

## **ABSTRACT**

Title of Dissertation:     **A ROBUST MODEL FOR ESTIMATING FREEWAY  
DYNAMIC ORIGIN-DESTINATION MATRICES**

Pei-Wei Lin, Doctor of Philosophy, 2006

Dissertation directed by: Professor Gang-Len Chang  
Department of Civil and Environmental Engineering

The purpose of this study is to develop an effective model and algorithm for estimating dynamic Origin-Destination demands for freeways. The primary challenge for this research subject lies in the fact that the number of unknown parameters is always more than the number of observable data, especially for a large network. Hence, the estimated O-D patterns may result in a large variance and insufficient reliability for use in practice. Besides, most existing approaches are grounded on the assumptions that a reliable initial O-D set is available and traffic volume data from detectors are accurate. However, in most highway network systems, both types of critical information are either unavailable or subjected to a significant level of measurement errors.

To deal with those critical issues, this study has developed a set of dynamic models and solution algorithms for estimating freeway dynamic O-D matrices. The first extended model formulations can capture the speed discrepancy among drivers with an embedded travel time distribution function and the derivable interrelations between time-varying ramp and mainline flows. These formulations also feature their best use of the

available mainline information and travel time function, and hence substantially increase the system observability with fewer parameters.

The second component is an iterative algorithm that can be used to provide a reliable estimate of the initial O-D set, which is often unavailable in practice. The proposed algorithm first divides the network into small sub-networks to reduce the number of unknown variables, and recursively compute the O-D proportions for each sub-network to well capture the relations between the O-D demands and the input information.

To deal with the constraints that the available data usually contain measurement errors, this research has developed an interval-based model for estimating dynamic freeway O-D demands. This component includes a set of formulations that converts each model input as an interval with its boundaries based on the prior knowledge.

This study has performed sensitivity analyses and explored their potential for real-world application with the I-95 freeway corridor in Maryland. The numerical results under various traffic scenarios have indicated the promising properties of the proposed models and algorithms.

**A ROBUST MODEL FOR ESTIMATING FREEWAY DYNAMIC  
ORIGIN-DESTINATION MATRICES**

by

Pei-Wei Lin

Dissertation submitted to the Faculty of the Graduate School of  
University of Maryland, College Park in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
2006

Advisory Committee:

Professor Gang-Len Chang, Chair  
Professor Martin Dresner  
Professor Ali Haghani  
Professor David Lovell  
Professor Paul Schonfeld

@ Copyright by

Pei-Wei Lin

2006

## **ACKNOWLEDGEMENTS**

I would like to express my deepest appreciation to my advisor, Dr. Gang-Len Chang, for his encouragement, inspiration, and guidance throughout my Ph.D. study at the University of Maryland at College Park. My career has been strongly influenced by his dedication to excellent research work.

My thanks also go to the members of my doctoral examination committee, Dr. Martin Dresner, Dr. Ali Haghani, Dr. David Lovell, Dr. Paul Schonfeld, and Dr. Richard Woo for their valuable comments and suggestions on my research work.

Very special thanks are to my colleagues, Nan Zou, Hua Xiang, Ying Liu, Xiaorong Lai, Jason Chou and Christy Fang for providing me both technical and spiritual support. I am also grateful to my best friends, Betty Lee and Yu Zhai, who let me feel like being in a family during the last two years of my Ph.D. study.

I profoundly appreciate my fiancé, Chun-Li Wang for his unconditional support and understanding throughout my studies. His support and encouragement was in the end what made this dissertation possible. Finally, I would like to express my deepest appreciation to my family for their unlimited encouragement and support. Especially, to my parents, you deserve my deepest gratitude and love for your dedication and the many years of support during my studies.

## Table of Contents

List of Tables .....	vi	
List of Figures.....	vii	
Chapter 1	Introduction.....	1
1.1	Background.....	1
1.2	Research Objectives.....	3
1.3	Organization of Dissertation.....	4
Chapter 2	Literature Review.....	7
2.1	Introduction.....	7
2.2	Model Formulations for Dynamic O-D Estimation.....	8
2.2.1	Dynamic Traffic Assignment Based Approaches.....	8
2.2.2	Dynamic Non-Traffic Assignment Based Approaches.....	12
2.3	Solution Algorithms for Dynamic O-D Estimation.....	17
2.3.1	Non-Recursive Estimation Algorithm.....	17
2.3.2	Recursive Estimation Algorithm.....	20
2.4	Closure.....	25
Chapter 3	Model Formulations of Dynamic O-D Estimation for Freeway Corridors... .....	27
3.1	Introduction.....	27

3.2	Nature of Problem and Basic Formulations.....	29
3.3	The Dynamic O-D Estimation Model with Travel Time Distribution .....	32
3.4	The State-Space Modeling and Kalman Filtering Estimation Algorithm.....	38
3.5	Numerical Examples and Sensitivity Analyses .....	43
3.5.1	Sensitivity Analysis for A Small Network.....	43
3.5.2	A Large Freeway Network Example .....	57
3.6	Closure .....	67
Chapter 4	An Algorithm for Estimating the Initial O-D Matrix in a Large Freeway Network.....	69
4.1	Introduction.....	69
4.2	Concepts and Rules of Network Decomposition .....	70
4.3	Step-By-Step Procedures of the Solution Algorithm for Initial O-D Sets	76
4.4	Numerical Example for Evaluating the Initial O-D Estimation Algorithm .....	79
4.5	Closure .....	90
Chapter 5	Modeling the Freeway Dynamic O-D Estimation System with Measurement Errors.....	92
5.1	Introduction.....	92
5.2	The Model Formulations to Account for Measurement Errors .....	93
5.3	An Interval-Based Solution Algorithm.....	96

5.4	Numerical Examples for Evaluating the Interval-Based Dynamic O-D Estimation Model.....	102
5.4.1	Data Generation and Experiment Designs.....	102
5.4.2	Evaluations and Comparisons of Estimation Results .....	105
5.5	Closure .....	116
Chapter 6	Conclusions and Future Research.....	117
6.1	Summary of Research Accomplishments .....	117
6.2	Future Research .....	120
Appendix A.	Interval Arithmetic.....	122
References	.....	123

## List of Tables

Table 3.1. The Input Time Series Of O-D Proportions.....	44
Table 3.2. The Absolute Error Statistics With Different Sets Of Initial Values.....	52
Table 3.3. The Absolute Error Statistics With Travel Time Variation.....	53
Table 3.4 The Geometry Information For Each Link.....	58
Table 3.5. Input O-D Demand For Simulation Experiments (Unit: Vph) .....	59
Table 3.6. Comparison Of Estimation Error Statistics .....	65
Table 4.1a. The Example Initial Value Sets And The Estimated Initial Value Sets From The Initial O-D Set Estimation Algorithm.....	82
Table 4.1b. The Example Initial Value Sets And The Estimated Initial Value Sets From The Initial O-D Set Estimation Algorithm.....	83
Table 4.2. The Average Absolute Error And Improvement Of The Initial O-D Sets.....	84
Table 4.3. The Taae Results Of The First 5 Cases.....	85
Table 4.4. The Average Absolute Error Statistics .....	90
Table 5.1. Scenarios Under Different Levels Of Measurement Errors In Flow Data ....	104
Table 5.2. Scenarios Under Different Levels Of Availability In Initial O-D Set .....	105
Table 5.3. Aae Statistics Of Estimation Results With Approach-1 – Experiment-1 .....	107
Table 5.4. Aae Statistics Of Estimation Results With Approach-2 – Experiment-1 .....	108
Table 5.5. The Overall Statistical Results For The Five Scenarios – Experiment-1 .....	109
Table 5.6. The Summary Of Estimation Errors From All Scenarios – Experiment-2....	113

## List of Figures

Figure 1.1. Interrelations between Primary Research Tasks.....	4
Figure 2.1. A Typical Freeway Corridor .....	12
Figure 3.1. The Assumed Distribution of Travel Times for Drivers Departing from Node i during Time Interval k to Node j .....	33
Figure 3.2. The Probability of Travel Time Distribution.....	34
Figure 3.3. The Relationship between Mainline Volumes and Entrance Volumes .....	35
Figure 3.4. Travel Times $t_{ij}(k)$ and $t_{ij}(k)$ .....	37
Figure 3.5. A Small Example Freeway Section for Model Sensitivity Test.....	44
Figure 3.6a. Graphical Estimation Results with Different Sets of Initial Values –X02 ...	45
Figure 3.6b. Graphical Estimation Results with Different Sets of Initial Values –X03 ...	46
Figure 3.6c. Graphical Estimation Results with Different Sets of Initial Values –X02 ...	46
Figure 3.6d. Graphical Estimation Results with Different Sets of Initial Values –X12 ...	47
Figure 3.6e. Graphical Estimation Results with Different Sets of Initial Values –X13 ...	47
Figure 3.6f. Graphical Estimation Results with Different Sets of Initial Values –X14....	48
Figure 3.7a. Graphical Absolute Errors with Different Sets of Initial Values – X02.....	48
Figure 3.7b. Graphical Absolute Errors with Different Sets of Initial Values – X03.....	49
Figure 3.7c. Graphical Absolute Errors with Different Sets of Initial Values – X04.....	49
Figure 3.7d. Graphical Absolute Errors with Different Sets of Initial Values – X12.....	50
Figure 3.7e. Graphical Absolute Errors with Different Sets of Initial Values – X13.....	50

Figure 3.7f. Graphical Absolute Errors with Different Sets of Initial Values – X14 .....	51
Figure 3.8a. Graphical Estimation Results with Travel Time Variations-X02.....	53
Figure 3.8b. Graphical Estimation Results with Travel Time Variations-X03 .....	54
Figure 3.8c. Graphical Estimation Results with Travel Time Variations-X04.....	54
Figure 3.8d. Graphical Estimation Results with Travel Time Variations-X12 .....	55
Figure 3.8e. Graphical Estimation Results with Travel Time Variations-X13.....	55
Figure 3.8f. Graphical Estimation Results with Travel Time Variations-X14.....	56
Figure 3.9. I-95 Freeway Corridor between two Beltways.....	57
Figure 3.10. A Graphical Illustration of the I-95 Freeway Corridor .....	58
Figure 3.11. Entry Traffic Volume Distribution.....	60
Figure 3.12. Exit Traffic Volume Distribution .....	60
Figure 3.13. Mainline Traffic Volume Distribution .....	61
Figure 3.14a. Travel Time Distribution for each O-D Pair – Origin 0 .....	61
Figure 3.14b. Travel Time Distribution for each O-D Pair – Origin 1 .....	62
Figure 3.14c. Travel Time Distribution for each O-D Pair – Origin 2 .....	62
Figure 3.14d. Travel Time Distribution for each O-D Pair – Origin 3 .....	63
Figure 3.14e. Travel Time Distribution for each O-D Pair – Origins 4 and 5 .....	63
Figure 3.14f. Travel Time Distribution for each O-D Pair – Origins 6 and 7 .....	64
Figure 3.15. The Graphical Estimation Result for O-D Pair 01 .....	66
Figure 3.16. The Graphical Estimation Result for O-D Pair 23 .....	66
Figure 3.17. The Graphical Estimation Result for O-D Pair 46 .....	67
Figure 4.1. Flow Chart of the Decomposition Algorithm for Initial O-D Estimation.....	72

Figure 4.2. A General Freeway Network.....	73
Figure 4.3. An Example of a Decomposed Sub-network with Rule-1 .....	74
Figure 4.4. An Example of a Decomposed Sub-network with Rule-2 .....	75
Figure 4.5. An Example of a Decomposed Sub-network with Rule-3 .....	75
Figure 4.6. Example Estimation Results for O-D pair b03 in Case 1 .....	86
Figure 4.7. Example Estimation Results for O-D pair b13 in Case 2.....	86
Figure 4.8. Example Estimation Results for O-D pair b38 in Case 3.....	87
Figure 4.9. Example Estimation Results for O-D pair b12 in Case 4.....	87
Figure 4.10. Example Estimation Results for O-D pair b26 in Case 5.....	88
Figure 5.1. Data Generation Procedures.....	103
Figure 5.2. The Graphical Illustration of Estimation Results (b01) – Experiment-1 .....	110
Figure 5.3. The Graphical Illustration of Estimation Results (b38) – Experiment-1 .....	111
Figure 5.4. The Graphical Illustration of Estimation Results (b03) – Experiment-2 .....	114
Figure 5.5. The Graphical Illustration of Estimation Results (b07) – Experiment-2 .....	114
Figure 5.6. The Graphical Illustration of Estimation Results (b24) – Experiment-2 .....	115
Figure 5.7. The Graphical Illustration of Estimation Results (b45) – Experiment-2 .....	115

## CHAPTER 1 INTRODUCTION

### 1.1 BACKGROUND

To contend with deteriorating traffic conditions and manage the available capacity of transportation networks, there has been an increasing demand of time-dependent Origin-Destination (O-D) information that can be used for short-term traffic controls and analysis of day-to-day network flow patterns. For instance, a variety of traffic control applications, such as real-time route guidance, dynamic traffic assignment and freeway corridor control, utilize time-dependent O-D demands as one of the essential input information.

Traditionally, O-D information is estimated mainly from field survey data. However, due to the concern of costs and time associated with a rigorous survey, transportation researchers over the last several decades have devoted considerable efforts in developing effective approaches for reliably estimating time-dependent O-D demands. Depending on the employed assumptions and solution algorithms, one may classify those proposed in the literature for time-dependent O-D estimation into the categories of assignment-based and non-assignment-based approaches. The former category of approaches models the relationship between link flows and path flows with a link/path incidence matrix derived from a dynamic traffic assignment model. All approaches in this category require a reliable prior time-dependent O-D set and a proper route choice

behavior mechanism to predict the distribution of dynamic network O-D patterns (e.g. Ashok and Ben-Akiva, 2002).

Research methods in the latter category were developed in response to the practical difficulties of having a reliable prior O-D set and an accurate route choice behavior model. Some researchers have explored the potential of using only observable information such as the time series of volume counts and travel times from network links to estimate the time-dependent O-D distribution over a target time period. With such a modeling methodology, one can reduce the dependency of the prior O-D information and circumvent the need to have a dynamic traffic assignment model (e.g., Chang and Wu, 1994). Most modeling approaches in this category intend to directly formulate temporal and spatial relationship between time-dependent O-D patterns and the observable flows, such as on-ramp, off-ramp and/or mainline flows. Due to the underdetermined nature of such formulations, most solution algorithms proposed in the literature use the recursive estimation procedures to yield the best approximation of the true O-D distribution patterns. Hence, depending on the available information level and the network structure, the estimated time-dependent O-D patterns in this category may result in a large variance, and insufficient reliability for use in practice. Over the past two decades, despite the significant progress made by transportation professionals along this line, developing a reliable and effectiveness model for estimating time-dependent network O-D patterns remains a challenging issue in the transportation community.

## 1.2 RESEARCH OBJECTIVES

Theoretically, a dynamic O-D model ready for use in practice should have the following desirable features:

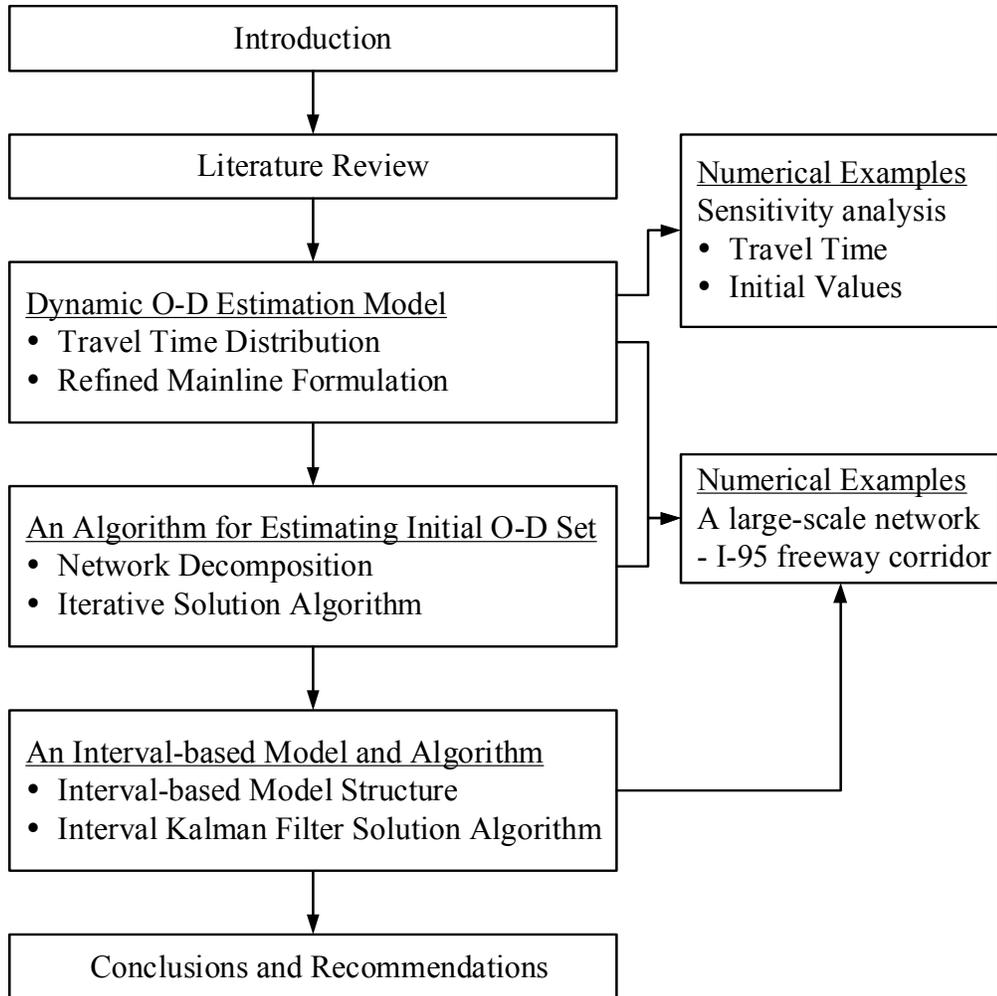
- 1 The proposed model can take full advantage of all available information and then well capture the dynamic interrelations between the estimated time-dependent O-D patterns and the observable information.
- 2 The formulations shall take into account real-world constraints, such as missing volume on some links or measurement errors of input data.
- 3 The solution algorithm shall be sufficiently robust and efficient in solving a network of realistic size.

Intending to embody all above desirable features in the proposed time-dependent network O-D model, this study has the following principal objectives:

- 4 Modeling the complex temporal and spatial interrelations between time-dependent O-D distributions and all observable information so as to increase the system observability with the minimal number of model parameters.
- 5 Developing a robust solution algorithm that can solve the proposed model formulations for a large network under commonly encountered constraints.
- 6 Demonstrating the applicability of the proposed model and algorithm for a large-scale freeway corridor with real-world system constraints, such as incomplete input information and measurement errors of the input data.

### 1.3 ORGANIZATION OF DISSERTATION

Based on the proposed research objectives, this study has organized the results of primary research tasks and key activities into 6 chapters. The interrelations among those tasks are illustrated in Figure 1.1.



**Figure 1.1. Interrelations between Primary Research Tasks**

Chapter 2 presents a comprehensive literature review of existing dynamic O-D estimation approaches, including both model formulations and solution algorithms. Advantages and limitations of those approaches are also addressed along with their potential enhancements in this chapter.

The primary focus of Chapter 3 is to develop a freeway dynamic O-D estimation model with an efficient solution algorithm that can circumvent those limitations identified in the literature review. The proposed estimation model can take into account the speed variance among vehicles, which have the same departure time, origin and destination, with an embedded travel time distribution function. This results in a substantial reduction of model parameters. To fully utilize both freeway mainline and ramp data, this study has further developed a set of mainline equations to capture the dynamic interrelations between the time-dependent O-D evolution patterns and the congestion level. Due to the nonlinear nature of the proposed formulations and concern with computing efficiency, this study has employed sequential extended Kalman filtering logic to develop a solution algorithm. Finally, extensive numerical analyses with both a small freeway network and the I-95 freeway corridor between Baltimore and Washington beltways have been conducted to test the sensitivity and reliability of the proposed model and its solution algorithm.

Chapter 4 presents an initial O-D estimation algorithm developed for use in refining randomly or arbitrarily generated initial O-D sets through the iterative estimation procedures so that it can yield a more accurate set of time-dependent O-D patterns over subsequent time intervals. Performance evaluation of the initial O-D estimation algorithm

with the example I-95 freeway corridor is also presented in this chapter. The carefully-designed experiment for evaluation contains 100 randomly generated initial O-D sets, and employs the proposed initial O-D algorithm to refine those 100 data sets for subsequent estimation of time-varying O-D distribution over the target time period.

Chapter 5 proposes an interval-based model and solution algorithm for estimating the time-varying O-D sets when the available traffic volume information actually vary within a range due to either measurement errors or some other factors. Performance evaluation of the proposed model and algorithm with the same I-95 freeway corridor is also presented in this chapter. This enhanced version of the dynamic O-D model and algorithm allow responsible agencies to circumvent the data deficiency embedded in most existing traffic surveillance systems. A set of rigorous numerical experiments has also been conducted and presented in this chapter.

Chapter 6 summarizes the contributions of this research and future directions, including the development of statistics for evaluating the reliability of a dynamic O-D estimation model, and an approach for solving the proposed model formulations when only partial sensor information is available.

## CHAPTER 2 LITERATURE REVIEW

### 2.1 INTRODUCTION

One of the most promising research directions for estimating O-D matrices is to directly formulate the relationship between O-D patterns and link counts because such information is readily available. Initially, most efforts on this regard were limited to developing methods applicable only for long-term transportation planning (Van Zuylen and Willumsen, 1980; Cremer and Keller, 1981; Masher, 1983; Bell 1983; Cascetta, 1984; Barbour and Fricker, 1994; Xu and Chan, 1993; Yang et al., 1991, 1992). However, these static methods did not take into account the time-varying nature of traffic flows and O-D trips. A comprehensive review of these static approaches can be found in the articles by Nguyen (1984), and Cascetta and Nguyen (1988).

There has been an increasing demand for time-varying O-D information that may be used in short-term traffic control and analysis due to the needs for improving traffic conditions and managing transportation networks. Since the actual number of variables to be estimated for either a static or dynamic system is always far over the available information, transportation researchers over the past two decades have explored various methods to tackle this difficult issue.

Nearly all O-D estimation approaches that are based on link flows consist of two primary steps. The first step is to formulate interrelationships between O-D patterns and link flows. Since the formulations for system equations generally do not have a unique

solution, the second step is to develop an algorithm to estimate the most feasible solution based on the employed assumptions and objective functions.

This review of the literature related to dynamic O-D estimation is organized as follows. Various model formulations for capturing the relationship between dynamic O-D matrices and links flows are presented in the next section. The estimation algorithms for these O-D formulations, including non-recursive and recursive estimation algorithms, are reviewed in Section 2.3. Finally, deficiencies of the existing approaches and potential idea for further developments are outlined in Section 2.4.

## **2.2 MODEL FORMULATIONS FOR DYNAMIC O-D ESTIMATION**

In the related literature review, recent studies for dynamic O-D estimation can be classified into two main categories: assignment and non-assignment based approaches (Wu, 1996).

### **2.2.1 Dynamic Traffic Assignment (DTA) Based Approaches**

In modeling dynamic O-D matrices for general networks, the DTA based approaches employ the assumptions that a reliable prior time-dependent O-D set and an accurate dynamic traffic assignment model predicting the route choice behavior are available. With such critical assumptions, the general formulations between dynamic O-D distributions and link flows can be described with the following equation:

$$y_{\ell}(k) = \sum_m \sum_{ij} P_{ij\ell}^m(k) T_{ij}(k-m) \quad (2.1)$$

where  $T_{ij}(k-m)$ : Demand departing from origin  $i$  at time interval  $k$  destined to  $j$ ;

$y_{\ell}(k)$ : Observed flows arriving on link  $\ell$  during time interval  $k$ ; and

$P_{ij\ell}^m(k)$ : Proportion of demand departing from origin  $i$  at time  $k-m$  to destination  $j$

that will arrive on link  $\ell$  at time  $k$ .

A comprehensive review of the DTA based approaches can be found in many existing publications (Ashok and Ben-Akiva, 1993, 2002; Tavana and Mahmassani, 2000; Peeta and Ziliaskopoulos, 2001). Some key research methods in this category are briefly reviewed below:

- Model of a Static Extension

To compute a dynamic O-D matrix, Willumsen (1984) first extended the static entropy maximization concept to a scenario of multiple time intervals. This class of approaches is known as the ‘‘Static Extension Model’’. In essence, statistical methods for static O-D estimation can be extended to dynamic cases (Cascetta et al., 1993; Casetta and Nguyen, 1988). The general model formulations are as follows:

$$\text{Min.} \sum_k \sum_{\ell} [y_{\ell}(k) - \sum_m \sum_{ij} P_{ij\ell}^m(k) T_{ij}(k-m)] \cdot [y_j(k) - \sum_m \sum_{ij} P_{ij\ell}^m(k) T_{ij}(k-m)] \quad (2.2)$$

This approach experienced problems of computational inefficiency and lack of consideration for the natural constraints. Keller and Ploss (1987) extended the static model to a dynamic model by utilizing the cross correlation between entry flows and exit flows as an estimate for the O-D proportion:

$$\bar{b}_{ij} = \frac{[\sum_k (q_i(k) - \bar{q}_i) \cdot (y_j(k) - \bar{y}_j)]^2}{[\sum_k (q_i(k) - \bar{q}_i)^2] \cdot [\sum_k (y_j(k) - \bar{y}_j)^2]} \quad (2.3)$$

where  $\bar{q}_i$ : The average value of  $q_i(k)$ ;

$q_i(k)$ : The entry flow at origin  $i$  during time period  $k$ ;

$\bar{y}_j$ : The average value of  $y_j(k)$ ; and

$y_j(k)$ : The exit flow at destination  $j$  during time period  $k$ .

It is reported that the accuracy of results decreases as the degree of correlation decreases.

- Auto-Regression Model

Another class of approaches employs an auto-regression assumption for the relationship between O-D pairs in successive intervals (Okutani, 1987; Okutani and Stephanedes, 1984; Ashok and Ben-Akiva, 1993, 2000, 2002). The state-space model is presented as follows:

$$T(k+1) = \sum_n A^n(k) \cdot T(k-n) + W(k) \quad (2.4)$$

$$Y(k) = \sum_m P^m(k) \cdot T(k-m) + V(k) \quad (2.5)$$

where  $T(k)=[\dots, T_{ij}(k), \dots]$  represents the time-varying O-D matrix;

$A^n(k)$  represents the corresponding auto-regression coefficient matrices;

$Y(k)=[\dots, y_\ell(k), \dots]$  represents the time varying link flows;

$P^m(k)=[\dots, P_{ij\ell}^m(k), \dots]$  represents the link use pattern of O-D trips resulting from a DTA application; and

$W(k)$  and  $V(k)$  represent random error and noise terms, respectively, of the dynamic system and are assumed to follow Gaussian distributions.

This dynamic O-D flow evolution model can be solved by the Kalman filtering algorithm to provide the recursive estimators. However, in order to calibrate the auto-regression model, sufficient historical data on the relationship between time intervals needs to be collected. A complete review of the auto-regression model can be found in Ashok (1996).

- AVI-based Models

In addition to the aforementioned models, some researchers have applied the Automatic Vehicle Identification (AVI) to dynamic O-D estimation models (Dixon and Rilett, 2000; Zhou and Mahmassani, 2005). They proposed that the proportions of some O-D pairs can be computed based on vehicles with onboard identification devices, thus, reducing some unknown parameters. However, this type of models is based on the premise that the demand pattern for the AVI data is equal to that for vehicles in the network.

In summary, all existing dynamic O-D estimation approaches based on the DTA concept have two major assumptions:

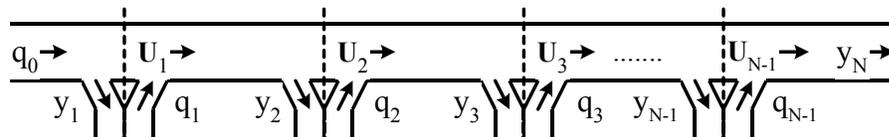
- Some reliable prior O-D information is required, and
- Accurate DTA parameters can be obtained.

This study intends to use the estimation approaches that only utilize the time series of available volume counts, and thus reduce the dependency on the prior O-D information. More in-depth review of these approaches is presented in the next section.

### 2.2.2 Dynamic Non-Traffic Assignment (Non-DTA) Based Approaches

Due to the practical difficulty in having a reliable prior O-D, most existing models classified in this category compute dynamic O-D matrices based on input/output flow relationship measured at network entrances and exits. Such modeling approaches use obvious relationships, such as the sum of O-D flows from the same origin equals the total trips entering the network from that origin, and the sum of O-D flows with the same destination is the total trips exiting the network at the target destination. A detailed description of the core formulation for this particular model category is presented below.

Consider a typical freeway corridor with link count information as shown in Figure 2.1, where detectors are placed at on-ramps, off-ramps, and mainline links. The information that is readily available for estimation of its time-dependent O-D flow proportion or dynamic O-D distribution is the time series of entering flow,  $q_i(k)$ , exiting flow,  $y_j(k)$ , and mainline flow,  $U_\ell(k)$ .



**Figure 2.1. A Typical Freeway Corridor**

Let  $b_{ij}(k)$  denote the proportion of vehicles entering from origin  $i$  to destination  $j$  during time interval  $k$ . By definition, it is subject to the following two natural constraints:

$$0 \leq b_{ij}(k) \leq 1, \quad 0 \leq i < j \leq N \quad (2.6)$$

$$\sum_{j=i+1}^N b_{ij}(k) = 1, \quad i = 0, 1, \dots, N-1 \quad (2.7)$$

The detailed reviews of the non-DTA based approaches are presented as follows:

- Simple Linear Model

For a small network, travel time from any origin to any destination is assumed to be negligible. The relation between entry and exiting flows is that the sum of O-D flows with the same destination is equal to the total trips exiting the network at the destination.

This relation can be formulated as follows (Cremer and Keller, 1981):

$$y_j(k) = \sum_{i=0}^{j-1} b_{ij}(k)q_i(k) \quad (2.8)$$

where  $q_i(k)$  : The number of vehicle trips entering the freeway from on-ramp  $i$  during time interval  $k$ .

$y_j(k)$  : The number of vehicle trips exiting the freeway from off-ramp  $j$  during time interval  $k$ .

$b_{ij}(k)$  : The proportion of  $q_i(k)$  heading toward destination node  $j$  during time interval  $k$ .

Note that the number of unknown variables for the example freeway as shown in Figure 2.1 is  $N \times (N+1)/2$ , and the number of equations for Equation (2.8) is  $N$ . Obviously, when  $N$  is greater than 1, the model is underdetermined as there are more unknown parameters  $\{b_{ij}(k)\}$  than system equations.

It should be noted that most models that are based on input/output flow, employ the assumption that travel time between origins and destinations is either constant or negligible. However, when the travel time is significantly longer so that it affects the input and output flow relationships, Equation (2.8) is no longer valid, and consequently, travel time factors must be explicitly captured in dynamic formulations.

- Linear Model with Travel Time Factors

In analyzing turning movements at intersections, Bell (1991a) modeled a travel time factor based on the platoon dispersion concept (Roberson, 1969). Bell's study assumed that the travel time needed for vehicles to pass through an intersection did not exceed one control time interval. Equation (2.9) was formulated based on the platoon dispersion relationship:

$$y_j(k) = (1 - \alpha_j) \cdot y_j(k-1) + \alpha_j \cdot q^T(k) \cdot b_j(k) \quad (2.9)$$

where  $\alpha_j$  is an additional smoothing parameter ( $0 \leq \alpha_j \leq 1$ ) and needs to be estimated.

This linear model captures the dynamic nature of small networks. However, the number of unknown variables has been increased to  $N \times (N+3)/2$ , and the number of system equations remains  $N$ .

- Linear Model with Freely Distributed Travel Times

The aforementioned model can only be applied to a small network, in which the longer travel time is within one control time interval. To contend with freely distributed travel times, Bell (1991b) further proposed an extended linear model. He introduced a new parameter,  $b_{ijm}$ , which denotes the proportion of trips from entrance  $i$  destined to exit  $j$  with a travel time of  $m$  intervals. Equation (2.10) is formulated based on the new parameter, and the new parameter is also subject to natural constraints (2.11) and (2.12).

$$y_j(k) = \sum_{m=0}^M \sum_{i=0}^{j-1} q_i(k-m) b_{ijm}(k), \quad j = 1, 2, \dots, N \quad (2.10)$$

$$\sum_{j=i+1}^N \sum_{m=0}^M b_{ijm}(k) = 1, \quad i = 0, 1, \dots, N-1 \quad (2.11)$$

$$0 \leq b_{ijm}(k) \leq 1, \quad 0 \leq i < j \leq N, \quad m = 0, 1, \dots, M \quad (2.12)$$

Equation (2.10) offers a more realistic formulation since the travel time for any O-D pair may distribute up to  $M$  control time intervals. However, if the travel time spans more than two time intervals, the system equations would involve too many parameters,  $b_{ijm}(k)$ .

- Non-linear Model with Mainline Traffic Flow

Chang and Wu (1994) proposed a freeway O-D estimation model by employing both mainline flow counts,  $U_\ell(k)$ , and ramp flow measurements,  $q_i(k)$  and  $y_j(k)$ , to construct a set of dynamic equations. To further capture the relationship between O-D flow proportions and traffic counts, they proposed a set of new variables,  $\theta_{ij}^-(k)$  and

$\theta_{ij}^+(k)$ , to represent the fraction of  $q_i(k-m)b_{ij}(k-m)$  trips that arrive at off-ramp  $j$  during time interval  $k$ . The model formulations are as follows:

$$y_j(k) = \sum_{m=0}^M \sum_{i=0}^{j-1} q_i(k-m)\theta_{ij}^m(k)b_{ij}(k-m) \quad j=1, 2, \dots, N \quad (2.13)$$

$$U_\ell(k) - q_\ell(k) = \sum_{m=0}^M \sum_{i=0}^{\ell-1} \sum_{j=\ell+1}^N q_i(k-m)\theta_{i\ell}^m(k)b_{ij}(k-m) \quad \ell=1, 2, N-1 \quad (2.14)$$

In this refined model, the number of unknown variables becomes  $(M+1) \times N \times (N+1)/2$  and the number of system equations increases to  $2N-1$ . To improve the operational efficiency, they also proposed an algorithm that aims to estimate an average O-D proportions,  $\bar{b}_{ij}(k)$ , over several consecutive time intervals as shown in the following equations:

$$y_j(k) = \sum_{i=0}^{j-1} \{q_i[k - t_{ij}^+(k)] \cdot \theta_{ij}^+(k) + q_i[k - t_{ij}^-(k)] \cdot \theta_{ij}^-(k)\} \cdot \bar{b}_{ij}(k), \quad j=1, 2, \dots, N \quad (2.15)$$

$$U_\ell(k) - q_\ell(k) = \sum_{i=0}^{\ell-1} \sum_{j=\ell+1}^N \{q_i[k - t_{i\ell}^+(k)] \cdot \theta_{i\ell}^+(k) + q_i[k - t_{i\ell}^-(k)] \cdot \theta_{i\ell}^-(k)\} \cdot \bar{b}_{ij}(k-m) \quad \ell=1, 2, N-1 \quad (2.16)$$

The number of unknown variables under these refined formulations is reduced to  $3N \times (N+1)/2$ . These formulations are based on the assumption that the speed of vehicles entering the freeway at the same time interval is distributed in a relatively small range.

## 2.3 SOLUTION ALGORITHMS FOR DYNAMIC O-D ESTIMATION

Since O-D flow proportions,  $b_{ij}(k)$ , cannot be solved uniquely due to the underdetermined nature of the above formulations, it is essential that some additional assumptions or estimation algorithms be employed to yield a feasible solution. The estimation algorithms can be classified into two categories: non-recursive and recursive computation methods.

### 2.3.1 Non-Recursive Estimation Algorithm

To overcome the underdetermined nature, some studies assume that certain relationships exist between O-D patterns during successive time intervals. As a result, the entire model can then be reformulated and solved with statistical methods such as generalized least squares and constrained least squares (Cremer, 1983; Cremer and Keller, 1981, 1984, 1987; Nihan and Davis, 1987, 1989). A brief description of these estimation algorithms is presented below.

- Least Squares Approaches

Given an observation of consecutive  $K$  intervals in a dynamic O-D system, one can construct a least-square estimate, such as an ordinary least squares estimate (Cremer and Keller, 1981, 1987), a constrained least squares estimate (Nihan and Davis, 1987), and a generalized constrained least squares estimate (Kessaci et al., 1989).

The ordinary least-squares approach is the most fundamental method, in which the O-D proportions are computed by minimizing the difference between detected traffic volumes and estimated traffic volumes as follow:

$$\text{Min.} \sum_{k=1}^K \sum_{j=i+1}^N [y_j(k) - b_j^T q(k)] \cdot [y_j(k) - b_j^T q(k)] \quad (2.17)$$

This approach does not take into account the natural constraints shown in Equations (2.6) and (2.7). Moreover, it requires computation of the matrix inversion. By taking into consideration of the natural constraints, the constrained least squares approach can yield relatively reliable estimation at the cost of an intensive computation burden (Nihan and Davis, 1987).

Kessaci et al. (1989) proposed a more generalized model for constrained least-square estimates:

$$\text{Min.} \sum_{k=1}^K \sum_{j=i+1}^N [y_j(k) - b_j^T q(k)] \cdot a_{jk} \cdot [y_j(k) - b_j^T q(k)] \quad (2.18)$$

However, this model is even more computationally intensive than all previous methods due to the considerably high number of matrix inversion operation required.

- Maximum Likelihood Approach

The Maximum Likelihood (ML) approach is based on the assumption that O-D flow proportions are equal to probabilities of flows entering from entrance  $i$  to exit  $j$  as shown in Equation (2.19).

$$\text{Max.} P[y(1), y(2), \dots, y(t) | \bar{b}(k)] \quad (2.19)$$

To compute the ML estimator, it is necessary to compute the deviations of these probabilities, which are unfortunately unavailable from the observed data. To contend with this issue, Nihan and Davis (1989) employed the Expectation Maximization (EM) algorithm proposed by Dempster et al. (1977) for computing the ML estimator. This

approach assumes that O-D proportions are constant over time and requires the computation of the covariance matrices during each time iteration.

- Fixed-point Approach

Since the ML approach requires quite intensive computational work, Nihan and Hamed (1992) later proposed a fixed point approach, which simplified the ML algorithm so that each cell of the matrix can be estimated separately as the following iteration function:

$$\hat{b}_{ij}(k+1) = \hat{b}_{ij}(k) + d / \sum_k q_i(k) \quad (2.20)$$

$$d = \sum_t \{b_{ij}(1-b_{ij})q_i(t)[y_j(t) - \sum_t b_{ij}q_i(t)] / \sum_i b_{ij}(1-b_{ij})q_i(t)\} \quad (2.21)$$

It is reported that this approach produced generally lower variances and more accurate estimates than with the least square-based approach (Nihan and Hamed, 1992).

- Correlation Approach

The correlation approach is based on static O-D estimation models (Ploss and Keller, 1986; Keller and Ploss, 1987). An initial estimate of a O-D matrix from the time series of input/output flows is first computed with a cross-correlation concept. Then, this prior O-D matrix is updated with the traffic counts of the current time interval. The O-D proportion is computed by the cross correlation between entry flows and exit flows:

$$b_{ij}(k) = \frac{[\sum_{t=k-T}^k \Delta q_i(t) \cdot \Delta y_j(t)]^2}{[\sum_{t=k-T}^k \Delta q_i^2(t) \cdot \sum_{t=k-T}^k \Delta y_j^2(t)]} \quad (2.22)$$

This approach was claimed (Ploss and Keller, 1986) to have a slight advantage over the recursive estimation algorithm by Cremer and Keller (1981).

### 2.3.2 Recursive Estimation Algorithm

All estimation algorithms in this category have the following recursive relation:

Given initial  $b_{ij}(0)$ , update each  $b_{ij}(k)$  recursively with the following:

$$\hat{b}(k+1) = \hat{b}(k) + G(k) \cdot [y(k+1) - q^T(k+1) \cdot \hat{b}(k)] \quad (2.23)$$

where  $G(k)$  is the gain term for updating the  $b_{ij}(k+1)$ . A variety of algorithms has been proposed for computing the  $G(k)$  matrix. Some key studies on this category are reviewed below:

- Stochastic Gradient (Gauss-Newton) Approach

Cremer and Keller (1987) proposed the stochastic gradient algorithm that is intended to minimize the expected prediction error variance for each time iteration.

$$V(b_j) = E[1/2[y_j(k) - q^T(k)b_j(k)]^2] \quad (2.24)$$

The gradient of the  $V(b_j)$  is then

$$\frac{\partial V(b_j)}{\partial b_j} = E\{-q(k) \cdot [y_j(k) - q^T(k)b_j(k)]\} \quad (2.25)$$

The estimator and gain term are:

$$\hat{b}_j(k) = \hat{b}_j(k-1) + G(k) \cdot q(k)[y_j(k) - q^T(k)\hat{b}_j(k-1)] \quad (2.26)$$

Generally, this approach shows faster convergence and can be more robust with respect to choices of initial values and other design parameters (Nihan and Davis, 1987).

However, this approach does not consider the natural constraints, and thus Nihan and Davis (1987) reported that it performs worse and could be improved if the natural constraints are taken into account.

- Recursive Least Squares Approaches

This type of approach, basically, is the recursive form of the aforementioned least squares method. Nihan and Davis (1987) were the pioneers credited with developing the recursive least squares algorithm. Its core concept is to recursively minimize the least-square estimator  $[y(k) - q^T(k)b]^2$ . The estimator and the gain term are defined as follows:

$$\hat{b}_j(k) = \hat{b}_j(k-1) + G_j(k)[y_j(k) - q^T(k)\hat{b}_j(k-1)] \quad (2.27)$$

$$P_j(k) = P_j(k-1) - \frac{P_j(k-1)q(k)q^T(k)P_j(k-1)}{1 + q^T(k)P_j(k-1)q(k)} \quad (2.28)$$

$$G_j(k) = \frac{P_j(k-1)q(k)}{1 + q^T(k)P_j(k-1)q(k)} \quad (2.29)$$

Kessaci et al. (1989) derived the generalized recursive least squares algorithm for their proposed generalized least-square model. The approach has the advantage in satisfying both sets of the natural constraints, i.e., Equations (2.6) and (2.7). However, due to the need to invert the matrix in the gain matrix, this estimation procedure is computationally more intensive than most other recursive algorithms.

To overcome this shortcoming, Nihan and Davis (1989) proposed the following two-stage process to incorporate the natural constraints in the estimation procedure:

Step 1 – Truncation. Modifying the recursive formulation to:

$$\hat{b}_j(k+1) = \hat{b}_j(k) + \alpha \cdot G_j(k) \cdot [y_j(k+1) - q^T(k+1) \cdot \hat{b}_j(k)] \quad (2.30)$$

where  $\alpha$  is chosen as the biggest scalar ( $0 \leq \alpha \leq 1$ ) to satisfy the inequality constraints, i.e.,  $0 \leq \hat{b}_{ij}(k) \leq 1$ .

Step 2 – To satisfy the equality constraint by using either

(a) Normalization:

$$\hat{b}_j(k) = \hat{b}_j(k) / \sum_j \hat{b}_{ij}(k), \text{ or} \quad (2.31)$$

(b) Projection:

$$\hat{b}_{ij}(k) = \hat{b}_{ij}(k) + [1 - \sum_j \hat{b}_{ij}(k)] / N \quad (2.32)$$

It is reported that this results in greater accuracy (Nihan and Davis, 1987). This two-step process can also be applied to other recursive algorithms.

- Bayesian Updating Approach

Maher (1983) was the first researcher who advocated the use of Bayesian statistical inference in O-D estimation. Instead of starting with a point estimate, he introduced a distribution of possible initial estimates in order to represent the degree belief in these prior probabilities. A posterior distribution of the possibilities was then produced from the prior distribution and observations using the Bayesian Theorem. Van der Zijpp (1996) proposed the Bayesian updating scheme using the multivariate normal assumption but in the truncated multivariate normal form for the subject probability distributions.

The advantage of this new procedure is that it deals with the inequality constraints in an appropriate statistical manner. This approach is an extended version of the Kalman filtering that incorporates the same inequality constraints. The core concept is to maximize the posterior distribution density of  $b_j(k)$  by solving the following optimization problem:

$$\text{Min. } \{[b_j(k) - \hat{b}_j(k)]^T S_j [b_j(k) - \hat{b}_j(k)] \mid 0 \leq \hat{b}_{ij}(k) \leq 1\} \quad (2.33)$$

where  $S_j$  is the estimated covariance matrix of  $\hat{b}_j(k)$ . Although this proposed method offers flexibility in the degree of belief on the prior estimate, the value,  $S_j$ , still needs to be determined in practice.

- Kalman Filtering Approach

The Kalman filtering approach is different from the aforementioned approaches, which are designed to minimize the distance between measured and predicted values. The method has been applied to O-D estimation problems by a number of researchers, namely Cremer and Keller (1987), Nihan and Davis (1987), Van der Zijpp and Hanerslag (1994), and Chang and Wu (1994). Prior to using any Kalman filtering method, one must provide the state and measurement equations. The former describes how the unknown parameters have evolved over time, while the latter captures the relationship between the unknown parameters and the measurements. In both equations, it is possible to specify uncertainty with noise terms. The following are examples of the state and measurement equations:

$$b(k+1) = b(k) + w(k) \quad (2.34)$$

$$y(k) = H^T(k)b(k) + v(k) \quad (2.35)$$

where Equation (2.34) is the state equation assumed to follow the random walk and Equation (2.35) is the measurement equation describing the relation between traffic flows and O-D proportions. Given the state and measurement equations, one can define the recursive estimator,  $\hat{b}(k)$ , with the Kalman filtering concept as follows (Kalman, 1960).

$$\hat{b}(k) = \hat{b}(k-1) + G(k) \cdot [y(k) - H^T(k)\hat{b}(k-1)] \quad (2.36)$$

$$G(k) = P(k-1)H(k)[H^T(k)P(k-1) + R(k)]^{-1} \quad (2.37)$$

$$P(k) = P(k-1) - P(k-1)H(k)[H^T(k)P(k-1)H(k) + R(k)]^{-1}H^T(k)P(k-1) \quad (2.38)$$

These equations define a recursion that begins with an initial estimate,  $\bar{b}(0)$ , an initial covariance matrix,  $P(0)$ , and the covariance matrices of those random terms,  $Q(k)$  and  $R(k)$ . In practice, these can be provided using historical information or available model experience. A theoretical analysis on this approach can be found in van der Zijpp and Hamerslag (1994).

Moreover, this estimate has been shown to be unbiased. If the noise terms and the initial state follow the Gaussian distribution, the Kalman filtering has been shown to produce unbiased estimates that yield minimum variance over all other estimators (Anderson and Moore, 1979).

A fundamental problem associated with the application of the Kalman filtering equations to estimating O-D proportions lies in the difficulty of ensuring that the natural inequality and equality constraints are satisfied. The scheme of normalization and truncation proposed by Nihan and Davis (1989) can be applied to deal with this problem.

## 2.4 CLOSURE

Existing approaches for estimating dynamic O-D matrices from time-series of traffic counts have been summarized and categorized into two classes based on input information and formulations. The main weakness of DTA-based approaches lies in that it requires both reliable prior O-D information and an accurate traffic assignment model. In addition, since the development of a reliable DTA remains an on-going research issue, the requirement for having a DTA pattern certainly limits the potential application of such O-D estimation approaches.

In contrast, the dynamic approaches with input/output flows are relatively promising since they need neither DTA nor a prior O-D matrix. Unfortunately, such approaches reported in the literature are effective mainly for special networks in which travel time is either negligible, constant, or with a limited range of variation.

In conclusion, to advance existing models for real-world applications, one needs to overcome the following three critical issues:

- The system equations for O-D estimation from traffic counts are clearly underdetermined as the number of equations is always far less than the number of O-D pairs.
- An efficient model for dynamic O-D estimation shall have the capability to deal with a large-scale freeway network.
- A practically useful model shall not be based on the unrealistic assumption that all entry and exiting flow counts are available, or that a reliable set of

prior O-D exists for model calibration. In reality, such information may be neither complete nor accurate at the desirable level.

This research is proposed to address these three critical issues, and is focused on developing a reliable and robust model that offers the potential for real-world applications.

## **CHAPTER 3 MODEL FORMULATIONS OF DYNAMIC O-D ESTIMATION FOR FREEWAY CORRIDORS**

### **3.1 INTRODUCTION**

Since the number of parameters to be estimated for both static and dynamic system is always greater than the available information, transportation researchers over the past two decades have studied various methods for contending with this difficult issue. Based on related literature reviewed in Chapter 2, one can classify recent studies for dynamic O-D estimation into two main categories: assignment and non-assignment based approaches. The former category of approaches is based on the assumption that a reliable prior O-D set and a dynamic traffic assignment model that predicts route choice behavior are available. Considering the practical difficulty in having the reliable prior O-D information, some researchers have developed various estimation approaches that utilize only the time series of available volume counts, thus reducing the dependency on prior O-D information. This study focuses on non-assignment-based methods, and aims to estimate dynamic freeway O-D distribution, based mainly on all observable link and ramp flow rates.

Note that if the travel time needed to traverse the network is constant, the dynamic O-D estimation for this type of scenario regresses back to the case of an isolated intersection or a small network, except when there is a constant time lag between each O-D pair. However, under congested conditions, the link travel time may vary substantially

with flow rates. This impact on the distribution of time-varying O-D patterns thus needs to be properly incorporated in the model formulation. Moreover, the measurable input and output flows at ramps is based on the assumption that no congestion exists on the freeway segment. To address these two critical issues, this study proposes a robust model that captures the speed variance among vehicles with the same departure time, origin, and destination using a specially derived travel time distribution function that substantially reduces model parameters. To fully utilize traffic flow information on the mainline freeway, this study refines the mainline equations to capture dynamic interrelations between the O-D evolution pattern and the congestion level.

The topics covered in this chapter are organized as follows: The basic relations between time-dependent O-D flows and time-series traffic measurements in a freeway corridor are formulated with a nonlinear dynamic system model as illustrated in Section 3.2. In Section 3.3, an enhanced model for the freeway corridor is developed using an embedded function that allows travel times to exceed one unit control interval thereby capturing varying travel times. The base solution algorithm with extended Kalman filtering procedures is illustrated in Section 3.4. Evaluations and results on the effectiveness of the proposed model and algorithm are presented in Section 3.5 using extensive simulation experiments. The last section of this chapter summarizes current studies and key contributions.

### 3.2 NATURE OF PROBLEM AND BASIC FORMULATIONS

Consider a freeway corridor of  $N$  segments from 0 to  $N-1$  as shown in Figure 2.1. The set of variables used in modeling the dynamic traffic flow and O-D relations is defined as follows:

$q_0(k)$  : The number of vehicles entering the upstream boundary of the freeway section during time interval  $k$ .

$q_i(k)$  : The number of vehicles entering freeway from on-ramp  $i$  during time interval  $k$ ,  $i = 1, 2, \dots, N-1$ .

$y_j(k)$  : The number of vehicles leaving freeway from off-ramp  $j$  during time interval  $k$ ,  $j = 1, 2, \dots, N-1$ .

$y_n(k)$  : The mainline volume at the downstream end of the freeway section during time interval  $k$ .

$U_i(k)$  : The number of vehicles crossing the upstream boundary of segment  $i$  during time interval  $k$ ,  $i = 1, 2, \dots, N-1$ .

$T_{ij}(k)$  : The number of vehicles entering the freeway from on-ramp  $i$  during time interval  $k$  that are destined to off-ramp  $j$  (i.e., the time-dependent O-D flow), where  $0 \leq i < j \leq N$ .

$t_0$  : The duration of one unit time interval.

$t_{ij}(k)$  : The average travel time from on-ramp  $i$  to off-ramp  $j$  departing during time interval  $k$ .

$\sigma_{ij}(k)$  : The standard deviation of the travel time for vehicles traveling from on-ramp  $i$  to off-ramp  $j$  departing during time interval  $k$ .

$b_{ij}(k)$  : The proportion of  $q_i(k)$  heading toward destination node  $j$  during time interval  $k$ .

$\theta_{ij}^m(k)$  : The fraction of  $T_{ij}(k-m)$  vehicles that arrive at off-ramp  $j$  during time interval  $k$ .

Exploiting the preceding notations and using Figure 2.1, one can establish the following relationships:

$$q_i(k) = \sum_{j=i+1}^N T_{ij}(k), \quad i = 0, 1, \dots, N-1 \quad (3.1)$$

$$T_{ij}(k) = q_i(k)b_{ij}(k), \quad 0 \leq i < j \leq N \quad (3.2)$$

Equations (3.1) and (3.2) are subject to the following constraints as previously discussed:

$$0 \leq b_{ij}(k) \leq 1, \quad 0 \leq i < j \leq N \quad (3.3)$$

$$\sum_{j=i+1}^N b_{ij}(k) = 1, \quad i = 0, 1, \dots, N-1 \quad (3.4)$$

Regarding speed variations among drivers, it is reasonable to assume that the departure time for vehicles from node  $i$  arriving at node  $j$  during time interval  $k$  are distributed among time intervals  $k, k-1, \dots, \text{and } k-M$ , where  $M$  is the maximum number of intervals required for vehicles to traverse the entire freeway section. The exit traffic volume,  $y_j(k)$ , can thus be stated as:

$$\begin{aligned}
y_j(k) &= \sum_{m=0}^M \sum_{i=0}^{j-1} T_{ij}(k-m) \theta_{ij}^m(k) \\
&= \sum_{m=0}^M \sum_{i=0}^{j-1} q_i(k-m) \theta_{ij}^m(k) b_{ij}(k-m), \quad j=1, 2, \dots, N
\end{aligned} \tag{3.5}$$

where  $\theta_{ij}^m(k)$ , a set of new time-dependent parameters satisfies the following constraints:

$$0 \leq \theta_{ij}^m \leq 1, \quad 0 \leq i \leq j \leq N, \quad m = 0, 1, \dots, M \tag{3.6}$$

$$\sum_{m=0}^M \theta_{ij}^m(k+m) = 1, \quad 0 \leq i < j \leq N \tag{3.7}$$

As discussed in the study by Chang and Wu (1994), Equation (3.5) is sufficient for capturing the dynamic relation between O-D patterns and link flows if the freeway is not congested and traffic flow is steady. Otherwise, the time varying traffic volume,  $U_\ell(k)$ , cannot be determined with only the entrance and exit flow data,  $q_i(k)$  and  $y_j(k)$ . Hence, the measurements of  $\{U_\ell(k)\}$  may actually provide additional valuable information for estimation. A set of constraints that utilizes the mainline traffic volume,  $U_\ell(k)$ , is given as follows (Chang and Wu, 1994):

$$\begin{aligned}
U_\ell(k) &= \sum_{m=0}^M \sum_{i=0}^{\ell-1} \sum_{j=\ell+1}^N q_i(k-m) \theta_{i\ell}^m(k) b_{ij}(k-m) + q_\ell(k) \\
U_\ell(k) - q_\ell(k) &= \sum_{m=0}^M \sum_{i=0}^{\ell-1} \sum_{j=\ell+1}^N q_i(k-m) \theta_{i\ell}^m(k) b_{ij}(k-m)
\end{aligned} \tag{3.8}$$

where  $\ell = 1, 2, \dots, N-1$ .

However, the system formulation contains a large number of the unknown parameters, i.e.,  $b_{ij}(k)$  and  $\theta_{ij}^m(k)$ . The number of the unknown parameters increases as

the required M value increases. Therefore, more refinement is necessary to ensure that the proposed model is computationally efficient and tractable. An innovative concept that utilizes a function to capture a potentially wide range of speed variations with a limited number of unknown parameters is proposed in the next section.

### 3.3 THE DYNAMIC O-D ESTIMATION MODEL WITH TRAVEL TIME DISTRIBUTION

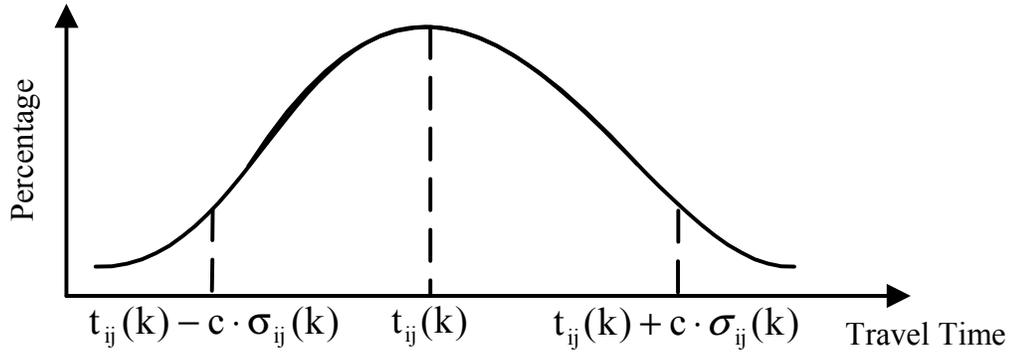
Assume that the travel time of drivers departing from node  $i$  during time interval  $k$  to node  $j$  follow a normal distribution, i.e.,  $N[\mu_{ij}(k), \sigma_{ij}^2(k)]$ , as shown in Figure 3.1, where:

$\mu_{ij}(k) = t_{ij}(k)$  : The average travel time of vehicles departing from node  $i$  during interval  $k$  to node  $j$

$\sigma_{ij}(k)$  : The standard deviation of the travel time of vehicles departing from node  $i$  during interval  $k$  to node  $j$

$\rho_{ij}^m(k)$  : The fraction of  $T_{ij}(k-m)$  vehicles departing from entry node  $i$  during time interval  $k$  that takes  $m$  time intervals to exiting node  $j$ .

$\rho_{i\ell_j}^m(k)$  : The fraction of  $T_{ij}(k-m)$  trips from entry node  $i$  during time interval  $k$  that takes  $m$  time intervals to mainline node  $\ell$ .



**Figure 3.1. The Assumed Distribution of Travel Times for Drivers Departing from Node i during Time Interval k to Node j**

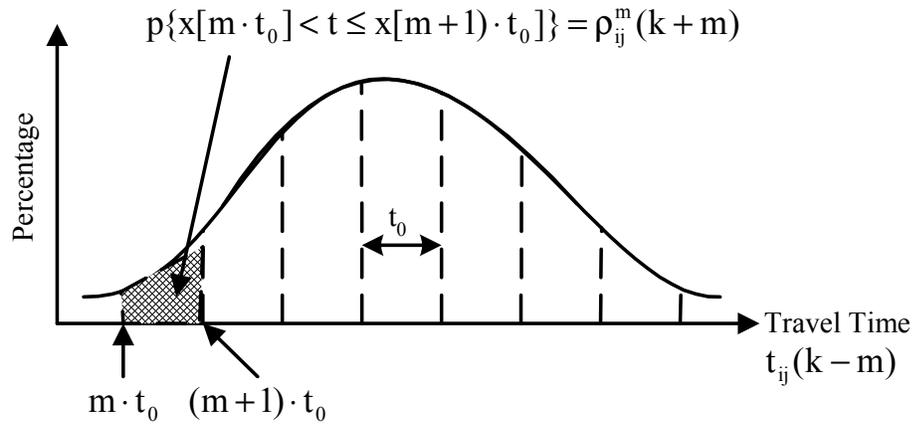
The use of normal distributions to approximate the travel time distribution of vehicles with the same O-D has previously been reported in literature by Bell (1991b), and Zhang and Maher (1998) as well as Grace and Potts (1964). Furthermore, Seddon (1972) examined the theoretical basis for the recurrence model and found that it corresponded to Pacey's (1956) diffusion model of platoon dispersion when the normal distribution for vehicle speeds was replaced with a shifted geometric distribution for travel times. However, this concept has not yet been applied to formulate the O-D estimation model.

As shown in Figure 3.2, since the travel time for an O-D pair departing during the same time interval follows a normal distribution,  $\rho_{ij}^m(k)$  can be replaced with a cumulative density function within a time interval,  $m$ , as follows:

$$\rho_{ij}^m(k) = \int_{m \cdot t_0}^{(m+1) \cdot t_0} f_{ij}(x) dx \quad (3.9)$$

$$\sum_{m=0}^M \rho_{ij}^m(k+m) = 1 \quad (3.10)$$

where  $0 \leq \rho_{ij}^m(k) \leq 1$ ,  $0 < i \leq j \leq N$ , and  $m = 0, 1, \dots, M$ .

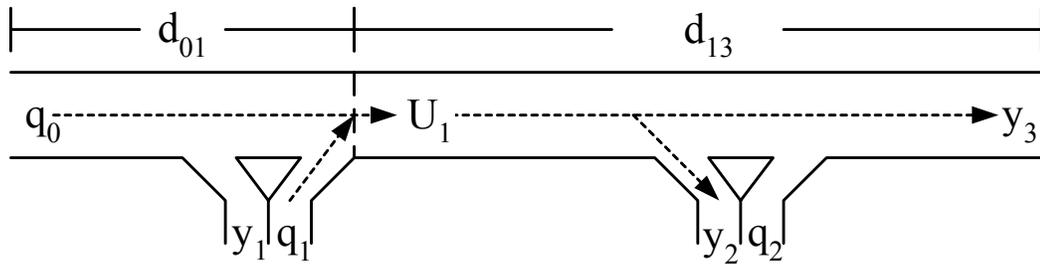


**Figure 3.2. The Probability Distribution of Travel Time**

By applying the above travel time distribution concept, Equations (3.5) can be rewritten as:

$$\begin{aligned} y_j(k) &= \sum_{m=0}^M \sum_{i=0}^{j-1} q_i(k-m) \rho_{ij}^m(k) b_{ij}(k-m) \\ &= \sum_{m=0}^M \sum_{i=0}^{j-1} q_i(k-m) \cdot \left[ \int_{m \cdot t_0}^{(m+1) \cdot t_0} f_{ij}(x) dx \right] \cdot b_{ij}(k-m), \quad 0 < i \leq j \leq N \end{aligned} \quad (3.11)$$

where  $f(x)$  is the density function of the travel time distribution with mean  $\bar{t}_{ij}(k)$  and standard deviation  $\sigma_{ij}(k)$ . To apply the travel time distribution concept to Equation (3.8), it is essential that the formulations are properly restructured to minimize the number of unknown parameters. Figure 3.3 illustrates an example of the relation between mainline volumes and entrance volumes.



**Figure 3.3. The Relation between Mainline Volumes and Entrance Volumes**

As shown in Figure 3.3, the mainline volume  $U_1$  is comprised of the vehicles traveling from Origin 0 to Destination 2 and Destination 3, and entrance volume  $q_1$ . If the speeds of vehicles are uniformly distributed, Equation (3.8) still holds. However, since normal distribution is applied to illustrate the travel time distribution, the fraction of trips in Equation (3.8) that arrive at the mainline segment is different from the fraction of trips that arrive at the off-ramp. A new parameter,  $\rho_{ij}^m(k)$ , is then introduced, which is defined as the fraction of  $T_{ij}(k)$  trips from on-ramp  $i$  during time interval  $k$  that takes  $m$  time intervals to mainline segment  $\ell$ . The previous equation can thus be modified as follows:

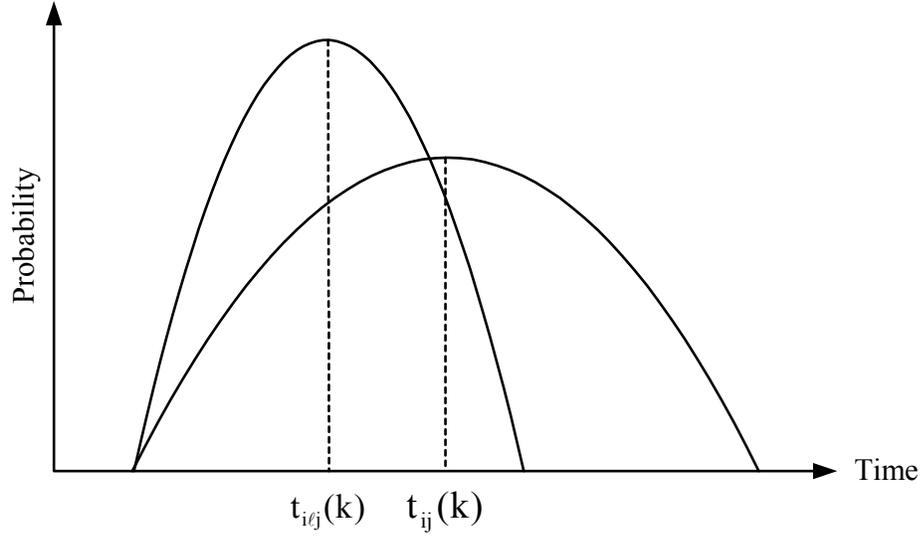
$$U_\ell(\mathbf{k}) - q_\ell(\mathbf{k}) = \sum_{m=0}^M \sum_{i=0}^{\ell-1} \sum_{j=\ell+1}^N q_i(\mathbf{k} - m) \rho_{i\ell j}^m(\mathbf{k}) b_{ij}(\mathbf{k} - m) \quad (3.12)$$

However, the contributions on this critical research issue would be more meaningful if additional observation constraints were identified without increasing the number of unknown parameters. Hence, this study has further assumed that the travel times for vehicles from origin  $i$  to mainline point  $\ell$ , i.e.,  $t_{i\ell}(\mathbf{k})$ , also follow a normal distribution, and one thus is able to construct the relation between the two fractions,  $\rho_{ij}^m(\mathbf{k})$  and  $\rho_{i\ell j}^m(\mathbf{k})$ . Assuming that the speed distribution for vehicles traveling from origin  $i$  to destination  $j$  remains unchanged, the relation between travel times  $t_{ij}(\mathbf{k})$  and  $t_{i\ell}(\mathbf{k})$  is the proportion of the corresponding distances as shown in Figure 3.3. Therefore, the distribution of travel time  $t_{i\ell}(\mathbf{k})$  can be presented as:

$$t_{i\ell}(\mathbf{k}) \sim N[\gamma_{i\ell j} \cdot \bar{t}_{ij}(\mathbf{k}), \gamma_{i\ell j}^2 \cdot \sigma_{ij}^2(\mathbf{k})] \quad (3.13)$$

$$\gamma_{i\ell j} = \frac{d_{i\ell}}{d_{ij}} \quad (3.14)$$

where  $\gamma_{i\ell j}$  is the ratio of the distance  $d_{i\ell}$ , to the distance  $d_{ij}$ . Figure 3.4 shows an example of travel times  $t_{ij}(\mathbf{k})$  and  $t_{i\ell}(\mathbf{k})$ .



**Figure 3.4. Travel Times  $t_{ij}(k)$  and  $t_{i'j}(k)$**

In addition to use normal distribution to demonstrate the variation of travel time, one can also estimate the average  $\bar{b}_{ij}(k)$  for consecutive intervals instead of solving the O-D flow distribution matrix for each small interval (Chang and Wu, 1994). Hence, all the  $b_{ij}(\cdot)$  terms in equations (3.11) and (3.12) can be replaced with  $\bar{b}_{ij}(k)$ :

$$\begin{aligned}
 y_j(k) &= \sum_{m=0}^M \sum_{i=0}^{j-1} [q_i(k-m) \rho_{ij}^m(k)] \cdot \bar{b}_{ij}(k) \\
 &= \sum_{m=0}^M \sum_{i=0}^{j-1} \left\{ q_i(k-m) \cdot \int_{m-t_0}^{(m+1)t_0} f_{ij}(x) dx \right\} \cdot \bar{b}_{ij}(k) \quad (3.15)
 \end{aligned}$$

$$\begin{aligned}
U_\ell(\mathbf{k}) - q_\ell(\mathbf{k}) &= \sum_{m=0}^M \sum_{i=0}^{\ell-1} \sum_{j=\ell+1}^N [q_i(\mathbf{k}-m) \rho_{ij}^m(\mathbf{k})] \cdot \bar{b}_{ij}(\mathbf{k}) \\
&= \sum_{m=0}^M \sum_{i=0}^{\ell-1} \sum_{j=\ell+1}^N \{q_i(\mathbf{k}-m) \cdot \int_{m \cdot t_0}^{(m+1) \cdot t_0} f_{ij}(x) dx\} \cdot \bar{b}_{ij}(\mathbf{k})
\end{aligned} \tag{3.16}$$

With the above enhancements, the average travel time for each O-D pair can be estimated with data provided by a surveillance system only leaving the O-D proportions,  $\bar{b}_{ij}(\mathbf{k})$ , and standard deviations,  $\sigma_{ij}(\mathbf{k})$  as the unknown set of parameters. An estimation algorithm based on the extended Kalman filtering concept is presented in the next chapter.

### 3.4 THE STATE-SPACE MODELING AND KALMAN FILTERING ESTIMATION ALGORITHM

In most existing approaches, dynamic O-D parameters,  $b_{ij}(\mathbf{k})$ , are assumed to follow the random walk process:

$$b_{ij}(\mathbf{k}+1) = b_{ij}(\mathbf{k}) + w_{ij}(\mathbf{k}), \quad 0 \leq i < j \leq N \tag{3.17}$$

where the random term,  $w_{ij}(\mathbf{k})$  is an independent Gaussian white noise sequence with zero mean and its covariance. By the same token, one can deduce a similar dynamic equation for  $\sigma_{ij}(\mathbf{k})$ :

$$\sigma_{ij}(\mathbf{k}+1) = \sigma_{ij}(\mathbf{k}) + v_{ij}(\mathbf{k}), \quad 0 \leq i < j \leq N \tag{3.18}$$

where each  $v_{ij}(\mathbf{k})$  is a Gaussian white noise sequence with zero mean.

It is quite complex to apply the dynamic relationships of  $\dot{b}_{ij}(k)$  to a matrix form because there exist time lag variables in Equations (3.15) and (3.16) that must be eliminated. However, with a simplified model proposed by Wu and Chang (1997) that estimates an average  $b_{ij}(k)$  for a consecutive number of the most recent time intervals, the time lag problem is no longer critical. To facilitate the formulation, the following variables are defined as:

$$\mathbf{b}(k) = [b_{ij}(k)]$$

$$\mathbf{W}(k) = [w_{ij}(k)]$$

Note that both  $\mathbf{b}(k)$  and  $\mathbf{W}(k)$  are column vectors of dimension  $N(N+1)/2$ .

Hence, the matrix form of Equation (3.17) is as follows:

$$\mathbf{b}(k+1) = \mathbf{b}(k) + \mathbf{W}(k) \quad (3.19)$$

where  $\mathbf{W}(k)$  is the corresponding white noise sequence with zero mean and the corresponding covariance matrix,  $\mathbf{D}_b = \text{diag}[d_b, \dots, d_b]$  is a  $N(N+1)/2$  dimensional matrix. Let  $\boldsymbol{\sigma}(k)$  denote the following  $N(N+1)/2$  dimension column vector:

$$\boldsymbol{\sigma}(k) = [\sigma_{01}(k), \sigma_{02}(k), \dots, \sigma_{0N}(k), \sigma_{12}(k), \dots, \sigma_{N-1,N}(k)]^T$$

and the vector  $\mathbf{V}(k)$  be the corresponding vectors in the following matrix form:

$$\boldsymbol{\sigma}(k+1) = \boldsymbol{\sigma}(k) + \mathbf{V}(k) \quad (3.20)$$

where  $\mathbf{V}(k)$  is an  $N(N+1)/2$  dimension Gaussian white noise vector with zero mean and covariance  $\mathbf{D}_\sigma$ , which is a constant semi-positive matrix.

With the above refinements for  $\mathbf{b}(\mathbf{k})$  and  $\boldsymbol{\sigma}(\mathbf{k})$ , Equations (3.15) and (3.16) can be restructured into the following matrix form:

$$\mathbf{Z}(\mathbf{k}) = \mathbf{H}[\boldsymbol{\sigma}_{ij}(\mathbf{k})] \cdot \mathbf{b}(\mathbf{k}) + \mathbf{e}(\mathbf{k}) \quad (3.21)$$

where  $\mathbf{Z}(\mathbf{k}) = [y_1(\mathbf{k}), y_2(\mathbf{k}), \dots, y_N(\mathbf{k}); U_1(\mathbf{k}) - q_1(\mathbf{k}), \dots, U_{N-1}(\mathbf{k}) - q_{N-1}(\mathbf{k})]^T$ ,  $\mathbf{Z}(\mathbf{k})$  is a column vector of dimension  $2N-1$ ,  $\mathbf{H}[\boldsymbol{\sigma}_{ij}(\mathbf{k})] = [\mathbf{h}_{rs}^k]$  is a matrix of dimension  $(2N-1) \times N(N+1)/2$  with its entries given by the corresponding coefficients in equations (3.15) and (3.16), and  $\mathbf{e}(\mathbf{k})$  is an observation noise vector of dimension  $(2N-1)$ , which can be defined as a Gaussian white noise with zero mean and its covariance matrix, where  $\mathbf{R} = \text{Var}[\mathbf{e}(\mathbf{k})] = \text{diag}[r_1, \dots, r_{2N-1}]$  is a diagonal positive definite matrix of dimension  $(2N-1) \times (2N-1)$ .

Through Equation (3.19)-(3.21), a canonical state-space system model has been established. Due to the nonlinear nature of the formulations and the concern of computing efficiency, this study has employed sequential extended Kalman filtering algorithm (Chui and Chen, 1999) and Gumbel distribution (an approximation of normal distribution) to develop a solution algorithm. A step-by-step description of the algorithm for estimating parameters,  $b_{ij}(\mathbf{k})$  and  $\boldsymbol{\sigma}_{ij}(\mathbf{k})$  is presented below:

Step 0: Initialization

- Link length  $L_i$ ,  $i = 0, 1, \dots, N-1$
- Length of each time interval,  $t_0$ , and the maximum number of intervals required to traverse the entire section  $M$
- Initial input mean speeds,  $V_i(m)$ ,  $m = -M, -M+1, \dots, 0$

- Initial input flows,  $q_i(m)$ ,  $m = -M, -M+1, \dots, 0$
- Initial travel times,  $t_{ij}(m) = L_i / V_i(m) + \dots + L_{j-1} / V_{j-1}(m)$ ,  $m = -M, -M+1, \dots, 0$
- $\text{Var}[\mathbf{e}(\mathbf{k})] = \text{diag} [r_1, r_2, \dots, r_{2N-1}]$

$$\begin{bmatrix} \mathbf{b}(\mathbf{0}) \\ \boldsymbol{\sigma}(\mathbf{0}) \end{bmatrix} = \mathbb{E} \begin{bmatrix} \mathbf{b}(\mathbf{0}) \\ \boldsymbol{\sigma}(\mathbf{0}) \end{bmatrix}, \mathbf{P}_0 = \text{Var} \begin{bmatrix} \mathbf{b}(\mathbf{0}) \\ \boldsymbol{\sigma}(\mathbf{0}) \end{bmatrix}$$

### Step 1: Compute Travel Time (mean value)

$$u_{ij}(\mathbf{k}) = t_{ij}(\mathbf{k})$$

### Step 2: Compute the Linearized Transformation Matrix

- $\mathbf{H}^{k-1} = [\mathbf{H}_{rs}^{k-1}]_{(2N-1) \times N(N+1)/2}$

$$H_{j, Ni+j-i(i+1)/2}^k = \sum_{m=0}^M q_i(k-m) \cdot [F_{m+1}(\boldsymbol{\sigma}_{ij}(k)) - F_m(\boldsymbol{\sigma}_{ij}(k))], \text{ for } 0 \leq i < j \leq N$$

$$H_{N+\ell, Ni+j-i(i+1)/2}^k = \sum_{m=0}^M q_i(k-m) \cdot [F_{m+1}(\boldsymbol{\sigma}_{ij}(k)) - F_m(\boldsymbol{\sigma}_{ij}(k))], \text{ for } 0 \leq i < \ell < j \leq N$$

- $\mathbf{J}^{k-1} = [\mathbf{J}_{rs}^{k-1}]_{(2N-1) \times N(N+1)/2}$

$$J_{j, Ni+j-i(i+1)/2}^k = \sum_{m=0}^M q_i(k-m) \cdot b_{ij}(k), \text{ for } 0 \leq i < j \leq N$$

$$J_{N+\ell, Ni+\ell-i(i+1)/2}^k = \sum_{m=0}^M q_i(k-m) \cdot \sum_{j \leq \ell} b_{ij}(k), \text{ for } 0 \leq i < \ell < j \leq N$$

- $\mathbf{J}_{rs}^k = 0$ , for the other entries of matrix  $\mathbf{J}^k$

- $\mathbf{J}_k = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_{2N-1} \end{bmatrix} = [\mathbf{H}^{k-1} \mathbf{J}^{k-1}]_{(2N-1) \times N(N+1)}$

where each  $f_i$  is a row vector of dimension  $N(N+1)$

$$\bullet \quad \mathbf{Z}'(\mathbf{k}) = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{2N-1} \end{bmatrix} = [y_1(\mathbf{k}), \dots, y_N(\mathbf{k}), U_1(\mathbf{k}) - q_1(\mathbf{k}), \dots, U_{N-1}(\mathbf{k}) - q_{N-1}(\mathbf{k})]^T$$

### Step 3: Initialization of sequential Kalman Filtering

- Set  $\mathbf{b}^0 = \mathbf{b}(\mathbf{k}-1)$ ,  $\boldsymbol{\sigma}^0 = \boldsymbol{\sigma}(\mathbf{k}-1)$
- $\mathbf{P}^0 = \mathbf{P}_{k-1} + \mathbf{D}$  where,  $\mathbf{D} = \begin{bmatrix} \mathbf{D}_b & \\ & \mathbf{D}_\sigma \end{bmatrix}$ ,  $\mathbf{D}_b = \text{diag}[d_b, \dots, d_b]$  is a covariance

matrix of  $W(\mathbf{k})$ , and  $\mathbf{D}_\sigma$  is a constant semi-positive matrix.

### Step 4: Sequential Kalman Filtering Iteration

For  $i = 1, 2, \dots, 2N-1$

- $\mathbf{g}^i = \mathbf{P}^{i-1} \mathbf{f}_i^T [\mathbf{f}_i \mathbf{P}^{i-1} \mathbf{f}_i^T + \mathbf{r}_i]^{-1}$
- $\mathbf{P}^i = \mathbf{P}^{i-1} - \mathbf{g}^i \mathbf{f}_i \mathbf{P}^{i-1}$
- $\delta^i = y_i(\mathbf{k}) - \mathbf{f}_i \mathbf{b}(\mathbf{k}-1)$
- Truncation:

Set

$$\begin{bmatrix} \mathbf{b}^i \\ \boldsymbol{\sigma}^i \end{bmatrix} = \begin{bmatrix} \mathbf{b}^{i-1} \\ \boldsymbol{\sigma}^{i-1} \end{bmatrix} + \alpha' \delta^i \mathbf{g}^i, \text{ where } \alpha' = \text{MAX}_{0 \leq \alpha \leq 1} [\alpha \mid 0 \leq [\mathbf{b}^{i-1}] + \alpha \delta^i \mathbf{g}^i \leq 1] \quad (3.22)$$

- Normalization:

For  $m = 1, 2, \dots, N-2$

$$\beta_m = \sum_{j=m+1}^N b_{mj}^i, b_{mj}^i = b_{mj}^i / \beta_m, j = m+1, \dots, N \quad (3.23)$$

#### Step 5: Prediction of the States

- Set  $\mathbf{P}_k = \mathbf{P}^{2N-1}$

$$\begin{bmatrix} \mathbf{b}(\mathbf{k}) \\ \boldsymbol{\sigma}(\mathbf{k}) \end{bmatrix} = \begin{bmatrix} \mathbf{b}^{2N-1} \\ \boldsymbol{\sigma}^{2N-1} \end{bmatrix}$$

$k=k+1$ , go to Step 1 for the next interval.

### **3.5 NUMERICAL EXAMPLES AND SENSITIVITY ANALYSES**

In order to evaluate the effectiveness of the proposed model with an embedded travel time distribution function, a small freeway network is designed for sensitivity analyses with respect to potential measurement errors in travel time and initial values. In addition, an example with a large freeway network is presented to demonstrate the advantage of incorporating mainline traffic volumes in the model formulation.

#### **3.5.1 Sensitivity Analysis for a Small Network**

To generate a meaningful data set for numerical analysis, an example freeway system using the presumed time series of O-D percentages was simulated with AIMSUN 4.0 (TSS, 2001), to produce time-dependent link traffic volumes. The traffic flow data was collected at an interval of two minutes over the entire simulation duration of one hour. Figure 3.5 illustrates the example freeway corridor, and Table 3.1 presents the presumed time series O-D proportions.

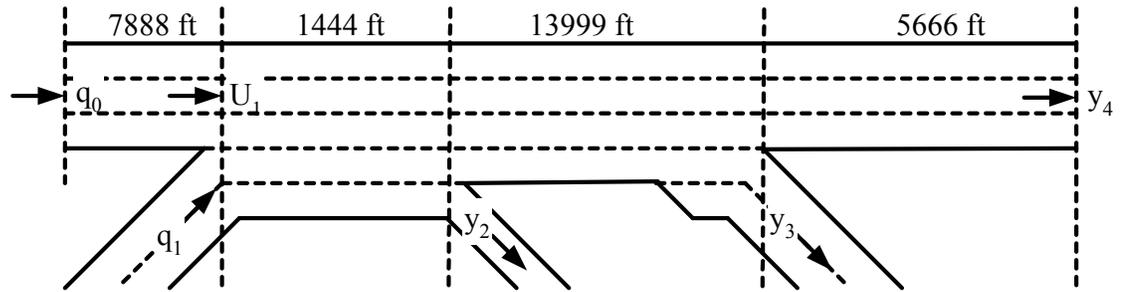


Figure 3.5. A Small Example Freeway Section for Model Sensitivity Test

Table 3.1. The Input Time Series of O-D Proportions

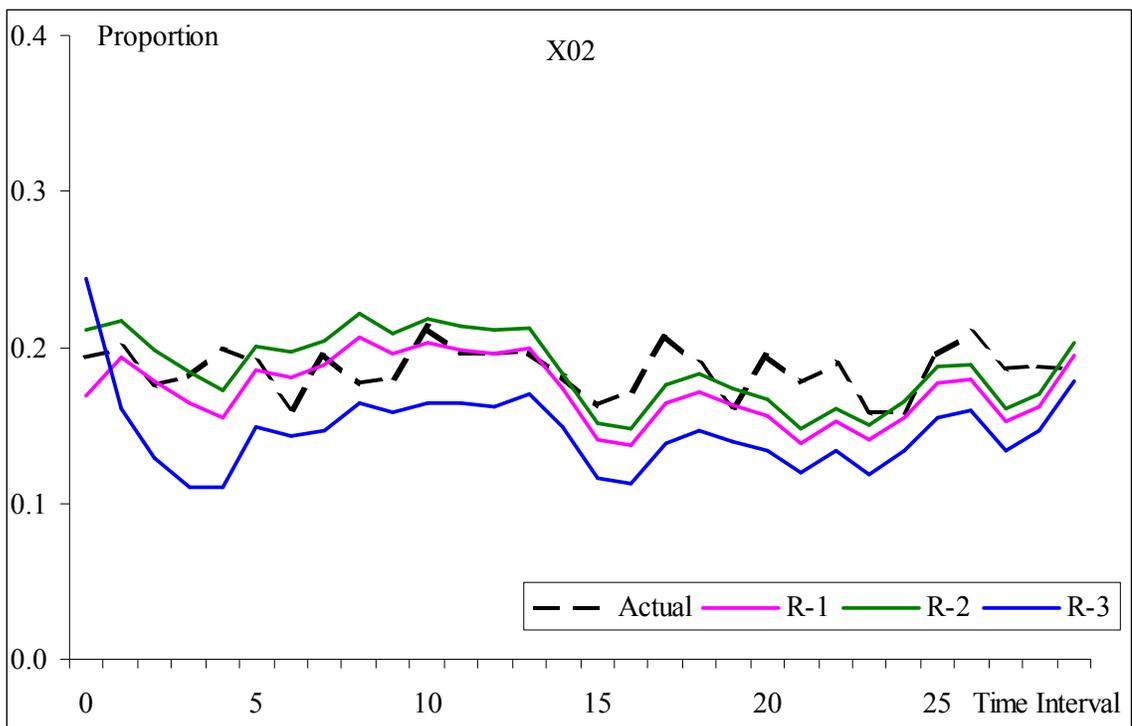
Interval \ O-D	b <sub>02</sub>	b <sub>03</sub>	b <sub>04</sub>	b <sub>12</sub>	b <sub>13</sub>	b <sub>14</sub>	Interval \ O-D	b <sub>02</sub>	b <sub>03</sub>	b <sub>04</sub>	b <sub>12</sub>	b <sub>13</sub>	b <sub>14</sub>
0	0.19	0.34	0.46	0.27	0.40	0.33	15	0.16	0.29	0.55	0.26	0.36	0.38
1	0.20	0.29	0.51	0.26	0.39	0.35	16	0.17	0.26	0.56	0.26	0.38	0.37
2	0.18	0.29	0.53	0.22	0.36	0.41	17	0.21	0.29	0.50	0.23	0.40	0.37
3	0.18	0.28	0.54	0.25	0.37	0.38	18	0.19	0.32	0.49	0.25	0.38	0.37
4	0.20	0.28	0.52	0.23	0.38	0.39	19	0.16	0.30	0.54	0.26	0.37	0.37
5	0.19	0.28	0.53	0.24	0.41	0.35	20	0.19	0.30	0.51	0.26	0.35	0.39
6	0.16	0.28	0.56	0.25	0.40	0.36	21	0.18	0.34	0.49	0.24	0.38	0.39
7	0.20	0.30	0.50	0.24	0.38	0.38	22	0.19	0.27	0.54	0.24	0.36	0.40
8	0.18	0.33	0.50	0.27	0.34	0.39	23	0.16	0.29	0.55	0.28	0.32	0.41
9	0.18	0.31	0.50	0.24	0.41	0.35	24	0.16	0.31	0.53	0.26	0.38	0.36
10	0.21	0.25	0.54	0.24	0.41	0.36	25	0.20	0.31	0.49	0.22	0.41	0.37
11	0.20	0.31	0.49	0.27	0.36	0.37	26	0.21	0.30	0.49	0.24	0.40	0.35
12	0.20	0.32	0.48	0.23	0.35	0.42	27	0.19	0.25	0.56	0.27	0.33	0.40
13	0.20	0.31	0.50	0.25	0.38	0.37	28	0.19	0.28	0.54	0.26	0.35	0.39
14	0.18	0.28	0.54	0.28	0.37	0.35	29	0.19	0.33	0.48	0.29	0.36	0.35

Note: b<sub>ij</sub> denotes the O-D fraction from ramp i to j.

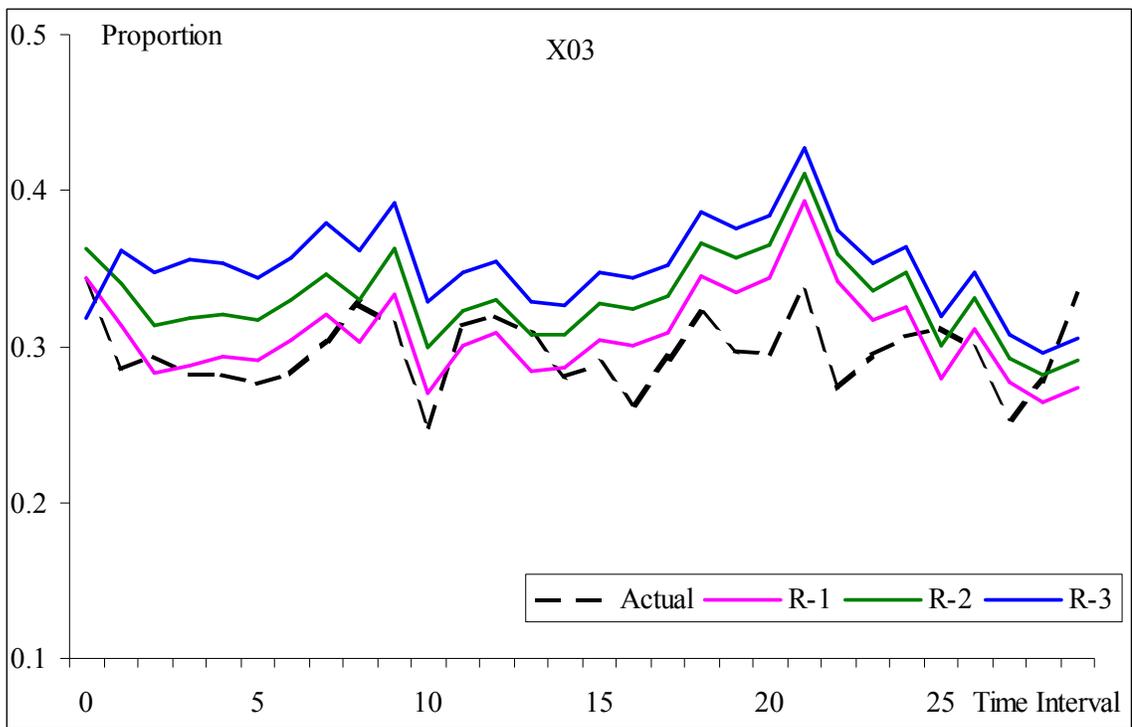
To test model performance under a different set of initial values, this study has generated the following three experimental sets:

- R-1 = [0.19 0.34 0.46 0.27 0.40 0.33] – the exact initial value set
- R-2 = [0.33 0.33 0.33 0.33 0.33 0.33] – the uniformed initial value set; and
- R-3 = [0.70 0.10 0.20 0.70 0.10 0.20] – the initial value set with a certain random variation

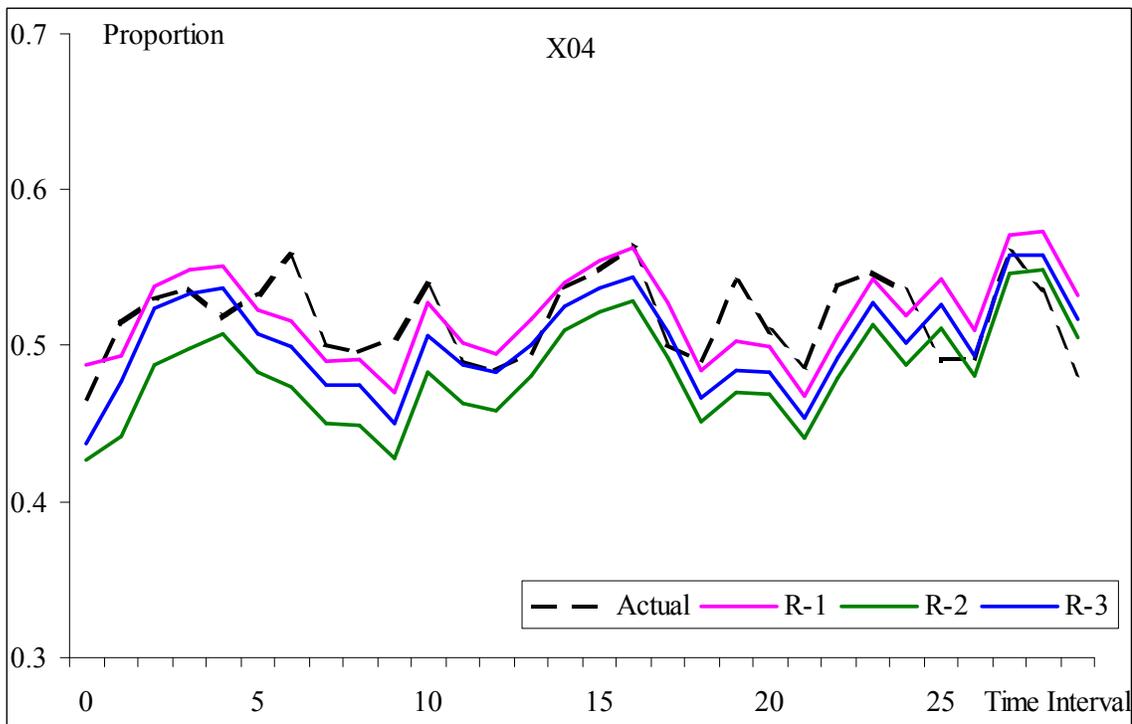
The graphical estimation results for these three sets of O-D proportions with comparison of the actual O-D proportions are reported in Figure 3.6 and the absolute error (AE) statistics are illustrated in Figure 3.7.



**Figure 3.6a. Graphical Estimation Results with Different Sets of Initial Values –X02**



**Figure 3.6b. Graphical Estimation Results with Different Sets of Initial Values –X03**



**Figure 3.6c. Graphical Estimation Results with Different Sets of Initial Values –X02**

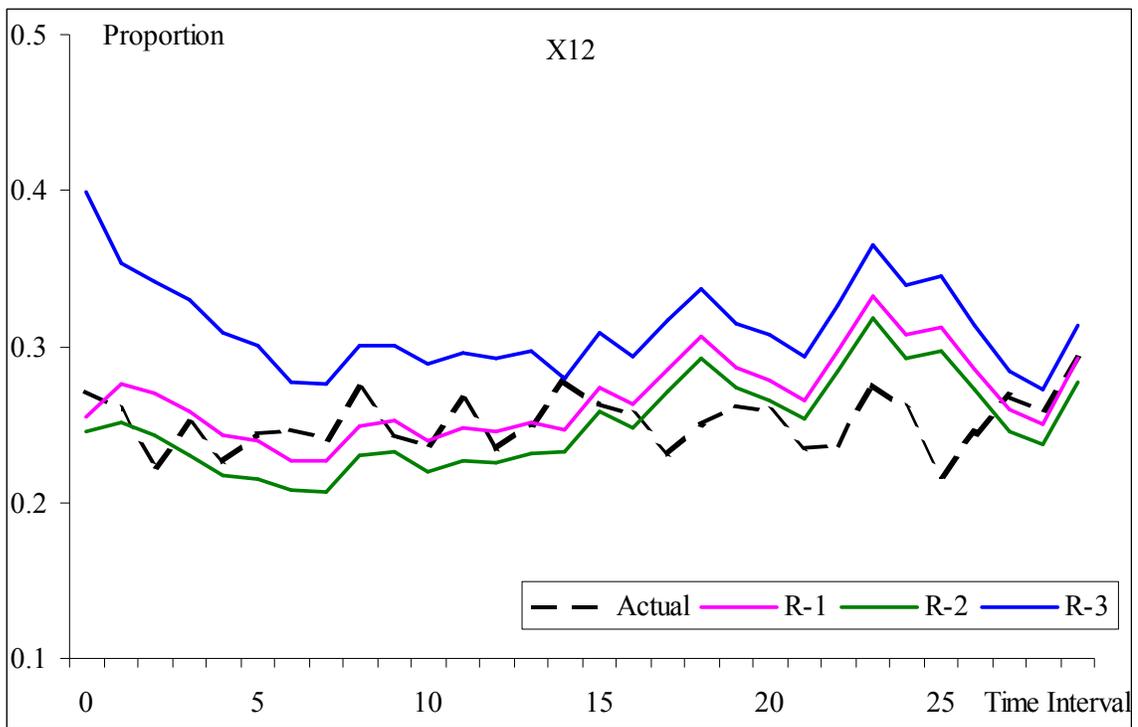


Figure 3.6d. Graphical Estimation Results with Different Sets of Initial Values –X12

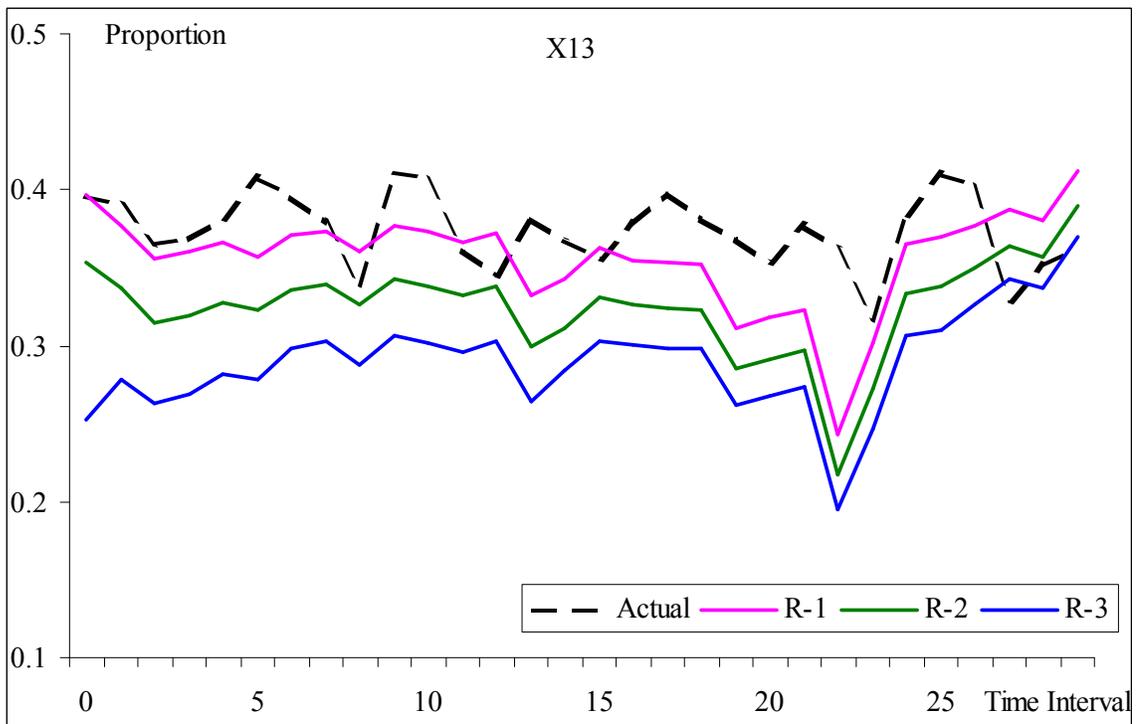
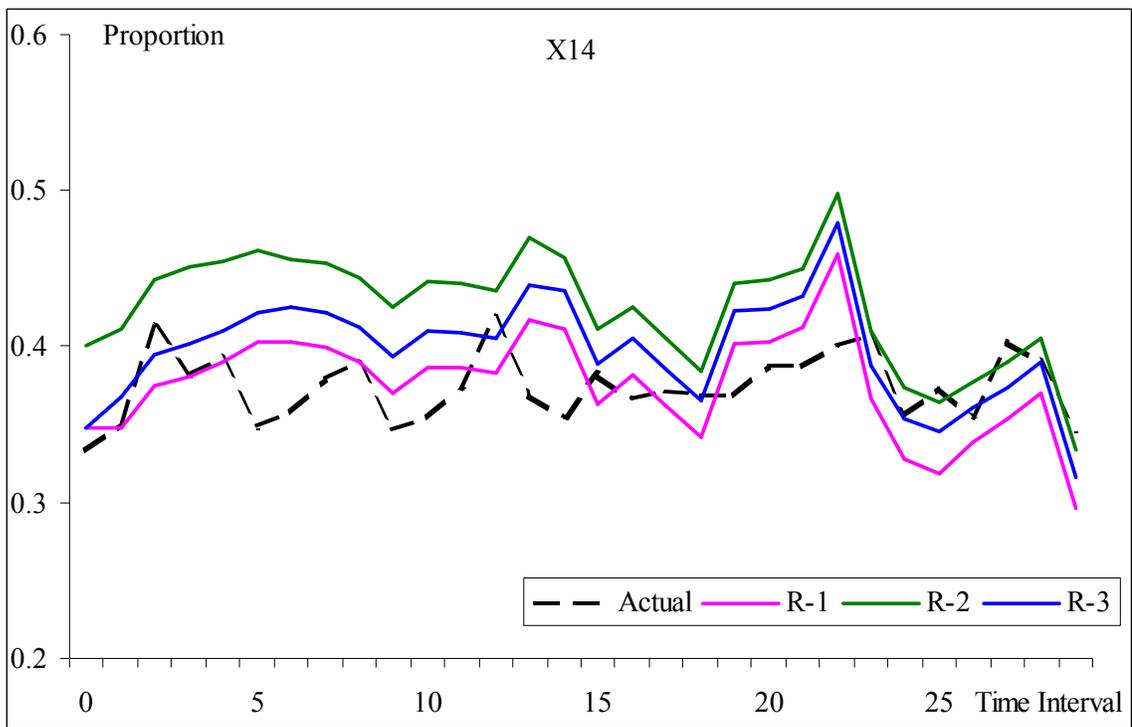
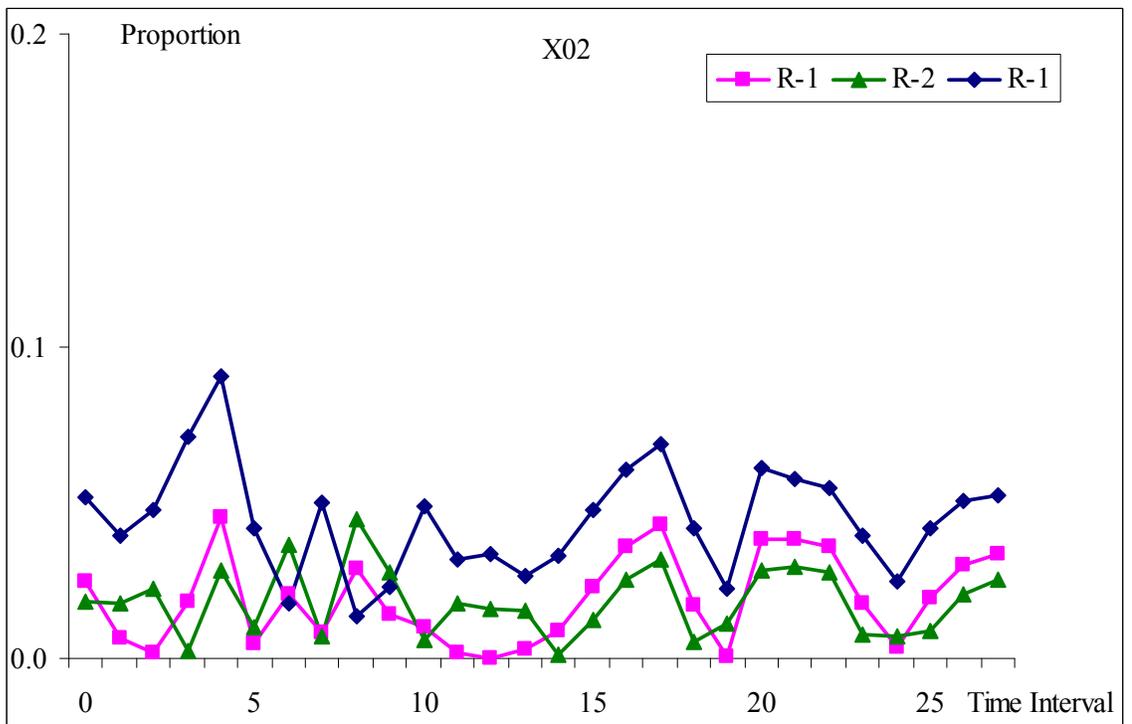


Figure 3.6e. Graphical Estimation Results with Different Sets of Initial Values –X13



**Figure 3.6f. Graphical Estimation Results with Different Sets of Initial Values –X14**



**Figure 3.7a. Graphical Absolute Errors with Different Sets of Initial Values – X02**

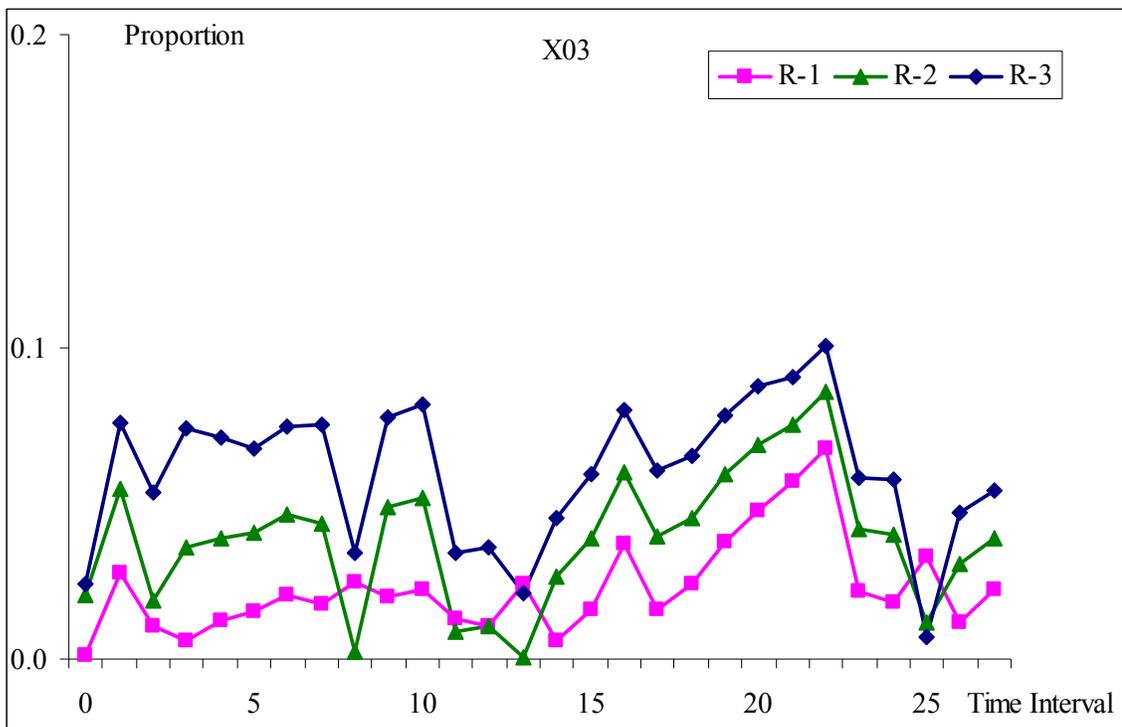


Figure 3.7b. Graphical Absolute Errors with Different Sets of Initial Values – X03

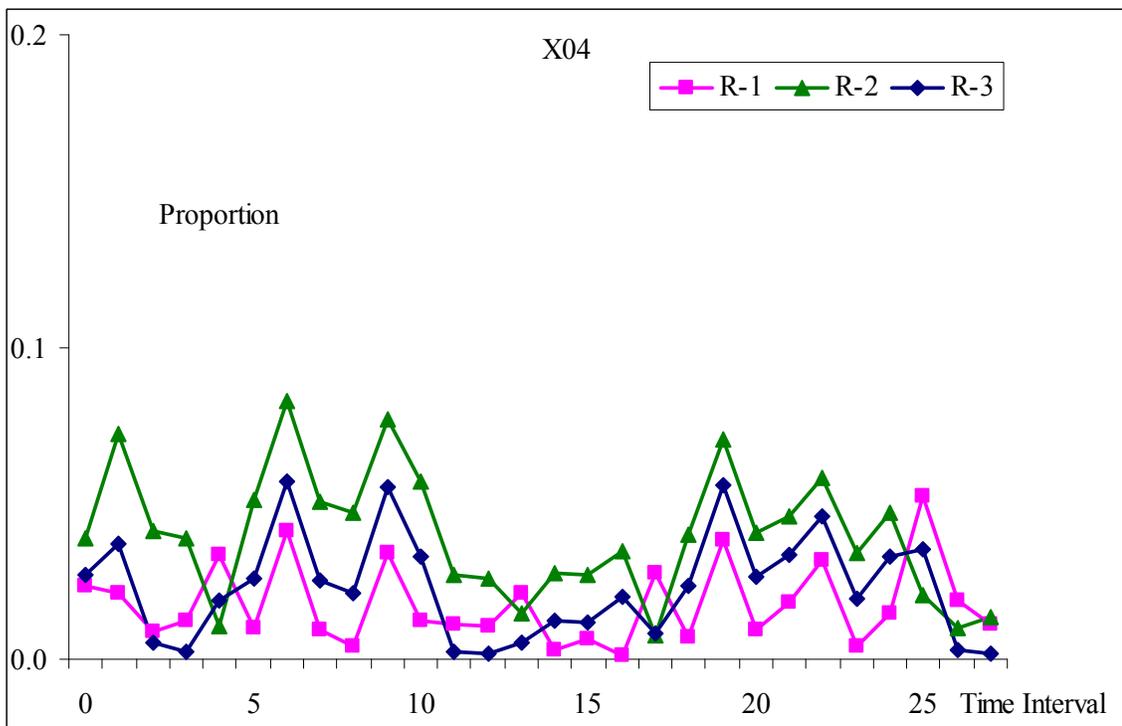


Figure 3.7c. Graphical Absolute Errors with Different Sets of Initial Values – X04

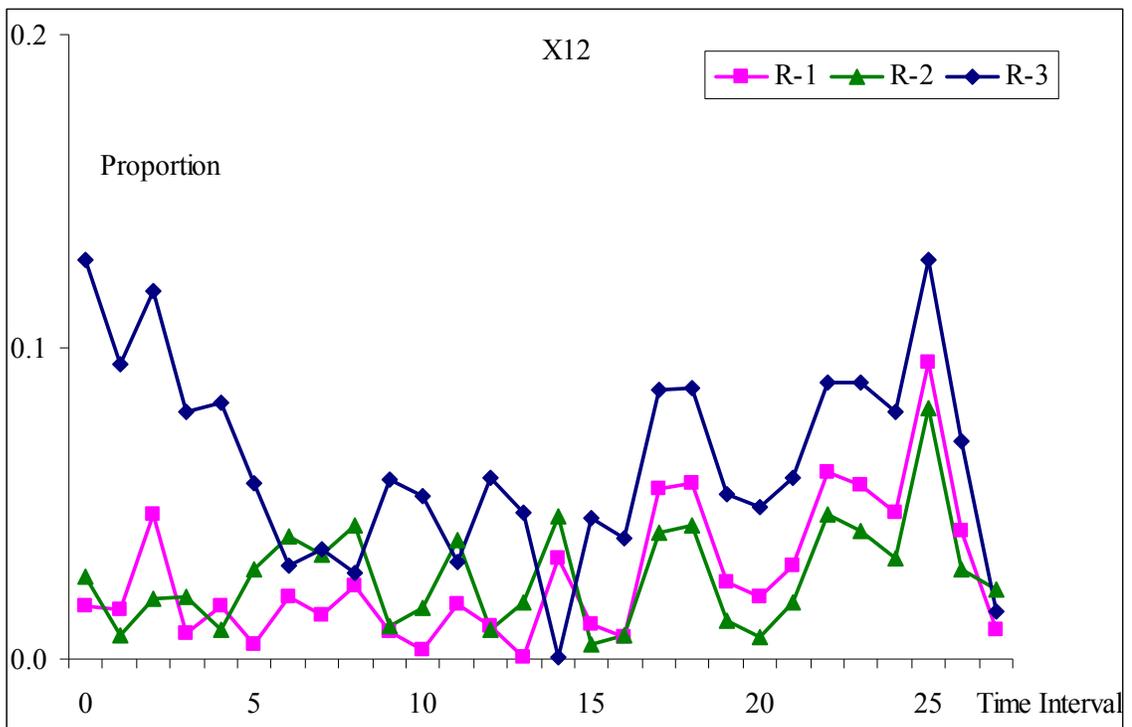


Figure 3.7d. Graphical Absolute Errors with Different Sets of Initial Values – X12

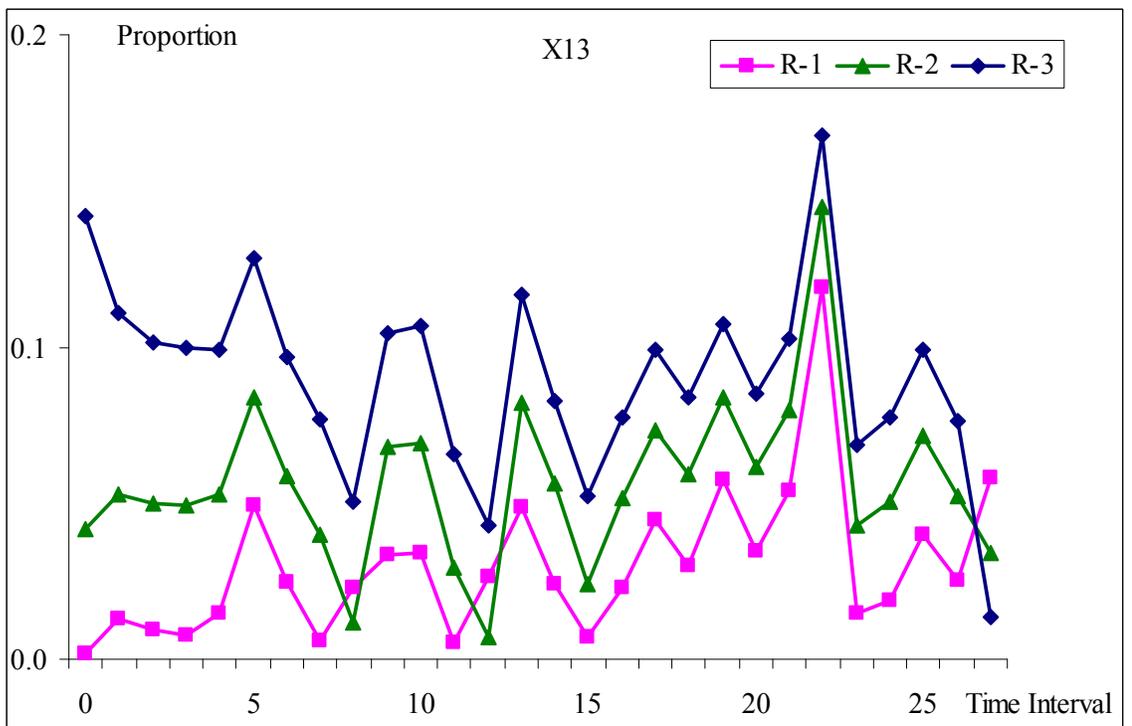
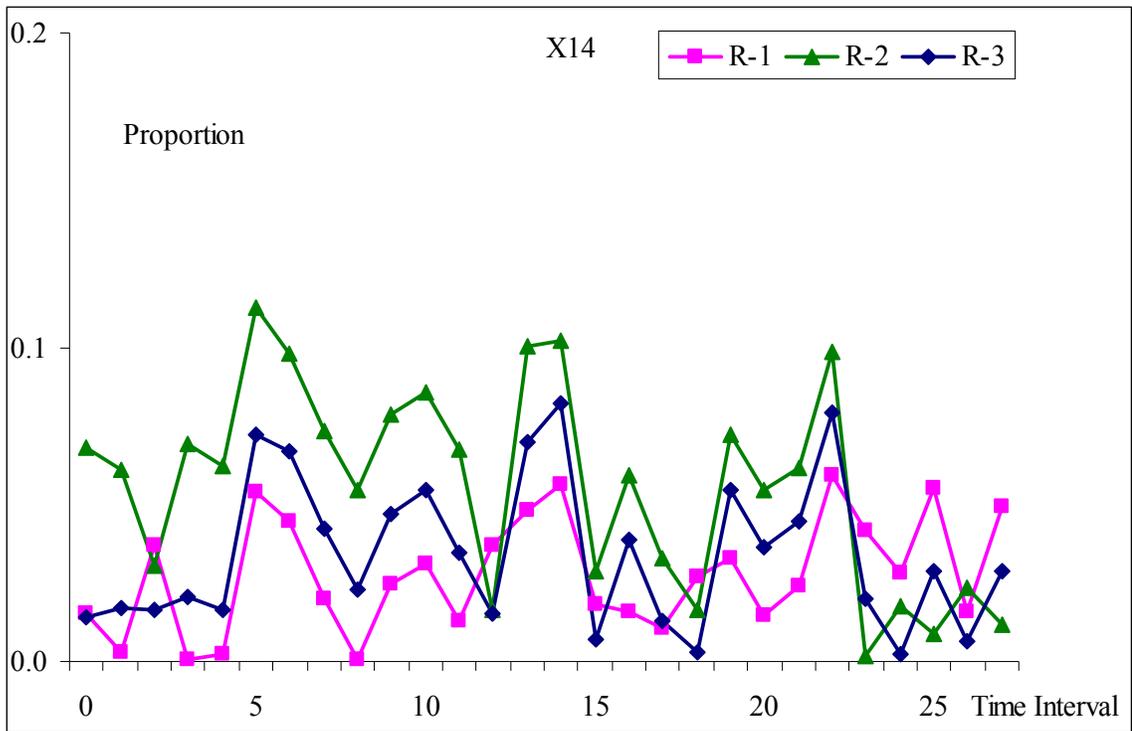


Figure 3.7e. Graphical Absolute Errors with Different Sets of Initial Values – X13



**Figure 3.7f. Graphical Absolute Errors with Different Sets of Initial Values – X14**

In Figures 3.6a – 3.6f, the patterns of the estimation results are the same as the actual O-D proportions when the initial value set is created from the correct initial O-D set (R-1). It takes longer for the other two sets of initial O-D values to evolve to the same pattern, especially for the initial O-D set with a larger random variation (R-3). In Figures 3.7a – 3.7f, the estimation errors for most estimated O-D sets are within 0.1.

Table 3.2 presents statistical results for the average absolute estimation error with different sets of initial values. With the reasonable range of initial value set (e.g., R-2), the estimation results with the proposed model are quite stable and vary only slightly.

**Table 3.2. The Absolute Error Statistics with Different Sets of Initial Values**

Abs. Error Initial Set	X02	X03	X04	X12	X13	X14	Avg.	Max.	Min.
R-1	0.0188	0.0237	0.0195	0.0254	0.0310	0.0283	0.0245	0.1190	0.0000
R-2	0.0181	0.0376	0.0382	0.0263	0.0539	0.0532	0.0379	0.1449	0.0005
R-3	0.0429	0.0579	0.0235	0.0606	0.0853	0.0328	0.0505	0.1676	0.0007

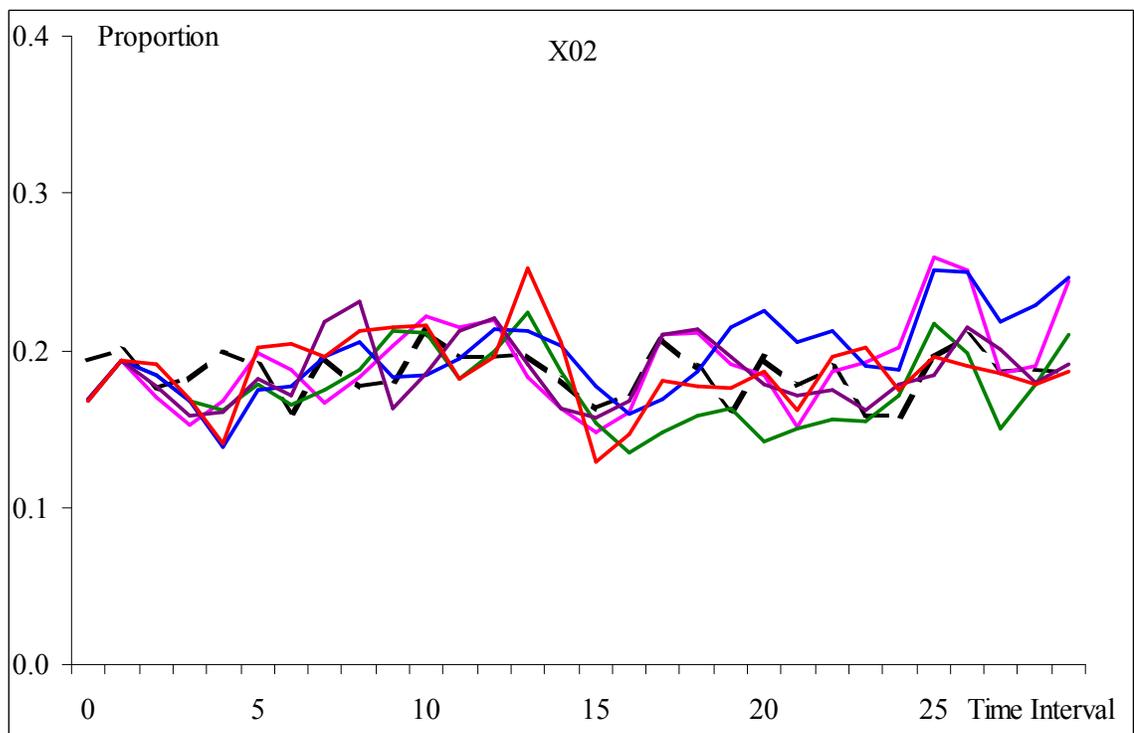
Aside from the set of initial values, the actual distribution of travel times is one of the most important factors that could influence estimated O-D proportions. The major advantage of the proposed model is that it allows travel times to vary within a certain range. To test the effectiveness of the proposed travel time formulation, the exact initial value set, R-1, is selected for executing the estimation, and 5 cases are generated with the average travel time randomly increased or decreased within the range of 10% from the average travel time.

Table 3.3 presents statistical results of the average absolute error associated with the 5 cases. It can be noted that the proposed model yields quite stable results, where the errors remain nearly constant even when the average travel times are subject to approximately 10% of variation.

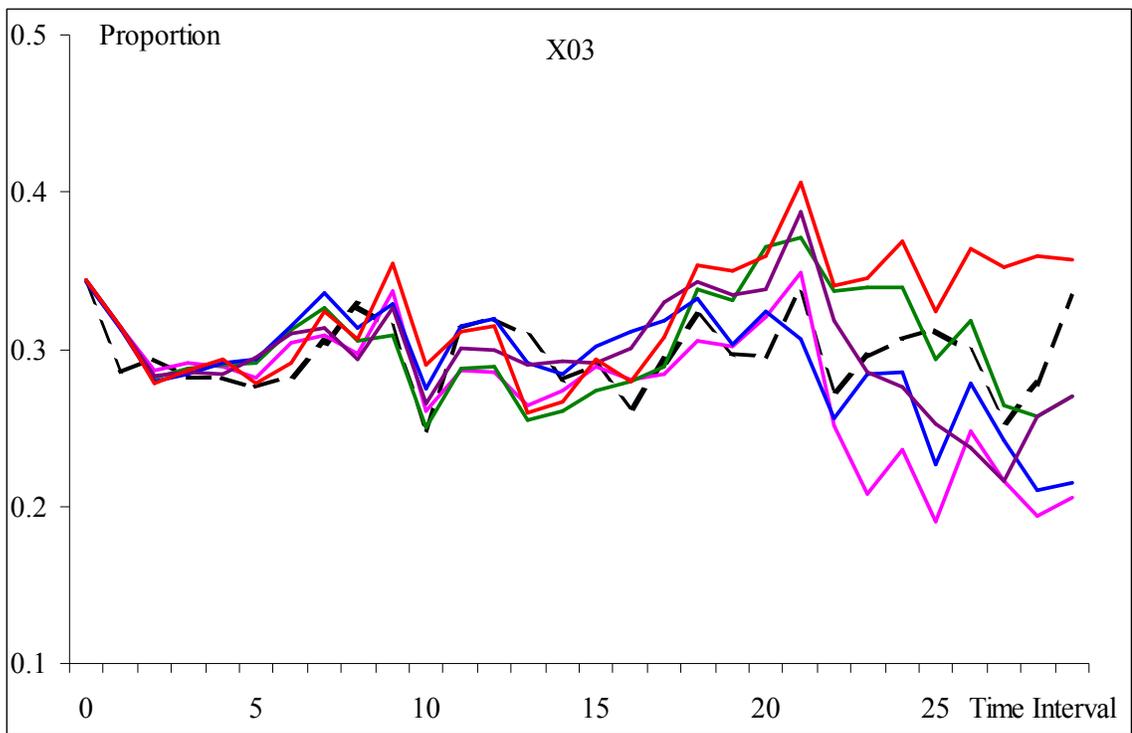
Figure 3.8 illustrates the graphical results for the 5 cases under different travel time scenarios, which generally follow the same pattern as the actual O-D distribution. The estimation errors, as expected, increase with the distance (or travel time) between each O-D pair. For example,  $b_{13}$  and  $b_{14}$  have relatively large estimation errors.

**Table 3.3. The Absolute Error Statistics with Travel Time Variation**

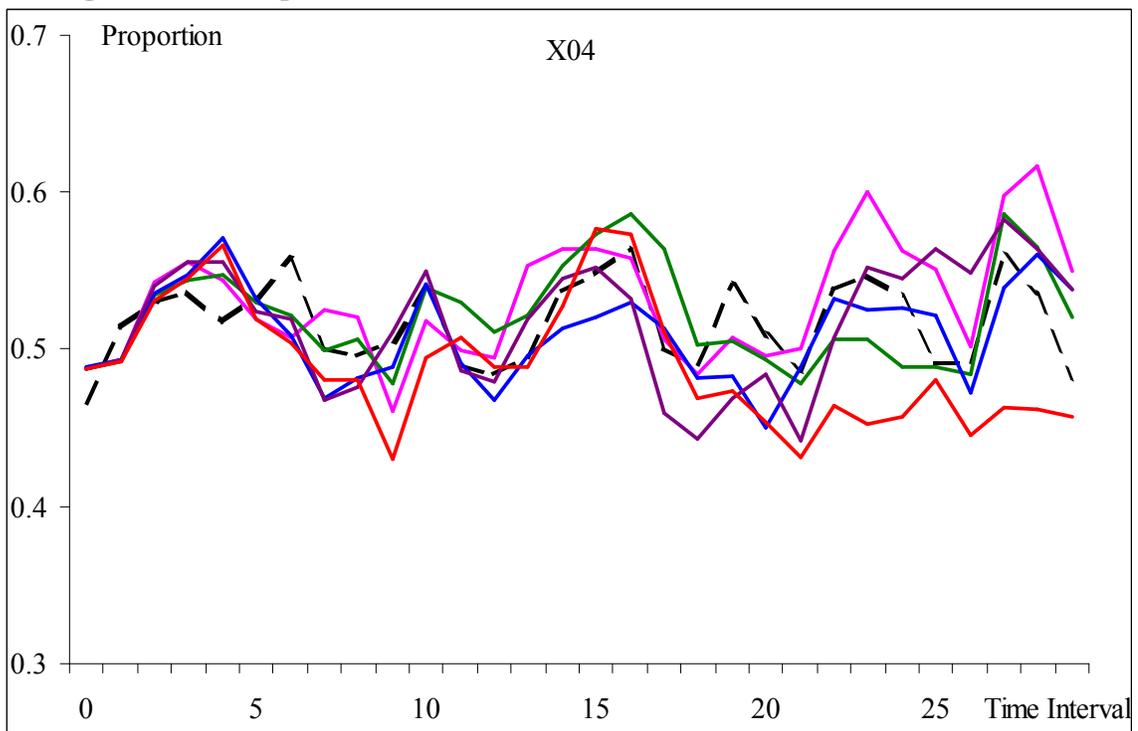
Abs. Error Case	X02	X03	X04	X12	X13	X14	Avg.	Max.	Min.
Case 1	0.0213	0.0316	0.0281	0.0287	0.0421	0.0301	0.0303	0.1885	0.0001
Case 2	0.0197	0.0247	0.0226	0.0283	0.0344	0.0271	0.0261	0.1080	0.0001
Case 3	0.0252	0.0242	0.0219	0.0291	0.0369	0.0333	0.0284	0.1511	0.0002
Case 4	0.0160	0.0262	0.0273	0.0265	0.0348	0.0420	0.0288	0.1086	0.0004
Case 5	0.0191	0.0326	0.0368	0.0251	0.0420	0.0479	0.0339	0.1457	0.0001
Average	0.0202	0.0279	0.0273	0.0275	0.0380	0.0361	0.0295	–	–



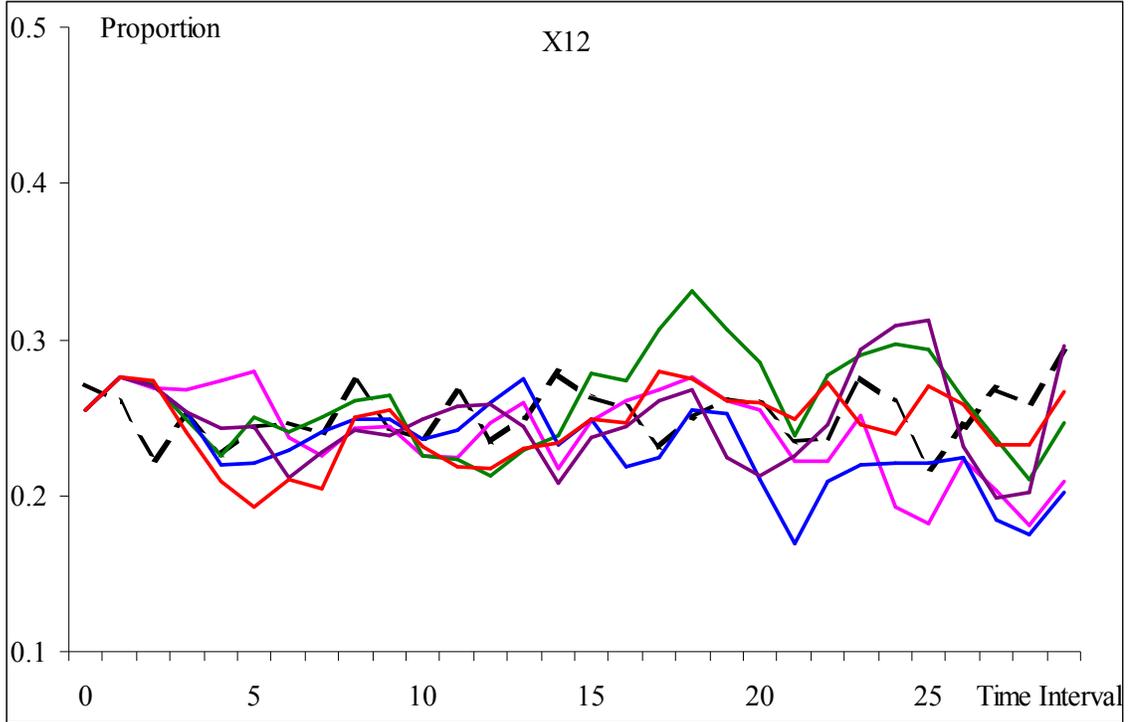
**Figure 3.8a. Graphical Estimation Results with Travel Time Variations-X02**



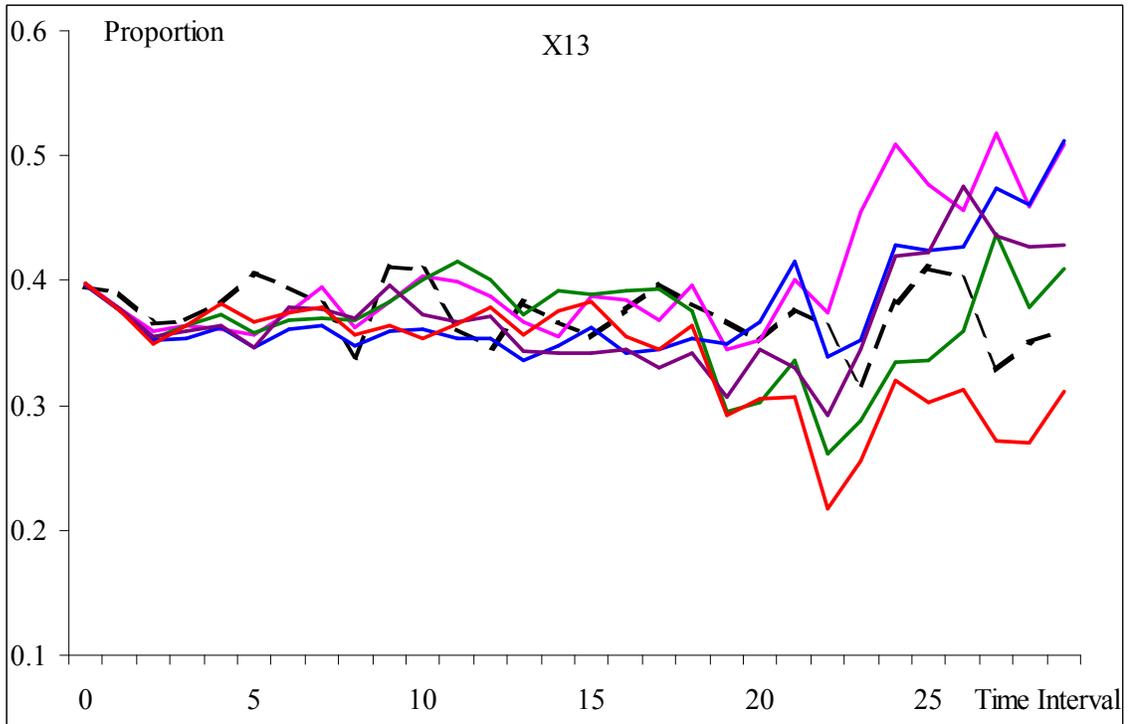
**Figure 3.8b. Graphical Estimation Results with Travel Time Variations-X03**



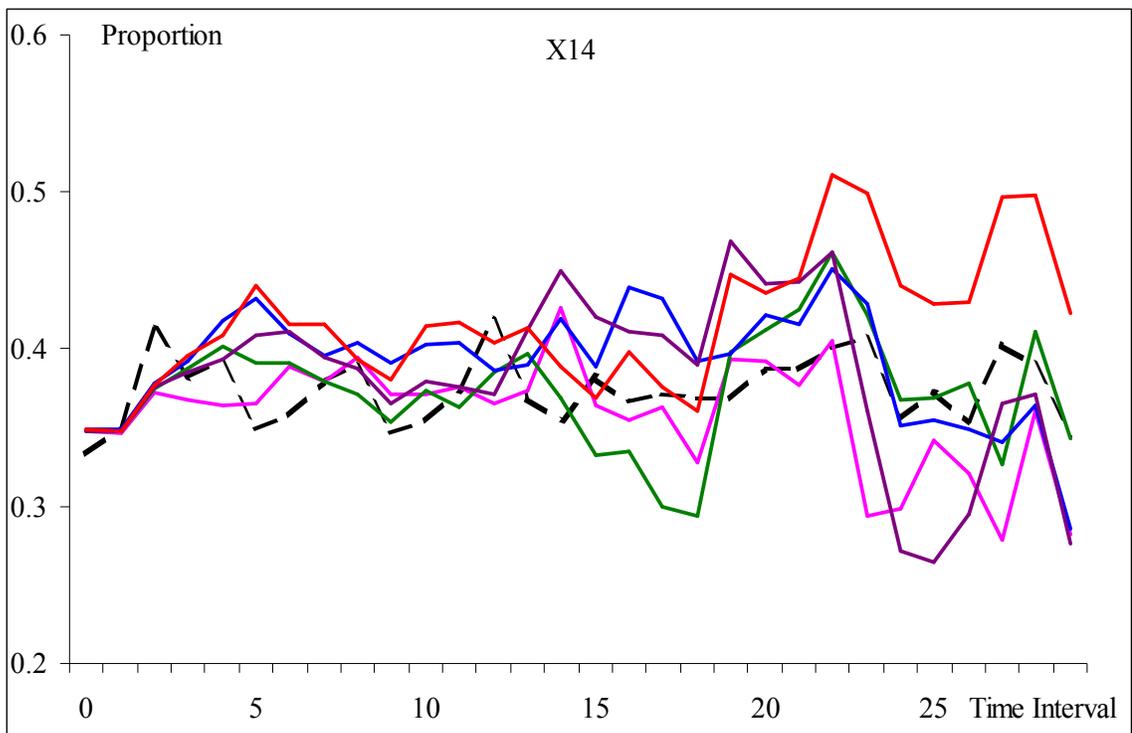
**Figure 3.8c. Graphical Estimation Results with Travel Time Variations-X04**



**Figure 3.8d. Graphical Estimation Results with Travel Time Variations-X12**



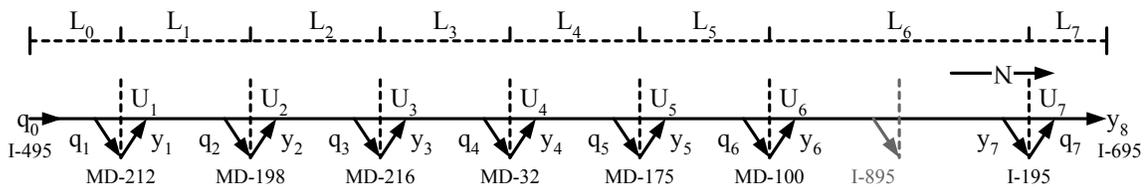
**Figure 3.8e. Graphical Estimation Results with Travel Time Variations-X13**



**Figure 3.8f. Graphical Estimation Results with Travel Time Variations-X14**



This freeway corridor consists of 7 main interchanges, 12 on-ramps, 14 off-ramps, and 120 O-D pairs. For purposes of practicability, each interchange is represented with only a pair of one on-ramp and one off-ramp, and the network is thus reduced to 7 pairs of on-ramps and off-ramps, and 36 O-D sets as shown in Figure 3.10, and Table 3.4 presents the geometry information for each link.



**Figure 3.10. A Graphical Illustration of the I-95 Freeway Corridor**

**Table 3.4 The Geometry Information for each Link**

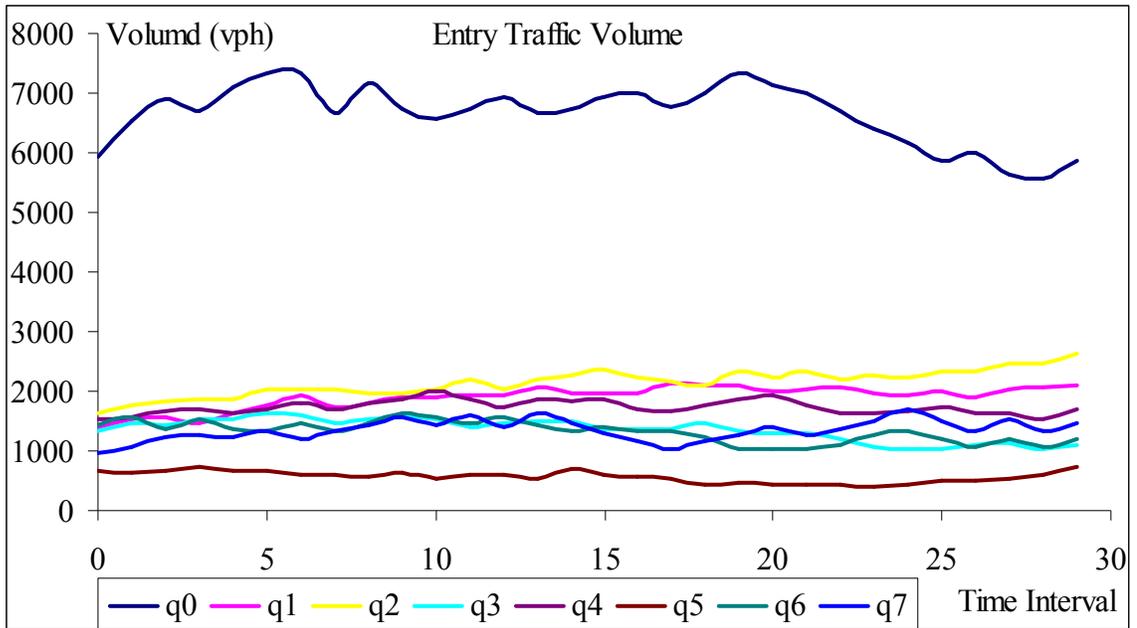
Link Information	L <sub>0</sub>	L <sub>1</sub>	L <sub>2</sub>	L <sub>3</sub>	L <sub>4</sub>	L <sub>5</sub>	L <sub>6</sub>	L <sub>7</sub>
Length (ft)	8,500	19,900	11,965	14,744	10,716	9,002	20,218	8,425
No. Lanes	4	4	4	4	4	4	4	4
Speed Limit (mph)	65	65	65	65	65	65	65	65

To generate a meaningful data set for numerical analysis, the example freeway system under the assigned time series O-D percentages was simulated with AIMSUN 4.0, to produce time-dependent link traffic volumes. For each scenario, the simulation was executed for one hour using the dynamic O-Ds at an interval of 2 minutes. Table 3.5 shows the aggregate input of O-D demands over the one-hour simulation with a unit interval of 2 minutes.

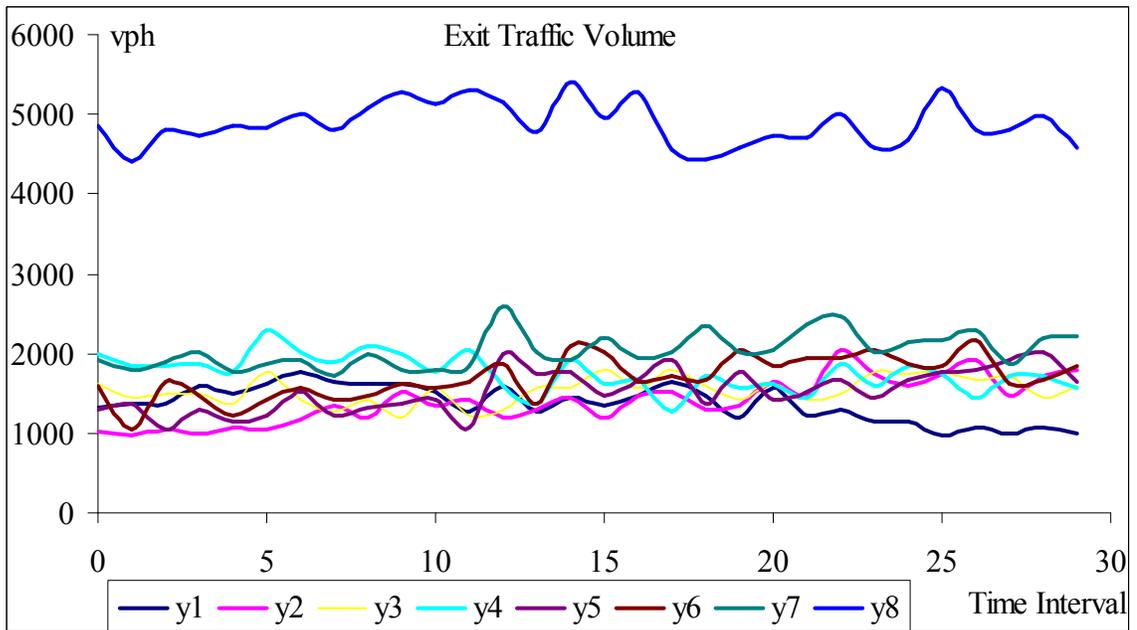
**Table 3.5. Input O-D Demand for Simulation Experiments (unit: vph)**

Origin \ Destination	1	2	3	4	5	6	7	8	Oi
0	1,371	1,190	749	827	692	560	570	375	6,334
1	–	212	303	186	214	226	264	210	1,615
2	–	–	484	395	196	348	224	311	1,958
3	–	–	–	344	263	170	218	390	1,385
4	–	–	–	–	166	206	306	980	1,658
5	–	–	–	–	–	185	134	248	567
6	–	–	–	–	–	–	318	1,031	1,349
7	–	–	–	–	–	–	–	1,338	1,338
Dj	1,453	1,022	1,359	1,517	1,138	1,244	1,916	4,144	13,793

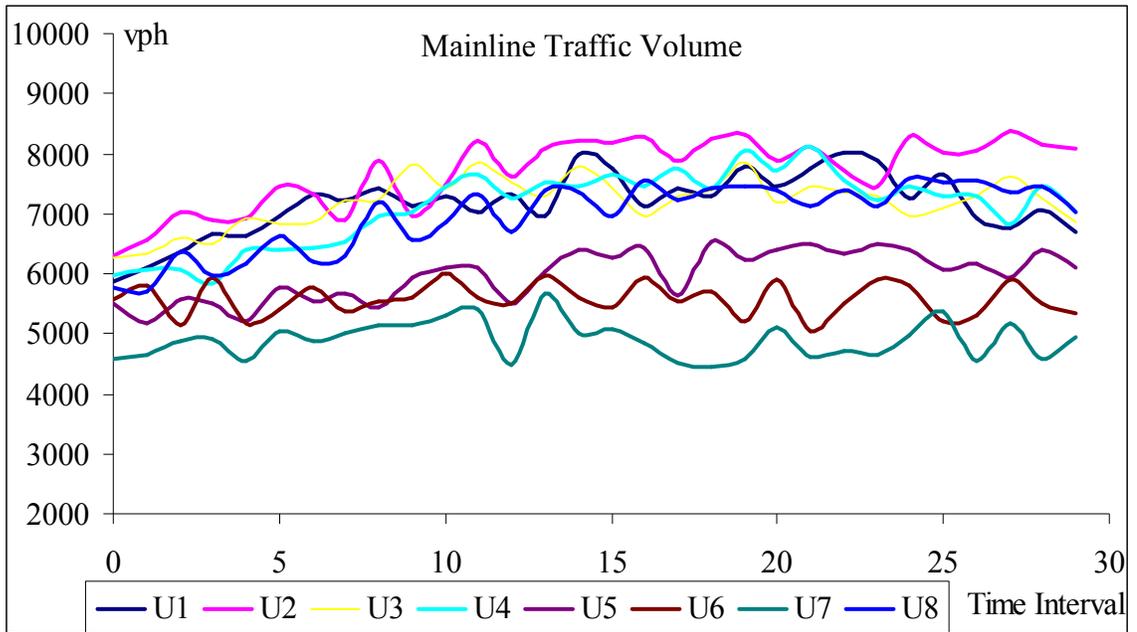
Following the simulation, one can obtain volume information for each link and the average value and variation of speed information for each O-D pair. Figures 3.11, 3.12 and 3.13 show the progression of volume for each entry, exit and mainline link, respectively. Figure 3.14 shows the travel time distribution for each O-D pair.



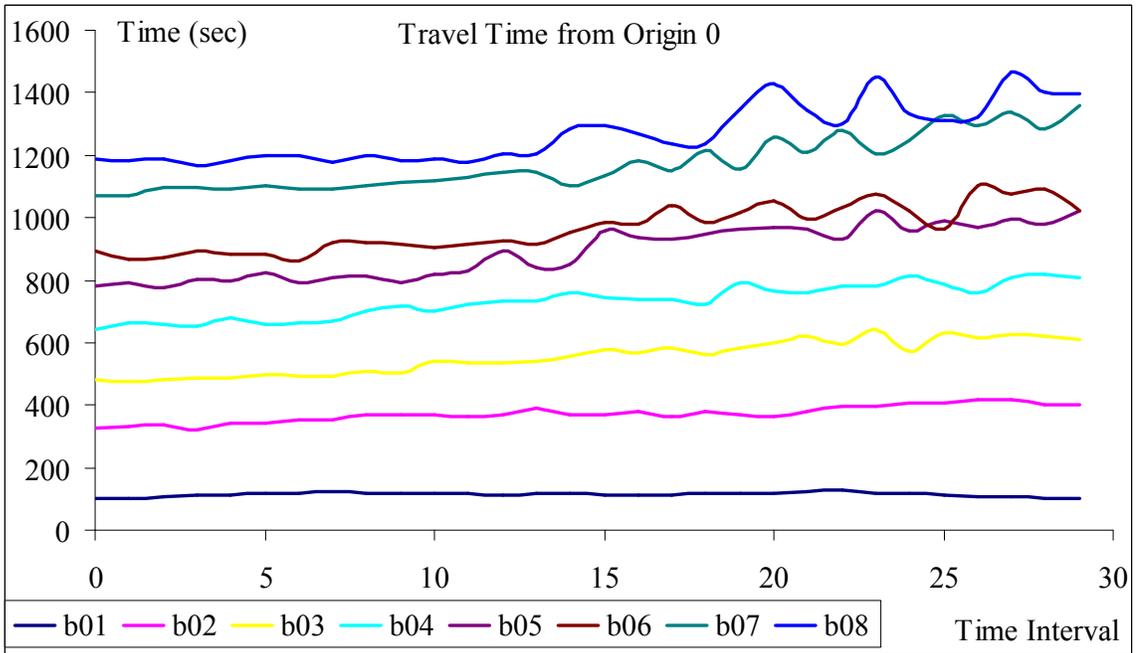
**Figure 3.11. Entry Traffic Volume Distribution**



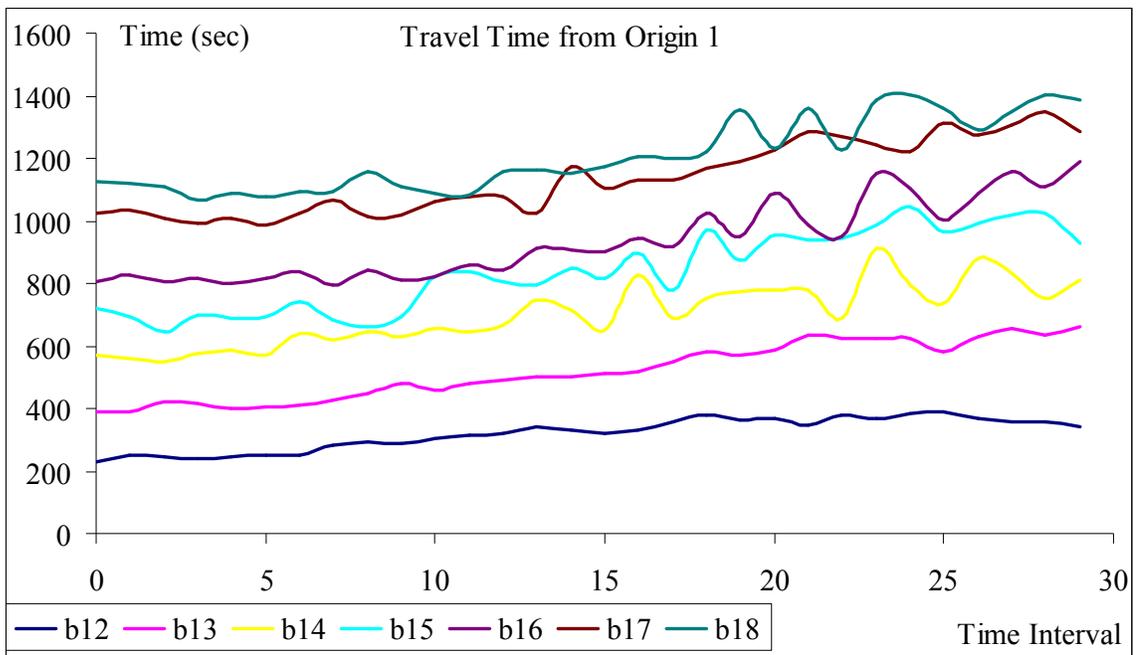
**Figure 3.12. Exit Traffic Volume Distribution**



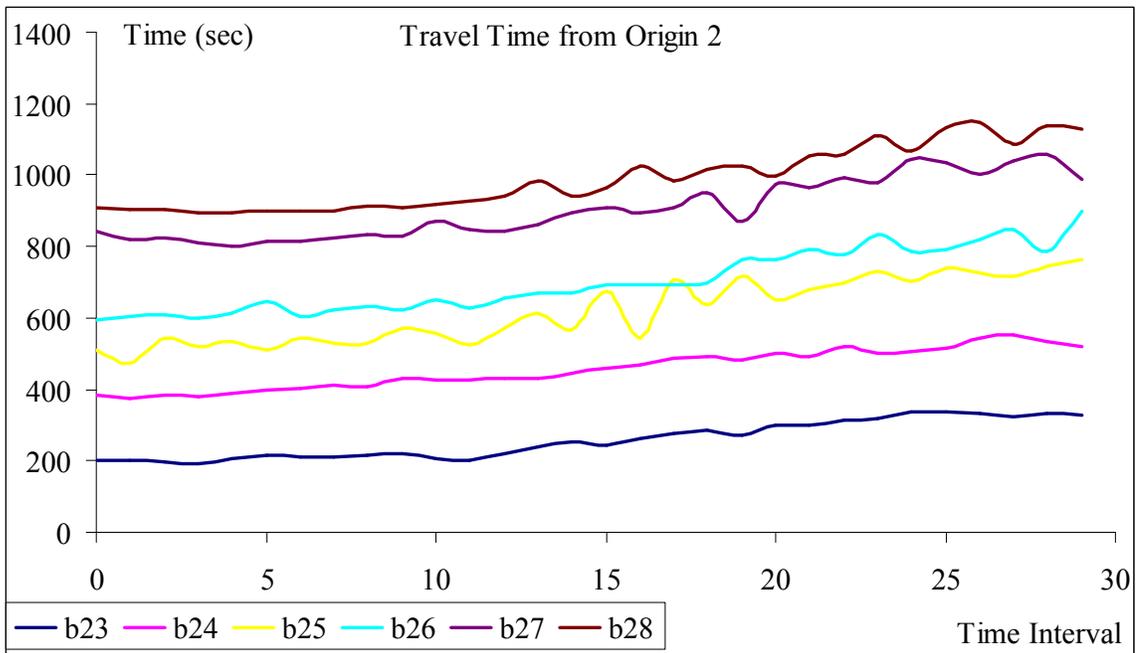
**Figure 3.13. Mainline Traffic Volume Distribution**



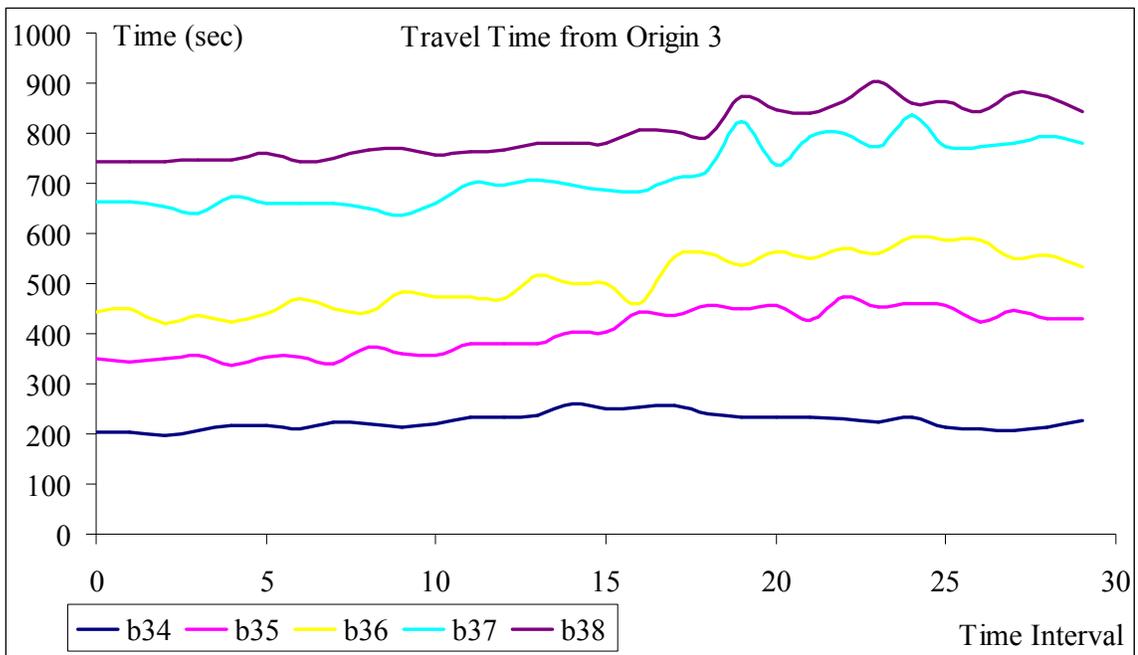
**Figure 3.14a. Travel Time Distribution for each O-D Pair – Origin 0**



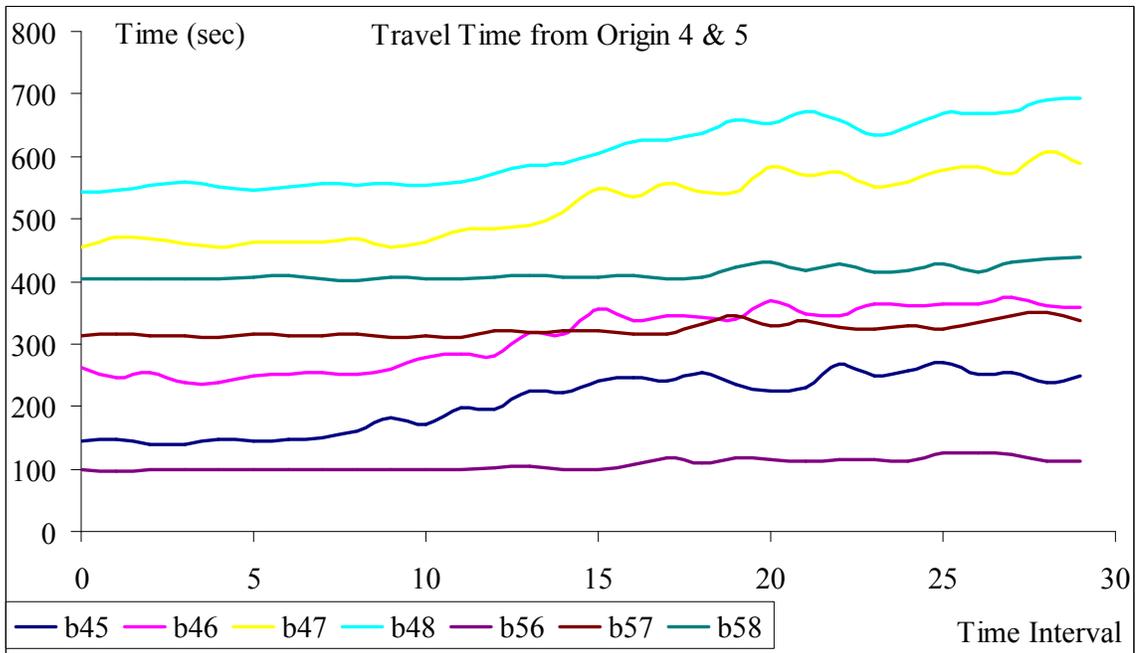
**Figure 3.14b. Travel Time Distribution for each O-D Pair – Origin 1**



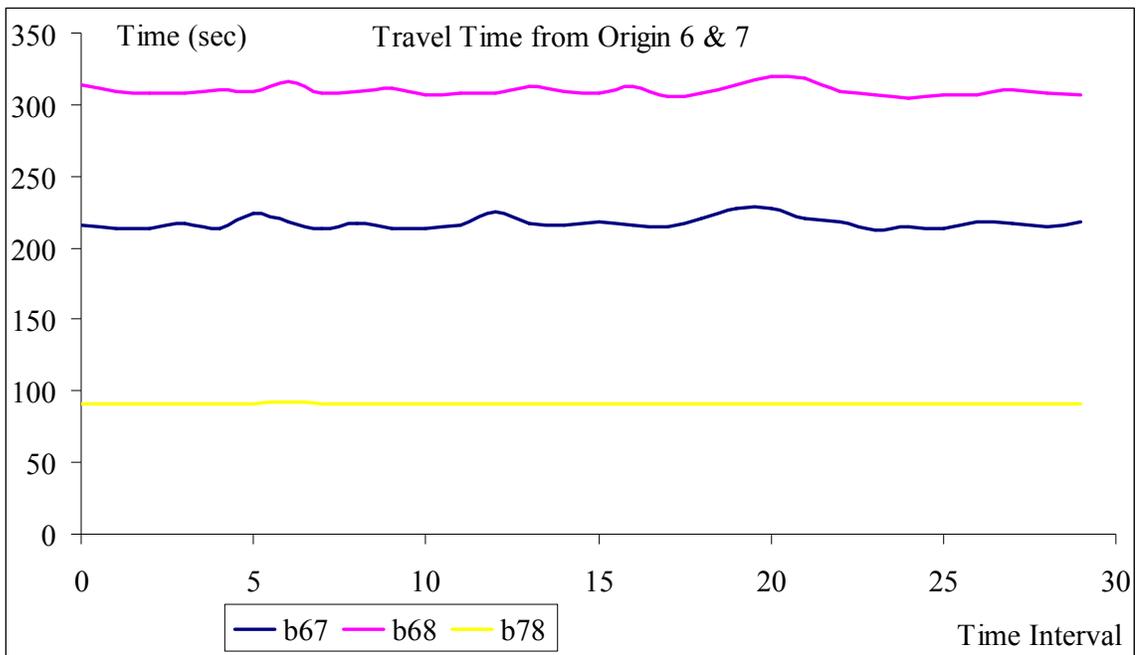
**Figure 3.14c. Travel Time Distribution for each O-D Pair – Origin 2**



**Figure 3.14d. Travel Time Distribution for each O-D Pair – Origin 3**



**Figure 3.14e. Travel Time Distribution for each O-D Pair – Origins 4 and 5**



**Figure 3.14f. Travel Time Distribution for each O-D Pair – Origins 6 and 7**

Using O-D proportions in the first time interval as the initial set of O-D values, this study demonstrates the estimation models with the entry/exit/mainline traffic volume and available travel time information. Table 3.6 compasses the aggregate absolute estimation errors from the origins under different scenarios:

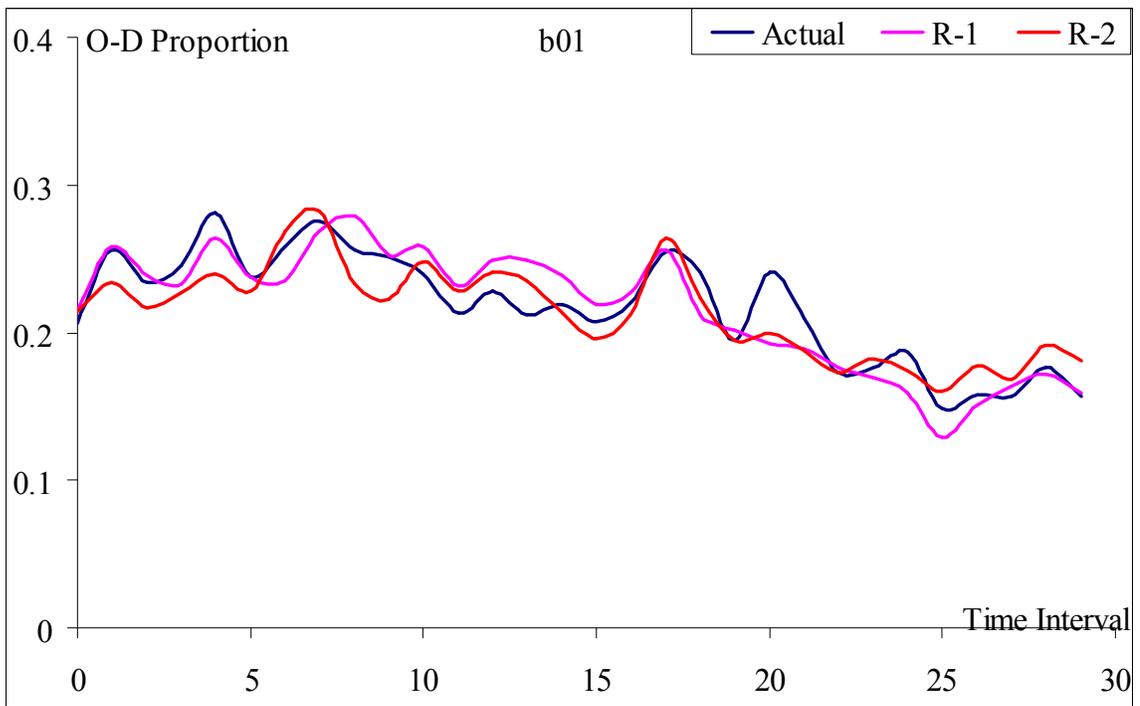
- R-1 – entry/exit volume and travel time information; and
- R-2 – entry/exit/mainline volume and travel time information.

As shown in Table 3.6, by adding the mainline information, the average estimation error for all O-D pairs is reduced from 0.0580 to 0.0543 (6.48%), which is not a substantial reduction. However, the standard deviation decreases significantly, i.e., from 0.0524 to 0.0463 (11.60%), and the maximum estimation error also drops to 0.2857

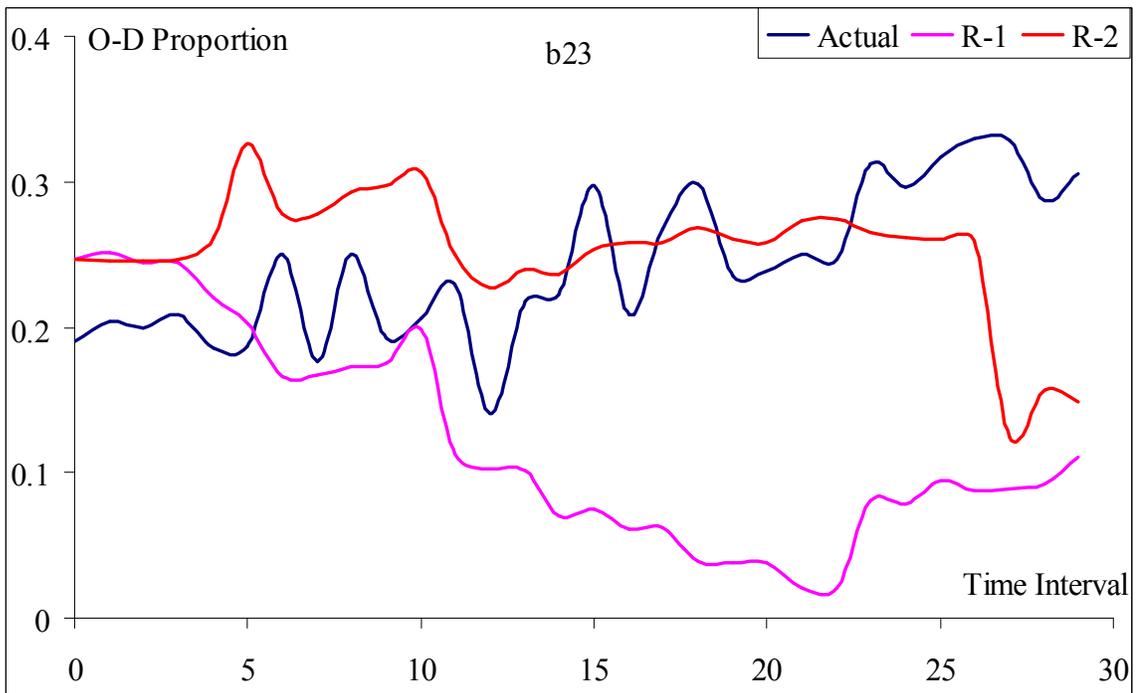
(15.96%). By adding the mainline information, it can be concluded that the proposed formulation can yield a more accurate estimation. Figures 3.15–3.17 illustrate the graphical estimation results for R-1 and R-2.

**Table 3.6. Comparison of Estimation Error Statistics**

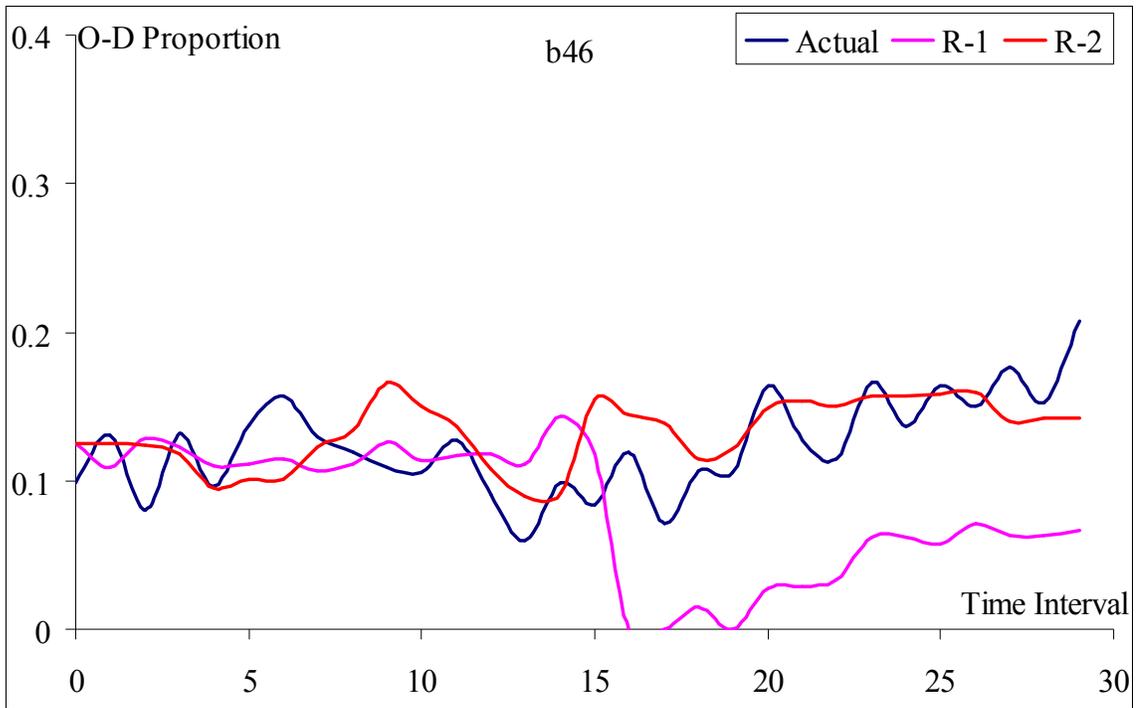
Abs. Errors	Origin 0		Origin 1		Origin 2		Origin 3	
	R-1	R-2	R-1	R-2	R-1	R-2	R-1	R-2
Avg.	0.0313	0.0317	0.0536	0.0487	0.0898	0.0666	0.0548	0.0710
Dev.	0.0267	0.0283	0.0469	0.0431	0.0769	0.0492	0.0424	0.0574
Max.	0.1128	0.1482	0.2203	0.2026	0.3400	0.2043	0.1969	0.2857
Min.	0.0001	0.0000	0.0003	0.0005	0.0015	0.0004	0.0014	0.0012
Abs. Errors	Origin 4		Origin 5		Origin 6		Overall	
	R-1	R-2	R-1	R-2	R-1	R-2	R-1	R-2
Avg.	0.0655	0.0537	0.0719	0.0705	0.0581	0.0621	<b>0.0580</b>	<b>0.0543</b>
Dev.	0.0454	0.0430	0.0468	0.0461	0.0378	0.0415	<b>0.0524</b>	<b>0.0463</b>
Max.	0.1818	0.1785	0.2269	0.1970	0.1282	0.1614	<b>0.3400</b>	<b>0.2857</b>
Min.	0.0004	0.0004	0.0020	0.0035	0.0007	0.0005	0.0001	0.0000



**Figure 3.15. The Graphical Estimation Result for O-D Pair 01**



**Figure 3.16. The Graphical Estimation Result for O-D Pair 23**



**Figure 3.17. The Graphical Estimation Result for O-D Pair 46**

### 3.6 CLOSURE

This chapter proposes a model for estimating the dynamic freeway O-D matrix with a measurable time series of ramp and mainline flows. The proposed model captures the speed variance among vehicles having the same departure time, origin and destination with a specially derived travel time distribution function that results in a substantial reduction in model parameters.

Extensive numerical analyses with respect to sensitivity of both the input measurement errors and the selection of initial parameters have revealed that the proposed model is sufficiently robust for real-world applications. To test the application

of the proposed model to a large-scale network, the study has constructed a simulator based on the I-95 freeway corridor in Maryland using the simulation program, AIMSUN 4.0 program, and performed a model applicability evaluation. The results indicate that the proposed model can yield reasonable estimates of dynamic O-D proportions for large freeway corridors.

One of the more critical issues in development of a dynamic O-D model that requires further study is how to better approximate the initial values of each O-D set from measurable information so that the estimation process with the recursive computing algorithm (such as extended Kalman filtering logic) can evolve more efficiently to a reliable and stable state.

## **CHAPTER 4 AN ALGORITHM FOR ESTIMATING THE INITIAL O-D MATRIX IN A LARGE FREEWAY NETWORK**

### **4.1 INTRODUCTION**

Since some essential data, such as initial O-D set information, for time-dependent O-D estimation may not be available in most real-world traffic networks, it is imperative that any developed system for such applications be sufficiently robust to accommodate to the potential missing information in the input data set.

As reported in the literature (Chang and Wu, 1994; Lin and Chang, 2005), either a historical or an arbitrary O-D demand is often used as an initial O-D set. For a small network, how to set a proper initial O-D demand may not be a critical issue, as the resulting discrepancies, after subsequent use of time-varying traffic information, are mostly insignificant. However, for a large network, a more reliable initial O-D set can generally yield better and more stable estimation results.

The quality of such initial O-D information may contribute to the estimation accuracy and the learning time for the computing algorithm to reach convergence, especially for large networks that contains a significant number of unknown system parameters. A failure to initialize the dynamic estimation process with a reliable O-D set may significantly degrade the quality of those subsequently computed O-D sets.

This critical issue of setting a reliable initial O-D set has not been addressed in the literature for time-dependent O-D estimation with recursive methods, and it is one of the

primary reasons that existing methods for time-dependent O-D estimation are not sufficiently effective for use in large real-world networks. To embody dynamic O-D models with such desirable properties, this study presents an estimation algorithm for systematically refining randomly or arbitrarily generated initial O-D sets so that an employed recursive estimation algorithm can yield more accurate sets of time-dependent O-D patterns over subsequent time intervals.

The remaining of the chapter is organized as follows. The core concept of the proposed estimation algorithm for refinement of an initial O-D set is presented in the following section. A step-by-step description of the initial O-D set estimation algorithm is presented in Section 4.3. Extensive numerical analyses for evaluating the effectiveness of the proposed formulations and solution algorithms are presented in Section 4.4. The key findings and conclusions are summarized in the last section.

## **4.2 CONCEPTS AND RULES OF NETWORK DECOMPOSITION**

The core concept of the estimation algorithm is to estimate the partial O-D proportions iteratively from the first sub-network until the estimates reach the steady state, and then incorporate the estimated parameters to the next sub-network. This procedure is incrementally repeated for each sub-network until all the O-D proportions have been estimated.

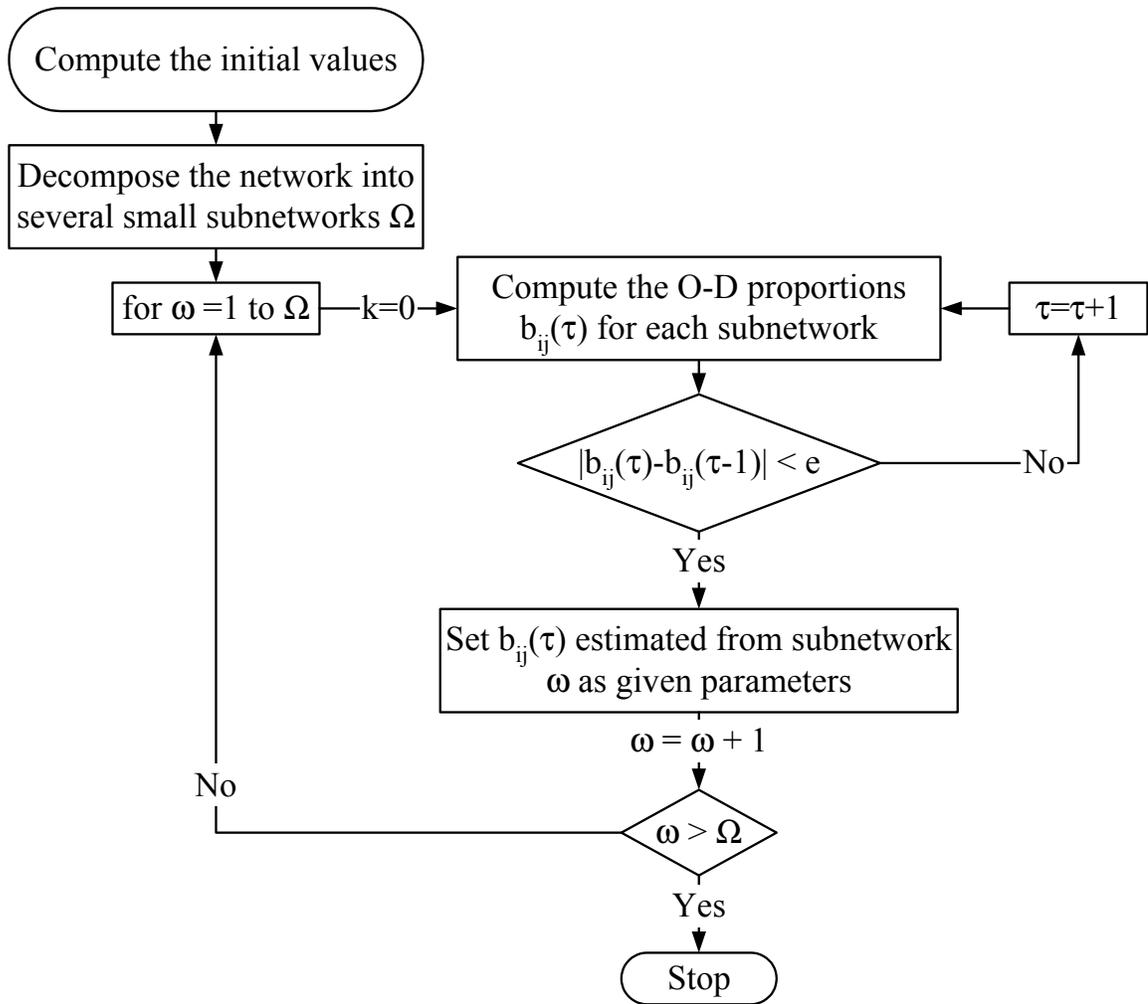
Note that, this estimation algorithm is applied only to compute the initial O-D set, based on input information observed over the first time interval. The time-dependent O-D

demands are then computed subsequently, based on this estimated initial O-D set and other time-series measurements of information, such as traffic volumes and travel times.

Figure 4.1 illustrates the flow chart of the decomposition algorithm. First of all, the entire freeway network is decomposed into  $\Omega$  sub-networks so that the number of unknown parameters is relatively low in each sub-network. One can then execute the base solution approach for each sub-network, such as Kalman filter or Generalized Least Square (GLS) approaches, based on the initial input information when  $k=0$  (e.g., travel time and volume) until it reaches convergency or meets the following relation:

$$|b_{ij}(\tau) - b_{ij}(\tau-1)| < \varepsilon, \text{ for } 0 \leq i < j \leq N \quad (4.1)$$

where  $\tau$  is the number of iterations and  $\varepsilon$  is a small number for the convergence examination, such as 0.0001.



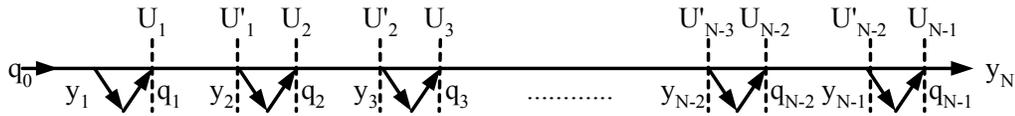
**Figure 4.1. Flow Chart of the Decomposition Algorithm for Initial O-D Estimation**

The initial O-D set computed with the above algorithm is more robust than those generated by other means (e.g., historical information) due primarily to the following reasons:

- The number of unknown parameters is relatively small in the sub-network due to the decomposed size of the network, and the replacement of some unknown parameters with estimated values from the previous sub-network.

- The recursive computing procedures that repeatedly employ actual travel volume and travel time information observed during the 1<sup>st</sup> time interval have captured, to some extent, the relations between the initial O-D set and the resulting volume distribution to the network.

In order to apply any solution approach in the proposed algorithm, one shall follow the following four rules of the network decomposition presented below with the illustrative freeway network shown in Figure 4.2.

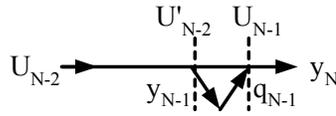


**Figure 4.2. A General Freeway Network**

**Rule-1:** In performing the decomposition, the following natural constraints shall still hold for each sub-network:

$$\sum_{j=i+1}^N b_{ij}(\tau) = 1 \quad i = 0, 1, \dots, N-1 \quad (4.2)$$

Equation (4.2) is the natural constraint for each sub-network. Figure 4.3 illustrates an example of a decomposed sub-network from Figure 4.2, where Equation (4.2) can be written as:  $b_{N-2,N-1}(\tau) + b_{N-2,N}(\tau) = 1$  and  $b_{N-1,N}(\tau) = 1$ .

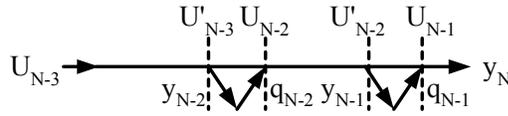


**Figure 4.3. An Example of a Decomposed Sub-network with Rule-1**

**Rule-2:** The first sub-network shall at least consist of two interchanges.

The purpose of the second rule is to ensure that at least one unknown parameter can be estimated from the sub-network. As shown in Figure 4.3, the estimated proportions of the O-D pairs from the sub-network are:  $b_{N-2,N-1}(\tau)$ ,  $b_{N-2,N}(\tau)$  and  $b_{N-1,N}(\tau)$ . Among these three O-D pairs, the first two are not the initial O-D pairs needed for computing in the O-D set for the entire network since their origin node is the mainline link. This type of parameters is defined as the “pseudo O-D parameters”, which denotes these O-D pairs with the mainline link as its origin node. The remaining unknown parameter in this sub-network is  $b_{N-1,N}(\tau)$ , which is obviously equal to 1 since, as shown in Figure 4.3, the only destination for Origin N-1 is N. Therefore, it is not necessary to estimate parameters for this sub-network.

Figure 4.4 shows an example of a decomposed sub-network following Rule-2, in which the unknown parameters are  $b_{N-2,N-1}(\tau)$ ,  $b_{N-2,N}(\tau)$  and  $b_{N-1,N}(\tau)$ , and the parameters  $b_{N-3,N-2}(\tau)$ ,  $b_{N-3,N-1}(\tau)$  and  $b_{N-3,N}(\tau)$  are the intermediate parameters.

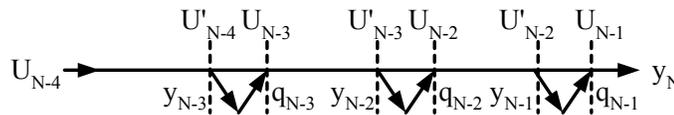


**Figure 4.4. An Example of a Decomposed Sub-network with Rule-2**

**Rule-3:** The next extended sub-network shall increase at least one on-ramp.

The core concept of the proposed algorithm is to compute the unknown parameters from the previous level of a decomposed sub-network, and then use these estimated parameters in the next level of the extended sub-network. By doing so, one can reduce the number of unknown parameters in the extended sub-network, and to improve its computing accuracy as well as efficiency. Hence, this rule is to ensure that every sub-network contains more unknown parameters than its previous sub-network.

Figure 4.5 illustrates an example of a decomposed sub-network generated subsequently after the sub-network shown in Figure 4.4. In this example, the unknown parameters in this sub-network are  $b_{N-3,N-2}(\tau)$ ,  $b_{N-3,N-1}(\tau)$  and  $b_{N-3,N}(\tau)$ , and the intermediate parameters are  $b_{N-4,N-3}(\tau)$ ,  $b_{N-4,N-2}(\tau)$ ,  $b_{N-4,N-1}(\tau)$  and  $b_{N-4,N}(\tau)$  since parameters,  $b_{N-2,N-1}(\tau)$ ,  $b_{N-2,N}(\tau)$  and  $b_{N-1,N}(\tau)$ , have been estimated from the previous sub-network.



**Figure 4.5. An Example of a Decomposed Sub-network with Rule-3**

**Rule-4:** Sub-network  $\omega-1$  shall be contained in Sub-network  $\omega$ , and the final sub-network should be the entire network.

This rule is to ensure that all initial O-D proportions can be estimated in different sub-networks. After executing all the estimation procedures for all sub-networks, the estimated O-D matrix shall be sufficiently robust to serve as a reliable initial O-D set for estimating the O-D parameters over the subsequent intervals.

### **4.3 STEP-BY-STEP PROCEDURES OF THE SOLUTION ALGORITHM FOR INITIAL O-D SETS**

Based on the flow chart and the proposed rules, one can proceed the entire initial O-D set estimation algorithm with the following steps:

#### **Step 0: Network decomposition and initialization**

This step is to decompose the entire target network into  $\Omega$  sub-networks ( $\omega = 0 \sim \Omega$ ) and to reset the inputs of traffic volumes and travel times for each iteration to the initial values, i.e.,  $q_i(\tau) = q_i(0)$ ,  $y_i(\tau) = y_i(0)$ ,  $U_i(\tau) = U_i(0)$ , for  $\tau = 1 \sim T$ , and set  $\omega = 1$ ,  $\tau = 0$ .

Based on the proposed rules, the first sub-network shall consist of two interchanges as shown in Figure 4.4, and the second sub-network shown in Figure 4.5 shall contain three interchanges, and so on. Finally, the last sub-network, as shown in Figure 4.2, is the entire network, which contains all N-1 interchanges.

### Step 1: The O-D proportion estimation for sub-networks

This step is to estimate the O-D proportions for Sub-network  $\omega$  using the solution algorithm presented in Chapter 3. In Sub-network 1 (see Figure 4.4), its O-D matrix consists of three unknown O-D pairs ( $b_{N-2,N-1}$ ,  $b_{N-2,N}$  and  $b_{N-1,N}$ ) and three intermediate O-D pairs ( $b_{N-3,N-2}$ ,  $b_{N-3,N-1}$  and  $b_{N-3,N}$ ). The following set of equations shows an example equation set for Sub-network 1, based on the exit traffic volumes:

$$\begin{aligned} y_{N-2}(\tau) &= \left[ \sum_m U_{N-3}(\tau-m) \rho_{N-3,N-2}^m(\tau) \right] \cdot b_{N-3,N-2}(\tau) \\ y_{N-1}(\tau) &= \left[ \sum_m U_{N-3}(\tau-m) \rho_{N-3,N-1}^m(\tau) \right] \cdot b_{N-3,N-1}(\tau) + \left[ \sum_m q_{N-2}(\tau-m) \rho_{N-2,N-1}^m(\tau) \right] \cdot b_{N-2,N-1}(\tau) \\ y_N(\tau) &= \left[ \sum_m U_{N-3}(\tau-m) \rho_{N-3,N}^m(\tau) \right] \cdot b_{N-3,N}(\tau) + \left[ \sum_m q_{N-2}(\tau-m) \rho_{N-2,N}^m(\tau) \right] \cdot b_{N-2,N}(\tau) \\ &\quad + \left[ \sum_m q_{N-1}(\tau-m) \rho_{N-1,N}^m(\tau) \right] \cdot b_{N-1,N}(\tau) \end{aligned}$$

### Step 2: The convergence evaluation

This step is to check whether the differences of the estimated O-D proportions between  $\tau-1$  and  $\tau$  iterations are within a small range  $\varepsilon$ . The procedures for the convergence checking are summarized below:

1. Set  $\varepsilon$  as a small number.
2. Check if  $|b_{ij}(\tau) - b_{ij}(\tau-1)| < \varepsilon$ .
3. If yes, go to step 3. Otherwise, let  $\tau = \tau+1$  and go to Step 1, and then execute the solution algorithm.

For example, one shall always check if the following inequalities are satisfied or not for Sub-network 1 (see Figure 4.4):

$$|b_{ij}(\tau) - b_{ij}(\tau-1)| < \varepsilon, \quad N-3 \leq i < j \leq N \quad (4.1)$$

**Step 3: The replacement of parameters**

This step is designed to replace the unknown parameters in Sub-network  $\omega+1$  with their estimates from Sub-network  $\omega$  so as to reduce the number of unknown parameters. The procedures are summarized below:

1. Set  $\omega = \omega+1, \tau = 0$ .
2. Set the O-D proportions estimated from Sub-network  $\omega$  as given parameters and use these estimated values to replace those unknown parameters in Sub-network  $\omega+1$ .
3. Check if  $\omega < \Omega$ . If yes, go to Step 1. Otherwise, stop.

By repeating the above procedures and setting the unknown O-D proportions,  $b_{N-2,N-1}$ ,  $b_{N-2,N}$  and  $b_{N-1,N}$  as given, one can have the equations for Exit Traffic Volume with Sub-network 2 as follows (see Figure 4.5):

$$y_{N-3}(k) = \left[ \sum_m U_{N-4}(k-m) \rho_{N-4,N-3}^m(k) \right] \cdot b_{N-4,N-3}(k)$$

$$y_{N-2}(k) = \left[ \sum_m U_{N-4}(k-m) \rho_{N-4,N-2}^m(k) \right] \cdot b_{N-4,N-2}(k) + \left[ \sum_m q_{N-3}(k-m) \rho_{N-3,N-2}^m(k) \right] \cdot b_{N-3,N-2}(k)$$

$$\begin{aligned}
y_{N-1}(k) &= \left[ \sum_m U_{N-4}(k-m) \rho_{N-4,N-1}^m(k) \right] \cdot b_{N-4,N-1}(k) + \left[ \sum_m q_{N-3}(k-m) \rho_{N-3,N-1}^m(k) \right] \cdot b_{N-3,N-1}(k) \\
&\quad + \left[ \sum_m q_{N-2}(k-m) \rho_{N-2,N-1}^m(k) \right] \cdot b_{N-2,N-1}(k) \\
y_N(k) &= \left[ \sum_m U_{N-4}(k-m) \rho_{N-4,N}^m(k) \right] \cdot b_{N-4,N}(k) + \left[ \sum_m q_{N-3}(k-m) \rho_{N-3,N}^m(k) \right] \cdot b_{N-3,N}(k) \\
&\quad + \left[ \sum_m q_{N-2}(k-m) \rho_{N-2,N}^m(k) \right] \cdot b_{N-2,N}(k) + \left[ \sum_m q_{N-1}(k-m) \rho_{N-1,N}^m(k) \right] \cdot b_{N-1,N}(k)
\end{aligned}$$

Note that one shall repeat the above steps until all parameters have been estimated in the last sub-network  $\Omega$ , and the estimated O-D set will be a robust initial set for estimating their subsequent time-dependent O-D sets.

#### 4.4 NUMERICAL EXAMPLE FOR EVALUATING THE INITIAL O-D ESTIMATION ALGORITHM

This section presents the numerical evaluation results of the proposed algorithm using the I-95 corridor between I-495 and I-695 in Maryland (see Figure 3.9). The I-95 corridor as presented in Chapter 3 is used to test the performance of the proposed estimation algorithm under various sets of initial values.

To conduct a comprehensive evaluation, this study has randomly generated 100 experimental initial O-D sets for testing the effectiveness of the proposed solution algorithm. All randomly generated initial O-D value sets shall satisfy the following natural constraints:

$$0 \leq b_{ij}(k) \leq 1, \quad 0 \leq i < j \leq N \quad (3.93)$$

$$\sum_{j=i+1}^N b_{ij}(k) = 1, \quad i = 0, 1, \dots, N-1 \quad (3.4)$$

Table 4.1 shows the true initial value set along with the first 5 examples of the generated initial value sets and their estimated results with the decomposition algorithm.

The estimation results of the initial O-D value sets shown in Table 4.1 reveal the following information:

- After executing the initial O-D set estimation algorithm, each set of randomly generated initial O-D values has converted to a new O-D set, which is closer to their true initial values for most O-D pairs. For example, in Case 1, the randomly generated value of  $b_{03}$  is 0.02. After revised with the decomposition algorithm, it has adjusted to 0.10, which is very close to the true initial value of 0.12.
- Regardless of their initial differences, all randomly generated initial O-D sets after revising by the decomposition algorithm tend to converge to a very similar set of O-D values, which is very close to the true initial O-D set. This seems to show the effectiveness of the proposed estimation algorithm. For example, in Case 4, the randomly generated initial O-D values for O-D pair  $b_{12}$  vary from 0.05 to 0.25. After executing the initial O-D set estimation algorithm, its estimated initial values for these 5 cases are all equal to 0.10, which is very close to its true initial O-D value of 0.13.

Table 4.2 shows the improvement of the initial O-D sets after being refined with the proposed estimation algorithm for these 5 cases. The discrepancy between the set of

true O-Ds and those randomly generated as well as revised with the proposed estimation algorithm are also shown in Table 4.2. The measurement of improvement is defined as below:

$$\text{Average Absolute Error} = \frac{\sum_{i,j} |\hat{b}_{ij}(0) - \bar{b}_{ij}(0)|}{N} \quad (4.3)$$

where  $N$  is the total number of O-D pairs in the network,  $\bar{b}_{ij}(0)$  is the true initial O-D value, and  $\hat{b}_{ij}(0)$  denotes either the randomly generated initial O-D value or the refined initial O-D value from the estimation algorithm.

**Table 4.1a. The Example Initial Value Sets and the Estimated Initial Value Sets from the Initial O-D Set Estimation Algorithm**

$b_{ij}$	$b_{01}$	$b_{02}$	$b_{03}$	$b_{04}$	$b_{05}$	$b_{06}$	$b_{07}$	$b_{08}$	$b_{12}$
True	0.21	0.19	0.12	0.13	0.11	0.09	0.09	0.06	0.13
Case 1	0.13 (0.10)	0.06 (0.07)	<b>0.02 (0.10)</b>	0.14 (0.11)	0.07 (0.06)	0.13 (0.07)	0.19 (0.15)	0.27 (0.34)	0.05 (0.10)
Case 2	0.11 (0.10)	0.17 (0.07)	0.14 (0.10)	0.20 (0.11)	0.15 (0.06)	0.04 (0.07)	0.14 (0.16)	0.06 (0.33)	0.15 (0.10)
Case 3	0.02 (0.10)	0.00 (0.07)	0.01 (0.10)	0.13 (0.11)	0.06 (0.06)	0.15 (0.07)	0.18 (0.15)	0.44 (0.34)	0.11 (0.10)
Case 4	0.16 (0.10)	0.05 (0.07)	0.16 (0.10)	0.20 (0.11)	0.16 (0.06)	0.09 (0.07)	0.13 (0.18)	0.06 (0.32)	<b>0.25 (0.10)</b>
Case 5	0.07 (0.10)	0.19 (0.07)	0.16 (0.10)	0.02 (0.11)	0.00 (0.06)	0.06 (0.07)	0.06 (0.15)	0.44 (0.35)	0.08 (0.10)
<hr/>									
$b_{ij}$	$b_{13}$	$b_{14}$	$b_{15}$	$b_{16}$	$b_{17}$	$b_{18}$	$b_{23}$	$b_{24}$	$b_{25}$
True	0.19	0.11	0.13	0.14	0.17	0.13	0.24	0.20	0.10
Case 1	0.08 (0.13)	0.01 (0.13)	0.08 (0.07)	0.22 (0.08)	0.22 (0.16)	0.34 (0.33)	0.12 (0.16)	0.15 (0.17)	0.08 (0.09)
Case 2	<b>0.08 (0.13)</b>	0.10 (0.13)	0.01 (0.07)	0.03 (0.08)	0.19 (0.17)	0.43 (0.32)	0.10 (0.16)	0.07 (0.17)	0.14 (0.09)
Case 3	0.20 (0.13)	0.02 (0.14)	0.04 (0.07)	0.24 (0.08)	0.01 (0.16)	0.40 (0.33)	0.03 (0.16)	0.16 (0.17)	0.27 (0.09)
Case 4	0.17 (0.13)	0.08 (0.13)	0.12 (0.07)	0.07 (0.08)	0.04 (0.19)	0.28 (0.30)	0.14 (0.16)	0.25 (0.17)	0.29 (0.09)
Case 5	0.17 (0.13)	0.16 (0.14)	0.15 (0.07)	0.17 (0.08)	0.11 (0.15)	0.16 (0.33)	0.05 (0.16)	0.30 (0.17)	0.15 (0.09)

Note: The number in each parenthesis shows the refined initial O-D value from the estimation algorithm.

**Table 4.1b. The Example Initial Value Sets and the Estimated Initial Value Sets from the Initial O-D Set Estimation Algorithm**

$b_{ij}$	$b_{26}$	$b_{27}$	$b_{28}$	$b_{34}$	$b_{35}$	$b_{36}$	$b_{37}$	$b_{38}$	$b_{45}$
True	0.18	0.12	0.16	0.25	0.19	0.12	0.16	0.28	0.10
Case 1	0.25 (0.10)	0.09 (0.09)	0.31 (0.39)	0.17 (0.28)	0.26 (0.14)	0.05 (0.15)	0.13 (0.15)	0.39 (0.28)	0.17 (0.19)
Case 2	0.19 (0.10)	0.20 (0.09)	0.30 (0.40)	0.06 (0.28)	0.01 (0.14)	0.13 (0.15)	0.05 (0.14)	0.76 (0.29)	0.31 (0.19)
Case 3	0.09 (0.10)	0.33 (0.09)	0.13 (0.39)	0.01 (0.28)	0.14 (0.14)	0.17 (0.15)	0.21 (0.15)	<b>0.48 (0.28)</b>	0.17 (0.19)
Case 4	0.01 (0.10)	0.06 (0.08)	0.25 (0.40)	0.10 (0.28)	0.11 (0.14)	0.15 (0.15)	0.06 (0.13)	0.58 (0.30)	0.30 (0.19)
Case 5	<b>0.03 (0.10)</b>	0.27 (0.09)	0.20 (0.39)	0.11 (0.28)	0.21 (0.14)	0.08 (0.15)	0.17 (0.15)	0.44 (0.28)	0.17 (0.19)
$b_{ij}$	$b_{46}$	$b_{47}$	$b_{48}$	$b_{56}$	$b_{57}$	$b_{58}$	$b_{67}$	$b_{68}$	$b_{78}$
True	0.12	0.19	0.59	0.33	0.24	0.43	0.24	0.76	1.00
Case 1	0.22 (0.21)	0.19 (0.20)	0.43 (0.39)	0.35 (0.27)	0.34 (0.25)	<b>0.31 (0.48)</b>	0.32 (0.29)	0.68 (0.71)	1.00 (1.00)
Case 2	0.33 (0.21)	0.18 (0.20)	0.17 (0.40)	0.41 (0.27)	0.40 (0.24)	0.19 (0.49)	0.39 (0.32)	0.61 (0.68)	1.00 (1.00)
Case 3	0.27 (0.21)	0.04 (0.21)	0.52 (0.39)	0.34 (0.27)	0.17 (0.26)	0.48 (0.48)	0.25 (0.27)	0.75 (0.73)	1.00 (1.00)
Case 4	0.25 (0.21)	0.06 (0.18)	0.40 (0.41)	0.06 (0.27)	0.03 (0.23)	0.91 (0.51)	0.34 (0.39)	0.66 (0.61)	1.00 (1.00)
Case 5	0.25 (0.21)	0.15 (0.21)	0.43 (0.38)	0.19 (0.27)	0.49 (0.26)	0.32 (0.48)	0.31 (0.26)	0.69 (0.74)	1.00 (1.00)

Note: The number in each parenthesis shows the refined initial O-D value from the estimation algorithm.

**Table 4.2. The Average Absolute Error and Improvement of the Initial O-D Sets**

Avg. Abs. Error	Randomly generated initial O-D set	Initial O-D set revised by decomposition algorithm	Improvement
Case 1	0.0827	0.0653	20.97%
Case 2	0.1159	0.0666	42.50%
Case 3	0.1034	0.0646	37.59%
Case 4	0.1118	0.0718	35.79%
Case 5	0.0890	0.0639	28.19%

As shown in Table 4.2, the average absolute errors of the initial O-D values in these five cases decrease from 20% up to 40% due to the revision with the initial O-D set estimation algorithm.

To further evaluate the effectiveness of the proposed algorithm, those 100 adjusted initial O-D sets are applied to estimate the time-dependent O-D matrices over subsequent time intervals. The estimation results with their originally generated O-D sets are used as the base line for comparison. The time-average absolute error (TAAE) statistics serve as the evaluation criterion:

$$\text{TAAE} = \frac{\sum_{k=0}^K |\hat{b}_{ij}(k) - \bar{b}_{ij}(k)|}{K} \quad (4.4)$$

Table 4.3 shows the TAAE results of the first 5 cases, which includes the average TAAE of all O-D pairs, deviation, maximum and minimum of the TAAE, and the average improvement percentage of the results with the proposed estimation algorithm.

**Table 4.3. The TAAE Results of the First 5 Cases**

AES	Random initial O-D				Refined with the estimation algorithm				Average Improve
	Average	Deviation	Max	Min	Average	Deviation	Max	Min	
Case 1	0.128	0.119	0.650	0.000	0.077	0.072	0.450	0.000	39.44%
Case 2	0.119	0.108	0.562	0.000	0.074	0.074	0.492	0.001	38.03%
Case 3	0.117	0.113	0.673	0.000	0.080	0.074	0.437	0.000	31.60%
Case 4	0.116	0.106	0.693	0.000	0.083	0.078	0.424	0.000	28.39%
Case 5	0.109	0.106	0.618	0.000	0.080	0.082	0.533	0.000	26.67%

As shown in Table 4.3, the proposed estimation algorithm not only has resulted in significant improvement on the estimated time-dependent O-D sets (e.g., ranging from 26% to 39%), but also has substantially reduced their deviations and the maximum of the TAAE. Figures 4.6-4.10 present the estimated results of one example O-D pair in each case with two different initial conditions: randomly generated initial O-D set (R-2), and the initial O-D set refined from the initial O-D set estimation algorithm (R-3), compared with the true time-dependent O-D demands (R-1).

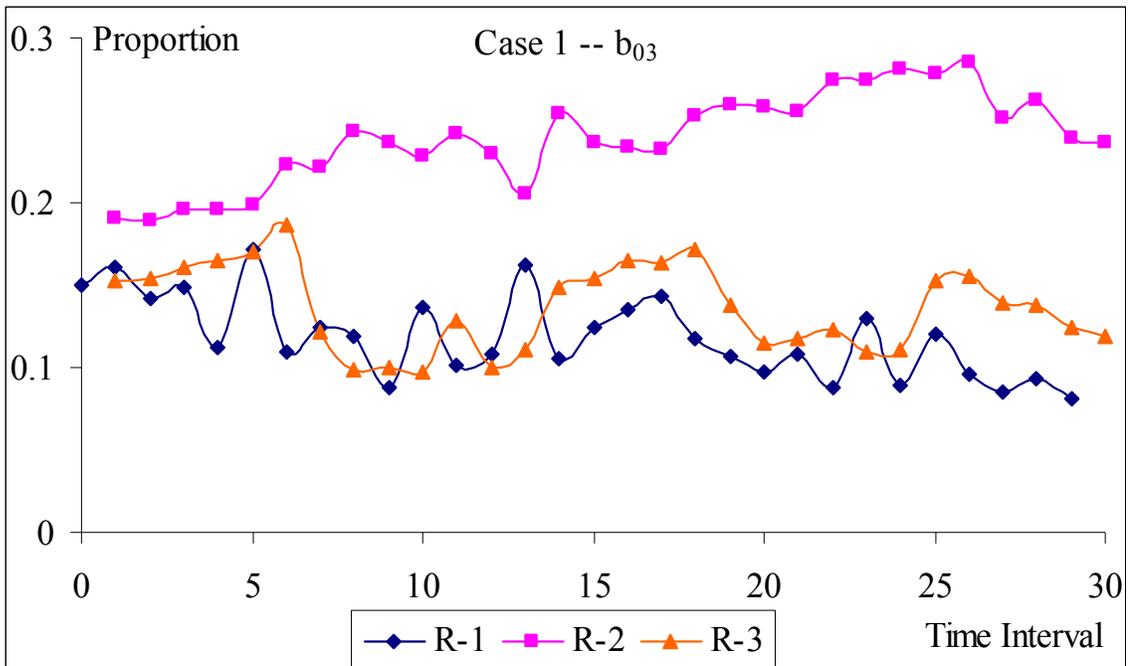


Figure 4.6. Example Estimation Results for O-D pair b<sub>03</sub> in Case 1

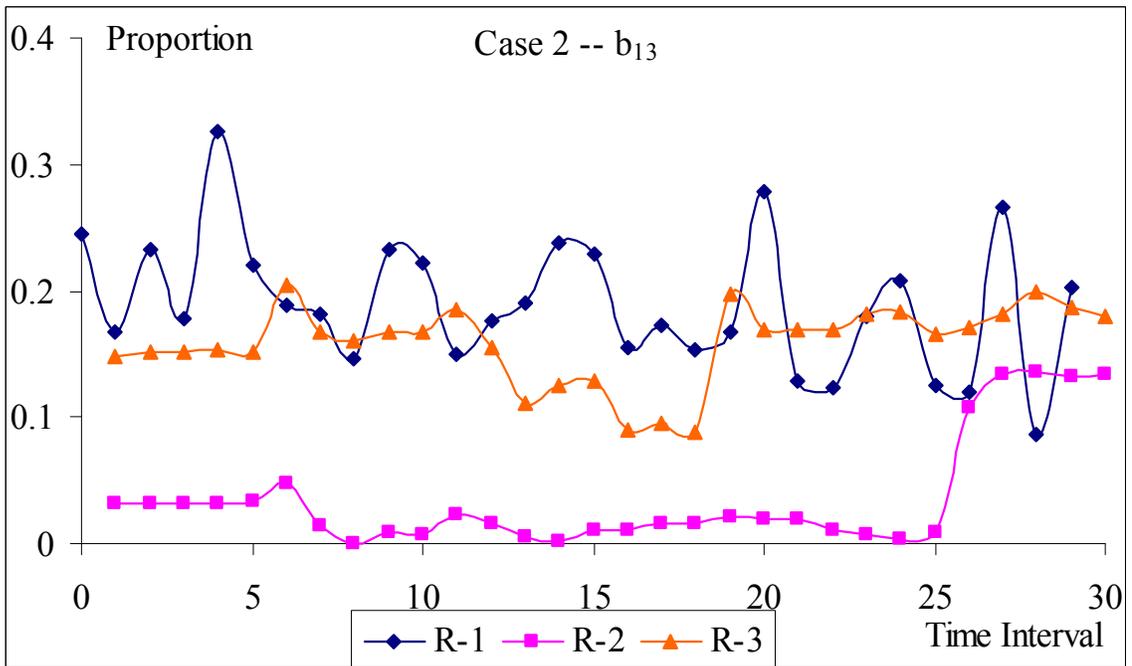


Figure 4.7. Example Estimation Results for O-D pair b<sub>13</sub> in Case 2

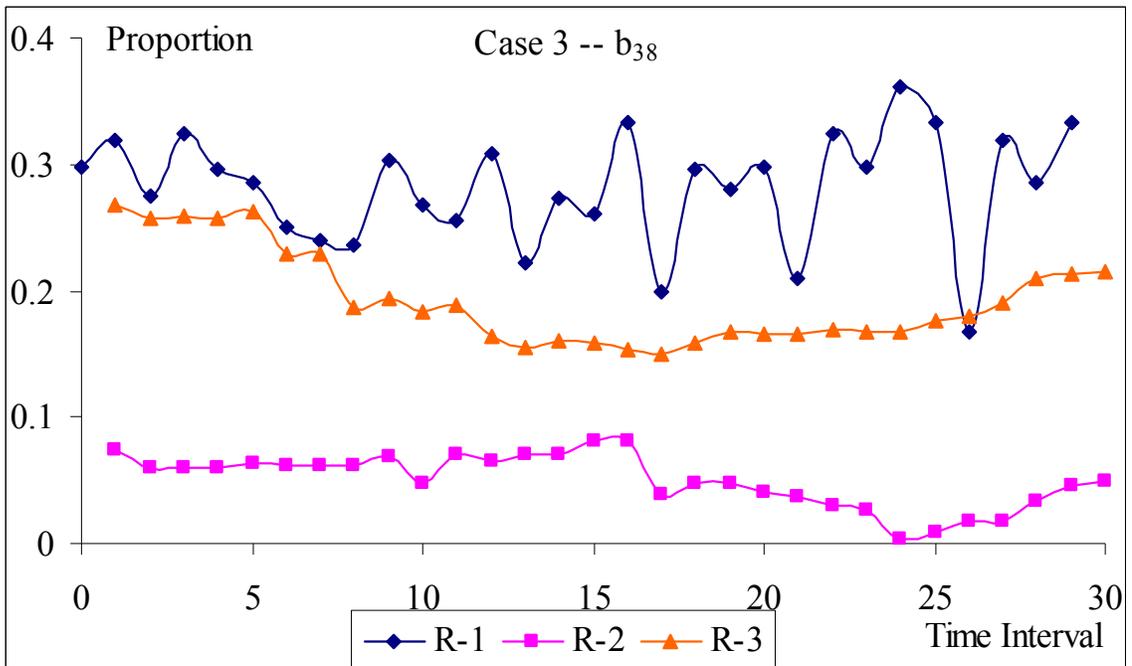


Figure 4.8. Example Estimation Results for O-D pair b<sub>38</sub> in Case 3

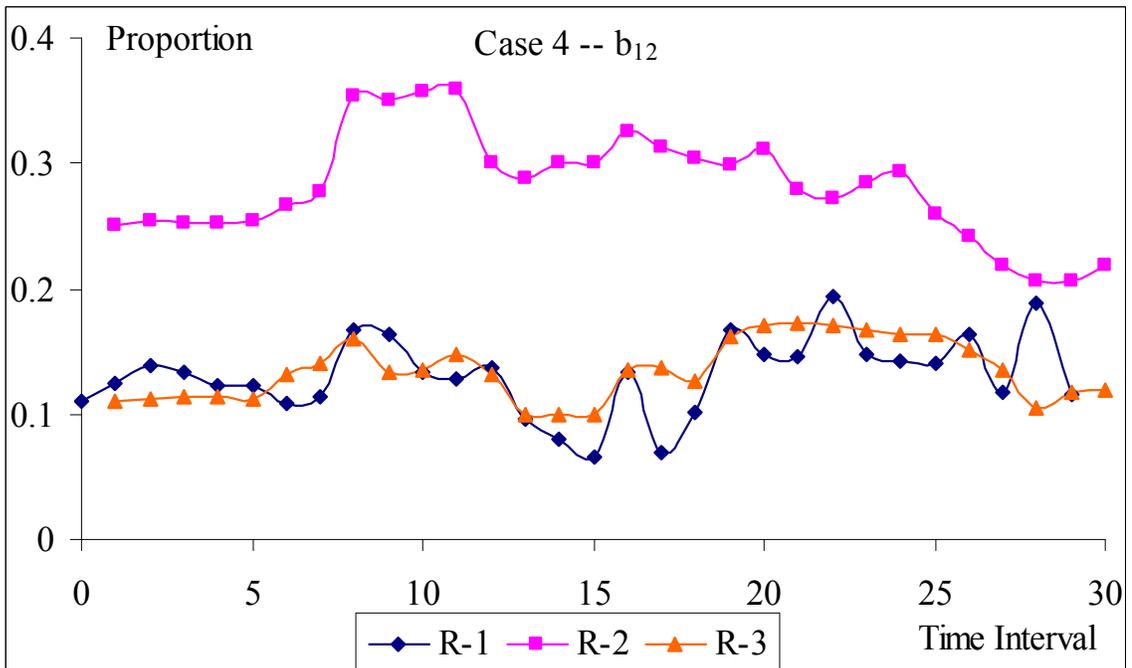
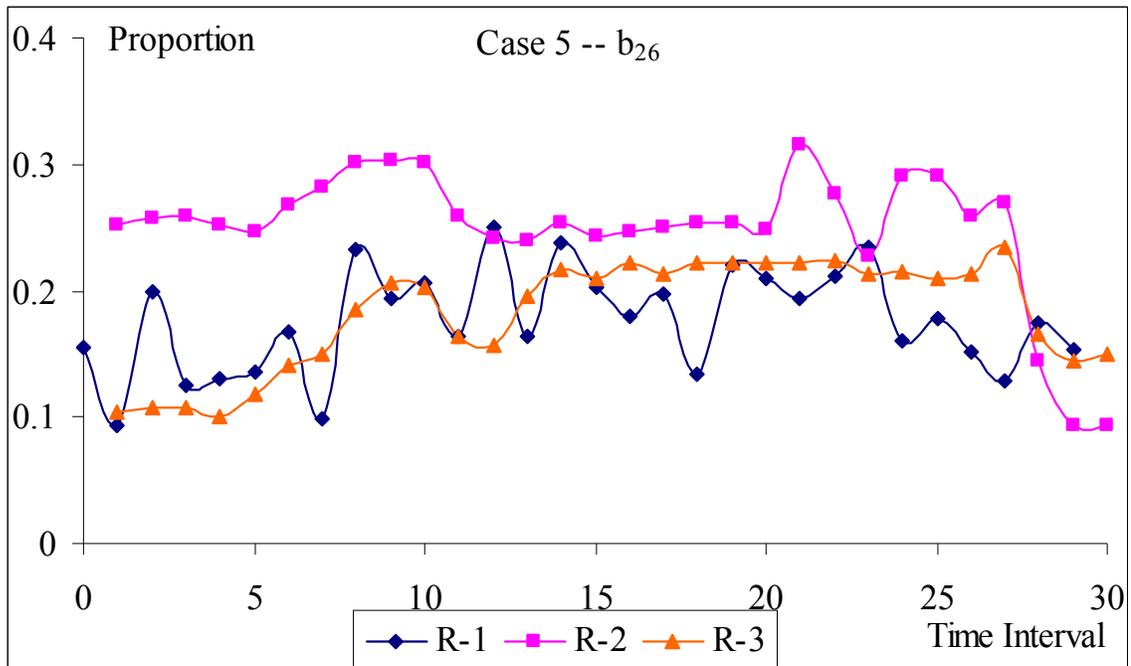


Figure 4.9. Example Estimation Results for O-D pair b<sub>12</sub> in Case 4



**Figure 4.10. Example Estimation Results for O-D pair b26 in Case 5**

Figure 4.6 presents an example of the estimation results of O-D pair  $b_{03}$  in Case 1, in which the estimation results with the initial O-D estimation algorithm exhibit the similar pattern as the one with the true initial O-D set. For this O-D pair, the randomly generated initial O-D is 0.02 and the TAAE based on this initial O-D is 0.1203, but the initial O-D after being refined with the proposed estimation algorithm is adjusted to 0.10, which is very close to the true initial O-D of 0.12 (see Table 4.1). The TAAE based on the refined initial O-D is 0.024, which has achieved about 80% improvement compared with the one with only the randomly generated set.

The estimated time-varying O-D pair  $b_{13}$  in Case 2 is presented in Figure 4.7. For this O-D pair, its initial O-D values after employing the proposed estimation algorithm is

adjusted from 0.48 to 0.28, while the true initial O-D value is 0.28. Although the estimated time-varying O-D patterns, based on the adjusted initial O-D values, exhibit some discrepancy from the true O-D demands (see Figures 4.7), the TAAE is still reduced from 0.2315 to 0.0899, about a 60% improvement. Similar findings also appear for O-D pair  $b_{38}$  in Case 3 as shown in Figures 4.8, in which its initial O-D value is adjusted from 0.08 to 0.13, very close to the true initial O-D value of 0.19. The resulting TAAE has been reduced from 0.1581 to 0.0570 (about 64% improvements).

As shown in Figures 4.9 and 4.10, the time-series O-D patterns estimated with the refined initial O-D sets exhibit very similar patterns as the true time-series O-D demands. In Figure 4.9, the initial O-D value for O-D pair  $b_{12}$  in Case 4 is revised from 0.25 to 0.10, which is close to the true initial O-D value of 0.13. The TAAE, based on this refined initial O-D, has achieved more than 80% improvement compared with the one based on the randomly generated initial O-D set. For the O-D pair  $b_{26}$  in Case 5 (see Figure 4.10), the initial O-D has been adjusted from 0.03 to 0.10, which results in the improvement of the TAAE for over 60%.

Table 4.4 reports the AES for these 100 randomly generated cases, where the proposed algorithm has resulted in significant improvement in 89 out of those 100 cases. The average improvement (in parameter accuracy) is about 12.02%, and the maximal improvement is 39.44%. For those without improvement, the TAAE only increases, on average, about 4%, which stay about the same level as those with randomly generated initial sets of O-D. The overall improvement in TAAE, due to the use of a better initial

O-D set from the decomposition algorithm, for those 100 randomly generated cases is about 10.25%.

**Table 4.4. The Average Absolute Error Statistics**

	AES						Improvement			No. Cases
	Random Initial O-D sets			Refined Initial O-D set			Avg.	Max.	Min.	
	Avg.	Max.	Min.	Avg.	Max.	Min.				
Cases with Improvement	0.1080	0.1449	0.0743	0.0948	0.1206	0.0735	12.02%	39.44%	0.42%	89
Cases without Improvement	0.0965	0.1154	0.0801	0.1004	0.1193	0.0850	-4.07%	-12.92%	-0.33%	11
Total	0.1068	0.1449	0.0743	0.0954	0.1206	0.0735	10.25%	N/A	N/A	100

#### 4.5 CLOSURE

The proposed initial O-D set estimation algorithm can effectively refine a randomly selected set of initial O-D, and yield a significantly better estimate for subsequent time-dependent O-D demands. The potential contribution of such an algorithm to estimation accuracy may increase with the network size and the deficiency level of available historical information. The effectiveness of the proposed algorithm is due mainly to the following reasons:

- The ratio between the number of unknown variables and the number of equations for each sub-network is relatively small compared with the one for the entire network.

- The iterative estimation procedures ensure that these estimates will converge to a set of values, which are more consistent with the observable travel time and volume information when compared with those preset with any other mean.
- The proposed algorithm has taken advantage of all observable information, including travel time information, on-ramp, off-ramp, and mainline volumes.

This study has also presented the performance evaluation results of the proposed estimation algorithm, based on the I-95 freeway corridor between I-495 and I-695 in Maryland. The results have revealed that the estimated time-varying O-D proportions for this example freeway corridor are more reliable due to the use of the algorithm for approximating the initial O-D set.

## **CHAPTER 5 MODELING THE FREEWAY DYNAMIC O-D ESTIMATION SYSTEM WITH MEASUREMENT ERRORS**

### **5.1 INTRODUCTION**

As reported in the literature, most existing approach for dynamic O-D estimation require the use of time-dependent ramp/mainline volumes as model inputs that are generally assumed to be available and contain no measurement errors. However, the assumption of having accurate link volume data is often subject to challenge, as most traffic volume data from detectors are constantly suffering from the hardware quality deficiency. Neglecting the impact of the data quality in the model formulations may contribute to significant estimation errors, especially for a large network as most dynamic O-D estimation models generally contain a large number of unknown system parameters.

Depending on the information availability and the network structure, the time-dependent O-D patterns estimated with those approaches may yield insufficient reliability for use in practice. To deal with such deficiencies embedded in the existing models, this study has developed an interval-based dynamic O-D model to account for measurement variances in network link volumes, which may vary significantly from their average levels. The proposed interval-based model recognizes the inevitable variation of day-to-day link volumes and the limitations of sensor technologies, allowing users to incorporate the reasonable upper and lower bounds of all input data in the model formulations and

parameter estimation. The dynamic O-D patterns estimated with the proposed methods are thus more likely to be reliable and relatively insensitive to data measurement errors.

The remaining of chapter is organized as follows. The next section presents the extended model formulations with measurement errors, based on the method proposed in Chapter 3. Section 5.3 illustrates an interval-based solution algorithm to contend with the input volume variance. Section 5.4 presents the detailed discussions of extensive numerical analyses for evaluating the effectiveness of the proposed model and solution algorithm. Finally, the summary of this chapter is presented in the last section.

## **5.2 THE MODEL FORMULATIONS TO ACCOUNT FOR MEASUREMENT ERRORS**

This section presents an interval-based model that can estimate time-dependent O-D demands under the circumstances that the traffic volumes may vary within a feasible range due to some measurement errors. Consider the following state-space system for estimating time-dependent O-D demands as presented in Chapter 3:

$$\mathbf{b}(\mathbf{k}+1) = \mathbf{b}(\mathbf{k}) + \mathbf{W}(\mathbf{k}) \quad (3.19)$$

$$\boldsymbol{\sigma}(\mathbf{k}+1) = \boldsymbol{\sigma}(\mathbf{k}) + \mathbf{V}(\mathbf{k}) \quad (3.20)$$

$$\mathbf{Z}(\mathbf{k}) = \mathbf{H}(\mathbf{k}) \mathbf{b}(\mathbf{k}) + \mathbf{e}(\mathbf{k}) \quad (3.21)$$

where Equations (3.19) and (3.20) are for illustrating the temporal evolution of the target system, and Equation (3.21) is for the measurement relation. One can solve such a system

with various existing solution approaches, such as the Kalman filtering or the least squares estimation approaches.

The matrices  $\mathbf{Z}(\mathbf{k})$  and  $\mathbf{H}(\mathbf{k})$  in Equation (3.21) consist of the time-series measurement of traffic volumes,  $q_i(\mathbf{k})$ ,  $y_i(\mathbf{k})$  and  $U_i(\mathbf{k})$ . If the traffic volume information contains some measurement errors, which may change progressively over time, most existing models and solution algorithms will not be applicable. The model formulations and the solution algorithm presented hereafter are proposed for tackling this critical issue.

Suppose that these two matrices of  $\mathbf{Z}(\mathbf{k})$  and  $\mathbf{H}(\mathbf{k})$  contain some level of uncertainty, which are only known to be bounded, one can then rewrite these two matrices as follows:

$$\mathbf{Z}^I(\mathbf{k}) = [\mathbf{Z}(\mathbf{k}) - |\Delta\mathbf{Z}(\mathbf{k})|, \mathbf{Z}(\mathbf{k}) + |\Delta\mathbf{Z}(\mathbf{k})|] \quad (5.1)$$

$$\mathbf{H}^I(\mathbf{k}) = [\mathbf{H}(\mathbf{k}) - |\Delta\mathbf{H}(\mathbf{k})|, \mathbf{H}(\mathbf{k}) + |\Delta\mathbf{H}(\mathbf{k})|] \quad (5.2)$$

where  $|\Delta\mathbf{Z}(\mathbf{k})|$  and  $|\Delta\mathbf{H}(\mathbf{k})|$  are the positive constant bounds for those unknowns, and  $\mathbf{Z}^I(\mathbf{k})$  and  $\mathbf{H}^I(\mathbf{k})$  are the interval matrices with their lower and upper bounds. Hence, Equation (3.21) can be written with the following interval expression:

$$\mathbf{Z}^I(\mathbf{k}) = \mathbf{H}^I(\mathbf{k}) \mathbf{b}^I(\mathbf{k}) + \mathbf{e}^I(\mathbf{k}) \quad (5.3)$$

with  $\mathbf{b}^I(\mathbf{k})$  denotes the interval matrix of unknown variables and  $\mathbf{e}^I(\mathbf{k})$  is the interval matrix of error terms. Equations (3.19) and (3.20) become:

$$\mathbf{b}^I(\mathbf{k}+1) = \mathbf{b}^I(\mathbf{k}) + \mathbf{W}^I(\mathbf{k}) \quad (5.4)$$

$$\boldsymbol{\sigma}^I(\mathbf{k}+1) = \boldsymbol{\sigma}^I(\mathbf{k}) + \mathbf{V}^I(\mathbf{k}) \quad (5.5)$$

Equation (5.3) along with Equations (5.4) and (5.5) forms an interval-based state space system, which formulates the observable volumes as intervals rather than constants.

Since the unknown variables,  $b_{ij}(k)$ , in the revised formulations, become intervals other than constant values, one needs to restructure these two natural constraints in Equations (3.3) and (3.4):

$$0 \leq b_{ij}(k) \leq 1, \quad 0 \leq i < j \leq N \quad (3.3)$$

$$\sum_{j=i+1}^N b_{ij}(k) = 1, \quad i = 0, 1, \dots, N-1 \quad (3.4)$$

To reformulate these two natural constraints, let the lower and upper bounds of the unknown variables be defined as  $\underline{b}_{ij}(k)$  and  $\bar{b}_{ij}(k)$ , respectively, with their boundaries lie within 0 and 1 (See Equations (5.6) and (5.7)).

$$0 \leq \underline{b}_{ij}(k) \leq 1, \quad 0 \leq i < j \leq N \quad (5.6)$$

$$0 \leq \bar{b}_{ij}(k) \leq 1, \quad 0 \leq i < j \leq N \quad (5.7)$$

With the above definitions, one can restructure Equations (3.4) as Equations (5.8) and (5.9). Note that Equation (5.8) is to ensure that the sum of the lower bound,  $\underline{b}_{ij}(k)$ , of all O-D proportions with the same origin lies between zero and one, where as Equation (5.9) is to let the sum of the upper bound,  $\bar{b}_{ij}(k)$ , of all O-D proportions with the same origin be equal or larger than one. If all the O-D proportions satisfy these two constraints, it is guaranteed that there exists at least one combination that can satisfy Equation (3.4).

$$0 \leq \sum_{j=i+1}^N \underline{b}_{ij}(k) \leq 1, \quad i = 0, 1, \dots, N-1 \quad (5.8)$$

$$\sum_{j=i+1}^N \bar{b}_{ij}(k) \geq 1, \quad i = 0, 1, \dots, N-1 \quad (5.9)$$

Since  $b_{ij}^I(k)$  is an interval, many possible combinations can satisfy this constraint. For example, two O-D proportions with the same origin are  $b_{01}=0.65$  and  $b_{02}=0.35$  and the sum of these two proportions equals 1. If these two O-D proportions are presented in an interval representation as  $b_{01}^I = [0.6, 0.8]$  and  $b_{02}^I = [0.3, 0.5]$ . It is not necessary that either the sum of the lower bound or the sum of the upper bound be equal to one as there exists many possible combinations. In such a case, the estimated O-D proportions are valid as long as there exist combinations that can satisfy the constraint. Hence, one can present the interval-based system as follows:

$$\mathbf{Z}^I(k) = \mathbf{H}^I(k) \mathbf{b}^I(k) + \mathbf{e}^I(k)$$

$$\mathbf{b}^I(k+1) = \mathbf{b}^I(k) + \mathbf{W}^I(k)$$

$$\boldsymbol{\sigma}^I(k+1) = \boldsymbol{\sigma}^I(k) + \mathbf{V}^I(k)$$

$$0 \leq \underline{b}_{ij}(k) \leq 1, 0 \leq \bar{b}_{ij}(k) \leq 1, 0 \leq i < j \leq N$$

$$0 \leq \sum_{j=i+1}^N \underline{b}_{ij}(k) \leq 1, \sum_{j=i+1}^N \bar{b}_{ij}(k) \geq 1, i = 0, 1, \dots, N-1$$

### 5.3 AN INTERVAL-BASED SOLUTION ALGORITHM

With Equations (5.3)–(5.5), one can construct an interval-based dynamic model for time-dependent O-D estimation that takes into account potential measurement errors. This set of formulations represents each traffic volume as an interval so it consists of the upper and lower bounds in the system equations. For such a dynamic system, the solution generated from standard algorithms, such as the standard Kalman filtering methods, may

not encompass all possible optimal solutions within the specified intervals. (i.e., Equations (5.3)–(5.5)). Hence, this study employs the *interval Kalman filtering scheme* (Chui and Chen, 1999) concept to derive a solution algorithm for the proposed interval-based dynamic system. The core logic of the interval Kalman filtering approach is summarized below. Some other related details interval arithmetic and interval analysis are available in Appendix A.

Initialization:  $\hat{\mathbf{b}}^1(\mathbf{0}) = E[\mathbf{b}^1(\mathbf{0})]$ ,  $\mathbf{P}^1(\mathbf{0}) = \text{Var}[\mathbf{b}^1(\mathbf{0})]$

For  $k=1, 2, \dots$

$$\mathbf{M}^1(\mathbf{k}-1) = \mathbf{P}^1(\mathbf{k}-1) + \mathbf{Q}(\mathbf{k}-1) \quad (5.10)$$

$$\mathbf{G}^1(\mathbf{k}) = \mathbf{M}^1(\mathbf{k}-1) [\mathbf{H}^1(\mathbf{k})]^T [\mathbf{H}^1(\mathbf{k}) \mathbf{M}^1(\mathbf{k}-1) [\mathbf{H}^1(\mathbf{k})]^T + \mathbf{R}_k]^{-1} \quad (5.11)$$

$$\hat{\mathbf{b}}^1(\mathbf{k}) = \hat{\mathbf{b}}^1(\mathbf{k}-1) + \mathbf{G}^1(\mathbf{k})[\mathbf{v}^1(\mathbf{k}) - \mathbf{H}^1(\mathbf{k})\hat{\mathbf{b}}^1(\mathbf{k}-1)] \quad (5.12)$$

$$\mathbf{P}^1(\mathbf{k}) = [\mathbf{I} - \mathbf{G}^1(\mathbf{k})\mathbf{H}^1(\mathbf{k})]\mathbf{M}^1(\mathbf{k}-1)[\mathbf{I} - \mathbf{G}^1(\mathbf{k})\mathbf{H}^1(\mathbf{k})]^T + [\mathbf{G}^1(\mathbf{k})]\mathbf{R}_k[\mathbf{G}^1(\mathbf{k})]^T \quad (5.13)$$

where  $\mathbf{Q}$  is the covariance matrix for the measurement errors, and  $\mathbf{R}$  is the covariance matrix of the random error.

Equation (5.11) involves a complex computation of the interval matrix inversion. To improve the computational efficiency, one can replace this inversion part with its worse-case inversion (Chen, et al., 1997). Then, Equation (5.11) is simplified as:

$$\mathbf{G}^1(\mathbf{k}) = \mathbf{M}^1(\mathbf{k}-1)[\mathbf{H}^1(\mathbf{k})]^T [\mathbf{H}(\mathbf{k})\mathbf{M}(\mathbf{k}-1)[\mathbf{H}(\mathbf{k})]^T + \mathbf{R}_k]^{-1} \quad (5.14)$$

As reported in the literature (Nihan and Davis, 1989), most studies employ the methods of truncation and normalization (as shown in Chapter 3) to tackle the difficulty of taking into account the natural constraints. In applying the method of truncation to this set of formulations, one shall compute the following equation:

$$\alpha' = \text{MAX}_{0 \leq \alpha \leq 1} [\alpha | 0 \leq [b^{i-1}] + \alpha \delta^i g^i \leq 1] \text{ and set } \begin{bmatrix} b^i \\ \sigma^i \end{bmatrix} = \begin{bmatrix} b^{i-1} \\ \sigma^{i-1} \end{bmatrix} + \alpha' \delta^i g^i \quad (3.22)$$

By the same token, in performing the normalization, it is necessary to compute the following relations:

$$\beta_m = \sum_{j=m+1}^N b_{mj}^i \text{ and } b_{mj}^i = b_{mj}^i / \beta_m, \quad j = m+1, \dots, N, \text{ for } m = 1, 2, \dots, N-2 \quad (3.23)$$

To incorporate the above concepts in the specified natural constraints (i.e., Equations (5.6)–(5.9)), one needs to modify the truncation and normalization process as follows:

$$\begin{bmatrix} b_n^I \\ \sigma_n^I \end{bmatrix} = \begin{bmatrix} b_{n-1}^I \\ \sigma_{n-1}^I \end{bmatrix} + \alpha' \cdot \delta_n^I \cdot g_n^I$$

where  $\alpha' = \text{MAX}_{0 \leq \alpha \leq 1} [\alpha | 0 \leq [b_{n-1}^I] + \alpha \cdot \delta_n^I \cdot g_n^I \leq 1]$  (5.15)

$$0 \leq \sum_{j=m+1}^N \underline{b}_{mj}^{n-1} \leq 1, \quad [\underline{b}_{mj}^{n-1}] = \text{lower bound}\{[b_{n-1}^I] + \alpha \cdot \delta_n^I \cdot g_n^I\}, \quad j = m+1, \dots, N \quad (5.16)$$

$$\sum_{j=m+1}^N \bar{b}_{mj}^{n-1} \geq 1, \quad [\bar{b}_{mj}^{n-1}] = \text{upper bound}\{[b_{n-1}^I] + \alpha \cdot \delta_n^I \cdot g_n^I\}, \quad j = m+1, \dots, N \quad (5.17)$$

Note that the purpose of the truncation is to find the largest step of improvement for the unknown variables,  $b$  and  $\sigma$ , for the next time interval so that the O-D proportions,  $b_{ij}$ , can still satisfy Equations (5.6) and (5.7). To incorporate this concept in developing the solution algorithm, it is essential that Equation (3.22) be restated as Equation (5.15), where  $\alpha$  is a value that represents the largest step of improvement so that all the possible values of  $b_{ij}$  can satisfy Equations (5.6) and (5.7). The normalization step is employed to satisfy Equations (5.8) and (5.9). Hence, one can formulate such relationship with Equations (5.16) and (5.17).

With all aforementioned reformulations, this research has presented an enhanced solution algorithm for solving the system formulations of Equations (5.3)–(5.5) and the interval-based natural constraints (5.6)–(5.9) as follows:

Step 0: Initialization

- Link length  $L_i$ ,  $i = 0, 1, \dots, N-1$
- Length of each time interval,  $t_0$ , and the maximal number of intervals required to traverse the entire section  $M$
- Initial input mean speeds,  $V_i(m)$ ,  $m = -M, -M+1, \dots, 0$
- Initial input flows,  $q_i^1(m)$ ,  $m = -M, -M+1, \dots, 0$
- Initial travel times,  $t_{ij}(m) = L_i/V_i(m) + \dots + L_{j-1}/V_{j-1}$ ,  $m = -M, -M+1, \dots, 0$
- $\text{Var}[e(k)] = \text{diag}[r_1, r_2, \dots, r_{2N-1}]$

$$\begin{bmatrix} \mathbf{b}^I(0) \\ \boldsymbol{\sigma}^I(0) \end{bmatrix} = \mathbb{E} \begin{bmatrix} \mathbf{b}^I(0) \\ \boldsymbol{\sigma}^I(0) \end{bmatrix}, \quad \mathbf{P}_0^I = \text{Var} \begin{bmatrix} \mathbf{b}^I(0) \\ \boldsymbol{\sigma}^I(0) \end{bmatrix}$$

Step 1: Compute the Travel Time (mean value)

$$u_{ij}(\mathbf{k}) = t_{ij}(\mathbf{k})$$

Step 2: Compute the Linearized Transformation Matrix

- $\mathbf{H}^I(\mathbf{k}) = [\mathbf{h}_{rs}^I(\mathbf{k})]_{(2N-1) \times N(N+1)/2}$

$$h_{j, Ni+j-i(i+1)/2}^I(\mathbf{k}) = \sum_{m=0}^M [q_i^I(\mathbf{k}-m) \cdot \int_{m \cdot t_0}^{(m+1) \cdot t_0} f_{ij}^I(x) dx], \quad \text{for } 0 \leq i < j \leq N$$

$$h_{N+\ell, Ni+j-i(i+1)/2}^I(\mathbf{k}) = \sum_{m=0}^M [q_i^I(\mathbf{k}-m) \cdot \int_{m \cdot t_0}^{(m+1) \cdot t_0} f_{i\ell j}^I(x) dx], \quad \text{for } 0 \leq i < \ell < j \leq N$$

- $\mathbf{J}^I(\mathbf{k}) = [\mathbf{j}_{rs}^I(\mathbf{k})]_{(2N-1) \times N(N+1)/2}$

$$j_{j, Ni+j-i(i+1)/2}^I(\mathbf{k}) = \sum_{m=0}^M q_i^I(\mathbf{k}-m) \cdot b_{ij}^I(\mathbf{k}), \quad \text{for } 0 \leq i < j \leq N$$

$$j_{N+\ell, Ni+\ell-i(i+1)/2}^I(\mathbf{k}) = \sum_{m=0}^M q_i^I(\mathbf{k}-m) \cdot \sum_{j \leq \ell} b_{ij}^I(\mathbf{k}), \quad \text{for } 0 \leq i < \ell < j \leq N$$

- $h_{rs}^I(\mathbf{k}) = 0$  and  $j_{rs}^I(\mathbf{k}) = 0$ , for the other entries of matrix  $\mathbf{H}^I(\mathbf{k})$  and  $\mathbf{J}^I(\mathbf{k})$

- $\mathbf{F}^I(\mathbf{k}) = \begin{bmatrix} \mathbf{f}_1^I \\ \mathbf{f}_2^I \\ \vdots \\ \mathbf{f}_{2N-1}^I \end{bmatrix} = [\mathbf{H}^I(\mathbf{k}) \mathbf{J}^I(\mathbf{k})]_{(2N-1) \times N(N+1)}$

where each  $\mathbf{f}_i^I$  is a row vector of dimension  $N(N+1)$

$$\bullet \quad \mathbf{Z}^I(\mathbf{k}) = \begin{bmatrix} z_1^I \\ z_2^I \\ \vdots \\ z_{2N-1}^I \end{bmatrix} = [y_1^I(\mathbf{k}), \dots, y_N^I(\mathbf{k}), U_1^I(\mathbf{k}) - q_1^I(\mathbf{k}), \dots, U_{N-1}^I(\mathbf{k}) - q_{N-1}^I(\mathbf{k})]^T$$

### Step 3: Initialization of the sequential Kalman Filtering

- Set  $\mathbf{b}_0^I = \mathbf{b}^I(k-1)$ ,  $\sigma_0^I = \sigma^I(k-1)$
- $\mathbf{P}_0^I = \mathbf{P}^I(k-1) + D$

### Step 4: Sequential Kalman Filtering Iteration

For  $n = 1, 2, \dots, 2N-1$

- $\mathbf{g}_n^I = \mathbf{P}_{n-1}^I (\mathbf{f}_n^I)^T [\mathbf{f}_n^I \mathbf{P}_{n-1}^I (\mathbf{f}_n^I)^T + \mathbf{r}_n]^I^{-1}$
- $\mathbf{P}_n^I = \mathbf{P}_{n-1}^I - \mathbf{g}_n^I \mathbf{f}_n^I \mathbf{P}_{n-1}^I$
- $\delta_{n-1}^I = y_n^I(\mathbf{k}) - \mathbf{f}_n^I \mathbf{b}^I(k-1)$
- $\alpha' = \text{MAX}_{0 \leq \alpha \leq 1} [\alpha \mid (1) \sim (3)]$

$$(1) \quad 0 \leq [\mathbf{b}_{n-1}^I] + \alpha \cdot \delta_{n-1}^I \cdot \mathbf{g}_n^I \leq 1$$

$$(2) \quad 0 \leq \sum_{j=m+1}^N \underline{b}_{mj}^{n-1} \leq 1, \quad j = m+1, \dots, N$$

$$(3) \quad \sum_{j=m+1}^N \bar{b}_{mj}^{n-1} \geq 1, \quad j = m+1, \dots, N$$

where  $\underline{b}_{mj}^{n-1}$  and  $\bar{b}_{mj}^{n-1}$  are the lower and upper bounds of  $b_{n-1,mj}^I$ ,  $[\mathbf{b}_{n-1}^I] = [b_{n-1,mj}^I]$ .

$$\text{set } \begin{bmatrix} \mathbf{b}_n^I \\ \sigma_n^I \end{bmatrix} = \begin{bmatrix} \mathbf{b}_{n-1}^I \\ \sigma_{n-1}^I \end{bmatrix} + \alpha' \cdot \delta_{n-1}^I \cdot \mathbf{g}_n^I$$

### Step 5: Prediction of the States

- Set  $P^I(k) = P_{2n-1}^I$

$$\begin{bmatrix} \underline{b}(k) \\ \underline{\sigma}(k) \end{bmatrix} = \begin{bmatrix} (\underline{b}^{2N-1} + \bar{b}^{2N-1})/2 \\ (\underline{\sigma}^{2N-1} + \bar{\sigma}^{2N-1})/2 \end{bmatrix}$$

$k=k+1$ , go to Step 1 for the next interval.

## **5.4 NUMERICAL EXAMPLES FOR EVALUATING THE INTERVAL-BASED DYNAMIC O-D ESTIMATION MODEL**

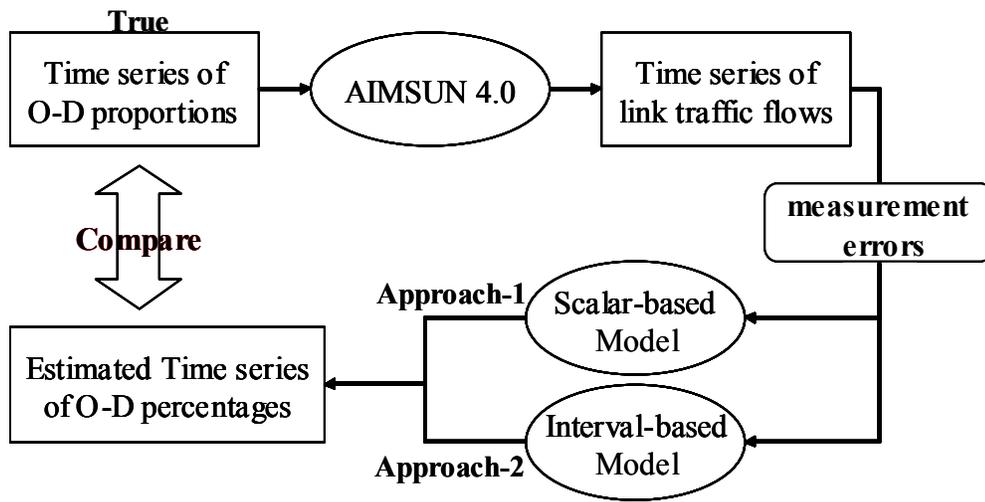
This section presents the evaluation results of the proposed interval-based model and algorithm using the I-95 northbound freeway corridor between two major beltways, I-495 and I-695 as presented in Chapter 3 (see Figure 3.9).

### **5.4.1 Data Generation and Experiment Designs**

This section describes the process of data generation and the design of experiments for evaluating the performance of the proposed interval-based model and algorithm.

Figure 5.1 illustrates the procedures for data generation for the performance evaluation. As mentioned in Chapter 3, the I-95 corridor was simulated with AIMSUN 4.0 using the assigned time series O-D percentages (see Table 3.4). Based on the simulation results, one can obtain the volume for each link and the average travel time for each O-D pair. The time series of traffic flows on both ramps and mainline links are

randomly increases or decreased (i.e.,  $\pm 5\%$  or  $\pm 10\%$ ) to represent the measurement errors caused by detection devices. As shown in Figure 5.1, a set of time-dependent O-D data estimated from Approach-1 (i.e., the scalar-based approach presented in Chapter 3) is set as the basis for comparing with Approach-2 (i.e., the proposed interval-based model and algorithm).



**Figure 5.1. Data Generation Procedures**

To conduct a comprehensive evaluation for the proposed model and algorithm, this study has developed two experiment designs as below.

**Experiment-1:** Comparing the performance of Approach-1 and Approach-2 under different levels of measurement errors in input flow data

This experiment is mainly designed to see if the interval-based approach (Approach-2) can overcome the deficiency of the input data under different levels of measurement errors. Table 5.1 presents five scenarios with the level of input measurement errors ranging from 5% to 30%. In these scenarios, the initial O-D set is pre-estimated from the initial O-D estimation algorithm proposed in Chapter 4.

**Table 5.1. Scenarios under Different Levels of Measurement Errors in Flow Data**

Scenario	Ramp/mainline Flow	Initial O-D Set
Scenario-1	Available with $\pm 5\%$ measurement errors	A reliable initial O-D matrix from the initial O-D estimation algorithm
Scenario-2	Available with $\pm 10\%$ measurement errors	A reliable initial O-D matrix from the initial O-D estimation algorithm
Scenario-3	Available with $\pm 15\%$ measurement errors	A reliable initial O-D matrix from the initial O-D estimation algorithm
Scenario-4	Available with $\pm 20\%$ measurement errors	A reliable initial O-D matrix from the initial O-D estimation algorithm
Scenario-5	Available with $\pm 30\%$ measurement errors	A reliable initial O-D matrix from the initial O-D estimation algorithm

**Experiment-2:** The performance evaluation of the interval-based approach with the initial O-D estimation algorithm

To further evaluate the compound effect of the interval-based algorithm and initial O-D estimation algorithm, the study designed the second set of experiments as shown in Table 5.2. Scenario-1 is the worst case without pre-estimated initial O-D set and Scenario-2 is the best case, which has the true initial O-D set. Scenario-3 has the initial

O-D set estimated from the initial O-D estimation algorithm proposed in Chapter 4, while the uniformed distributed initial O-D set is used in Scenario-1.

**Table 5.2. Scenarios under Different Levels of Availability in Initial O-D set**

Scenario \ Data	Ramp/mainline volume	Initial O-D matrix
Scenario-1	Available with $\pm 10\%$ measurement errors	Not available (uniformly distributed O-D set)
Scenario-2	Available with $\pm 10\%$ measurement errors	A True Initial O-D Matrix
Scenario-3	Available with $\pm 10\%$ measurement errors	A reliable initial O-D matrix from the initial O-D estimation algorithm

#### 5.4.2 Evaluations and Comparisons of Estimation Results

To evaluate the model performance, the average absolute error (AAE) statistics for the estimated O-D proportions serves as the evaluation criterion, which has been defined in Chapter 4 as:

$$AAE = \frac{\sum_{k=0}^K |\hat{b}_{ij}(k) - \bar{b}_{ij}(k)|}{K} \quad (4.4)$$

The comparisons of the estimation results from the proposed two sets of experiments (see Tables 5.1 and 5.2). The evaluation information includes the average AAE statistics of estimation results for each scenario, and the percentage of improvement modeled by Approach-2 (i.e., the interval-based approach) over Approach-1 (i.e., the scalar-based approach). Examples of graphical results for the time-dependent O-D demands are also depicted in this section.

**Experiment-1:** Comparing the performance of Approach-1 and Approach-2 under different levels of measurement errors in input flow data

Tables 5.3 and 5.4 present the TAAE statistics of the O-D proportions estimated with Approach-1 (i.e., scalar-based approach) and Approach-2 (i.e., interval-based approach), respectively, under different levels of input measurement errors. Table 5.5 summarizes the average AAE of the estimated O-D proportions in each scenario, and the improvement by Approach-2 over Approach-1, including the average, maximal and minimal values of the AAE statistics.

As shown in Tables 5.3 and 5.4, the estimation errors with Approach-1 increase from 0.0772 to 0.1091 (more than 40%) due to the increase in the measurement errors in the input flow data. In contrast, while the estimation results with Approach-2 are not sensitive to the input measurement errors in these five scenarios. For example, as shown in Table 5.4, the average estimation error is 0.0748 in the scenario of 30% measurement errors, compared with 0.0720 in the scenario of 5% measurement errors. As shown in Table 5.5, in the scenarios of measurement errors less than 10 % (i.e., Scenario-1 and Scenario-2), the improvement with Approach-2 (i.e., the interval-based approach) is not significant. However, when the measurement errors are more than 10 % (i.e., Scenario-3 to Scenario-5), Approach-2 has yielded substantially better estimation results, ranging from 27.58% to 45.88% improvement.

Overall, the interval-based approach substantially improves the estimation results, especially when the measurement errors in the flow data are significant. To further illustrate the estimation results, this study has depicted the example graphical results of

the estimated time-dependent O-D proportions in Figures 5.2 and 5.3. It is obvious that, over the entire computation period, the estimation results from the interval-based approach are closer to the true O-D values than those by Approach-1.

**Table 5.3. AAE Statistics of Estimation Results with Approach-1 – Experiment-1**

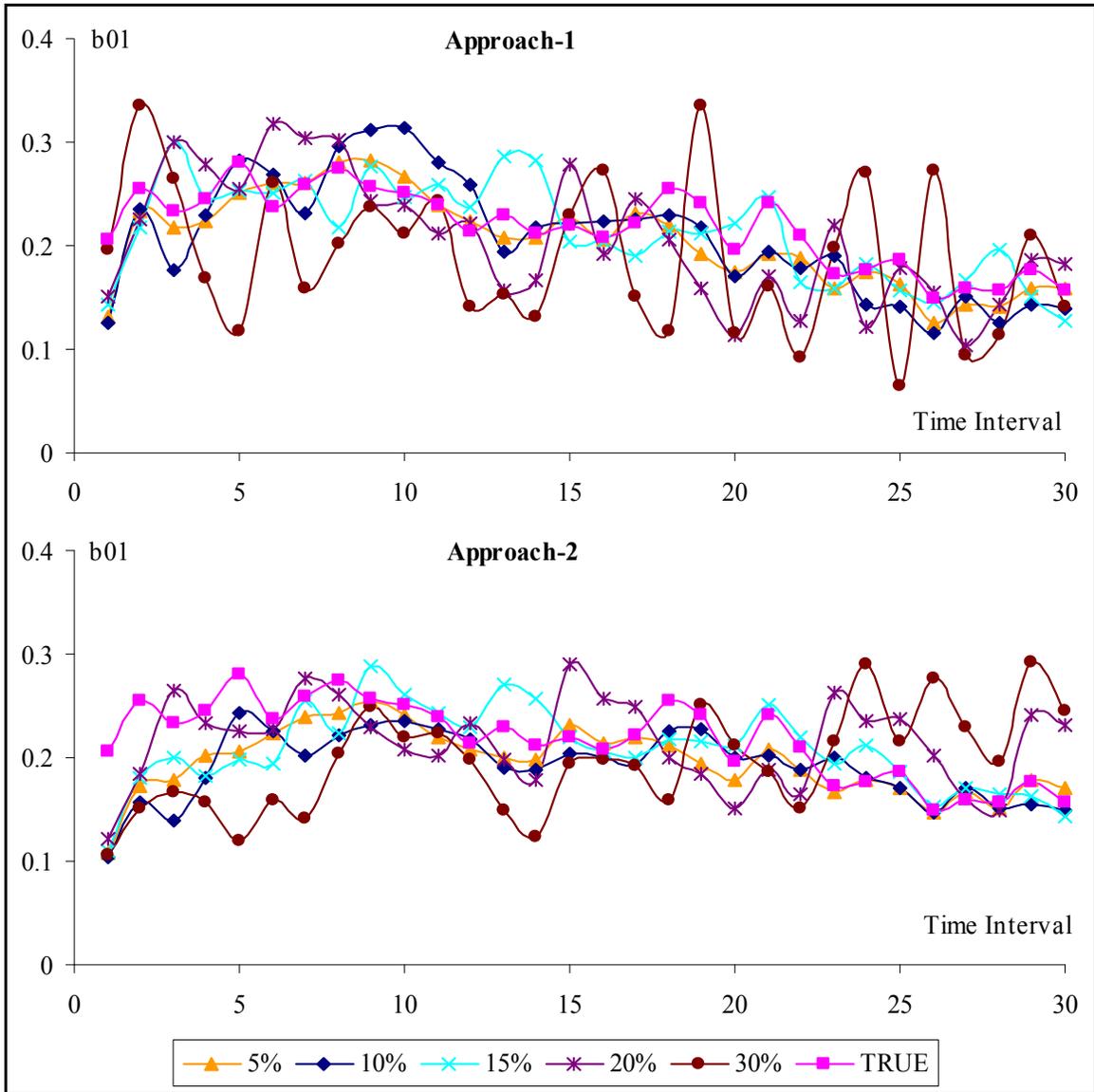
AAE	<b>b01</b>	<b>b02</b>	<b>b03</b>	<b>b04</b>	<b>b05</b>	<b>b06</b>	<b>b07</b>	<b>b08</b>	<b>b12</b>
5%	0.0193	0.0237	0.0640	0.0367	0.0240	0.0356	0.0373	0.0430	0.0307
10%	0.0298	0.0287	0.0402	0.0297	0.0396	0.0335	0.0266	0.0532	0.0235
15%	0.0275	0.0431	0.0497	0.0358	0.0317	0.0388	0.0302	0.0497	0.0479
20%	0.0408	0.0317	0.0506	0.0256	0.0364	0.0338	0.0456	0.0632	0.0691
30%	0.0674	0.0433	0.0666	0.0671	0.0425	0.0495	0.0505	0.0672	0.0860
AAE	<b>b13</b>	<b>b14</b>	<b>b15</b>	<b>b16</b>	<b>b17</b>	<b>b18</b>	<b>b23</b>	<b>b24</b>	<b>b25</b>
5%	0.0840	0.0700	0.0556	0.0751	0.0930	0.1623	0.1650	0.0838	0.0602
10%	0.0752	0.0586	0.0730	0.0843	0.0907	0.1684	0.1334	0.0675	0.0402
15%	0.1320	0.1071	0.0616	0.0708	0.0710	0.1217	0.1186	0.0912	0.0502
20%	0.0994	0.0902	0.0708	0.0947	0.0507	0.1509	0.1481	0.0615	0.0502
30%	0.0614	0.1048	0.0741	0.0761	0.0701	0.0833	0.1459	0.0908	0.0641
AAE	<b>b26</b>	<b>b27</b>	<b>b28</b>	<b>b34</b>	<b>b35</b>	<b>b36</b>	<b>b37</b>	<b>b38</b>	<b>b45</b>
5%	0.0996	0.0769	0.2535	0.0692	0.0533	0.0485	0.1029	0.0442	0.0659
10%	0.0368	0.0610	0.2420	0.0613	0.0700	0.0360	0.0766	0.0717	0.0548
15%	0.0657	0.0836	0.2596	0.0739	0.0557	0.0593	0.0732	0.0959	0.0652
20%	0.0575	0.0812	0.2220	0.1067	0.0747	0.0896	0.0762	0.0795	0.0722
30%	0.0991	0.0737	0.3655	0.1205	0.0849	0.0530	0.0658	0.1801	0.1897
AAE	<b>b46</b>	<b>b47</b>	<b>b48</b>	<b>b56</b>	<b>b57</b>	<b>b58</b>	<b>b67</b>	<b>b68</b>	<b>Avg.</b>
5%	0.0934	0.0713	0.1266	0.0608	0.0826	0.1125	0.2332	0.2332	0.0831
10%	0.0942	0.0659	0.0784	0.0774	0.0857	0.1205	0.2248	0.2248	0.0772
15%	0.0641	0.1110	0.1262	0.0426	0.0898	0.1123	0.2698	0.2698	0.0860
20%	0.0707	0.0859	0.1283	0.0627	0.0880	0.0988	0.2659	0.2659	0.0872
30%	0.1189	0.0975	0.2823	0.1730	0.0839	0.1852	0.2223	0.2223	0.1091

**Table 5.4. AAE Statistics of Estimation Results with Approach-2 – Experiment-1**

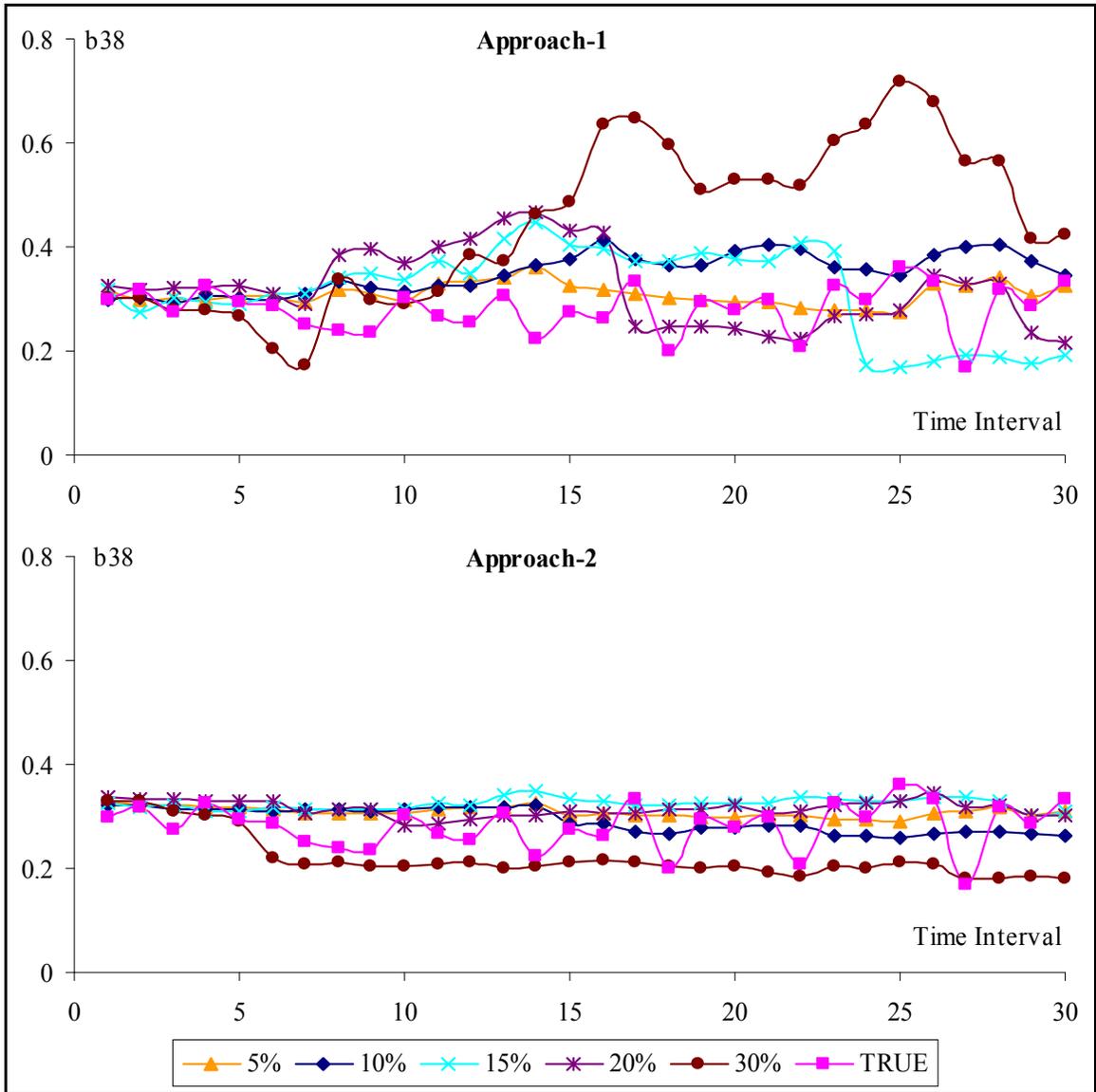
<b>AAE</b>	<b>b01</b>	<b>b02</b>	<b>b03</b>	<b>b04</b>	<b>b05</b>	<b>b06</b>	<b>b07</b>	<b>b08</b>	<b>b12</b>
5%	0.0246	0.0257	0.0484	0.0201	0.0171	0.0403	0.0184	0.0551	0.0227
10%	0.0298	0.0356	0.0295	0.0226	0.0256	0.0467	0.0211	0.0664	0.0377
15%	0.0277	0.0358	0.0321	0.0217	0.0269	0.0212	0.0220	0.0383	0.0406
20%	0.0430	0.0239	0.0372	0.0269	0.0292	0.0255	0.0329	0.0499	0.0322
30%	0.0651	0.0421	0.0336	0.0443	0.0263	0.0259	0.0348	0.0343	0.0500
<b>AAE</b>	<b>b13</b>	<b>b14</b>	<b>b15</b>	<b>b16</b>	<b>b17</b>	<b>b18</b>	<b>b23</b>	<b>b24</b>	<b>b25</b>
5%	0.0804	0.0465	0.0500	0.0609	0.0484	0.1347	0.1291	0.0483	0.0421
10%	0.0661	0.0470	0.0620	0.0557	0.0452	0.1098	0.0884	0.0504	0.0367
15%	0.0869	0.0701	0.0655	0.0680	0.0567	0.0924	0.0565	0.0617	0.0258
20%	0.0788	0.0628	0.0655	0.0594	0.0422	0.0982	0.1138	0.0629	0.0266
30%	0.0570	0.0813	0.0577	0.1017	0.0529	0.0863	0.0526	0.0451	0.0328
<b>AAE</b>	<b>b26</b>	<b>b27</b>	<b>b28</b>	<b>b34</b>	<b>b35</b>	<b>b36</b>	<b>b37</b>	<b>b38</b>	<b>b45</b>
5%	0.0690	0.0523	0.1973	0.0589	0.0655	0.0630	0.0816	0.0399	0.0549
10%	0.0903	0.0553	0.1574	0.0620	0.0550	0.0687	0.0797	0.0448	0.0641
15%	0.0349	0.0597	0.1761	0.0613	0.0647	0.0386	0.0633	0.0479	0.0578
20%	0.0370	0.0573	0.1685	0.0500	0.0584	0.0588	0.0656	0.0410	0.0401
30%	0.0672	0.0557	0.1325	0.0819	0.0402	0.0345	0.0616	0.0680	0.1112
<b>AAE</b>	<b>b46</b>	<b>b47</b>	<b>b48</b>	<b>b56</b>	<b>b57</b>	<b>b58</b>	<b>b67</b>	<b>b68</b>	<b>Avg.</b>
5%	0.1167	0.0520	0.1295	0.0662	0.0748	0.1141	0.2226	0.2226	0.0720
10%	0.1177	0.0529	0.1414	0.0786	0.0728	0.1190	0.2170	0.2170	0.0714
15%	0.0537	0.0751	0.1199	0.0666	0.0778	0.1068	0.2365	0.2365	0.0674
20%	0.0804	0.0520	0.0733	0.0842	0.0804	0.1150	0.2244	0.2244	0.0673
30%	0.1413	0.0836	0.2095	0.0966	0.0725	0.1194	0.1967	0.1967	0.0748

**Table 5.5. The Overall Statistical Results for the Five Scenarios – Experiment-1**

AAE		Average	Max.	Min.
Scenatio-1: With 5% Measurement Errors in Flow Data	Approach-1	0.0831	0.2535	0.0193
	Approach -2	0.0720	0.2226	0.0171
	Improvement	<b>15.32%</b>	<b>13.87%</b>	<b>12.85%</b>
Scenatio-2: With 10% Measurement Errors in Flow Data	Approach-1	0.0772	0.2420	0.0235
	Approach -2	0.0714	0.2170	0.0211
	Improvement	<b>8.09%</b>	<b>11.50%</b>	<b>11.32%</b>
Scenatio-3: With 15% Measurement Errors in Flow Data	Approach-1	0.0860	0.2698	0.0275
	Approach -2	0.0674	0.2365	0.0212
	Improvement	<b>27.58%</b>	<b>14.05%</b>	<b>29.97%</b>
Scenatio-4: With 20% Measurement Errors in Flow Data	Approach-1	0.0872	0.2659	0.0256
	Approach -2	0.0673	0.2244	0.0239
	Improvement	<b>29.64%</b>	<b>18.49%</b>	<b>7.21%</b>
Scenatio-5: With 30% Measurement Errors in Flow Data	Approach-1	0.1091	0.3655	0.0425
	Approach -2	0.0748	0.2095	0.0259
	Improvement	<b>45.88%</b>	<b>74.46%</b>	<b>64.18%</b>



**Figure 5.2. The Graphical Illustration of Estimation Results (b01) – Experiment-1**



**Figure 5.3. The Graphical Illustration of Estimation Results (b38) – Experiment-1**

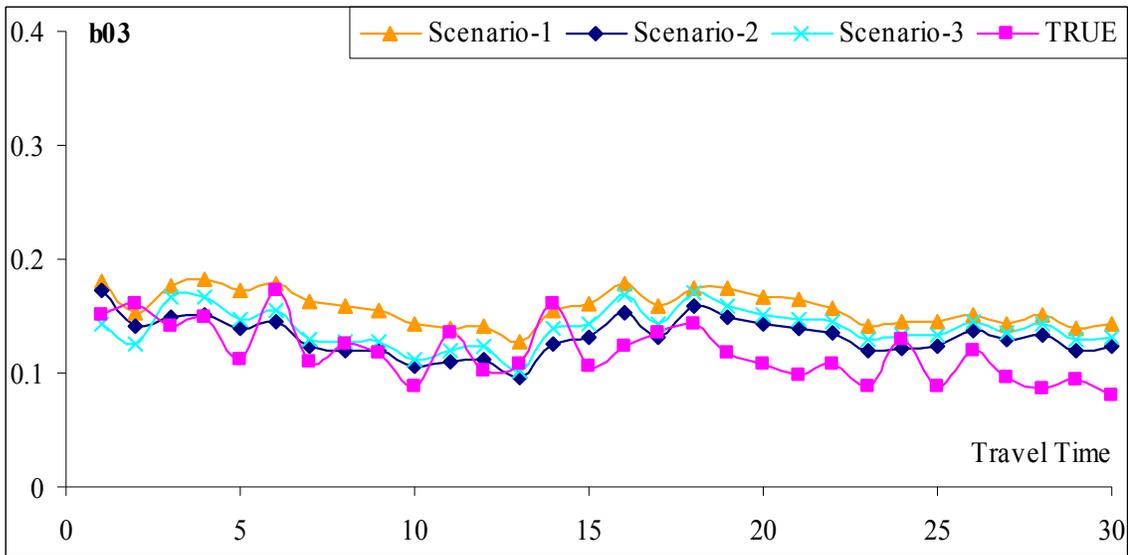
**Experiment-2:** The performance evaluation of the interval-based approach with the initial O-D estimation algorithm

This experiment is mainly designed to evaluate the compound effect of the interval-based model and the initial O-D estimation algorithm on the estimation accuracy. To do so, this study has compared the estimation results with the interval-based approach (Approach-2) under different data quality levels of the initial O-D set (see Table 5.2).

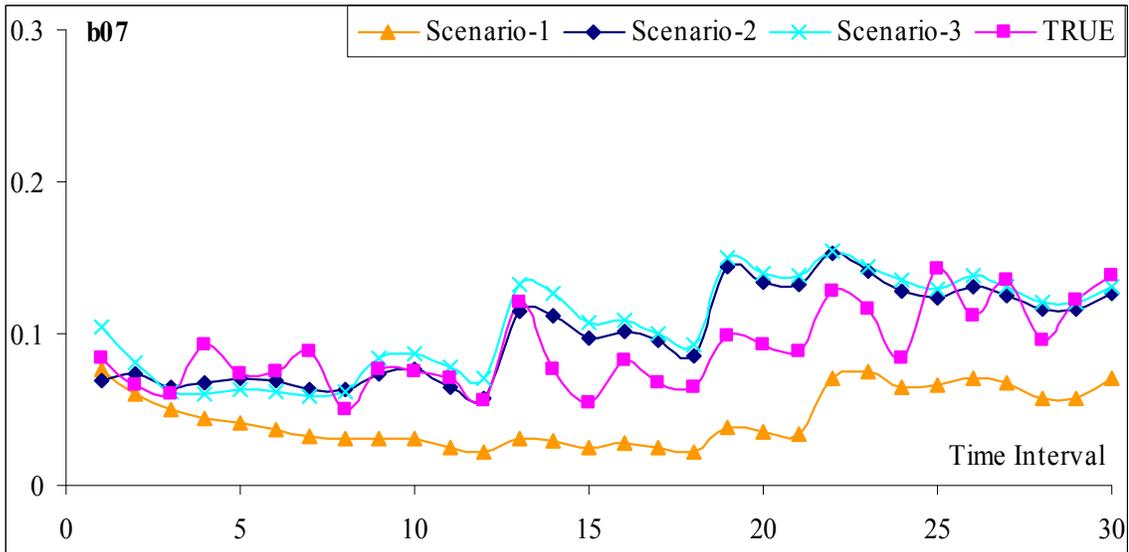
Table 5.6 presents the overall statistical results for those three scenarios and Figures 5.4 and 5.5 illustrate the examples of the time-dependent O-D estimation results. As shown in Table 5.6, the statistical results reveal that Scenario-2 yields the best results scenario with the average AAE value of 0.0474, and Scenario-1 produces the worst estimates among the three scenarios as it uses the uniformly distributed initial O-D set as an input. In Scenario-3, the initial O-D set is generated with the initial O-D estimation algorithm. Hence, as expected, its estimation quality lies between Scenario-1 and Scenario-2. Figures 5.4 to 5.7 provide some further numerical evidences of such a relation.

**Table 5.6. The Summary of Estimation Errors from All Scenarios – Experiment-2**

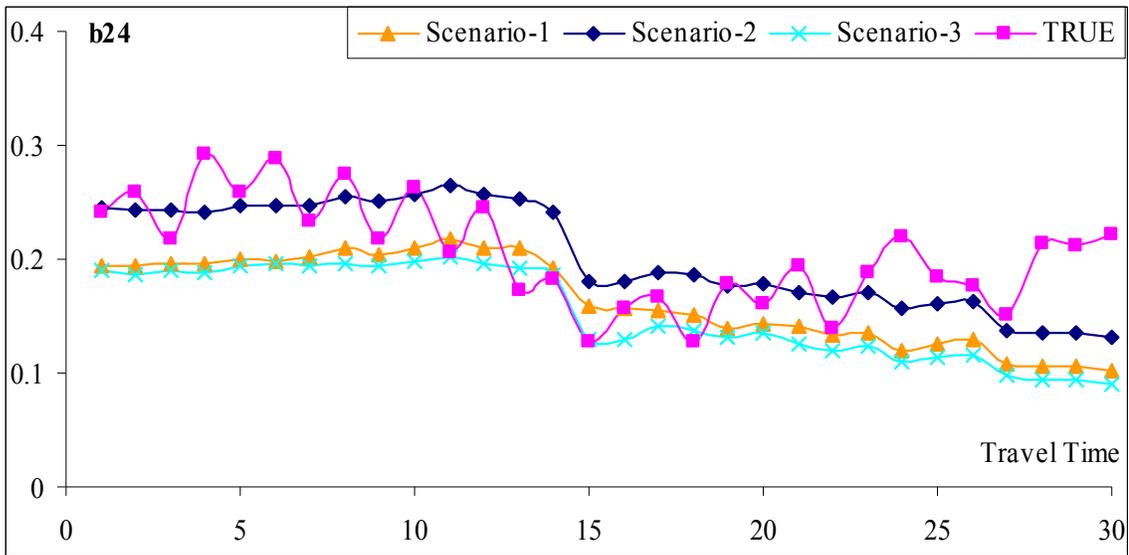
AAE	<b>b01</b>	<b>b02</b>	<b>b03</b>	<b>b04</b>	<b>b05</b>	<b>b06</b>	<b>b07</b>	<b>b08</b>	<b>b12</b>
Scenario-1	0.0147	0.0290	0.0396	0.0366	0.0295	0.0549	0.0446	0.0584	0.0455
Scenario-2	0.0161	0.0307	0.0228	0.0197	0.0361	0.0378	0.0191	0.0117	0.0319
Scenario-3	0.0243	0.0347	0.0281	0.0214	0.0261	0.0429	0.0236	0.0513	0.0379
AAE	<b>b13</b>	<b>b14</b>	<b>b15</b>	<b>b16</b>	<b>b17</b>	<b>b18</b>	<b>b23</b>	<b>b24</b>	<b>b25</b>
Scenario-1	0.0514	0.0527	0.0371	0.0408	0.0502	0.0328	0.0849	0.0488	0.0396
Scenario-2	0.0591	0.0793	0.0392	0.0515	0.0716	0.0377	0.0941	0.0351	0.0351
Scenario-3	0.0579	0.0512	0.0602	0.0565	0.0453	0.0912	0.0897	0.0551	0.0361
AAE	<b>b26</b>	<b>b27</b>	<b>b28</b>	<b>b34</b>	<b>b35</b>	<b>b36</b>	<b>b37</b>	<b>b38</b>	<b>b45</b>
Scenario-1	0.0751	0.0354	0.0373	0.0501	0.0451	0.0827	0.0349	0.0828	0.1101
Scenario-3	0.1073	0.0311	0.0324	0.0510	0.0633	0.0751	0.0451	0.0403	0.0224
Scenario-3	0.0939	0.0598	0.1761	0.0655	0.0486	0.0671	0.0821	0.0434	0.0861
AAE	<b>b46</b>	<b>b47</b>	<b>b48</b>	<b>b56</b>	<b>b57</b>	<b>b58</b>	<b>b67</b>	<b>b68</b>	<b>Avg.</b>
Scenario-1	0.1439	0.0665	0.3104	0.0589	0.1012	0.1254	0.2430	0.2430	<b>0.0733</b>
Scenario-2	0.0296	0.0543	0.0716	0.0648	0.0673	0.0934	0.0646	0.0646	<b>0.0474</b>
Scenario-3	0.1129	0.0554	0.1552	0.0817	0.0742	0.1246	0.2052	0.2052	<b>0.0714</b>



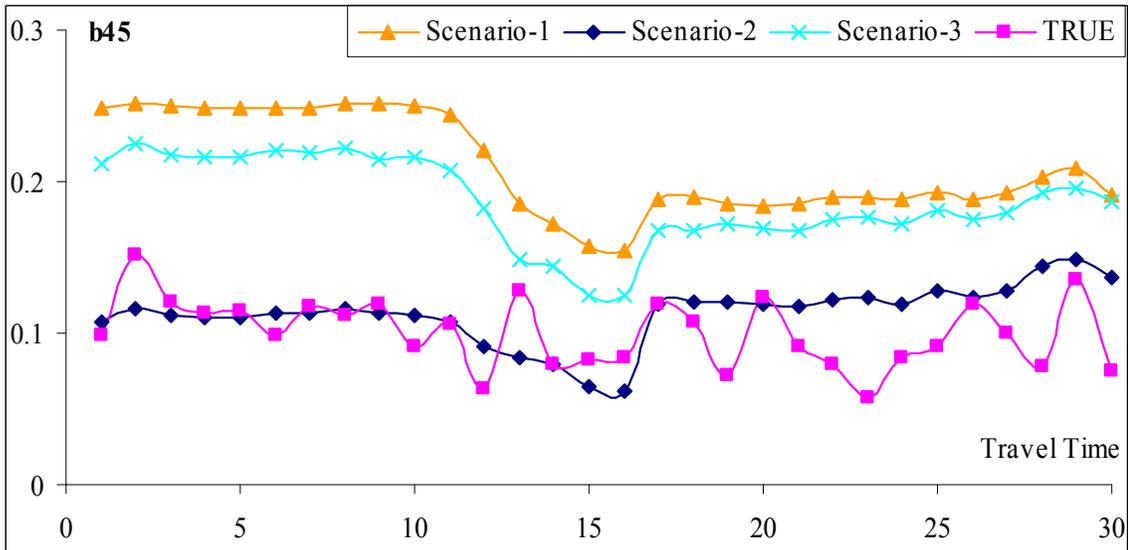
**Figure 5.4. The Graphical Illustration of Estimation Results (b03) – Experiment-2**



**Figure 5.5. The Graphical Illustration of Estimation Results (b07) – Experiment-2**



**Figure 5.6. The Graphical Illustration of Estimation Results (b24) – Experiment-2**



**Figure 5.7. The Graphical Illustration of Estimation Results (b45) – Experiment-2**

## 5.5 CLOSURE

This chapter has developed an interval-based model and its solution algorithm for contending with the inevitable measurement errors embedded in most traffic data due to hardware deficiencies. The interval-based model for dynamic O-D estimation has captured the variance of all required input data in the formulations, and allows the users to set the reasonable upper and lower boundaries for each available input. To solve the proposed interval-based model, the study has further presented a recursive solution algorithm developed with the Interval Kalman filter.

This chapter has also presented the performance evaluation results of the proposed interval-based model using the I-95 freeway corridor as Chapter 3. The results from the first experiment have revealed the fact that the time-dependent O-D estimates from the interval-based model are relatively insensitive to the input measurement errors. A second experiment has also been contented to evaluate the compound effect of the interval-based model and the initial O-D estimation algorithm proposed in Chapter 4. The extensive numerical results presented in this chapter have indicated the potential for an integrated application of both algorithms in the estimation of large dynamic freeway O-D distributions.

## **CHAPTER 6 CONCLUSIONS AND FUTURE RESEARCH**

### **6.1 SUMMARY OF RESEARCH ACCOMPLISHMENTS**

This research has mainly investigated potential technical issues associated with the estimation of dynamic O-D patterns for a large freeway corridor. Based on the deficiencies of existing models and limitations of data quality in real-world traffic surveillance systems, this study has focused on the following three aspects:

- Reformulating state-of-the-art formulations for dynamic O-D estimation with a travel time distribution function that can best take advantage of observable dynamic interrelations between system components with a reduced number of unknown variables;
- Constructing a recursive algorithm for approximating the initial O-D set, which is generally not available but is an essential start point for any algorithm to execute the dynamic O-D estimation; and
- Developing an interval-based model and solution algorithm to account for the measurement errors due to traffic sensor deficiencies or the available data quality. The primary accomplishments of this study include:

### Formulating a Dynamic O-D Estimation Model with an Embedded Travel Time Distribution Function

Most existing freeway dynamic O-D estimation models suffer from two main limitations: insufficient observations and unrealistic assumptions. Chapter 3 has presented a model and its solution algorithm for estimating the time-dependent O-D demands in a large freeway corridor to overcome these deficiencies. The proposed model can reflect the speed variance among vehicles having the same departure time, origin and destination with an embedded travel time distribution function, and consequently provide a set of formulations with fewer unknown variables. Its solution algorithm developed with the sequential extended Kalman filtering logic has proven to be sufficiently efficient and reliable for use in practice.

### Developing an Initial O-D Set Estimation Algorithm for Estimating a Reliable Initial O-D Set for a Large Freeway Network

This algorithm is developed for addressing the commonly encountered issue in real-world applications, that is, to compute the initial O-D set, which is often not available but an essential input of most existing dynamic models for estimating subsequent time-dependent O-D distributions. The key concept of the proposed algorithm is to decompose a large freeway corridor into several small segments so as to reduce the number of unknown variables in each freeway segment. The algorithm can capture the relations between the estimated initial O-D proportions and the flow information due to the iterative use of the input information observed during the first time interval. Furthermore, the replacement of the unknown variables estimated from the previous

segment can also reduce the number of unknown variables and improve the estimation accuracy and efficiency.

#### Proposing an Interval-based Algorithm for Estimating Dynamic O-D Patterns with Variance in the Available Information

This study has developed an interval-based model to tackle the input measurement errors due to the deficiencies of surveillance systems. Since the accurate traffic flows are likely to lie within intervals rather than constant values, the proposed model has formulated each input variable with a lower and upper bounds. To accommodate the unique modeling features, the study has further derived an interval-based solution algorithm with the Interval Kalman filter method. This set of interval-based model and solution algorithm can yield the estimation results that are less sensitive to the measurement errors embedded in input data.

#### Conducting Extensive Model Sensitivity Analyses and Performance Evaluation with a Real-World Freeway Corridor

This study has presented the sensitivity analyses of the dynamic O-D estimation model proposed in Chapter 3 with respect to the key system factors and parameters, including the initial O-D set and travel time variance. The numerical analyses have revealed that the proposed dynamic O-D estimation model is sufficiently effective for tackling the travel time variance with a certain range.

This study has also used the I-95 freeway corridor between Baltimore and DC beltways to evaluate the performance of the proposed models and algorithms and to demonstrate its applicability. The results of extensive experiments have indicated that the

dynamic O-D estimation model proposed in Chapter 3 can yield reasonable time-dependent O-D estimates with a reliable initial O-D set. The numerical experimental results have also evidenced that it is necessary to employ the interval-based model and algorithm if the available input data set is subjected to some degree of measurement errors.

## **6.2 FUTURE RESEARCH**

In spite of the contributions made in this research, many key issues remain to be investigated so as to improve the applicability of the time-dependent O-D estimation model. Examples of these essential research issues are summarized below.

- **Reliability Analysis for the Dynamic O-D Estimation Model**

One of the critical issues remains to be investigated in the area of time-dependent O-D estimation is to estimate the interrelation between model reliability and the available information. Ideally, one should have some effective statistics or procedures to evaluate if the available detector information is sufficient to render a target level of estimation quality. Such information will serve as the basis to evaluate the number of detectors needed to generate the data set for a reliable O-D estimation.

The proposed statistics for reliability analysis may offer the potential guidelines for determining the detector locations and spacing under the resource constraints. For instance, one shall be able to use the developed statistics to best approximate the number

of detectors needed and the distribution of their locations with a target level of reliability and the intended estimation model.

- The Dynamic O-D Estimation with Only Partial Information

As mentioned in the previous chapters, most existing approaches for estimating dynamic O-D matrices require a set of initial/prior O-D and ramp/mainline volumes as model inputs. Depending on the available information and the network structure, the O-D patterns estimated with those approaches may result in a large variance, and thus insufficient reliability for use in practice. Besides, many of those essential data, such as on/off ramp volumes, may not be available in most real-world freeway networks.

Note that the detectors in most freeway systems are mostly deployed on the mainline segments, not at the on-ramps and off-ramps. However, nearly all existing O-D estimation models assume that a complete set of on-ramp and off-ramp volume information is always available. To deal with these commonly encountered constraints, one may focus on developing a model to estimate the ramp volumes, based on the available mainline detected volumes and historical travel time information if available.

## APPENDIX A. INTERVAL ARITHMETIC

Let  $[\underline{x}, \bar{x}]$  be an interval which is a closed and bounded subset in  $\mathbb{R} = (-\infty, \infty)$ . This interval computation shall obey the following properties (Chui and Chen, 1999):

- Equality:  $[\underline{x}_1, \bar{x}_1] = [\underline{x}_2, \bar{x}_2]$  if and only if  $\underline{x}_1 = \underline{x}_2$  and  $\bar{x}_1 = \bar{x}_2$ .
- Inequality:  $[\underline{x}_1, \bar{x}_1] < [\underline{x}_2, \bar{x}_2]$  if and only if  $\bar{x}_1 < \underline{x}_2$ , and  $[\underline{x}_1, \bar{x}_1] > [\underline{x}_2, \bar{x}_2]$  if and only if  $\underline{x}_1 > \bar{x}_2$ .
- Intersection:  $[\underline{x}_1, \bar{x}_1] \cap [\underline{x}_2, \bar{x}_2] = [\max\{\underline{x}_1, \underline{x}_2\}, \min\{\bar{x}_1, \bar{x}_2\}]$ .
- Union:  $[\underline{x}_1, \bar{x}_1] \cup [\underline{x}_2, \bar{x}_2] = [\min\{\underline{x}_1, \underline{x}_2\}, \max\{\bar{x}_1, \bar{x}_2\}]$ , only if  $[\underline{x}_1, \bar{x}_1] \cap [\underline{x}_2, \bar{x}_2] \neq \emptyset$ .
- Disjoint:  $[\underline{x}_1, \bar{x}_1] \cap [\underline{x}_2, \bar{x}_2] = \emptyset$ .
- Inclusion:  $[\underline{x}_1, \bar{x}_1] \subseteq [\underline{x}_2, \bar{x}_2]$  if and only if  $\underline{x}_2 \leq \underline{x}_1$  and  $\bar{x}_1 \leq \bar{x}_2$ .

In addition to the foregoing properties, the interval arithmetic operations are defined differently from a single value:

- Addition:  $[\underline{x}_1, \bar{x}_1] + [\underline{x}_2, \bar{x}_2] = [\underline{x}_1 + \underline{x}_2, \bar{x}_1 + \bar{x}_2]$
- Subtraction:  $[\underline{x}_1, \bar{x}_1] - [\underline{x}_2, \bar{x}_2] = [\underline{x}_1 - \bar{x}_2, \bar{x}_1 - \underline{x}_2]$
- Reciprocal Operation: If  $0 \notin [\underline{x}, \bar{x}]$ , then  $[\underline{x}, \bar{x}]^{-1} = [1/\bar{x}, 1/\underline{x}]$ . If  $0 \in [\underline{x}, \bar{x}]$ , then  $[\underline{x}, \bar{x}]^{-1}$  is undefined.
- Multiplication:  $[\underline{x}_1, \bar{x}_1] \cdot [\underline{x}_2, \bar{x}_2] = [\underline{y}, \bar{y}]$   
 where  $\underline{y} = \min\{\underline{x}_1 \underline{x}_2, \underline{x}_1 \bar{x}_2, \bar{x}_1 \underline{x}_2, \bar{x}_1 \bar{x}_2\}$  and  $\bar{y} = \max\{\underline{x}_1 \underline{x}_2, \underline{x}_1 \bar{x}_2, \bar{x}_1 \underline{x}_2, \bar{x}_1 \bar{x}_2\}$
- Division: if  $0 \notin [\underline{x}_2, \bar{x}_2]$ , then  $[\underline{x}_1, \bar{x}_1] / [\underline{x}_2, \bar{x}_2] = [\underline{x}_1, \bar{x}_1] \cdot [\underline{x}_2, \bar{x}_2]^{-1}$ . Otherwise, it is undefined.

## REFERENCES

- Anderson, B.D.O., and Moore, J.B., *Optimal Filtering*, Prentice Hall, 1979.
- Ashok, K., and Ben-Akiva, M., "Alternative Approaches for Real-Time Estimation and Prediction of Time-Dependent Origin-Destination Flows," *Transportation Science*, Vol. 34, No. 1, pp. 21-36, 2000.
- Ashok, K., and Ben-Akiva, M.E., "Dynamic Origin-Destination Matrix Estimation for Real-Time Traffic Management Systems," *Transportation and Traffic Theory*, pp. 465-484, 1993.
- Ashok, K., and Ben-Akiva, M.E., "Estimation and Prediction of Time-dependent Origin-Destination Flows with a Stochastic Mapping to Path Flows and Link Flows," *Transportation Science*, Vol. 36, pp. 184-198, 2002.
- Ashok, K., *Estimation and Prediction of Time-Dependent Origin-Destination Flows*, Ph.D. Dissertation, MIT, 1996.
- Barbour R., and Fricker, J. D., "Estimating an Origin-Destination Table Using a method based on Shortest Augment Paths," *Transportation Research B*, Vol. 28, No. 2, pp. 77-89, 1994.
- Bell, M. G. H., "The Estimation of an Origin-Destination Matrix from Traffic Counts," *Transportation Science*, Vol. 17, No. 2, pp. 198-217, 1983.
- Bell, M.G.H., "The Estimation of Origin-Destination Matrices by Constrained Generalized Least Squares," *Transportation Research B*, Vol. 25, No. 1, pp. 13-22, 1991a.

- Bell, M.G.H., "The Real-Time Estimation of Origin-Destination Flows in the Presence of Platoon Dispersion," *Transportation Research B*, Vol. 25, No.2/3, pp. 115-125, 1991b.
- Cascetta, E. "Estimation of Trip Matrices from Traffic Counts and Survey Data: A Generalized Least Squares Estimator," *Transportation Research B*, Vol. 18 No. 4/5, pp. 289-299, 1984.
- Cascetta, E. Inaudi, D., and Marguis, G., "Dynamic Estimators of Origin-Destination Matrices using Traffic Counts," *Transportation Science*, Vol. 27, pp. 363-373, 1993.
- Cascetta, E., and Nguyen, S., "A Unified Framework for Estimating or Updating Origin-Destination Matrices from Traffic Counts," *Transportation Research B*, Vol. 22, No. 6, pp. 437-455, 1988.
- Chang, G-L, and Tao, X., "An Integrated Model for Estimating Time-Varying Network Origin-Destination Distributions," *Transportation Research A*, Vol. 33, No. 5, pp. 381-399, 1999.
- Chang, G-L, and Tao, X., "Estimation of Dynamic O-Ds for Urban Networks," *Transportation and Traffic Flow Theory*, Vol. 13, pp. 1-20, 1996.
- Chang, G-L, and Wu, J., "Recursive Estimations of Time-Varying Origin-Destination Flows from Traffic Counts in Freeway Corridors," *Transportation Research B*, Vol. 28, No. 2, pp. 141-160, 1994.
- Chen, G., Wang, J., and Ghieh, L. S., "Interval Kalman Filtering," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 33, No. 1, pp. 250-258, 1997.

- Chui, C. K, and Chen, G., *Kalman Filtering with Real-time Applications*, Springer-Verlag, New York, 1999.
- Cremer, M. “Determining the Time Dependent Trip Distribution in a Complex Intersections for Traffic Responsive Control,” *Preprint of the 4th International IFAC/IFIP/IFORS Conference on Control in Transportation Systems*, Baden-Baden, Germany, 1983.
- Cremer, M., and Keller, H., “A New Class of Dynamic Methods for the Identification of Origin-Destination Flows,” *Transportation Research B*, Vol. 21, No. 2, pp. 117-132, 1987.
- Cremer, M., and Keller, H., “A Systems Dynamics Approach to the Estimation of Entry and Exit O-D Flows,” *the Proceeding of the 9th International Symposium on Transportation and Traffic Theory*, pp. 431-450, 1984.
- Cremer, M., and Keller, H., “Dynamic Identification of O-D flows from Traffic Counts at Complex Intersections,” *the Proceeding of the 8th International Symposium on Transportation and Traffic Theory*, Toronto University, Canada, 1981.
- Dempter, A. P., Laird, N. M., and Dubin, D. B., “Maximum Likelihood from Incomplete Data via the EM Algorithm,” *Journal of the Royal Statistical Society*, Vol. 39, pp. 1-38, 1977.
- Dixon, M. P., and Rilett, L. R., “Real-Time OD Estimation Using Automatic Vehicle Identification and Traffic Count Data,” *Computer–Aided Civil and Infrastructure Engineering*, Vol. 17, No. 1, pp. 7-21, 2002.

- Grace, M. J., and Plotts, R. B., "A Theory of the Diffusion of Traffic Platoons," *Operations Research*, Vol. 12, pp. 255-275, 1964.
- Hazelton, M., "Estimation of Origin-Destination Matrices from Link Flows on Uncongested Networks," *Transportation Research B*, Vol. 34, No. 7, pp. 549-566, 2000.
- Kaman, R.E., "A New Approach to Linear Filtering and Prediction Problem," *Journal of Basic Engineering (ASME)*, Vol. 82, pp. 35-45, 1960.
- Keller, H., and Ploss, G. "Real-time Identification of O-D Network flow from Counts for Urban Traffic Control," *Transportation and Traffic Theory*, pp. 267-284, 1987.
- Kessaci, A., Farges, J., Henry, J., "On-line Estimation of Turning Movements and Saturation Flows in PRODYN," *the Proceedings of the 6th IFAC/IFIP/IFORS Conference on Control in Transportation*, Paris, France, 1989.
- Lin, P. W., and Chang, G-L, "A Robust Model for Estimating Freeway Dynamic Origin-Destination Matrix," Accepted in *Transportation Research Record*, 2005.
- Lo, H. P., and Chan, C. P., "Simultaneous Estimation of an Origin-Destination Matrix and Link Choice Proportions Using Traffic Counts," *Transportation Research A*, Vol. 37, No. 9, pp. 771-788, 2003.
- Maher, M.J., "Inferences on Trip Matrices from Observations on Link Volumes: A Bayesian Statistical Approach," *Transportation Research B*, Vol. 17, No. 6, pp. 435-447, 1983.

- Nihan, N. L., and Hamed, M. M., "Fixed-Point Approach to Estimating Freeway Origin-Destination Matrices and the Effect of Erroneous Data on Estimate Precision," *Transportation Science*, Vol. 23, pp. 77-90, 1992.
- Nihan, N.L., and Davis, G. A., "Application of Prediction-error Minimization and Maximum Likelihood to Estimate Intersection O-D Matrices from Traffic Counts," *Transportation Science*, Vol. 23, pp. 77-90, 1989.
- Nihan, N.L., and Davis, G. A., "Recursive Estimation of Origin-Destination Matrices from Input/Output Counts," *Transportation Research B*, Vol. 21, No. 2, pp. 149-163, 1987.
- Okutani, I., "The Kalman Filtering Approaches in Some Transportation and Traffic Problems," *Transportations and Traffic Theory*, 1987.
- Okutani, I., and Stephanedes, Y. J., "Dynamic Prediction of Traffic Volume Through Kalman Filtering Theory," *Transportation Research B*, Vol. 18, No., 1, pp.1-11, 1984.
- Pacey, G. M., "The progress of a bunch vehicles released from a traffic signal," *Note for Road Research Laboratory*, No. RN/2665/GMP, Road Research Laboratory, Growthorne, UK, 1956.
- Peeta, S., and Ziliaskopoulos, A., "Foundations of Dynamic Traffic Assignment: the Past, the Present and the Future," *Network and Spatial Economics*, Vol. 1:2, pp. 201-230, 2001.

- Ploss, G., and Keller, H., "Dynamic Estimation of Origin and Destination Flows from Traffic Counts in Network," *the Proceedings of the International Conference on Transportation System Studies*, Delhi, 1986.
- Roberson, D. I., "TRANSYT: A Traffic Network study Tool," *Road Research Laboratory Report*, No. LR253, Crowthorne, England, 1969.
- Roberson D. I., "TRANSYT Method for Area Traffic Control," *Traffic Engineering and Control*, Vol. 11, pp. 276-281, 1969.
- Seddon, P. A., "The Prediction of Platoon Dispersion in the Combination Methods of Linking Traffic Signals," *Transportation Research*, Vol. 6, pp. 125-130, 1972.
- Sherali, H. D., and Park, T., "Estimation of Dynamic Origin-Destination Trip Tables for a General Network," *Transportation Research B*, Vol. 35, No. 3, pp. 217-235, 2001.
- Sherali, H. D., Narayanan, A., and Sivanadan, R., "Estimation of Origin-Destination Trip Tables based on a Partial Set of Traffic Link Volumes," *Transportation Research B*, Vol. 37, No. 9, pp. 815-836, 2003.
- Tavana, H. and Mahmassani, H., "Estimation of Dynamic Origin-Destination Flows from Sensor Data Using Bi-level Optimization Method," *the Proceeding of the 80th Annual Meeting of the Transportation Research Board*, 2000.
- TSS-Transportation Simulation Systems, *GETRAM Operation Traffic Simulation Environment – TEDI Version 4.0 User Manual*, 2001.

- Van der Zijpp, N. J., and Hamerslag, R., "An Improved Kalman Filtering approach to Estimate Origin-Destination Matrices for freeway Corridors," *Transportation Research Record*, Vol. 1443, 1994.
- Van der Zijpp, N. J., *Dynamic Origin-Destination Matrix Estimation on Motorway Networks*, Ph.D. Thesis, Delft University of Technology, 1996.
- Van Zuylen, H. J., and Willumsen, L. G., "The Most Likely Trip Matrix Estimated from Traffic Counts," *Transportation Research B*, Vol. 14, pp. 281-293, 1980.
- Willumsen, K. G., "Estimating Time-Dependent Trip Matrices from Traffic Counts," *the Proceeding of the 9th International Symposium on Transportation and Traffic Theory*, pp. 397-411, 1984.
- Wu, J., "A Real-Time Origin-Destination Matrix Updating Algorithm for On-Line Application," *Transportation Research B*, Vol. 31, No. 5, pp. 381-396, 1997.
- Wu, J., and Chang, G-L, "Estimation of Time-Varying Origin-Destination Distributions with Dynamic Screenline Flows," *Transportation Research B*, Vol. 30, No. 4, pp. 277-290, 1996.
- Xu, W., and Chan, Y., "Estimating an Origin-Destination Matrix with fuzzy weights," *Transportation Planning and Technology*, Vol. 17, pp. 145-163, 1993.
- Yang, H., Iida, Y., and Sasaki, T., "An analysis of the reliability of an Origin-Destination Trip Matrix Estimated from Traffic Counts," *Transportation Research B*, Vol. 25, No.5, pp. 351-363, 1991.

- Yang, H., Meng, Q., and Bell, M.G.H., "Simultaneous Estimation of the Origin-Destination Matrices and Travel-Cost Coefficient for Congested Networks in a Stochastic Use Equilibrium," *Transportation Science*, Vol. 35, pp. 107-123, 2001.
- Yang, H., Sasaki, T., Iida, Y., and Asakura, Y., "Estimation of Origin-Destination Matrices from Link Traffic Counts on Congested Networks," *Transportation Research B*, Vol. 26, No. 6, pp. 417-434, 1992.
- Zhang, X, and Maher, M. J., "The Evaluation and Application of a Fully Disaggregate Method for Trip Matrix Estimation with Platoon Dispersion," *Transportation Research B*, Vol. 32, No. 4, pp. 216-276, 1998.
- Zhou, X., and Mahmassani, H. S., "Dynamic Origin-Destination Demand Estimation Using Automatic Vehicle Identification Data," *the Proceeding of the 84th Annual Meeting of the Transportation Research Board*, 2005.