Observability of Origin-Destination Matrices for Dynamic Traffic Assignment

by

Ashish Gupta

B.Tech. in Civil Engineering (2003)
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

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2005

Author .....
Department of Civil and Environmental Engineering
June 30, 2005

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

Accepted by .....
Andrew Whittle
Chairman, Department Committee on Graduate Studies

Observability of Origin-Destination Matrices for Dynamic Traffic Assignment

by

Ashish Gupta

Submitted to the Department of Civil and Environmental Engineering
on June 30, 2005, in partial fulfillment of the
requirements for the degree of
Master of Science in Transportation

Abstract

The estimation of dynamic Origin-Destination (O-D) matrices from aggregated sensor counts is one of the most important and well-researched problems in Dynamic Traffic Assignment (DTA) systems. In practice, more often than not, number of sensors are far less than the number of potential O-D pairs, and hence this problem is modeled in an optimization framework as function of historical estimates of O-D flows. However, in the absence of reliable historical O-D flows, it is critical that O-D estimation module is *observable*. Observability is defined as a property of the system by which it is possible to uniquely determine the (initial) state (O-D flows) of the system eventually by making regular indirect measurements of the state.

In DTA systems, observability implies that given enough sensor data, it is possible to uniquely determine O-D flows without any prior information about them. This thesis develops a methodology to verify the observability property of the O-D estimation model given limited sensor coverage on the network. A case study involving a large-scale network from Los Angeles, California is used to demonstrate the feasibility of the proposed approach. A comprehensive off-line calibration exercise for the same network is then used to verify the validity of the conclusion.

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

Acknowledgments

I would like to express my deepest gratitude and acknowledgment for my thesis advisor Prof. Moshe Ben-Akiva for his valuable guidance and constant support during my work at Intelligent Transportation Systems Program at MIT. It has been a privilege to work with and learn from a great transportation researcher and professor such as him.

I also want to put on record my appreciation for all professors of Department of Civil Engineering at MIT, particularly Prof. Nigel Wilson and Prof. Odoni, for making my education at MIT a pleasant experience. I am also grateful to Leanne Russell and Cynthia Stewart for their help in all matters pertaining to my status as a student and as a researcher at MIT.

I would also like to thank my coworkers at ITS Lab for their help in various stages of my work. A very special thanks is due to Rama for his regular guidance, constructive comments and help with Linux & Latex. I also thank Yang Wen for his insightful tips into DynaMIT code; Charisma for her tips on MITSIM, sharing with me my love of music(!) & get-together that she diligently organized for the whole lab; Costas for his help with Linux and MySQL; Tomer for his generous suggestions on anything and everything under the sun; Bhanu for his great friendship; and Deniel for our discussions on DynaMIT and 1.202 coursework that we shared. I also wish Anita and Gunwoo good luck for their future studies.

A special thanks to Federal Highway Administration (FHWA), Los Angeles Department of Transportation (LADoT) and Volpe Center for Transportation for their financial support without which my study at MIT would not have been feasible. I also wish to acknowledge Dr. Scott Smith for his contributions towards the goal of this thesis.

I cannot but thank enough my friends at MIT for making this a lifetime experience for me. I would like to specially mention Lakshman, Jaykumar, Vinod, Hemant, Abhinandan, Jaffery, Travis, Mithila, Sreeja, Saikat, Smeet, Chintan, Hetal and Vivek for everything they have (or have not) done for me!

Before I finish, I must convey my heartfelt gratitude to my family for their constant and unquestioned support. I thank the God for helping me achieve this milestone of my life and always being there for me. Finally, I would like to apologize and thank again to anyone whom I have inadvertently missed in this vote of thanks.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Ever growing traffic congestion and the externalities associated with it have forced transportation engineers and researchers around the world to drift from solutions involving increasing supply to solutions involving managing existing supply and demand. Intelligent Transportation Systems (ITS) have emerged as a key field in this direction providing the tools and technologies to ease our travels by "intelligently" managing all components of transportation, including planning, design, operation, maintenance, pollution and congestion. One of the applications of ITS is to provide pre-trip and en-route travel information to enable informed decision making and divert traffic away from the congested regimes, both in space and in time. A necessary requirement of such systems is the ability to predict future traffic, so as to proactively respond to drivers' concerns and avoid potentially undesirable network conditions such as long delays, congestions, queuing and excessive pollution. Dynamic Traffic Assignment (DTA) systems are the product of research in this area which attempt to model, simulate and forecast network conditions under varying traffic demand and influencing factors such as weather, special events, construction activities, accidents, etc.

DTA systems model complex and dynamic interactions between transportation demand and supply to effectively anticipate network reactions to expected or unexpected demand fluctuations. One of the key factors influencing reliable deployment of such systems is appropriate calibration. An effective calibration captures the charac-

teristics specific to demand-supply interactions in the study area under consideration in the form of estimated values of parameters of numerous mathematical models which DTA systems are comprised of. Real-time applications of DTA systems usually utilize inputs from the traffic surveillance system, along with calibrated set of historical database, to estimate and predict the demand on the network, represented in the form of Origin-Destination (O-D) flows.

Given a set of route-choice parameters, estimation of O-D matrix from sensor counts is simply a solution of linear system of equations mapping O-D flows to sensor counts, with assignment matrix[1] as mapping function. However, in practice, this problem is compounded by the following observations:

- Route-choice parameters are the part of calibration exercise and hence unknown beforehand,
- Number of sensors are almost always much less than the number of potential O-D pairs, and
- Sensor counts are not completely error-free

Various methods have been proposed in the literature [5, 10, 16, 17] to solve O-D estimation problem[2]. Most common and accepted approach is to model the problem as nonlinear optimization framework trying to minimize some function of difference between estimated and observed quantities. All these formulations, however, rely critically on the availability of reliable set of time dependent historical O-D matrices. Therefore in situations where such information is not available, or if available, is not reliable, the need for an alternate formulation which is independent of historical (starting) O-D matrix is of significant importance.

[1] Assignment Matrix is a time-dependent matrix, of dimension equal to number of sensors by number of O-D pairs, of fractions of O-D flows which are counted by the sensors.

[2] For comprehensive literature review on O-D estimation, the reader is referred to Ashok [3].

1.1 Motivation

Real-time applications of Dynamic Traffic Assignment systems utilize real-time traffic surveillance input with calibrated set of parameters and historical information to estimate and predict state of the network. The predicted O-D and link flows are then used to generate consistent and reliable route guidance. Thus, the credibility of such systems relies on the ability to accurately estimate and anticipate traffic under congested flow conditions. It is notable that drivers' reactions to route-guidance may merely cause spatial and temporal shifts from predicted network state, thus invalidating the very prediction that influenced the guidance in the first place. Therefore consistency of DTA systems is pivotal to successful application in the field. Consistency implies that predicted network state matches exactly what drivers would experience on the network. In order to ensure that the system behaves consistently and in expected manner when deployed at site, calibration of the models against field data collected from site of actual deployment cannot be overemphasized.

The motivation for this thesis, therefore, arises from the unavailability of historical information on Origin-Destination patterns and flow, which are essential ingredients of existing calibration exercises. Since the estimation of a set of historical O-D matrices is crucial to identify the demand and congestion pattern in the network, and because the number of sensors are almost never enough to uniquely determine the O-D flows, unavailability of prior knowledge in this regard makes the calibration dependent on *assumed* starting O-D flows. It is not difficult to spot the deficiency of such an approach, since even the best guesses are expected to be considerably far from demand patterns in the field. Also, different starting points might yield completely different final results, making it difficult, if not impossible, to select historical O-D flows for on-field deployment.

Thus it is critical to identify and verify the uniqueness of estimated O-D patterns, before embarking on a full-fledged calibration of DTA system with field data.

1.2 Problem Definition and Thesis Focus

Though O-D information is indirectly available from archived sensor counts, under-specified system of linear equations renders the direct computation infeasible. However, if the system is *observable*[3] then it is possible to uniquely estimate O-D flows irrespective of the starting point. Specifically,

> If *every potential* O-D flow is counted by *at least* one sensor, then given *sufficiently enough* data—consisting of observed sensor counts—it is possible, in theory, to estimate O-D flows *uniquely* and *independent of* starting O-D matrix.

The explanation of several terms emphasized above is in order here. Since we seek to reproduce the true O-D matrix without any information a-priori, only information we have regarding the O-D flows must come from the observed surveillance data. Hence it is important that significant fraction of flow for each O-D pair must be counted on the network, preferably by multiple sensors, but at least by one sensor. In other words, we simply cannot estimate O-D pairs for which we neither have prior information nor are they observed by any of the traffic surveillance systems. Efficiency of the estimation is also a complex function of number of sensors counting any particular O-D pair and the fraction of O-D pair's flow that are counted. More sensors counting an O-D pair reduces the linear dependency among observed counts, and each additional observation has stabilizing effect on the estimated value. Similarly, higher the fraction of O-D flow counted, more reliable is the estimation. This is because the estimation model becomes less and less dependent on the errors in the route choice model and on inherent randomness of the phenomenon.

In the absence of any a-priori knowledge on the pattern of O-D flows, calibration must start with considering all possible O-D pairs, that is, from all possible origins to all possible destinations within the scope of the network. Now that we want to extract O-D information "hidden" in the observed counts, and that the number of such counts are significantly lower than the number of potential O-D flows, it is

[3]Comprehensive treatment of Observability is presented in the Chapter 2.

obvious that multiple intervals of data is required so that converged set of O-D flows is independent of starting O-D matrix. In theory, it means that *eventually* it is possible to achieve same set of estimated O-D flows irrespective of the O-D matrix we start with. In practice, however, it implies that we can achieve acceptable level of accuracy in the estimation of O-D matrices after calibration with few intervals of data. A rough estimate of the number of intervals of data required to achieve observability is the ratio of number of potential O-D pairs and number of usable sensors on the network.

This thesis attempts to overcome the dependency of calibration framework on the set of historical O-D flows by developing methodology to test if it is possible to estimate the same O-D matrix irrespective of the starting conjecture on the magnitude or the pattern of the O-D flows. A case study involving practical-size network from Los Angeles is used to verify the conclusion.

1.3 Literature Review

Literature on observability in O-D estimation is very limited and often linked with the sensor location problem. Since estimation of the O-D matrices in DTA systems heavily relies on sensor counts to update or to estimate flows in the network, sensor location problem seeks to answer questions such as:

- Given a transportation network, what is the minimum number, and the corresponding locations, of the sensors which will maximize the efficiency of O-D estimation process. Alternatively stated, what is the minimum number of sensors which will result in some predefined level of accuracy in O-D estimation. One more variant of this question could also assume some given sensor coverage and seek to optimize the locations of additional set of sensors.

- Given the maximum number of sensors to locate, what are the optimal links so as to maximize the efficiency of O-D estimation process. This question can again be paraphrased to include the threshold of acceptance or prior sensor coverage.

Both of these questions are applicable in different scenarios in planning applications. Second question raised above is actually an indirect way of representing financial constraints. Because of their mutual dependence, the sensor location problem and the observability issue are closely related.

Though Ashok and Ben-Akiva [4] proposed a methodology of using Kalman Filter and Extended Kalman Filter to estimate O-D flows efficiently and in real time, the observability problem has received little attention. The authors, however, do mention that under conditions of observability, the influence of initial value of the state vector in the proposed state-space model would disappear with time, and that such a feature is desirable in DTA systems. They also list several factors affecting the observability of the estimation problem. While the ratio of the number of sensors to the number of O-D pairs to be estimated is one important factor, the authors also emphasize the degree of linkage between O-D flows and counts in the form of the assignment matrix, and the degree of linkage between O-D flows over time in the form of the transition matrix.

For example, if an entire column of the assignment matrix is zero then it implies that the corresponding O-D flow never gets measured. In this case it is not possible to estimate the flow, unless it is related to other measured O-D flows in some known way. Similarly, in cases where entries corresponding to a particular O-D pair in assignment matrix are small enough that effect of that pair towards sensor counts cannot be unambiguously identified, estimation of O-D flows is often subject to high variance. Linkage of O-D flows over time in form of transition matrix is important in simultaneous estimation[4] procedures where an O-D matrix is repeatedly estimated as many times as the degree of the corresponding state-space model.

Yang *et al.* [27] propose a measure of efficiency and reliability of estimated O-Ds in form of "Maximum Possible Relative Error" (MPRE) index. Since the true O-D matrix is unknown, the authors claim that MPRE bounds possible relative error in

[4]Since information of O-D flows in an interval is available not only from sensor counts for that interval but also from sensor counts from previous intervals, simultaneous estimation procedure re-estimates an O-D matrix multiple times. Though this is theoretically advantageous over sequential estimation process, computational costs are often prohibitive to use in real-time context. Detailed discussion on simultaneous estimation is presented in [6].

the estimated O-D matrix and forms basis of comparison among various estimation techniques. Based on the assumptions of perfect route choice model and the error-free sensor observations, authors develop a quadratic optimization model to compute MPRE as,

$$\begin{aligned} \text{Maximize } \phi_\nu(\lambda) &= \lambda \cdot \lambda' \\ \text{subject to } PHT &= 0 \\ \lambda_i &\geq -1 \; \forall \; i = 1, 2, \ldots, n. \end{aligned} \tag{1.1}$$

where $\sqrt{\phi_\nu(\lambda)}$ is defined as MPRE, P is known assignment matrix mapping O-D flows to link flows, T is the vector of estimated O-D, and

$$H = \begin{bmatrix} \lambda_1 & & 0 & 0 \\ & \lambda_2 & & 0 \\ 0 & & \ddots & \\ 0 & 0 & & \lambda_n \end{bmatrix} \tag{1.2}$$

where n is the number of O-D pairs, and

$$\lambda_i = \frac{T_i^* - T_i}{T_i} \tag{1.3}$$

where T_i^* is unknown true i^{th} O-D value and T_i is its estimated equivalent.

Authors note that "if the trips between an OD pair (or several OD pairs) are not observed at any traffic counting points, the Maximum Possible Relative Error in the estimated OD matrix will become infinite", thereby proposing the OD Covering Rule. The **OD Covering Rule** states that the traffic counting points on the network must be located so that the trips between any OD pair can be observed. The authors also propose an index of reliability of the estimated O-D as

$$Re(T) = \frac{1}{1 + \text{MPRE}} \qquad (\text{MPRE} \geq 0) \tag{1.4}$$

such that $Re(T) = 1$ for MPRE $= 0$ and $Re(T) = 0$ as MPRE $= +\infty$. The paper concludes by observing that the number and locations of counting points have significant impact on the reliability of O-D matrix, and that as number of links counted increases in the network, MPRE tends to approach real estimation error.

These findings are also confirmed by Yang *et al.* [28] in a seminal paper where authors extend previous work in this area by proposing three more rules apart from the O-D Covering Rule mentioned above. Their rule 1, **O-D Covering Rule** reiterates that the traffic counting points on the network should be located so that certain portion of trips between *any* O-D pair are observed. The second rule, **Maximum Flow Fraction Rule**, specifies that sensor should be located such that the maximum possible fraction of trips between any particular O-D pair can be captured. Paraphrasing this for certain number of links, **Maximal Flow Intercepting Rule** says that chosen set of links should intercept as many O-D pairs as possible. And finally, **Link Independence Rule** requires that counted traffic on chosen links should be mutually independent. The authors prioritize O-D Covering Rule and Link Independence Rule as fundamental rules that must be satisfied while other two rules can be incorporated in the objective function for maximizing their compliance. Assuming known route choice probabilities and turning fractions, [28] also develop a mixed integer program and suggest some heuristics to solve it.

Bianco *et al.* [8] attempt to approach sensor location problem from a graph theoretical perspective and propose a two-stage procedure that "first derives the complete traffic flow vector on a transport network and then produces a reliable estimate of Origin/Destination (O/D) trip matrix" based on minimal-cost set of traffic counts. Authors propose a heuristic algorithm based on the combined cutset principle of graph theory and claim that "O/D estimation error is always bounded, O/D Covering Rule being satisfied or not". They, however, also assume the perfect knowledge of turning fractions, and that the sensors are located at the nodes, instead of on the links.

All these approaches critically depend on the prior knowledge of the true route choice parameters, turning fractions and error-free observations, and solve sensor location problem for static O-D estimation while ignoring traffic dynamics. However,

since route choice parameters are the outcome of calibration process, it is not possible in practice to know their true values a-priori.

Related works in locating transportation system control and communication structures focus on locating Automatic Vehicle Identification Readers ([1]), image sensors ([14]) and Variable Message Signs ([18] and [19]). However various specific characteristics of the sensors in terms of the performance and different criteria for the optimality in O-D estimation necessitates the exclusive study of observability and sensor location problem in context of Dynamic Traffic Assignment Systems.

1.4 Thesis Outline

The rest of this thesis is organized as follows. Chapter 2 introduces the concept of observability in detail and presents the theoretical and empirical procedure involved in testing observability. This chapter also describes the comprehensive calibration methodology treating both demand and supply calibration of DTA systems. The following chapter presents a case study expounding the methodologies presented in previous chapter. We present results of observability tests and off-line calibration on the Los Angeles network. Finally, Chapter 4 summarizes the findings of this study and proposes further topics of research that can be explored in related areas.

Chapter 2

Observability and Calibration Framework

This chapter is divided into two main parts. In the first part, we introduce the concept of observability in the context of dynamic control systems and discuss the implications of observability in answering questions pertaining to changing system states over time. We then proceed to identify observable systems and further describe how this concept is relevant for Dynamic Traffic Assignment systems. Subsequently we will elaborate on the steps involved in verifying observability in the context of DTA system and present both theoretical and empirical methodologies for the same.

In the second part of this chapter we will dwell upon the calibration of demand and supply parameters and outline a framework for systematic calibration process. The primary objective of calibration process for a DTA system is to identify the values of parameters and inputs which, when used in field with real-time data, will replicate reality as closely as possible. Since various models and parameters in simulation based DTA systems do not often have closed mathematical form, calibration generally involves starting with an initial set of parameters and running the DTA system to obtain simulated measures of performance, which are then compared against observed data from field and some sort of objective function is evaluated. Initial parameters' estimates are then modified based on some form of search algorithm and a new objective function is computed. This iterative process continues until a

predefined threshold of level of convergence is achieved.

2.1 Observability

Observability is a well known concept applicable to Dynamic Control Systems. Observability is the property of such systems and is based on system configuration and dynamics. Before we introduce observability, however, it would be prudent to briefly describe such systems and their characteristics.

2.1.1 Dynamic Control Systems

Control Systems are physical systems which have a characteristic set of parameters (or variables) representing the state of the system at any instant, and whose change of state from one instant to another is governed by certain set of rules and processes, which may or may not be influenced by external factors (or forces). The state of the system is represented by a vector of values of those parameters of the system which uniquely and unambiguously identify the condition of the system at any point in time. All state variables of a system are mutually linearly independent, and are the "smallest possible subset of system variables that can represent the entire state of system at any given time" ([26]). The minimum number of such variables defines the degree of that control system.

These systems are most conveniently represented by what is known as a state-space formulation. A state-space model typically consists of two equations: the measurement equation (2.1) and the transition equation (2.2). Measurement equation, on one hand, maps the state of the system at any instant to the vector of measurements for the system at that instant, while transition equation governs the transition of the system from one system state to another:

$$\mathbf{y}(t) = \mathcal{C}(t)\mathbf{x}(t) + \mathcal{D}(t)\mathbf{u}(t) \tag{2.1}$$

$$\dot{\mathbf{x}}(t) = \mathcal{A}(t)\mathbf{x}(t) + \mathcal{B}(t)\mathbf{u}(t) \tag{2.2}$$

For a system with n inputs, m outputs and p state variables, $\mathcal{A}$ is a $p \times p$ **state matrix**, $\mathcal{B}$ is a $p \times n$ **input matrix**, $\mathcal{C}$ is a $m \times p$ output or **measurement matrix**, and $\mathcal{D}$ is a $m \times n$ **feed forward matrix**. Also, $\mathbf{x}$ is a $p \times 1$ **state vector**, $\mathbf{y}$ is a $m \times 1$ output or **measurement vector** and $\mathbf{u}$ is a $n \times 1$ input or **control vector**. $\dot{\mathbf{x}}$ is the derivative of $\mathbf{x}$ with respect to time t.

Based on the nature of these equations (2.1, 2.2), state-space models can be classified as linear or nonlinear models. Further, such systems can be classified as static or dynamic, depending on whether the parameters of measurement and transition matrices (i.e. matrices $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$) remain unchanged over time. If the state variables can take values in continuous domain, the system is categorized as a continuous system, otherwise it is classified as a discrete system. An alternative version of the transition equation might also involve a random forcing function $\mathbf{w}(t)$ and a corresponding matrix. Note that, while a set of state variables are sufficient to describe the complete state of the system at any instant, the state vector is not a unique set of variables. Any other set $\mathbf{x}'(t)$ related to $\mathbf{x}(t)$ by a nonsingular transformation

$$\mathbf{x}'(t) = \mathcal{E}(t)\mathbf{x}(t)$$

can also be considered a state vector. However, some vectors are preferred over others because they are easier to interpret and deal with.

We conclude our introduction of control systems by providing a simple example[1] involving a common spring-mass-damper system (Figure 2-1) from physics. Consider a block of mass m rolling frictionlessly on a plane and connected to the rigid left wall by a spring with spring coefficient k and a damper with damping coefficient c. The whole setup is subjected to acceleration $\mathbf{w}(t)$ to the right as shown in Figure 2-1. Newton's Second Law of motion yields

$$m\ddot{\mathbf{x}} + c\dot{\mathbf{x}} + k\mathbf{x} = m\mathbf{w}(t) \tag{2.3}$$

where $\mathbf{x}$ is the displacement measured positive from left. We can represent this system

[1]This example is adopted from [13].

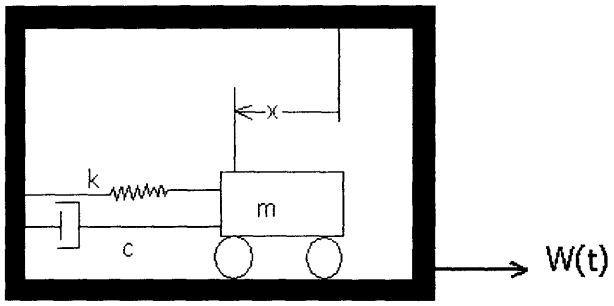


Figure 2-1: Spring-Mass-Damper System as Second-Order Control System ([13])

as state-space formulation by defining state as

$$\overline{\mathbf{x}} = \begin{bmatrix} \mathbf{x} \\ \dot{\mathbf{x}} \end{bmatrix} \tag{2.4}$$

The complete equation for system dynamics then can be written as:

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \ddot{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -k/m & -c/m \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \dot{\mathbf{x}} \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{w}(t) \end{bmatrix} \tag{2.5}$$

Now that we are acquainted with the definitions and terms of a control system, we proceed to define observability and discuss its properties in the next section.

2.1.2 Definition of Observability

Let us rewrite the measurement equation (2.1) for discrete control systems without an input vector such that measurements are indirectly dependent on the state of system, plus some measurement noise as in equation (2.6),

$$\mathbf{y}_k = \mathcal{C}_k \mathbf{x}_k + \mathbf{v}_k \tag{2.6}$$

where k is a step indicator. Then, given a sequence of measurements $\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_k$, observability defines our ability to determine $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_k$ from these measurements.

Consider the discrete deterministic n-order system, with transition equation as

$$\mathbf{x}_{k+1} = \Phi \mathbf{x}_k \tag{2.7}$$

and n noise-free measurements given by

$$\mathbf{y}_k = \mathcal{C}\mathbf{x}_k, \forall\ k = 0, 1, 2, \ldots, n-1 \tag{2.8}$$

Then we may write,

$$\begin{aligned}
\mathbf{y}_0 &= \mathcal{C}\mathbf{x}_0 \\
\mathbf{y}_1 &= \mathcal{C}\mathbf{x}_1 = \mathcal{C}\Phi\mathbf{x}_0 \\
\mathbf{y}_2 &= \mathcal{C}\mathbf{x}_2 = \mathcal{C}\Phi\mathbf{x}_1 = \mathcal{C}\Phi^2\mathbf{x}_0 \\
&\vdots \\
\mathbf{y}_{n-1} &= \mathcal{C}\mathbf{x}_{n-1} = \mathcal{C}\Phi^{n-1}\mathbf{x}_0
\end{aligned} \tag{2.9}$$

And, therefore,

$$\begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{n-1} \end{bmatrix} = \begin{bmatrix} \mathcal{C} \\ \mathcal{C}\Phi \\ \vdots \\ \mathcal{C}\Phi^{n-1} \end{bmatrix} \mathbf{x}_0 = \Gamma\mathbf{x}_0 \tag{2.10}$$

Here Φ^n is n^{th} power of matrix Φ. Thus, if $\mathbf{x}_0$ is to be estimated uniquely from measurements $\mathbf{y}_0$, $\mathbf{y}_1$, ..., $\mathbf{y}_k$, the matrix Γ must be nonsingular. Or alternately stated, the given system is observable *if and only if* the matrix

$$\Gamma = \begin{bmatrix} \mathcal{C} \\ \mathcal{C}\Phi \\ \vdots \\ \mathcal{C}\Phi^{n-1} \end{bmatrix} \tag{2.11}$$

is of full rank n. Similarly, for a control system defined by the set of equations

(2.1, 2.2), for known input vector $\mathbf{u}$, observability implies that

$$\text{rank} \begin{bmatrix} \mathcal{C} \\ \mathcal{C}\mathcal{A} \\ \vdots \\ \mathcal{C}\mathcal{A}^{n-1} \end{bmatrix} = n \tag{2.12}$$

Now that we have laid the mathematical background for observability, we are ready to formally define observability. As per Gelb [13],

> "A system is observable at time $t_1 \geq t_0$ if it is possible to determine $\mathbf{x}(t_0)$ by observing $\mathbf{y}(t)$ in the interval (t_0, t_1). If *all* states $\mathbf{x}(t)$ corresponding to all $\mathbf{y}(t)$ are observable, the system is *completely observable.*"

Alternatively, [23] defines *un*observability as

> "A state q of a finite dimensional dynamic system is said to be *unobservable* over $[0, T)$ if, with $x(0) = q$ and for every $u(t)$ over $[0, T)$, we get the same $y(t)$ as we would with $x(0) = 0$, i.e. an unobservable initial condition cannot be distinguished from the zero initial condition. The *dynamic system* is called *unobservable* if it has an unobservable state, otherwise it is called *observable.*"

In other words, observability property holds for a system if it is possible to uniquely identify initial state of the system given system characteristics (measurement and transition equations) and sufficient indirect measurements data. Obviously, if a system is observable at any time t, it is observable at all times.

2.1.3 Review of Sequential O-D Estimation Techniques

This section reviews the common O-D estimation methodologies before we proceed to define observability in context of O-D estimation. The O-D estimation algorithm

in DTA systems can be generalized in following mathematical form,

$$\begin{aligned} \mathbf{x}_h &= argmin[\mathbf{F}_1(\mathbf{x}_h, \mathbf{x}_h^a) + \mathbf{F}_2(\sum_{p=h-p'}^{h} \mathbf{a}_h^p \mathbf{x}_p, \mathbf{y}_h)] \\ \mathbf{a}_h^p &= g(\mathbf{x}_p, \beta, \gamma, tt^{eq}) \\ tt^{eq} &= h(\beta, \gamma, \mathbf{x}_p) \end{aligned} \tag{2.13}$$

where $\mathbf{x}_h$ is a vector of estimated O-D flows for interval h, and $\mathbf{x}_h^a$ is corresponding a-priori (also, target or seed) vector. $\mathbf{y}_h$ is a vector of observed counts, and $\mathbf{a}_h^p$ is the assignment matrix which maps contributions of O-D flows departing in interval p to sensor counts measured in interval h. This assignment matrix itself is an outcome of the network supply simulator and hence, a function of O-D flows in the interval p ($\mathbf{x}_p$) along with route choice parameters (β), supply parameters (γ), and equilibrium travel time (tt^{eq}). Functions $\mathbf{F}_1$ and $\mathbf{F}_2$ measure the "distance" between observed and estimated quantities. Furthermore, fixed point nature of this process is completed by tt^{eq} being function of route choice and supply parameters and O-D flows.

We briefly describe the iterative approach to demand calibration[2] here which starts with calibration of route choice parameters. In absence of disaggregate route choice survey data, we attempt to calibrate path choice parameters from aggregate sensor information. Hence procedure starts with the assumption of default values of route choice parameters, along with a-priori O-D flows and historical travel times—or in absence, with free flow travel times—to simulate a set of time dependent assignment matrices. These assignment matrices and corresponding set of O-D flows are passed to O-D estimation module as described by equations (2.13) to generate new set of O-D flows. This process continues iteratively until convergence at two level (Figure 2-2) of assignment matrix and route choice parameters is achieved.

[2]For more detailed description of O-D estimation problem and demand calibration, please refer to [6] and [7].

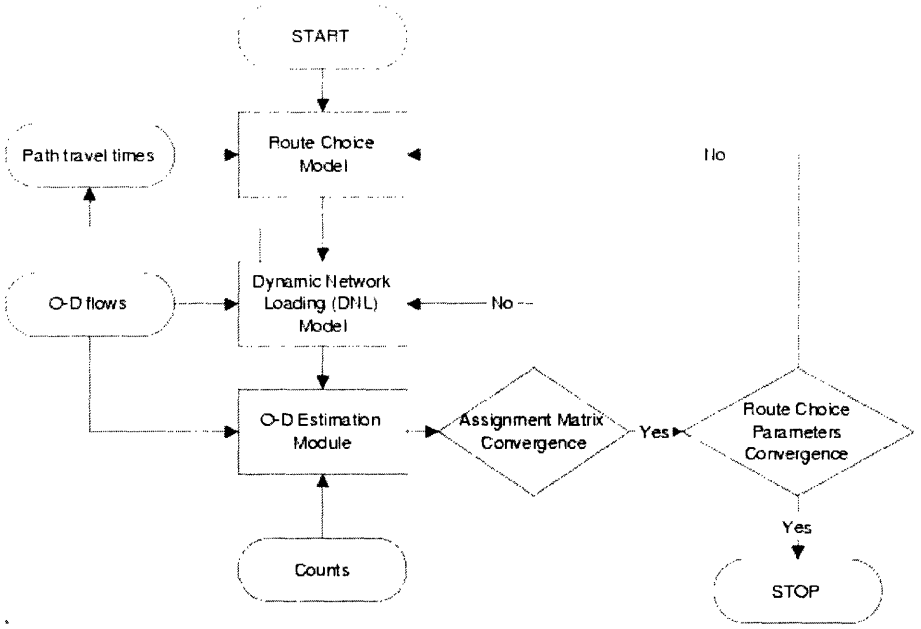


Figure 2-2: Iterative Calibration Framework

Generalized Least Square Formulation

We present here the Generalized Least Square (GLS) formulation of O-D estimation algorithm presented by equations (2.13) where functions $\mathbf{F}_1$ and $\mathbf{F}_2$ are replaced by least square distance between corresponding observed and simulated quantities, and optimization takes form of GLS Minimization problem. O-D estimation then can be interpreted as two-part objective function which seeks to minimize the (weighted) discrepancy between:

- Estimated O-D flows and a-priori O-D flows, and
- Simulated sensor counts and observed sensor counts

First part of the objective function, which attempts to keep estimated O-D flows as close to their target values as possible, is termed as *direct measurement*, with associated error vector. By definition, a direct measurement provides a preliminary estimate of an O-D flow. Mathematically speaking, direct measurements are expressed as:

$$\mathbf{x}_h^a = \mathbf{x}_h + \mathbf{u}_h \tag{2.14}$$

Here $\mathbf{u}_h$ is the error between the estimated and a-priori O-D flows, which we seek to minimize.

Second part of the objective function, which attempts to keep simulated sensor counts as close to observed counts as possible, is termed as *indirect measurement.* Mathematically, we attempt to minimize error $\mathbf{v}_h$ associated with indirect measurements in following equation:

$$\mathbf{y}_h = \hat{\mathbf{y}}_h + \mathbf{v}_h \tag{2.15}$$

where $\mathbf{y}_h$ are observed sensor counts and $\hat{\mathbf{y}}_h$ are corresponding simulated values. However, since simulated sensor counts are nothing but mapping of estimated O-D matrices on the network through assignment matrices, equation (2.15) can also be written as

$$\mathbf{y}_h = \sum_{p=h-p'}^{h} \mathbf{a}_h^p \mathbf{x}_p + \mathbf{v}_h \tag{2.16}$$

where p' is the maximum number of intervals required by any vehicle to cross the whole network. Note that the sensor counts in interval h are contributed by flows from all previous intervals starting from $h - p'$ up to current interval h. However, in a sequential estimation procedure, we already have the best available O-D estimates from previous interval, and hence equation (2.16) can be modified to the following form,

$$\mathbf{y}_h - \sum_{p=h-p'}^{h-1} \mathbf{a}_h^p \hat{\mathbf{x}}_p = \mathbf{a}_h^h \mathbf{x}_h + \mathbf{v}_h \tag{2.17}$$

where $\hat{\mathbf{x}}_p$ is the best estimate so far of the flows in interval p.

Combining equations (2.14) and (2.17), and reformatting in matrix algebra, we obtain:

$$\begin{bmatrix} \mathbf{y}_h - \sum_{p=h-p'}^{h-1} \mathbf{a}_h^p \hat{\mathbf{x}}_p \\ \mathbf{x}_h^a \end{bmatrix} = \begin{bmatrix} \mathbf{a}_h^p \\ \mathbf{I}_{n_{O-D}} \end{bmatrix} \mathbf{x}_h + \begin{bmatrix} \mathbf{v}_h \\ \mathbf{u}_h \end{bmatrix} \tag{2.18}$$

where n_{O-D} is the number of O-D pairs in the network. If the variance-covariance matrix of errors in $\mathbf{v}_h$ be $\mathbf{V}_h$, and same for errors in $\mathbf{u}_h$ be $\mathbf{W}_h$, then GLS solution

for formulation represented by (2.18) can be written as

$$\hat{\mathbf{x}}_h = argmin[(\mathcal{Y}_h - \mathbf{A}_h\mathbf{x}_h)'\Omega_h^{-1}(\mathcal{Y}_h - \mathbf{A}_h\mathbf{x}_h)] \tag{2.19}$$

where $\mathcal{Y}_h$ and $\mathbf{A}_h$ are appropriate augmented vectors/matrices from (2.18), and

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

Note that sequential O-D estimation procedure is dependent on a-priori estimates of O-D flows. These a-priori flows can be estimated by various methods ([6]), which may or may not be dependent on historical estimates of O-D flows. Nevertheless, there is obvious need for starting O-D matrix which will trigger our sequential GLS estimation for first interval.

Equivalent State-Space Formulation

O-D estimation problem formulated as GLS problem above can also be looked as equivalent state-space formulation as discussed in [3]. Using concept of deviations[3] proposed by [4], and using autoregressive factors to estimate a-priori O-D matrix from previous historical and estimated O-D matrices, we can rewrite equation (2.14) as

$$\partial\mathbf{x}_h^a = \partial\mathbf{x}_h + \mathbf{u}_h \tag{2.20}$$

where

$$\partial\mathbf{x}_h^a = \sum_{p=h-q'}^{h-1} \mathbf{f}_h^p \partial\hat{\mathbf{x}}_p \tag{2.21}$$

[3]Historical database of O-D flows posses wealth of structural (spatial and temporal) information synthesized from several days. Use of deviations from historical flows ($\mathbf{x}_h - \mathbf{x}_h^H$) instead of historical flows themselves as state variables provides a way to utilize the structural information in O-D estimation formulation. Also, as our variables are now deviations which can be both positive or negative, such formulation possesses various desirable properties which are very useful in estimation. One example of such property is that error in direct measurement equation in GLS formulation is more justified to be assumed to be normally distributed.

and $\partial \mathbf{x}_h$ denotes deviation in $\mathbf{x}_h$ from historical O-D ($\mathbf{x}_h^H$) and $\partial \mathbf{x}_h^a$ denotes its (deviations') a-priori O-D estimate. $\mathbf{f}_h^p$ represents an auto-regressive matrix encapsulating the effect of deviations in O-D flow departing in past interval p to those in interval h. q' is the degree of auto-regression.

Similarly, we can rewrite equation (2.16) as follows:

$$\mathbf{y}_h - \mathbf{y}_h^H = \sum_{p=h-p'}^{h} \mathbf{a}_h^p (\mathbf{x}_p - \mathbf{x}_p^H) + \mathbf{v}_h \tag{2.22}$$

where

$$\mathbf{y}_h^H = \sum_{p=h-p'}^{h} \mathbf{a}_h^p \mathbf{x}_p^H \tag{2.23}$$

Defining deviations in the O-D flows from historical estimates (i.e. $\partial \mathbf{x}_h$) as the state of the system, and using (2.20) as transition equation and (2.22) as measurement equation, we completely define O-D estimation as a state-space formulation. Ashok [3] notes that since sensor observations at time h contain information about not only $\mathbf{x}_h$ but also from prior O-D flows, each O-D matrix should be estimated multiple times to fully extract the information from the surveillance system. The author proposes an augmented state-space formulation using lagged O-D flows, and appropriately modified transition and measurement equations. We reproduce below only the results of such a formulation, and the reader is referred to [3] and [4] for a detailed discussion in this regard.

We can define vectors,

$$\mathbf{X}_h = [\mathbf{x}'_h \quad \mathbf{x}'_{h-1} \quad \dots \quad \mathbf{x}'_{h-s}]'$$
$$\mathbf{X}_h^H = \left[\mathbf{x}'^H_h \quad \mathbf{x}'^H_{h-1} \quad \dots \quad \mathbf{x}'^H_{h-s}\right]'$$

Correspondingly, the augmented state of the system is defined as

$$\begin{aligned} \mathcal{X}_h &= \mathbf{X}_h - \mathbf{X}_h^H \\ \Rightarrow \mathcal{X}_h &= [\partial \mathbf{x}'_h \quad \partial \mathbf{x}'_{h-1} \quad \dots \quad \partial \mathbf{x}'_{h-s}]' \end{aligned}$$

where $s = \max(p', q' - 1)$ is the order of this system.

Hence, the augmented transition equation can be written as:

$$\mathcal{X}_{h+1} = \mathbf{\Phi}_h \mathcal{X}_h + \mathbf{W}_{h+1} \tag{2.24}$$

where

$$\mathbf{\Phi}_h = \begin{bmatrix} & \mathbf{F}_h & \\ \mathbf{I}_{n^s_{O-D} \times n^s_{O-D}} & & \mathbf{0}_{n^s_{O-D} \times n_{O-D}} \end{bmatrix} \tag{2.25}$$

and,

$$\mathbf{F}_h = \begin{bmatrix} \mathbf{f}^h_{h+1} & \mathbf{f}^{h-1}_{h+1} & \dots & \mathbf{f}^{h-s}_{h+1} \end{bmatrix} \tag{2.26}$$

and n_{O-D} is the number of O-D pairs to be estimated, and n^s_{O-D} is its s^{th} power. $\mathbf{W}_{h+1}$ is an appropriately augmented matrix of error covariances.

Similarly, the augmented measurement equations are written as:

$$\mathcal{Y}_h = \mathbf{A}_h \mathcal{X}_h + \mathbf{v}_h \tag{2.27}$$

where

$$\mathbf{A}_h = \begin{bmatrix} \mathbf{a}^h_h & \mathbf{a}^{h-1}_h & \dots & \mathbf{a}^{h-s}_h \end{bmatrix} \tag{2.28}$$

and

$$\mathcal{Y}_h = \mathbf{y}_h - \mathbf{y}^H_h \tag{2.29}$$

Historical O-D flows ($\mathbf{X}^H_h$) are calculated as average of estimated flows ($\hat{\mathbf{X}}_h$) for multiple days using sequential GLS procedure described above. Transition matrix $\mathbf{\Phi}_h$ is obtained by linear in parameters regression using deviations in estimated flows from its historical value as dependent variable and similar deviations in previous intervals as independent variables as discussed in [3].

2.1.4 Observability in O-D Estimation

Following our discussion of observability in Section 2.1.2 and noting the state-space formulation presented in the previous section, it is clear that the O-D estimation

problem is observable *if and only if*

$$\text{rank}\begin{bmatrix} \mathbf{A}_0 \\ \mathbf{A}_1\mathbf{\Phi}_0 \\ \mathbf{A}_2\mathbf{\Phi}_0\mathbf{\Phi}_1 \\ \vdots \\ \mathbf{A}_{s-1}\prod_{r=0}^{s-2}\mathbf{\Phi}_r \end{bmatrix} = s \tag{2.30}$$

We now make an important observation. Note that both $\mathbf{A}_h$ and $\mathbf{\Phi}_h$ are dependent on the estimated O-D flows and are the outcome of calibration process, since correct auto-regressive factors and assignment matrices cannot be known unless correct demand is estimated. Hence, analytical verification of observable system is, unfortunately, not possible unless we have fully calibrated system! While we started examining observability of the system since we did not have prior information on O-D flows, let alone having a calibrated system, analytical formulation is intractable at least at this stage.

We, however, realize that while it may not be possible to explicitly calculate all measurement and state transition matrices, this information is "hidden" indirectly in the estimated O-D matrices after each interval of calibration. This is because $\mathbf{x}_h$ is estimated using simulated assignment matrix with $\hat{\mathbf{x}}_{h-1}$ as target flows. Thus, in principle, system can be proven to be observable if after sequential calibration from different starting O-D matrices over multiple intervals, we prove that the estimated O-D matrix is independent of the starting O-D matrix. Apart from its conceptual simplicity, this method verifies observability by first principles. As we start the calibration with different a-priori O-D matrices, and as we move from one interval to the next, we expect the estimated O-D flows to converge to the same values after several intervals.

Sequential O-D estimation process for an interval is briefly described as follows.

We start by calculating a-priori (target) O-D matrix as

$$\mathbf{x}_h^a = \begin{cases} \mathbf{x}_0 & \text{for } h = 1 \\ \hat{\mathbf{x}}_{h-1} & \text{otherwise} \end{cases} \tag{2.31}$$

where $\mathbf{x}_h^a$ and $\hat{\mathbf{x}}_h$ are a-priori and estimated O-D matrices respectively for interval h and $\mathbf{x}_0$ is the starting O-D matrix. Next we estimate O-D $\mathbf{X}_h$ using assignment matrix from the previous interval (or an assignment matrix based on free flow travel times in case of the first interval) using the GLS formulation. We then use a simulation tool, such as DynaMIT to calculate link travel times as experienced by drivers, given the input demand. Assignment matrix is then recalculated from the new set of link travel times and the latest route choice parameters. O-D flows for the current interval are then re-estimated. We continue this iteratively until convergence of O-D flows and assignment matrix is achieved. We proceed to the next interval $(h+1)$ with estimated flow $\hat{\mathbf{x}}_h$ as target OD.

Once we are able to identify the state (O-D flows) of DTA system, we can proceed to detailed calibration. Following sections outline the supply and demand calibration procedures which are demonstrated through a Case Study presented in the next chapter.

2.2 Supply Calibration

Calibration of parameters of supply simulation is specific to the type of model adopted for supply simulation. We use the DynaMIT[4] system as an example. DynaMIT implements mesoscopic supply simulator which models the movement of vehicles based on pre-specified speed-density relationships. Each segment in DynaMIT is comprised of two parts: a moving part, where cars move based on a calibrated relationship, and a queuing part, which keeps track of length, duration and extent of queuing. DynaMIT can handle specific lane groups for turning movement (defined through lane connections). Supply calibration is discussed in detail by Kunde [20].

[4] A brief introduction to DynaMIT is presented in Appendix A.

2.2.1 Speed-Density Relation

Speed-Density relation in DynaMIT takes the following mathematical form:

$$\begin{aligned} v &= \max(v_{min}, v_{calc}) \\ v_{calc} &= \begin{cases} v_{max}, & \text{if } k \leq k_{min} \\ v_{max}\left[1 - \left(\frac{k-k_{min}}{k_{jam}}\right)^{\beta}\right]^{\alpha}, & \text{otherwise} \end{cases} \end{aligned} \tag{2.32}$$

where v is the speed of vehicle (in mph), v_{max} is the maximum speed on that segment which is free-flow speed, k is current segment density (in vehicle/mile/lane), k_{min} is the minimum density beyond which traffic is no longer free flow, and k_{jam} is the jam density. α and β are segment-specific model coefficients.

Key parameters of calibration here are six-element vectors for each segment (v_{max}, k_{min}, k_{jam}, α, β, v_{min}). However, in reality, sufficient information for each segment is hardly available for these parameters to be calibrated individually (primary because not all segments are quipped with usable detectors). Even for the segments which have working detectors, data for sufficient periods of time for all ranges of densities are not available to fine-tune the parameters in (2.32). Hence calibration is usually performed by pooling together data for segments with "similar" expected characteristics.

2.2.2 Capacities

Another important input to supply simulation are segment capacities. As DynaMIT is a mesoscopic traffic simulation tool, movement of vehicles from one segment to the next is governed by available capacity and physical space considerations, and any constraint on either would cause vehicles to queue. Hence an important calibration step involves the computation of segment and lane-group capacities which truly capture allowable turning movements, section geometry, traffic control strategies, and other pertinent factors.

Initial estimates of segment capacities are calculated based on recommendations of Highway Capacity Manual [24].

2.3 Demand Calibration

This section outlines the demand calibration framework. We identify the contribution of various models to the O-D estimation process and list the set of parameters which are to be calibrated from demand side. The main components of demand calibration process are:

- Path choice set generation,
- Historical database of time varying O-D matrices and travel times,
- O-D prediction model, and
- Route choice model

2.3.1 Path Choice Set Generation

Generation of a path choice set for each O-D pair is critical to all further calibration process. As contributions of O-D flows to sensors are contingent upon the route set and route choice probabilities, and this information is critical in estimation and identifying observability in O-D estimation process, path choice set should include all feasible and reasonable paths between each O-D pair. Also, route-choice parameters are important to correctly identify the order of preference in selection and relative percentages of traffic on various paths originating from an origin. Section A.2.1 discusses various parameters in path choice set generation process as implemented in DynaMIT.

2.3.2 Historical Database

The sequential calibration procedure for O-D flows using a GLS approach has been detailed in Section 2.1.3. The following subsection identifies an important issue in calculating assignment matrices which are inputs to the GLS O-D estimation formulation.

Analytical Assignment Matrix

The O-D estimation problem outlined previously requires a set of time-dependent assignment matrices as input, along with observed counts and historical O-D flows. We can obtain these assignment matrices in two ways: by simulation, or by analytical calculation. A simulated assignment matrix (AM) is obtained by simulating the movement of vehicles, including route choice decisions, on the network and then counting the number of vehicles belonging to an O-D pair i departing in interval p which cross sensor j in interval h to provide the $(i,j)^{th}$ fraction of assignment matrix $\mathbf{a}_h^p$. An analytical assignment matrix, on the other hand, is computed by mathematically calculating the travel times of vehicles from origin to the sensor, and using route choice probabilities to estimate the fractions. Table 2.1 compares the advantages and disadvantages of both methods of computation.

As analytical AM is no-doubt superior to the simulated one, we have calculated assignment matrix for O-D estimation module based on calibrated route-choice parameters and simulated travel times. The approach used for this computation follows the proposed methodology by Ashok [3], and is presented below.

Let each O-D pair r be connected by a set of paths K_r, and F_h^k be the flow along path[5] k departing in interval h such that

$$x_{rh} = \sum_{k \in K_r} F_h^k \tag{2.33}$$

If q_{kh} denotes the fraction[6] of vehicles choosing path k in interval h, we have

$$F_h^k = x_{rh} q_{kh} \tag{2.34}$$

Noting that link flows are contribution from many paths from many O-D pairs using

[5]Since each path has a unique correspondence with its O-D pair, we drop the subscript for O-D pair for simplicity of notation.

[6]Obtained through appropriate route-choice model.

	Simulated AM	Calculated AM
1	AM is simulated with given demand, and hence function of input O-D flows.	AM is calculated and hence not directly a function of input O-D flows (but via experienced travel times).
2	Since O-D flows are unknown beforehand, AM is less reliable. Even with iterative solution technique, it is still dependent on estimated O-D flows.	Though travel times are unknown too, experienced travel times are more stable and less fluctuating from one interval to next, thus number of iterations for convergence are less.
3	Since route-choice probabilities are simulated, AM fractions are just one realization of the random experiment.	Since route-choice probabilities are calculated, AM fractions are the mean of random experiment.
4	Simulated probabilities do not capture paths with very low probabilities.	Calculated probabilities capture all paths with proportionate flow.
5	Zero starting O-D flows do not contribute to assignment fractions, thus estimation algorithm keeps zero O-D flow pairs zero for full estimation duration.	Since AM is independent of O-D flow volumes, all pairs are captured, providing GLS estimation an opportunity to make zero flow pairs to non-zero in next iteration.
6	In cases where a-priori O-D flows are unknown, dependence of AM on O-D flows can lead to situation where some flows are estimated zero, and which can never be positive again.	Analytical AM estimates all flows in each iteration with proper representation, and is independent of starting O-D matrix.
7	AM is likely to be very sparse.	AM is likely to be relatively denser than simulated AM thereby requiring more computational efforts in estimation. Trial estimates showed that while simulated AM had 0.06% of non-zero entries, the number for corresponding analytical AM was 7.03%.
8	Requires additional run-time for the simulation, and computational efforts are proportional to the traffic on the network.	Requires additional run-time for the calculation which is independent of traffic on the network, thereby proving to be more useful in congested periods.
9	AM is based on simulated departure times for vehicles from origin.	AM is based on assumption of uniform departure of vehicles from origin within estimation interval.

Table 2.1: Comparison of Simulated and Calculated Assignment Matrix

that link, we get

$$y_{lh} = \sum_{p=h-p'}^{h} \sum_{k=1}^{K} \alpha_{lh}^{kp} F_p^k + v_{lh} \tag{2.35}$$

where y_{lh} is the observed flows on link l in the interval h, K is the number of paths for all O-D pairs, and α_{lh}^{kp} is defined as contribution of k^{th} path flow departing the origin during interval p towards the flow across detector l during the interval h. We can now obtain assignment fraction for contribution of O-D flow r in interval p towards counts on link l in interval h as

$$a_{lh}^{rp} = \sum_{k \in K_r} \alpha_{lh}^{kp} q_{kp} \tag{2.36}$$

Computation of α_{lh}^{kp} requires an assumption to be made about departure instants of the vehicles from their origins. Assuming uniform departure intervals for all vehicles during the estimation interval as suggested by Cascetta *at al.* [11] and Ashok [3], we obtain:

$$\alpha_{lh}^{kp} = \begin{cases} 1 & \text{if } (h-1)H < \eta_{1l}^{kp} < \eta_{2l}^{kp} < hH \\ (hH - \eta_{1l}^{kp})/(\eta_{2l}^{kp} - \eta_{1l}^{kp}) & \text{if } (h-1)H < \eta_{1l}^{kp} < hH < \eta_{2l}^{kp} \\ H/(\eta_{2l}^{kp} - \eta_{1l}^{kp}) & \text{if } \eta_{1l}^{kp} < (h-1)H < hH < \eta_{2l}^{kp} \\ (\eta_{2l}^{kp} - (h-1)H)/(\eta_{2l}^{kp} - \eta_{1l}^{kp}) & \text{if } \eta_{1l}^{kp} < (h-1)H < \eta_{2l}^{kp} < hH \\ 0 & \text{otherwise} \end{cases} \tag{2.37}$$

where η_{1l}^{kp} and η_{2l}^{kp} represent times when first and last vehicle from the interval p using path k, cross the sensor at link l. Since we also need link travel times for the intervals later than the current estimation interval to calculate η_{2l}^{kp}, best estimates of available travel times from previous interval or previous days of calibration are used in the calculation of the assignment matrix.

2.3.3 Estimating Auto-regressive Factors

Prediction capabilities of DynaMIT require estimation of future O-D flows from the estimated and historical flows. Using concept of deviations proposed by Ashok and Ben-Akiva [4], we can write deviation of future O-D flow from its historical value as linear combination of deviations of estimated O-D flows from their historical counterpart as follows,

$$\mathbf{x}_h - \mathbf{x}_h^H = \sum_{p=h-q'}^{h-1} \mathbf{f}_h^p (\mathbf{x}_p - \mathbf{x}_p^H) \tag{2.38}$$

where $\mathbf{f}_h^p$ is a matrix of auto-regressive coefficients in the above regression equation and q' is the degree of auto-regression, a parameter to be estimated.

Factors of auto-regressive matrix $\mathbf{f}_h^p$ can be obtained by regression of the form:

$$\hat{x}_{r,h} - x_{r,h}^H = \sum_{p=h-q'}^{h-1} \{f_{r,h}^{1,p}(\hat{x}_{1,p} - x_{1,p}^H) + \ldots + f_{r,h}^{n_{OD},p}(\hat{x}_{n_{OD},p} - x_{n_{OD},p}^H)\} + w_{r,h} \tag{2.39}$$

Note that each day of observations yields only one data point for regression and n_{OD} such regression for each interval h are required to obtain entire set of auto-regressive matrices. In absence of large data set encompassing multiple days of observations, some simplifying assumptions can be made about structure of auto-regressive matrices.

For example, one could assume that the auto-regressive structure is independent of the particular values of h and p but only depends on the difference $h - p$. It could also be reasonable sometimes to assume that deviations in flow between r^{th} O-D pair are mostly affected by previous deviations in r^{th} O-D pair only. With these assumptions, equation (2.39) can be reduced to the following simpler form:

$$\hat{x}_{r,h} - x_{r,h}^H = \sum_{p=h-q'}^{h-1} f_{r,h}^{r,p}(\hat{x}_{r,p} - x_{r,p}^H) + w_{r,h} \tag{2.40}$$

where $\hat{x}_{r,h}$ is simply the r^{th} estimated O-D pair for the interval h.

2.3.4 Estimating Error Covariance Matrices

The GLS formulation presented in Section 2.1.3 assumes prior information on error-covariance matrices $\mathbf{V}_h$ and $\mathbf{W}_h$. However, in practice such data in not often available a-priori and has to be estimated during the calibration process. While we can initialize the estimation with variance-covariance matrices based on their interpretation as inverse of weights[7] assigned to various O-D flows and sensor observation in GLS formulation, error-covariance matrices can be estimated as follows once the calibration for all days is over.

Let $e_h^{k,y}$ be the vector of differences between observed sensor counts and those simulated by the DTA system using estimated O-D matrices for the intervals upto h on the k^{th} day, i.e.

$$e_h^{k,y} = \mathbf{y}_h - \sum_{p=h-p'}^{h} \mathbf{a}_h^p \hat{\mathbf{x}}_p \tag{2.41}$$

and let $e_h^{k,x}$ be the vector of differences between the a-priori and the estimated O-D flows for interval h on k^{th} day, i.e.

$$e_h^{k,x} = \mathbf{x}_h^a - \hat{\mathbf{x}}_h \tag{2.42}$$

then if $e_{i,h}^x$ and $e_{i,h}^y$ denote the n-dimensional row vectors of residuals for n days of calibration associated with i^{th} direct and indirect measurements respectively, $(i,j)^{th}$ element of matrices $\mathbf{W}_h$ and $\mathbf{V}_h$ can be approximated by

$$W_{ijh} = e_{ih}^x e_{jh}^{x\prime}/n \tag{2.43}$$

$$V_{ijh} = e_{ih}^y e_{jh}^{y\prime}/n \tag{2.44}$$

where e' is the vector transpose of e. Note that each day of data would provide only one observation for estimation of $\mathbf{V}_h$ and $\mathbf{W}_h$. We can simplify the estimation by assuming independence of error-covariance matrices on interval h thereby yielding one data point per interval instead of one point per day. We may further approximate,

[7] Refer [6] for discussion on this subject.

depending on the the specifics of the case in hand, that error-covariance matrices are diagonal in structure. This will not only simplify the computation but will also considerably reduce the run-time of estimation process as the GLS solution involves inverting the error-covariance matrix (equation (2.19)).

2.3.5 Route Choice Model Calibration

Choice probabilities of various paths in the choice set are evaluated using a discrete choice model, which can model route choice prior to embarking on the trip, and after starting the trip, separately taking into account different sets of alternatives[8] available at different stages of the trip and the difference in perception of relative utility of paths. Though it would be preferable to calibrate parameters in utility specification model from disaggregate data involving individual decisions of drivers in different categories in varying situations, aggregate nature of available data for calibration such as sensor counts makes calibration of all the parameters difficult. Usual approach in such cases is to estimate a smaller set of most suitable parameters while keeping other fixed to default values.

2.4 Joint Calibration and Model Validation

The calibration parameters and approach presented in preceding sections depicts the calibration problem as separate demand and supply side calibration. However, in practice, as both state estimation (Section A.1.1) and prediction and guidance generation (Section A.1.2) modules involve demand-supply interactions, calibration is a joint and iterative exercise concerning all parameters such that the model best replicates reality. Both supply and demand parameters are regularly updated as the calibration progresses based on queue lengths, capacity constraints, sensor reliability, etc.

[8]For example, while changing route, mode, departure time and/or destination are possible before the start of the journey; only change in route, and possibly, destination, is a feasible alternative en-route.

Once the calibration is performed to a satisfactory extent, the next important step is the validation of calibrated parameters. This is critical as it is possible that the model replicates the data with which it is calibrated, yet fails to continue to predict given fresh inputs. For a DTA system which is designed to reside in a Traffic Management Center and assist transportation analysts with prediction-based guidance, it is expected that model will continue to perform satisfactorily with new real-time information from the surveillance system. Validation of model parameters is performed by measuring the performance of the calibrated model with a completely new set of data. Discrepancies between model prediction and system behavior during validation may necessitate re-calibration of partial or complete set of parameters.

2.5 Summary and Conclusion

In this chapter we introduced the concept of observability as applicable to general dynamic control systems. We identified characteristics of such systems, and presented a formulation to verify the observability condition in dynamic systems. We then defined observability as a system property by which, given sufficient regular measurements of the quantity of interest, it is possible to uniquely determine the state of the system irrespective of any assumptions about starting state of the system. Observability as applicable to O-D estimation module in Dynamic Traffic Assignment systems is discussed thereafter. Subsequently, an empirical setup is developed to test observability for large-scale networks.

This chapter also highlights the demand and calibration framework adopted for the calibration of DTA systems. We describe various inputs and parameters to be calibrated for both demand and supply models, and discuss the superiority of an analytically computed assignment matrix over a simulated one. A comprehensive discussion on calibration of DynaMIT[9] and the results are presented in Chapter 3.

[9]DynaMIT is introduced in Appendix A.

Chapter 3

Case Study

In Chapter 2, we discussed the theoretical background for observability and its application in the context of O-D estimation. We also outlined a general calibration framework for both supply and demand parameters for real-time application of DTA system. In this chapter we present empirical results of observability tests and calibration on a large-scale network in downtown Los Angeles using DynaMIT as our calibration system. We begin by briefly describing the study network and the data availability, and then proceed to verify observability followed by calibration and validation results for the parameters described in the Chapter 2.

3.1 The Los Angeles South Park Study Area

The network and the data used for calibration and validation are taken from the South Park area of the Los Angeles metropolis. This heavily traveled network forms part of the downtown area of Los Angeles, and falls in District 7, Los Angeles/Ventura in California, United States. The study area attracts high volumes of mixed traffic due to the presence of numerous shopping malls, restaurants, office buildings and entertainment and conference facilities.

3.1.1 Network Description

The study network (Figure 3-1) is comprised of two major freeways and a dense network of arterial streets. Interstate I-110 (Harbor Freeway) runs North-South while I-10 (Santa Monica Freeway) runs East-West through the center of the network. Both freeways cross in the middle of network and are widely connected to arterial streets through a set of ramps. Major arterials marking the border of study network include St. Hoover Street on west, St. Adams Blvd on south, Olympic Blvd on north and Grand Avenue on the east. Being part of downtown area, this region attracts high freeway and street traffic. Many events, meetings, conferences, and exhibits scheduled in Convention Center and Staples Center, which are located in the middle of study region, throughout the year also contribute to temporal traffic surge within the network. This network is represented in DynaMIT by a set of 243 nodes interconnected

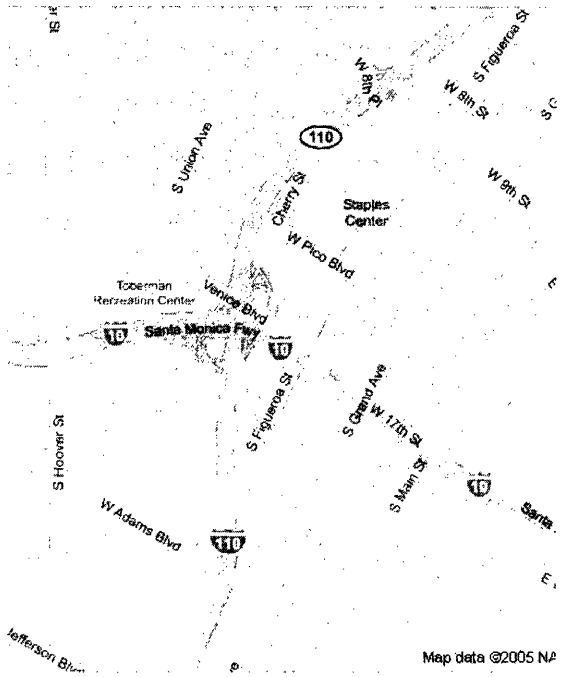


Figure 3-1: South Park Study Area

by 606 directed links. These links represent the physical links (directional streets) on the network, and are further subdivided into 740 segments to model changing section geometry within a link. Nodes are classified into intersection nodes and loader nodes.

Among more than 200 arterial intersections in the network, about half are known to be signalized. Signals along the major arterial segments are synchronized.

3.1.2 Data Description and Analysis

The data required by DynaMIT can be classified into following categories:

- Network information
- Surveillance information
- Incident log
- Signal timing and phasing charts
- Historical (seed) O-D

This data was collected from variety of sources and agencies. Following sections provide brief overview of the available data for South Park network.

Network Information

The network data comprising of network geometry, nodes, links and lane groups was provided by Los Angeles Department of Transportation (LADoT) in form of maps of arterial intersections within the region. Latitude and longitude coordinates of the nodes were supplemented by closeup screenshots of each intersection and the study area. Freeway alignment information was indirectly available in form of latitude-longitude coordinates and post-miles of sensors located on freeways.

Surveillance Information

Surveillance data on the network was collected from embedded loop-detectors. Volumes and occupancies for lane specific arterial sensors were collected and provided by LADoT for the month of September 2004 with 30 second resolution for 721 sensors. Corresponding information for freeway sensors was obtained from archived data by Freeway Performance Measurement System (PeMS, [12]) project undertaken at the

University of California at Berkeley in cooperation with California Department of Transportation (CalTrans). Archived link specific volumes, speeds and densities were obtained for freeway and ramps sensors with 5 minute resolution.

Incident Log

Freeway incident log was also obtained from PeMS [12] website which reads details of incidents continuously from the website of California Highway Patrol [9] and archives them. Start and end dates and times and location in words along with the reason for incident was available for each incident. Arterial incident log was obtained from LADoT for the duration of study period which included end time, ID of nearest detector and intersection, and a binary variable indicating type of the incident among two reported types.

Signal Plans

Signal timings and phasing charts for the signalized arterial intersections were made available by LADoT.

Historical/Seed O-D

As outlined in the Section 2.1.3, historical O-D flows along with sensor counts are used for estimation of time dependent O-D matrices. However, no historical or seed O-D information (static or dynamic) was available for this network. The estimation of historical O-D flows is thus a part of calibration.

Sensor Data Analysis

In this section we describe the preliminary analysis of data and prepare the groundwork for input preparation (Section 3.1.3) to DynaMIT. Arterial sensor data for this network were available at 30 second interval while freeway and ramp data were available at 5 minute interval. Hence, all arterial sensor data were aggregated for 5 minute time intervals. Furthermore, as DynaMIT requires link specific sensor counts, all lane

specific detectors on a link were aggregated to link based sensor counts. Following screening process was then adopted, in order, to identify usable arterial sensors for calibration. Those sensors not included in calibration where:

1. not all lanes in a link were counted

2. data was not available for at least one comprising lane based sensor

3. one or more comprising sensors were deemed to behave erratically

4. sensor was found to be inconsistent with the neighboring sensors

Thus finally 189 out of available 330 link based arterial sensors were found usable. Similarly processing the data reported from freeway and ramp sensors, 14 useful sensors were identified for calibration—12 on the freeways and 2 on the ramps.

Occupancies reported by arterial sensors were converted to densities assuming average car length of 5 meter as:

$$K = \left(\frac{52.8}{L_{veh} + L_{sens}} \right) Occup \tag{3.1}$$

where K is density in vehicle/lane/mile, L_{veh} is average length of vehicles in feet (assumed 16.4 ft), L_{sens} is diameter of loop-detector in feet (6 ft), and $Occup$ is occupancy as percentage of time the detector is occupied by vehicles.

Flow counted by sensors were compared for day to day variations among the weekdays and no systematic or significant variation was observed as shown in sample sensor plots in Figures 3-2 to 3-7. Sensor plots for weekends, however, showed difference in traffic pattern between Saturdays and Sundays, particularly in the morning and early afternoon (Figures 3-8, 3-9, and 3-10). September 6^{th} 2004, Monday, was observed to be significantly different than the rest of the weekdays because of Labor Day holiday.

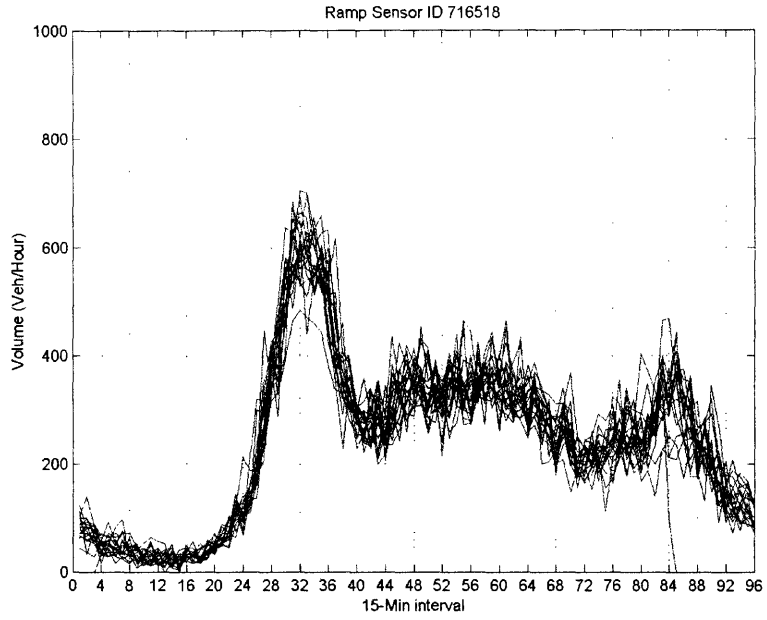


Figure 3-2: Variation Across Weekdays - Freeway Sensor ID 716518

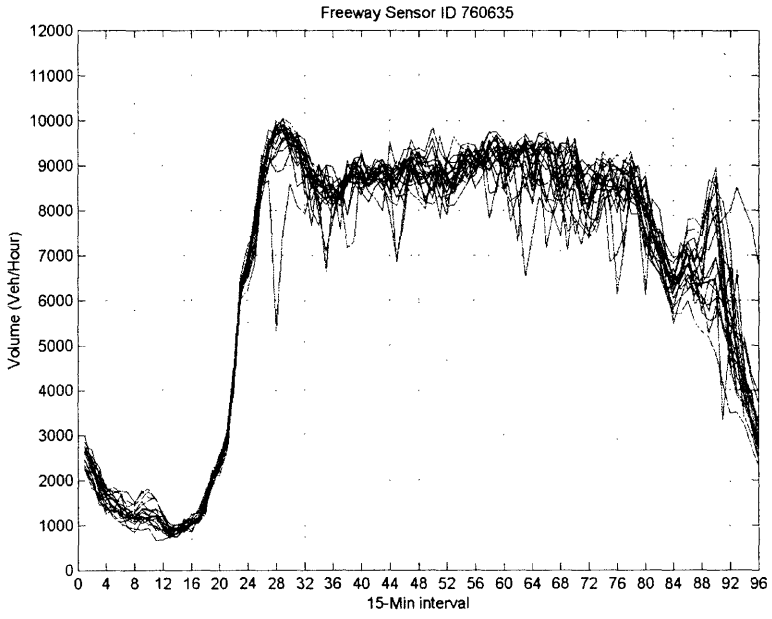


Figure 3-3: Variation Across Weekdays - Freeway Sensor ID 760635

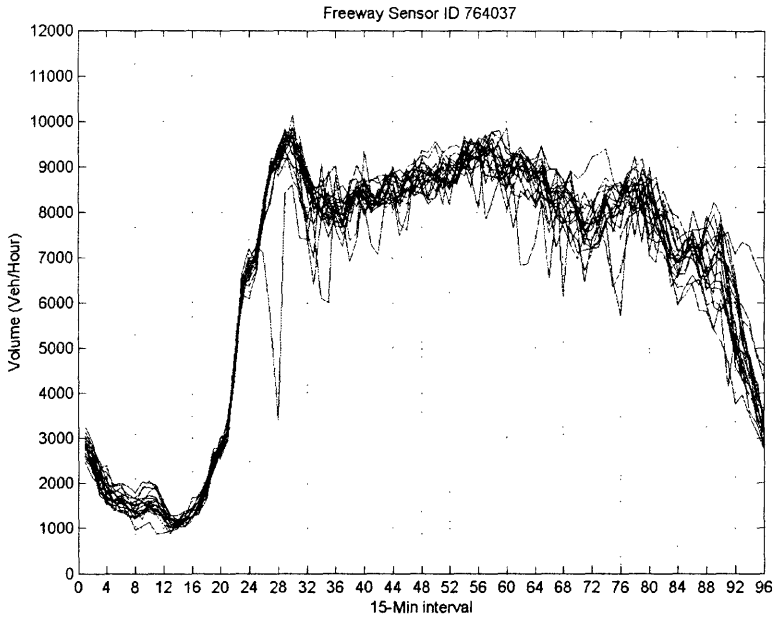


Figure 3-4: Variation Across Weekdays - Freeway Sensor ID 764037

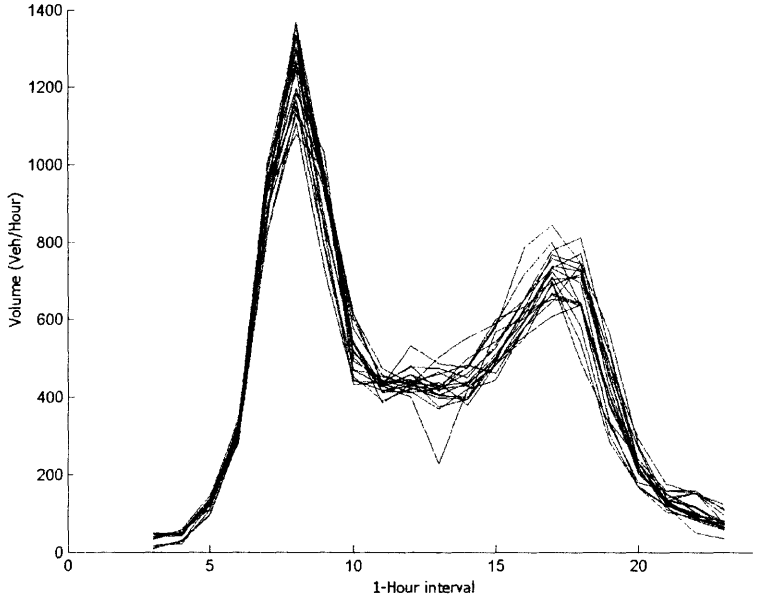


Figure 3-5: Variation Across Weekdays - Arterial Sensor ID 9056

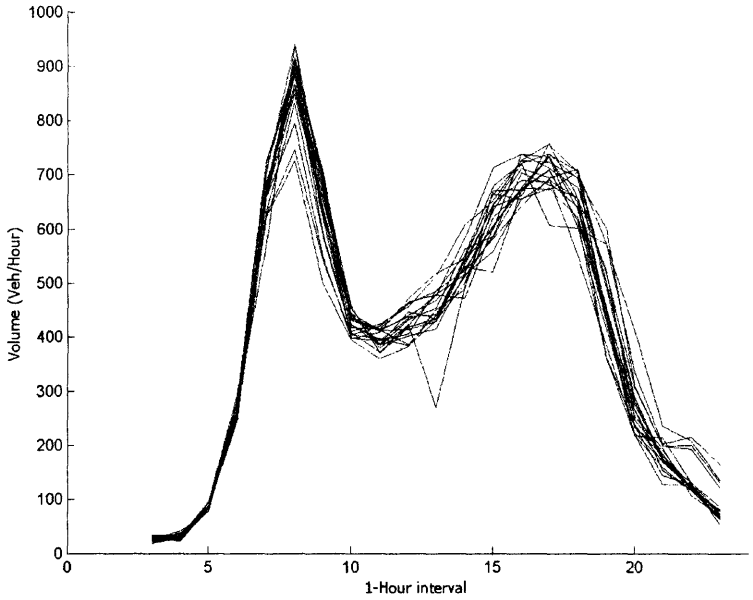


Figure 3-6: Variation Across Weekdays - Arterial Sensor ID 5075

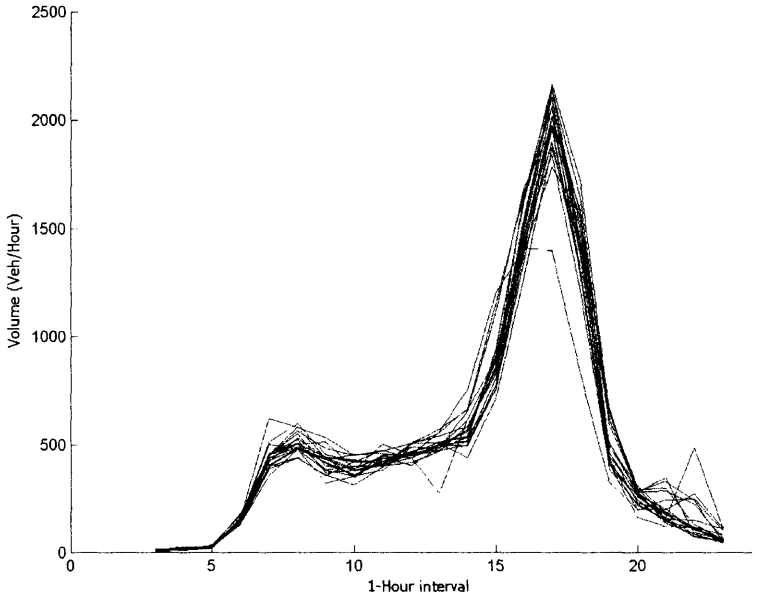


Figure 3-7: Variation Across Weekdays - Arterial Sensor ID 3105

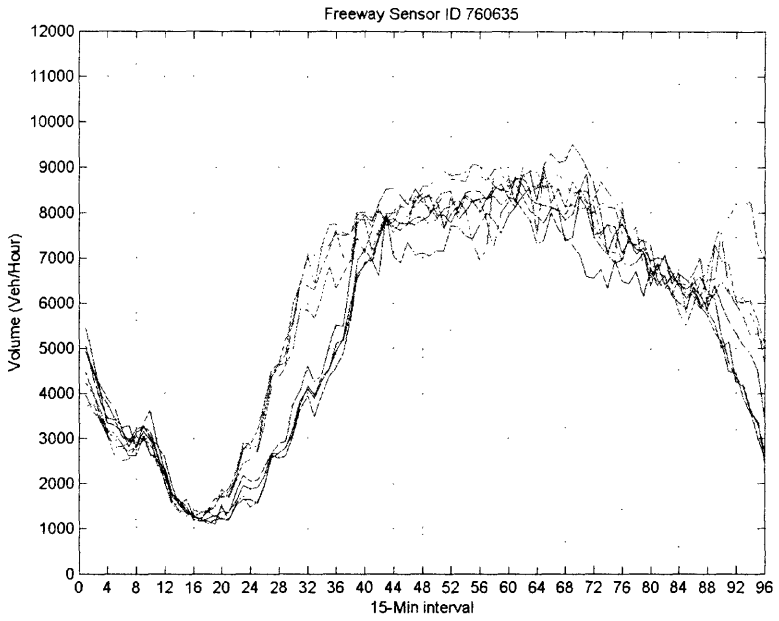


Figure 3-8: Variation Across Weekends - Freeway Sensor ID 760635

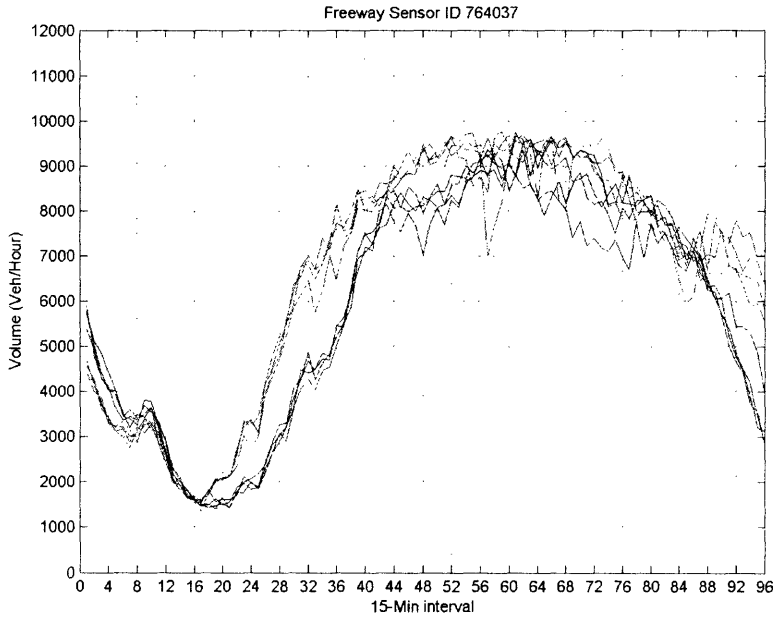


Figure 3-9: Variation Across Weekends - Freeway Sensor ID 764037

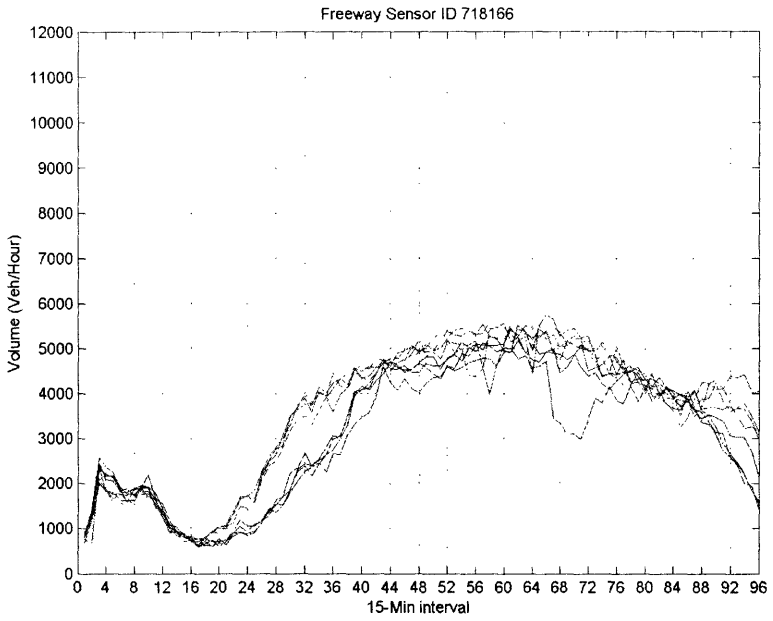


Figure 3-10: Variation Across Weekends - Freeway Sensor ID 718166

3.1.3 Input Preparation

DynaMIT has the ability to handle a variety of information pertaining to traffic and demographic characteristics which are used in demand and supply simulations, and prediction based guidance generation. We will briefly describe the input files and format required by the DynaMIT and preparation of analyzed data to suit this for calibration.

Network Data

Network information is conveyed to DynaMIT through a **Network file** that describes locations of nodes, links, segments, lanes and sensors. Link sections are subdivided into segments to model changing section geometry such as curvature, number of lanes or lane groups. Link connectivity information specifies permissible and prohibited vehicle movements, while sensor locations on segments designate measurement points.

Demand Information

DynaMIT requires time-dependent starting O-D flows in **Demand file** which are updated using the latest count data during the O-D estimation process. Time interval for input O-D flows should match the time interval over which sensor counts have been aggregated. We have used constant 15 minutes time slices for the purpose of this analysis. DynaMIT also uses time dependent historical travel times on various links of network which are result of long term collective experience of the users and are used in path choice modeling as **Link Time file**. In absence of any such information in this case study, free flow travel times for each link were assumed for the initial intervals, where it is a reasonable assumption owing to very low flow at 3:00 AM. Input travel times for later intervals were taken from estimated link times from previous interval, and were updated based on simulated travel times. A set of **error variance-covariance matrix**, used in the GLS based estimation process, are also an outcome of the calibration process.

Supply Information

The **Supply Parameter file** defines the parameters that characterizes vehicle dynamics over segments of the network. Calibrated values of coefficients used in speed-density relationship (Section 2.2.1) and capacities were stored for each segment in the network which are used by DynaMIT's supply simulator. Speed-density parameters to be calibrated for DynaMIT are maximum (free-flow) speed, minimum speed, minimum density over which free-flow speed breaks down and maximum density over which speed drops to minimum speed, along with exponents α and β of the model.

Car User Groups Information

One of the key features of DynaMIT's modeling capabilities is its ability to model different categories of drivers with different traffic characteristics. DynaMIT can classify drivers based on the purpose of trip, information source and the value of time, and can model their perception about route choice utilities at different stages of their

trips. A trip can be defined to have purpose of work (WO), leisure (LE) and other (OT). Similarly, information availability can be subdivided into In-Vehicle (IV) information, no In-Vehicle information (NI), HOV-Guided (HV) and HOV-Guided with In-Vehicle information (IH). All possible combination of trip purpose and information source combined with low (L), medium (M) and high (H) value of time provide total 36 categories of trips between each origin-destination pair. This data is tabulated into **SocioEco file** as input to DynaMIT. Different values of coefficients in utility specification of path choice model are supported for different categories of drivers. Also, different value of travel time prior to and after starting the trip can be modeled with different set of parameters in **Behavioral Parameters file**. While socioeconomic characteristics of the drivers are known in advance, values of parameters in behavioral parameters file are subject of the calibration.

Socio-economic data, however, is rarely available for real-life applications because of high costs associated with its collection and dissemination. This problem is handled by assuming single value of time for all drivers. We also assume that majority of traffic is commuter traffic with average value of time (i.e. purpose = WO and VOT = M). This assumption is reasonable, particularly in the congested peak periods when the usefulness of real-time estimation and prediction is of most importance, since morning and evening peaks are usually characterized by, and caused by, commuter traffic.

Traffic Surveillance Information

DynaMIT needs interval by interval volume counts at all sensor locations for estimation of O-D flows and calibration of other parameters. This sensor information is provided for every 15 minute interval from the archived and processed sensor data in form of **Sensor file**. Locations, start times, durations and severities (in form of reduced capacities) of the incidents during simulation horizon are supplied as **Incident file**.

Prediction Model Parameters

Predictive capabilities of DynaMIT lies in forecasting future O-D flows and consequent network state. A set of **auto-regressive factors** are required for this application, which form input to DynaMIT's demand simulator. However, these files are generated as part of the calibration exercise.

3.2 Supply Calibration

Parameters to be calibrated for supply simulation fall in two main categories:

- Parameters of the speed-density relationships expressed by equation (2.32), and
- Lane group capacities on freeway and arterial segments, and at intersections

In this section, we will present the results of calibrating supply parameters as discussed in Section 2.2.

3.2.1 Speed-Density Parameters

Calibration of six-element vector (Section 2.2.1) of speed-density parameters follows approach suggested by Kunde [20]. Los Angeles South Park network has total of 740 segments, and in contrast we have only 203 usable sensors. Furthermore, all arterial sensors report occupancy as number of seconds occupied in 30 second interval. Density values were, therefore, deduced from these occupancies using equation (3.1). Hence, given the limit on amount of sensor data, individual calibration of each segment was not possible. Thus segments were grouped based on similar expected characteristics[1], and data from all sensors corresponding to that category were pooled to calibrate speed-density parameters for all segments of that category.

Based on availability of data and expected similarity among segments, following segment groups were identified for separate set of parameter calibration:

[1]Type of segment among arterial streets, ramps, and freeways, and the number of lanes were used to group the sensor data.

1. Group A1 ⇒ Arterial street segments with only one lane
2. Group A2 ⇒ Arterial street segments with only two lanes
3. Group An ⇒ Arterial street segments with three or more lanes
4. Groups F1–F11 ⇒ Freeway segments, individual calibration for each freeway sensor
5. Group R ⇒ Ramp segments

Within each group, speed and density data from all relevant sensors were collected and pooled to plot speed-density curve. Curve of the form given by equation (2.32) was then fit to minimize the errors in the least square sense.

Results of the calibration of speed density parameters are presented here for all 15 segment categories listed above. Table 3.1 shows final calibrated parameters, and Figures 3-11 to 3-25 shows corresponding fitted curves.

Segment Group	v_{max} (mph)	k_{jam} (veh/lane/mi)	α	β	K_{min} (veh/lane/mi)	V_{min} (mi/hr)
A1	30.00	400	3.55	0.63	14.00	8
A2	30.00	400	3.50	0.58	16.00	8
An	35.00	400	3.70	0.60	18.00	8
F1	74.00	400	1.90	0.61	25.12	15
F2	72.00	400	3.60	0.60	28.00	15
F3	71.50	400	6.80	0.85	16.00	15
F4	70.00	280	3.50	0.80	16.00	15
F5	74.60	400	19.23	0.89	24.84	15
F6	67.00	400	6.71	0.85	28.17	15
F7	72.00	380	4.11	0.70	27.00	15
F8	75.00	230	3.60	0.80	25.00	15
F9	73.00	400	12.45	0.98	27.60	15
F10	72.00	400	9.46	1.00	25.00	15
F11	72.00	400	10.11	0.86	29.50	15
R	55.00	400	3.60	0.80	28.00	10

Table 3.1: Calibrated Speed-Density Parameters

Note that since we had only two ramp sensors, data required for calibration of speed-density parameters was inadequate for ramps. However, since ramps usually

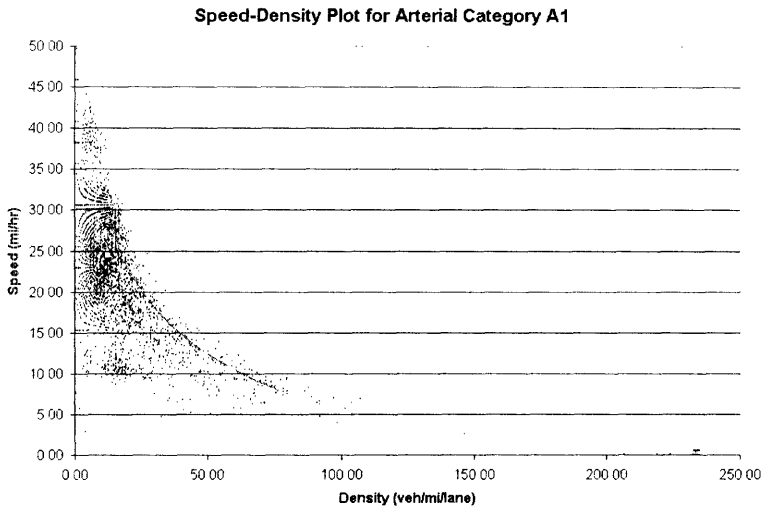


Figure 3-11: Speed-Density Curve for Segment Group A1

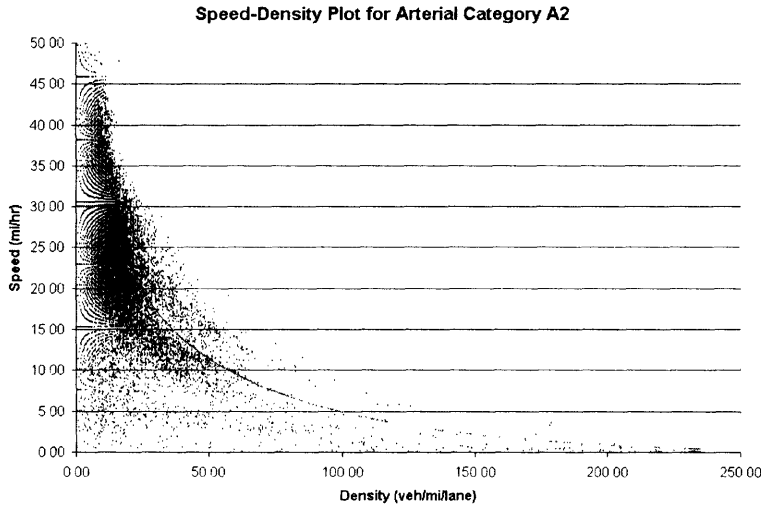


Figure 3-12: Speed-Density Curve for Segment Group A2

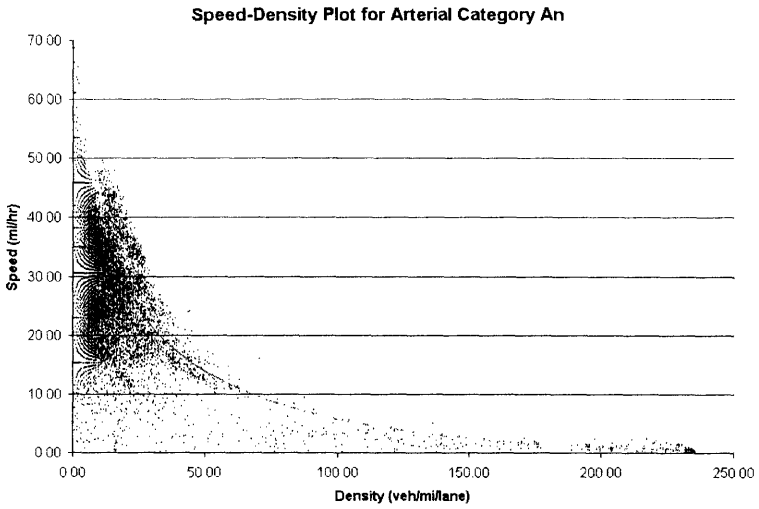


Figure 3-13: Speed-Density Curve for Segment Group An

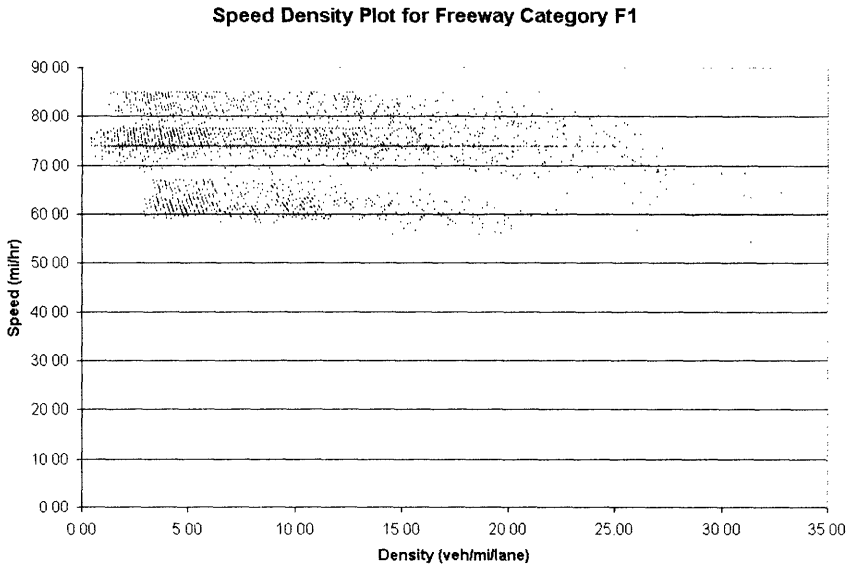


Figure 3-14: Speed-Density Curve for Segment Group F1

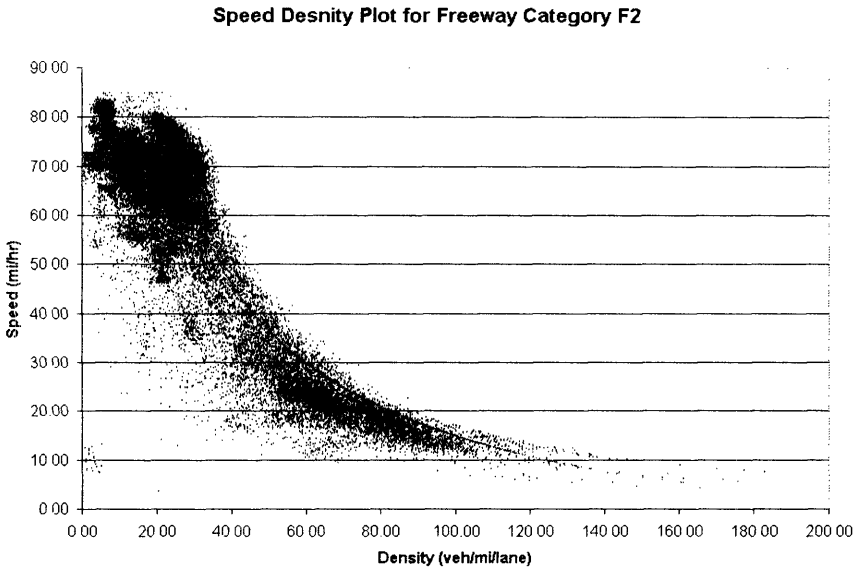


Figure 3-15: Speed-Density Curve for Segment Group F2

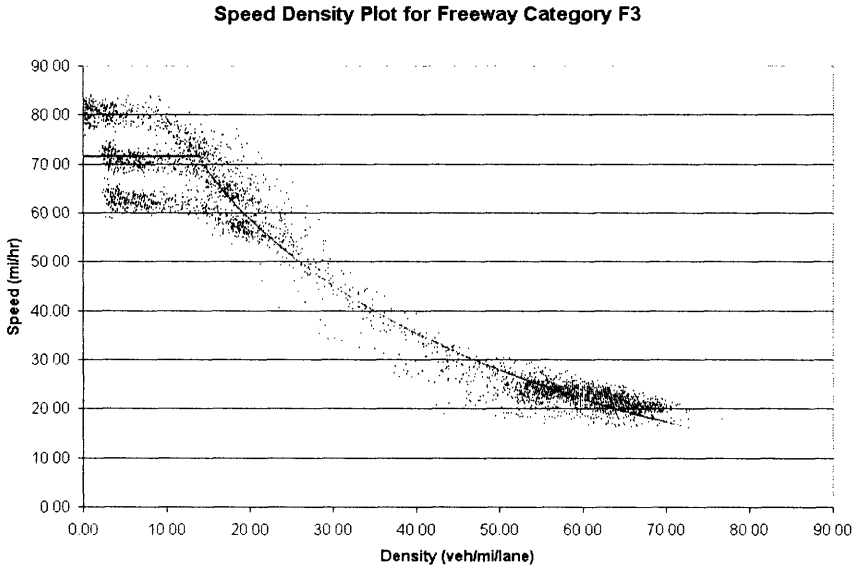


Figure 3-16: Speed-Density Curve for Segment Group F3

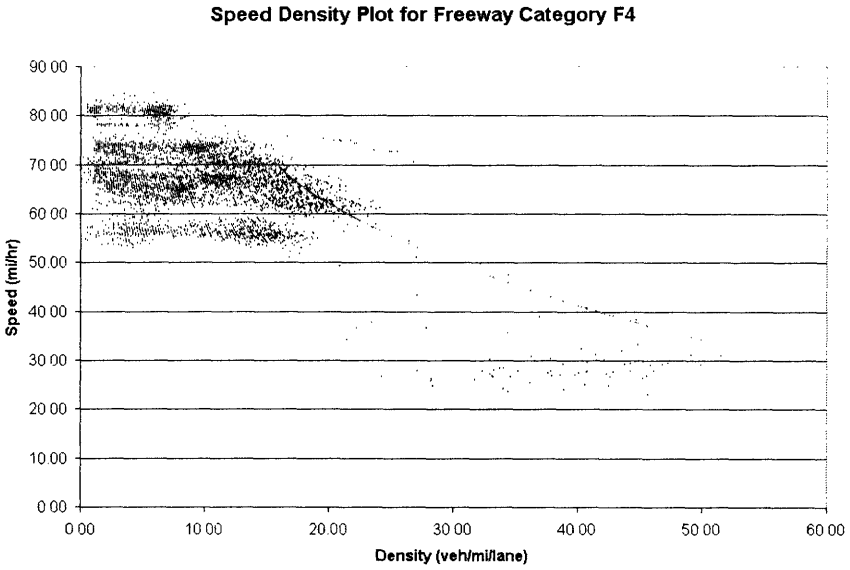


Figure 3-17: Speed-Density Curve for Segment Group F4

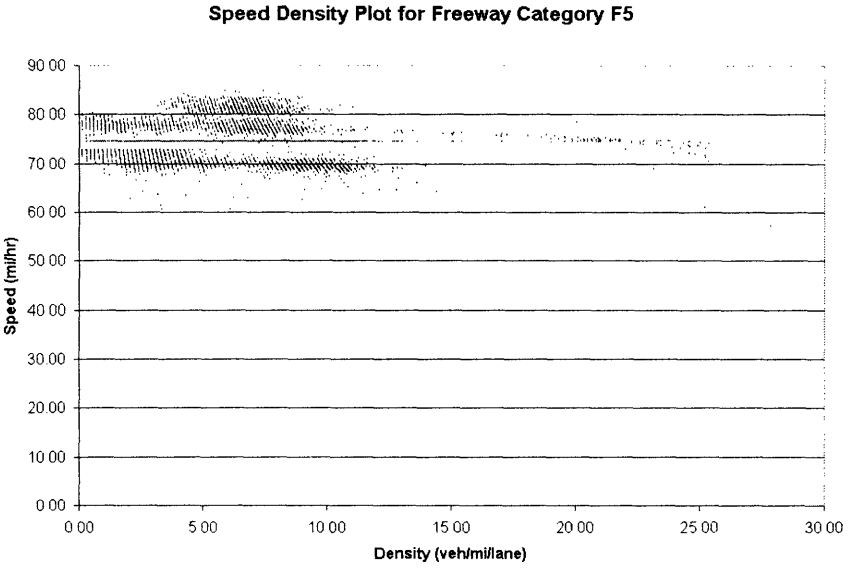


Figure 3-18: Speed-Density Curve for Segment Group F5

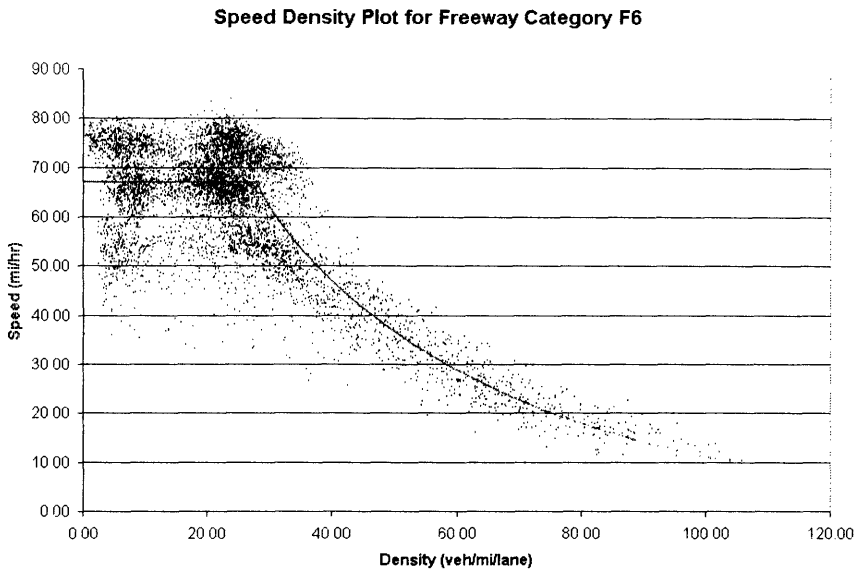


Figure 3-19: Speed-Density Curve for Segment Group F6

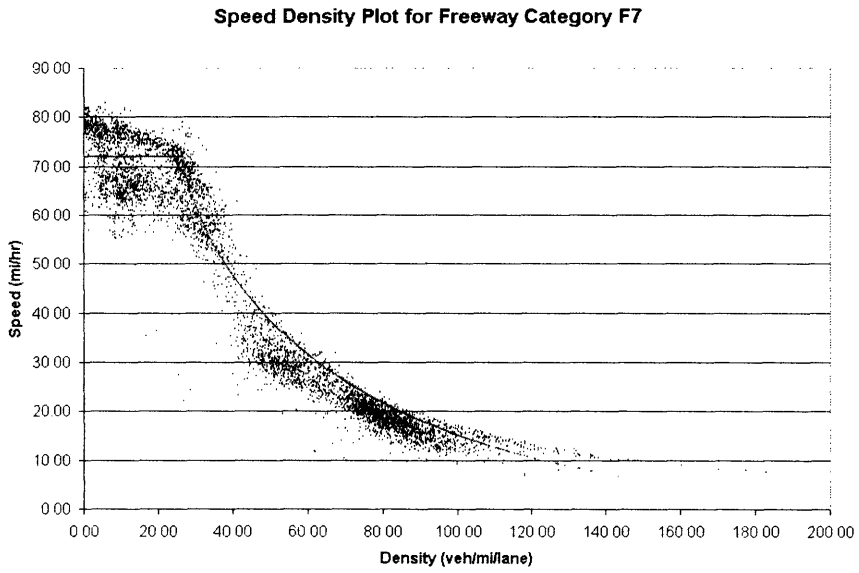


Figure 3-20: Speed-Density Curve for Segment Group F7

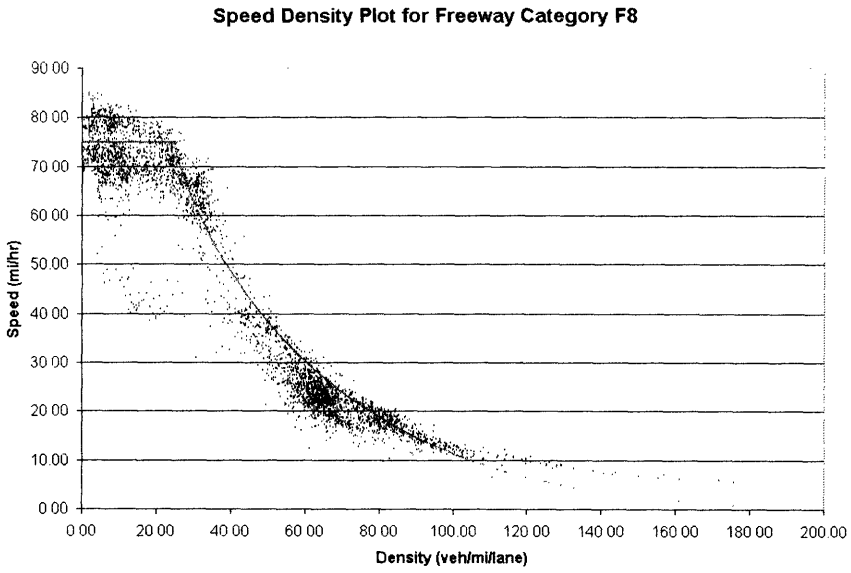


Figure 3-21: Speed-Density Curve for Segment Group F8

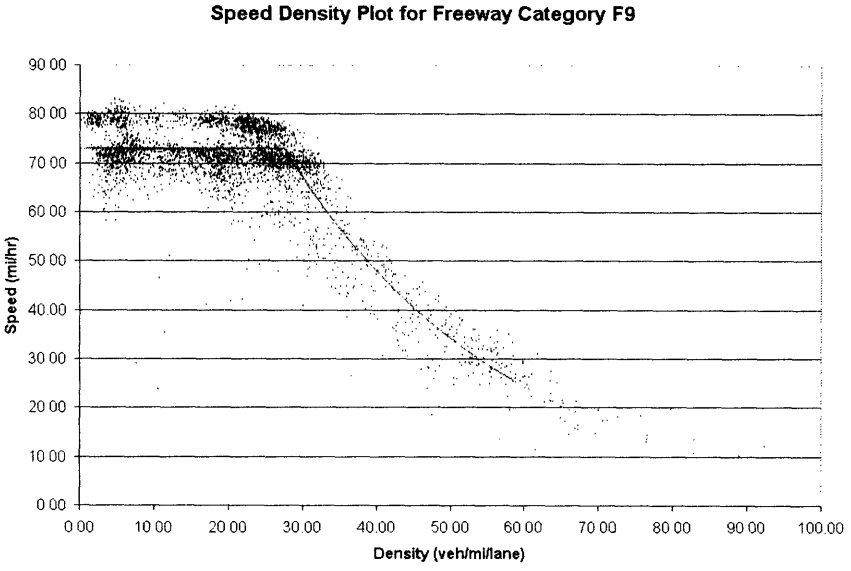


Figure 3-22: Speed-Density Curve for Segment Group F9

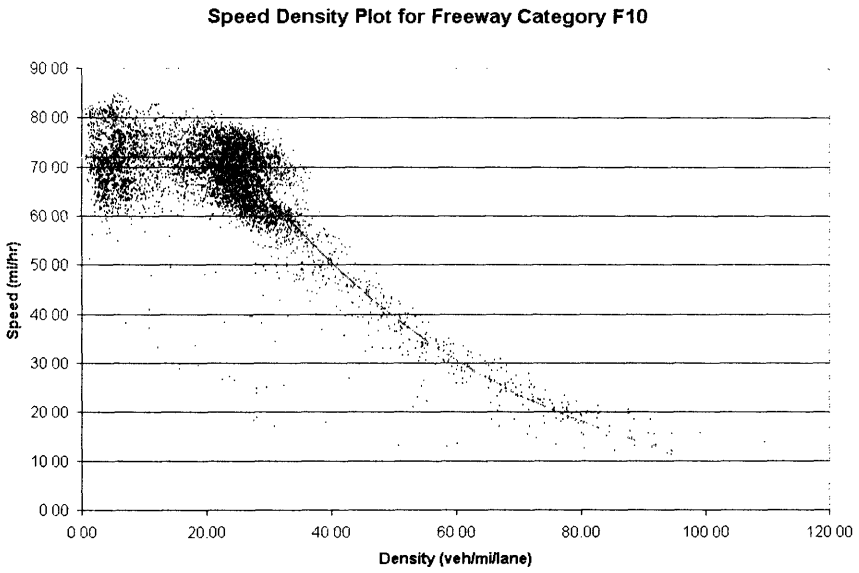


Figure 3-23: Speed-Density Curve for Segment Group F10

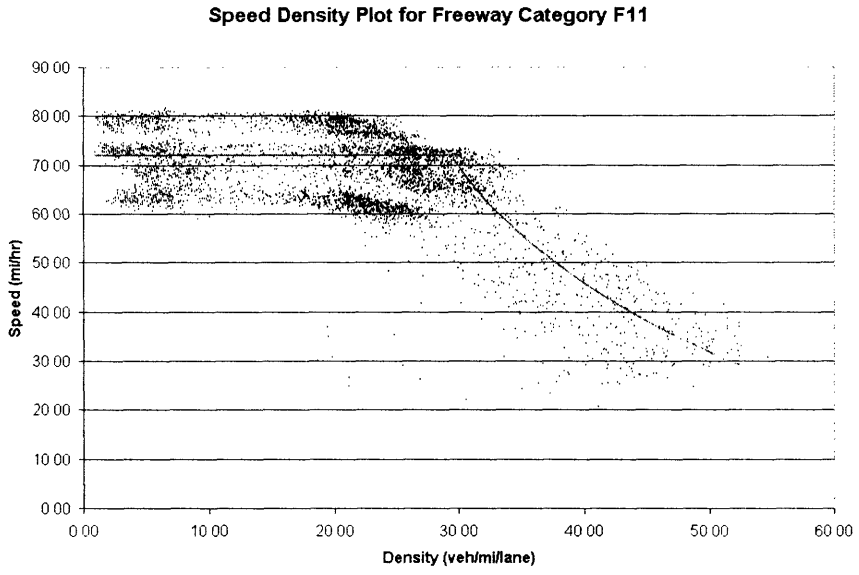


Figure 3-24: Speed-Density Curve for Segment Group F11

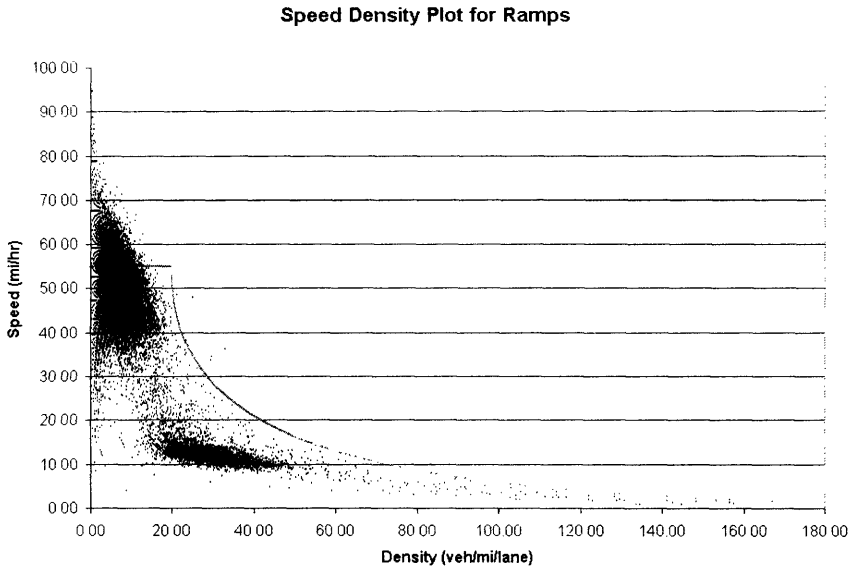


Figure 3-25: Speed-Density Curve for Segment Group R

have maximum speeds lower than those on freeways and higher than those on arterials, and are characterized by curvy and inclined geometry with merging and diverging conflicts, ramp specific calibration was necessitated. Therefore, parameters estimated above are partially based on speeds and densities observed on the ramps, and partially on parameters calibrated for freeways.

3.2.2 Capacities

Initial capacities of the segments were calculated based on the recommendations of Highway Capacity Manual (HCM, [24]). Best estimates of capacities from HCM, however, require information such as composition of heavy vehicles, percentage of experienced and unexperienced drivers, terrain slope, geometry and pedestrian movements. In absence of such detailed input, some simplifying assumptions were made and default coefficients values were used. Following capacities were initially assumed for various segment types:

- Freeway capacity = 2400[2] vehicle/hour/lane

[2]Figure 3-2, pp 3-4, HCM, 1998

- Ramp capacity = 2200[3] vehicle/hour/lane

- Lane group capacity at signalized intersection = $s_i \frac{g_i}{C}$; where s_i and g_i are saturation flow rate (vph) and effective green time (sec) for lane group i respectively, and C is cycle time (sec)

Saturation flow rate is a function of a host of parameters like lane width, approach grade, number of buses parked, right and left turns etc. Default saturation flow was assumed to be 1200[4] vehicle/hour. Effective green and cycle times were computed from signal phasing charts provided by LADoT. For the unsignalized intersections, or the intersections for which no signal plans were available, capacities were computed considering them unsignalized intersections with average arterial capacity per lane as recommended by Highway Capacity Manual [24].

3.3 Demand Calibration

Calibration exercise discussed here directly follows from discussion presented in Section 2.3 and is based on [6]. Following subsections prepare the ground work for the demand calibration such as generating suitable path choice set, defining period of study and state major simplifying assumptions. Thereafter, we proceed to present calibrated set of parameters including estimated O-D flows and auto-regressive factors.

3.3.1 Path Choice Set Generation

Foremost step in the demand calibration is the generation of a suitable set of paths for each O-D pair which we seek to estimate. The path generation process attempts to capture all reasonable and feasible paths for each O-D pair. As our study network is a dense grid network, there exist innumerable paths, particularly for arterial O-D pairs. Thus, calibration of path set generation parameters attains even more importance.

[3]Table 5-6, pp 5-14, HCM, 1998
[4]Equation (9-10), pp 9-14, HCM, 1998

Also, estimation of O-D matrices is critically dependent on the fractions and the number of O-D pairs getting counted by sensors on the field, and hence the path choice set should include all practical paths for all O-D pairs. For this study, a suitable path set was obtained using 20 random draws to complement link-elimination based shortest paths (Section A.2.1) from every link on the network to every destination node. Recognizing the fact that between two paths with the same travel time, the path with a higher freeway percentage is usually preferred, a freeway "bias" factor of 0.6 was used to force the path generation algorithm to include paths with longer freeway sections. To eliminate many roundabout and unreasonable paths, all paths longer than the shortest path by more than 20% were eliminated between each O-D pair. Manual inspection of various O-D pairs confirmed that nearly all practicle paths were included in the path choice set, which finally contained total of 44,224 O-D paths for 3745 O-D pairs.

3.3.2 Period of Study

Since DynaMIT will be deployed for real-time traffic estimation and prediction based guidance generation on the site, period of off-line calibration encompasses whole of 24 hour day, barring 3 hours from midnight to 3:00 AM as arterial data was not available for the said period. Thus this calibration will produce all required parameters for real-time estimation and prediction from 3:00 AM every day until next midnight. The 21 hour calibration period was divided into 84 interval of 15 minutes each. 15 minute interval was decided considering computational efficiency and the fact that sensor counts do not exhibit significantly large variations for shorter interval.

3.3.3 Simplifying Assumptions

Several simplifying assumptions were made in order to accommodate practical considerations (such as run time requirements for real-time updating) and data availability while estimating the model parameters. The error covariance matrices $\mathbf{V}_h$ and $\mathbf{W}_h$ were assumed to possess a diagonal structure implying that the sensor measurement

errors and direct O-D measurement errors are uncorrelated. Also, the structure of error-covariance matrices was found to be practically invariable for high flow period between morning 6:00 AM and evening 8:00 PM. Similarly, no significant difference was observed in error covariance matrices for periods between 3:00 AM–6:00 AM and 8:00 PM–12:00 AM. The auto-regressive factors $\mathbf{f}_h^p$ were also assumed to be diagonal suggesting that deviation of O-D pair r from its historical values depends only on the deviations of same O-D pair from its historical values in previous intervals. The O-D flows themselves were grouped in high, medium, low and very low[5] to estimate common set of auto-regressive factors within each group.

3.3.4 Error Statistics

Quantification of error in model performance and system output is important to identify the degree of satisfaction of calibration and evaluate the need for re-calibration. A good measure of error also helps in identifying the direction of improvement during calibration process. We have adopted following error statistics for measuring the discrepancies between observed (y_i) and simulated ($\hat{y}_i$) quantities.

- Scale = $\sqrt{\frac{\sum_i {y_i}^2}{\sum_i \hat{y}_i^2}}$
- Root Mean Square Error (RMSE) = $\sqrt{\frac{\sum_i (y_i - \hat{y}_i)^2}{N}}$
- Root Mean Square Normalized (RMSN) = $\sqrt{\frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i {y_i}^2}}$

We want to keep scale as close to one as possible while RMSE and RMSN should both be close to zero.

[5]Since O-D pairs with flow less than 20veh/hr were grouped as *very low*, auto-regressive factors are likely to be unstable. However errors introduced by those factors are also likely to be insignificantly small.

3.3.5 Calibrating Route Choice Parameters

Route choice model in DynaMIT uses Path-Size Logit model to calculate the probabilities of selecting various routes with following utility specification:

$$V_i = \beta_1 tt_{Ai} + (\beta_2 \beta_1) tt_{Fi} + \beta_3 N_{on-off} \tag{3.2}$$

where V_i is the utility of path i which is linear function of two travel time components and a penalty term. tt_{Ai} is the arterial component of travel time for path i and tt_{Fi} is the freeway component for the same path. N_{on-off} is the number of times a path changes link type (classified among freeway, arterial and ramp types) only to come back again to the original link type. This factor was introduced as it was observed that availability of many alternative paths between a freeway-to-freeway O-D pair were forcing an unreasonably high proportion of drivers to use part-arterial-based path. The penalty factor β_3 penalizes those paths which involve traveling from freeway to arterial to freeway again, or vice versa. β_1 is the coefficient of arterial travel time, and β_2 is freeway bias. Thus effective travel time coefficient for the freeway component of travel time is $\beta_1 \times \beta_2$.

Optimal route choice parameters were determined by using various combination of parameters and comparing the error in fit between simulated and observed counts. Results of several of these runs are summarized in Table 3.2. β_1 and β_3 are expected to be negative numbers, while β_2 is between 0 and 1.

Iteration	β_1	β_2	β_3	RMSE	RMSN
1	-0.005	0.5	-5.0	56.37	0.1444
2	-0.005	0.6	-5.0	56.09	0.1436
3	-0.005	0.3	-6.0	55.21	0.1414
4	-0.007	0.5	-4.0	61.52	0.1575
5	-0.010	0.5	-4.0	60.27	0.1543
6	-0.010	0.5	-5.0	54.37	0.1392
7	-0.010	0.8	-5.0	58.74	0.1504

Table 3.2: Error in Fit to Counts with various Route Choice Parameters

The optimal combination of (β_1, β_2, β_3) was found to be (-0.010, 0.50, -5.0).

In effect, penalty for changing link types was estimated to be equal to (β_3/β_1) 500 seconds of arterial travel time. A high value of 45 for path size exponent γ was selected to best fit the counts[6].

3.3.6 Starting O-D Matrix: Verifying Observability

In the absence of any information on the pattern of possible O-D pairs, we start by considering all possible O-D pairs. In this case, with 60 origins and 66 destination, there are total of 3908 potential O-D pairs that are required to be estimated. It is very likely that most of these will be zero (or negligibly small) in the final estimated matrix, yet we cannot rule out any pair beforehand lest we loose critical pairs in making uninformed decisions. However, as discussed in Section 1.2, it is not possible to estimate O-D pairs which are not counted by any sensor on the network. Thus we begin by screening all those O-D pairs which are not counted, based on the calibrated set of route choice parameters. After screening out 163 O-D pairs, we continue calibration process with 3745 pairs.

We start by randomly and independently generating three starting O-D matrices. Three cases are thus labeled as follows:

1. **Case D:** Deterministic starting OD, i.e $\mathbf{x}_{i,0} = 1 \; \forall \; i$

2. **Case U:** Starting OD uniformly distributed between 70 and 130, i.e $\mathbf{x}_{i,0} \sim$ U[70,130] $\forall \; i$

3. **Case N:** Starting OD normally distributed with mean 20 and standard deviation 5, i.e $\mathbf{x}_{i,0} \sim$ N(20,5) $\forall \; i$

Note that all the values are for a 15-minute estimation interval. While selecting the starting O-D flows, it was ensured that the three O-D flows are completely different both in terms of scale and the dispersion.

Results of the estimation process are presented below. Figure 3-26 compares the starting O-D flows in 'D' and 'N' cases, while subsequent figures (Figure 3-27 to 3-30)

[6]This is consistent with the experiences of [6] and [25].

compares estimated flows in 'D' and 'N' cases after each iteration. We note that as we use more and more intervals of sensor counts, the estimated O-D flows tend to converge.

Similar comparison statistics are presented for 'N' and 'U' cases. Figure 3-31 compares starting O-D matrices in this case, while following figures (Figure 3-32, 3-33, 3-34) compare estimated O-D flows after each estimation interval.

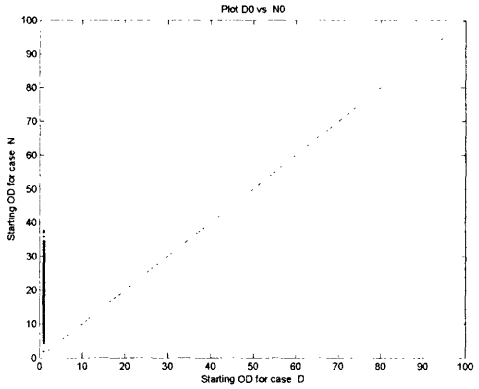


Figure 3-26: Comparing Starting O-D for 'D' & 'N' cases

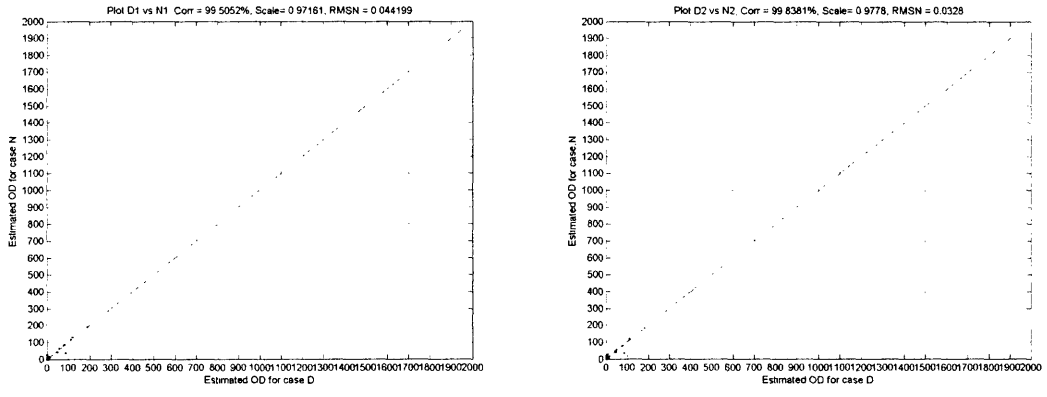


Figure 3-27: Interval 1–2: Comparing Estimated O-D for 'D' & 'N' cases

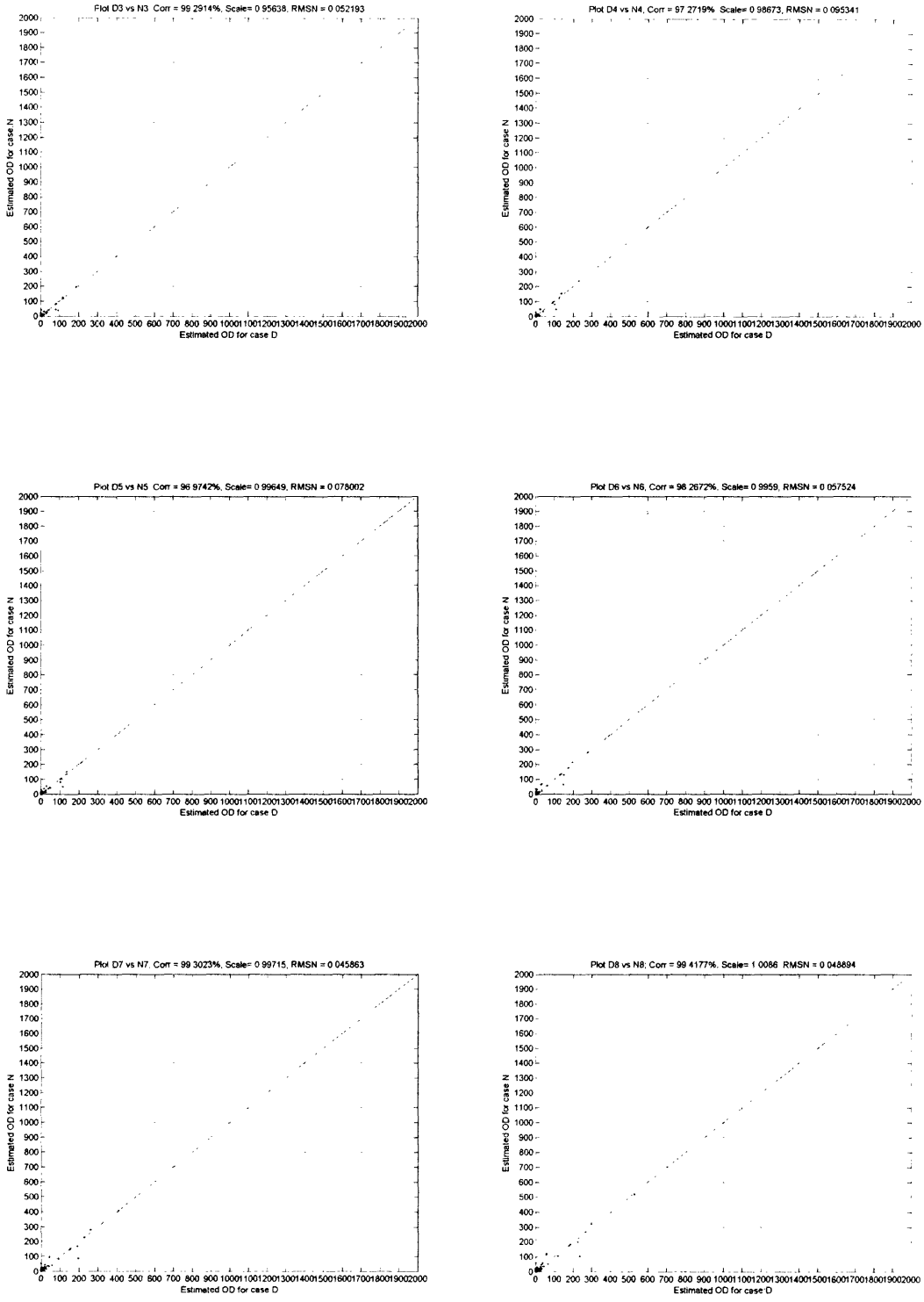


Figure 3-28: Interval 3–8: Comparing Estimated O-D for 'D' & 'N' cases

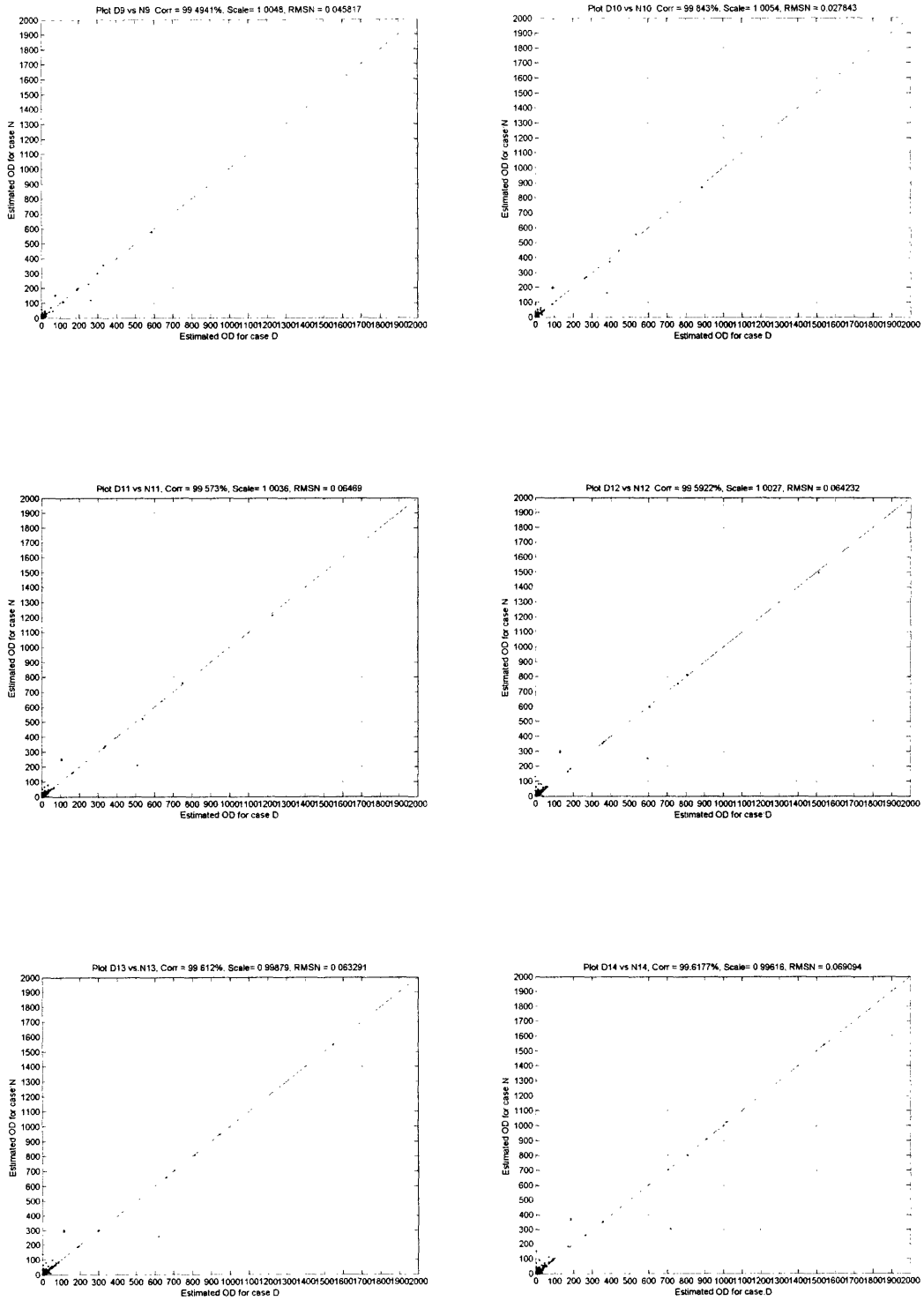


Figure 3-29: Interval 9–14: Comparing Estimated O-D for 'D' & 'N' cases

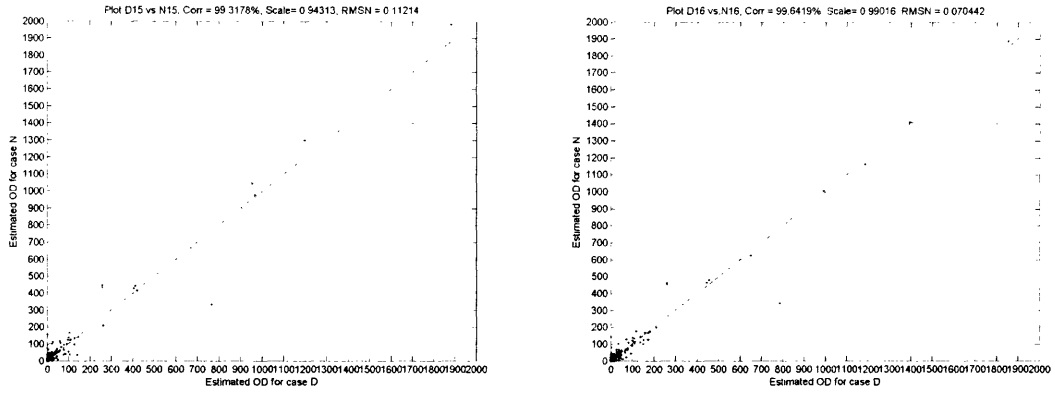


Figure 3-30: Interval 15–16: Comparing Estimated O-D for 'D' & 'N' cases

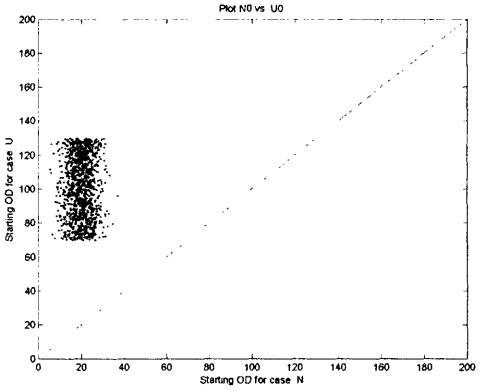


Figure 3-31: Comparing Starting O-D for 'N' & 'U' cases

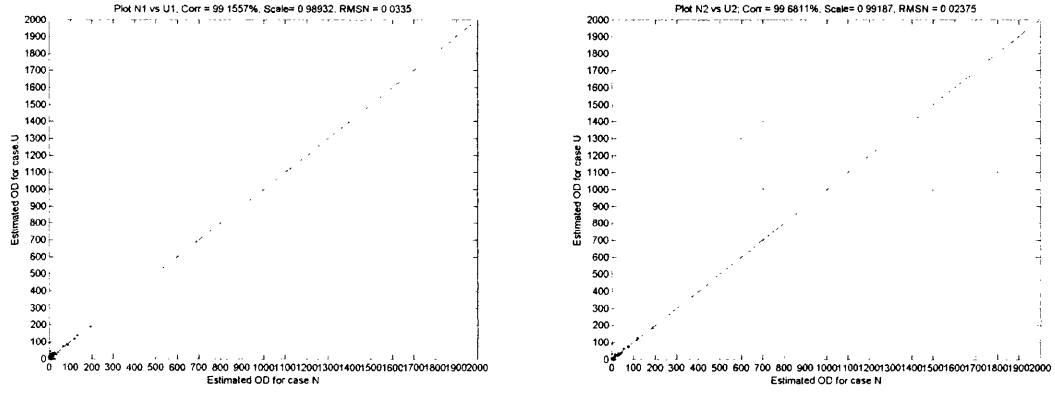


Figure 3-32: Interval 1–2: Comparing Estimated O-D for 'N' & 'U' cases

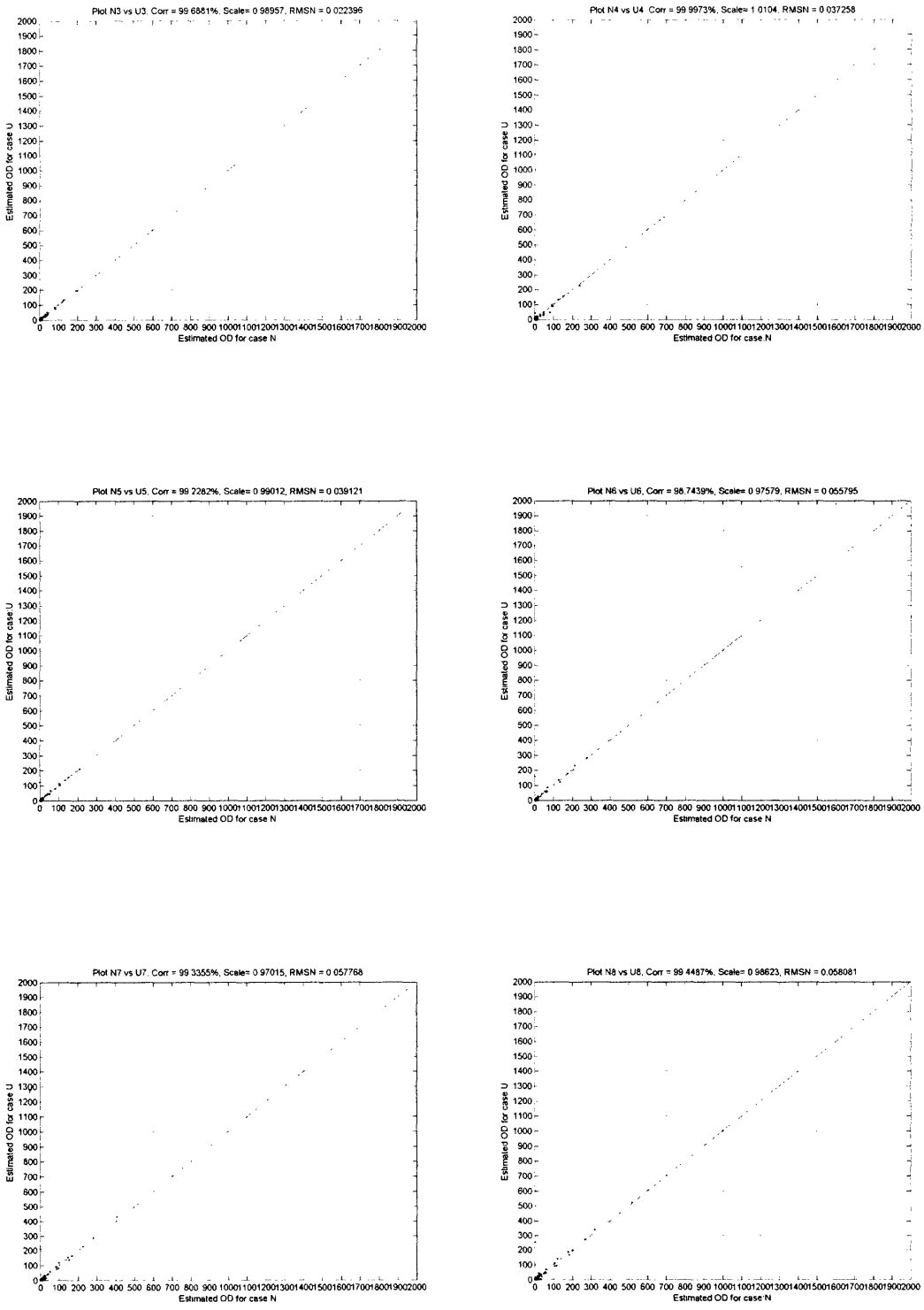


Figure 3-33: Interval 3–8: Comparing Estimated O-D for 'N' & 'U' cases

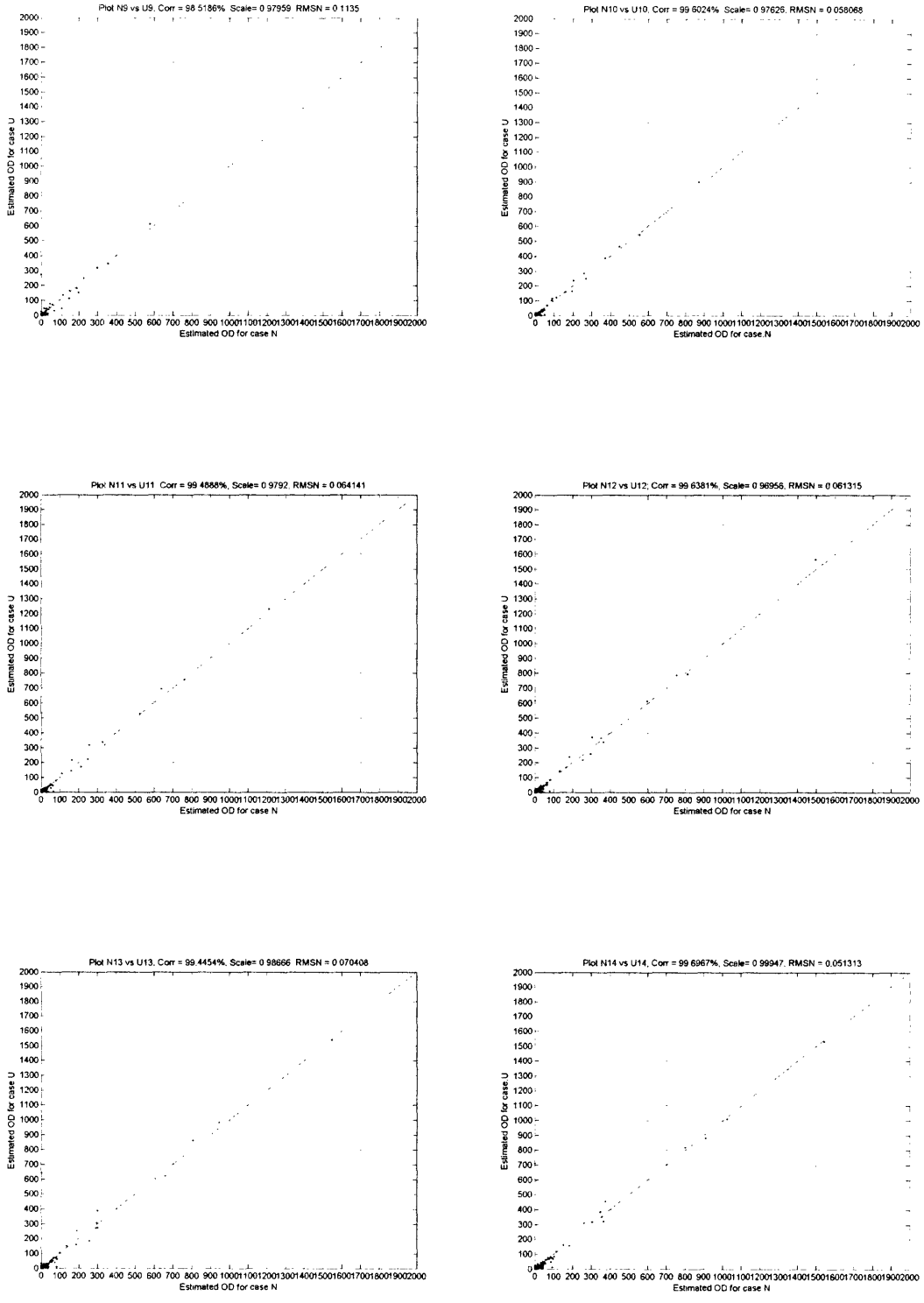


Figure 3-34: Interval 9–14: Comparing Estimated O-D for 'N' & 'U' cases

Table 3.3 summarizes error statistics comparing estimated O-D flows across various intervals for 'D' and 'N' case. Similar results for 'N' and 'U' cases are tabulated in Table 3.4.

Interval	Scale	RMSN
1	0.971	0.044
2	0.978	0.033
3	0.956	0.052
4	0.987	0.095
5	0.996	0.078
6	0.996	0.057
7	0.997	0.046
8	1.009	0.049
9	1.005	0.046
10	1.005	0.028
11	1.004	0.065
12	1.003	0.064
13	0.999	0.063
14	0.996	0.069
15	0.943	0.112
16	0.990	0.070

Table 3.3: Comparison of Estimated O-D flows for case 'D' and 'N'

Before we close the section, a final note is in order here. We observe that the two points (O-D pairs) are consistently off-diagonal in all the intervals in case 'D' and 'N'. This is because these particular O-D pairs are major freeway-to-freeway O-D pairs which are not counted along their main freeway path. Thus information about these two O-D pairs is captured by minuscule (less than 0.1%) fraction of traffic which uses arterial streets. This reconfirms our initial assertion that each O-D flow must be counted by at least one sensor to be estimated, and preferably by more than one sensor with significant fractions, to be estimated reliably.

As evident from the statistics and comparison plots, we note that estimated O-D flows indeed converge to a unique set of O-D flows irrespective of different starting points. Finally, 1129 O-D pairs out of 3745 pairs were found to have non-zero flows. As the O-D estimation in this network is proven observable, once a unique set of non-zero O-D pairs were identified, starting O-D matrix (or seed O-D matrix) for

Interval	Scale	RMSN
1	0.989	0.034
2	0.992	0.024
3	0.990	0.022
4	1.010	0.037
5	0.990	0.039
6	0.976	0.056
7	0.970	0.058
8	0.986	0.058
9	0.980	0.113
10	0.976	0.058
11	0.980	0.064
12	0.970	0.061
13	0.987	0.070
14	0.999	0.051

Table 3.4: Comparison of Estimated O-D flows for case 'N' and 'U'

each interval was taken from the best estimated O-D matrix from previous intervals for the first two days[7] of calibration, and was generated from historical matrix[8] from third day onwards. Mathematically,

$$\begin{aligned} \mathbf{x}_h^{ad} &= \hat{\mathbf{x}}_{h-1}^{d} \qquad \text{for } d = 1, 2 \\ \mathbf{x}_h^{ad} &= \frac{\mathbf{x}_h^H}{\mathbf{x}_{h-1}^H} \hat{\mathbf{x}}_{h-1}^{d} \qquad \text{for } d > 2 \end{aligned} \tag{3.3}$$

where $\mathbf{x}_h^{ad}$ is the a-priori O-D matrix for interval h for day d and $\hat{\mathbf{x}}_h^d$ and $\mathbf{x}_h^H$ are the corresponding estimated and historical O-D matrices respectively.

3.3.7 O-D Estimation Calibration

As all the weekdays were observed to be similar in terms of day-to-day and within-day demand patterns, a single historical database of O-D matrices and other parameters was estimated for all weekdays from Monday to Friday. Five days were selected

[7]Since assuming estimates of only one day of calibration as historical is likely more erroneous, given that variance of mean is inversely proportional to number of data points.

[8]Average of estimated O-D matrices for interval h for all previously calibrated days was designated as historical O-D matrix for interval h.

spanning the entire month of September and all days of week (Monday–Friday) to carry out the sequential calibration process.

We started from the first day of calibration, which was taken as September 1^{st}, 2004, Wednesday. We started with a seed O-D matrix and after about 16 intervals (15-minutes each) found that the O-D matrix was observable. In absence of any initial estimates of error covariance, a Feasible Generalized Least Squares[9] (FGLS) approach was adopted with very high weights to sensor counts and very small weights to seed O-D flows. This is because we wanted to extract all possible information from sensor counts, while we had little reason to trust our initial estimate of seed O-D flows. Seed O-D flows for later intervals were generated as described by equation (3.3). Generalized Least Square formulation as outlined in Section 2.1.3 was used to estimate O-D matrix interval by interval. Once first day was completely estimated, error covariance matrices were calculated as described in Section 2.3.4.

Second day was taken as Friday, September 24^{th}, 2004. Calibration procedure for the second day was exactly similar to the first day, except—calculated error covariance matrix was used in GLS formulation. At the end of second day, we recalculated error covariance matrix, as well as updated historical O-D flows as an average of estimated flows from first two day.

This approach was repeated with third day (September 13^{th}, Monday), fourth day (September 21^{st}, Tuesday) and fifth day (September 9^{th}, Thursday), every time recalculating the error covariance matrix and updating historical O-D matrices and experienced travel times. During fourth and fifth day, it was observed that number of iterations required to converge to estimated O-D were very few, and only one in many cases, thus signifying that our database of historical O-D is approaching toward true historical O-D flows.

Table 3.5 presents the estimated error statistics for the AM Peak from 7:00 AM to 9:00 AM for 5 days for which estimation was carried out. Similar results are presented for PM Peak from 5:00 PM to 7:00 PM for 5 days of calibration in Table 3.6. For the sake of brevity of the report, graphical comparisons of simulated and observed

[9]Reader is referred to [15] for detailed theoretical treatment of the subject.

counts for the same periods but only for 3 of the 5 days are presented in Figures 3-35 to 3-58. Detailed error statistics for full calibration period for all 5 days are presented in Appendix B.

Interval Ending	Day 1	Day 2	Day 3	Day 4	Day 5
7:15	0.1558	0.1512	0.1221	0.1301	0.1693
7:30	0.1337	0.1700	0.1374	0.1414	0.1575
7:45	0.1661	0.1919	0.1637	0.1494	0.1743
8:00	0.1696	0.1878	0.1866	0.1621	0.1778
8:15	0.1601	0.1860	0.1605	0.1711	0.1877
8:30	0.1562	0.1815	0.1693	0.1516	0.1848
8:45	0.1495	0.1865	0.1563	0.1729	0.1919
9:00	0.1285	0.1811	0.1648	0.1649	0.1873

Table 3.5: RMSN Statistics for AM Peak Calibration for 5 days

Interval Ending	Day 1	Day 2	Day 3	Day 4	Day 5
17:15	0.1429	0.2089	0.1582	0.1715	0.1876
17:30	0.1442	0.2227	0.1599	0.1674	0.1817
17:45	0.1534	0.2165	0.1527	0.1761	0.1744
18:00	0.1630	0.2034	0.1474	0.1713	0.1721
18:15	0.1538	0.1771	0.1480	0.1655	0.1680
18:30	0.1520	0.1816	0.1431	0.1624	0.1700
18:45	0.1523	0.1752	0.1484	0.1700	0.1654
19:00	0.1466	0.1617	0.1336	0.1645	0.1603

Table 3.6: RMSN Statistics for PM Peak Calibration for 5 days

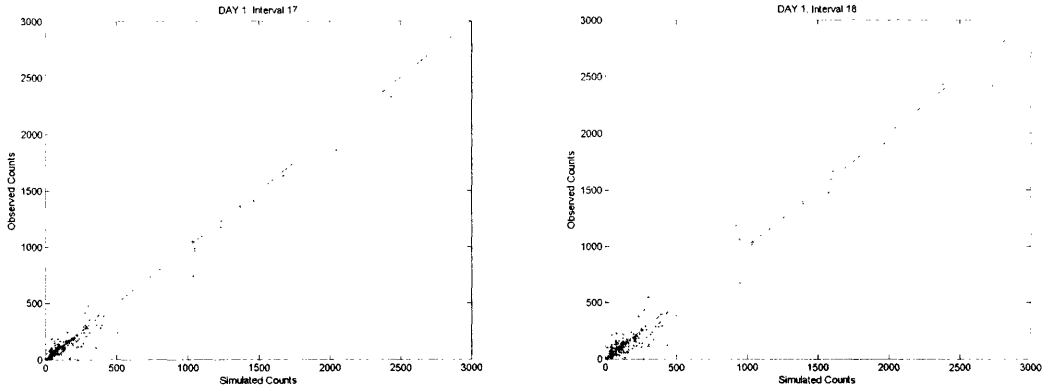


Figure 3-35: Calibration Results for Day 1: 7:00 AM–7:30 AM

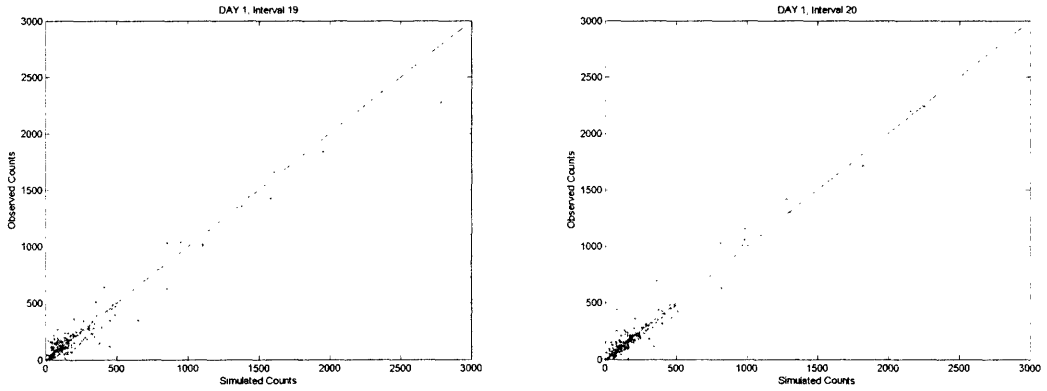


Figure 3-36: Calibration Results for Day 1: 7:30 AM–8:00 AM

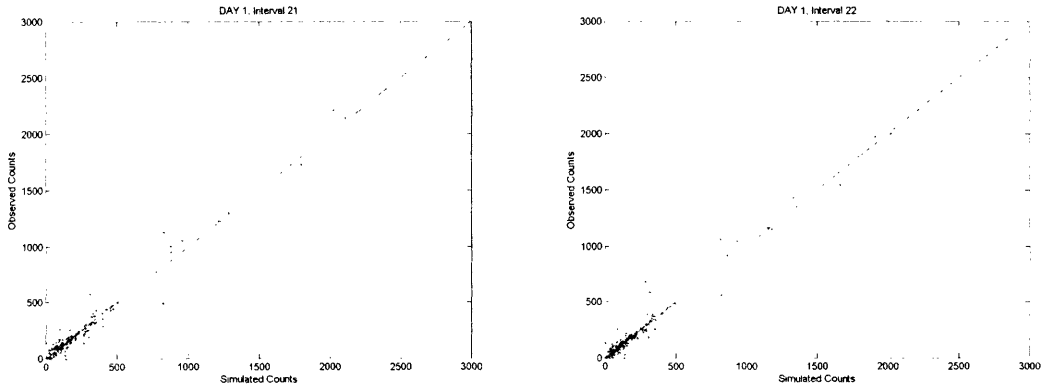


Figure 3-37: Calibration Results for Day 1: 8:00 AM–8:30 AM

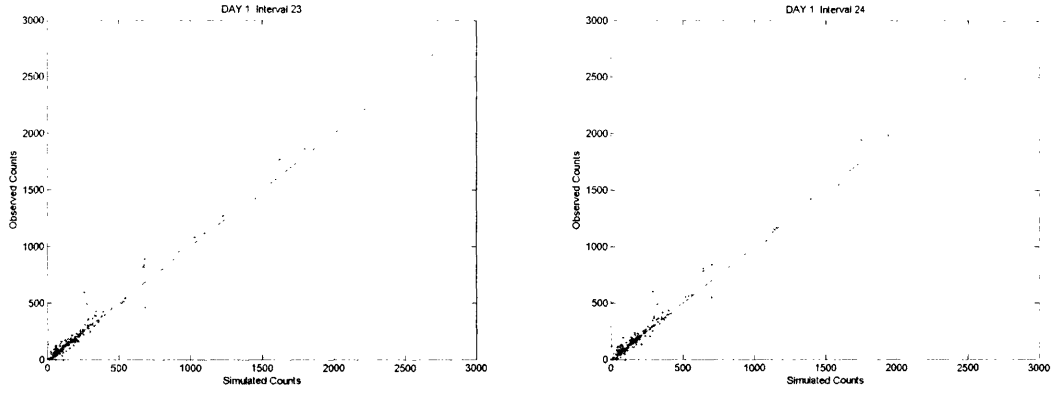


Figure 3-38: Calibration Results for Day 1: 8:30 AM–9:00 AM

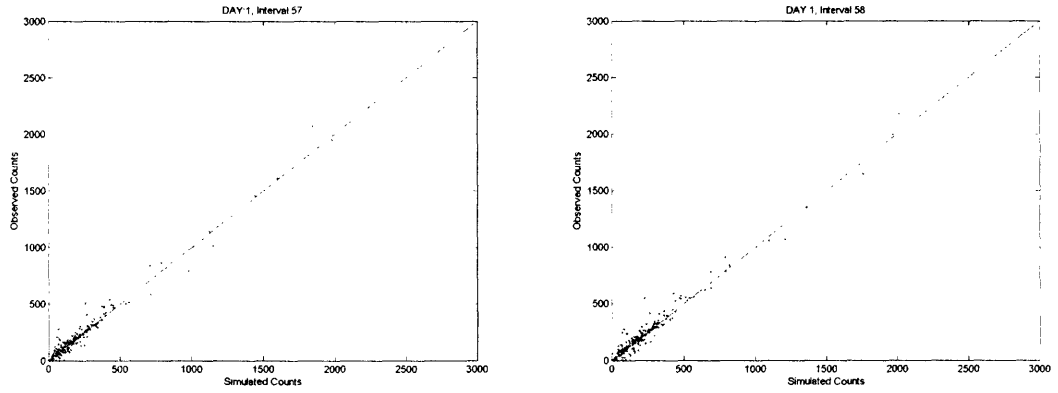


Figure 3-39: Calibration Results for Day 1: 5:00 PM–5:30 PM

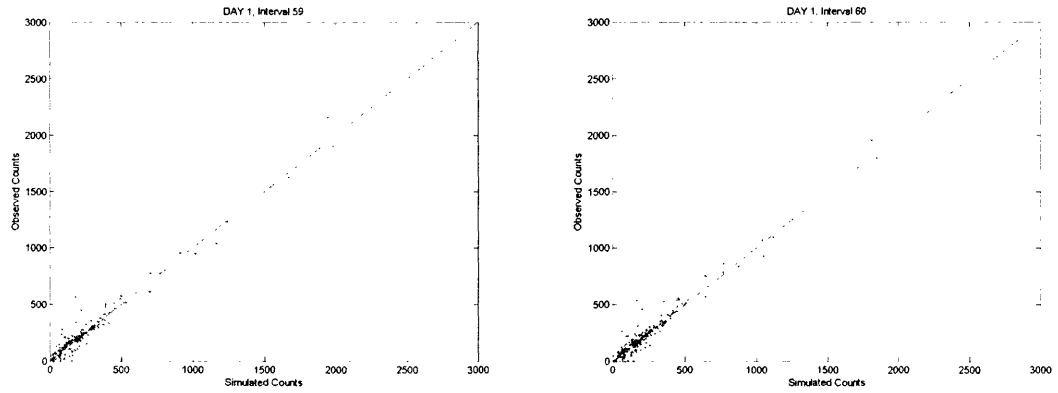


Figure 3-40: Calibration Results for Day 1: 5:30 PM–6:00 PM

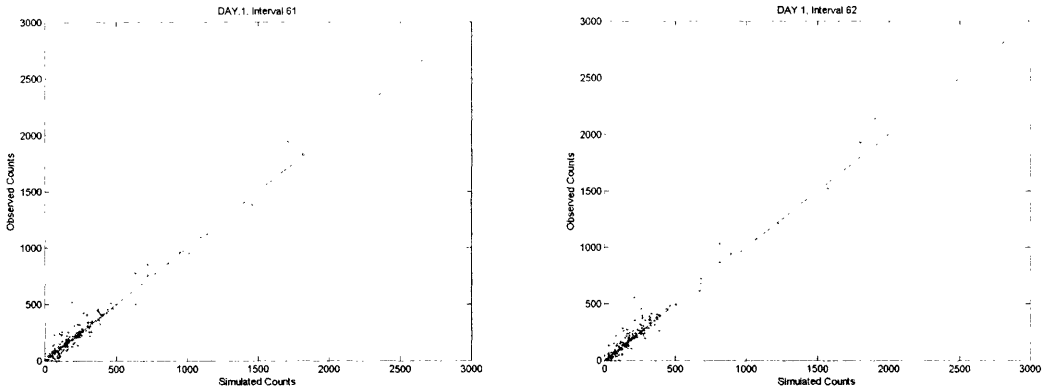


Figure 3-41: Calibration Results for Day 1: 6:00 PM–6:30 PM

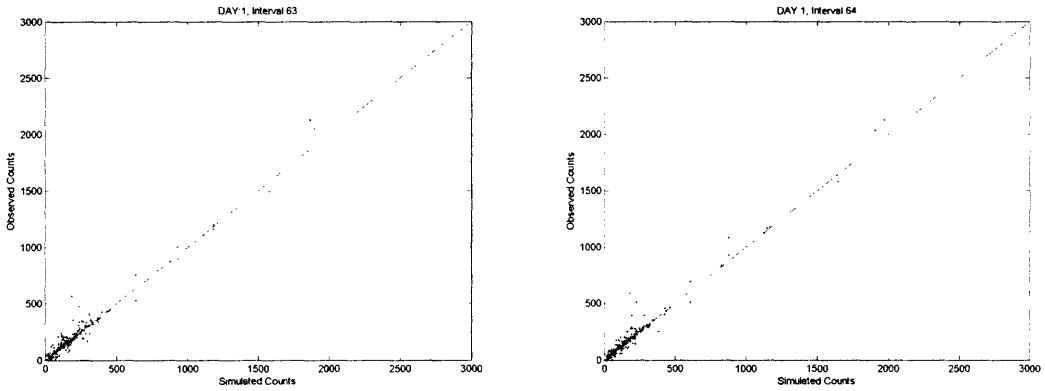


Figure 3-42: Calibration Results for Day 1: 6:30 PM–7:00 PM

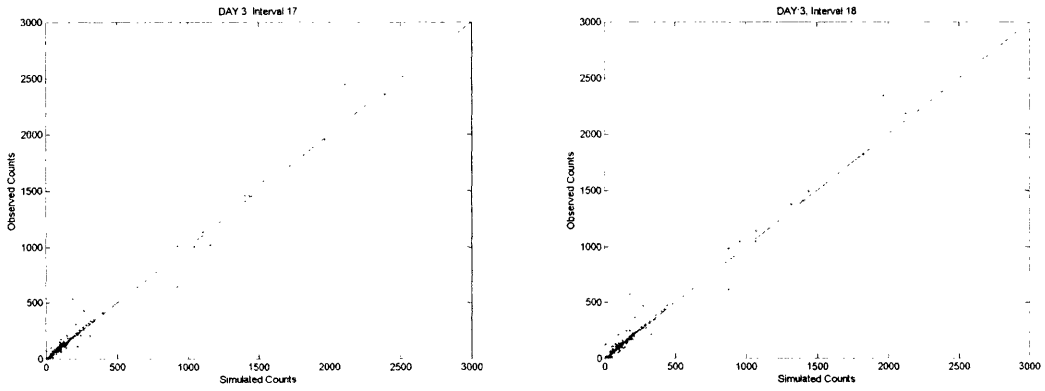


Figure 3-43: Calibration Results for Day 3: 7:00 AM–7:30 AM

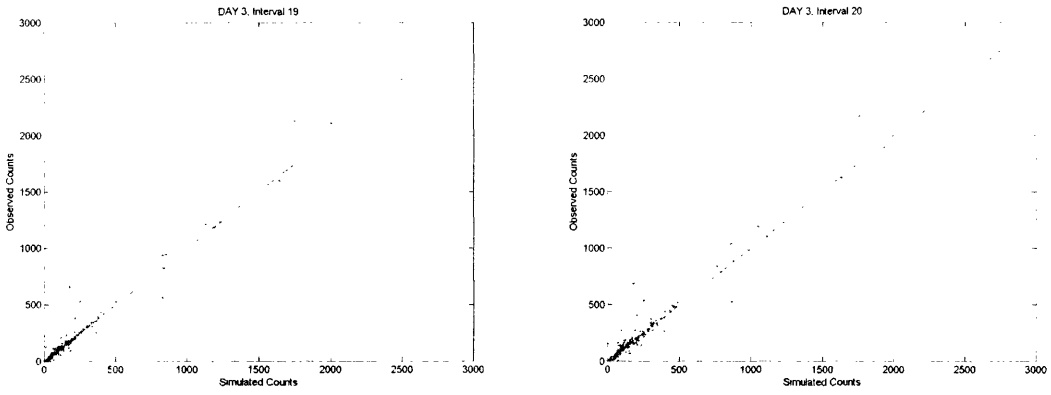


Figure 3-44: Calibration Results for Day 3: 7:30 AM–8:00 AM

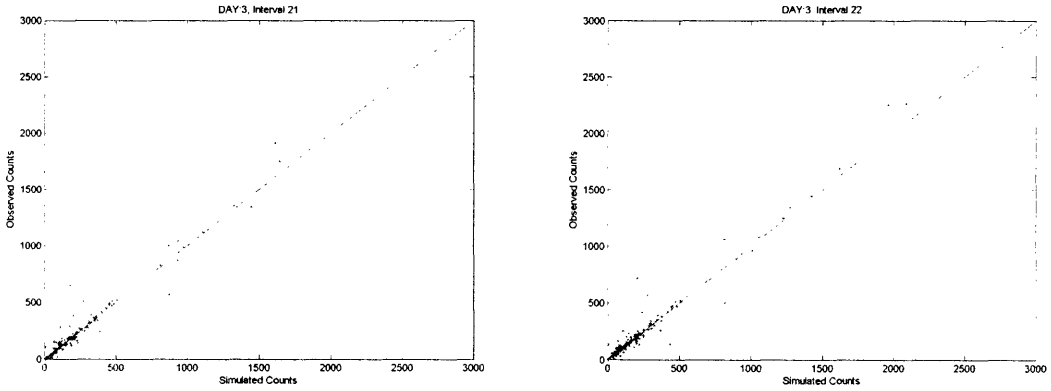


Figure 3-45: Calibration Results for Day 3: 8:00 AM–8:30 AM

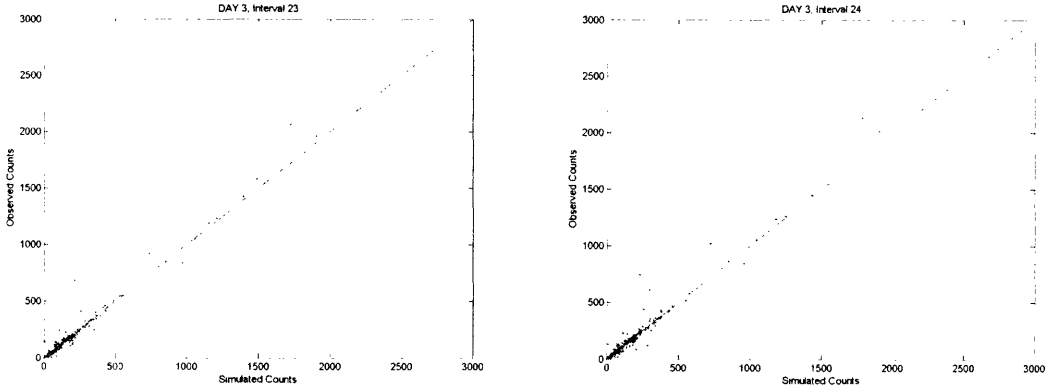


Figure 3-46: Calibration Results for Day 3: 8:30 AM–9:00 AM

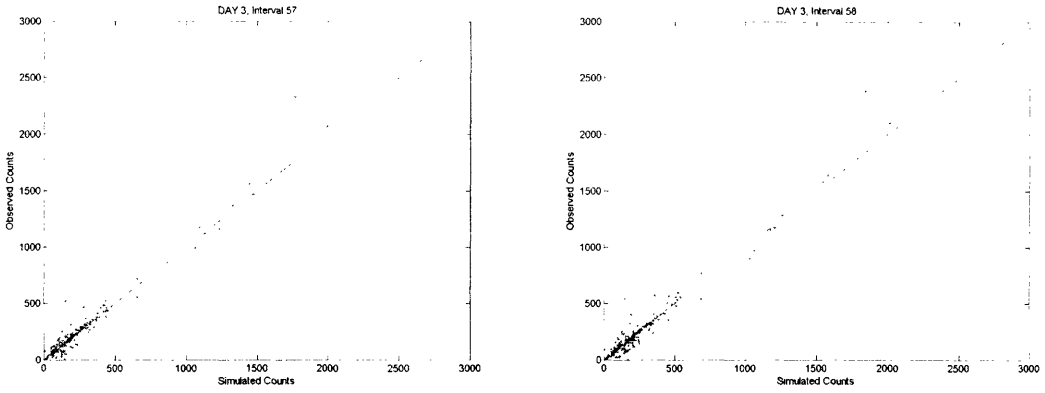


Figure 3-47: Calibration Results for Day 3: 5:00 PM–5:30 PM

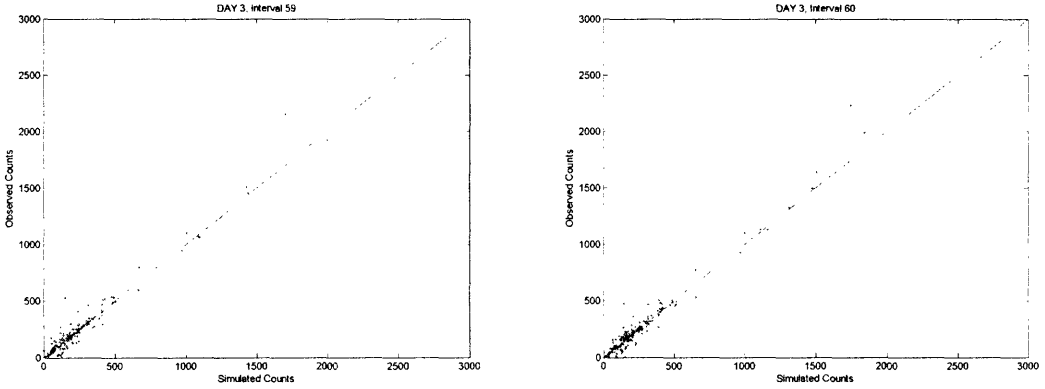


Figure 3-48: Calibration Results for Day 3: 5:30 PM–6:00 PM

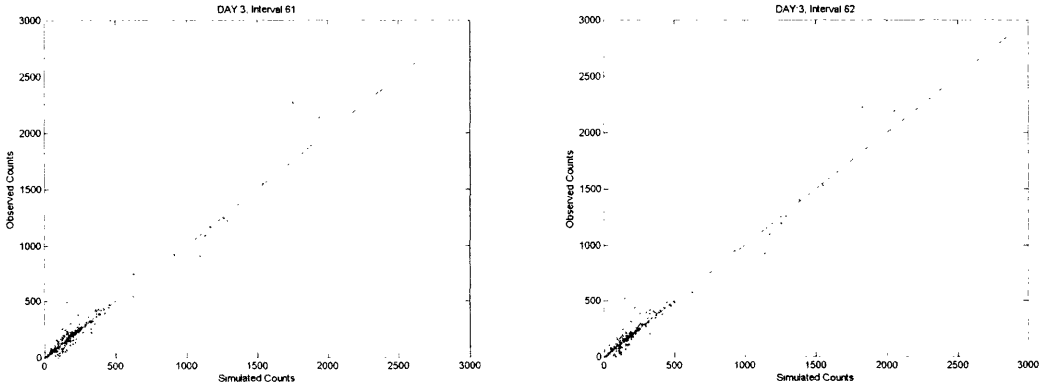


Figure 3-49: Calibration Results for Day 3: 6:00 PM–6:30 PM

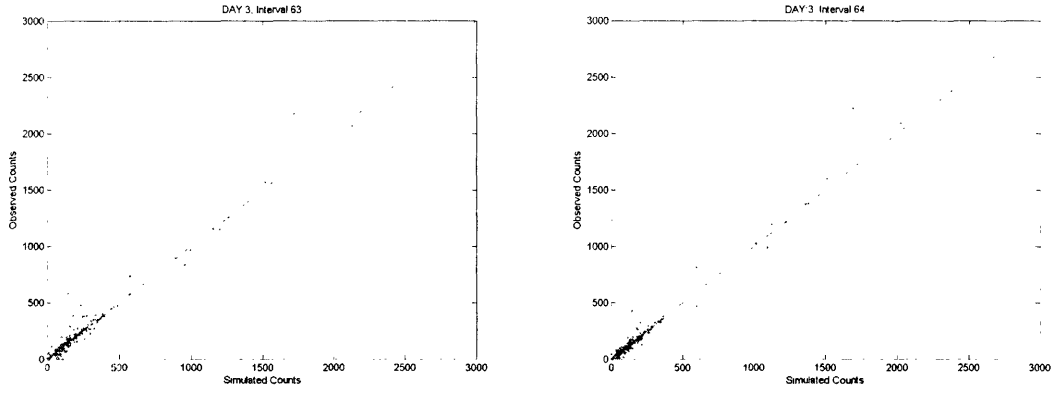


Figure 3-50: Calibration Results for Day 3: 6:30 PM–7:00 PM

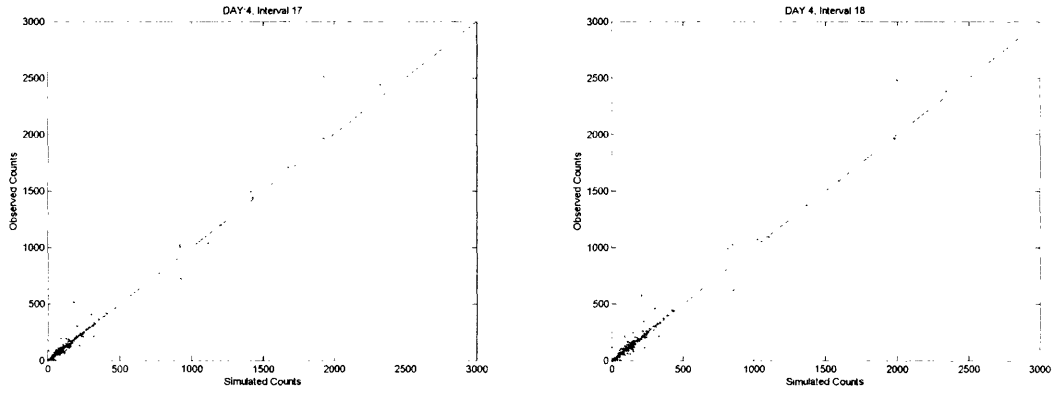


Figure 3-51: Calibration Results for Day 4: 7:00 AM–7:30 AM

The performance of calibration was also verified by comparing time-varying estimated counts with observed counts from the field. Some sample sensors are presented in Figures 3-59 to 3-66.

Having calibrated for 5 days, we next proceeded to calculate auto-regressive factors which will form basis for DynaMIT's predictive capabilities. We started by calculating deviations of estimated O-D flows for all 5 days with corresponding historical O-D flows, and then regressed them as proposed in equation (2.40). O-D pairs were grouped into four categories: high ($>$ 200 veh/hr), medium (50–200 veh/hr), low (20–50 veh/hr) and very low ($<$ 20 veh/hr). Similarly study period were also grouped

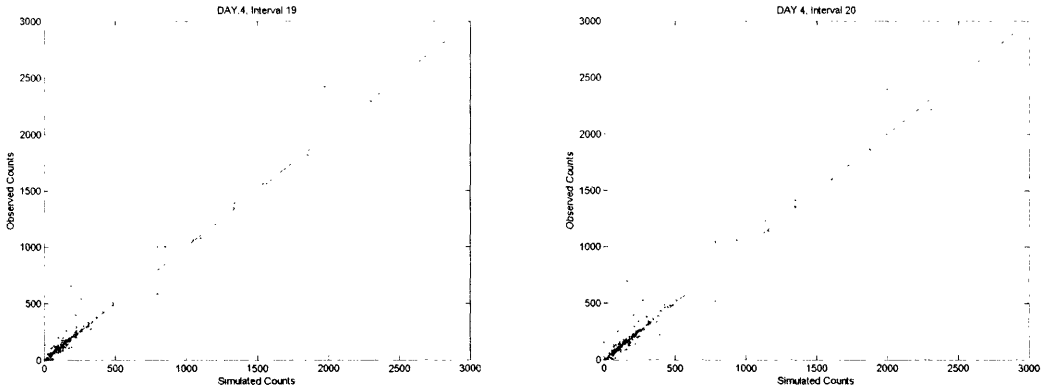


Figure 3-52: Calibration Results for Day 4: 7:30 AM–8:00 AM

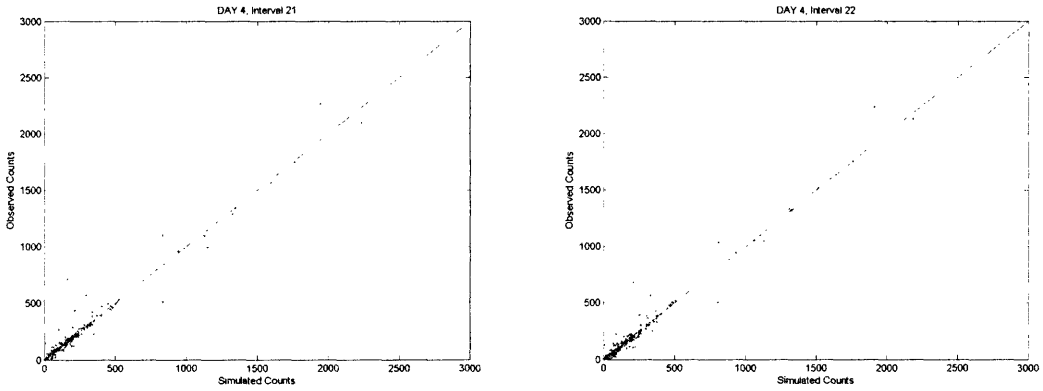


Figure 3-53: Calibration Results for Day 4: 8:00 AM–8:30 AM

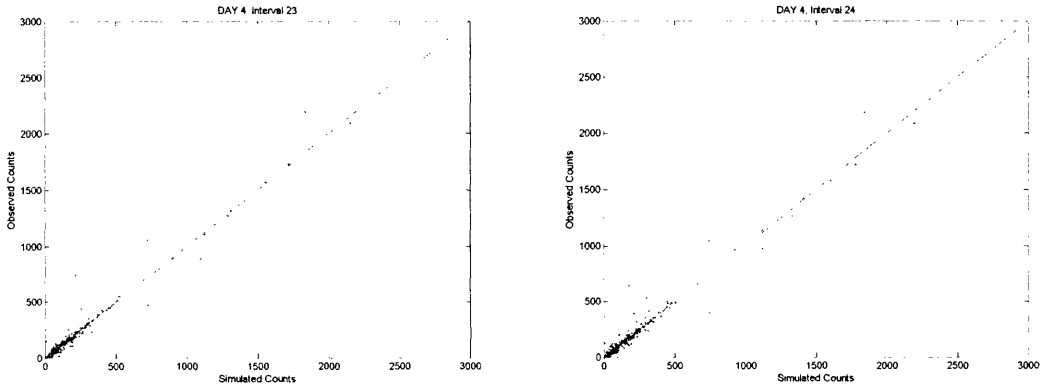


Figure 3-54: Calibration Results for Day 4: 8:30 AM–9:00 AM

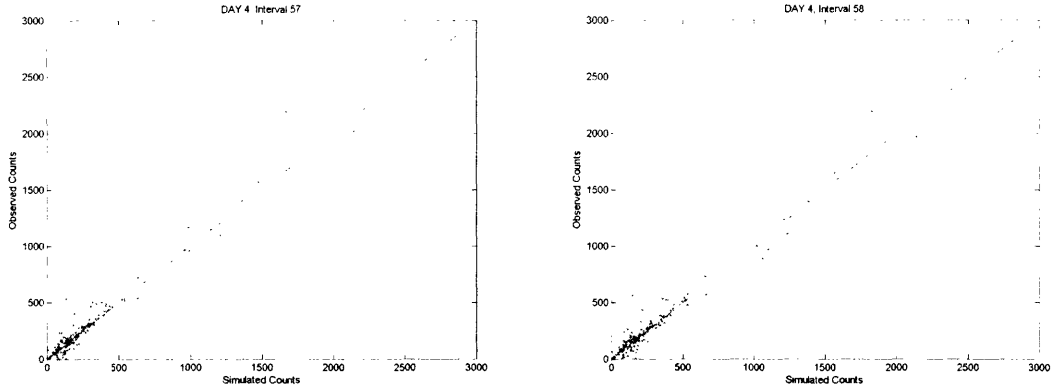


Figure 3-55: Calibration Results for Day 4: 5:00 PM–5:30 PM

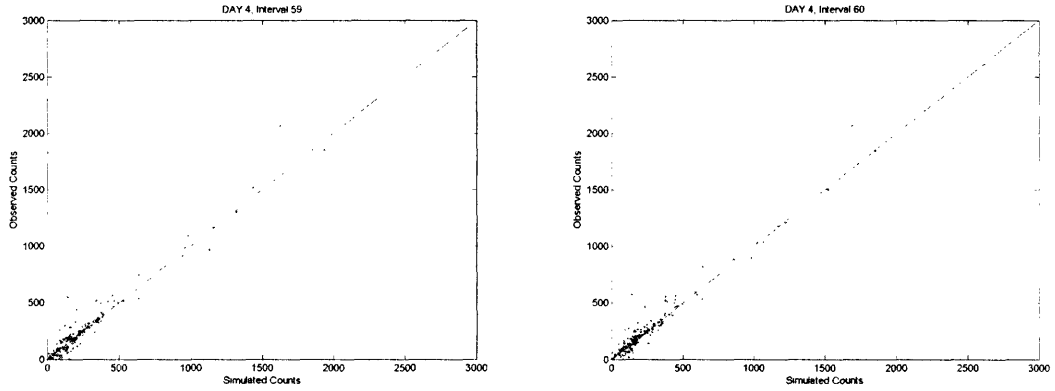


Figure 3-56: Calibration Results for Day 4: 5:30 PM–6:00 PM

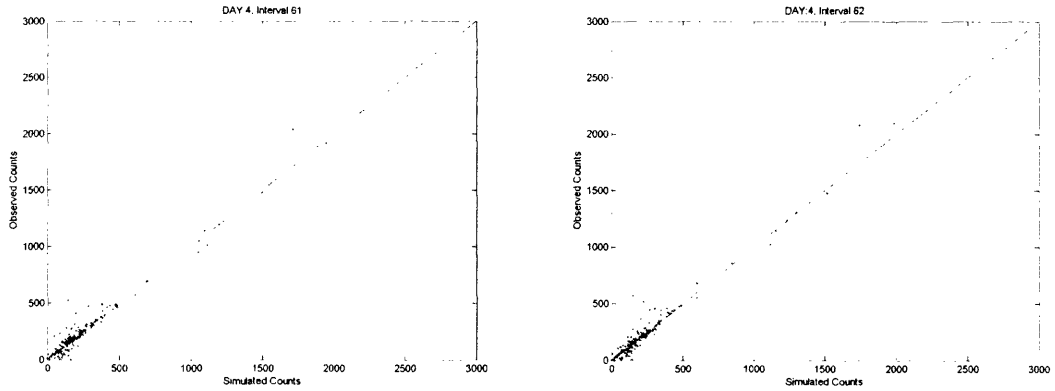


Figure 3-57: Calibration Results for Day 4: 6:00 PM–6:30 PM

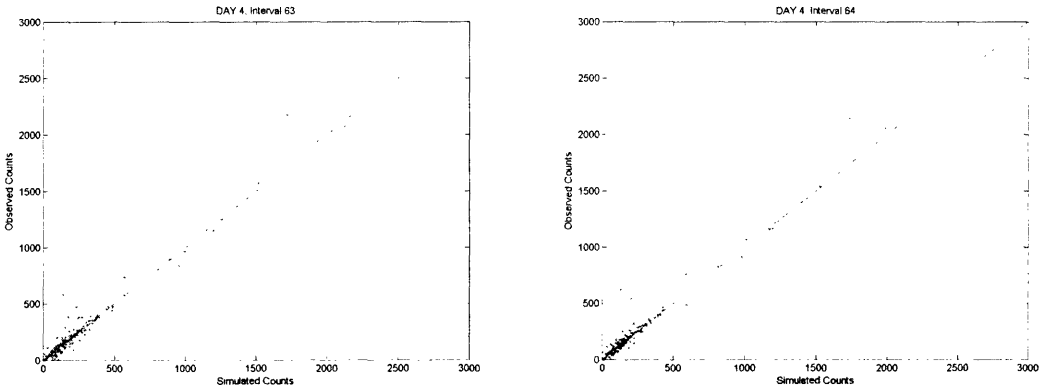


Figure 3-58: Calibration Results for Day 4: 6:30 PM–7:00 PM

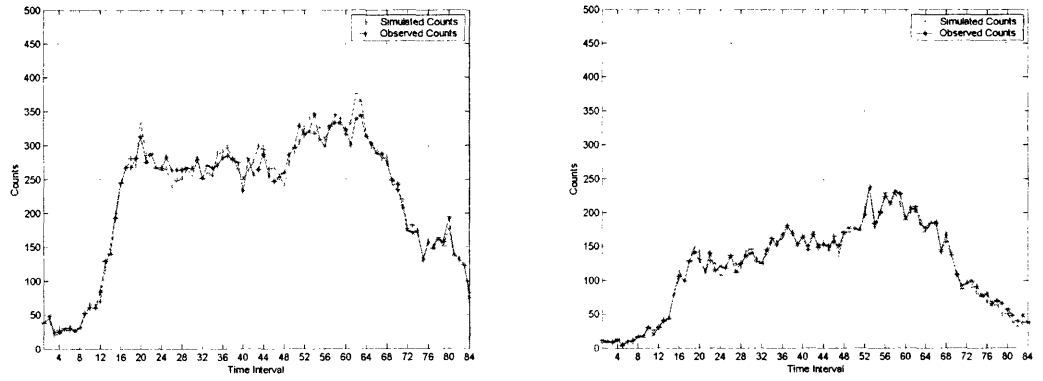


Figure 3-59: Day 2: Arterial Sensor 3015 (left) and 3017 (right)

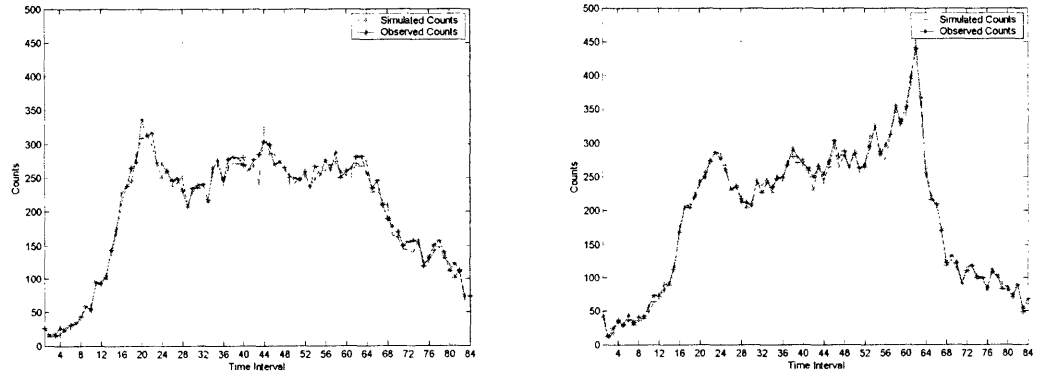


Figure 3-60: Day 2: Arterial Sensor 5007 (left) and 7011 (right)

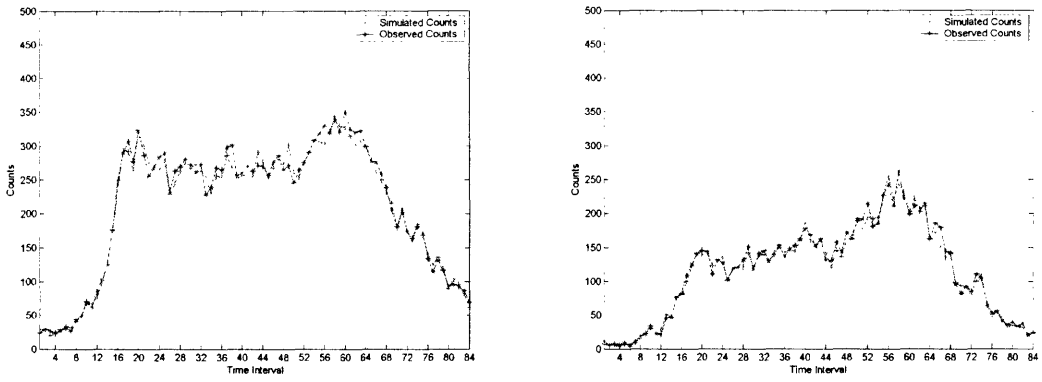


Figure 3-61: Day 5: Arterial Sensor 3015 (left) and 3017 (right)

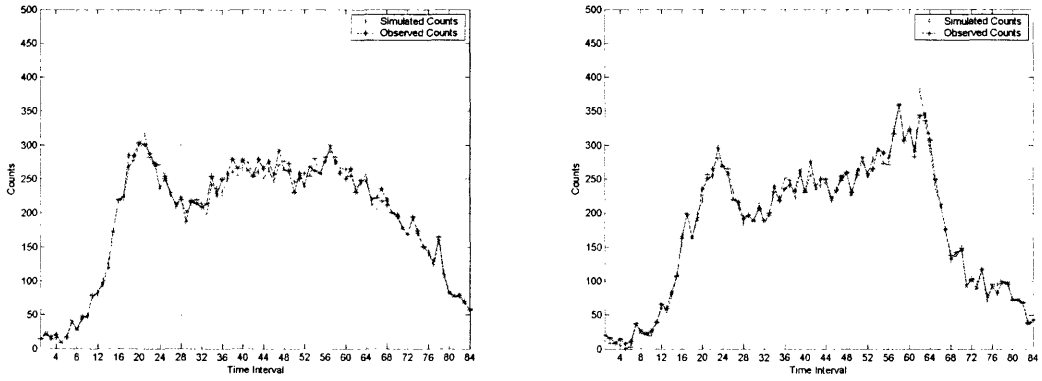


Figure 3-62: Day 5: Arterial Sensor 5007 (left) and 7011 (right)

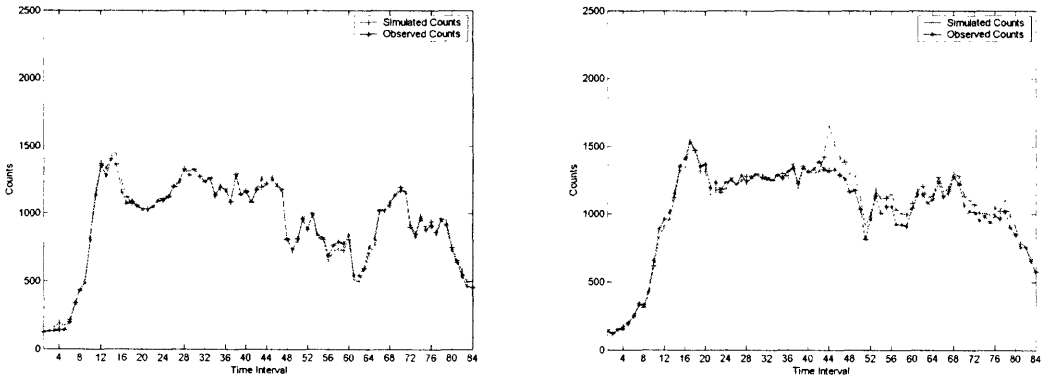


Figure 3-63: Day 2: Freeway Sensor 718165 (left) and 718166 (right)

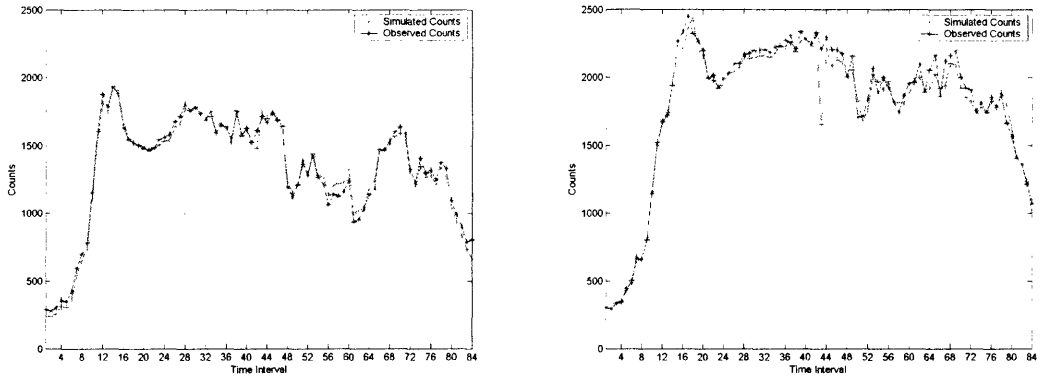


Figure 3-64: Day 2: Freeway Sensor 764032 (left) and 764037 (right)

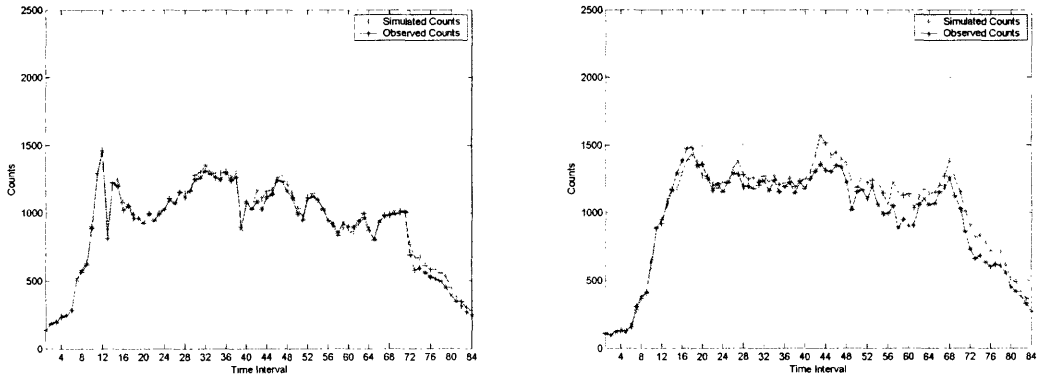


Figure 3-65: Day 5: Freeway Sensor 718165 (left) and 718166 (right)

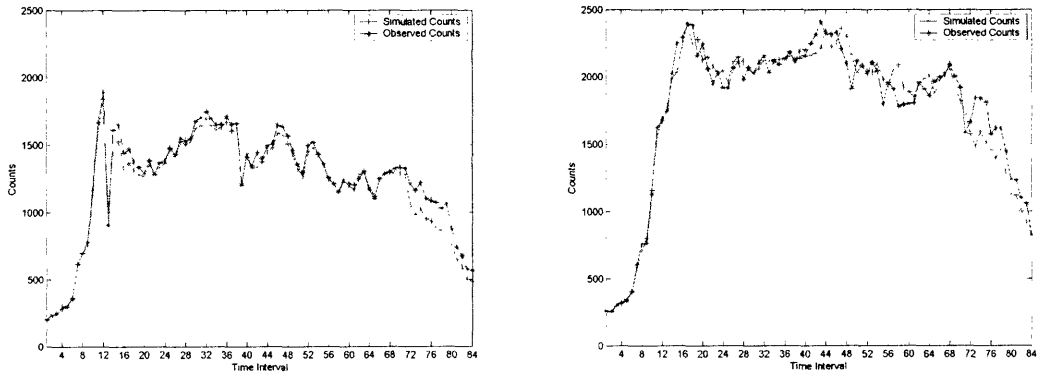


Figure 3-66: Day 5: Freeway Sensor 764032 (left) and 764037 (right)

into 3 periods and separate set of factors were calibrated for each group of O-D flow in that time period. Calibrated AR factors are presented in Table 3.7.

O-D Group	Time Group	AR Degree	$f_{r,h}^{r,h-1}$	$f_{r,h}^{r,h-2}$	$f_{r,h}^{r,h-3}$	$f_{r,h}^{r,h-4}$
> 200	4:00 AM–8:00 AM	4	0.7296	0.2195	0.1495	-0.1827
	8:00 AM–7:00 PM	4	0.5322	0.0763	0.1371	0.1832
	7:00 PM–0:00 AM	3	0.7678	-0.1415	0.2708	-
50–200	4:00 AM–8:00 AM	4	0.9020	0.1286	-0.0055	-0.0128
	8:00 PM–7:00 PM	4	0.5855	0.2294	0.0563	0.0615
	7:00 PM–0:00 AM	3	0.6603	0.1383	0.0696	-
20–50	4:00 AM–8:00 AM	4	0.7216	0.0280	0.0159	0.1735
	8:00 AM–7:00 PM	4	0.6232	0.1618	0.1042	0.0485
	7:00 PM–0:00 AM	3	0.6937	0.1613	-0.0239	-
< 20	4:00 AM–8:00 AM	4	0.5490	0.0532	0.0948	0.0841
	8:00 AM–7:00 PM	4	0.6656	0.1720	0.0466	0.0358
	7:00 PM–0:00 AM	3	0.6604	0.1144	0.0502	-

Table 3.7: Calibrated Auto-Regressive Factors for Weekdays

Having calibrated both demand and supply parameters for DynaMIT for weekdays, we move on to verify estimation and prediction capabilities of the calibrated model through validation tests. Since calibration procedure for weekends is precisely the same as weekdays, complete calibration results for Saturdays and Sundays are presented in Appendix C.

3.4 Validation of Calibration Results

The calibrated system was tested for its estimation and prediction performance using another day of data which was not part of calibration process. We used data from September 29^{th}, Wednesday and September 7^{th}, Tuesday for this purpose.

3.4.1 Validation of Estimation Capabilities

Validation of DynaMIT's estimation capabilities tests if the calibrated set of historical O-D matrices, historical travel times, and variance-covariance matrices continue to perform satisfactorily when supplied with fresh input, as would be the case in on-site

application. Process of validation is exactly same as the calibration, except that it is performed with new day's sensor data. Table 3.8 reports the error statistics for both days of validation process for AM Peak period as compared to calibration statistics of Day 5, while results for PM Peak are listed in Table 3.9.

Interval Ending	Day 5	Sept 29	Sept 07
7:15	0.1693	0.1678	0.1110
7:30	0.1575	0.1553	0.1245
7:45	0.1743	0.1630	0.1479
8:00	0.1778	0.1770	0.1613
8:15	0.1877	0.1890	0.1521
8:30	0.1848	0.1921	0.1569
8:45	0.1919	0.1779	0.1634
9:00	0.1873	0.1831	0.1525

Table 3.8: RMSN Statistics for Validation of Estimation Capabilities for AM Peak

Interval Ending	Day 5	Sept 29	Sept 07
17:15	0.1876	0.1674	0.1530
17:30	0.1817	0.2001	0.1737
17:45	0.1744	0.1853	0.1707
18:00	0.1721	0.1731	0.1671
18:15	0.1680	0.1735	0.1751
18:30	0.1700	0.1862	0.1614
18:45	0.1654	0.1883	0.1950
19:00	0.1603	0.1922	0.1989

Table 3.9: RMSN Statistics for Validation of Estimation Capabilities for PM Peak

We observe that errors during validation runs are comparable to those obtained during calibration process, and are within acceptable limit. Complete validation results for full day of data are presented in Appendix B. Having verified that DynaMIT's estimation is stabilized, we proceed to test prediction capability next.

3.4.2 Validation of Prediction Capabilities

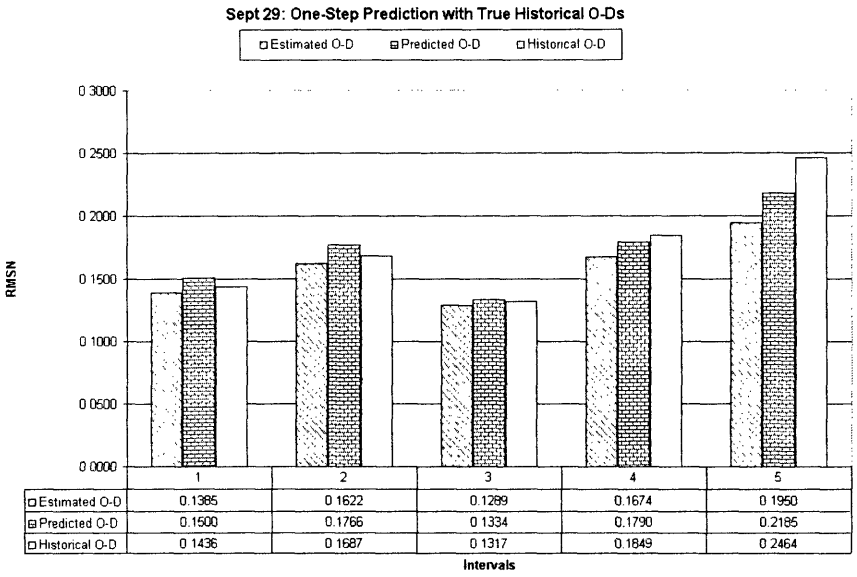
Prediction tests for calibrated auto-regressive (AR) parameters were performed using data from September 29, Wednesday and September 30, Thursday, 2004. We started

by estimating O-D flows for selected time intervals for both of these days. 5 intervals were picked uniformly during the day to test effectiveness of the calibrated AR factors. We then predicted deviations in the O-D flows for next 15 minute period (one-step prediction) using estimated O-D flows so far for each time period using equation (2.40). Similarly, prediction for next 15 minute (two-step prediction) were performed using set of estimated O-D flows from past intervals, and just calculated predicted O-D flow for previous interval. In the same manner three- and four-step predictions (up to one hour in future) were carried out. The predicted set of O-D flows were then simulated and the simulated counts were compared with the observed counts. To benchmark the prediction and estimation errors, historical O-D flows for prediction period were simulated and compared against observed counts.

Here historical flows represent the performance of DynaMIT if real-time inputs from the field are not utilized at all. Estimated flows, on the other hand, represent the situation when DynaMIT updates and re-estimates current network condition based on historical flows and surveillance data. Since estimated O-D flows are optimized to match day specific sensor data and encompass not only historical information but also real-time sensor information, they are our best estimates of true O-D flows on the network in that interval. Finally, predicted O-D flows are predictions of future O-D flows based on latest available estimates of O-D flows.

Figures 3-67 to 3-74 compare error statistics for one-, two-, three- and four-step predictions for both days, among estimated O-D flows, predicted O-D flows, and historical O-D flows. As we would have expected, RMSN corresponding to estimated O-D flows is always lowest among the three since estimated flows are the best estimates of true O-D flows. However, we note that predicted O-D flows do not have consistent advantage over their historical counterpart, as one would have expected. This is because the errors with historical flows are themselves satisfactorily low, thus validating that our historical flows are pretty good estimates of time-dependent O-D flows. As auto-regressive formulation multiplies deviations by AR factors to calculate predicted flows, and as historical flows are close to true flows themselves, flows predicted based on estimates of auto-regressive factors sometimes seems to perform

worse than historical flows. Nevertheless, error statistics corresponding to predicted flows are also within acceptable range.



	1	2	3	4	5
Estimated O-D	0.1385	0.1622	0.1289	0.1674	0.1950
Predicted O-D	0.1500	0.1766	0.1334	0.1790	0.2185
Historical O-D	0.1436	0.1687	0.1317	0.1849	0.2464

Figure 3-67: Sept 29: One Step Prediction with True Historical O-D

Figure 3-75 compares the RMSN with steps of predictions for five study intervals selected. We observe that errors in predicting farther in the future (in form of RMSN of three- or four-step prediction) are not always increasing as intuition would have suggested. This is because predicted O-D flows are usually close to historical O-D flows, and since our historical flows are very good by themselves, predicted flows also tend be acceptable. To test this hypothesis, and to test the validity of our prediction model, we recalculated predicted flows for selected intervals, but with *wrong* set of historical flows. For the purpose of this study, historical O-D flows for morning periods are taken from evening intervals and vice-versa, ensuring that our historical data base is very bad to start with.

Results of this tests for September 29 are shown in Figures 3-76 to 3-79. We observe that predicted O-D flows always perform better than historical flows for one-, two-step predictions. However, as we predict farther, predicted flows tend perform worse than the historical flows. These results validate our prediction model that

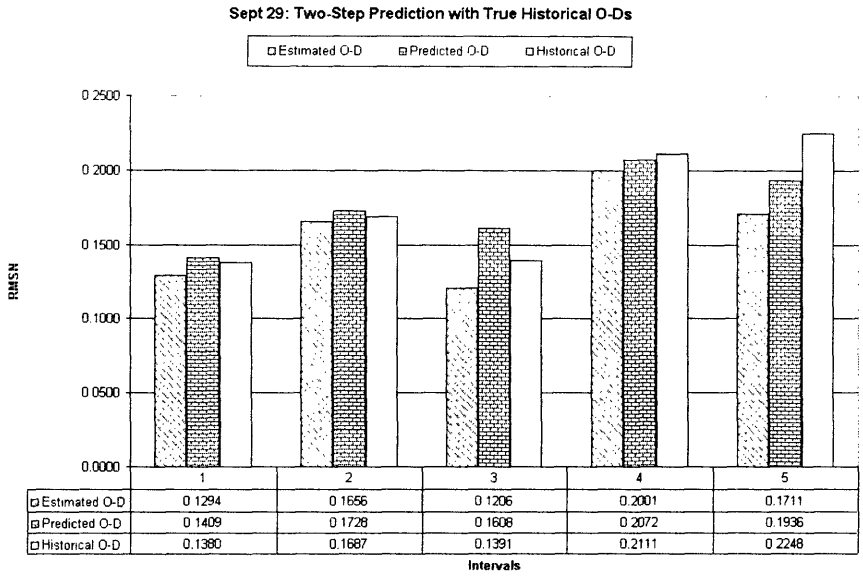


Figure 3-68: Sept 29: Two Step Prediction with True Historical O-D

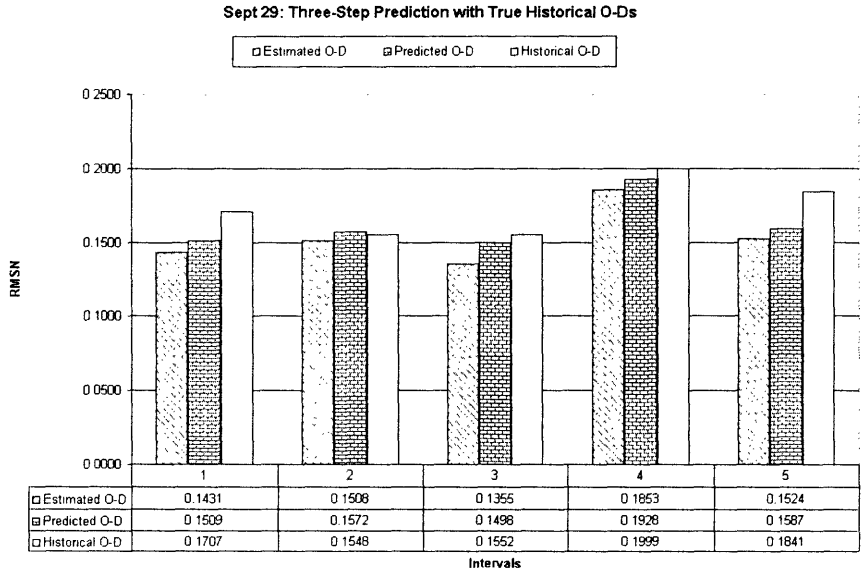
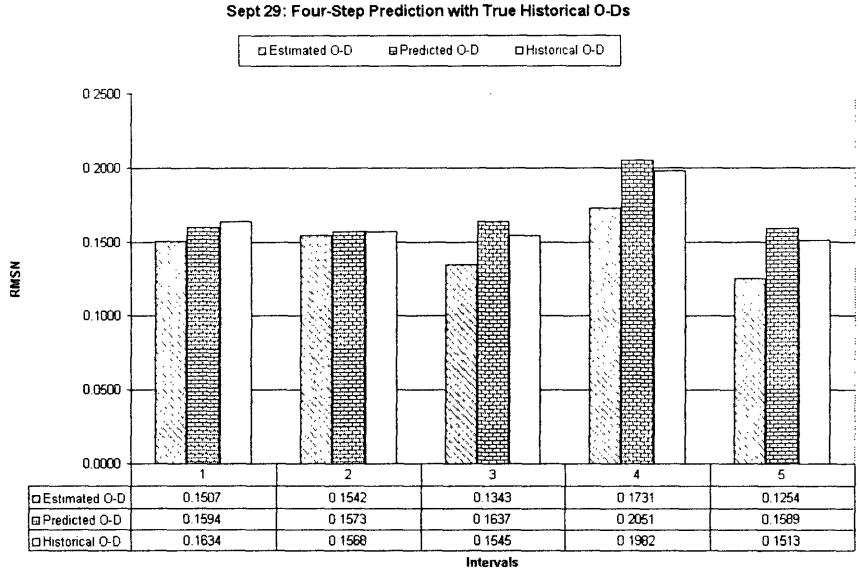
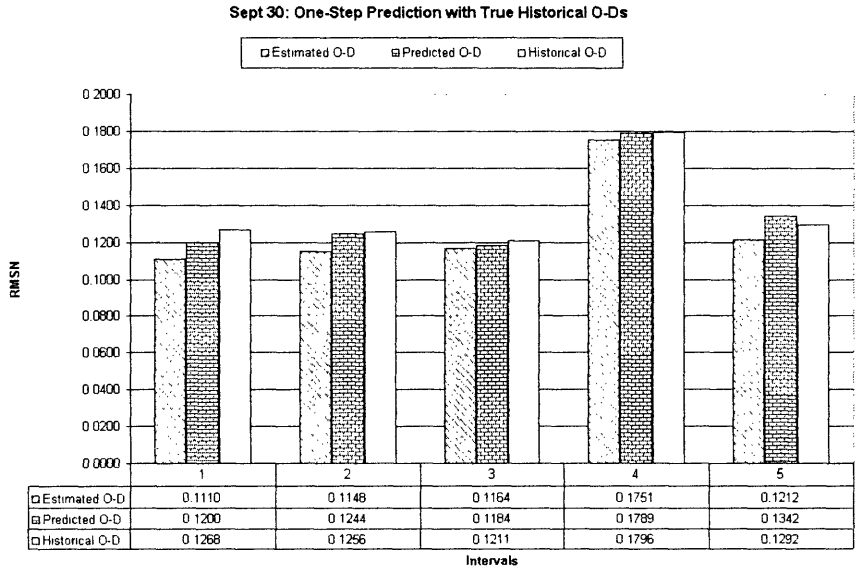


Figure 3-69: Sept 29: Three Step Prediction with True Historical O-D



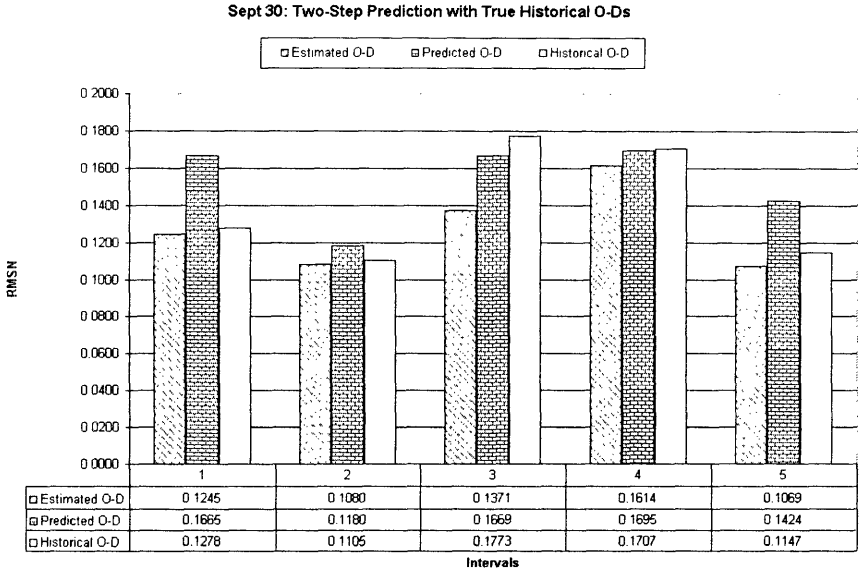
	1	2	3	4	5
Estimated O-D	0.1507	0.1542	0.1343	0.1731	0.1254
Predicted O-D	0.1594	0.1573	0.1637	0.2051	0.1589
Historical O-D	0.1634	0.1568	0.1545	0.1982	0.1513

Figure 3-70: Sept 29: Four Step Prediction with True Historical O-D



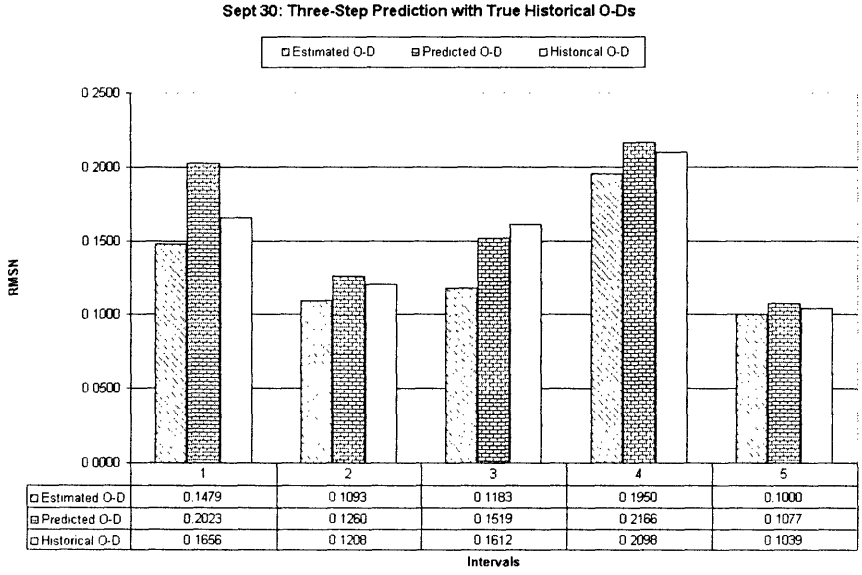
	1	2	3	4	5
Estimated O-D	0.1110	0.1148	0.1164	0.1751	0.1212
Predicted O-D	0.1200	0.1244	0.1184	0.1789	0.1342
Historical O-D	0.1268	0.1256	0.1211	0.1796	0.1292

Figure 3-71: Sept 30: One Step Prediction with True Historical O-D



	1	2	3	4	5
Estimated O-D	0.1245	0.1080	0.1371	0.1614	0.1069
Predicted O-D	0.1665	0.1180	0.1669	0.1695	0.1424
Historical O-D	0.1278	0.1105	0.1773	0.1707	0.1147

Figure 3-72: Sept 30: Two Step Prediction with True Historical O-D



	1	2	3	4	5
Estimated O-D	0.1479	0.1093	0.1183	0.1950	0.1000
Predicted O-D	0.2023	0.1260	0.1519	0.2166	0.1077
Historical O-D	0.1656	0.1208	0.1612	0.2098	0.1039

Figure 3-73: Sept 30: Three Step Prediction with True Historical O-D

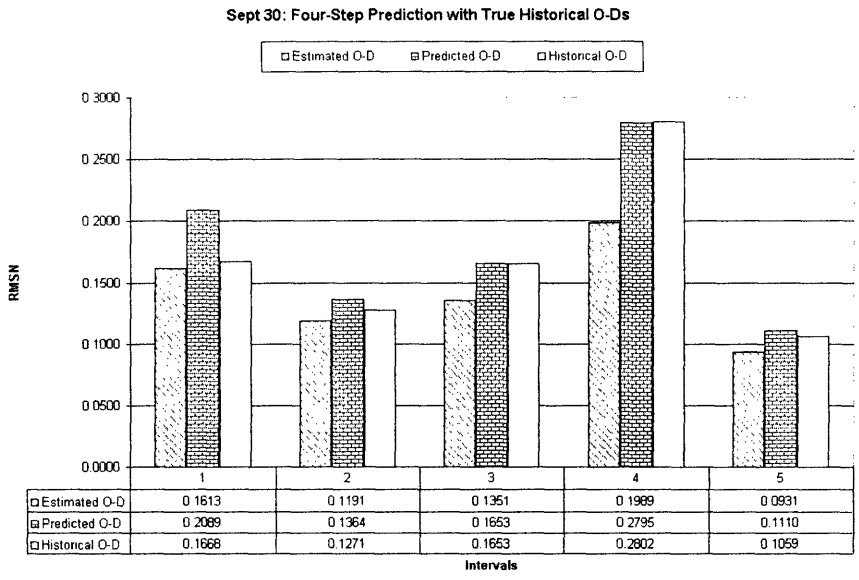


Figure 3-74: Sept 30: Four Step Prediction with True Historical O-D

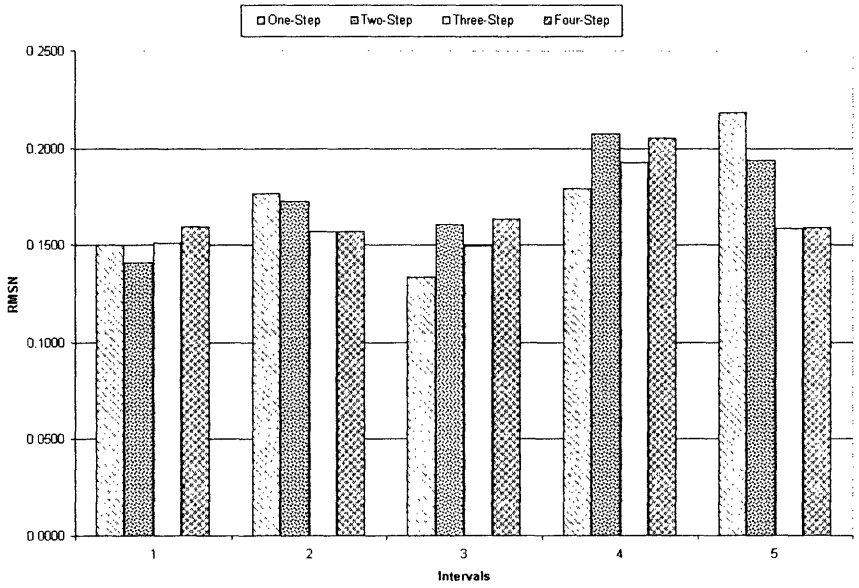
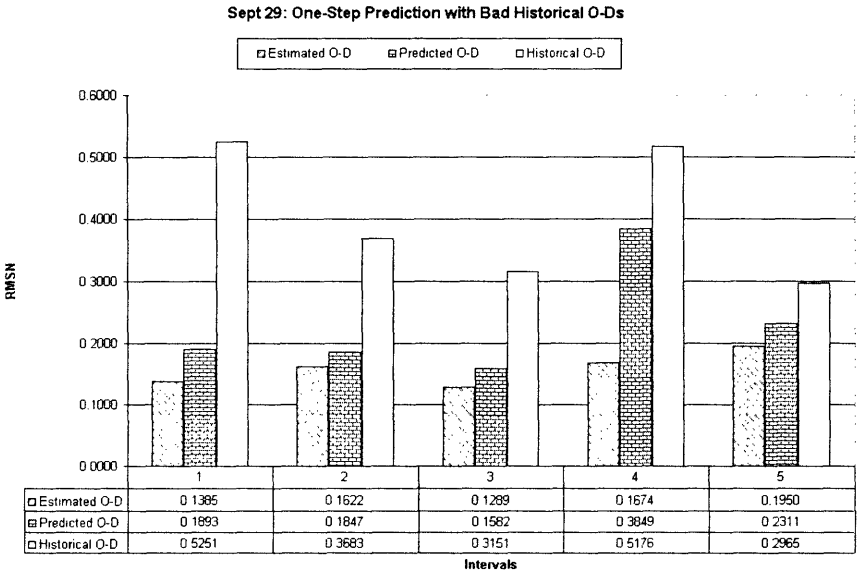


Figure 3-75: Sept 29: Prediction Errors with Steps using True Historical O-D

prediction offers significant improvement over historical flows, even when historical flows are very bad. They also confirm that the reason we could not isolate a trend in the performance of predicted flows over historical flows when we used the good historical flows is because data base of historical was quite close to observed flows. Note that the predicted flows perform worse than the historical flows in the fifth interval of testing. This is because this interval corresponded to 7–8 PM for which "bad" historical flows were taken from 7–8 AM. Since prediction based on deviations assumes that trend in deviations observed in previous intervals will continue in future, and as 7–8 AM is period of increasing morning traffic while 7–8 PM is period of decreasing evening traffic, we note that prediction module tend to overestimate the flows, thereby increasing the discrepancy and the RMSN.

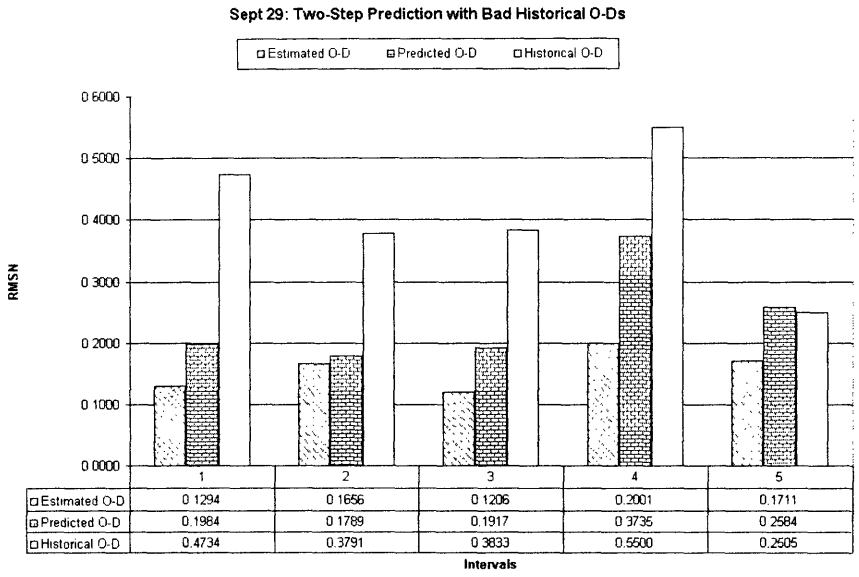


	1	2	3	4	5
Estimated O-D	0.1385	0.1622	0.1289	0.1674	0.1950
Predicted O-D	0.1893	0.1847	0.1582	0.3849	0.2311
Historical O-D	0.5251	0.3683	0.3151	0.5176	0.2965

Figure 3-76: Sept 29: One Step Prediction with Bad Historical O-D

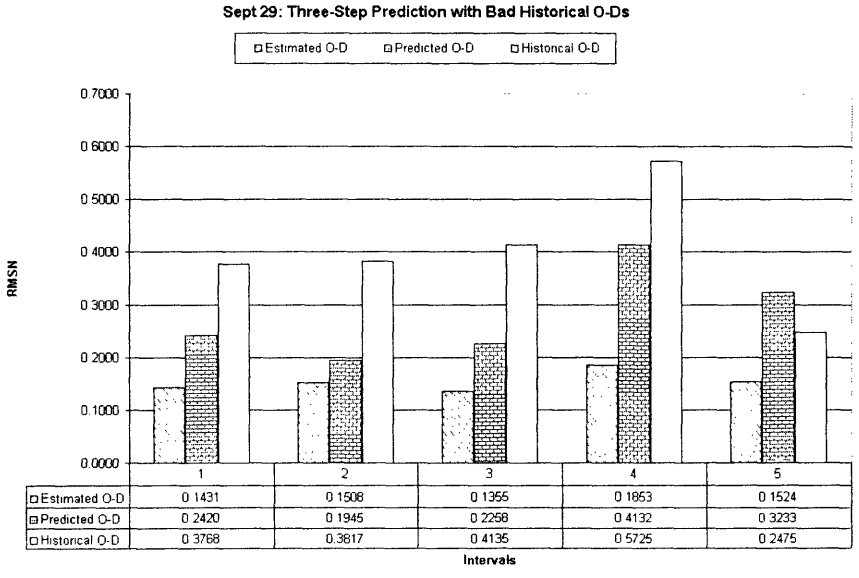
Finally, we note that prediction errors tend to increase as we predict farther in the future, as shown in Figure 3-80, as we would have expected. These findings are also consistent with Ashok [3].

Table 3.10 compares the estimated and predicted deviations (from historical flows i.e. $x_{h,r} - x^H_{h,r}$) in selected large O-D pairs. First four rows for each O-D pair present



	1	2	3	4	5
Estimated O-D	0.1294	0.1656	0.1206	0.2001	0.1711
Predicted O-D	0.1984	0.1789	0.1917	0.3735	0.2584
Historical O-D	0.4734	0.3791	0.3833	0.5500	0.2505

Figure 3-77: Sept 29: Two Step Prediction with Bad Historical O-D



	1	2	3	4	5
Estimated O-D	0.1431	0.1508	0.1355	0.1853	0.1524
Predicted O-D	0.2420	0.1945	0.2258	0.4132	0.3233
Historical O-D	0.3768	0.3817	0.4135	0.5725	0.2475

Figure 3-78: Sept 29: Three Step Prediction with Bad Historical O-D

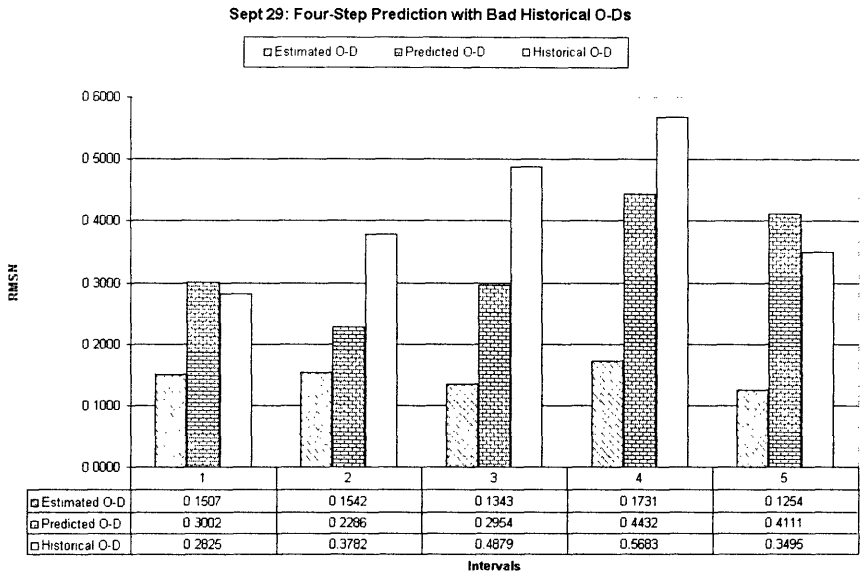


Figure 3-79: Sept 29: Four Step Prediction with Bad Historical O-D

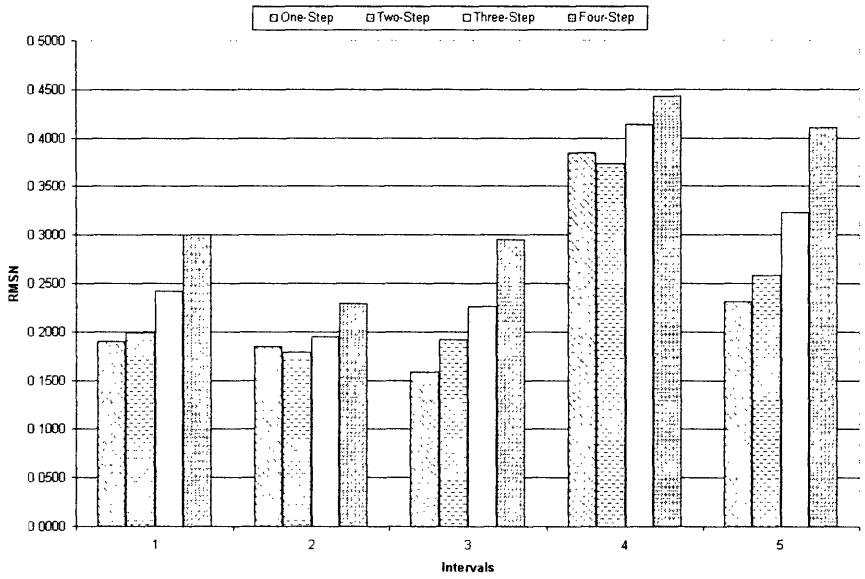


Figure 3-80: Sept 29: Prediction Errors with Steps using Bad Historical O-D

the estimated deviations which were used to predict deviations upto four-steps. The estimated deviations from prediction intervals benchmark the performance of prediction model.

We note that since significant fraction of predicted deviation comes from estimated deviation corresponding to immediately preceding interval, the difference in sign of estimated deviation of interval $h-1$ from other preceding intervals ($h-2$, $h-3$, ...), results in wrong sign of predicted deviation for interval h. This further results in next set of predicted deviations going in "wrong" (than the estimated deviations) direction due to recursive nature of auto-regressive process. Eventually, we have a scenario where historical flows (with deviations of zero) are closer to estimated flows. We note this behavior in O-D pair 691 and 692. This phenomenon, which is observed in some of other O-D pairs as well, explains why many a times predicted O-D flows perform worse than the historical flows.

In such cases, a modified auto-regressive process is adopted for prediction, where deviations on the right side of equation (2.38) are replaced by their moving averages (MA). With 3-point moving average process, modified auto-regressive formulation takes the following form:

$$\partial \mathbf{x}_h = \sum_{p=h-q'}^{h-1} \mathbf{f}_h^p \cdot \left(\frac{\partial \mathbf{x}_p + \partial \mathbf{x}_{p-1} + \partial \mathbf{x}_{p-2}}{3} \right) \tag{3.4}$$

where all terms have their usual meaning, and

$$\partial \mathbf{x}_h = \mathbf{x}_h - \mathbf{x}_h^H$$

Equation (3.4) can be interpreted that if the deviations during previous intervals have been positive (negative) *on average* then we expect the deviation for the next interval to also be positive (negative). Table 3.11 summarizes the estimated and predicted deviations in case of 3-point MA AR[10] process along with AR process represented by equation (2.40)[11].

[10]Moving Average Auto-Regressive

[11]Note that this AR process can also be construed as 1-point moving average AR process.

Steps	Estimated Deviation	Predicted Deviation
O-D pair: 692		
-	13.63	-
-	35.76	-
-	68.01	-
-	-48.22	-
1	10.57	-18.14
2	31.17	-10.69
3	1.20	-0.74
4	44.94	-9.43
O-D pair: 691		
-	-0.08	-
-	-1.55	-
-	-0.15	-
-	1.13	-
1	-2.35	0.53
2	-13.32	0.44
3	-3.03	0.42
4	-3.12	0.44
O-D pair: 674		
-	-0.22	-
-	-4.48	-
-	-2.16	-
-	-4.03	-
1	-3.54	-2.99
2	-5.10	-3.14
3	-2.77	-2.81
4	-2.12	-2.98
O-D pair: 931		
-	72.79	-
-	33.42	-
-	20.05	-
-	70.11	-
1	19.84	58.34
2	50.73	51.57
3	-9.12	45.27
4	0.77	47.31

Table 3.10: Estimated vs Predicted Deviations

Steps	Est. Dev.	1-point MA AR		3-point MA AR	
		Pred. Dev.	OD Diff.	Pred. Dev.	OD Diff.
O-D pair: 692					
1	10.57	-18.14	28.71	18.41	7.84
2	31.17	-10.69	41.86	13.03	18.14
3	1.20	-0.74	1.94	1.58	0.38
4	44.94	-9.43	54.37	11.14	33.80
O-D pair: 691					
1	-2.35	0.53	2.88	-0.22	2.13
2	-13.32	0.44	13.76	0.20	13.52
3	-3.03	0.42	3.45	0.32	3.35
4	-3.12	0.44	3.56	0.09	3.21
O-D pair: 674					
1	-3.54	-2.99	0.55	-3.30	0.24
2	-5.10	-3.14	1.96	-3.96	1.14
3	-2.77	-2.81	0.04	-2.83	0.06
4	-2.12	-2.98	0.86	-2.86	0.74
O-D pair: 931					
1	19.84	58.34	38.50	39.27	19.43
2	50.73	51.57	0.84	43.59	7.14
3	-9.12	45.27	54.39	46.63	55.75
4	0.77	47.31	46.54	40.23	39.46

Table 3.11: 1-point vs 3-point Moving Average AR Process

Columns titled as 'OD Diff.' show the absolute difference between estimated and predicted OD flows (which is same as the absolute difference between estimated and predicted deviations). We observe that predicted deviations using 3-point MA AR process are more closer, in general, to estimated deviations as the difference in OD flows is smaller in case of 3-point MA AR process more often than the other. This justifies our approach to use 3-point MA AR process as our prediction model. Table 3.12 documents the corresponding auto-regressive factors. Since MA AR process is based on average performance of previous deviations, one degree is sufficient[12] to give acceptable model.

OD Group	Time Group	AR Degree	$f_{r,h}^{r,h-1}$
> 200	4:00 AM–8:00 AM	1	1.0000
	8:00 AM–7:00 PM	1	0.8985
	7:00 PM–0:00 AM	1	0.8977
50–200	4:00 AM–8:00 AM	1	1.0000
	8:00 AM–7:00 PM	1	0.9157
	7:00 PM–0:00 AM	1	0.8091
20–50	4:00 AM–8:00 AM	1	0.8498
	8:00 AM–7:00 PM	1	0.9147
	7:00 PM–0:00 AM	1	0.7603
< 20	4:00 AM–8:00 AM	1	0.6904
	8:00 AM–7:00 PM	1	0.9105
	7:00 PM–0:00 AM	1	0.7740

Table 3.12: Calibrated MA AR Factors for Weekdays

3.5 Summary and Conclusion

In this chapter we calibrated and validated complete DynaMIT system for South Park Los Angeles network. We started by describing the network and data availability. Then we moved on to analyze data. Thereafter, we started calibration process for weekdays using 5 selected days of data. We proceeded to calibrating speed-density relationships by categorizing segments into various groups and estimating parameters

[12]MA AR factors, for example, are 0.89, 0, 0, 0.03 for degree four, for high flow AM Peak case, with similar prediction performance.

for each group. We also calculated capacities of all segments and intersections before proceeding to demand side calibration. On demand side, we begun by identifying optimum set of path choice set generation parameters and route choice parameters. Then we presented a case study demonstrating the feasibility of approach for verifying observability to large scale transportation network. We started by selecting three completely independent and randomly generated initial flows and conclusively proved observability of O-D estimation process by showing that all three matrices converged to exactly the same final state after couple of iterations. Thereafter we proceeded to estimate set of historical O-D flows and estimated error covariance matrix and experienced link travel times. We then presented a set of autoregressive factors which are used by DynaMIT to predict O-D flows. We finally presented validation results confirming that calibration methodology can be applied to large network to estimate robust set of parameters. Consistent estimation and calibration results despite of no information on a-priori O-D widens the scope and applications of this methodology for many networks and cases where prior information is not available or unreliable.

Chapter 4

Conclusion

In this chapter, we begin by summarizing contributions of this research to O-D estimation and model calibration in DTA context, and propose some future directions that can be pursued as extension to this work.

4.1 Research Contribution

This report outlines the detailed and comprehensive off-line calibration of a complex DTA model for a large scale network using multiple days of data and in absence of any prior information regarding Origin-Destination flow magnitudes or patterns. Contributions of this work towards the literature in this area can be classified into two broad categories.

Firstly, we develop and explore theoretical implications and requirements of the observability as it applies to O-D estimation process. Comprehensive theoretical treatment of observability was presented and conditions for O-D estimation module to be observable in a particular application were discussed. Experimental setup to test observability was developed and used to successfully demonstrate the application of test methodology on a practical size network. We conclude that if all O-D pairs are counted by at least one sensor on the network along at least one of its path, it is possible given sufficient data, to prove observability. This also makes O-D estimation process independent of starting historical or planning O-D information,

thereby widening the field of potential applications in future on different networks. This is of crucial significance as often transportation planning studies which yields static O-D flows estimates are very expensive and undertaken once in two or three decades. Thus, planning data is usually not very reliable in fast changing demographies. Observability completely eliminates our dependence on any O-D information a-priori.

Second specific contribution of this report is comprehensive off-line calibration using multiple days of data for both demand and supply parameters. Consistent estimation and prediction results successfully demonstrate the robustness of the calibration methodology to a large network area. Also, this is one of the first attempts to calibrate DynaMIT for round the clock run in real-time. Detailed treatment as demonstrated in supply calibration of each individual freeway sensor, and in estimating different auto-regressive factors for different set of O-D pairs for different intervals within a day also mark some significant contributions of this study.

In a small yet significant step, this thesis also demonstrates superiority of analytically calculated assignment matrix over a simulated assignment matrix. Not only is this theoretically advantageous, it is also computationally beneficial in terms of reducing run-time to simulate assignment matrix.

4.2 Future Research

The focus of this thesis was limited to demonstrating observability and subsequently calibrating South Park Los Angeles area transportation network for real time application of DynaMIT. There are several issues that can form natural extension to this work as discussed below.

- One of the obvious extension, as also discussed in Section 1.3 of this report, is to consider implications of sensor coverage on the network, and its impact on observability. In the present case study, a rich data set including multiple days of counts was available, but this is not the case in many applications. In such cases it may not be affordable to use some intervals of data to estimate

unique O-D matrix. Study on various degrees of sensor coverage with its effect on number of iterations (or intervals of data) it takes to achieve predefined level of observability could enlighten their interdependence more clearly.

- As has been repeatedly emphasized in this report, observability cannot be achieved unless every potential O-D pair is counted on the network. Hence critical planning decision for transportation authorities would be to locate surveillance infrastructures so as to trade off budget and efficiency of estimation and prediction capabilities. Hence sensor location problem raised earlier can be discussed in DTA context.

- O-D calibration in this report follows sequential approach due to computational time constraints. However, a more simultaneous approach using augmented state vector as discussed in [3] using Kalman Filtering techniques is likely to yield more efficient parameter estimates.

- While calibration methodology proposed here deals with major incidents, effect of other critical factors as weather, scheduled event, unexpected demand fluctuations (such as in emergency), planned construction of lane closings, etc. are not accounted for. An extension to this study could include analysis of DynaMIT's estimation and prediction capabilities under such scenarios.

4.3 Summary and Conclusion

This report presents a theoretical discussion and empirical framework for understanding the implication of and testing the O-D estimation model's compliance with observability property. We prove that effect of starting O-D diminishes as the number of intervals tend to increase. We then proceed to calibrate state-of-the-art decision support tool DynaMIT using multiple days of data for on line deployment for round the clock operations. Validation studies show encouraging results regarding estimation and prediction capabilities of the system. We conclude by pointing out specific contribution of this report and suggest direction for future work in this area.

Appendix A

The DynaMIT System

DynaMIT (**Dy**namic **N**etwork **A**ssignment for the **M**anagement of **I**nformation to **T**ravelers) is a state-of-the-art real-time simulation software for the traffic estimation and prediction. DynaMIT acts as a decision support tool at Traffic Management Centers (TMCs) by providing real-time traveler information and route guidance. Sponsored by the Federal Highway Administration (FHWA), DynaMIT is the result of intense research at the Intelligent Transportation Systems Program at the Massachusetts Institute of Technology. We briefly highlight here the functionalities and the capabilities of DynaMIT, before we proceed to identify various models and parameters used in DynaMIT. DynaMIT is used for observability and calibration case studies presented in Chapter 3.

A.1 Overall Framework

DynaMIT is a mesoscopic traffic simulation tool which can be used in real-time as well as in planning applications. DynaMIT-R, the real-time version of DynaMIT, and DynaMIT-P, the planning version of DynaMIT, can both model variety of traffic and network characteristics and provide real-time forecasts of state of the network in near future. The key to DynaMIT's functionality is its detailed network representation coupled with comprehensive treatment of traffic dynamics in form of various mathematical models for traveler behavior and network response. This section very briefly

introduces the overall framework of DynaMIT and reader is referred to [21] and [22] for more extensive treatment of the subject.

DynaMIT is composed of several complex models and algorithms and supports following two important functionalities (Figure A-1):

- Estimation of current state of the network, using effective integration of real-time input with historical database and calibrated parameters, and
- Generation of prediction based guidance for future time horizon

Estimation and prediction continues in real-time in rolling horizon mode within DynaMIT. As DynaMIT estimates current state of the network based on off-line information such as historical O-D flows, travel times and other model parameters, combined with real-time information in form of sensor counts and incident detections, it continuously updates its historical database and knowledge of network conditions. DynaMIT ensures that prediction and guidance generated by it are consistent and unbiased to maintain credebility and compliance. While *unbiasedness* guarantees that information provided to the travelers is best and latest available knowledge of current and anticipated network conditions, *consistency* ensures that drivers' experience of network state matches with anticipated condition by DynaMIT.

A.1.1 State Estimation

State estimation module of DynaMIT is responsible for estimation of up-to-date state of the network in terms of number of vehicles, O-D flows, congestions, queue lengths, incident, flows, speeds and densities. Input data for state estimation process comes from both real-time surveillance systems and archived historical database. DynaMIT reads detailed network representation and typical conditions for network as calibrated from historical database and simulates the traffic flow on the network with latest available demand and supply characteristics. It then compares the simulated state with real time information from sensor counts and modifies its best estimate of the state of the network. This iterative process (Figure A-2) between demand and supply

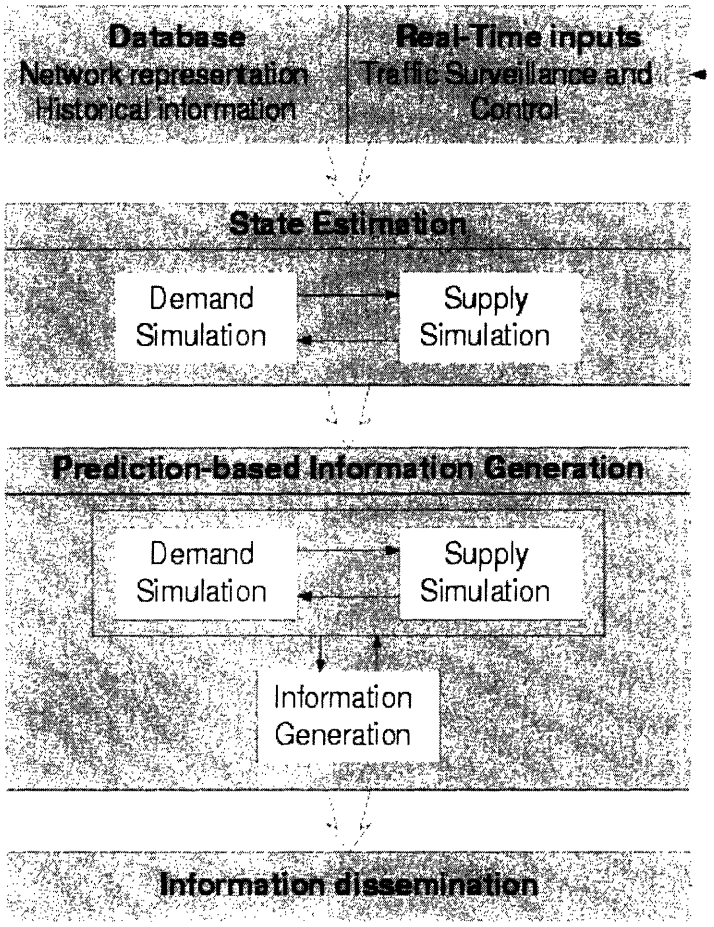


Figure A-1: The DynaMIT Framework

simulator components of the DynaMIT continues until desired level of convergence is achieved. Degree of discretization of time steps in simulation can be controlled by specifying appropriate input to advance and update intervals.

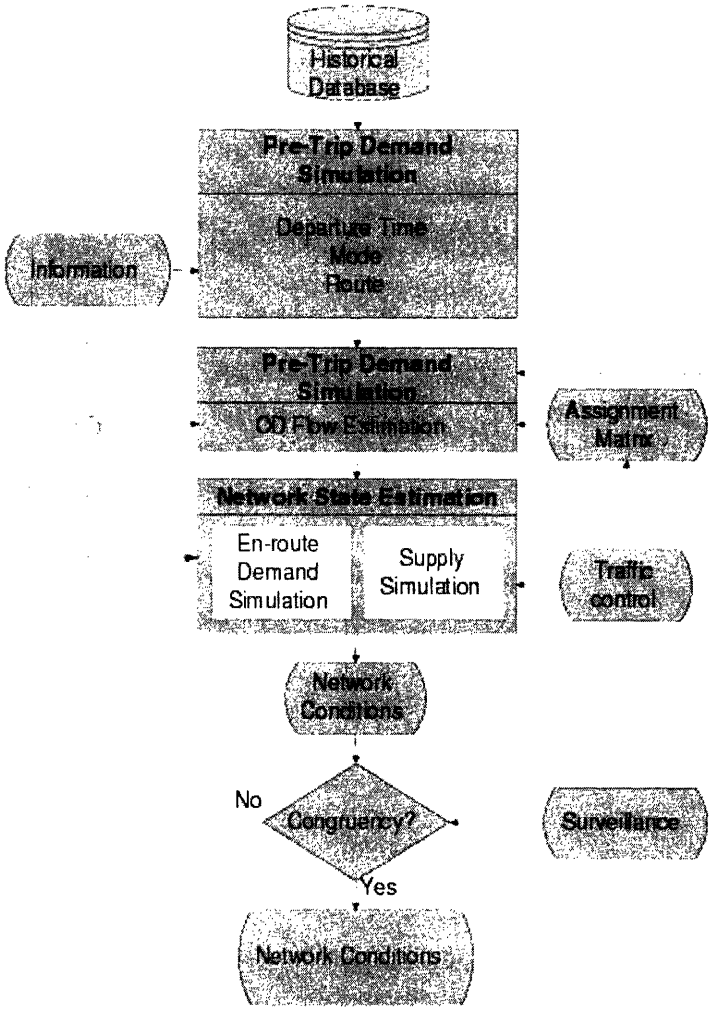


Figure A-2: State Estimation Module in DynaMIT

A.1.2 Prediction and Guidance Generation

The prediction and guidance generation module of DynaMIT undertakes the generation of unbiased and consistent guidance to drivers based on predictive assessment of the future state of the network. Current state of the network estimated in state

estimation module becomes the starting point of the prediction run. Best available O-D information is then forecasted using calibrated set of auto-regressive factors for future and adjusted to take into account any change in trip characteristics (viz. departure time, route, mode, etc.) made by users based on pre-trip guidance. This updated flow is then simulated and state of the network is estimated by supply simulator. A consistency check is then performed, and iterative and interactive process (Figure A-3) between demand simulator, supply simulator and guidance generation algorithm continues until consistency is achieved.

A.2 Parameters for Calibration

In this section, we will enumerate various model parameters to be calibrated in DynaMIT.

A.2.1 Demand Simulator Parameters

Demand simulator in DynaMIT is a key component which explicitly simulates pre-trip departure time, mode and route choice decisions for each disaggregated drivers based on discrete choice probabilities. These drivers are then aggregated based on departure time and O-D to produce the flows used by demand estimation module. However, the O-D flows on a particular day are likely to be different from historical flows for variety of reasons, including but not exhausting, 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. The O-D Estimation module uses updated historical O-D flows, real-time measurements of actual link flows on the network, and estimates of assignment fractions to estimate the O-D flows for the current estimation interval.

Route Choice Parameters

DynaMIT relies on multinomial path size discrete choice formulation to model the route choice by an individual driver. This process comprises of following three steps,

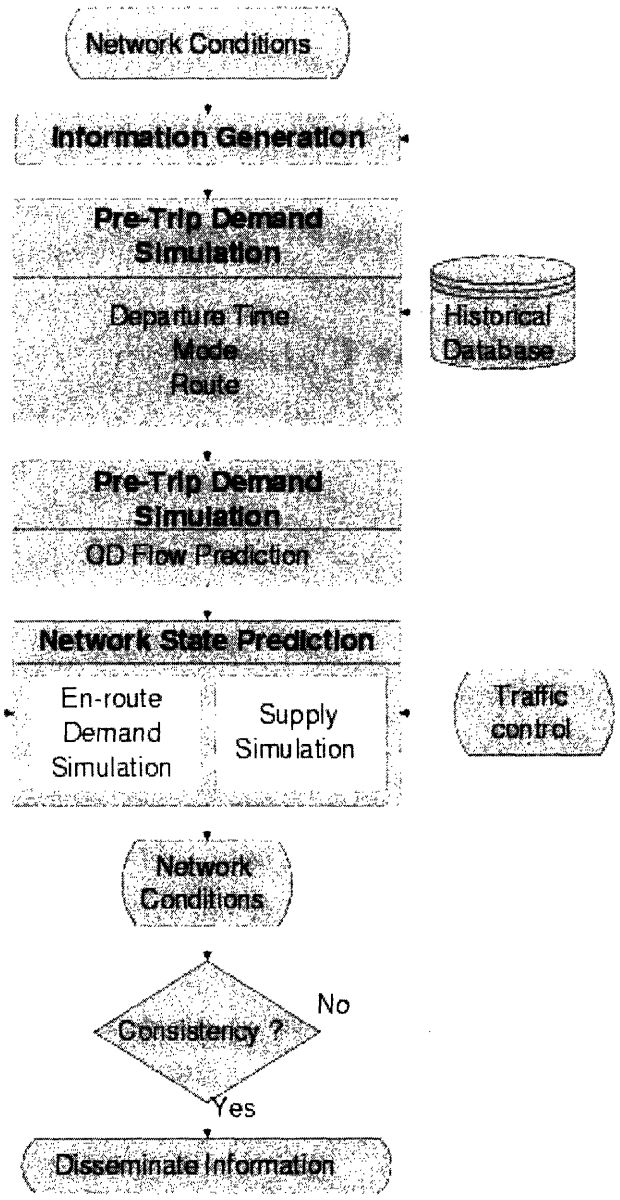


Figure A-3: Prediction and Guidance Module in DynaMIT

first of which is called once for a network while other two are called as many times as route choice decisions—habitual, pre-trip or en-route—are made.

1. Path choice set generation for each O-D pair in demand file,
2. Computation of utilities for all alternatives available to a driver for an O-D pair, and
3. Selection of one of the alternative based on simulated probabilities

As selection of a good set of paths is critical to all further calibration process, path choice set generation step in DynaMIT ensures that all reasonable and practical paths are captured. Path choice set generation is a three step process in itself which includes:

1. **Shortest path computation:** Generation of shortest paths from each link to each destination,
2. **Link elimination:** Generation of shortest paths from each link to each destination while every other link is removed from the network one at a time, and
3. **Random perturbation:** Generation of shortest paths from each link to each destination based on randomly perturbed impedances of links in the network.

This process generates sufficient alternative paths for the drivers and ensures that there is at least one path available for each O-D pair even in case of incident on a link on the network. DynaMIT further screens unreasonably long and duplicated paths.

Once a reasonable set of paths are computed, their utilities are evaluated for each user during path choice decision making. Currently DynaMIT supports parameters for path length, path travel time, number of intersections, trip purpose and number of left turns for linear in parameter utility specification model. This list can, however, easily be augmented for any specific characteristics that may be observed in the field. Subsequently, probability of selection of a particular path is computed using Path-Size Logit model, as specified below:

$$P_n(i) = \frac{e^{V_i + \ln PS_i}}{\sum_{j \in C_n} e^{V_j + \ln PS_j}} \tag{A.1}$$

where $P_n(i)$ is the probability of user n choosing path i, V_i is the utility of path i, PS_i is the size of path i, and C_n denotes the choice set for individual n. The size of a path is defined as proposed by Ramming [25],

$$PS_{in} = \sum_{a \in \Gamma_i} \left(\frac{l_a}{L_i} \right) \frac{1}{\sum_{j \in C_n} \frac{L_i^\gamma}{L_j^\gamma} \delta_{aj}} \tag{A.2}$$

where l_a is the length of link a, L_i is the length of path i and δ_{aj} is binary coefficient which is 1 if link a is a part of path j and 0 otherwise. The inner summation is computed over all paths in choice set C_n, while the outer summation is computed over all links a in path Γ_i. The parameter γ is an exponent in the model.

To summarize, following parameters need to be calibrated for route choice model:

- Parameters in path choice set generation process, viz. the number of random perturbations, shortest-path factor[1] (≥ 1.0) and freeway bias[2] ($\geq 0, \leq 1$)
- Parameters in utility model, viz. Value-Of-Time coefficients, freeway-on-off penalty[3] and freeway bias.
- Path-Size exponent γ

O-D Estimation Parameters

Current version of DynaMIT implements sequential GLS based O-D estimation module as presented in Section 2.1.3. External inputs to module are calibrated set of historical O-D matrices, time-dependent link travel times, and auto-regressive factors; while important internally generated inputs are time-dependent assignment matrices. Following parameters must be calibrated for O-D estimation process:

[1]Paths longer than shortest-path factor times the shortest path length for an O-D pair are not included in the choice set.

[2]Freeway travel time component of a path is discounted by this factor to model lower "discomfort" associated with freeway travel time against same travel time on streets for most drivers.

[3]This penalizes paths which include getting on the freeway and then getting off again, or vice versa, for shorter distances.

- Historical database of O-D flows, $\mathbf{x}_h^H$
- Error-covariance matrices associated with direct and indirect measurement errors, $\mathbf{V}_h$ and $\mathbf{W}_h$
- Matrices $\mathbf{f}_h^p$ of autoregressive factors

A.2.2 Supply Simulator Parameters

Supply Simulator, also called the Network State Estimator, in DynaMIT simulates movements of individual drivers on the network based on pre-specified speed-density relationships and network characteristics such as capacities and control strategies. Outputs of supply simulation are volumes, speeds and densities on all segments of the network per time interval. This interval can be specified beforehand and often is smaller than the estimation interval. Supply simulator also models en-route path choice decisions in response to guidance provided, and correctly identifies buildup of queues and spill backs in case of congested networks. Following supply parameters need to be calibrated in DynaMIT,

- Six-element vectors of speed-density parameters for each segment (v_{max}, k_{min}, k_{jam}, α, β, v_{min})
- Segment and intersection capacities

A.3 Summary and Conclusion

In this chapter, we described the functionalities and capabilities of DynaMIT as decision support tool in Traffic Management Centers. We emphasized the ability of DynaMIT to estimate the condition of network in real-time and provide forecasted state in near future with consistency and unbiasedness. We then briefly described the structure of various mathematical models built in the DynaMIT and concluded the chapter by identifying parameters which need to be calibrated, both for demand side and supply side.

Appendix B

Calibration and Validation Results for Weekdays

Calibration and validation results are tabulated below.

Interval	Scale	RMSE	RMSN
3:15	1.0444	6.48	0.1510
3:30	0.9871	6.43	0.1533
3:45	1.0257	6.34	0.1400
4:00	0.9909	5.02	0.0952
4:15	1.0310	7.60	0.1361
4:30	1.0171	6.14	0.0852
4:45	0.9977	8.83	0.0918
5:00	1.0040	9.61	0.0827
5:15	0.9995	12.46	0.0965
5:30	1.0042	13.79	0.0733
5:45	0.9955	17.62	0.0687
6:00	1.0298	20.79	0.0695
6:15	1.0210	24.57	0.0779
6:30	1.0822	47.99	0.1351
6:45	1.0593	41.08	0.1079
7:00	0.9846	55.51	0.1430
7:15	0.9438	59.99	0.1558
7:30	1.0329	53.36	0.1337
7:45	1.0599	64.81	0.1661
8:00	1.0178	66.49	0.1696
8:15	1.0387	62.06	0.1601
8:30	1.0086	57.05	0.1562
8:45	1.0500	50.71	0.1495
9:00	1.0476	46.01	0.1285
9:15	1.0265	41.14	0.1146
9:30	1.0275	44.54	0.1231
9:45	1.0148	35.60	0.0954
10:00	1.0183	34.95	0.0943
10:15	1.0073	32.69	0.0897
10:30	0.9979	37.95	0.1102
10:45	1.0186	33.73	0.0938
11:00	1.0250	35.82	0.0955
11:15	1.0337	36.36	0.1017
11:30	1.0207	35.60	0.0945
11:45	1.0244	34.60	0.0895
12:00	1.0247	33.03	0.0866
12:15	1.0219	41.94	0.1097
12:30	1.0178	42.41	0.1086
12:45	1.0199	37.97	0.1032
13:00	1.0240	44.11	0.1157
13:15	1.0192	41.11	0.1114
13:30	1.0163	41.27	0.1073

Interval	Scale	RMSE	RMSN
13:45	1.0217	41.37	0.1061
14:00	1.0232	43.31	0.1099
14:15	1.0131	44.22	0.1158
14:30	1.0299	44.17	0.1210
14:45	1.0408	47.90	0.1264
15:00	1.0319	45.98	0.1195
15:15	1.0303	45.77	0.1178
15:30	1.0096	43.32	0.1148
15:45	1.0302	43.92	0.1139
16:00	1.0227	47.64	0.1224
16:15	1.0218	46.46	0.1179
16:30	1.0066	46.16	0.1191
16:45	1.0283	48.38	0.1308
17:00	1.0394	53.11	0.1431
17:15	1.0255	52.45	0.1429
17:30	1.0290	56.04	0.1442
17:45	1.0272	57.66	0.1534
18:00	1.0255	58.50	0.1630
18:15	1.0407	53.50	0.1538
18:30	1.0628	55.77	0.1520
18:45	1.0478	55.00	0.1523
19:00	1.0440	52.21	0.1466
19:15	1.0490	58.63	0.1604
19:30	1.0474	52.32	0.1467
19:45	1.0404	51.69	0.1467
20:00	1.0398	49.54	0.1468
20:15	1.0408	38.84	0.1202
20:30	1.0166	38.71	0.1261
20:45	1.0331	29.90	0.0986
21:00	1.0116	28.56	0.0997
21:15	1.0259	25.46	0.0894
21:30	1.0252	24.58	0.0831
21:45	1.0326	24.85	0.0873
22:00	1.0307	31.05	0.1102
22:15	1.0123	25.63	0.0962
22:30	1.0151	29.08	0.1111
22:45	1.0114	24.90	0.1065
23:00	1.0282	18.17	0.0842
23:15	1.0124	14.19	0.0731
23:30	1.0175	15.09	0.0807
23:45	1.0545	34.52	0.1839
00:00	1.1015	19.51	0.1325

Table B.1: Calibration Results for Day 1

Interval	Scale	RMSE	RMSN
3:15	1.0460	6.86	0.1416
3:30	1.0309	6.67	0.1426
3:45	1.0281	7.01	0.1401
4:00	1.0313	7.98	0.1480
4:15	1.0284	8.61	0.1366
4:30	1.0406	9.36	0.1274
4:45	1.0081	10.81	0.1124
5:00	1.0132	11.11	0.1054
5:15	1.0061	14.51	0.1163
5:30	1.0170	23.03	0.1256
5:45	1.0241	32.11	0.1266
6:00	1.0214	34.68	0.1200
6:15	1.0225	39.90	0.1335
6:30	1.0254	50.56	0.1474
6:45	1.0121	53.05	0.1423
7:00	1.0471	66.54	0.1743
7:15	1.0047	58.99	0.1512
7:30	1.0141	66.12	0.1700
7:45	1.0257	73.25	0.1919
8:00	1.0206	73.56	0.1878
8:15	1.0382	69.23	0.1860
8:30	1.0307	66.89	0.1815
8:45	1.0267	67.76	0.1865
9:00	1.0263	66.02	0.1811
9:15	1.0138	57.48	0.1593
9:30	1.0167	58.93	0.1623
9:45	1.0111	57.45	0.1575
10:00	1.0063	52.35	0.1439
10:15	1.0085	51.50	0.1400
10:30	1.0075	53.17	0.1420
10:45	1.0161	54.15	0.1464
11:00	1.0059	52.28	0.1431
11:15	0.9920	54.54	0.1501
11:30	1.0030	54.38	0.1489
11:45	1.0116	56.90	0.1534
12:00	1.0149	58.82	0.1566
12:15	1.0025	57.30	0.1509
12:30	1.0102	64.03	0.1690
12:45	1.0351	62.78	0.1667
13:00	1.0142	62.24	0.1644
13:15	1.0334	58.36	0.1614
13:30	0.4314	71.98	1.5382
13:45	1.1380	94.14	0.2497
14:00	1.0429	67.94	0.1783
14:15	1.0500	63.78	0.1696
14:30	1.0595	64.14	0.1709
14:45	1.0475	55.08	0.1507
15:00	1.0567	60.36	0.1734
15:15	1.0504	58.87	0.1686
15:30	1.0698	70.91	0.2102
15:45	1.0545	61.64	0.1900
16:00	1.0684	68.78	0.2024
16:15	1.0609	66.11	0.1802
16:30	1.0589	76.48	0.2144
16:45	1.0749	72.49	0.2048
17:00	1.0673	77.06	0.2134
17:15	1.0686	72.79	0.2089
17:30	1.0733	80.40	0.2227
17:45	1.0756	76.78	0.2165
18:00	1.0645	74.72	0.2034
18:15	1.0401	63.58	0.1771
18:30	1.0349	65.77	0.1816
18:45	1.0290	60.17	0.1752
19:00	1.0222	55.04	0.1617
19:15	1.0237	53.79	0.1541
19:30	1.0095	51.73	0.1537
19:45	1.0217	51.49	0.1516
20:00	1.0079	49.85	0.1415
20:15	1.0023	42.82	0.1263
20:30	1.0045	35.93	0.1129
20:45	0.9933	38.20	0.1238
21:00	0.9965	42.78	0.1448
21:15	0.9720	50.45	0.1835
21:30	0.9811	50.94	0.1772
21:45	0.9911	42.90	0.1568
22:00	0.9796	46.37	0.1670
22:15	0.9884	47.09	0.1729
22:30	0.9952	46.99	0.1623
22:45	0.9648	45.22	0.1737
23:00	0.9764	44.36	0.1821
23:15	0.9820	43.34	0.1967
23:30	0.9820	45.34	0.2152
23:45	0.9714	36.69	0.1956
00:00	1.0319	37.06	0.2196

Table B.2: Calibration Results for Day 2

Interval	Scale	RMSE	RMSN
3:15	1.0188	4.74	0.1299
3:30	1.1140	7.58	0.1861
3:45	1.0581	6.38	0.1364
4:00	0.9968	4.77	0.0997
4:15	1.0281	6.75	0.1343
4:30	0.9983	5.24	0.0791
4:45	1.0193	6.08	0.0659
5:00	1.0182	9.03	0.0833
5:15	1.0220	9.97	0.0760
5:30	1.0177	13.46	0.0709
5:45	1.0203	15.76	0.0595
6:00	1.0144	15.50	0.0528
6:15	1.0348	24.32	0.0778
6:30	1.0129	21.09	0.0631
6:45	1.0305	31.34	0.0856
7:00	1.0490	53.55	0.1391
7:15	1.0253	47.88	0.1221
7:30	1.0548	52.64	0.1374
7:45	1.0593	58.83	0.1637
8:00	1.0587	67.17	0.1866
8:15	1.0284	56.70	0.1605
8:30	1.0525	66.24	0.1693
8:45	1.0362	57.36	0.1563
9:00	1.0468	61.67	0.1648
9:15	1.0091	52.26	0.1428
9:30	0.9860	46.23	0.1390
9:45	1.0236	41.89	0.1214
10:00	1.0370	44.72	0.1198
10:15	1.0205	38.86	0.1103
10:30	1.0310	37.25	0.1029
10:45	1.0423	43.08	0.1180
11:00	1.0203	34.83	0.0954
11:15	1.0353	39.05	0.1084
11:30	1.0241	36.26	0.0984
11:45	1.0139	37.74	0.1035
12:00	1.0329	36.72	0.0994
12:15	1.0179	35.66	0.0987
12:30	1.0242	35.51	0.0988
12:45	1.0476	41.47	0.1125
13:00	1.0226	33.19	0.0917
13:15	1.0471	41.39	0.1097
13:30	1.0665	52.08	0.1379
13:45	0.9841	57.84	0.1544
14:00	1.0037	43.15	0.1114
14:15	1.0387	41.13	0.1058
14:30	1.0452	45.59	0.1141
14:45	1.0296	39.32	0.1014
15:00	1.0367	45.93	0.1215
15:15	1.0376	41.86	0.1111
15:30	1.0375	44.43	0.1168
15:45	1.0361	49.59	0.1275
16:00	1.0489	55.39	0.1440
16:15	1.0387	49.36	0.1281
16:30	1.0356	54.19	0.1428
16:45	1.0508	54.74	0.1427
17:00	1.0496	65.25	0.1640
17:15	1.0454	62.23	0.1582
17:30	1.0428	64.87	0.1599
17:45	1.0316	58.56	0.1527
18:00	1.0347	57.75	0.1474
18:15	1.0313	55.87	0.1480
18:30	1.0281	54.82	0.1431
18:45	1.0495	56.21	0.1484
19:00	1.0323	48.12	0.1336
19:15	1.0331	42.16	0.1169
19:30	1.0306	43.83	0.1242
19:45	1.0132	37.06	0.1089
20:00	1.0033	38.17	0.1217
20:15	1.0338	39.39	0.1283
20:30	1.0346	30.00	0.1016
20:45	0.9944	26.42	0.0983
21:00	0.9891	24.84	0.0971
21:15	1.0074	23.03	0.0910
21:30	0.9918	26.02	0.0961
21:45	1.0138	24.51	0.0929
22:00	0.9967	22.75	0.0915
22:15	0.9998	18.82	0.0808
22:30	1.0057	21.62	0.0895
22:45	1.0188	20.86	0.0947
23:00	1.0173	18.29	0.0935
23:15	1.0729	34.73	0.1625
23:30	0.9955	17.76	0.1029
23:45	0.9911	14.11	0.1041
00:00	1.0632	14.79	0.1252

Table B.3: Calibration Results for Day 3

Interval	Scale	RMSE	RMSN
3:15	1.0579	6.50	0.1507
3:30	1.0547	6.33	0.1447
3:45	1.0644	6.70	0.1524
4:00	1.0359	7.56	0.1501
4:15	1.0054	5.97	0.1209
4:30	1.0414	7.47	0.1174
4:45	1.0273	9.72	0.1048
5:00	1.0477	10.28	0.0929
5:15	1.0216	13.16	0.1027
5:30	1.0304	16.34	0.0855
5:45	1.0350	21.18	0.0797
6:00	1.0261	21.55	0.0714
6:15	1.0306	26.47	0.0896
6:30	1.0308	28.23	0.0822
6:45	1.0355	37.80	0.1006
7:00	1.0558	52.00	0.1347
7:15	1.0641	52.25	0.1301
7:30	1.0607	56.65	0.1414
7:45	1.0608	58.88	0.1494
8:00	1.0487	66.71	0.1621
8:15	1.0197	67.79	0.1711
8:30	1.0191	59.51	0.1516
8:45	1.0347	66.50	0.1729
9:00	1.0226	63.51	0.1649
9:15	0.9935	58.42	0.1639
9:30	1.0337	52.87	0.1427
9:45	1.0208	47.20	0.1278
10:00	1.0274	44.11	0.1271
10:15	1.0212	43.40	0.1286
10:30	1.0331	42.37	0.1169
10:45	1.0297	42.96	0.1177
11:00	1.0259	42.15	0.1163
11:15	1.0174	42.45	0.1195
11:30	1.0176	43.49	0.1193
11:45	1.0217	43.32	0.1212
12:00	1.0285	47.61	0.1267
12:15	1.0206	40.50	0.1094
12:30	1.0155	43.64	0.1178
12:45	1.0302	42.99	0.1157
13:00	1.0140	44.06	0.1222
13:15	1.0340	45.93	0.1233
13:30	1.0574	55.17	0.1456
13:45	1.0234	51.47	0.1410
14:00	1.0300	48.30	0.1270
14:15	1.0511	52.22	0.1383
14:30	1.0549	53.40	0.1404
14:45	1.0467	53.50	0.1436
15:00	1.0376	46.56	0.1243
15:15	1.0390	50.06	0.1366
15:30	1.0424	54.17	0.1529
15:45	1.0767	55.41	0.1495
16:00	1.0329	51.15	0.1413
16:15	1.0612	57.88	0.1496
16:30	1.0510	62.60	0.1618
16:45	1.0536	59.09	0.1579
17:00	1.0554	61.94	0.1692
17:15	1.0526	65.07	0.1715
17:30	1.0237	65.03	0.1674
17:45	1.0560	66.13	0.1761
18:00	1.0521	64.23	0.1713
18:15	1.0422	60.51	0.1655
18:30	1.0450	60.62	0.1624
18:45	1.0485	62.48	0.1700
19:00	1.0447	59.02	0.1645
19:15	1.0579	59.64	0.1649
19:30	1.0456	57.61	0.1569
19:45	1.0133	45.17	0.1278
20:00	1.0310	38.89	0.1142
20:15	1.0092	35.23	0.1101
20:30	1.0540	34.31	0.1080
20:45	1.0106	28.76	0.0992
21:00	1.0362	27.67	0.1002
21:15	1.0491	30.54	0.1113
21:30	1.0300	28.74	0.0970
21:45	1.0430	32.81	0.1173
22:00	1.0754	40.55	0.1339
22:15	1.0787	47.85	0.1512
22:30	1.0401	42.81	0.1408
22:45	1.0019	24.56	0.0985
23:00	1.0002	21.73	0.1027
23:15	1.0224	19.35	0.1068
23:30	1.0129	13.51	0.0811
23:45	1.0242	13.38	0.0893
00:00	1.0082	13.60	0.1053

Table B.4: Calibration Results for Day 4

Interval	Scale	RMSE	RMSN
3:15	1.0092	4.07	0.1059
3:30	1.0138	4.55	0.1133
3:45	1.0189	4.79	0.1061
4:00	1.0110	6.07	0.1228
4:15	1.0074	4.68	0.0904
4:30	1.0238	5.60	0.0901
4:45	1.0228	8.41	0.0852
5:00	1.0076	10.73	0.0935
5:15	1.0151	12.37	0.0975
5:30	1.0266	20.24	0.1078
5:45	1.0356	25.37	0.0964
6:00	1.0187	29.26	0.1002
6:15	1.0298	39.43	0.1416
6:30	1.0338	44.05	0.1332
6:45	1.0546	52.93	0.1452
7:00	1.0373	57.37	0.1560
7:15	1.0559	65.30	0.1693
7:30	1.0590	60.38	0.1575
7:45	1.0233	64.79	0.1743
8:00	1.0545	68.32	0.1778
8:15	1.0421	70.82	0.1877
8:30	1.0330	66.38	0.1848
8:45	1.0530	70.39	0.1919
9:00	1.0178	66.64	0.1873
9:15	1.0468	69.53	0.1901
9:30	1.0285	59.74	0.1662
9:45	1.0336	55.89	0.1546
10:00	0.9988	58.64	0.1668
10:15	1.0359	56.03	0.1573
10:30	1.0276	57.85	0.1628
10:45	1.0226	48.40	0.1380
11:00	1.0373	50.74	0.1408
11:15	1.0158	51.92	0.1497
11:30	1.0234	51.73	0.1431
11:45	1.0210	56.68	0.1585
12:00	1.0294	54.10	0.1490
12:15	1.0280	52.05	0.1419
12:30	1.0155	52.85	0.1457
12:45	1.0479	47.51	0.1384
13:00	1.0155	48.92	0.1384
13:15	1.0450	48.60	0.1362
13:30	1.0447	60.24	0.1632
13:45	1.0264	49.28	0.1323
14:00	1.0038	50.80	0.1352
14:15	1.0267	53.73	0.1417
14:30	1.0547	53.53	0.1383
14:45	1.0371	49.81	0.1309
15:00	1.0132	51.54	0.1413
15:15	1.0054	49.94	0.1456
15:30	1.0457	52.54	0.1479
15:45	1.0310	57.62	0.1609
16:00	1.0393	58.35	0.1620
16:15	1.0620	60.45	0.1628
16:30	1.0461	59.93	0.1641
16:45	1.0258	59.94	0.1681
17:00	1.0370	62.67	0.1734
17:15	1.0405	71.09	0.1876
17:30	0.9949	65.75	0.1817
17:45	1.0304	64.75	0.1744
18:00	1.0291	63.08	0.1721
18:15	1.0292	59.52	0.1680
18:30	1.0347	62.11	0.1700
18:45	1.0456	59.34	0.1654
19:00	1.0238	55.40	0.1603
19:15	1.0230	47.18	0.1424
19:30	1.0174	46.12	0.1374
19:45	1.0167	39.54	0.1200
20:00	1.0095	39.71	0.1198
20:15	0.9951	34.54	0.1123
20:30	1.0048	35.52	0.1212
20:45	0.9619	37.73	0.1432
21:00	1.0130	43.38	0.1744
21:15	1.0970	54.33	0.2185
21:30	1.0610	58.63	0.2246
21:45	1.0740	53.06	0.2123
22:00	1.0760	50.02	0.2141
22:15	1.0943	59.86	0.2581
22:30	1.0997	61.52	0.2593
22:45	1.1130	56.27	0.2675
23:00	1.0505	39.67	0.2254
23:15	1.0801	39.90	0.2435
23:30	1.0475	35.55	0.2320
23:45	1.0435	31.44	0.2297
00:00	1.0098	24.74	0.2160

Table B.5: Calibration Results for Day 5

Interval	Scale	RMSE	RMSN
3:15	1.0790	11.60	0.2497
3:30	1.0494	7.35	0.1590
3:45	1.0412	8.30	0.1726
4:00	1.0824	7.68	0.1602
4:15	1.0809	9.09	0.1644
4:30	1.0511	9.93	0.1435
4:45	1.0563	13.07	0.1372
5:00	1.0411	19.47	0.1752
5:15	1.0384	21.11	0.1591
5:30	1.0448	28.77	0.1499
5:45	1.0595	41.15	0.1515
6:00	1.0222	36.38	0.1248
6:15	1.0552	42.99	0.1385
6:30	1.0527	45.77	0.1294
6:45	1.0704	61.01	0.1431
7:00	1.0295	55.99	0.1507
7:15	1.0645	65.99	0.1678
7:30	1.0451	59.80	0.1553
7:45	1.0512	62.52	0.1630
8:00	1.0462	69.65	0.1770
8:15	1.0444	72.31	0.1890
8:30	1.0400	72.74	0.1921
8:45	1.0278	65.98	0.1779
9:00	1.0511	67.52	0.1831
9:15	1.0178	58.53	0.1622
9:30	1.0339	59.56	0.1656
9:45	1.0302	54.34	0.1508
10:00	1.0282	57.31	0.1542
10:15	1.0228	50.79	0.1440
10:30	1.0337	57.63	0.1628
10:45	1.0315	53.42	0.1490
11:00	1.0059	44.68	0.1298
11:15	1.0359	51.16	0.1463
11:30	1.0376	46.37	0.1280
11:45	1.0168	53.83	0.1519
12:00	1.0069	50.35	0.1402
12:15	1.0191	52.94	0.1447
12:30	1.0276	52.26	0.1412
12:45	1.0349	49.07	0.1335
13:00	1.0197	46.23	0.1288
13:15	1.0464	48.52	0.1325
13:30	1.0789	59.83	0.1621
13:45	1.0625	47.29	0.1255
14:00	1.0265	41.82	0.1103
14:15	1.0643	49.70	0.1289
14:30	1.0440	46.97	0.1206
14:45	1.0655	53.19	0.1355
15:00	1.0622	51.72	0.1343
15:15	1.0486	48.39	0.1273
15:30	1.0577	52.88	0.1392
15:45	1.0701	58.56	0.1555
16:00	1.0570	52.08	0.1384
16:15	1.0535	58.60	0.1491
16:30	1.0727	60.45	0.1564
16:45	1.0779	64.25	0.1663
17:00	1.0533	64.28	0.1672
17:15	1.0698	64.42	0.1674
17:30	1.0740	77.96	0.2001
17:45	1.0658	70.73	0.1853
18:00	1.0624	65.87	0.1731
18:15	1.0674	64.88	0.1735
18:30	1.0684	69.59	0.1862
18:45	1.0656	67.84	0.1883
19:00	1.0567	68.79	0.1922
19:15	1.0581	66.82	0.1950
19:30	1.0353	57.47	0.1711
19:45	1.0347	52.58	0.1524
20:00	0.9990	38.21	0.1254

Table B.6: Validation of Estimation Capabilities: Day 6

Interval	Scale	RMSE	RMSN
3:15	1.0035	6.04	0.1628
3:30	1.0271	5.51	0.1492
3:45	1.0313	6.77	0.1660
4:00	0.9962	4.69	0.1036
4:15	1.0551	6.15	0.1161
4:30	0.9742	5.99	0.0975
4:45	1.0363	7.38	0.0764
5:00	1.0199	9.64	0.0871
5:15	1.0141	10.83	0.0849
5:30	1.0271	14.90	0.0789
5:45	1.0316	18.76	0.0731
6:00	1.0143	19.17	0.0680
6:15	1.0321	23.08	0.0766
6:30	1.0521	29.34	0.0834
6:45	1.0796	48.35	0.1268
7:00	1.0262	41.83	0.1102
7:15	1.0370	46.54	0.1110
7:30	1.0882	56.37	0.1245
7:45	1.0580	58.31	0.1479
8:00	1.0642	63.38	0.1613
8:15	1.0404	59.08	0.1521
8:30	1.0261	58.38	0.1569
8:45	1.0428	59.86	0.1634
9:00	1.0388	55.45	0.1525
9:15	1.0376	53.05	0.1449
9:30	1.0440	47.45	0.1289
9:45	1.0345	43.30	0.1199
10:00	1.0265	41.20	0.1129
10:15	1.0214	40.63	0.1148
10:30	1.0402	38.97	0.1080
10:45	1.0387	40.49	0.1093
11:00	1.0295	42.90	0.1191
11:15	1.0443	38.31	0.1075
11:30	1.0133	42.80	0.1187
11:45	1.0385	44.32	0.1228
12:00	1.0472	45.94	0.1207
12:15	1.0321	38.74	0.1032
12:30	1.0390	39.86	0.1071
12:45	1.0222	36.25	0.1006
13:00	1.0448	41.44	0.1081
13:15	1.0531	46.45	0.1193
13:30	1.0982	57.25	0.1471
13:45	1.0506	45.84	0.1189
14:00	1.0456	44.41	0.1155
14:15	1.0585	45.36	0.1165
14:30	1.0557	45.04	0.1151
14:45	1.0506	45.36	0.1178
15:00	1.0430	43.26	0.1148
15:15	1.0529	43.48	0.1164
15:30	1.0697	53.46	0.1371
15:45	1.0550	45.82	0.1183
16:00	1.0709	52.57	0.1351
16:15	1.0700	53.84	0.1345
16:30	1.0559	55.80	0.1433
16:45	1.0790	56.19	0.1452
17:00	1.0688	61.30	0.1563
17:15	1.0573	60.39	0.1530
17:30	1.0485	67.92	0.1737
17:45	1.0457	61.66	0.1707
18:00	1.0505	61.07	0.1671
18:15	1.0602	65.01	0.1751
18:30	1.0471	59.52	0.1614
18:45	1.0309	66.57	0.1950
19:00	0.9601	58.67	0.1989
19:15	0.9813	49.83	0.1586
19:30	1.0556	51.85	0.1504
19:45	1.0519	42.86	0.1214
20:00	1.0222	44.69	0.1382

Table B.7: Validation of Estimation Capabilities: Day 7

Appendix C

Calibration Results for Weekend Days

Calibration results for Saturday and Sunday are tabulated below.

Interval	Scale	RMSE	RMSN
3:15	1.0138	14.78	0.1969
3:30	1.0252	15.28	0.2138
3:45	1.0171	15.02	0.2296
4:00	1.0243	13.85	0.2284
4:15	1.0217	13.04	0.2389
4:30	1.0189	14.82	0.2285
4:45	1.0078	13.84	0.1987
5:00	1.0197	14.14	0.1946
5:15	1.0006	13.77	0.1735
5:30	0.9965	14.77	0.1607
5:45	1.0085	19.70	0.1663
6:00	1.0065	18.59	0.1512
6:15	1.0172	17.05	0.1298
6:30	0.9993	20.87	0.1279
6:45	1.0029	26.77	0.1390
7:00	1.0107	31.49	0.1490
7:15	1.0078	28.78	0.1339
7:30	1.0045	31.17	0.1302
7:45	1.0010	42.61	0.1520
8:00	1.0068	42.36	0.1443
8:15	1.0039	39.71	0.1468
8:30	1.0125	47.50	0.1576
8:45	1.0190	50.34	0.1599
9:00	0.9948	46.20	0.1403
9:15	1.0040	43.87	0.1424
9:30	1.0104	42.25	0.1355
9:45	1.0231	75.02	0.2155
10:00	1.0130	83.71	0.2097
10:15	1.0088	81.58	0.2141
10:30	1.0056	74.26	0.2113
10:45	1.0190	77.23	0.2152
11:00	1.0291	80.73	0.2156
11:15	1.0315	75.42	0.2048
11:30	1.0259	76.66	0.2081
11:45	1.0195	60.73	0.1660
12:00	1.0179	95.38	0.2493
12:15	1.0188	75.38	0.1995
12:30	1.0086	82.95	0.2084
12:45	1.0658	82.04	0.2043
13:00	1.0097	80.13	0.2057
13:15	1.0212	68.49	0.1789
13:30	1.0267	71.03	0.1825
13:45	1.0388	81.22	0.2085
14:00	0.9942	94.01	0.2147
14:15	1.0105	88.92	0.2312
14:30	0.9816	99.20	0.2237
14:45	1.0081	80.93	0.2150
15:00	1.0193	81.90	0.2158
15:15	1.0126	81.67	0.2188
15:30	0.9845	71.67	0.1923
15:45	1.0077	82.22	0.2150
16:00	1.0057	64.53	0.1712
16:15	1.0210	78.72	0.2114
16:30	1.0011	81.37	0.2161
16:45	1.0037	75.87	0.2091
17:00	1.0030	72.06	0.2055
17:15	0.9849	74.35	0.2028
17:30	0.9838	69.94	0.1883
17:45	0.9980	62.40	0.1752
18:00	0.9748	78.63	0.2026
18:15	0.9910	72.11	0.2184
18:30	0.9412	68.15	0.2140
18:45	0.9897	53.47	0.1683
19:00	0.9874	61.01	0.1826
19:15	1.0011	58.49	0.1826
19:30	0.9946	59.58	0.1911
19:45	1.0022	59.42	0.1900
20:00	0.9933	67.60	0.1849
20:15	1.0051	67.83	0.2193
20:30	1.0018	60.73	0.1949
20:45	0.9952	69.60	0.1942
21:00	0.9861	65.90	0.2202
21:15	1.0044	56.70	0.1922
21:30	0.9961	65.00	0.2163
21:45	1.0063	61.18	0.1982
22:00	1.0121	54.27	0.1900
22:15	0.9754	65.46	0.2157
22:30	0.9673	67.38	0.2161
22:45	1.0076	58.10	0.1942
23:00	1.0146	55.70	0.1819
23:15	0.9906	60.80	0.2011
23:30	0.9789	60.94	0.2065
23:45	0.9764	57.25	0.2138
00:00	1.0697	53.12	0.2043

Table C.1: Calibration Results for Saturday 1

Interval	Scale	RMSE	RMSN
3:15	0.9937	13.22	0.1847
3:30	0.9771	10.92	0.1588
3:45	0.9729	11.68	0.1883
4:00	0.9948	10.14	0.1765
4:15	1.0059	10.08	0.1724
4:30	0.9941	10.49	0.1700
4:45	1.0092	12.07	0.1731
5:00	1.0033	10.16	0.1387
5:15	1.0026	11.23	0.1488
5:30	0.9938	12.08	0.1262
5:45	0.9968	13.81	0.1146
6:00	1.0023	15.94	0.1236
6:15	0.9890	15.31	0.1261
6:30	0.9856	17.11	0.1089
6:45	1.0035	22.44	0.1225
7:00	1.0022	25.55	0.1210
7:15	1.0135	27.48	0.1307
7:30	1.0096	34.52	0.1350
7:45	1.0010	41.72	0.1475
8:00	1.0038	40.09	0.1344
8:15	1.0044	33.45	0.1239
8:30	0.9965	34.92	0.1234
8:45	0.9974	38.67	0.1297
9:00	0.9996	41.73	0.1281
9:15	1.0011	38.69	0.1273
9:30	0.9848	39.06	0.1262
9:45	0.9966	47.92	0.1407
10:00	0.9843	55.52	0.1604
10:15	1.0110	47.87	0.1460
10:30	1.0049	45.36	0.1320
10:45	0.9945	43.89	0.1275
11:00	0.9976	44.50	0.1271
11:15	1.0009	46.52	0.1306
11:30	0.9591	77.57	0.2111
11:45	1.0035	42.98	0.1183
12:00	1.0080	53.07	0.1420
12:15	1.0092	51.01	0.1313
12:30	0.9949	56.51	0.1461
12:45	1.0372	66.03	0.1684
13:00	1.0057	63.02	0.1627
13:15	1.0250	47.34	0.1271
13:30	0.9919	52.17	0.1402
13:45	1.0295	56.52	0.1480
14:00	0.9828	85.67	0.2261
14:15	0.9526	74.98	0.2008
14:30	1.0274	69.72	0.1923
14:45	0.9961	54.06	0.1480
15:00	0.9912	62.38	0.1701
15:15	0.9847	50.85	0.1406
15:30	1.0263	56.96	0.1593
15:45	0.9796	49.56	0.1385
16:00	0.9986	56.44	0.1688
16:15	0.9749	66.66	0.1829
16:30	0.9954	57.36	0.1627
16:45	0.9980	44.10	0.1207
17:00	1.0330	45.97	0.1299
17:15	1.0017	42.60	0.1226
17:30	0.9822	45.88	0.1303
17:45	0.9631	60.00	0.1689
18:00	0.9907	52.16	0.1499
18:15	1.0115	36.25	0.1050
18:30	1.0228	41.61	0.1203
18:45	0.9922	36.27	0.1067
19:00	1.0094	34.55	0.1049
19:15	0.9961	37.15	0.1174
19:30	1.0135	36.45	0.1136
19:45	0.9712	66.06	0.2152
20:00	0.9872	34.05	0.1118
20:15	0.9792	36.55	0.1196
20:30	0.9921	39.09	0.1320
20:45	0.9600	39.75	0.1368
21:00	0.9869	39.20	0.1424
21:15	0.9858	33.86	0.1298
21:30	1.0180	34.98	0.1230
21:45	0.9773	37.53	0.1352
22:00	0.9866	40.52	0.1441
22:15	0.9733	42.79	0.1550
22:30	1.0070	37.14	0.1277
22:45	0.9249	76.00	0.2848
23:00	0.9835	42.94	0.1601
23:15	0.9984	45.54	0.1763
23:30	0.9915	31.82	0.1280
23:45	0.9898	36.03	0.1545
00:00	1.0508	27.71	0.1328

Table C.2: Calibration Results for Saturday 2

Interval	Scale	RMSE	RMSN
3:15	1.0200	17.54	0.1921
3:30	1.0233	13.32	0.1806
3:45	1.0178	12.99	0.1910
4:00	1.0100	10.96	0.1831
4:15	1.0328	11.54	0.2019
4:30	1.0325	10.47	0.1688
4:45	1.0153	9.37	0.1604
5:00	1.0277	8.02	0.1302
5:15	1.0036	8.19	0.1318
5:30	1.0212	8.58	0.1178
5:45	1.0216	10.96	0.1252
6:00	1.0102	11.26	0.1283
6:15	1.0121	10.38	0.1199
6:30	1.0155	15.06	0.1503
6:45	0.9826	23.07	0.1988
7:00	0.9982	19.16	0.1471
7:15	1.0179	17.56	0.1314
7:30	1.0224	22.70	0.1385
7:45	1.0110	23.78	0.1338
8:00	1.0188	28.42	0.1456
8:15	1.0054	22.09	0.1218
8:30	1.0081	29.43	0.1476
8:45	1.0059	33.18	0.1550
9:00	1.0065	39.36	0.1676
9:15	1.0027	33.29	0.1395
9:30	1.0076	35.69	0.1323
9:45	1.0201	42.07	0.1350
10:00	1.0087	51.23	0.1546
10:15	1.0104	46.32	0.1441
10:30	1.0224	50.97	0.1558
10:45	1.0093	57.71	0.1709
11:00	1.0250	79.96	0.2280
11:15	1.0069	82.13	0.2057
11:30	1.0116	74.62	0.2127
11:45	1.0032	78.44	0.2259
12:00	0.9960	69.72	0.2064
12:15	0.9914	60.26	0.1831
12:30	0.9972	60.20	0.1836
12:45	0.9920	58.98	0.1813
13:00	1.0062	64.75	0.1953
13:15	1.0100	72.69	0.2061
13:30	1.0086	71.15	0.2081
13:45	0.9932	75.70	0.2044
14:00	1.0130	74.75	0.2079
14:15	0.9940	78.39	0.1993
14:30	0.9969	79.11	0.2198
14:45	1.0031	82.86	0.2259
15:00	1.0040	76.29	0.2160
15:15	1.0124	70.53	0.1967
15:30	1.0051	69.28	0.1841
15:45	0.9952	76.23	0.2055
16:00	0.9935	77.71	0.2183
16:15	1.0159	71.77	0.1960
16:30	1.0125	67.63	0.1813
16:45	0.9907	75.83	0.2054
17:00	0.9957	80.15	0.2004
17:15	1.0183	74.90	0.2067
17:30	0.9914	79.73	0.2189
17:45	1.0010	68.33	0.1933
18:00	1.0074	65.89	0.1885
18:15	0.9767	67.00	0.1962
18:30	1.0077	69.71	0.2114
18:45	1.0120	63.58	0.1895
19:00	1.0094	60.46	0.1818
19:15	1.0221	64.35	0.1918
19:30	1.0076	55.38	0.1713
19:45	1.0273	62.52	0.1927
20:00	1.0160	60.29	0.1885
20:15	0.9927	50.15	0.1583
20:30	0.9930	51.03	0.1655
20:45	1.0131	49.36	0.1655
21:00	0.9836	40.76	0.1445
21:15	0.9966	44.50	0.1588
21:30	1.0093	39.18	0.1373
21:45	1.0149	35.03	0.1285
22:00	1.0027	37.98	0.1477
22:15	1.0243	32.41	0.1301
22:30	1.0186	36.32	0.1502
22:45	1.0117	32.54	0.1520
23:00	1.0367	27.65	0.1439
23:15	1.0458	26.04	0.1489
23:30	1.0400	24.49	0.1554
23:45	1.0341	21.30	0.1483
00:00	0.9234	22.13	0.1829

Table C.3: Calibration Results for Sunday 1

Interval	Scale	RMSE	RMSN
3:15	0.9818	15.04	0.1882
3:30	0.9644	11.11	0.1558
3:45	0.9637	11.86	0.1853
4:00	0.9774	10.35	0.1691
4:15	0.9837	8.53	0.1566
4:30	0.9967	7.50	0.1379
4:45	0.9885	7.56	0.1426
5:00	1.0014	8.25	0.1324
5:15	0.9887	7.66	0.1404
5:30	0.9878	10.22	0.1565
5:45	0.9897	11.16	0.1511
6:00	0.9785	10.88	0.1460
6:15	0.9864	9.61	0.1291
6:30	0.9766	12.26	0.1354
6:45	1.0004	14.15	0.1288
7:00	0.9653	18.05	0.1546
7:15	0.9899	18.75	0.1519
7:30	0.9753	23.27	0.1595
7:45	1.0030	29.03	0.1651
8:00	0.9913	26.86	0.1466
8:15	1.0057	21.73	0.1295
8:30	0.9980	22.80	0.1252
8:45	0.9870	23.00	0.1170
9:00	0.9994	24.12	0.1113
9:15	0.9938	27.03	0.1221
9:30	0.9952	30.98	0.1277
9:45	0.9989	36.18	0.1271
10:00	0.9880	36.03	0.1208
10:15	0.9951	37.24	0.1271
10:30	0.9875	42.68	0.1376
10:45	1.0139	41.29	0.1259
11:00	0.9856	43.08	0.1327
11:15	0.9798	44.14	0.1404
11:30	0.9998	41.02	0.1249
11:45	0.9995	45.88	0.1367
12:00	0.9944	40.62	0.1252
12:15	0.9963	48.01	0.1392
12:30	0.9946	45.60	0.1310
12:45	0.9944	50.56	0.1406
13:00	0.9862	48.95	0.1380
13:15	0.9950	46.47	0.1323
13:30	1.0075	45.25	0.1311
13:45	0.9918	47.29	0.1383
14:00	0.9964	43.51	0.1241
14:15	0.9882	43.51	0.1240
14:30	0.9945	47.05	0.1327
14:45	0.9985	47.91	0.1334
15:00	1.0006	43.16	0.1221
15:15	1.0056	43.59	0.1178
15:30	1.0020	51.70	0.1399
15:45	0.9967	46.94	0.1343
16:00	1.0071	46.13	0.1314
16:15	0.9998	47.03	0.1361
16:30	0.9948	54.68	0.1560
16:45	1.0310	47.16	0.1345
17:00	1.0075	41.90	0.1210
17:15	0.9891	59.12	0.1770
17:30	1.0068	59.29	0.1752
17:45	1.0007	54.96	0.1560
18:00	0.9931	43.30	0.1326
18:15	1.0050	46.16	0.1422
18:30	0.9955	44.64	0.1339
18:45	0.9816	47.77	0.1437
19:00	1.0049	42.25	0.1337
19:15	0.9937	46.17	0.1506
19:30	1.0025	43.16	0.1394
19:45	0.9764	46.95	0.1521
20:00	0.9791	38.19	0.1255
20:15	0.9555	43.60	0.1494
20:30	0.9697	42.77	0.1443
20:45	0.9680	47.54	0.1713
21:00	0.9660	42.97	0.1556
21:15	0.9639	43.26	0.1605
21:30	0.9626	41.67	0.1567
21:45	0.9614	37.49	0.1454
22:00	0.9973	33.49	0.1355
22:15	0.9764	29.16	0.1255
22:30	0.9936	26.10	0.1196
22:45	0.9772	31.32	0.1525
23:00	0.9719	25.15	0.1393
23:15	1.0004	21.83	0.1310
23:30	0.9795	21.13	0.1342
23:45	0.9859	16.90	0.1229
00:00	1.0032	20.41	0.1733

Table C.4: Calibration Results for Sunday 2

O-D Group	Time Group	AR Degree	$f_{r,h}^{r,h-1}$	$f_{r,h}^{r,h-2}$	$f_{r,h}^{r,h-3}$	$f_{r,h}^{r,h-4}$
> 200	4:00 AM–8:00 AM	4	0.7109	0.2558	-0.1817	0.3351
	8:00 AM–1:00 PM	4	0.4497	0.2598	0.0639	0.0414
	1:00 PM–7:00 PM	4	0.1607	0.0659	0.0138	0.0435
	7:00 PM–0:00 AM	4	0.0627	0.0588	0.0554	0.0074
50–200	4:00 AM–8:00 AM	4	0.6155	0.2125	0.0245	0.3125
	8:00 AM–1:00 PM	4	0.1323	0.3339	0.1216	0.2462
	1:00 PM–7:00 PM	4	0.1486	0.2247	0.0627	0.0438
	7:00 PM–0:00 AM	0	0.0000	0.0000	0.0000	0.0000
20–50	4:00 AM–8:00 AM	4	0.5519	0.2464	0.1499	0.1159
	8:00 AM–1:00 PM	4	0.2354	0.1390	0.0673	0.2592
	1:00 PM–7:00 PM	4	0.1959	0.1181	0.0939	0.0748
	7:00 PM–0:00 AM	4	0.0128	0.0023	0.106	0.0036
< 20	4:00 AM–8:00 AM	4	0.4916	0.1967	0.1297	0.1236
	8:00 AM–1:00 PM	4	0.2053	0.1939	0.0713	0.1485
	1:00 PM–7:00 PM	4	0.1655	0.1578	0.0923	0.1148
	7:00 PM–0:00 AM	4	0.3999	-0.0906	0.0457	0.0043

Table C.5: Calibrated Auto-Regressive Factors: Saturday

O-D Group	Time Group	AR Degree	$f_{r,h}^{r,h-1}$	$f_{r,h}^{r,h-2}$	$f_{r,h}^{r,h-3}$	$f_{r,h}^{r,h-4}$
> 200	4:00 AM–8:00 AM	4	0.7873	0.1080	-0.0005	0.2083
	8:00 AM–1:00 PM	4	0.6017	0.1667	0.1665	0.0756
	1:00 PM–7:00 PM	4	0.0012	0.0035	0.0026	0.4014
	7:00 PM–0:00 AM	4	0.3768	0.2841	0.1840	0.0023
50–200	4:00 AM–8:00 AM	4	0.5753	0.3180	0.1600	0.0022
	8:00 AM–1:00 PM	4	0.5529	0.1857	0.1440	0.0604
	1:00 PM–7:00 PM	4	0.2105	0.1525	0.0726	0.0306
	7:00 PM–0:00 AM	4	0.2410	0.2952	0.0476	0.1561
20–50	4:00 AM–8:00 AM	4	0.3479	0.2493	0.1068	0.2145
	8:00 AM–1:00 PM	4	0.5042	0.1629	0.0678	0.1341
	1:00 PM–7:00 PM	4	0.3798	0.0961	0.0841	0.0000
	7:00 PM–0:00 AM	4	0.3432	0.2358	0.0856	0.0830
< 20	4:00 AM–8:00 AM	4	0.2529	0.1714	0.1519	0.1366
	8:00 AM–1:00 PM	4	0.3151	0.1862	0.1323	0.1423
	1:00 PM–7:00 PM	4	0.2533	0.1162	0.0435	0.0456
	7:00 PM–0:00 AM	4	0.2733	0.2267	0.0787	0.0675

Table C.6: Calibrated Auto-Regressive Factors: Sunday

OD Group	Time Group	AR Degree	$f_{r,h}^{r,h-1}$
> 200	4:00 AM–8:00 AM	1	1.0000
	8:00 AM–1:00 PM	1	0.8140
	1:00 PM–7:00 PM	1	0.2495
	7:00 PM–0:00 AM	1	0.1781
50–200	4:00 AM–8:00 AM	1	1.0000
	8:00 AM–1:00 PM	1	0.7349
	1:00 PM–7:00 PM	1	0.4523
	7:00 PM–0:00 AM	0	-
20–50	4:00 AM–8:00 AM	1	1.0000
	8:00 AM–1:00 PM	1	0.5964
	1:00 PM–7:00 PM	1	0.4392
	7:00 PM–0:00 AM	1	0.0258
< 20	4:00 AM–8:00 AM	1	0.9428
	8:00 AM–1:00 PM	1	0.5577
	1:00 PM–7:00 PM	1	0.4700
	7:00 PM–0:00 AM	1	0.3153

Table C.7: Calibrated MA AR Factors: Saturday

OD Group	Time Group	AR Degree	$f_{r,h}^{r,h-1}$
> 200	4:00 AM–8:00 AM	1	0.8452
	8:00 AM–1:00 PM	1	1.0000
	1:00 PM–7:00 PM	0	-
	7:00 PM–0:00 AM	1	0.8452
50–200	4:00 AM–8:00 AM	1	1.0000
	8:00 AM–1:00 PM	1	0.9459
	1:00 PM–7:00 PM	1	0.4499
	7:00 PM–0:00 AM	1	0.6968
20–50	4:00 AM–8:00 AM	1	0.8338
	8:00 AM–1:00 PM	1	0.8478
	1:00 PM–7:00 PM	1	0.5518
	7:00 PM–0:00 AM	1	0.7204
< 20	4:00 AM–8:00 AM	1	0.6550
	8:00 AM–1:00 PM	1	0.7295
	1:00 PM–7:00 PM	1	0.4334
	7:00 PM–0:00 AM	1	0.6105

Table C.8: Calibrated MA AR Factors: Sunday

Bibliography

[1] Anthony, C., Piya, C. and Surachet, P., "A Multi-Objective Model for Locating Automatic Vehicle Identification Readers", presented at the 83^{rd} annual meeting of the *Transportation Research Board*, 2004.

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

[3] Ashok, K., "Estimation and Prediction of Time Dependent Origin Destination Flows", Ph.D. thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, 1996.

[4] Ashok, K. and Ben-Akiva, M., "Alternative Approaches for Real-Time Estimation and Prediction of Time-Dependent Origin-Destination Flows", *Transportation Science*, Vol. 34, No. 1, 2000, pp. 21–36.

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

[6] Balakrishna, R., "Calibration of the Demand Simulator in a Dynamic Traffic Assignment System", Master's thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, 2002.

[7] Balakrishna, R., Koutsopoulos, H. N. and Ben-Akiva, M., "Calibration and Validation of Dynamic Traffic Assignment Systems", accepted for presentation and

publication by the 16^{th} *International Symposium on Transportation and Traffic Theory*, 2005.

[8] Bianco, L., Confessore, G. and Reverberi, P., "Optimal Location of Traffic Counting Points for Transport Network Control", *IFAC Transportation Systems*, Chania, Greece, 1997.

[9] California Highway Patrol Incident Log, *http://cad.chp.ca.gov/*.

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

[11] Cascetta, E., Inaudi, D. and Marquis, G., "Dynamic Estimation of Origin-Destination Matrices using Traffic Counts", *Transportation Science*, Vol. 27(4), 1993, pp. 363–373.

[12] Freeway Performance Measurement System, *http://pems.eecs.berkeley.edu/Public/*.

[13] Gelb, A. (ed), "Applied Optimal Estimation", MIT Press, Cambridge, MA, 1974.

[14] Gentili, M. and Mirchandani, P., "Locating Image Sensors on Traffic Networks", presented in *TRISTAN V: The fifth triennial symposium in transportation analysis*, 2004.

[15] Green, W., "Econometric Analysis", 4^{th} edn, Prentice Hall, 2000.

[16] Hazelton, M. L., "Estimation of Origin-Destination Matrices from Link Counts on Un-congested Networks", *Transportation Research B*, Vol. 34, 2000, pp. 549–566.

[17] He, R., Miaou, S., Ran, B. and Lan, C., "Developing an On-Line Calibration Process for an Analytical Dynamic Traffic Assignment Model", presented at the 78^{th} annual meeting of the *Transportation Research Board*, 1998.

[18] Huynh, N., Chiu, Y. and Mahamassani, H. S., "Determining Optimal Locations for Variable Message Signs under Stochastic Incident Scenarios", presented at the 80^{th} annual meeting of the *Transportation Research Board*, 2000.

[19] Huynh, N., Chiu, Y. and Mahamassani, H. S., "Finding Near Optimal Locations for Variable Message Signs for Real-Time Network Traffic Management", presented at the 82^{nd} annual meeting of the *Transportation Research Board*, 2002.

[20] Kunde, K., "Calibration of Mesoscopic Traffic Simulation Models for Dynamic Traffic Assignment", Master's thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, 2002.

[21] Massachusetts Institute of Technology, "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, 2000.

[22] Massachusetts Institute of Technology, "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, 2000.

[23] Massachusetts Institute of Technology, Lecture notes for 6.241 "Dynamic Systems and Control", Fall 2004.

[24] National Research Council, "Highway Capacity Manual", *Special Report 209*, 3^{rd} ed., Washington, D.C., 1998

[25] Ramming, S., "Network Knowledge and Route Choice", Ph.D. thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, 2001.

[26] State Space (control)-Wikipedia, *http://en.wikipedia.org/wiki/State_space_(controls)*, April 5, 2005.

[27] Yang, H., Iida, Y. and Sasaki, T., "An Analysis of the Reliability of an Origin-Destination Trip Matrix Estimated from Traffic Counts", *Transportation Research B*, Vol. 25, 1991, pp. 351–363.

[28] Yang, H. and Zhou, J., "Optimal Traffic Counting Locations for Origin-Destination Matrix Estimation", *Transportation Research B*, Vol. 33, No. 2, 1998, pp. 109–126.