A LINEAR PROGRAMMING APPROACH FOR SYNTHESIZING ORIGIN-DESTINATION (O-D) TRIP TABLES BASED ON A PARTIAL SET OF LINK TRAFFIC VOLUMES

by

Arvind Narayanan

Thesis Submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Civil Engineering

APPROVED:

Dr. R. Sivanandan
(Co-Chairman)

Dr. H. D. Sherali
(Co-Chairman)

Dr. Antoine G. Hobeika

December 15, 1995

Blacksburg, Virginia

c.2

A Linear Programming Approach for Synthesizing Origin-Destination(O-D) Trip Tables Based on a Partial Set of Link Traffic Volumes

by
Arvind Narayanan
Dr. R. Sivanandan (Co-Chairman)
Dr. H. D. Sherali (Co-Chairman)
Civil Engineering

(ABSTRACT)

This research effort is motivated by the need to quickly obtain origin-destination (O-D) trip information for an urban area, without expending the excessive time and effort usually accompanying survey-based methods. The proposed approach aims to exploit the information contained even in a *"partial set"* of available link volumes to estimate an O-D trip table.

Recently, a new approach to synthesize a trip table from observed link flows on the network was developed at Virginia Tech. This approach employs a linear programming formulation and is based on a non-proportional assignment, user-equilibrium principle. The model is designed to determine a traffic equilibrium network flow solution that reproduces the link volume data, if such a solution exists. If such alternate solutions exist, then it is designed to find that which most closely resembles a target trip table. A modified column generation technique is employed to solve the problem. The methodology also accommodates a specified prior or target trip table, and drives the solution toward a tendency to match this table using user controlled parameters. The limitation of this approach is that it needs the specification of a complete set of link flows for its accurate operation. Such a requirement limits the applicability of this model to real networks, since link volumes are not always available on all the links of a network.

This research work enhances the above linear programming methodology, adding the capability to estimate OD trip tables even when only a *"partial set"* of

link traffic counts are available. The proposed approach formulates a sequence of linear programs to approximate a fundamentally nonlinear optimization problem that is employed to estimate origin-destination flows, given incomplete network flow information. The research suggests techniques for terminating a given linear program in the sequence, as well as criteria for terminating the sequences of such LPs, and also develops a procedure for continually updating the cost vector from one linear program to the next. Modifications in the column generation technique, necessary to solve the revised model formulation, are also developed.

The enhanced model is evaluated and compared with the maximum entropy approach, which is a popular approach for OD table estimation. These models are evaluated through tests on an artificial network and a real network. The tests aim to evaluate these models using various sets of link volumes and prior table information. The results indicate that the linear programming approach performs better than the maximum entropy approach for most cases. Conclusions and recommendations for future research are also presented.

ACKNOWLEDGMENTS

I would like to thank Dr. H. D. Sherali and Dr. R. Sivanandan for their guidance, patience and encouragement through out this research effort. I would also like to express my gratitude to Dr. A. G. Hobeika for his support and for serving on my thesis committee.

Id like to thank my friends for making my stay in Blacksburg very enjoyable. TPS and Vijay as room mates contributed to a lot of fun. Jay brought in a nice TV. I recommend Nithya's coffee and hospitality. Her husband -Jay, has a nice set of Tee Shirts which are accessible if you have a little charm and don't care too much about fit. Srikanth the Chef Jonnalagaddas cooking was often a topic of mid flight conversation between me and Roof as we played basketball. Working with Srikanth has been a lot of fun and may his nagging skills continue to grow.

The Center has been a great place to work and I would like to thank everyone involved. I would also like to place on record a deep sense of gratitude for the Centers ex photocopying machine which jammed more frequently than it copied..

I would like to thank my parents, sister, brother in law, despoki, nandita, jay and nithya for all their support.

TABLE OF CONTENTS

1. INTRODUCTION 1

1.1 Origin Destination Trip Tables 1
1.2 Establishing OD Trip Tables 1
1.3 OD Trip Table Estimation Using Link Traffic Volumes 2
1.4 Linear Programming Approach for OD Table Estimation 2
1.5 Research Objectives 3
1.6 Organization of this Thesis 4

2. LITERATURE REVIEW 6

2.1 Introduction 6
2.2 Overview of Existing Approaches 7
2.3 The Maximum Entropy Approach 9
2.4 The Linear Programming Approach 11

3. MODEL FORMULATION AND SOLUTION PROCEDURE 13

3.1 Introduction 13
3.2 Notation 14
3.3 Background 17
3.4 Problem Statement 19
3.5 Approach 20
3.6 Model Formulation 22
3.6.1 Update Criterion for the Route Cost Vector 24
3.6.2 Link Cost Resetting Procedure 28
3.6.3 Cost Vector $(c_B)^r$ Updating Procedure 29
3.6.4 Dual Variable Updating Procedure 30
3.7 Solution Technique for LPMLV 30
3.8 Summary of the Column Generation Algorithm (CGA) for Solving LPMLV 37

4. EVALUATION OF MODEL USING ARTIFICIAL NETWORK 41
4.1 Introduction 41
4.2 Evaluation Procedure 42
4.2.1 Evaluation Factors 42
4.2.1.1 Available Link Volumes 42
4.2.1.2 Target Trip Tables 43
4.2.2 Comparison of Results 44
4.2.2.1 Replication of Observed Link Volumes 44
4.2.2.2 Closeness of Estimated Table to True Trip Table 45
4.2.2.3 Computer Resources for Tests 45
4.3 Test Networks 46
4.3.1 Artificial Network 46
4.3.1.1 Target Trip Tables 48
4.4 Test Results for Artificial Network 48
4.4.1 Case 1: Structural Table as Target 48
4.4.2 Case 2 - No Information Based Trip Tables as Targets 52
4.4.2.1 Case 2a: Partial No Information Based Trip Table (45% Cells with Info.) 53
4.4.2.2 Case 2b: Partial No Information Based Trip Table (64% Cells with Info.) 55
4.4.2.3 Case 2c: Partial No Information Based Trip Table (82% Cells with Info.) 58
4.4.2.4 Case 2d: Partial No Information Based Trip Table (100% Cells with Info.) 60
4.4.3 Relatively Small Error Target Trip Tables 63
4.4.3.1 Case 3a: Partial Relatively Small Error Trip Table (45% Cells with Info.) 63
4.4.3.2 Case 3b: Partial Relatively Small Error Trip Table (64% Cells with Info.) 65
4.4.3.3 Case 3c: Partial Relatively Small Error Trip Table (82% Cells with Info.) 68
4.4.3.4 Case 3d: Relatively Small Error Target Trip Table (100% Cells with Info.) 70
4.4.4 Correct Trip Table Based Target Tables 72
4.4.4.1 Case 4a: Partial Correct Trip Table (45% Cells with Info.) 72
4.4.4.2 Case 4b: Partial Correct Trip Table (64% Cells with Info.) 75
4.4.4.3 Case 4c: Partial Correct Trip Table (82% Cells with Info.) 77
4.4.4.4 Case 4d: Correct Trip Table (100% Cells with Info.) 80
4.5 Number of Objective Function Resets and Iterations for LPMLV 81
4.6 Comparison with a No Impedance Updating Procedure 83
4.7 Conclusions 84

5. EVALUATION USING A REAL NETWORK 89
5.1 Introduction 89
5.2 Purdue University Network 89
5.2.1 Target Trip Tables 90
5.3 Test Results 91
5.3.1 Case 1: Structural Table 91
5.3.2 Case 2a: Partial No Prior Information Table (60% Cells with Info.) 93
5.3.3 Case 2b: Partial No Prior Information Table (80% Cells with Info.) 95
5.3.4 Case 2c: Partial No Prior Information Table (100% Cells with Info.) 97
5.3.5 Case 3a: Relatively Small Error Trip Table (60% Cells with Info., $\psi = 0.8$) 99
5.3.6 Case 3b: Relatively Small Error Trip Table (80% Cells with Info., $\psi = 0.8$) 101
5.3.7 Case 3c: Relatively Small Error Trip Table (100% Cells with Info., $\psi = 0.8$) 103
5.3.8 Case 4a: Relatively Small Error Trip Table (60% Cells with Info., $\psi = 0.9$) 105
5.3.9 Case 4b: Relatively Small Error Trip Table (80% Cells with Info., $\psi = 0.9$) 107
5.3.10 Case 4c: Relatively Small Error Trip Table (100% Cells with Info., $\psi = 0.9$) 109
5.4 Conclusions 111

6. CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH 117
6.1 Conclusions 117
6.2 Recommendations for Further Research 119

REFERENCES 120

List of Tables

Table 4-1: Structural Target Trip Table 48
Table 4-2: Performance of the Linear Programming Approach (Structural Table as a Target) 49
Table 4-3: Performance of the Maximum Entropy Approach (Structural Table as a Target) 50
Table 4-4: No Information Based Trip Table (45% Cells with Info.) 52
Table 4-5: Performance of the Linear Programming Approach (No Information Based Trip Table (45% Cells with Info.) as a Target) 53
Table 4-6: Performance of the Maximum Entropy Approach (No Information Based Trip Table (45% Cells with Info.) as a Target) 54
Table 4-7: No Information Based Trip Table (64% Cells with Info.) 55
Table 4-8: Performance of the Linear Programming Approach (No Information Based Trip Table (64% Cells with Info.) as a Target) 56
Table 4-9: Performance of the Maximum Entropy Approach (No Information Based Trip Table (64% Cells with Info.) as a Target) 57
Table 4-10: No Information Based Trip Table (82% Cells with Info.) 58
Table 4-11: Performance of the Linear Programming Approach (No Information Based Trip Table (82% Cells with Info.) as a Target) 59
Table 4-12: Performance of the Maximum Entropy Approach (No Information Based Trip Table (82% Cells with Info.) as a Target) 59
Table 4-13: No Information Based Trip Table (100% Cells with Info.) 60
Table 4-14: Performance of the Linear Programming Approach (No Information Based Trip Table (100% Cells with Info.) as a Target) 61
Table 4-15: Performance of the Maximum Entropy Approach (No Information Based Trip Table (100% Cells with Info.) As Target) 62
Table 4-16: Relatively Small Error Trip Table (45% Cells with Info.) 63
Table 4-17: Performance of the Linear Programming Approach (Relatively Small Error Trip Table (45% Cells with Info.) as a Target 64
Table 4-18: Performance of the Maximum Entropy Approach Relatively Small Error Trip Table (45% Cells with Info.) as a Target) 64
Table 4-19: Relatively Small Error Target Trip Table (64% Cells with

Info.) 66
Table 4-20: Performance of the Linear Programming Approach (Relatively Small Error Trip Table (64% Cells with Info.) as a Target) 66
Table 4-21: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table (64% Cells with Info.) as a Target) 67
Table 4-22: Relatively Small Error Trip Table (82% Cells with Info.) 68
Table 4-23: Performance of the Linear Programming Approach (Relatively Small Error Trip Table (82% Cells with Info.) as a Target) 69
Table 4-24: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table (82% Cells with Info.) as a Target) 69
Table 4-25: Relatively Small Error Target Trip Table(100% Cells with Info.) 70
Table 4-26: Performance of the Linear Programming Approach (Relatively Small Error Trip Table (100% Cells with Info.) as a Target) 71
Table 4-27: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table (100% Cells with Info.) as a Target) 71
Table 4-28: (Correct Trip Table Based (45% Cells with Info.) Target) 72
Table 4-29: Performance of the Linear Programming Approach (Correct Trip Table (45% Cells with Info.) as a Target) 73
Table 4-30: Performance of the Maximum Entropy Approach (Correct Trip Table (45% Cells with Info.) as a Target) 74
Table 4-31: Partial Correct Trip Table (64% Cells with Info.) 75
Table 4-32: Performance of the Linear Programming Approach Correct Trip Table (64% Cells with Info.) as a Target 76
Table 4-33: Performance of the Maximum Entropy Approach Correct Trip Table (64% Cells with Info.) as a Target 76
Table 4-34: Correct Trip Table (82% Cells with Info.) 77
Table 4-35: Performance of the Linear Programming Approach (Correct Trip Table (82% Cells with Info.) as a Target) 78
Table 4-36: Performance of the Maximum Entropy Approach (Correct Trip Table (82% Cells with Info.) as a Target) 79
Table 4-37: Correct Trip Table (100% Cells with Info.) 80
Table 4-38: Performance of the Linear Programming Approach (Correct Trip Table as a Target) 80
Table 4-39: Performance of the Maximum Entropy Approach (Correct Trip Table as a Target) 81
Table 4-40. Number of Objective Function Resets and Iterations for

LPMLV 82
Table 4-41: Comparison of ϕ statistic: No Link Cost Update Implementation vs. LPMLV 84
Table 5-1: Performance of the Linear Programming Approach (Structural Table as a Target) 91
Table 5-2: Performance of the Maximum Entropy Approach (Structural Table as a Target) 92
Table 5-3: Performance of the Linear Programming Approach (No Prior Information Table (60% Cells with Info.) as a Target) 93
Table 5-4: Performance of the Maximum Entropy Approach (No Prior Information Table (60% Cells with Info.) as a Target) 94
Table 5-5: Performance of the Linear Programming Approach (No Prior Information Table (80% Cells with Info.) as a Target) 95
Table 5-6: Performance of the Maximum Entropy Approach (No Prior Information Table (80% Cells with Info.) as a Target) 96
Table 5-7: Performance of the Linear Programming Approach (No Prior Information Table (100% Cells with Info.) as a Target) 97
Table 5-8: Performance of the Maximum Entropy Approach (No Prior Information Table (100% Cells with Info.) as a Target) 98
Table 5-9: Performance of the Linear Programming Approach (Relatively Small Error Trip Table (60% Cells with Info., $\psi = 0.8$) as a Target) 99
Table 5-10: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table (60% Cells with Info., $\psi = 0.8$) as a Target) 100
Table 5-11: Performance of the Linear Programming Approach (Relatively Small Error Trip Table(80% Cells with Info., $\psi = 0.8$) as a Target) 102
Table 5-12: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table(80% Cells with Info., $\psi = 0.8$) as a Target) 102
Table 5-13: Performance of the Linear Programming Approach (Relatively Small Error Trip Table(100% Cells with Info., $\psi = 0.8$) as a Target) 104
Table 5-14: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table (100% Cells with Info.,

$\psi = 0.8$) as a Target) 104

Table 5-15: Performance of the Linear Programming Approach (Relatively Small Error Trip Table (60% Cells with Info., $\psi = 0.9$) as a Target) 106

Table 5-16: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table (60% Cells with Info., $\psi = 0.9$) as a Target) 106

Table 5-17: Performance of the Linear Programming Approach (Relatively Small Error Trip Table (80% Cells with Info., $\psi = 0.9$) as a Target) 108

Table 5-18: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table (80% Cells with Info., $\psi = 0.9$) as a Target) 108

Table 5-19: Performance of the Linear Programming Approach (Relatively Small Error Trip Table (100% Cells with Info., $\psi = 0.9$) as a Target) 109

Table 5-20: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table (80% Cells with Info., $\psi = 0.9$) as a Target) 110

List of Figures

Fig 4-1	Test Network-1 Charactersitics	47
Fig 4-2	Artificial Network - LP Model PHI (TT) Values	85
Fig 4-3	Artificial Network - THE Model PHI (TT) Values	86
Fig 4-4	Artificial Network - THE Model %MAE (Vol) Values	87
Fig 5-1	Purdue Network - LP Model PHI (TT) Values	112
Fig 5-2	Purdue Network - THE Model PHI (TT) Values	113
Fig 5-3	Purdue Network - LP Model %MAE (Vol) Values	114
Fig 5-4	Purdue Network - THE Model %MAE (Vol) Values	115

1. Introduction

1.1 Origin Destination Trip Tables

An origin-destination (OD) trip table is a two dimensional matrix of elements whose cell values represent the number of trips made between various OD zone pairs in a given region. An OD table thus captures the travel patterns of trip makers within a region.

The knowledge of such an OD matrix is a basis for several transportation planning and traffic engineering studies, as well as for the purposes of route guidance and traffic diversion. An important component of transportation planning involves an understanding of the impact of a particular plan on the travel behavior of motorists. In order to acquire such an understanding it is often very necessary to know the travel patterns within the area. Similarly, many traffic engineering operations require the knowledge of OD patterns. Effective diversion strategies can be devised during an incident, only with the knowledge of the destinations of travelers.

1.2 Establishing OD Trip Tables

Interest in more effective transportation planning for urban areas led to several approaches for OD trip table estimation. These approaches focus on establishing OD tables through extensive surveys of the travelers. These surveys include home interviews, license plate surveys, road side surveys etc. These survey techniques are expensive, time consuming and labor intensive. Most of these methods are conducted through sampling and thus make it impossible to determine a “real” trip table. Even if all the trips are recorded on a particular day, the OD table so determined may not be stable over time, due to day to day variations.

In addition, there are other problems associated with conventional methods for OD estimation. One of the disadvantages of conventional techniques, is that they are too long drawn and elaborate to permit the planners to capture the changes in trip pattern that arise due to changes in the factors that influence travel patterns. For instance, it is prohibitive to carry out conventional OD estimation surveys whenever the land use changes affect the travel patterns, thus outdating the previous OD trip table. This will necessitate re-surveying, requiring further investment of capital, manpower and time. In addition, several small planning agencies are often unable to carry out these surveys due to budgetary limitations. Nevertheless, these agencies require trip table information for many planning and management purposes. These reasons led to the development of cheaper and quicker techniques for OD estimation.

1.3 OD Trip Table Estimation Using Link Traffic Volumes

The reasons outlined in section 1.2 form the motivation for the development of theoretical approaches that estimate OD trip tables from link traffic volumes. These approaches have evolved since the 1970's. These approaches attempt to predict travel patterns using measurements of the link volumes in a network. They were formulated with the idea of exploiting information contained in the link volumes. They aim at deriving a trip table that is consistent with the observed set of link volumes. The approaches developed utilize different sets of techniques ranging from parameter calibration that began in the 1970's, to fuzzy logic and linear programming approaches developed in the 1990's. These approaches are described in greater detail in the literature review described in Chapter 2.

1.4 Linear Programming Approach for OD Table Estimation

One of the recent approaches developed at Virginia Tech, employs a linear programming based formulation for synthesizing a trip table from observed link flows on the network (Sivanandan [1991] and Sherali et. al [1994]). This approach is based on non-proportional assignment and is motivated by user equilibrium. The model is designed to determine a traffic equilibrium network flow solution which reproduces the link volume data exactly, if such a solution indeed does exist. If such alternate solutions exist, then it is designed to find that which most closely resembles a specified target trip table. However, it recognizes that due to incomplete information, the traffic may not conform to an equilibrium flow pattern, and permits for inconsistencies in the link flow data. Accordingly, the model permits violations in the equilibrium conditions, and also accommodates an inconsistent set of link counts. The model also employs a specified seed trip table and allows for controlling the output solution trip table to be as close to the seed table as desired.

However, this model needs the specification of a complete set of link flows for its accurate operation. This thesis enhances the model to estimate OD trip tables when only a partial set of link volume counts is available.

1.5 Research Objectives

The linear programming model developed at Virginia Tech (Sivanandan, [1991] & Sherali et al [1994]) assumes the knowledge of link volumes on all the links of a network. Such an assumption is necessary for the complete specification of a linear objective function, that is one in which the cost vector is not a function of the decision variables. Also, the link volumes are used to define the solution space for the linear program. Such an assumption places severe limitations on the applicability of this model for most real networks. This is because link volume data is seldom available for all links within a network. Many planning and traffic agencies operate under budgetary constraints that do not permit them to collect volume counts on all the links in a network. For the above reasons, the enhancement of the linear programming approach is a major objective of this research. The validation of this model with tests on real and artificial networks is another important objective of this project.

Till date, several approaches have been developed for the purposes of OD estimation. The maximum entropy approach is popular and well known among these. This approach has been incorporated in several transportation software packages. This approach has been tested by many researchers including those at Virginia Tech, and in addition this model lends itself to the estimation of OD trip tables with only a partial set of link volumes. Thus another objective of this research is to evaluate the performance of the enhanced linear programming model with that of the maximum entropy model.

There are two main factors that usually influence the output of OD estimation models. The input of known link volumes, is obviously a very important factor, since the solution procedure is primarily designed to replicate these volumes. In addition, these models employ a seed trip table that is used to reduce the degree of underspecification of the problem, and thereby guide the solution. The study of the influence of the seed table (prior table) on the OD estimation models, is another research objective.

In light of the above discussion, the specific objectives of this research are to:

1. Enhance the linear programming model developed by Sivanandan [1991] to synthesize a trip table even in the presence of a partial set of available link volumes. While retaining the original structure of the model, the enhancement was achieved by modifying the formulation and accordingly accounting for the changes required to the solution procedure.

2. Evaluate the performance of the enhanced linear programming approach and compare with the maximum entropy model. The evaluation included sensitivity of the models to various extents of available link volume information and quality of prior (target) trip table information.

1.6 Organization of this Thesis

The organization of this thesis is as follows. Chapter 2 presents a literature review of OD estimation models. Chapter 3 describes the model development and solution procedure. The validation of this model and the maximum entropy approach is described in detail in chapters 4 and 5. The conclusions and recommendations for further research are discussed in chapter 6.

2. LITERATURE REVIEW

2.1 Introduction

The early 1970's saw the dawn of theoretical approaches for estimating O-D trip tables using link traffic counts. The interest in these approaches was kindled due to a need for a shift in planning philosophy from long term to intermediate and short term, necessitated by limitations in budget, time and man power resources. Since then, different approaches to accomplish this task have evolved, incorporating various desirable features and refinements. Many of these are non linear, and are based on the general framework of trip table estimation utilizing link volumes.

Earlier approaches to this problem relied on linear or nonlinear regression analysis to construct demand models assuming a gravity type flow pattern, for estimating trip table entries. These models, however, required data on zone specific variables like population, employment etc. One group of models was based on the network traffic equilibrium approach, so that it could account for congestion effects. Yet another group, attempted to extract a most likely trip table consistent with link volumes, through maximum entropy and/or minimum information approaches. Other group of models utilize statistical techniques to produce further estimates based on prior information. The interest in the problem continues, as is evident by reports of enhancements and refinements to the above procedures.

This chapter first gives the reader an overview of the literature. Then it describes one of the popular models - the maximum entropy approach in greater detail. Then the linear programming model developed by Sivanandan[1991] and Sherali et al[1994] is described.

2.2 Overview of Existing Approaches .

The need for obtaining trip tables through cheaper and quicker means was fulfilled in the early 1970's, which saw the dawn of theoretical approaches for synthesizing trip tables from easily available information regarding traffic volumes on links of the road network. In most urban areas, such data is collected rather inexpensively on a regular basis using detectors. Essentially, the existing approaches for synthesizing trip tables from link counts fall into the broad category of one of two types of approaches, namely, parameter calibration techniques, and matrix estimation methods (O'Neill, 1987). Parameter calibration approaches use linear or nonlinear regression analysis to construct demand models assuming a gravity-type flow pattern, in order to estimate the trip table entries. However, these methods require zonal data for calibrating the parameters of the demand models, and are therefore of limited practical use, because zonal data are not only less easily available, but also become outdated relatively soon. Willumsen [1978], Nguyen [1984], and O'Neill [1987] review the related literature for this class of methods. On the other hand, matrix estimation techniques rely only on link traffic counts, and prior information in the form of a trip table, and are easy to implement. There are three types of approaches within this category, namely, statistical estimation methods, models based on maximum-entropy/minimum-information theory, and network equilibrium based techniques.

The statistical estimation methods produce future estimates based on prior information, employing Bayesian inference techniques such as in Maher [1983], or using least squares estimation models such as in Carey et al. [1981], Cascetta [1984], McNeil and Hendrickson [1985] O'Neill [1987], and Bell [1991]. In contrast, the other two types of matrix estimation techniques are based on mathematical programming concepts, and are hence more relevant to our approach. The first of these two approaches attempts to determine the most likely O-D trip table that is consistent with the information contained in the link traffic volume data, while either maximizing the entropy or using a minimum information based objective function with respect to a prior trip table. Willumsen [1978] presents the elements of the maximum entropy approach following Wilson [1970], while Van Zuylen [1978] discusses the principles of information minimization. Models following both these philosophies are further developed and analyzed in Van Zuylen and Willumsen [1980]. These models are based on a proportional-assignment assumption in which some exogenously determined coefficients specify the proportion of trips between each O-D pair that use a particular link, for each link in the network. Many researchers, such as Hall et al. [1980], Van Zuylen [1981], Van Vliet and Willumsen [1981], Willumsen [1982, 1984], Bell [1983a, b], Nguyen [1984], and Hamerslag and Immers [1988], conducted tests or have proposed improvements and enhancements to this type of model. Also, Fisk [1988] has shown how to combine the maximum entropy model and the user-optimal assignment into one problem. In a more recent approach, Brenninger-Göthe, Jörnsten, and Lundgren [1989] suggest a nonpreemptive multiobjective technique in which efficient points are sought that compromise between the separation of the solution from the observed traffic counts versus its separation from the prior target O-D matrix data. This separation is measured using an entropy function.

In contrast with the proportional-assignment assumption, whenever network congestion effects are prominent, a non proportional-assignment approach based on the traffic equilibrium principles of Wardrop [1952] becomes more appropriate. Nguyen [1977, 1984] develops such an approach via two models for deriving O-D trip tables from traffic counts. In the first model, a standard traffic equilibrium solution is sought (see Dafermos and Sparrow, 1969) subject to the side constraint that the total equilibrium cost computed by using observed path impedances should equal the known total observed system cost. Nguyen suggests that this model is suitable only for small networks in which there are but a few O-D pairs. In the second model, the foregoing side constraint is accommodated into the objective function such that the Karush-Kuhn-Tucker optimality conditions ensure that for each O-D path flow, the equilibrium based impedance equals the observed impedance using the available link traffic data. Hence, both models attempt to regenerate observed costs, rather than observed flows. Turnquist and Gur [1979] present an iterative, heuristic solution technique for this second model based on the Frank-Wolfe algorithm, and Gur, Turnquist, Schneider, LeBlanc, and Kurth [1980] imbed a similar approach in their well known LINKOD system. This LINKOD model has been extensively tested and validated by Han et al. [1981], Han and Sullivan [1983], and Dowling and May [1984]. For all these techniques, there can exist several O-D trip tables that reproduce the observed traffic volumes, and so target trip tables are usually specified to guide the solution toward more likely flow patterns. Hence, a distribution assumption is utilized in order to extract the most likely trip table. Nguyen [1984] suggests an alternative approach for selecting the most likely trip matrix, and Fisk and Boyce [1983] propose a combined distribution and assignment technique based on the model presented by Erlander, Nguyen, and Stewart [1979], which also relies on link flow data. In particular, Fisk [1989] has shown that the network equilibrium approach, the maximum entropy approach, and the combined distribution-assignment-formulation can be expected to produce the same results under congested network conditions, when observed link volumes correspond to an equilibrium flow pattern. Cascetta and Nguyen [1988] have also presented a unified framework for estimating or updating origin-destination matrices from traffic counts.

2.3 The Maximum Entropy Approach

Willumsen (Van Zuylen and Willumsen, 1980) proposed a maximum entropy approach for OD estimation. This approach is popular and used by many existing planning agencies, and is detailed below. It is based on Wilson's (1970) application of the concept of entropy to the O-D trip matrix. Here, the most likely trip matrix is defined as the one having the greatest number of micro-states associated with it. Attempting to maximize the number of ways of selecting a trip matrix, Willumsen formulates the problem as:

$$\max F(T,T^*) = -\sum_i \sum_j t_{ij}(\ln t_{ij} / t_{ij}^* - 1) \qquad 2.1$$

subject to

$$v_a = \sum_i \sum_j p_{ij}^a t_{ij}$$

$t_{ij}^* =$ prior (or) old trip matrix

$$v_a^* = \sum_i \sum_j t_{ij}^* p_{ij}^a$$

where, T represents the matrix of interest, and T^* represents a reference matrix, v_a^* is the volume of traffic on link a, p_{ij}^a is the proportion of trips between origin i and destination j using link a

The derived table would be the most likely that is consistent with information contained in the link flows.

Both the above approaches are shown to reduce to a multiproportional problem. In particular, the maximum entropy approach reduces to solving the following optimality conditions:

$$t_{ij} = t_{ij}^* \prod_a x_a^{p_{ij}^a}$$

where,

$$x_a = (\sum_{ij} t_{ij})^{1/L} . e^{-\lambda_a}$$

and where L denotes the number of counted links, and λ_a is the Lagrange multiplier corresponding to the count on link a constraint. Van Zuylen and Willumsen [1980] also give an algorithm for solving the above problem based on Murchland's [1977] algorithm for the multiproportional problem. It is further indicated that the coverage of the problem has not been satisfactorily proved. Counts on all the links are not necessary. However, a complete set of counts is expected to yield better results. It is to be noted that both the above formulations require information on link usage proportions.

O'Neill [1987] summarizes the conditions to be satisfied by the maximum entropy model, in order for an estimated trip table to reproduce observed volumes fully, as: (1) consistency of link volumes (flow conservation), (2) consistency of prior trip table with observed flows and route choice proportions, and (3) consistency of route choice assumption with observed flows.

Many researchers (Hall et al. [1980], Van Zuylen [1981], Van Vilet and Willumsen [1981], Willumsen [1982], Bell [1983], and Nguyen [1984] etc.) have conducted tests or have proposed improvements on this type of model. Of particular interest, is the attempt to consider the effects of congestion, through equilibrium assignment. Willumsen [1982] proposed and tested a heuristic model that includes the equilibrium principle. Bromage [1988], while at Central Transportation Planning Staff (CTPS), Boston, programmed the maximum entropy model, incorporating a capacity restraint procedure for the assignment step. This program was enhanced by Beagan [1990], also of CTPS, to include an equilibrium assignment option, as proposed by Hall et al. [1980]. Further improvements to the model were carried out to the model by Ed Bromage in 1994. The improved versions of the model were in fact used in this research to test the maximum entropy approach.

2.4 The Linear Programming Approach

In contrast with the nonlinear network equilibrium approaches described so far, the linear programming approach of Sivanandan[1991] and Sherali et al [1994] employs the non proportional-assignment assumption, and that finds a user equilibrium solution which reproduces the observed link flows whenever such a solution exists, but one which is a linear programming model. The model recognizes that due to incomplete information, although the individual user is driven by the choice of a least impedance path, the actual flow may not exactly conform to a user equilibrium solution. Moreover, due to inherent inconsistencies in the link traffic data, there might not exist a trip table that can exactly duplicate the link flows. Accordingly, these features are accommodated into the model through suitable artificial variables and objective penalties.

Variations in these penalties can also be used to reflect biases in selecting amongst alternative viable O-D trip tables. However, if there does exist a user equilibrium solution that reproduces the link flows, the model, with suitable penalty parameters, will determine such a solution along with the corresponding O-D trip table. Additionally, due to the potentially large number of alternative paths to be considered between the different O-D pairs, we develop an efficient column generation technique that utilizes shortest-path sub problems in order to determine an optimal solution to the linear programming model. The model is also designed to handle the situation in which a prior target trip table is specified, and it is required to find a solution that, in addition to the foregoing considerations, has a tendency toward reproducing this table as closely as possible. This modification guides the alternative solutions of the original model toward a prescribed distribution, and also permits additional positive path flows at optimality.

Sivanandan[1991] reports tests conducted with this model. These tests indicate that the model performs better than the maximum entropy and LINKOD approaches both in terms of accuracy and computational efficiency. However one of the drawbacks with this model is that it requires the specification of all the link volumes in the network. This limits the application of the model to real networks because link volumes are usually not available for all the links in a network. The present research aims at overcoming this defect and the next chapter describes the formulation and solution procedure of the enhanced model to account for cases in which only a partial set of link counts are available.

3. MODEL FORMULATION & SOLUTION PROCEDURE

3.1 Introduction

A linear programming approach (LP(TT)) for synthesizing OD tables from link volumes has been developed at Virginia Tech by Sivanandan [1991] and Sherali et al. [1994]. This model aims to find a trip table that replicates the observed link volumes to the extent possible, while extracting a user equilibrium solution, if it exists, and being as close as desired to a specified prior trip table. This approach has been tested by Sivanandan [1991] and compared with other existing models, revealing that the linear programming model is superior in terms of accuracy as well as computational efficiency.

Link volume information plays a very important role in the OD estimation process of LP(TT). They are used to define the solution space for the linear program. More importantly, link volume information for the entire network is required for the complete specification of the linear objective function, so that the cost vector is not a function of the decision variables. Such a requirement on the availability of link volume counts places severe limitations on the applicability of this model for most real networks, as explained in Chapter 1. This necessitates the enhancement of the linear programming model to account for cases where only a partial set of link volumes might be available.

This chapter deals with the development of a model to accordingly enhance the foregoing approach. The concept of the approach rests on employing a sequence of linear optimization problems to approximate a solution to a fundamentally nonlinear problem.

This chapter is organized to first expose the reader to the notation used in the remaining sections. Then the background to the problem is presented. This is followed by a description of the proposed approach. Then the model formulation is presented and the solution procedure is described.

3.2 Notation

Given an urban road network for a particular region, let $G(N, A)$ represent the underlying digraph. Here, N is the set of nodes representing either traffic intersection points where flow is conserved, or zones where trips are generated and/or where trips terminate, and A denotes the set of corresponding directed links or arcs representing the roadways existing between designated pairs of nodes. Let A_v represents the set of links in the traffic network under consideration for which traffic volume information is available, and let A_m represent the set of links for which volumes are missing, where $A \equiv A_v \cup A_m$. Let OD denote the set of origin-destination (O-D) pairs that comprise the trip table to be estimated, where O is the set of possible source or origin nodes, and D is the set of potential sink or destination nodes. Note that a node representing a given zone is typically both an origin as well as a destination node.

Let T_{ij} represent the trip interchange from origin i to destination j of OD. Let us consider the implicit enumeration of all possible (n_{ij}) paths p_{ij}^k, $k = 1,2,....,n_{ij}$, between each O-D pair $(i,j) \in OD$, where p_{ij}^k is a vector that has one component for each link $a \in A$, this component being unity if the corresponding link belongs to that particular path, and being zero otherwise. Denote x_{ij}^k as the flow on the path p_{ij}^k for each $k = 1,2,...n_{ij}^k$, $((i,j) \in OD)$, and let x denote the vector of components x_{ij}^k.

It is also well known that due to errors in measurement and due to approximations in network representation, there might arise errors or inconsistencies in the set of observed volumes. To account for these inconsistencies, two nonnegative artificial variables y^+ and y^-, with respective components y_a^+ and y_a^- for each link $a \in A_v$, are introduced. Suppose that a prior trip table is specified, having associated O-D flows $Q_{ij} > 0$ for each $(i,j) \in \overline{OD} \subseteq OD$, where $\overline{OD}$ might represent some significant or key OD pairs. Let $\overline{f_a}$ represent the set of observed volumes on links $a \in A_v$, and let $(p_{ij}^k)_a$ denote the component of (p_{ij}^k) that corresponds to link $a \in A_v$. For each $(i,j) \varepsilon \overline{OD}$, the deviation of the O-D trip interchange $T_{ij} = \sum_{k=1}^{n_{ij}} x_{ij}^k$ from the target trip table value Q_{ij} is measured by the difference of two nonnegative ("artificial") variables Y_{ij}^+ and Y_{ij}^-. The effective absolute value $(Y_{ij}^+ + Y_{ij}^-)$ of this deviation is penalized in the objective function via a penalty parameter $M_\sigma \geq 0$.

Based on the observed flows $\bar{f}_a, a \varepsilon A$, we can compute a corresponding time-based link impedance $\bar{c}_a \equiv c_a(\bar{f}_a)$, where $c_a(\overline{f}_a)$ is a link travel time/cost function, for example, of the following form as suggested by the Bureau of Public Roads (BPR) [1964]:

$$c_a(f_a) = c_a^F . [1 + 0.15(f_a / u_a)^4] \forall a \varepsilon A \qquad (3.1)$$

Here, f_a is the flow on the link $a \varepsilon A$, c_a^F is the free flow travel time/cost on this link, and u_a is its flow capacity. Since we only need to consider links $a \varepsilon A$ for which $\bar{f}_a > 0$, and since multiplying $\overline{c_a}$ for all $a \varepsilon A$ by a constant leaves the problem invariant, we will assume for the sake of our development that $c_a \geq 1$ and is integral for each $a \varepsilon A$. Let us denote $\bar{c}$ as the vector of components $\bar{c}_a$, $a \varepsilon A$. (This would require the determination of a scale factor to integerize $\bar{c}_a \forall a \varepsilon A$, and then using the scaled form of (3.1), we could round up any value computed via (3.1) for any estimated flows on the links $a \in A_m$.)

Now, if the observed flow pattern indeed represents a network user equilibrium solution corresponding to some interchange of traffic between the designated O-D pairs, then by Wardrop's (first) principle [1952], all the routes between any O-D pair that have positive flows should have equal travel costs, and this cost must not exceed the travel cost on any other unused route between this O-D pair. Notationally, let $c_{ij}^k \equiv \bar{c}.p_{ij}^k$ denote the time impedance or cost on route k between O-D pair (i,j) for each $k = 1,\ldots,n_{ij},(i,j) \varepsilon OD$, and let $c_{ij}^* = \min\{c_{ij}^k, k = 1,\ldots,n_{ij}\}$. Furthermore, let us define

$$K_{ij} = \{k \varepsilon \{1,\ldots,n_{ij}\} : c_{ij}^k = c_{ij}^*\}, \text{ and let } \overline{K}_{ij} = \{1,\ldots,n_{ij}\} - K_{ij}. \qquad (3.2)$$

Then, according to Wardrop's principle, if an equilibrium solution exists that reproduces the observed flows, we should be able to find a solution to (1) satisfying the condition that $x_{ij}^k > 0$ only if $k \varepsilon K_{ij}$, for each $(i, j) \in OD$.

3.3 Background

The linear programming model (LP(TT)) developed by Sivanandan [1991] and Sherali et al. [1994] for synthesizing OD trip tables from link counts has the following form.

LP(TT):

$$\textit{Minimize} \qquad \sum_{(i,j)\varepsilon OD} \sum_{k=1}^{n_{ij}} \hat{c}_{ij}^k x_{ij}^k + Me.(y^+ + y^-) + M_\sigma \sum_{(i,j)\varepsilon \overline{OD}} (Y_{ij}^+ + Y_{ij}^-) \quad (3.3.a)$$

$$\textit{subject to} \qquad \sum_{(i,j)\varepsilon OD} \sum_{k=1}^{n_{ij}} p_{ij}^k x_{ij}^k + y^+ - y^- = \bar{f} \quad (3.3.b)$$

$$\sum_{k=1}^{n_{ij}} x_{ij}^k + Y_{ij}^+ - Y_{ij}^- = Q_{ij} \forall (i, j) \varepsilon \ \overline{OD} \quad (3.3.c)$$

$$x \ge 0, y^+ \ge 0, y^- \ge 0, Y^+ \ge 0, Y^- \ge 0 \quad (3.3.d)$$

where $\hat{c}_{ij}^k = \begin{cases} \bar{c}.\, p_{ij}^k \ \forall\, k \in K_{ij}, \\ M_1 \bar{c}.\, p_{ij}^k \ \forall\, k \in \overline{K}_{ij}, \end{cases}$ (3.3.e)

and where, $M_1 > 1$.

This model estimates the OD trip table by minimizing the objective function of Equation (3.3.a) subject to the observed volumes constraint of Equation (3.3.b), and the trip table constraints of Equation (3.3.c). Sherali et al. show that the value of the penalty parameter M can be set to $M \geq 1 + \bar{c}_* + (|A|!)\overline{C}_{total}$, where $\overline{C}_{total} \equiv \sum_{a \varepsilon A_v} \bar{c}_a \bar{f}_a$ and $\bar{c}_* \equiv maximum\{\bar{c}_a : a \varepsilon A\}$, to ensure that $y^* = 0$ in any optimal solution, given that such a solution exists. It is also shown that a value of $M \geq 1 + \bar{c}_* + \overline{C}_{total}$ ensures that $e.y^* < 1$. The value chosen for M_σ must naturally tradeoff the penalty imposed for deviations from the targeted trip table values with the remainder of the objective function. Since, the parameter M penalizes similar deviations in the accounted link flows from the observed ones, the authors set $M_\sigma = \sigma M$, where $0 \leq \sigma \leq 1$. This value of σ is selected to reflect the relative degree of importance in minimizing the trip table deviations versus the link flow deviations, ranging from unimportant ($\sigma = 0$) to equally important ($\sigma = 1$).

Sherali et al. report tests carried out with LP(TT) on an artificial test network that was developed by Gur et al. These tests revealed that the model is accurate as well as computationally efficient. The model was faster and more accurate than the maximum entropy and LINKOD models. These encouraging results gave an additional impetus to enhance this model further by extending it to account for missing link volume information.

3.4 Problem Statement

The network equilibrium formulation of LP(TT) requires a very accurate estimate of the cost vector $\hat{c}_{ij}^{k}$ which is a function of the link costs. Since link costs are a monotonically increasing function of link volumes, the model LP(TT) requires an accurate specification of the link volumes. In practice, link volume information is not available for all the links in a network. Thus the applicability of the model LP(TT) to such a network is then questionable.

Link volume information plays a two-fold role in model LP(TT). The link volume f_a for a link is used to determine the link cost c_a for that link. Thus it is a factor that determines the route cost terms ($\hat{c}_{ij}^{k}$) in the objective function. Secondly, the values for f_a are used to formulate the constraints for the model. Thus one may state that link volume information is used to determine the objective gradient and solution space for the linear program LP(TT). With this discussion in mind, then, it is easy to see that in the event of missing or information on a set of link volume, the performance of the model is seriously limited.

The objective of the present research is to enhance the linear programming model (LP(TT)) to account for missing link volumes in a network. However, it is inevitable that the solution space for any variant of the model LP(TT) expands due to the loss of one constraint for every missing link volume. Thus the focus of this research is to effectively enhance the model to accommodate partial link volumes counts, while accounting for the functional dependence of the objective function on the missing link volumes.

3.5 Approach

If one were able to obtain some knowledge of the link costs, say by actually measuring the travel time on all the links $a \in A_m$, then one may be able to estimate the link cost vector $\bar{c}$ to a reasonable degree of accuracy. However, such an approach is often not viable, due to extensive manpower requirements and the cost of the operation that is involved.

Another possible approach for estimating the cost vector $\bar{c}$ is through determining the missing link volumes by assigning a prior trip table developed for the network. If such a table is available, and if it truly represents the prevailing trip pattern, then it is indeed possible to obtain a good measure of these missing link volumes by assigning this table to the network. It is to be noted that the assignment technique must correctly reflect driver route-choice behavior, and that information about network characteristics such as capacity, free flow travel time, etc., be known with good confidence. However, this approach may not always be appropriate because the prior trip table may not be representative of current trip patterns and may have been outdated due to changes in land-use over time. The assignment of such a table will then produce erroneous link volumes that are inconsistent with the actual volumes. Also, if some significant cell values of the trip table are missing, then this assignment technique cannot estimate the missing link volumes satisfactorily, and the subsequent calculations of the link costs on this network are bound to be incorrect.

A more fundamental approach to the problem of incomplete link volume data specification, is to let the link cost component be a function of the sum of unknown OD path flows x_{ij}^{k}'s that utilize the directed arc a, $a \in A_m$. This would, however, introduce nonlinear terms in the objective function of LP(TT), because the route cost coefficients in the objective function are now a function of the decision variables . This would transform the linear program into a nonlinear optimization problem. Such a nonlinear formulation may be represented by:

$$\textit{Minimize} \quad \sum_{(i,j)\varepsilon OD} \sum_{k=1}^{n_{ij}} \hat{c}_{ij}^{k} x_{ij}^{k} + Me.(y^{+} + y^{-}) + M_{\sigma} \sum_{(i,j)\varepsilon \overline{OD}} (Y_{ij}^{+} + Y_{ij}^{-}) \tag{3.4.a}$$

$$\textit{Subject to:} \quad \sum_{(i,j)\in OD} \sum_{k=1}^{n_{ij}} (p_{ij}^{k})_a x_{ij}^{k} + y_a^{+} - y_a^{-} = \overline{f}_a \ \forall \ a \in A_v \tag{3.4.b}$$

$$\sum_{k=1}^{n_{ij}} x_{ij}^{k} + Y_{ij}^{+} - Y_{ij}^{-} = Q_{ij} \forall (i,j) \varepsilon \ \overline{OD} \tag{3.4.c}$$

$$x \geq 0, y^{+} \geq 0, y^{-} \geq 0, Y^{+} \geq 0, Y^{-} \geq 0, \tag{3.4.d}$$

where $\bar{c}_a = f(\sum_{i,j \in OD} \sum_{k=1}^{n_{ij}} (p_{ij}^{k})_a x_{ij}^{k}) \forall a \in A_m$, $\hat{c}$ is given by (3.3.e), (3.4 e)

and where $\bar{x}$ is optimal to the linear program (3.4.a) -(3.4.d).

Note that the data in the objective function is itself dependent on the solution to this problem, hence introducing the stated nonlinearity in the problem. Hence, in essence, we are attempting to determine a set of flows $\bar{x}$ such that with the objective function computed using this flow vector $\bar{x}$, the linear program (3.4.a) - (3.4.d) reproduces $\bar{x}$ as an optimal solution.

The present approach attempts to use a sequence of linear programs to approximate the nonlinear optimization problem given by Equations (3.4.a - 3.4.e). Each linear program is a modified form of the model LP(TT), thus allowing us to take advantage of the solution procedure developed for this model.

3.6 Model Formulation

As stated in the previous section the nonlinear program described by Equations 3.4.a-3.4.e is approximated by a sequence of linear programs. Consider the following optimization problem proposed for the r^{th} variant of the linear program.

LPMLV:

$$Minimize \sum_{(i,j)\varepsilon OD} \sum_{k=1}^{n_{ij}} (\hat{c}_{ij}^k)^r x_{ij}^k + M \sum_{a \in A_v} (y_a^+ + y_a^-) + M_\sigma \sum_{(i,j)\varepsilon \overline{OD}} (Y_{ij}^+ + Y_{ij}^-) \tag{3.5.a}$$

Subject to:

$$\sum_{(i,j)\in OD} \sum_{k=1}^{n_{ij}} (p_{ij}^k)_a x_{ij}^k + y_a^+ - y_a^- = \bar{f}_a \ \forall \ a \in A_v \tag{3.5.b}$$

$$\sum_{k=1}^{n_{ij}} x_{ij}^k + Y_{ij}^+ - Y_{ij}^- = Q_{ij} \ \forall (i,j) \in \overline{OD} \tag{3.5.c}$$

$$x \geq 0, y^+ \geq 0, y^- \geq 0, Y^+ \geq 0, Y^- \geq 0 \tag{3.5.d}$$

where $$(\hat{c}_{ij}^k)^r = \begin{cases} (\bar{c})^r . p_{ij}^k \ \forall \, k \in (K_{ij})^r, \\ M_1 (\bar{c})^r . p_{ij}^k \ \forall \, k \in (\overline{K}_{ij})^r \end{cases} \tag{3.5.e}$$

and where $(\bar{c})^r$ is based on some assumed set of flows.

Here, $(\hat{c}_{ij}^k)^r$ is the route cost vector for the r^{th} linear program, defined by the r^{th} choice $(\bar{c})^r$ of the link impedance vector. As in Section 3.2 of this chapter, here again, we assume for the sake of development that $(\bar{c}_a)^r$ is integral and ≥ 1 $\forall \, a \in A$. This vector is updated during the solution of the model, as described in the sections to follow. The value of M_1 can be set to 2 as suggested by Sherali et al. [1994]. Note that there are no constraints in this linear program for links whose volumes are not known. The proposed methodology solves the r^{th} linear program for a certain number of iterations and then after updating the link costs for $a \in A_m$ and the route cost vector, proceeds to formulate and solve the $r+1^{th}$ program. Since the constraints for the r^{th} and $r+1^{th}$ program remain the same, a basic feasible solution for the r^{th} program continues to be basic feasible for the $r+1^{th}$ program. Hence to proceed to the $r+1^{th}$ linear program we only need to update the cost vector in the objective function along with the associated dual variables.

In the formulation suggested above, two important issues need to be addressed. The first issue concerns the update rate of the objective function, and the second concerns the update procedure to be employed.

3.6.1 Update Criterion for the Route Cost Vector

The motivation for the update criterion employed by the proposed methodology stems from an observation that there exists a bound on the decrease in the objective function value for any iteration of LPMLV.

Lemma 3.1: The maximum value of the marginal rate of decrease in the objective function for any iteration of model LPMLV is given by $(\sum_a |\bar{\pi}_a| + \max|\bar{\mu}_{ij}|)$.

Proof: The entering column of LPMLV can be only one of the following:

1. A unit vector corresponding to an entering artificial variable.
2. $\left[\left[p_{ij}^k\right]_{A_v}^t, e_{ij}^t\right]^t$, where e_{ij} is a unit vector, with zeros for all rows, and one for the row with the trip table constraint (if it exists) for the interchange $ij \in \overline{OD}$, and the superscript t denotes the transpose operator.

Hence, the marginal rate of decrease in the objective function when a variable enters by pivoting at row i is either $\pm\bar{\pi}_a - M,$ or $\pm\bar{\mu}_{ij} - M_\sigma$ or $((p_{ij}^k)_{A_v}.\bar{\pi} + \bar{\mu}_{ij} - (\hat{c}_{ij}^k)^r)$.

Since, $|\bar{\pi}_a| \leq \sum_{a \in A_v} |\bar{\pi}_a|$, $(p_{ij}^k)_{A_v}.\bar{\pi} \leq \sum_{a \in A_v} |\bar{\pi}_a|$, $\bar{\mu}_{ij} \leq max\ |\bar{\mu}_{ij}|$, and M, M_σ, and $(\hat{c}_{ij}^k)^r$ are non-negative, it follows that the maximum value of the marginal rate of decrease in the objective function is given by $(\sum_{a}|\overline{\pi}_a| + \max|\overline{\mu}_{ij}|)$. This completes the proof.

Lemma 3.2 The maximum value of any variable in any basic feasible solution to LPMLV is bounded by $R_{max} = (\ |A_v| + |\overline{OD}| - 1\ )(\ |A_v|\ )(\ \upsilon + \upsilon|\overline{OD}|\ )$where $\upsilon = \max\{\overline{f}_a, a \in A_v, Q_{ij}, (i,j) \in \overline{OD}\}$.

Proof: Note that due to linear dependency y_a^+ and y_a^- cannot be simultanesouly basic in any BFS, and similarly, nor can Y_{ij}^+ and Y_{ij}^-, for any $a \in A_v$, or $(i,j) \in \overline{OD}$. Furthermore if y_a^- or Y_a^- are basic, then they are determined via their respective equations in (3.5.b) and (3.5.c), while the other basic variables are determined by the remaining active constraints. However, the latter active constraints have nonnegative coefficients of 0 or 1 on the LHS, and have $\overline{f}_a$ and Q_{ij} values on the RHS. Hence each of $x_{ij}^k, y_a^+, Y_{ij}^+$ are bounded above by $\upsilon = \max\{\overline{f}_a, a \in A_v, Q_{ij}, (i,j) \in \overline{OD}\}$.Consider an explicit enumeration of all the constraints of model LPMLV. Adding all the constraints, one may represent the sum of the artificial variables y_a^- and Y_{ij}^- by Equation 3.6.a below.

$$\sum_{a \in A_v} y_a^- + \sum_{(i,j) \in \overline{OD}} Y_{ij}^- = \sum_{(i,j) \in OD} \sum_{k=1}^{n_{ij}} x_{ij}^k [p_{ij}^k . W] + \sum_{(i,j) \in \overline{OD}} x_{ij}^k + \sum_{a \in A_v} y_a^+ + \sum_{(i,j) \in \overline{OD}} Y_{ij}^+ - \sum_{a \in A_v} \overline{f}_a - \sum_{(i,j) \in \overline{OD}} Q_{ij} \tag{3.6.a}$$

where W is a vector of components $w_a, a \in A$, such that $w_a = \begin{cases} 1 \forall a \in Av \\ 0 \forall a \in A_m \end{cases}$.

Since $\overline{f}_a \geq 0, \forall a \in A_v$, and $Q_{ij} \geq 0, \forall (i,j) \in \overline{OD}$, we obtain the following relationship:

$$\sum_{a \in A_v} y_a^- + \sum_{(i,j) \in \overline{OD}} Y_{ij}^- \leq \sum_{(i,j) \in OD} \sum_{k=1}^{n_{ij}} x_{ij}^k [p_{ij}^k . W] + \sum_{(i,j) \in \overline{OD}} x_{ij}^k + \sum_{a \in A_v} y_a^+ + \sum_{(i,j) \in \overline{OD}} Y_{ij}^+ \; . \qquad (3.6.b)$$

Noting first that the maximum value of $[p_{ij}^k . W] = |A_v|$, and that the variables on the RHS of Equation (3.6.b) are bounded by υ, and further that the maximum number of basic variables among the $x_{ij}^k, y_a^+, Y_{ij}^+$, in order to have atleast one basic variable from the variables y_a^- and Y_{ij}^-, is $|A_v| + |\overline{OD}| - 1$, it immedietely follows that

$$\sum_{a \in A_v} y_a^- + \sum_{(i,j) \in \overline{OD}} Y_{ij}^- < (|A_v| + |\overline{OD}| - 1)(|A_v|)\upsilon \; + \upsilon|\overline{OD}| . \qquad (3.6.c)$$

From 3.6.c, and since each of $x_{ij}^k, y_a^+, Y_{ij}^+$ are bounded above by υ, it follows that the maximum value of any variable in any BFS for LPMLV is bounded by $R_{max} = (|A_v| + |\overline{OD}| - 1)(|A_v|)\upsilon \; + \upsilon|\overline{OD}|$. This completes the proof.

Lemma 3.3 The maximum decrease in the objective function for any iteration of LPMLV is bounded by $(\sum_a |\overline{\pi}_a| \; + \max|\overline{\mu}_{ij}| \;)^* R_{max}$

Proof: By Lemma 3.1 we know that $\sum_a |\overline{\pi}_a| \; + \max|\overline{\mu}_{ij}|$ is the bound on the marginal decrease of the objective function. From Lemma 3.2 we know $R_{\max}$ to be the bound on the largest value of any variable in any BFS to LPMLV. The maximum decrease in the objective function, then has to be bounded by the product of maximum rate of decrease in the objective function and the largest change in any non-basic variable that could occur when pivoting for any basis for LPMLV. This completes the proof.

Remark 3.1 If we were to reset the objective function of the r^{th} linear program of LPMLV, as soon as the objective function value decreased by an amount in excess of $(\sum_a |\bar{\pi}_a| + \max|\bar{\mu}_{ij}|)^* R_{max}$, then we could ensure that at least two iterations would be performed before resetting the objective function.

Remark 3.2 If one were to implement model LPMLV to initialize with an all artificial basis, the objective reset criteria of $(\sum_a |\bar{\pi}_a| + \max|\bar{\mu}_{ij}|)^* R_{max}$ is likely to be very high during the initial stages of the program. This is because the dual variables assume very large values due to the high cost of the basic variables, that are mostly artificial variables. Having such a high value of the reset criteria would then require that several iterations be completed before we reset the objective function. (Tests, conducted with this criteria confirmed this.) This would then imply that we would be ignoring the cost variations for links $a \in A_m$, (these are likely to fluctuate widely in the initial stages of LPMLV) during this process. This motivates us to have a reset criteria that permits us to update model LPMLV at a faster rate, during the initial stages of the problem, and at a less frequent rate when the link costs for $a \in A_m$ no longer vary significantly from iteration to iteration.

In order to address the concerns discussed in Remark 3.2 the proposed methodology aims to adopt an objective resetting criteria that is small enough to capture the link volume fluctuations for $a \in A_m$ that occur during the initial stages of the program, and once these link volumes have somewhat stabilized, it resets the objective function less frequently. To accomplish this, the methodology proposed scales the criteria of Lemma 3.3 by multiplying it with $S_r \leq 1$. This value is continually increased every time we reset the objective. While any monotonically increasing function S_r of r, that has a limiting value of 1, would be appropriate, the following function was adopted

$$S_r = \frac{1}{1 + se^{-r}}$$

where s is a factor ≥ 1, and r is the number of times the objective function has been reset.

3.6.2 Link Cost Resetting Procedure

The motivation to reset the objective function stems from a need to have the link costs for links $a \in A_m$, reflect the link volumes they carry. The objective function resetting procedure addresses this need.

Consider, for the sake of illustration, a case where the objective function is reset based on the link flows (x_{ij}^k) of the basic feasible solution at the beginning of the $r+1^{th}$ linear program. Such an approach has to be adopted with a great deal of caution because it does not guarantee the convergence of the program as the new objective may have a lower value for the basis corresponding to the previous iteration. Hence, we could have a situation wherein we continuously oscillate between the basis of the r^{th} and $r+1^{th}$ linear program. In fact, such an approach was tested (with the number of iterations before updating being fixed at 1) and failed as expected.

The above discussion serves to illustrate the impact of the objective resetting procedure on the convergence criterion. The proposed methodology for resetting the objective computes the link costs (for $a \in A_m$) based on the average of the link volumes associated with all the bases traversed by LPMLV between the r^{th} and $r+1^{th}$ resets and the previous values of the costs. Such an approach avoids infinite oscillations because, the moving average tends to smooth out the cost vector.

3.6.3 Updating Procedure for the Cost Vector $(c_B)^r$

Once the link cost vector has been determined at the beginning of the $r+1^{th}$ iteration, the path cost vector $(\hat{c}_{ij}^k)^r$ may be updated after recomputing the path costs $(p_{ij}^k.(\bar{c})^r)$for all the legitimate variables in the basis. It is important to note that even if the path cost of some legitimate variable remains the same, its cost in the objective function may change because the set of user equilibrium paths $(K_{ij})^r$ may change after resetting the link costs. Thus, after every reset of the objective function we need to check if $x_{ij}^k \in (K_{ij})^r$, and if so, then each $(\hat{c}_{ij}^k)^r$ = $(p_{ij}^k.(\bar{c})^r)$, else, $(\hat{c}_{ij}^k)^r$ = 2*$(p_{ij}^k.(\bar{c})^r)$. Once the path costs are reset, then one may easily reset $(c_B)^r$.

3.6.4 Dual Variable Updating Procedure

Once the cost vector is updated, the dual variables can be easily updated since B^{-1} remains the same after the reset. Thus we only need to multiply $(c_B)^r$ with B^{-1} to obtain the updated set of dual variables.

3.7 Solution Technique for LPMLV

The solution technique for model LPMLV is designed along the lines of the solution technique adopted for its predecessor model LP(TT). The procedure is elaborated here with the aim of pointing out the differences between the two models' solution techniques, as well as for the sake of completeness.

An explicit statement of the linear programming model LPMLV requires the enumeration of all possible paths between each O-D pair $(i, j) \in OD$. While this can be readily accomplished in theory, it is computationally prohibitive except for small sized networks. As an alternative, one can consider the explicit generation of only "efficient" paths in the sense of Dial [1971], that is, paths for which each link has its initial (FROM) node closer to the origin than its final (TO) node, and has its TO node closer to the destination than its FROM node. This approach, however, was found unsatisfactory based on tests conducted by Sivanandan

[1991], and so was not pursued. To overcome the foregoing difficulties, an algorithm is presented here in which the columns of LP are selectively generated only as and when needed within the framework of the revised simplex method, in order to optimally solve the problem. Such an algorithm is known as a column generation procedure (see Lasdon, 1970), but as explained below, requires a non-standard modification for our problem.

The procedure initializes with a basic feasible solution to LPMLV. One may use the artificial variables y^+ and Y^+ as the set of initial basic variables. Alternatively, an advanced start procedure as described by Sivanandan [1991] may be considered. To describe the main loop of this algorithm, suppose that at any iteration, we have a basic feasible solution $(\bar{x},\bar{y},\bar{Y})$ to LPMLV represented by a revised simplex tableau, having $\bar{\pi}$ and $\bar{\mu}$, as the corresponding vectors of simplex multipliers or complementary dual basic solution values $c_B B^{-1}$ associated with the constraints [3.5c] and [3.5d], respectively, where B is the available basis matrix and c_B is the vector of basic variable cost coefficients (for convenience in presentation, let $\bar{\mu}_{ij} \equiv 0 \forall (i,j) \varepsilon OD - \overline{OD}$). In order to ascertain the optimality of $(\bar{x},\bar{y},\bar{Y})$, we need to price the nonbasic variables, that is, we need to compute their reduced costs. For the artificial variables y_a^+ and y_a^- for

$a \in A_v$, and Y_{ij}^+, Y_{ij}^- for $(i,j) \in \overline{OD}$, this reduced cost is given by $(M - \overline{\pi}_a)$, $(M + \overline{\pi}_a)$, $(M_\sigma - \overline{\mu}_{ij})$, and $(M_\sigma + \overline{\mu}_{ij})$, respectively. If any of these is negative, then the corresponding artificial variable is pivoted into the basis, and the main step is repeated. Hence, suppose that none of these artificial variables are enterable into the basis.

We now need to price the x-variables. However, the columns of the x-variables are not all available in explicit form, and so we need to implicitly determine if any x_{ij}^k variable has a negative reduced cost. If we can conclude that no such variable exists, then we can declare optimality.

Toward this end, consider the following two shortest (simple) path problems defined for each $(i,j) \in OD$ (see Bazaraa et al., 1990, for a discussion on shortest path problems and related solution algorithms).

$$SP_{ij}^I : Minimize\{(M_2(\overline{c})^r - \overline{L}).p_{ij}^k : k = 1,\ldots n_{ij}\} \equiv (M_2(\overline{c})^r - \overline{L}).p_{ij}^{k*} \tag{3.7.a}$$

where $M_2 = \sum_{a \in A_v} |\overline{\pi}_a| + 1$, and $\overline{L}_a = \overline{\pi}_a \; \forall \, a \in A_v ; = 0 \; \forall \, a \in A_m$.

$$\text{Let } v_{ij}^1 = ((\overline{c})^r - \overline{L}).p_{ij}^{k*} - \overline{\mu}_{ij} \tag{3.7.b}$$

$$SP_{ij}^{II} : Minimize\{(M_1(\overline{c})^r - \overline{L}).p_{ij}^k : k = 1,\ldots n_{ij}\} \equiv (M_1(\overline{c})^r - \overline{L}).p_{ij}^{k**} \tag{3.8.a}$$

and let $v_{ij}^{2}=(M_1(\bar{c})^r-\bar{L}).p_{ij}^{k**}-\bar{\mu}_{ij}$. (3.8.b)

Note that SP_{ij}^{I} and SP_{ij}^{II} seek the shortest simple paths from i to j using respective link cost vectors $(M_2(\bar{c})^r-\bar{L})$ and $(M_1(\bar{c})^r-\bar{L})$. Lemmas 3.4 and 3.5 below provide the motivation for these problems.

Lemma 3.4. SP_{ij}^{I} defined by (3.7) is a shortest path problem with nonnegative cost coefficients and its solution yields a path $k^* \in (K_{ij})^r$, where, as described earlier, $(K_{ij})^r$ refers to the set of equilibrium paths from i to j with link costs for missing link volumes set after the r^{th} reset. Moreover, if $v_{ij}^{I} \geq 0$ in (3.7.b), then no x_{ij}^{k} for $k^* \in (K_{ij})^r$ is enterable into the current basis for LPMLV. Otherwise, $x_{ij}^{k^*}$ is enterable with a reduced cost of $v_{ij}^{I} < 0$.

Proof. Since $(\bar{c}_a)^r \geq 1 \forall a \in A$, we have by the definition of M_2 in (3.7.b) that $(M_2(\bar{c})^r-\bar{\pi})>0$. Furthermore, for any $k \in (K_{ij})^r$ and any $\bar{k} \in (\bar{K}_{ij})^r$, since $(\bar{c})^r$ is integral, we have from (3.2) that $\bar{t}.(p_{ij}^{\bar{k}}-p_{ij}^{k}) \geq 1$. Hence, we get

$$(M_2(\bar{c})^r-\bar{L}).p_{ij}^{\bar{k}}-(M_2(\bar{c})^r-\bar{L}).p_{ij}^{k} \geq M_2-\bar{L}(p_{ij}^{\bar{k}}-p_{ij}^{k}) \geq M_2-\sum_{a\in A}\left|\bar{L}_a\right| \geq 1 \quad \text{and so}$$

$k^* \in (K_{ij})^r$ in (3.7.a).

Now, for any $k \in (K_{ij})^r$, the reduced cost for x_{ij}^{k} is given from (3.5) by

$$(\bar{c})^r.p_{ij}^k - \bar{L}.p_{ij}^k - \bar{\mu}_{ij} \equiv ((\bar{c})^r - \bar{L}).p_{ij}^k - \bar{\mu}_{ij}. \qquad 3.9$$

Hence, if $v_{ij}^1 < 0$, then $x_{ij}^{k^*}$ is enterable into the current basis for LPMLV. On the other hand, if $v_{ij}^1 \geq 0$, then from (3.5), (3.9) and the fact that $(\bar{c})^r.p_{ij}^k = (\bar{c})^r.p_{ij}^{k^*}$ by the definition of $(K_{ij})^r$ in (3.2), we obtain that the reduced cost for $x_{ij}^k, k \in (K_{ij})^r$, is given by

$$(\bar{c})^r.p_{ij}^k - \bar{L}.p_{ij}^k - \bar{\mu}_{ij} = (\bar{c})^r.p_{ij}^k - \bar{L}.p_{ij}^k \;-\; (\bar{c})^r.p_{ij}^{k^*} - \bar{L}.p_{ij}^{k^*} \;+\; v_{ij}^1$$

$$= M_2(\bar{c})^r.p_{ij}^k - \bar{L}.p_{ij}^k \;-\; M_2(\bar{c})^r.p_{ij}^{k^*} - \bar{L}.p_{ij}^{k^*} \;+\; v_{ij}^1 \;\geq v_{ij}^1 \geq 0,$$

and hence no x_{ij}^k, $k \in (K_{ij})^r$ is enterable into the basis. This completes the proof. □

The above Lemma provides an implicit way of pricing the variables x_{ij}^k for $k \in (K_{ij})^r, (i,j) \in OD$. The following result addresses the pricing of the remaining variables x_{ij}^k for $k \in (\overline{K}_{ij})^r, (i,j) \in OD$.

Lemma 3.5. Suppose that $v_{ij}^1 \geq 0$ in (3.7) and that we solve SP_{ij}^{II} and obtain k^{**} and v_{ij}^2 as in (3.8.b)

(i) If $k^{**} \in (\overline{K}_{ij})^r$, and $v_{ij}^2 < 0$, then $x_{ij}^{k^{**}}$ is enterable into the current basis.

(ii) If $k^{**} \in (\overline{K}_{ij})^r$, and $v_{ij}^2 \geq 0$, then no x_{ij}^k for $k \in (\overline{K}_{ij})^r$ is enterable into the current basis.

(iii) If $k^{**} \in (K_{ij})^r$ then no $x_{ij}^k, k = 1,..n_{ij}$, is enterable into the current basis.

Proof. Noting that the reduced cost for any x_{ij}^k for $k \in (\overline{K}_{ij})^r$, is given by

$$M_1(\overline{c})^r . p_{ij}^k - \overline{L} . p_{ij}^k - \overline{\mu}_{ij} \equiv (M_1(c_{ij}^k)^r - \overline{L}) . p_{ij}^k - \overline{\mu}_{ij} \quad (3.10)$$

cases (i) and (ii) follow directly from (3.7.b). Hence, consider case (iii).

Since $v_{ij}^1 \geq 0$, we have by Lemma 3.4 and (3.9) that $v_{ij}^3 = ((\overline{c})^r - \overline{L}) . p_{ij}^{k**} - \overline{\mu}_{ij} \geq 0$. Hence, for any $k \in (\overline{K}_{ij})^r$, we obtain from (3.8) and (3.10) that the reduced cost for x_{ij}^k is given by

$$(M_1(\overline{c})^r - \overline{L}) p_{ij}^k - \overline{\mu}_{ij} \geq (M_1(\overline{c})^r - \overline{L}) . p_{ij}^{k**} - \overline{\mu}_{ij} \geq ((\overline{c})^r - \overline{L}) . p_{ij}^{k**} - \overline{\mu}_{ij} = v_{ij}^3 \geq 0.$$

Moreover, since we also have by Lemma 3.4 that no x_{ij}^k, $k \in (K_{ij})^r$, is enterable into the basis, this completes the proof. □

Remark (3.3). In order to use Lemma 3.5, we need to check whether the path k^{**} belongs to $(K_{ij})^r$ or to $(\overline{K}_{ij})^r$. From (3.2), we simply need to check if $c_{ij}^{k**} = (\overline{c})^r . p_{ij}^{k**}$ equals $c_{ij}^* = min\{(\overline{c})^r . p_{ij}^k, k = 1, \ldots, n_{ij}\}$ or exceeds it. The quantities c_{ij}^* may be computed once at the start of **each reset** of the algorithm by solving $|O|$ shortest path problems having nonnegative link coefficients $\overline{t}_a$, $a \in A$, one for determining the shortest paths from each $i \in D$ to all $j \in D$. This can be accomplished in $O(|O||N|^2)$ time.

Remark (3.4). The shortest path problem SP_{ij}^{II} has mixed-sign cost coefficients, and so might contain negative cost circuits that are reachable from node *i*. In such a case, the problem of finding a shortest simple (no loop) path is NP-complete. Hence, an enumerative branch-and -bound routine would need to be adopted to solve SP_{ij}^{II} whenever the shortest path routine encounters a negative cost circuit as in Bazaraa et al. [1990]. Alternatively, to avoid excessive effort for large sized networks, the following heuristic may be adopted. Consider the formulation of a minimum cost network flow programming problem, with the link cost vector given by $(M_1(\bar{c})^r - \bar{L})$, and with a supply of $|D|$ at node *i*, a demand of one at each node $j \in D$, and an upper bound of $|D|$ on the flow on each link $a \in A$. Assuming without loss of generality that all nodes in D are reachable from node *i*, since this a feasible and bounded network flow programming problem, an optimal solution can be obtained very efficiently (we used the routine RELAXTII developed by Bertsekas and Tseng [1988], for this purpose). In the case of negative cost circuits, the solution might include loop flows within such circuits in addition to path flows from *i* to each of the nodes $j \in D$, whence the corresponding paths might only be near optimal shortest simple paths. Nonetheless, these simple paths may be identified, along with any existing loop flows, using a backtracing routine as presented by Sherali et al. [1994]. Assuming these to be the shortest simple paths, we can now apply Lemma 3.5.

3.8 Summary of the Column Generation Algorithm (CGA) for Solving LPMLV

Initialization *(Starting Basic Feasible Solution)*: For each $i \in O$, find the shortest paths to all $j \in D$ using the observed costs for links with known volumes and free flow costs for links with missing volumes to comprise the link cost vector. Hence, determine $(c_{ij}^{*})^{r} \equiv \min imum\{(c_{ij}^{k})^{r} : k = 1,\ldots n_{ij}\}\}$ for each $(i, j) \in OD$. Let $M_1 = 2$, select a value for *M* as discussed in Section 3.3 and let $M_\sigma = M\sigma$, where $0 \le \sigma \le 1$ is selected as in Section 3.3. Find an initial basic feasible solution $(\bar{x}, \bar{y}, \bar{Y})$ to LP(TT) as outlined above, or simply use an all artificial basis for this purpose. Let $(\bar{\pi}, \bar{\mu})$ be the corresponding set of simplex multipliers (complementary dual solution) $c_B B^{-1}$, where *B* is the basis matrix corresponding to $(\bar{x}, \bar{y}, \bar{Y})$ and c_B is the vector of objective coefficients for the basic variables. Construct the initial revised simplex tableau, initialize the counters K=1 (outer loop counter), $r = 1$ (update counter), s=1 (termination counter), $K(r) = K$, and proceed to Step 1.

Step 1. *(Computing Objective Function resetting Criterion):* Compute the bound on the maximum decrease in the objective function as

$$dz^{criterion} = S_r\left(\sum_a |\bar{\pi}_a| \; + \; \max|\bar{\mu}_{ij}| \;\right) * R_{max}$$

Set $z^r = c_B B^{-1} b$.

Step 2. *(Pricing Artificial Variables $y^{\pm}$ and $Y^{\pm}$)*: Compute $v^{0+} = M - max\{\overline{\pi}_a : a \in A_v\}$, $v^{0-} = M + min\{\overline{\pi}_a : a \in A_v\}$ $V^{0+} = M_\sigma - max\{\overline{\mu}_{ij} : (i,j) \in \overline{OD}\}$ and $V^{0-} = M_\sigma + min\{\overline{\mu}_{ij} : (i,j) \in \overline{OD}\}$. If $v^0 = minimum\{v^{0+}, v^{0-}, V^{0+}, V^{0-}\} \geq 0$, then proceed to Step 2. Otherwise, if $v^0 = v^{0+} \equiv M - \overline{\pi}_{a*}$, set the entering variable $z_K = y^+_{a*}$, its reduced cost $v_K = v^0$, and its column $P_K = e_{a*}$ (which is a unit vector of length $|A| + |\overline{OD}|$ with the element one corresponding to position $a * \varepsilon A_v$), and go to Step 5. Similarly, if $v^0 = v^{0-} \equiv M + \overline{\pi}_{a*}$, set $z_K = y^-_{a*}, v_K = v^0, P_K = -e_{a*}$, and go to Step 4. On the other hand, if $v^0 = V^{0+} \equiv M_\sigma - \overline{\mu}_{(ij)*}$, then set $z_k = Y^+_{(ij)*}, v_K = v^0$, and $P_K = e_{(ij)*}$ (which is again a unit vector of length $|A_v| + |\overline{OD}|$ with the element one corresponding to position $(ij)^* \in \overline{OD}$), and proceed to Step 4. Finally, if $v^0 = V^{0-} \equiv M_\sigma + \overline{\mu}_{(ij)*}$, then

position $(ij)^* \in \overline{OD}$), and proceed to Step 4. Finally, if $v^0 = V^{0-} \equiv M_\sigma + \overline{\mu}_{(ij)*}$, then

set $z_K = Y^-_{(ij)*}, v_D = v^0, P_K = -e_{(ij)*}$, and go to Step 5.

Step 3. *(Pricing x^k_{ij} Variables for $k\varepsilon\ (K_{ij})^r, (i,j)\varepsilon\ OD$)*: For each $i\varepsilon\ O$, solve the shortest path problem SP^I_{ij} given by (3.7.a - 3.7.c) for all $j \in D$. Compute $(ij)^* = \arg\min\{v^1_{ij} : (i,j)\varepsilon OD\}$. If $v^1_{(ij)*} \equiv ((\overline{c})^r - \overline{\pi}).p^{k*}_{(ij)*} - \overline{\mu}_{(ij)*} < 0$, put $z_K = x^{k*}_{(ij)*}, v_K = v^1_{i(ij)*}$, $P_K = p^{k*}_{(ij)*} + e_{(ij)^*}$, and go to Step 5. Otherwise, proceed to Step 4.

Step 4. *(Pricing x^k_{ij} Variables for $k\varepsilon\ (\overline{K}_{ij})^r, (i,j)\varepsilon OD$ and Termination Check)*: For each $i \in O$, solve the shortest (simple) path problems SP^{II}_{ij} given by (3.7a) for all $j \in D$ as in Remark 3.4. Compute $(ij)^* = arg\,min\{v^2_{ij} : (i,j)\varepsilon OD$ and $k^{**}\varepsilon \overline{K}_{ij}$ in (3.8)}. If $v^2_{(ij)*} < 0$, put $z_K = x^{k**}_{(ij)*}, v_K = v^2_{(ij)*}, P_K = p^{k**}_{(ij)*}$, and proceed to

Step 5. Ohterwise, if *(ij)** does not exist or if $v^2_{(ij)*} \geq 0$, the current solution is optimal for the current cost vector ($(c_B)^r$ of LPMLV). Set termination counter s=s+1. If s > 3 terminate. Else go to step 8.

Step 5. *(Update Primal and Dual Solutions)*: Determine the updated column $B^{-1}P_K$ of the (entering) nonbasic variable z_K having a reduced cost $v_K < 0$, pivot z_K into the basis, and hence update the basic feasible solution, the basis inverse, and the dual vector $(\overline{\pi}, \overline{\mu})$ of simplex multipliers.

Step 6. (Compute Moving Average of Link Volumes for Links with Unknown Volumes): First compute the volumes for links with unknown volumes, corresponding to the present corner point as $f_a^K = \sum \{x_{ij}^k (p_{ij}^k . e_a) : x_{ij}$ is basic$\}$ $\forall a \in A_m$. Compute the grand moving average for the link $a \in A_m$ as $f_{av}^K = f_{av}^{K-1}(K - K(r)) + f_a^K / K - K(r) + 1$.

Step 7. *(Checking whether objective has to be reset):* Compute $dz = z^r - z^K$, where z^r is the objective value at the last reset, and z^K its current value. If $dz < dz^{criterion}$ then go to step 2.

Step 8. *(Resetting Link cost vector):* Set $r = r + 1$ and let $K(r) = K$. Set $(c_a)^r$ to be a function of f_{av}^K, $\forall \mathrm{a} \in \mathrm{A_m}$.

Step 9. *(Computing $(c_B)^r$ and Resetting Dual Variable Values):* For each $i \in O$, find the shortest paths to all $j \varepsilon D$ using the observed costs for links with known volumes and $(c_a)^r$ for links with missing volumes to comprise the link cost vector. Hence, determine $(c_{ij}^*)^r \equiv \min\{(\overline{c})^r . p_{ij}^k : k = 1, \ldots n_{ij}\}\}$ for each $(i, j) \in OD$.

Scan the basis row by row. If the variable in row i is an artificial variable then $(c_B^i)^r = M$, or M_σ, as the case might be. Else we now have to compute the costs for the legitimate variables. For $x_{ij}^k \in B$, if $(\bar{c})^r . p_{ij}^k = (c_{ij}^*)^r$, then $(c_B^i)^r = (c_{ij}^*)^r$ else $(c_B^i)^r = M_1(\bar{c})^r . p_{ij}^k$. Set $(\bar{\pi}, \bar{\mu}) = (c_B)^r B^{-1}$. Set $z^r = z^K$. Reset $S_r = \dfrac{1}{1+se^{-r}}$.

Set K=K+1. Compute the objective function resetting criterion $dz^{criterion}$. Go to Step 2.

4. EVALUATION OF MODEL USING ARTIFICIAL NETWORK

4.1 Introduction

The previous chapter dealt with the mathematical formulation and subsequent development of the solution technique for a linear programming based approach for determining traveler origin-destination trip tables using incomplete network link flow specification. This chapter deals with the validation of this model. The validation procedure includes a comprehensive testing of the model's performance, sensitivity to the influencing input factors, and comparative evaluation with the maximum entropy approach.

The model LPMLV can be used to determine an OD trip table that tends to conform to a user equilibrium solution (if it exists) which reproduces the partial set of link flows to as close an extent as possible while being as consistent as possible with target information on trip patterns provided through a target table. The target trip table is primarily used by the model to guide the solution and to discriminate among the alternate optimal solutions for the user equilibrium objective.

The influence of missing link volume information is presented in detail in this section. As discussed earlier, the link volume counts are the most important input for the OD estimation models, because the solution trip table is designed primarily to be consistent with these measurements. As the number of link volume measurements increases, the output trip table is expected to be consistent with a greater number of measures of travel patterns, and if these measurements are correct and meaningful, one would expect the output trip table to get increasingly better. Since in practice, link volumes on all links are not always known, a knowledge of the model sensitivity to this aspect is very useful in modeling the travel characteristics of real networks. This chapter presents an extensive discussion for the sensitivity of the linear programming and maximum entropy approaches. Like most other approaches for OD estimation these models are expected to have their respective output tables being influenced by the target trip table, thus making it necessary to study the influence of the target trip tables on them. Presented in this chapter are test results on the sensitivity of these models to the target trip table provided.

4.2 Evaluation Procedure

The evaluation procedure is designed to validate the LPMLV model, and to measure its performance in estimating OD tables vis-à-vis the maximum entropy model. In addition, the sensitivity of the linear programming and maximum entropy models to varying target tables and available link volumes are also studied. In order to do so, the evaluation procedure requires a "correct" trip table and a complete set of link volumes. The percentage of link volumes that are assumed to be known, are however varied for each case. Similarly, the target trip tables used are different for each case. A spectrum of target tables, ranging from those that are very close to the correct trip table to those that are far different were used.

Detailed explanations of the evaluation procedure, measures of effectiveness and test results are presented in the sections that follow.

4.2.1 Evaluation Factors

4.2.1.1 Available Link Volumes

Any basic feasible solution to the linear programming approach represents a trip table that satisfies the constraints defined by the link volumes. The program then identifies the trip table that most closely replicates the available link volumes and which comes closest to representing a user equilibrium solution, while being as close as possible to the target trip table. The maximum entropy approach first assigns the target trip table and continually modifies this table. Then it picks the table that is the closest to the observed link volumes.

Thus the extent of available link volumes serve to largely determine the solution space for the linear programming and maximum entropy approaches. This input serves to drive the solution in both these models. The networks used for purposes of evaluation had known volume data for all the links. However, a partial set of these link volume counts were used as input for each model, and the performance of the models were evaluated.

4.2.1.2 Target Trip Tables

While, both linear programming and maximum entropy approaches use a target trip table to guide the solution, they differ in the mechanism by which the target table is used. The maximum entropy approach modifies the target trip table cell values, and maximizes a non-linear entropy objective function so as to replicate observed link volumes. The output table thus obtained is one with the maximum number of microstates associated with it. The solution of the maximum entropy approach is known to be sensitive to the correctness of the ratio of each trip interchange to the sum of all trip interchanges specified in the target trip table. Also, the traveler route choice proportions as determined by the assignment of the target trip table, and hence the value of each trip interchange, influence the solution table of the maximum entropy approach. The linear programming approach, on the other hand, constrains the output trip table to be as close as possible to the target trip table. Thus the target trip table serves to reduce the underspecification problem by a different mechanism, for each of the two approaches discussed above. Also, it must be noted that the choice of penalty parameter for prior trip table deviations may cause some of the output trip table solutions to not conform to a set of user equilibrium path flows.

The tests described in the following section employ a set of target trip tables, that vary from being very close to the "true" trip table, to being very different from it. Both approaches are tested with a given target trip table for a set of cases of available link volumes, and their performances are evaluated using a set of measures of effectiveness.

4.2.2 Comparison of Results

In order to compare the test results from the two approaches, two measures of closeness are used in judging the results. The first is based on the replication of the link volumes by the solution trip table when assigned to the network. This is accomplished by comparing the output link volumes obtained from an equilibrium assignment with the observed volumes. For both the linear programming and maximum entropy approaches, the link volumes corresponding to the output trip tables are computed within the software programs developed for these approaches. The second measure is the closeness of the estimated trip table to the "correct" trip table. These two criteria are obvious choices since the objective of trip table estimation is to determine a trip table that replicates observed link volumes when assigned, and is as close as possible to the "correct" trip table.

4.2.2.1 Replication of Observed Link Volumes

The most important measure of the quality of the trip table is its ability to replicate observed volumes on the links of the network. The link volumes corresponding to the output trip table can be obtained for both approaches a byproduct of the estimation process. The Percentage Root Mean Square Error (% RMSE) and Percentage Mean Absolute Error (% MAE) are chosen as measures of error rate to compare the observed volumes with the assigned volumes. These measures are defined as follows:

$$\%\ RMSE = \sqrt{\frac{\sum_{a \in A_v} (V^a_{assign} - V^a_{obs})^2}{n}} * \frac{100 * n}{\sum_{a \in A_v} V^a_{obs}}$$

$$\% \ MAE = \frac{\sum_{a \in A_v} |V^a_{assign} - V^a_{obs}|}{n} * \frac{100 * n}{\sum_{a \in A_v} V^a_{obs}}$$

where

V^a_{assign} = equilibrium assigned volume on link a,

V^a_{obs} = observed volume on link a, and

n= number of links with available volumes.

4.2.2.2 Closeness of Estimated Table to True Trip Table

There are various measures of closeness for comparing trip matrices. Smith and Hutchinson (1981) evaluate different goodness of fit statistics for trip distribution matrices and conclude that the phi-statistic (ϕ) is one of the most appropriate to test the goodness of fit of alternate trip distribution models. The percentage mean absolute error statistics has also been reported as an useful indicator. Consequently the percentage RMSE, MAE and (ϕ) are used in the following analysis for trip table comparisons. These measures of closeness are defined below:

$$\% \ RMSE = \sqrt{\frac{\sum (t_{ij} - t^*_{ij})^2}{n_{OD}}} * \frac{100 * n_{OD}}{\sum t^*_{ij}}$$

$$\% \ MAE = \frac{\sum |t_{ij} - t^*_{ij}|}{n_{OD}} * \frac{100 * n_{OD}}{\sum t^*_{ij}}$$

$$\phi = \sum max(1, t^*_{ij}) \left| ln \frac{max(1, t^*_{ij})}{max(1, t_{ij})} \right|$$

where t^*_{ij} is the “correct” number of trips for interchange (i, j), t_{ij} is the estimated number of trips for interchange (i, j), and n_{OD} is the number of feasible O-D interchanges.

4.2.2.3 Computer Resources for Tests

The tests reported in the following section were obtained using an IBM-PC 486 with a clock speed of 66 MHz.

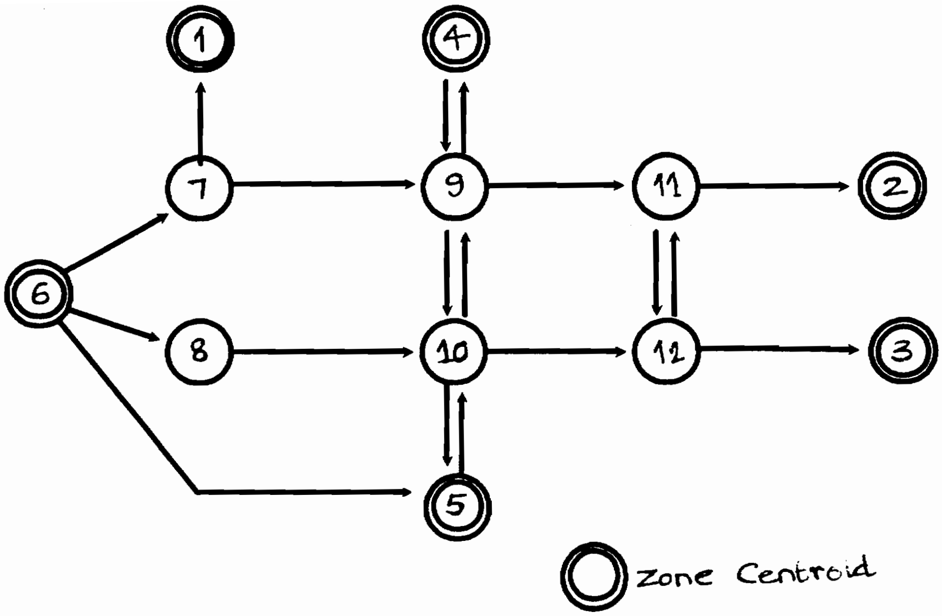
4.3 Test Networks

This thesis presents results on tests carried out on two networks. The first network is one that was artificially created for the purposes of OD estimation tests, and the validation results are reported in this chapter. Chapter 5 deals with test results for a real network. The real network represents a one square mile area in the town of West Lafayette, Indiana, and is described in detail in the next chapter.

Presented below are descriptions and results of the tests conducted on the sample network.

4.3.1 Artificial Network

The artificial network chosen for the tests is a hypothetical network called the "corridor network", reflecting a travel corridor (Gur et al., 1980). This network has been used extensively for tests while developing the LINKOD system of models. The network consists of 6 zones, 6 intersection nodes and 18 links. Although it is a hypothetical network, it has multiple routes between O-D pairs and multiple equilibrium solutions. Due to its small size it is ideal for performing extensive tests, without having to expend too much computational effort. This network is shown in Figure 4-1.



Link Number	Begining Node	Ending Node	Observed Volume
1	4	9	2400
2	5	10	2000
3	6	5	100
4	6	7	5000
5	6	8	500
6	7	1	500
7	7	9	4500
8	8	10	500
9	9	4	2000
10	9	10	1500
11	9	11	4900
12	10	5	1600
13	10	9	1500
14	10	12	900
15	11	2	4800
16	11	12	300
17	12	3	1000
18	12	11	200

Fig 4-1. Test Network-1 Characteristics

4.3.1.1 Target Trip Tables

The target trip tables used in the tests conducted on this network are primarily of four types. Using these four types of trip tables 15 trip tables for the tests were derived. These trip tables are described below, in reverse order of their closeness to the true trip table.

The structural trip table represents one which has 0-1 values for its cells to indicate merely if that trip interchange is feasible or not. The no target information trip table has a uniform value of 983 for all feasible interchanges. The third type of trip table used as a target was a relatively close trip table. This trip table has cell values that are quite close to the actual trip table. The correct trip table for the network was another type used in the tests.

The tests reported here, use target tables that extract different extents of information the four basic target tables described above. The extent to which information from a given type of target trip table is used varies from nearly 50%-100% of the cells of a particular table type being used as a target value.

4.4 Test Results for Artificial Network

4.4.1 Case 1: Structural Table as a Target

The target table used for this case is shown in Table 4-1 below.

Table 4-1: Target Trip Table for Case 1

From\To	1	2	3	4	5	6
4	0	1	1	0	1	0
5	0	1	1	1	0	0
6	1	1	1	1	1	0

Table 4-2 below shows the measures of errors that describe the performance of model LPMLV with the structural trip table as target. The RMSE of the deviations from the correct trip table values vary from 31%-128%, which is to be expected when the target table contains almost no information. The mean absolute error percentages for trip table deviations also are high, as expected. While increasing the availability of link volume information, in general, seems to decrease the measures of error rates, the test results show that a statement on the monotonicity of the decrease cannot be made. While the decrease of the measures of error rates, with increasing knowledge of link volumes, is perceptible for the MAE and RMSE statistics, it is not as dramatically visible as the φ statistic's variation. The largest error rate as measured by this statistic, occurs when only 50% of the link volumes are available. Increasing the availability of link volumes rapidly decreases this error rate, as seen in Table 4-2. Another observation that is noteworthy, is that the available link volumes are perfectly replicated for all six cases of available volumes, as indicated by the zero values for the MAE% and RMSE %.

Table 4-2: Performance of the Linear Programming Approach (Structural Table as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	127.56	92.96	35472	0	0
60	103.99	68.97	21332	0	0
70	102.57	76.96	21608	0	0
80	97.09	69.00	8590	0	0
90	102.98	76.00	9192	0	0
100	30.22	19.04	2140	0	0

In contrast with the performance of the linear programming approach, the maximum entropy approach performs much better for the structural target table input, as seen in Table 4-3 below:. The RMSE, MAE and φ for trip table deviations are significantly lower for most cases, now ranging only between 43%-52%, 39%-46% and 3728-4536 respectively. However, unlike the linear programming approach, the replication of observed volumes is not perfect for any case of available link volume proportions . In fact, for the 50% and 60% available link volume cases, the RMSE and MAE for link volume deviations are significantly high, and range between 28%-39% and 22%-37% respectively.

Table 4-3: Performance of the Maximum Entropy Approach (Structural Table as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	48.56	39.87	3728	38.97	37.02
60	51.56	45.91	4536	28.47	22.43
70	48.75	43.08	4191	4.77	3.66
80	49.09	43.67	4233	5.93	4.31
90	51.29	44.31	4265	2.82	1.73
100	43.73	38.69	3754	2.86	1.71

The 50% link volume case is particularly noteworthy. For this case, the deviations of the output trip table from the correct trip table are one of the least, among the other cases of available link volumes. However, the replication of link volumes is the worst among the other cases, with the RMSE and MAE being as high as 39% and 37% respectively. Also note that there is no monotonic behavior of any of the error statistics with respect to available link volume percentages.

The above results lead to the following conclusions. The maximum entropy model yields superior results for the 70%, 80%, and 90% available link volume cases. This conclusion is based on the facts that the trip table error statistics are significantly lower for the maximum entropy approach and that its link volume replication is quite good. However, for the 50% and 60% available link volume cases, this model is unable to replicate observed link volumes satisfactorily. In contrast, the linear programming approach performs exceptionally well for all the cases of available link volumes, with respect to link volume replication. The model replicates link volumes exactly with zero error. However, the error statistics for the output trip tables show that its performance in general, is not as good as that for the maximum entropy model. It must be added here, that the applicability of either model for the 50% and 60% link volumes is questionable because of the high values for trip table errors for the linear programming approach and unsatisfactory replication of observed link volumes by the maximum entropy approach. Thus in the context of the two performance criteria, namely the closeness of the output trip table to the "correct" trip table, and replication of observed link volumes, neither model is consistently superior to the other. On the other hand, if the models are to be purely judged on the basis of the quality of output trip table, then the maximum entropy model results are superior for most cases. However, note that for the 100% link volume availability the linear programming approach is superior.

4.4.2 Case 2 - No Information Based Trip Tables as Targets

A no information based target trip table is one in which all feasible interchanges carry an equal number of trips. Four no information trip tables were used as target trip tables for the tests discussed in this section. Each feasible interchange in all the four trip tables carry the same number of trips. The tables only differ in the number of cells that contain such information. The number of cells with information varied from 5,7,9 and 11 (which corresponds to the total number of feasible trip interchanges). The idea behind the tests is of course to test the sensitivity of these models to the extent of information provided in the trip table. The trip tables with 5 cells approximately represents a case where 45% of the cells contain information. And the other tables correspond to 64%, 82%, and 100% cells containing information. These target tables, and the tests conducted with them are described in greater detail in the sections to follow.

4.4.2.1 Case 2a: Partial No Information Based Trip Table (45% Cells with Info.)

The target trip table for case 2a is shown below:

Table 4-4: No Information Based Trip Table (45% Cells with Info.)

From\To	1	2	3	4	5	6
4	0	983	0	0	0	0
5	0	983	0	983	0	0
6	0	983	0	983	0	0

Table 4-5, shows the performance of the linear programming approach with the no information target table shown in Table 4-4. The target trip table for case 2a is shown below.

Similar to the case of the structural table, the performance of the model, is generally seen to improve with increasing information on link volumes. The RMSE error rate for the output trip table ranges from (64%-103.67%), which is lower than it was for Case 1, and seems to improve with an increase in link volume information. The MAE and ϕ deviations follow a similar trend while varying from 48%-84% and 4294-10920, respectively.

Table 4-5: Performance of the Linear Programming Approach (No Information Based Trip Table (45% Cells with Info.) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	103.67	84.34	10920	0	0
60	63.6	53.32	8812	0	0
70	63.73	53.32	8502	0	0
80	67.00	53.32	8609	0	0
90	64.02	47.66	4294	0	0
100	64.02	47.66	4294	0	0

An anomaly in the performance of the model can be seen to occur with 80% available volumes. One would expect the error statistics with 80% volumes to be lesser than those obtained with 70% available link volumes. However, it is seen that they are much higher than those obtained with 70% available volumes. In fact, the φ value with 80% link volumes is even higher than the corresponding value obtained with only 50% link volumes. Note however, that the replication of available link volumes is perfect for all the cases shown in Table 4-5. One possible explanation for the anomalous behavior could be the existence of multiple optimal solutions to the objective of model LPMLV.

When the link volume specification is not complete, there are lesser restrictions on varying Y_{ij}^{+} and Y_{ij}^{-} to obtain an output trip table while keeping $e.y$, the route cost component of LPMLV's objective, and $\sum_{ij \in \overline{OD}} Y_{ij}^{+} + Y_{ij}^{-}$, constant. This may increase the number of feasible bases that can be optimal.

The performance of the maximum entropy approach for this case of target information, is shown in Table 4-6.

Table 4-6: Performance of the Maximum Entropy Approach
(No Information Based Trip Table (45% Cells with Info.) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)		RMSE % (Vol)	MAE % (Vol)
50	79.79	61.08	16463	24.27	20.35
60	84.77	69.76	16888	30.52	26.11
70	64.47	56.76	13107	29.5	24.24
80	40.71	37.49	4673	6.36	4.39
90	45.5	39.57	5669	9.92	6.34
100	60.06	53.07	6775	9.74	5.45

One would expect the maximum entropy approach to yield better results for this case of the target trip table than it did for the structural target trip table case. However, the results are to the contrary. For every case of available link volumes, the structural table input yielded lesser values for trip table errors. The RMSE, MAE and φ statistics for trip table errors now range between 41%-85%, 37%-70%, and 4673-16463, respectively. These error measurements are much higher than the corresponding values obtained for Case 1. Also, the link volume replication error statistics are higher for this case except when only 50% link volume information is available.

Several interesting observations can be made, comparing the results of the linear programming and maximum entropy approaches. First, note that the replication of available volumes is much superior for the linear programming approach. In contrast, for the maximum entropy approach, the RMSE and MAE statistics for link volume replication range from 6%-31% and 4%-26%, respectively. The trip table error statistics for this case do not clearly indicate the superior performance of one model over another. The RMSE (TT) and MAE (TT) are lower for the maximum entropy approaches for all but the cases of 60% and 70% volume availability. In contrast, the φ error statistic is lower for the linear programming approach for all cases except when only 80% of the link volumes are known.

If one were to consider the three cases of available volumes (50%,60%,70%) then one may conclude that the linear programming approach performs better although the RMSE(TT) and MAE(TT) are higher than the corresponding values for the maximum entropy approach. Such a claim is justifiable because the error statistics for link volume replication, and φ, are much higher for the maximum entropy approach. The φ statistic, according to Smith and Hutchinson [1981], is a much better measure than RMSE(TT) and MAE(TT) for trip table replication. Hence the above conclusion is in order. On the other hand for the 80%,90% and 100% available link volumes, one may conclude that the maximum entropy approach performs better despite the φ, RMSE (Vol) and MAE (Vol) being slightly higher than the linear programming approach. Such a conclusion can be made, because the RMSE(TT) and MAE(TT) are much lower for the maximum entropy approach.

However, the applicability of either model for such a target trip table still remains questionable in the light of the high error values obtained.

4.4.2.2 Case 2b: Partial No Information Based Trip Table (64% Cells with Info.)

The target trip table for Case 2b is shown in Table 4-7 below:

Table 4-7: No Information Based Trip Table (64% Cells with Info.)

From\To	1	2	3	4	5	6
4	0	983	983	0	0	0
5	0	983	0	983	0	0
6	0	983	0	983	983	0

The addition of two more cells of uniform value 983 improves the performance of model LPMLV as reflected in the error rate values shown in Table 4-8 below. While there are some cases when a particular error rate is higher than the corresponding value in Case 2a, there is no instance when all the three measures for trip table errors are higher for Case 2b. On the other hand, the 80% volume test with this target table is a representative case where all the measures of errors are lower for Case 2b. The RMSE variation for trip table deviations for this test vary from 58% to 107%. The MAE and φ variations range from 36%-84% and 3262-7206, respectively. While the RMSE and MAE statistics for the output trip table are not low, it is only to be expected since the target trip table really does not contain any significant information on travel patterns. Also, note again that there is no monotonic decrease in the error statistics with increasing link volume information. As explained for Case 2a, the anomaly may be attributed to the existence of multiple optimal solutions.

Table 4-8: Performance of the Linear Programming Approach (No Information Based Trip Table (64% Cells with Info.) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	106.6	83.66	7206	0	0
60	62.37	52.98	6242	0	0
70	62.19	52.64	5116	0	0
80	66.96	52.64	5218	0	0
90	63.79	46.98	4428	0	0
100	57.99	36.34	3262	0	0

The performance of the maximum entropy approach for the target trip table in Table 4-7 is shown below:

Table 4-9: Performance of the Maximum Entropy Approach (No Information Based Trip Table (64% Cells with Info.) as a Target)

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	66.6	53.08	4428	33.05	29.17
60	55.24	55.11	12739	24.25	20.03
70	62.2	53.34	12714	28.62	21.67
80	40.71	33.39	4379	8.1	5.72
90	37.30	32.70	3201	4.24	2.65
100	48.33	43.14	4116	4.47	2.64

The measures of trip table errors for the maximum entropy approach, are lower for this case than for Case 2a. In fact, for every case of percentage available volumes these errors are lesser than the previous case. The RMSE, MAE and φ statistics for trip table errors now range between 33%-54%, 27%-45%, and 3201-12739, respectively. These error measurements are lower than the corresponding values of 36%-70%, 31%-58%, and 4673-16463 obtained for Case 2a. The link volume replication error statistics are not always lower for Case 2b. However, except for the 50% and 80% link volume availability case, the link volume errors are lower for Case 2b.

A comparison of the two approaches allows us to draw conclusions similar to that made for the previous case . For the 80%-100% volume availability range, one may conclude that the performance of the maximum entropy approach is superior. Although the available link volume replication for the maximum entropy approach for these three cases is not exact, the measures of errors RMSE(Vol) and MAE(Vol) are less than 9% and 6%, respectively. The error statistics for trip table closeness are lower for the maximum entropy approach for these three cases of available link volumes.

For the 60% and 70% link volume available cases, all the error statistics are much lower for the linear programming approach, than for the maximum entropy approach. For the 50% link volume availability case, the correct trip table replication error statistics for the linear programming approach are over 1.5 times the corresponding values for the maximum entropy approach. However, the RMSE(Vol) and MAE(Vol) for the maximum entropy approach are over 33% and 29% respectively.

4.4.2.3 Case 2c: Partial No Information Based Trip Table (82% Cells with Info.)

The target trip table for this case is shown below:

Table 4-10: No Information Based Trip Table (82% Cells with Info.)

From\To	1	2	3	4	5	6
4	0	983	983	0	983	0
5	0	983	983	983	0	0
6	0	983	0	983	953	0

Compared to case 2b, this target table, in general, seems to improve the performance of the model LPMLV, as seen in Table 4-11 below. The RMSE of the trip table deviations now lie between 48%-104% as compared to Case 2b where the variation was between 62%-106%. Similarly, the MAE and φ of trip table deviations now lie between 36%-78% and 3203-6409, respectively, which are lower than the corresponding ranges obtained with Case 2b. However, the improvement in performance cannot be claimed for every case of available link volumes. For the 90% and 100% volume cases, the performance for both cases (2b and 2c) are identical. In other words, the addition of two new constraints did not change the optimality of the solution for these three cases. In this case again, if we disregard the results with 80% available link volume, one may claim that the error rates decrease with increasing available link volumes.

Table 4-11: Performance of the Linear Programming Approach (No Information Based Trip Table (82% Cells with Info.) as a Target)

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	104.31	78.00	7206	0	0
60	58.18	46.98	4313	0	0
70	58.18	46.98	4313	0	0
80	63.25	46.98	7224	0	0
90	63.79	46.98	4428	0	0
100	47.99	36.34	4428	0	0

The performance of the maximum entropy approach for the same target table is shown in Table 4-12 below.

Table 4-12: Performance of the Maximum Entropy Approach (No Information Based Trip Table (82% Cells with Info.) as a Target)

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	47.81	39.41	3731	23.74	20.56
60	51.51	45.91	4536	12.79	10.08
70	48.64	42.92	4060	5.32	4.13
80	49.09	43.66	4231	5.93	4.31
90	52.08	44.88	4320	2.8	1.73
100	49.09	44.31	4265	2.99	1.71

One may note that the measures of error for link volume replication are lesser than the corresponding values obtained for Case 2b. The trip table error statistics show interesting trends. The range for the RMSE(TT), MAE(TT) and φ

statistics are lower varying from 47%-51%, 39-46%, and 3731-4536, respectively. Compared with Case 2b, one may note that while all three measures of error decrease considerably for the 50%, 60% and 70% available volume cases, they increase for the 80%, 90% and 100% cases.

For Case 2c, all measures of error for trip table comparisons, in general, are lower for the maximum entropy approach, as compared to the linear programming model. While the link volume errors are 0 for the linear programming approaches, they are quite small for the maximum entropy approach, except for the 50% available link volume case. The RMSE (Vol) and MAE (Vol) for the maximum entropy approach range from 3%-13% and 2%-10% when the available link volumes are higher than 50%. These measures are as high as 24% and 21%, respectively, when the link volumes available equal 50%.

The results, thus, suggest that the maximum entropy approach performs much better for this case of the target trip table except for the 50% link volume availability case.

4.4.2.4 Case 2d: Partial No Information Based Trip Table (100% Cells with Info.)

The target trip table for Case 2d is shown below:

Table 4-13: No Information Based Trip Table (100% Cells with Info.)

From\To	1	2	3	4	5	6
4	0	983	983	0	983	0
5	0	983	983	983	0	0
6	983	983	983	983	953	0

The target table shown in Table 4-13 has a uniform cell value of 983 for all the feasible trip interchanges. In a sense the trip table is complete because now all feasible trip interchanges have a target value. Table 414 below shows the test results obtained. The values for the errors are the lowest as compared to cases (2a, 2b, 2c) for all cases of available link volumes (except 70%). The anomaly may again be attributed to the presence of multiple optimal solutions. The RMSE deviations range between 46%-66%, which correspond to the lowest upper bound values for the ranges among all the cases described so far. Similarly the MAE and φ values have the smallest range and the least lower and upper limits of the range. In general, one may infer a gradual improvement in the quality of results for the linear programming approach, can be seen from Case 2a to Case 2d. This conclusion is logical, and is consistent with the belief that the output table must improve with increasing percentage of target trip information.

Table 4-14: Performance of the Linear Programming Approach (No Information Based Trip Table (100% Cells with Info.) as a Target)

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	65.85	56.24	4246	0	0
60	61.41	44.98	4008	0	0
70	61.41	44.48	5503	0	0
80	65.45	41.67	3890	0	0
90	63.71	41.67	3908	0	0
100	46.23	31.02	2701	0	0

The performance of the maximum entropy approach for Case 2d is shown in Table 4-15 below:

Table 4-15: Performance of the Maximum Entropy Approach (No Information Based Trip Table (100% Cells with Info.) As Target)

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	48.56	39.87	3908	24.28	20.35
60	51.51	45.91	4536	12.79	10.08
70	48.75	43.08	4191	4.77	3.66
80	49.10	43.67	2432	5.99	4.36
90	51.29	44.31	2465	2.82	1.73
100	43.73	38.69	3753	2.86	1.71

The maximum entropy approach shows results that are close to that obtained for Case 2c. The error statistics for link volume replication remain more or less the same. Similarly the RMSE(TT) and MAE (TT) remain nearly the same. However, the φ statistic shows an interesting trend. The values for this statistic are slightly higher for the 50% 60% and 70% available volume cases. On the other hand these error values have considerably decreased for the 80%, 90% and 100% link volume cases.

The comparison of the maximum entropy and linear programming approaches for Case 2d, is similar to the relative performance of these models in Case 2c. Except for the MAE(TT) and φ for the 60% link volume availability case, all the measures of effectiveness for all link volume availability cases are lower for the maximum entropy approach. Thus one may conclude the superiority of the maximum entropy approach for Case 2d.

4.4.3 Relatively Small Error Target Trip Tables

These trip tables have cell values that are relatively close to the correct trip table. Again we used varying extents of information in the target tables, ranging from 45% of cells carrying relatively small error information to 100% cells with such information.

4.4.3.1 Case 3a: Partial Relatively Small Error Trip Table (45% Cells with Info.)

The target trip table for this case is shown below:

Table 4-16: Relatively Small Error Trip Table (45% Cells with Info.)

From\To	1	2	3	4	5	6
4	0	806	0	0	0	0
5	0	1512	0	0	0	0
6	0	2520	0	2016	0	0

Case 3a. produces the least values for all the error statistics, among all the test cases described so far. The trip table cells are relatively closer to the true trip interchanges. The RMSE and MAE of trip table deviations are quite low, ranging from 26%-30% and 16%-21% respectively. However, there is no monotonically decreasing pattern with respect to available volumes, observed for any of the error statistics. The φ statistic, though, shows a pattern that comes close to monotonic behavior. The anomaly in the performance of these measures of effectiveness occurs for this test when the percentage of available link volumes is 70. The anomaly can be attributed to the existence of multiple optimal solutions. The replication of observed volumes by the output trip table continues to remain perfect as reflected by the zero values for the RMSE and MAE measures for these parameters.

Table 4-17: Performance of the Linear Programming Approach (Relatively Small Error Trip Table (45% Cells with Info.) as a Target

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	27.67	16.88	2663	0	0
60	26.96	16.92	2327	0	0
70	27.20	17.16	2383	0	0
80	29.02	21.16	2644	0	0
90	26.23	16.4	1659	0	0
100	26.23	16.4	1659	0	0

The performance of the maximum entropy approach for case 3a is shown in the table below.

Table 4-18: Performance of the Maximum Entropy Approach Relatively Small Error Trip Table (45% Cells with Info.) as a Target)

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	107.91	83.03	16588	41.22	31.8
60	81.31	60.64	13184	29.47	21.13
70	30.27	23.46	3991	12.6	9.46
80	29.45	21.08	3453	5.29	3.6
90	31.75	21.74	2347	3.57	2.23
100	32.19	22.46	2465	4.10	2.41

One would expect the maximum entropy approach to yield better results for this case of the target trip table than it did for Case 2a. As expected, the results for

Case 3a are better for every case of available link volumes except the 50% link volume availability case. The RMSE, MAE and φ statistics for trip table errors now range between 29%-108%, 21%-83%, and 2347-16588 respectively. The link volume error statistics are reasonably low for all cases except when the link with known volume constitute 50% and 60% of the network.

Comparing the results of the linear programming and maximum entropy approaches we first note that note that the replication of available volumes continues to be much superior for the linear programming approach. In contrast for the maximum entropy approach, the RMSE and MAE statistics for link volume replication range from 2%-32% and 4%-41%, respectively. The trip table error statistics for this case clearly indicate the superior performance of the linear programming approach. The RMSE (TT) and MAE (TT) are lower for the linear programming approach for all available link volume cases. The φ error statistic is also lower for the linear programming approach for all cases.

If one were to consider the three cases of available volumes (50%,60%,70%) then one may conclude that the linear programming approach performs far better than the maximum entropy approach for these cases. Such a claim is justifiable because its error statistics are much lower than the maximum entropy approach. On the other hand for the 80%, 90% and 100% available link volumes, error statistics are comparable for the two approaches.

Thus one may conclude general superiority of the linear programming model over the maximum entropy model.

4.4.3.2 Case 3b: Partial Relatively Small Error Trip Table (64% Cells with Info.)

The target trip table for Case 3b is shown in Table 4-19 below:

Table 4-19: Relatively Small Error Target Trip Table (64% Cells with Info.)

From\To	1	2	3	4	5	6
4	0	806	504	0	0	0
5	0	1512	0	0	0	0
6	0	2520	0	2016	605	0

The use of the target trip table shown in Table 4-19, results in a performance of the linear programming approach that is much better than the previous case. As seen in Table 4-20, all the error statistics for the linear programming approach are now very low. The RMSE for trip table deviations range between 0%-13% which is much lower than that obtained for Case 3a. The MAE and φ for trip table deviations are also very low, now ranging from 0%-8.2% and 802-0%, respectively. For the 100% volume case the exact replication of the correct trip table and observed link volumes is obtained. Also, note that for every case of link volume availability all the error statistics for trip table deviation are lower for Case 3b as compared to Case 3a.

Table 4-20: Performance of the Linear Programming Approach (Relatively Small Error Trip Table (64% Cells with Info.) as a Target)

	Measures of Error				
Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	12.95	8.28	802	0	0
60	13.34	8.24	795	0	0
70	12.69	7.78	755	0	0
80	9.28	4.72	473	0	0
90	1.32	0.8	63	0	0
100	0	0	0	0	0

The performance of the maximum entropy approach is shown in

Table 4-21 below:

Table 4-21: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table (64% Cells with Info.) as a Target)

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	69.99	57.77	13583	34.07	30.05
60	59.91	39.09	11611	24.26	19.94
70	53.26	38.10	11636	28.54	21.61
80	29.74	21.08	3561	8.41	5.99
90	31.87	21.81	2351	3.63	2.33
100	33.79	23.69	2593	4.91	3.05

For the 50 and 60% link volume availability, the measures of trip table errors for the maximum entropy approach, are much lower for this case than for Case 3a. However, for the 70% link volume availability case, the errors are lower for Case 3a. For every other case percentage available volumes these errors are higher, but in the same range as in the previous case. The RMSE, MAE and φ statistics for trip table errors now range between 30%-70%, 21%-58%, and 2351-11611 respectively. The comparison of link volume replication error statistics for Case 3a and 3b follows the same trend as the trip table error statistics. The link volume error statistics now range between 3%-34.07% and 2%-30% for the RMSE(Vol) and MAE(Vol), respectively.

A comparison of the two approaches indicates the superior performance of the linear programming approach for all cases of link volume availability. In fact, all the five measures of errors are much lower for the linear programming approach for every case of known link volumes.

4.4.3.3 Case 3c: Partial Relatively Small Error Trip Table (82% Cells with Info.)

The target trip table for this case is shown in Table 4-22 below:

Table 4-22: Relatively Small Error Trip Table (82% Cells with Info.)

From\To	1	2	3	4	5	6
4	0	806	504	0	1109	0
5	0	1512	504	0	0	0
6	0	2520	0	2016	605	0

With the target table shown in Table 4-22, the linear programming approach continues to perform well. As seen in Table 4-23, all the error statistics for the linear programming approach are low. The RMSE for trip table deviations ranges between 0.6%-14%. The MAE and φ for trip table deviations are also very low, now ranging from 0.36%-9% and 29-848, respectively. However, for every case of link volume availability all the error statistics for trip table deviation are higher than for the previous case. This may be attributed to two plausible reasons. The first is the existence of multiple optima to the LPMLV formulation. The second reason is that the additional information in the two trip interchanges that we have added is not correct but only relatively correct. Thus it would seem that there is a tradeoff between the reduction in underspecification with the addition of two cells in the target table, and constraining the output trip table values for these interchanges to be close to the incorrect cell values.

Table 4-23: Performance of the Linear Programming Approach (Relatively Small Error Trip Table (82% Cells with Info.) as a Target)

	Measures of Error				
Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	13.77	8.7	848	0	0
60	13.75	8.42	820	0	0
70	13.42	8.18	795	0	0
80	9.28	4.72	473	0	0
90	1.45	1.16	92	0	0
100	0.6	0.36	29	0	0

The performance of the maximum entropy approach is shown in Table 4-24 below:

Table 4-24: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table (82% Cells with Info.) as a Target)

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	29.13	23.19	2379	12.58	11
60	30.77	22.92	2503	7.91	5.93
70	30.56	21.12	2332	4.09	3.11
80	30.34	21.44	2355	4.24	3.03
90	30.92	21.03	2308	2.79	1.72
100	30.26	21.89	2478	7.09	4.41

The maximum entropy approach shows a significant improvement in its performance as compared to the previous case (Case 3b). All the error statistics are now quite low. The RMSE, MAE and φ statistics for trip table errors now range between 29%-31%, 21%-23%, and 2308-2503 respectively. These error measurements are lower than the corresponding values obtained for Case 3b. The link volume replication error statistics for Case 3c are also low. The RMSE(Vol) and MAE(Vol) now range between 4%-13% and 3%-11% respectively.

Similar to the conclusion reached for Case 3a, the linear programming approach continues to remain superior over the maximum entropy approach, for all cases of link volume availability.

4.4.3.4 Case 3d: Relatively Small Error Target Trip Table (100% Cells with Info.)

The target trip table for Case 3d is shown in Table 4-25 below:

Table 4-25: Relatively Small Error Target Trip Table(100% Cells with Info.)

From\To	1	2	3	4	5	6
4	0	806	504	0	1109	0
5	0	1512	504	0	0	0
6	504	2520	0	2016	605	0

The performance of the linear programming approach for the target table shown in Table-25 is similar to its performance in Case 3c, as seen in Table 4-26 below. While the RMSE, MAE and φ variations are still low they are not much better than Case 3b. In fact, except for the 100% volume availability case the results of Case 3b are much better. As explained earlier this could be attributed to the trade off between the reduction in the degree of underspecification due to the addition of information and constraining the solution to be close to the wrong information specified. However, note here again that the replication of observed link volumes is exact.

Table 4-26: Performance of the Linear Programming Approach (Relatively Small Error Trip Table (100% Cells with Info.) as a Target)

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	13.77	8.7	848	0	0
60	13.69	8.34	811	0	0
70	13.42	8.18	795	0	0
80	12.62	7.7	749	0	0
90	12.86	8.02	779	0	0
100	0.0	0.0	0	0	0

For the 100% available link volume case the output trip table matches the correct trip table exactly. This is reflected in the 0 values for all the three measures for deviation of the output trip table from the correct trip table.

The performance of the maximum entropy approach for this case is shown in the table below:

Table 4-27: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table (100% Cells with Info.) as a Target)

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	29.14	23.19	2379	12.58	11
60	30.64	22.85	2496	8.57	7.47
70	30.36	21.38	2382	3.93	3.26
80	30.32	21.47	2360	4.14	3.42
90	30.92	21.04	2308	2.79	1.72
100	30.24	21.87	2478	7.09	4.41

The performance of the maximum entropy approach continues to remain good for Case 3d. The measures of errors, are low as that found for the previous case. The RMSE, MAE and φ statistics for trip table errors now range between 29%-31%, 21%-23%, and 2308-2496 respectively. The errors for link volume replication are also low, with the RMSE(Vol) and MAE(Vol) ranging from 3%-13% and 2%-11% respectively.

However, comparison of the linear programming and maximum entropy approaches clearly indicates the superiority of model LPMLV. Every error statistics is much lower for the linear programming approach, for every case of available link volumes. Thus, here again we confirm the superiority of the linear programming approach over the maximum entropy approach.

4.4.4 Correct Trip Table Based Target Tables

Trip tables with "correct" trip interchange values are seldom available as target tables, in most practical applications. However, we employ such target tables from a perspective of theoretical interest to see how these models perform. Here, again we used varying extents of information in the target tables, ranging from 45% of cells carrying correct trip information to 100% cells with such information.

4.4.4.1 Case 4a: Partial Correct Trip Table (45% Cells with Info.)

The target trip table used for this case is shown in Table 4-28 below:

Table 4-28: (Correct Trip Table Based (45% Cells with Info.) Target)

From\To	1	2	3	4	5	6
4	0	600	0	0	0	0
5	0	1700	0	0	0	0
6	0	2500	0	2000	0	0

The performance of the linear programming approach with a target table having 5 cells with correct information, is shown in Table 4-29 below. The RMSE and MAE of trip table deviations are quite low, ranging from 19%-28% and 8%-16% respectively. However, there is no monotonically decreasing pattern with respect to available volumes, observed for any of the error statistics. In fact some irregularities may be observed.

Table 4-29: Performance of the Linear Programming Approach (Correct Trip Table (45% Cells with Info.) as a Target)

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	28.14	12.00	1840	0	0
60	28.14	12.00	1840	0	0
70	18.76	8.00	934	0	0
80	18.76	8.00	934	0	0
90	26.33	16.00	1599	0	0
100	26.33	16.00	1599	0	0

One might expect a comparison of these error statistics with those obtained for Case 3a to indicate that the values for Case 4a are lower. However, except for the 70% and 80% available cases, the error statistics are lower for Case 3a. This anomaly can be attributed to the existence of multiple optimal solutions. The replication of observed volumes by the output trip table continues to remain perfect as reflected by the zero values for the RMSE(Vol) and MAE(Vol) statistics.

The performance of the maximum entropy approach for this case is shown below:

Table 4-30: Performance of the Maximum Entropy Approach (Correct Trip Table (45% Cells with Info.) as a Target)

	Measures of Error				
Avail. Vols.	RMSE % (TT)	MAE % (TT)		RMSE % (Vol)	MAE % (Vol)
50	107.86	82.99	18236	41.19	31.78
60	80.78	60.29	12756	29.5	21.26
70	31.22	24.00	4069	12.61	9.47
80	29.93	21.01	3537	5.3	3.6
90	31.78	21.79	2352	3.63	2.28
100	32.09	22.26	2414	2.78	1.76

The RMSE, MAE and φ statistics for trip table errors now range between 29%-107%, 21%-83%, and 2352-18236, respectively. The link volume error statistics are reasonably low for all cases except when the links with known volume constitute 50% and 60% of the network.

Comparing the results of the linear programming and maximum entropy approaches first note that the replication of available volumes continues to be much superior for the linear programming approach. In contrast for the maximum entropy approach, the RMSE and MAE statistics for link volume replication range from 3%-41% and 2%-32%, respectively. Trip table error statistics for this case, clearly indicate the superior performance of the linear programming approach. The RMSE (TT) and MAE (TT)are lower for the linear programming approach for cases link volume availability. The φ error statistic is also lower for the linear programming approach for all cases.

Similar to the patterns observed for Case 3a, if one were to consider the three cases of available volumes (50%,60%,70%) then one may conclude that the linear programming approach performs far better than the maximum entropy approach for these cases. Such a claim is justifiable because the error statistics for link volume replication, and φ are much lower for the maximum entropy approach. On the other hand, for the 80%,90% and 100% available link volumes, the φ, RMSE (Vol) and MAE (Vol) statistics are comparable for the two approaches.

Thus, similar to Case 3a, one may conclude that the linear programming model performs much superior than the maximum entropy model for the 50, 60 and 70% link volume availability cases. For the 80, 90 and 100% link volume while the linear programming model continues to perform better, the maximum entropy model yields comparable results.

4.4.4.2 Case 4b: Partial Correct Trip Table (64% Cells with Info.)

Table 4-31: Partial Correct Trip Table (64% Cells with Info.)

From\To	1	2	3	4	5	6
4	0	600	700	0	0	0
5	0	1700	0	0	0	0
6	0	2500	0	2000	600	0

The linear programming approach reproduces the link volumes and the correct trip table for the target trip table shown in Table 4-31 below. These results are an obvious improvement over Case 4a. and clearly indicate that with a good target information table, the linear programming model performs very well even when the link volume information is unavailable for 50% of the links.

Table 4-32: Performance of the Linear Programming Approach Correct Trip Table (64% Cells with Info.) as a Target

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	0	0	0	0	0
60	0	0	0	0	0
70	0	0	0	0	0
80	0	0	0	0	0
90	0	0	0	0	0
100	0	0	0	0	0

The performance of the maximum entropy approach for this case is shown in Table 4-33 below:

Table 4-33: Performance of the Maximum Entropy Approach Correct Trip Table (64% Cells with Info.) as a Target

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	58.27	47.32	18236	41.19	31.78
60	58.78	39.71	12756	29.5	21.26
70	53.26	38.10	4069	12.61	9.47
80	27.56	18.24	3537	5.3	3.6
90	31.98	22.07	2352	3.63	2.28
100	33.10	22.41	2414	2.78	1.76

In contrast to the performance of the linear programming model, the performance of the maximum entropy approach is not as good as one would expect it to be when the trip table provided is quite good. In fact, some surprising results have been reported. When the percentage of available volumes is 50% and 60%, the performance of the maximum entropy model turns out to be much better with a structural table input as compared to Case 4b when 7 cells had the correct trip information. Also, one would expect better replication of the link volumes, but as the statistics indicate the performance is not so. In fact, for the 50% link volume availability case the RMSE and MAE are as high as 29.5%-41.19% respectively. Such a high value for the error, with 7 of 11 cells bearing correct information, is inexplicable.

However, the measures of error for closeness to observed volume do show a clear monotonic improvement as the link volume information availability increases. The measures of error for trip table deviations, on the other hand, do not monotonically improve with increasing link volumes.

Comparing the linear programming and maximum entropy approaches for Case 4b, again clearly indicates the superiority of the linear programming approach. As found for Case 3c, 3d, and 4a, every error statistic is much lower for the linear programming approach for all cases of available link volumes.

4.4.4.3 Case 4c: Partial Correct Trip Table (82% Cells with Info.)

The target trip table for this case is shown in Table 4-34 below:

Table 4-34: Correct Trip Table (82% Cells with Info.)

From\To	1	2	3	4	5	6
4	0	600	700	0	1100	0
5	0	1700	300	0	0	0
6	0	2500	0	2000	600	0

The performance of the linear programming approach for the target trip table shown in Table 4-34 below, is identical to the performance in the previous case. The replication of the link volumes is perfect and the output trip table exactly matches the correct trip table for the network, as reflected in the 0 values for the RMSE, MAE and φ values for the trip table deviation errors, and similar values for the RMSE and MAE of link volume replication errors.

Table 4-35: Performance of the Linear Programming Approach (Correct Trip Table (82% Cells with Info.) as a Target)

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	0	0	0	0	0
60	0	0	0	0	0
70	0	0	0	0	0
80	0	0	0	0	0
90	0	0	0	0	0
100	0	0	0	0	0

The performance of the maximum entropy approach for this case is shown in Table 4-36 below:

Table 4-36: Performance of the Maximum Entropy Approach (Correct Trip Table (82% Cells with Info.) as a Target)

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	25.36	15.77	1632	9.31	9.31
60	24.23	12.97	1388	8.3	8.25
70	25.46	14.62	1539	8.27	8.27
80	28.98	17.66	1976	2.57	2.57
90	31.08	20.96	2298	2.8	2.8
100	30.66	21.46	2350	4.03	4.03

The performance of the maximum entropy approach for Case 4c, is in general much superior than the other cases described so far. The upper bound of the range of the RMSE, MAE and φ statistics for trip table deviation are much lower than the other cases. They now vary only between 24%-31%, 13%-21% and 1388-2350, respectively. However, it may be noted that the measures of error for trip table deviation show an increasing trend with increasing link volume availability while the statistics for link volume replication errors are decreasing. This result is contrary to what one may expect.

While the results obtained, for this case, with the maximum entropy approach are good, they are no where as good as those obtained with the linear programming approach. The comparison of the two approaches yields the same conclusion as that arrived for Cases 3b, 3c, 4a and 4b, with the linear programming approach performing much better than the maximum entropy approach.

4.4.4.4 *Case 4d: Correct Trip Table (100% Cells with Info.)*

The target trip table for this case is the correct trip table for the network and is shown in Table 4-37 below:

Table 4-37: Correct Trip Table (100% Cells with Info.)

From\To	1	2	3	4	5	6
4	0	600	700	0	1100	0
5	0	1700	300	0	0	0
6	0	2500	0	2000	600	0

As seen in Table 4-38 below, the linear programming model continues to replicate the observed volumes and correct trip table exactly.

Table 4-38: Performance of the Linear Programming Approach (Correct Trip Table as a Target)

	Measures of Error				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	0	0	0	0	0
60	0	0	0	0	0
70	0	0	0	0	0
80	0	0	0	0	0
90	0	0	0	0	0
100	0	0	0	0	0

The performance of the maximum entropy approach for this case is shown in Table 4-39 below:

Table 4-39: Performance of the Maximum Entropy Approach (Correct Trip Table as a Target)

	Measures of Error				
Avail. Vols.	RMSE % (TT)	MAE % (TT)	ϕ	RMSE % (Vol)	MAE % (Vol)
50	25.12	15.49	1609	9.31	9.75
60	24.25	13.00	1390	8.25	7.33
70	25.37	14.63	1540	8.27	7.32
80	28.98	17.66	1976	2.57	1.62
90	31.08	20.96	2298	2.8	1.72
100	30.66	21.46	2350	4.03	2.69

Since the correct trip table is used for the target trip table one would expect the maximum entropy approach to perform much better than it did for the previous cases. As expected, the results obtained with this trip table are indeed the best results, among the tests conducted so far. However the output trip table is still not very close to the actual table. As seen in Table 4-39 the statistics for trip table error measurements are quite high, considering that the correct trip table has been used as the target trip table. This is in contrast with the performance of the linear programming approach, which had zero measures of error for the same target trip table. Thus for this Case too the linear programming method produces much higher quality of results than the maximum entropy approach.

4.5 Number of Objective Function Resets and Iterations for LPMLV

The linear programming model LPMLV triggers updating procedures for the impedances of links with unknown volumes, whenever the objective function falls below a implicitly defined threshold, as described in Chapter 3. For the case studies presented in this Chapter, the number of such updates (r), and the total number of iterations completed before termination are shown in Table 4-40.

Table 4-40. Number of Objective Function Resets and Iterations for LPMLV

Target Trip Table	% Available Volumes									
	50%		60%		70%		80%		90%	
	r	k	r	k	r	k	r	k	r	k
Structural	8	56	8	61	8	35	9	39	9	49
n5	7	25	8	34	8	30	9	29	9	51
n7	8	31	8	34	8	30	9	29	9	51
n9	8	34	8	93	9	76	9	45	9	50
n11	8	39	9	107	9	46	8	39	9	52
s5	8	33	8	36	9	44	9	40	9	51
s7	8	28	8	97	9	201	9	169	9	48
s9	8	52	8	53	9	37	9	33	9	52
s11	8	54	9	55	9	45	9	50	9	72
c5	8	43	8	28	8	77	9	35	9	52
c7	8	61	8	77	9	252	9	174	9	53
c9	8	59	8	82	9	35	9	39	9	55
c11	8	83	9	113	9	56	9	58	9	72

Legend:

n5, n7, n9, n11: no information based trip table having 5, 7, 9 and 11 cells with information, respectively.

s5, s7, s9, s11: small error based trip table having 5, 7, 9 and 11 cells with information, respectively.

c5, c7, c9, c11: correct information based trip table having 5, 7, 9 and 11 cells with information, respectively.

r = number of updates

k = number of iterations

Note that for the all artificial start implementation of Model LPMLV, the initial value of the objective function increases with increasing knowledge of available link volumes. Also, note that the penalty for link volume deviation is the largest objective cost coefficient. Furthermore, the final output trip table for all the case runs described so far replicates available link volumes perfectly. Hence the difference between the initial and termination values of the objective function is likely to be increase as knowledge of available link volumes increases. Thus, the objective function resetting mechanism adopted triggers a greater number of resets when the knowledge of link volumes increases, as seen in Table 4-40.

4.6 Comparison of Updating Procedure with a No Impedance Updating Methodology

To investigate the influence of the impedance updating procedures for links with unknown volumes, the results of model LPMLV are compared with a case in which no cost updating was performed. (Such a no impedance updating implementation of the model can easily be adopted by setting the scaling factor for the objective functions resetting criteria to be a very large number.) The tests used the smaller error trip tables based target tables. Table 4-41 shows the comparison.

Table 4-41: Comparison of ϕ statistic: No Link Cost Update Implementation vs. LPMLV

Target Table	Available Volumes				
	50%	**60%**	**70%**	**80%**	**90%**
	5596	5652	2367	2405	1663
s5	2669	2327	2383	2644	1659
	802	795	795	791	63
s7	802	795	765	473	63
	856	811	811	811	92
s9	848	820	795	473	92
	856	811	811	811	92
s11	848	811	795	745	779

Legend:

No Cost Update
LPMLV

The results, in general, show that the cost vector updating scheme for objective function indeed produces a better quality trip table.

4.7 Conclusions

The test results discussed in this chapter are graphically encapsulated in Figures 4.2-4.4. Each figure shows the variation of one of the error statistic (represented in the z axis) with respect to the different available link volume and prior information cases (represented in the xy plane) discussed so far, for each model.

The variation of the ϕ statistic for the linear programming and maximum entropy approaches are depicted in Figures 4-2 and 4-3, respectively. In general an improving trend is observed for the ϕ statistic when better quality target table is

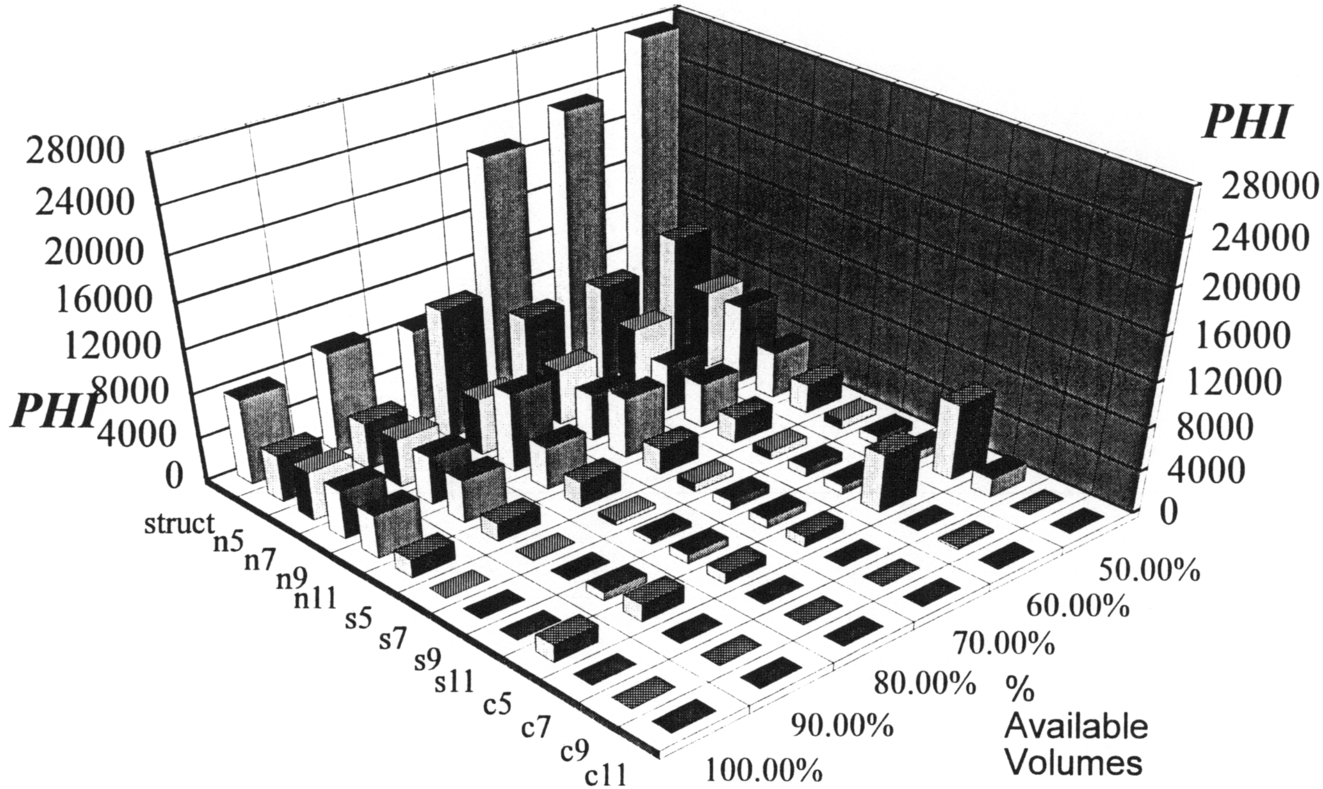


Fig 4-2 Artificial Network - LP Model PHI (TT) Values

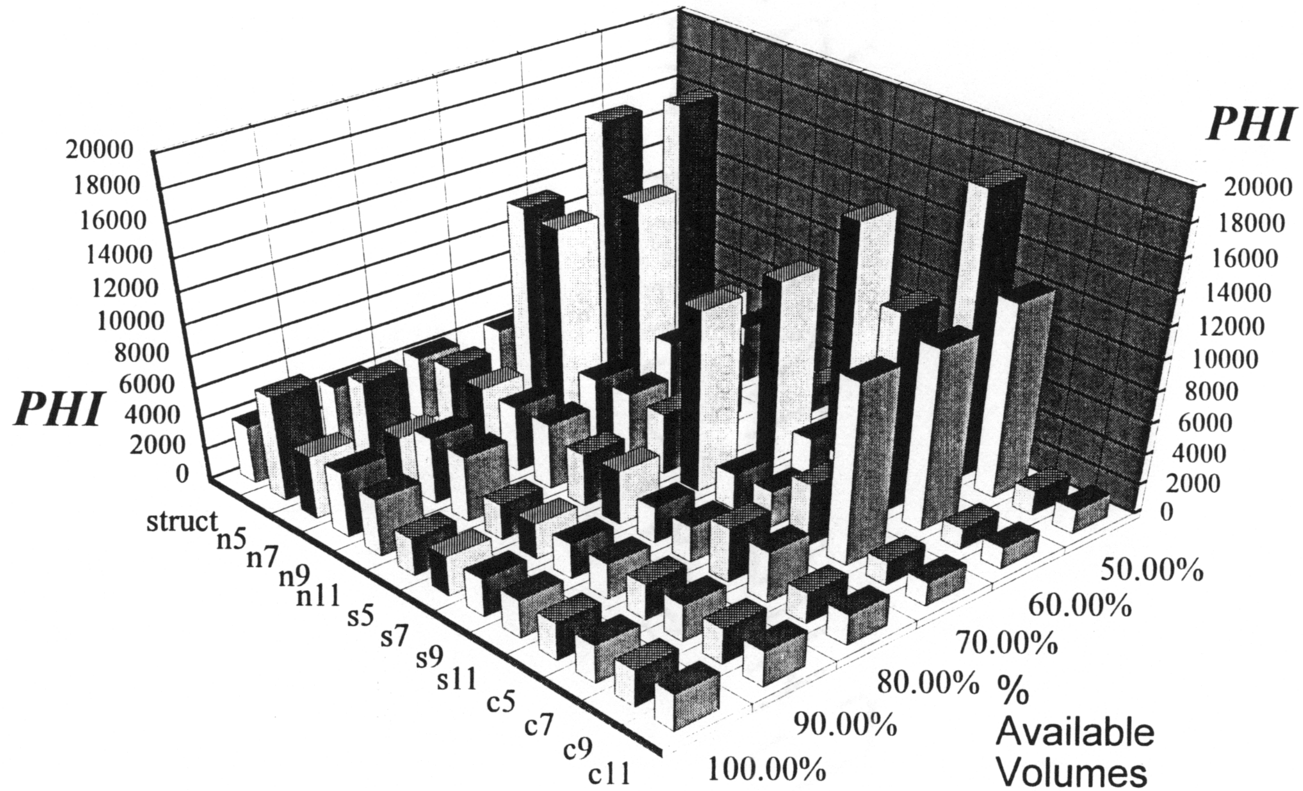


Fig 4-3 **Artificial Network - THE Model PHI (TT) Values**

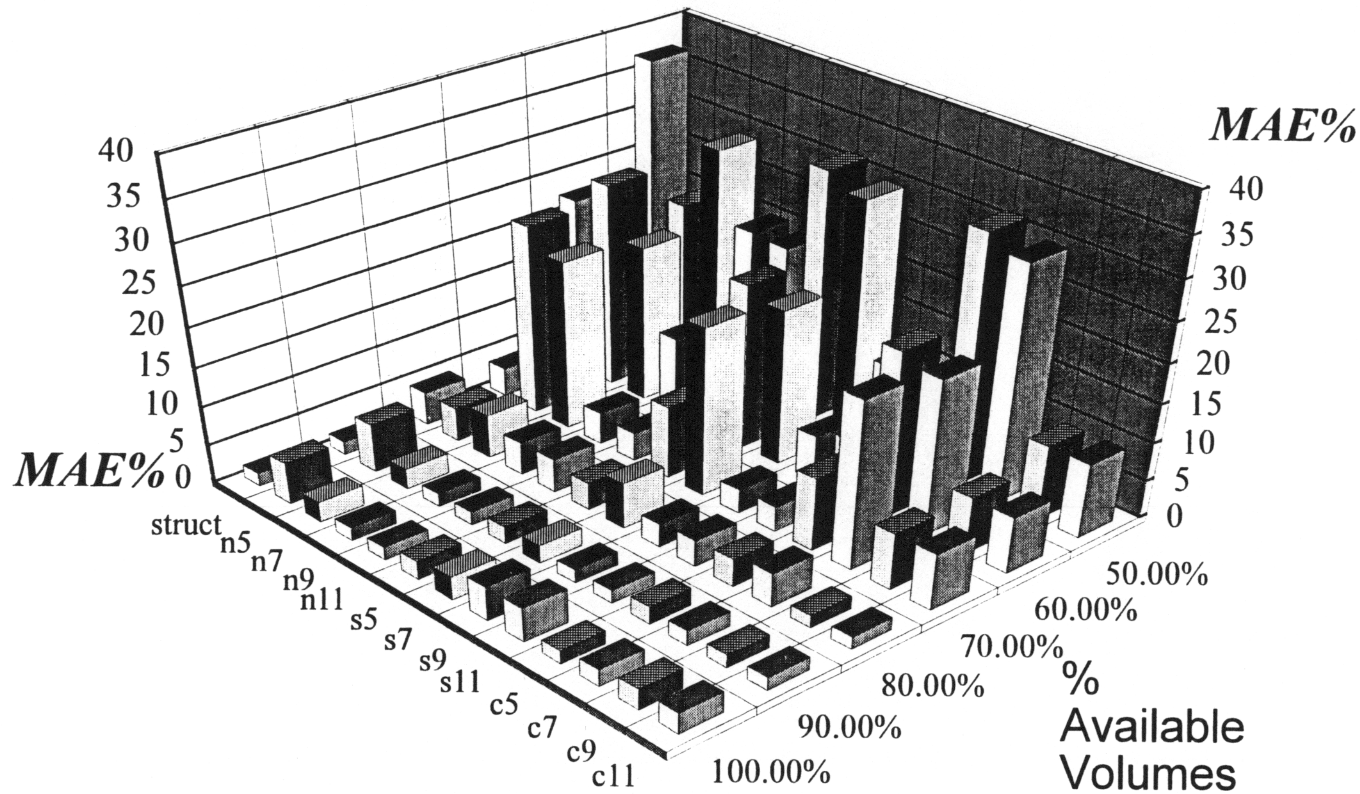


Fig 4-4 **Artificial Network - THE Model %MAE (Vol) Values**

supplied ranging from no prior information (n5,n7,n9,n11) to correct target tables (c5,c7,c9,c11). For each of these 13 target trip tables, in general, both models are seen to improve their performance with increasing available link volume percentages. Thus the ϕ statistic is seen to generally decrease for both models with the improvement in target table information, and increase in percentage of available volumes.

The variation of the link volume replication error for the maximum entropy approach as measured by the MAE(Vol) statistic is depicted in Figure 4-4. For the maximum entropy approach, this statistic is seen to improve with increase in percentage available link volumes and better target information. The linear programming approach perfectly replicates the available link volumes, and the MAE(Vol) is zero for all cases of percentage available link volumes and target trip tables.

The figures also serve to illustrate the general superiority of the linear programming approach over the maximum entropy approach. Such an inference is drawn based on the lower values of the error statistics for the linear programming model, as compared to the corresponding values obtained for the maximum entropy approach.

5. Evaluation using Real Network

5.1 Introduction

The previous chapter presented the validation of the linear programming approach through tests on an artificial test network. The results indicated that the linear programming approach performed very well and was found to be much superior than the maximum entropy approach for most cases. This chapter presents tests to validate this approach using a real network.

5.2 Purdue University Network

The network chosen is the village network of Purdue University, West Lafayette, Indiana. It has 16 zones, 43 nodes and 130 links. The network covers an area of one square mile and the small size of the network lends itself well for the purpose evaluating OD estimation approaches. The network characteristics like capacity, free flow speed, etc., required to test the linear programming and maximum entropy approaches were available for every link of the network. A trip table that was obtained by synthesizing five different approaches, including a license plate survey, and believed to be reasonably good by researchers at Purdue University (Fricker & Barbour), was available. However, this table on user equilibrium assignment did not replicate the observed volumes, and implied that this trip table is inconsistent with the observed link volumes.

One of the characteristics of this network was that the volume to capacity ratios were quite low for every link, thus making the updated travel cost on every link equal to its free flow cost. To obtain a partial set of link volumes, some of these link volumes were assumed to be unknown. While estimating the solution trip table using the linear programming approach, the additional information that the actual travel costs are the same as the free flow cost was exploited to increase the computational efficiency of the solution procedure. The link cost update procedures and the objective function cost vector resetting procedures for links with unknown volumes, as described in Chapter 3, are no more necessary and were thus skipped in the models solution procedure. This added to the linear programming model's computational efficiency.

5.2.1 Target Trip Tables

The target trip tables used in the tests conducted on this network are primarily of four types. Each of these four types contain cell values for all possible trip interchanges. Using these four types of trip tables 10 trip tables for the tests were derived. These derived tables were partial trip tables that were obtained by removing some information from the four types of tables. The extents to which information was removed vary from 0%-50% of the total number of feasible interchanges. The four types of tables that form the basis for the derived tables are described below.

The structural trip table represents one which has 0-1 values for its cells to indicate merely if that trip interchange is feasible or not. The no prior information trip table has a uniform value for all feasible interchanges. The value contained in the cells corresponding to the feasible interchanges, is equal to 27 which represents the average true trip interchange factored by a value of 0.8. The factor of 0.8 was used in order for the prior trip table to resemble an outdated table, which is likely to have lesser total travel in the network. The third and fourth types of trip table used as a seed, were relatively close to the assumed "true" trip table. They were generated by introducing errors into the true trip table using the approach described below:

$$P_{ij} = C_{ij}(\psi + \beta_{ij}) \qquad (5.1)$$

$$-\beta_{\max} < \beta < \beta_{\max}$$

where P_{ij} is the target table's ij^{th} cell, C_{ij} is the corresponding element in the "correct trip table", ψ is the mean cell ratio of target trip table cell to correct trip table cell, β_{ij} is a normally distributed cell value error and $\beta_{\max}$ is the bound on this error. ψ was set to be the same for all cell values in a given table. For the third type of target tables we used ψ =0.8 and β_{max} =0.2. For the fourth type of target tables we used ψ =0.9 and β_{max} =0.2.

5.3 Test Results

5.3.1 Case 1: Structural Table

This case in which the target trip table cell values are either 0 or 1, the trip table constraints for the linear programming approach aim to keep the output trip table to be as close to either 0 or 1 (as the case may be). Since such a target table does not carry much information, the linear programming approach does not perform very well. The RMSE statistic for trip table replication, for all four cases of available link volumes are indeed very high with the minimum being as high as 140% (Table 5-1). The MAE statistic for trip table replication is also very high, ranging from 80%-89%. However, as explained earlier such a performance is only to be expected.

Table 5-1: Performance of the Linear Programming Approach (Structural Table as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	156.09	84.71	13713	4.96	2.88
70	157.14	88.52	14663	8.94	4.89
90	155.53	85.09	14742	9.57	5.53
100	140.37	79.94	12611	10.1	5.37

The replication of observed volumes, however turns out to be very good. The % RMSE error for link volume replication varies from 10.1% (when all link volumes are specified) to 4.96% (when only 50% of link volumes are specified. The relatively low values for these measures indicates that the output trip table replicates the observed volumes well. Since the link volume error statistics are low, one may attribute the high values for trip table error statistics to the poor quality of the target trip table.

The maximum entropy approach, on the other hand, performs marginally better than the linear programming approach. The performance of the maximum entropy approach for this case is shown in Table 5-2 below:

Table 5-2: Performance of the Maximum Entropy Approach (Structural Table as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	134.23	64.19	6872	25	23.1
70	127.36	60.95	5938	26.87	21.04
90	131.66	67.98	7250	26.87	21.04
100	124.41	65.81	7258	21.69	13.64

When compared to the linear programming approach the RMSE, MAE and φ for trip table deviations are significantly lower for the maximum entropy approach, ranging between 124%-134%, 61%-68% and 5938-7258, respectively. However, the replication of observed volumes is not as low as those obtained with the linear programming approach . As compared to the linear programming approach, the RMSE and MAE for link volume deviations are higher for the maximum entropy approach, and range between 22%-27% and 14%-23%, respectively.

It must be added here, that the applicability of either model for structural information target table case is questionable because of the high values for trip table errors for both approaches.

5.3.2 Case 2a: Partial No Prior Information Table (60% Cells with Info.)

This table has 60% of the cells having the value equal to 80% of the sum of all true trip interchange values.

Table 5-3: Performance of the Linear Programming Approach (No Prior Information Table (60% Cells with Info.) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	167.13	81.17	10160	5.13	2.4
70	138.05	81.17	10520	8.91	4.63
90	132.83	70.81	9887	8.47	4.63
100	124.06	66.61	9648	10	5.4

The linear programming model shows a substantial improvement, with the increased information provided by the target trip table for Case 2a. The RMSE(TT) and MAE(TT) values decrease for all cases except the 50% available link volume case. However, the values for these statistics continue to remain high as seen in Table 5-2. This can be attributed to the poor quality of the target table and also to the inconsistency between the observed volumes and the trip patterns that are assumed to be correct.

Note that a monotonically decreasing trend may be observed for the RMSE(TT) and MAE(TT) errors, as the available link volumes increase. However, the RMSE(Vol) and MAE (vol.) show the reverse trend, i.e., they increase as the percentage of available link volumes increase. This could be attributed to errors or due to the fact that the link volumes are not consistent with OD flows, and inconsistencies with observed link volume data.

The performance of the maximum entropy approach for this case is shown in Table 5-4 below:

Table 5-4: Performance of the Maximum Entropy Approach (No Prior Information Table (60% Cells with Info.) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	ϕ	RMSE % (Vol)	MAE % (Vol)
50	208.99	94.35	11848	24.75	22.45
70	210.00	92.19	11459	27.81	21.74
90	185.16	89.07	11642	24.96	17.53
100	177.75	88.81	11612	21.42	13.41

The maximum entropy approach performs poorly for Case 2a, as reflected by the error statistics described in Table 5-4. In fact, the performance of this approach is much worse than it was for Case 1. The table by and large reflects a trend of decreasing values for error statistics with increasing link volume information.

The RMSE(TT) and MAE(TT) values are very high ranging from 178%-209% and 89%-94% respectively. The link volume replication error statistics though lower than that obtained for Case I, are still quite high with the RMSE(Vol) and MAE (Vol) ranging from 21%-28% and 13%-22%, respectively.

Comparing the results of the two approaches, the superiority of the linear programming approach is clear. This conclusion can be drawn based on trip table error statistics and link volume replication statistics, which are lower for the linear programming approach.

5.3.3 Case 2b: Partial No Prior Information Table (80% Cells with Info.)

This table has 80% of the cells having the value equal to 80% of the sum of all true trip interchange values.

The measures of errors continue to remain high for the linear programming approach, as is reflected in Table 5-5 below. As explained earlier this is only to be expected as the target trip table does not contain any information that is close to the actual trip table, and the "correct" trip table itself does not replicate the observed link volumes. As compared to case 2a, the φ statistic shows a clear improvement for all four cases of available link volumes. The RMSE(TT) and MAE(TT) remain in the same range as for case 2a. The replication of observed volumes continues to remain good, with the RMSE(Vol) and MAE(Vol.) varying between 5%-10% and 3%-6%, respectively.

Table 5-5: Performance of the Linear Programming Approach (No Prior Information Table (80% Cells with Info.) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	179.59	80.13	9048	5.08	2.59
70	130.61	80.13	9086	8.79	4.71
90	128.98	69.60	9392	6.94	4.63
100	137.70	69.5	9396	10.42	5.4

The performance of the maximum entropy approach for Case 2b is shown in the Table 5-6 below:

Table 5-6: Performance of the Maximum Entropy Approach (No Prior Information Table (80% Cells with Info.) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	230.00	86.46	10005	25.02	23.03
70	223.49	86.43	9524	27.71	21.82
90	170.72	82.89	9821	26.12	17.78
100	159.59	80.86	9873	21.63	13.61

The measures of trip table errors for the maximum entropy approach, are much lower for this case than for Case 2a. In fact, for every case of percentage available volumes these errors are lesser than the previous case. The RMSE, MAE and φ statistics for trip table errors now range between 160%-230%, 81%-86%, and 9821-10005, respectively. These error measurements are lower than the corresponding values of 178%-209%, 81%-86%, and 11459-11848 obtained for Case 2a. However, the link volume replication errors are not always lower for Case 2b. In fact, except the 70% link volume availability case both RMSE(Vol) and MAE(Vol) are lower for Case 2a.

Comparison of the two approaches allows us to make conclusions similar to that made for the previous case . For all cases of available link volumes, it is seen that the linear programming approach has lower values for all measures of error. However, while this observation permits us to conclude the superiority of the linear programming approach over the maximum entropy approach, questions about the applicability of either approach still remain because of the high values for trip table and link volume error statistics obtained with either approach. However, one cannot conclude that either approach is not applicable because the "correct trip table" which is the basis for error computations is itself inconsistent with the observed link volumes, and the assumption that this table is "*correct*" or "*true*" must be treated with some reservation!

5.3.4 Case 2c: Partial No Prior Information Table (100% Cells with Info.)

This table has 100% of the cells having the value equal to 80% of the sum of all true trip interchange values. Thus the target table for Case 2c has an uniform cell value of 27 for all the feasible trip interchanges. The trip table is complete because now all feasible trip interchanges have a target value. Table 5-7 below, shows the test results obtained with the linear programming approach. As compared to Case 2a and 2b, the values for the errors are the lower for all cases of available link volumes. The RMSE deviations range between 106%-119%, which correspond to the lowest upper and lower bound values for the ranges among all the cases described so far. Similarly, the MAE and φ values have the smallest range and the least lower and upper limits of the range. The link volume replication errors continue to be low. In general, a gradual improvement in the quality of results for the linear programming approach can be seen from Case 2a to Case 2c. Though the trip table error statistics have definitely improved, they are still very high.

Table 5-7: Performance of the Linear Programming Approach (No Prior Information Table (100% Cells with Info.) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	118.78	65.11	8721	5.00	2.38
70	113.65	65.11	8344	9.01	4.88.
90	107.67	62.03	8364	12.42	6.36
100	105.67	61.39	7304	9.38	4.8

The performance of the maximum entropy approach for Case 2c is shown in the Table 5-8 below:

Table 5-8: Performance of the Maximum Entropy Approach (No Prior Information Table (100% Cells with Info.) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	131.94	63.14	6602	24.51	22.67
70	127.43	61.07	5808	26.72	21.13
90	133.85	69.82	7351	26.39	17.98
100	125.00	66.23	7300	21.73	13.68

The maximum entropy approach shows results that are better than that obtained for Case 2b. The RMSE(TT) and MAE (TT) now vary between 125%-134% and 61%-70% which is much lower than the values obtained for cases 2a and 3b. The φ statistic too, continues to remain low. The values for this statistic are also lower than those obtained for cases 2a and 2b. The link volume replication errors, however, remain more or less the same as those obtained for previous cases.

The comparison of the maximum entropy and linear programming approaches for Case 2c is similar to the relative performance of these models in Case 2a. Except for the MAE(TT) for the 50% and 60% link volume availability cases, all the other measures of effectiveness for all link volume availability cases are lower for the linear programming approach. Thus one may conclude the superiority of the linear programming approach for Case 2c.

The limitations and questions regarding the applicability of either of these models, as described earlier, continue to apply.

5.3.5 Case 3a: Relatively Small Error Trip Table (60% Cells with Info., $\psi = 0.8$)

This trip table has 60% of its cells obtained using $\psi = 0.8$ in equation 5-1, and the remaining cells are zero.

This case produces the least values of all the error statistics for the linear programming approach, among all the test cases described so far, as seen in Table 5-9 below. The trip table cells, however, are still not very close to the "correct" trip interchanges, as is reflected in the high values of RMSE(TT), MAE(TT) and ϕ. The RMSE and MAE of trip table deviations are still high, ranging from 103%-106% and 54%-61%, respectively. There is no monotonically decreasing pattern with respect to available volumes, observed for any of the error statistics. The replication of observed volumes by the output trip table continues to remain good as reflected by the low values for the RMSE and MAE measures for these parameters. The RMSE(Vol) and MAE(Vol) now range between 3%-11% and 3%-5%, respectively.

Table 5-9: Performance of the Linear Programming Approach (Relatively Small Error Trip Table(60% Cells with Info., $\psi = 0.8$) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	105.92	54.87	7517	3.09	1.75
70	103.09	54.87	7173	8.00	4.05
90	105.09	54.39	8145	10.88	5.78
100	107.13	60.12	8837	8.99	4.95

The performance of the maximum entropy approach for case 3a is shown in the table below.

Table 5-10: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table (60% Cells with Info., $\psi = 0.8$) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	211.09	92.79	12522	23.87	21.35
70	228.21	84.23	10252	27.08	20.69
90	186.11	87.86	11953	25.37	17.06
100	184.48	87.48	11917	21.27	13.34

One would expect the maximum entropy approach to yield better results for this case of the target trip table than it did for Case 2a. However, this is not found to be entirely true. In fact the RMSE(TT) is higher than Case 2a for every case of available link volume and the ϕ statistic is lower for Case 3a only for the 70% link volume available case. On the other hand, the MAE(TT) values are lower in Case 3a for all cases of available link volumes. The RMSE, MAE and φ statistics for trip table errors now range between 183%-211%, 84%-93%, and 10251-12522, respectively. The link volume error statistics have not changed much from the previous cases.

Comparing the results of the linear programming and maximum entropy approaches, we first note that the replication of available volumes continues to be much superior for the linear programming approach, ranging between 3%-11% and 1%-6% for RMSE(Vol) and MAE(Vol), respectively, and 27%-23% and 22%-13%, respectively, for the maximum entropy approach. Trip table error statistics for this case clearly indicate the superior performance of the linear programming approach. The RMSE (TT) and MAE (TT) are lower for the linear programming approach for all cases of link volume availability. The φ error statistic is also lower for the linear programming approach for all cases.

Thus one may conclude that the linear programming model performs much superior than the maximum entropy model for Case 3a.

5.3.6 Case 3b: Relatively Small Error Trip Table (80% Cells with Info., $\psi = 0.8$)

This trip table has 80% of its cells obtained using $\psi = 0.8$ in Equation 5-1, and the remaining cells are zero

The use of the target trip table for Case 3b yields interesting results for the linear programming approach. The lowest measures of errors are obtained for the 50% link volume availability case. This could be attributed to the inconsistency of the "correct" trip table with the observed volumes. The trip table error statistics are lower for Case 3b than for Case 3a, except for the 70% available link volume case. The RMSE for trip table deviations ranges between 87%-106% which is lower than in the corresponding range of 103%-108% obtained for Case 3a. The MAE and φ for trip table deviations now vary between 46%-56% and 6279-7988, respectively.

Table 5-11: Performance of the Linear Programming Approach (Relatively Small Error Trip Table(80% Cells with Info., $\psi = 0.8$) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	86.74	45.93	6279	4.93	2.14
70	106.40	45.93	8117	7.62	4.1
90	105.00	55.67	7437	7.85	4.27
100	103.92	54.08	7988	10.41	4.95

The performance of the maximum entropy approach for this case is shown in Table 5-12 below.

Table 5-12: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table(80% Cells with Info., $\psi = 0.8$) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	225.91	84.45	10700	25.47	23.19
70	228.61	84.23	10252	26.94	20.72
90	167.58	79.14	9890	25.23	16.87
100	162.83	79.33	9922	21.36	13.30

For 90% and 100% available link volumes, the measures of trip table errors for the maximum entropy approach are much lower for this case than for Case 3a.

For the other two available link volume cases, Case 3a and Case 3b yield nearly the same results. The RMSE(TT), MAE(TT) and ϕ statistics for Case 3b, range between 162%-225%, 79%-85% and 9890-10700, respectively.

Comparison of the two approaches indicates the superior performance of the linear programming approach for all cases of link volume availability. In fact, all the five measures of errors are much lower for the linear programming approach for every case of available link volumes. As discussed for all the cases till now, the applicability of either model is questionable, given the high values of the error statistics.

5.3.7 Case 3c: Relatively Small Error Trip Table (100% Cells with Info., $\psi = 0.8$)

This trip table has 100% of its cells obtained using $\psi = 0.8$ in Equation 5-1, and the remaining cells are zero.

For this target table, the linear programming approach performs much better than all the other cases reported thus far. As seen in Table 5-13 below, all the error statistics for the linear programming approach are lower than those obtained for the previous cases. The RMSE for trip table deviations ranges between 76%-103%. The MAE and φ for trip table deviations are also relatively lower, now ranging from 38%-50% and 4667-6936 respectively. The link volume replication errors continue to remain low.

Table 5-13: Performance of the Linear Programming Approach (Relatively Small Error Trip Table(100% Cells with Info., $\psi = 0.8$) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	76.76	38.2	4667	3.89	2.14
70	90.46	38.2	5841	7.84	4.1
90	103.13	50.48	6936	8.03	4.27
100	101.02	49.97	6945	7.9	4.1

Note the nearly monotonically increasing trend of the trip table error statistics with increasing knowledge of available link volumes. This trend further supports the belief that the available link volumes mey be inconsistent with the "correct" trip table.

The results obtained for Case 3c with the maximum entropy approach are shown in Table 5-14 below:

Table 5-14: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table (100% Cells with Info., $\psi = 0.8$) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	129.91	61.04	6807	24.68	22.52
70	123.77	57.51	5965	26.14	19.97
90	142.63	64.19	7126	25.27	16.94
100	138.66	65.05	7529	21.34	13.26

The maximum entropy approach shows a significant improvement in its performance as compared to the previous case (Case 3b). All the error statistics are much lower than the corresponding values obtained for cases 3a and 3b. The RMSE, MAE and φ statistics for trip table errors now range between 123%-142%, 58%-65%, and 5965-7529, respectively. The link volume replication error statistics for this case are in the same range as for the previous case. The RMSE(Vol) and MAE(Vol) now range between 21%-26% and 13%-23%, respectively

Comparison of the two approaches once again leads to similar conclusions as reached for many of the previous cases. The linear programming approach is seen to perform much better than the maximum entropy approach with significantly lower values of all measures of errors for every available link volume case. However, again note that the error statistics have high values for both the models.

5.3.8 Case 4a: Relatively Small Error Trip Table (60% Cells with Info., $\psi = 0.9$)

This trip table has 60% of its cells obtained using $\psi = 0.9$ in Equation 5-1, and the remaining cells are zero.

The performance of the linear programming approach, with a target table having its cells that may be considered to be relatively close to the "correct" cell values, is shown in Table 5-15 below. The RMSE and MAE of trip table deviations are the least for target tables in which only 60% of the cells have information, among the cases described so far. The RMSE, MAE and φ statistics for trip table errors now range between 91%-109%, 47%-58%, and 6948-8557, respectively. The link volume replication error statistics, continue to remain low with the RMSE(Vol) and MAE(Vol) varying between 4%-11% and 2%-6%, respectively.

Table 5-15: Performance of the Linear Programming Approach (Relatively Small Error Trip Table(60% Cells with Info., $\psi = 0.9$) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	105.16	49.62	6992	3.96	2.35
70	90.97	49.62	7067	8.69	4.43
90	108.69	58.02	8557	9.71	5.38
100	92.41	46.60	6949	10.61	5.93

Comparing these error statistics with those obtained for Case 3a indicate that the values for Case 4a are lower, except for the 90% available link volume case. The inconsistency of the "correct" trip table cell values with the observed link volumes is a possible reason for the above observation .

The test results obtained with the maximum entropy approach for this case are shown in the following table.

Table 5-16: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table(60% Cells with Info., $\psi = 0.9$) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	232.49	95.90	13016	25.05	21.78
70	239.27	94.21	12621	28.53	21.26
90	185.71	87.51	11935	25.34	17.02
100	183.57	87.60	11907	21.27	13.34

The RMSE, MAE and φ statistics for trip table errors now range between 183%-239%, 88%-96%, and 11907-13016, respectively. The link volume error statistics vary in the same range as observed for previous cases. An interesting observation can be made when we compare cases 3a and 4a. The error values turn out to be higher for 4a for all instances of available link volumes except the 90% availability case.

For this case again, the linear programming approach is seen to perform much better than the maximum entropy approach, with lower values for all measures of errors.

5.3.9 Case 4b: Relatively Small Error Trip Table (80% Cells with Info., $\psi = 0.9$)

This trip table has 80% of its cells obtained using $\psi = 0.9$ in Equation 5-1, and the remaining cells are zero.

From the output of the linear programming approach, as depicted in Table 5-17 below, several observations can be made. The RMSE(TT) and MAE(TT) now very between 78%-107% and 38%-56%, respectively. When comparing these ranges with the previous case, one may note that the upper and lower limits of the range are lower for case 4b as one would expect. However, one may see that Case 4b does not yield better results compared to every corresponding case of 4a. To illustrate this consider the 70% and 100% available link volume cases, for which the target table of the previous case yields lesser RMSE(TT) and MAE(TT) errors. As mentioned earlier, this may well be occurring because of the inconsistencies between the observed volumes and the assumed "correct" trip table that was used to generate the target trip tables.

Table 5-17: Performance of the Linear Programming Approach (Relatively Small Error Trip Table(80% Cells with Info., $\psi = 0.9$) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	77.84	37.88	6067	3.32	2.07
70	103.35	37.88	7894	7.88	4.00
90	102.84	55.76	8616	8.81	5.24
100	106.84	52.49	8155	9.51	4.86

The test results obtained with the maximum entropy approach are described in Table 5-18 below:

Table 5-18: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table(80% Cells with Info., $\psi = 0.9$) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	238.99	85.84	11483	25.02	22.54
70	237.81	72.11	10896	27.18	20.74
90	167.44	79.30	9878	25.2	16.84
100	163.09	79.30	9891	21.35	13.29

In contrast to the linear programming approach the maximum entropy approach records a significant improvement in its output for Case 4b, as compared to case 4a. Except for the instance where 50% link volumes are available all the trip table error statistics are lower for case 4b. The RMSE, MAE and φ statistics for trip table errors now range between 163%-239%, 72%-85%, and 9891-11482, respectively. Also, note that a clear monotonically decreasing behavior of the error statistics is observed, as the knowledge of available volumes increases.

Once again, a comparison of the two approaches reveals that the linear programming approach performs much better than the maximum entropy approach, with all the error statistics recording a lower value for the LPMLV model. Again, note that the error statistics are very high for both approaches.

5.3.10 Case 4c: Relatively Small Error Trip Table (100% Cells with Info., $\psi = 0.9$)

This trip table has 100% of its cells obtained using $\psi = 0.9$ in Equation 5-1. The target table for this case is expected to be closest to the "correct" trip table. And hence one would expect the best results from both models. The linear programming model shows very interesting results. First, it must be noted that the lowest values of the error statistics for all cases of available link volumes are obtained with this target table. The RMSE, MAE and φ statistics for trip table errors now range between 63%-93%, 30%-49%, and 4041-6602, respectively, which represent the least values of the upper and lower bounds among all the cases tested.

Table 5-19: Performance of the Linear Programming Approach (Relatively Small Error Trip Table(100% Cells with Info., $\psi = 0.9$) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	62.57	30.22	4041	4.17	2.27
70	77.07	49.39	5715	7.72	4.03
90	90.87	47.07	6602	9.73	5.11
100	93.17	45.96	6381	9.06	4.62

Observe that as the percentage of observed link volumes increases all trip table errors increase. This again could be attributed to the fact that the "correct" trip table is inconsistent with the available link volumes.

The maximum entropy approach performances also improves as seen in the Table 5-20 below:

Table 5-20: Performance of the Maximum Entropy Approach (Relatively Small Error Trip Table(80% Cells with Info., $\psi = 0.9$) as a Target)

	Measures of Error Rates				
% Avail. Vols.	RMSE % (TT)	MAE % (TT)	φ	RMSE % (Vol)	MAE % (Vol)
50	135.47	59.67	6285	23.92	21.7
70	129.40	54.81	5549	25.91	19.77
90	130.70	61.14	6897	25.62	16.94
100	127.24	61.30	6686	21.32	13.28

The results of the maximum entropy approach show significant improvement when compared to case 4b. However, these results are not necessarily the best among all the cases tested so far.

For this test, the RMSE, MAE and φ statistics for trip table errors now range between 127%-135%, 55%-61%, and 5549-6897, respectively, which are much lesser than the previous case. Note again that, for this approach too, we do not get very good results with a target trip table that is close to the true trip table. This offers further ground to suspect the inconsistency of the observed link volumes with the reported "correct" trip table.

The comparison of the linear programming and maximum entropy approaches, once again shows that the linear programming approach yields superior results.

5.4 Conclusions

The test results discussed in this chapter are graphically encapsulated in Figures 5.1-5.4..

The variation of the ϕ statistic for the linear programming and maximum entropy approaches are depicted in Figures 5-1 and 5-2, respectively. The ϕ statistic is seen to have a maximum value occurring with the structural table (struct). A gradual improvement is seen with the no prior information based target tables (n60, n80, n100). Further improvement is seen when we employ the relatively small error tables (s8_60, s8_80, s8_100) and (s9_60, s9_80, s9_100) which correspond to 60%, 80% and 100% of total trip interchange cell values information obtained using $\psi = 0.8$ and $\psi = 0.9$, respectively. For each of these 10 target trip tables, the linear programming model is seen to improve its performance with decreasing available link volume percentages. This could be attributed to errors due to the fact that the link volumes may not be consistent with OD flows and inconsistencies with observed link volume data. For the maximum entropy approach mixed trends are seen for the variation of the ϕ statistic with percentage available link volumes. In general, the linear programming approach has lower values for the ϕ statistic, as compared to the maximum entropy approach.

The variation of the link volume replication error for the linear programming and maximum entropy approaches as measured by the MAE(Vol) statistic are depicted in Figures 5-3 and 5-4, respectively. As seen in the figures, the linear programming approach has significantly lower values for this error statistic for every test case.

The figures also serve to illustrate the general superiority of the linear programming approach over the maximum entropy approach. Such an inference is again based on the lower values of the error statistics for the linear

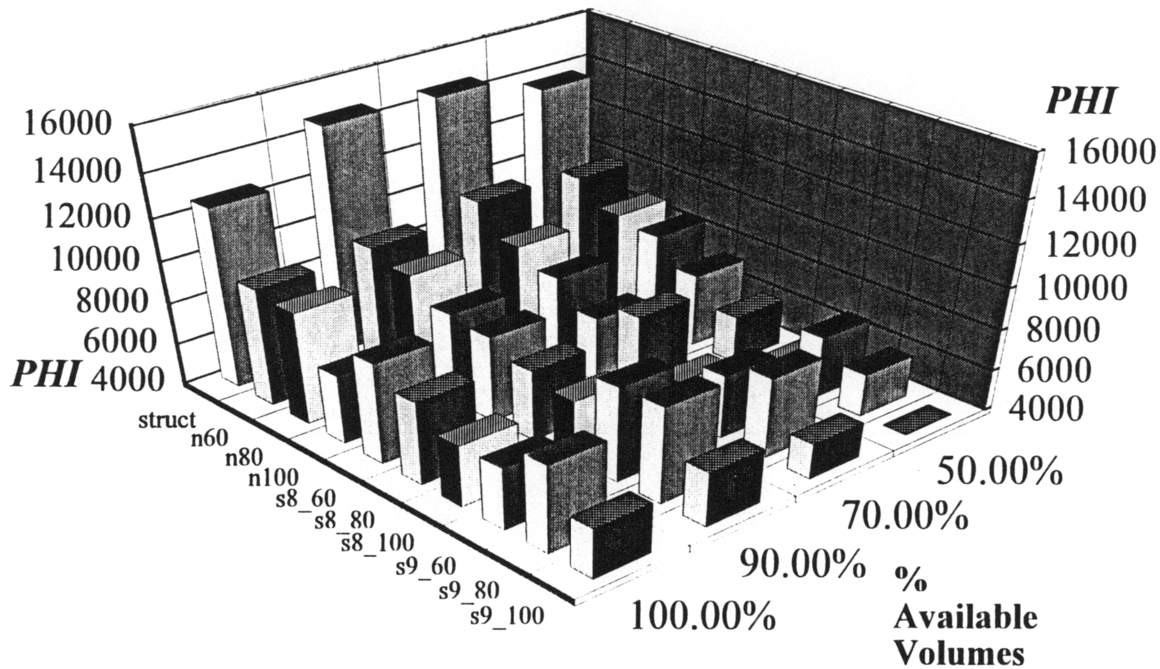


Fig 5-1 Purdue Network - LP Model PHI (TT) Values

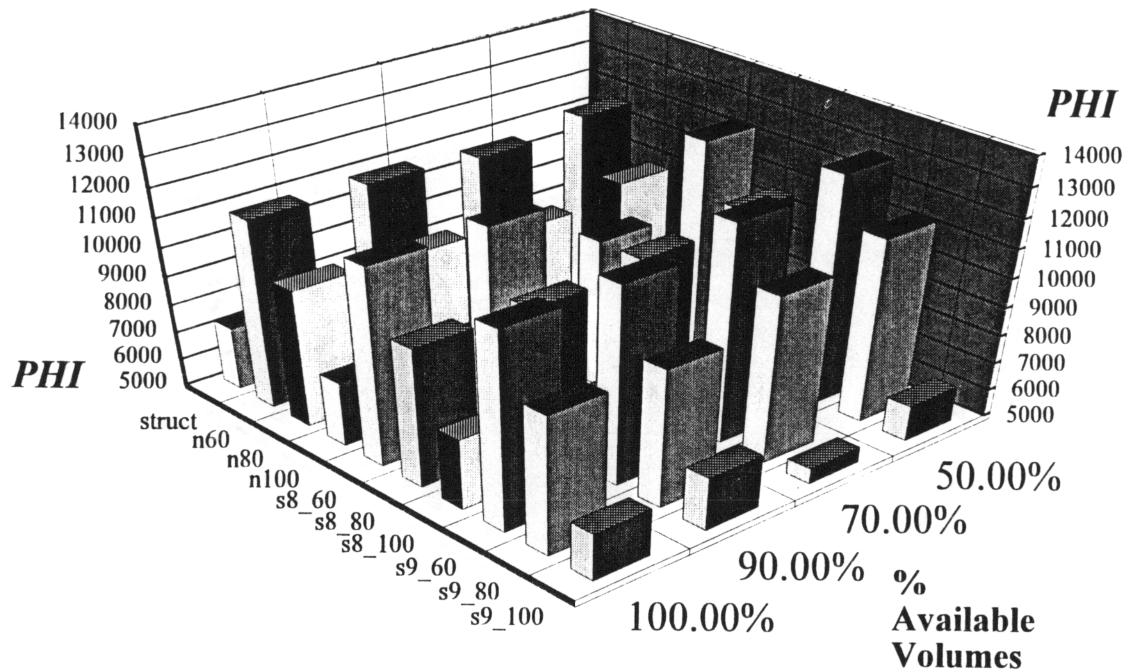


Fig 5-2 Purdue Network - THE Model PHI (TT) Values

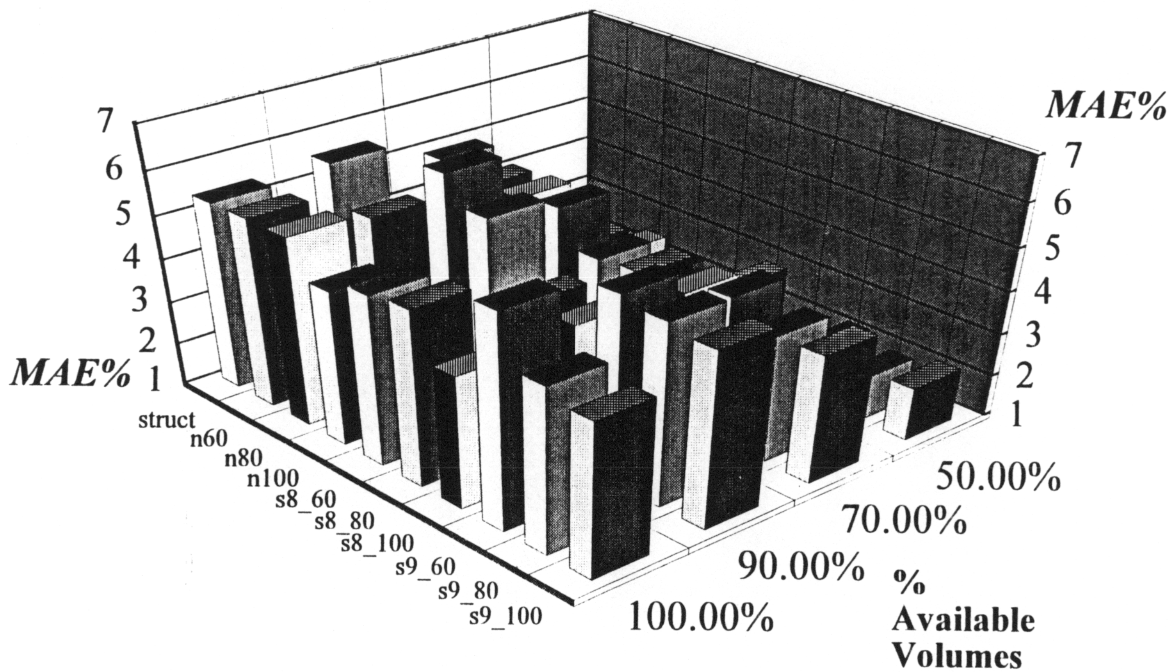


Fig 5-3 Purdue Network - LP Model %MAE (Vol) Values

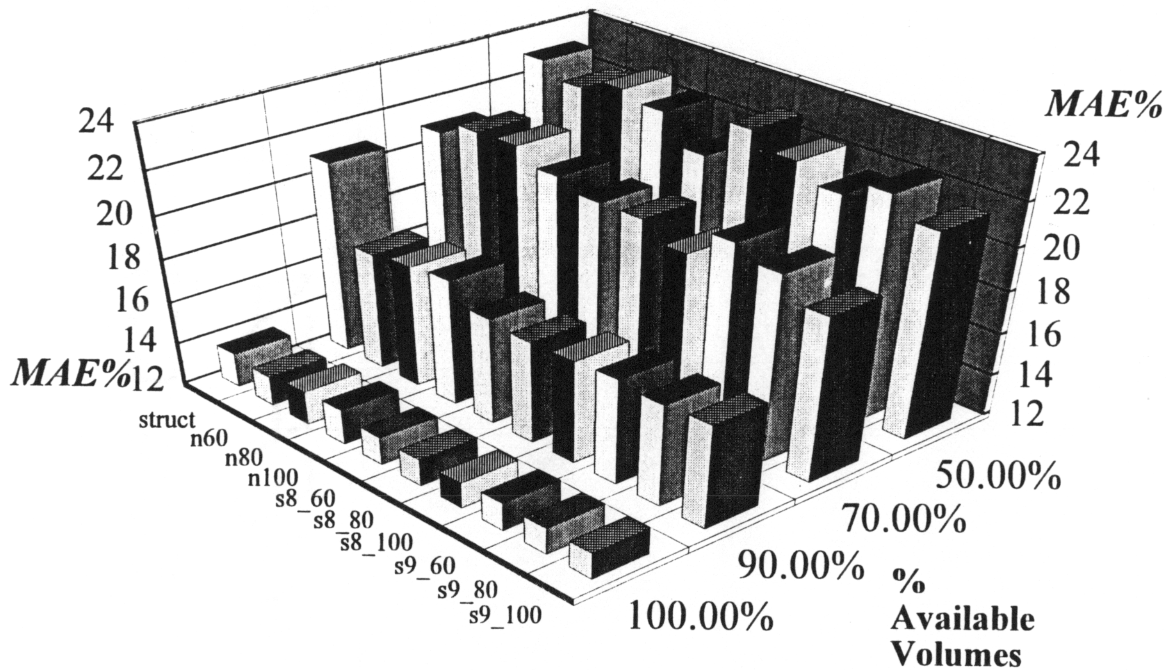


Fig 5-4 **Purdue Network - THE Model %MAE (Vol) Values**

programming model, as compared to the corresponding values obtained for the maximum entropy approach.

6. CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

In this research, a linear programming model for estimating O-D trip tables developed at Virginia Tech has been enhanced to accommodate a partial set of available link volume data. Its development and solution algorithm have also been detailed. The procedure uses an efficient modified column generation technique to derive a solution. Extensive tests to validate the model, using both artificial and real networks, have also been conducted. Conclusions on the model and recommendations for enhancing it are presented below.

6.1 Conclusions

The approach developed uses a sequence of linear programs to approximate a fundamentally nonlinear optimization problem that is employed to estimate origin-destination flows, given incomplete network flow information. The procedure utilizes shortest path network flow programming subproblems in order to determine a path decomposition of flow that will reproduce the observed flows as closely as possible, and that is driven by user equilibrium principles. The approach is very efficient, provided that the suggested heuristic procedure is adopted to approximately solve the shortest simple path problems, whenever the underlying network has negative cost circuits. This method is designed to accommodate the case in which it is required to produce a solution that has a tendency to match a specified, prior trip table, perhaps from among several optimal solutions to the model.

Extensive tests to validate the model using artificial and real networks are reported in Chapters 4 and 5, respectively. These results are also compared with the maximum entropy approach which is a popular approach for performing similar functions. Sensitivity tests of both the models to various extents of available link volume information as well as to different sets of prior trip tables were also carried out to evaluate the performance of these models.

The tests on the artificial network show that the linear programming approach produces good results for almost all cases of prior information (except the structural target table) and available link volume percentages. The errors for link volume replication and trip table errors turn out to be quite low. In comparison with the maximum entropy approach, the linear programming approach produces better quality results for most cases of prior information and available link volumes. In particular, the performance of the linear programming approach turns out to be much superior than the maximum entropy approach for the cases where the percentage available link volumes vary from 50%-70%. For the 80%-100% available link volume cases, although the performance of both models are comparable, the linear programming model continues to be marginally superior.

For the real network of Purdue University, the performances of both the approaches are not as good as that obtained for the artificial network. The performance of both the models for this network must be viewed with some caution, since the reliability of the available "correct" trip table cannot be proved. In addition, the inconsistency of this table with the observed link volumes further raises the question of reliability of the data. However, these tests have given the opportunity to address several issues relevant to the application of synthetic O-D estimation models to real networks. Again, in general, the linear programming approach proves superior to the maximum entropy approach except when we use a structural target trip table.

The tests indicate, in general, that the performance of both models improve as the quality of the target trip table improves. This suggests that these methods can be usefully applied in practical cases where an old trip table needs to be updated. While the tests do not rule out the possibility of applying these models to other scenarios, it serves to provide some understanding on some of the pitfalls that one might expect. In particular, the tests clearly highlight the limitation of applying these models when the lack of a good target table is coupled with the unavalibility of a significant amount of link volume information.

6.2 Recommendations for Further Research

Tests on a larger real network for which the *"true"* trip table has been established to a good degree of accuracy, need to be done in order to fully test the proposed model. This can lead to conlusions on the *true* performances of the models, and is a litmus test for validating any synthetic O-D estimation model. Also, the implementation of an advanced start procedure is expected to benefit the performance of the linear programming model. Furthermore, additional tests on how well the linear programming approach approximates the underlying nonlinear model would establish the accuracy of the sequential linear programming approach.

REFERENCES

Barbour R. and D. J. Fricker [1990], “Balancing Link Counts at Nodes Using a Variety of Criteria: An application in Local Area Traffic Assignment”, Transportation Research Record, 1220: 33-39.

R. Barbour, and J. D. Fricker [1993], "Estimating an Origin Destination Table Using A Method Based on Shortest Augmenting Paths," Accepted for publication in Transportation Research-A, forthcoming.

Bazaraa, M. S., J. J. Jarvis, and H. D. Sherali [1990], *Linear Programming and Network Flows*, Second Edition, John Wiley & Sons, Inc., New York.

Beagon, D. F. [1990], “Maximum Entropy Matrix Estimation (ME2),” Memorandum, Center Transportation Plannubg Staff, Boston, Massachusetts.

Brenninger-Gothe, M., K. O. Jornsten, and J. T. Lundgren [1989], “Estimation of Origin-Destinition Matrix from Traffic Counts Using Multiobjective Programming Formulations,” Transportation Research, 23(B): 257-265.

Bromage, E. J.[1988], *The Highway Emulator - Release 3.0* , Traffic Analysis and Design Group, Central Transportation Planning Staff.

Carey, M., C. Hendrickson and K. Siddharthan [1981], "A Method for Direct Estimation of Origin/Destination Trip Matrices," Transportation Science, 15(1):32-49.

Cascetta, E.[1984], "Estimation of Trip Matrices From Traffic Counts and Survey Data: A Generalized Least Squares Estimater," Transportation Research-B, 18B(4/5): 289-299

Chin, S., H. Hwang and T. Pei [1994], "Using Neural Networks to Synthesize Origin-Destination Flows in a Traffic Circle", Paper No. 940353, Transportation Research Board, Washington, D.C.

Dowling, R. G and A. D. May [1984], "Comparison of Small Area O-D Estimation Techniques," unpublished.

Erlander, S., S. Nguyen, and N. F. Stewart [1979], " On the Calibration of the Combined Distribution--Assignment Model," Transportation Research, 13(B): 259-267.

Fisk, C. S. [1988], "On Combining Maximum Entropy Trip Estimation with User Optimal Assignment," Transportation Research, 22(B): 69-79.

Fisk, C. S. [1989], "Trip Matrix Estimation from Link Traffic Counts: The Congested Network Case," Transportation Research, 23(B): 331-336.

Fisk, C. S. and D. E. Boyce [1983], "A note on Trip Matrix Estimation From Link Traffic Count Data," Transportation Research, 17(B): 245-250.

Gur, J., M. Turnquist, M. Schneider, L. Leblanc and D. Kurth [1980], *Estimation of an Origin Destination Trip Table on Observed Link Volumes and Turning Movements.- Volume 1: Technical Report,* FHWA, Rept. FHWA/RD-80/034

Gaudry, M. and L, Lamarre, (1979), "Estimating Origin-Destination Matrices from Traffic Counts: A Simple Linear Intercity Model for Quebec," The Logistics and Transportation Review, 15: 631-642.

Han, A. F. and R. G. Dowling, E. C. Sullivan, and A. D. May [1981], "Deriving Origin-Destinition Information from Routinely Collected Traffic Counts," Volume II: Trip Table Synthesis for Multipath Networks, RR81-9, Institute for Transportation Studies, University of California, Berkeley.

Han, A. F. and E. C.. Sullivan [1983], “Trip Table Synthesis for CBD Networks: Evalution of the LINKOD Model,” Transportation Research Record 944: 106-112.

Hogberg, P. [1976], “Estimation of Parameters in Models for Traffic Prediction: A Non-Linear Regression Approach,” Transportation Research, 10: 263-265.

Holm, T. Jensen, S. K. Nielsen, A. Christensen, B. Johnsen and G. Ronby [1976], "Calibrating traffic models on traffic census results only," Traffic Engineering and Control, April:137-140.

Kennington, J. L., and R. V. Helgason [1980], *Algorithms for Network Flow Programming*, John Wiley and Sons, New York.

Kurth, D. M. Schneider and Y. Gur [1979], “Small-Area Trip Distribution Model,” Transportation Research Record 728:35-40, Washington, D. C.

Lasdon, L. S. [1970], *Optimization Theory for Large Systems*, MacMillan.

Low, D. E. [1972], " A New Approach to Transportation Systems Modeling," Traffic Quarterly, 26:391-404

Maher, M. J. [1983], "Inferences on Trip Matrices From Observations on Link Volumes: A Bayesian Statistical Approach," Transportation Research, 17B(6):435-447.

Murchland, J.. [1977], “The Multi-proportional Problem,” University College, London, Reseach Note JDM 263 (Unpublished).

Nguyen, S. [1977], "Estimating an OD Matrix from Network Data: A Network Equilibrium Approach," University of Montreal Publication No. 60.

Nguyen, S. [1984], "Estimation Origin-Distinition Matrices from Observed Flows," in *Transportation Planning Models,* Editor: M. Florian, Elsevier Science Publishers B.V., Netherlands, 363-380.

O'Neill, W. A., [1987], "Origin-Destinition Trip Table Estimation Using Traffic Counts," Ph.D dissertation, University of New York ar Buffalo, New York.

Robillard, P., [1975], "Estimating the O-D Matrix from Observed Link Volumes," Transportation Research, 9: 123-128

Sherali, D., R. Sivanandan and A.G. Hobeika [1994], " A Linear Programming Approach for Synthesizing Origin Destination (O-D) Trip Tables from Link Traffic Volumes," Transportation Research-B, 28B: 213-233.

Sherali, D., R. Sivanandan, A.G. Hobeika and A. Narayanan[1994], " Estimating Missing Link Volumes in a Traffic Network-A Linear Programming Approach," Presentation at the TRB Annual Meeting, Washington, D.C.

Shewey P. J. H.[1983], "An Improved Algorithm for Matching Partial Registration Numbers", Transportation Research, 17B, 391-397.

Sivanandan [1991], " A Linear Programming Approach for Synthesizing Origin - Destination Trip Tables From Link Traffic Volumes," Ph.D. Dissertation, Virginia Polytechnic Institute & State University, Blacksburg, Virginia.

Turnquist, M., and Y. Gur, [1979], "Estimation of Trip Tables from Observed Link Volumes," Transportation Research Record 730: 1-6, Washington D. C.

Van Vliet, D. and L. G. Willumsen [1981], "Validation of the ME2 Model for Estimmating Trip Matrices from Traffic Counts," Proceeding of the Eighth International Symposium on Transportation and Traffic Theory, Editors: V. F. Hurdle, E. Hauer and G. M.. Stuart, University of Toronto Press, 640-655.

Van Zuylen, H. J., [1978], "The Information Minimization Method: Validity and Applicability to Transport Planning," in *New Developments in Modeling Travel Demand and Urban Systems*, G. R. H. Jansen et al., eds. Saxon, Farnboroug

Van Zuylen, H. J., and L. G. Willumsen,[1980], "The Most Likely Trip Matrix Estimated from Traffic-Counts," Transportation Reseach-B,14B:291-293.

Walting, D. P., and M. J. Maher[1992], "A Statistical Procedure for Estimating a Mean Origin-Destination Matrix from a Partial Registration Plate Survey", Transportation Research, B, 26B, 171-193.

Walting, D. P., and D. R. Grey[1991], "Analysis of Partial Registration Plate Data Using a Model with Poisson Input and Output", Proceedings of the International Conference on Mathematics in Transport Planning and Control (held by the Institute of Mathematics and Its Applications at the University of Wales College of Cardiff, September 1989), Oxford University Press.

Wills, M. [1977], "Non Linear Estimation of Origin-Destination Flows and Gravity Model Parameters from Traffic Counts on Links," Department of Geography, University of British Columbia, Mimeographed Notes.

Willumsen, L. G., [1978], "Estimation of an O-D Matrix from Traffic Counts - A Review," Working Paper 99, Institute of Transportation Studies, University of Leeds, England.

Xu, W. and Y. Chan [1993], "Estimating an Origin-Destination Matrix With Fussy Weights, Part II: Case Studies", Tansportation Planning and Technology, 17: 145-163.

Yang, H., T. Akiyama, and T. Sasaki [1992], A Neural Network Approach to the Identification of Real Time Origin-Distination from Traffic Counts," Proceedings of the International Conference on Artificial Intelligence Application in Transportation Engineering, 253-269.

VITA

Arvind Narayanan was born on May 1, 1970, in Tirunelveli, India. He obtained a Bachelor's degree in Civil Engineering from the Indian Institute of Technology, Madras. He joined the Masters program in Civil Engineering, specializing in Transportation Engineering, at Virginia Tech in Fall 1992. He joined the Virginia Tech Center for Transportation Research as a Research Associate in May 1994. He intends to pursue a doctoral program in Operations Research.