Calibration of the Demand Simulator in a Dynamic Traffic Assignment System

by

Ramachandran Balakrishna

B.Tech in Civil Engineering (1999)
Indian Institute of Technology, Madras, India

Submitted to the Department of Civil and Environmental Engineering in partial fulfillment of the requirements for the degree of

Master of Science in Transportation Systems

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2002

Author ....
Department of Civil and Environmental Engineering
May 24, 2002

Certified by ....
Moshe E. Ben-Akiva
Edmund K. Turner Professor
Department of Civil and Environmental Engineering
Thesis Supervisor

Certified by ....
Haris N. Koutsopoulos
Operations Research Analyst
Volpe National Transportation Systems Center
Thesis Supervisor

Accepted by ....
Oral Buyukozturk
Chairman, Department Committee on Graduate Studies

Calibration of the Demand Simulator in a Dynamic Traffic Assignment System

by

Ramachandran Balakrishna

Submitted to the Department of Civil and Environmental Engineering
on May 24, 2002, in partial fulfillment of the
requirements for the degree of

Master of Science in Transportation Systems

Abstract

In this thesis, we present a methodology to jointly calibrate the O-D estimation and prediction and driver route choice models within a Dynamic Traffic Assignment (DTA) system using several days of traffic sensor data. The methodology for the calibration of the O-D estimation module is based on an existing framework adapted to suit the sensor data usually collected from traffic networks. The parameters to be calibrated include a database of time-varying historical O-D flows, variance-covariance matrices associated with measurement errors, a set of autoregressive matrices that capture the spatial and temporal inter-dependence of O-D flows, and the route choice model parameters. Issues involved in calibrating route choice models in the absence of disaggregate data are identified, and an iterative framework for jointly estimating the parameters of the O-D estimation and route choice models is proposed. The methodology is applied to a study network extracted from the Orange County region in California. The feasibility and robustness of the approach are indicated by promising results from validation tests.

Thesis Supervisor: Moshe E. Ben-Akiva
Title: Edmund K. Turner Professor
Department of Civil and Environmental Engineering

Thesis Supervisor: Haris N. Koutsopoulos
Title: Operations Research Analyst
Volpe National Transportation Systems Center

Acknowledgments

I would like to express my heartfelt gratitude to my thesis supervisors Professor Moshe Ben-Akiva and Doctor Haris Koutsopoulos for their constant support and guidance. I have learned so much from them. I am especially thankful to Haris, who selflessly committed his time and energy to helping me with the details of my research.

Several people have helped shape my thoughts through useful insights and enlightening discussions. I thank Dr Kalidas Ashok, Prof. Shlomo Bekhor, Prof. Michel Bierlaire, Prof. Denis Bolduc, Dr Jon Bottom, Dr John Bowman, fellow student Sarah Bush, Prof. Ennio Cascetta, Prof. Michael Florian, Prof. Eiji Hato, Prof. Ikki Kim, Prof. Frank Koppelman, Dr Scott Ramming, office-mate Tomer Toledo and Dr Joan Walker for their valuable inputs and suggestions.

I am grateful to my fellow people at the MIT ITS Program for their help, friendship and cameraderie at various stages. I thank Bruno and Didier for ensuring my smooth transition into the Lab; Manish, who put up with my early rising; Srini for his constant teasing and the cricket sessions; Tomer for lending a patient ear every time I felt like voicing my opinions; Patrick, who briefly relieved me of my sysad duties; Yosef and Constantinos, who helped me with much of my struggle with Unix; Angus, Akhil, Atul, Dan, Deepak, Jens, Kunal, Marge and Zhili for making the Lab a lively place to work in.

I also wish to thank my professors at IIT Madras: Prof. Kalyanaraman, who introduced me to the exciting world of research, and whose fine example and high standards I have always strived to emulate; Dr Dilip Veeraraghavan, whose discussions helped me find direction and balance in my life.

I thank the DTA project and the Federal Highway Administration for their financial support, Oak Ridge National Laboratories and the University of California, Irvine for their support and the data used in this research.

I cannot thank enough the administrative staff at both the Center for Transportation Studies and the Department of Civil and Environmental Engineering, especially Sydney Miller, Leanne Russell, Julie Bernardi, Cynthia Stewart and James Riefstahl, who cheerfully accommodated all my requests and questions.

My special thanks to the Pakodas, and apartment-mates Fr-OO-t, Paddu, Torpy, Umang, Manish, Sriram, Prahladh and Jeff, who were responsible for much of my life experiences; Lakshmi, Vinay, Arvind and PJ for their unending hospitality; Vinay for his invaluable violin lessons.

I express my indebtedness to my grandparents, parents and sister for their love, encouragement and inspiration, and for standing by me all through my life.

And finally, I give thanks to God for all my cherished memories, and for giving me the courage and strength to surmount the mountains, both big and small, that lay along the long road leading to this moment.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Physical and economic constraints are causing urban and suburban congestion relief solutions to move away from building more roads. Increasing attention is being focused on Advanced Traffic Management Systems (ATMS) and Intelligent Transportation Systems (ITS). Emerging Dynamic Traffic Management Systems (DTMS) attempt to optimize the utilization of existing system capacity by performing two basic functions. Firstly, such systems provide pre-trip and en-route information to drivers about anticipated network conditions for the duration of the proposed trip. In addition, the systems assist traffic control systems to adapt and adjust in real-time to dynamic traffic conditions.

A desirable feature of such systems is the ability to *predict* future traffic. Lack of knowledge about projected traffic conditions could render guidance or control initiatives irrelevant or outdated by the time they take effect. Dynamic Traffic Assignment (DTA) is a critical component of such traffic prediction systems.

A DTA system models complex interactions between supply and demand in a transportation network. While the supply component captures traffic dynamics through the movement of *packets* on the road network according to aggregate traffic relationships, the demand component estimates and predicts point-to-point network demand and models driver behavior with respect to choice of departure time, mode and route. The demand simulator also models drivers' response to information. The DTA system is designed to reside within a Traffic Management Center (TMC), and aid the operator in locating congestion before it happens. This information will be valuable in initiating preventive measures through a wide range of control strategies that include modifications to signal operations, diversion of traffic, and dissemination of route guidance information.

1.1 Motivation for Demand Calibration

Real-time applications of a Dynamic Traffic Assignment system typically utilize inputs from a traffic surveillance system to estimate and predict O-D flows (Figure 1-1). These predicted flows are then used as a basis to generate route guidance that may be disseminated to equipped drivers as traveler information. It is important to note that driver reaction to this guidance could merely cause spatial and temporal shifts in the predicted congestion, thereby invalidating the very prediction that influenced the guidance. The credibility of such systems therefore relies heavily on their ability to accurately estimate and predict traffic conditions under congested regimes, and to generate *consistent* route guidance. Consistency indicates that the predicted network state (in terms of congestion and travel times) matches that actually experienced by the drivers on the network.

Driver behavior and the underlying trip patterns are typically characteristic of the demographical and geographic section under study. Decisions regarding choice of route or departure time can depend on a host of observable variables (such as socio-economic characteristics, route attributes and value of time) and latent variables (like network knowledge and "aggressiveness") that capture local conditions. The origin-destination flow patterns might be a function of the spatial distribution of residential and commercial zones. Accurate estimation and prediction of O-D flows and driver behavior therefore play key roles in the evaluation of driver response to guidance in the context of real-time traffic management. In order to ensure that the system reacts and behaves in a realistic manner when deployed at site, it is critical to calibrate these models against field data collected from the site of actual deployment. The calibration of the O-D estimation/prediction and route choice models is therefore critical to

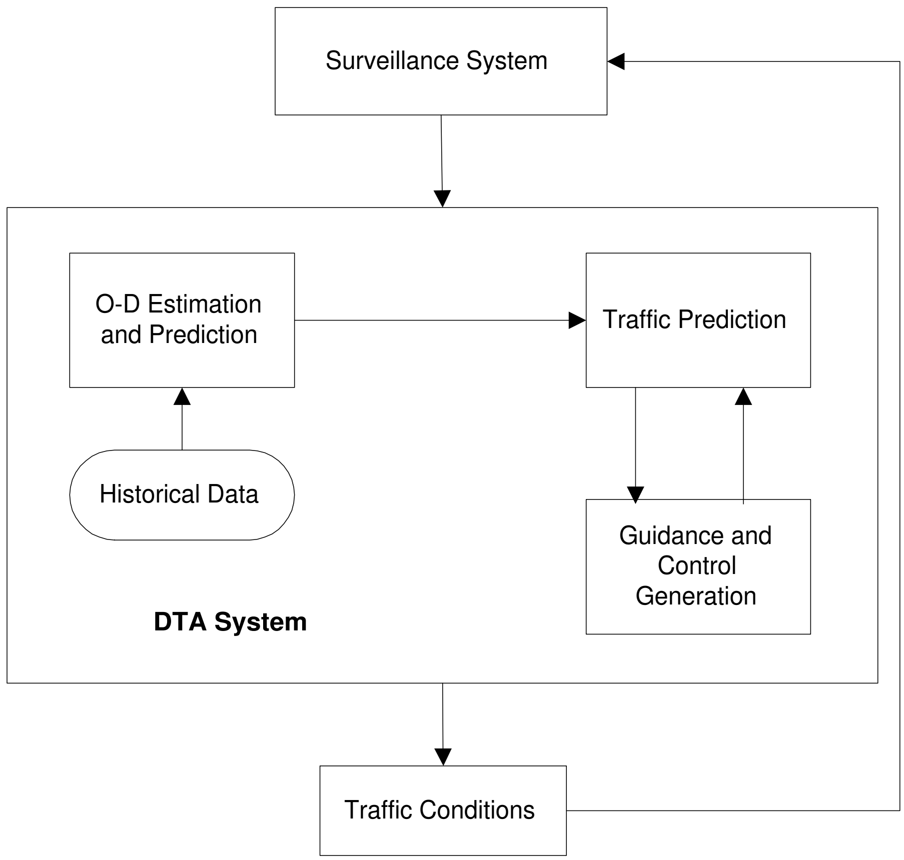

Figure 1-1: Dynamic Traffic Assignment

establishing the credibility of the guidance system. Further, the close linkage between the route choice model and the O-D estimation and prediction model necessitates a joint calibration approach that will yield a consistent set of parameters across both models. This thesis focuses on the joint calibration of the demand simulator, comprised of the driver route choice and O-D estimation and prediction models.

1.2 Problem Definition and Thesis Objective

Figure 1-2 summarizes the input and output requirements of the calibration process. For the present, the calibration process can be visualized as a black box that uses available O-D flow estimates and several days of surveillance data as inputs. The outputs from the black box are the parameters in the route choice and O-D estimation and prediction models.

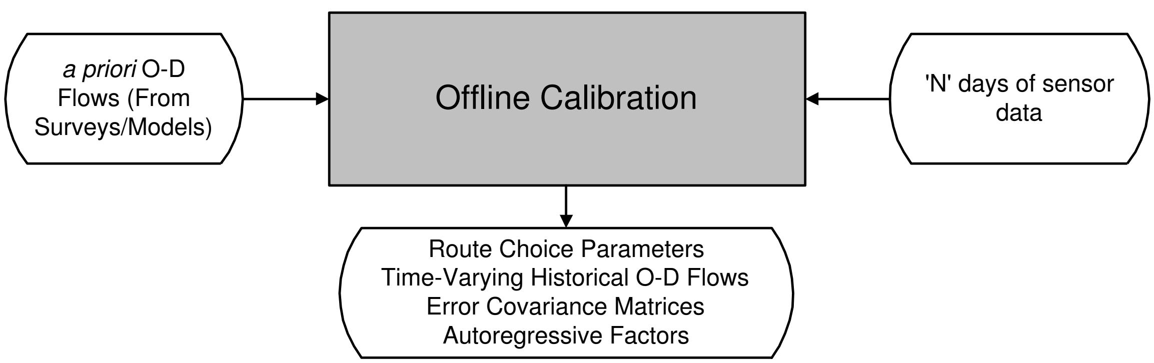

Figure 1-2: The Calibration Problem

Initial O-D flow estimates may be obtained through surveys, planning studies, or demand forecasting models. The quality of these estimates can vary significantly based on the methods used to generate them. Often, a good set of initial flow estimates is not available, and has to be extracted from the surveillance data. The real-time data collected by the surveillance system can include time-dependent traffic counts, occupancies (a proxy for densities), speeds and queue lengths on links equipped with sensors. While the counts are typically used by the O-D estimation and prediction model, the speeds, occupancies and queue lengths will be used to ensure the accuracy of the supply-side model parameters that govern the movement of vehicles on the network.

The parameters in the route choice model are theoretically estimated from detailed disaggregate data obtained through travel surveys. Ideally, such surveys yield data pertaining to a host of attributes and socio-economic characteristics that might help explain individual route choice decisions. However, such a rich data set is seldom available. In most cases, we need to identify alternative methods of calibrating the route choice model parameters from available aggregate data. This data is often a manifestation of individual discrete choice decisions. The calibration approach should be able to accurately reproduce the observed aggregate data by adjusting the parameters in the individual driver behavior models. In this thesis, we propose a calibration methodology to jointly calibrate the parameters in the O-D estimation/prediction and route choice models using aggregate sensor data.

Each additional day of data contains valuable information about O-D flow patterns and driver behavior. The calibration framework takes advantage of the availability of several days of sensor data to generate the best parameter estimates. Various alternative approaches to using the available data are discussed. Demand patterns might also vary by day of the week, weather conditions, special events and incidents. Given data collected over an extended period of time, we would like to generate stratified O-D flow databases that cover such a wide range of demand conditions. Finally, recalling the real-time nature of the DTA system, we require a convenient means of updating the historical database as each additional day of observations are recorded.

1.3 Literature Review

Literature on the calibration of DTA systems is limited, and often relates to simple networks under uncongested traffic flow regimes. He, Miaou, Ran and Lan (1999) attempt to list the major sources of error in a DTA system, and lay out frameworks for the offline and online calibration of the system. The proposed frameworks treat the calibration of the dynamic travel time, route choice, flow propagation and O-D estimation models sequentially. The authors consider a modified Greenshields model to explain dynamic travel time variations on freeway links, and split the travel times on arterials into a cruise time component and a delay component[1]. The suggested calibration approach aims to minimize the "distance" between the analytically computed travel times and those measured by detectors. Further, the maximum likelihood estimation procedure suggested for the calibration of the route choice model relies heavily on the availability of adequate survey or detector data about travelers' route choices. This assumption would fail in several real cases, where only aggregate network performance measures (such as link counts) are available. A procedure similar to that adopted for the dynamic travel time models is applied for the flow propagation model, where link inflows and outflows are matched against detector data. While such a detailed level of model calibration might be preferred, the lack of such a rich disaggregate detector data set would often render the approach infeasible. In addition, the paper does not include the O-D estimation module, which constitutes a critical part of the demand simulator within a DTA system.

In a subsequent paper, He and Ran (2000) suggest a calibration and validation approach that focuses on the route choice and flow propagation components of a DTA system. This paper again assumes prior knowledge of time-dependent O-D matrices, and further simplifies the demand process by imposing the temporal independence of O-D flows between all O-D pairs. Also, the assumption of discrete data availability to allow a Maximum Likelihood Estimation of the route choice model is still a restriction on the practical applicability of the proposed approach. The two approaches reviewed thus far fail to address the overall problem of jointly calibrating the O-D estimation and route choice models.

An approach by Hawas (2000) uses an *ad hoc* ranking of DTA system components to determine the order in which these components must be calibrated. However, a linear relationship between components is inherently assumed, and cyclic data ex-

[1]Delays on arterial links are attributed to queuing at intersections.

change through feedback between components (as in the demand-supply interaction within a DTA system) is ignored. Further, the paper uses the components of a traffic simulator, and not a DTA system, as a case study.

The problem of O-D estimation itself has been the focus of several theoretical studies. Cascetta (1984) and Cascetta, Inaudi and Marquis (1993) extend a GLS-based static O-D estimation approach to the dynamic context[2]. The tests, however, were performed on a linear network with no route choice. Also, no guidelines were provided for estimating the error covariance matrices that are critical estimation inputs.

Subsequent work by Ashok (1996), Ashok and Ben-Akiva (1993) and Ashok and Ben-Akiva (2000a) outlines a real-time O-D estimation and prediction framework based on a Kalman Filtering algorithm working with *deviations* of O-D flows from their historical values[3]. The approach is flexible so as to allow the explicit handling of measurement errors and stochasticity in the assignment matrices. However, the theoretical development is tested on small scale networks with minimal or no route choice. Moreover, the initialization of the Kalman Filter requires historical flows, error covariances and autoregressive matrices, that are unknown at the beginning of the calibration process. While some approaches to estimating these quantities are suggested, they have neither been tested on real-sized networks, nor studied in detail.

Van der Zijpp and Lindveld (2001) present a framework for O-D estimation that attempts to capture the phenomenon of peak-spreading in congested traffic networks by integrating both route and departure time choice into the estimation process. The core of this approach involves the estimation of a dynamic O-D matrix, whose cells are also associated with preferred departure times. A switch from the preferred departure time is associated with a certain disutility which is dictated by a schedule delay function. The estimation problem is modeled as a Space-Time Extended Net-

[2]A detailed review of dynamic O-D estimation using Generalized Least Squares (GLS) is presented in Section 2.4.

[3]Deviations attempt to capture the wealth of information about spatial and temporal relationships between O-D flows contained in the historical estimates. The historical database would be synthesized from several days of data

work (STEN) built by replicating the physical network as many times as there are departure time periods. The replicated virtual origin nodes are duplicated further to represent the preferred departure interval, with links between the two sets of extra nodes depicting the realized departure intervals. The authors outline an iterative solution scheme based on the Frank-Wolfe user optimal assignment algorithm to identify a STEN that is consistent with the observed link flows. In a separate discussion, the authors mention some of the limitations and drawbacks of the methodology. Firstly, travel delays on links are approximated by a function of link inflows alone:

$$t_a(q_a) = t_a^0 \left(1 + a \left(\frac{q_a}{q^{cap}} \right)^{\beta} \right) \tag{1.1}$$

where t_a is the delay on link a, t_a^0 is the delay under free-flow conditions, q_a is the inflow for link a and q^{cap} is the capacity of the link. While this function does not explicitly capture congestion effects due to queuing, the authors suggest that a bottleneck model might be approximated by employing a large value for the constant β. However, the Frank-Wolfe algorithm was found to be incapable of handling the sensitive link delay functions resulting from large values of β. Other experimental observations point to the possibility of the iterative algorithm either cycling indefinitely between equivalent STENs, or terminating at a local minimum or saddle point. In addition, the numerical tests were performed on a small artificial network with assumed "true" O-D flow, preferred departure time and scheduled delay data. The route choice model was also assumed to be known.

Hazelton (2000) presents a Bayesian approach to the static O-D estimation problem. Assuming Poisson-distributed O-D flows, the paper outlines a Maximum Likelihood Estimation procedure to estimate the unknown O-D flows from several days of link observations. The full likelihood expression is also simplified to yield a tractable estimation method that incorporates measurement errors in the observed link counts. The paper, however, suffers from several limitations. Firstly, the focus remains on static O-D estimation, and is therefore not directly suited to DTA scenarios. Secondly, the theoretical underpinnings are based on uncongested network conditions, which would be unrealistic in most real-time DTA applications. Lastly, the numerical tests described in the paper were performed on a small and simple grid network with 6 nodes and 14 directed links, along with synthetic data based on a known matrix of true O-D flows.

In a recent paper, Cascetta and Postorino (2001) evaluate different iterative solution algorithms for the fixed-point O-D estimation problem on congested networks. However, the paper does not focus on model calibration, and its scope is limited to static O-D estimation. As the paper provides key insights into the solution of fixed-point problems, we now review the concepts presented therein. Cascetta et al state the fixed-point formulation of the O-D estimation problem as:

$$\mathbf{f}^* = \mathbf{f}(\mathbf{t}^*) \tag{1.2}$$

$$\mathbf{t}^* = argmin[\mathbf{F}_1(\mathbf{x}, \bar{\mathbf{t}}) + \mathbf{F}_2(\mathbf{H}^*(\mathbf{c}(\mathbf{f}^*))\mathbf{x}, \bar{\mathbf{f}})] \tag{1.3}$$

The function $\mathbf{f}(.)$ in Equation 1.2 maps estimated O-D flows $\mathbf{t}^*$ to estimated link flows $\mathbf{f}^*$. Equation 1.3 is the solution step that generates $\mathbf{t}^*$ by optimizing over the variable $\mathbf{x}$, with $\mathbf{F}_1(.)$ and $\mathbf{F}_2(.)$ forming two terms in the objective function. While $\mathbf{F}_1(.)$ measures the distance between the O-D flow $\mathbf{x}$ and the target flow $\bar{\mathbf{t}}$, the function $\mathbf{F}_2(.)$ measures the distance between the actual vehicle counts $\bar{\mathbf{f}}$ measured at sensors, and the link flows resulting from the assignment of the demand vector $\mathbf{x}$. In addition, the function $\mathbf{H}^*(.)$ assigns the O-D vector $\mathbf{x}$ first to paths flows (using link cost function $\mathbf{c}(.)$) and then to link flows.

The paper briefly looks at existence and uniqueness conditions for the composed fixed point problem outlined above. While these conditions are in general difficult to analyze mathematically, the authors state sufficient conditions that require both $\mathbf{F}_1(.)$ and $\mathbf{F}_2(.)$ to be continuous and convex, while at least one of them is strictly convex. Three iterative algorithms were tested for accuracy and speed of convergence.

Functional Iteration provides a simple iterative framework where the latest O-D flow estimates are fed back into the assignment model for the next iteration. The *Method of Successive Averages*, or MSA, is an alternative scheme that assigns weights to the results from each iteration:

$$\mathbf{t}^k = \frac{1}{k}\mathbf{x}^k + \frac{k-1}{k}\mathbf{t}^{k-1} \tag{1.4}$$

where the "filtered" flows $\mathbf{t}^k$ are obtained by weighting the latest estimates $\mathbf{x}^k$, and the filtered flows from the previous iteration $\mathbf{t}^{k-1}$. The iteration counter k is used to generate the weights. While this method has the advantage of utilizing the estimates from all iterations, a potential drawback is that newer estimates are allotted lower weights. Consequently, "bad" initial estimates might be weighted higher, thereby lowering the speed of convergence. The *Method of Successive Averages with Decreased Reinitializations*, or MSADR, is a modification of the MSA algorithm that reinitializes the algorithm's "memory" with a frequency that decreases with the number of iterations. This acceleration step causes the algorithm to skip intermediate values and focus more on the later iterations that are closer to the solution. While the MSADR and MSA algorithms produce the same sequence of solutions in the limit, the convergence of the MSA (and hence of the MSADR) is difficult to establish. These algorithms are therefore suggested as heuristic approaches to solve the fixed point problem of interest. The authors mention the *Baillon Algorithm* as a fourth approach, but restrict themselves to the first three schemes for numerical testing.

The paper concludes that all three algorithms tested converged to the same fixed point even under stringent stopping tests. Functional iteration and MSADR outperformed MSA on speed of convergence. While these findings are indicative of the suitability of iterative schemes in solving fixed-point problems, they suffer from several limitations:

- The tests were based on a small and simple grid network.
- The study used synthetic surveillance data. O-D flow and sensor count measurements were generated artificially from known values.

- The Logit route choice model used a fixed travel time coefficient.

- The model estimated static O-D flows. While these might suffice for certain limited planning applications, they would be inadequate in a real-time DTMS scenario.

Conclusion

While there has been a profusion of theoretical works on DTA systems, the literature indicates that little work has been done on the calibration of their model components. Much of the results have been based on small and simple uncongested networks with little or no route choice, using synthetic data generated by adding a random term to some assumed "true" values. A few studies involving congested networks have either focused only on static O-D estimation, or have used simple networks consisting entirely of freeways. Above all, there seem to have been no previous attempts at jointly calibrating the parameters of the route choice model and the O-D estimation and prediction model. This thesis presents a framework for the joint calibration of the route choice and O-D estimation and prediction models within a DTA system. The proposed methodology is then applied to a case study from a large-sized network comprising of both freeways and arterials. The estimation and prediction abilities of the calibrated system are validated against real sensor data obtained from the field.

1.4 Thesis Outline

This thesis is organized as follows. In Chapter 2, we undertake a detailed explanation of the input and output requirements of the demand calibration process for a DTA system, and present the proposed calibration methodology. Chapter 3 explores issues pertaining to setting up a historical database of O-D flows and other inputs to the demand simulator. In chapter 4, we introduce DynaMIT, a state-of-the-art DTA-based prediction and guidance system, and DynaMIT-P, an application of DTA for planning. This chapter briefly describes the features, models and applications of the

DynaMIT and DynaMIT-P systems that form a basis for our case studies. Chapter 5 describes case studies that demonstrate the feasibility and robustness of the proposed calibration methodology when applied to large-scale networks, and presents validation results that evaluate the estimation and prediction capabilities of the calibrated DynaMIT system. We conclude by stating some of the contributions of this research in Chapter 6, and outline directions for future research.

Chapter 2

Calibration Methodology

The dynamic traffic assignment system should possess the ability to accurately replicate traffic conditions as observed by the surveillance system. The primary objective of calibration is therefore to identify DTA model parameters and inputs that match the offline surveillance data available. In this chapter, we cast the calibration problem in an optimization framework, and discuss some important properties characterizing the problem and its solution. We later present an iterative approach to the joint calibration of the demand models within a DTA system.

Figure 2-1 summarizes the overall framework for the calibration of a DTA system. We first identify a set of parameters that requires calibration, and assign initial values to the same. The next step involves running the DTA system with the chosen parameters to generate simulated measures of system performance. This output is compared against the observed data obtained from the surveillance system, to evaluate an objective function. Our parameter estimates are now modified, and a new optimization problem is solved. This process is continued iteratively until a predefined set of convergence criteria is satisfied.

The calibration framework can be viewed as a large optimization problem, with the final objective of matching various simulated and observed quantities:

$$\underset{\beta,\,\gamma,\,\mathbf{x}_p}{Minimize} \left\| \text{simulated quantity} - \text{observed quantity} \right\|$$

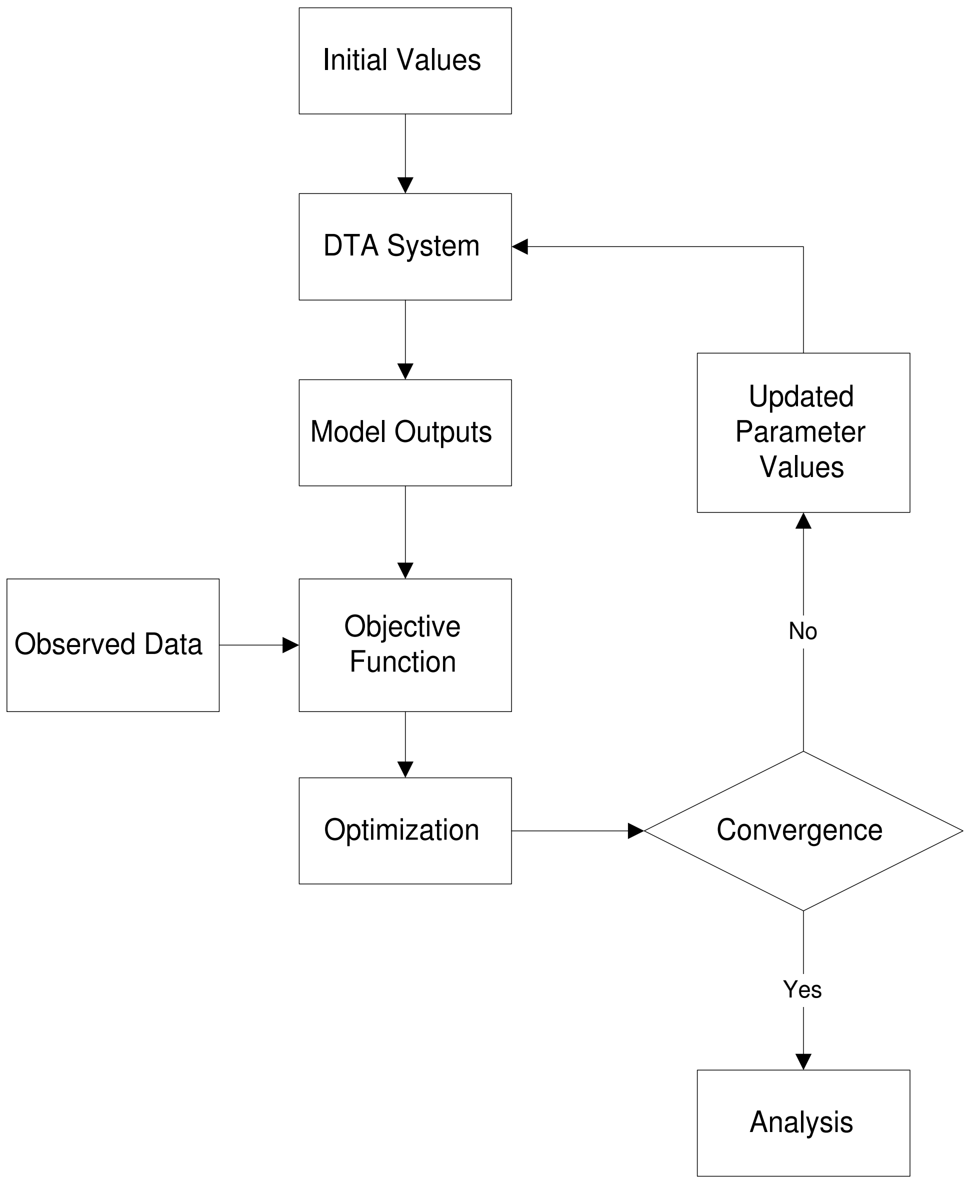

Figure 2-1: General Calibration Framework

The minimization can be over a combination of several network performance measures such as flows, speeds or densities. The simulated quantities are obtained through the following steps:

$$
\begin{aligned}
\mathbf{x}_h &= argmin[\mathbf{F}_1(\mathbf{x}_h, \mathbf{x}_h^a) + \mathbf{F}_2(\sum_{p=h-p'}^{h} \mathbf{a}_h^p \mathbf{x}_p, \mathbf{y}_h)] && (2.1)\\
\mathbf{a}_h^p &= g(\mathbf{x}_p, \beta, \gamma, tt_l^{eq}) && (2.2)\\
tt_l^{eq} &= h(\beta, \gamma, \mathbf{x}_p) && (2.3)
\end{aligned}
$$

where β represents the route choice parameters, and γ the parameters in the supply simulator. Equation (2.1) forms the basic O-D estimation step in the calibration framework, and is itself an optimization problem with a two-part objective function. The function $\mathbf{F}_1$ measures the distance of the estimated flows $\mathbf{x}_h$ from their target values $\mathbf{x}_h^a$, while $\mathbf{F}_2$ measures the distance of the measured counts $\mathbf{y}_h$ from their simulated values. The assignment matrices $\mathbf{a}_h^p$ required by the O-D estimation module are outputs of the dynamic network loading model, and are functions of the as yet unknown O-D flows, the equilibrium travel times on each link (tt_l^{eq}), the route choice parameters β and supply-side parameters γ (Equation (2.2)). Finally, the equilibrium travel times are themselves a function of the O-D flows, route choice parameters and supply-side parameters (Equation (2.3)). Equations (2.1) to (2.3) therefore capture the fixed-point nature of the calibration problem.

We now take a deeper look at the demand calibration process. Let us begin with the O-D estimation module, which forms an important part of the demand simulator within any DTA system. This module attempts to estimate point-to-point flows in the network based on link-specific traffic counts measured by the surveillance system. A key input for O-D estimation is the assignment matrix, the dimensions of which depend on the number of independent link sensors in the network and the number of O-D pairs to be estimated. Each element in the assignment matrix, therefore, represents the fraction of an O-D flow associated with a particular link, and is derived by simulating path flows in a dynamic network loading module. Path flows in turn are obtained through a route choice model combined with knowledge of the O-D flows themselves, which we set out to estimate (Figure 2-2). The solution to the calibration problem is therefore a fixed point between the route choice fractions, the assignment fractions and the O-D flows, and necessitates an iterative solution approach.

2.1 Iterative Approach to Demand Calibration

The proposed iterative approach is illustrated in Figure 2-3. We use the route choice model as the starting point in our iterative framework. Ideally, the parameters in the route choice model should be estimated separately using a rich travel survey data set. Such a survey would typically contain information regarding individual drivers' socio-economic characteristics and preferences, and attributes associated with the alternatives[1]. In the absence of such disaggregate data, we need a method of estimating these parameters from aggregate data such as vehicle counts at sensors. The first step is to assume a set of initial values for the parameters in the route choice model. In order to generate an assignment matrix, we need additional information about path travel times and the O-D flows themselves. For a large network, the effort required to collect and store time-dependent path travel times throughout the period of interest can be prohibitive. We therefore simulate the assignment fractions using an initial seed O-D matrix. The route choice model uses the latest available travel time estimates to compute path choice probabilities[2]. Driver route choices are then simulated to generate time-dependent assignment matrices. The assignment matrix and corresponding O-D flows are passed on to the O-D estimation module, the output of which is a new set of O-D flows.

This process raises two issues relating to convergence. The first issue deals with the **assignment matrix**. The newly estimated O-D flows (generated by the O-D estimation module) can be re-assigned to the network, using the original route

[1]The alternatives in this case being the different possible combinations of routes and departure times.

[2]In the absence of initial travel time estimates, we might start the process with free-flow travel times.

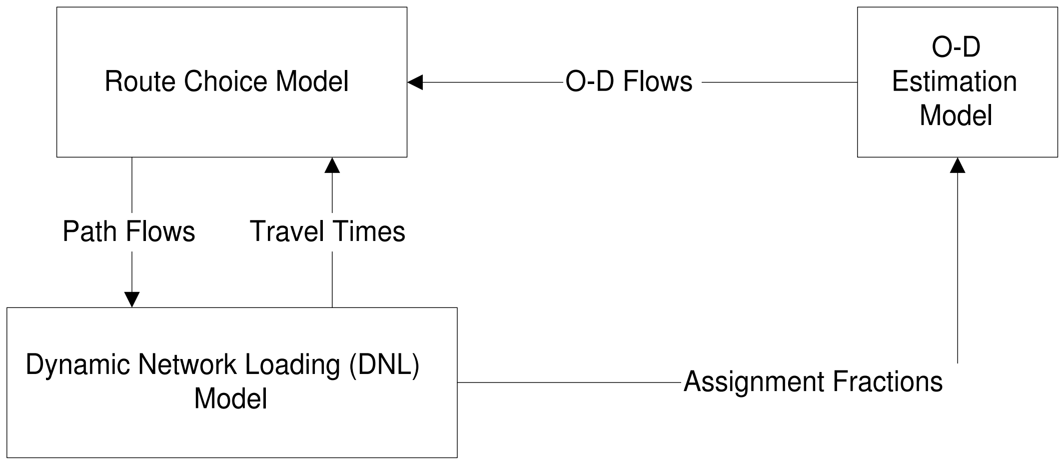

Figure 2-2: The Fixed Point Problem

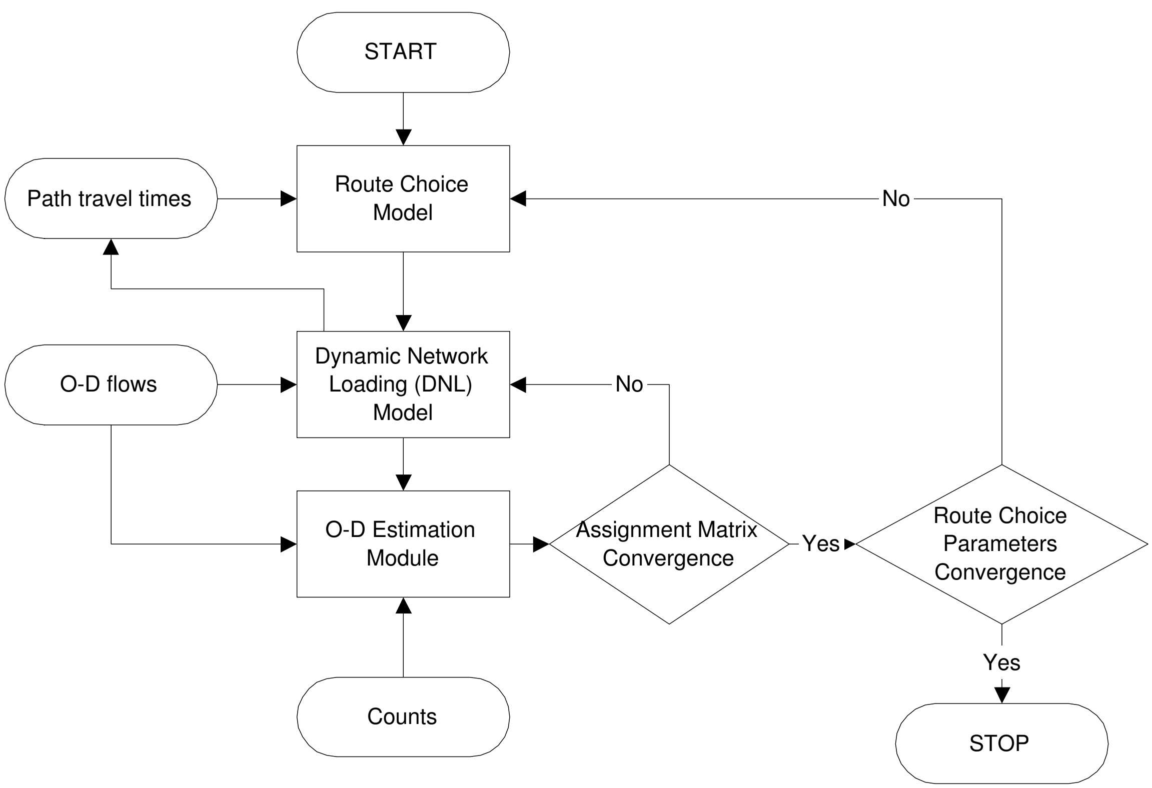

Figure 2-3: Iterative Calibration Framework

choice model. The resulting assignment matrix is very likely to be different from the matrix used as input to the O-D estimation module. It is therefore essential to ensure consistency by iterating locally on the assignment matrix. A smoothing algorithm such as a weighted average[3] (of the latest matrix estimate with estimates from previous iterations) can be used as input for the subsequent iteration. A suitable termination rule should be determined in order to provide a stopping criterion. The second issue relates to the **route choice parameters**. The estimated O-D flows obtained thus far were based on the initial values assumed for the parameters in the route choice model. In the absence of disaggregate data, we are forced to employ aggregate approaches to estimate these parameters. Observing that link flows are indirect manifestations of driver route choice decisions, we iterate on the route choice model parameters until we are able to replicate fairly well the link flows observed at the sensors.

2.1.1 Iterative schemes

Cascetta and Postorino (2001) present three different algorithms[4] to iteratively solve the fixed point problem associated with static O-D estimation in congested networks. Their tests indicate that functional iterations and MSADR perform better than the MSA in terms of speed of convergence. No significant difference was noticed in the final solutions obtained. It is hard, however, to extend the findings to a larger dynamic problem. The suitability of the various algorithms will have to be evaluated through numerical testing.

The general structure of the iterative solutions consists of three basic steps:

- Re-assign the current estimated O-D flows to obtain a new assignment matrix.
- Compute a new assignment matrix using a smoothing algorithm.
- Use the smoothed assignment matrix to estimate new O-D flows.

[3]Typically, newer estimates are assigned higher weights than estimates from previous iterations

[4]Functional Iteration, Method of Successive Averages (MSA) and MSA with decreasing reinitializations (MSADR)

The steps are repeated until a suitable set of termination rules is satisfied. The following sections provide a brief overview of the three iterative schemes.

Functional iteration represents the simplest form of the generic iterative algorithm outlined previously, with the latest assignment matrix estimate being used for the next iteration. Convergence is measured by the difference between successive O-D flow estimates across all O-D pairs. A simple modification in the above procedure can allow for the averaging of successive O-D flow estimates while computing the current estimate. Yet another algorithm can be visualized by averaging successive assignment matrices. However, this approach would involve the explicit evaluation and storage of several assignment matrices, and can be a burden on computational efficiency and storage.

The **Method of Successive Averages (MSA)** provides a way of assigning weights to successive O-D flow estimates, thereby speeding up the convergence of the iterative solution:

$$\mathbf{x}^{k^*} = \frac{1}{k}\mathbf{x}^k + \frac{k-1}{k}\mathbf{x}^{k-1} \tag{2.4}$$

where $\mathbf{x}^{k^*}$ is the new flow vector after iteration k, $\mathbf{x}^k$ is the estimate generated in iteration k and $\mathbf{x}^{k-1}$ is the estimate from the previous iteration. Such a weighted approach is often known as *smoothing*, and has been found to significantly improve convergence properties. It should be noted that the weight assigned to a new estimated flow vector decreases as the iterations progress towards a solution, implying that initial solutions are weighed more heavily than those closer to the true solution. Consequently, the algorithm can slow down dramatically in the final iterations before convergence. This drawback of the classical MSA algorithm is remedied in part by its variant, the **MSA with Decreasing Reinitializations (MSADR)**, which erases the memory of the algorithm at predefined intervals. The frequency with which the memory is reinitialized decreases as the algorithm gets closer to the true solution it seeks. The effect of such an approach would be to skip intermediate solution steps, thereby ensuring faster convergence.

2.2 The Route Choice Model

The choice of route and departure time are important aspects in the modeling of incidents, response to information and the evaluation of guidance strategies. The route choice model captures key aspects of driver behavior in the context of dynamic traffic assignment, and attempts to replicate driver choice decisions using a set of explanatory variables. These variables can include route attributes such as travel times and cost components, as well as driver characteristics such as income level.

Route choices can be modeled at two levels. Pre-trip route choices are made *before* the driver embarks on the trip. At this stage, the driver has access to traffic information and guidance from a variety of sources such as radio traffic bulletins, the Internet, or a dedicated traveler information provider. The important aspect to note here is that drivers have the ability to change their departure time depending on the current information. Hence apart from changing the chosen route, the driver can also decide to leave sooner or later, in a bid to beat the rush. Other alternatives that are feasible at this point are a change of mode (say to public transit options), or to cancel the trip itself. Once the driver has embarked on the trip, however, several of the alternatives discussed above become unavailable. For example, drivers can no longer change their departure times. We now model en-route driver decisions that capture the phenomenon of path switching in response to real-time information. Such information could be available, for example, through in-vehicle units (for equipped drivers) or through Variable Message Signs (VMS) installed at specific locations on the network.

The route choice model operates on each driver's choice set. This set contains all feasible combinations of departure times and routes. The choice probabilities for all alternatives within the choice set are evaluated using a discrete choice model[5]. It further simulates this choice through random draws, and determines the chosen path. The aggregation of all such chosen paths will result in path flows by O-D pair, which are used by the Dynamic Network Loading model discussed next.

[5]A detailed discussion of route choice model parameters is presented in Section 4.4.1.

2.3 The Dynamic Network Loading Model

The path flows output by the route choice model are used by the dynamic network loading (DNL) model to create time-dependent assignment matrices. This step involves the assigning of path flows to the network, and counting the vehicles as they cross traffic sensors. The departure times of the counted vehicles are also tracked, as the matrices $\mathbf{a}_h^p$ depend on the current interval h as well as the departure interval p.

Various analytical approaches with different underlying assumptions have been suggested for the DNL problem. Cascetta and Cantarella (1991a) approximate the behavior of all vehicles on path k in time interval h to be part of the same packet. In a modified continuous packet approach, the vehicles in a packet are assumed to be continuously spread between the "head" and the "tail" of the packet. Cascetta and Cantarella (1991b) suggest a further variation that allows packets to "contract" or "expand", to help account for differing average speeds. All these approaches yield very complex expressions for the assignment matrix fractions, and introduce a high degree of non-linearity in the objective functions. Simulation provides us with a more tractable alternative. The path and departure time records for each vehicle can be tracked in a microscopic traffic simulator and later processed to obtain time-dependent assignment matrices. A more aggregate mesoscopic traffic simulator can also be used to achieve the same result.

2.4 The O-D Estimation Module

The O-D matrix estimation problem involves the identification of a set of O-D flows that, when assigned to the transportation network, results in link flows that closely match observed flows. The input data can be obtained from various sources, derived from different processes and characterized by different error component structures. We therefore need a unified approach that extracts the maximum information out of the available data in a statistically efficient manner.

2.4.1 Inputs and Outputs

Figure 2-4 illustrates the basic inputs and outputs for an O-D estimation model. This model obtains initial O-D flow estimates from a historical database (the generation of this database is a part of our calibration exercise). In addition, the module gets a vector of link counts from the surveillance system[6]. In an offline application, such as calibration, link counts for the entire analysis period would be available to the model. However, in a real-time context, at the end of each interval h, only the counts measured during that interval will be reported by the surveillance system. Finally, the model either gets, or computes, a set of assignment matrices[7]. By comparing the link counts measured by the sensors with the counts obtained by assigning the estimated or *a priori* O-D flows to the network, the O-D estimation model updates these estimates. While the offline estimation can attempt to estimate flows for all T intervals simultaneously, the real-time estimation necessarily proceeds sequentially through the intervals.

2.4.2 Preliminary Definitions

We begin by introducing some basic definitions that will be followed henceforth in this thesis. Consider $\mathcal{T}$, the analysis period of interest, to be divided into T subintervals $1 \ldots T$ of equal length. The network is represented by a directed graph that includes a set $\mathcal{N}$ of consecutively numbered nodes and a set of numbered links $\mathcal{L}$. The network is assumed to have n_{LK} links, of which n_l are equipped with sensors. There are n_{OD} O-D pairs of interest.

We denote by x_{rh} the number of vehicles between the rth O-D pair that departed from their origin in time interval h, and by x^H_{rh} the corresponding best historical estimate. A historical estimate typically is the result of estimation conducted during previous days. Further, let the corresponding $(n_{OD} * 1)$ vector of all O-D flows and their corresponding historical estimates be denoted as $\mathbf{x}_h$ and $\mathbf{x}^H_h$ respectively. Let

[6]More generally, the inputs are a set of direct and indirect measurements, as described shortly.

[7]As mentioned in Chapter 1, an assignment matrix maps origin-destination flows into link flows.

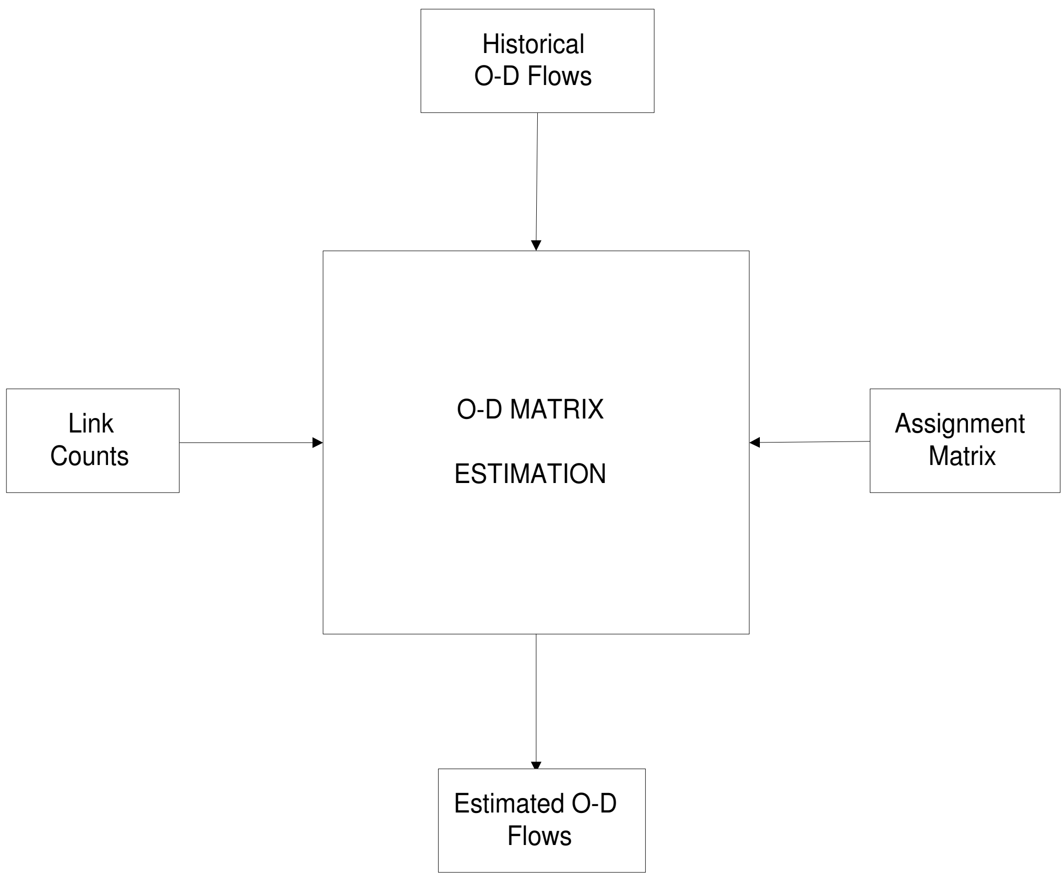

Figure 2-4: Overview of Inputs and Outputs

$\hat{\mathbf{x}}_h$ represent the estimate of $\mathbf{x}_h$. Finally, denote by y_{lh} the observed traffic counts at sensor l during time interval h, and by $\mathbf{y}_h$ the corresponding $(n_l * 1)$ vector.

Traffic measurements fall into two broad categories:

- Indirect measurements
- Direct measurements

Indirect measurements contain information about the quantities we wish to estimate, but are not direct observations of the quantities themselves. Our objective is to estimate O-D flows. Observations of link counts at traffic sensors, for example, would contain information about the O-D flows that caused them. In fact, link counts constitute the most common type of indirect measurements. Link flows can be related to the O-D flows through a set of assignment matrices:

$$\mathbf{y}_h = \sum_{p=h-p'}^{h} \mathbf{a}_h^p \mathbf{x}_p + \mathbf{v}_h \tag{2.5}$$

where $\mathbf{y}_h$ are the link counts observed in time interval h, $\mathbf{x}_p$ is a matrix of O-D flows departing in interval p, $\mathbf{a}_h^p$ is an assignment matrix that relates flows departing in interval p to counts observed in interval h, and $\mathbf{v}_h$ is a random error term. p' is the maximum number of intervals required by a vehicle to clear the network. Let the variance-covariance matrix of $\mathbf{v}_h$ be denoted as $\mathbf{V}_h$.

The indirect measurements represent n_l equations in n_{OD} unknown variables. Typically, n_{OD} is much larger than n_l. In order to be able to estimate the n_{OD} unknown flows, we supplement the indirect measurements with direct measurement "observations".

Direct measurements provide preliminary estimates of the unknown O-D flows. We therefore express a direct measurement as follows:

$$\mathbf{x}_h^a = \mathbf{x}_h + \mathbf{u}_h \tag{2.6}$$

where $\mathbf{x}_h^a$ is the starting estimate of matrix $\mathbf{x}_h$, and $\mathbf{u}_h$ is a random error term. We now have a total of $(n_l + n_{OD})$ equations to solve for the n_{OD} unknown flows. Let the variance-covariance matrix of $\mathbf{u}_h$ be denoted as $\mathbf{W}_h$.

2.4.3 Deviations

It should be noted that Equation (2.6) only provides a means of including additional observations on the O-D flows. The *a priori* estimates $\mathbf{x}_h^a$ can have a variety of functional forms (Ashok (1996)). Assume that we possess a historical database of O-D flows synthesized from several days of data. Such a database would contain valuable information about the spatial and temporal relationships between O-D flows. In order to accommodate this structural information into the new estimates, we work with *deviations* of current O-D flows from their historical values (Figure 2-5):

$$\mathbf{x}_h^a = \mathbf{x}_h^H + \sum_{p=h-q'}^{h-1} \mathbf{f}_h^p(\mathbf{x}_p - \mathbf{x}_p^H) \tag{2.7}$$

where $\mathbf{f}_h^p$ is a matrix that relates deviations in time interval p to the current time interval h. Stated as deviations,

$$\partial\mathbf{x}_h^a = \sum_{p=h-q'}^{h-1} \mathbf{f}_h^p \partial\mathbf{x}_p \tag{2.8}$$

The measurement equations get modified as:

$$\mathbf{y}_h - \sum_{p=h-p'}^{h} \mathbf{a}_h^p \mathbf{x}_p^H = \sum_{p=h-p'}^{h} \mathbf{a}_h^p(\mathbf{x}_p - \mathbf{x}_p^H) + \mathbf{v}_h \tag{2.9}$$

$$\mathbf{y}_h - \mathbf{y}_h^H = \sum_{p=h-p'}^{h} \mathbf{a}_h^p \partial\mathbf{x}_p + \mathbf{v}_h \tag{2.10}$$

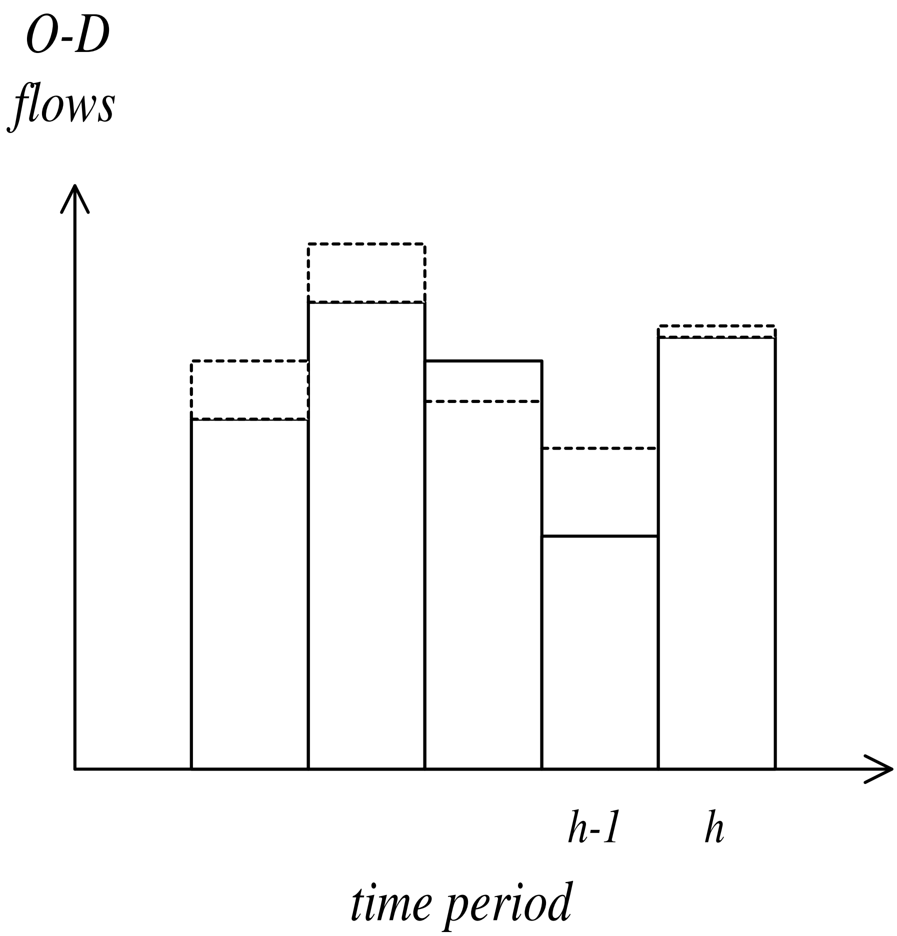

Figure 2-5: O-D Flow Deviations

where $\mathbf{y}_h^H = \sum_{p=h-p'}^{h} \mathbf{a}_h^p \mathbf{x}_p^H$. We are now ready to synthesize O-D flow estimates by combining the information contained in the direct and indirect measurements. The following sections review two approaches to O-D estimation from the literature. The first is a generalized least squares approach, while the second uses a Kalman Filtering technique.

2.4.4 Generalized Least Squares (GLS) Approach

The classical GLS estimator minimizes the "distance" between the estimated O-D flows and the *a priori* flows, while simultaneously ensuring that the measured link flows are as close as possible to the newly assigned link flows (the link flows obtained by assigning the newly estimated O-D flows to the network). The indirect measurements are first adjusted for link flow contributions from previous intervals:

$$\mathbf{y}_h - \sum_{p=h-p'}^{h-1} \mathbf{a}_h^p \hat{\mathbf{x}}_p = \mathbf{a}_h^h \mathbf{x}_h + \mathbf{v}_h \tag{2.11}$$

The direct measurement equations are maintained as is:

$$\mathbf{x}_h^a = \mathbf{x}_h + \mathbf{u}_h \tag{2.12}$$

The combined system of indirect and direct measurements can be expressed in matrix algebra as:

$$\begin{bmatrix} \mathbf{y}_h - \sum_{p=h-p'}^{h-1} \mathbf{a}_h^p \hat{\mathbf{x}}_p \\ \mathbf{x}_h^a \end{bmatrix} = \begin{bmatrix} \mathbf{a}_h^p \\ \mathbf{I}_{n_{OD}} \end{bmatrix} \mathbf{x}_h + \begin{bmatrix} \mathbf{v}_h \\ \mathbf{u}_h \end{bmatrix}$$

or

$$\mathcal{Y}_h = \mathbf{A}_h \mathbf{x}_h + \epsilon_h \tag{2.13}$$

The error covariances $\mathbf{V}_h$ and $\mathbf{W}_h$ associated with the indirect and direct measurements can be represented by a single covariance matrix:

$$\Omega_h = \begin{bmatrix} \mathbf{V}_h & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_h \end{bmatrix}$$

The above structure for Ω_h assumes that the measurement errors associated with the indirect and direct measurements are not correlated. This would be a reasonable assumption in most cases, as the two types of measurements are obtained from two entirely different measurement processes. The flow estimates are then given by:

$$\hat{\mathbf{x}}_h = argmin[(\mathcal{Y}_h - \mathbf{A}_h \mathbf{x}_h)' \Omega_h^{-1} (\mathcal{Y}_h - \mathbf{A}_h \mathbf{x}_h)] \tag{2.14}$$

Ω_h^{-1} can be interpreted as a weighting matrix. The weights for each measurement (indirect or direct) reflect our confidence in the measurement process. For example, we might discount the information contained in the *a priori* O-D flows because they were derived from a static planning study. The corresponding high variance in the direct measurement errors would manifest themselves as a low weight. On the other hand, we also capture the effect of measurement errors in the counts recorded by traffic sensors. A good sensor would therefore be associated with a lower error variance, thereby implying a higher weight. The variability of the errors is thus captured through the covariance matrix Ω_h.

The estimator we have just described is a *sequential* estimator, in that it treats the T time intervals one at a time. In other words, the estimates $\hat{\mathbf{x}}_{h-1}, \hat{\mathbf{x}}_{h-2}, \ldots$ obtained for past intervals $h-1, h-2, \ldots$ are kept constant when estimating flows for current interval h. Expanding Ω_h, we can state the optimization problem as:

$$\hat{\mathbf{x}}_h = argmin[(\mathbf{x}_h - \mathbf{x}_h^a)'\mathbf{W}_h^{-1}(\mathbf{x}_h - \mathbf{x}_h^a)$$
$$+(\mathbf{y}_h - \sum_{p=h-p'}^{h-1} \mathbf{a}_h^p \hat{\mathbf{x}}_p - \mathbf{a}_h^h \mathbf{x}_h)'\mathbf{V}_h^{-1}(\mathbf{y}_h - \sum_{p=h-p'}^{h-1} \mathbf{a}_h^p \mathbf{x}_p - \mathbf{a}_h^h \mathbf{x}_h)] \tag{2.15}$$

While this approach has advantages from a computational standpoint, we might be able to obtain more consistent estimates by estimating the flows for all T time intervals *simultaneously*[8]. The equivalent of equation (2.14) would then be:

$$(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_T) = argmin[\sum_{h=1}^{T}(\mathcal{Y}_h - \mathbf{A}_h \mathbf{x}_h)'\Omega_h^{-1}(\mathcal{Y}_h - \mathbf{A}_h \mathbf{x}_h)] \tag{2.16}$$

where the minimization is over $\mathbf{x}_i \geq 0 \quad \forall i = 1, 2, \ldots, T$. Again, we can expand Equation (2.16) as:

$$(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_T) = argmin \sum_{h=1}^{T}[(\mathbf{x}_h - \mathbf{x}_h^a)'\mathbf{W}_h^{-1}(\mathbf{x}_h - \mathbf{x}_h^a)]$$
$$+\sum_{h=1}^{T}[(\mathbf{y}_h - \sum_{p=h-p'}^{h} \mathbf{a}_h^p \mathbf{x}_p)'\mathbf{V}_h^{-1}(\mathbf{y}_h - \sum_{p=h-p'}^{h} \mathbf{a}_h^p \mathbf{x}_p)] \tag{2.17}$$

It should be noted that Equations (2.16) and (2.17) represents a much bigger optimization problem when compared with Equations (2.14) and (2.15). In a practical application, computational considerations might drive us towards the sequential estimator.

2.4.5 Kalman Filter Approach

The Kalman Filter approach (Ashok and Ben-Akiva (2000b)) provides a convenient O-D estimation and prediction method for real-time applications as well as offline estimation. This approach views the O-D estimation problem as a state-space model[9]

[8]Simultaneous estimation gains significance in offline applications such as model calibration.
[9]The reader is referred to Gelb (1974) for an exhaustive coverage of state-space models.

which consists of two equation systems:

$$MeasurementEquation : \mathcal{Y}_h = \mathbf{A}_h \mathcal{X}_h + \mathbf{v}_h \tag{2.18}$$

$$TransitionEquation : \mathcal{X}_{h+1} = \Phi_h \mathcal{X}_h + \mathbf{W}_{h+1} \tag{2.19}$$

where

$$\mathcal{X}_h = \left[\begin{array}{cccc} \partial \mathbf{x}'_h & \partial \mathbf{x}'_{h-1} & \cdots & \partial \mathbf{x}'_{h-s} \end{array} \right]'$$

The transition equation is based on an autoregressive process similar to Equation (2.7). The state variables $\mathcal{X}_h$ used in the above formulation correspond to an augmented state approach that combines the advantages of both simultaneous and sequential estimators while reducing the heavy computational load associated with simultaneous estimation. The term $\mathbf{A}_h$ is a suitably constructed matrix of assignment fractions.

Assume that the initial state of the system $\mathcal{X}_0$ has known mean $\bar{\mathcal{X}}_0$ and variance $\mathbf{P}_0$. The system of measurement and transition equations can then be combined into a sequence of steps that easily integrate a rolling horizon:

$$\Sigma_{0|0} = \mathbf{P}_0 \tag{2.20}$$

$$\Sigma_{h|h-1} = \Phi_{h-1} \Sigma_{h-1|h-1} \Phi'_{h-1} + \mathcal{Q}_{h-1} \tag{2.21}$$

$$\mathbf{K}_h = \Sigma_{h|h-1} \mathbf{A}'_h (\mathbf{A}_h \Sigma_{h|h-1} \mathbf{A}'_h + \mathbf{R}_h)^{-1} \tag{2.22}$$

$$\Sigma_{h|h} = \Sigma_{h|h-1} - \mathbf{K}_h \mathbf{A}_h \Sigma_{h|h-1} \tag{2.23}$$

$$\hat{\mathcal{X}}_{0|0} = \bar{\mathcal{X}}_0 \tag{2.24}$$

$$\hat{\mathcal{X}}_{h|h-1} = \Phi_{h-1} \hat{\mathcal{X}}_{h-1|h-1} \tag{2.25}$$

$$\hat{\mathcal{X}}_{h|h} = \hat{\mathcal{X}}_{h|h-1} + \mathbf{K}_h (\mathcal{Y}_h - \mathbf{A}_h \hat{\mathcal{X}}_{h|h-1}) \tag{2.26}$$

The term $\hat{\mathcal{X}}_{h|h-1}$ represents a one-step prediction of the state $\mathcal{X}_h$, and signifies the best knowledge of the deviations $\mathcal{X}_h$ prior to obtaining the link counts for interval h. This prediction is achieved through the autoregressive process described by Equation

(2.25). $\Sigma_{h|h-1}$ denotes the variance of $\hat{\mathcal{X}}_{h|h-1}$, and depends on the uncertainty in both $\hat{\mathcal{X}}_{h-1|h-1}$ as well as the error in the autoregressive process (Equation (2.21)). The role of the Kalman gain matrix $\mathbf{K}_h$ is explained intuitively by Equation (2.26). $\mathbf{K}_h$ helps update the one-step prediction $\hat{\mathcal{X}}_{h|h-1}$ with the information contained in the new measurements $\mathcal{Y}_h$.

2.4.6 GLS vs the Kalman Filter

The Kalman Filter yields a convenient and elegant sequential methodology to update the O-D flow estimates and error covariance matrices with data from successive time intervals. The advantages of simultaneous estimation can also be easily integrated into the estimation framework through the technique of state augmentation, and is well suited to function in a rolling horizon[10]. In addition, the transition equation inherently embeds a prediction step based on a rich autoregressive process. The Kalman Filter approach is therefore a joint O-D estimation and prediction model, and is an obvious choice within a real-time DTA-based traffic prediction and guidance system.

While the advantages are many, the Kalman Filter might suffer from some limitations in a calibration framework. The inputs to the Kalman Filter algorithm include a historical database of O-D flows, covariances of the initial flows, error covariance matrices and autoregressive matrices. As several of these input parameters are likely to be unavailable in the initial stages of a calibration application, we need alternative ways of initializing the Kalman Filter before it can be deployed on site. Further, Equations (2.20) to (2.26) represent repeated matrix operations involving matrices of very large dimensions, and could potentially render an iterative calibration scheme computationally expensive and infeasible.

A practical solution to the above problem might be suggested by Ashok (1996), which presents a derivation of the Kalman Filter algorithm from a GLS approach and establishes the equivalence of the two methods for discrete stochastic linear processes.

[10]The concept of a rolling horizon in real-time DTA systems is covered in detail in Chapter 4.

Chapter 3 explores this approach in detail, and addresses issues related to initializing the calibration process, and the availability of data.

2.5 Supply-Side Calibration

While the focus of this thesis remains the calibration of the model parameters on the demand side of a dynamic traffic assignment system, one should note that a DTA system models equilibrium in the traffic network by modeling the interactions between the demand and supply components. Indeed, the assignment matrix, which is a crucial input to the O-D estimation module, is an output of a supply simulation that captures traffic dynamics and simulates driver decisions in the context of capacity constraints. It is therefore necessary to place the problem of demand calibration within an even larger overall calibration framework.

The demand calibration framework presented here assumes the availability of a calibrated supply simulator. The DTA calibration process should capture the close interaction between the demand and supply components, and should attempt to estimate the parameters in both classes of models jointly. While we do not present detailed discussions on the calibration of the supply model parameters in this work, some results on this topic may be found as part of the case studies in Chapter 5.

2.6 Model Validation

A good indication of a valid model is its ability to predict future system behavior. Model validation is therefore a critical final step in the calibration framework. The calibration process might yield consistent and efficient parameter estimates that fit the observed data. However, it is necessary to ensure that the calibrated system continues to perform in a satisfactory manner when faced with fresh information from the surveillance system. This issue is particularly important in the context of a Traffic Management Center, where the calibrated DTA system must possess the robustness to deliver accurate estimations and predictions in the face of real-time fluctuations in the measurements. These fluctuations can consist of either day-to-day variations in traffic flows, or can represent significant deviations from normal conditions due to accidents, incidents, weather or special events.

It is often too expensive or difficult to conduct separate data collection efforts for calibration and validation. In such situations, we are forced to utilize the available data wisely to achieve both objectives. We might then divide the available data into two sets. While the first set would be used to estimate our model parameters, the second might be used as the basis for model validation. Such an approach might also help eliminate some of the effects specific to a particular data set.

It should be noted that discrepancies between model predictions and system output after implementation might indicate the necessity to re-calibrate the model with new data.

2.7 Conclusion

In this chapter, we formulated the overall DTA calibration process as an optimization problem, and presented an integrated framework for the joint calibration of the demand models within a DTA system. We discussed the roles played by the components within the demand and supply simulators, and analyzed the inputs and outputs at each calibration step. We concluded with a note on model validation and the larger calibration picture. The next chapter will explore some of the issues relating to starting the calibration process in the absence of reliable initial parameter estimates, and focus on the use of multiple days of surveillance data to generate efficient estimates. Subsequent chapters will present case studies that demonstrate the feasibility and robustness of the proposed calibration methodology.

Chapter 3

Creating a Historical Database

In Chapter 2, we presented an iterative approach to the calibration of the demand models within a DTA system. However, the component steps in this process had several input requirements that might not be readily available. In this chapter, we present approaches to starting the calibration process in such cases, and outline simplifications that would allow us to estimate the missing inputs from a single day of data. We further discuss ways of obtaining efficient parameter estimates from the entire set of available data spanning several days of surveillance records, and conclude with a section detailing the process of updating our historical database of O-D flows, error covariances and autoregressive factors with each additional day of estimations.

We begin by briefly reviewing the input requirements of the O-D estimation module that forms the core of the demand calibration routine.

3.1 The O-D Estimation Module

The inputs to the O-D estimation module are the *a priori* O-D flows $\mathbf{x}_h^a$, the sensor counts $\mathbf{y}_h$, the assignment matrices $\mathbf{a}_h^p$ and the variance-covariance matrices ($\mathbf{V}_h$ and $\mathbf{W}_h$) associated with the direct and indirect measurement errors. The outputs are the estimated O-D flows. The new O-D flows can be estimated based on Cascetta's

GLS formulation (Cascetta (1984) and Cascetta et al. (1993)).

$$(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_T) = argmin \sum_{h=1}^{T} [(\mathbf{x}_h - \mathbf{x}_h^a)' \mathbf{W}_h^{-1} (\mathbf{x}_h - \mathbf{x}_h^a)] + \sum_{h=1}^{T} [(\mathbf{y}_h - \sum_{p=h-p'}^{h} \mathbf{a}_h^p \mathbf{x}_p)' \mathbf{V}_h^{-1} (\mathbf{y}_h - \sum_{p=h-p'}^{h} \mathbf{a}_h^p \mathbf{x}_p)] \quad (3.1)$$

$$\hat{\mathbf{x}}_h = argmin[(\mathbf{x}_h - \mathbf{x}_h^a)' \mathbf{W}_h^{-1} (\mathbf{x}_h - \mathbf{x}_h^a) + (\mathbf{y}_h - \sum_{p=h-p'}^{h-1} \mathbf{a}_h^p \hat{\mathbf{x}}_p - \mathbf{a}_h^h \mathbf{x}_h)' \mathbf{V}_h^{-1} (\mathbf{y}_h - \sum_{p=h-p'}^{h-1} \mathbf{a}_h^p \mathbf{x}_p - \mathbf{a}_h^h \mathbf{x}_h)] \quad (3.2)$$

Equation (3.1) solves for flows in several time intervals simultaneously. This involves the estimation of $(n_{OD} * T)$ unknown O-D flows, where n_{OD} is the number of O-D pairs, and T is the number of time intervals in the period of interest. The large number of unknown quantities increases the computational burden tremendously. Equation (3.2) sequentially estimates flows in successive time intervals. Note that both approaches assume a knowledge of $\mathbf{x}_h^a$, $\mathbf{V}_h$ and $\mathbf{W}_h$. The following sections address the issue of generating these matrices.

3.2 Generating *a priori* O-D flow estimates $(\mathbf{x}_h^a)$

The *a priori* O-D flows provide direct measurement observations of the true flows. Apart from supplementing the set of indirect measurements, these observations help capture temporal relationships and spatial travel patterns among the O-D flows. It is therefore imperative that we include these direct measurements, even if the number of indirect measurements is adequate to solve for the O-D flows. Equation (2.6) allows us to compute the *a priori* flows in a variety of ways. The most straightforward possibility is to use flows from a historical database as starting guesses. Since our aim, however, is to generate these historical flows, we need to look elsewhere for our direct measurements. In a strict sense, the direct measurements are not real observations. We simply attempt to create realistic measures that might add to the information already available. While several formulations for $\mathbf{x}_h^a$ can be visualized, we look at a few possibilities. The most straightforward option for a direct measurement is to use available O-D flow estimates:

$$\mathbf{x}_h^a = \mathbf{x}_h^H \tag{3.3}$$

In the absence of such time-dependent historical flow estimates, an alternative is to use the flows estimated for interval $h-1$ as the starting point for interval h (Cascetta et al. (1993)):

$$\mathbf{x}_h^a = \hat{\mathbf{x}}_{h-1} \tag{3.4}$$

However, such a formulation does not account for any patterns in the underlying demand process. Let us for the present assume that we have good historical O-D flow estimates[1]. Ashok (1996) hypothesizes that the ratio of interval-over-interval O-D flows is stable on a day-to-day basis, and suggests:

$$\mathbf{x}_h^a = (\mathbf{x}_h^H / \mathbf{x}_{h-1}^H)\hat{\mathbf{x}}_{h-1} \tag{3.5}$$

We might further hypothesize that the O-D flows for interval h may be influenced by flows in several preceding intervals. This prompts us to consider a modification of (3.5) that spans several prior intervals:

$$\mathbf{x}_h^a = \sum_{p=h-q'}^{h-1} \tau_p (\mathbf{x}_h^H / \mathbf{x}_p^H)\hat{\mathbf{x}}_p \ , \ \sum_p \tau_p = 1 \tag{3.6}$$

[1]We will soon revisit the case when good historical flow estimates are unavailable.

Another method of capturing temporal dependence between the O-D flows is to build an autoregressive process of degree q' into the *a priori* estimates (Ashok (1996)):

$$\mathbf{x}_h^a = \sum_{p=h-q'}^{h-1} \mathbf{f}_h^p \hat{\mathbf{x}}_p \tag{3.7}$$

The $(n_{OD} * n_{OD})$ matrix $\mathbf{f}_h^p$ in Equation (3.7) captures the effects of $\mathbf{x}_p$ on $\mathbf{x}_h$, and helps accommodate temporal effects in O-D flows. A big drawback of this method is that the wealth of information contained in the historical estimates is ignored. By moving to *deviations*, we can integrate this information into the current estimation problem:

$$\mathbf{x}_h^a = \mathbf{x}_h^H + \sum_{p=h-q'}^{h-1} \mathbf{f}_h^p (\hat{\mathbf{x}}_p - \mathbf{x}_p^H) \tag{3.8}$$

The new target O-D flows $\mathbf{x}_h^a$ now contain information about flows in past intervals. The matrix $\mathbf{f}_h^p$ is now an $(n_{OD} * n_{OD})$ matrix of effects of $(\mathbf{x}_p - \mathbf{x}_p^H)$ on $(\mathbf{x}_h - \mathbf{x}_h^H)$. This formulation captures correlation over time among deviations, which arise from unobserved factors (such as weather conditions or special events) that are themselves correlated over time.

An important question that needs to be answered pertains to starting the process. It should be noted that historical O-D estimates or autoregressive factors might not be available while processing the first few days of data. While Equation 3.8 is clearly one of the best ways to obtain the target O-D flows, we need to explore other ways of obtaining $\mathbf{x}_h^a$ until we have sufficient data to estimate $\mathbf{x}_h^H$ and $\mathbf{f}_h^p$.

Consider the first day of data. We clearly need some reasonable flow estimates for the first interval. Let us denote by $\mathbf{x}_h^S$ a set of seed matrices of O-D flows. These matrices represent the flow estimates available at the beginning of the calibration process, and can be dynamic or static. A dynamic set of seed matrices will allow us to use a modification of Equation 3.3 to start the process:

$$\mathbf{x}_1^a = \mathbf{x}_1^S \tag{3.9}$$

The target flows for subsequent intervals can be generated by a process similar to Equation 3.5, but with the historical values replaced by the corresponding seed matrices. In general, the target flows could be expressed as:

$$\mathbf{x}_h^a = (\mathbf{x}_h^{H,\ d-1} / \mathbf{x}_{h-1}^{H,\ d-1}) \hat{\mathbf{x}}_{h-1}^d \tag{3.10}$$

where d is the day index. When $d = 1$, the seed flows $\mathbf{x}_h^S$ are used in place of the flows from day $d - 1$.

A static seed matrix, on the other hand, is more difficult to handle. Such a matrix would contain O-D flows cumulated over the entire period of interest (the AM peak, for example). We might then follow different approaches. One way is to break down the static flows into time-dependent flows:

$$x_{r,h}^S = (PF)_r x_r^{static} \tag{3.11}$$

where the starting (seed) O-D flow $x_{r,h}^S$ between O-D pair r departing during time interval h is obtained by scaling the corresponding static flow x_r^{static} by a peak factor $(PF)_r$. One way of obtaining the peak factors is to consider the fraction of flow recorded by sensors in interval h over the corresponding flow aggregated across the peak period:

$$(PF)_r = \frac{y_{s,h}}{\sum_{h' \in peak} y_{s,h'}} \tag{3.12}$$

where $y_{s,h}$ represents the flow counted at sensor location s during time interval h. The summation in Equation (3.12) is carried out over the time horizon of the static O-D matrix available. One might also have to determine an "average" peak factor based on several independent sensors that would measure flows arising from O-D pair

r. This procedure would depend on several factors such as network structure and spread, sensor locations, primary O-D flows and the number of alternative routes available to each O-D pair. An alternative approach would be to approximate from the static matrix a seed matrix of O-D flows $\mathbf{x}_1^S$ for the first time interval, and use Equation 3.4 for the remaining intervals in the study period. Once we process a few days of data in this manner, we would be able to generate some preliminary estimates of $\mathbf{V}_h$, $\mathbf{W}_h$ and $\mathbf{f}_h^p$, which would allow us to employ Equation 3.8 henceforth. Sections 3.3 and 3.4 discuss the issues involved in estimating these parameters.

3.3 Estimating Autoregressive Factors

The matrix $\mathbf{f}_h^p$ represents the effect of deviations in time interval p on the deviations in the current interval h. Let us denote the elements of this matrix by $\{f_{r,h}^{r',p}\}$, which captures the effect of the deviation in the r'th O-D flow of interval p on the deviation in the rth O-D flow of the current interval. Estimation of the matrix would be done element by element for each interval. The factor $\{f_{r,h}^{r',p}\}$ could be estimated through a regression of the form:

$$x_{r,h} - x_{r,h}^H = \sum_{p=h-q'}^{h-1} (f_{r,h}^{1,p}(x_{1,p} - x_{1,p}^H) + \cdots + f_{r,h}^{n_{OD},p}(x_{n_{OD},p} - x_{n_{OD},p}^H)) + w_{r,h} \quad (3.13)$$

where $w_{r,h}$ is the error term. It should be noted that n_{OD} such regressions would be needed in order to obtain the entire $\mathbf{f}_h^p$ matrix. Moreover, one would have to obtain such a matrix for each time interval h. In other words, each day of O-D flow data would yield exactly one observation for calibration. As a consequence, a large data set encompassing sufficient days of usable traffic records is critical to the calibration process. In the absence of such a large data set, one could effect simplifications to the structure of the autoregressive matrices and manage to estimate the factors. For example, it might be reasonable to assume that the structure of the autocorrelation remains constant with h. This would imply that the elements of $\mathbf{f}_h^p$ would depend only on the difference $h-p$, and not on the individual values of h and p. We can now write equations such as (3.13) for each interval within one day and have enough observations to estimate the elements of the matrix.

Some situations might allow a further simplification through the assumption that deviations in flow between O-D pair r would be affected primarily by the preceding deviations in flow *between O-D pair r alone*, and that contributions from other O-D pairs would be insignificant in comparison. The resulting regression would assume a much simpler form:

$$x_{r,h} - x^{H}_{r,h} = \sum_{p=h-q'}^{h-1} f^{r,p}_{r,h}(x_{r,p} - x^{H}_{r,p}) + w_{r,h} \tag{3.14}$$

and the resulting $\mathbf{f}^{p}_{h}$ matrix would be diagonal. However, one should be careful in using this approximation. While it is expected that $\mathbf{f}^{p}_{h}$ will be sparse, the relationship among the various O-D pairs will depend on several factors that include the structure of the freeway and arterial network, land use and geographical patterns. It might also be reasonable in some applications to assume that the O-D flows of interest belong to specific groups. An obvious basis for such a classification would be the division of the O-D flows into low, medium and high flow groups based on their magnitudes.

It should be noted that the autoregressive factors are often not readily available, and are an output of the calibration process. While the estimation process outlined in this section suggests approaches to obtaining these factors, it assumes a knowledge of the historical O-D flows $x^{H}_{r,h}$. The absence of reliable historical flow estimates in the initial stages of calibration applications might be handled by applying Equation (3.5) on the first few days of data. The autoregressive factors could then be estimated periodically, after designating the latest O-D flow estimates as historical.

3.4 Estimating Error Covariance Matrices

The GLS estimators represented in Equations (2.17) and (2.15) assume prior knowledge of the error covariance matrices $\mathbf{V}_h$ and $\mathbf{W}_h$. In most cases, however, this information is initially not available, and has to be estimated. Feasible Generalized Least Squares (or FGLS) provides a method by which the estimation of the error covariances is integrated with the O-D estimation step. While we present a brief review of the FGLS procedure, the reader is referred to Greene (2000) for a more comprehensive theoretical treatment.

Figure 3-1 outlines a sequential implementation of FGLS estimation that handles one day of data at a time. Further, the time intervals within each day are processed sequentially. The initialization step involves the definition of default covariance matrices for the error components. Theoretically, the FGLS procedure sets these unknown covariance matrices to the identity matrix, thereby reducing the estimation to one of Ordinary Least Squares (OLS). While such a naive initialization can be employed in general, we might be able to obtain a better starting point by incorporating some additional information about the sensor counts and target O-D flows. We illustrate this point by drawing from our discussion in Chapter 2 regarding the interpretation of the covariance matrix Ω_h. The assumption of a diagonal structure for $\mathbf{V}_h$ and $\mathbf{W}_h$ would result in a diagonal Ω_h, the elements along the diagonal representing the variances of the corresponding errors. The matrix Ω_h^{-1} would then be interpreted as a weight matrix associated with the indirect and direct measurements. Our default covariance matrices could now be constructed to reflect our *a priori* knowledge about the reliability of our measurements. In a calibration application, for example, it might be reasonable to assume that the counts reported by the surveillance system are more reliable than the seed O-D flows available. We might then assign higher default weights to the counts so as to discount the effect of the seed O-D flows. It should be noted that better covariance estimates would be obtained as the calibration process progresses across the multiple days of data available, thereby eliminating the need for the initialization step beyond the first few days[2].

Once past the initialization step, we proceed through the intervals in day i sequentially, starting from interval 1. The target O-D flows are computed for each interval using the approaches suggested in Section 3.2, and the Dynamic Network Loading model is used to generate an initial set of assignment matrices. The observed counts, simulated assignment matrices and latest covariance estimates are used as inputs by the O-D estimation model to generate new estimates of the O-D flows for each interval. At this stage, it is critical to ensure that the assignment matrices obtained by assigning the new estimated flows is comparable to those input to the O-D estimation model. We therefore iterate over the O-D flows until the assignment fractions converge, and proceed to the next time interval.

Once the flows for all intervals in day i have been estimated, one is in a position to estimate the error variances. This could be achieved by first computing the error (or residual) between measured and estimated quantities. The error in the sensor counts would be the difference between the actually observed counts and those simulated by the DTA system:

$$\mathbf{y}_h - \sum_{p=h-p'}^{h} \mathbf{a}_h^p \mathbf{x}_p$$

Likewise, the error in the O-D flows could be computed as the difference between the estimated flows and their corresponding target values:

$$\mathbf{x}_h^a - \mathbf{x}_h$$

If we denote the number of O-D pairs by n_{OD} and the number of links with sensors by n_l, we have a total of $(n_{OD} + n_l)$ variances to estimate. Let us first hypothesize that we have n data points for each of the $(n_{OD} + n_l)$ error terms. Denote by $\mathbf{e}_{ih}^x$ and $\mathbf{e}_{ih}^y$ the n-dimensional row vectors of residuals associated with the i^{th} direct and indirect measurements respectively. The $(i, j)^{th}$ elements of the matrices $\mathbf{W}_h$ and $\mathbf{V}_h$

[2]The latest covariance matrix is obtained from the historical database.

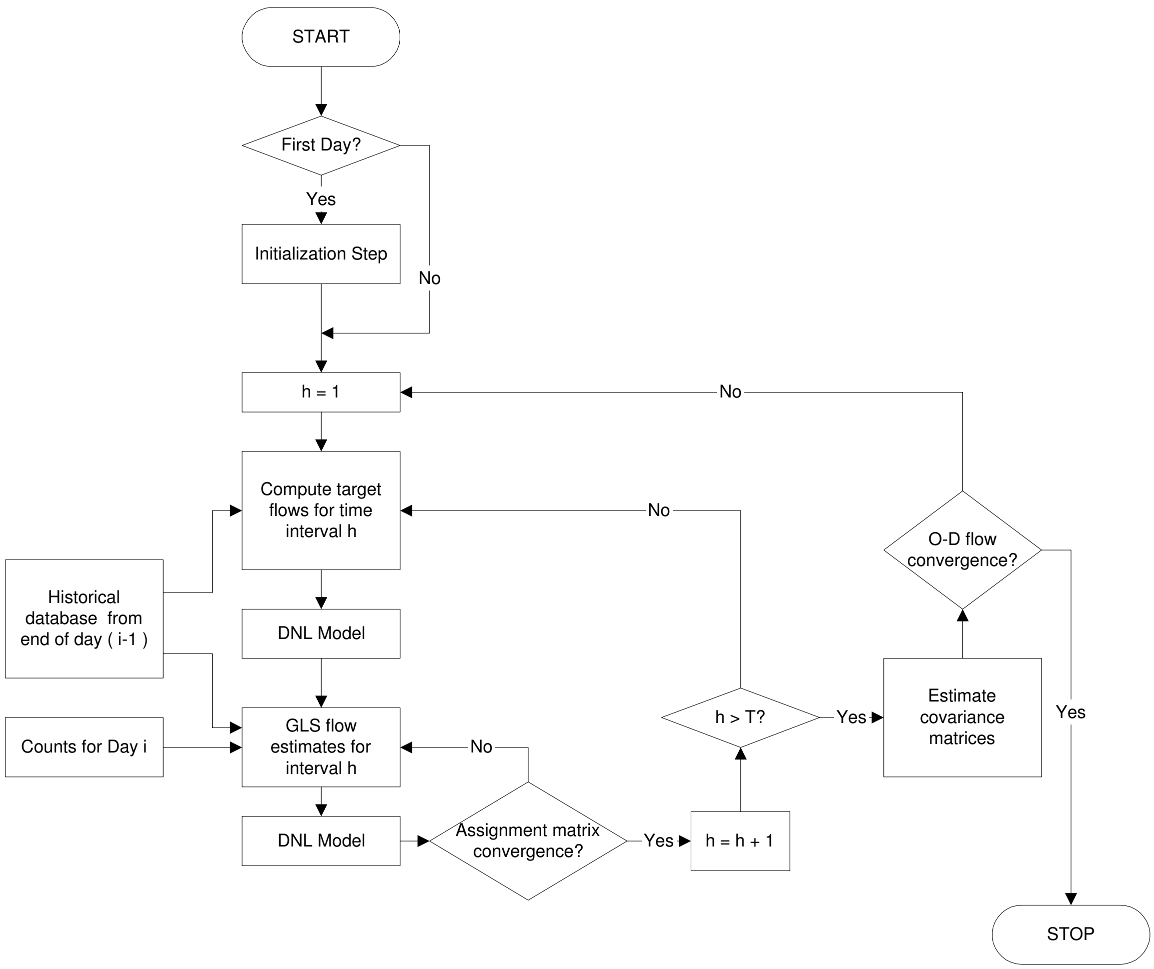

Figure 3-1: Sequential O-D and Covariance Estimation

could then be approximated by:

$$W_{ijh} = \mathbf{e}^x_{ih}\mathbf{e}^{x\,\prime}_{jh}/n$$
$$V_{ijh} = \mathbf{e}^y_{ih}\mathbf{e}^{y\,\prime}_{jh}/n$$

It should be noted that each day of data would provide exactly one observation for the error term associated with each of the measurement. While estimates of the error variances based on a single observation[3] can be inefficient, we might increase n through simplifying assumptions similar to those in Section 3.3. The first assumption would be to remove the dependence of $\mathbf{W}_h$ and $\mathbf{V}_h$ on h, thereby assuming that the error covariance matrices are invariant across the time intervals within the study period[4]. Each time interval would now yield an additional observation for every error term, thereby increasing the efficiency of the covariance estimates.

Another practical simplifying approach would be to assume that the measurements fall within a finite number of classes. This would allow us to separate the residuals into groups and estimate a single variance for each group. The criteria used for grouping the residuals could vary across the different types of measurements. The O-D flows, for example, could be grouped into high, medium and low categories based on their estimated values. A similar grouping could be used for the counts as well, though it might be more intuitive to segregate the counts based on the type and location of the sensor[5].

The algorithm represented by Figure 3-1 would terminate when the estimated O-D flows converge across iterations.

[3]A single observation would imply that $n = 1$.

[4]It might be more reasonable to expect constant covariances across a peak period. A more general approach using multiple days of data is presented in Section 3.5.

[5]One example would be to compute a different variance for freeway and arterial sensors.

3.5 Updating the Historical Database

The preceding sections discussed approaches to estimating the unknown autoregressive factors and error covariances. We presented several simplifying assumptions that would allow us to obtain efficient parameter estimates using estimation results from a single day of surveillance observations. We now look at the broader picture involving several days of data, and outline a framework for updating our historical database with the results of each additional day of estimations.

Figure 3-2 summarizes the essence of the update process. The calibration process is initialized by loading a set of default parameters into the constituent models. Subsequently, the available data encompassing several days is processed sequentially as described below. The surveillance data recorded on day i is read, and the O-D flows for each of the constituent time intervals is estimated according to the procedure outlined in Figure 3-1.

Iterations are now performed to identify an optimal set of parameters for the route choice model, so as to achieve the best fit for the observed sensor counts. The search algorithm used in this step can range from a simple line or grid technique (in the case of one or two parameters respectively), or can involve a more complex algorithm when dealing with a larger set of parameters. It is not necessary, however, to re-calibrate the route choice model parameters after each day of data. In fact, such a step might lead to widely fluctuating parameter estimates from day to day, and could have adverse effects on the final estimates output by the calibration process. A more reasonable approach would be to calibrate the route choice parameters once at the beginning of the process, and regularly re-calibrate them after a specific number of days of data have been processed. An efficient strategy at this stage might be to increase the frequency of re-calibrations as the number of days processed increases, thereby allowing the initial estimates to stabilize before intensive calibration begins. This logic is reflected in the dashed arrow in Figure 3-2 that allows the calibration process to bypass the route choice parameter calibration step for certain days.

A critical issue at this stage pertains to the methodology of updating the historical

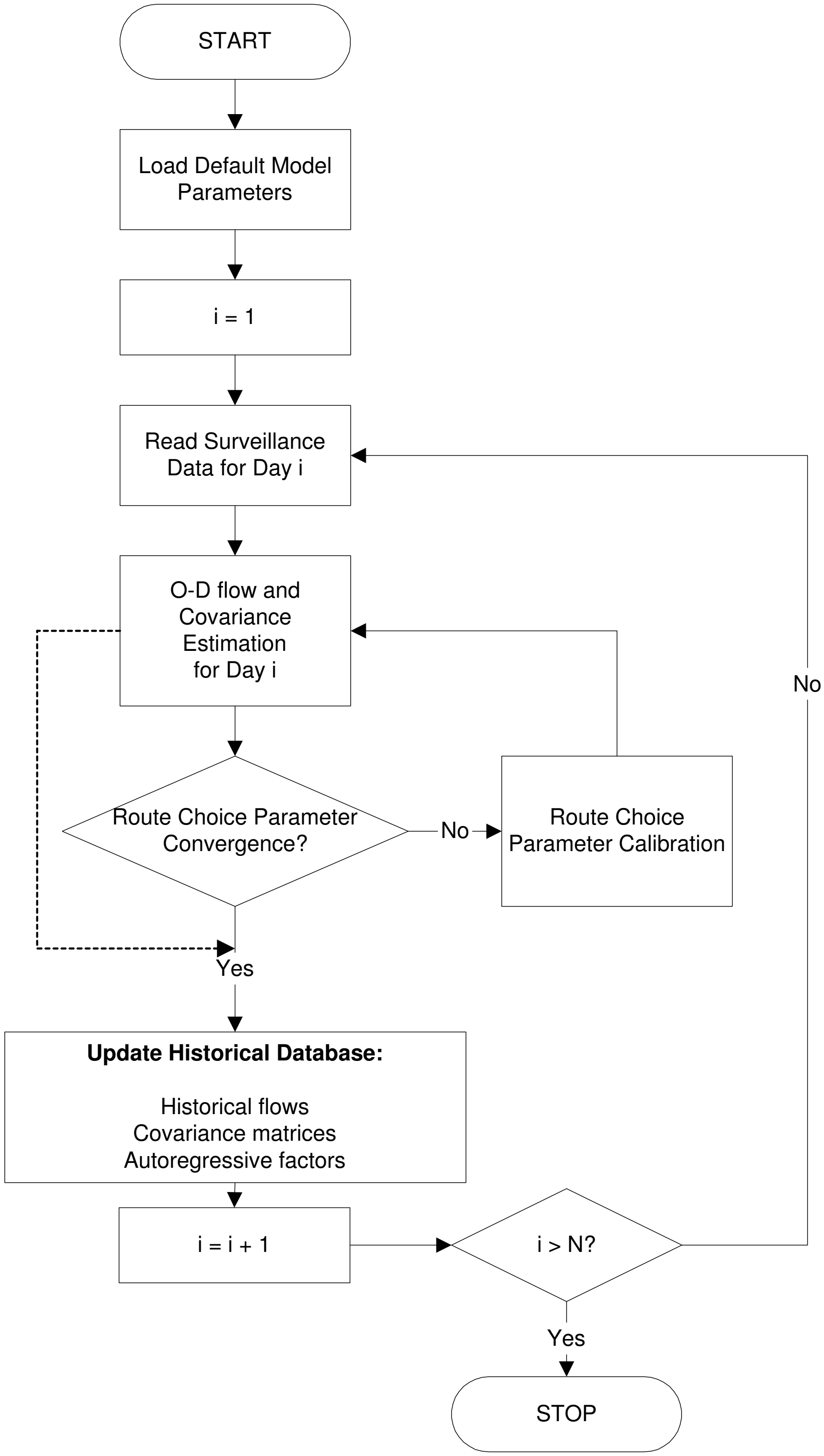

Figure 3-2: Parameter Update Methodology

database with the results of our latest day of estimations, and involves the following operations:

- Creation of new historical O-D flows.
- Creation of new error covariance matrices.
- Re-evaluation of the autoregressive factors.

The process of updating the historical O-D flows is significant in two different applications. From a calibration perspective, we aim to construct a database of O-D flows that would be stratified by various criteria such as the day of the week, type of weather and special events. Also, once calibrated, the DTA system would reside within a traffic management center and estimate and predict O-D flows in real-time. In both cases, we need to update our best historical flow estimates with the information contained in every new day of estimations. This updating process could be achieved in several ways. The simplest approach is to designate the latest flow estimates (those obtained during the last day) as our historical flows, as these estimates would encapsulate all prior history. One could also construct alternative updating methods that use estimates from several days. For example, a moving average of the last few estimates would reflect day-to-day trends in the O-D flows. A third alternative is to use a smoothing formula that might assume the following form:

$$x_{rh}^{H,d} = x_{rh}^{H,d-1} + \alpha(\hat{x}_{rh}^{d} - x_{rh}^{H,d-1}) \tag{3.15}$$

where $x_{rh}^{H,d}$ represents the historical flow between O-D pair r and departure time h after d days, $\hat{x}_{rh}^{d}$ is the corresponding estimate on day d, and α is a scalar between zero and one.

Updating the historical error covariance matrices based on new covariance estimates is a little more involved. In order to obtain statistically efficient estimates of the error covariances, one should ideally utilize the information contained in all days of data upto the current day. This could be achieved by re-estimating the O-D flows for days $1, 2, \ldots, d$ using the updated historical flows after day d as target flows for each day. The resulting residuals from the d days of estimations would then form the basis for computing the error covariances. The estimates obtained after processing all available days of data would represent an efficient synthesis of the information contained in the surveillance data set.

Once the historical O-D flows and error covariances have been updated, the approach to re-estimating the autoregressive factors is similar to that described in Section 3.3. As with the covariances, the efficiency of the autoregressive estimates could also be increased by using the information in more than one day of observations. Thus regressions similar to Equation (3.13) can be run using d days of data, again using the updated historical flows after day d to compute the O-D flow deviations.

3.6 Conclusion

In this chapter, we presented approaches to the creation of a historical database comprising of O-D flows, error covariance matrices and autoregressive factors. We outlined methods for estimating these parameters, and addressed the issue of starting the calibration process in the absence of reliable initial parameter estimates. Finally, we described a sequential framework that attempts to increase the efficiency of our estimates by jointly using multiple days of observations. In the remaining chapters, we use DynaMIT, a state-of-the-art DTA system, as an example to identify the primary calibration variables and explore the methodology proposed here through detailed case studies.

Chapter 4

The DynaMIT System

In previous chapters, we have laid out a methodology to calibrate the demand simulator within a general Dynamic Traffic Assignment system. In this chapter, we discuss the features and functionalities of one such application, and proceed to list the parameters we might wish to calibrate. This would form the basis for the case studies presented in subsequent chapters.

DynaMIT (Dynamic Network Assignment for the Management of Information to Travelers) is a state-of-the-art real-time computer system for traffic estimation and prediction, and the generation of traveler information and route guidance. DynaMIT supports the operation of Advanced Traveler Information Systems (ATIS) and Advanced Traffic Management Systems (ATMS) at Traffic Management Centers (TMC). Sponsored by the Federal Highway Administration (FHWA) with Oak Ridge National Laboratories (ORNL) as the program manager, DynaMIT is the result of intense research and development at the Intelligent Transportation Systems Program at the Massachusetts Institute of Technology.

4.1 Features and Functionality

The key to DynaMIT's functionality is its detailed network representation, coupled with models of traveler behavior. Through an effective integration of historical databases with real-time inputs from field installations (surveillance data and control logic of traffic signals, ramp meters and toll booths), DynaMIT is designed to efficiently achieve:

- Real time estimation of network conditions.
- Rolling horizon predictions of network conditions in response to alternative traffic control measures and information dissemination strategies.
- Generation of traffic information and route guidance to steer drivers towards optimal decisions.

To sustain users' acceptance and achieve reliable predictions and credible guidance, DynaMIT incorporates *unbiasedness* and *consistency* into its core operations. Unbiasedness guarantees that the information provided to travelers is based on the best available knowledge of current and anticipated network conditions. Consistency ensures that DynaMIT's predictions of expected network conditions match what drivers would experience on the network.

DynaMIT has the ability to trade-off level of detail (or resolution) and computational practicability, without compromising the integrity of its output. Its important features include:

- A microscopic demand simulator that generates individual travelers and simulates their pre-trip and en-route decisions (choice of departure time and route) in response to information provided by the ATIS.
- Estimation and prediction of origin-destination flows.
- Simulation of different vehicle types and driver behaviors.
- A mesoscopic supply simulator that explicitly captures traffic dynamics related to the development and dissipation of queues, spillbacks, and congestion.
- Time-based supply simulation that simulates traffic operations at a user-defined level of detail that facilitates real time performance. The level of detail could be determined by the choice of time steps and the level of aggregation of vehicles into homogeneous packets.

- Traveler information and guidance generation based on predicted traffic conditions to account for driver over-reaction to incident congestion. The system iterates between predicted network state, driver response to information and the resulting network state, towards the generation of consistent and unbiased information strategies.
- Adaptable to diverse ATIS requirements.
- Distinguishes between informed and uninformed drivers.
- Generates detailed vehicle trajectories by simulating each individual trip.
- Optimally uses historical, surveillance and O-D data to generate reliable O-D estimates in real-time. The system records the results from previous O-D estimations to update O-D databases.
- Uses a rolling horizon to achieve efficient and accurate real-time estimations and predictions.
- Handles real time scenarios including incidents, special events, weather conditions, highway construction activities and fluctuations in demand.
- Integrated with the MITSIMLab microscopic traffic simulator for offline evaluation and calibration.
- Ready for integration with external global systems (such as a TMC) using an external distributed CORBA interface, which allows for future adaptability and expansion.

4.2 Overall Framework

DynaMIT is composed of several detailed models and algorithms to achieve two main functionalities:

- Estimation of current network state using both historical and real-time information.

- Generation of prediction-based information for a given time horizon.

The estimation and prediction phases operate in a rolling horizon. This concept is illustrated with a simple example (Figure 4-1).

It is now 8:00am. DynaMIT starts an execution cycle, and performs a state estimation using data collected during the last 5 minutes. When the state of the network at 8:00 is available, DynaMIT starts predicting for a given horizon, say one hour, and computes a guidance strategy which is consistent with that prediction. At 8:07, DynaMIT has finished the computation, and is ready to implement the guidance strategy on the real network. This strategy will be in effect until a new strategy is generated. Immediately following that, DynaMIT starts a new execution cycle. Now, the state estimation is performed for the last 7 minutes. Indeed, while DynaMIT was busy computing and implementing the new guidance strategy, the surveillance system continued to collect real-time information, and DynaMIT will update its knowledge of the current network conditions using that information. The new network estimate is used as a basis for a new prediction and guidance strategy. The process continues rolling in a similar fashion during the whole day.

The overall structure with interactions among the various elements of DynaMIT is illustrated in Figure 4-2. DynaMIT utilizes both off-line and real-time information. The most important off-line information, in addition to the detailed description of the network, is a database containing historical network conditions. This database might combine directly observed data and the results of off-line models. The historical database contains time-dependent data, including origin-destination matrices, link travel times and other model parameters. Clearly, the richer the historical database, the better the results. Such a rich historical database requires substantial efforts towards data collection and careful calibration.

Real-time information is provided by the surveillance system and the control system. DynaMIT is designed to operate with a wide range of surveillance and control systems. The minimum real-time information required by DynaMIT is time-dependent link flows, incident characteristics (location, starting time, duration and severity), and traffic control strategies.

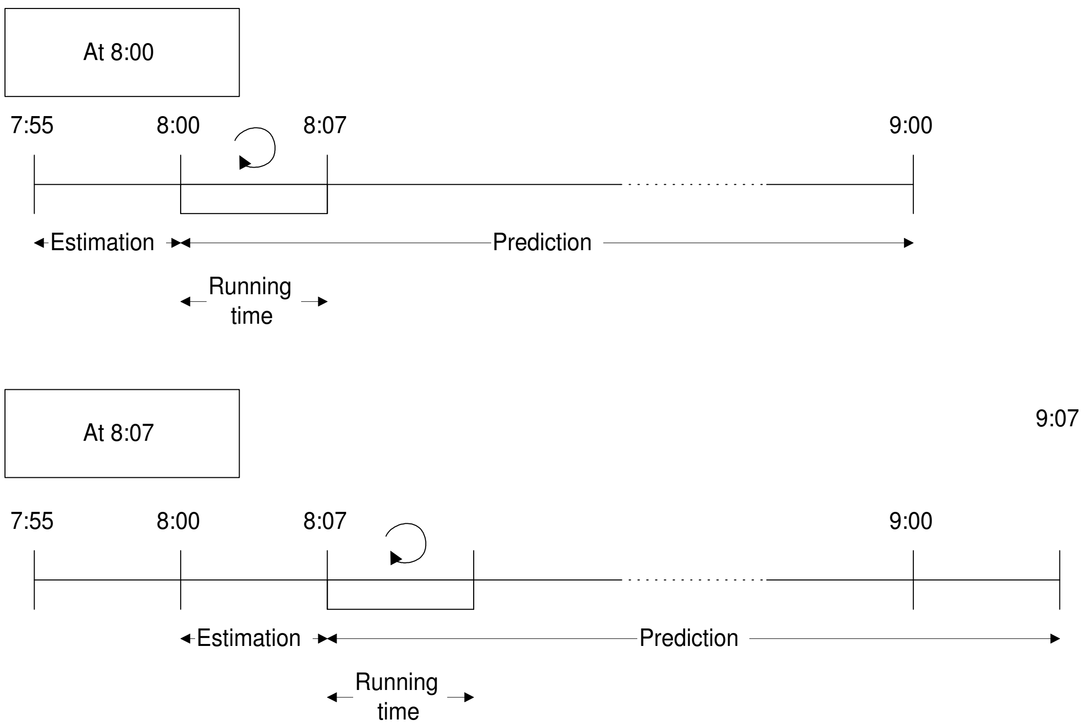

Figure 4-1: The Rolling Horizon

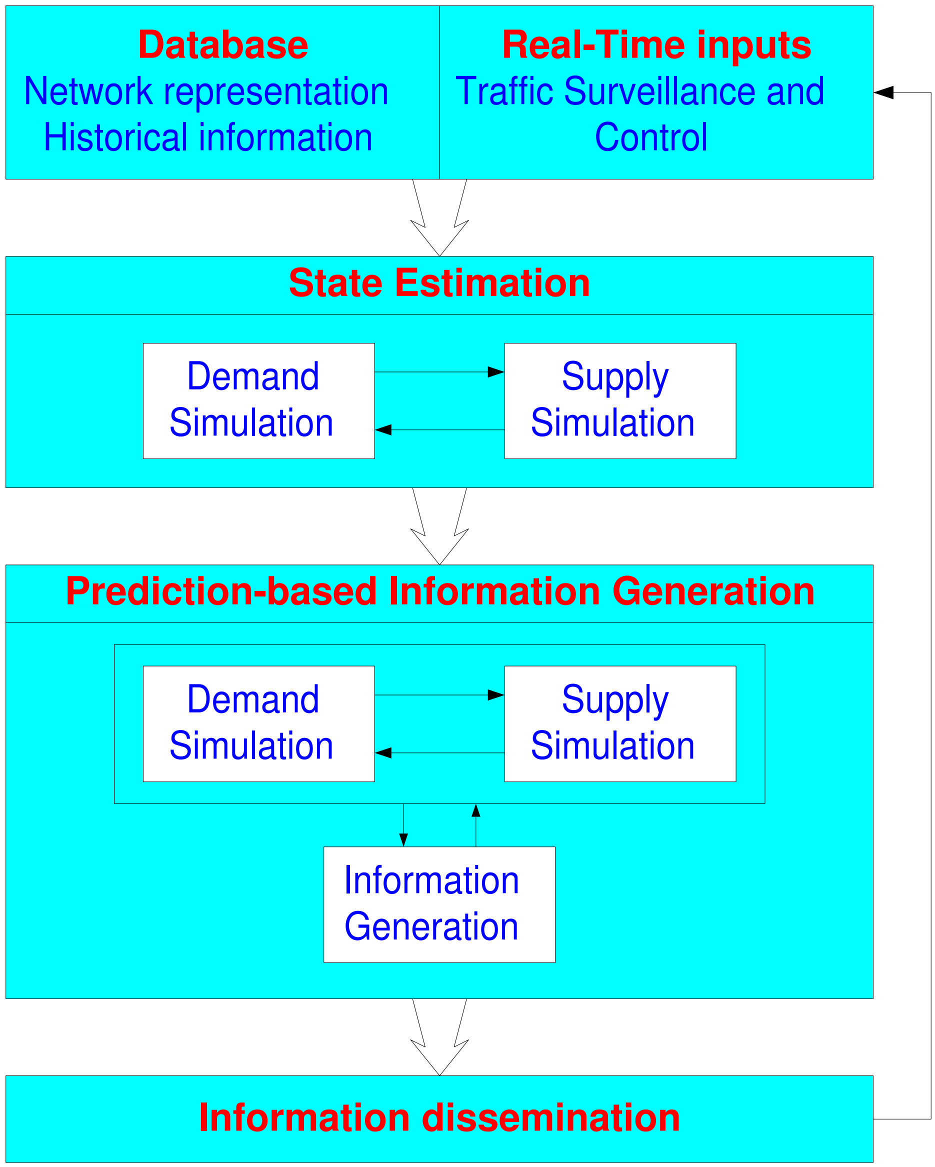

Figure 4-2: The DynaMIT Framework

4.2.1 State Estimation

The state estimation module provides estimates of the current state of the network in terms of O-D flows, link flows, queues, speeds and densities. This step represents an important function of DTA systems, since information obtained from the traffic sensors can vary depending on the type of surveillance system employed. In an ideal system where there is two-way communication between the traffic control center and every vehicle in the network, perfect information about the vehicle location and possibly its origin and destination can be obtained. While such perfect systems are possible in the future, most existing surveillance systems are limited to vehicle detectors located at critical points in the network. The information provided by these traffic sensors therefore must be used to infer traffic flows, densities and queue lengths in the entire network.

The main models used by the State Estimation module are:

- A demand simulator that combines real-time O-D estimation with user behavior models for route and departure time choice.

- A network state estimator (also known as the supply simulator) that simulates driver decisions and collects information about the resulting traffic conditions.

The demand and supply simulators interact with each other in order to provide demand and network state estimates that are congruent and utilize the most recent information available from the surveillance system (Figure 4-3).

4.2.2 Demand Simulation

Demand estimation in DynaMIT is sensitive to the guidance generated and information provided to the users, and is accomplished through an explicit simulation of pre-trip departure time, mode and route choice decisions that ultimately produce the O-D flows used by the O-D estimation model. The pre-trip demand simulator updates the historical O-D matrices by modeling the reaction of each individual to

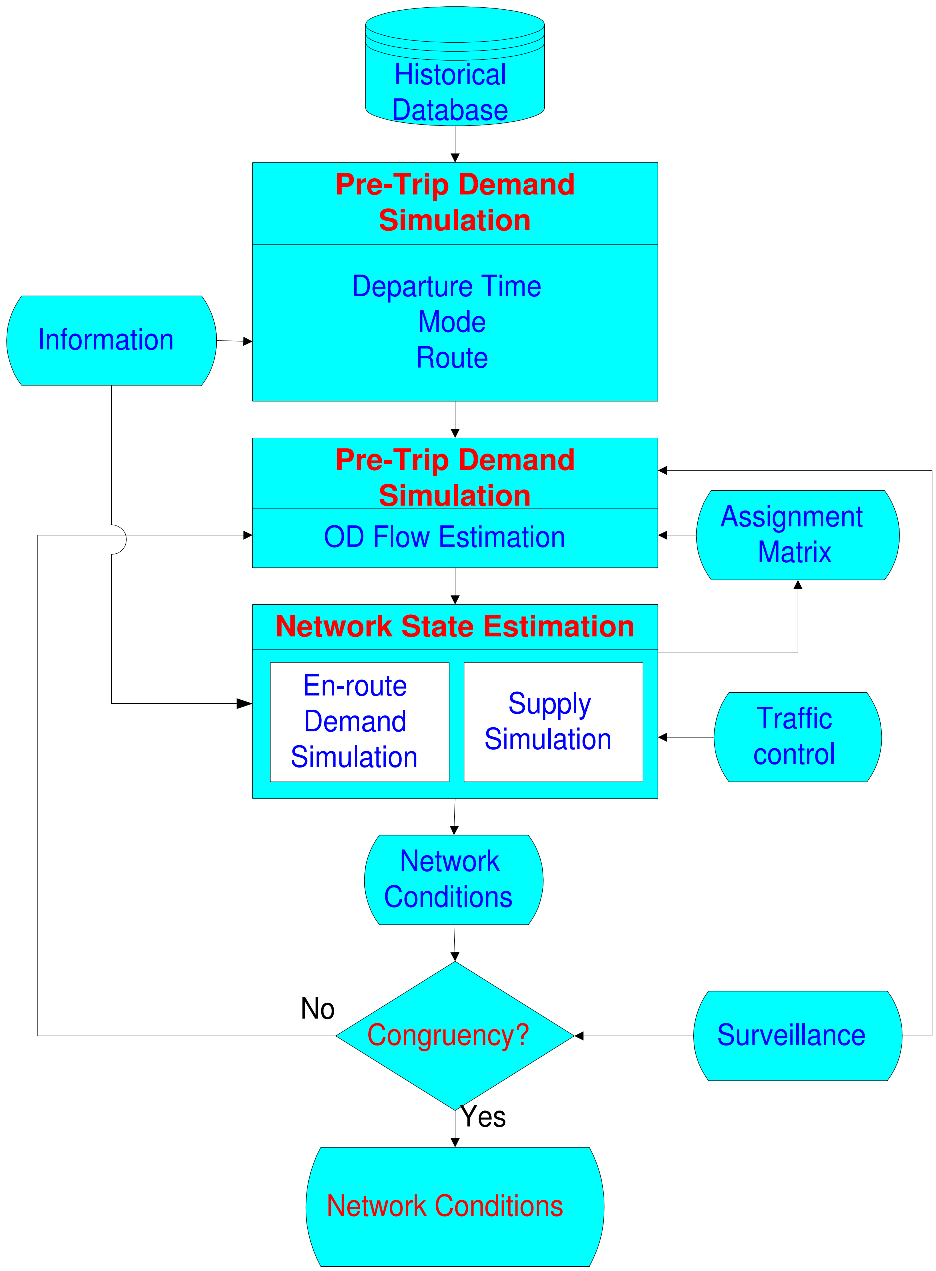

Figure 4-3: State Estimation in DynaMIT

guidance information. The consequent changes are then aggregated to obtain updated historical O-D matrices. However, these updated historical O-D flows require further adjustments to reflect the actual travel demand in the network. Reasons for the divergence of actual O-D flows from historical estimates include capacity changes on the network (such as the closure of roads or lanes), special events that temporarily attract a large amount of trips to a destination, and other day-to-day fluctuations. Consequently, one of the requirements for dynamic traffic modeling is the capability to estimate (and predict) O-D flows in real time. The O-D model uses updated historical O-D flows, real- time measurements of actual link flows on the network, and estimates of assignment fractions (the mapping from O-D flows to link flows based on route choice fractions and travel times) to estimate the O-D flows for the current estimation interval.

Note on O-D Smoothing

The fixed point nature of the O-D estimation procedure, coupled with the real-time requirements of a prediction-based DTA system, necessitates the use of an efficient solution scheme that would result in quick iterative convergence. The O-D estimation module within DynaMIT utilizes an algorithm similar to the Method of Successive Averages with Decreasing Reinitializations (MSADR)[1] to compute the target O-D flows for successive iterations. Stated mathematically,

$$\mathbf{x}^{k^*} = \hat{\mathbf{x}}^{k-1} + \alpha_k(\hat{\mathbf{x}}^k - \hat{\mathbf{x}}^{k-1}) \tag{4.1}$$

$$= \alpha_k\hat{\mathbf{x}}^k + (1-\alpha_k)\hat{\mathbf{x}}^{k-1} \tag{4.2}$$

where $\mathbf{x}^{k^*}$ is the new target O-D flow vector, $\hat{\mathbf{x}}^k$ and $\hat{\mathbf{x}}^{k-1}$ are the estimated flows from iterations k and $k-1$ respectively. The weighting parameter α_k is computed so

[1]See Section 2.1.1, or Cascetta and Postorino (2001).

as to accelerate the convergence of the iterative algorithm:

$$\alpha_k = \left[\frac{1}{ae^{-ka}}\right]\left[\frac{1}{\sum_{j=1}^{k}\frac{1}{ae^{-ja}}}\right] \tag{4.3}$$

where k is the iteration counter. The parameter a in the above expression assumes a default value of 1.0.

4.2.3 Supply Simulation

The network state estimator utilizes a traffic simulation model that simulates the actual traffic conditions in the network during the current estimation interval. The inputs to this model include the travel demand (as estimated by the demand simulator), updated capacities and traffic dynamics parameters, the control strategies implemented and the traffic information and guidance actually disseminated. The driver behavior model captures the responses to ATIS in the form of en route choices.

4.2.4 Demand-Supply Interactions

One of the inputs to the O-D estimation model is a set of assignment matrices. These matrices map the O-D flows from current and past intervals to link flows in the current interval. The assignment fractions therefore depend on the time interval, and also on the route choice decisions made by individual drivers. The flows measured on the network are a result of the interaction between the demand and supply components. It may be necessary to iterate between the network state estimation and the O-D estimation models until convergence is achieved. The output of this process is an estimate of the actual traffic conditions on the network, and information about origin-destination flows, link flows, queues, speeds and densities.

4.2.5 Prediction and Guidance Generation

The prediction-based guidance module (Figure 4-4) consists of several interacting steps:

- Pre-trip demand simulation
- O-D flow prediction
- Network state prediction
- Guidance generation

The O-D prediction model uses as input the aggregated historical demand adjusted by the pre-trip demand simulator to account for departure time, mode and route choices in response to guidance, and provides the required estimates of future O-D flows. The network state prediction function undertakes the important task of traffic prediction for a given control and guidance strategy and predicted set of O-D flows, using the current network condition estimated by the state estimation module as a starting point. The performance of the network in the prediction horizon is evaluated using a traffic simulation model and en-route behavioral models.

The traffic information and guidance generation function uses the predicted traffic conditions to generate information and guidance according to the various ATIS in place. Traffic control is loosely coupled with DynaMIT in the current version of the system. Control strategies are assumed to be generated outside the DTA system, using the predictions as an input.

The generated traffic information and guidance must be consistent and unbiased. Under such conditions, there would be no better path that a driver could have taken based on the provided information. An iterative process is employed in order to obtain guidance that satisfies these requirements. Each iteration consists of a trial strategy, the state prediction (comprising both demand prediction and network state prediction) under the trial strategy, and the evaluation of the predicted state for consistency. Since, in general, the updated historical O-D flows depend on future guidance and information, the update of the historical O-D flows (using the departure time and mode choice models) and the O-D prediction models are included in the iteration. This general case represents the situation where pre-trip guidance is available to the drivers. In the special case where only en-route guidance is available, the pre-trip

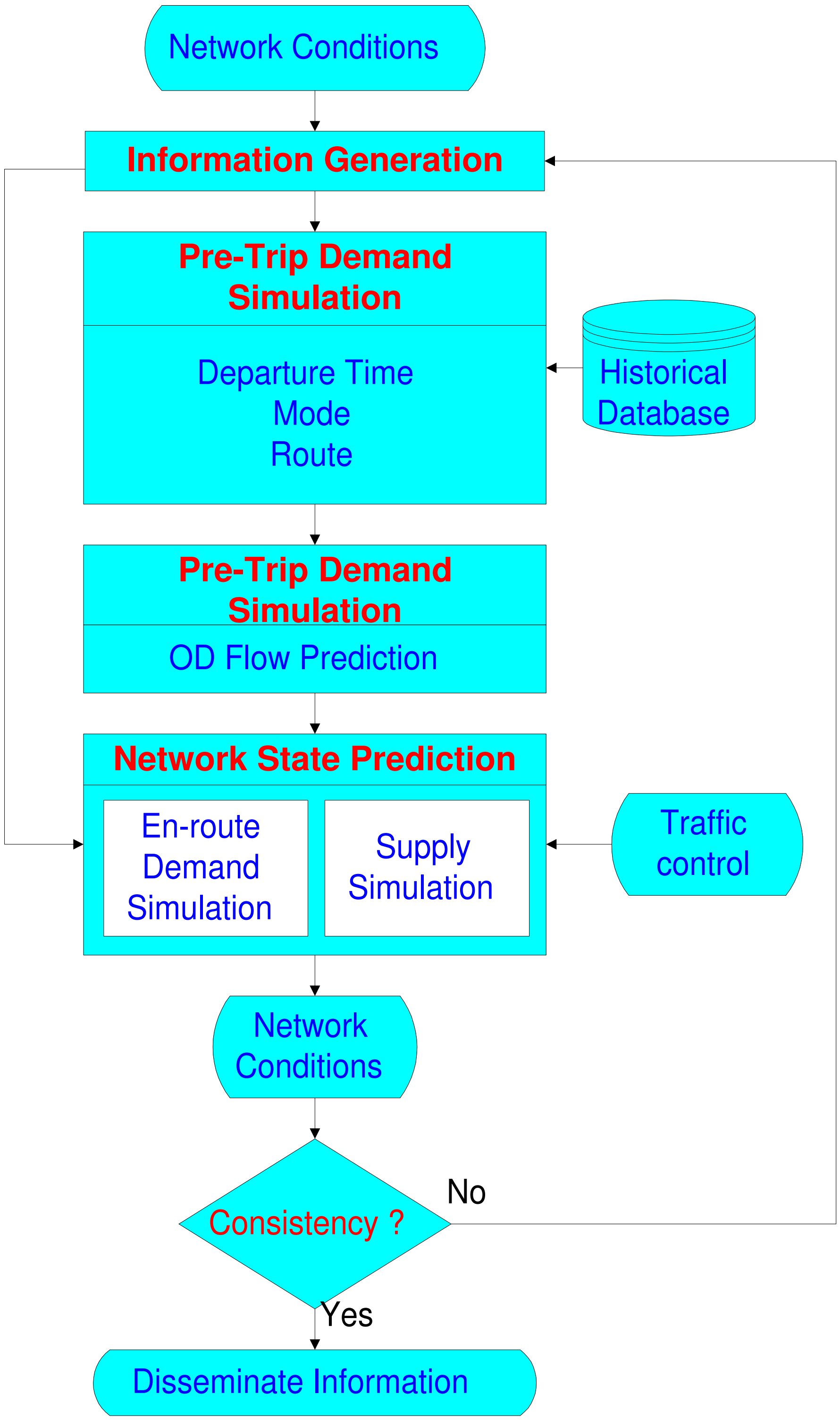

Figure 4-4: Prediction and Guidance Generation in DynaMIT

demand simulator is bypassed in the iterations. The initial strategy could then be generated from the prediction and guidance generation of the previous period.

4.3 DynaMIT for Planning

Apart from its real-time applications, DTA has the potential to significantly improve the transportation planning process for networks with congested facilities. DynaMIT-P is a DTA-based planning tool developed at MIT that is designed to assist planners in making decisions regarding proposed investments and operational changes in local and regional transportation networks. DynaMIT-P efficiently adapts the modules contained in the real-time DynaMIT system to suit offline planning applications.

Travel-related choices vary with regard to the time horizon over which the decisions are made. Individuals make long-term, short-term and within-day travel decisions (Figure 4-5). Long-term mobility decisions could include choices related to residential location and auto ownership. Short-term (or day-to-day) travel decisions include choice of trip frequency, destination, departure time, mode and route. Adjustments in short-term decisions are made in response to changes in long-term decisions (like auto ownership) and changes in the network. Individuals form habitual travel patterns that they follow regularly. Within-day decisions capture deviations from the habitual travel patterns. These deviations could be in response to real-time information, unusual weather conditions, incidents, or other special events.

DynaMIT-P focuses on modeling the short-term and within-day travel decisions, assuming that the long-term decisions are given. The inputs to DynaMIT-P include the potential users of the system, their demographic characteristics, residential location, etc. The output is the performance of the transportation system in terms of consumption of resources and benefits. Several important characteristics distinguish DynaMIT-P from traditional planning approaches:

- Microsimulation ensures accurate depiction of individual traveler behavior.
- Detailed modeling of spillbacks, queue formation and dissipation captures the

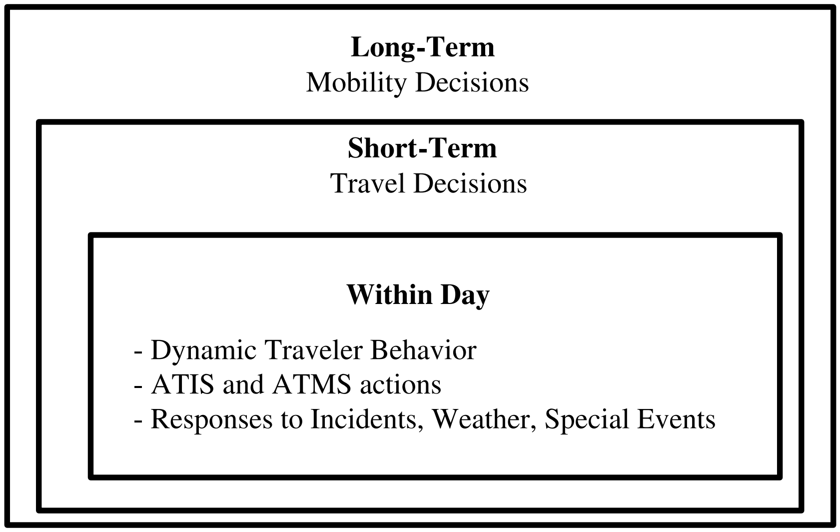

Figure 4-5: Framework for Travel Behavior

essence of network dynamics.

- Sensitivity to ATMS/ATIS facilitates the evaluation of ITS strategies.
- Time-dependent interactions between the demand and supply components presents a realistic picture of equilibrium.

DynaMIT-P employs three main components to achieve the functionality described above:

- The supply simulator
- The demand simulator
- The day-to-day learning model

The supply simulator is a mesoscopic traffic simulation model. For a given set of travelers and control strategies, it predicts the performance of the network by measuring time-dependent flows, travel times, queue lengths, etc. The simulator is designed to operate at different levels of granularity, depending on the requirements of each application. The main elements of the demand simulator are the O-D matrix estimation and the behavioral models. The O-D estimation model takes link counts and historical O-D flows as inputs, and produces an updated time-dependent O-D matrix to match the observed counts. The behavioral models are used to predict the travel behavior of individual travelers as a function of network level of service characteristics, perceptions and past experiences, information access and socioeconomic characteristics. Driver behavior is modeled using the path-size logit model (PS-Logit, Ramming Ramming (2001)), which is an extension of C-Logit (Cascetta, Nuzzolo, Russo and Vitetta (1997)). This model accounts for the degree of overlap among alternative routes while simulating individual route choice. The day-to-day learning model updates travelers perceptions of travel times based on past experiences and expectations, according to the following model:

$$\bar{T}_k^t = \lambda T_k^{t-1} + (1-\lambda)\bar{T}_k^{t-1} \tag{4.4}$$

where $\bar{T}_k^t$ is the expected time-dependent travel time along path k on day t, and T_k^t is the time-dependent travel time experienced along path k on day t. λ captures the learning rate, and may vary across market segments. The value of λ lies between 0 and 1, and is affected by the use of ATIS.

The demand and supply simulators interact with the learning models in a systematic way to capture both the day-to-day and within-day (short-term) demand-supply interactions (Cantarella and Cascetta (1995)). The structure of the short-term dynamics module is shown in Figure 4-6. The model is based on an iterative process between the demand and supply simulators. The main input to short-term dynamics is an O-D matrix of potential travelers. The demand simulator uses the corresponding behavioral models to update their frequency, destination, departure time, mode, and route, choices. The travelers are then loaded onto the supply simulator and a new network performance is obtained. Based on the learning model, travelers update their decisions in response to the observed level of service and network performance. When convergence between supply and demand is reached, the process ends. The output of the short-term dynamics component is the travelers' habitual travel behavior.

The purpose of the within-day dynamics model is to evaluate the performance of the transportation network in the presence of stochastic factors such as unusual weather, incidents, and special events (music, sports, etc.), which could substantially affect traffic conditions as compared to the usual conditions. The habitual travel behavior, obtained from the short-term dynamics, is input to the within-day model. Figure 4-7 summarizes the interactions among the different elements of the within-day dynamics component.

The outputs from both the short-term and within-day behavior components are used to generate the desired resource consumption and benefits (such as total savings in travel delays, costs, revenues, air pollution, safety, fuel consumption, etc.) DynaMIT-P's open system of demand models, detailed representation of network dynamics, and flexible structure make it a useful tool for a host of planning applications:

- Impact studies of Work Zone Activity, and minimum-impact work zone scheduling
- Special Events
- High-occupancy Vehicle (HOV) and High-occupancy Toll (HOT) facilities
- Congestion Pricing strategies
- Effectiveness of ATMS and ATIS

While both DynaMIT and DynaMIT-P capture the interaction between the O-D estimation and route choice model components within the DTA system, DynaMIT-P is better suited for calibration purposes. This stems primarily from DynaMIT-P's functionality to compute equilibrium network travel times. Network equilibrium for a given travel demand level results from a balance between the demand and supply elements. The fixed point nature of the calibration problem constrains us to require equilibrium at each calibration stage in order to ensure that the route choice fractions, assignment matrices and estimated O-D flows are consistent in each iteration.

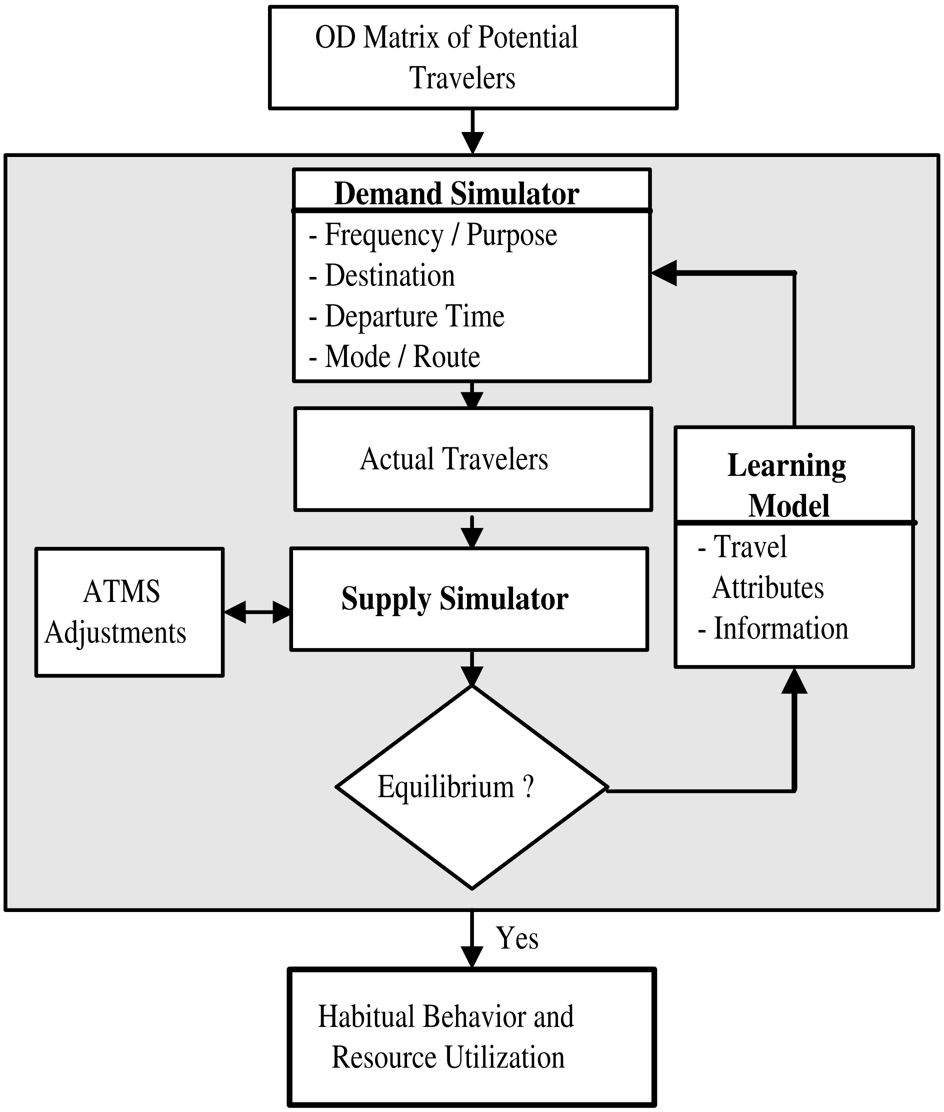

Figure 4-6: Short-Term Dynamics

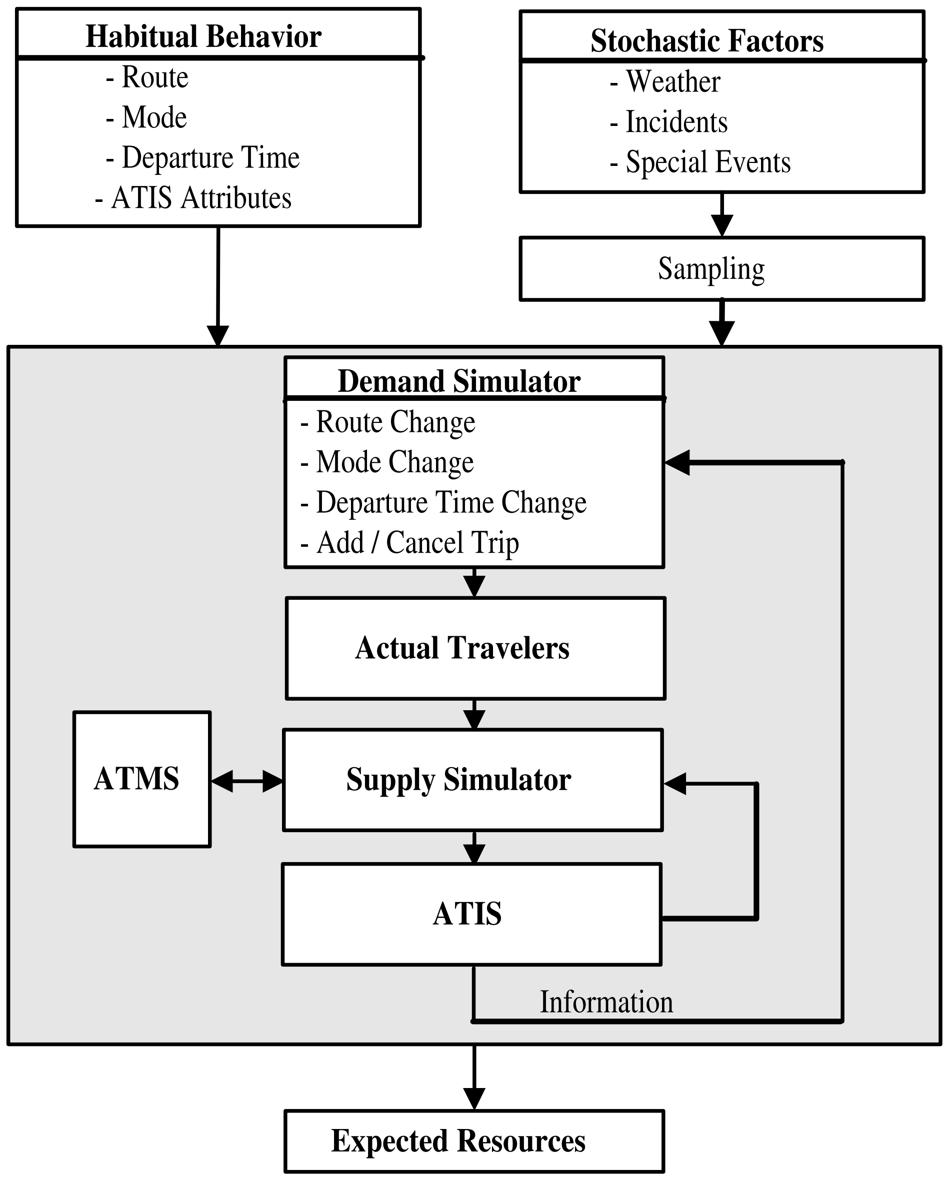

Figure 4-7: Within-Day Dynamics

4.4 Calibration Variables

In this section, we enumerate the various model parameters that need to be calibrated within the DynaMIT system. We separate the calibration variables into demand-side and supply-side parameters, and briefly review the role played by each in the overall DTA framework.

4.4.1 Demand Simulator Parameters

The demand simulator is primarily comprised of the driver behavior model and the O-D estimation and prediction model.

Parameters in the Route Choice Model

DynaMIT relies on discrete choice models to model driver decisions. This process involves three steps:

- Generate a choice set of feasible route alternatives for each O-D pair.
- Compute the utility of each alternative, and evaluate a choice probability for each alternative.
- Simulate the computed probability through a random draw to determine the "chosen" alternative.

The choice set generation step involves the computation of a good set of feasible paths connecting every O-D pair of interest. This is a critical step that controls the outcome of the entire calibration exercise. While the shortest set of paths between every O-D pair might capture driver behavior under uncongested traffic conditions, rising congestion levels can increase the attractiveness of previously unchosen routes. Also, incidents can cause blocks in the shortest route and force drivers onto less attractive paths. A good set of paths for each O-D pair is therefore essential in traffic estimation and prediction.

DynaMIT employs three steps in its path generation algorithm. The **shortest path computation** step generates the shortest path connecting each link in the network to all defined destination nodes. This set represents the most probable paths chosen by drivers under uncongested conditions. A **link elimination step** augments the paths from the shortest path set with alternative paths. This step involves the elimination of each link in the network and the subsequent re-computation of the shortest path, and ensures that an incident on any link will still leave alternative paths open for every O-D pair. A further **random perturbation** step is performed in order to obtain a richer path set. The impedances of the links are perturbed randomly to simulate varying travel times. Another set of shortest paths are now computed, and appended to the existing set. The number of random perturbations performed could be controlled by the user. The algorithm also screens the final path set for uniqueness, and eliminates unreasonably long paths.

Once the choice set has been defined, DynaMIT is ready to compute the probability with which an individual might choose each of the available alternatives. Utility theory along with the Path-Size Logit (PS-Logit) model is employed to evaluate these probabilities. Stated mathematically,

$$P_n(i) = \frac{e^{V_i + lnPS_i}}{\sum_{j \in C_n} e^{V_j + lnPS_j}} \tag{4.5}$$

where $P_n(i)$ is the probability of individual n choosing alternative i, V_i is the utility of alternative i, PS_i is the size of path i, and C_n denotes the choice set for individual n.

The utility V_i of each path is a function of several explanatory variables including attributes of the alternatives (such as travel times) and socio-economic characteristics of the individual driver. The size of a path is defined as (Exponential Path-Size formulation, Ramming (2001)):

$$PS_{in} = \sum_{a \in \Gamma_i} \left(\frac{l_a}{L_i} \right) \frac{1}{\sum_{j \in C_n} \frac{L_i^{\gamma}}{L_j^{\gamma}} \delta_{aj}} \tag{4.6}$$

where l_a is the length of link a, L_i is the length of path i and δ_{aj} takes the value 1 if link a is a part of path j (and is zero otherwise). The inner summation is computed over all paths in choice set C_n, while the outer summation is over all links a in path Γ_i. The parameter γ is an exponent in the model.

We list below the key calibration parameters for the route choice model:

- Parameters in the path choice set generation algorithm.

- Parameters in the path utility specification

- Path-size exponent γ.

The aggregate nature of the observed surveillance data[2] necessitates that the calibration of the disaggregate route choice model parameters be undertaken at an aggregate level. Given this constraint, it might be advantageous in some applications to estimate just a few factors that would scale initial parameter estimates uniformly. This approach obviously requires the availability of a good default set of route choice model parameters. An alternative approach would be to focus on the calibration of a subset of key route choice coefficients while holding all other parameters in the model fixed at their initial values.

Parameters in the O-D Estimation/Prediction Model

The current version of DynaMIT employs a sequential GLS-based O-D estimation module similar to Equation (2.14). The external inputs to the model include link counts and the historical database of O-D flows. The historical flows are coupled with the concept of flow deviations in order to effectively capture the information contained in the past estimates. The key inputs generated internally are the time-dependent assignment matrices. While the matrices $\mathbf{a}_h^p$ are generated by the supply simulator, the historical database would have to be created offline. The O-D estimation and prediction algorithm is also based on an autoregressive process that captures spatial

[2] Counts and speeds, for example, might be cumulated or averaged over a time interval covering several minutes.

and temporal correlations between the O-D flows. We list below the key calibration parameters for the O-D estimation and prediction model:

- The historical database of O-D flows, $\mathbf{x}_h^H$.
- The variance-covariance matrix $\mathbf{V}_h$ associated with the indirect measurement errors.
- The variance-covariance matrix $\mathbf{W}_h$ associated with the direct measurement errors.
- The matrices $\mathbf{f}_h^p$ of autoregressive factors.

4.4.2 Supply Simulator Parameters

The focus of this thesis is the calibration of the demand simulator within a DTA system. However, the close link between the demand and the supply components necessitates a joint calibration of all the associated model parameters. Though we will not undertake a detailed discussion involving the supply models, we present here a brief overview of the relevant supply-side parameters that would enter the overall calibration framework.

The supply simulator obtains aggregate measures of network performance by simulating the movement of drivers on the road network. Detailed mesoscopic models capture traffic dynamics and accurately model the build-up and dissipation of lane-specific queues and spillbacks. The links in the network are subdivided into segments to capture changing section geometries. Further, the lanes within each segment are grouped into lane groups to account for turning-movement-specific capacities at diversion and merge points and intersections.

Each segment contains a moving part (with vehicles moving at certain speeds), and a queuing part. The movement of vehicles in the moving part are governed by macroscopic speed-density relationships that take the following form:

$$v = v_{max} \left[1 - \left(\frac{k - k_{min}}{k_{jam}} \right)^{\beta} \right]^{\alpha} \tag{4.7}$$

where v is the speed of the vehicle (in mph), v_{max} is the speed on the segment under free-flow traffic conditions, k is the current segment density (in vehicles/mile/lane), k_{min} is the minimum density beyond which free-flow conditions begin to break down, k_{jam} is the jam density, and α and β are segment-specific coefficients. In addition, the speeds computed using Equation 4.7 are subject to a segment-specific minimum speed v_{min}.

The movement of vehicles from one segment to the next is governed by a host of capacity calculations. The primary quantities of interest are the input and output capacities of the various segments. These capacities are compared with the available physical space on the downstream segments before allowing vehicles to cross segment boundaries. A constraint on either capacity or space would cause vehicles to queue. An important calibration step is therefore the computation of segment (more specifically, lane group) capacities that truly approximate the allowed turning movements, signal logic and sectional geometry of the network.

We list below the key calibration variables on the supply side:

- Segment-specific speed-density parameters (v_{max}, k_{min}, k_{jam}, β, α and v_{min}).
- Lane group capacities on freeway and arterial segments.
- Lane group capacities at intersections, based on signal control logic.

4.5 Conclusion

In this chapter, we reviewed the features and functionalities of DynaMIT and DynaMIT-P, two state-of-the-art DTA systems for traffic estimation, prediction and planning applications. We outlined the critical models within the demand and supply modules of the DTA system, and identified key variables that need to be calibrated before the systems can be employed for real-time traffic management or planning applications. These variables were classified into demand and supply side parameters to facilitate a sequential approach to overall DTA system calibration. In the following chapter, we discuss the calibration of DynaMIT and DynaMIT-P using the methods and concepts outlined in Chapter 2, and present results from demand calibration and validation studies carried out on a real-sized traffic network.

Chapter 5

Case Studies

In Chapter 2, we presented an iterative scheme for the joint calibration of the route choice and O-D estimation/prediction model parameters within a DTA system. The objective of this chapter is to utilize data recorded by a real traffic surveillance system to test the performance of the proposed approach. The results presented in this chapter are derived from case studies based on a large-sized study network in Irvine, California. The DynaMIT-P system discussed in Chapter 4 is used as the calibration system.

We begin by briefly describing the study network and the data available for the calibration study, and proceed to outline the calibration approach as applied to this data set. We then discuss several assumptions that were employed to work around practical constraints and considerations, and present results from calibration and validation tests that verify the performance and feasibility of the methodology we proposed in Chapter 2. We conclude by summarizing our results and findings.

5.1 The Irvine Dataset

The data used in this research was collected from Irvine, a part of District 12 in Orange County, California, USA. We describe briefly the main features of this network and the surveillance data recorded.

5.1.1 Network Description

The study network (Figure 5-1) is comprised of three major freeways and a dense network of arterial segments. The I-5 and I-405 Interstates, along with State Route 133, define a wedge that is criss-crossed by several major arterials. The city of Irvine is located in Orange County, just outside Los Angeles. It therefore lies along the heavily traveled corridor connecting Los Angeles and San Diego. It is a major commercial and business center, and serves as an important regional airport. Irvine is also home to several schools and universities. The city therefore attracts a varied mix of commuters and travelers.

Figure 5-1: The Irvine Network

The network is represented as a set of 298 nodes connected by 618 directed links. These links represent the physical links on the network, and are further subdivided into 1373 segments to model changing link section geometry. Almost all of the 80 intersections within the study area are signalized, and are controlled by vehicle-actuated signal logic. A high fraction of the signals along the primary arterials (Barranca Parkway, Alton Parkway and Irvine Center Drive) are coordinated to minimize the number of stops.

5.1.2 Data Description and Analysis

The available data was derived primarily from four sources:

- PARAMICS network files
- O-D flows from OCTAM planning study
- Time-dependent detector data
- Signal timing and coordination plans

Information regarding the network was contained in a set of input files created for the PARAMICS traffic simulation system. These files included descriptions of network geometry, link and lane connectivity, sensor locations and signal phase timing plans.

The OCTAM planning study generated a static matrix of O-D flows covering the morning peak period. While this matrix contained every possible origin-destination combination from 61 zones, several of these flows were zero. In this study, a subset of 655 primary O-D pairs with non-zero flows was extracted from the static matrix. A review of the static O-D matrix showed that freeway-based O-D pairs contributed to a major proportion of the total demand. Figures 5-2 and 5-3 indicate the primary O-D pairs. The thickness of the lines connecting the origins and destinations is a measure of the magnitude of the O-D flow.

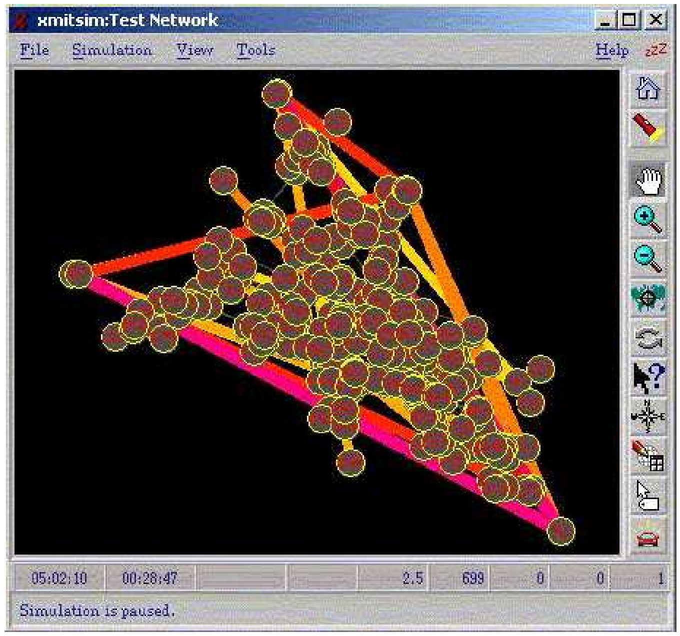

Figure 5-2: Primary O-D Pairs

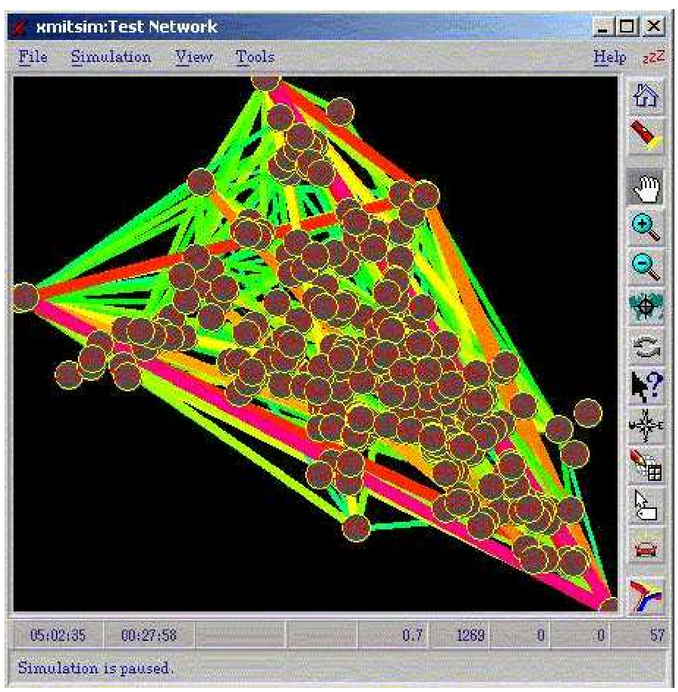

Figure 5-3: Primary O-D pairs

Time-varying freeway and arterial detector data recorded over 5 working days was available from California Department of Transportation (Caltrans). This data consisted of counts and occupancies measured by lane-specific sensors on freeway links and lane-group-specific sensors on arterial links. While the freeway detectors reported data every 30 seconds, the arterial detectors aggregated the data by 5 minute time slices. It should be noted that only 68 out of 225 sensors reported usable data. The remaining detectors were either situated outside the study area, or their data files were inconsistent. The 68 usable sensors were split into 30 situated on freeway and ramp links, and 38 on arterial links.

The data collected by the surveillance system consisted of vehicle counts and occupancies. The counts were aggregated into common intervals of length 15 minutes. Time-varying densities were approximated from the occupancies by assuming a mean vehicle length of 5 metres. Further analysis indicated that there was not much day-to-day variability in the sensor data received, even though these 5 days fell on different days of the week (Figures 5-4 and 5-5).

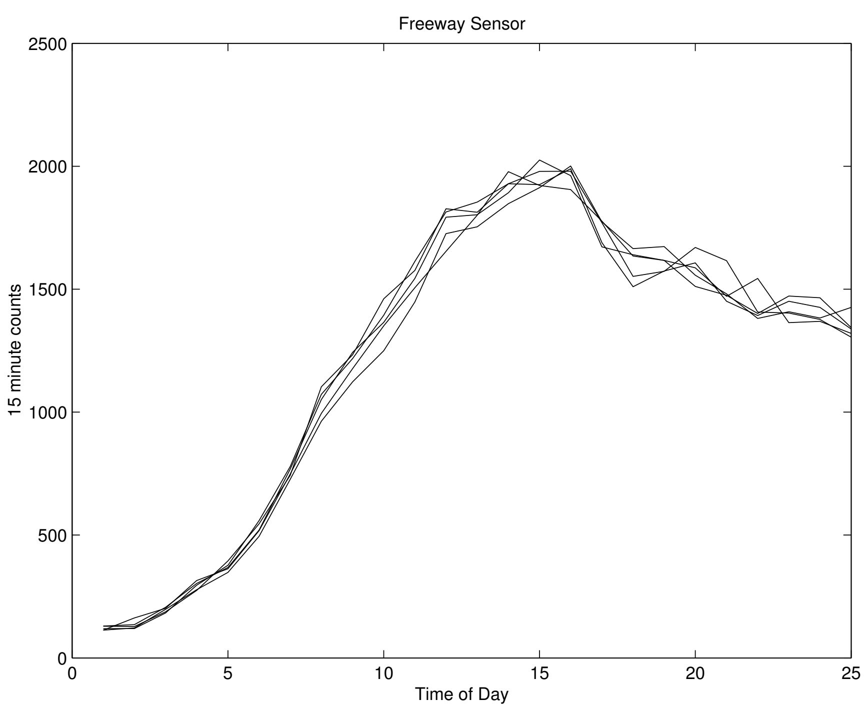

Figure 5-4: Counts Variation Across Days: Freeway Sensor

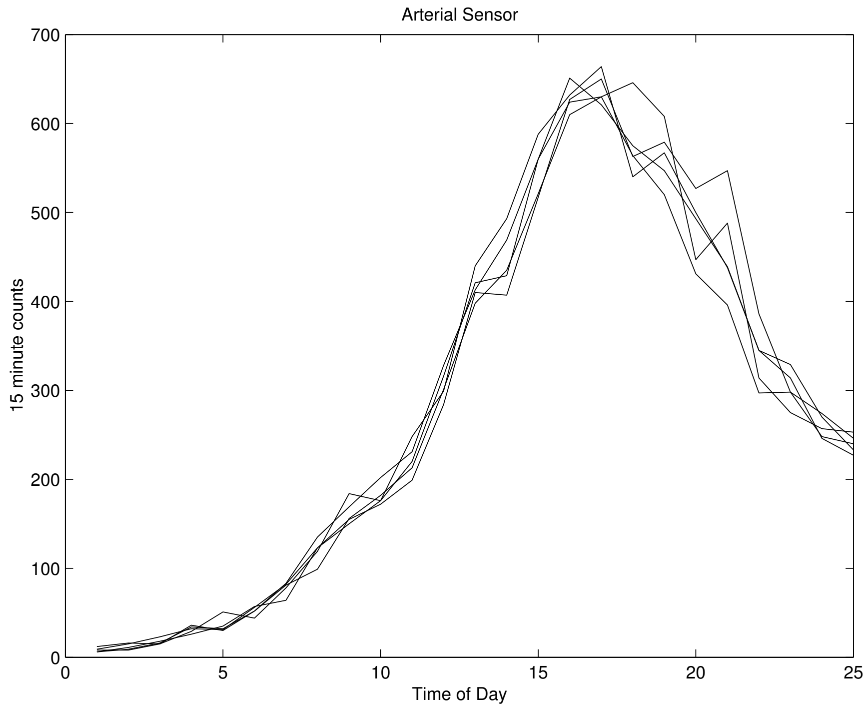

Figure 5-5: Counts Variation Across Days: Arterial Sensor

Signal timing and coordination charts from the City of Irvine specified the details regarding signal phasing, timing, actuation and coordination.

5.1.3 DynaMIT Input Files

Let us briefly review the input files required by the DynaMIT system. The **network file** describes the locations of nodes, links, segments and sensors, and defines the connectivity between individual lanes in the network. The network information contained in the PARAMICS data set was parsed and translated into a format understood by DynaMIT. Changing section geometry within links was modeled by dividing the links into smaller segments. The positions of freeway, arterial and ramp sensors were located using latitude and longitude information.

The **O-D demand file** specifies the time-varying origin-destination flows for each O-D pair in the network that has a non-zero flow between it. We specify a different block of O-D flows for each time interval in our study period. The time intervals are assumed to be of constant duration.

The **supply parameter file** contains segment-specific relationships that are used by DynaMIT's supply simulator while simulating the movement of vehicles on the network. Specifically, this file contains the characteristics discussed in Equation 4.7.

DynaMIT also requires information regarding aggregate driver characteristics that are used while generating the population of drivers. These include trip purpose, information source and value of time, and are defined for each origin-destination pair that appears in the demand file. The **socio-economic characteristics** are used by DynaMIT's route choice models. Trip purpose can fall within the categories WORK, LEISURE and OTHER. Guided drivers receive information regarding updated travel times during their trips. This source could be an in-vehicle device. Unguided vehicles would not have access to updated network information during their trip, except in the vicinity of Variable Message Signs (VMS) and Changeable Message Signs (CMS). The value of time reflects the importance that drivers place on trip travel time, and is a critical component of the route choice model. The flexibility to handle time-varying socio-economic characteristics is built into the DynaMIT system.

A file containing **time-varying traffic sensor counts** is used by the O-D estimation and prediction module within DynaMIT. The data recorded by the detectors is specified in this file after aggregating all sensors over a common time interval duration.

DynaMIT uses historical estimates of **link travel times** while assigning initial routes to drivers. These travel times are the result of an equilibrium between the demand and supply in the transportation network. It should be noted that these time-dependent inputs are also an output of our calibration exercise.

In addition to the inputs discussed above, DynaMIT's demand simulator requires **error covariance and autoregressive matrices**, that are to be estimated by our calibration process.

5.2 Supply Side Calibration

The calibration of the supply simulator parameters within DynaMIT falls outside the scope of this thesis, and is reported in Kunde (2002). In this section, we present a brief review of the essential details of this process for completeness. Calibration of the supply module entails the estimation of lane group capacities and the parameters in the segment-specific speed-density relationships. Lane group output capacities at intersections are complex functions of the signal control logic and timing data. Capacities were initially estimated by approximating the signal logic based on the recommendations of the Highway Capacity Manual.

The speed-density relationships were estimated by fitting curves through the plots of observed speeds against densities. The 1373 segments were separated into eleven categories based on the type of segment (freeway, arterial, ramp, weaving section) and the different types of relationships observed in the sensor reports.

A subnetwork extracted from the Irvine test network was used to further refine the supply model parameter estimates. Dynamic speeds (computed from flow and density measurements) were used as the criterion during this step. Figure 5-8 shows a sample comparison between simulated and "observed" speeds. It should be noted that these supply parameter estimates were merely initial values. They were modified at various stages of the demand calibration process in order to obtain consistent estimates across both demand and supply models.

5.3 Demand Side Calibration

We now describe the demand calibration process on the Irvine data set. This discussion is split across several sections. We first present the details that prepare the ground for the actual calibration exercise. This includes the generation of the path choice set, definition of the period of study, generation of initial seed O-D flows, and a statement of the critical simplifying assumptions. We then proceed to outline the steps in the actual calibration process before presenting our major findings.

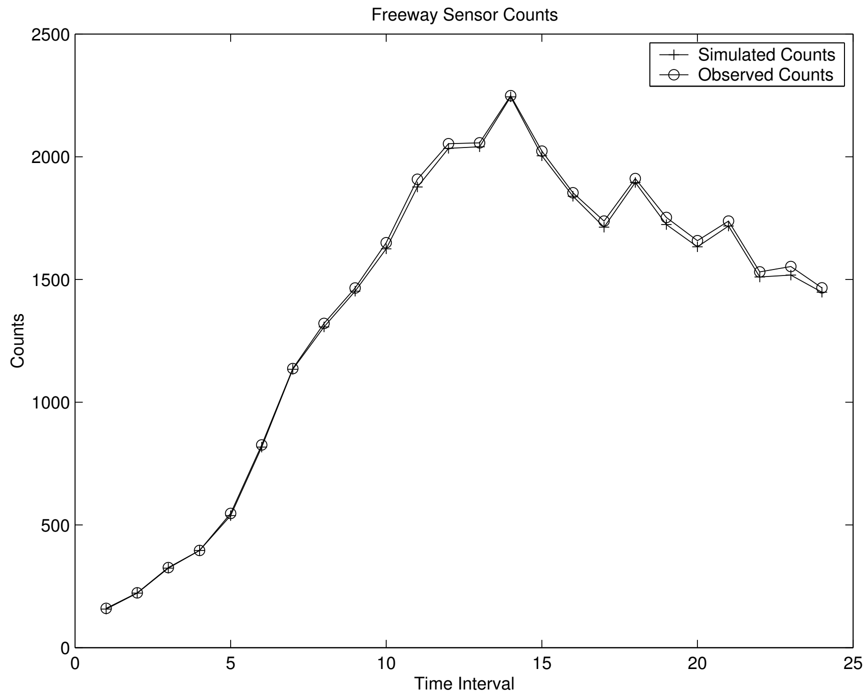

Figure 5-6: Subnetwork Calibration: Freeway Sensor

5.3.1 Path Choice Set Generation

The first step in the calibration process was the generation of a good set of paths for each O-D pair of interest. Optimum parameters for the path generation algorithm were identified so as to capture most of the feasible paths for every O-D pair. A suitable path set was obtained by using 100 random draws to complement the set of link-elimination-based shortest paths from every link to a destination node. Recognizing the need to replicate most of the freeway-based paths, an internal freeway "bias" of 0.6 was used to force the path generation algorithm into preferring paths with longer freeway sections. The random draws helped augment this set with arterial paths. All unique paths shorter than twice the shortest distance were included in the final path set. Manual inspection confirmed that most of the practical alternatives had been selected in the path generation stage. The final set contained a total of 9036 origin-destination paths.

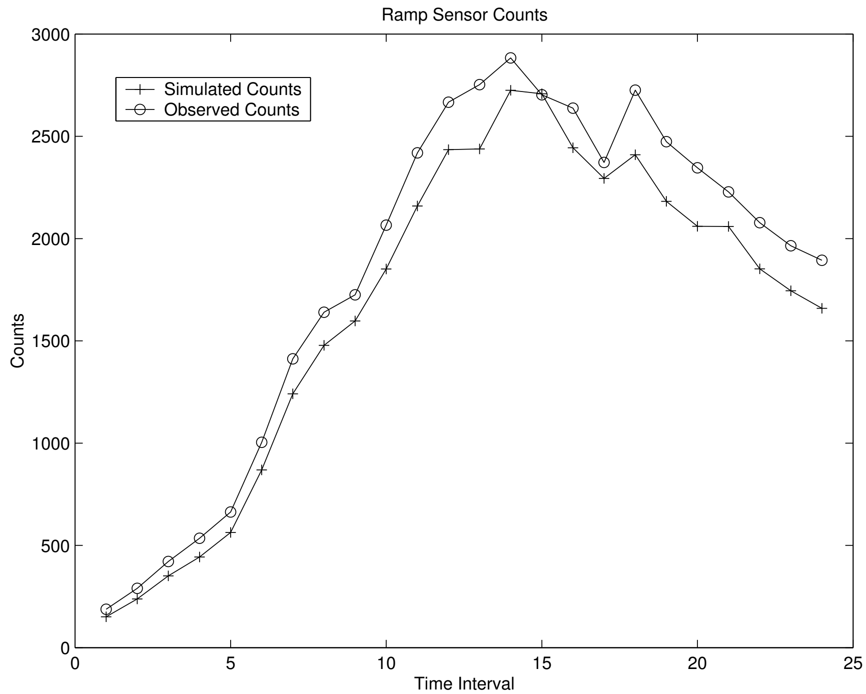

Figure 5-7: Subnetwork Calibration: Ramp Sensor

5.3.2 Defining the Period of Study

The period of interest was defined as the AM peak, from 6:00 AM to 8:30 AM[1]. However, it had to be ensured that the first interval we estimated had minimal "interference" from vehicles departing in prior time intervals[2]. Calibration was begun at 4:00 AM, when sensors indicated that the network was predominantly empty. The study interval was divided into equal subintervals of length 15 minutes. This discretization was based on probe vehicle data that indicated maximum travel times of the order of 10 minutes along the freeways and major arterials.

5.3.3 Generating Seed O-D Flows

A reasonable starting estimate of the O-D flows was constructed manually to start the O-D estimation process. The planning matrix of static O-D flows was distributed

[1] Flows were found to drop beyond 8:30 AM.

[2] Estimation of $\mathbf{x}_h$ requires knowledge of $\mathbf{a}_h^p$, which is not known initially.

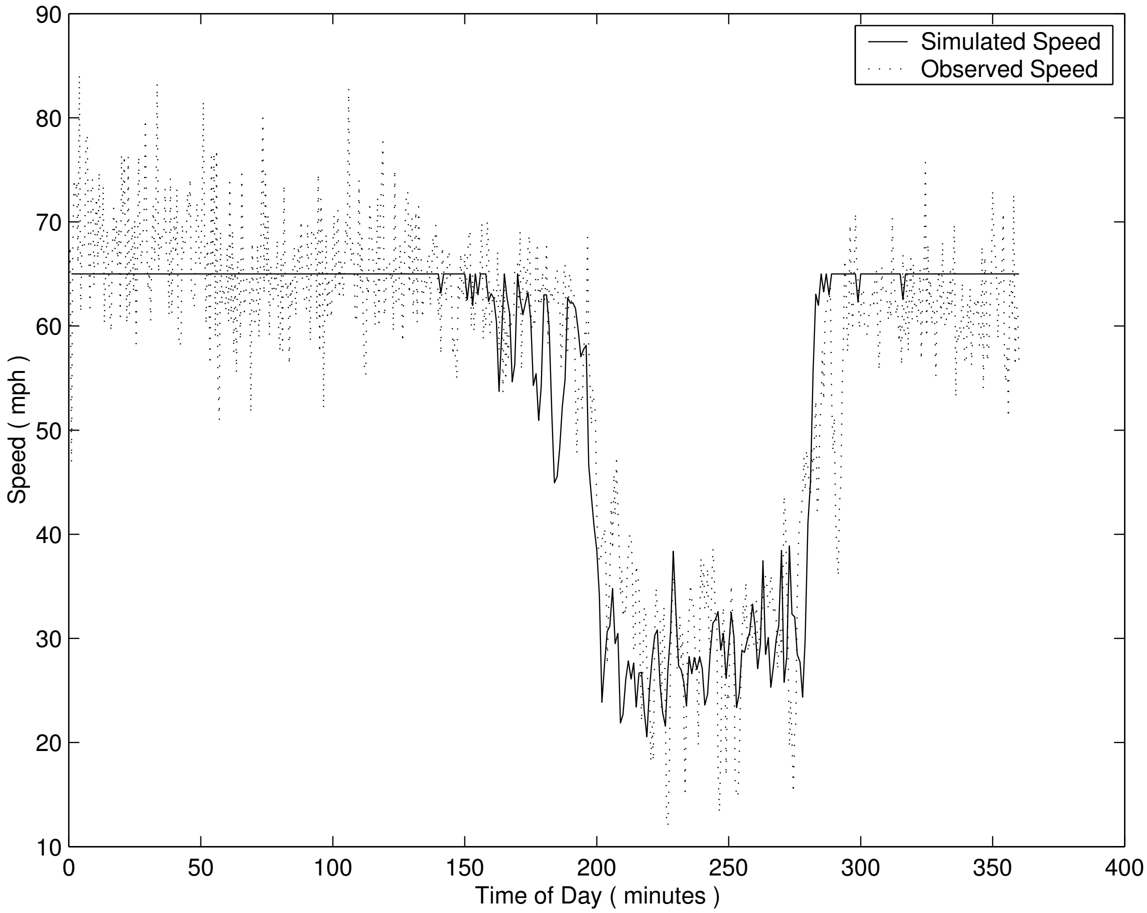

Figure 5-8: Subnetwork: Speed Comparison

across the time intervals within the study period. The fraction of flow assigned to a particular interval was proportional to a *peak factor* that was computed for each sensor as the ratio of the counts measured in the current interval to that measured across the peak period. The peak factor for each O-D flow was approximated by considering an average of the primary sensors that measured the O-D flow under study.

5.3.4 Simplifying Assumptions

Several assumptions were made in order to accommodate practical considerations while estimating the model parameters with the limited data available. The error covariance matrices $\mathbf{V}_h$ and $\mathbf{W}_h$ were assumed to possess a diagonal structure, in order to ensure that we had enough observations to estimate the elements of these matrices from just one day of data. The structure of the error covariances was also assumed to remain constant across the peak period, thereby further increasing the number of observations. The autoregressive matrices $\mathbf{f}_h^p$ were assumed to be diagonal, meaning that the deviations of flow between O-D pair r from their historical values depend on prior interval flow deviations between O-D pair r alone. The O-D flows themselves were further grouped into high, medium and low categories, and common autoregressive factors were estimated for each flow class.

5.3.5 Error Statistics

The following error statistics were used in analyzing the results:

1. Root Mean Square (RMS) Error = $\sqrt{\frac{\sum_i (y_i - \hat{y}_i)^2}{N}}$
2. Root Mean Square Normalized (RMSN) Error = $\sqrt{\frac{N \sum_i (y_i - \hat{y}_i)^2}{\sum_i y_i}}$
3. Weighted Root Mean Square (WRMS) Error = $\sqrt{\frac{\sum_i w_i (y_i - \hat{y}_i)^2}{N}}$
4. Weighted Root Mean Square Normalized (WRMSN) Error = $\frac{\sqrt{N \sum_i w_i (y_i - \hat{y}_i)^2}}{\sum_i y_i}$

where w_i, the weight associated with measurement i, is the inverse of the variance of the corresponding measurement error. The weighted statistics were computed in order to discount the effect of large measurement errors on the objective function value.

5.3.6 Calibration Approach

Three days of data were set aside for model calibration. The calibration process utilized these different days of data sequentially, and attempted to refine the parameter estimates with each additional day of data. The seed O-D flows computed from the static planning matrix were used as target flows for the O-D estimation module. These direct measurements were treated as in Equation 3.9. In the absence of any initial error covariance estimates, a weighted least squares approach was adopted, with the sensor counts being assigned higher weights than the target O-D flows. This step was employed in order to extract all the O-D flow information from the sensor counts, and downplay the importance of the seed O-D flows. The weights used for O-D flows were not kept constant across the O-D pairs. Low flows (as indicated by the planning matrix) were assigned high weights in order to maintain them in an acceptable range, and capture the fact that such flows normally exhibit low variability. Flows for each time interval were estimated sequentially using the GLS formulation outlined in Section 2.4.

Several iterations were performed in conjunction with the O-D estimation step to determine an optimal set of parameters for the route choice model. The utility V_i of each path is a function of two travel time components:

$$V_i = \beta_1 tt_{Ai} + (\beta_2 \beta_1) tt_{Fi} \tag{5.1}$$

where tt_{Ai} and tt_{Fi} are the arterial and freeway components of the travel time along path i, β_1 is the coefficient of arterial travel time, and β_2 is the freeway bias[3]. Different

[3]Freeway bias attempts to capture the driver's preference for a freeway section over an arterial section of the same impedance.

combinations of the parameters β_1 and β_2 were tried , to cover a wide range of route choice situations. Obviously, one would expect β_1 to carry a negative sign, while β_2 would be a positive number between 0 and 1. Table 5.1 summarizes the last few iterations of this search process. Iterations 1 through 6 represent (β_1,β_2) of (-0.03,0.77), (-0.04,0.75), (-0.035,0.75), (-0.03,0.75), (-0.04,0.78) and (-0.04,0.80) respectively. The optimal value for (β_1, β_2) was found to be (-0.04, 0.80) (as seen from Figure 5-9), and was identified after the third day of data was processed. A very high value of 40 for the Path-Size exponent γ was also selected for best model fit[4].

Iteration	RMS	RMSN	WRMS	WRMSN
1	9.77E+03	1.3012	87.7477	0.0117
2	1.01E+04	1.3440	61.0514	0.0081
3	9.68E+03	1.2891	53.0597	0.0071
4	9.89E+03	1.3175	52.0614	0.0069
5	9.76E+03	1.3002	49.2280	0.0066
6	9.67E+03	1.2874	47.9872	0.0064

Table 5.1: Error in Fit to Counts for Varying Route Choice Parameters

Once the flows for all intervals had been estimated using the first day of data, the error covariances were estimated from the resulting residuals, as described in Section 3.4. The O-D flows were then re-estimated for day 1 using the new covariance estimates. The resulting O-D flows and simulated counts were used to re-estimate the error covariances again before moving to day 2.

Data from day 2 was processed in a manner similar to day 1, with a slight difference. The O-D flow estimates obtained from day 1 were designated as historical flows for day 2, and Equation 3.5 was used to compute the target flows for subsequent time intervals. The day 1 estimates for the first interval were used as target flows for the first interval in day 2. The error covariances were estimated before proceeding to day 3. The treatment of data from day 3 was identical to that from day 2.

Figures 5-10 to 5-14 graphically depict the comparison of estimated and observed sensor counts during the off-peak period leading to the AM peak. Having thus con-

[4]Such a high value of γ is consistent with the findings of Ramming (2001), who reports best fit as $\gamma \to \infty$.

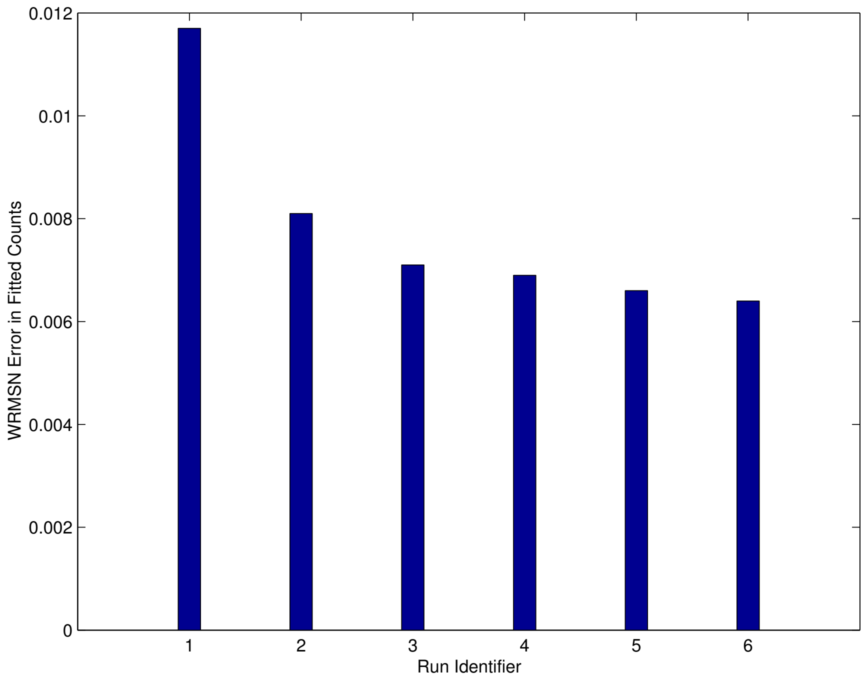

Figure 5-9: Route Choice Parameter Estimation

firmed that we have a reasonable estimate of the network state at the commencement of the peak period of interest, we focus on the peak period for statistical comparisons. Similar visual comparisons for intervals in the peak period are presented in Figures 5-15 to 5-17 after processing the third day of data. Tables 5.2 and 5.3 summarize the error statistics that measure the performance of the calibration algorithm at various stages[5]. While the RMSN errors were found to be stable across the four estimation iterations, the progress in the estimates was indicated by the Weighted RMSN (WRMSN) errors, that accounted for the variability in sensor measurements. Table 5.3 clearly indicates an improvement in fit with every additional day of surveillance data processed.

Analysis of the calibration algorithm's performance was also evaluated by comparing the time-varying sensor counts reported by the supply simulator against that recorded by the surveillance system (Figure 5-21). It was found that the counts on

[5]More detailed numerical results are presented in Appendix A

		Initialization	day 1	day 2	day 3
6:00		0.09	0.11	0.13	0.11
6:15		0.17	0.13	0.16	0.15
6:30		0.09	0.09	0.14	0.11
6:45		0.11	0.11	0.13	0.11
7:00		0.13	0.13	0.11	0.10
7:15		0.15	0.14	0.11	0.14
7:30		0.13	0.12	0.14	0.11
7:45		0.14	0.14	0.15	0.22
8:00		0.16	0.13	0.16	0.21
8:15		0.23	0.16	0.18	0.23

Table 5.2: RMSN Errors from Four Estimations

		Initialization	day 1	day 2	day 3
6:00		231.87	23.36	11.37	9.66
6:15		150.92	16.77	17.44	16.63
6:30		230.51	14.92	10.35	9.23
6:45		159.67	15.00	10.63	10.36
7:00		195.26	50.82	7.59	7.31
7:15		156.34	23.14	13.19	7.39
7:30		108.79	13.30	21.04	13.83
7:45		157.03	16.32	7.81	15.97
8:00		107.66	13.94	9.41	9.77
8:15		132.07	35.54	8.44	18.53

Table 5.3: WRMSN Errors from Four Estimations

freeway links were replicated with a high degree of accuracy, while the counts on arterial and ramp links were subject to a small degree of stochasticity. This error could be largely attributed to approximations made while converting the vehicle-actuated signal control logic at intersections into segment-specific capacities. Recent research in this area has subsequently led to the integration of a capacity translator in DynaMIT, that allows for the online adjustment of segment capacities at intersections. This addition was however, not included in this study.

A further check involving the O-D flows was performed to ascertain the variation of O-D flows with time. A comparison of the dynamic estimated flows for one of the most heavily traveled O-D pairs (Figure 5-22) against the initial seed flows indicate that the dynamic seed flows obtained from the static planning flows is indeed a good approximation of the actual flows, and provide a reasonable starting point for the calibration process.

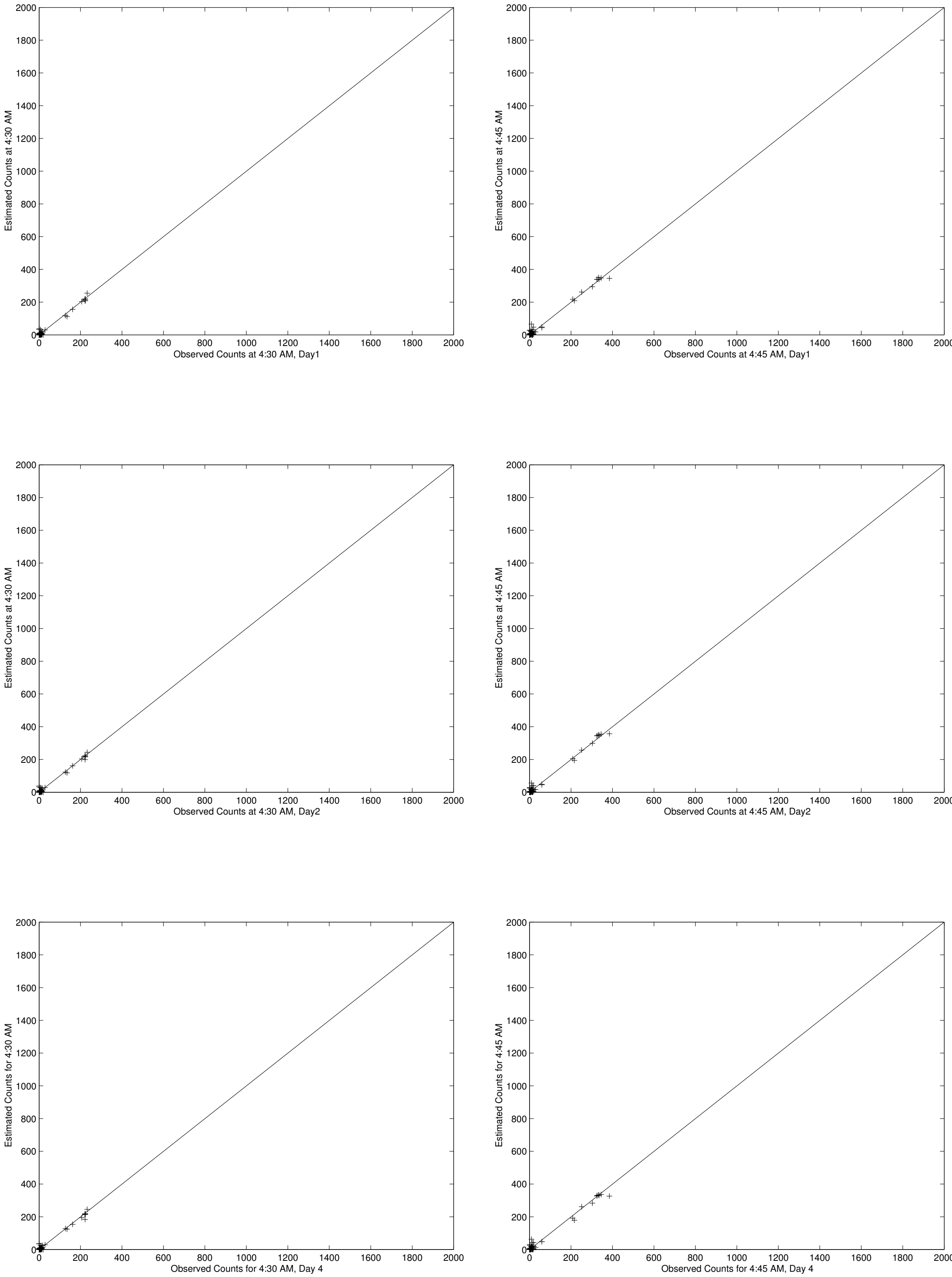

Figure 5-10: Days 1, 2 and 4: Counts from 4:15 AM to 4:45 AM

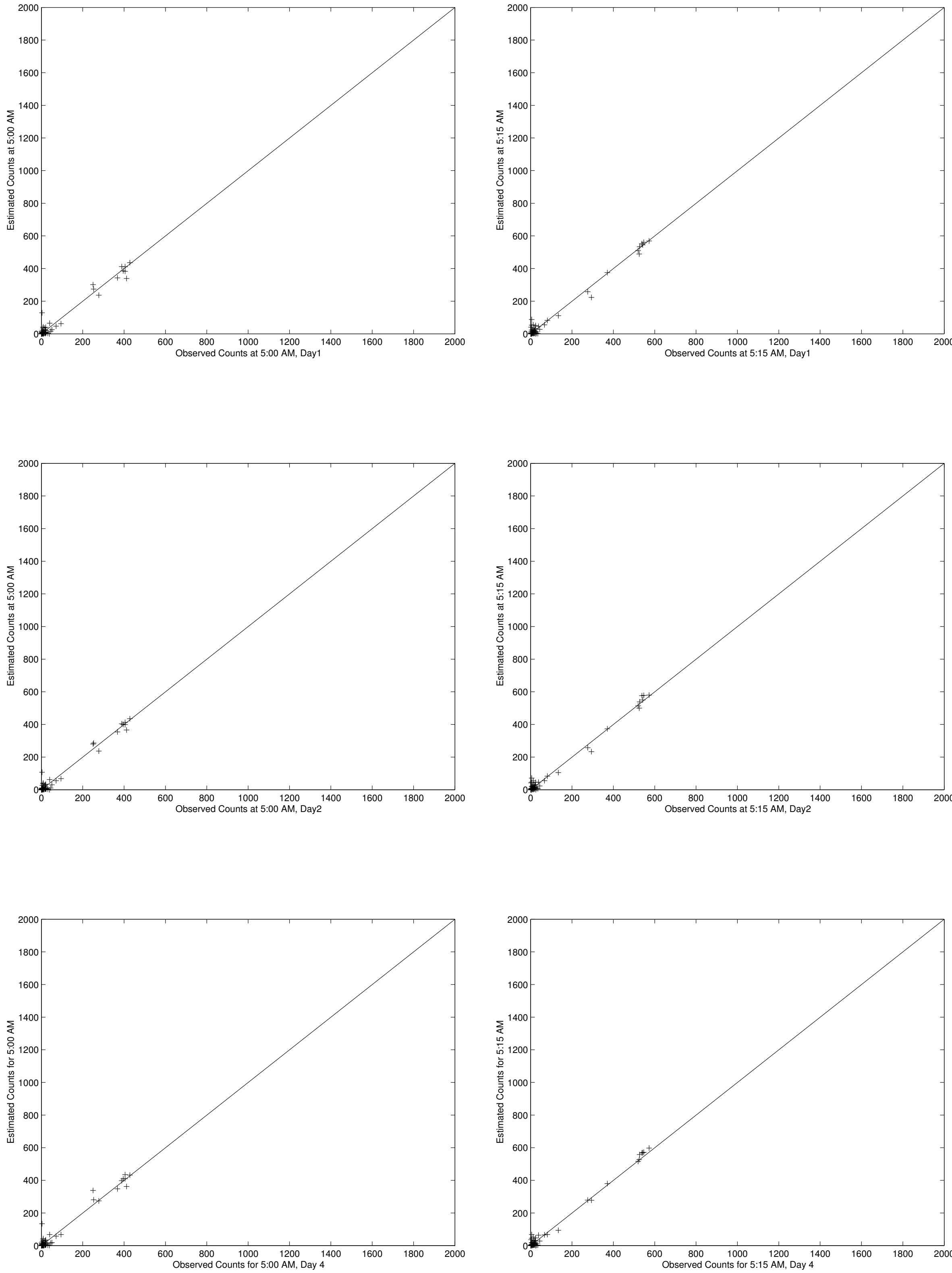

Figure 5-11: Days 1, 2 and 4: Counts from 4:45 AM to 5:15 AM

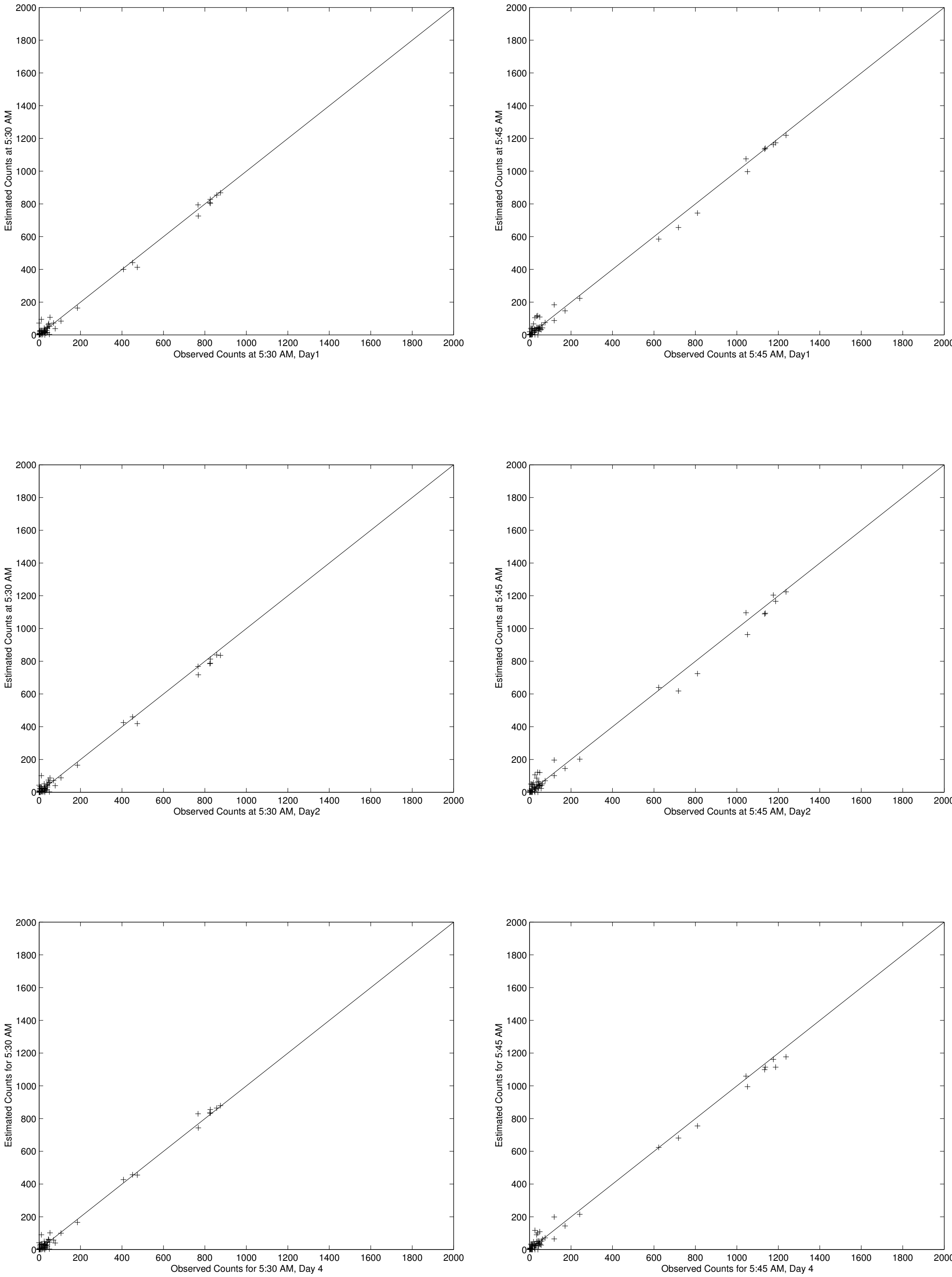

Figure 5-12: Day 1: Counts from 5:15 AM to 5:45 AM

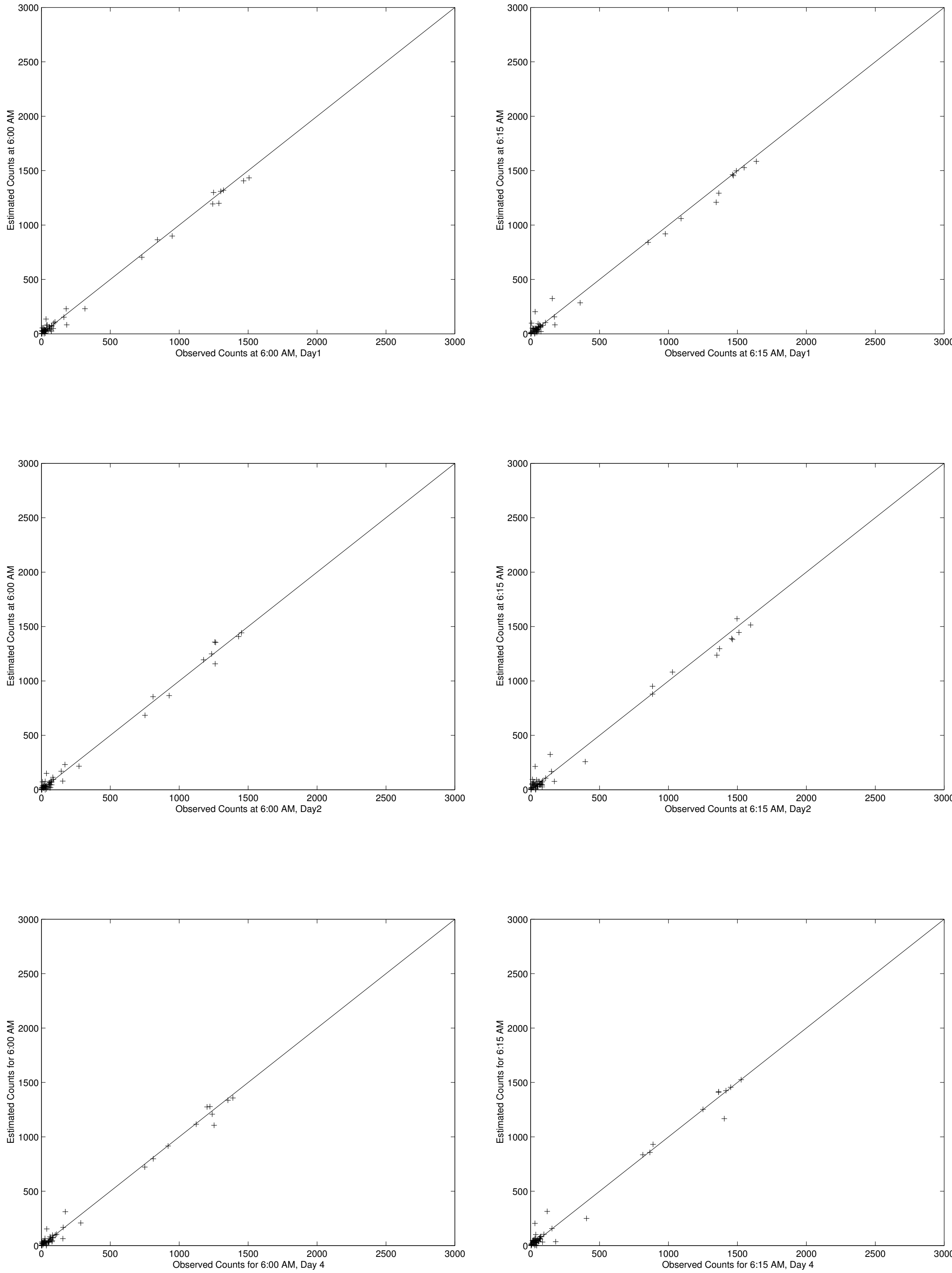

Figure 5-13: Day 2: Counts from 5:45 AM to 6:15 AM

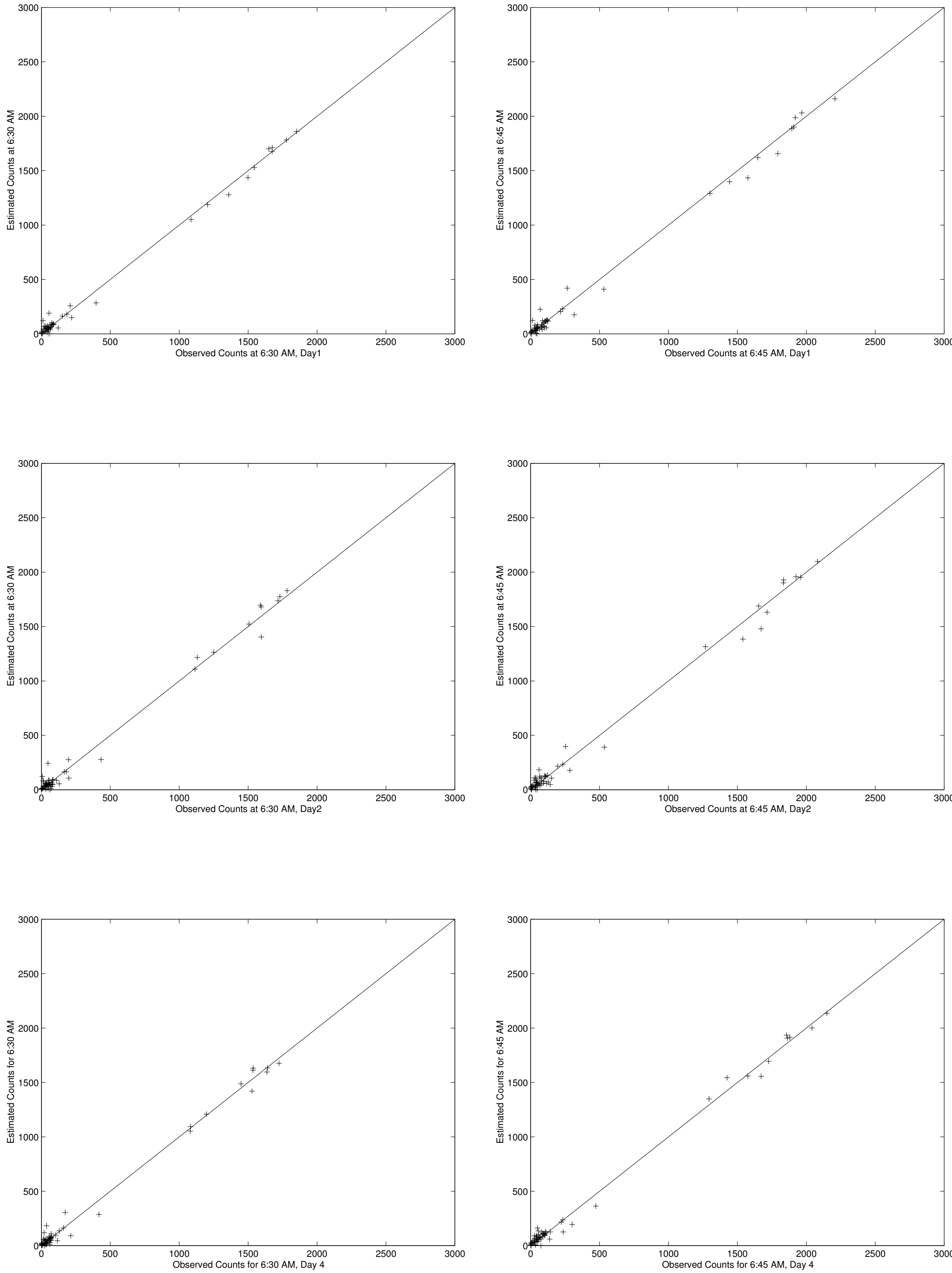

Figure 5-14: Day 2: Counts from 6:15 AM to 6:45 AM

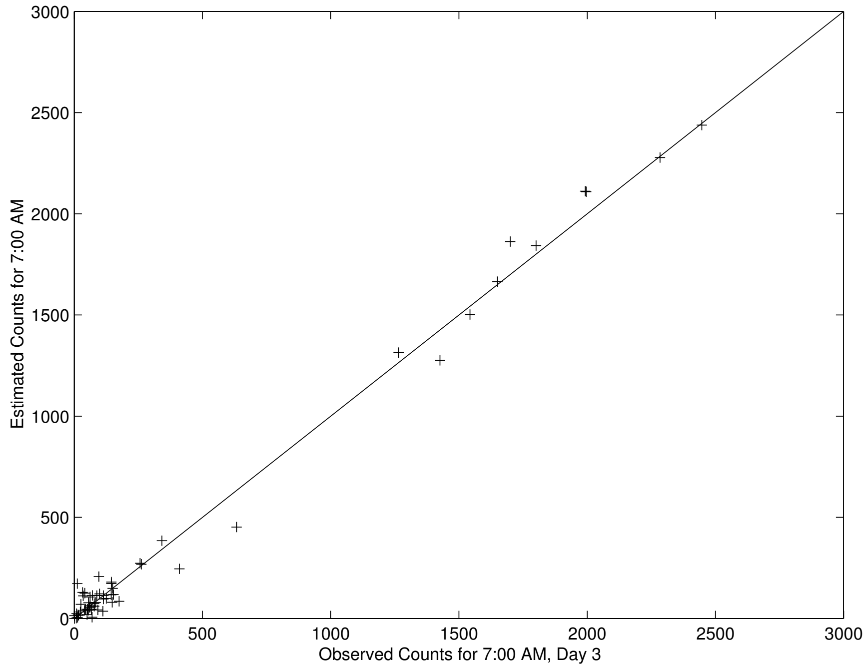

Figure 5-15: Estimated Counts for 6:45 to 7:00 AM

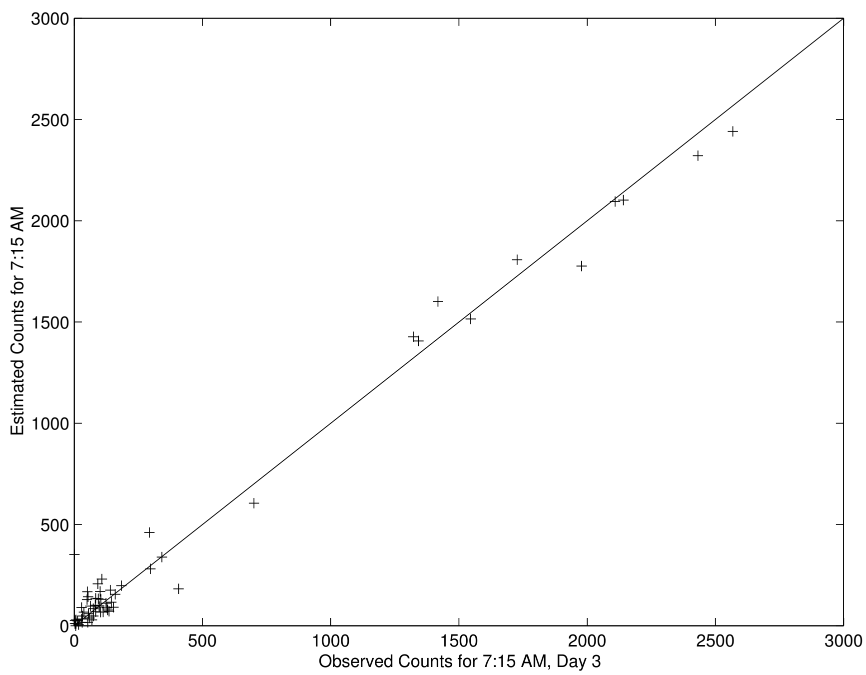

Figure 5-16: Estimated Counts for 7:00 to 7:15 AM

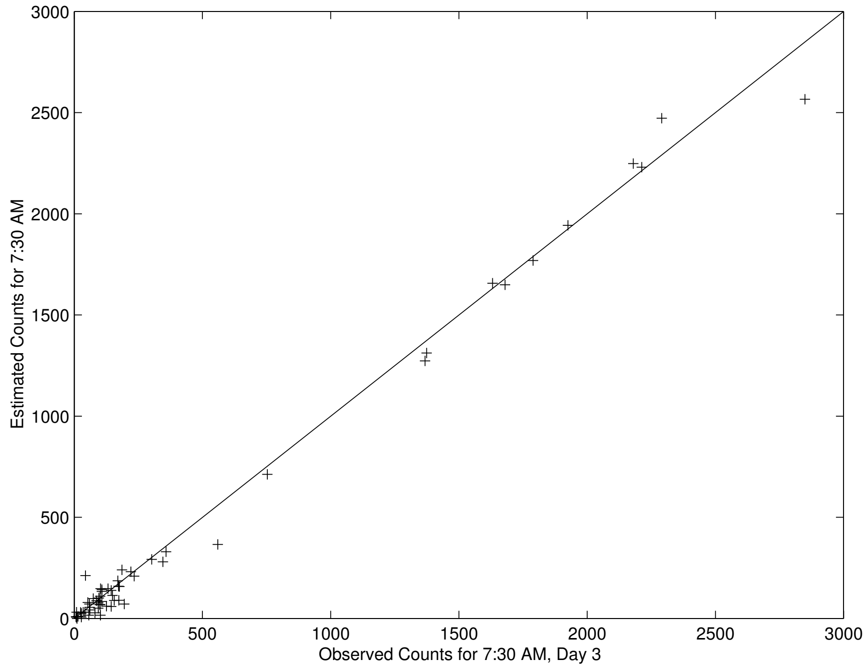

Figure 5-17: Estimated Counts for 7:15 to 7:30 AM

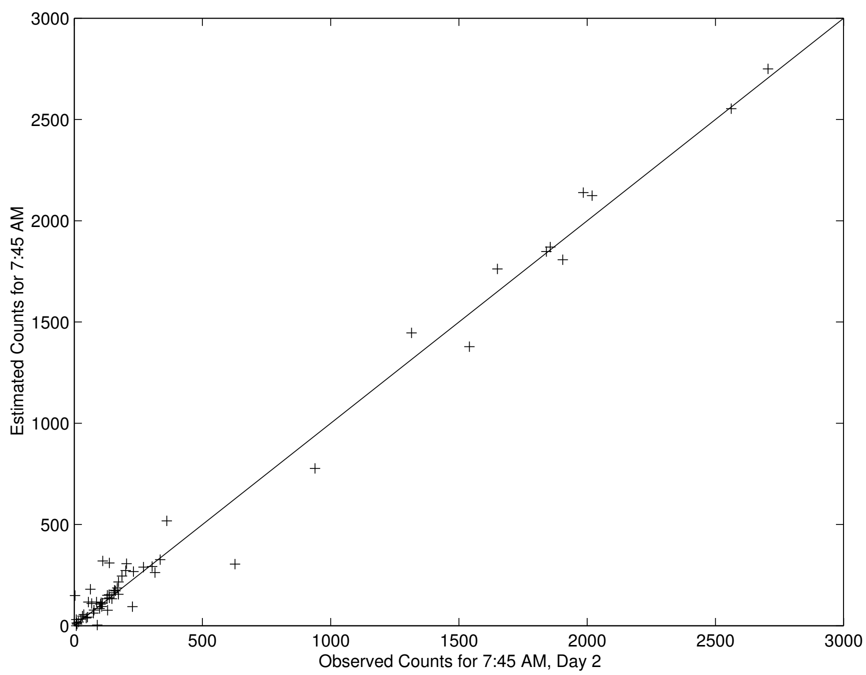

Figure 5-18: Estimated Counts for 7:30 to 7:45 AM

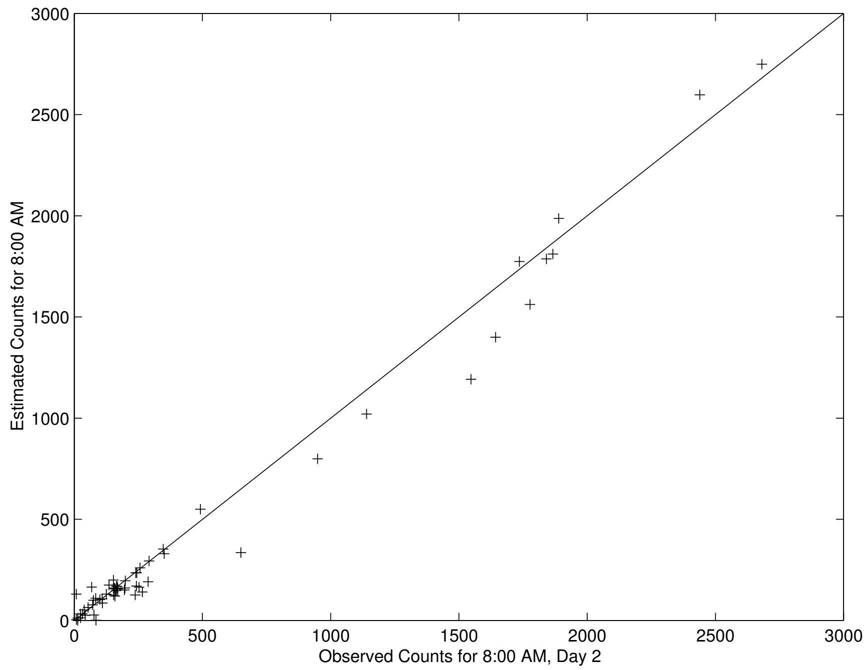

Figure 5-19: Estimated Counts for 7:45 to 8:00 AM

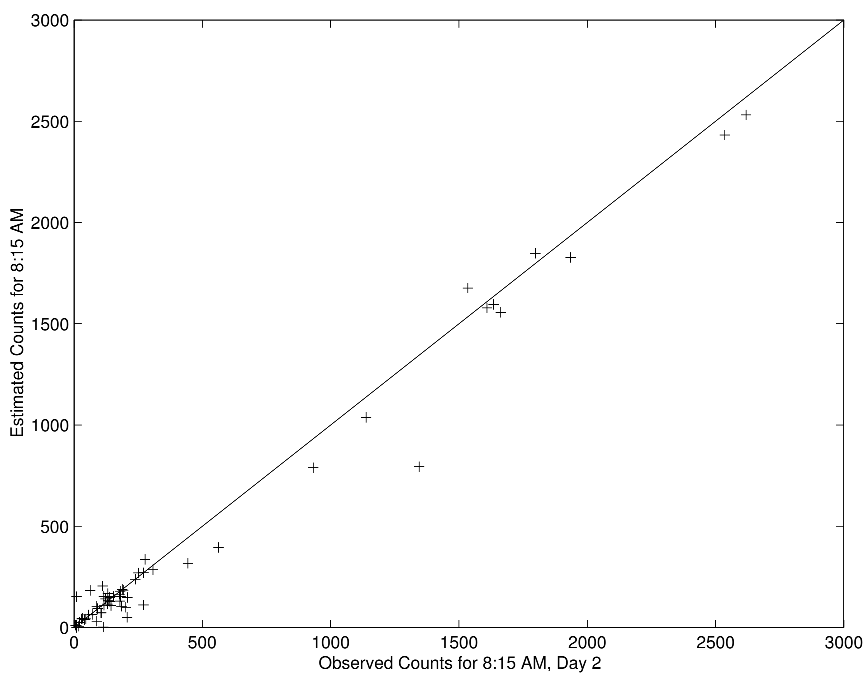

Figure 5-20: Estimated Counts for 8:00 to 8:15 AM

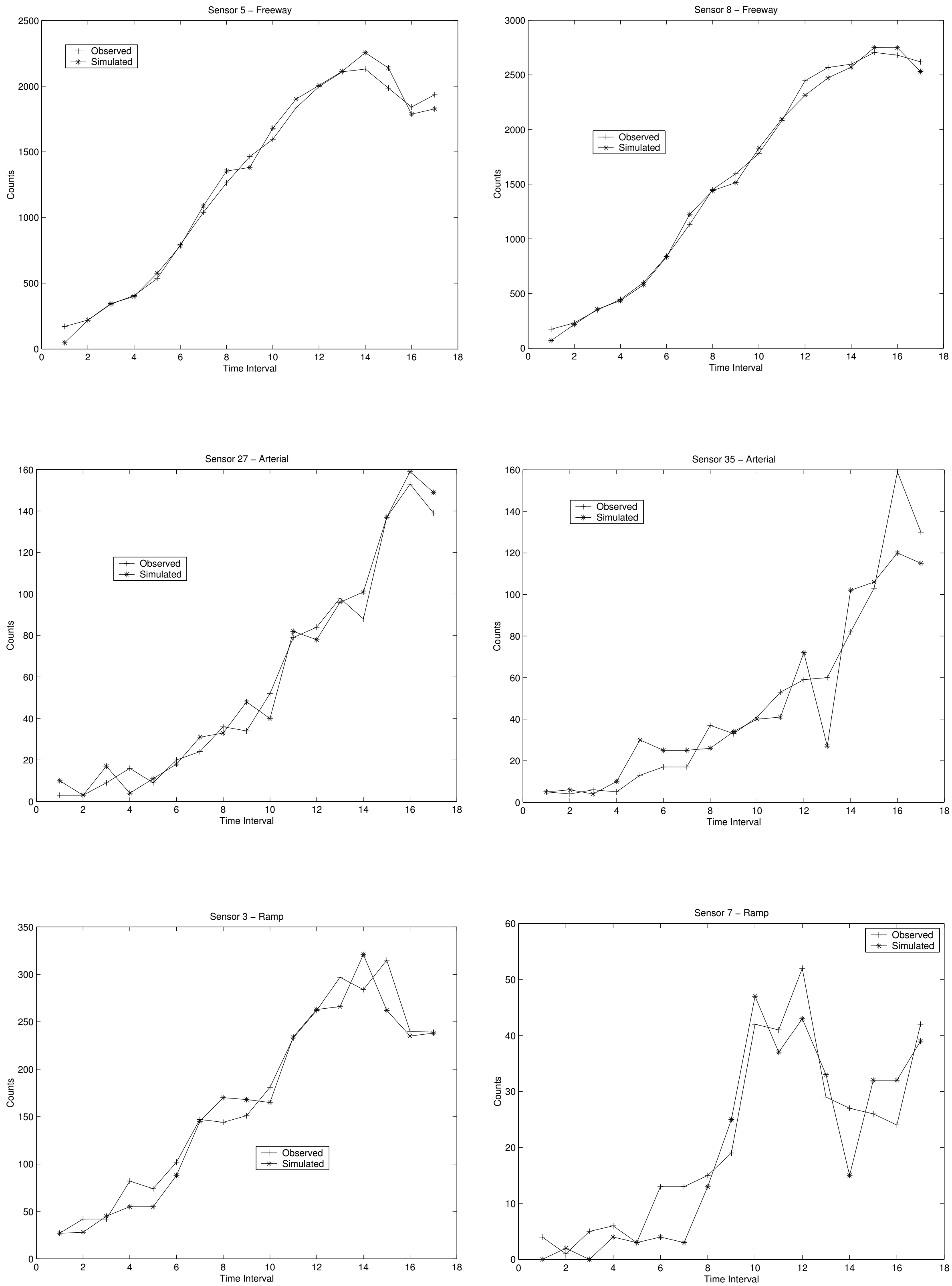

Figure 5-21: Comparison of Time-Varying Freeway, Arterial and Ramp Sensor Counts

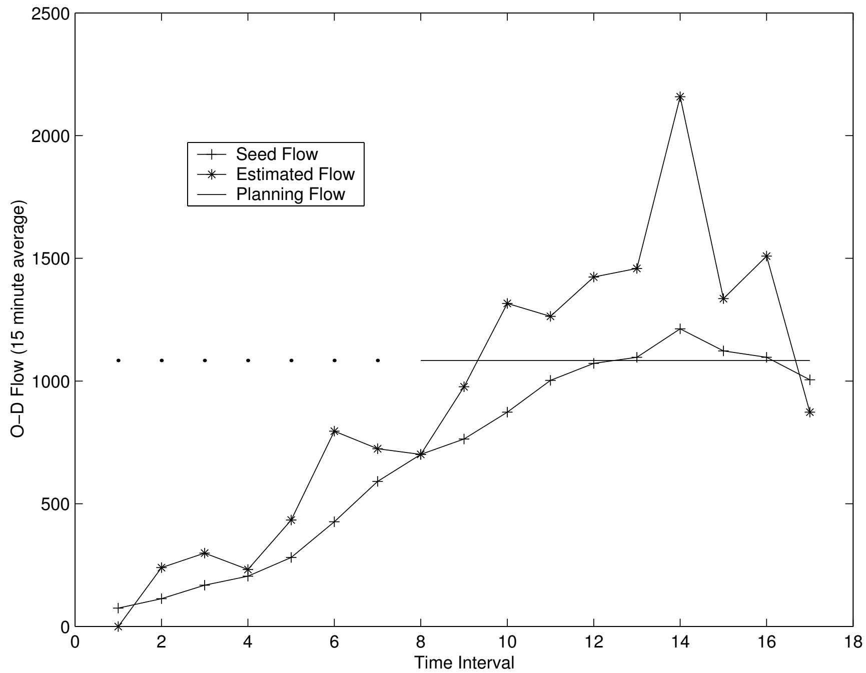

Figure 5-22: Comparison of O-D Flows

Having used the first three days of data to obtain the flow and error covariance estimates, we proceeded to estimate an autoregressive process that would form the basis for DynaMIT's predictive capability. The deviations of the day 3 estimates from their historical values[6] were computed, and regressions similar to Equation 3.13 were run. An autoregressive process of degree 4 was found to fit the data best, with an additional degree causing only a marginal improvement in fit (by reducing the RMS error from 30.04 to 29.84). Separate parameters were estimated for O-D pairs falling within low, medium and high flow brackets[7]. The estimated factors are summarized below:

$$\begin{bmatrix} \rho_{h-1}^{hi} & \rho_{h-2}^{hi} & \rho_{h-3}^{hi} & \rho_{h-4}^{hi} \\ \rho_{h-1}^{med} & \rho_{h-2}^{med} & \rho_{h-3}^{med} & \rho_{h-4}^{med} \\ \rho_{h-1}^{low} & \rho_{h-2}^{low} & \rho_{h-3}^{low} & \rho_{h-4}^{low} \end{bmatrix} = \begin{bmatrix} 0.4186 & 0.2840 & 0.0103 & -0.1841 \\ 0.3408 & 0.2280 & -0.0642 & 0.7881 \\ 0.2076 & 0.1124 & 0.0547 & 0.0359 \end{bmatrix} \tag{5.2}$$

[6] The estimates from day 2 were used as historical flows here.

[7] Historical flows higher than 600 vehicles/hour were classified as high, while those lower than 280 vehicles/hour were designated low.

The predictive ability of the calibrated DynaMIT system was evaluated through validation tests.

5.4 Validation of Calibration Results

The calibrated system was tested for its estimation and prediction performance using data that was not a part of the original calibration exercise. Data from day 4 was used for this purpose.

5.4.1 Validation of Estimation Capability

The process of validating the estimation capability of the calibrated DynaMIT system was similar to the actual calibration process. The best historical O-D flow estimates, error covariances and route choice parameters were used along with a new day of traffic surveillance data to re-estimate the O-D flows. The resulting simulated sensor counts were compared against those recorded by the field sensors.

	RMSN		WRMSN	
	day 3	day 4	day 3	day 4
6:00	0.11	0.12	9.66	8.77
6:15	0.15	0.16	16.63	12.26
6:30	0.11	0.12	9.23	8.34
6:45	0.11	0.10	10.36	6.97
7:00	0.10	0.12	7.31	6.29
7:15	0.14	0.13	7.39	4.72
7:30	0.11	0.20	13.83	7.47
7:45	0.22	0.22	15.97	8.40
8:00	0.26	0.26	9.77	8.91
8:15	0.42	0.42	18.53	12.84

Table 5.4: Validation of Estimation Results

Table 5.4 summarizes the results of the first validation test. The first two columns indicate the RMSN error in counts from the last calibration run, and the validation run respectively. These errors were found to be fairly stable. The weighted RMSN errors (contained in columns 3 and 4) showed a further slight negative trend, thereby indicating that the five estimation iterations performed in total have moved in the right direction with reference to replicating the observed link counts.

Having verified DynaMIT's ability to estimate O-D flows, we next moved to evaluating its predictive ability.

5.4.2 Validation of Prediction Capability

The validation of DynaMIT's prediction capability is critical for the evaluation of the performance of the DTA-based guidance and traveler information system. For this test, the calibrated system was run from 4:00 AM through 7:30 AM. During this run, DynaMIT estimated flows for each time interval until 7:30 AM. The autoregressive process was now applied on the flow deviations to generate one, two, three and four-step O-D flow predictions (covering a one-hour prediction horizon), which were then simulated by DynaMIT's supply simulator to generate simulated sensor counts. The comparison of simulated counts with those actually observed by the surveillance system is summarized in Table 5.5. As expected, the fit to counts deteriorated as the prediction interval moved away from the current interval. Figures 5-23 and 5-24 graphically indicate a slight reduction in fit as the prediction horizon increases, thereby highlighting the importance of a rolling horizon approach to traffic prediction.

	RMSN		WRMSN	
	day 4	Validation	day 4	Validation
7:30	0.20	0.14	7.47	9.25
7:45	0.22	0.12	8.40	18.54
8:00	0.26	0.15	8.91	17.53
8:15	0.42	0.16	12.84	19.16
8:30	-	0.25	-	21.61

Table 5.5: Validation of Prediction Results

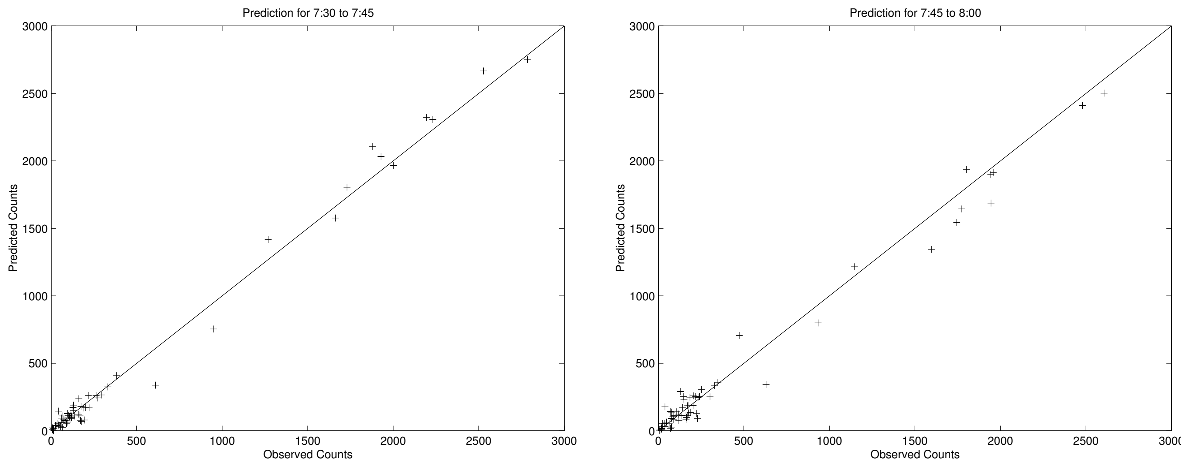

Figure 5-23: Predicted Counts for 7:30 AM to 8:00 AM

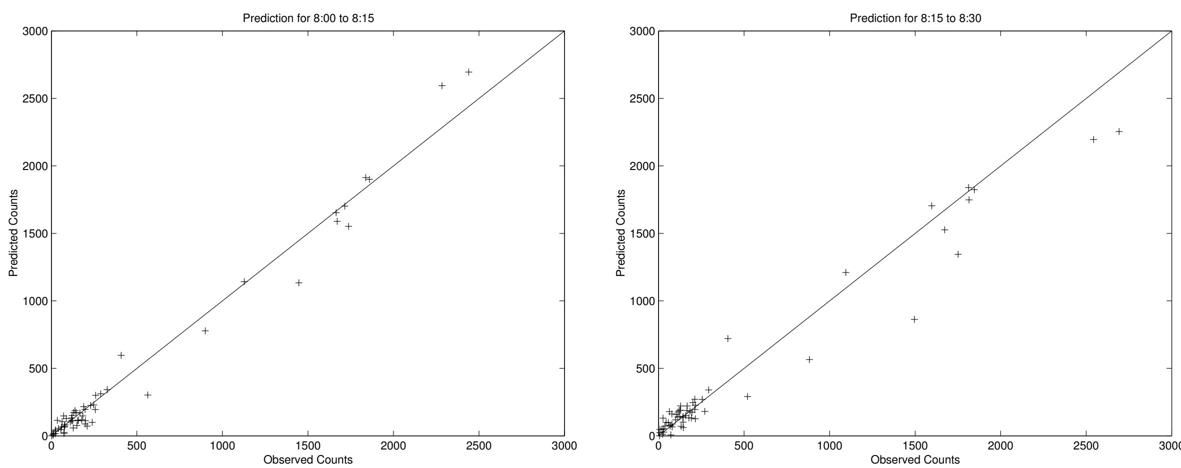

Figure 5-24: Predicted Counts for 8:00 AM to 8:30 AM

5.5 Summary and Conclusion

In this chapter, we calibrated the DynaMIT/DynaMIT-P DTA system using data collected by a traffic surveillance system in Irvine, California. We further validated the estimation and prediction capabilities of the calibrated system. The results indicate that the calibration methodology proposed in earlier chapters can indeed be applied to large urban networks with complex routes to obtain robust parameter estimates.

Chapter 6

Conclusion

We begin this chapter with an assessment of the contributions of this research to the state of the art of DTA model calibration methods, and conclude with some suggestions for future research directions.

6.1 Research Contribution

This research represents significant extensions to existing literature related to the calibration and validation of Dynamic Traffic Assignment Systems. Specifically,

- The methodology presented here is one of the first attempts at the joint calibration of route choice and O-D estimation and prediction models. It extends existing O-D estimation approaches to handle the needs of a calibration process, and provides an integrated framework for the joint calibration of individual driver route choice models in the absence of disaggregate data.

- This thesis undertakes a detailed outline of the component steps in the demand calibration process, and analyzes the inputs and outputs at each stage. The critical issue of starting the calibration process from static O-D data in the absence of reliable historical estimates is also discussed and implemented.

- This work addresses the issue of using several days of data to update the historical database of O-D flows, error covariances and autoregressive factors.

- The case studies in this research involve a real-sized urban network consisting of both freeways and arterials, and demonstrate the practical feasibility and robustness of the proposed calibration approach. More importantly, the calibrated DTA system has been validated through evaluations of its estimation and prediction capabilities, with very encouraging results.

6.2 Directions for Further Research

The focus of this thesis was on a general framework for the calibration of the demand models within a DTA system. We have discussed the component steps in the process, and the factors involved in obtaining the inputs and outputs at each step. We addressed issues relating to starting the process for the first few days of data, when several key inputs are likely to be either unavailable, or unreliable at best. We further presented a framework for setting up a historical database of O-D flows, error covariances and autoregressive factors, and outlined some approaches to updating this database with multiple days of surveillance data. There are several issues, however, that could be pursued as further research.

6.2.1 Updating Model Parameter Estimates

While this work has considered the within-day variability in O-D flows, a common underlying process was assumed to govern all four days of data used in this research. A topic for further research could include the analysis of correlations across multiple days of data, and efficient estimation methods to handle the same. Such effects would be important especially when the calibration objective is expanded to include the creation of a stratified historical database covering a wide range of demand conditions and traffic scenarios. Criteria for such stratification schemes could include day of the week, weather conditions, special events, scheduled workzone operations and incidents.

Yet another issue is the importance of the starting seed O-D matrix. Clearly, the seed plays a critical role in the calibration process, particularly when the assignment matrices are obtained from simulation models. With a sufficiently large data set encompassing several months of data, one would expect the starting seed flows to have a negligible effect on the final flow estimates. The limited size of the data set available for this research, however, prevented us from exploring this idea in greater detail.

6.2.2 Simultaneous estimation

DynaMIT employs a sequential approach to O-D estimation, primarily from considerations of real-time computational efficiency. The calibration process, however, is offline, and a simultaneous approach would be expected to yield more efficient estimates. One might therefore consider replacing the sequential GLS O-D estimator with alternative estimation methods such as a simultaneous GLS estimator, or the Kalman filter-based offline smoothing technique developed by Ashok (1996). This approach, however, would require the computation and storage of assignment matrices from all past intervals that influence counts in the current interval, and can be computationally prohibitive. The development of efficient storage schemes might yield a more robust calibration methodology.

6.2.3 Effect of number of sensors

The size of the O-D estimation problem depends on the number of sensors on the network. In the perfect scenario, there would be as many sensors as O-D pairs, so that the unknown O-D flows can be obtained in a simple manner. Unfortunately, such is rarely the case. The number of sensors is often far lower than the number of flows we wish to estimate. It would therefore be interesting to study the effect of network coverage (in terms of the number of independent sensors) on the robustness of the O-D flow estimates. One could also attempt to identify an optimum coverage level to ensure consistent and robust estimation.

6.2.4 Handling incidents

An important aspect of traffic estimation and prediction is the handling of incidents. An extension of this study could include the analysis of DynaMIT's estimation and prediction capabilities under unexpected demand fluctuations, and its ability to generate consistent route guidance to minimize the impacts of the disturbance. A requirement for such a study is the availability of a rich data set encompassing several months of surveillance data, including detailed records of accidents and incidents that caused capacity reductions in the network.

6.2.5 Driver behavior models

The driver behavior models in a DTA system would be composed of several network-related attributes that might explain the choices made by travelers. In fact, drivers in the real world might cancel their trips or switch modes based on information provided to them by the ATIS. Under congested conditions, drivers might also change their departure times in order to beat the rush. Such deviations could be expected to seriously impact the O-D estimation module. Recent enhancements to DynaMIT's behavioral models have added the functionality to model these complex decisions. Future work could study the effect of such driver behavior on the efficiency of the final estimates, and on the quality of DynaMIT's predictions.

6.3 Conclusion

An integrated approach to the calibration of demand models within a Dynamic Traffic Assignment system was presented in this thesis. The route choice and O-D estimation and prediction models were jointly calibrated for a large-scale urban network, using data collected by a real traffic surveillance system. Validation studies indicated encouraging results regarding the estimation and prediction capabilities of the calibrated system. Further work involving larger data sets and more complex DTA models would significantly extend the state of the art of real-time traffic simulation model calibration and validation.

Appendix A

Tables

		RMS	RMSN	WRMS	WRMSN
6:00		27.71	0.09	75223.04	231.87
6:15		59.59	0.17	52672.55	150.92
6:30		35.85	0.09	93143.59	230.51
6:45		51.88	0.11	76971.29	159.67
7:00		64.74	0.13	100490.56	195.26
7:15		79.48	0.15	83111.06	156.34
7:30		77.16	0.13	65180.36	108.79
7:45		80.89	0.14	93435.59	157.03
8:00		96.78	0.16	64204.88	107.66
8:15		128.71	0.23	73878.42	132.07

Table A.1: Error Statistics for Day 1 Data (initialization)

		RMS	RMSN	WRMS	WRMSN
6:00		35.54	0.11	7577.61	23.36
6:15		46.35	0.13	5854.21	16.77
6:30		38.05	0.09	6030.49	14.92
6:45		53.39	0.11	7230.90	15.00
7:00		65.27	0.13	26153.72	50.82
7:15		72.77	0.14	12302.60	23.14
7:30		73.17	0.12	7967.53	13.30
7:45		85.47	0.14	9712.34	16.32
8:00		78.30	0.13	8314.34	13.94
8:15		89.87	0.16	19882.89	35.54

Table A.2: Error Statistics for Day 1 Data

		RMS	RMSN	WRMS	WRMSN
6:00		38.04	0.13	3401.00	11.37
6:15		55.73	0.16	6072.21	17.44
6:30		57.16	0.14	4105.82	10.35
6:45		60.12	0.13	5017.08	10.63
7:00		55.47	0.11	3843.48	7.59
7:15		60.65	0.11	6990.40	13.19
7:30		79.04	0.14	11514.79	21.04
7:45		85.47	0.15	4561.13	7.81
8:00		93.50	0.16	5559.81	9.41
8:15		100.06	0.18	4810.94	8.44

Table A.3: Error Statistics for Day 2 Data

		RMS	RMSN	WRMS	WRMSN
6:00		37.71	0.11	3183.22	9.66
6:15		56.14	0.15	6194.73	16.63
6:30		48.35	0.11	4034.87	9.23
6:45		59.19	0.11	5444.57	10.36
7:00		58.14	0.10	4053.42	7.31
7:15		80.18	0.14	4259.06	7.39
7:30		68.10	0.11	8313.14	13.83
7:45		133.17	0.22	9705.44	15.97
8:00		129.67	0.21	5953.61	9.77
8:15		131.48	0.23	10495.00	18.53

Table A.4: Error Statistics for Day 3 Data

		RMS	RMSN	WRMS	WRMSN
6:00		39.81	0.12	2886.65	8.77
6:15		57.76	0.16	4340.39	12.26
6:30		47.63	0.12	3342.41	8.34
6:45		46.88	0.10	3435.29	6.97
7:00		65.21	0.12	3367.01	6.29
7:15		71.71	0.13	2556.83	4.72
7:30		117.26	0.20	4400.79	7.47
7:45		136.35	0.22	5263.17	8.40
8:00		162.64	0.26	5514.74	8.91
8:15		244.45	0.42	7469.00	12.84

Table A.5: Error Statistics for Day 4 Data

Bibliography

Antoniou, C.: 1997, *Demand Simulation for Dynamic Traffic Assignment*, Master's thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Ashok, K.: 1992, *Dynamic Trip Table Estimation for Real Time Traffic Management Systems*, Master's thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Ashok, K.: 1996, *Estimation and Prediction of Time-Dependent Origin-Destination Flows*, PhD thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Ashok, K. and Ben-Akiva, M.: 1993, Dynamic Origin-Destination Matrix Estimation and Prediction for Real-Time Traffic Management Systems, *in* C. Daganzo (ed.), *Transportation and Traffic Theory*, pp. 465–484.

Ashok, K. and Ben-Akiva, M.: 2000a, Alternative Approaches for Real-Time Estimation and Prediction of Time-Dependent Origin-Destination Flows, *Transportation Science* **34**(1), 24.

Ashok, K. and Ben-Akiva, M.: 2000b, Dynamic Origin-Destination Matrix Estimation and Prediction for Real-Time Traffic Management Systems.

Ashok, K. and Ben-Akiva, M.: 2000c, Estimation and Prediction of Time-Dependent Origin-Destination Flows with a Stochastic Mapping to Path Flows and Link Flows.

Barton-Aschman Associates Inc., Parsons Transportation Group: 1999, Orange County Transportation Analysis Model OCTAM III, *Technical report*. Prepared for Orange County Transportation Authority.

Ben-Akiva, M. E. and Lerman, S.: 1985, *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge, MA, USA.

Bottom, J.: 2000, *Consistent Anticipatory Route Guidance*, PhD thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Brandriss, J.: 2001, *Estimation of Origin-Destination Flows for Dynamic Traffic Assignment*, Master's thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA.

Cantarella, G. and Cascetta, E.: 1995, Dynamic Processes and Equilibrium in Transportation Networks: Towards a Unifying Theory, *Submitted to Transportation Science* **29**(4).

Cascetta, E.: 1984, Estimation of Trip Matrices from Traffic Counts and Survey Data: A Generalised Least Squares Estimator, *Transportation Research* **18B**(4/5), 288–299.

Cascetta, E. and Cantarella, G.: 1991a, A Day-to-Day and Within-Day Dynamic Stochastic Assignment Model, *Transportation Research* **25A**.

Cascetta, E. and Cantarella, G.: 1991b, Modelling Dynamics in Transportation Networks. Part II. Within-Day Dynamic Models, *Submitted to Transportation Science* .

Cascetta, E., Inaudi, D. and Marquis, G.: 1993, Dynamic Estimators of Origin-Destination Matrices using Traffic Counts, *Transportation Science* **27**(4).

Cascetta, E., Nuzzolo, A., Russo, F. and Vitetta, A.: 1997, A Modified Logit Route Choice Model Overcoming Path Overlapping Problems: Specification and Some

Calibration Results for Interurban Networks, *Transportation and Traffic Theory* .

Cascetta, E. and Postorino, M.: 2001, Fixed Point Approaches to the Estimation of O/D Matrices using Traffic Counts on Congested Networks, *Transportation Science* .

Florian, M., Astarita, V., Er-Rafia, K., Mahut, M. and Velan, S.: 2001, A Comparison of Three Methods for Dynamic Network Loading, *Transportation Research Board* .

Gelb, A. (ed.): 1974, *Applied Optimal Estimation*, MIT Press.

Greene, W.: 2000, *Econometric Analysis*, 4th edition edn, Prentice Hall.

Hawas, Y.: 2000, Calibrating Simulation Models for ATIS/ATMS Applications, *Submitted for publication in Transportation Research* .

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

He, R., Miaou, S., Ran, B. and Lan, C.: 1999, Developing an On-Line Calibration Process for an Analytical Dynamic Traffic Assignment Model, *78th Annual Meeting of the Transportation Research Board* .

He, R. and Ran, B.: 2000, Calibration and Validation of a Dynamic Traffic Assignment Model, *79th Annual Meeting of the Transportation Research Board* .

Kunde, K.: 2002, *Calibration of the Supply Simulator in a Dynamic Traffic Assignment System (forthcoming)*, Master's thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Massachusetts Institute of Technology: 2000a, Development of a Deployable Real-Time Dynamic Traffic Assignment System, Executive Summary: DynaMIT and DynaMIT-P, Version 0.90, *Technical report*, Submitted to Oak Ridge National Laboratories.

Massachusetts Institute of Technology: 2000b, DTA System Enhancement and Evaluation at Traffic Management Center, Task P: Framework for the Use of DynaMIT in Transportation Planning Applications (DynaMIT-P), *Technical report*, Submitted to Oak Ridge National Laboratories.

Ramming, S.: 2001, *Network Knowledge and Route Choice*, PhD thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Transportation Research Board: 2000, *Highway Capacity Manual: Special Report 209*, third edn, National Research Council, Washington, D.C.

Van der Zijpp, N. J. and Lindveld, C. D. R.: 2001, Estimation of O-D Demand for Dynamic Assignment with Simultaneous Route and Departure Time Choice, *Transportation Research Board* .