	=	Set of all network links ($A = M \cup U$)
C_a	=	Capacity of link a
d_a	=	Dual variable of capacity constraint on link a
f_k^{rs}	=	Estimated flow on path k connecting origin r and destination s
K_{rs}	=	Set of paths connecting origin r and destination s
M	=	Set of measured links
o_{rs}^+	=	Dual variable of target trip table for O-D pair rs (upper limit)
o_{rs}^-	=	Dual variable of target trip table for O-D pair rs (lower limit)
q_{rs}	=	O-D flows from origin r to destination s
RS	=	Set of O-D pairs
$t_a(\cdot)$	=	Link cost function
U	=	Set of unmeasured links

u_a^+	=	Dual variable of observation constraint on link a (upper limit)
u_a^-	=	Dual variable of observation constraint on link a (lower limit)
v_a	=	Observed flow on link a
x_a	=	Estimated flow on link a
z_{rs}	=	A Priori O-D flow from origin r to destination s

3.2 Mathematical Formulation

The PFE model is based on an equivalent formulation of the logit-based SUE problem proposed by Fisk (1980) and, with the notation presented earlier, it has the following form.

$$\text{Minimize: } \frac{1}{q} \sum_{rs} \sum_k f_k^{rs} (\ln f_k^{rs} - 1) + \sum_a \int_0^{x_a} t_a(w) dw, \quad (3.1)$$

$$\text{subject to: } x_a = \sum_{rs} \sum_k f_k^{rs} d_{ka}^{rs}, \quad \forall a \in A, \quad (3.2)$$

$$(1 - e_a) \cdot v_a \leq x_a \leq (1 + e_a) \cdot v_a, \quad \forall a \in M, \quad (3.3)$$

$$x_a \leq C_a, \quad \forall a \in U, \quad (3.4)$$

$$q_{rs} = \sum_{k \in K_{rs}} f_k^{rs}, \quad \forall rs \in RS, \quad (3.5)$$

$$(1 - e_{rs}) \cdot z_{rs} \leq q_{rs} \leq (1 + e_{rs}) \cdot z_{rs}, \quad \forall rs \in RS, \quad (3.6)$$

where θ is the dispersion parameter in the logit model; f_k^{rs} is the flow on path k connecting O-D pair rs ; $t_a(\cdot)$ is the travel time on link a ; x_a is the estimated traffic volume on link a ; d_{ka}^{rs} is the path-link indicator, 1 if link a is on path k between O-D pair rs and 0 otherwise; v_a is the observed traffic volume on link a ; e_a is the measurement error allowed for traffic count on link a ; C_a is the capacity of link a ; q_{rs} is the estimated travel demand between O-D pair rs ; z_{rs} is the *a priori* travel demand for O-D pair rs ; e_{rs} is the measurement error allowed for target trip table; and M , U , A , and RS are sets of network links with measurements and without measurement, set of all network links ($A = M \cup U$), and set of O-D pairs respectively.

The objective function (3.1) has two terms, which are entropy and user equilibrium terms; the entropy term seeks to spread trips onto multiple paths, while the user equilibrium term tends to cluster trips on the minimum-cost paths. Equation (3.2) constrains the sum of path flows from all O-D pairs to produce the total flow on link a . Equation (3.3) constrains the estimated link flows within a predefined confidence level of the measured traffic counts. The confidence level essentially accounts for measurement error of traffic counts. A more reliable count will use a smaller tolerance to constrain the estimated flows within a narrower range, while a less reliable count will use a larger tolerance to allow for a larger range of estimated flows. For unobserved links, equation (3.4) constrains the estimated link flows to be less than or equal to its link

capacity. Equation (3.5) sums up the estimated path flows to obtain O-D flows. Likewise, O-D flow estimates can also be constrained by the confidence level assigned to a target trip table as in (3.6). Introducing confidence levels to equations (3.3) and (3.6) allows for a more flexible estimation of the O-D trip table.

Similar to the logit-based SUE model (this similarity is in regards to the logarithmic term), path flows can be derived analytically as a function of path cost and dual variables associated with constraints (3.3), (3.4) and (3.6), as follows:

$$f_k^{rs} = \exp \left(q \cdot \left(- \sum_a t_a(x_a) d_{ka}^{rs} + \sum_{a \in M} (u_a^- d_{ka}^{rs} + u_a^+ d_{ka}^{rs}) + \sum_{a \in U} d_a d_{ka}^{rs} + o_{rs}^+ + o_{rs}^- \right) \right), \forall k \in K_{rs}, rs \in RS. \quad (3.7)$$

For constraints (3.3) and (3.6), there are two dual variables for each constraint, the lower (u_a^-, o_{rs}^-) and upper (u_a^+, o_{rs}^+) limits. These dual variables are zero if the estimated link flows and O-D flows are within an acceptable range defined by the measurement error, and non-zero if they are binding at one of the limits. The dual variables associated with constraint (3.4) can be interpreted as queuing delay (Bell and Iida, 1997). Queuing delays exist if the estimated link flows for unmeasured links reach the available capacity, or zero otherwise. The remainder of the dual variables can be interpreted as corrections of the cost functions (e.g., link travel time or O-D travel time), which are used to match the estimated flow pattern to the observed flow pattern.

3.3 Solution Procedure

3.3.1 Iterative Balancing Technique

The solution procedure of PFE is primarily based on the iterative balancing technique. Path flows are sequentially scaled to fulfill one constraint at a time by adjusting the dual variables. The solution procedure can be summarized as follows:

Step 0: Initialization. Set $n = 0$, u_a^n, u_a^n, d_a^n , and $x_a^n = 0$ for all links. Set o_{rs}^n, o_{rs}^n , and $q_{rs}^n = 0$ and $K_{rs}^n = \emptyset$ for all O-D pairs. Set $n = n + 1$.

Step 1: Updating Link Costs. $t_a^n = t_a(x_a^{n-1}) - d_a^{n-1} - u_a^{n-1} - u_a^{n-1}$.

Step 2: Finding the Shortest Path (k_{rs}^n). If $k_{rs}^n \notin K_{rs}^{n-1}$ then update the path set and reset all dual variables and all flows (link, path, and O-D flows) to the initial state ($n = 0$); otherwise, terminate.

Step 3: Updating Dual Variables.

For all measured links $a \in M$.

Compute link cost excluding the dual variables. Compute path flows and flows on link a using Equations (3.7) and (3.2) respectively. If $x_a^n > 0$ then update the dual variables:

$$u_{a^+}^n = \text{Min} \left\{ 0, u_{a^+}^{n-1} + \frac{1}{q} \ln \left(\frac{(1 + e_a) \cdot v_a}{x_a^n} \right) \right\}, \quad (3.8)$$

$$u_{a^-}^n = \text{Max} \left\{ 0, u_{a^-}^{n-1} + \frac{1}{q} \ln \left(\frac{(1 + e_a) \cdot v_a}{x_a^n} \right) \right\}, \quad (3.9)$$

otherwise, set the dual variables

$$u_{a^-}^n = ? , \quad (3.10)$$

where η is the upper limit on the changes of dual variables (a large positive number).

For all unmeasured links $a \in U$.

Compute link cost excluding the dual variables. Compute path flows and flows on link a using Equations (3.7) and (3.2) respectively. If $x_a^n > 0$ then update the dual variables:

$$d_a^n = \text{Min} \left\{ 0, d_a^{n-1} + \frac{1}{q} \ln \left(\frac{C_a}{x_a^n} \right) \right\}. \quad (3.11)$$

For all target O-D flows.

Compute link costs excluding the dual variables. Compute path flows and O-D flows using Equations (3.7) and (3.5) respectively. Update the dual variables:

$$o_{rs^+}^n = \text{Min} \left\{ 0, o_{rs^+}^{n-1} + \frac{1}{q} \ln \left(\frac{(1 + e_{rs}) \cdot z_{rs}}{q_{rs}^n} \right) \right\}, \quad (3.12)$$

$$o_{rs^-}^n = \text{Min} \left\{ 0, o_{rs^-}^{n-1} + \frac{1}{q} \ln \left(\frac{(1 + e_{rs}) \cdot z_{rs}}{q_{rs}^n} \right) \right\}. \quad (3.13)$$

Step 4: Testing Convergence.

Determine the maximum change (x) considering all dual variables,

$$x = \text{Max} \left\{ \text{Max}_{\forall a} \left\{ \left| u_{a^+}^n - u_{a^+}^{n-1} \right|, \left| u_{a^-}^n - u_{a^-}^{n-1} \right|, \left| d_a^n - d_a^{n-1} \right| \right\}, \text{Max}_{\forall rs} \left\{ \left| o_{rs^+}^n - o_{rs^+}^{n-1} \right|, \left| o_{rs^-}^n - o_{rs^-}^{n-1} \right| \right\} \right\}, \quad (3.14)$$

Then, proceed according to one of the following criteria.

If $?_0 < x < ?^2$, set all dual variables of the next iteration ($n+1$) equal to those of the current iteration (n). Set $n = n+1$ and go to step 3.

If $x \geq ?$, set all dual variables of the next iteration ($n+1$) equal to those of the current iteration (n). Set $n = n+1$ and go to step 1, otherwise terminate.

² η_0 is a convergence tolerance (e.g., 10^{-10}) and η is the upper limit defined earlier.

The iterative balancing scheme iterates between the solutions of primal and dual problems of program (3.1) – (3.6) using equations (3.7) through (3.13) until convergence is achieved (e.g., no change in the dual variables). If the estimated flow on the unobserved link exceeds its capacity, the dual variable (queuing delay) will be scaled up until the estimated flow is less than the capacity. Similarly, if the estimated link flow or O-D flow exceeds its upper limit, its dual variable must be increased. On the other hand, the dual variable needs to be scaled down when the estimated link flow or O-D flow is below the lower limit.

3.3.2 Column Generation

For a given iteration, link costs are updated based on the new estimated flows and the dual variables. The shortest path is then determined for each O-D pair accordingly. If new paths are found, they are added to the current path set. Step 2 of the procedure presented above is known as the column generation procedure, which circumvents the need to enumerate all possible paths for general networks. In addition, the augmented link costs (including dual variables) will force the path finding to build the path set such that each observed link is on at least one of these paths. Thus, all link observation constraints are included. If the algorithm is terminated at step 2 (i.e., no new path is found), this set of traffic counts might be highly inconsistent and the chosen error bounds are not sufficient to resolve the inconsistency problem.

3.4 Examining O-D Trip Table Estimated by PFE

3.4.1 Inconsistencies of Traffic Counts

Although the original PFE does not require all links to be measured, traffic counts on all observed links must, however, be consistent. This means that the total observed traffic volume entering and exiting any intermediate node must be equal. One source of the inconsistency problem occurs when conservation of flow at intermediate nodes cannot be maintained; it then becomes impossible for PFE to estimate an O-D trip table that can duplicate all traffic counts exactly. Another source is the capacity constraint used to restrict estimated link flows on unobserved links. In practice, it is sometimes observed that the total observed traffic volumes entering a node is greater than the capacity of all exiting links combined. If traffic flows are not observed on these exiting links, their capacities are instead used to constrain the problem. Thus, there is no path flow pattern (O-D flow) that can satisfy both observation and capacity constraints at the same time. If such a situation occurs, the dual variables of link observation constraints tend to be positive or negative infinity, while those of capacity constraints tend to be positive infinity (infinite link delay). This will cause numerical problems when implementing PFE (overflow or underflow of numerical values). This difficulty is also applied when considering a target trip table (e.g., target O-D demand is inconsistent with flow entering and leaving the sink/source node).

The original PFE handles this issue by allowing imperfect duplication within acceptable ranges, through the specification of measurement errors (equations 3.3 and 3.6). However, the specification of measurement errors, associated with every single traffic count, is laborious and

thus impractical. In addition, some of the experiments (see section 3.4.3) indicate that improper selection of error bounds could lead to the underestimation of link flows and total demand utilizing the network. The heuristic procedure presented in section 3.5.2 is proposed to systematically select and readjust these error bounds.

3.4.2 Effects of Number and Locations of Traffic Counts

During the past several years, extensive studies on the O-D estimation topics have demonstrated that the quality of O-D estimates is highly dependent on the quantity and quality of available traffic data. The quality of input data is visualized, mostly, in terms of its reliability (errors inherited from the data collection). Yang et al. (1991) examined the reliability of the O-D estimates by taking into account, not only the number of traffic counts, but their locations as well. This, in a way, implies the importance of the locations of traffic counts used in the O-D estimation process. In general, it is intuitive to expect that the accuracy of O-D estimate can be increased with more traffic information. However, the same number of traffic counts may contain different amounts of information.

In a general road network, even though the traffic volumes are observed on all network links; it may not be possible to estimate a unique O-D trip table. This is due to the under-determinate nature of the O-D estimation problem in which the number of unknowns is usually much higher than the number of linearly independent equations (in the context of system of linear equations). This implies that it may not be cost-effective to collect traffic data from all network links, since we can infer the traffic volumes on the missing links if all linearly independent links are observed. Nonetheless, the number of linearly independent link counts, which is equal to the number of network links minus the number of internal nodes, is still numerous for networks of realistic size. Lam and Lo (1990) suggested a heuristic procedure based on the feedback process to identify the traffic counts that are useful to the O-D estimation. The benefit of selected traffic counts is revealed in the reduction of the estimation errors defined by the root mean squares error (RMSE). Yang et al. (1991) proposed the O-D covering rule in which the locations of traffic counts should at least intercept a part of total trips between all O-D pairs. Yang et al. (2001) further modified the set-covering formulation, mostly studied in the location theory (see Toregas and ReVelle (1973) for example), to assess the minimum number of traffic counting stations required to intercept the total flows in the network. Bianco et al. (2001) also proposed solving the O-D matrix estimation problem in two stages: (i) select the locations of affordable traffic counts and infer the missing traffic counts to obtain a full coverage, and (ii) estimate the O-D matrix using the full coverage data from the first stage.

To study the effect of locations of traffic counts, we use the variant of the set-covering formulation proposed by Yang et al. (2001) to determine the minimum number and locations of traffic counts to be used in the experiments.

$$\text{Minimize} \quad Z = \sum_{a \in A} x_a \quad (3.15)$$

$$\text{Subject to:} \quad \sum_{rs \in RS} \Phi_{rs}(p_{rs}(\mathbf{x})) = |RS|, \quad (3.16)$$

$$x_a = \begin{cases} 1, & \text{if link } a \text{ is selected} \\ 0, & \text{otherwise} \end{cases}, \forall a \in A. \quad (3.17)$$

In order to determine whether the O-D pair is intercepted or not, a shortest path algorithm is applied to a set of specially defined link travel times called *virtual* link travel times used in Yang et al. (2001). If a link is selected as a counting station, its travel time is triggered to 1; otherwise, it remains 0. The activation function (Φ) indicates that if the minimum travel time between O-D pair rs , (p_{rs}) is zero, the O-D pair is not covered by this set of counting stations. In other words, for O-D pair rs , there is at least one path that can avoid passing through these counting stations. Apparently, the travel time on the shortest path indicates the number of counting stations along the path. The summation of Φ over the set of O-D pairs (RS) indicates the number of O-D pairs being separated. For this formulation, the objective (equation 3.15) is to minimize the number of counting stations in order to separate all O-D pairs required in equation (3.16). x_a is the binary variable (0 and 1) indicating whether link a is selected as the counting station or not.

Hypothetic Network

For illustration purposes, the simple network displayed in Figure 3-1 is used. This simple network consists of 5 nodes, 7 links, and 2 O-D pairs. Table 3-1 provides the characteristics of the network links. Traffic counts, also displayed in Figure 3-1, were obtained by assuming that there are 45 travelers leaving from node 1, 20 and 25 of them are respectively destined to nodes 4 and 5, and they follow the SUE principle with a dispersion parameter of 0.10. Link cost functions are based on the standard Bureau of Public Road (BPR) function given in equation (3.18), with 0.15 for α and 4.0 for β . This cost function is used throughout this study.

$$t_a(x_a) = t_a^0 \cdot \left(1 + a \cdot \left(\frac{x_a}{C_a} \right)^b \right). \quad (3.18)$$

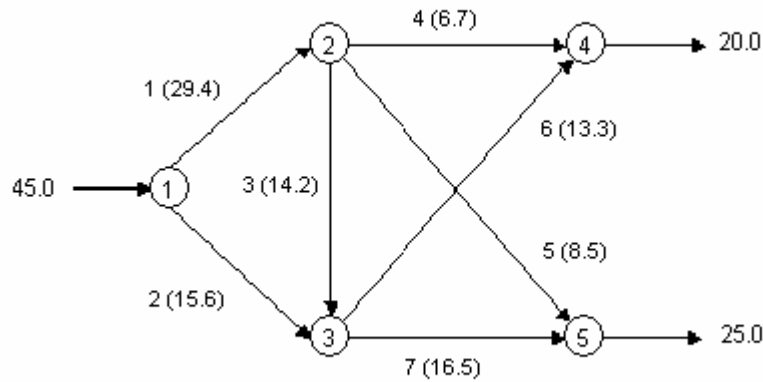


Figure 3-1 Synthetic traffic counts for hypothetical network.

Table 3-1 Link Characteristics of Hypothetic Network

Link No.	Node		Free Flow Time, t_a^0	Capacity, C_a
	To	From		
1	1	2	4.0	30.0
2	1	3	5.2	30.0
3	2	3	1.0	20.0
4	2	4	5.0	20.0
5	2	3	5.0	20.0
6	3	4	4.0	20.0
7	3	5	4.0	20.0

In order to examine the effects of the number and locations of traffic counts, we do not use the full formulation of PFE. Instead, we assume that there is no measurement error in the traffic counts (i.e., $e_a = 0$ for all links). Hence, constraint (3.3) achieves equality. Estimated link flows must match the observed traffic counts exactly. In addition, we further assume that the *a priori* trip table is not available; hence constraint (3.6) is not used. This section investigates the minimum number of link observations required for PFE to correctly capture the total demand and/or individual O-D demand (spatial distribution). It should be noted that such a requirement is dependent on the network topology and the number of O-D pairs to be estimated. Table 3-2 presents the results using different sets of observations to estimate the O-D trip table.

Table 3-2 O-D Demand Estimates with Different Sets of Traffic Counts

Number of Observations	Combination of Observed Links	O-D Demand		
		OD (1,4)	OD (1,5)	Total
7	1,2,3,4,5,6,7	20.00	25.00	45.00
1	1	15.10	15.10	30.20
1	2	8.61	8.61	17.23
2	1,2	22.50	22.50	45.00
2	1,3	15.10	15.10	30.20
2	4,6	20.00	1.21	21.21
3	1,2,4	22.07	22.93	45.00
3	1,2,6	21.00	24.00	45.00
3	1,4,6	20.00	10.75	31.75
4	1,2,4,6	20.00	25.00	45.00
4	1,2,5,7	20.00	25.00	45.00
4	4,5,6,7	20.00	25.00	45.00

When all seven observations are available, both total and individual O-D demands can be obtained correctly. This provides a benchmark for comparing with other sets of traffic counts. It should be noted that this might not be the case for other network topologies where the number of independent links in the network is less than the number of O-D pairs. In this case, unique individual O-D estimates are difficult to obtain. Thus, the estimation problem is generally solved using an optimization procedure with a distance measure as the objective in order to select the most appropriate O-D trip table that is consistent with the under-specified linear constraints (observation constraints). For this network, it is easy to verify that there are five linearly independent links (the total number of network links minus the number of internal nodes; Bell and Iida, 1997). In addition, only four of them are required to define the two, linearly independent equations for the two unknowns (O-D pairs). As we shall see in Table 3-2 that for this network, any combination of the four linearly independent observations will give the same estimate with the correct total and individual O-D demands.

For this network, using only one observation is clearly not sufficient to capture the total demand. However, each link contains different amounts of information. For example, using link 1 produces better estimates of the total demand as well as individual O-D demands compared to those produced using link 2. This suggests that link 1 contains a higher quality of information. In addition, this can be explained by considering the path set for this network, which is provided in Table 3. Since link 1 is used more (four paths) and it contains a higher flow compared to other links, it has a higher contribution to the total demand estimate. When two observations are used, only the combination of links 1 and 2 can correctly capture the total demand. It is easy to see from the network topology why other combinations (e.g., links 1 and 3 or links 4 and 6) cannot estimate the total demand correctly. Since there is only one origin (node 1) and when traffic volumes are observed on links 1 and 2, travel demand originated from node 1 is totally captured. As can be seen from Table 3-3, with these two observations (links 1 and 2), traffic flows on all paths are also observed. However, the spatial distribution of the total demand is not correct. Similar results are also observed using three observations. To obtain the correct individual O-D demands, PFE requires at least 4 observations. There are three combinations that can achieve this as reported in Table 3-2. Since the network contains only one origin (node 1) and two destinations (node 4 and node 5), it is clear that if all entry flows to both destinations (links 4, 5, 6, and 7) or the total demand (links 1 and 2) with all entry flows to one of the destinations (e.g., links 4 and 6 or links 5 and 7) are observed, both total and individual O-D demands would be estimated correctly.

Table 3-3 Path Set of Hypothetic Network

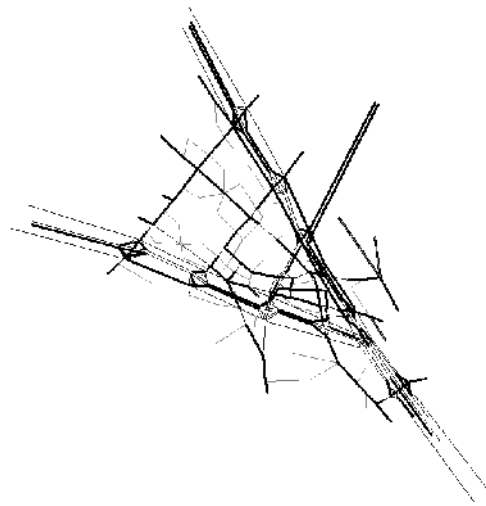
O-D Pair	Path	Link Sequence
(1,4)	1	1-4
	2	1-3-6
	3	2-6
(1,5)	1	1-5
	2	1-3-7
	3	2-7

Irvine Network

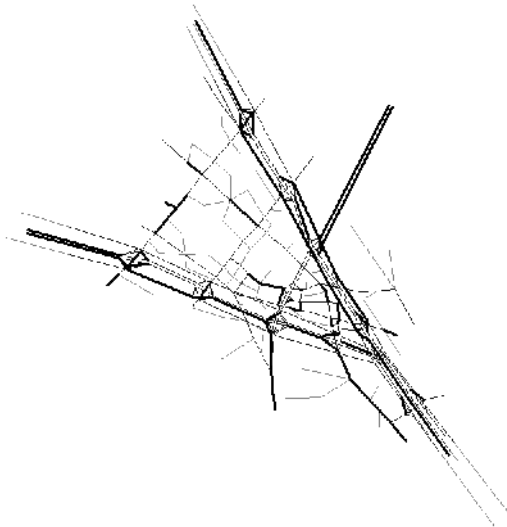
Using the findings in the previous section, one can expect that the total demand of a network should be correctly estimated when traffic counts are collected on all links that are directly connected to all origins and destinations (centroid connectors). However, since centroid connectors do not physically exist (hence traffic counts), the set-covering formulation presented earlier comes to play a role selecting traffic counts for general road networks. Here, PFE is performed using three different sets (A, B, and C) of traffic counts. The first two sets contain the synthesized traffic counts³ in which the number and locations are determined to cover all O-D pairs and only freeway O-D respectively. The locations of traffic counts required for sets A and B are depicted in Figure 3-2(a) and 3-2(b). The third set of traffic counts, depicted in Figure 3-2(c), contains the same number of observations as set A does, but they are randomly selected. For the fourth set (set D), a priori trip interchanges for the freeway O-D pairs (17 O-D pairs) are also used in addition to the traffic counts of set A. In practice, this type of information recently becomes available through the advancement in traffic surveillance and electronic toll/fare collection and can be used to increase the accuracy and reliability of O-D trip table estimates. As a result, there are 186 observations (38% of network links) for sets A and C, 99 observations (20% of network links) for set B, and 186 observations plus 17 prior O-D demand for set D. In this section, we further assume that these traffic counts contain no measurement errors.

The estimation results are summarized in Table 3-4 in terms of total demand estimates, estimation errors, RMSE for target O-D demands, and the TDS measure presented in Chapter 2. It should be noted that link counts (constraints) could be reproduced exactly by PFE; the corresponding RMSE is zero. In addition, as reported in Table 3-4, the column generation procedure used in the PFE is capable of generating sufficient number of paths (close to the number of paths generated by the SUE principle) required by PFE to match the traffic counts. The O-D estimates using the traffic count sets A and C, emphasize the importance of the locations of the traffic counts on the overall quality of O-D trip table estimates. Although both sets A and C contain the same amount of traffic counts, the qualities (i.e., information contained in the data set) are significantly different. As expected, PFE can closely capture the total demand of this network when matching the traffic counts in set A, while it underestimates the demand when matching traffic counts in set C.

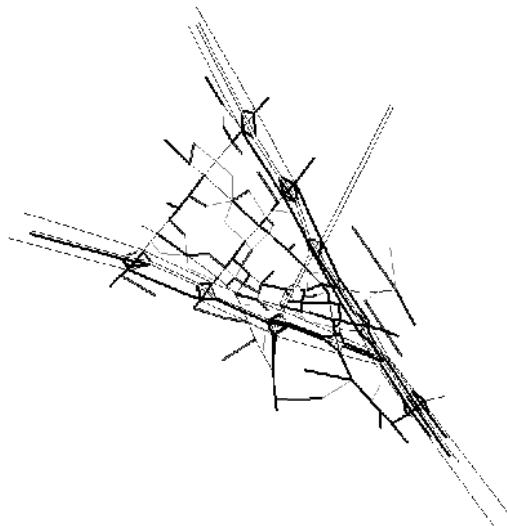
³ Traffic counts were generated by loading O-D trip table extracted from OCTAM model to the network according to Logit model with $\theta = 0.01$. (See Chapter 5 for SUE traffic assignment models and Chapter 6 for detailed descriptions of OCTAM model)



a) Observation Set A



b) Observation Set B



c) Observation Set C

Figure 3-2 Locations of different sets of traffic counts for Irvine network.

Table 3-4 Summary of Estimation Results for Irvine Network

Description	OCTAM	Observation			
		Set A	Set B	Set C	Set D
No. of Link Observations	-	186	99	186	186(+17 ^a)
No. of Paths Generated	2,011 ^b	2,004	2,062	2,029	2,107
Total Demand	47,522.00	47,522.00	46,631.32	43,410.99	47,522.00
(% Error)	-	0.00	-1.87	-8.65	0.00
RMSE – Link Flow Estimates	-	0.00	0.00	0.00	0.00
RMSE – O-D Flow Estimates	-	47.34	46.88	61.21	36.24
No. of O-D pairs not covered	-	0	149	54	0
$\phi_{\max}$	-	47,522.00	∞	∞	47,522.00
$\phi_{\min}$	-	47,522.00	43,067.17	41,401.02	47,522.00
$\phi_{\max} - \phi_{\min}$	-	0.00	∞	∞	0.00

a - number of prior trip interchanges

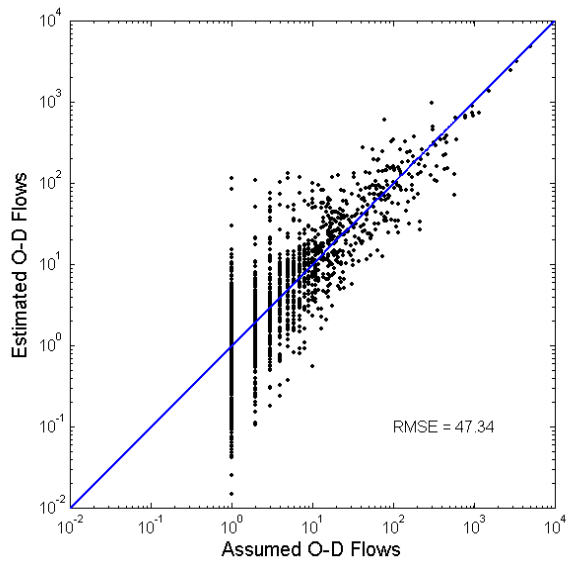
b - number of paths obtained when assigning the assumed O-D trip table according to the SUE principle

When traffic counts in set B are used, the total demand of the network is certainly underestimated; a certain portion of the demand between other O-D pairs (internal-to-external demand, external-to-internal demand, or internal-to-internal demand) is not fully captured. However, since the freeway-O-Ds always involve high traffic volumes, the total demand estimated by this set of traffic counts is quite close to the true total demand.

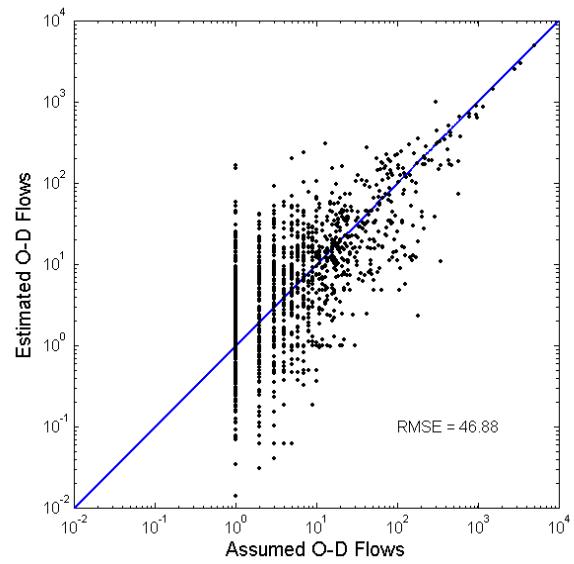
In Table 3-5, the estimation results for the freeway O-D pairs are reported together with their estimation errors. We do not report the results for observation set D because the O-D estimates are exactly the same as the target trip interchanges provided in the second column. In case B, even though the number of observations is less than that of case A, the spatial distribution of the freeway O-D pairs appears much better. For example, the maximum absolute percentage error reduces from 115.16 to 9.63 for O-D pair (42,67). This is because the information provided by observation set B is better for estimating freeway O-D demands; traffic flows on all freeways are completely observed and they usually involve high traffic volumes. The comparison of the true and estimated O-D flows is depicted in Figures 3-3(a) through 3-3(d) for each set of observations (in log scale), where each point represents an O-D pair comparison and the data points along the 45-degree line represent perfect matches. Figures 3-4(a) through 3-4(d) provide similar comparisons, but focus only on the freeway O-Ds. From Figures 3-3, it is observed that PFE can estimate high O-D demands quite well, but it has problem in estimating low O-D demands. The estimates for O-D pairs with low demand could deviate substantially from the true values. It is worth pointing out that the RMSE only represents the aggregated quality of O-D estimates (on average). As displayed, the data points in Figure 3-3(b) appear more scattered than those in Figure 3-3(a); thus the overall quality of observation set A should be better. However, the resultant RMSE indicates that the O-D estimates obtained from observation set B is slightly better than those obtained from observation set A. This is because the RMSE measure penalizes more heavily the O-D pairs with higher demands.

Table 3-5 Estimation Results for Freeway O-D pairs of Irvine Network

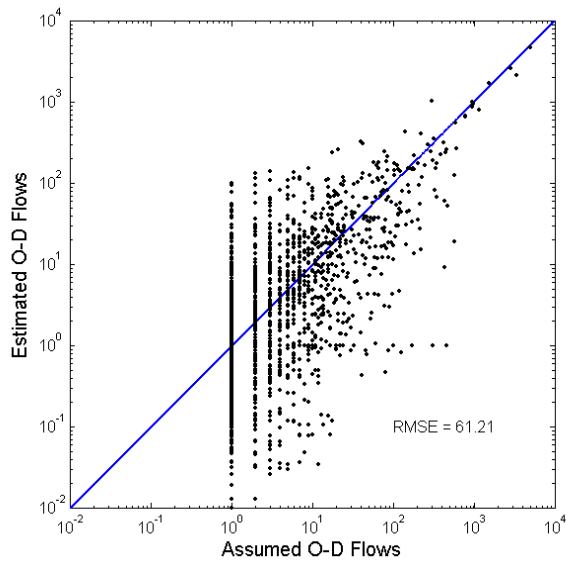
O-D Pairs	OCTAM	Set A		Set B		Set C		Remark
		PFE	% Error	PFE	% Error	PFE	% Error	
(40,54)	5,084.00	4,836.60	-4.87	4,899.35	-3.63	4,706.07	-7.43	On I-5 S to San Diego
(40,55)	780.00	676.60	-13.26	653.07	-16.27	676.94	-13.21	From I-5 S to SR-133
(40,67)	419.00	217.95	-47.98	262.70	-37.30	320.14	-23.59	From I-5 S to SR-133
(42,54)	2,829.00	2,443.49	-13.63	2,498.48	-11.68	2,624.19	-7.24	From I-405 S to San Diego on I-5 S
(42,55)	946.00	882.85	-6.68	887.50	-6.18	989.57	4.61	From I-405 S to SR-133
(42,67)	122.00	262.49	115.16	133.74	9.63	140.52	15.18	From I-405 S to SR-133
(44,49)	3,392.00	3,139.08	-7.46	3,010.33	-11.25	2,144.15	-36.79	On I-5 N to Los Angeles
(44,51)	1,553.00	1,353.28	-12.86	1,427.76	-8.06	1,719.24	10.70	From I-5 N to I-405 N
(44,55)	973.00	676.58	-30.46	631.46	-35.10	940.87	-3.30	From I-5 N to SR-133



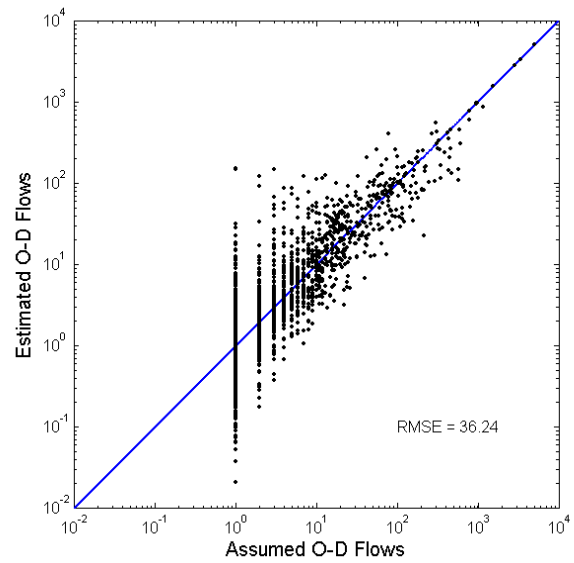
a) Observation Set A



b) Observation Set B

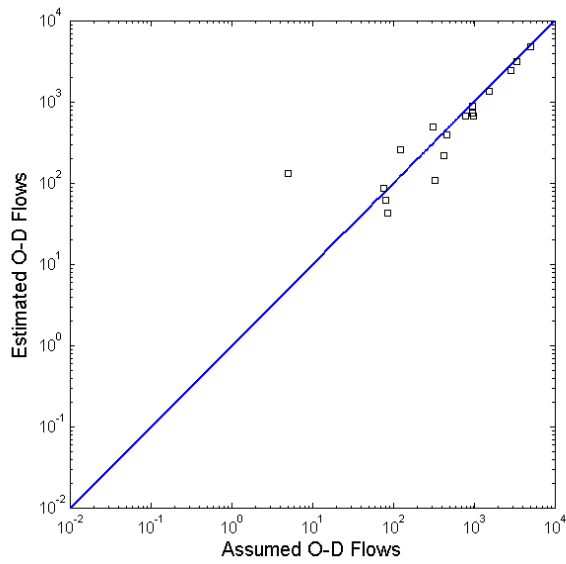


c) Observation Set C

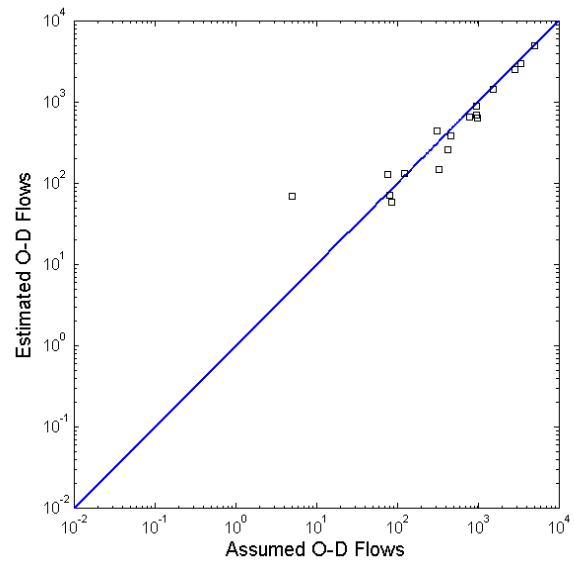


d) Observation Set D

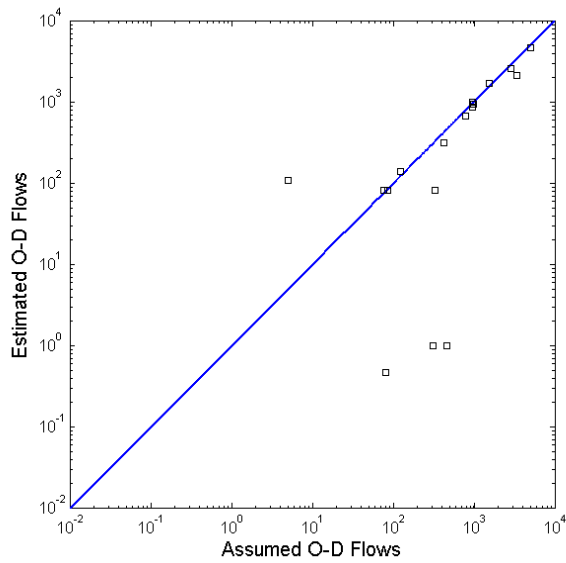
Figure 3-3 Comparison of true and estimated O-D demands with different sets of traffic counts.



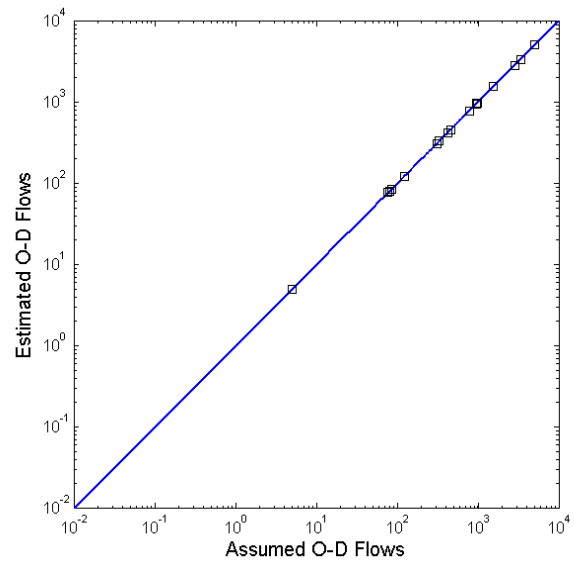
a) Observation Set A



b) Observation Set B



c) Observation Set C



d) Observation Set D

Figure 3-4 Comparison of true and estimated freeway O-D using different sets of traffic counts.

3.4.3 Effect of Measurement Errors of Traffic Counts

Although complete traffic counts can be identified and collected using the method discussed earlier, they always associate with the measurement errors. This is due to many factors such as the malfunction of detection instruments in data collection or the unseen error in data processing. In addition to its impact on the accuracy of O-D estimates, these measurement errors also create inconsistencies among observations; another issue to be considered when using the PFE and other O-D estimation methods with a similar structure. If inconsistency among observations exists at certain nodes (the total inflow does not equal the total outflow), it is difficult for PFE to match all inconsistent observations unless such a definition of matching is relaxed through the pre-specified error bounds (Equations 3.3 and 3.6).

To specify a proper error bound, the assistance from experienced traffic engineers is required. In practice, this task is quite time-consuming especially for networks of a realistic size. It is possible to set a uniform error bound across all observations, but the resultant error bounds of some observations may be too loose. This frequently leads to the underestimation of travel demand utilizing the network, as we shall see from the numerical experiment. Instead of using identical error bounds for all observations, individual error bounds can be automatically readjusted along with the iterative balancing procedure.

As proposed in this study, the information derived from the PFE (e.g., dual variables) can be used to guide this adjustment process. When PFE has a difficulty to match some observations or inconsistencies among observations exist, the dual variables associated with the particular observation constraints tend to be either positive or negative infinity. In a way, this information informs us of the necessity to relax the matching requirements of such observations. Since there is no general rule in selecting links to adjust the error bounds, the following provides a heuristic to assist the adjustment procedure.

Step 1. Set $i = 1$ and perform the simulation of PFE as observations are error free, $e_a^i = 0, \forall a \in M$.

Step 2. If the PFE converges, then terminate; otherwise go to step 3.

Step 3. Rank the observations according to the absolute values of their dual variables in the ascending order, set p to the absolute value of dual variable at α -percentile, and update error bounds for all observed links as follows.

i. If $|x_a - v_a|/v_a > e_a^i$, and $\text{Max}\left\{|u_a^+|, |u_a^-|\right\} \geq p$, then

$$e_a^{i+1} = \text{Min}\left\{i \cdot e_{\min}, \text{roundup}\left(|x_a - v_a|/v_a\right)\right\},$$

ii. If $|x_a - v_a|/v_a > e_a^i$, and $\text{Max}\left\{|u_a^+|, |u_a^-|\right\} < p$, then

$$e_a^{i+1} = \text{Min} \left\{ i \cdot e_{\min}, \text{round} \left(|x_a - v_a| / v_a \right) \right\},$$

iii. Otherwise $e_a^{i+1} = e_a^i$.

Step 4. Set $i = i + 1$, perform the simulation of PFE using the adjusted error bounds, and go to step 2.

$e_{\min}$ is the maximum adjustment allowed at each iteration. The error bounds of each observation will be expanded, upon necessity, with a proper adjustment factor based on the current information. It is necessary to set an upper limit ($e_{\min}$) for the adjustment in order to prevent the over-relaxation of the error bound at the earlier iterations. It is reasonable to believe that if the adjustments allowed, at each iteration, are fairly small ($e_{\min} \Rightarrow 0$), the link flow estimates will have the smallest deviation from the observations. It is worth noting that if $e_{\min}$ is chosen too small, the convergence rate of the algorithm could be very slow. On the other hand, the algorithm is expected to converge faster with a larger adjustment factor, but the error bounds of some observations may be too loose and may affect the accuracy of the link flow estimates. Hence, it is crucial to choose a proper adjustment factor to balance the speed of convergence and the quality of estimates.

In this section, we investigate the effect of the measurement errors on the quality of O-D estimates and demonstrate the potential of the proposed heuristic used to facilitate the selection of error bounds. To create some inconsistencies in the traffic counts, we assume that the variation of traffic counts follows a normal distribution with a mean of zero and a standard deviation equal to one-third of the actual flows used in the previous section. Then 100 samples of traffic counts were generated based on the assumed probability distribution and the average traffic volumes are used as the observations. The inconsistent traffic counts (varied between -9% and $+9\%$) on the exact same locations as observation set A are used here. The results are presented for two cases in which the error bounds for link measurements are uniformly set across all observations and they are adjusted using the proposed heuristic. The results as well as all associated statistics, namely the RMSE of O-D flow estimates, the number of iterations required for the iterative balancing to converge, the number of paths found, the objective value, and the total demand estimates, are summarized in Table 3-6.

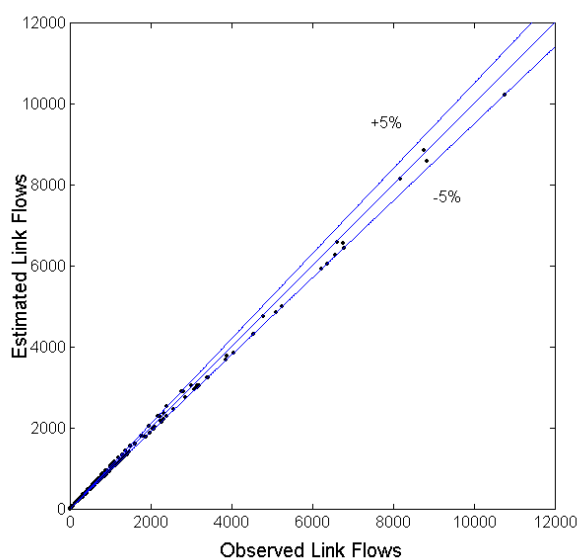
For the case of uniform error bounds, we incrementally increase the error bounds for all links until the PFE converges. Convergence is found when the error bounds are set larger-than-or-equal-to 5 percent. Interestingly, the over-relaxations of some observation constraints are able to compensate the measurement errors of some other observations, since the maximal possible errors introduced to the data set is indeed $\pm 9\%$. In other words, the major errors inherited with some observations are transferred to observations with minor or no error. Using $e_{\min}$ of 0.02 and α of 85 in the heuristic procedure, PFE is performed three times with a total of 48,712 iterations and the maximum error bound selected by this process is 6 percent, which is also less than the maximum error introduced into the data set. Figures 3-5(a) through 3-5(d) present the comparisons between estimated and observed link flows when the error bounds are uniformly set at 5 percent, 7 percent, 10 percent, and by the heuristic respectively.

Table 3-6 Effect of Error Bounds Selection on the Estimation Results

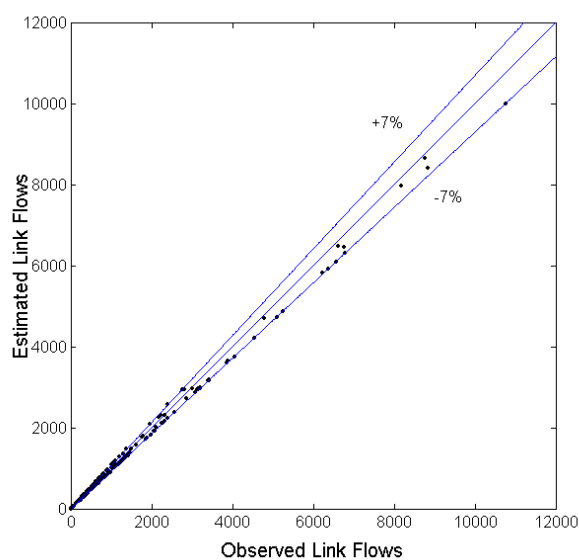
Error Bound	Converge?	RMSE-OD	PFE Obj.	# Paths	# Iterations	Total Demand ^a
0%	No	-	-	-	20,000	-
3%	No	-	-	-	20,000	-
5%	Yes	46.29	18,598,521	2,053	1,743	45,269
7%	Yes	49.74	17,919,947	2,049	1,493	44,223
10%	Yes	55.19	16,988,351	2,038	1,177	42,785
Heuristic	Yes	43.36	19,875,884	2,071	48,712	47,118

a: Total demand from OCTAM model is 47,522.

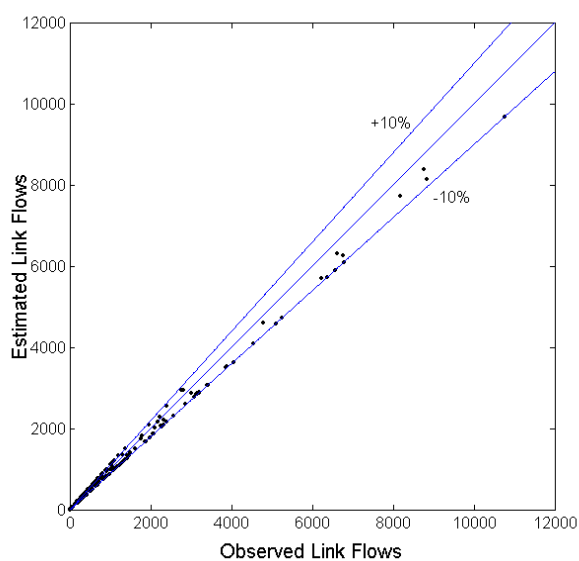
Figures 3-5(a) through 3-5(d) indicate that the estimates in all cases are acceptable since PFE produces link flow estimates within the pre-specified ranges. However, it can be observed that the link flow estimates are frequently biased toward the lower bounds of acceptable ranges. This implies that link flow estimates, as well as the corresponding total demands (see also the values in Table 3-6), are likely to be underestimated when the error bounds are unnecessarily large (> 5 percent). This characteristic can be explained in accordance with the objective function of PFE. As long as all observation constraints are satisfied, PFE is likely to go for the lower path flow estimates, which produce a lower total cost (UE objective) and entropy value. When the error bounds are properly adjusted, by taking the dual variables into account, such underestimation is partly preventable. On the other hand, the link flow estimates are least deviated from the observations as shown in Figure 3-5(d); a large portion of points representing a pair of estimated and observed link flows sit almost on the 45-degree line. Moreover, the corresponding total demand estimate is closer to the known value.



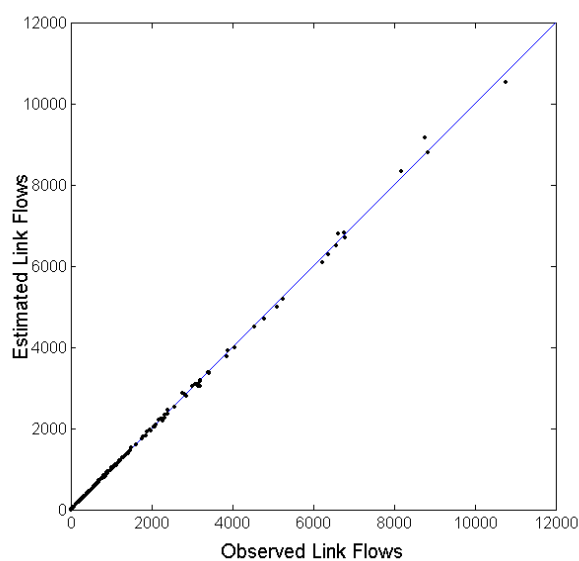
a) 5 % of uniform error bound



b) 7% of uniform error bound



c) 10% of uniform error bound



d) Non-uniform error bound

Figure 3-5 Comparison of observed link flows and estimated link flows at different selected error bounds.

3.5 Derivation of Intersection Turning Movements by PFE

Because a unique set of path flows is readily available from the PFE, it is possible to trace these paths to derive other useful information at different spatial levels. For example, the sum of all path flows from all O-D pairs gives the total demands utilizing the network, the sum of all path flows emanating from a given origin gives a total trip production, and the sum of path flows terminating at a given destination gives a total trip attraction. Flows between an O-D pair can be obtained by simply adding up the flows on all paths connecting that O-D pair. The aggregated link flows are obtained by adding up all path flows passing through a given link. For the turning movements at an intersection, the orientations of links connected to intersection are needed so as to determine the individual turning movement (e.g., left, right, or through movement) from the used paths without a need to expand the network (for representing turning movements).

In order to derive turning movements at the intersection, from path flows, orientations of links connected to the intersection are needed. The orientations of links are defined as in Figure 3-6.

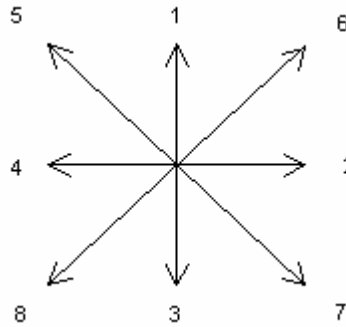


Figure 3-6 Convention of link orientation.

The orientations of links along each path are used to determine the turning movement (e.g., left, right, or through movement). For example, if a particular path precedes the intersection in direction 1 and leaves the intersection in direction 2, the flow on this path is contributing to the amount of flows on northbound right-turn (NBR). According to the convention of the link orientation given below, the flow on turning movements constituted by the leading direction g^+ and the leaving direction g^- at intersection i , t_{g^+,g^-}^i , can be computed as follows:

- Step 0.* Set $t_{g^+,g^-}^i = 0, \forall i, g^+, g^-$, and $rs = 1$.
- Step 1.* Set $k = 1$.
- Step 2.* Set $\ell = 1$.
- Step 3.* For path k , if the ending node of link $a_{k\ell}$ is intersection i , then set $g^+ = \text{orien}(a_{k\ell}), g^- = \text{orien}(a_{k(\ell+1)})$, and $t_{g^+,g^-}^i = t_{g^+,g^-}^i + f_k^{rs}$.
- Step 4.* If $\ell < |A_k|$, set $\ell = \ell + 1$ and go to step 3, otherwise go to step 5.
- Step 5.* If $k < |K_{rs}|$, set $k = k + 1$ and go to step 2, otherwise go to step 6.
- Step 6.* If $rs < |RS|$, set $rs = rs + 1$ and go to step 1, otherwise terminate.

where

- I : Set of intersection nodes
- G_i : Set of movements or directional movements at intersection i
- A_k : Set of links along path k
- f_k^{rs} : Flows on path k connecting origin r and destination s
- t_{g^+,g^-}^i : Flows on movement constituted by leading direction g^+ and leaving direction g^- at intersection i
- $\text{orien}(\cdot)$: Orientation of link

By incorporating the above procedure into PFE, flows on each turning movement, at each intersection can be obtained as part of the estimation results (together with path flows and O-D flows).

3.5.1 A Simple Isolated Intersection

For the standard four-leg intersection as depicted in Figure 3-7, there are three movements for each approach (NB, SB, EB and WB), with a total of twelve movements. Traffic counts are also provided in Figure 3-7. For this network, the turning movements can be treated as a small O-D trip table without route choice. For example, traffic flows on the eastbound right-turn, which is 157 vehicles, are essentially the travel demand from node 1 to node 3. The total demand for this intersection is 5,894 vehicles. The link characteristics of this intersection are provided in Table 3-7.

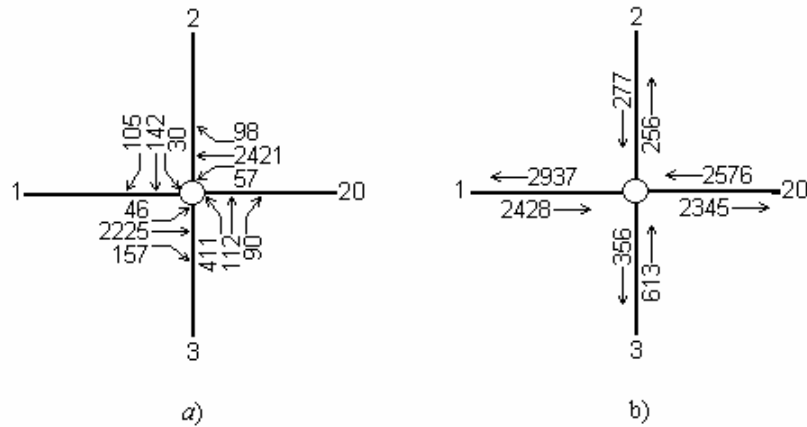


Figure 3-7 Intersection layout, a) turning-movement counts, and b) entry and exit flows.

Table 3-7 Link Characteristics of an Isolated Intersection

Node		Orientation	Capacity (vph)	Speed (mph)	Distance (mile)	Traffic Count (vph)
From	To					
1	19	2	3,600	35	0.1313	2,428
2	19	3	1,800	25	0.1347	277
3	19	1	1,800	25	0.1313	613
20	19	4	3,600	35	0.1237	2,576
19	1	4	3,600	35	0.1313	2,937
19	2	1	1,800	25	0.1347	256
19	3	3	1,800	25	0.1330	356
19	20	2	3,600	35	0.1237	2,345

Since delay on each movement is different, the probability of selecting each movement should not be equal. In order to account for this non-uniformity, the delay of each movement is assumed to be different but constant according to Table 3-8. The delays for prohibited movements (e.g., direction 1 and then direction 3) are set to a very large number in order to prevent such illegal movements. These movement delays are also included in the path travel time calculation. The dispersion parameter (θ) for this network is assumed to be 30 hr^{-1} , which is the inverse of the average O-D travel time.

Table 3-8 Intersection Turning Movement Delays (Turn Penalties)

Approach	Delay (hour)			
	1 (NB)	2 (EB)	3 (SB)	4 (WB)
1 (NB)	0.024	0.024	∞	0.017
2 (EB)	0.020	0.030	0.030	∞
3 (SB)	∞	0.017	0.024	0.024
4 (WB)	0.030	∞	0.020	0.030

Two sets of traffic counts, (1) entry flows only and (2) entry and exit flows, are used to examine the quality of information contained in traffic counts to the accuracy of turning movement estimates. It is assumed that the observations were obtained without measurement error and, they are consistent. Furthermore, the true turning movement volumes are as observed. Accuracy of the estimates, therefore, can be defined by the root mean square error (RMSE). PFE aims to reproduce the observed flows for links with measurement and to produce less-than-capacity flows for links without measurement. The estimation results using the first and second sets of traffic counts are respectively presented in Tables 3-9 and 3-10 with their corresponding RMSEs.

Table 3-9 Estimation Results using Entry Flows only ($RMSE_{OD} = 1,314.04$)

From/To	1	2	3	20	Total
1	-	915	709	805	2,428
2	89	-	79	110	277
3	245	168	-	200	613
20	846	717	1,013	-	2,576
Total	1,180	1,800	1,800	1,114	5,894

Table 3-10 Estimation Results using Entry and Exit Flows ($RMSE_{OD} = 160.95$)

From/To	1	2	3	20	Total
1		171	194	2,063	2,428
2	164		8	105	277
3	425	11		177	613
20	2,348	74	154		2,576
Total	2,937	256	356	2,345	5,894

The results are presented in an O-D trip table format, which indicates the amount of travel demands from one node to the others. Considering the first set of traffic counts, PFE can match the observations as well as the total demand of this network perfectly. However, due to insufficient information regarding the exit flows, the best estimates on exit links are constrained by their capacities. Besides, the amount of observed flows on the entry links seems to be uniformly distributed among all movements (due to the nature of objective function, maximum entropy). As can be seen, the estimated and true O-D flows (turning movements) are rather different as indicated by a RMSE of 1,314.04.

By including the exit flows into the estimation, PFE is able to match both entry and exit flows. The quality of the O-D trip table is generally improved as indicated by the reduction in RMSE from 1,314.04 to 160.95. However, the true spatial pattern of the O-D trip table cannot be captured due to the under-determinate nature of the problem (i.e., the number of unknowns is less than the number of linearly independent equations). Although, the entropy function is used as the objective function to select the most probable flow pattern, it is apparent that the true O-D matrix is seldom the most probable one. In such a situation, a better turning movement delay representation, at the intersection, may help us to obtain a better estimate.

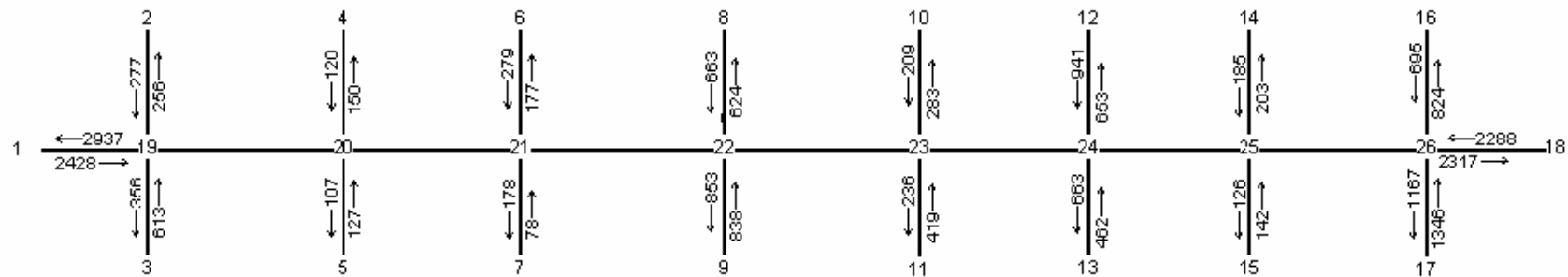
3.5.2 *A Signalized Arterial*

This example considers an arterial street connected by a series of signalized intersections. The layout of this arterial and the traffic counts are provided in Figure 3-8. Traffic counts were collected during the evening peak hour. Data preprocessing was required to remove errors and inconsistencies in order to avoid getting erroneous results. This network involves 8 signalized intersections, 50 links, 18 external stations, and 306 O-D pairs. The total demand of this network is 12,110 vehicles. The dispersion parameter of this network is assumed to be 8 hr^{-1} . In addition, it is assumed that travel delay patterns at all intersections are identical to the one presented in the previous example. Traffic counts on links originating from and destined to the external stations, 36 out of 50 links, are used as the observations. The O-D demand estimates are presented in Table 3-11.

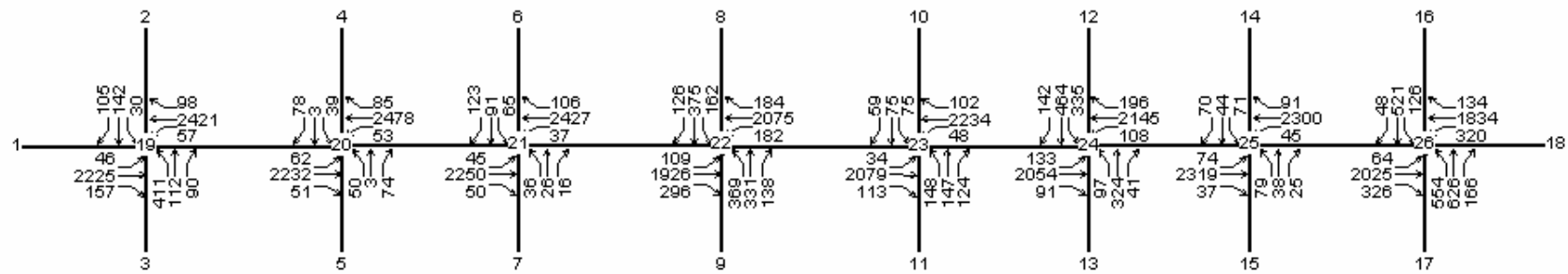
As can be seen from Table 3-11, the total demand of this network can be estimated correctly since all entry and exit flows are provided. Similar to the previous example, the observations on entry and exit links represent trip productions at origins and trip attractions at destinations in the sense of a trip distribution procedure. These numbers are exactly matched as well since they are used as observations. Because of the unavailability of the true O-D trip table, we do not report the RMSE for the O-D estimates; instead we report the RMSE for the estimated turning movements. The estimated turning movements are depicted in Figure 3-9.

For this case study, it is found that the turning movement estimates on the major street are fairly acceptable while those on the minor streets are not, especially for the through movement. In general, the through movement estimates on the minor streets are underestimated while their left- and right-turn counterparts are overestimated. This is due to the fact that there is only one path traversing the through movement on any minor street (e.g., path from node 3 to node 2 in Figure 3-8) while there is at least one path traversing the other movements (from minor street to major

street). In addition, since the main traffic stream is on the major street, the contribution of flows from one minor street (through movement) to match the observation on the opposite minor street is always dominated by the flows from the major street especially when the turn penalties at intersection are likely to be mis-specified. The comparison of the observed and estimated turning movements is shown in Figure 3-10. Each point represents the volume of an estimated-observed turning movement pair. The data points along the 45-degree line represent a perfect match. Although PFE can reproduce link traffic counts perfectly, the spatial distribution of turning movements (also pattern of O-D demands) is difficult to capture with the limited amount of information provided.



a)



b)

Figure 3-8 Layout of Arterial Network, a) entry and exit flow observations, and b) turning movement counts.

Table 3-11 O-D Trip Table Estimated using Observations on Entry and Exit Links

From/To	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Total
1		178	243	87	59	83	76	238	308	79	62	152	133	33	18	104	148	426	2,428
2	196		8	3	2	3	3	9	11	3	2	5	5	1	1	4	5	15	277
3	451	12		6	4	6	6	18	23	6	5	11	10	2	1	8	11	32	613
4	71	2	3		1	2	2	5	7	2	1	3	3	1	0	2	3	10	120
5	78	2	4	2		2	2	5	7	2	1	3	3	1	0	2	3	9	127
6	133	4	6	3	3		5	17	23	6	5	11	10	2	1	8	11	31	279
7	39	1	2	1	1	1		5	6	2	1	3	3	1	0	2	3	8	78
8	260	7	12	6	5	10	11		71	19	15	37	32	8	4	25	36	103	663
9	357	10	16	9	7	14	15	65		24	19	45	40	10	5	31	44	127	838
10	61	2	3	1	1	2	3	12	18		6	15	13	3	2	10	15	42	209
11	129	4	6	3	2	5	6	25	38	13		28	25	6	3	19	27	79	419
12	221	6	10	5	4	9	10	43	65	24	23		77	20	11	63	90	259	941
13	114	3	5	3	2	5	5	22	34	13	12	39		9	5	29	42	120	462
14	31	1	1	1	1	1	1	6	9	3	3	11	12		3	15	22	63	185
15	25	1	1	1	0	1	1	5	7	3	3	9	9	3		11	16	45	142
16	93	3	4	2	2	4	4	18	27	10	10	34	35	13	8		106	322	695
17	202	6	9	5	4	8	9	39	59	22	21	74	76	27	18	141		626	1,346
18	473	14	22	11	9	19	20	92	139	52	48	172	178	64	42	349	584		2,288
Total	2,937	256	356	150	107	177	178	624	853	283	236	653	663	203	126	824	1,167	2,317	12,110

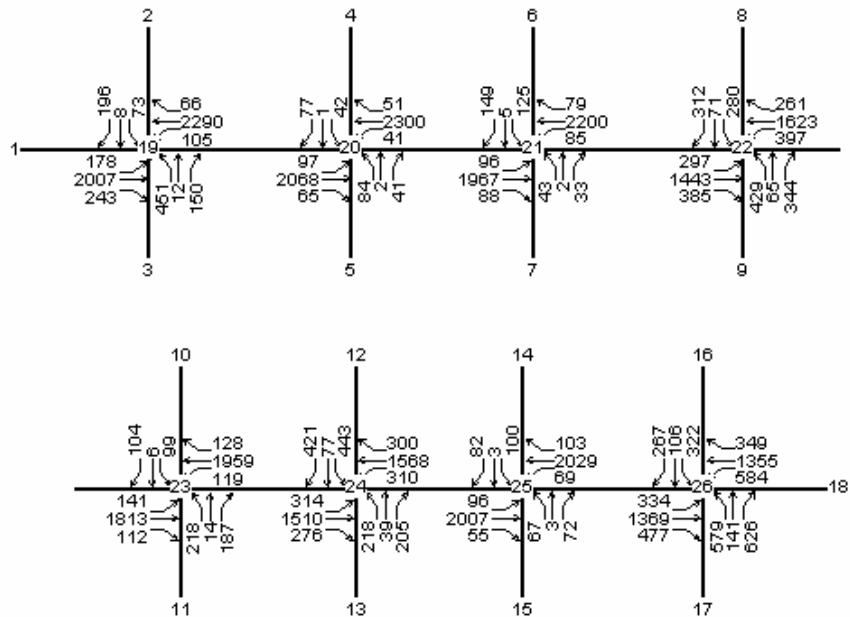


Figure 3-9 Estimated turning movements for arterial network ($RMSE_{turn} = 209.048$).

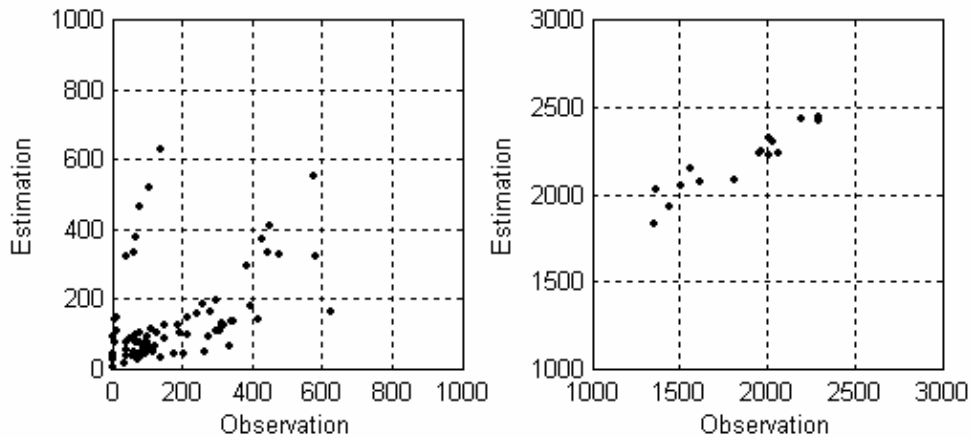


Figure 3-10 Comparison of derived and observed turning movements.

3.5.3 Mixed Freeway and Surface Streets

The City of Santa Rosa is located approximately 60 miles north of San Francisco. The study area, depicted in Figure 3-10, occupies the southeast quadrant of the city, where major changes in land use and transportation network are planned. Several major residential projects throughout the area are being proposed. In addition, a major arterial street will connect the east and south ends of the study area. There are 23 traffic analysis zones (TAZs) including the external stations, which result in a total of 506 O-D pairs.

Traffic counts over a period of 24 hours were collected on average weekdays with electronic devices. Link volumes of the evening peak hour (5pm to 6pm) are used for this study. Due to limitation in the number of devices, not all links were collected on the same day. In addition, link volumes on some of the smaller local streets were not collected. There are 91 out of 174 network links with traffic counts, which is approximately 52 percent of the total number of links. Turning movement counts are collected later and used to validate the turning movement estimates. There is no information relative to movement delay. The dispersion parameter for this network is assumed to be 18 hr^{-1} , which is the inverse of the average O-D travel time. Since the traffic counts are highly inconsistent throughout the network (e.g., the summation of inflow is not equal summation of outflow at nodes), the error bound for all links are arbitrarily set at ± 20 percent, which is slightly larger than the normal value.

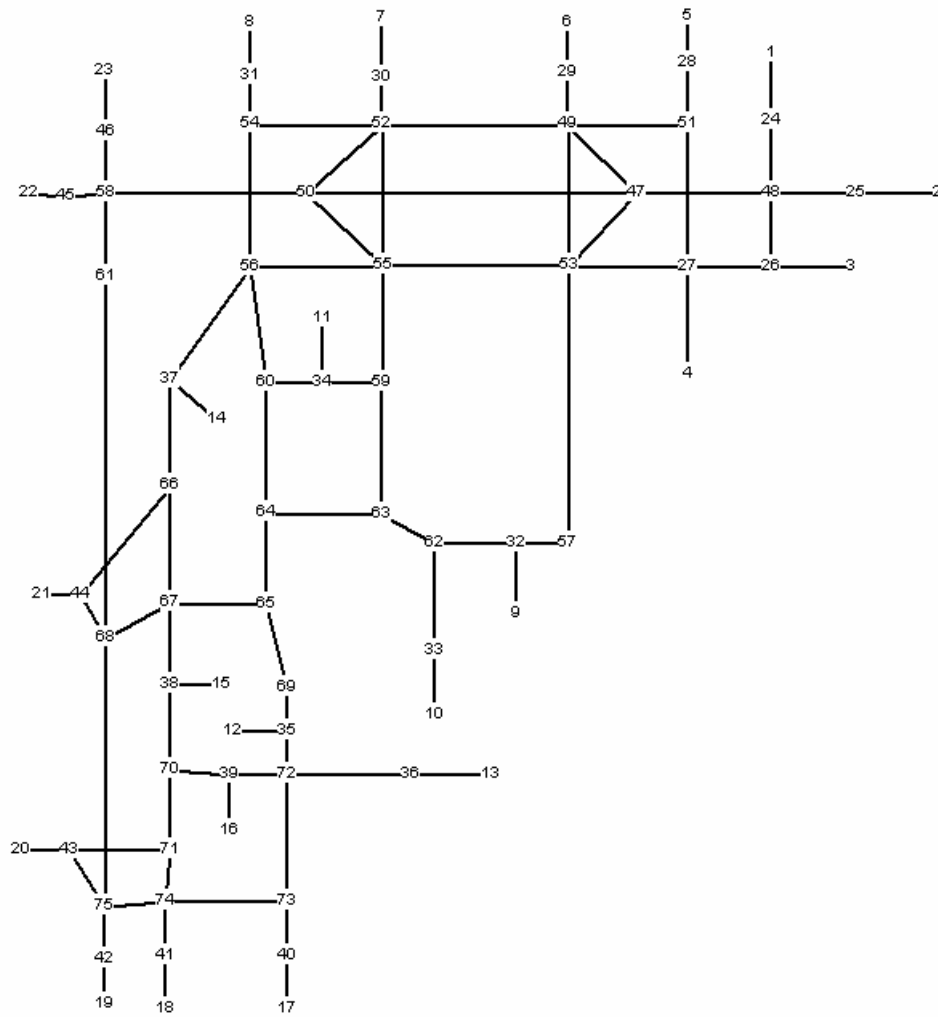


Figure 3-11 Layout of mixed freeway and surface streets, City of Santa Rosa, CA, USA.

The accuracy, in terms of link flow estimates, is presented in Figure 3-12. As expected, PFE can produce the traffic flows within the acceptable error bounds (± 20 percent). The estimates for the turning movements of the intersections with observations are presented in Table 3-12. In general, the turning movement estimates derived from PFE are acceptable for movements with higher volumes, particularly the through movements. However, PFE has a difficulty to estimate the left- and right-turn movements, as well as, those with low volumes. Using this set of traffic counts, the turning movements are derived with a RMSE of 150.187.

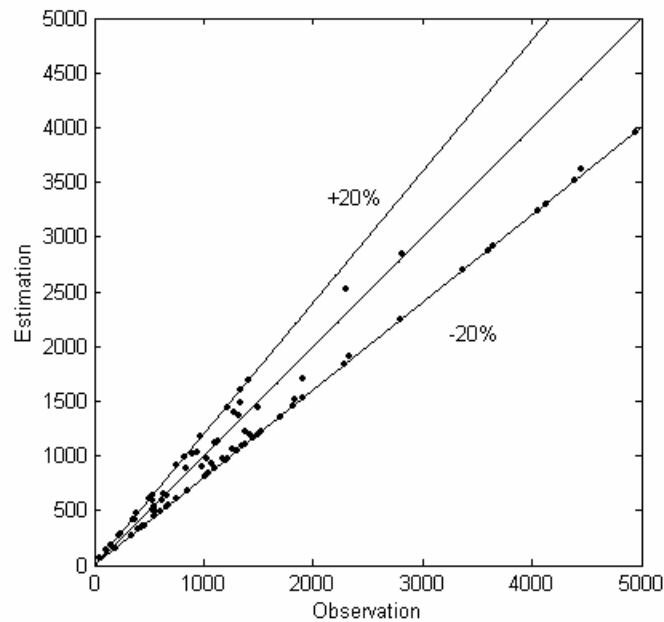


Figure 3-12 Comparison between estimated and observed traffic counts.

Table 3-12 Comparisons of Turning Movement Estimations (RMSE = 150.187)

Intersection		NBL	NBT	NBR	SBL	SBT	SBR	EBL	EBT	EBR	WBL	WBT	WBR
62	Estimation	143		42					312	139	18	203	
	Turning Movement Counts	83		71					918	108	88	568	
63	Estimation				112		253	204	340			284	62
	Turning Movement Counts				34		380	197	489			291	25
64	Estimation		478	447	97	546					421		117
	Turning Movement Counts		483	587	41	612					574		54
72	Estimation	109	729	3	26	772	333	293	30	185	9	27	27
	Turning Movement Counts	115	921	21	4	889	254	379	24	127	11	22	19
73	Estimation	220	756			773	192	84		51			
	Turning Movement Counts	181	1,045			944	79	26		95			

A close examination of the actual turning movement counts and PFE outputs reveals potential reasons, the areas to improve the quality of the turning movement estimates, as well as the O-D trip table. First, the network representation as well as intersection signal timings are crucial to the accuracy of the estimation. Zones and streets serving up to a certain number of flows should be included in the network representation. For example, in this case study, several residential zones and the streets that connect them to a major arterial are not included in the network representation, which inevitably creates inconsistencies in the estimation of turning movements

at intersections upstream and downstream of these residential zones. In turn, these aforementioned factors imply the importance of path set used in the estimation. Different path sets, generated either by a path enumeration or a column generation, certainly lead to different estimates of which their accuracies might be totally different. Second, traffic counts on links connecting to network gateways (i.e., end nodes) are also important to the quality of the estimation (estimating the total demand utilizing the network), since they ensure the accuracy of the estimation along major roadways. Third, the fact that link traffic counts used were not collected at the same time also represents an important source of errors (inconsistency problem) in this case study.

4 SENSITIVITY ANALYSIS FOR PFE MODEL

4.1 Derivation of Sensitivity Expression

Let $\mathbf{u}$ be a vector of perturbations including Δv_a (change of traffic counts) for all $a \in \mathbf{M}$ and ΔC_a (change of link capacity) for all $a \in \mathbf{U}$. The PFE has the following mathematical formulation.

$$\text{Minimize: } Z = \frac{1}{q} \sum_k f_k(?) \cdot (\ln f_k(?) - 1) + \sum_a \int_0^{\sum_k f_k(?) \cdot d_a^k} t_a(w, ?) dw, \quad (4.1)$$

Subject to:

$$\sum_k f_k(?) \cdot d_a^k = v_a(?), \quad \forall a \in \mathbf{M}, \quad (4.2)$$

$$\sum_k f_k(?) \cdot d_a^k \leq C_a(?), \quad \forall a \in \mathbf{U}, \quad (4.3)$$

$$f_k(?) \geq 0, \quad \forall k. \quad (4.4)$$

Given $\mathbf{u}$ and $\mathbf{d}$ are the vectors of dual variables associated with constraints (4.2) and (4.3). $\mathbf{d}$ must be non-negative. The Lagrangian of this problem can be written as:

$$\begin{aligned} \Gamma(\mathbf{f}, \mathbf{u}, \mathbf{d}, ?) = & \frac{1}{q} \sum_k f_k(?) \cdot (\ln f_k(?) - 1) + \sum_a \int_0^{\sum_k f_k(?) \cdot d_a^k} t_a(w, ?) dw + \\ & \sum_{a \in \mathbf{M}} u_a(?) \cdot \left(\sum_k f_k(?) \cdot d_a^k - v_a(?) \right) + \sum_{a \in \mathbf{U}} d_a(?) \cdot \left(\sum_k f_k(?) \cdot d_a^k - C_a(?) \right). \end{aligned} \quad (4.5)$$

At the optimality, the following Karoush-Kuhn-Tucker (KKT) conditions are satisfied.

$$\nabla_{\mathbf{f}} \Gamma(\mathbf{f}(?), \mathbf{u}(?), \mathbf{d}(?), ?) = 0, \quad (4.6)$$

$$\sum_k f_k(?) \cdot d_a^k - v_a(?) = 0, \quad \forall a \in \mathbf{M}, \quad (4.7)$$

$$d_a(?) \cdot \left(\sum_k f_k(?) \cdot d_a^k - C_a(?) \right) = 0, \quad \forall a \in \mathbf{U}. \quad (4.8)$$

where

$$\nabla_{f_k} \Gamma = \frac{1}{q} f_k(?) + \sum_{a \in \mathbf{A}} t_a(x_a, ?) \cdot d_a^k + \sum_{a \in \mathbf{M}} u_a(?) \cdot d_a^k + \sum_{a \in \mathbf{U}} d_a(?) \cdot d_a^k, \quad \forall k \in \mathbf{K}. \quad (4.9)$$

$$x_a(?) = \sum_k d_a^k \cdot f_k(?), \quad \forall a \in \mathbf{A}. \quad (4.10)$$

The purpose of sensitivity analysis is to analyze the behavior of a solution of the problem $\mathbf{f}(\mathbf{u})$, when $\mathbf{u}$ is subject to a small perturbation, given Equations (4.7) and (4.8) are linearly independent. For any small perturbation, $\mathbf{f}$, $\mathbf{u}$, and $\mathbf{d}$ must be adjusted to maintain the optimality (KKT condition). Let $\mathbf{H}$ be a set of equations on the left hand side of (4.6), (4.7), and (4.8) and $\mathbf{y}(\mathbf{u})$ be a vector of solutions including $\mathbf{f}(\mathbf{u})$, $\mathbf{u}(\mathbf{u})$, and $\mathbf{d}(\mathbf{u})$. At the optimality,

$$\nabla_{\mathbf{y}} \mathbf{H} = 0. \quad (4.11)$$

Since $\mathbf{y}(\mathbf{u})$ is also a function of $\mathbf{u}$, the chain rule is applied. Equation (4.11) becomes:

$$\nabla_{\mathbf{y}} \mathbf{H} + \nabla_{\mathbf{y}} \mathbf{H} \times \nabla_{\mathbf{y}} \mathbf{y} = 0. \quad (4.12)$$

From equation (4.10), $\nabla_{\mathbf{y}} \mathbf{y}$ is the element of interest; it represents the rate of change of the solution vector (including path flows $\mathbf{f}$) with respect to a unit change of perturbation. It is essentially the sensitivities of path flows and dual variables with respect to a small perturbation. Equation (4.12) can be rearranged to obtain the sensitivity expression as follow.

$$\nabla_{\mathbf{y}} \mathbf{y} = \nabla_{\mathbf{y}} \mathbf{H}^{-1} \times (-\nabla_{\mathbf{y}} \mathbf{H}). \quad (4.13)$$

The following gives a more detailed expression for each element of Equation (4.13).

Finding $\nabla_{\mathbf{y}} \mathbf{H}$

$$\text{Let } \nabla_{\mathbf{y}} \mathbf{H} = \begin{bmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} & \mathbf{J}_{13} \\ \mathbf{J}_{21} & \mathbf{J}_{22} & \mathbf{J}_{23} \\ \mathbf{J}_{31} & \mathbf{J}_{32} & \mathbf{J}_{33} \end{bmatrix}, \text{ then}$$

$$\mathbf{J}_{11} = \left[\nabla_{f_k f_\ell} \Gamma \right]_{K \times K} = \text{diag} \left(\frac{1}{q \cdot f_k(?)} \right) + ? \cdot \text{diag} \left(\nabla_{x_a} t_a(x_a, ?) \right) \cdot ?^T,$$

where $\text{diag}(\cdot)$ is a diagonal matrix, and $? = \begin{bmatrix} d_a^k \end{bmatrix}$ is the path-link incidence matrix.

$$\mathbf{J}_{12} = \left[\nabla_{f_k u_a} \Gamma \right]_{K \times M} = \begin{bmatrix} d_a^k, \forall a \in M \end{bmatrix} = ?_m,$$

$$\mathbf{J}_{13} = \left[\nabla_{f_k d_a} \Gamma \right]_{K \times U} = \begin{bmatrix} d_a^k, \forall a \in U \end{bmatrix} = ?_u,$$

where Δ_m and Δ_u are the path-link incidence matrices considering observed and capacitated links respectively.

$$\mathbf{J}_{21} = \left[\nabla_{f_\ell} \left(\sum_k d_a^k \cdot f_k(?) - v_a(?) \right) \right]_{M \times K} = [d_a^\ell, \forall a \in M] = \mathbf{?}_m^T.$$

$$\mathbf{J}_{22} = \left[\nabla_{u_a} \left(\sum_k d_a^k \cdot f_k(?) - v_a(?) \right) \right]_{M \times M} = \mathbf{0}.$$

$$\mathbf{J}_{23} = \left[\nabla_{d_a} \left(\sum_k d_a^k \cdot f_k(?) - v_a(?) \right) \right]_{M \times U} = \mathbf{0}.$$

$$\mathbf{J}_{31} = \left[\nabla_{f_\ell} \left(d_a(?) \cdot \left(\sum_k d_a^k \cdot f_k(?) - C_a(?) \right) \right) \right]_{U \times K} = [d_a(?) \cdot d_a^\ell] = \text{diag}(d_a(?)) \cdot \mathbf{?}_u^T,$$

$$\mathbf{J}_{32} = \left[\nabla_{u_a} \left(d_a(?) \cdot \left(\sum_k d_a^k \cdot f_k(?) - C_a(?) \right) \right) \right]_{U \times M} = \mathbf{0}.$$

$$\mathbf{J}_{33} = \left[\nabla_{d_a} \left(d_a(?) \cdot \left(\sum_k d_a^k \cdot f_k(?) - C_a(?) \right) \right) \right]_{U \times U} = \left[\sum_k d_a^k \cdot f_k(?) - C_a(?) \right] = \mathbf{0}.^4$$

Then,

$$\nabla_y \mathbf{H} = \begin{bmatrix} \text{diag}\left(\frac{1}{q \cdot f_k(?)}\right) + \mathbf{?} \cdot \text{diag}\left(\nabla_{x_a} t_a(x_a, ?)\right) \cdot \mathbf{?}^T & \mathbf{?}_m & \mathbf{?}_u \\ \text{---} & \mathbf{?}_m^T & \mathbf{0} & \mathbf{0} \\ \text{---} & \mathbf{d} \cdot \mathbf{e}^T \cdot \mathbf{?}_u^T & \mathbf{0} & \mathbf{0} \end{bmatrix}. \quad (4.14)$$

⁴ Note that only active constraints are included in the sensitivity expression. Therefore, when the inequality constraints are binding, they are equal to zero.

Finding $\nabla_{\mathbf{?}} \mathbf{H}$

Note that the expression of this matrix is dependent on the parameters being perturbed (u). In the application of PFE, the change of observed flows on link $a \in \mathbf{M}$, and the change of capacity of link $a \in \mathbf{U}$, are investigated. Therefore, $\nabla_{\mathbf{?}} \mathbf{H}$ can be determined as follows.

Given that $v_a(\mathbf{?}) = v_a + \Delta v_a$, and $C_a(\mathbf{?}) = C_a + \Delta C_a$.

$$\text{Let } \nabla_{\mathbf{?}} \mathbf{H} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \\ \mathbf{B}_{31} & \mathbf{B}_{32} \end{bmatrix}, \text{ then}$$

$$\mathbf{B}_{11} = \left[\nabla_{f_k \Delta v_a} \Gamma \right]_{\mathbf{K} \times \mathbf{M}} = \mathbf{0}.$$

$$\mathbf{B}_{12} = \left[\nabla_{f_k \Delta C_a} \Gamma \right]_{\mathbf{K} \times \mathbf{U}} = \left[\nabla_{\Delta C_a} t_a(x_a, \mathbf{?}) \cdot d_a^k, \forall a \in \mathbf{U} \right] = \mathbf{?}_u \cdot \text{diag} \left(\nabla_{\Delta C_a} t_a(x_a, \mathbf{?}) \right).$$

$$\mathbf{B}_{21} = \left[\nabla_{\Delta v_a} \left(\sum_k d_a^k \cdot f_k(\mathbf{?}) - v_a(\mathbf{?}) \right) \right]_{\mathbf{M} \times \mathbf{M}} = -\mathbf{I}.$$

$$\mathbf{B}_{22} = \left[\nabla_{\Delta C_a} \left(\sum_k d_a^k \cdot f_k(\mathbf{?}) - v_a(\mathbf{?}) \right) \right]_{\mathbf{M} \times \mathbf{U}} = \mathbf{0}.$$

$$\mathbf{B}_{31} = \left[\nabla_{\Delta v_a} \left(d_a(\mathbf{?}) \cdot \left(\sum_k d_a^k \cdot f_k(\mathbf{?}) - C_a(\mathbf{?}) \right) \right) \right]_{\mathbf{U} \times \mathbf{M}} = \mathbf{0}.$$

$$\mathbf{B}_{32} = \left[\nabla_{\Delta C_a} \left(d_a(\mathbf{?}) \cdot \left(\sum_k d_a^k \cdot f_k(\mathbf{?}) - C_a(\mathbf{?}) \right) \right) \right]_{\mathbf{U} \times \mathbf{U}} = \left[-d_a(\mathbf{?}), \forall a \in \mathbf{U} \right] = -\text{diag}(d_a(\mathbf{?})).$$

Then,

$$\nabla_{\gamma} \mathbf{H} = \begin{bmatrix} \mathbf{0} & \gamma_u \cdot \text{diag}\left(\nabla_{\Delta C_a} t_a(x_a, \gamma)\right) \\ -\mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\text{diag}\left(d_a(\gamma)\right) \end{bmatrix}. \quad (4.15)$$

4.2 Determination of $\nabla_y \mathbf{H}^{-1}$ and Sensitivities of Path Flow Estimates

For any real-world applications, only some parts of the inverted matrix are used for further analysis. Therefore, due to the computational burden and the storage requirement, there is no need to compute the inversion of a full matrix; only the required sub-matrices may need to be computed. In the application of PFE, the sensitivities of dual variables regarding the perturbations are rarely needed. Only the sensitivities of path flows are usually used to quantify the uncertainty of the O-D flow estimated (error and uncertainty analysis). Again, considering the sensitivity expression for solution vector $\mathbf{y}(\mathbf{u})$.

$$\nabla_{\gamma} \mathbf{y} = \nabla_y \mathbf{H}^{-1} \times (-\nabla_{\gamma} \mathbf{H}),$$

$$\begin{bmatrix} \nabla_{\Delta v} \mathbf{f} & \nabla_{\Delta C} \mathbf{f} \\ \nabla_{\Delta v} \mathbf{u} & \nabla_{\Delta C} \mathbf{u} \\ \nabla_{\Delta v} \mathbf{d} & \nabla_{\Delta C} \mathbf{d} \end{bmatrix} = - \begin{bmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} & \mathbf{J}_{13} \\ \mathbf{J}_{21} & \mathbf{J}_{22} & \mathbf{J}_{23} \\ \mathbf{J}_{31} & \mathbf{J}_{32} & \mathbf{J}_{33} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \\ \mathbf{B}_{31} & \mathbf{B}_{32} \end{bmatrix}.$$

$$\text{Let } \begin{bmatrix} \tilde{\mathbf{J}}_{11} & \tilde{\mathbf{J}}_{12} & \tilde{\mathbf{J}}_{13} \\ \tilde{\mathbf{J}}_{21} & \tilde{\mathbf{J}}_{22} & \tilde{\mathbf{J}}_{23} \\ \tilde{\mathbf{J}}_{31} & \tilde{\mathbf{J}}_{32} & \tilde{\mathbf{J}}_{33} \end{bmatrix} \text{ be the inversion of } \nabla_y \mathbf{H}.$$

Note from the derivation provided earlier that $\mathbf{B}_{11}$, $\mathbf{B}_{22}$, and $\mathbf{B}_{31}$ are null matrices. Therefore, the sensitivities of path flow estimates with respect to observed link flows and link capacities are:

$$\nabla_{\gamma v} \mathbf{f} = -\tilde{\mathbf{J}}_{12} \cdot \mathbf{B}_{21} = -(\tilde{\mathbf{J}}_{12}) \cdot (-\mathbf{I}) = \tilde{\mathbf{J}}_{12}. \quad (4.16)$$

$$\nabla_{\gamma C} \mathbf{f} = -\left(\tilde{\mathbf{J}}_{11} \cdot \mathbf{B}_{12} + \tilde{\mathbf{J}}_{13} \cdot \mathbf{B}_{32}\right). \quad (4.17)$$

As can be seen from Equations (4.16) and (4.17), only $\tilde{\mathbf{J}}_{11}$, $\tilde{\mathbf{J}}_{12}$, and $\tilde{\mathbf{J}}_{13}$ need to be computed. In order to find the inversion by partitioning, $\nabla_{\mathbf{y}} \mathbf{H}$ needs to be rearranged as follow.

$$\nabla_{\mathbf{y}} \mathbf{H} = \left[\begin{array}{c|c|c} \mathbf{J}_{11} & \mathbf{J}_{12} & \mathbf{J}_{13} \\ \hline \mathbf{J}_{21} & \mathbf{J}_{22} & \mathbf{J}_{23} \\ \hline \mathbf{J}_{31} & \mathbf{J}_{32} & \mathbf{J}_{33} \end{array} \right] = \begin{bmatrix} \mathbf{D} & \mathbf{E} \\ \mathbf{F} & \mathbf{G} \end{bmatrix},$$

where $\mathbf{D}$ and $\mathbf{G}$ are the square matrices while the others are not necessary. Based on the new configuration of the matrix, the corresponding elements ($\tilde{\mathbf{D}}, \tilde{\mathbf{E}}, \tilde{\mathbf{F}}$, and $\tilde{\mathbf{G}}$) of the inversion matrix are given by (Press et al., 1992):

$$\tilde{\mathbf{D}} = \mathbf{D}^{-1} + (\mathbf{D}^{-1} \cdot \mathbf{E}) \cdot (\mathbf{G} - \mathbf{F} \cdot \mathbf{D}^{-1} \cdot \mathbf{E})^{-1} \cdot (\mathbf{F} \cdot \mathbf{D}^{-1}).$$

$$\tilde{\mathbf{E}} = -(\mathbf{D}^{-1} \cdot \mathbf{E}) \cdot (\mathbf{G} - \mathbf{F} \cdot \mathbf{D}^{-1} \cdot \mathbf{E})^{-1}.$$

$$\tilde{\mathbf{F}} = -(\mathbf{G} - \mathbf{F} \cdot \mathbf{D}^{-1} \cdot \mathbf{E})^{-1} \cdot (\mathbf{F} \cdot \mathbf{D}^{-1}).$$

$$\tilde{\mathbf{G}} = (\mathbf{G} - \mathbf{F} \cdot \mathbf{D}^{-1} \cdot \mathbf{E})^{-1}$$

Clearly, $\tilde{\mathbf{D}}$ represents $\tilde{\mathbf{J}}_{11}$ while $\tilde{\mathbf{E}}$ represents $\tilde{\mathbf{J}}_{12}$ and $\tilde{\mathbf{J}}_{13}$. Let relate $\tilde{\mathbf{D}}$ and $\tilde{\mathbf{E}}$ to the original matrix $\mathbf{J}$ as follows.

$$\begin{aligned} \tilde{\mathbf{J}}_{11} &= \mathbf{J}_{11}^{-1} + \left(\mathbf{J}_{11}^{-1} \cdot \left[\mathbf{J}_{12} \mid \mathbf{J}_{13} \right] \right) \cdot \left(- \left[\frac{\mathbf{J}_{21}}{\mathbf{J}_{31}} \right] \cdot \mathbf{J}_{11}^{-1} \cdot \left[\mathbf{J}_{12} \mid \mathbf{J}_{13} \right] \right)^{-1} \cdot \left(\left[\frac{\mathbf{J}_{21}}{\mathbf{J}_{31}} \right] \cdot \mathbf{J}_{11}^{-1} \right) \\ &= \mathbf{J}_{11}^{-1} - \left[\mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12} \mid \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{13} \right] \cdot \left[\frac{\mathbf{J}_{21} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12}}{\mathbf{J}_{31} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12}} \mid \frac{\mathbf{J}_{21} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{13}}{\mathbf{J}_{31} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{13}} \right]^{-1} \cdot \left[\frac{\mathbf{J}_{21} \cdot \mathbf{J}_{11}^{-1}}{\mathbf{J}_{31} \cdot \mathbf{J}_{11}^{-1}} \right] \end{aligned} \quad (4.18)$$

$$\begin{aligned} \left[\tilde{\mathbf{J}}_{12} \mid \tilde{\mathbf{J}}_{13} \right] &= - \left(\mathbf{J}_{11}^{-1} \cdot \left[\mathbf{J}_{12} \mid \mathbf{J}_{13} \right] \right) \cdot \left(- \left[\frac{\mathbf{J}_{21}}{\mathbf{J}_{31}} \right] \cdot \mathbf{J}_{11}^{-1} \cdot \left[\mathbf{J}_{12} \mid \mathbf{J}_{13} \right] \right)^{-1} \\ &= \left[\mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12} \mid \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{13} \right] \cdot \left[\frac{\mathbf{J}_{21} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12}}{\mathbf{J}_{31} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12}} \mid \frac{\mathbf{J}_{21} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{13}}{\mathbf{J}_{31} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{13}} \right]^{-1} \end{aligned} \quad (4.19)$$

From Equations (4.18) and (4.19), there is a common matrix to be inverted. The same technique (*inversion by portioning*) used earlier can also be applied here.

$$\text{Let } \left[\begin{array}{c|c} \mathbf{F} & \mathbf{S} \\ \hline \mathbf{O} & ? \end{array} \right] = \left[\begin{array}{c|c} \mathbf{J}_{21} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12} & \mathbf{J}_{21} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{13} \\ \hline \mathbf{J}_{31} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12} & \mathbf{J}_{31} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{13} \end{array} \right], \text{ then}$$

$$\tilde{\mathbf{F}} = \mathbf{F}^{-1} + \underbrace{(\mathbf{F}^{-1} \cdot \mathbf{S})}_a \cdot \underbrace{(? - \mathbf{O} \cdot \mathbf{F}^{-1} \cdot \mathbf{S})^{-1}}_b \cdot \underbrace{(\mathbf{O} \cdot \mathbf{F}^{-1})}_c.$$

$$\tilde{\mathbf{S}} = -\underbrace{(\mathbf{F}^{-1} \cdot \mathbf{S})}_a \cdot \underbrace{(? - \mathbf{O} \cdot \mathbf{F}^{-1} \cdot \mathbf{S})^{-1}}_b.$$

$$\tilde{\mathbf{O}} = -\underbrace{(? - \mathbf{O} \cdot \mathbf{F}^{-1} \cdot \mathbf{S})^{-1}}_b \cdot \underbrace{(\mathbf{O} \cdot \mathbf{F}^{-1})}_c.$$

$$\tilde{?} = \underbrace{(? - \mathbf{O} \cdot \mathbf{F}^{-1} \cdot \mathbf{S})^{-1}}_b.$$

From four equations above, there are three common elements, which are a , b , and c . They can be rewritten in terms of the original matrix $\mathbf{J}$ as follows.

$$a. (\mathbf{J}_{21} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12})^{-1} \cdot (\mathbf{J}_{21} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{13}),$$

$$b. \left((\mathbf{J}_{31} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{13}) - (\mathbf{J}_{31} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12}) \cdot (\mathbf{J}_{21} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12})^{-1} \cdot (\mathbf{J}_{21} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{13}) \right)^{-1},$$

$$c. (\mathbf{J}_{31} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12}) \cdot (\mathbf{J}_{21} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12})^{-1}.$$

Now Equations (4.18) and (4.19) can be rewritten as:

$$\begin{aligned} \tilde{\mathbf{J}}_{11} &= \mathbf{J}_{11}^{-1} - \left[\mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12} \mid \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{13} \right] \cdot \left[\begin{array}{c|c} \tilde{\mathbf{F}} & \tilde{\mathbf{S}} \\ \hline \tilde{\mathbf{O}} & \tilde{?} \end{array} \right] \cdot \left[\begin{array}{c} \mathbf{J}_{21} \cdot \mathbf{J}_{11}^{-1} \\ \hline \mathbf{J}_{31} \cdot \mathbf{J}_{11}^{-1} \end{array} \right] \\ &= \mathbf{J}_{11}^{-1} - \mathbf{J}_{11}^{-1} \cdot \left(\mathbf{J}_{12} \cdot \tilde{\mathbf{F}} \cdot \mathbf{J}_{21} + \mathbf{J}_{13} \cdot \tilde{\mathbf{O}} \cdot \mathbf{J}_{21} + \mathbf{J}_{12} \cdot \tilde{\mathbf{S}} \cdot \mathbf{J}_{31} + \mathbf{J}_{13} \cdot \tilde{?} \cdot \mathbf{J}_{31} \right) \cdot \mathbf{J}_{11}^{-1} \end{aligned} \quad (4.20)$$

$$\tilde{\mathbf{J}}_{12} = \left[\mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12} \mid \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{13} \right] \cdot \left[\begin{array}{c} \tilde{\mathbf{F}} \\ \hline \tilde{\mathbf{O}} \end{array} \right] = \mathbf{J}_{11}^{-1} \cdot \left(\mathbf{J}_{12} \cdot \tilde{\mathbf{F}} + \mathbf{J}_{13} \cdot \tilde{\mathbf{O}} \right). \quad (4.21)$$

$$\tilde{\mathbf{J}}_{13} = \left[\mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12} \mid \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{13} \right] \cdot \begin{bmatrix} \tilde{\mathbf{S}} \\ \tilde{\mathbf{?}} \end{bmatrix} = \mathbf{J}_{11}^{-1} \cdot \left(\mathbf{J}_{12} \cdot \tilde{\mathbf{S}} + \mathbf{J}_{13} \cdot \tilde{\mathbf{?}} \right). \quad (4.22)$$

By substituting $\tilde{\mathbf{J}}_{11}$, $\tilde{\mathbf{J}}_{12}$, and $\tilde{\mathbf{J}}_{13}$ back into Equations (4.16) and (4.17), the sensitivities of path flow estimates with respect to small perturbations of observed link flows and link capacities are finally obtained. In the case that there is no active capacity constraint, the related elements, primarily $\mathbf{J}_{i3}$, $\mathbf{J}_{3k}$, and $\mathbf{B}_{i2}$, do not exist. Equation (4.21) cannot be used. By excluding all elements related to the capacity constraint, $\tilde{\mathbf{J}}_{12}$ has the following expression:

$$\tilde{\mathbf{J}}_{12} = \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12} \cdot \mathbf{F}^{-1} = \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12} \cdot \left(\mathbf{J}_{21} \cdot \mathbf{J}_{11}^{-1} \cdot \mathbf{J}_{12} \right)^{-1}. \quad (4.23)$$

4.3 Determination of the Independence Constraints

One of the requirements for conducting sensitivity analysis is that the active constraints, Equations (4.2) and (4.3), must be linearly independent. One way to identify the independent constraints is to use Gaussian elimination with partial pivoting. After solving the PFE, active inequality (capacity) constraints are revealed. These constraints are combined with the observation constraints, which are always considered active, and used for sensitivity analysis. However, since some of them may be linearly dependent, the dual variables previously obtained might not be unique and the sensitivity analysis cannot be performed right away. This will only be the case for capacity constraints, since their dual variables are used in the sensitivity expressions derived earlier. The dependent constraints must be removed first and the solution (e.g., path flows and dual variables) needs to be re-estimated.

Consider the observation and active capacity constraints,

$$\begin{bmatrix} \mathbf{?}_m \\ \mathbf{?}_u \end{bmatrix}^T \cdot \mathbf{f} = \begin{bmatrix} \mathbf{v} \\ \mathbf{C} \end{bmatrix}. \quad (4.24)$$

Linearly independent constraints (links) are basically the independent columns of path-link incidence matrix considering only observed and capacitated links. The independent columns are simply the columns with pivot elements. If there is no pivot element in the column, it is linearly dependent⁵ with other columns. The basic concept of Gaussian elimination is to eliminate all elements below the pivot position (by making them zero) using elemental row operations (subtracting and interchanging). The procedure can be demonstrated using the following path-link incidence matrix (6 links and 8 paths), see hypothetical network in Figure 4-1.

⁵ This column can be generated by linear combination of other independent columns.

$$? = \begin{bmatrix} \hat{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

It is clear that the first pivot position is 1, which is in the first row and the first column, as indicated by the “hat” sign. Note that if the element in the first row and the first column is zero, the row interchange is needed to bring the nonzero element to the first row. After identifying the pivot element, other nonzero elements below the pivot position are eliminated as follow.

$$? = \begin{bmatrix} \hat{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{array}{l} \\ R_4 - R_1 \\ \\ R_6 - R_1 \\ R_7 - R_1 \\ \end{array}.$$

The next pivot element is 1, which is in the second row and the second column. This procedure is repeated for all columns (move down and to the right). The “hat” sign indicates all pivot elements. All intermediate steps and final pivoted matrix are given below.

$$? = \begin{bmatrix} \hat{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \hat{1} & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{array}{l} \\ \\ -(R_3 - R_2) \\ \\ -(R_5 - R_2) \\ \\ -(R_8 - R_2) \end{array} \Rightarrow ? = \begin{bmatrix} \hat{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \hat{1} & 1 & 0 & 0 & 1 \\ 0 & 0 & \hat{1} & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{array}{l} \\ \\ \\ \\ \\ -(R_8 - R_3) \end{array},$$

$$\begin{aligned}
? &= \begin{bmatrix} \hat{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \hat{1} & 1 & 0 & 0 & 1 \\ 0 & 0 & \hat{1} & 0 & 0 & 1 \\ 0 & 0 & 0 & \hat{1} & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad -(R_7 - R_4) \Rightarrow ? = \begin{bmatrix} \hat{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \hat{1} & 1 & 0 & 0 & 1 \\ 0 & 0 & \hat{1} & 0 & 0 & 1 \\ 0 & 0 & 0 & \hat{1} & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{matrix} R_5 = R_6, \\ R_6 = R_5 \end{matrix} \\
? &= \begin{bmatrix} \hat{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \hat{1} & 1 & 0 & 0 & 1 \\ 0 & 0 & \hat{1} & 0 & 0 & 1 \\ 0 & 0 & 0 & \hat{1} & 1 & 0 \\ 0 & 0 & 0 & 0 & \hat{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad R_7 - R_5 \Rightarrow ? = \begin{bmatrix} \hat{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \hat{1} & 1 & 0 & 0 & 1 \\ 0 & 0 & \hat{1} & 0 & 0 & 1 \\ 0 & 0 & 0 & \hat{1} & 1 & 0 \\ 0 & 0 & 0 & 0 & \hat{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & \hat{1} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{matrix} \\ \\ \\ \\ \\ R_8 - R_6 \end{matrix}
\end{aligned}$$

As can be seen, since there is a pivot element in each column, all six columns (links or constraints) are linearly independent. In addition, the number of linearly independent columns also represents a *Rank* of the matrix. Let m be the number of columns and n be the number of independent columns. If n is less than m , there might be more than one combination of n independent columns (links), at most $C_{n,m}$ combinations.

The problem then is which combination of n links is to be selected. One possible way is to pick the links associated with higher traffic volume; therefore, the total flow produced by the selected set is maximized. This is because the derived sensitivity information will be used for the error and uncertainty analysis and it is always assumed that the larger error is inherited with a higher value of measurement (input). To select the set of links with the maximal traffic volume, all columns must be sorted by the corresponding observed flows in the descending order from left to right. Then after performing the Gaussian elimination, the first n columns with pivot elements are the linearly independent links containing the maximal observed traffic volume.

4.4 Variance of Estimation Results

4.4.1 First Test Network

The network, as depicted in Figure 4-1, was used by Bell and Iida (1997) to demonstrate the application of PFE with constant link cost. It consists of 6 links and 4 O-D pairs: 1-2, 1-3, 2-4, and 2-4. There are a total of 8 paths connecting these 4 O-D pairs. It was assumed that traffic counts are available on links 1 and 2; 10 and 10 units of flows respectively. The estimated path flows are provided in Table 4-1, while the dual variables of constraints associated with network links are provided together with link characteristics in Table 4-2.

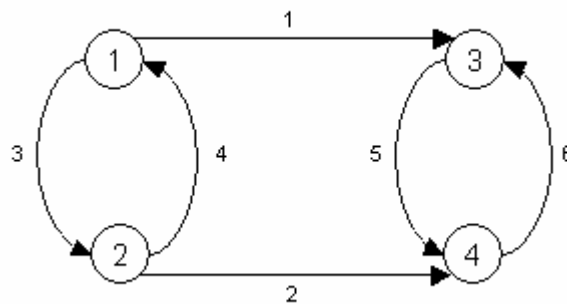


Figure 4-1 Test network with 6 link and 4 O-D pairs.

Table 4-1 Estimated Path Flows

Path	1	2	3	4	5	6	7	8
Links	1	2,3,6	2	1,4,5	2,3	1,5	1,4	2,6
Flow	4.3572	1.1326	4.3572	1.1326	1.8674	2.6428	1.8674	2.6428

Table 4-2 Link Flows, Link Capacities, and Dual Variables

Link	Capacity	Flow	Dual variable
1 (measured)		10.0000	19.7183
2 (measured)		10.0000	19.7183
3	3.0000	3.0000	6.4730
4	3.0000	3.0000	6.4730
5	10.0000	3.7754	0.0000
6	10.0000	3.7754	0.0000

As can be seen from Table 4-2, the capacity constraints of links 3 and 4 are active. Four constraints might be involved in conducting the sensitivity analysis (two observed links and two capacitated links). For this network, it can be shown that there are six linearly independent links (as already shown in the previous section). Therefore, all four constraints will be used and the elements required for computing the sensitivities of path flow estimates can be obtained as follows.

$$\mathbf{?}_m = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{?}_u = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \mathbf{d} = \begin{bmatrix} 6.4730 \\ 6.4730 \end{bmatrix}, \text{ and } \mathbf{f} = \begin{bmatrix} 4.3572 \\ 1.1326 \\ 4.3572 \\ 1.1326 \\ 1.8674 \\ 2.6428 \\ 1.8674 \\ 2.6428 \end{bmatrix}.$$

Since link costs are constant, $\nabla_{x_a} t_a(\cdot) = 0$ and $\nabla_{c_a} t_a(\cdot) = 0$.

$$\mathbf{J}_{11} = \begin{bmatrix} 0.2295 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.8829 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2295 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.8829 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5355 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.3784 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5355 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.3784 \end{bmatrix},$$

$$\mathbf{J}_{12} = \mathbf{?}_m, \mathbf{J}_{13} = \mathbf{?}_u, \mathbf{J}_{21} = \mathbf{?}_m^T,$$

$$\mathbf{J}_{31} = \begin{bmatrix} 0 & 6.4730 & 0 & 0 & 6.4730 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6.4730 & 0 & 0 & 6.4730 & 0 \end{bmatrix},$$

$$\mathbf{B}_{32} = \begin{bmatrix} -6.4730 & 0 \\ 0 & -6.4730 \end{bmatrix},$$

$$\tilde{\mathbf{F}} = \begin{bmatrix} 0.1429 & 0 \\ 0 & 0.1429 \end{bmatrix}, \tilde{\mathbf{S}} = \begin{bmatrix} 0 & -0.0221 \\ -0.0221 & 0 \end{bmatrix},$$

$$\tilde{\mathbf{O}} = \begin{bmatrix} 0 & -0.1429 \\ -0.1429 & 0 \end{bmatrix}, \tilde{\mathbf{q}} = \begin{bmatrix} 0.0736 & 0 \\ 0 & 0.0736 \end{bmatrix}.$$

Substituting these elements into Equations (4.16), (4.17), (4.20), (4.21), and (4.22), the sensitivities of path flow estimates to the perturbations of traffic counts on links 1 and 2, and the perturbations of capacities of links 3 and 4 can be obtained as shown in Table 4-3. This sensitivity information is exactly the same as given by Bell and Iida (1997). This verifies the derived sensitivity expression.

Table 4-3 Sensitivities of Path Flows to Link Observations and Link Capacities

Path	Links	Measurement		Capacity	
		Link 1	Link 2	Link 3	Link 4
1	1	0.6225	0	0	-0.6225
2	2,3,6	0	0	0.3775	0
3	2	0	0.6225	-0.6225	0
4	1,4,5	0	0	0	0.3775
5	2,3	0	0	0.6225	0
6	1,5	0.3775	0	0	-0.3775
7	1,4	0	0	0	0.6225
8	2,6	0	0.3775	-0.3775	0

4.4.2 Second Test Network

The network depicted in Figure 4-2 was used for the second example. It consists of 7 links, 5 nodes and 2 O-D pairs: 1-4 and 1-5. It was assumed that traffic counts are available on all links and they are consistent. The estimated path flows are provided in Table 4-4. Table 4-5 displays the estimated link flows as well as the dual variables and link characteristics.

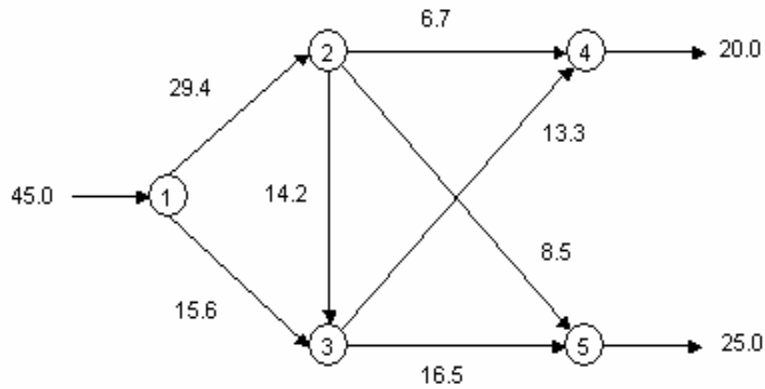


Figure 4-2 Test network with 7 links and 2 O-D pairs.

Table 4-4 Estimated Path Flows

Path	O-D	Links	Flows
1	1-4	1-4	6.7000
2	1-4	2-6	6.9624
3	1-4	1-3-6	6.3376
4	1-5	1-5	8.5000
5	1-5	2-7	8.6376
6	1-5	1-3-7	7.8624

Table 4-5 Link Flows, Dual Variables, and Link Characteristics

Link	t_0	Capacity	Flows	Dual variables	
				1*	2**
1	4	25	29.4	-30.4889	0.0000
2	5.2	25	15.6	-30.5082	-0.0282
3	1	15	14.2	-0.0288	-0.0377
4	5	15	6.7	1.2904	-29.1985
5	5	15	8.5	-1.1367	-31.6256
6	4	15	13.3	1.4138	-29.0662
7	4	15	16.5	-1.2497	-31.7298

* The order of constraints: Links 1, 2, 3, 4, 5, 6, 7

** The order of constraints: Links 7, 6, 5, 4, 3, 2, 1

For this network, it can be shown that there are five linearly independent links ($n = 5$ and $m = 7$). The resultant dual variables shown in Table 5 may not be unique; it is dependent on the order of constraints that the iterative balancing technique aims to satisfy. It was also shown in Table 4-5 that if the order of constraints were reversed, another set of dual variables would be obtained. In this case, there could be up to 21 combinations of 5 independent links. The combination representing the maximal observed flows are links 1, 2, 3, 5, and 7. The new results in terms of path flow and link flow estimates remain unchanged; only the dual variables do change. The new values of dual variables, which are unique, are provided in Table 4-6.

Table 4-6 Dual Variables of Independent Constraints with the Maximal Observed Flows

Link	1	2	3	4	5	6	7
Dual Variables	-29.1985	-29.0944	0.0946	0.0000	-2.4271	0.0000	-2.6636

Five independent links with the maximal observed flows were further used to conduct the sensitivity analysis. As can be seen from Table 4-6, the capacity constraints (links 4 and 6) are not active; the dual variables are zero. Therefore, only $\tilde{\mathbf{J}}_{12}$ needs to be computed using Equation (4.23). All required elements for Equation (4.23) are given below.

$$\mathbf{J}_{11} = \begin{bmatrix} 1.6665 & 0 & 0.1561 & 0.1561 & 0 & 0.1561 \\ 0 & 1.5781 & 0.1115 & 0 & 0.0303 & 0 \\ 0.1561 & 0.1115 & 1.8795 & 0.1561 & 0 & 0.1901 \\ 0.1561 & 0 & 0.1561 & 1.3690 & 0 & 0.1561 \\ 0 & 0.0303 & 0 & 0 & 1.4010 & 0.2130 \\ 0.1561 & 0 & 0.1901 & 0.1561 & 0.2130 & 1.6749 \end{bmatrix}^6,$$

Note that the BPR cost function is used in this example, therefore,

$$\nabla_{x_a} t_a(\cdot) = 0.6 \cdot t_o \cdot x_a^3 / C_a^4,$$

$$\nabla_{C_a} t_a(\cdot) = -0.6 \cdot t_o \cdot x_a^4 / C_a^5.$$

⁶ Dispersion parameter θ is 0.10.

$$\mathbf{J}_{12} = \mathbf{J}_m = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix}, \text{ and } \mathbf{J}_{21} = \mathbf{J}_m^T.$$

By substituting these elements into Equation (4.23), the sensitivities of path flow estimates to the observations on links 1, 2, 3, 5, and 7 are given in Table 4-7.

Table 4-7 Sensitivities of Path Flow Estimates to Link Observations

Path	Links	Link				
		1	2	3	5	7
1	1-4	1.0000	0.0000	-1.0000	-1.0000	0.0000
2	2-6	0.0000	0.7362	0.2899	0.0000	-0.5235
3	1-3-6	0.0000	0.2638	0.7101	0.0000	-0.4765
4	1-5	0.0000	0.0000	0.0000	1.0000	0.0000
5	2-7	0.0000	0.2638	-0.2899	0.0000	0.5235
6	1-3-7	0.0000	-0.2638	0.2899	0.0000	0.4765

The sensitivities computed above can be used to approximate the new values of path flows when there are small changes in link observations. The linear approximation is given by:

$$\mathbf{f}(\Delta \mathbf{v}) = \mathbf{f}(0) + \nabla_{\Delta \mathbf{v}} \mathbf{f}(0) \cdot \Delta \mathbf{v}. \quad (4.24)$$

Assume that link perturbations are about +5 percent of link measurements (simultaneously perturbed); $\Delta \mathbf{v} = [1.4700, 0.7800, 0.7100, 0.3350, 0.4250, 0.6650, 0.8250]$. Table 4-8 compares the estimated path flows using the linear approximation with the actual path flows re-estimated by PFE. It can be seen that the linear approximation can provide a reasonable estimate without the need to resolve PFE with the perturbed link flows. This can also be applied for the O-D flow estimates.

Table 4-8 Estimated Path Flows with Simultaneously Perturbed Link Observations

Path	Links	f(0)	f_actual	f_estimated
1	1-4	6.7000	7.0350	7.0350
2	2-6	6.9624	7.3105	7.3105
3	1-3-6	6.3376	6.6545	6.6545
4	1-5	8.5000	8.9250	8.9250
5	2-7	8.6376	9.0695	9.0695
6	1-3-7	7.8624	8.2555	8.2555

This sensitivity information obtained previously can be further used to determine the uncertainties of model outputs (outputs of PFE). These may include path flows, O-D flows, and the total demand estimated. If the variances of link observations (model inputs) are assumed to be the same as the means (Poisson distribution) and each observation is independent, the variance of aforementioned outputs can be computed as follows:

Path flows

$$\mathbf{Var}(\mathbf{f}) = \nabla_{\mathbf{v}} \mathbf{f} \cdot \mathbf{Var}(\mathbf{v}) \cdot \nabla_{\mathbf{v}} \mathbf{f}^T. \quad (4.25)$$

$$\mathbf{Var}(\mathbf{f}) = \begin{bmatrix} 52.1000 & -4.1159 & -10.0841 & -8.5000 & 4.1159 & -4.1159 \\ -4.1159 & 14.1688 & 10.0688 & 0.0000 & -2.6847 & -5.9529 \\ -10.0841 & 10.0688 & 11.9937 & 0.0000 & -5.9528 & -1.9096 \\ -8.5000 & 0.0000 & 0.0000 & 8.5000 & 0.0000 & 0.0000 \\ 4.1159 & -2.6847 & -5.9528 & 0.0000 & 6.8006 & 1.8370 \\ -4.1159 & -5.9529 & -1.9096 & 0.0000 & 1.8370 & 6.0255 \end{bmatrix}.$$

O-D flows

$$\mathbf{Var}(\mathbf{q}) = \nabla_{\mathbf{v}} \mathbf{q} \cdot \mathbf{Var}(\mathbf{v}) \cdot \nabla_{\mathbf{v}} \mathbf{q}^T, \quad (4.26)$$

where

$$\nabla_{\mathbf{v}} \mathbf{q} = \nabla_{\mathbf{f}} \mathbf{q} \cdot \nabla_{\mathbf{v}} \mathbf{f}. \quad (4.27)$$

The first term of Equation (4.27) is essentially the O-D path incidence matrix, $\mathbf{L}$, while the second term is the sensitivity information obtained earlier.

$$\mathbf{v} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

$$\nabla_{\mathbf{v}} \mathbf{q} = \begin{bmatrix} 1.00 & 1.00 & 0.00 & 1.00 & 1.00 \\ 0.00 & 0.00 & 0.00 & 1.00 & 1.00 \end{bmatrix}.$$

Then, the variance-covariance of O-D flow estimates using this set of independent observations is:

$$\mathbf{Var}(\mathbf{q}) = \begin{bmatrix} 70.00 & -25.00 \\ -25.00 & 25.00 \end{bmatrix}.$$

Total Demand of Network (D)

$$\mathbf{Var}(\mathbf{D}) = \nabla_{\mathbf{v}} \mathbf{D} \cdot \mathbf{Var}(\mathbf{v}) \cdot \nabla_{\mathbf{v}} \mathbf{D}^T, \quad (4.28)$$

where

$$\nabla_{\mathbf{v}} \mathbf{D} = \nabla_{\mathbf{q}} \mathbf{D} \cdot \nabla_{\mathbf{f}} \mathbf{q} \cdot \nabla_{\mathbf{v}} \mathbf{f}. \quad (4.29)$$

The first term of Equation (4.29) is a unity row vector of length |RS|, a total number of O-D pairs in the network. The second and the third terms are the same as explained previously.

$$\nabla_{\mathbf{v}} \mathbf{D} = \begin{bmatrix} 1 & 1 \end{bmatrix}.$$

$$\mathbf{Var}(\mathbf{D}) = \begin{bmatrix} 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 70.00 & -25.00 \\ -25.00 & 25.00 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 45.00.$$

5 DEVELOPMENT OF THE BI-LEVEL ESTIMATOR

For general networks, with both route choice and congestion, the proportional assignment assumption is no longer valid due to the dependency between route choice and O-D trip table (Bell and Iida, 1997). It then becomes necessary to incorporate a route choice model into the estimation process. This issue leads to an estimation approach based on bi-level programming, in which route choice proportions are endogenously determined while estimating the O-D trip table (Fisk, 1988; Fisk, 1989; Yang et al., 1992; Maher et al., 2001). In the bi-level programming approach, the upper-level problem uses one of the statistical techniques, such as the generalized least squares, maximum entropy, and maximum likelihood, etc. to select the most appropriate O-D trip table, whereas the lower-level problem determines route choice proportions (e.g., deterministic user equilibrium or stochastic user equilibrium) compatible with the estimation.

In general, the bi-level O-D estimation problem can be defined through the following mathematical program.

$$\text{Min}_q \quad \omega_1 \cdot F_1(q, \bar{q}) + \omega_2 \cdot F_2(x(q), \bar{v}) \quad (5.1)$$

subject to:

$$q \geq 0, \quad (5.2)$$

$$x = M(q), \quad (5.3)$$

where

$\bar{v}$ represents a vector of observed link flows on a subset of network links;

$\bar{q}$ denotes a vector of the target O-D demands;

$F_1(q, \bar{q})$ and $F_2(x, \bar{v})$ are functions of the generalized distances between target and estimated O-D demands, and observed and estimated link flows respectively;

ω_1 and ω_2 are the weights or beliefs given to different sources of data;

M represents the traffic assignment problem (lower-level problem), which is the relationship between link flows and O-D flows;

x is the solution of the traffic assignment problem (link flows) for any fixed q .

Most of bi-level programming methods explicitly account for the inconsistencies of traffic counts by removing the requirement of exact duplication. In other words, the deviations between the estimated and observed link flows are allowed. The difficulty of solving this bi-level program comes from the fact that we do not have the explicit functional form of $M(q)$. This makes a majority of the traditional optimization techniques unable to be applied directly because the feasibility of intermediate solutions cannot be ensured.

In practice, bi-level O-D estimation is solved using the iterative procedure. The upper and lower problems are iteratively optimized until the decisions of both levels are stable. Among the efficient techniques, the sensitivity analysis-based (SAB) algorithm might be most commonly

used and has been successfully applied to many bi-level transportation problems. The basic idea of the SAB algorithm is to use the sensitivity information derived at the current solution of the lower-level problem to approximate the responses of $M(q)$ when solving the upper-level problem. After this transformation, the upper-level problem can be solved by using any optimization technique suitable for the nature of its objective function (linear, nonlinear, or quadratic function).

This chapter is dedicated to the development of the bi-level estimator of which its estimation results will later be compared with the estimation results of PFE for validation purpose.

5.1 Lower Level Problem: Route Choice Models

Typically, the lower problem of the bi-level estimator is a network equilibrium problem that represents traveler route choice behavior. Up to now, the most popular route choice models applied are the deterministic user equilibrium, DUE (Beckmann et al., 1956) and the stochastic user equilibrium, SUE (Daganzo and Sheffi, 1977). DUE model is based on strong assumptions that all travelers are perfectly aware of the travel times on the network and are always capable of identifying the shortest travel time route. To relax the assumption that travelers possess the perfect knowledge of network travel times, some researchers (Daganzo and Sheffi, 1977; Fisk, 1980; Sheffi and Powell, 1982) have proposed different SUE models (i.e., DN-SUE), allowing travelers to select routes based on their perceived travel times. These two models are well documented and interested readers can refer to Sheffi (1985).

Traffic assignment is an essential and fundamental step in the transportation planning process. It predicts the vehicular flows on the transportation network by assigning travel demands, given in terms of an O-D trip table, to routes in a network according to some route choice model. The DUE model is perhaps the most widely used route choice model in transportation planning practices. This UE route choice model is based on the well-known Wardrop's First Principle, which states:

“The journey times on all the routes actually used are equal, and less than those which would be experienced by a single vehicle on any unused route” (Wardrop, 1952).

The criterion used in the route choice decision process is to minimize user travel time. All travelers are assumed to be rational, possess perfect knowledge of network travel times, and are capable of identifying the minimum travel time route. Given these assumptions, all travelers make route choices that result in a stable equilibrium traffic flow pattern, such that, there is no incentive for anyone to change his/her route (Sheffi, 1985).

The assumptions of perfect knowledge of network travel times and the ability to always identify the minimum travel time route are rather unrealistic, since travelers do not always have correct perception of the network travel times, rarely are all possible routes in the network known, and certainly users do not always select the route based solely on the minimum travel time criterion. To relax some of these assumptions, a random error term is incorporated in the route choice decision process to simulate travelers' imperfect perceptions of travel times such that they do not

always choose the minimum travel time route. The random error term here is interpreted as the perception error of network travel times due to the travelers' imperfect knowledge of network conditions. 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* route travel time, which can be obtained by adding up the *perceived* travel times of all links belonging to the route. This extension of the UE model is known as the SUE model and has the following definition.

“At SUE, no motorists can improve his or her *perceived* travel time by unilaterally changing routes” (Daganzo and Sheffi, 1977; Sheffi, 1985).

Route choice models proposed under this approach can have different specifications according to the modeling assumptions on the random error term. The two 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 model has a closed-form probability expression and an equivalent mathematical programming formulation (Fisk, 1980), which can be solved using both path enumeration techniques (Ben-Avika et al., 1984; Cascetta et al., 1997, 2002) and column generation techniques (Bell et al., 1993; Maher, 1998). The logit-based SUE model can be implemented in the link- or path-based domains. The link-based algorithms do not require path storage and often use Dial's STOCH algorithm (Dial, 1971) or Bell's alternative (1994) as the stochastic loading step, while the path-based algorithms require explicit path storage in order to directly compute the logit route choice probabilities. The drawbacks of the logit model are: (1) inability to account for overlapping (or correlation) among routes and (2) 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 model, on the other hand, does not have such drawbacks, since it handles the overlapping and identical variance problems between routes 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). Recently, Rosa and Maher (2002) extended the probit link-based SUE method to a path-based method that can handle one or more user classes. Other specifications of the random error term include uniform (Burrell, 1968), gamma (Bovy and Stern, 1990), and lognormal (Von Falkenhausen, 1966; Cantarella and Binetti, 1998).

Recently, there are renewed interests to improve the logit-based route choice model due to the importance of the route choice model in Intelligent Transportation Systems (ITS) applications, particularly the applications of advanced traveler information systems (ATIS). Several modifications or generalizations of the logit structure have been proposed to relax the IID assumptions of the logit model. Among these extended logit models include the C-logit (Cascetta et al., 1997), path-sized logit (Ben-Akiva and Bierlaire, 1999), and paired combinatorial logit (Bekhor and Prashker, 1999) model, which all aim to resolve the overlapping problem while maintaining the analytical tractability of the logit choice probability function. For

all these models, with the exception of the paired combinatorial logit (PCL), the correction term to the utility of alternative routes are included. This correction term will reduce the perceived utility of a particular route when it highly overlaps with the others. In the PCL model, each pair of alternative routes can have a similarity relationship that is completely independent of the similarity relationship of other pairs of alternatives.

5.1.1 Deterministic User Equilibrium

The DUE flow pattern can be obtained by solving the following program.

$$\text{Minimize: } Z_1 = \sum_a \int_0^{x_a} t_a(w) dw. \quad (5.4)$$

Subject to:

$$\sum_{k \in K_{rs}} f_k^{rs} = q_{rs}, \quad \forall r, s, \quad (5.5)$$

$$x_a = \sum_{rs} \sum_k f_k^{rs} d_{ka}^{rs}, \quad \forall a, \quad (5.6)$$

$$f_k^{rs} \geq 0, \quad \forall k, r, s. \quad (5.7)$$

The first-order condition of this program given below shows that travel times (c_k^{rs}) of paths carrying flow (positive flow) are equal to the minimum O-D travel time (p_{rs}) while those of paths not carrying flow are greater than or equal to p_{rs} . p_{rs} is equivalent to the dual variable of the flow conservation constraint, equation 5.5. Therefore, the solution of this mathematical program corresponds to the equilibrium condition in which no traveler can unilaterally switch routes to improve their travel time. At the optimality,

$$f_k^{rs} (c_k^{rs} - p_{rs}) = 0, \quad \forall k, r, s, \quad (5.8)$$

$$c_k^{rs} - p_{rs} \geq 0, \quad \forall k, r, s \quad (5.9)$$

Equations (5.5) and (5.7).

5.1.2 Logit-Based Stochastic User Equilibrium

Logit Model (Fisk, 1989)

$$\text{Minimize: } Z = Z_1 + Z_2 \quad (5.10)$$

$$Z_1 = \sum_a \int_0^{x_a} t_a(w) dw, \quad (5.10a)$$

$$Z_2 = \frac{1}{q} \sum_{rs} \sum_k f_k^{rs} \ln f_k^{rs}, \quad (5.10b)$$

subject to:

Equations (5.5) – (5.7).

It can be shown that the optimality condition of this mathematical program provides the logit probability expression of path flow as follows:

$$P_k^{rs} = \frac{\exp(-q \cdot c_k^{rs})}{\sum_{\forall j \in K_{rs}} \exp(-q \cdot c_j^{rs})}, \forall k, r, s. \quad (5.11)$$

To properly account for the dependency among alternative routes, several modifications or generalizations of the logit structure have been proposed. The correction term, which represents the correlation between paths (overlapping), has been included to the utility of alternative routes. This minor modification maintains the analytical tractability of the logit model and according to Cascetta et al. (1997), the following correction term, so-called commonality factor, can be introduced.

$$CF_k = b_0 \ln \sum_{j \in K_{rs}} \left(\frac{L_{kj}^{rs}}{\sqrt{L_k^{rs} \cdot L_j^{rs}}} \right)^g, \forall k, r, s, \quad (5.12)$$

where L_{kj}^{rs} is the length of paths k and j have in common. β_0 and γ are the parameters to be calibrated. The value of commonality factor is zero in the case of independent paths (no overlapping; hence, equivalent to logit model). If the commonality of paths grows larger (highly overlapped), the utility of path will greatly reduce, thus becoming less attractive. The probability expression of the C-logit model is given by:

$$P_k^{rs} = \frac{\exp(-q \cdot c_k^{rs} - CF_k)}{\sum_{\forall j \in K_{rs}} \exp(-q \cdot c_j^{rs} - CF_j)}, \forall k, r, s. \quad (5.13)$$

Later, Ben-Akiva and Bierlaire (1999) proposed another form of correction term considering the contribution of a link to all relevant paths (path-size). The correction term is given by:

$$PS_k^{rs} = \sum_{a \text{ on } k} \left(\frac{l_a}{L_k^{rs}} \right) \cdot \left(\frac{1}{\sum_{\forall j \in K_{rs}} d_{ka}^{rs}} \right), \forall k, r, s, \quad (5.14)$$

which leads to the following probability expression (path-size logit).

$$P_k^{rs} = \frac{PS_k^{rs} \cdot \exp(-q \cdot c_k^{rs})}{\sum_{\forall j \in K_{rs}} PS_j^{rs} \cdot \exp(-q \cdot c_j^{rs})}, \forall k, r, s. \quad (5.15)$$

However, there is no equivalent mathematical formulation given for either one of these extended logit models until recently (Zhou and Chen, 2003). Based on the probability expression (assumed to hold at the optimality as for the logit model), the mathematical formulation of C-logit model can be derived backward as shown below. The existence of a mathematical program will benefit the development of efficient path-based solution algorithms. The mathematical formulation of the path-size logit can be obtained in a similar way.

C-logit Model

$$\text{Minimize: } Z = Z_1 + Z_2 \quad (5.16)$$

$$Z_1 = \sum_a \int_0^{x_a} t_a(w) dw, \quad (5.16a)$$

$$Z_2 = \frac{1}{q} \sum_{rs} \sum_k f_k^{rs} (\ln f_k^{rs} + CF_k), \quad (5.16b)$$

subject to:

Equations (5.5) – (5.7).

Path-size Logit Model

$$\text{Minimize: } Z = Z_1 + Z_2 \quad (5.17)$$

$$Z_1 = \sum_a \int_0^{x_a} t_a(w) dw, \quad (5.17a)$$

$$Z_2 = \frac{1}{q} \sum_{rs} \sum_k f_k^{rs} (\ln f_k^{rs} - \ln PS_k), \quad (5.17b)$$

subject to:

Equations (5.5) – (5.7).

Paired combinatorial Logit (Bekhor and Prashker, 1999)

The paired combinatorial logit (PCL) model proposed by Chu (1989), was further developed to demonstrate the structure, properties, and estimation by Koppleman and Wen (2000), and adapted to model route choice by Bekhor and Prashker (1999) and Gliebe et al. (1999). Contrasts to the simple structure of the logit model, the PCL model has a hierarchical structure that decomposes the choice probability into two levels represented by the marginal, $P(kj)$, and conditional, $P(k/kj)$, probabilities. Thus, the PCL choice probability can be expressed as:

$$P(k) = \sum_{j \neq k} P(kj) \cdot P(k / kj), \quad (5.18)$$

The PCL equivalent mathematical formulation presented here is due to Bekhor and Prashker (1999). They show that the marginal and conditional probabilities of the PCL model can be represented by two entropy terms in the objective function as follows:

$$\text{Minimize:} \quad Z = Z_1 + Z_2 + Z_3 \quad (5.19)$$

$$Z_1 = \sum_a \int_0^{x_a} t_a(w) dw, \quad (5.19a)$$

$$Z_2 = \frac{1}{q} \sum_{rs} \sum_k \sum_{j \neq k} b_{kj} f_{k|kj}^{rs} \ln \left(\frac{f_{k|kj}^{rs}}{b_{kj}} \right), \quad (5.19b)$$

$$Z_3 = \frac{1}{q} \sum_{rs} \sum_{k=1}^{|K_{rs}|-1} \sum_{j=k+1}^{|K_{rs}|} (1 - b_{kj}) (f_{k|kj}^{rs} + f_{j|kj}^{rs}) \ln \left(\frac{f_{k|kj}^{rs} + f_{j|kj}^{rs}}{b_{kj}} \right), \quad (5.19c)$$

subject to:

$$\sum_k \sum_{j \neq k} f_{k(kj)}^{rs} = q^{rs}, \quad \forall r, s, \quad (5.20)$$

Equations (5.6) – (5.7).

$f_{k(kj)}^{rs}$ is the flow on route k (of the route pair kj) between origin r and destination s . β_{kj} is a measure of dissimilarity between paths k and j , while θ is a dispersion parameter accounting for perception error of travelers. When the dissimilarity is one, this formulation collapses to Fisk's formulation. In addition, $f_{k|kj}^{rs}$ becomes f_k^{rs} because the selection of path pair kj and path k is independent. The PCL model accounts for the overlapping problem in route choice by incorporating a similarity index into the objective function (i.e., Z_2 and Z_3). This enables the PCL model to treat the similarity effect separately from the congestion effect (Z_1) and the stochastic effect (Z_2 in Fisk's formulation). Hence, the PCL SUE formulation is theoretically capable of accounting for congestion, stochastic, and similarity effects.

5.1.3 Path Based Algorithms for Solving Traffic Assignment Problem

In practice, the link-based algorithms, which do not require path storage, have been used to solve the traffic assignment problems: the well-known Frank-Wolfe (FW) algorithm for DUE and the efficient Dial's STOCH algorithm (Dial, 1971) or Bell's alternative (1994) for SUE. Their popularities come from their modest memory requirement when dealing with large-scale networks and from their simplicity. However, this marked advantage also excludes planners from analyzing the useful path flow solutions, which are becoming increasingly important in certain transportation applications such as optimal routing in route guidance system,

environmental impact studies, fuel consumption, or even O-D trip table estimation, etc. Given the computing and storage powers of current computer technology, storing paths of traffic assignment models in practice becomes possible.

Notation

n	Iteration counter
x_a	Flow on link a
y_a	Auxiliary flow on link a
$t_a(x_a)$	Travel time of link a
l_a	Length of link a
q_{rs}	Demand between origin r and destination s
d_{ka}^{rs}	Path-link incidence indicator
$\overline{k}_{rs}$	Shortest path between origin r and destination s
K_{rs}	Path set between origin r and destination s
f_k^{rs}	Flow on route k from origin r and destination s
$f_{k(kj)}^{rs}$	Flow on route k (of the route pair kj) between origin r and destination s
h_k^{rs}	Auxiliary flow on route k from origin r and destination s
a	Step size
e	Tolerance
E	Maximum percentage change in link flow
q	Dispersion coefficient
b_{kj}	Measure of dissimilarity between paths k and j
$ K_{rs} $	Number of paths between origin r and destination s

The solution procedure given below can be applied for all traffic assignment models presented earlier. The only difference is the direction finding step in which travel demands are loaded onto the network either deterministically or probabilistically. This step presents different route choice behaviors defined by the objective functions of their mathematical programs.

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

1. Set $n = 0$, $x_a^n = 0$, $t_a^n = t_a(x_a^n)$, $\forall a$, and $K_{rs} = f$, $\forall r, s$
2. Set iteration counter $\Rightarrow n = 1$

3. Solve the shortest path problem for all origins $\overline{k_{rs}^n}$, and $K_{rs} = K_{rs} \cup \overline{k_{rs}^n}, \forall r, s$
4. Perform All-or-Nothing traffic assignment $\Rightarrow f_k^{rs}(n) = q_{rs}, \forall k, r, s$
5. Assign path flows to links $\Rightarrow x_a^n = \sum_{rs} \sum_{k \in K_{rs}} f_k^{rs}(n) d_{ka}^{rs}, \forall a$

Step 1. Direction Findings. Compute auxiliary link and path flows.

6. Increase iteration counter $\Rightarrow n = n + 1$
7. Update link travel time $\Rightarrow t_a^n = t_a[x_a^{n-1}], \forall a$
8. Solve the shortest path problem $\Rightarrow \overline{k_{rs}^n}, \forall r, s$.
9. Determine whether $\overline{k_{rs}^n}$ exist in the path set K_{rs} or not,
 If $k_{rs}^n \notin K_{rs}$, then $K_{rs} = K_{rs} \cup \overline{k_{rs}^n}, \forall r, s$
10. Update path costs $\Rightarrow c_k^{rs}(n) = \sum_a t_a^n d_{ka}^{rs}, \forall k, r, s$
11. Compute route choice probabilities, and auxiliary path flows
 - a. *User equilibrium (UE)*
 - b. *Multinomial logit (MNL)*
 - c. *C-Logit*
 - d. *Path size logit*
 - e. *Paired combinatorial logit (PCL)*

Note: See the detailed procedure for each method in the next section.

$$h_k^{rs} = P_k^{rs}(n) \cdot q_{rs}, \forall k, r, s.$$

12. Assign path flows to obtain auxiliary links $\Rightarrow y_a^n = \sum_{rs \in RS} \sum_{k \in K_{rs}} h_k^{rs} d_{ka}^{rs}, \forall a$

Step 2. Line Search.

13. Perform line search to determine step size, α .
 - a. *Method of successive average (MSA)*
 - b. *Bisection*
 - c. *Quadratic Interpolation*
 - d. *Armijo's rule*

Note: See the detailed procedure for each method in the subsequent section.

Step 3. Update Solutions.

14. Update link flows $\Rightarrow x_a^n = x_a^{n-1} + a[y_a^n - x_a^{n-1}], \forall a$.

15. Update path flows $\Rightarrow f_k^{rs}(n) = f_k^{rs}(n-1) + a[h_k^{rs}(n) - f_k^{rs}(n-1)], \forall k, r, s$.

Step 4. Convergence Testing.

16. Determine maximum percentage change of flow differences.

$$E = \max_a \left| \frac{x_a^n - x_a^{n-1}}{x_a^{n-1}} \right| \times 100$$

17. Check the convergence, if $E < \epsilon$ (e.g., 10^{-6}), then stop; otherwise go to step 1.

5.1.4 Direction Finding (item 11)

a. User Equilibrium (UE)

$$P_k^{rs}(n) = 1.0, \forall k \in \bar{k}_{rs}^n, r, s$$

$$P_k^{rs}(n) = 0.0, \forall k \in K_{rs}, \text{ and } k \notin \bar{k}_{rs}^n, \forall r, s$$

b. Multinomial Logit (MNL-SUE)

$$P_k^{rs}(n) = \frac{\exp(-q \cdot c_k^{rs}(n))}{\sum_{\forall j \in K_{rs}} \exp(-q \cdot c_j^{rs}(n))}, \forall k, r, s$$

c. C-Logit (Cascetta et al., 1997)

a. Compute the commonality factor (CF_k) of any path k :

$$CF_k = b_0 \ln \sum_{j \in K_{rs}} \left(\frac{L_{kj}^{rs}}{\sqrt{L_k^{rs} \cdot L_j^{rs}}} \right)^g, \forall k, r, s,$$

b. Compute route choice probabilities:

$$P_k^{rs}(n) = \frac{\exp(-q \cdot c_k^{rs}(n) - CF_k)}{\sum_{\forall j \in K_{rs}} \exp(-q \cdot c_j^{rs}(n) - CF_j)}, \forall k, r, s.$$

d. **Path-size Logit** (Ben-Akiva and Bierlaire, 1999)

1. Compute size (PS_k) of any path k :

$$PS_k^{rs} = \sum_{a \text{ on } k} \left(\frac{l_a}{L_k^{rs}} \right) \cdot \left(\frac{1}{\sum_{\forall j \in K_{rs}} d_{ka}^{rs}} \right), \forall k, r, s.$$

2. Compute route choice probabilities:

$$P_k^{rs}(n) = \frac{PS_k^{rs} \cdot \exp(-q \cdot c_k^{rs}(n))}{\sum_{\forall j \in K_{rs}} PS_j^{rs} \cdot \exp(-q \cdot c_j^{rs}(n))}, \forall k, r, s.$$

e. **Paired Combinatorial Logit (PCL)** (Bekhor and Prashker, 1999)

1. Compute similarity and dissimilarity indices between paths k and j :

$$s_{kj}^{rs} = s_{jk}^{rs} = \frac{L_{kj}^{rs}}{\sqrt{L_k^{rs} \cdot L_j^{rs}}}, \quad (5.21)$$

$$\text{and set } \beta_{kj}^{rs} = 1 - s_{kj}^{rs}, \forall k, j = k+1, \dots, |K_{rs}(n)|, r, s. \quad (5.22)$$

2. Compute route choice probabilities:

$$P_{kj}^{rs}(n) = \frac{b_{kj} \left[\exp\left(-\frac{q \cdot c_k^{rs}(n)}{b_{kj}}\right) + \exp\left(-\frac{q \cdot c_j^{rs}(n)}{b_{kj}}\right) \right]^{b_{jk}}}{\sum_{m=1}^{|K_{rs}|-1} \sum_{l=m+1}^{|K_{rs}|} b_{ml} \left[\exp\left(-\frac{q \cdot c_m^{rs}(n)}{b_{ml}}\right) + \exp\left(-\frac{q \cdot c_l^{rs}(n)}{b_{ml}}\right) \right]^{b_{ml}}}. \quad (5.23)$$

$$P_{k|kj}^{rs}(n) = \frac{\exp\left(-\frac{q \cdot c_k^{rs}(n)}{b_{kj}}\right)}{\exp\left(-\frac{q \cdot c_k^{rs}(n)}{b_{kj}}\right) + \exp\left(-\frac{q \cdot c_j^{rs}(n)}{b_{kj}}\right)}. \quad (5.24)$$

$$P_k^{rs}(n) = \sum_{j \neq k} P_{kj}^{rs}(n) \cdot P_{k|kj}^{rs}(n)$$

5.1.5 Line Search or Step size Selection (item 13)

a. Method of successive average (MSA)

$\alpha = 1/n$, where n is the iteration counter.

b. Bisection

The step size (α) is obtained by solving the following optimization problem.

$$\min_{0 \leq a \leq 1} Z(a) = Z\{x + a(y - x), f + a(h - f)\}^7. \quad (5.26)$$

Step 1. Set $a = 0$, and $b = 1$.

Step 2. Set $\alpha = (a + b)/2$, and compute the first derivative of Z with respect to α , $\nabla Z_a(a)$.

Step 3. If $\nabla Z_a(a) \leq 0$, then set $a = \alpha$, otherwise set $b = \alpha$.

Step 4. If the difference between a and b is smaller than the specified tolerance (e.g., 10^{-5}), set $\alpha = (a + b)/2$ and stop, otherwise go to step 2.

c. Quadratic interpolation

Step 1. Set $A = \nabla Z_a(a = 0)$, and $B = \nabla Z_a(a = 1)$.

Step 2. Approximate the step size based on: $\alpha = -A/(-A + B)$.

d. Armijo's rule

Step 1. Select the scalars $s > 0$, $\beta \in \{0, 1\}$, and $s \in \{0, 0.5\}$. Set $m = 0$.

Step 2. Set $m = m + 1$.

Step 3. Set $a = b^m s$. If the following condition is satisfied, then stop; otherwise go to step 2.

$$Z(x, f) - Z(x + a(y - x), f + a(h - f)) \geq -s \cdot a \cdot \nabla Z_a(a = 0)(y - x, h - f). \quad (5.27)$$

⁷ See the functional form of Z for each route choice model presented earlier.

5.2 Upper Level Problem: Statistical Inferences

The upper level problem, of the bi-level estimator, is the conventional trip table estimation problem in which the link use portion is assumed given. The problem is to find the trip table that is least deviated from the target trip table and able to produce the small differences between estimated link flows and observed link flows defined by some distance measures (e.g., Euclidian distance). Different specifications of the distance measure lead to different types of models.

5.2.1 Maximum Entropy Estimator

For this family of models, the entropy function can be used for the measure of distance or error. F_1 and F_2 defined at the beginning of this chapter are given by:

$$F_1(q, \bar{q}) = \sum_{rs} q_{rs} \cdot \left(\log \left(\frac{q_{rs}}{\bar{q}_{rs}} \right) - 1 \right), \quad (5.28)$$

$$F_2(x, \bar{v}) = \sum_{a \in M} x_a \cdot \left(\log \left(\frac{x_a}{\bar{v}_a} \right) - 1 \right). \quad (5.29)$$

With the assumption of constant link-use proportion, any solution techniques for convex programming can be applied to obtain the best combination of q and x . It should be noted that the combination of q and x is best only for this fixed link-use proportion.

5.2.2 Generalized Least Square

This model minimizes the weighted sum of Euclidian distances (L_2 norm) between estimated and observed values. F_1 and F_2 are given by:

$$F_1(q, \bar{q}) = (\bar{q} - q)^T Q^{-1} (\bar{q} - q), \quad (5.30)$$

$$F_2(x, \bar{v}) = (\bar{v} - x(q))^T V^{-1} (\bar{v} - x(q)), \quad (5.31)$$

where Q and V are the variance-covariance matrices for target O-D flows and observed link flows respectively. Oftentimes, these two matrices can be used to represent the weights or beliefs of these observations. If the non-negativity constraints, equation 5.2, are assumed inactive and the link-use proportion is fixed, the solution to this problem (q) can be obtained analytically as given below. This is one practical advantage of this model since it requires only matrix manipulation and inversion to obtain the O-D flow estimates.

$$q = (Q^{-1} + M^T V^{-1} M)^{-1} \cdot (Q^{-1} \bar{q} + M^T V^{-1} \bar{v}). \quad (5.32)$$

M represents the relationship between O-D flow and link flow estimates, link-use proportion, which can be obtained by traffic assignment. There are evidences that the non-negativity constraints are occasionally violated (Bell, 1991) and they may have an impact on the quality of estimates. Thus, it is necessary to consider the non-negativity constraints into the estimation. However, the inclusion of non-negativity constraints prevents the existence of a closed form solution, as shown previously. The iterative balancing must be applied to remove the negative O-D flows by adjusting the associated Lagrange multipliers (m).

$$q = (Q^{-1} + M^T V^{-1} M)^{-1} \cdot (Q^{-1} \bar{q} + M^T V^{-1} \bar{v} + m). \quad (5.33)$$

If constraints are inactive, the values of Lagrange multipliers remain zero and equation 5.33 collapses to equation 5.32. If some of constraints are violated, the values of corresponding multipliers need to be adjusted so that those constraints are satisfied.

5.3 Solution Techniques

5.3.1 Iterative Procedure

The algorithms for solving the bi-level program are generally based on the concept of iterative procedure. The upper and lower problems are iteratively solved until the decisions of both levels are stable. When the upper-level problem is solved, the solution of lower-level problem can either be fixed or readjusted along with the upper-level decisions according to the so-called influence factors, which represent changes of lower-level decision variables when upper-level decision variables change. It was observed that these two strategies might end up with solutions that have different characteristics (see more discussion on this issue in Yang and Bell, 1998). Among the iterative techniques, the SAB algorithm might be the most commonly used techniques and has been successfully applied to many bi-level transportation applications, including some of the studies mentioned earlier (Wong and Yang, 1997; Gao and Song, 2002). The basic idea of the SAB algorithm is to use the sensitivity information derived at the current solution of the lower-level problem to approximate the responses of nonlinear constraints, equation 5.3, when solving the upper-level problem. After this transformation, the upper-level problem can be solved using any optimization technique suitable for its nature. Despite some successes of the SAB algorithm, there is evidence that the size of most realistic problems may cause SAB algorithm to be computationally inefficient (e.g., involvement of matrix inversion) and, more importantly, unable to find the global optimum. As reported by Gao and Song (2002), the effectiveness of the algorithm is highly dependent on the initial solution when the size of the problem (number of decision variables) is somewhat large. In addition, when there are several decision variables involved, the problem is unlikely uni-modal. Although the convergence of the algorithm has been witnessed, the global optimality cannot be guaranteed with certainty.

In the context of O-D trip table estimation, the iterative procedure can be summarized as follows.

Upper Level: Solve for q using suitable optimization technique, or using equation 5.32 for the generalized least square model.

Lower Level: With q from the upper level, evaluate the link-use proportion (M) or the response of link flows with respect to O-D flows.

5.3.2 Genetic Algorithms

Due to practical issues, mentioned earlier for the iterative techniques, a Genetic Algorithms (GA) based approach has been proposed as one of the alternatives for solving the transportation problem with a bi-level structure (Yin, 2000). GA is a powerful technique developed by John Holland (1975) over the course of 1960's and 1970's. David Goldberg provided significant contributions that increased the popularity of this algorithm, since he was able to solve a difficult problem involving the control of gas-pipeline transmission as his dissertation (Goldberg, 1989). GA is the stochastic search that mimics the process of natural selection and when compared with the iterative approach, GA has the following advantages.

- It does not require derivative information and does not require the problem to process certain mathematical properties, such as continuity, differentiability, uni-modal, convexity, etc., which may not be satisfied by many real-world problems.
- It simultaneously searches from multiple starting points and, most likely, is able to locate the optimal solution.
- It optimizes a large number of parameters with extremely complex cost surfaces (objective function) and it can jump out of a local minimum with the help of its operator.
- It can provide a set of optimum parameters, not just a single solution (if any).

In the context of O-D estimation, GA searches for the optimal trip table by the aid of genetic operators including reproduction, crossover, and mutation. They stochastically manipulate a set of feasible trip tables, referred to as the population, in order to explore the domain of good solutions. The population consists of the number of coded strings, which represent solution vectors of the problem (O-D demands to be estimated). Regarding its fit to the problem, each chromosome (O-D estimate) has a different chance to contribute its traits to the next set of trial solutions; a so-called offspring. The fitness of each chromosome can be evaluated through the objective function of the problem (objective function of the upper level problem).

The existing solutions can be manipulated to generate new solutions by the aid of genetic operators. Similar to the mathematical operators, these are used to manipulate the value of the variables to generate new solutions from the current populations. These GA operators are briefly

described below (see Goldberg; 1989, and Gen and Cheng; 2000 for more detailed description of GA operators).

Reproduction selects the chromosomes from a population pool on the basis of their fitness for mating purposes. The fitness of a chromosome implies the number of times that each chromosome will be in the mating pool. The most commonly used selection schemes are roulette wheel and tournament selection. For the roulette wheel selection, each chromosome is characterized as a segment in the roulette wheel. The fitter chromosome occupies a larger portion of the roulette wheel, reflecting a higher chance to be selected. A number between 0 and 1 is randomly generated to determine the selection of chromosomes. For the tournament selection, instead of using a uniform random number to identify which chromosome is being selected, two chromosomes are randomly compared and the fitter chromosome enters the mating pool

Crossover is the means of exchanging genetic materials between two parent chromosomes in order to produce two new offspring, each of which inherits some parts of genetic material from their parents. Crossover occurs with a constant probability (P_c), which implicitly indicates the expected number of chromosomes in the mating pool undergoing crossover. The approach to perform crossover can be as simple as a single point crossover, the slightly more complicated multi-point crossover, or a uniform crossover. For the single point crossover, a random crossover point is selected to exchange the genetic materials of the two parents in order to produce two new offspring. For the multi-point crossover, several successive crossover points are chosen. The genetic materials between these points are exchanged between the two parents. For the uniform crossover, a crossover mask, which consists of 0s and 1s and has the same length as a chromosome, determines the positions between the parent chromosomes to exchange genetic materials. The value of each bit in the crossover mask indicates the parent who will supply genetic material to the offspring.

Mutation is the means of alternating the value of genetic units, hopefully, to introduce new genetic material to the population. All new offspring are subjected to the mutation operator with a predefined mutation rate (P_m). Genes are then randomly selected to change their values. Mutation allows GA to explore new regions of the solution space and prevents premature convergence to a sub-optimal solution.

In the context of O-D estimation, GA can be applied to estimate the most appropriate O-D trip table as follows:

Step 0. Define the parameters of GA (population size, P_c , P_m , etc.)

Step 1. Generate vectors of O-D demands of which the value of each element falls within the feasible range defined by the lower and upper bounds of O-D demands. The number of O-D demand vectors is equal to the population size.

Step 2. Evaluate the fitness of each O-D vector by performing traffic assignment (e.g., DUE or SUE). Compute the distance measure (using equations presented earlier either entropy

function or generalized least square) between the estimated and observed values. This value defines the fitness of the solutions.

Step 3. Perform GA operators to generate the new O-D vectors.

Step 4. Check the stopping criterion (e.g., number of generation). If it is satisfied, stop; otherwise go to step 2.

6 DATA COLLECTION

6.1 Network Data

The associated network data were extracted from the Orange County Transportation Analysis Model (OCTAM). The OCTAM 3.0 has 2,940 traffic analysis zones (TAZ) within Orange County and 1,282 external stations (including cordon stations). The extracted network, as depicted in Figure 6-1, is composed of 163 nodes (67 nodes are the TAZ), 496 links, and 1,547 O-D pairs. There are 39 internal nodes (TAZ) for which there exist the one-to-one correspondences with OCTAM internal zones, as shown in Table 6-1. Nodes 40 through 67 were created as the external stations where the OCTAM network was cut. This network data was converted into the format required by PFE.

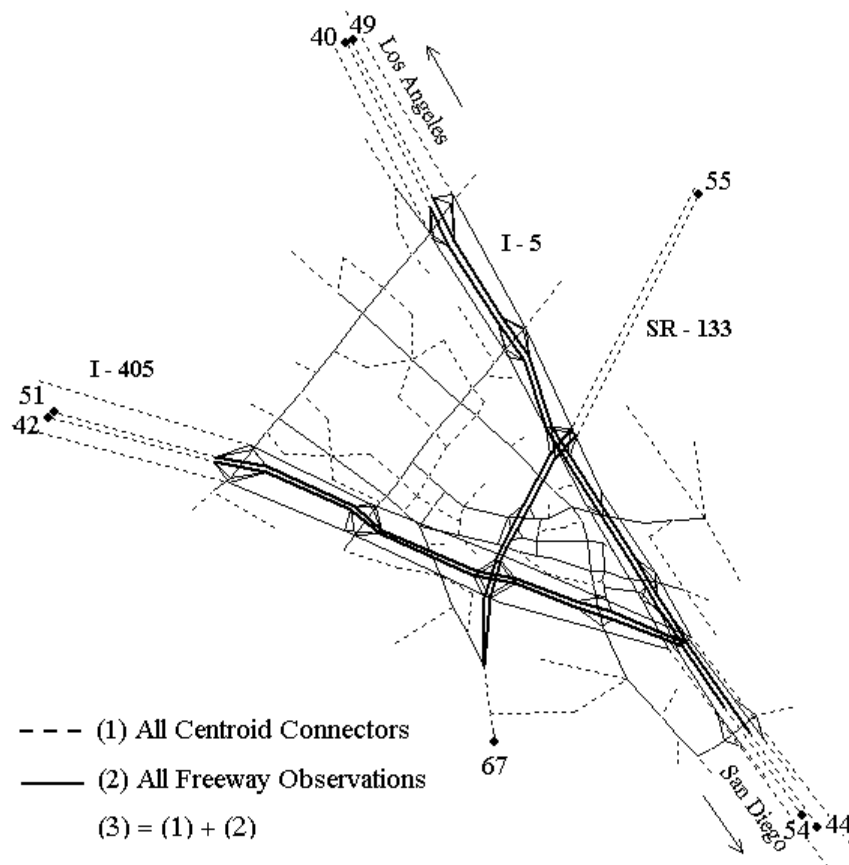


Figure 6-1 Irvine Network, Orange County, California.

Table 6-1 Node Numbering Systems between OCTAM and PFE

Original OCTAM Zone ID (Centroid)	PFE Zone ID	Original OCTAM Zone ID (Centroid)	PFE Zone ID
1298	1	1328	21
1297	2	1326	22
1314	3	1332	23
1315	4	1115	24
1318	5	1118	25
1320	6	1336	26
1322	7	1337	27
1114	8	1338	28
1323	9	1333	29
1321	10	1339	30
1319	11	1259	31
1317	12	1340	32
1316	13	1334	33
1324	14	1335	34
1325	15	1121	35
1327	16	1120	36
1329	17	1123	37
1116	18	1122	38
1331	19	1341	39
1330	20		

6.2 O-D Demand Data From OCTA Model

In addition, the associated demand data were extracted from the OCTA model, which contains data for the whole county. The OCTA model incorporates the best “state-of-the-practice” modeling components that are consistent with the new Southern California Regional Transportation Model recently released by the Southern California Association of Governments (SCAG). The OCTAM 3.0 was developed and validated for Base Year 1991 conditions, and re-validated for the Year 1998 to better reflect the current highway and transit data. The OCTAM 3.0 model incorporates a new sophisticated set of mode choice models consistent with those now incorporated in the new SCAG regional model. The five different trip purposes include:

1. Home-based work
2. Home-based school
3. Home-based other
4. Work-based other
5. Other-based other

All five trip-purposes in the COTAM have the same mode choice nesting structure. Choice between toll and free highway modes are also incorporated to reflect the opportunity to travel by

the toll facilities in Orange County, which include the SR-91 express lanes, the San Joaquin Transportation Corridor, the Foothill Transportation Corridor and the Eastern Transportation Corridor. The OCTAM splits demand into four time categories, they are:

1. 6 am – 9 am
2. 9 am – 3 pm
3. 3 pm – 7 pm
4. 7 pm – 6 am

6.3 Traffic Counts

For the experiments conducted in several previous chapters, traffic counts for the Irvine network were generated by assigning travel demands from OCTAM to the network according to the SUE principle with 0.01 of dispersion parameter. Only the PM peak (3 pm – 7 pm) demand was considered. The experiments were developed using different sets of synthetic traffic counts (see for example Figure 6-1).

In addition, the actual traffic counts on the surface streets in the City of Irvine were obtained during January and February 2003. They include turning counts at 29 intersections as listed in Table 6-1. All intersections are numbered sequentially, as in the table, and depicted in the original map as shown in Figure 6-2. The intersections with a number are those with traffic counts and the intersections with the question mark (?) are those that may need additional data. However, these traffic counts have yet to be used for estimation purposes.

6.4 Travel Time Patterns

Travel time patterns collected from MOU 4120 (GPS/GIS Technologies for traffic surveillance and management: A Testbed Implementation Study) will be used to validate path travel estimated by PFE.

Table 6-2 List of Intersections with Traffic Counts

No.	Symbol	Description	Date	Time
1	SCA	Sand Canyon and Water Works Way	01/28/03	10:57-11:52
2	BSC	Sand Canyon and Barranca	01/28/03	7:23-8:18
3	SCIC	Sand Canyon and Irvine Center Drive	01/28/03	7:00-x:xx
4	SCOC	Sand Canyon and Oak Canyon	01/30/03	6:57-7:52
5	JNB405	Jeffrey and NB 405	01/30/03	19:08-20:03
6	AG	Alton and Gateway	01/31/03	7:58-8:43
7	JTM	Jeffrey and The Meadows	01/31/03	7:53-x:xx
8	BJ	Jeffrey and Barranca	02/03/03	7:00-8:00
9	UM	University / Michalson	02/03/03	7:02-7:57
10	AA	Alton and Ada	02/04/03	8:13-8:58
11	ATW	Alton and Technology	02/04/03	6:56-7:41
12	ICG	Irvine Center Drive and Gateway	02/05/03	6:50-7:45
13	ANB5	Alton and NB I-5	02/05/03	8:00-8:55
14	SCHOS	Sand Canyon and Hospital	02/06/03	6:49-7:49
15	AS	Alton and Sand Canyon	02/06/03	7:29-8:24
16	ICH	Irvine Center Drive and Hubble & Lion City	02/06/03	7:00-7:55
17	AM	Alton and Muirlands / Barranca	02/10/03	7:37-8:32
18	ICJ	Jeffrey and Irvine Center Drive	02/14/03	8:07-8:52
19	AM	Alton and Barranca / Muirlands	02/17/03	7:45-8:10
20	ATE	Alton and Technology East	02/17/03	8:16-9:11
21	ICDC	Irvine Center Drive and Charter	02/17/03	7:29-8:24
22	ICB	Irvine Center Drive and Barranca	02/17/03	7:40-8:25
23	USB405	Jeffrey / University and 405 SB	02/18/03	8:01-8:56
24	JQ	Jeffrey and Qual Creek	02/20/03	7:55-8:49
25	AE	Alton and Enterprise	01/21/03	7:40-x:xx
26	AI	Alton and Irvine Center Drive	02/21/03	7:40-8:25
27	ICDI	Irvine Center Drive and Discovery	02/28/03	7:43-8:38
28	ICP	Irvine Center Drive and Pacific	02/26/03	8:24-9:19
29	ICSB405	Irvine Center Drive and 405 SB	02/26/03	7:16-8:11

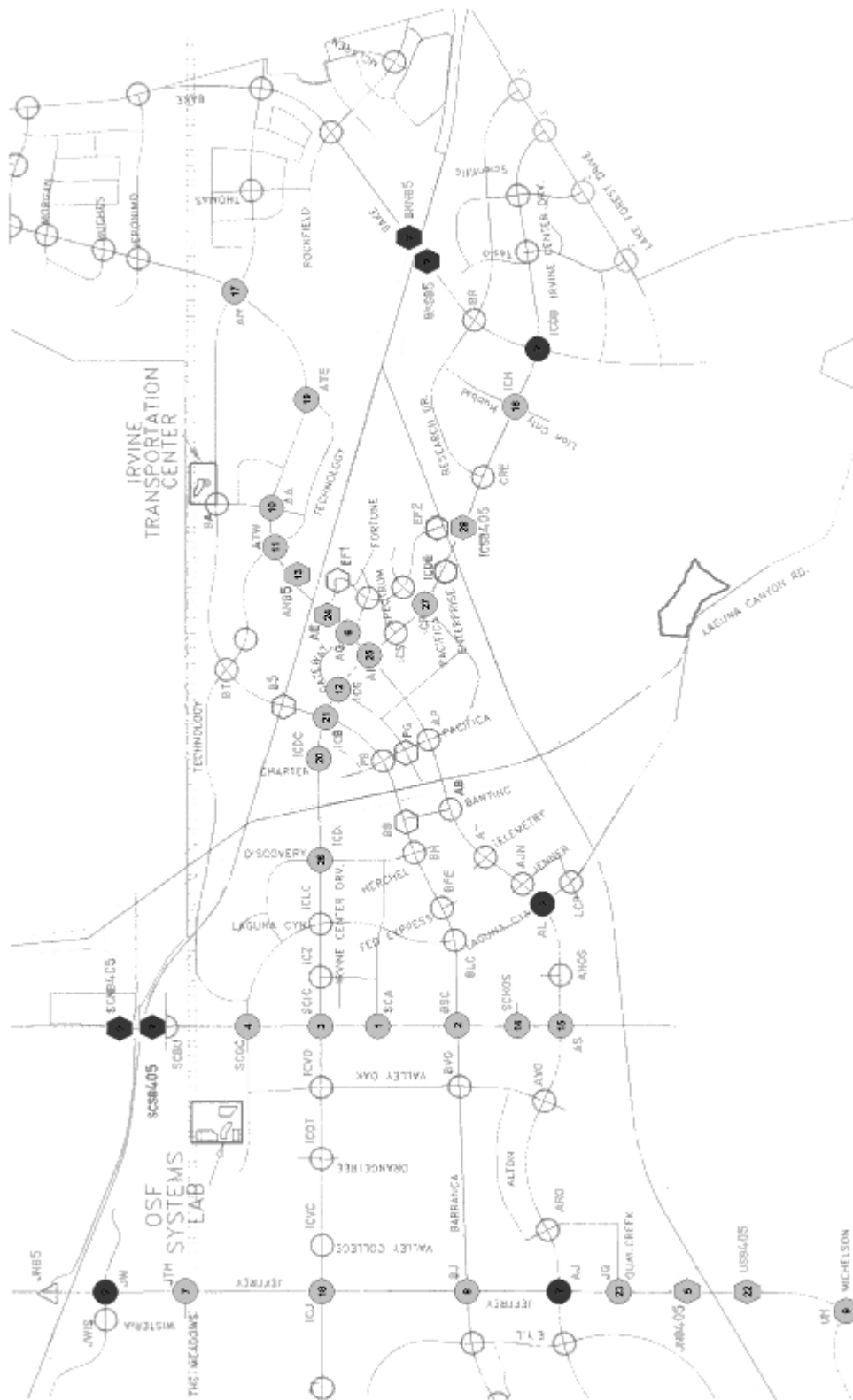


Figure 6-2 Intersection map for Irvine network.

7 DEVELOPMENT OF GRAPHICAL USER INTERFACE

7.1 Introduction

A software tool for estimating origin-destination (O-D) trip tables from traffic counts in transportation networks is currently being developed for the California Partners for Advanced Transit and Highways program (PATH TO 4135) at the Institute of Transportation Studies of the University of California Irvine. The O-D estimation software will be implemented for both off-line transportation planning applications and on-line traffic management applications that best meets Caltrans' needs. The software will contain a graphical user interface (GUI), from which it is possible to specify user adjustments of inputs for the generation of a full O-D trip table. These user-defined inputs include parameters (e.g., dispersion parameter, confident level of measurements), predefined paths (e.g. fixed route vehicles from buses or delivery vehicles, habitual drivers), point measurements (e.g. link counts, link travel times), and section-related measurements (e.g., subpath information from vehicle re-identification). The GUI will also allow flexible display of the outputs (e.g. flows from certain O-D pairs, travel times from certain paths). The architecture of the system is outlined in the subsequent section.

7.2 System Architecture

The O-D estimator and the GUI will be built as a custom application in TransCAD, a Geographic Information System (GIS) for transportation planning. TransCAD is the only Geographic Information System (GIS) designed specifically for transportation professionals to store, display, manage, and analyze transportation data. TransCAD custom applications are executed like any other WindowsTM programs. Using TransCAD as the platform to implement the O-D estimation provides the following advantages that are perfectly suited for the design objectives of the O-D estimator:

1. TransCAD is a full-scale GIS that can be used to create and customize maps, build and maintain geographic data sets, and perform many different types of spatial analysis. It provides the O-D estimator with the capability to utilize data from various sources (e.g., Global Positioning Systems) and to present results as thematic maps with geographical accuracy.
2. TransCAD contains special extensions for transportation planning tasks, including application modules for routing, travel demand forecasting, and logistics. These modules can pre-process data for the O-D estimator. Furthermore, the results of O-D estimation can be post-processed in TransCAD for various off-line transportation planning applications.
3. TransCAD can be jointly used with TransCAD for the Web, a separate software package that provides Internet access to TransCAD custom applications residing on a server. If TransCAD for the Web is used, the extension would enable the O-D estimator to be used

in on-line traffic management applications. Users would be able to remotely input data through a Web page and the results can also be displayed and retrieved over the Internet.

TransCAD includes the Geographic Information System Developer's Kit (GISDK) that will be used to build the O-D estimator GUI, including command menus, toolbars, toolboxes and dialog boxes. The GUI will be linked to the estimator program so the application will respond to user actions in a way best suited for the O-D estimation tasks.

7.3 Implementing the PFE GUI (Prototype)

The PFE GUI (prototype) can be implemented in a Windows-based workstation running TransCAD in two modes: testing and full deployment. In the testing mode, developers compile the GUI's GISDK source code into a temporary database then execute the GUI within TransCAD window. Testing mode is ideal for developers to test the program before the development is fully completed. The full deployment mode enables users to launch the GUI from a desktop icon similar to typical Windows-based programs. Before either mode can be implemented, files and folders required for the PFE program need to be placed under the TransCAD program folder.

7.3.1 Required Files and Folders

To run PFE GUI, the executable PFE needs to be copied and pasted into the TransCAD program directory (e.g., C:\Program Files\TransCAD). In addition, two folders containing PFE's input and output files must be copied and pasted into the TransCAD program directory as well. The two folders should be named: Network and Result (misspelling of the folder names may result in run time error during PFE execution). The Network folder contains network data files. A PFE network is divided into three separate text files. File names of the three files are the same except for the last digit (e.g., hyang1.txt, hyang2.txt, and hyang3.txt). The Result folder should be empty before a PFE run. After a PFE execution is completed, various files will be created and saved into this folder. Figure 7-1 shows the location of the PFE required folders and file in a Windows Explorer view.

7.3.2 Running PFE

Running PFE will first ask users to verify if all the required files and folders are present in the TransCAD program folder. If users answer "No", users will be asked to prepare these files and the program will not be executed. If users respond with "Yes", the PFE program will be executed and the results generated. In the case when the files are not present but the users select "Yes", the program will issue an error message.

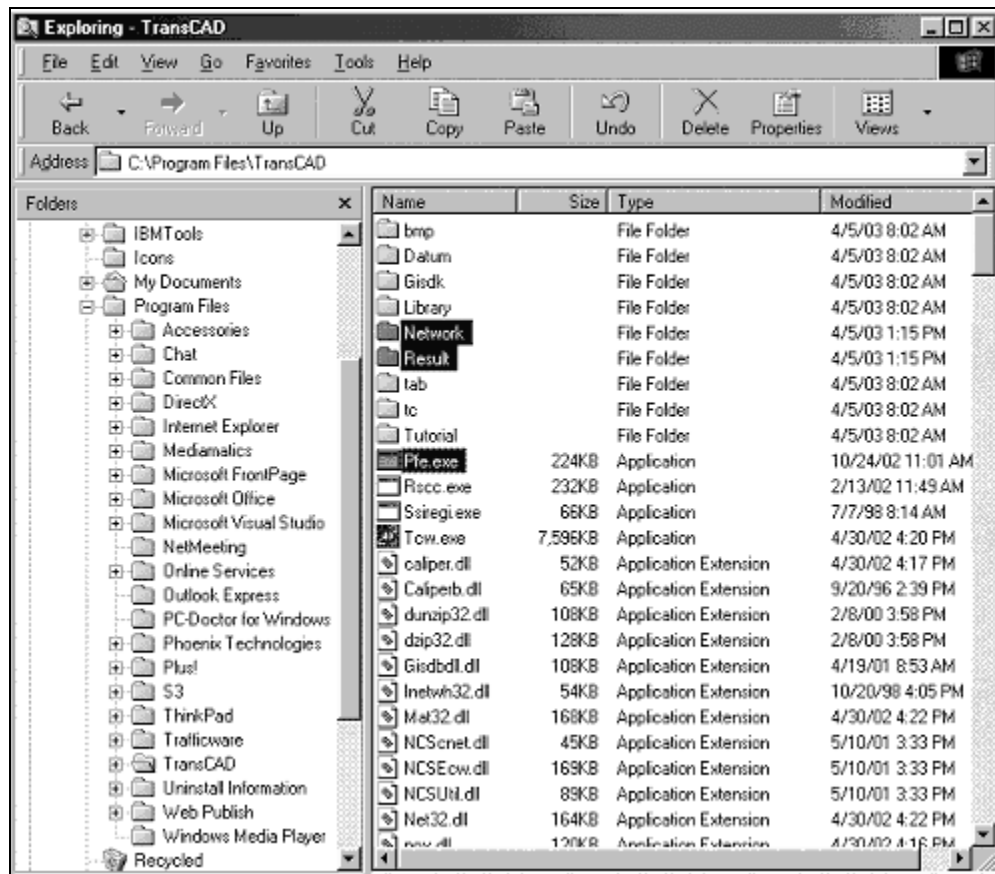


Figure 7-1 Location of PFE required folders and PFE executable file.

7.3.3 Displaying PFE results

After a successful PFE run, five comma-delimited text files will be generated. They are:

1. [Network file].geo
2. Linksum.csv
3. Odsum.csv
4. Pathsum.csv
5. Errsum.csv

The first file is a text/geography file importing the PFE network to TransCAD. The name of the file is the PFE network file name entered to the DOS command line during PFE execution. Linksum.csv contains link flow estimates, resulting from the PFE run. Odsum.csv contains the estimates of O-D demand. Pathsum.csv keeps estimates of path flows. Errsum.csv keeps track of estimation errors during the PFE estimation process. When launched, PFE GUI lets users identify locations of the five files and convert these files to various file formats in TransCAD for the purposes of graphical presentation and subsequent analysis as shown in Figure 7-2. The

network will be saved in TransCAD's proprietary GIS data format. Estimates of O-D demand will be saved as a Matrix file, which facilitates efficient matrix manipulation. Path flows will be saved as a Route System file, a file structure specially designed for managing and displaying transit or delivery routes. Figure 7-3 also displays the outputs of PFE in the TransCAD's environment.

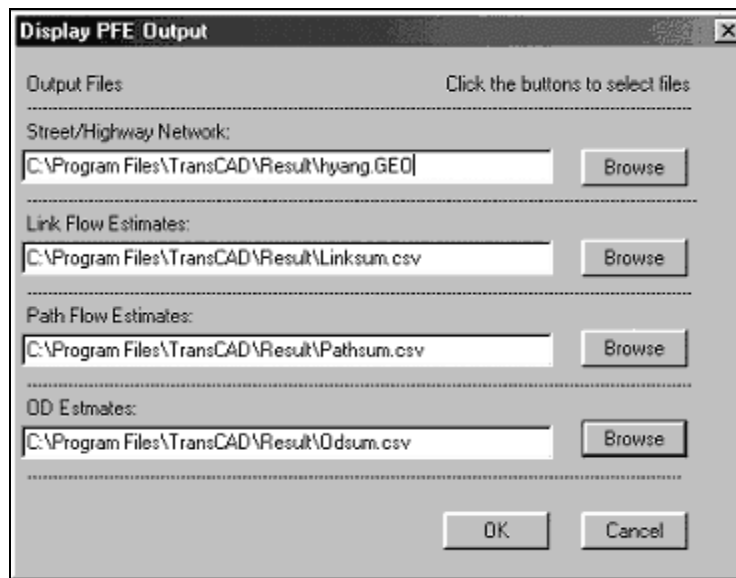


Figure 7-2 Display PFE outputs dialog box.

7.4 Enhancement of PFE GUI

Since the delivery of the prototype GUI, PFE has undergone several improvements that significantly increase the software's value as a practical tool for transportation planning. The improvements were identified through real-world applications of the software. In these applications, it was found that inconsistencies in the traffic counts are common and practically unavoidable; this suggests that the error bound set for each link volume estimation needs to be a controllable variable throughout the estimation process. A heuristic procedure has been developed to facilitate the process of defining error bounds to efficiently handle the inconsistencies among the observed traffic volumes. Currently, the heuristic procedure has yet to be integrated into the GUI.

Other areas of improvement have also been identified. First, PFE uses a proprietary network data format specifically designed for the software's unique estimation procedure. The software tool of PFE can greatly benefit from a graphical network editor, which would enable the creation and modification of network data in its own format. Furthermore, a data conversion component

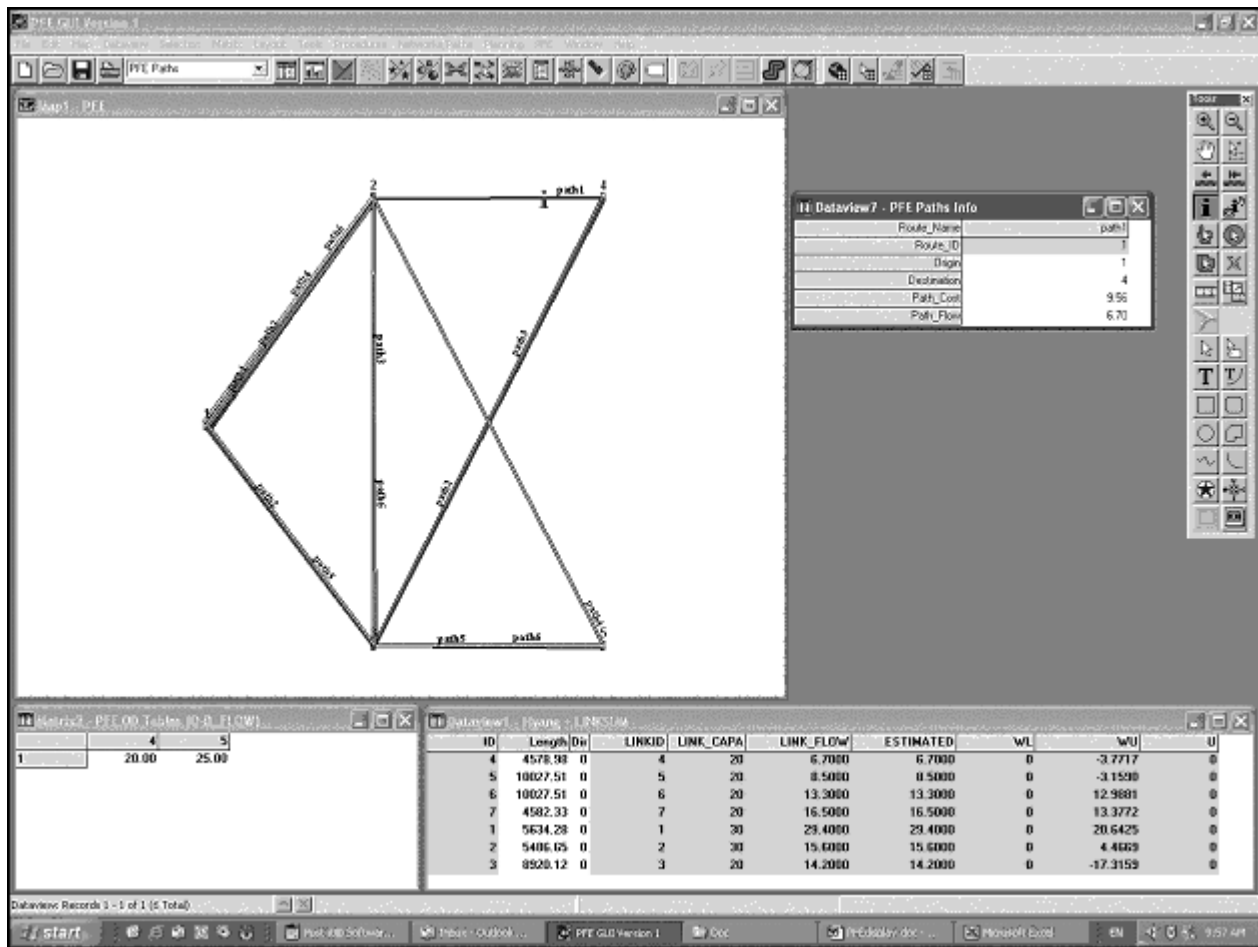


Figure 7-3 Views of PFE outputs

that transforms network data of other formats to the PFE format will enhance the software's practical value. Finally, diagnostic tools such as scatter plots and sensitivity analysis for evaluating PFE outputs can be integrated in the GUI. The enhanced PFE GUI will contain the following four components:

7.4.1 Network Editor

To further streamline the data preparation process for PFE, a graphical environment and utilities will be developed to build and edit networks and allow them to be saved directly in the PFE format.

7.4.2 Data Conversion

The PFE O-D estimation has its own network data format. In order to use existing network data from various agencies, a data converter will be developed and integrated in the GUI. The outputs of PFE can also be converted to different formats for post-processing in other software packages.

7.4.3 Estimation Control

Currently, PFE does not accept inputs other than the network data during the estimation. In order to implement the new error bound heuristic into the software, substantial programming necessary to build a GUI component for estimation control, is required.

7.4.4 Diagnostic Tools

Once a set of O-D estimation is completed, it is difficult to diagnose the precision of the estimation without seeing the results in a graphical environment. Diagnostic tools that display the estimation results graphically will be developed. The tools are centered on a scatter plot that displays estimated versus observed volumes on a two-dimensional space. That is, observed values are on the Y-axis while the estimated counterparts are on the X-axis. Each observation in the scatter plot will be linked to a graphical representation of the traffic network, facilitating visual identification of problematic estimates. These observations will also be linked to a table view for generation of summary statistics and sensitivity analysis.

8 CONCLUSIONS AND FURTHER RESEARCHES

8.1 Findings and Conclusions

Phase one of this project has primarily explored the PFE for estimating steady-state O-D trip tables and developed several promising algorithms. The PFE has also been applied to estimate the O-D trip table for the Testbed Irvine network in Orange County, California. This network consists of three major freeways (I-5, I-405, and SR-133), and several arterials in the City of Irvine. In this application, the Irvine network and associated demand data were extracted from the Orange County Transportation Analysis Model (OCTAM), which contains data for the whole county. The extracted network is composed of 163 nodes, 496 links connecting 39 traffic analysis zones (TAZ), 28 external stations, and 1,547 O-D pairs. Using both simple and real networks, we have demonstrated that PFE has the capability to correctly estimate the total and individual O-D demands when proper information is provided. Based on the preliminary results, it has been determined that the selection of observed links plays an important role in the O-D estimation problem, as each observation contributes different amount and quality of information. Though the quality of estimates can be improved with the usage of more traffic counts, this may be impractical when the budget for data collection is scarce. From the small network, the results show that if there is at least one observation on each path, the total demand of the network can be correctly captured. In addition, higher observed traffic volumes appear to contribute more to the quality of O-D estimation. The most difficult task observed thus far is the estimation of the spatial patterns of O-D demands. Even when all network links are measured, the individual O-D demands may not be estimated correctly.

Currently, we have completed the following tasks important to obtaining the research goal:

8.1.1 *Software*

- Fortran codes for the steady-state PFE algorithm
- A prototype of PFE GUI integrated with TransCAD
- PFE GUI 1.0 User's Guide

8.1.2 *Data Collection*

- Irvine network coded in PFE (163 nodes, 496 links, and 1,547 O-D pairs)
- An initial O-D trip table extracted from the OCTA model
- Freeway traffic counts for the Irvine network
- Selected traffic counts on the surface streets

8.2 Tasks to be completed in Phase Two

Phase two of the PFE project focuses on developing a time-dependent version of the PFE for deriving time-dependent O-D trip table. The adaptations of the PFE mentioned for the steady-state case will be extended to the time-dependent case. Significant effort in phase two will be devoted to developing fast algorithms that achieve near, real-time speed for online

implementation. State-of-the-art, path-based algorithms for the static traffic assignment problem, primarily the gradient projection (GP) algorithm will be adapted to solve the time-dependent PFE problem. Note that we have already successfully implemented and tested both algorithms for the static traffic assignment problem on various realistic large-scale networks, including the ADVANCE network which consists of 2,552 nodes, 7850 links, 447 zones, and 137,417 O-D pairs (See Chen and Lee, 1999; Chen et al., 2000).

Implementation of the time-dependent PFE is essentially comprised of two major components. The first component deals with the problem of finding the time-dependent shortest paths within the road network. These are then used to construct time-space networks for solving a nonlinearly constrained large-scale optimization problem.

Several state-of-the-art, time-dependent shortest path algorithms will be explored and the most efficient algorithm will be coded and tested for later use in the proposed O-D trip table estimation scheme. These estimation schemes will include the label correcting algorithm with deque implementation, the decreasing order of time (DOT) algorithm, and the Chrono-SPT algorithm. Secondly, the two PFE solvers suggested by Bell et al. (1996), the method of successive averages (MSA) and the method of iterative balancing will be implemented as is. However, it is unclear at this stage how efficient these algorithms will be when estimating O-D trip tables in large-scale networks, since they are not extensively tested over large networks. Hence one of the main tasks in the second phase is to test the convergence properties of these algorithms and evaluate their computational efficiency over large networks. In the event that these algorithms do not perform satisfactorily, we will refine these algorithms to improve their computational efficiency or, develop new solution algorithms based on more powerful optimization techniques. Moreover, efforts will be made to extend Bell's time-dependent PFE (TDPFE) to include traffic measurements other than link traffic counts, such as link/path travel time and cordon flow counts, so as to improve the reliability of the O-D estimates.

After the development of the core algorithms of the TDPDE, a thorough testing of codes will first be performed to assure reliability and robustness of the developed codes. Next we will develop a set of criteria for evaluating the quality and reliability of the time-dependent O-D estimates. Certainly, link counts and travel times will be major measures to judge the quality of the estimated trip tables. Beyond these two measures, much uncertainty exists. For example are there other measures that can give us clues on how good our estimates are? Can trip length distributions be used to probe the quality of O-D estimates? How can we use cordon counts and license plate matching, which gives partial but nonetheless very useful information on O-D patterns for part of the network, to assess and improve the quality of the O-D estimates? All these questions will be examined and studied towards the end of this project.

9 REFERENCE

- Ashok, K. and Ben-Akiva, M.E. (1993) "Dynamic origin-destination matrix estimation and prediction for real-time management systems." *Proceedings of the 12th International Symposium on Transportation and Traffic Theory*, Berkeley, California.
- Ashok, K. and Ben-Akiva, M.E. (2000) "Alternative approaches for real-time estimation and prediction of time-dependent origin-destination flows." *Transportation Science* 34, 21-36.
- Barbour, R. and Fricker, J.D. (1994) "Estimation of an origin-destination table using a method on shortest augmenting paths." *Transportation Research* 28B, 77-89.
- Bazaraa, M. and Shetty, C. (1979) *Nonlinear Programming: Theory and Algorithms*. John Wiley & Sons, New York, NY.
- 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.
- Bell, M.G.H. (1983a) "The estimation of an origin-destination matrices from traffic counts." *Transportation Science* 17(2), 198-217.
- Bell, M.G.H. (1983b) "The estimation of an origin-destination flows and their confidence intervals from measurements of link volumes: A computer program." *Traffic Engineering & Control* 24, 202-205.
- Bell, M.G.H. (1984a) "The estimation of junction turning volumes from traffic counts: the role of prior information." *Traffic Engineering & Control* 25(5), 279-283.
- Bell, M.G.H. (1984b) "Variance and covariance for origin-destination flows when estimated by log-linear models" *Transportation Research* 18B, 497-507.
- Bell, M.G.H. (1984c) "Log-linear models for the estimation of origin-destination matrices from traffic counts: An approximation." *Proceedings of the 9th International Symposium on Transportation and Traffic Theory*, J. Volmuller and R. Hamerslag, (eds.), VNU Science Press, Utrecht, The Netherland.
- Bell, M.G.H. (1991a) "The estimation of origin-destination matrices by constrained generalized least squares" *Transportation Research* 25B, 13-22.
- Bell, M.G.H. (1991b) "The real-time estimation of origin-destination flows in the presence of platoon dispersion" *Transportation Research* 25B, 115-125.

- Bell, M.G.H. (1994) "A simple alternative to Dial's logit assignment algorithm." *Transportation Research* 29B, 369-386.
- Bell, M.G.H. and Grosso, S. (1998) "The path flow estimator as a network observer." *Traffic Engineering & Control*, 540-549.
- Bell, M.G.H. and Iida, Y. (1997) *Transportation Network Analysis*. John Wiley & Sons.
- Bell, M.G.H., Lam, W.H.K., and Iida, Y. (1996) "A time-dependent multiclass path flow estimator" *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, Lyon, France.
- 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.
- Bell, M.G.H. and Shield, C.M. (1995) "A log-linear model for path flow estimation." *Proceedings of the 4th International Conference on the Applications of Advanced Technologies in Transportation Engineering*, Capri, 695-699.
- Bell, M.G.H., Shield, C.M., Busch, F., and Kruse, G. (1997) "A stochastic user equilibrium path flow estimator" *Transportation Research* 5C, 197-210.
- Yang, H, and Bell, M.G.H. (1998) "Models and algorithms for road network design: A review and some new developments." *Transportation Reviews* 18(3), 257-278.
- 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. Halled, Kluwer Publishers.
- Bertsekas, D. (1976) "On the Goldstein-Levitin-Polyak gradient projection method." *IEEE Transactions on Automatic Control*, 21 (2), 174-183.
- Bianco, L., Confessore, G., and Reverberi, P. (2001) "A network based model for traffic sensor location with implications on O/D matrix estimates." *Transportation Science* 35(1), 50-60.
- Bierlaire, M. and Toint, P.L. (1995) "MEUSE: An origin-destination matrix estimator that exploits structure." *Transportation Research* 29B, 47-60.
- Bierlaire, M. (2002) "The total demand scale: a new measure of quality for static and dynamic origin-destination trip tables." *Transportation Research*, 36B, 755-851.

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

Brenninger-Gothe, M., Jornsten, K.O., and Lundgren, J.T. (1989) “Estimation of origin-destination matrices from traffic counts using multi-objective programming formulations.” *Transportation Research* 23B, 257-269.

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

Carey, M. (1992) “Non convexity of the dynamic traffic assignment problem.” *Transportation Research* 26B, 127-133.

Cascetta, E. (1984) “Estimation of origin-destination matrices from traffic counts and survey data: A generalized least squares estimator.” *Transportation Research*, 18B, 289-299.

Cascetta, E., Inaudi, D., and Marquis, G. (1993) “Dynamic estimation of origin-destination matrices using traffic counts.” *Transportation Science* 27, 363-373.

Cascetta, E. and Nguyen, S. (1988) “A unified framework for estimating or updating origin/destination matrices from traffic counts.” *Transportation Research* 22B, 437-455.

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 route enumeration: comparison of models and algorithms. *Proceedings of International Federation of Automatic Control: Transportation Systems*, Chania, Greece. 1078-1084.

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

Chang, G.L. and Tao, X. (1999) “An integrated model for estimating time-varying network origin-destination distributions.” *Transportation Research*, 33A, 381-399.

Chang, G.L. and Wu, J. (1994) “Recursive estimation of time-varying origin-destination flows from traffic counts in freeway corridors.” *Transportation Research*, 28B, 141-160.

Chang, G.L. and Wu, J. (1996) “Estimation of dynamic network O-D with screenline flows.” *Transportation Research*, 30B, 277-290.

- Chen, A. and Lee, D.-H. (1999) "Path-based algorithms for large scale traffic equilibrium problem: a comparison between DSD and GP." *Paper presented at the 78th Annual Meeting of the Transportation Research Board.*
- Chen, A., Lee, D.-H., and Jayakrishnan, R. (2000) "Computational Study of State-of-the-Art Path-Based Traffic Assignment Algorithms," submitted to *ASCE Journal of Transportation Engineering.*
- Chu, C. (1989) "A Paired Combinatorial Logit Model for Travel Demand Analysis." *Proceedings of the Fifth World Conference on Transportation Research 4*, Ventura, CA, 1989, 295-309.
- Cremer, M. and Keller, H. (1981) "Dynamic identification of flows from traffic counts at complex intersection." *Proceedings of the 8th International Symposium on Transportation and Traffic Theory*, V.F. Hurdle et al., (eds.), University of Toronto Press, Toronto.
- Cremer, M. and Keller, H. (1984) "A systems dynamic methods to the estimation of entry and exit O-D flows." *Proceedings of the 9th International Symposium on Transportation and Traffic Theory*, J. Volmuller and R. Hamerslag, (eds.), VNU Science Press, Utrecht, The Netherland.
- Cremer, M. and Keller, H. (1987) "A new class of dynamic methods for the identification of origin-destination flows." *Transportation Research*, 21B, 117-132.
- Daganzo, C. and Sheffi, Y. (1977) "On stochastic models of traffic assignment." *Transportation Research*, 14B, 243-255.
- Davis, G.A. (1993) "A statistical theory for estimation of origin-destination parameters from time-series of traffic counts." *Proceedings of the 12th International Symposium on Transportation and Traffic Theory*, Daganzo, C.F (eds.), Elsevier Science Publishing Co.
- Dial, R. (1971) "A probabilistic multi-path assignment model that obviates path enumeration." *Transportation Research*, 5, 83-11.
- Fisk, C. (1980) "Some developments in equilibrium traffic assignment." *Transportation Research*, 14B, 243-255.
- Fisk, C. (1988) "On combining maximum entropy trip matrix estimation with user equilibrium." *Transportation Research*, 22B, 69-73.
- Fisk, C. (1989) "Trip matrix estimation from link traffic counts: the congested network case." *Transportation Research*, 23B, 245-250.
- Fisk, C. and Boyce, D.E. (1983) "A note on trip matrix estimation from traffic count data." *Transportation Research*, 17B, 245-250.

Florian, M. and Hearn, D. (1995) "Network equilibrium models and algorithms." Handbooks in Operations Research and Management Science, Volume 8, Network Routing (M.O. Ball et al. Editors). Elsevier Science. Netherlands.

Gao, Z, and Song, Y. (2002) "A reserve capacity model of optimal signal control with user-equilibrium route choice." *Transportation Research* 36B, 313-323.

Gen, M. and Cheng, R. (2000) "*Genetic Algorithms and Engineering Optimization*." John Wiley and Sons, Inc., New York.

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.

Goldberg, D. (1989) "*Genetic Algorithms in Search, Optimization, and Machine Learning*." Addison-Wesley, Reading, MA.

Gur, Y., Turnquist, M., Schneider, M., LeBlanc, L., and Kurth, D. (1980) "Estimation of an origin-destination trip table based on observed link volumes and turning movements." Volume 1: Technical report, FHWA, FHWA/RD-8-034.

Hauer, E., Pagitsas, E., and Shin, B.T. (1981) "Estimating turning flows from automatic counts." *Transportation Research Record* 759, 1-7.

Hazelton, M.L. (2000) "Estimation of origin-destination matrices from link flows on uncongested networks." *Transportation Research*, 34B, 549-566.

Hendrickson, C. and McNeil, S. (1985) "Estimation of origin-destination matrices with constrained regression." *Transportation Research Record* 976, 25-32.

Janson, B.N. (1993) "Most likely O-D link uses from equilibrium assignment." *Transportation Research*, 27B, 333-350.

Janson, B.N. and Southworth, F. (1992) "Estimating departure times from traffic counts using dynamic assignment." *Transportation Research*, 26B, 3-16.

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.

Jeffreys, M. and Norman, M. (1977) "On finding realistic turning flows at road junctions." *Traffic Engineering & Control*, 18(1), 19-25.

Keller, H. and Ploss, G. (1987) "Real time identification of O-D network flow from counts for urban traffic control." *Proceedings of the 10th International Symposium on Transportation and Traffic Theory*, N.H. Gartner and N.H.M. Wilson (eds.), Elsevier Science Publishing Co.

- Khattak, A.J. and De Palma, A. (1997) "The impact of adverse weather conditions on the propensity to change travel decisions: A survey of Brussels commuters." *Transportation Research*, 31A, 181-203.
- Koppelman, F.S. and Wen, C.H. (2000) "The paired combinatorial logit model: properties, estimation and application." *Transportation Research* 34B, 75-89.
- Lam, W.H.K. and Lo, H.P. (1990) "Accuracy of O-D estimates from traffic counts." *Traffic Engineering & Control* 31, 358-367.
- Lam, W.H.K. and Lo, H.P. (1991) "Estimation of origin-destination matrix from traffic counts: a comparison of entropy maximizing and information minimizing models." *Transportation Planning and Technology* 16, 166-171.
- Lan, C.J. and Davis, G.A. (1999) "Real-time estimation of turning movement proportions from partial counts on urban networks." *Transportation Research*, 7C, 305-327.
- Larsson, T. and Patriksson, M. (1992) "Simplicial decomposition with disaggregated representation for the traffic assignment problem." *Transportation Science*, 26, 4-17.
- Larsson, T., and Patriksson, M. (1995) "An augmented Lagrangian dual algorithm link capacity side constrained traffic assignment problems" *Transportation Research* 29B, 433-455.
- LeBlanc, L. and Farhangian, K. (1982) "Selection of a trip table which reproduces observed link flows." *Transportation Research*, 16B, 83-88.
- Liu, S. and Fricker, J.D. (1996) "Estimation of a trip table and the θ parameter in a stochastic network." *Transportation Research*, 30A, 287-305.
- Lo, H.P., Zhang, N., and Lam, W.H.K. (1996) "Estimation of an origin-destination matrix with random link choice proportions: a statistical approach." *Transportation Research*, 30B, 309-324.
- Madanat, S.M., Hu, S.R., and Krogmeier, J (1996) "Dynamic Estimation and Prediction of Freeway O-D Matrices with Route Switching Considerations and Time-Dependent Model Parameters." *Transportation Research Record* 1537, 98-105.
- Maher, M. (1983) "Inferences on trip matrices from observations on link volumes: a Bayesian statistical approach." *Transportation Research*, 17B, 435-447.
- 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.

- Maher, M., Zhang, X., and Van Vliet, D. (2001) "A bi-level programming approach for trip matrix estimation and traffic control problems with stochastic user equilibrium link flows." *Transportation Research*, 35B, 23-40.
- McNeil, S. and Hendrickson, C. (1985) "A regression formulation of the matrix estimation problem." *Transportation Science* 19, 287-292.
- Mekky, A. (1979) "On estimating turning flows at road junctions." *Traffic Engineering & Control* 20(10), 486-487.
- Mirchandani, P. and Soroush, H. (1987) "Generalized traffic equilibrium with probabilistic travel times and perceptions." *Transportation Science* 21, 133-152.
- Nanne, J., Zijpp, V. D., and Hamerslag, R. (1994) "Improved Kalman Filtering Approach for Estimating origin-Destination Matrices for Freeway Corridors." *Transportation Research Record* 1443, 54-63
- Nguyen, S. (1977) "Estimating an OD matrix from network data: A network equilibrium approach." In *Transportation Planning Models*, Florian, M. (Ed), 363-380.
- Nguyen, S. (1984) "Estimating origin-destination matrices from observed flow." Publication 87 CRT, University of Montreal.
- Nihan, N.L. and Davis, G.A. (1987) "Recursive estimation of origin-destination matrices from input/output counts." *Transportation Research*, 21B, 149-163.
- Nihan, N.L. and Davis, G.A. (1989) "Application of prediction-error minimization and maximum likelihood to estimate intersection O-D matrices from traffic counts." *Transportation Science* 23, 77-90.
- Nihan, N.L. and Hamed, M.M. (1992) "Fixed-point approach to estimating freeway origin-destination matrices and the effect of erroneous data on estimate precision." *Transportation Research Record* 1357, 18-28.
- Norman, M., Hoffman, N., and Harding, F. (1979) "Non-iterative methods for generating a realistic turning flow matrix for a junction." *Traffic Engineering & Control* 20(12), 587-589.
- Okutani, I. (1987) "The Kalman filtering approaches in some transportation and traffic problems." *Proceedings of the 10th International Symposium on Transportation and Traffic Theory*, N.H. Gartner and N.H.M. Wilson (eds.), Elsevier Science Publishing Co.
- Patriksson, M. (1994) *The Traffic Assignment Problem: Models and Methods*. VSP, Utrecht, The Netherlands.
- Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T. *Numerical Recipes in Fortran*. 2nd Eds. Cambridge University Press, 1992.

- Robillard, P. (1975) "Estimation of the O-D matrix from observed link volumes." *Transportation Research* 9, 123-128.
- Schaefer, M.C. (1988) "Estimation of Intersection Turning Movements from Approach Counts." *ITE Journal* 58(10), 41-46.
- Sheffi, Y. (1985) *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Prentice-Hall, Englewood Cliffs, New Jersey.
- Sheffi, Y. and Barnhart, C. (1987) "XENT: Extended traffic assignment model." *Transportation Engineering*, 113(4), ASCE, 450-462.
- Sheffi, Y. and Powell, W. (1982) "An algorithm for the equilibrium assignment problem with random link times." *Networks* 12(2), 191-207.
- Sherali, H.D., Arora, N., and Hobeika, A.G. (1997) "Parameter optimization methods for estimating dynamic origin-destination trip-tables." *Transportation Research*, 31B, 141-157.
- Sherali, H.D. and Park, T. (2001) "Estimation of dynamic origin-destination trip tables for a general network." *Transportation Research*, 35B, 217-235.
- Sherali, H.D., Sivanandan, R., and Hobeika, A.G. (1994) "A linear programming approach for synthesizing origin-destination trip tables from link traffic volumes." *Transportation Research*, 28B, 213-234.
- Spiess, H. (1987) "A maximum likelihood model for estimating origin-destination matrices." *Transportation Research*, 21B, 395-412.
- Sun, C., Jayakrishnan, R. and Tsai, W.K. (1994) "Computational study of a path-based algorithm and its variants for static traffic assignment." *Transportation Research Record* 1537, 106-115.
- Tatneni, M., Edwards, H., and Boyce, D. (1998) "Comparison of the disaggregate simplicial decomposition and Frank-Wolfe algorithms for user-optimal route choice." Forthcoming in *Transportation Research Record*, 1998.
- Toregas, C. and ReVelle, C. (1973) "Binary logic solutions to a class of location problems." *Geographical Analysis*, 5, 145-155.
- Turnquist, M. and Gur, Y. (1979) "Estimation of trip tables from observed link volumes." *Transportation Research Record* 730, 1-6.
- Van Aerde, M., et al. (1993) "QUEENSOD: A method for estimating time varying origin-destination demands for freeway corridors/networks." Transportation Research Board Annual Meeting, Washington, D.C.

- Van Zuylen, H.J. (1979) "The estimation of turning flows on a junction." *Traffic Engineering and Control* 20(11), 539-541.
- Van Zuylen, H.J. and Willumsen, L.G. (1980) "The most likely trip estimated from traffic counts." *Transportation Research*, 14B, 281-293.
- Von Falkenhausen, H. (1976) "Traffic assignment by a stochastic model." *Proceedings of the Fourth International Conference on Operational Research*, 415-421.
- Wardrop, J.G. (1952) "Some Theoretical Aspects of Road Traffic Research." *Proceedings, Institution of Civil Engineers*, Part II, 1, 325-378.
- Willumsen, L.G. (1984) "Estimating time-dependent trip matrices from traffic counts." *Proceedings of the 9th International Symposium on Transportation and Traffic Theory*, The Netherlands, Delft University.
- Wong, S.C, and Yang, H. (1997) "Reserve capacity of a signal-controlled road network." *Transportation Research* 31B, 397-402.
- Wu, J. (1997) "A real-time origin-destination matrix updating algorithm for on-line applications." *Transportation Research*, 30B, 277-290.
- Wu, J., and Chang, G.L. (1996) "Estimation of Time-Varying Origin-Destination Distributions with Dynamic Screenline Flows." *Transportation Research*, 30B, 277-290.
- Yang, H. (1995) "Heuristic algorithms for the bi-level origin-destination matrix estimation algorithm." *Transportation Research*, 29B, 231-242.
- Yang, H., Gan, L., and Tang, W.H. (2001) "Determining cordons and screen lines for origin-destination trip studies." *Proceedings of the 3rd Eastern Asia Society for Transportation Studies*, Vol. 3, No. 2, Hanoi, Vietnam, October, 2001, 85-99.
- Yang, H., Iida, Y., and Sasaki, T. (1991) "An analysis of the reliability of an origin-destination trip matrix estimated from traffic counts." *Transportation Research*, 25B, 351-363.
- Yang, H., Iida, Y., and Sasaki, T. (1992) "Estimation of origin-destination matrices from traffic counts on congested networks." *Transportation Research*, 26B, 417-434.
- Yang, H., Iida, Y., and Sasaki, T. (1994) "The equilibrium-based origin-destination matrix estimation problem." *Transportation Research*, 28B, 23-33.
- Yang, H., Sasaki, T., Iida, Y., and Asakura, Y. (1992) "Estimation of origin-destination matrices from traffic counts on congested networks." *Transportation Research*, 26B, 417-434.

Yang, H., Meng, Q., and Bell, M.G.H. (2000) "Simultaneous estimation of the origin-destination matrices and travel-cost coefficient for congested networks in a stochastic equilibrium." *Transportation Science* (in press).

Yin, Y. (2000) "Genetic Algorithms based approach for bilevel programming models." *Journal of Transportation Engineering* 126(2), 115-120.

Yu, P. and Davis, G.A. (1994) "Estimating freeway origin-destination patterns using automatic traffic counts." *Transportation Research Board Preprint No. 940112*. Washington, D.C.

Zhou, Z. and Chen, A. (2003) "Stochastic user equilibrium problem: A comparison between length-based and congestion-based C-logit model." *Proceedings of the 8th Conference of Hong Kong Society for Transportation Studies*, 244-253.