Report BD-432

Final Report

Accuracy Standards and Data Coverage Requirements for Model Validation in FSUTMS

Prepared by

Chang-Jen Lan, Assistant Professor
Jibing Li, Research Assistant
Xiaojun Gu, Research Assistant

Department of Civil, Architectural, and Environmental Engineering
University of Miami
1251 Memorial Dr., 325 McArthur
Coral Gables, Miami, Florida 33146
Phone: (305) 284-6447, Fax: (305) 284-3492
E-mail: clan@miami.edu

and

Min-Tang Li, Senior Research Associate
Rax Jung, Research Associate
Albert Gan, Associate Professor

Department of Civil and Environmental Engineering
Florida International University
10555 West Flagler Street, EC 3680
Miami, Florida 33174
Phone: (305) 348-3116, Fax: (305) 348-2802
E-mail: gana@fiu.edu

in cooperation with the

State of Florida Department of Transportation
605 Suwannee Street, MS 30
Tallahassee, Florida 32399

October 2005

DISCLAIMER

The opinions, findings and conclusions expressed in this publication are those of the authors and not necessarily those of the State Florida Department of Transportation.

ACKNOWLEDGEMENTS

This research was funded by the Research Center of the Florida Department of Transportation (FDOT). The authors wish to thank the project manager, Mr. Shi-Chiang Li, AICP, of the Office of Planning and Environmental Management at the FDOT District 4 Office, for his guidance and advice throughout the course of the research. Mr. Thomas Rossi from Cambridge Systematics, Inc., is also greatly appreciated for reviewing the draft and providing useful comments.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Accuracy Standards and Data Coverage Requirements for Model Validation in FSUTMS		5. Report Date October 2005 6. Performing Organization Code
7. Author(s) Chang-Jen Lan, Jibing Li, Xiaojun Gu, Min-Tang Li, Rax Jung, and Albert Gan		8. Performing Organization Report No.
9. Performing Organization Name and Address Dept. of Civil, Architectural and Environmental, University of Miami, 1251 Memorial Dr., 325 McArthur, Coral Gables, Fl 33146		10. Work Unit No. (TRAIS) 11. Contract or Grant No. BD-432
12. Sponsoring Agency Name and Address State of Florida Department of Transportation 605 Suwannee Street, MS 30, Tallahassee, FL 32399		13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code

15. Supplementary Notes
Mr. Shi-Chiang Li, AICP, of the Florida Department of Transportation District 4 Office served as the project manager.

16. Abstract
Model validation is an important step in travel demand modeling. It serves to ensure that the calibrated model produces outputs that are consistent with the observed data. In the Florida Standard Urban Transportation Model Structure (FSUTMS), various consistency checks for the validation of the distribution and assignment models are suggested for highway networks. These checks involve the comparison of predicted link volumes with ground counts at screenlines, cordon lines, and/or cutlines. Statistics such as system-wide vehicle-mile-traveled (VMT), vehicle-hour-traveled (VHT), link volume by facility type, area type, facility size (in term of number of lanes), and volume group Root Mean Square Error (RMSE) are major criteria used for validation. Errors between model outputs and what are observed from the actual networks may come from such sources as measurement, sampling, model misspecification, prediction, and aggregation process. In this study, sampling and model misspecification errors were considered in deriving accuracy standards for highway networks. The required number of count stations to be sampled for each type of facility to achieve a specific accuracy standard at a certain level of confidence was also derived and a general rule of thumb was recommended for validation purposes. For transit networks, the accuracy standards were derived to account for statistical variability, aggregation errors, and prediction errors. Errors in modal share subjected to aggregation and prediction variability were derived for both multinomial logit and nested logit models. However, no results are presented due to a lack of data support.

17. Key Words	18. Distribution Statement
Accuracy Standard, Allowable Error Limits, Model Validation, Data Coverage Requirement, Sampling Error, Aggregation Error, Prediction Error	No restrictions

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	144	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

EXECUTIVE SUMMARY

Model validation is an important step in travel demand modeling. It serves to ensure that the calibrated model produces outputs that are consistent with the observed data. In the Florida Standard Urban Transportation Model Structure (FSUTMS) highway modeling arena, various consistency checks for the validation of the distribution and assignment models are suggested for highway networks. They include: (1) the comparison of the gravity model trip length outputs with those from the trip survey data from which the model was calibrated; (2) the comparison of predicted link volumes with ground counts at screenlines, cordon lines, and/or cutlines; and (3) the comparison between the derived statistics from the subject model with those from case studies of other areas with similar characteristics. Statistics such as system-wide vehicle-mile-traveled (VMT), vehicle-hour-traveled (VHT), link volume by facility type, area type, facility size (in terms of number of lanes), and volume group Root Mean Square Error (RMSE) are major criteria for validation.

Errors between model outputs and what are observed from the networks may come from different sources such as measurement, sampling, model misspecification, prediction, and aggregation process. In highway networks, sampling and model misspecification errors were considered in deriving the allowable error limits (accuracy standards). Accuracy standards to account for sampling errors were adapted from the Guide to Urban Traffic Volume Counting (GUTVC) developed by the Federal Highway Administration (FHWA). The guide provides a uniform method of estimating the VMT based on random selection, and sampling of highway links that are stratified by volume, facility type, and area type. All the available detector data collected from permanent and portable count stations in the State of Florida in year 2000 were used to compile the necessary statistical information. However, for cost consideration, only 7-day ADT count samples were used to construct the accuracy standards. The t and F tests on the first two moments of statistics indicated that the 7-day ADT count samples can be used to sufficiently represent the AADT. The size reduction is considered fairly cost-effective, given the sampling schedule of the weekly data is chosen appropriately. In addition, depending on the time of day, i.e., morning rush hours, mid-day, and afternoon rush hours, accuracy standards were developed for 68%, 85%, and 95% levels of significance. The results show that there is no significant difference in allowable error limits for different times of day. Except for several facility types with a limited sample size (less than 50), almost all error limits are below 10% (average is 5%) at the 95% level of confidence. One-way and toll facilities were slightly higher than others, indicating that the variability of traffic counts observed on these facilities was in general higher.

On the other hand, allowable error limits to account for model misspecification errors were derived based on the perception error associated with travelers in choosing routes. The common belief is that the perception error in route choice will significantly drive traffic flow pattern away from the user equilibrium condition. To segregate this error from other sources, deterministic simulations were conducted by varying the setting of the dispersion factors in the Stochastic User Equilibrium (SUE) assignment method while keeping all other model parameters unchanged. Although not used in FSUTMS, the SUE assignment method was used in this study to derive the assignment error for its capability to model the perception error in route choice and to handle the condition in which travelers, in general, do not have perfect information on travel time over the network. The resulting assignment errors are an exponentially decreasing function of link assigned volume. Finally, the overall standards are derived by adding a 5% sampling error on the model misspecification error, assuming that the sampling error is independent of the model misspecification error. As a comparison, the FHWA standards are about 10% higher than the proposed standards for ADTs equal to 15,000 vpd or higher, and are substantially lower for ADTs below 15,000 vpd. The proposed accuracy standards meet the Michigan DOT standards quite closely, except when the average assigned volume falls between 2,500 vpd and 15,000 vpd. The current practices by both FHWA and Michigan DOT seem to underestimate the assignment error by 10 to 15 percent within the abovementioned volume range compared to the proposed standards.

In addition, the current area-wide RMSE standard of 35-50% seems arbitrary. The standard should be a composite measure of the actual (or estimated) distribution of the roadway AADT groups and the corresponding allowable error in each volume group. In general, for rural study areas, the allowable error should be relatively high since the majority of the roadway ADTs are low. On the other hand, the allowable error should be relatively low for urban areas because of high ADTs on the majority of the urban roadways. It is extremely difficult to assign a reasonable accuracy standard if the information on distribution of the roadway ADTs in the study area is not available. Posting accuracy standards by a coarse classification of area types, such as the one used by Reno, will not be reasonable if the actual distribution of roadway ADTs are significantly different from that of the designated area types. Accordingly, a worksheet was designed and proposed in this study for planners in different jurisdictional areas to develop their own standards. Finally, this study recommends eliminating the facility-type and size-of-facility specific accuracy standards, since they have created some contradictions to the standards by volume-group. Current practices show that different standards were applied onto different roadway facilities even though they have carried similar traffic volumes. While this may seem to contradict to the initial intent of this study to accommodate the two-digit codes for facility type, it is suggested that these extra categories be eliminated to simplify and standardize the structure of the current accuracy standards for highway networks.

The other dimension of problem encountered in transportation planning is to determine how many counts need to be sampled for each type of facility to achieve a specific accuracy standard at a certain level of confidence. Based on the current accuracy standards, the required sample sizes are computed for each functional class area-wide at 68%, 85%, and 95% levels of confidence. The results show that the required sample size increases significantly as the level of confidence increases. However, the required sample size remains at an affordable range of 10 and 60. Except for one-way facility, which may require a larger sample such as 50, the use of 30 count stations is a rule of thumb for general sampling purposes.

For transit networks, the accuracy standards were derived to account for statistical variability, aggregation errors, and prediction errors. Errors in modal share subjected to aggregation and prediction variability were derived for both multinomial logit and nested logit models based on a second-order Taylor expansion theory. Variability of the modal share was also derived to establish confidence intervals at a user-specified level of significance as an alternative form of accuracy standards. Due to the limitation of data sources in both quantity and quality in supporting estimation of the multinomial and/or nested logit models, however, the accuracy standards to account for aggregation and prediction errors could not be developed in this study.

On the other hand, errors due to statistical variability are significantly lower than the current accuracy standards in terms of daily ridership, which is expected since no other sources of errors were considered.

TABLE OF CONTENTS

DISCLAIMER……………………………………………………………………..…………………………i
ACKNOWLEDGEMENTS……………………………………………………………..…………i
EXECUTIVE SUMMARY……………………………………………………..………...iii
TABLE OF CONTENTS…………………………………………………………………………vi
LIST OF FIGURES…………………………………………………………..………………viii
LIST OF TABLES………………………………………………………..…………………ix

CHAPTER ONE
INTRODUCTION……………………………………………………..……………………. 1
1.1. Motivation……………………………………………………………………………….1
1.2. Overview of the Model Validation Process………………………………………..………2
1.3. Sources of Error……………………………………………………………………………3
1.4. Objectives…………………………………………………………………..…...……… 4
1.5. Report Organization……………………………………………………….......………4

CHAPTER TWO
LITERAURE REVIEW ON ACCURACY STANDARDS FOR HIGHWAY
NETWORKS…………………………………………………………………………...……… 5
2.1. Current Practices…………………………………………………………………...………5
2.2. Calibration and Validation on Trip Generation……………………………………………… 9
2.3. Calibration and Validation on Trip Distribution…………………………………….………13
2.4. Calibration and Validation on Modal Split……………………………………......………17
2.5. Calibration and Validation on Highway Network……………………………………..……… 22
2.5.1. Current Practices………………………………………………………………..………22
2.5.2. Other Related Research Work…………………………………………………….……… 32

CHAPTER THREE
LITERAURE REVIEW ON ACCURACY STANDARDS FOR TRANSIT
NETWORKS…………………………………………………………………………….……… 38
3.1 Transit Validation Process……………………………………………………………..……… 38
3.2 Stepwise Validation…………………………………………………………………….……… 39
3.3 Transit Accuracy Standards……………………………………………………….……… 41
3.4 Summary…………………………………………………………………………..……… 43

CHAPTER FOUR
ACCURACY STANDARDS FOR HIGHWAY NETWORKS…………………………………45
4.1. Accuracy Standards for Assigned Volume……………………………………….……… 45
4.1.1. Accuracy Standards Due to Data Statistical Variation……………………....……… 46
4.1.2. Accuracy Standard Due to Model Specification Error…………………………...……… 60
4.2. Required Sample Size for Assigned Volume………………………………………….………66
4.3. Accuracy Standards for Vehicle Mile Traveled……………………………………….………69
4.4. Required Sample Size for Vehicle Mile Traveled…………………………………….………. 70

CHAPTER FIVE

ACCURACY STANDARDS FOR TRANSIT NETWORKS……………………………..……… 85
5.1 Introduction……………………………………………………………………..………85
5.2 Background……………………………………………………………………..………86
5.3 Errors in Modal Split…………………………………………………………...……… 90
5.4 Statistical Variability of Transit Ridership………………………………………………… 104

CHAPTER SIX
CONCLUSIONS AND FUTURE WORKDS………………………………………………….121
6.1 Conclusions……………………………………………………………………..……..121
6.2 Future Works…………………………………………………………………...………123

REFERENCES……………………………………………………………………………… 124
APPENDICES……………………………………………………………………………… 130
Appendix I. Hypothesis Testing on the Weekly ADT………………………………….………130
Appendix II. Derivatives for the MNL and NL models……………………………..……….143

LIST OF FIGURES

Figure 2.1 Error Limits as a Function of AADT for Individual Links…………………..………25
Figure 2.2 Maximum Desirable Deviations Proposed by CALTRAN…………………..……… 26
Figure 4.1 Coefficient of Variation for All Facilities (Morning rush hours)………….………49
Figure 4.2 Coefficient of Variation for All Facilities (Midday)………………………………… 49
Figure 4.3 Coefficient of Variation for All Facilities (Afternoon rush hours)…………..……… 50
Figure 4.4 Assignment Errors as a Function of Dispersion Factors……………………..……. 61
Figure 4.5 Assignment Errors versus Average Assigned Volumes…………………...………63
Figure 5.1 The Nested Logit Structure of the Modal Choice in FSUTMS………………………92
Figure 5.1 Relationship between CV and Weekday Average Ridership for Jacksonville……...119
Figure 5.2 Relationship between CV and Weekday Average Ridership for Miami……………119
Figure 5.3 Relationship between CV and Weekday Average Ridership for Tampa…………... 119

LIST OF TABLES

Table 2.1 Accuracy Standards for Trip Distribution Validation Used in FSUTMS………………7
Table 2.2 Accuracy Standards for Traffic Assignment Used in FSUTMS……………....……… 8
Table 2.3 Specific Accuracy Standards in Traffic Assignment as a Function of Facility Type, Facility Size, and ADT……………………………………..………..…………….. 8
Table 2.4 Typical Distributions of VMT in Urban Area by Facility Type…...................………24
Table 2.5 Accuracy Standards on Volume by AADT (TMIP, 1997)……………...……..……… 24
Table 2.6 Percent RMSE by Facility and Area Types in Reno (TMIP, 1997)………......……… 24
Table 2.7 Validation Standards and Criteria for Small Jurisdictions in California……...……… 27
Table 2.8 Suggested Traffic Counting Accuracy Standards……..………………………………29
Table 3.1 Estimated Accuracy of Parameters for Public Transit Loading (Robbins, 1978)……. 42
Table 3.2 Initial FDOT Transit Validation Standards (COMSIS Corporation, 1981)…..……….43
Table 3.3 Updated FDOT Transit Validation Standards (Schimpeler-Corradino Assoc., 1984)……………………………………………………………………………… 43
Table 4.1 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 7-day (Week) Sample Size (Morning Rush Hours)……………. 51
Table 4.2 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 7-day (Week) Sample Size (Mid-day)…………………………. 52
Table 4.3 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 7-day (Week) Sample Size (Afternoon Rush Hours)…………...53
Table 4.4 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 5-day Sample Size (Morning Rush Hours)…………………….. 54
Table 4.5 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 5-day Sample Size (Mid-day)………………………………….. 55
Table 4.6 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 5-day Sample Size (Afternoon Rush Hours)……………………56
Table 4.7 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 3-day Sample Size (Morning Rush Hours)……………………...57
Table 4.8 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 3-day Sample Size (Mid-day)…………………………………...58
Table 4.9 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 3-day Sample Size (Afternoon Rush Hours)……………………59
Table 2.5a Proposed Accuracy Standards on Volume by AADT……………………..…….... 63
Table 2.5b Proposed Accuracy Standards Worksheet on the Area-wide Allowable Error…….... 64
Table 2.2a Proposed Modified Accuracy Standards for Traffic Assignment Used in FSUTMS……………………………………………………………………….64
Table 4.10 Required Sample Size Based on Current Accuracy Standards (Morning Rush Hours)………………………………………………………………………….. 66
Table 4.11 Required Sample Size Based on Current Accuracy Standards (Mid-day)………….. 67
Table 4.12 Required Sample Size Based on Current Accuracy Standards (Afternoon Rush Hours)……………………………………………………………………….……..68
Table 4.13 Accuracy Standards for Facility Type at 68%, 85% and 95% Confidence Level……………………………………………………………………………70
Table 4.14a Sample Size Computation for Freeway and Expressway at Different Confidence Levels (Relative Error = 0.05, Number of Mileages)…………………...74

Table 4.14b Sample Size Computation for Freeway and Expressway at Different Confidence Levels (Relative Error = 0.15, Number of Mileages)…………………...74
Table 4.14c Sample Size Computation for Freeway and Expressway at Different Confidence Levels (Relative Error = 0.25, Number of Mileages)…………………...74
Table 4.15a Sample Size Computation for Divided Arterial at Different Confidence Levels (Relative Error = 0.05, Number of Mileages)…………………………..…… 75
Table 4.15b Sample Size Computation for Divided Arterial at Different Confidence Levels (Relative Error = 0.15, Number of Mileages)…………………………..…… 75
Table 4.15c Sample Size Computation for Divided Arterial at Different Confidence Levels (Relative Error = 0.25, Number of Mileages)…………………………..…… 75
Table 4.16a Sample Size Computation for Undivided Arterial at Different Confidence Levels (Relative Error = 0.05, Number of Mileages)…………………………..…… 76
Table 4.16b Sample Size Computation for Undivided Arterial at Different Confidence Levels (Relative Error = 0.15, Number of Mileages)…………………………..…… 76
Table 4.16c Sample Size Computation for Undivided Arterial at Different Confidence Levels (Relative Error = 0.25, Number of Mileages)…………………………..…… 76
Table 4.17a Sample Size Computation for One-Way Facility at Different Confidence Levels (Relative Error = 0.05, Number of Mileages)……..………………………… 77
Table 4.17b Sample Size Computation for One-Way Facility at Different Confidence Levels (Relative Error = 0.15, Number of Mileages)……..………………………… 77
Table 4.17c Sample Size Computation for One-Way Facility at Different Confidence Levels (Relative Error = 0.25, Number of Mileages)……..………………………… 77
Table 4.18a Sample Size Computation for Collectors at Different Confidence Levels (Relative Error = 0.05, Number of Mileages)……..…………………………………78
Table 4.18b Sample Size Computation for Collectors at Different Confidence Levels (Relative Error = 0.15, Number of Mileages)……..…………………………………78
Table 4.18c Sample Size Computation for Collectors at Different Confidence Levels (Relative Error = 0.25, Number of Mileages)……..…………………………………78
Table 4.19a Practical Sample Size (Number of Counts) for Freeway and Expressway (Relative Error = 0.05)……………………………………………………………80
Table 4.19b Practical Sample Size (Number of Counts) for Freeway and Expressway (Relative Error = 0.15)……………………………………………………………80
Table 4.19c Practical Sample Size (Number of Counts) for Freeway and Expressway (Relative Error = 0.25)……………………………………………………………80
Table 4.20a Practical Sample Size (Number of Counts) for Divided Arterial (Relative Error = 0.05)……..…………………………………………………...81
Table 4.20b Practical Sample Size (Number of Counts) for Divided Arterial (Relative Error = 0.15)……..…………………………………………………...81
Table 4.20c Practical Sample Size (Number of Counts) for Divided Arterial (Relative Error = 0.25)……..…………………………………………………...81
Table 4.21a Practical Sample Size (Number of Counts) for Undivided Arterial (Relative Error = 0.05)……..…………………………………………………...82
Table 4.21b Practical Sample Size (Number of Counts) for Undivided Arterial (Relative Error = 0.15)……..…………………………………………………...82
Table 4.21c Practical Sample Size (Number of Counts) for Undivided Arterial (Relative Error = 0.25)……..…………………………………………………...82

Table 4.22a Practical Sample Size (Number of Counts) for One-Way Facility (Relative Error = 0.05)……..…………………………………………..……….83
Table 4.22b Practical Sample Size (Number of Counts) for One-Way Facility (Relative Error = 0.15)……..…………………………………………..……….83
Table 4.22c Practical Sample Size (Number of Counts) for One-Way Facility (Relative Error = 0.25)……………………………………………………… 83
Table 4.23a Practical Sample Size (Number of Counts) for Collectors (Relative Error = 0.05)………………………………………………………84
Table 4.23b Practical Sample Size (Number of Counts) for Collectors (Relative Error = 0.15)………………………………………………………84
Table 4.23c Practical Sample Size (Number of Counts) for Collectors (Relative Error = 0.25)………………………………………………………84
Table 5.1 % RMSEs due to Classification……..…………………………………...………89
Table 5.2 % RMS for Three Scales of Geographic Classifications……..…………….………89
Table 5.3 % RMS for Three Scales of Geographic Classifications……..…………….………90
Table 5.4 Modes in Various MNL Mode Choice Models in FSUTMS……..…………..……….91
Table 5.5 Weekday Average Transit Ridership of Jacksonville for Years 2003-2004.....…….. 107
Table 5.6 Weekday Average Transit Ridership of Miami for Years 2002-2003……………… 109
Table 5.7 Weekday Average Transit Ridership of Tampa for Years 2003-2004……………… 112
Table 5.8 System-wide Error Percentage Resulting from Transit Ridership Variability……… 114
Table 5.9 Route Errors Percentage Resulting from Transit Ridership Variability in Jacksonville……..……………………………..…………………………………115
Table 5.10 Route Errors Percentage Resulting from Transit Ridership Variability in Miami…………………………..…………………………………...………116
Table 5.11 Route Errors Percentage Resulting from Transit Ridership Variability in Tampa……..…………………………………………………………………...118
Table 5.12 Comparison between the FDOT Accuracy Standard and Transit Ridership Variation……..……………………………………………….…………………..120
Table I.1 Freeway and Expressway Weekly and Annual Morning-ADT t-test and F-test……..…………………………………………………….…………... 127
Table I.2 Divided Arterial Weekly and Annual Morning-ADT t-test and F-test……………… 128
Table I.3 Undivided Arterial Weekly and Annual Morning-ADT t-test and F-test…....……… 130
Table I.4 Collectors Weekly and Annual Morning-ADT t-test and F-test…………...……… 130
Table I.5 Toll Facility Weekly and Annual Morning-ADT t-test and F-test………...………130
Table I.6 Freeway and Expressway Weekly and Annual Midday-ADT t-test and F-test……..……………………………………………………..…………...131
Table I.7 Divided Arterial Weekly and Annual Midday-ADT t-test and F-test……...………132
Table I.8 Undivided Arterial Weekly and Annual Midday-ADT t-test and F-test……..………134
Table I.9 Collectors Weekly and Annual Midday-ADT t-test and F-test…………....………134
Table I.10 Toll Facility Weekly and Annual Midday-ADT t-test and F-test………..……… 134
Table I.11 Freeway and Expressway Weekly and Annual Afternoon-ADT t-test and F-test……..……………………………………………………….………135
Table I.12 Divided Arterial Weekly and Annual Afternoon-ADT t-test and F-test……………136
Table I.13 Undivided Arterial Weekly and Annual Afternoon-ADT t-test and F-test…………138
Table I.14 Collectors Weekly and Annual Afternoon-ADT t-test and F-test………..………138
Table I.15 Toll Facility Weekly and Annual Afternoon-ADT t-test and F-test……..……… 138

CHAPTER ONE

INTRODUCTION

1.1. Motivation

Model validation is an important step in travel demand modeling. It serves to ensure that the calibrated model produces outputs that are consistent with the observed data. In the Florida Standard Urban Transportation Model Structure (FSUTMS) highway modeling arena, various consistency checks have been suggested for the validation of the distribution and assignment models for highway networks. They include: (1) the comparison of the gravity model trip length outputs with those from the trip survey data from which the model was calibrated; (2) the comparison of predicted link volumes with ground counts at screenlines, cordon lines, and/or cutlines; and (3) the comparison between the derived statistics from the subject model with those from case studies of other areas with similar characteristics. Statistics such as system-wide vehicle-mile-traveled (VMT), vehicle-hour-traveled (VHT), link volume by facility type, area type, facility size (in terms of number of lanes), and volume group Root Mean Square Error (RMSE) are major criteria used for validation. In the transit modeling arena, accuracy standards for validating transit assignments have been extremely limited.

In FSUTMS, specific accuracy standards are prescribed in *Urban Transportation Planning Model Update-Phase II, Task C, Develop Standard Distribution and Assignment Models* (COMSIS, 1981) for each validation criterion to determine if further model diagnoses are needed. These accuracy standards are established as the critical benchmarks for gauging the model prediction capability and establishing the minimal levels for highway demand forecast model acceptance, primarily for determining number of lanes for new roads. While these standards have been used for more than two decades, they are inadequate when the models are used in applications beyond determining the number of lanes. Consequently, the current standards need to be amended. Further, new model validation standards should also be established for the following reasons:

1. The constant threshold values need to be enhanced to recognize the dependency of the accuracy of the observed data, to which the model outputs are compared, on the underlying network and traffic conditions, the number of count samples, and the desired level of statistical confidence. Statistically, the range of error allowance should be in proportion to the statistical variability inherent in the sampled counts in combination with the prediction variability associated with the planning models, and in reversal proportion to the number of available count samples. Failure to account for these elements will unavoidably result in unreasonable estimates of error allowance, leading to validation that is either too strict or too lenient. In these regards, research efforts are clearly needed to develop well-defined statistical methodology and procedure on consistency checking to support more reliable model validation.

2. The current accuracy standards are available for only a subset of latest facility type classifications. In the late 1990's, FSUTMS highway network coding scheme was

greatly expanded by adopting the “two-digit” codes. However, new validation standards have yet to be established to complement the new coding scheme.

3. Validation standards for transit assignment model need to be established anew in meeting the emphasis of transit and multi-modal planning. As some of the major Florida urbanized areas are reaching the “built-out” status with fewer right-of-way for highway expansion, the importance of transit demand forecasts has become more and more prominent. Currently, there exist no specific FSUTMS transit model calibration standards for determining the fleet size and headway required for validating transit patronage, loading at the levels of mode, route, corridor, and screenline basis. The development of accuracy standards for the validation of transit models is overdue and should be addressed promptly.

1.2. Overview of the Model Validation Process

To produce reliable results, travel demand models need to be well built before they are applied. The model-building process consists of model estimation, calibration, and validation, which are defined below based on “*The Travel Model Improvement Program: Model Validation and Reasonableness Checking Manual*” (Barton-Aschman Associates, etc., 1997):

- **Model Estimation**: Statistical estimation procedures are used to find the values of the model parameters (coefficients) which maximize the likelihood of fitting observed travel data, such as a household travel survey or on-board transit survey. The focus is on correctly specifying the form of the model and determining the statistical significance of the variables. For example, the initial cross-classification of a trip production model or the logit estimation of level-of-service coefficients in a mode choice model is developed in the estimation phase. If local data are not available, then this initial step is often skipped and the coefficients are borrowed from another urban area.

- **Model Calibration:** After the model parameters are estimated, calibration is used to adjust the parameter values until the predicted travel matches the observed travel demand levels in the region. For example, calibration of the mode-specific constants in a mode choice model ensures that the estimated shares match the observed shares by mode (and often by mode of access).

- **Model Validation:** In order to test the ability of a model in predicting the future behavior, model validation requires comparing the model predictions with information other than that used in estimating the model. This step is typically an iterative process that is linked to model calibration. It involves checking the model results against the observed data and adjusting the model parameters until the model outputs fall within an acceptable range of error. If the only way that a model will replicate the observed data is through the use of unusual parameters and procedures or localized "quick-fixes", then it is unlikely that the model can reliably forecast future conditions.

In the model application stage, although the model may replicate the base-year conditions, the application of the model to the future-year conditions and policy options requires checking of the reasonableness of the projections. Therefore, there is a link between application and validation as well. The sensitivity of the models in response to system or policy changes is often the main issue in model application. Typically the calibration and validation processes focus on the overall outputs of the travel forecasting models, such as mode shares, overall transit ridership, transit boardings for a specific line, or traffic volumes, without detailed checking of results from individual model components. This "all-too-common" approach to model validation might be used under the justification that traffic counts or transit boarding are the only historical data available or because time constraints preclude detailed checking of certain interim model steps.

The approach advocated in the *Model Validation and Reasonableness Checking* Manual is to apply reasonableness checks during the processes of calibrating each individual model component. After each component has been validated, the overall set of models is validated to ensure that each component is properly interfaced and that modeling error is not propagated by chaining the models together. These types of model validations are described as follows:

1. *Individual model validations* are used as part of calibration to show that each component reasonably reproduces observed travel characteristics. For example, trip generation models should be checked to ensure that trip productions and attractions estimated on a district and regional basis are reasonably similar to the observed number of trips; trip distribution models are checked to ensure that they reasonably reproduce the observed average trip lengths by trip purpose; etc.

2. *Validation of the overall set of models* tests the effects of compounding errors. For example, suppose that the trip production model produced too few trips from a zone that was relatively close to a large attractor of trips. If these trip generation results are inputs to the trip distribution model, they would have a tendency to increase trip lengths because of the error in trip production modeling. The overall measures of model performance, such as regional VMT and screenline volumes, should be reviewed with the possibility of error propagation in mind.

1.3. Sources of Error

Even when models reasonably reproduce their portions of regional travel, they are not without error. Error is inherent in all computational models since they are abstractions of real-world travel behavior. Simplifications of reality are unavoidable in order to make the models tractable and practically useful. Sources of error resulting from model building process of travel models include:

- *Measurement Errors* - inherent in the process of measuring data in the base year, such as instrumental inaccuracy, survey questions, network coding and digitizing errors, errors resulting from poor data quality control, etc.

- *Sampling Errors* - such as bias introduced in the process of selecting the set of observations from the population.

- *Specification Errors* - due to improper structure of the model, such as improper functional form and omission of relevant variables.

- *Prediction Errors (Transfer Errors)* - when a model or parameters developed for one context or region is applied in a different one.

- *Aggregation Errors* - arising from the need to forecast for groups of individuals at the household or region level while modeling is done at the level of the individual.

A major concern for validation of travel models is the error inherent in the collection of input data or historical data used for validation. Problems with input data or validation data can lead to erroneous corrections to models that, ultimately, will damage model performance, credibility, and results. For example, if daily traffic counts collected at screenlines are low due to the use of incorrect collection methods, the analyst may attempt to increase the auto occupancy rates or lower the trip rates in order to match the screenline counts. This suggests that a course of action for responding to models that do not pass the validation test is to check for errors first and then consider adjustments to parameters. This study will explicitly consider the combination of different sources of errors in developing the accuracy standards for highway and transit networks.

1.4. Objectives

The main goal of this project is to set up and/or modify the accuracy standards of the validation task for the FSUTMS model. A set of procedures for validating this model that reflects the current household travel survey and transportation data available will be used. This research study seeks to accomplish the following objectives:

1. Develop a refined statistical methodology on consistency checking to support future model validation task,
2. Review and enhance current validation accuracy standards used in FSUTMS using the latest count data collected from different types of facilities in different areas,
3. Devise size requirements of count sample at the desired level of statistical significance for those validation standards to facilitate consistency checking in FSUTMS, and
4. Develop validation standards for the transit assignments.

1.5. Report Organization

The rest of this report is organized as follows. Chapters 2 and 3 provide a detailed review of the existing literature on model calibration and validation for highway and transit networks, respectively. Chapter 4 is devoted to developing accuracy standards and sample size requirements for highway networks. Chapter 5 discusses sources of errors associated with model choice and derives the mathematical form for those errors from which the accuracy standards for transit networks can be derived.

CHAPTER TWO

LITERAURE REVIEW ON ACCURACY STANDARDS FOR HIGHWAY NETWORKS

2.1. Current Practices

Like other travel demand forecasting models, the Florida Standard Urban Transportation Model Structure (FSUTMS) is not immune to under- and over-assignments of trips to individual links or screenlines due to numerous causes. In fact, some of the errors can be so large that the growth in trips for the forecast year is not sufficient to overcome the magnitude of trip under-assignments, resulting in forecast trips that are lower than those of the base year for some links. This often leaves planers to suspect if their models were correct (Li and Lapkowski, 1999). Various means of model calibration and validation were therefore developed to reduce the degree of assignment errors and to obtain more accurate trip forecasts. Acceptable tolerance limits, or accuracy standards, of these assignment errors for different facility types are given in *Urban Transportation Planning Model Update-Phase II, Task C, Develop Standard Distribution and Assignment Models* (FDOT, 1981).

While it is recognized that demand models are inherently susceptible to under- and over-assignments, it is often less recognized that the yardsticks, i.e., accuracy standards, for measuring those "off-the-center" discrepancies might be out of adjustment. A review of current practices used in FSUTMS will facilitate attainment of knowledge on correctness and possible deficiency of accuracy standards such that further amendment or enhancement can be exerted. In addition, literature on the calibration, validation and the corresponding accuracy standards on the planning model will also be extensively reviewed in this Chapter. The information is useful for keeping up to date the state-of-the-art of targeted research topics, identifying major sources of error affecting model predictive capability, and developing methodology and procedure for model validation.

<u>Current Validation Accuracy Standards in FSUTMS</u>

Once components of four-step planning models have been developed and calibrated, validation of the models is needed to forecast trips and flow pattern in a future time horizon based on a comparison of their outputs with the observed travel and count data from a year other than the calibration year, for which *sufficient ground counts and other socioeconomic data are available*. Task C report (FDOT, 1981) is devoted specifically to calibration and validation accuracy standards in Trip Distribution and Traffic Assignment models.

For the Trip Distribution model, the following three types of "checks" were recommended:

- Mean trip length and frequency distribution between the Gravity model output and observed origin-destination data,
- Estimated traffic volumes compared to ground counts at screenlines, cordon lines, and cutlines, and

- Derived statistics such as VMT or VHT per capita and per vehicle compared to other studies conducted in areas of similar characteristics. (reasonableness check)

The validation standards on Trip Distribution identified in the Task C report are listed in Table 2.1. Note that the volume-to-count ratio statistics at screenline, cordon lines and cutlines are also used in the Trip Distribution Table since aggregation of assigned volume at check points along those lines could also be used to check if the directional trip movement from the sub-areas resulted from the distribution model were reasonable. The same set of control statistics are also used in Traffic Assignment Table, which indicates that the volume-to-count ratios at screenline, cordon line and cutlines are in fact important statistics for checking the entire model chain. According to the Task C report, especially for screenline volumes, it gives directional trip movement information from subareas, which indicates:

- Destination choice estimation (distribution)
- Route choice estimation (assignment), and
- Trip making choice estimation (generation)

Any systematic under- or over-estimated volumes at screenlines, cordon lines, or cutlines indicate that the inherent errors come from these models either individually or collectively.

On the other hand, for the Traffic Assignment model, area-wide and facility-specific validation accuracy standards are summarized in Tables 2.2 and 2.3, respectively. Major control statistics include assigned VMT (VHT) to count VMT (VHT), volume-to-count ratio, and root mean square error (RMSE). The ratio of Assigned VMT (VHT) to count VMT (VHT) is evaluated at area-wide level or broken down by facility type (FT), area type (AT), and number of lanes (NL) combination. The volume-to-count ratio standards are the same with the Trip Distribution model, evaluated at screenlines and cutlines. The RMSE measure is defined as follows:

$$RMSE(\%) = \frac{\sqrt{\sum_{i=1}^{N}\left(v_{ci} - v_{ai}\right)^2 \Big/ (N-1)}}{\sum_{i=1}^{N} v_{ci} \Big/ N} \times 100 \qquad (2.1)$$

where v_{ai} and v_{ci} denote assigned volumes and counts, respectively, and N denotes the total umber of count stations located within the whole area or the volume group. As one can observe from the tables, the acceptable error limit increases as the size of travel population (assigned volume, ADT, etc.) decreases, meaning that error is more tolerable with lower travel volumes. In addition, a wide range of variation exists in the error limit as a function of trucking proportion. All the tabulated accuracy standards listed in Table 2.3 assume only 5% trucking activity and no parking on both freeways and arterials. For freeways, the acceptable error limits will be reduced by 8% if trucking activity is increased to 10%, and be increased by 10% if no trucking is assumed. For arterials, the variation is -6% with 10% trucking activity increase and +5% with no trucking activity. The error limits also vary with the size of living population and area type (residential area versus CBD) according to the following schedules:

- For residential area, increase by 11%
- For CBD, decrease by 9%

- For living population changing from 750,000 to 500,000, decrease by 5%
- For living population changing from 750,000 to 1,000,000, increase by 5%

Table 2.1. Accuracy Standards for Trip Distribution Validation Used in FSUTMS

Control Statistic	Accuracy Standard	Source
Mean trip length (by purpose)	3%	OD data
Trip length frequency distribution (by purpose)	Visual inspection	OD data
Intra-zonal trip percentage (by purpose)	5%	OD data
Volume-count ratio (screenline)	10% (for volume ≥ 50,000 vpd) 20% (for volume < 50,000 vpd)	Ground counts
Volume-count ratio (cordonline)	10% (for volume ≥ 50,000 vpd) 20% (for volume < 50,000 vpd)	Ground counts
Volume-count ratio (cutline)	10% (for volume ≥ 50,000 vpd) 20% (for volume < 50,000 vpd)	Ground counts
Person-hour of travel/person	0.75 – 0.85 hours/day	UMTA (1978)
Person-mile/person	10 – 15 mile per capita (12 – 19 in the future)	UMTA (1978)
Average VMT/trip	4 – 6 miles/trip	UMTA (1978)
Average VMT/vehicle	24 – 28 VMT/vehicle-per-day	UMTA (1978)
Average VMT/capita	10 – 15 VMT/capita-per-day	UMTA (1978)

Table 2.2. Accuracy Standards for Traffic Assignment Used in FSUTMS

Control Statistic	Accuracy Standard	Source
Assigned VMT/Count VMT (Area)	5%	Ground counts
Assigned VHT/Count VHT (Area)	5%	Ground counts
Volume-count ratio (screenline, cordonline, cutline)	10% (AADT ≥ 50,000 vpd) 20% (AADT < 50,000 vpd)	Ground counts
Assigned VMT/Count VMT (Facility type, area type, no. of lane combination)	15% (≥ 100,000) 25% (< 100,000)	Ground counts
Assigned VHT/Count VHT (Facility type, area type, no. of lane combination)	15% (≥ 20,000) 25% (< 20,000)	Ground counts
Root Mean Square Error (%) (Area)	35% - 50%	Ground counts
Root Mean Square Error (%) (Link volume group)	25% ≥ (AADT ≥ 50,000 vpd) 30 - 100% (AADT < 50,000 vpd) > 100% (AADT < 3,000 vpd)	Ground counts

Table 2.3. Specific Accuracy Standards in Traffic Assignment as a Function of Facility Type, Facility Size, and ADT

Facility Type	Number of Lanes	ADT Range (Mean)	Accuracy Standard
Freeway	8 6 4	80,000 – 105,000 (92,800) 55,000 – 80,000 (67,600) 30,000 – 55,000 (42,600)	13% 18% 29%
Divided Arterial	8 6 4	37,000 – 47,000 (42,400) 27,000 – 37,000 (31,800) 16,000 – 27,000 (21,200)	13% 17% 25%
Undivided Arterial	4 2	9,000 – 18,000 (13,600) 2,000 – 8,000 (4,850)	34% 56%
One-way Street	4 3 2	18,000 – 24,000 (21,200) 13,000 – 18,000 (15,900) 8,000 – 13,000 (10,600)	13% 17% 25%

2.2. Calibration and Validation on Trip Generation

Based on socioeconomic and demographic data, the trip generation model estimates trip productions and attractions. A trip production is a trip end associated with home for home-based trips and the origin for the non-home-based trips. A trip attraction is a trip end not associated with home for home-based trips and the destination for the non-home-based trips (TMIP, 1997). Trips are usually classified by trips. Current standard trip purposes used in FSUTMS include home-based work (HBW), home-based shopping (HBSHP), home-based social-recreational (HBSR), home-based school (HBSCH), home-based others (HBO), non-home-based work (NHBW), and non-home-based others (NHBO). For calibration and validation purpose, in addition to the total person trips per household, purpose-specific breakdowns of the production rates are also critical and need to be verified.

The primary data sources for calibration typically include the National Household (Personal) Transportation Survey (NHTS, NPTS). Census Transportation Planning Package (CTPP) at either the TAZ (Traffic Analysis Zone) or CT (Census Tract) level could be used as supplementary source for the same purpose (TMIP, 1997). The generation model can be validated using the socioeconomic data from these sources. Several possible validation tests suggested by TMIP (1997) are:

Trip Production

- Number of households by socioeconomic subgroups
- Total person-trip productions per household, per capita, or at regional level
- Number or percentage of person-trip productions by purpose
- Automobile ownership per household

Trip Attraction

- HBW person trip attractions per total employment
- HBSCH trips per school employment
- HBSHP tips per retail employment

Traditionally, two approaches used to construct the trip production model include cross-classification and multivariate regression analysis. Multivariate regression method is inferior due to the fact that all explanatory variables are correlated with dependent variable in the linear fashion and they are often inter-correlated with each other. The cross-classification technique is better in its capability to handle nonlinear relationships. Different approaches might require different data source and procedures for calibration and validation. For example, cross-classification is calibrated using disaggregate household data, while regression analysis is usually conducted at the zonal level that calls for aggregate data (TMIP, 1997). However, it should be noted that the regression analysis could be applied at the household level to explain the trips produced from household as a function of various socioeconomic variables (Lan and Hu, 2000).

Although the model validation manual (TMIP, 1997) gives detailed information on validation tests, no accuracy standards are provided. To develop accuracy standards, various possible sources of errors, including sampling error, bias, model specification error, etc., need to be explicitly accounted for in the adopted modeling approach. The following section reviews previous research related to model calibration and validation on trip generation.

Data collection is a very costly, labor-intensive and time-consuming work, especially in small cities where a high sample rate is required. If one can use the current existing data and models that have been calibrated from other areas of the similar characteristics, significant savings in time and budget can be realized in the demand forecasting process. Therefore, there is a need to test the transferability of different trip generation models. **Caldwell and Demetsky (1980)** conducted a study to determine the suitability and adequacy of a FHWA method for transportation planning in Virginia. In the study, cross-classification, disaggregate regression and aggregate regression trip generation models and trip rates are developed for three cities in Virginia, which are chosen on the basis of certain similarities and on the availability of data.

This study uses two explanatory variables, automobile ownership and household size. Income is not used as one of the independent variables in this study, considering the unavailability of its information in Virginia. The regression model that uses the preceding two independent variables are calibrated with the household data averaged for the traffic zones (aggregated data) from the origin-destination survey, thus they can be compared with the cross-classification trip rates. In addition, these regression models were also calibrated with the aggregated predictions from the individual household observations for each zone. The distribution of predicted trips in the study areas in both cases were found to be significantly different from the reported trips, which means that the models did not perform to the satisfactory degree of accuracy even with the use of disaggregate data and models.

In addition to comparing actual trip productions reported in the O-D survey with trip productions predicted using the calibrated models, the calibrated models are compared with models transferred from the other two study areas. The analysis shows that cross-classification models can be transferred among cities but attention should be paid as to selection of cities of similar characteristics. The cross-classification models were also evaluated with expanded base-year planning data aggregated to the zonal level. The results shows that the disaggregate cross-classification models did not perform well with the aggregate data. Overall, models calibrated with aggregated zonal data appear to perform better than models calibrated on disaggregate household data when used with forecast aggregated zonal data.

Three area-wide trip attraction forecasting methods are used for the analysis of trip attraction, including the standard method recommended by FHWA, linear regression method, and the trip rate method based on specific land uses. The study shows that the procedures for predicting trip attractions can be fairly inaccurate and the use of specific land-use trip rates was a new method to improve these methods because it predicts trip ends very well. However, it also has two disadvantages due to: (1) it requires a directional factor to split trip ends to productions and attractions, and (2) it also requires very specific land use forecasts.

The analysis in the paper also leads one to think of establishing a standard method to improve the synthetic trip generation analysis procedures in the near future so that the transferability issue can be properly addressed. The emphasis should be placed on the development of a prototype model for application in groups of cities. A determination should be made based on how accurately the disaggregate data can be forecast ultimately.

If models developed for one urban area can be transferred to another, the cost of transportation planning studies can be reduced by as much as 80 percent. **Goode and Heimbach (1983)** conducted a research to develop vehicle travel prediction models for small urban areas (populations less than 250,000) to test the transferability of vehicle trip generation models from one urban area to another with acceptable degree of accuracy. The model was based on three separate elements: (1) the comprehensive 1977 traffic counts for the street system of the Fayetteville urban area, which was used to test the accuracy of each model to duplicate existing traffic volumes on the Fayetteville street system; (2) the 1969 Fayetteville origin-destination study, which uses trip rate analysis for productions and regression analysis for attractions; and (3) a composite model developed from the Greensboro-High Point-Gastonia transportation studies, which uses cross-classification model for productions and regression analysis for attractions.

The independent variables used in these models are household size, automobiles per household, income per household, number of households and employment. These models predict vehicle trips per day as a function of these independent variables. The validity of each set of calibrated trip generation models was measured by the ability of the trip generation, trip distribution and traffic assignment process to duplicate traffic volumes on the street system of urban area of Fayetteville, North Carolina. Two kinds of models were used to determine how well they duplicate comprehensive traffic counts, and both of them were assigned to the same street network and use the same traffic assignment procedure.

One kind of model is origin-destination trip generation model based on the comprehensive home interview origin-destination traffic survey in Fayetteville, which is trip rate analysis and regression analysis. The independent variables used in these models are housing condition, race, employment and number of households. The dependent variable is vehicle trip per day. The trip distribution model is the gravity model, which relates trip interchanges between two zones in terms of the total trips produced in the zone of production, the total trips destined to the zone of attraction, and measures of the spatial separation of the two zones. The traffic assignment adopts the FHWA assignment program and uses the all-or-nothing assignment concept.

Another model is the synthesized trip generation model, which uses an external cordon traffic interview survey and internal socioeconomic and travel characteristic developed from an aggregate sample of internal origin-destination surveys at the cities of Greensboro, Gastonia, and High Point. The independent variables used in these models are household size (persons per household), income (income per household), and automobile availability (automobiles owned per household). They are related to the socioeconomic characteristics of the dwelling units. The dependent variable is the average automobile trips generated per household per day for each of the independent variable used in cross classification analysis.

These two trip generation models were evaluated for their ability to duplicate the actual average daily traffic volumes on the networks. The criterion, percent root-mean-square error (RMSE) of estimated trips from actual trips, instead of the commonly used Chi-Square test, is used to test the accuracy of the models. The RMSE provides a better understanding of how much error existed in each respective assignment and it can be expressed as:

$$\text{RMSE} = \sqrt{\sum_{i=1}^{n}\left(f_{0i} - f_{ei}\right)^2 / n} \tag{2.2}$$

where f_0 = the number of observed trips or the actual trips at the i^{th} household,

f_e = the number of trips estimated or expected at the i^{th} household, and

n = the number of links for which comparison are made.

In addition, two techniques for evaluating the accuracy of these models were used: screen line comparison and link volume comparison. Four screen lines were used to compare the assigned volumes for each of the trip generation with the comprehensive traffic. Link volume comparison was used to compare the two traffic assignments by calculating the percent RMSE of each assignment by link section. The low resulting RMSE shows that the synthesized models adequately duplicated traffic volumes in urban areas of similar size. Therefore, it was concluded that the trip generation models were transferable from one urban area to another urban area of similar size.

Trip generation is the first stage in travel demand estimation, and it bears great importance in influencing the final outcome of the transportation planning models results. The development of a person-based afternoon peak-period travel model has received scant attention in the transportation literature. **Kumar and Levinson (1993)** developed an afternoon-period travel model for both work and non-work trip to better replicate travel demand and capture the intermediate stops that characterize many of the trips from work to home. This model is based on the three main data sources: the 1987-1988 Metropolitan Washington Council of Governments (MWCOG) Household Travel Survey, the Montgomery County Planning Department'1987 Census-Update Survey, and the Montgomery County Trip Generation Study conducted from 1986 to 1988. The household travel survey was used to estimate the trip-generation coefficients and rates, the census update survey was used to correct the demographical bias in household travel survey, and the trip generation study was used to validate the model against biased site-specific trip generation rates.

The variables used in this model to determine the trip generation were age, household size and dwelling size. In order to estimate travel demand during a whole day, the authors classified trips into seven one-way movement trips by origin and destination activities rather than by production and attraction. In addition, chained work trips, which form a significant component of afternoon travel, were simplified as one-stop trips to analyze the afternoon travel behavior properly. Normalization procedures were developed to ensure that all ends of a chained trip were properly accounted for. The basic equation for normalization is given as:

$$p_i = p_i \frac{\sum_{j=1}^{J} q_j}{\sum_{i=1}^{I} p_i} \tag{2.3}$$

For chained trip purposes, normalization requires the following two equations:

$$p_i' = p_i \frac{\sum_{k=1}^{K} r_k'}{\sum_{i=1}^{I} p_i}, \; r_k' = r_k \frac{\sum_{j=1}^{J} q_j}{\sum_{k=1}^{K} r_k} \tag{2.4}$$

where i, j, k = origin, destination and intermediate zones, respectively;

p_i, q_j, r_k = trip generated in origin, destination and intermediate zones, respectively; and

p_i, r_k = adjusted trips generated in origin and intermediate zones, respectively

The authors also defined three main trip ends to estimate the trip-generation factors. They are conventionally defined "work" and "home", and "other", which includes all trip ends other than home or work. For the home-end trips, a separate person-based trip generation model of afternoon peak-period used age, household size and dwelling size to determine the trip generation. For the non-home-end trip generation, ordinary least squares (OLS) method was used to relate trips to employment by type and population characteristics. The variables used to estimate trip rates for the work-end were employment in offices, retail and other. A regression analysis was conducted for each trip purpose. One example of the regression models has the following form:

$$T_i = B_1 \times OFFEMP_i + B_2 \times RETEMP_i + B_3 \times OTHEMP_i \tag{2.5}$$

where T_i = person trips attracted per worker in i^{th} zone,

$OFFEMP_i$ = office employment in i^{th} zone,

$RETEMP_i$ = retail employment in the i^{th} zone,

$OTHEMP_i$ = other employment in the i^{th} zone, and

B_1, B_2, B_3 = model coefficients

Validation of these models was accomplished by comparing the trip-generation coefficients estimated by the 1987-88 Household Travel Survey and those obtained from the Montgomery County Trip Generation Study. This comparison shows that the relatively short and less regular shopping trips were underreported in nature. The person trip generation rates for the non-home end of non-work trip were used from the trip generation study to correct the model.

2.3. Calibration and Validation on Trip Distribution

To maintain consistency between the modeled and observed trip pattern, the OD flow tables need to be calibrated in the trip distribution step if the observed OD flow data are available. However, the OD flow data are difficult to observe and collect. The traditional household travel survey could be used to provide some data to establish a broader picture regarding travel patterns, but its applicability is usually restricted to modeling due to its limited sample size. As an alternative, the trip length distribution data are often used for calibrating the gravity model.

In the calibration process, the deterrence (friction) function parameters are adjusted to reproduce the observed base-year data, such as trip length distribution, OD flows, or traffic counts. Depending on the type of data available, the objective and the methodology of the calibration process are distinct but interconnected. First, if the OD flows were directly available, the friction model parameters could be estimated based on the maximum-likelihood approach by assuming a multinomial distribution for the sampled OD flow frequencies (Hyman, 1969; Evans, 1971), or they can also be estimated using the least-squares (LS) methods (Kirby, 1974). Second, if direct observations of OD flows were not available, a common approach is to estimate OD flows from traffic counts. Even if no data on trip-end or trip length frequency distribution are available, calibration of gravity model can be performed to match with the observed traffic counts. A review of estimation algorithms for this purpose can be found in van der Zijpp *et al.* (1998).

If the trip length frequency distribution is available, the iterative procedure suggested by the Federal Highway Administration (1977) for gravity model calibration could possibly be used to reproduce the observed trip length distribution. As described in NCHRP Report 365 (Martin and McGuckin, 1998), the calibration process iteratively adjusts friction factors using the ratio of the observed and calculated frequencies at each trip length increment, and then uses the adjusted friction factors to fine-tune the functional relationship of the impedance function through the LS method. This iterative process is performed until the calculated trip length distribution sufficiently estimates the observed counterpart. Martin and McGuckin (1998) provided typical friction model parameters by trip purposes that were more applicable and transferable to smaller urban areas. In cases where the distribution of OD trips were not consistent at the interchange level with the observed pattern, further adjustment to the socioeconomic factors with resulting friction factors could produce closer estimations between the OD patterns and the trip length distributions (Papacostas and Prevedouros, 2001).

Several validation tests are suggested by TMIP (1997), including:

- Average trip length (duration) by purpose
- Percent of intrazonal trips by purpose
- Region-wide Trip interchanges and trip movements against observed screenline or cutline counts
- Stratification of trip length and/or trip interchanges by income class

In what follows, previous research on model calibration and validation for trip distribution or its integration with other processes was reviewed. **Morrall (1971)** proposed a unified trip distribution model to enhance trip distribution and post-distribution modal split models. The model includes (1) a function allowing work trip ends to be split into captive and non-captive trip makers based on employment and dwelling unit opportunities; (2) direct estimation of the future travel impedance factor function; and (3) estimation of the future modal split travel impedance factor function based on information on land use and transport characteristics. The model was calibrated with data from the 1964 Regional Transportation Study conducted in Metropolitan Toronto. Results from model validation produced good agreement between the estimated and observed travel demand patterns.

Levinson and Kumar (1994) applied a multimodal trip distribution model at the strategic planning level to alternative high-occupancy vehicle alignments, for choosing alignments for further study and right-of-way preservation. The proposed model was estimated and validated in the metropolitan Washington, D.C., region. Within the modeling framework, a methodology for measuring accessibility, which was used as a measure of effectiveness for networks, based on the impedance curves in the distribution model was also presented.

With the increasing concern for environment in recent years and the response of managing travel demand, there is a need to evaluate the feasibility of introducing high-occupancy vehicle (HOV) and transit facilities. It is important to explicitly account for different distribution characteristics of mode other than the single-occupancy vehicle (SOV). Levinson and Kumar (1994) tried to fill this gap by estimating a multimodal trip distribution function for metropolitan Washington, D.C. region.

Several different formulations of trip distributions over the years were reviewed. They include the Fratar (growth) model, the gravity model, the intervening opportunities model, the combined destination choice and mode choice model, and the logit-based mode choice model. The authors pointed out that one of the key drawbacks to the application of many early models was the inability to take into account congested travel time on the road network in determining the probability of making a trip between two locations. The authors used the travel time feedback method in their model.

The data source for the estimation of the trip distribution model consists of detailed person travel surveys conducted by the metropolitan Washington Council of Governments for 1968 and 1987-1988. These two surveys defined three primary travel modes, including transit, automobile and walking. Seven trip purposes were defined in the context of origin and destination, including home to work, work to home, home to other, other to home, other to work, work to other and other to other. For estimation, trip purposes were grouped into three categories, work, non-work, and chained work. In addition, impedance functions were applied in the estimation for application in the gravity model, with the dependent variable being the number of trips per unit area in each 5-minute time band. Travel time and mathematical transforms of travel time serve as independent variables. The impedance function is expressed as:

$$f(C_{ijm}) = e^{(a \cdot t + b \cdot t^{0.5} + c \cdot t^2 + d)}) \tag{2.6}$$

where $f(C_{ijm})$ is the impedance function for travel time t, and *a*, *b*, *c* and *d* are the calibration coefficients. The multimodal impedance function (f_{ij}) is expressed as follows:

$$f_{ij} = \sum_{m=1}^{M} P_{jim} \cdot f(C_{ijm}) \tag{2.7}$$

subject to

$$\sum_{m=1}^{M} P_{ijm} = 1 \tag{2.8}$$

where P_{ijm} = probability of using mode *m* on a trip from *i* to *j* (from mode choice model),

C_{ijm} = travel time from *i* to *j* using mode *m*, and

$f(C_{ijm})$ = friction (impedance) function (negative exponential).

In model estimation the average density of opportunities available in each 5-minute time band is assumed to be uniform. The number of opportunities is estimated by assuming 5-min radius time contours, i.e., the first circle (0-5 minutes) has an area of 25π minute square, the second circle (5-10 minutes) has an area of 100π -25π = 75π minute squared, and so on. In the true travel time contours the number of opportunities could be estimated by a geographical information system.

A doubly constrained gravity model was used to derive Equation (2.7). The impedance matrix for work trips was balanced against each of the origin and destination vectors to obtain the trip table for trip purposes. These all-mode trip tables were multiplied by the mode choice probabilities to obtain vehicle trips by class (SOV, HOV) and transit person trip tables (walk access, automobile access), which were then assigned. In the feedback procedure, vehicle trips were assigned for a single iteration, producing a new origin-destination travel (skim) times. The new impedances were used to update the modal probabilities and then the impedance matrices. This process continued, with the new demand assigned to the congested network, until it reached convergence.

The travel time (C_{ij}), multimodal impedance functions (f_{ij}), and demand to be assigned (T_{ij}) are updated after each iteration of route assignment to ensure consistency between the input and output travel times. Since the authors applied the travel time feedback method along with balancing procedure, the distributions of model output, observed 1988 and observed 1968 fit very well with each other. This can be visualized clearly from the work trip time distribution curves, and can also be proven by the **Friedman nonparametric method**. The chi-square test statistics is 6.3 with a significance of 0.042.

There are three sources of error between the applied model and the surveys. The first is the variation existed in flows from area to area, which are influenced by some socioeconomic factors. These factors will affect the mode choice but are not directly considered in the distribution model. The second is the inaccuracy in the estimates of impedance matrices for the various modes. The third is the travel cost, which is highly correlated with time. They are not explicitly accounted for in the distribution model. The authors adjusted the model to match the observed data.

The multimodal trip distribution impedance functions were tested in a transportation planning model with feedback between different components to produce consistent results. This method has the advantage that it accounts for the effect of changes in transportation supply on demand better than a conventional gravity model that uses only automobile impedance. Because transportation planning process places heavier emphases on multiple modes, models need to account for all of the modal choices.

A method for evaluating network using multiple modes was developed in this paper to support transportation planning and decision making. The methodology of measuring accessibility to test the relative impacts of different networks is in contrast to evaluating traffic volumes or total travel times on each alternative. It was applied at the strategic planning level to alternative high-occupancy vehicle alignments to select alignments for further study and right-of-way preservation. The benefits were defined as the accessibility between homes and jobs provided by the network given a fixed land-use pattern. The accessibility is expressed in the following form:

$$A_{jm} = \sum_{i=1}^{I} [f(C_{ijm}) \cdot EMP_i] \tag{2.9}$$

where $A_{jm} =$ accessibility index for residential zone *j* by mode *m*,
$f(C_{ijm}) =$ friction factor between zones i and *j* by mode *m*, and
$EMP_i =$ employment in zone i.

The accessibility measures were further used to evaluate the entire network. This method does not guarantee the optimal solution, but it lays the groundwork for quantifying the impacts of each alignment on a consistent basis, particularly in an attempt to rank the benefit-cost ratio of the alignments. The results of the application of this method prove its benefit in analyzing real-world problems.

2.4. Calibration and Validation on Modal Split

Modal split model deals with the traveler decision on what mode (including non-motorized model) to take as a function of the characteristics of trip-makers, trip purpose, and the service characteristics of available modes. Two types of modal choice model are popularly used, including multinomial logit and nested logit models. Interested readers should refer to Ben-Akiva and lerman (1985) for a detailed review of the differences between these two modeling approaches. A number of variables are required as the model inputs, including transit time, transit fare, number of transfers, highway travel time, auto costs, auto-ownership, parking fee, household characteristics, and land use pattern. During calibration, the model parameters associated with these explanatory variables need to be estimated using the observed data. Reasonableness checks need to be performed on these parameters as part of calibration work. Model validation on modal splits can be performed in either disaggregate or aggregate mode. TMIP (1997) suggested the following disaggregate validation tests:

- Household characteristics
- Trip-maker characteristics
- Zonal characteristics
- Trip characteristics

For aggregate validation tests, the following checks are suggested:

- Transit ridership, highway vehicle, and auto occupancy counts at screenlines by time of day
- Total patronage by transit mode
- Average auto occupancy by trip purpose
- Percent of single occupancy vehicle (SOV) by trip purpose
- HBW transit trips
- Mode share by area types, such as CBD

In the following, previous researches on model calibration and validation for modal splits are reviewed. **Daly and Zachary (1975)** proposed the formulation and calibration procedure for several modal-split models. The authors analyze the predictability of these models in the proportions of the travelers for given journeys on a sound, logical, and statistical basis, which shows clearly the corresponding effect of the necessary assumptions. The authors also discuss the premises needed to set up the modal choice models and present the procedure on parameters estimation and validation techniques of the model. At last, this paper discusses the application of these models to urban transportation planning and the potentiality to extend the modeling and calibration techniques to other aspects of the planning process.

Ganek and Saulino (1976) discuss a disaggregate mode choice model involving three travel modes: drive alone, car pool and transit. The coefficients of variables in the model are analyzed whether they have influence on the mode choice. The socioeconomic variables that are statistically significant are analyzed, which are work location, cars per driver, and sex. The model and the coefficients are applied to estimate the various subpopulations of commuters. The author presents the difference between the determinants of mode choice for CBD worker and those of non-CBD workers as well as the difference in the cost and time coefficients among travel corridors and income classes. And finally this model is validated by successfully predicting the mode choices for three kinds of commuters.

Train (1978) discusses the validation test of a work trip mode choice model built for worker before Bay Area Rapid Transit (BART) opened for service. Two methods are presented to validate the model by comparing the actual and predicted value of mode shares and the parameters of the model before and after BART opened. The reasons that may account for the differences in the two comparisons are also analyzed.

Willis and Lee (1980) studied the stability over time in mode-choice attitudes based on two surveys in the Orlando, Florida, urbanized area. A mode-share model is developed and calibrated using the data collected from the first survey conducted in the area in 1973, according to mode-share relationships found in Minneapolis area. The questionnaire used in this survey is designed to provide input into a mode-share model so that future patronage of alternative transit systems could be determined. The second survey accomplished in 1978 is designed to duplicate the first survey to the maximum extent possible, and its purpose is to determine whether opinions toward transit had changed since the original survey in 1973.

The difference in response to the income question in the two surveys appeared significant. In these two surveys, a sampling procedure is developed to ensure that socioeconomic groups would be proportionately represented. The authors determine the proportion of household in each of the income groups by using statistical error-limit analysis. The formula is expressed as follows:

$$N = \pi(1-\pi)(Z_\alpha / E_{\max}) \tag{2.10}$$

where N = sample size,
π = percentage of group that will respond in a certain way to a question,
Z_α = confidence point in the normal distribution, and
$E_{\max}$ = maximum acceptable error level

The comparison of the two surveys showed that the attitudes toward transit use in the Orlando area remains stable over time, at least for the purpose of short-range planning (three to five years). This is particularly significant when the time frame of the two surveys in Orlando is considered. The results of the two surveys also imply that there is stability over time in mode-choice attitudes, even over a period when significant changes occurred in socioeconomic factors generally related to mode choice. These results would therefore tend to support similar stability in other areas over longer period of time. This could be particularly important to other areas in which updating of the existing mode-choice survey is being considered.

At last, the authors also pointed out the following concerns. First, the results are limited to specific events in Orlando area between 1973 and 1978, therefore additional research should be done to further verify the stability because of the extremely limited database related to stability of trip-making characteristics over time. Second, this research should, where possible, consider longer period of time and, ideally, should come from areas in which existing or historic use of modes other than the automobile is significant.

Travel-demand forecasting constitutes the most critical element of the urban transportation planning process. The traditional approach to transit demand analysis has been criticized as being oriented toward larger cities and being insensitive to the needs and attributes of small and medium-sized urban areas. **Cynecki, Khasnabis, and Flak (1982)** developed and applied a one-step modal-split process designed specifically for small and medium-sized urban areas. Further, the model is oriented toward the database commonly available for small and medium-sized urban areas and is responsive to the needs of smaller urban areas.

The model is developed by using the logit approach which incorporates modal-choice decisions through the use of explanatory variables in a set of mathematical formulations. Probabilistic equations are developed to reflect characteristics based on the relative attractiveness of the candidate modes, as expressed below:

$$P\left(\frac{i-j}{m}\right) = \frac{\exp\left[-U\left(\frac{i-j}{m}\right)\right]}{\sum_{m=1}^{N}\exp\left[-U\left(\frac{i-j}{m}\right)\right]} \tag{2.11}$$

where $P\left(\frac{i-j}{m}\right)$ is the proportion of the total person trips from zone *i* to zone *j* by using mode *m*, and N is the total number of travel modes. $U\left(\frac{i-j}{m}\right)$ is the utility or disutility value of a trip from *i* to *j* by using mode *m* as described below:

$$U\left(\frac{i-j}{m}\right) = F_c\left(\frac{i-j}{m}\right) + F_t\left(\frac{i-j}{m}\right) + F_s\left(\frac{i-j}{m}\right) \tag{2.12}$$

where $F_c\left(\frac{i-j}{m}\right) =$ function of the out-of-pocket cost in making the trip from *i* to *j* by mode *m*,

$F_t\left(\frac{i-j}{m}\right) =$ function of the travel time in making the trip from *i* to *j* by mode *m*, and

$F_s\left(\frac{i-j}{m}\right) =$ function of the socioeconomic characteristic of the tripmaker or land use characteristics associated with trips from *i* to *j* by mode *m*.

In addition, the following constraint must hold true:

$$\sum_{m=1}^{N} P\left(\frac{i-j}{m}\right) = 1.00 \tag{2.13}$$

Each of the three utility or disutility functions (F_c, F_t, and F_s) can be developed as a linear or nonlinear combination of independent variables. A linear combination of the following form as used in this study:

$$F\left(\frac{i-j}{m}\right) = (\alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_n X_n) \tag{2.14}$$

where $F\left(\frac{i-j}{m}\right) =$ impedance function (time, cost, distance, etc.) for trips from *i* to *j* by using mode *m*,

$X_i =$ individual elements within the impedance function (e.g., in-vehicle time, waiting time, out-of-pocket cost, parking cost, etc.), and

$\alpha_i =$ coefficient to be estimated as part of the model calibration.

The model developed is multimodal in nature and does not require lengthy step-by-step process of branching and submodal split. The authors built the model in four main steps and applied it in the area of Genesee County, Michigan. The model is calibrated by using a sample of the total trips. For the sake of brevity, only one work-trip model and one non-work-trip model is discussed in detail in this paper. Each model studied the following five modes: (a) automobile, drive alone; (b) automobile, one passenger; (c) automobile, two passengers; (d) automobile, three or more passengers, and (e) transit, bus service. The results indicated that the logit model is a valid approach to travel-demand modeling for multimodal analysis by using aggregate approach and that the potential of applying this approach in other urban areas is quite high, although further calibration and validation efforts are needed before a more widespread application is practiced.

The paper also indicates that, unlike the traditional modal-split, i.e., diversion-curve-type, models, the resource requirement for these models are nominal and thus can be used for transportation planning purpose in small and medium-sized urban areas. The model is also sensitive to changes in transportation system attributes as well as in trip-maker characteristics and can be applied to test air quality, energy, and other impacts of transportation strategies implemented in smaller urban areas.

Alpern and Levinsohn (1984) discussed the building of modal split model based on the 1980 Census Urban Transportation Planning Package (UTPP). In spite of providing journey-to-work travel data only, the 1980 Census Urban Transportation Planning Package (UTPP) is an available source of travel information for all urban areas in the U.S. Without the need for new data collection, the authors analyze the advantages of the UTPP data in development and calibration of work trip modal split models. It is also pointed out that due to the possible differences between the split of worker by usual mode used and the split of home-based work trip by mode used on an average day, the models should be carefully validated against independent data sources when calibrated with UTPP.

Testing a disaggregate choice model by comparing predicted and observed market shares has much intuitive appeal, but it is not the only way in which these models can be tested. In practice, the decision whether differences between predictions and observations are excessive is made judgmentally, thereby raising the possibility that a correct or approximately correct model will be rejected because of the effects of random sampling errors, which are not relevant to the questions of whether the model under consideration is correct. A large number of test procedures based on likelihood ratio tests, Lagrangian multiplier tests and the likelihood ratio index goodness-of-fit are also available. It is worthwhile to consider how one might choose among these tests in practical model development. **Horowitz (1984)** tried to find a C test in this paper to test the statistical significance of differences between predictions and observations for disaggregate choice models.

The author derived the expression for this C test and gives the five easily programmable steps for implementing the test, which can easily be programmed for implementation by computer. And in addition, the comparison between the C test and other tests is analyzed. An important difference between the C test and the other tests is that the other tests require the analyst to specify an alternative model against which the model under consideration is to be tested. In fact, these tests attempt to determine whether the alternative model fits the available data better than the model under consideration does, in which case the model under consideration is rejected as being incorrect. In contrast, the C test does not require specification of an alternative model. In fact, it tests the model under consideration against all alternatives simultaneously.

In the mean time, it should also be realized that the ability of a test against a specific alternative to identify an erroneous model and the relative power of the C test depend on the choice of the alternative. This lead to the conclusion that the C test and other tests against specific alternatives are complements, rather than substitutes, and that both types of tests should be carried out during the process of developing empirical models. A practical approach to this begins with formulating several alternatives to the model under consideration and the current model should be tested against the alternatives using likelihood ratio tests or other appropriate test procedures. If the current model is not rejected in test against specific alternatives, a C test should be performed. Therefore, the C test amounts to a test of the current model against all remaining alternatives and may have a higher probability of detecting errors in the current model than do the tests against specific alternatives if the specific alternatives are themselves highly erroneous.

In general freight, there are many factors that may influence mode choice decision. **Wilson, Bisson and Kobia (1986)** studied the shippers of general freight commodities in the Atlantic provinces of Canada. The authors discuss about the data collection method and explain the survey data in details. A linear logit model is presented to analyze the factors that have effect on the selection of various modes for good shipments, and the relationship between the utility of each mode the explanatory variables is analyzed based on this model. The study proves that logit analysis is superior to the conventional use of waybill data in validity and power, because it would improve the quality of the result and gain insight into the shipper mode choice decision process. In addition, using the suggested model forms and data in further research would improve the results and provide a greater understanding of the mode choice decision process.

Tretvik and Widlert (1998) investigated the transferability between two cities in neighboring Scandinavian countries of a complicated nested structure for travel-to-work decisions by household. They analyzed the instability problems of model specification and parameter estimation in disaggregate choice models. The comparative analysis shows that the transfer-ability of the model is feasible with only a minimum of updating based on mode choice, car, allocation, and trip frequency. In addition, the model in which all parameters are re-estimated using local data behaved in an almost complete accordance with the best local model in all validation tests. The model with recalibrated constant terms only showed just minor discrepancies with the best local model in prediction tests.

2.5. Calibration and Validation on Highway Network

2.5.1. Current Practices

In the four-step process, calibrations of traffic assignment process concern about the following aspects, including link delay-volume function, model parameters governing the assignment method, model parameters governing flow propagation (such as free-flow speed and capacity), and time-of-day/directional split factors if time-of-day model is performed. In recent years, various delay-volume functions have been proposed, including the modified BPR (Bureau of Public Roads) function calibrated for different facility types, Akçelik’s formulation (Akçelik, 1991), and Conical function (Spiess, 1990). Parameters associated with different models are subjected to calibration if flow-speed data were available. Free-flow speed and capacity are also two primary parameters that govern link delay and flow propagation, and need to be calibrated with caution since they might cause systematic discrepancy between estimated volume and ground counts. The time-of-day/directional split factors are typically used to convert production-attraction trips to time-of-day OD trips (TMIP, 1997). Several validation tests are suggested by TMIP manual for time-of-day factors, including:

- Percent of trips by time-of-day by purpose
- Percent of trips by time-of-day by mode
- Percent of trips by time-of-day by direction

Volume measurements and estimates are important in estimating travel trends; developing travel models, and establishing improvement plans. Reliable measurements of vehicle miles of travel (VMT) in urban areas provide an important benchmark for urban transportation studies. Similar to current practices used in FSUTMS, the validation tests for highway assignment are presented at system-wide, corridor, and link-specific levels. Specific tests suggested are:

- VMT per region, per household, per person
- VMT by facility type and urban area population
- Volume by screenlines
- Volume for all links with counts
- Link speeds by facility type and area type

Typical distributions of VMT by facility type in urban areas were given in Fleet and DeCorla-Souza (1990), as listed in Table 2.4. Also reasonable ranges of VMT per household are 30 – 40 miles per day for small urban areas and 40 – 60 miles per day for large urban areas. At the person level, reasonable ranges of VMT are 10 – 16 miles per day for small urban areas and 17 – 24 miles per day for large urban areas (Ismart, 1990).

In term of assigned volume, Ismart (1990) reported that region-wide error limit should be less than 5%. By facility type (functional classes of roadway), the suggested error limits are:

- Freeways: 7%
- Principle Arterials: 10%
- Minor Arterials: 15%
- Collectors and Frontage Roads: 25%

As also quoted in TMIP manual (1997), Michigan Department of Transportation (DOT) has set 5% and 10% allowance for volume discrepancy on screenlines and cutlines, respectively, which are tighter than the current practices used in FSUTMS (10 – 20% depending on ADT). For individual links, Ismart (1990) and Michigan DOT (1993) have proposed allowable error limits according to level of AADT. Comparison is listed in Table 2.5. A continuous error limit curve as a function of AADT is also furnished in Ismart (1990), as shown in Figure 2.1. The coefficient of Determination, R^2, and RMSE are also used as criteria for gauging volume discrepancy. As suggested by Ismart (1990), the area-wide R^2 should be at least 0.88. Montana DOT suggests that area-wide aggregate RMSE should be less than 30%, again slightly tighter than the current practice in FSUTMS (35 – 50%). In Reno area, it is suggested to set the allowance by combination of facility and area types, as opposed to by area and by volume group practices in FSUTMS. However, in principle, both sets of standards share similarity as to larger allowance being assigned to small volume groups. Their accuracy standards are listed in Table 2.6.

Table 2.4 Typical Distributions of VMT in Urban Area by Facility Type

Facility Type	Urban Area Population		
	Small (5 – 200k)	Medium (200k – 1M)	Large (> 1M)
Freeways/Expressways	18 – 23%	33 – 38%	40%
Principal Arterials	37 – 43%	27 – 33%	27%
Minor Arterials	25 – 28%	18 – 22%	18 – 22%
Collectors	12 – 15%	8 – 12%	8 – 12%

Table 2.5 Accuracy Standards on Volume by AADT (TMIP, 1997)

AADT	Allowable Error Limits	
	FHWA	Michigan DOT
< 1,000	60	200
1,000 – 2,500	47	100
2,500 – 5,000	36	50
5,000 – 10,000	29	25
10,000 – 25,000	25	20
25,000 – 50,000	22	15
> 50,000	21	10

Table 2.6 Percent RMSE by Facility and Area Types in Reno (TMIP, 1997)

Facility Type	Area Type					
	1	2	3	4	5	All
Freeway	11.6	18.1	21.9	-	11.3	18.3
Major Arterial	22.5	37.8	42.2	43.1	43.3	36.8
Minor Arterial	25.9	44.1	52.4	28.4	60.1	43.9
Collector	-	53	88.9	115.3	70.1	77.5
Ramp	24.2	63.5	47.6	80.6	131	74.8
All	21.3	37.2	37.8	43.7	38.7	36.8

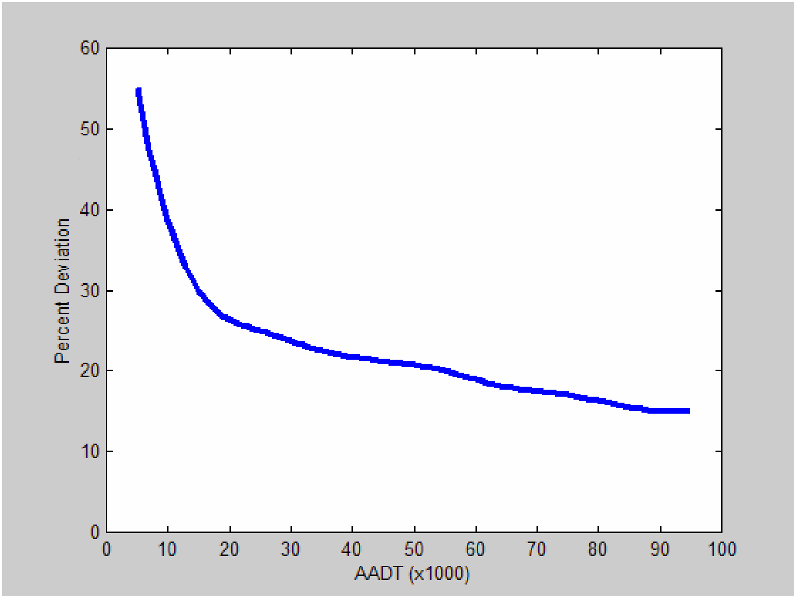

Figure 2.1 Error Limits as a Function of AADT for Individual Links

CALTRAN (JHK & Associates and Dowling Associates, 1992) proposed maximum desirable deviation (error) for link and screenline volumes. The deviation measure is defined as:

$$\frac{\text{Base - Year Assignment Volume} - \text{Base - Year Count}}{\text{Base - Year Count}}$$

Figure 2.2 depicts the maximum desirable deviation as a function of daily assigned volumes for individual links and screenlines. One can find that the link-specific maximum deviation proposed by CALTRAN essentially identical to what is proposed by Ismart (1990). The suggested allowable deviation for different facilities is then defined as certain percentage of link coverage should meet the maximum desirable deviation. For example, 75% of freeway and principle arterials should meet the standards. Screenlines are generally required to meet the standards at full (100%) coverage.

Based on the similar principle, the County Costa Transit Authority (CCTA) in the San Francisco Bay area has developed a set of accuracy standards on link volumes and turning volumes at intersections during rush hours. They are:

- 75% of all freeway links must be within 20% of traffic counts
- 50% of all freeway links must be within 10% of traffic counts
- 75% of all major arterial links must be within 30% of traffic counts
- 50% of all major arterial links must be within 15% of traffic counts
- 50% of all intersection major turning movements must be within 20% of traffic counts
- 30% of all intersection secondary turning movements must be within 20% of traffic counts

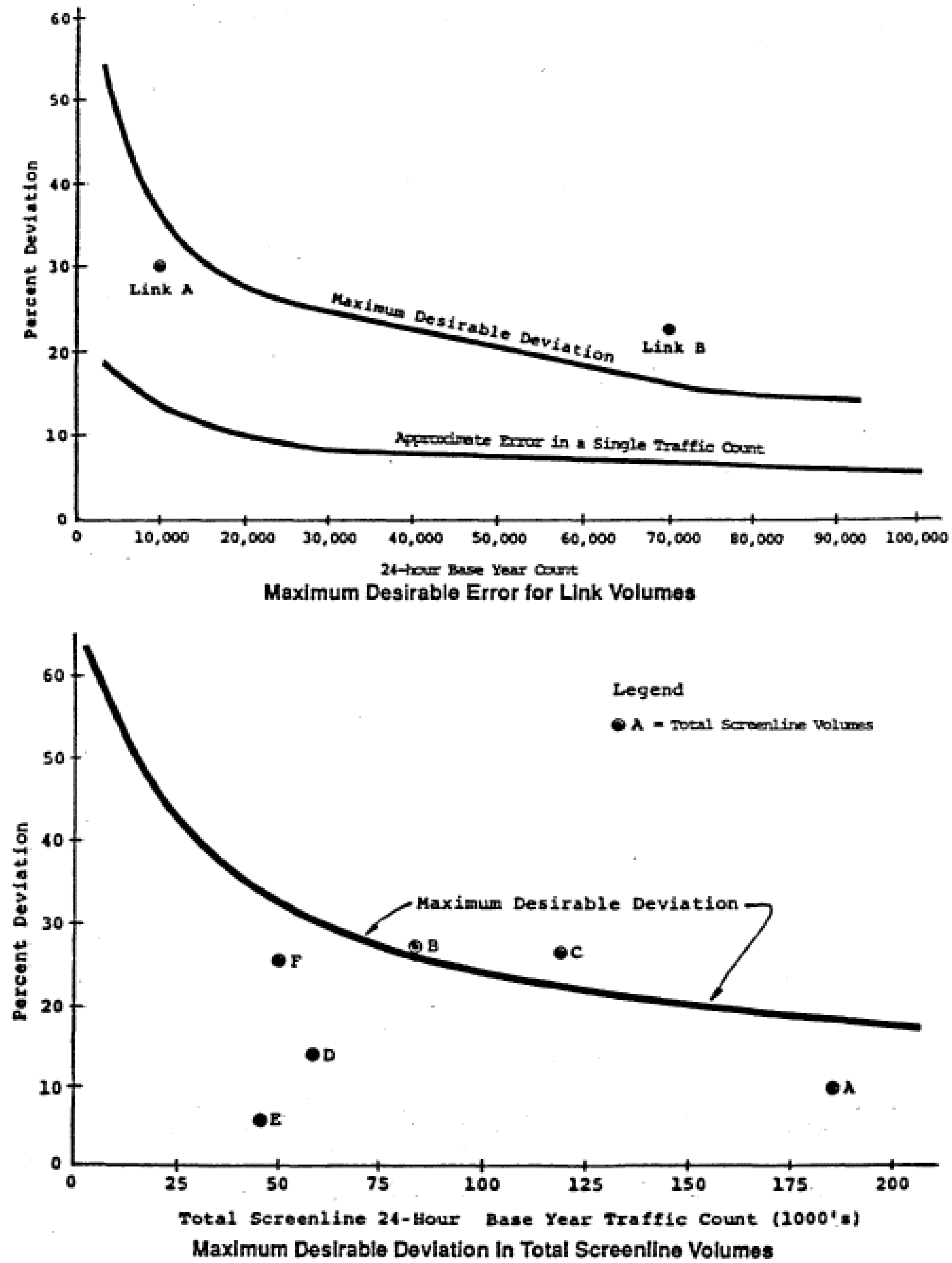

Figure 2.2 Maximum Desirable Deviations Proposed by CALTRAN

As indicated in Figures 2.1 and 2.2, the acceptable error increases substantially with base-year count below 10,000 vpd. However, in some small jurisdictions in California (Chung and Henry, 1999) where over 90% of the calibrated links have daily volume less than 10,000, planners believe it is possible to achieve more stringent standards for low-volume links than CALTRAN guidelines. For long-range planning purpose, they consider the CALTRAN guidelines are adequate but mandate 80% of calibrated links should fall within maximum deviation standards. For more detailed planning or site impact projects, more restrict standards and criteria (see Table 2.7) are adopted.

Table 2.7 Validation Standards and Criteria for Small Jurisdictions in California

Base-Year ADT	CALTRAN	Proposed for Long-Range Planning	Proposed for Specific Planning
< 1,000	65	65	50
1,000	65	65	50
3,000	53	53	40
5,000	46	46	33
7,500	40	40	28
10,000	36	36	25
Validation Criteria	75% within max. deviation	80% within max. deviation	90% within max. deviation

Overall, it can be concluded that the accuracy standards on validation of highway assignment vary across different studies by the Federal, state, and county levels. This is not surprising since the allowable error limits depends not only on the abovementioned factors, but also on number of count data available, level of significance desirable, and the statistical variability inherent in the sampled counts in combination with the prediction variability associated with the planning models. A unified statistical methodology is required to quantify these error allowances systematically.

Tracing back to these literatures, it can be found that most of these abovementioned accuracy standards at the nation, state, and county levels in fact adapt from the **Guide to Urban Traffic Volume Counting** (**GUTVC**) developed by the Federal Highway Administration (FHWA) (Levinson and Roark, 1975). This guide provides a uniform method of calculating estimation of VMT. It shows how sampling procedures can be used to estimate VMT within an urban area. The method is based on random selection, and sampling of highway links that are stratified by volume, facility type, and area type. Based on statistical sampling theory, a desired confidence interval can be assumed and a sample size calculated. The selected links are then matched with a randomly selected day for sampling.

VMT Calculation

GUTVC presents formulas and methods used to calculate the VMT, which can be described below. The mean vehicle-miles per mile of roadway should be completed for each volume group. This is obtained from the formula:

$$\overline{V_h} = \frac{\sum_i^m V_{hj} \cdot l_{hj}}{\sum l_{hj}} = \frac{\sum_i^m V_{hj} \cdot l_{hj}}{n_h} \tag{2.15}$$

Where

V_{hj} = the volume on link j of stratum h;

l_{hj} = the length of link j of stratum h;

$\overline{V}$ = the mean volume of stratum h;

m = the number of links measured = (number of miles of roadway measured in stratum h)÷(the mean length of link in stratum h) (round upward to the nearest integer); and

Σl_{hj} = number of miles of roadway measured in stratum h.

The total estimated vehicle-miles of travel for the h^{th} volume group is obtained by multiplying the mean volume obtained from the sample by the miles of roadway in that stratum.

$$V_h = N_h \cdot \overline{V_h} \tag{2.16}$$

where N_h = total miles of road in stratum h;

$\overline{V}$ = the mean volume of stratum h; and

V_h = total VMT in stratum h.

The total estimated vehicle-miles of travel in all strata are the sum of VMT in the individual strata. In addition, the overall average VMT per mile is obtained, $\overline{X}$ as follows:

$$\overline{X} = \frac{\sum^{R} V_h}{\sum^{R} N_h} = \frac{\sum^{R} N_h \overline{V_h}}{\sum^{R} N_h} \tag{2.17}$$

where R is the number of strata.

To obtain a measurement of statistical reliability of the VMT estimates and provide a basis for subsequent modifications to sample-size calculations and data-collection processes, the variance in estimated VMT is calculated by using the formula below:

$$S_x^2 = \left[\sum_{h=1}^{R} \frac{N_h}{n_h^2 n_{h-1}}\right]\left[n_h \sum_{j=1}^{n} X_{hj}^2 - (\sum_{j=1}^{n} X_{hj})^2\right] \tag{2.18}$$

where S_x = variance of estimated VMT;

N_h = total number of links in stratum h;

n_h = number of links sampled in stratum h;

X_{hj} = VMT for link j of stratum h; and

R = number of strata.

The coefficient of variation, C, in the mean is computed as follows:

$$C = \frac{S_x}{\overline{X}} \tag{2.19}$$

In addition, the desired errors in the sample size and VMT calculation and the desired levels of confidence should be specified. As sample size is increased, there is an increase in the accuracy of the estimate, as well as the costs of obtaining the basic information. These two factors, cost and accuracy, weigh heavily in assessing various program to achieve desired results, and in establishing accuracy requirements of alternative methodologies. GUTVC suggests ranges in confidence levels and relative errors, as is shown in the Table below. As one can see, the 68 percent confidence level is suggested for traffic counting purposes, except for the screen-line and cordon volumes. The ranges are intended as guidelines. They should be modified to detect the rates of travel growth in each urban area. Specific urban areas or states may desire to modify accuracy requirements based on experience, field testing and rates of travel growth.

Table 2.8 Suggested Traffic Counting Accuracy Standards

Type of Urban Volume Count	**Relative Error** (percent)	**Confidence Interval** (percent)
Vehicle-Miles of Travel Weekday		
Area-wide (all facility)	5	68
Major geographic sub-areas	5-15	68
Freeways	5	68
Arterial-Collectors	5	68
Local Streets	10-15	68
Vehicle-Miles of Travel Weekend		
Area-wide (all facility)	10	68
Major geographic sub-areas	10-20	68
Freeways	10	68
Arterial-Collectors	10	68
Local Streets	25	68
Screen-line and Cordon Volumes	5	95
Link Volumes		
Freeway	10	68
Arterials-Collectors	10	68
Local Streets	15-20	68

Sample Size Calculation

Determination of the required sample size is also an important task to validation. It gives planners an idea how many data points need to be collected from counting stations. GUTVC uses the standard sampling theory to calculate the sample size. Two basic techniques are used in determining the sample size for each subpopulation. Simple Random Sampling is used for local streets (with volumes under 2,000 vehicles per day) and for arterial streets and freeways where stratification by lanes and volumes is not possible. The following formula is used to determine the sample size for each geographic area from which to collect local traffic data:

$$n = \frac{Z^2 (C_S^2 + C_t^2)}{d^2 + \frac{C_S^2 Z^2}{N}} \tag{2.15}$$

if ignoring the finite population correction factor ($\frac{C_S^2 Z^2}{N}$),

$$n = \frac{Z^2 (C_1^2 + C_2^2)}{d^2} \tag{2.16}$$

where n = sample size;
Z = normal variate = 1.0 for 68 percent confidence, 1.45 for 85 percent confidence and 1.96 for 95 percent confidence;
C_s = spatial coefficient of variation;
C_t = temporal coefficient of variation; and
d = accuracy standard in traffic assignment as a function of facility type, facility size, and ADT.

Another technique, stratified random sampling, is utilized wherever freeway and arterial street volumes can be further grouped into volume strata. The narrower is the stratum, the lower is its variance and the required sample size. The sample size can be computed from the following formula:

$$n = \frac{\left(\sum W_h S_h\right)^2}{\frac{E^2}{Z^2} + \frac{\sum W_h S_h^2}{N}} \tag{2.17}$$

if ignoring the finite population correction factor ($\frac{\sum W_h S_h^2}{N}$),

$$n \approx \frac{Z^2 (\sum W_h S_h)^2}{E^2} \tag{2.18}$$

where W_h = weight of stratum h;
S_h = composite standard deviation of vehicle miles per mile in stratum;
E = absolute error in average vehicle miles per mile = (relative error) × (total vehicle miles) ÷ (total miles of roadway);
Z = normal variate = 1.0 for 68 percent confidence, 1.45 for 85 percent confidence and 1.96 for 95 percent confidence; and
N = total number of miles.

In response to FHWA's request to use this method in VMT estimation, the Florida Department of Transportation (FDOT) and Georgia Department of Transportation (GDOT) conducted studies to test the applicability of this method by comparing it with their own VMT estimation methods.

The study data FDOT choose is the historical traffic count from Brevard County in the state of Florida. As suggested in the GUTVC, FDOT's study area is categorized by area types (rural, central area and central business district) and facility types (local, arterial, and freeway-expressway). Before this study, FDOT employs sample counts and gasoline consumption information to arrive at VMT estimates. Their method can be described as follows: VMT of the county level is found for state highway (including federal route) system only; total VMT—that is, VMT for all facility types—is found for the entire state by use of gasoline consumption data; VMT by county is found by applying counts to the length of highway section from which they are taken to arrive at VMT for the sample link.

In FDOT's test, only the arterials are stratified by volume because sufficient historical data are not available for local and freeway links. The range of those volume strata is 5000 average daily traffic (ADTs). Simple random sampling technique is used in calculation for local streets and freeway/expressway, and stratified random sampling technique is applied for arterial streets. To remain consistent with the values suggested in the GUTVC, FDOT uses a confidence interval of 68 percent and the following relative errors by functional class: local streets, 15 percent; arterial streets, 5 percent; and freeway-expressway, 10 percent. Using formulas above, the required sample size and VMT for each area and facility is calculated.

The comparison shows that there is just relative small difference between VMT calculated by methods presented in the GUTVC and FDOT's VMT estimation program, which shows GUTVC's method is good. But FDOT also points out that there are some deficiencies in the application of the GUTVC's technique because of the high relative error calculated in the case study. They suggest that the use of correct and complete database for sample-size calculation would do the most to reduce relative error, which in the mean time can reduce the costs of data collection and can also be developed for future use. FDOT also find that by sectoring the area geographically, travel cost is reduced substantially, accuracy in VMT is not badly affected, and a useful form of VMT data is created. In addition, the GUTVC's method provides a general breakdown of VMT by area and facility type, which is not available by FDOT's method, and it also provides flexibility in the choice of area size, from which one may select a sample.

In 1977, recognizing the need for a uniform method of calculating estimations of VMT, GDOT contracted with FHWA to test the GUTVC's method. Before this test, GDOT has for a number of years provided a statewide estimate of vehicle miles of travel, and this statistics is based on traffic information collected by Georgia's coverage counting program. In this test, GDOT select the historical traffic data from the Savannah area in the state of Georgia. Each street is stratified by volume according to its functional classification. Since historical traffic data are available for freeway, arterial and collector systems in the area, stratified random sample methods are used in the determining the sample size for these groups. Because of the lack of available of historic data, simple random sampling methods are used for local system. During the analysis, GDOT utilizes numeric code to identify the geographic subarea, facility type, volume stratum, specific location, and mileage of each link in the study network, which can be compiled as a link network data file. This greatly facilitates their work to be conducted by computer.

Following GUTVC' procedure, GDOT calculated the sample size and VMT for each geographic subarea and each functional classification. A compassion of the total VMT estimation with the VMT produced annually through Georgia's coverage counting programming is made, which shows a difference of only 3.48 percent, and a comparison is also made between the two methodologies in the cost of field data collection, which suggests that GUTVC' method cost 111.78 percent more than Georgia's conventional method. GDOT find that the GUTVC' approach of sampling links in various strata selected randomly to account for spatial and temporal variations provides a uniform systematic method for computing VMT. And this idea of uniform systematic method is a positive approach to resolving a problem what will increase in magnitude as programs become more dependent on the VMT statistic.

Moreover, as is shown by the comparison, the temporal variation can be addressed in a less costly manner than that outlined in GUTVC with only a minor variation in results, because factors of temporal variation are readily available from continuous-count and seasonal-control programs that are currently maintained in most states. This study also shows that using counting locations that are randomly selected by functional classification is desirable to allow for spatial variation in computing VMT.

In summary, the technique used in GUTVC -- sampling traffic counts to obtain urban VMT is feasible. The more knowledge is available in advance of sampling, the more efficient is the overall sampling program. The separation analysis of sample size and VMT calculation by facility type and geographic area really reduce the variation within parts of the total sample. Hence, the stratified random sampling has merit over simple random sampling in estimating sample size and VMT.

2.5.2. Other Related Research Work

In the following, previous research on model calibration and validation for highway assignment is reviewed. Aiming at the unsatisfactory analysis about the development of efficient algorithm that produce equilibrium flows in the assignment of traffic to a road network, **Florian and Nguyen (1976)** at University of Montreal, Quebec, Canada, carried out a pilot implementation of the developed algorithms applied on the road network of the City of Winnipeg. They made an attempt at validating the behavioral hypothesis imbedded in the equilibrium methods for traffic assignment and presented their thoughts on some important steps that must be accomplished in order to bridge the gap between the theoretical and practical situation in this area.

The assignment problem is formulated as the following multi-commodity problem with convex costs on a network of nodes $n, n \in N$ and arcs $a, a \in A$.

$$\text{Min. } \sum_a \int_0^{v_a} s_a(v)dv \tag{2.19}$$

$$\text{S.t. } \sum_k h_{k,pq} = D_{pq}, all(p,q)$$

$$v_a = \sum_p \sum_q \sum_k \delta_{ak,pq} h_{k,pq}, all(a) \qquad (2.20)$$
$$h_{k,pq} \geq 0$$

where $s_a, a \in A$ is used to denote the link volume delay functions, $\delta_{ak,pq} = 1$, if a belongs to path k for O-D pair (p,q); $\delta_{ak,pq} = 0$, otherwise.

The volume decay function has the analytical form as follows:

$$s_a(v_a) = d_a \left[\delta + \alpha \left(\frac{v_a}{l_a} - \gamma \right) + \sqrt{\alpha^2 \left(\frac{v_a}{l_a} - \gamma \right)^2 + \beta} \right] \qquad (2.21)$$

where d_a is the length of the link, l_a is the number of lanes of the link, and $\alpha, \beta, \gamma, \delta$ are constants that vary for each of the curves.

The database the authors use comes from a transportation study of the City of Winnipeg, which was completed in 1966. In addition, a road network model of the city and an OD survey study is provided for their research. Two different algorithms are used to produce equilibrium flows for carrying out the assignments. One is an algorithm suggested by Murchland (1969) as an adaptation of the Frank and Wolfe algorithm for convex programming, which is considerably simple and requires the computation of shortest routes and one-dimensional minimization of the convex function. A modification is introduced in the Dijkstra-type shortest route algorithm to prevent the construction of a path that contains consecutive dummy links. This algorithm ensures that no anomalies are introduced in the resulting assignment due to the modification to the shortest route method. Another algorithm is developed by Nguyen (1974) as an adaption of the convex simplex method, which uses a mechanism of locating and tracing simplex-like cycles that improve the solution by modifying the flow on the arcs that belong to the cycle.

The assignment is carried out and the results are analyzed based on some plots, including the predicted volumes against the observed times, the computed times from the delay curves against the observed times. Shortest routes from each origin to all destinations are computed using only arcs that carried flow. The lengths of these routes, computed from the delay curves, are plotted against the observed lengths of these routes. Some refinements should be made to correct the inconsistencies and error of the transcription in the data, which includes: (1) the link and node diagram is checked toe ensure that network aggregation and abstraction are properly accomplished; (2) the volume-delay curve assigned to a link and the number of lanes are changed for some links; (3) some of the original data on observed volumes and observed travel times are checked and special situations are detected where the observed value is no longer representative of the actual situation; and (4) some of the elements of the origin-destination matrix are modified after a careful analysis of several suitably selected screen lines.

These plots show a good agreement between the predicted and observed values of the relevant parameters and thus demonstrate the suitability of the method for planning purposes. However, there still exist some differences between predicted and observed values of the relevant parameter, partly due to the limitation of the model to explain all route choice behavior as a function of time alone, and partly due to the way in which observed temporal values relate to predictions made in a static model. There are inherent variations in the observed volumes, peculiarities of aggregation in the network, and possible systematic biases in the volume delay curves. Some of the other observed differences might be due to trips that have intermediate destinations. In the mean time, the results are sensitive to the minor difference of the network and the specification of link parameters. This intrigues interest in developing automatic or semi-automatic network aggregation procedures.

Stover, Benson, and Buechler (1976) investigated the sensitivity issue of the results of traffic assignments to the inputs from proceeding modeling phases. In their study, the assignment results produced by different trip matrices were used to evaluate the sensitivity of various commonly used measures of assignment accuracy. The study found that, due to the aggregative nature of the assignment procedure, many of the differences observed at the zonal and zonal interchange levels tend to be washed out in the assignment results. This give an important practical implication that much of the precision in the preceding trip generation and trip distribution phases can be sacrificed, while still producing reasonably accurate assignment results. In other words, it implies that there might exist many-to-one relationship between trip tables and consistent assignment outputs with the observed ground counts. Therefore, a sketch planning process should produce assignment results of sufficient accuracy for valid evaluation and comparison of system alternatives.

EMME (Equilibre multimodal-multimodal equilibrium) is a two-mode transportation planning technique that is suggested by **Florian et al. (1979)**. It may be characterized as an integrated two-mode traffic equilibrium method. This method combines a zonal aggregate-demand model with an equilibrium-type road assignment and a transit assignment method. Its theoretical properties have been studied before, and herein the authors focus its validation and application by using data from the city of Winnipeg, Manitoba, Canada. The equilibrium-type route-choice model for travel by private automobiles in congested urban areas is validated in the Winnipeg road network, and the transit-assignment model is essentially a shortest-route choice coupled with the diversion mechanism among sections served by common lines.

In order to perform this work, some preliminary work has been done required by EMME, which includes the subdivision, speed-delay study in the summer of 1976 and departure code definition of the city of Winnipeg. In addition, bus travel times measured in the speed-delay study are used to recalibrate the volume-delay curves, which are used in the road assignment to calibrate the bus-automobile travel-time relationship. The authors use a new simplified BPR formula to recalibrate the volume-delay curves by using the 1976 data, which can fit the observed data better than the conventional function. It can be expressed in the form of:

$$S_a(v_a) = d_a t_0[1+\alpha({}^{v_a}\!/\!{}_{c_a})^{\beta}] \tag{2.22}$$

where d_a = the link length,

v_a = the link volume,

c_a = the practical capacity of the link

t_0 can be obtained from the road network data.

By using the departure codes, the total automobile origin-destination matrix can be obtained and a set of adjustment factors that serve to add other-purpose trip and truck trips, i.e.,

$$g_{pq}^{au} = (g_{pq} * {r_{pq}}/{\gamma_{pq}}) * f_{pq} \tag{2.23}$$

where $(p,q) =$ an origin-destination pair of zones,

g_{pq} = the total person work trips between q and p,

r_{pq} = the proportion of trips by automobile,

γ_{pq} = the automobile occupancy, and

f_{pq} = the factor for other trips and truck trips.

This matrix is then be assigned to the road network and the calibration of the road network is achieved by a trial-and error process that compares the observed link volumes with the link volumes predicted by the traffic-assignment model.

The coded transit network must be validated to ensure that it can represent adequately all possible passenger movements and transit vehicle movements. The tools used in validation are: (1) EMME data bank programs; (2) network generation programs; (3) graphical display of the network; (4) manual checks of the data; and (5) analysis of the complete printout of the transit assignment. The calibration of the transit network estimates the value of certain parameters of the transit path algorithm in order to produce feasible paths between the various origin-destination pairs. For each origin-destination pair the algorithm selects the path with minimum impedance from origin to destination based on the impedance function below:

$$\text{IMP} = \text{WALK* (access + egress time)} + \text{WAIT*} \sum_{\ell=1}^{n} w_\ell + \sum_{\ell=1} T_l + n * \text{WPEN} \tag{2.24}$$

where IMP denotes the path impedance,

WALK = the weight of walking time (access-egress) used in the calculation of the path impedance,

WAIT = the weight of waiting time used in the calculation of the path impedance,

WPEN = a constant penalty added to the impedance due to passenger waiting for the bus,

w_ℓ = the waiting time of the ℓ^{th} line, and

T_ℓ = the in-vehicle time spent on the ℓ^{th} line, which is assumed to have a wait of 1.0 in the impedance calculations.

Then the modal-split function is calibrated with the origin-destination survey data. Several functional forms and transformation of variables are chosen for the calibration, such as logistic function and “dogit” function, but none of them give a satisfactory result. The unavoidable error on individual origin-destination pair leads the subdivision of the origins into subgroups. Finally, four modal-split models are obtained. The base-year calibration is accomplished by using the bimodal model, which includes the simultaneous use of the vehicle assignment, the transit assignment and the modal-split function. Each is calibrated independently and then used jointly in the computations. The demand differences are not large and within the variability range that is acceptable in calibration of transportation models.

Overall, one can conclude that the EMME model is feasible. There are several ways to use EMME to simulate the impact of contemplated improvement scenario. One may use the single-mode assignment modules and thus simulate the impact of the scenario without changing the modal share of the demand. The other way to use EMME is to simulate the impact of each scenario with a full bimodal run, which is unique to EMME. In addition, cost of implementing the model is not expensive. However, there still exist some problems. The main obstacles are the quality of the available data and the calibration of the demand model. If good quality of data becomes accessible, one would succeed with great chance in calibrating a satisfactory modal-split model.

Traffic assignment is the process of allocating a set of present or future trip interchanges to a specified transportation network. Its results have wide application, which includes many planning and design decision-making processes. Therefore, it is very important for one to produce high-quality forecast, especially with the help of microcomputer technologies in this work. **Easa (1991)** gives an overview of the traffic-assignment elements and a synthesis of practical problems faced by the users.

The paper focuses on the five basic components of the application of computerized traffic-assignment models, there are: (1) preparing the network; (2) establishing the OD demands; (3) identifying a traffic-assignment technique; (4) calibrating and validating a model; and (5) forecasting. It also analyzes the practical problems related to traffic assignment models, which can be classified into four categories: network representation, system-to-subarea data translation, model-calibration and forecasting.

Network-representation problems include coding that might allow illogical movements, coding of special movements or operational strategies, and network aggregation. Other data-related problems arise when it is necessary to establish traffic data for subareas based on available system-level forecasts. The representation can be simplified if the conventional shortest-route algorithm with movement prohibitions is implemented in the traffic-assignment model. This can ensure that illogical movements are eliminated, which is particularly important for micro-macro network coding and coding of centroids and connectors. If the model has a turn penalty or prohibition capability, illogical movements can be eliminated without special network coding.

Origin-destination (OD) demand matrix can be established in two situations. For short-range analysis, the base year OD matrix can be adjusted to represent the horizon year by using any of the growth-factor methods. For long-range analysis, the base year OD matrix is used along with the land-use, economic, population and transportation network data to estimate the future OD matrix.

This paper emphasizes the application of user-equilibrium technique in traffic-assignment. This technique utilizes mathematical optimization, which not only guarantees convergence for convex link-performance functions, but is also capable of accommodating large-scale networks. The stochastic multinomial probit technique of traffic assignment is conceptually the most promising, but the technique is not yet readily available to practitioners. For congested networks, the equilibrium mathematical optimization technique can produce results comparable with those of the stochastic equilibrium technique.

In model calibration and validation, the acceptable deviation of the assigned link volumes from actual counts should vary according to the level of the observed link volume. *The acceptable deviation should be higher for low-volume links and lower for high-volume links.* The reason for this is that large deviations on low-volume facilities would not have major design implications. If the measures predicted by the traffic-assignment model are not sufficiently close to the observed measures of the base year, one or more of the following elements must be modified, including base-year OD matrix, network representation, assignment technique, and/or link-performance function.

Forecasting is to use the validated model to evaluate the impacts of proposed improvements or changes to the study area. Forecasting problems include checking the accuracy of traffic-volume forecasts, updating an old OD matrix, establishing traffic volumes for a different year, and applying traffic-assignment technique in site-impact studies. Some preliminary checks should be performed on the system-level traffic forecasts to ensure the accuracy and reasonableness of the traffic-assignment model, even though it has been calibrated and validated. These checks include: (1) examining land-use data assumptions; (2) comparison of trip end summaries with land-use data; (3) comparison of forecasted traffic growth with historical growth trends. These checks, in addition to the base-year validation, should be performed to determine the overall accuracy of the traffic forecasts prior to applying the project-level refinement.

Finally, the author emphasizes that research should continue to develop more efficient methods for solving complex assignment problems, such as stochastic network equilibrium, and to investigate the assumptions involved in the route-choice analysis.

CHAPTER THREE

LITERAURE REVIEW ON ACCURACY STANDARDS FOR TRANSIT NETWORKS

Transit network validation has traditionally been a low priority because transit modes account for only a very small portion of regional travel in many areas. For small and medium size urban areas, it is not unusual for the modelers to disregard transit trips completely if transit patronage is insignificant. However, as some of the major urbanized areas, especially in Florida, are reaching the "built-out" status with very limited right-of-way for highway expansion, the importance of transit demand forecasts will become more prominent. Validation standards for transit assignment model will thus have to be established anew in meeting the emphasis of transit and multi-modal planning.

While a number of reports have described the calibration and validation of regional travel demand models, few have provided the acceptable model errors for validating a transit model. A TMIP major document in this area, entitled Model Validation and Reasonableness Checking Manual, gives specific guidelines for each step of the traditional four-step modeling process and suggests the performance standards for validation of each model step. However, thresholds for validation of transit models have not been included.

3.1 Transit Validation Process

As in the highway modeling, each step in the transit modeling process may contribute to the overall error. While there is a potential for the errors to offset each other, there is no guarantee that they will. Each step in the model chain incorporates the results from the previous steps and should be validated separately to reduce the compounding of errors. The performance of transit models depends not only on the proper calibration of parameters, but also on careful review of socioeconomic inputs and transportation system characteristics. Consequently, individual model validations should be applied to make sure that each component reasonably reproduces the observed travel characteristics. After each component has been validated, validation of the overall set of models is performed to ensure that each component is properly interfaced and to test the effects of compounding errors.

A number of measures have been used to compare the observed and estimated values from a given model output over a number of observations. The more common measures include (Barton-Aschman Associates, Inc. and Cambridge Systematics, Inc., 1997):

1. Absolute difference: It is calculated as the estimated value minus the observed value. The sign (positive or negative) may be an important indicator for revealing systematic biases.

2. Relative difference: Values are normalized to remove scaling effects. It can be expressed as a percentage difference or as a ratio and are calculated as follows:

$$\text{Percentage Difference} = \frac{\text{Estimated - Observed}}{\text{Observed}} \times 100$$

$$\text{Ratio} = \frac{\text{Estimated}}{\text{Observed}}$$

3. R-squared: The coefficient of determination R^2 estimates the correlation between the observed and estimated values, i.e., how well a regression fits the relationship. R^2 can range from 0 to 1, with a value of 0 for no correlation and 1 for perfect correlation. Acceptable values of R^2 can vary depending on the type of comparison being made, but it would ideally explain more than half of the variation ($R^2 > 0.5$).

4. Percent RMSE: The percent root mean square error defined as follows is usually used to measure the variance between observed and estimated values.

$$\%\text{RMSE} = \frac{\sqrt{\sum_i \left(\text{Estimated}_i - \text{Observed}_i\right)^2 \Big/ \left(\text{Number of Samples-1}\right)}}{\sum_i \text{Observed}_i \Big/ \text{Number of Samples}} \times 100$$

3.2 Stepwise Validation

The assignment of transit trips to the network is the primary output of the transit modeling process. The execution of transit assignment requires completion of all previous steps in the model chain. Validation of the trip assignment step is built upon the successful validation of each of the previous steps. This validation is not meaningful if significant errors persist in the earlier steps. The input for transit assignment model includes the coded networks and the person trip tables. It is critical that transit network data be checked prior to other steps in validation. FDOT (1997) specified that the data to be validated include overall fleet size, peak vehicle requirements, systemwide total vehicle hours of service, systemwide total vehicle miles of service, route miles, and round trip times for individual routes. Any discrepancies (between the actual route data and the network data) that are greater than about five percent or one-half mile should be resolved.

Most travel demand forecasting software can now display transit network for error checking, such as zone to zone skims. At system-level checks for transit networks, minimum and maximum headways and range of walk or auto access times to stations/bus stops should be used as the validation check. Bus speeds should also be checked to ensure that they are less than or equal to auto speeds.

Evaluation of the performance of the mode choice model focuses on the proportional shares of total trips estimated by the model to be highway or transit, as well as the total numbers of transit person trips made on an average weekday. The application of mode choice is determined by the complexity of the study area's transit system. Mode choice models require a number of inputs including transit travel time (out-of-vehicle, in-vehicle, walk time, wait time), number of transfers, highway travel time, transit fare, auto costs, household income and/or auto ownership, household size, number of workers, and land use characteristics. All of these inputs should be reviewed for reasonableness and compared with observed values or parameters estimated in other regions.

To validate models at the aggregate level, mode shares by trip purpose should be subdivided into submode shares by purpose. The validation checks for mode choice models can include (Barton-Aschman Associates, Inc. and Cambridge Systematics, Inc., 1997):

- Transit ridership, highway vehicle, and auto occupancy counts at screenlines by time of day.
- Home-based work trips by mode and origin and destination district.
- Total patronage by transit mode.
- Counts of transit patrons by access mode at major stations serving transfers between auto and feeder bus and express transit services.
- Average auto occupancies by trip purpose.
- Percent single occupant vehicles by trip purpose.
- Home-based work transit trips as a percent of total transit trips.
- Mode shares to/from area types or major districts.
- Average auto occupancies to/from area types or major districts.

The validation of the transit assignment is the final validation of the complete transit model set. Typically, the most crucial and extensive validation efforts were based on the assigned transit ridership. This process is accomplished by comparing the base-year transit usage resulting from travel survey with the results obtained from the model prediction. The validation procedure of the transit assignment model is essentially the same as the highway assignment model, except that transit ridership counts would replace traffic counts (JHK & Associates and Dowling Associates, 1992). The transit validation can include analysis of transit boardings for the following comparisons (Barton-Aschman Associates, Inc. and Cambridge Systematics, Inc., 1997):

- Observed vs. estimated boardings for region, by mode, by time of day, and by trip length.
- Observed vs. estimated transfers per trip.
- Observed vs. estimated screenline volumes.
- Observed vs. estimated boardings by route or group of routes.
- Observed vs. estimated district-to-district transit trips.

It should be noted that both highway and transit counts are only estimates of traffic volumes and transit ridership, respectively. These count data along with the other input information should be tested for reasonableness during validation. It should be recognized that ridership on individual transit route are subject to the day-to-day inherent variability, and the error in counting ridership for any specified period of time is often large.

The procedure described above presents a general practice for validating transit models, although many regions may have their own tests. In general, the process used to validate a travel model is dependent on the purpose of the model, available data resources, model structure, and desired level of accuracy. For example, the North Central Texas Council of Government (NCTCOG) recommended that the following list be examined for the base validation year, when it comes to subsequent statements of transit forecastability:

- Observed vs. modeled similarities in peak and off-peak transit speeds used in skimming for mode choice.
- Model-derived vs. actual bus and train VMT for the peak and off-peak periods.
- Modeled vs. observed weekday riders by bus route and rail route (route-level RMSE and percent error).
- Modeled vs. observed weekday riders by route groups (e.g., by quadrants, radial corridors, or other meaningful sub-areas).
- Modeled vs. observed weekday rail station boardings (station-level RMSE and percent error).
- Reasonableness of modeled vs. observed mode of access distributions to individual rail stations (e.g., access by park-and-ride, walking, transfer from a bus, or transfer from another rail route).
- Sensitivity checks of the "elasticity reasonableness" of the calibrated model parameters (e.g., how much does the transit ridership change if CBD parking costs, transit fares, auto operating costs, in-vehicle roadway/transit travel times, transit headways, etc. are separately increased/decreased by 10% in separate model runs).

3.3 Transit Accuracy Standards

Model validation requires the checking of the differences between the estimated volumes and the observed patronages at levels ranging from system-wide (e.g., across screenlines), major movements (e.g., across cutlines), to individual links. Currently, there are no absolute measures or thresholds that can be achieved to say if a model is truly validated. The desired level of accuracy expected of a model depends on the time and resources available and the intended application of the model.

Accuracy standards for highway model can be set on the basis of the travel volumes. It assumes that the maximum desirable traffic assignment should not affect the number of lanes required to handle the estimated volume (Tekiele, 1993). However, a similar basis for developing the accuracy standards for a transit modeling procedure is not apparent. Although absolute criteria for assessing the validity of transit model cannot be precisely defined, a number of target values can be developed statistically. These values can be expected to provide some guidance for evaluating the relative performance of a particular models. It is important to recognize that a travel demand model is constituted with a series of steps with many built-in assumptions, that the model is just an approximation of real traveler behavior and cannot be expected to be able to perform exactly at a micro-level. Overly optimistic about matching the modeled volumes to ground counts should be avoided (Ismart, 1990).

A comprehensive literature search on the accuracy standards of transit modeling did not produce any theoretical or conceptual approaches for determining the appropriate levels of accuracy. Further, little has been done to dictate some suggested specifications of accuracy. As with the highway assignment performance, there are a number of transit assignment performance standards that need to be met to successfully calibrate the model. Table 3.1 gives the estimated accuracy of some typical assigned flows in the transit modeling process. In general, links with higher volumes tend to achieve a higher level of accuracy while those with lower volumes can be expected to have a lower level of accuracy.

Table 3.1 Estimated Accuracy of Parameters for Public Transit Loading (Robbins, 1978)

Parameter	Typical Magnitude	95% Confidence Limit
Average Rural Link	< 500 passengers	Extremely inaccurate
Average Urban Link	5,000 passengers	> ±46%
Important Urban Link	10,000 passengers	> ±33%
Major Urban Link	20,000 passengers	> ±23%

Note: These figures do not allow for errors in forecasting input parameters.

The comparisons between model performance and actual counts at the following levels were recommended during the Phase II development of the FSUTMS Model Updates (COMSIS Corporation, 1981):

- Areawide Transit Usage
- Cutline Comparisons
- Screenline Comparisons
- Route Comparisons

Areawide and screenline comparisons are basic and should be accomplished in both validation and calibration. Screenline comparisons are used to check major directional movements across the entire study area. Cutline comparisons are most valid as an assignment check where routes are grouped into major corridors and comparisons made against ground counts. Route checks can be made for calibration by assigning survey and model data and comparing the results on a route- by-route basis.

The standards of accuracy established were based upon experience, review of both highway and transit results, and intuition. Table 3-2 gives the first set of transit validation standards for FSUTMS. Overall, for all routes, the daily transit trip assignments should be within ± 3 percent of the existing ridership. For individual lines the standards are based on the number of daily riders on each route. These standards, however, were found to be too low and would normally be met by even a relatively poor model. Accordingly, the standards were tightened during Phase III development of FSUTMS Model Updates to provide a better means to distinguish between an acceptable model from an unacceptable one (Schimpeler-Corradino Associates, 1984). These standards, as given in Table 3-3, were developed based on the assumption that the standard modal-split models have been updated to the local area by adjusting the model parameters and were believed to be more closely responsive to actual needs.

Table 3.2 Initial FDOT Transit Validation Standards (COMSIS Corporation, 1981)

Measure	Acceptable Error	Acceptable Range Estimated/Actual
Total Area Transit Trips	± 3%	0.97 – 1.03
Trips Entering the Central Area	± 5%	0.95 – 1.05
Sectors of the Central Area Boundary (perhaps 4-5 corridors)	± 20%	0.80 – 1.20
Total Area Transit Average Trip Length	± 5%	0.95 – 1.05
Cutlines and /or Routes		
< 1,000 Passengers/Day	-100% – +150%	0.00 – 2.50
1,000 – 2,000 Passengers/Day	± 90%	0.10 – 1.90
2,000 – 5,000 Passengers/Day	± 70%	0.30 – 1.70
5,000 – 10,000 Passengers/Day	± 45%	0.55 – 1.45
10,000 – 20,000 Passengers/Day	± 35%	0.65 – 1.35
> 20,000 Passengers/Day	± 30%	0.70 – 1.30

Table 3.3 Updated FDOT Transit Validation Standards (Schimpeler-Corradino Assoc., 1984)

Measure	Acceptable Error	Acceptable Range Estimated/Actual
Total Area Transit Trips	± 1%	0.99 – 1.01
Trips Entering the Central Area	± 2.5%	0.975 – 1.025
Screenlines	± 10%	0.90 – 1.10
Total Area Transit Average Trip Length	± 5%	0.95 – 1.05
Corridors or Cutlines		
< 1,000 Passengers/Day	±100%	0.00 – 2.00
1,000 – 2,000 Passengers/Day	± 65%	0.35 – 1.65
2,000 – 5,000 Passengers/Day	± 35%	0.65 – 1.35
5,000 – 10,000 Passengers/Day	± 25%	0.75 – 1.25
10,000 – 20,000 Passengers/Day	± 20%	0.80 – 1.20
> 20,000 Passengers/Day	± 15%	0.85 – 1.15

3.4 Summary

Travel demand model validation has traditionally been focused on highway modeled volumes. These models are typically calibrated to a base year in accordance with the validation standards for acceptable level of error as set forth by the Federal Highway Administration, or by locally adjusted standards on top of that. Although a number of highway model performance standards manual/guidelines have been developed in many regions, the validation standards for transit model are currently in the fairly early stages of development.

Transit assigned volumes provides a best index to evaluate the performance of the transit modeling process. However, if the assigned transit ridership shows significant deviation from the observed values, it could be the possible effect of compounding error from the previous steps in model chain. Thus, it is important to validate each individual model component as thoroughly as possible prior to performing transit assignments. The error inherent in the collection of input data or validation data is always a concern and should be checked first. Key model parameters are then modified with reasonable values until the model replicates observed values with an acceptable range of error. The accuracy standards can be built upon afterward to identify the acceptable model error.

The only existing accuracy standards with specific threshold values have been established by the FDOT under the Phase III development of FSUTMS Model Updates in early 1980. The threshold values used in these standards were set based mainly on experience and intuition and there is no theoretical basis to support the use of these values.

CHAPTER FOUR

ACCURACY STANDARDS FOR HIGHWAY NETWORKS

Most of the current accuracy standards at the nation, state, and county levels adapt from the Guide to Urban Traffic Volume Counting (GUTVC) developed by the Federal Highway Administration (FHWA) (Levinson and Roark, 1975). This guide provides a uniform method of calculating estimation of VMT based on random selection, and sampling of highway links that are stratified by volume, facility type, and area type. Using the sampling theory, the required sample size at a desired confidence interval can also be calculated.

The accuracy standards suggested in GUTVC are in fact referred from previous research works, which are primarily based on empirical experiences and subjective judgment from actual planning model implementation. No further theoretical justification is provided in the GUTVC report. However, it was mentioned that those standards should be modified to reflect the rates of travel growth in each urban area. Specific urban areas or states may modify accuracy standards based on experience, field tests, and rates of travel growth. This study intends to investigate the use of statistical procedures to establish theoretical accuracy standards, and compare against the current counterparts. In addition, the required sample sizes for achieving current accuracy standards on assigned volumes and VMT are also calculated based on the detector counts available within the State.

4.1. Accuracy Standards for Assigned Volume

Validation of the four-step transportation planning model often employs area-wide Root Mean Square Error (RMSE) between ground counts and link assigned volumes as an indictor for gauging the model performance. Let v_i and $\hat{v}_i$ be ground counts and assigned volumes on links. The RMSE is usually computed as:

$$E\,(\%) = \frac{\sqrt{\dfrac{\sum_i (\hat{v}_i - v_i)^2}{N}}}{\dfrac{\sum_i v_i}{N}} \times 100 \tag{4.1}$$

This assignment error may come from various sources, such as:

1. Statistical variation of ground counts,
2. Recording error with ground counts,
3. Model misspecification errors,
4. Errors accumulated from demand forecasting steps prior to assignment, and
5. Incorrect or insufficient network geometry.

When these error causes compound together, their effects on the final assignment error are either additive or cancelled out each other. Therefore, it is rather difficult to differentiate and quantify their individual effect. Without loss of generality, statistical variation and model misspecification are selected for analysis purpose. The data recording error is assumed screened or corrected in the data collection process. Error propagated from steps prior to assignment is deemed not quantifiable or estimable unless full rounds of simulation are executed over the entire forecasting process (Zhao and Kockelman, 2001). Finally, network geometry is case dependent and would be extremely difficult for one to establish a quantitative relationship with the assignment error.

4.1.1. Accuracy Standards Due to Data Statistical Variation

To establish accuracy standards for the link assigned volume, all the available detector data collected from permanent and portable count stations in State of Florida in year 2000 were used to compile necessary statistical information. Continuous daily traffic (DT) data were available from permanent stations, with possibly a small portion of data missing due to detector malfunctioning or other reasons. By examining the detector data within the year, one can find that the data pattern repeat itself according to day of week, and the variation of data within a specific detector station is less than the counterpart among detector stations. To obtain representative degree of variation, one needs to eliminate the within-station data similarity to a great extent by performing sample size reduction.

To that end, for a specific count station, the mean and standard deviation computed from weekly DT were compared with the counterparts calculated from the whole year. The *t* and *F* tests were used to verify the hypotheses that mean and standard deviation of weekly DT are equal to the annual counterparts. Assuming that the weekly AADT data and the annual AADT data are both normally distributed, the pooled *t*-test is used to test if the weekly AADT data and the annual AADT data have the same mean at the 95% level of significance. The *t* value can be calculated as:

$$t = \frac{|x_1 - x_2|}{s_p \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \tag{4.2}$$

where $s_p = \sqrt{\dfrac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}}$,

x_1 = mean value of weekly AADT,
x_2 = mean value of annual AADT,
n_1 = sample size of weekly AADT (= 7), and
n_2 = sample size of annual AADT.

The pooled *t*-test has a freedom of n_1+n_2-2. In the same way, since n_1+n_2-2 is sufficiently large, the critical value for *t*-test is 1.96. As listed in Tables I.1 through I.15 in the Appendix, one can see that all the computed *t* values are less than the critical value to support the null hypothesis, i.e., H_0: $x_1 = x_2$.

One can also test if they have the same variance at a significance level of 95 percent with the *F* test. Let σ_1^2 and σ_2^2 be the variances from the weekly AADT data and annual AADT data, the hypotheses can be set as H_0: $\sigma_1^2 = \sigma_2^2$ versus H_1: $\sigma_1^2 \neq \sigma_2^2$. For each continuous count, use the following formula to compute the *f* value.

$$f = \frac{s_1^2}{s_2^2} \tag{4.3}$$

where s_1 = standard deviation of weekly AADT, and
s_2 = standard deviation of annual AADT.

The *F* test has a degree of freedom equal to n_1+n_2-2. Since n_1+n_2-2 is sufficiently large, the critical region, $f_{(0.05,\, n1,\, n2)}$, is approximately 2.10. Again, as listed in Table I.1 through I.15, one can see that the computed *F* test value is less than the critical value to support the null hypothesis. It is therefore concluded that the weekly AADT data and the annual AADT data have similar mean and variance at the 95 percent level of significance, suggesting that the weekly DT data are sufficient to represent the annual DT data. It implies that, with proper sampling schedule, one could place a portable counter for a week to obtain daily traffic measure that is representative of the whole year. Based on a similar testing procedure, the 5-day and 3-day DT data are also proven sufficiently representative of the annual DT data.

In the following exercise, all the continuous detector counts from permanent count stations are evaluated with same weights as the counts from portable count stations. A total of seven day's daily traffic volumes that pass the *t* and *F* tests are retrieved from each permanent counter and combined with portable count data to form the primary database for computing accuracy standards. Accuracy standards are composed according to *facility type* and *size of facility* with the categorization consistent with the current accuracy standard. Types of facility include freeway/expressway, divided arterial, undivided arterial, one-way facility, toll facility, and collector. Depending on the facility type, the size of facility ranges from 1 to 5 lanes per direction. Accuracy standard is computed according to the following equation:

$$d = \frac{ZC_v}{\sqrt{n}} \times 100 \tag{4.4}$$

where n = sample size;
Z = normal variate = 1.0 for 68 percent confidence, 1.45 for 85 percent confidence and 1.96 for 95 percent confidence;
C_v = coefficient of variation; and
d = allowable error limit (accuracy standard) as a function of facility type and size of facility size (in percent of the mean ADT.

Factors affecting the accuracy standards include coefficient of variations and the sample size. As the ratio of standard deviation to mean, coefficient of variation reflects estimableness of the mean ADT. When the ratio approaches one, the mean ADT is basically unpredictable using any predictor other than the mean itself. In the dataset, this indicator behaves as a decreasing function of volume level. As shown in Figures 4.1 through 4.3, it is indicated that the traffic pattern is fairly stable at medium to high volume levels since the coefficient of variation rapidly drops below 10% when the volume level goes beyond 10,000 vpd for all facilities. The trends are similar among the three times of day. In addition, one can find that the coefficient of variation is significantly lower when grouped by the mean level of traffic carried than when grouped by the facility type.

As the sample size increases, the accuracy of the estimate will be improved, resulting in a decrease in the allowable error limit. The trade-off is the costs of data collection. These two factors, including cost and accuracy, should be considered at the same time to assess the number of detector stations to achieve the desired results (Levinson and Roark, 1975). As demonstrated earlier, the 7-day ADT count samples can be used to represent the AADT sufficiently based on the first two moments of statistics. This is considered very cost-effective when the sampling schedule of the week data is chosen appropriately. In this study, all the count samples after size reduction will be used to derive the allowable error limit. Depending on the time of day, i.e., morning rush hours, mid-day, and afternoon rush hours, the accuracy standards are established at the 68%, 85%, and 95% levels of significance. The results based on the 7-day (week), 5-day, and 3-day sample sizes are listed in Tables 4.1 through 4.3, Tables 4.4 through 4.6, and Tables 4.7 through 4.9, respectively.

Under the current infrastructure and spatial density of detector information, the allowable error limits can vary in a wide range, from less than 1% to almost 40%. One could always enhance the sample size to reduce the error limits for those facilities of different size with limited samples. When compared with the current accuracy standards, the proposed error limits due to detector samples should be regarded as the lowest possible bounds. The actual errors in the assigned volumes also come from other sources such as model specification error. The specification errors associated with the traditional four-step planning models will be extremely difficult to quantify, since they may be compounded or canceling out each other when propagating through the individual model components sequentially. As a comparison with Tables 4.1 through 4.3, it seems that all the current accuracy standards (see Table 2.3) exceed the proposed lowest bounds except for 4-lane one-way facility (13% versus 16% at 68% level of significance).

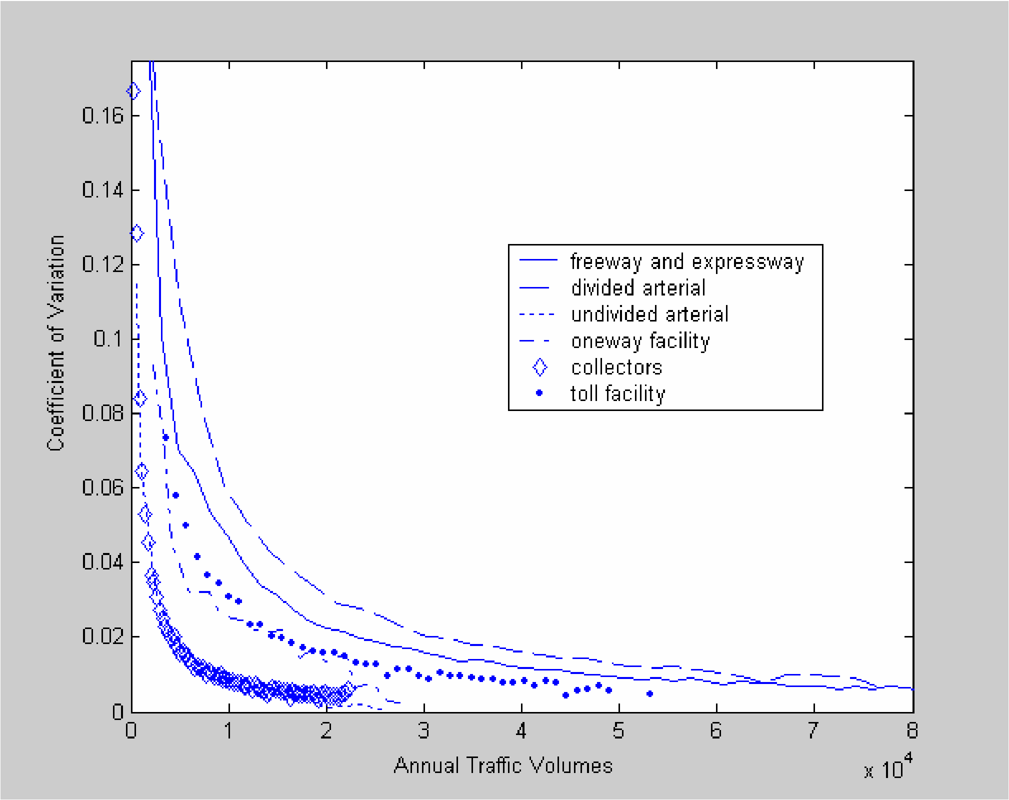

Figure 4.1 Coefficient of Variation for All Facilities (Morning rush hours)

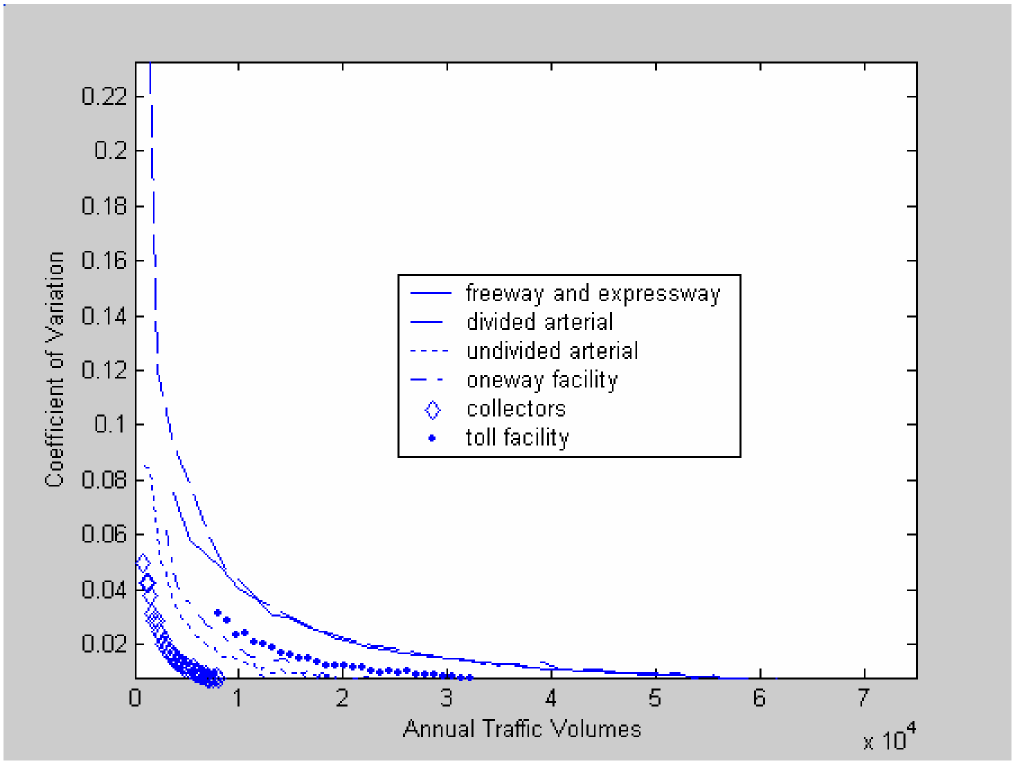

Figure 4.2 Coefficient of Variation for All Facilities (Midday)

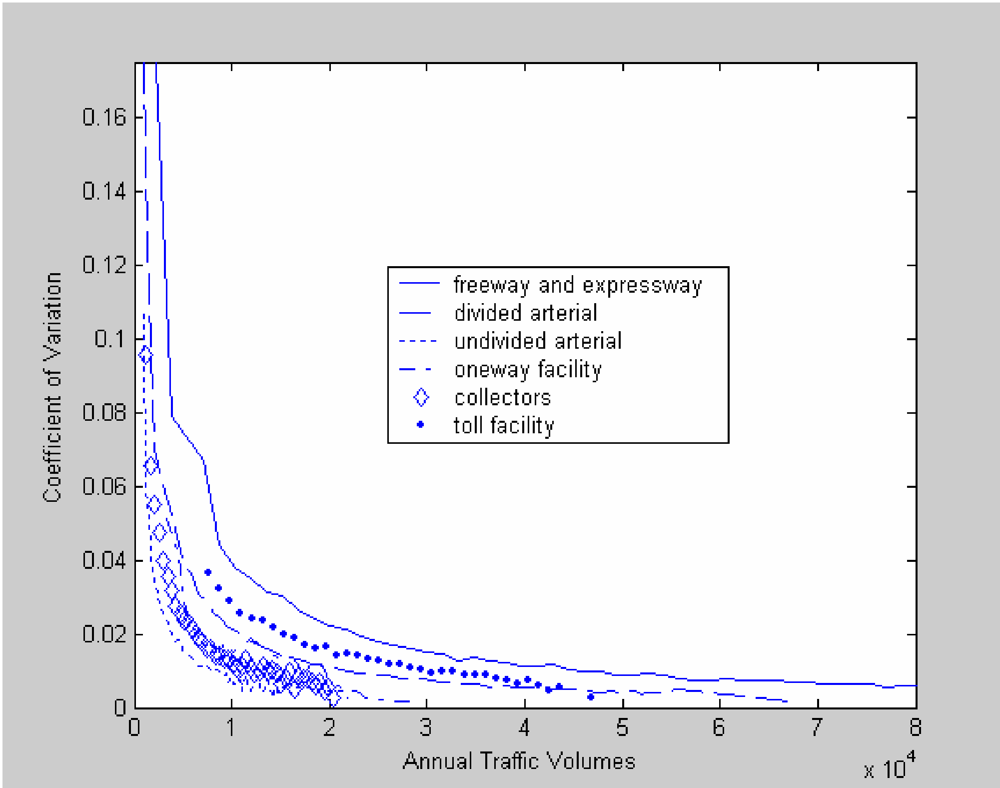

Figure 4.3 Coefficient of Variation for All Facilities (Afternoon rush hours)

Table 4.1 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 7-day (Week) Sample Size (Morning Rush Hours)

Facility Type	**No. of Lanes**[1]	**ADT Ranges (Mean)**	**Sample Size *n***	C_v	**Allowable Error Limit (%)**		
					68%[2]	**85%**	**95%**
Freeway/ Expressway	≥5	15707~169065 (117190)	143	0.2875	2.40	3.49	4.71
	4	15291~126375 (72080)	273	0.3749	2.27	3.29	4.44
	3	6074~125117 (52905)	639	0.4753	1.88	2.73	3.69
	2	4022~91856 (25435)	625	0.5713	2.29	3.31	4.48
Divided Arterial	4	5216~49024 (25516)	167	0.3684	2.85	4.13	5.59
	3	4179~78792 (22775)	2400	0.4600	0.94	1.36	1.84
	2	648~74815 (15315)	4695	0.5153	0.75	1.09	1.47
	1	409~22001 (7830)	780	0.4888	1.75	2.54	3.43
Undivided Arterial	3	2912~32877 (13619)	20	0.6652	14.87	21.57	29.15
	2	1599~26482 (10189)	344	0.4806	2.59	3.76	5.08
	1	658~19896 (6078)	709	0.5245	1.97	2.86	3.86
One way facility	4	6041~29343 (15663)	12	0.5553	16.03	23.24	31.42
	3	2755~27767 (15541)	74	0.3877	4.51	6.54	8.83
	2	2454~32868 (11679)	118	0.5320	4.90	7.10	9.60
	1	448~8503 (4347)	8	0.5576	19.71	28.58	38.64
Toll Facility	3	5658~45392 (31937)	28	0.4194	7.93	11.49	15.54
	2	3803~39529 (19296)	84	0.5162	5.63	8.17	11.04
Collectors	2	1913~23803 (9169)	185	0.4928	3.62	5.25	7.10
	1	393~16200 (4876)	422	0.5502	2.68	3.88	5.25

Note: 1. In one direction
2. Level of significance

Table 4.2 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 7-day (Week) Sample Size (Mid-day)

Facility Type	**No. of Lanes**[1]	**ADT Ranges (Mean)**	**Sample Size *n***	C_v	**Allowable Error Limit (%)**		
					68%[2]	**85%**	**95%**
Freeway/ Expressway	≥5	16786~158354 (116730)	143	0.2538	2.12	3.08	4.16
	4	17516~ 118409 (72266)	273	0.3135	1.90	2.75	3.72
	3	8167~ 111878 (52817)	639	0.4018	1.59	2.30	3.12
	2	5727~79280 (25377)	625	0.4848	1.94	2.81	3.80
Divided Arterial	4	8234~49728 (25707)	167	0.2766	2.14	3.10	4.20
	3	4951~81108 (22762)	2400	0.3733	0.76	1.11	1.49
	2	800~48365 (15220)	4695	0.4177	0.61	0.88	1.20
	1	477~21717 (7819)	780	0.4333	1.55	2.25	3.04
Undivided Arterial	3	2990~23367 (13457)	20	0.6234	13.94	20.21	27.32
	2	1700~25068 (10210)	344	0.4393	2.37	3.43	4.64
	1	19~22072 (6072)	709	0.4648	1.75	2.53	3.42
One way facility	4	7292~29343 (15663)	12	0.4857	14.02	20.33	27.48
	3	3043~28700 (15541)	74	0.3884	4.51	6.55	8.85
	2	2454~28862 (11659)	118	0.5321	4.90	7.10	9.60
	1	795~8503 (4633)	8	0.4879	17.25	25.02	33.81
Toll Facility	3	23735~39410 (30506)	28	0.1504	2.84	4.12	5.57
	2	8863~34158 (18180)	84	0.3993	4.36	6.32	8.54
Collectors	2	1858~21196 (9122)	185	0.4363	3.21	4.65	6.29
	1	1093~13341 (4876)	422	0.4151	2.02	2.93	3.96

Note: 1. In one direction
2. Level of significance

Table 4.3 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 7-day (Week) Sample Size (Afternoon Rush Hours)

Facility Type	**No. of Lanes**[1]	**ADT Ranges (Mean)**	**Sample Size** ***n***	C_v	**Allowable Error Limit (%)**		
					68%[2]	**85%**	**95%**
Freeway/ Expressway	≥5	15805~164776 (116070)	143	0.2670	2.23	3.24	4.38
	4	16792~127710 (72950)	273	0.3290	1.99	2.89	3.90
	3	6865~113299 (53142)	639	0.4210	1.67	2.41	3.26
	2	4575~71303 (25343)	625	0.4970	1.99	2.88	3.90
Divided Arterial	4	5867~45767 (25709)	167	0.3140	2.43	3.52	4.76
	3	1575~77133 (22747)	2400	0.3893	0.79	1.15	1.56
	2	752~53912 (15245)	4695	0.4333	0.63	0.92	1.24
	1	496~20768 (7822)	780	0.4464	1.60	2.32	3.13
Undivided Arterial	3	2673~29126 (13434)	20	0.6459	14.44	20.94	28.31
	2	1612~28759 (10208)	344	0.4546	2.45	3.55	4.80
	1	918~18094 (6066)	709	0.4758	1.79	2.59	3.50
One way facility	4	7602~29343 (15663)	12	0.4527	13.07	18.95	25.62
	3	2985~27218 (15541)	74	0.3890	4.52	6.56	8.86
	2	2454~31691 (11669)	118	0.5301	4.88	7.08	9.56
	1	902~8503 (4489)	8	0.5025	17.77	25.76	34.82
Toll Facility	3	19202~51209 (31961)	28	0.2689	5.08	7.37	9.96
	2	9491~36994 (18647)	84	0.4105	4.48	6.49	8.78
Collectors	2	1716~23074 (9123)	185	0.4583	3.37	4.89	6.60
	1	1030~13599 (4918)	422	0.4555	2.22	3.21	4.35

Note: 1. In one direction
2. Level of significance

Table 4.4 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 5-day Sample Size (Morning Rush Hours)

Facility Type	No. of Lanes[1]	ADT Ranges (Mean)	Sample Size n	C_v	Allowable Error Limit (%)		
					68%[2]	85%	95%
Freeway/ Expressway	≥5	15707~169065 (117823)	137	0.2829	2.41	3.49	4.72
	4	15291~127470 (71932)	259	0.3718	2.31	3.35	4.53
	3	5360~12913 (53262)	591	0.4653	1.91	2.77	3.74
	2	4022~91856 (25474)	573	0.5716	2.38	3.45	4.66
Divided Arterial	4	5494~49024 (25560)	159	0.3612	2.86	4.15	5.61
	3	4174~78792 (22700)	2359	0.4576	0.94	1.36	1.84
	2	648~70564 (15503)	4467	0.5032	0.75	1.09	1.47
	1	409~22001 (7817)	777	0.4943	1.77	2.57	3.47
Undivided Arterial	3	2912~32877 (13619)	20	0.6652	14.87	21.56	29.15
	2	1687~26482 (10237)	340	0.4785	2.59	3.76	5.08
	1	658~19896 (6116)	689	0.5232	1.99	2.89	3.90
One way facility	4	6041~29343 (15663)	12	0.5553	16.03	23.24	31.42
	3	2755~27767 (15541)	74	0.3877	4.51	6.54	8.84
	2	2454~32868 (11679)	118	0.5320	4.90	7.11	9.60
	1	448~8503 (4347)	8	0.5576	19.71	28.58	38.63
Toll Facility	3	5438~45392 (32332)	20	0.4189	9.37	13.59	18.37
	2	4072~38207 (19321)	60	0.5152	6.65	9.64	13.03
Collectors	2	1913~23803 (9177)	177	0.4893	3.67	5.32	7.19
	1	399~15726 (4930)	398	0.5286	2.65	3.84	5.19

Note: 1. In one direction
2. Level of significance

Table 4.5 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 5-day Sample Size (Mid-day)

Facility Type	No. of Lanes[1]	ADT Ranges (Mean)	Sample Size *n*	C_v	Allowable Error Limit (%)		
					68%[2]	85%	95%
Freeway/ Expressway	≥5	16786~158354 (116425)	137	0.2602	2.22	3.22	4.35
	4	17516~118409 (71759)	259	0.3081	1.98	2.87	3.88
	3	8167~111878 (52868)	591	0.4082	1.68	2.44	3.29
	2	5727~79280 (25025)	573	0.4946	2.07	3.00	4.06
Divided Arterial	4	8234~49728 (25879)	159	0.2845	0.26	0.38	0.51
	3	4951~81108 (22641)	2359	0.3742	0.77	1.12	1.51
	2	799~48240 (15418)	4467	0.4075	0.61	0.88	1.20
	1	477~21717 (7816)	777	0.4390	1.57	2.28	3.08
Undivided Arterial	3	2990~23367 (13457)	20	0.6234	13.94	21.21	27.32
	2	1700~25068 (10261)	340	0.4370	2.37	3.44	4.65
	1	19~22072 (6080)	689	0.4689	1.79	2.60	3.51
One way facility	4	7292~29343 (15663)	12	0.4857	14.02	20.33	27.48
	3	3043~28700 (15541)	74	0.3884	4.51	6.54	8.84
	2	2454~28862 (11659)	118	0.5321	4.90	7.11	9.60
	1	795~8503 (4633)	8	0.4879	17.25	25.01	33.81
Toll Facility	3	24825~32571 (29315)	20	0.1212	2.91	4.22	5.70
	2	8758~32891 (17720)	60	0.3928	5.07	7.35	9.94
Collectors	2	1858~21196 (9105)	177	0.4430	3.33	4.83	6.53
	1	974~13341 (4851)	298	0.4113	2.06	2.99	4.04

Note: 1. In one direction
2. Level of significance

Table 4.6 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 5-day Sample Size (Afternoon Rush Hours)

Facility Type	No. of Lanes[1]	ADT Ranges (Mean)	Sample Size *n*	C_v	Allowable Error Limit (%)		
					68%[2]	85%	95%
Freeway/ Expressway	≥5	15805~164776 (116386)	137	0.2688	2.30	3.34	4.51
	4	16792~127710 (72424)	259	0.3303	2.05	2.97	4.02
	3	6865~114566 (53338)	591	0.4212	1.73	2.51	3.39
	2	4575~71303 (25071)	573	0.5075	2.12	3.07	4.16
Divided Arterial	4	5867~45767 (25824)	159	0.3177	2.52	3.65	4.94
	3	1574~77133 (22631)	2359	0.3898	0.80	1.16	1.57
	2	753~51333 (15472)	4467	0.4252	0.64	0.93	1.25
	1	496~20768 (7822)	777	0.4521	1.62	2.35	3.18
Undivided Arterial	3	2673~29126 (13434)	20	0.6459	14.44	20.94	28.30
	2	1612~28759 (10266)	340	0.4511	2.44	3.54	4.78
	1	918~18094 (6098)	689	0.4787	1.82	2.64	3.57
One way facility	4	7602~29343 (15663)	12	0.4527	13.07	18.95	25.63
	3	2985~27218 (15541)	74	0.3890	4.52	6.55	8.86
	2	2454~31691 (11669)	118	0.5301	4.88	7.08	9.56
	1	902~8503 (4489)	8	0.5025	17.77	25.77	34.83
Toll Facility	3	18330~45996 (32062)	20	0.2424	5.42	7.86	10.62
	2	8830~36909 (18705)	60	0.4184	5.40	7.83	10.58
Collectors	2	1716~23074 (9130)	177	0.4712	3.54	5.13	6.94
	1	1028~12871 (4922)	398	0.4494	2.25	3.26	4.41

Note: 1. In one direction
2. Level of significance

Table 4.7 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 3-day Sample Size (Morning Rush Hours)

Facility Type	No. of Lanes[1]	ADT Ranges (Mean)	Sample Size *n*	C_v	Allowable Error Limit (%)		
					68%[2]	85%	95%
Freeway/ Expressway	≥5	15707~169065 (117395)	131	0.2886	2.52	3.65	4.94
	4	15291~124951 (71256)	245	0.3702	2.37	3.44	4.65
	3	6631~118053 (53086)	541	0.4601	1.97	2.86	3.86
	2	4022~91856 (25078)	517	0.5865	2.58	3.74	5.06
Divided Arterial	4	5713~49024 (24991)	151	0.3657	2.97	4.31	5.82
	3	3728~78792 (22485)	2307	0.4564	0.95	1.38	1.86
	2	648~73831 (15382)	4267	0.4953	0.78	1.13	1.53
	1	409~22001 (7865)	757	0.4925	1.79	2.60	3.51
Undivided Arterial	3	2912~32877 (13619)	20	0.6652	14.87	21.56	29.15
	2	1917~26482 (10277)	336	0.4775	2.60	3.77	5.10
	1	658~19896 (6108)	669	0.5257	2.03	2.94	3.98
One way facility	4	6041~29343 (15663)	12	0.5553	16.03	23.24	31.42
	3	2755~27767 (15541)	74	0.3877	4.51	6.54	8.84
	2	2454~32868 (11679)	118	0.5320	4.90	7.11	9.60
	1	448~8503 (4347)	8	0.5576	19.71	28.58	38.63
Toll Facility	3	4302~45392 (32277)	12	0.4351	12.56	18.21	24.62
	2	3966~37752 (16672)	36	0.5283	8.80	12.76	17.25
Collectors	2	1913~23803 (9116)	169	0.4883	3.76	5.45	7.37
	1	520~14268 (4881)	374	0.4982	2.58	3.74	5.06

Note: 1. In one direction
2. Level of significance

Table 4.8 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 3-day Sample Size (Mid-day)

Facility Type	No. of Lanes[1]	ADT Ranges (Mean)	Sample Size n	C_v	Allowable Error Limit (%)		
					68%[2]	85%	95%
Freeway/ Expressway	≥5	16786~158354 (117100)	131	0.2631	2.30	3.34	4.51
	4	17516~118409 (71296)	245	0.3250	2.08	3.02	4.08
	3	8167~111878 (53306)	541	0.4035	1.73	2.51	3.39
	2	5727~79280 (24895)	517	0.5145	2.26	3.28	4.43
Divided Arterial	4	8234~49728 (25423)	151	0.2953	2.40	3.48	4.70
	3	4951~81108 (22517)	2307	0.3778	0.79	1.15	1.55
	2	799~47728 (15384)	4267	0.4047	0.62	0.90	1.22
	1	477~21717 (7866)	757	0.4391	1.60	2.32	3.14
Undivided Arterial	3	2990~23367 (13457)	20	0.6234	13.94	20.21	27.32
	2	1700~25068 (10301)	336	0.4362	2.38	3.45	4.66
	1	19~22072 (6120)	669	0.4682	1.81	2.62	3.55
One way facility	4	7292~29343 (15663)	12	0.4857	14.02	20.33	27.48
	3	3043~28700 (15541)	74	0.3884	4.51	6.54	8.84
	2	2454~28862 (11659)	118	0.5321	4.90	7.11	9.60
	1	795~8503 (4633)	8	0.4879	17.25	25.01	33.81
Toll Facility	3	29789~40972 (35080)	12	0.0786	2.27	3.29	4.45
	2	9021~30831 (19146)	36	0.3292	5.49	7.96	10.76
Collectors	2	1858~21196 (9098)	169	0.4472	3.44	4.99	6.74
	1	1053~13341 (4895)	374	0.4099	2.12	3.07	4.16

Note: 1. In one direction
2. Level of significance

Table 4.9 Accuracy Standards as a Function of Facility Type, Size of Facility (One-Way), and ADT Based on 3-day Sample Size (Afternoon Rush Hours)

Facility Type	No. of Lanes[1]	ADT Ranges (Mean)	Sample Size *n*	C_v	Allowable Error Limit (%)		
					68%[2]	85%	95%
Freeway/ Expressway	≥5	15805~164776 (116782)	131	0.2728	2.38	3.45	4.66
	4	16792~127710 (71624)	245	0.3353	2.14	3.10	4.19
	3	6865~113299 (53325)	541	0.4184	1.79	2.60	3.51
	2	4575~71303 (24856)	517	0.5280	2.32	3.36	4.55
Divided Arterial	4	5867~45767 (25372)	151	0.3295	2.68	3.89	5.25
	3	1574~77133 (22467)	2307	0.3927	0.82	1.19	1.61
	2	753~53762 (15387)	4267	0.4227	0.65	0.94	1.27
	1	496~20768 (7866)	757	0.4521	1.64	2.38	3.31
Undivided Arterial	3	2673~29126 (13434)	20	0.6459	14.44	20.94	28.30
	2	1612~28759 (10306)	336	0.4501	2.46	3.57	4.82
	1	918~18094 (6121)	669	0.4783	1.85	2.68	3.63
One way facility	4	7602~29343 (15663)	12	0.4527	13.07	18.95	25.62
	3	2985~27218 (15541)	74	0.3890	4.52	6.55	8.86
	2	2454~31691 (11669)	118	0.5301	4.88	7.08	9.56
	1	902~8503 (4489)	8	0.5025	17.77	25.77	34.83
Toll Facility	3	23956~46000 (34168)	12	0.1962	5.66	8.21	11.09
	2	10749~31262 (18928)	36	0.3459	5.76	8.35	11.29
Collectors	2	1716~23074 (9052)	169	0.4714	3.63	5.26	7.11
	1	1030~12597 (4891)	374	0.4275	2.21	3.20	4.33

Note: 1. In one direction
2. Level of significance

4.1.2. Accuracy Standard Due to Model Specification Error

The model misspecification errors may stem from specifications of the delay-volume function for different facility types, perception error associated with travelers in choosing routes, and so on. Here the intention is to quantify the effect of route choosing perception errors on the assignment error, since it is believed that the perception error in route choice will significantly drive the traffic flow pattern away from the user equilibrium condition. To segregate this error from other sources, deterministic simulations were conducted by varying the setting of the dispersion factors in the Stochastic User Equilibrium (SUE) assignment method while keeping all other model parameters unchanged. It is noted that, although not used in FSUTMS, the SUE assignment method is adopted to derive the assignment error for its capability to model the perception error in route choice and to handle the condition in which travelers, in general, do not have perfect information regarding travel time over the network. (Sheffi, 1985)

Deviations of the flow assignment pattern associated with range of dispersion factors from its nominal value will be recorded and analyzed. Then, depending on the information available from model calibration, the upper limit on the assignment error can be analyzed. If the variance (or standard deviation) of the dispersion factor (parameter) is known, a statistical analysis based on the normality assumption can be conducted. On the other hand, if the unit change in the dispersion factor is known, then the sensitivity analysis can be used to determine the upper limit.

<u>Based on Statistical Analysis</u>

The assignment error can be empirically fitted as

$$E = \frac{a}{1 + b\left|\ln \tilde{\theta}\right|^{c} + d \cdot \overline{V}} \tag{4.5}$$

where

$\tilde{\theta} = \theta/\theta_0$,

θ_0 = the reference dispersion factor that generates the assigned flow pattern best matches with ground counts, and

$\overline{V}$ = the average assigned link volume due to the reference dispersion factor.

The relationship between E and $\ln(\theta/\theta_0)$ is depicted as shown in Figure 4.4. Note that the assignment error is systematically higher when OD flows are lower, because that the corresponding average assigned link volume ($\overline{V}$) is also lower. Using the ordinary least-squares method, model parameters are fitted as a = 1.5772, b = 56.2777, c = -2.2381, and d = 0.00044. The corrected R^2 = 0.796.

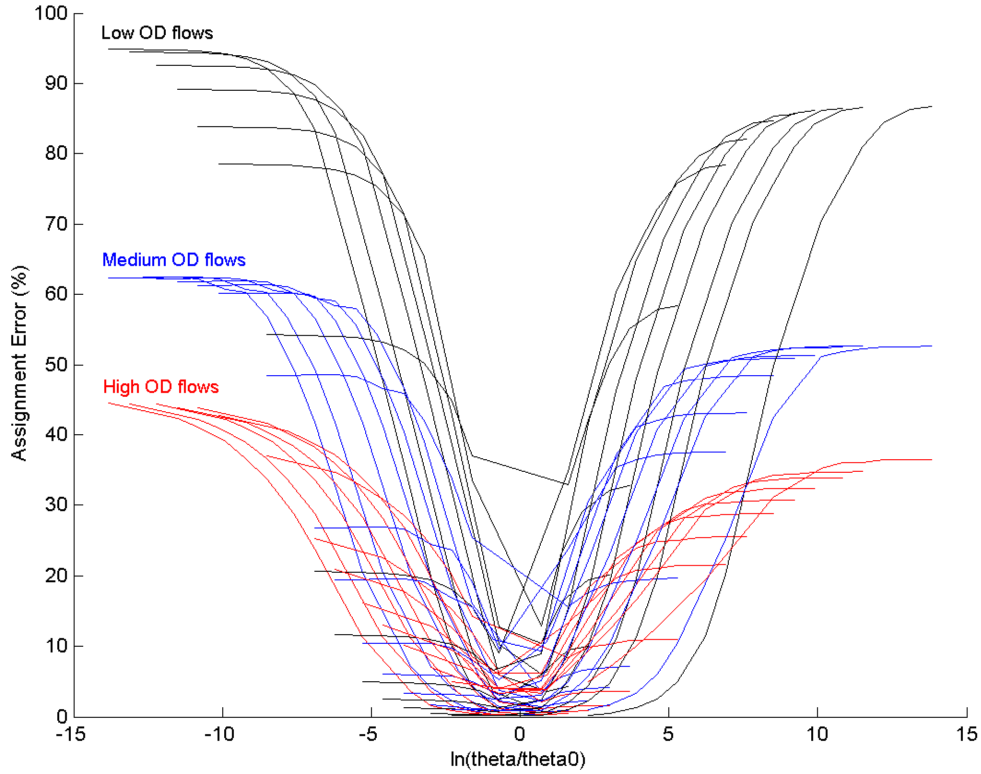

Figure 4.4. Assignment Errors as a Function of Dispersion Factors

Since θ is a strictly positive parameter, without loss of generality one could assume that ($\bar{\theta} = \theta_0$)

$$\theta \sim Ga(\alpha, \beta) = Ga\left(\frac{\sigma_\theta^2}{\theta_0}, \frac{\theta_0^2}{\sigma_\theta^2}\right)$$

The mean and variance of E, i.e., $\bar{E}$ and σ_E^2, can be evaluated using the Monte-Carlo simulation method. The upper limit on the assignment error can be determined as the upper confidence bound at (1-α) level of confidence, i.e.,

$$\bar{E} + z_\alpha \sigma_E \tag{4.6}$$

Based on the assumption of θ_0 = 0.1 and σ_θ = 0.5, the assignment errors at the 95% level of significance as a function of the average link assigned volumes are calculated. Further simulation implementations indicate that:

1. The assignment error is relatively insensitive to the dispersion shape of θ unless $\sigma_\theta \leq 0.1$

which is rarely the case in practice, and
2. The assignment error is relatively insensitive to the desired level of significance.

To account for data variability, 5% assignment error is added (see Table 4.1) assuming the error due to statistical variation of ground counts is independent of the model misspecification error. The overall standard is plotted against the current standard in Figure 4.5. As a comparison, the FHWA standards are about 10% higher than the proposed ones for AADTs being 15,000 vpd or higher, and are substantially lower for AADTs 15,000 vpd or less. The proposed accuracy standards meet the current MDOT standards quite closely, except when the average assigned volume falls in the range of 2,500 vpd and 15,000 vpd. As shown in Table 2.5a, the current practices by both FHWA and MDOT seem to underestimate the assignment error by 10-15 percent within which range compared to the proposed standards.

On the other hand, the area-wide RMSE standard (see Table 2.2) seems arbitrary (35-50%), which should be a composite measure of the actual (or estimated) distribution of the roadway AADT groups and the corresponding allowable error in each volume group. In general, if the study area is rural, the allowable error should be higher since the majority of the roadway AADTs are low. In the urban area, however, the allowable error should be lower since the majority of the roadways have higher AADTs. It is extremely difficult to assign a reasonable accuracy standard if the information on distribution of the roadway AADTs in the study area is not available. Deriving accuracy standards by a coarse classification of area types, such as the one used by Reno (see Table 2.6), will not be reasonable if the actual distribution of roadway AADTs are significantly different from those designated area types. Therefore, a worksheet as shown in Table 2.5b is designed for planners in different jurisdictional areas to develop their own standards. An example is provided as follows for illustration purpose. Assume that 8% of roadways have AADT between 1,000 and 2,500 vpd, 25% of them fall in the range from 2,500 to 5,000 vpd, and so on. (see column 2 in Table 2.5b) The sum of the column 4 which is the product of mid AADT in each category and the corresponding distribution will give the mean area-wide AADT. Column (5) gives the allowable error weighted by column (4), whose sum will yield the weighted average allowable error by the marginal distribution of AADT. Finally, the ratio of the sum of column (5) to the sum of column (4) gives the area-wide allowable error.

Finally, this study recommends eliminating the facility-type specific accuracy standards, such as Tables 2.3 and 2.6, since they have created confusions with and, sometimes, contradictions to the standards by volume-group. Current practices show that different standards were applied onto different roadway facilities even though they have carried similar traffic volumes. Elimination of this extra category will greatly simplify and standardize the structure of the current accuracy standards on highway networks.

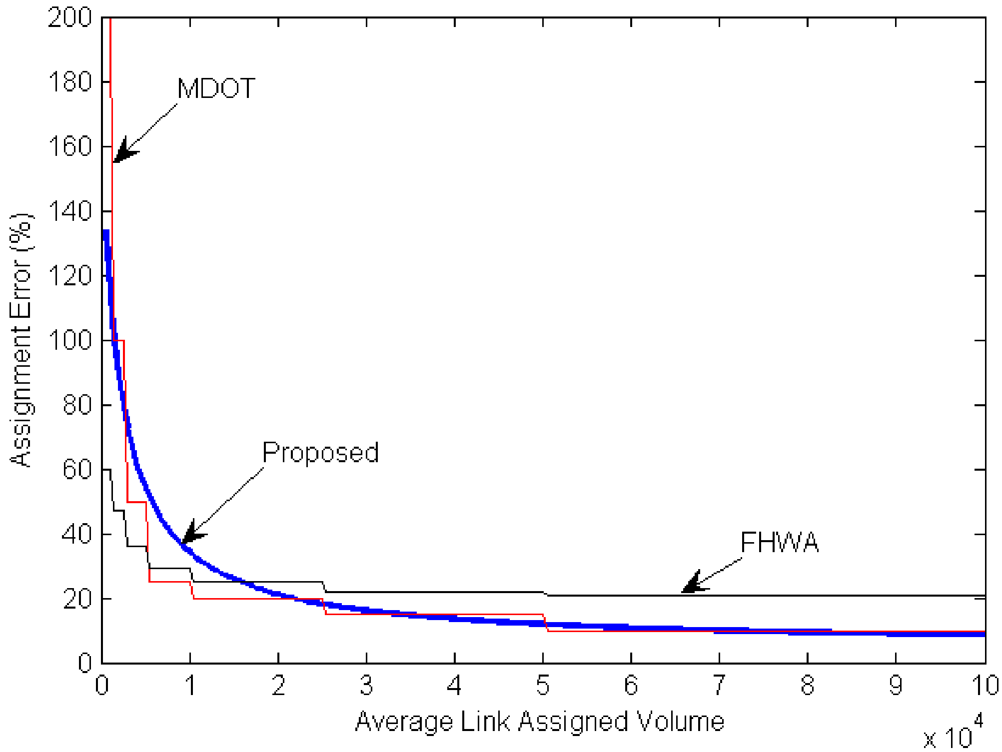

Figure 4.5 Assignment Errors versus Average Assigned Volumes

Table 2.5a Proposed Accuracy Standards on Volume by AADT

AADT	Allowable Error Limits (%)		
	FHWA	Michigan DOT	Proposed
< 1,000	60	200	150
1,000 – 2,500	47	100	100
2,500 – 5,000	36	50	65
5,000 – 10,000	29	25	45
10,000 – 15,000	25	20	35
15,000 – 25,000			25
25,000 – 50,000	22	15	15
> 50,000	21	10	10

Table 2.5b Proposed Accuracy Standards Worksheet on the Area-wide Allowable Error

AADT	Mean AADT (vpd) (1)	Distribution of Roadway AADT (2)	Proposed Allowable Error (3)	Col. (1) × Col. (2) (4)	Col. (3) × Col. (4) (5)
< 1,000	500	0	1.5	0	0
1,000 – 2,500	1750	0.08	1.0	140	140
2,500 – 5,000	3750	0.25	0.65	937.5	609.4
5,000 – 10,000	7500	0.3	0.45	2250	1012.5
10,000 – 15,000	12500	0.2	0.35	2500	875
15,000 – 25,000	20000	0.17	0.25	3400	850
25,000 – 50,000	37500	0	0.15	0	0
> 50,000	75000	0	0.1	0	0
Sum	-	1.0	-	9227.5 (6)	3487 (7)
Area-wide Allowable Error = (7) ÷ (6) = 3487 ÷ 9227.5 = 0.38 (38%)					

Table 2.2a Proposed Modified Accuracy Standards for Traffic Assignment Used in FSUTMS

Control Statistic	**Original Accuracy Standard**	**Modified Accuracy Standard**
Assigned VMT/Count VMT (Area)	5%	Same
Assigned VHT/Count VHT (Area)	5%	Same
Volume-count ratio (screenline, cordonline, cutline)	10% (AADT ≥ 50,000 vpd) 20% (AADT < 50,000 vpd)	Same
Assigned VMT/Count VMT (Facility type, area type, no. of lane combination)	15% (≥ 100,000) 25% (< 100,000)	Same
Assigned VHT/Count VHT (Facility type, area type, no. of lane combination)	15% (≥ 20,000) 25% (< 20,000)	Same
Root Mean Square Error (%) (Area)	35% - 50%	See Table 2.5b
Root Mean Square Error (%) (Link volume group)	25% ≥ (AADT ≥ 50,000 vpd) 30-100% (AADT < 50,000 vpd) >100% (AADT < 3,000 vpd)	See Table 2.5a

Based on Sensitivity Analysis

By definition, sensitivity of the dispersion factor on the assignment error can be expressed as

$$S = \frac{\partial E/E}{\partial \tilde{\theta}/\tilde{\theta}} = \frac{\partial E}{\partial \tilde{\theta}}\frac{\tilde{\theta}}{E} \tag{4.7b}$$

where

$$\begin{aligned}\frac{\partial E}{\partial \tilde{\theta}} &= \frac{-abc\left[\ln \tilde{\theta}\right]^{c-1}\left(1/\tilde{\theta}\right)}{\left(1+b\left[\ln \tilde{\theta}\right]^{c} + d\cdot \overline{V}\left(\theta_0\right)\right)^2}, \text{ if } \tilde{\theta} \geq 1 \\ &= \frac{abc\left[\ln\left(1/\tilde{\theta}\right)\right]^{c-1}\left(1/\tilde{\theta}\right)}{\left(1+b\left[\ln\left(1/\tilde{\theta}\right)\right]^{c} + d\cdot \overline{V}\left(\theta_0\right)\right)^2}, \text{ if } \tilde{\theta} \leq 1\end{aligned} \tag{4.8}$$

Therefore,

$$\begin{aligned}S &= \frac{-bc\left[\ln \tilde{\theta}\right]^{c-1}}{1+b\left[\ln \tilde{\theta}\right]^{c} + d\cdot \overline{V}\left(\theta_0\right)}, \text{ if } \tilde{\theta} \geq 1 \\ &= \frac{bc\left[\ln\left(1/\tilde{\theta}\right)\right]^{c-1}}{1+b\left[\ln\left(1/\tilde{\theta}\right)\right]^{c} + d\cdot \overline{V}\left(\theta_0\right)}, \text{ if } \tilde{\theta} \leq 1\end{aligned} \tag{4.7b}$$

The sensitivity can be interpreted as the unit change in the assignment error as a function of the unit change in the dispersion factor. In practice, if the average unit change in the dispersion factor (denoted as $\Delta\theta$) is known from model calibration process, the average unit change in the assignment error (denoted as ΔE) can be estimated as $|S|\Delta\theta$. It follows that the upper limit on the assignment error can be determined as:

$$E + \Delta E = E + |S|\Delta\theta \tag{4.9}$$

4.2. Required Sample Size for Assigned Volume

The primary purposes of calculating required sample size are to determine how many counts need to be sampled for each type of facility to achieve a specific accuracy standard at a certain level of confidence. The required sample size can be calculated as

$$n = \left(\frac{ZC_v}{d} \right)^2 \tag{4.10}$$

In the following exercise, the required sample sizes are calculated for 68%, 85%, and 95% level of significance and the results are listed in Tables 4.10 through 4.12. One can find that, to achieve the current standards, the required sample size is considered economically affordable. As a comparison, samples provided by the current infrastructure system are sufficiently large to support the needs under the current accuracy standards except for 4-lane one-way facility. Upon obtaining this information, simple sampling method can be used to select detector locations from stratified volume levels in each type of facility.

Table 4.10 Required Sample Size Based on Current Accuracy Standards (Morning Rush Hours)

Facility Type	No. of Lanes	ADT Ranges (Mean)	Current Accuracy Standards	C_v	Required Sample Size		
					68%	85%	95%
Freeway	≥10	15707~169065(117190)	-	0.2875	-	-	-
	8	15291~126375 (72080)	13%	0.3749	8.3	17.5	32.0
	6	6074~125117 (52905)	28%	0.4753	2.9	6.1	11.1
	4	4022~91856 (25435)	29%	0.5713	3.9	8.2	14.9
Divided Arterial	8	5216~49024 (25516)	13%	0.3684	8.0	16.9	30.6
	6	4179~78792 (22775)	17%	0.4600	7.3	15.4	28.1
	4	648~74815 (15315)	25%	0.5153	4.3	8.9	16.3
	2	409~22001 (7830)	-	0.4888	-	-	-
Undivided Arterial	6	2912~32877 (13619)	-	0.6652	-	-	-
	4	1599~26482 (10189)	34%	0.4806	2.0	4.2	7.6
	2	658~19896 (6078)	56%	0.5245	0.9	1.8	3.4
One way facility	8	6041~29343 (15663)	-	0.5553	-	-	-
	6	2755~27767 (15541)	-	0.3877	-	-	-
	4	2454~32868 (11679)	13%	0.5320	16.8	35.2	64.4
	2	448~8503 (4347)	25%	0.5576	5.0	10.5	19.1
Toll Facility	6	5658~45392 (31937)	-	0.4194	-	-	-
	4	3803~39529 (19296)	-	0.5162	-	-	-
Collectors	4	1913~23803 (9169)	-	0.4928	-	-	-
	2	393~16200 (4876)	-	0.5502	-	-	-

Table 4.11 Required Sample Size Based on Current Accuracy Standards (Mid-day)

Facility Type	**No. of Lanes**	**ADT Ranges (Mean)**	**Current Accuracy Standards**	**C_v**	**Required Sample Size**		
					68%	**85%**	**95%**
Freeway	≥10	16786~158354(116730)	-	0.2538	-	-	-
	8	17516~ 118409 (72266)	13%	0.3135	5.8	12.2	22.4
	6	8167~ 111878 (52817)	28%	0.4018	2.1	4.3	7.9
	4	5727~79280 (25377)	29%	0.4848	2.8	5.9	10.7
Divided Arterial	8	8234~49728 (25707)	13%	0.2766	4.5	9.5	17.4
	6	4951~81108 (22762)	17%	0.3733	4.8	10.1	18.5
	4	800~48365 (15220)	25%	0.4177	2.8	5.9	10.7
	2	477~21717 (7819)	-	0.4333	-	-	-
Undivided Arterial	6	2990~23367 (13457)	-	0.6234	-	-	-
	4	1700~25068 (10210)	34%	0.4393	1.7	3.5	6.4
	2	19~22072 (6072)	56%	0.4648	0.7	1.5	2.7
One way facility	8	7292~29343 (15663)	-	0.4857	-	-	-
	6	3043~28700 (15541)	-	0.3884	-	-	-
	4	2454~28862 (11659)	13%	0.5321	16.8	35.2	64.4
	2	795~8503 (4633)	25%	0.4879	3.8	8.0	14.6
Toll Facility	6	23735~39410 (30506)	-	0.1504	-	-	-
	4	8863~34158 (18180)	-	0.3993	-	-	-
Collectors	4	1858~21196 (9122)	-	0.4363	-	-	-
	2	1093~13341 (4876)	-	0.4151	-	-	-

Table 4.12 Required Sample Size Based on Current Accuracy Standards (Afternoon Rush Hours)

Facility Type	No. of Lanes	ADT Ranges (Mean)	Current Accuracy Standards	C_v	Required Sample Size		
					68%	85%	95%
Freeway	≥10	15805~164776(116070)	-	0.2670	-	-	-
	8	16792~127710 (72950)	13%	0.3290	6.4	13.5	24.6
	6	6865~113299 (53142)	28%	0.4210	2.3	4.8	8.7
	4	4575~71303 (25343)	29%	0.4970	2.9	6.2	11.3
Divided Arterial	8	5867~45767 (25709)	13%	0.3140	5.8	12.3	22.4
	6	1575~77133 (22747)	17%	0.3893	5.2	11.0	20.2
	4	752~53912 (15245)	25%	0.4333	3.0	6.3	11.5
	2	496~20768 (7822)	-	0.4464	-	-	-
Undivided Arterial	6	2673~29126 (13434)	-	0.6459	-	-	-
	4	1612~28759 (10208)	34%	0.4546	1.8	3.8	6.9
	2	918~18094 (6066)	56%	0.4758	0.7	1.5	2.8
One way facility	8	7602~29343 (15663)	-	0.4527	-	-	-
	6	2985~27218 (15541)	-	0.3890	-	-	-
	4	2454~31691 (11669)	13%	0.5301	16.6	35.0	63.9
	2	902~8503 (4489)	25%	0.5025	4.0	8.5	15.5
Toll Facility	6	19202~51209 (31961)	-	0.2689	-	-	-
	4	9491~36994 (18647)	-	0.4105	-	-	-
Collectors	4	1716~23074 (9123)	-	0.4583	-	-	-
	2	1030~13599 (4918)	-	0.4555	-	-	-

4.3. Accuracy Standards for Vehicle Mile Traveled

To establish accuracy standards for the vehicle miles traveled (VMT), all the available count samples along with their link coverage lengths need to be utilized. Unfortunately, the link coverage lengths are not available in the current count database and need to be manually coded. To that end, the GIS-based Florida Transportation Information system (FDOT, 2002) was used to facilitate the mileage finding. For counters on freeway sections, the mileage associated with a counter starts and/or ends at interchanges since traffic might merge and/or diverge from mainlines to some extends. If there exist two or more counters between two interchanges, the mileage will be evenly divided. On surface streets, the mileage associated with a counter starts and/or ends at intersections, since traffic on the subject link could make turns at downstream intersection and traffic from upstream intersections could turn into the subject link. Both activities distort the traffic flow patterns among the subject link and other adjacent links. The mileage will also be evenly divided if there exist more than one counter on the subject link.

Upon the completion of data preparation, the accuracy standard can be calculated as follows:

$$d = \frac{Z}{\bar{x}}\sqrt{\frac{\left(\sum W_h s_h\right)^2 - \frac{n^*}{N^*}\left(\sum W_h s_h^2\right)}{n^*}} \tag{4.11}$$

where n^* = total number of miles in all the available samples;

Z = normal variate = 1.0 for 68% level of significance; 1.45 for 85% level of significance and 1.96 for 95% level of significance;

s_h = Composite standard deviation of stratum h ($=\sqrt{s_1^2 + s_2^2}$);

s_1^2 = Spatial variance;

s_2^2 = Temporal variance;

d = Accuracy standard in traffic assignment as a function of facility type, size of facility, and ADT (desired relative error);

$\bar{x}$ = Mean volume (mean VMT per mile); and

N^* = Total number of miles area-wide.

Since $n^* << N^*$, the population correction term, i.e., $\frac{n^*}{N^*}\left(\sum W_h s_h^2\right)$, can be omitted. Therefore, the formula becomes:

$$d = \frac{\left(\sum W_h s_h\right)Z}{\bar{x}\sqrt{n^*}} \tag{4.12}$$

The accuracy standards for specific facility types are listed in Table 4.13. As shown, one can see that all facilities have tighter standards than the current counterpart except one-way facility due to the limited sample size.

Table 4.13 Accuracy Standards for Facility Type at 68%, 85% and 95% Confidence Level

Facility Type	Accuracy Standards		
	68%	85%	96%
Freeway/Expressway	1.1%	1.5%	2.1%
Divided Arterial	0.5%	0.8%	1.1%
Undivided Arterial	1.8%	2.6%	3.5%
One-way facility	6.1%	8.8%	12%
Collectors	2.4%	3.5%	4.8%

4.4. Required Sample Size for Vehicle Mile Traveled

Reliable estimates of urban vehicle miles of travel (VMT) are important to assess the effectiveness of safety programs, to develop financing and improvement procedures, and to evaluate urban traffic models. The most desirable method for determining urban VMT, weekday VMT particularly, would be to make representative traffic counts on every section of urban street. This, of course, would result in extremely expensive counting programs and would be practical only in cities that make extensive counts for traffic operation purposes. Accordingly, random sampling procedures were developed to provide a cost-effective means for estimating VMT on urban streets and highways.

The sampling procedures show how the required sample size can be determined. After the sample data are collected, the variability (error) in the sample should be determined. This will provide a basis for modifying the sample survey design in the forthcoming years. The sampling procedures are designed to yield estimates of weekday urban VMT within a specified precision at the 68, 85 or 95 percent level of confidence. They may also be used to estimate average annual VMT. The sampling procedures are based on sampling 24-hour, 48-hour or 5-day link volume counts. A sampling program for estimating urban weekday VMT includes the following steps.

1. Establish Geographic Sub-Areas

The urban area should be subdivided into the analysis units for which specific information is desired. A small urban area (population under 250,000) would be probably be subdivided into 2 or 3 divisions which distinguish between central city and suburbs. Intermediate size urban areas (population 250,000 to 1,000,000) may be subdivided into 3 to 5 sub-areas. Larger urban areas (population over 1,000,000) could be subdivided into 5 to 7 sub-areas.

2. Classify Facility Types

Stratification by type of facility is essential since each basic group represents a distinctive population. The facility types employed herein are classified into six categories, which are freeway and expressway, divided arterial, undivided arterial, collectors, one-way facility, toll facility.

3. Further Stratify Locations in Each Class to the Extent That Prior Information is Available.

This stratification should be done on the basis of the best information available. Some suggestions for different facility types are:

(a) Local streets have a volume group of 0 to 2,000 vehicles per day,
(b) Arterial streets are stratified by 5,000 volume strata, and
(c) Freeways and expressways are stratified by number of lanes.

4. Identify Links and Sampling Periods

The sampling-period, and links within each of the sub-populations (and their respective stratifications) should be identified. Links should represent sections of roads with homogeneous volume. Links within each stratum should be of relative uniform length. Where links vary significantly in length, they should be further subdivided into road sections of approximately equal lengths. A maximum 10 to 20 percent variation in link length represents a desirable objective. Suggested link-lengths for different facility type are as follows:

(a) Local streets:	0.25 to 0.5 mile
(b) Arterial streets:	0.5 to 1.0 mile
(c) Freeways:	1.0 mile

The traffic assignment network is overlapped on a street map to obtain local links. Links should be clearly coded according to: (1) geographic sector, (2) facility type, (3) volume strata, (4) specific location, and (5) length.

5. Establish Desired Precision Levels

The desired errors in the VMT estimates and the desired levels of confidence should be specified. The confidence level herein are established to be 68%, 85% and 95%, and the following relative error are employed in each functional class: 5%, 15% and 25%. These values are used in sample size calculations.

6. Obtain Estimate of Spatial and Temporal Variations

A community should use its own estimates for spatial and temporal variances where these estimates are available. In the absence of previous traffic counts from which to derive variation estimates, the variations tabulated on pages 28 and 29 of GUTVC can be employed.

7. Estimate the Theoretical Sample Size (the number of links to be sampled)

After the sub-populations have been defined, links identified, precision established and basic statistical parameters estimated, sample size should be determined for each sub-population.

If there is no prior information available regarding the distribution of links by volume in each road class, it is necessary to use simple random sampling and to assume broad coefficients of variation for each road class. The required number of links can be computed from the following formula:

$$n = \frac{Z^2(s_1^2 + s_2^2)}{(d\overline{x})^2 + \frac{s_1^2}{N}Z^2} \qquad (4.13)$$

However, if n<<N, then the finite population correction factor $\frac{s_1^2}{N}Z^2$ can be omitted. The formula becomes:

$$n = \frac{Z^2(s_1^2 + s_2^2)}{(d\overline{x})^2} \qquad (4.14)$$

where n = Sample number of miles to count in a functional class, (= mean vehicle miles per mile);

Z = Normal variate= 1.0 for 68% confidence internal, 1.45 for 85 confidence internal and 1.96 for 95% confidence interval;

s_1^2 = Spatial variance;

s_2^2 = Temporal variance;

d = Accuracy standard in traffic assignment as a function of facility type, facility size and ADT (desired relative error);

$\overline{x}$ = Mean volume (mean VMT per mile); and

N = Number of miles in a functional class.

If the information regarding the distribution of links by volume in each road class is available and the link volume can be further grouped into volume strata, stratified random sampling should be used for each road class. The narrower is the stratum, the lower is its variance and hence the sample size requirements. The required number of links can be computed from the following formula:

$$n = \frac{\left(\sum W_h s_h\right)^2}{\frac{(d\bar{x})^2}{Z^2} + \frac{\left(\sum W_h s_h^2\right)}{N}} \tag{4.15}$$

For $n << N$ the finite population correction factor $\frac{\sum W_h s_h^2}{N}$ can be omitted. Therefore the formula becomes:

$$n = \frac{Z^2 (\sum W_h s_h)^2}{(d\bar{x})^2} \tag{4.16}$$

After the sample size, n, is determined, the number of miles to count in stratum h can be computed according to the following formula:

$$n_h = n \frac{W_h s_h}{\sum W_h s_h} \tag{4.17}$$

The number of links to count in any stratum h can be computed as follows:

$$n_{lh} \cong \frac{n_h}{l_h}$$

This number should be rounded to the next highest integer. Additional variables are denoted as follows:

s_h = Composite standard deviation of stratum $h = \sqrt{s_1^2 + s_2^2}$,
N_h = Number of miles of in stratum h in a functional class,
W_h = Weight of stratum $h = \frac{N_h}{N}$,
$\bar{x}$ = Mean volume = $\sum W_h \bar{x}_h$,
$\bar{x}_h$ = Midpoint of volume range in stratum h,
l_h = Average length of links in stratum h, and
n_{lh} = Number of links to count in stratum h.

The use of simple random sampling and stratified random sampling techniques within each of the different facility types depends on: (1) the facility typed involved, and (2) the degree of information available regarding the traffic volumes within each roadway class.

Table 4.14a Sample Size Computation for Freeway and Expressway at Different Confidence Levels (Relative Error = 0.05, Number of Mileages)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
<=4	24759	614.68	0.478	12007	18.1	36.1	61.1
6	54326	426.2	0.332	20846	21.7	43.4	73.6
8	69459	161.34	0.126	22013	8.7	17.4	29.4
>=10	115170	83.04	0.064	31837	6.4	12.9	21.9
Total		1285.26	1.00	---	54.9	109.8	186.0
Required number of counts					31	62	105

Table 4.14b Sample Size Computation for Freeway and Expressway at Different Confidence Levels (Relative Error = 0.15, Number of Mileages)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
<=4	24759	614.68	0.478	12007	2.1	4.4	7.9
6	54326	426.2	0.332	20846	2.5	5.3	9.5
8	69459	161.34	0.126	22013	1.0	2.1	3.8
>=10	115170	83.04	0.064	31837	0.8	1.5	2.8
Total		1285.26	1.00	---	6.4	13.3	24.1
Required number of counts					4	8	14

Table 4.14c Sample Size Computation for Freeway and Expressway at Different Confidence Levels (Relative Error = 0.25, Number of Mileages)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
<=4	24759	614.68	0.478	12007	0.7	1.6	2.9
6	54326	426.2	0.332	20846	0.9	1.9	3.5
8	69459	161.34	0.126	22013	0.4	0.7	1.4
>=10	115170	83.04	0.064	31837	0.3	0.6	1.0
Total		1285.26	1.00	---	2.3	4.8	8.8
Required number of counts					2	3	5

Table 4.15a Sample Size Computation for Divided Arterial at Different Confidence Levels (Relative Error = 0.05, Number of Mileages)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	7853	562.56	0.115	3398	3.5	7.3	13.2
4	15150	3046.2	0.620	5908	33.3	69.1	123.8
6	22502	1236.8	0.252	7855	18.0	37.3	66.8
>=8	26814	63.46	0.013	6625	0.8	1.6	2.9
Total		4909.02	1.00	---	55.6	115.3	206.7
Required number of counts					45	92.5	166

Table 4.15b Sample Size Computation for Divided Arterial at Different Confidence Levels (Relative Error = 0.15, Number of Mileages)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	7853	562.56	0.115	3398	0.4	0.8	1.5
4	15150	3046.2	0.620	5908	3.7	7.9	14.3
6	22502	1236.8	0.252	7855	2.0	4.2	7.7
>=8	26814	63.46	0.013	6625	0.1	0.2	0.3
Total		4909.02	1.00	---	6.2	13.1	23.8
Required number of counts					5	11	19

Table 4.15c Sample Size Computation for Divided Arterial at Different Confidence Levels (Relative Error = 0.25, Number of Mileages)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	7853	562.56	0.115	3398	0.1	0.3	0.5
4	15150	3046.2	0.620	5908	1.3	2.8	5.2
6	22502	1236.8	0.252	7855	0.7	1.5	2.8
>=8	26814	63.46	0.013	6625	0.0	0.1	0.1
Total		4909.02	1.00	---	2.1	4.7	8.6
Required number of counts					2	4	7

Table 4.16a Sample Size Computation for Undivided Arterial at Different Confidence Levels (Relative Error = 0.05, Number of Mileages)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	6088	438.44	0.747	2648	41.9	77.5	118.8
4	10102	134.4	0.229	4376	21.2	39.3	60.2
>=6	15055	13.88	0.024	7475	3.8	6.9	10.6
Total		586.7	1.00	---	66.9	123.7	189.6
Required number of counts					64	117	179

Table 4.16b Sample Size Computation for Undivided Arterial at Different Confidence Levels (Relative Error = 0.15, Number of Mileages)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	6088	438.44	0.747	2648	5.2	10.8	19.3
4	10102	134.4	0.229	4376	2.7	5.5	9.8
>=6	15055	13.88	0.024	7475	0.5	1.0	1.7
Total		586.7	1.00	---	8.4	17.3	30.8
Required number of counts					8	17	29

Table 4.16c Sample Size Computation for Undivided Arterial at Different Confidence Levels (Relative Error = 0.25, Number of Mileages)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	6088	438.44	0.747	2648	1.9	4.0	7.2
4	10102	134.4	0.229	4376	1.0	2.0	3.7
>=6	15055	13.88	0.024	7475	0.2	0.4	0.6
Total		586.7	1.00	---	3.1	6.4	11.5
Required number of counts					3	6	11

Table 4.17a Sample Size Computation for One-Way Facility at Different Confidence Levels (Relative Error = 0.05, Number of Mileages)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	5368	2.61	0.042	2309	0.6	0.7	0.8
4	11321	34.73	0.552	6001	19.6	24.7	27.7
6	14831	18.57	0.295	6351	11.1	13.9	15.6
8	18858	7.00	0.111	8575	5.6	7.1	7.9
Total		62.91	1.00	---	36.8	46.4	52.0
Required number of counts					65	82	92

Table 4.17b Sample Size Computation for One-Way Facility at Different Confidence Levels (Relative Error = 0.15, Number of Mileages)

No of Lanes	Mean Volume	N_h	W_h	s_h	*N(68%)*	*N(85%)*	*n(95%)*
2	5368	2.61	0.042	2309	0.1	0.2	0.4
4	11321	34.73	0.552	6001	4.7	8.5	12.7
6	14831	18.57	0.295	6351	2.7	4.8	7.2
8	18858	7.00	0.111	8575	1.3	2.4	3.7
Total		62.91	1.00	---	8.8	15.9	24.0
Required number of counts					16	28	43

Table 4.17c Sample Size Computation for One-Way Facility at Different Confidence Levels (Relative Error = 0.25, Number of Mileages)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	5368	2.61	0.042	2309	0.1	0.1	0.2
4	11321	34.73	0.552	6001	1.8	3.6	6.1
6	14831	18.57	0.295	6351	1.1	2.1	3.5
8	18858	7.00	0.111	8575	0.5	1.1	1.8
Total		62.91	1.00	---	3.5	6.9	11.6
Required number of counts					7	13	21

Table 4.18a Sample Size Computation for Collectors at Different Confidence Levels (Relative Error = 0.05, Number of Mileages)

No of Lanes	Mean Volume	N_h	W_h	s_h	*N(68%)*	*n(85%)*	*n(95%)*
2	5027	189.86	0.718	1861	26.9	45.6	64.1
4	9200	71.08	0.269	4023	21.8	37.0	51.9
6	8062	3.36	0.013	2840	0.7	1.2	1.7
Total		264.3	1.00	---	49.4	83.8	117.7
Required number of counts					51	87	122

Table 4.18b Sample Size Computation for Collectors at Different Confidence Levels (Relative Error = 0.15, Number of Mileages)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	5027	189.86	0.718	1861	3.7	7.5	13.1
4	9200	71.08	0.269	4023	3.0	6.1	10.6
6	8062	3.36	0.013	2840	0.1	0.2	0.4
Total		264.3	1.00	---	6.8	13.8	24.1
Required number of counts					7	15	25

Table 4.18c Sample Size Computation for Collectors at Different Confidence Levels (Relative Error = 0.25, Number of Mileages)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	5027	189.86	0.718	1861	1.4	2.8	5.1
4	9200	71.08	0.269	4023	1.1	2.3	4.1
6	8062	3.36	0.013	2840	0.0	0.1	0.1
Total		264.3	1.00	---	2.5	5.2	9.3
Required number of counts					3	6	10

8. Estimate the Practical Sample Size

The resultant sample sizes tabulated above are computed from formulas for each functional class area-wide. They are the minimum theoretical sample size requirement. In practice, sample size should include at least six miles per strata and 30 per facility type on an area-wide basis (GUTVC, 1975). To achieve these requirements, the sample size for each facility type should be adjusted accordingly. An attention should be paid is that, in practice, every freeway mile would be counted for random sampling and every other mile for stratified sampling. Using these criteria, the minimum practical sample size requirements are adjusted and listed in the following tables.

Table 4.19a Practical Sample Size (Number of Counts) for Freeway and Expressway (Relative Error = 0.05)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
<=4	24759	614.68	0.478	12007	11	21	35
6	54326	426.2	0.332	20846	13	25	42
8	69459	161.34	0.126	22013	6	12	28
>=10	115170	83.04	0.064	31837	6	12	21
Total		1285.26	1.00	---	36	70	126

Table 4.19b Practical Sample Size (Number of Counts) for Freeway and Expressway (Relative Error = 0.15)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
<=4	24759	614.68	0.478	12007	8	17	30
6	54326	426.2	0.332	20846	10	22	38
8	69459	161.34	0.126	22013	6	12	23
>=10	115170	83.04	0.064	31837	6	12	21
Total		1285.26	1.00	---	30	63	112

Table 4.19c Practical Sample Size (Number of Counts) for Freeway and Expressway (Relative Error = 0.25)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
<=4	24759	614.68	0.478	12007	8	17	30
6	54326	426.2	0.332	20846	10	22	38
8	69459	161.34	0.126	22013	6	11	21
>=10	115170	83.04	0.064	31837	6	12	20
Total		1285.26	1.00	---	30	62	109

Table 4.20a Practical Sample Size (Number of Counts) for Divided Arterial (Relative Error = 0.05)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*N(95%)*
2	7853	562.56	0.115	3398	6	13	23
4	15150	3046.2	0.620	5908	27	56	100
6	22502	1236.8	0.252	7855	15	30	54
>=8	26814	63.46	0.013	6625	6	12	22
Total		4909.02	1.00	---	54	111	199

Table 4.20b Practical Sample Size (Number of Counts) for Divided Arterial (Relative Error = 0.15)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	7853	562.56	0.115	3398	6	12	23
4	15150	3046.2	0.620	5908	12	26	47
6	22502	1236.8	0.252	7855	6	13	24
>=8	26814	63.46	0.013	6625	6	12	18
Total		4909.02	1.00	---	30	63	112

Table 4.20c Practical Sample Size (Number of Counts) for Divided Arterial (Relative Error = 0.25)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	7853	562.56	0.115	3398	6	12	23
4	15150	3046.2	0.620	5908	12	26	47
6	22502	1236.8	0.252	7855	6	13	24
>=8	26814	63.46	0.013	6625	6	12	18
Total		4909.02	1.00	---	30	63	112

Table 4.21a Practical Sample Size (Number of Counts) for Undivided Arterial (Relative Error = 0.05)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	6088	438.44	0.747	2648	40	74	113
4	10102	134.4	0.229	4376	20	38	57
>=6	15055	13.88	0.024	7475	6	12	17
Total		586.7	1.00	---	66	124	187

Table 4.21b Practical Sample Size (Number of Counts) for Undivided Arterial (Relative Error = 0.15)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	6088	438.44	0.747	2648	15	32	56
4	10102	134.4	0.229	4376	9	19	33
>=6	15055	13.88	0.024	7475	6	12	17
Total		586.7	1.00	---	30	63	106

Table 4.21c Practical Sample Size (Number of Counts) for Undivided Arterial (Relative Error = 0.25)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	6088	438.44	0.747	2648	15	32	56
4	10102	134.4	0.229	4376	9	18	33
>=6	15055	13.88	0.024	7475	6	12	17
Total		586.7	1.00	---	30	62	106

Table 4.22a Practical Sample Size (Number of Counts) for One-Way Facility (Relative Error = 0.05)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	5368	2.61	0.042	2309	6	7	8
4	11321	34.73	0.552	6001	35	44	49
6	14831	18.57	0.295	6351	20	25	28
8	18858	7.00	0.111	8575	10	13	14
Total		62.91	1.00	---	71	89	99

Table 4.22b Practical Sample Size (Number of Counts) for One-Way Facility (Relative Error = 0.15)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	5368	2.61	0.042	2309	6	7	8
4	11321	34.73	0.552	6001	12	22	33
6	14831	18.57	0.295	6351	6	11	17
8	18858	7.00	0.111	8575	6	11	14
Total		62.91	1.00	---	30	51	72

Table 4.22c Practical Sample Size (Number of Counts) for One-Way Facility (Relative Error = 0.25)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	5368	2.61	0.042	2309	6	7	8
4	11321	34.73	0.552	6001	12	22	33
6	14831	18.57	0.295	6351	6	11	17
8	18858	7.00	0.111	8575	6	11	10
Total		62.91	1.00	---	30	51	68

Table 4.23a Practical Sample Size (Number of Counts) for Collectors (Relative Error = 0.05)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	5027	189.86	0.718	1861	28	47	66
4	9200	71.08	0.269	4023	23	38	54
6	8062	3.36	0.013	2840	6	11	15
Total		264.3	1.00	---	57	96	135

Table 4.23b Practical Sample Size (Number of Counts) for Collectors (Relative Error = 0.15)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	5027	189.86	0.718	1861	13	27	48
4	9200	71.08	0.269	4023	11	23	40
6	8062	3.36	0.013	2840	6	11	15
Total		264.3	1.00	---	30	61	103

Table 4.23c Practical Sample Size (Number of Counts) for Collectors (Relative Error = 0.25)

No of Lanes	Mean Volume	N_h	W_h	s_h	*n(68%)*	*n(85%)*	*n(95%)*
2	5027	189.86	0.718	1861	13	22	48
4	9200	71.08	0.269	4023	11	23	40
6	8062	3.36	0.013	2840	6	6	6
Total		264.3	1.00	---	30	51	94

CHAPTER FIVE

ACCURACY STANDARDS FOR TRANSIT NETWORKS

5.1 Introduction

Traditionally, travel demands are forecasted at the zonal level based on aggregate interzonal flows that are calibrated from zonal average trip attributes and socioeconomic characteristics. However, interzonal travel demands are the result of aggregating individual travel decisions and traditional aggregate demand models do not make the connection between the two (McFadden and Reid, 1978). Consequently, disaggregate models have been developed to estimate aggregate forecasts. These models offer substantial advantages over the traditional models while remain sufficiently practical for applications. For example, the choice models derived for mode choice in the four-step travel demand modeling process were generally calibrated using the data collected from sampled individuals at the disaggregate level. The mode an individual selected for a given trip is the model output. However, predictions for person trips at zonal levels are the true aggregate level forecasts of interest to transportation professionals. The core of the problem of aggregating across individuals lies in predicting the proportion of the population choosing each alternative mode. The linkage between an individual trip maker and the number of trips one makes for a certain purpose is also an obstacle to better estimating mode share.

The common practice in developing a mode choice model in Florida has been to transfer model coefficients from other cities (Abdel-Aty and Abdelwahab, 2001). The model is then implemented in the following manner: (1) adjusting the modal bias coefficients (i.e., constants of the utility equation) to replicate the transit ridership data, and (2) examining the validation results to identify any additional adjustments to coefficients or other parameters which were appropriate. An iterative process is then used to calibrate the constants. Since forecasts of individual choice behavior are aggregated into geographic, socioeconomic, or supply market groups of interest for the purpose of transportation planning and policy analysis, such models are subject to aggregation error (McFadden et al., 1977). Depending on the geographic level of aggregation, the aggregation error could be substantial. In addition, coefficients in the mode choice models are statistical estimates and thus are subject to the so-called “forecast errors”.

Predictions of future travel behavior are based on hypotheses about the factors that influence travel behavior and the structure of those influences (Koppelman, 1976). These hypotheses are carried through the model formulation and prediction process in the steps of model specification, data collection, estimation of model parameters, and prediction of future travel behavior. Up to present, these hypotheses are still difficult to validate due to the lack of intensive data required. However, with the advanced technology in transportation planning and Geographic Information System (GIS), more specific data have become available to assist the task of validation. The following section describes the backgrounds of the aforementioned aggregation and forecast errors.

5.2 Background

In the travel demand modeling process, disaggregate models that predict individual choice behaviors are commonly used to obtain group predictions (Koppelman, 1975a; Koppelman, 1976; McFadden and Reid, 1975; Parody, 1977; Ben-Akiva and Lerman, 1985; Daly and Ortúzar, 1990). Disaggregate choice models are developed based on the assumptions that travel behavior represents an individual's choice response to the stimulus of a set of available alternatives. These models relate the probability of choosing one out of a set of available alternatives to the estimated utility of each alternative for the individual decision maker. A theoretically consistent aggregation procedure is then applied to determine the proportion of the prediction group expected to choose an alternative. Koppelman identified the following two major error sources in the modeling practice (Koppelman, 1976):

- forecast errors associated with the calibrated disaggregate choice model, and
- aggregation errors due to the procedure used to estimate aggregate predictions.

Errors in the choice model and errors in variables interact to produce errors in the prediction of individual choice probabilities. The errors are propagated through the aggregation procedure to produce errors in aggregate prediction. The aggregation procedure itself also introduces error directly into the aggregate prediction.

More detailed error classifications can be found in (Daly and Ortúzar, 1990). The types of errors that may enter a travel demand forecast model were defined as follows (Daly and Ortúzar, 1990):

- measurement errors arise in the actual measurement process in the base year,
- sampling errors introduced by the models that are estimated using finite datasets,
- computational errors inherit from the iterative process that is designed to search for optimal solutions,
- specification errors that take place in the model built-up process,
- transfer errors that occur when a model developed in one context is applied in a different context, and
- data aggregation.

The first five types of errors may be interpreted as the forecast errors defined by Koppelman. The definitions of both forecast and aggregation errors as well as the findings and conclusions from selected literature are presented in the following subsections.

Forecast Errors

Forecast errors in a disaggregate choice model may be the result of misspecification of the utility function and errors in the measurement of the independent variables. Examples of misspecification errors include the case in which the true explanatory variables in the utility function are either excluded or replaced by other variables. Errors in the independent variables include those introduced by imperfect sample data, such as reporting incorrect values for a specific survey question. In addition, the errors introduced by borrowing coefficients from the model calibrated on one data set collected in one area during one time period and then applying the model to a different time and/or place could also be classified as misspecification errors. Furthermore, coefficients in the choice model are sample estimates used to compute the choice probabilities of the population and are thus subject to sampling errors. For example, if the systematic utilities in a choice model are linear in their parameters, the asymptotic variance of the systematic utilities is given by (Ben-Akiva and Lerman, 1985)

$$\operatorname{var}\left[\hat{\boldsymbol{\beta}}'\mathbf{x}_{in}\right] = \mathbf{x}'_{in}\sum\nolimits_{\boldsymbol{\beta}}\mathbf{x}_{in} \tag{5.1}$$

where

$\hat{\boldsymbol{\beta}}$	=	estimated coefficients for variables in the systematic utility function;
$\mathbf{x}_{in}$	=	measurable characteristics included in the systematic utility function for individual n for mode i;
$\sum_{\boldsymbol{\beta}}$	=	asymptotic variance-covariance matrix of the coefficient estimates;
C_n	=	choice set for individual n.

Likewise,

$$\operatorname{cov}\left[\hat{\boldsymbol{\beta}}'\mathbf{x}_{in}, \hat{\boldsymbol{\beta}}'\mathbf{x}_{jn}\right] = \mathbf{x}'_{in}\sum\nolimits_{\boldsymbol{\beta}}\mathbf{x}_{jn} \tag{5.2}$$

Thus,

$$\operatorname{cov}\left[\hat{P}_n(i)\right] = \sum_{j\in C_n}\sum_{i\in C_n}\frac{\partial P_n(i)}{\partial\left(\boldsymbol{\beta}'\mathbf{x}_{in}\right)}\frac{\partial P_n(j)}{\partial\left(\boldsymbol{\beta}'\mathbf{x}_{jn}\right)}\operatorname{cov}\left[\hat{\boldsymbol{\beta}}'\mathbf{x}_{in}, \hat{\boldsymbol{\beta}}'\mathbf{x}_{jn}\right] \tag{5.3}$$

Applying these two equations with 874 samples collected in 1968 from Washington, D.C., Koppelman found that the standard error of the estimate of the probability for a mode choice was about 11% of the estimate itself (Koppelman, 1975a; Koppelman, 1975b). It was believed that most of the errors associated with the choice model are unknown to general transportation practitioners since they were not involved in the model calibration process and were lacking in original sample data, i.e., $\mathbf{x}_{in}$, that were used for the model calibration. As a result, forecast errors introduced by sampling errors are generally ignored. For the same reason, forecast errors are not investigated in this study.

<u>Aggregation Errors</u>

Aggregation errors, also known as aggregation biases (Landau 1981), are those introduced during the approximation procedure applied to estimate the values of the explanatory variables for each individual in the prediction group. A variety of alternative procedures with less extensive and intensive input data requirement have been proposed in the past. Each procedure reduces the problem of aggregating forecasts across individuals by making some simplifying assumptions about the choice model, the population, or both (Ben-Akiva and Lerman 1985). These procedures can be categorized as follows (Koppelman 1976, Ben-Akiva and Lerman 1985):

- Average individual, also known as native procedure, that predicts the expected choice share by using average or representative values of variables for the entire prediction group,
- Classification procedure that predicts the expected choice shares for individuals classes by using average or representative values of variables for each class,
- Statistical differential that predicts aggregate shares in terms of the moments, typically the first and second central moments, of the distribution of explanatory variables,
- Summation-integration procedure that takes sums or integrals of the attributions in the population, and
- Sample enumeration procedures that estimate expected shares by averaging the choice probabilities for a sample of the prediction group.

Sample enumeration uses a random sample of the population as representative of the entire population. The predicted share of the samples choosing alternative *i* is then used as an estimate for the fraction of population choosing alternative *i*, *W*(*i*):

$$\hat{W}(i) = \frac{1}{N_s}\sum_{n=1}^{N_s} P(i \mid \mathbf{x}_n) \tag{5.4}$$

where N_s is the number of decision makers in the sample. Monte Carlo simulation is typically adopted in a sample enumeration procedure to reduce the computational burden of producing a forecast. Monte Carlo method calculates the probability for each individual in the sample choosing a given alternative and then assigns the individual to the corresponding alternative with the highest value. The proportion of individuals in the sample that are assigned to alternative *i* is then used as an estimate of the share of the population choosing *i*.

The last three procedures listed above require detailed information that is generally not available to transportation analysts. In addition, the classification procedure was found to work extremely well even when a relatively small number of classes was used. Table 5.1 gives, at different levels of geographic aggregation, the weighted root mean square errors (RMSEs) due to classification for work trip modal shares for downtown workers (Koppelman, 1975a). The RMSE was weighted to account for the variation in number of observations across classes. The RMSE (in percentage) is calculated as follows:

$$RMSE = \left\{ \sum_{g=1}^{G} \frac{N_g}{N_T} \sum_{i \in C_g} O_g(i) \left[\frac{W_g(i) - O_g(i)}{W_g(i)} \right]^2 \right\}^{1/2} \times 100 \tag{5.5}$$

where

$O_g(i)$	=	sample enumeration forecast of the share of travelers in group g using alternative i;
N_g	=	number of decision makers in group g;
N_T	=	total number of decision makers in the population;
$W_g(i)$	=	the share of the population in group g choosing alternative i; and

C_g = choice set for group g.

As Table 5.1 shows, Koppelman's results were relatively low in magnitude and insensitive to geographic scale. Consequently, the classification procedure is typically implemented in the current practice.

Table 5.1 % RMSEs due to Classification

		Demographic Classification		
		No drive-alone Auto available < 1 Auto available ≥ 1	Auto available ≤ 0.25 0.25 < Auto available < 1 Auto available ≥ 1	No drive-alone Drive along
Level of Geographic Aggregation	45 Districts	3.3	9.9	5.2
	10 Groups	2.9	8.4	5.2
	4 Rings	2.8	7.9	5.3

Koppelman presented an analysis of error propagation for the binary logit model (Koppelman, 1976). He concluded that aggregation error by the naive and classification procedures was small compared to forecast error. He showed that aggregate share prediction based on disaggregate choice models was relatively accurate and the forecast error was higher than the aggregation error in both magnitude and insensitivity.

McFadden *et al.* investigated the errors introduced by different levels of geographic and socioeconomic aggregation for work trips modal shares on a sample of 771 workers drawn from about half of the San Francisco Bay area (McFadden et al., 1977; Reid, 1977). In the data, the overall exact mode shares were 55.6% auto alone, 17.4% bus-with-walk-access, 3.9% bus-with-drive-access, and 23.1% shared ride. The following percent root mean square (% RMS) of the choice shares was applied to measure the aggregation error:

$$\left(\sum_{j=1}^{J} \left(\frac{\hat{P}_j - P_j}{P_j} \right)^2 P_j \right)^{0.5} \tag{5.6}$$

where

J = number of alternative modes,

$\hat{P}_j$ = aggregate share of alternative j estimated by the tested method, and

P_j = aggregate share by enumeration.

The resulting errors from the average values of the explanatory variables in classes, as defined by the residential origin of the trips at the indicated geographic scale, are shown in Table 5.2. The numbers in parentheses at each geographic classification are the number of cells at the corresponding scale.

Table 5.2 % RMS for Three Scales of Geographic Classifications

	Region (1)	Cities (17)	TAZ (200)
% RMS	40.0	17.9	13.8

The results in Table 5.2 shows that geographic classification alone may not be adequate since the errors were significantly large. Five methods of aggregate prediction of regional choice shares were subsequently applied and the results are given in Table 5.3. It can be concluded from these results that aggregation errors could be quite large unless the groups within the population with different choice sets were treated separately.

Table 5.3 % RMS for Three Scales of Geographic Classifications

	Naive	Statistical Differentials	Classification by City (17)	Classification by Auto Ownership	Classification by Utility Scale (4)
% RMS	40.0	121.0	17.9	21.7	3.1

5.3 ERRORS IN MODAL SPLIT

In this section, the background of both aggregation error and prediction error in the multinomial logit (MNL) and nested logit models (NL) that are incorporated in FSUTMS is described. The existing MNL models and in FSUTMS are first presented, followed by the method that could be used to approximate the aggregation error and prediction error.

Multinomial Logit (MNL) Model in FSUTMS

In FSUTMS, MNL mode choice models include the single-path (SP), multi-path (MP), and multi-period/multi-path (MP/MP) applications. A simplified representation of the binary logit formula is given as follows (FDOT, 1997):

$$MS_t = A + \frac{1 - A - B}{1 + e^{C(\Delta u + D)}} \tag{5.7}$$

where

MS_t	=	mode split to transit mode,
A	=	transit captive fraction,
B	=	highway captive fraction,
C	=	slope parameter which determines elasticity of the logit curve,
D	=	point of symmetry of the logit curve, and
Δu	=	transit disutility minus auto disutility.

In FSUTMS, a default system-wide captive rate is specified in the MODESP.SYN, MODEMP.SYN, and MODEMPMP.SYN files for each transportation mode available in a given urban network (FDOT, 1997). After the captive highway and transit trips for each trip purpose in a TAZ are calculated, the logit model is applied to estimate the choice probability for a given mode by using a vector of representative values for the explanatory variable incorporated in the mode choice model. The existing practice eliminated the trips made by captive individuals and only considers the trips made by individuals with choices for highway and transit modes in the logit model. The explanatory variables estimated at aggregate TAZ level are then used to calculate the proportion of trips for each available mode. The modes considered in the applications of mode choice models are given in Table 5.4.

Table 5.4 Modes in Various MNL Mode Choice Models in FSUTMS

	SP	MP	MP/MP
Modes	1. Drive along auto 2. Two person auto 3. Three+ person auto 4. Transit	1. Drive Alone Auto 2. Two Person Auto 3. Three+ Person Auto 4. Local Bus 5. Line Haul Walk Access 6. Line Haul Drive Alone Auto Access 7. Line Haul Shared Ride Auto Access	1. Drive Alone Auto plus truck/taxi plus IE[1] 2. 2 Person Auto 3. 3+ Person Auto 4. EE[2] 5. Local bus HBW[3] 6. Local bus HBO[4] and NHB[5] 7. LHWAHBW[6] 8. LHWAHBNW[7] 9. LHAAHBW[8] 10. LHAANW[9]

Notes:

1. Internal-External;
2. External-External;
3. Home-Based Work;
4. Home-Based Other;
5. None-Home Based;
6. Line haul transit with walk access for HBW trips;
7. Line haul transit with walk access for Home-based Non-work trips;
8. Line haul transit with auto access for HBW trips; and
9. Line haul transit with auto access for Home-based Non-work trips.

The utility function for the three MNL model applications for the trips by mode M and trip purpose P is given as follows:

$$U(M, P) = AK(P)\times WK(M, P) + BK(P)\times WT(M, P) + CK(P)\times IVT(M, P) + DK(P)\times PC(M, P) + EK(P)\times OC(M, P) + K(M, P) \quad (5.8)$$

where

K = constant for each mode and purpose,
WK = walk time,
WT = wait time,
IVT = in-vehicle time,
PC = parking cost,
OC = other cost (auto operating cost and transit fare),
U = utility, and
AK, BK, CK, DK, EK = purpose specific utility coefficients.

The coefficients used in the utility function for each purpose and each mode are specified in the MODESP.SYN, MODEMP.SYN, or MODEMPMP.SYN files

Nested Logit (NL) Model in FSUTMS

The nested logit model can be graphically expressed as a structure of four-tier hierarchy. (Abdel-Aty and Abdelwahab, 2001) As depicted in Figure 5.1, the following nice modes are available in FSUTMS (in parenthesis is the abbreviation used later for model description):

1. Drive alone (DA)
2. Shared ride (SD)
3. Walk-access to local bus (LW)
4. Walk-access to express bus (EW)
5. Walk-access to metro rail (MW)
6. Walk-access to tri rail (TW)
7. Auto-access to express bus (ED)
8. Auto-access to metro rail (MD)
9. Auto-access Walk to tri rail (TD)

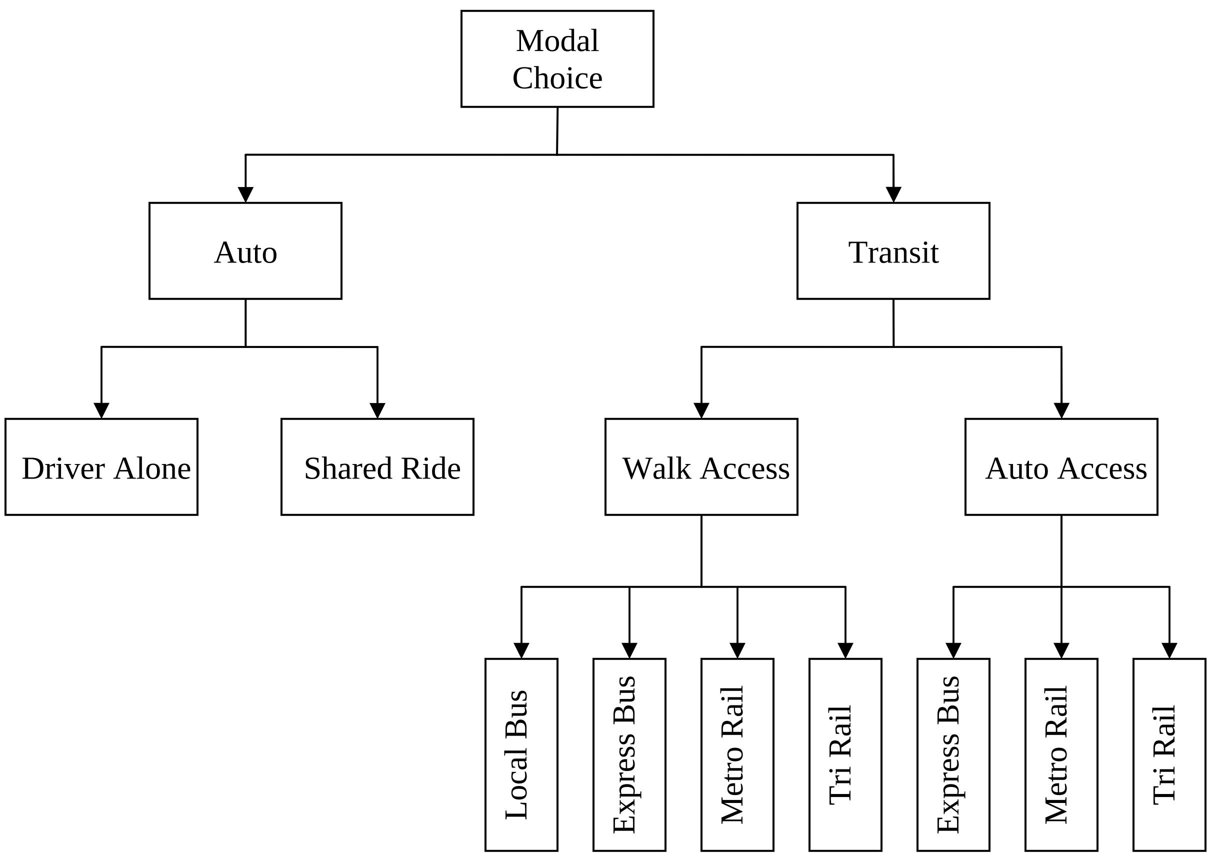

Figure 5.1 The Nested Logit Structure of the Modal Choice in FSUTMS

Let **x** be the set of explanatory variables and β be the vector of the model parameters associated with each mode. The utility functions for each mode in the lowest (fourth) hierarchy can then be expressed as:

<u>For Transit</u>

$U_{LW} = \mathrm{x}'\beta_{\mathrm{LW}}$, $U_{EW} = \mathrm{x}'\beta_{\mathrm{EW}}$, $U_{MW} = \mathrm{x}'\beta_{\mathrm{MW}}$, $U_{TW} = \mathrm{x}'\beta_{\mathrm{TW}}$, $U_{ED} = \mathrm{x}'\beta_{\mathrm{ED}}$, $U_{MD} = \mathrm{x}'\beta_{\mathrm{MD}}$, and $U_{TD} = \mathrm{x}'\beta_{\mathrm{TD}}$

<u>For Auto</u>

$U_{DA} = \mathrm{x}'\beta_{\mathrm{DA}}$ and $U_{SD} = \mathrm{x}'\beta_{\mathrm{SD}}$

The conditional probability of each mode given different access modes, i.e., Walk (WK) versus Drive (DV), under the transit system (TR) is:

$$P_{LW|WK|TR} = \frac{e^{U_{LW}}}{e^{U_{LW}} + e^{U_{EW}} + e^{U_{MW}} + e^{U_{TW}}}$$

$$P_{EW|WK|TR} = \frac{e^{U_{EW}}}{e^{U_{LW}} + e^{U_{EW}} + e^{U_{MW}} + e^{U_{TW}}}$$

$$P_{MW|WK|TR} = \frac{e^{U_{MW}}}{e^{U_{LW}} + e^{U_{EW}} + e^{U_{MW}} + e^{U_{TW}}}$$

$$P_{TW|WK|TR} = \frac{e^{U_{TW}}}{e^{U_{LW}} + e^{U_{EW}} + e^{U_{MW}} + e^{U_{TW}}}$$

$$P_{ED|DV|TR} = \frac{e^{U_{ED}}}{e^{U_{ED}} + e^{U_{MD}} + e^{U_{TD}}}$$

$$P_{MD|DV|TR} = \frac{e^{U_{MD}}}{e^{U_{ED}} + e^{U_{MD}} + e^{U_{TD}}} \text{, and}$$

$$P_{TD|DV|TR} = \frac{e^{U_{TD}}}{e^{U_{ED}} + e^{U_{MD}} + e^{U_{TD}}}$$

The inclusive values for the “Access” (third) hierarchy, i.e., Walk versus Drive, are

$$I_{WK} = \ln(e^{U_{LW}} + e^{U_{EW}} + e^{U_{MW}} + e^{U_{TW}}) \text{, and } I_{DV} = \ln(e^{U_{ED}} + e^{U_{MD}} + e^{U_{TD}})$$

Based on these derived inclusive values, the conditional probabilities of access modes given the transit system can be calculated as:

$$P_{WK|TR} = \frac{e^{\tau_{WK} I_{WK}}}{e^{\tau_{WK} I_{WK}} + e^{\tau_{DV} I_{DV}}}, \text{ and } P_{DV|TR} = \frac{e^{\tau_{DV} I_{DV}}}{e^{\tau_{WK} I_{WK}} + e^{\tau_{DV} I_{DV}}}$$

where τ_{WK} and τ_{DV} are parameters associated with inclusive values. One could consider τ_{WK} and τ_{DV} as the elements in variable **x** and I_{WK} and I_{DV} as the elements in parameter vector, β_{WK} and β_{DV}, respectively. This leads to:

$$U_{WK} = x'\beta_{WK} = \tau_{WK} I_{WK}, \text{ and } U_{DV} = x'\beta_{DV} = \tau_{DV} I_{DV}$$

Similarly, the inclusive value for the next higher hierarchy can be written as I_{TR} = $\ln(e^{\tau_{WK} I_{WK}} + e^{\tau_{DV} I_{DV}})$. Once these parameters are estimated, the share of each mode can be calculated as follows:

$$P_{LW} = P_{LW|WK|TR} \cdot P_{WK|TR} \cdot P_{TR} = \frac{e^{U_{LW}}}{e^{U_{LW}} + e^{U_{EW}} + e^{U_{MW}} + e^{U_{TW}}} \cdot \frac{e^{\tau_{WK} I_{WK}}}{e^{\tau_{WK} I_{WK}} + e^{\tau_{DV} I_{DV}}} \cdot P_{TR}$$

$$P_{EW} = P_{EW|WK|TR} \cdot P_{WK|TR} \cdot P_{TR} = \frac{e^{U_{EW}}}{e^{U_{LW}} + e^{U_{EW}} + e^{U_{MW}} + e^{U_{TW}}} \cdot \frac{e^{\tau_{WK} I_{WK}}}{e^{\tau_{WK} I_{WK}} + e^{\tau_{DV} I_{DV}}} \cdot P_{TR}$$

$$P_{MW} = P_{MW|WK|TR} \cdot P_{WK|TR} \cdot P_{TR} = \frac{e^{U_{MW}}}{e^{U_{LW}} + e^{U_{EW}} + e^{U_{MW}} + e^{U_{TW}}} \cdot \frac{e^{\tau_{WK} I_{WK}}}{e^{\tau_{WK} I_{WK}} + e^{\tau_{DV} I_{DV}}} \cdot P_{TR}$$

$$P_{TW} = P_{TW|WK|TR} \cdot P_{WK|TR} \cdot P_{TR} = \frac{e^{U_{TW}}}{e^{U_{LW}} + e^{U_{EW}} + e^{U_{MW}} + e^{U_{TW}}} \cdot \frac{e^{\tau_{WK} I_{WK}}}{e^{\tau_{WK} I_{WK}} + e^{\tau_{DV} I_{DV}}} \cdot P_{TR}$$

$$P_{ED} = P_{ED|DV|TR} \cdot P_{DV|TR} \cdot P_{TR} = \frac{e^{U_{ED}}}{e^{U_{ED}} + e^{U_{MD}} + e^{U_{TD}}} \cdot \frac{e^{\tau_{DV} I_{DV}}}{e^{\tau_{WK} I_{WK}} + e^{\tau_{DV} I_{DV}}} \cdot P_{TR}$$

$$P_{MD} = P_{MD|DV|TR} \cdot P_{DV|TR} \cdot P_{TR} = \frac{e^{U_{MD}}}{e^{U_{ED}} + e^{U_{MD}} + e^{U_{TD}}} \cdot \frac{e^{\tau_{DV} I_{DV}}}{e^{\tau_{WK} I_{WK}} + e^{\tau_{DV} I_{DV}}} \cdot P_{TR}$$

$$P_{TD} = P_{TD|DV|TR} \cdot P_{DV|TR} \cdot P_{TR} = \frac{e^{U_{TD}}}{e^{U_{ED}} + e^{U_{MD}} + e^{U_{TD}}} \cdot \frac{e^{\tau_{DV} I_{DV}}}{e^{\tau_{WK} I_{WK}} + e^{\tau_{DV} I_{DV}}} \cdot P_{TR}$$

For the higher hierarchies, the probability terms are listed as follows:

$$P_{DA|HY} = \frac{e^{U_{DA}}}{e^{U_{DA}} + e^{U_{SD}}}, \; P_{SD|HY} = \frac{e^{U_{SD}}}{e^{U_{DA}} + e^{U_{SD}}}$$

$$I_{HY} = \ln(e^{U_{DA}} + e^{U_{SD}}), \; I_{Tr} = \ln(e^{U_{TR}})$$

$$P_{TR} = \frac{e^{U_{TR}}}{e^{U_{TR}} + e^{U_{HY}}},\ P_{HY} = \frac{e^{U_{HY}}}{e^{U_{TR}} + e^{U_{HY}}},\ \text{where } U_{TR} = \mathrm{x}'\beta_{\mathrm{TR}},\ \text{and } U_{HY} = \mathrm{x}'\beta_{\mathrm{HY}}$$

$$P_{DA} = P_{DA|HY} \cdot P_{HY} = \frac{e^{U_{DA}}}{e^{U_{DA}} + e^{U_{SD}}} \cdot P_{HY}$$

$$P_{SD} = P_{SD|HY} \cdot P_{HY} = \frac{e^{U_{SD}}}{e^{U_{DA}} + e^{U_{SD}}} \cdot P_{HY}$$

Aggregation Errors

Modal split models adopted in FSUTMS were originally calibrated at the individual trip maker level. When implemented, however, variables aggregated at the TAZ level are used to estimate the mode share. As mentioned, the current practice is subject to the errors occurred in the aggregation and forecast process. This section presents the theoretical background of aggregation errors on the total number of trips estimated for a given mode when perfect model parameters and variables are used, i.e., no prediction error exists. In other words, the probability of selecting a given mode for the trips originated at production TAZ i to travel to attraction TAZ j has to be precisely quantified. Denote the total expected number of trips choosing a given alternative mode m in study area T as $N_T(m)$, it follows that:

$$N_T(m) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{k_i}\sum_{l=1}^{l_k} P\left(m \middle| \mathbf{x}_{ijkl}\right) \tag{5.9}$$

where

k = index for an individual in TAZ i;
l = index for trips made by an individual k in TAZ i;
k_i = number of trip makers in TAZ i;
l_k = number of trips made by individual k in TAZ i;
n = number of internal TAZs in study area T; and
p_k = number of trips made by individual k in TAZ i.

In Equation (5.9), $P(m|\mathbf{x}_{ijkl})$ is the probability for a trip maker k in TAZ i to choose mode m when making trip l to destination TAZ j. Consequently, the vector $\mathbf{x}_{ijkp}$ should contain the attributes associated with an individual, k, in TAZ i for making a trip, p, to destination TAZ j. However, in the traditional four-step sequential modeling, the attributes associated with an individual are not available. Instead, aggregated attributes at TAZ level, $\mathbf{x}_{ij}$, are used. Given that the number of trips allocating to zone pair i and j is a scale value without variation and

$$E\left[P\left(m \middle| \mathbf{x}_{ij}\right)\right] = \int_{\mathbf{x}_{ij}} P\left(m \middle| \mathbf{x}_{ij}\right) P\left(\mathbf{x}_{ij}\right) d\mathbf{x}, \tag{5.10}$$

Equation (5.9) becomes

$$N_T(m) \cong \sum_{i=1}^{n}\sum_{j\neq i}^{n} T_{ij}(m) = \sum_{i=1}^{n}\sum_{j\neq i}^{n} T_{ij} \times E\left(P\left(m\middle|\mathbf{x}_{ij}\right)\right) \cong \sum_{i=1}^{n}\sum_{j\neq i}^{n} T_{ij} \times P\left(m\middle|\tilde{\mathbf{x}}_{ij}\right). \tag{5.11}$$

where

T_{ij} = number of interchanging trips between TAZs i and j;
$\mathbf{x}_{ij}$ = a vector of aggregate zonal attributes between TAZs i and j;
$\tilde{\mathbf{x}}_{ij}$ = a vector of representative values.

In the current practice, $\tilde{\mathbf{x}}_{ij}$ is used to estimate the zonal overall probability of selecting alternative mode m. The discrepancy in Equation (5.11) is:

$$\begin{aligned} Err_m &= \sum_{i=1}^{n}\sum_{j\neq i}^{n} T_{ij} \times E\left(P\left(m\middle|\mathbf{x}_{ij}\right)\right) - \sum_{i=1}^{n}\sum_{j\neq i}^{n} T_{ij} \times P\left(m\middle|\tilde{\mathbf{x}}_{ij}\right) \\ &= \sum_{i=1}^{n}\sum_{j\neq i}^{n} T_{ij} \times \left\{E\left(P\left(m\middle|\mathbf{x}_{ij}\right)\right) - P\left(m\middle|\tilde{\mathbf{x}}_{ij}\right)\right\} \end{aligned} \tag{5.12}$$

By expanding $P\left(m\middle|\mathbf{x}_{ij}\right)$ as a second-order Taylor's series around $\bar{\mathbf{x}}_{ij}$, Equation (5.12) becomes (Ben-Akiva and Lerman 1985)

$$\begin{aligned} P\left(m\middle|\mathbf{x}_{ij}\right) &\cong P\left(m\middle|\bar{\mathbf{x}}_{ij}\right) + \sum_{r=1}^{R} \frac{\partial}{\partial x_r} P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\bar{\mathbf{x}}_{ij}} \left(x_{ijr} - \bar{x}_{ijr}\right) \\ &\quad + \frac{1}{2}\sum_{r=1}^{R}\sum_{r'=1}^{R} \frac{\partial^2}{\partial x_r \partial x_{r'}} P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\bar{\mathbf{x}}_{ij}} \left(x_{ijr} - \bar{x}_{ijr}\right)\left(x_{ijr'} - \bar{x}_{ijr'}\right) \end{aligned} \tag{5.13}$$

Thus,

$$\begin{aligned} E\left[P\left(m\middle|\mathbf{x}_{ij}\right)\right] &\cong \int_{\mathbf{x}_{ij}} P\left(m\middle|\bar{\mathbf{x}}_{ij}\right) P\left(\mathbf{x}_{ij}\right) d\mathbf{x} + \sum_{r=1}^{R} \int \frac{\partial}{\partial x_r} P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\bar{\mathbf{x}}_{ij}} \left(x_{ijr} - \bar{x}_{ijr}\right) P\left(\mathbf{x}_{ij}\right) d\mathbf{x} \\ &\quad + \frac{1}{2}\sum_{r=1}^{R}\sum_{r'=1}^{R} \int \frac{\partial^2}{\partial x_r \partial x_{r'}} P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\bar{\mathbf{x}}_{ij}} \left(x_{ijr} - \bar{x}_{ijr}\right)\left(x_{ijr'} - \bar{x}_{ijr'}\right) P\left(\mathbf{x}_{ij}\right) d\mathbf{x} \end{aligned} \tag{5.14}$$

$$\begin{aligned} E\left[P\left(m\middle|\mathbf{x}_{ij}\right)\right] &\cong P\left(m\middle|\bar{\mathbf{x}}_{ij}\right) + \sum_{r=1}^{R} \frac{\partial}{\partial x_r} P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\bar{\mathbf{x}}_{ij}} \int \left(x_{ijr} - \bar{x}_{ijr}\right) P\left(\mathbf{x}_{ij}\right) d\mathbf{x} \\ &\quad + \frac{1}{2}\sum_{r=1}^{R}\sum_{r'=1}^{R} \frac{\partial^2}{\partial x_r \partial x_{r'}} P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\bar{\mathbf{x}}_{ij}} \int \left(x_{ijr} - \bar{x}_{ijr}\right)\left(x_{ijr'} - \bar{x}_{ijr'}\right) P\left(\mathbf{x}_{ij}\right) d\mathbf{x} \\ &= P\left(m\middle|\bar{\mathbf{x}}_{ij}\right) + \frac{1}{2}\sum_{r=1}^{R}\sum_{r'=1}^{R} \frac{\partial^2}{\partial x_r \partial x_{r'}} P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\bar{\mathbf{x}}_{ij}} \operatorname{cov}\left(x_{ijr}, x_{ijr'}\right) \end{aligned} \tag{5.15}$$

As a result, if $\tilde{\mathbf{x}}_{ij}$ is equal to $\bar{\mathbf{x}}_{ij}$, Equation (5.12) may be simplified to:

$$
\begin{aligned}
Err_m &= \sum_{i=1}^{n}\sum_{j\neq i}^{n} T_{ij} \times \left\{ P\left(m\middle|\bar{\mathbf{x}}_{ij}\right) + \frac{1}{2}\sum_{r=1}^{R}\sum_{r'=1}^{R} \frac{\partial^2}{\partial x_r \partial x_{r'}} P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\bar{\mathbf{x}}_{ij}} \operatorname{cov}\left(x_{ijr}, x_{ijr'}\right) - P\left(m\middle|\tilde{\mathbf{x}}_{ij}\right) \right\} \\
&= \sum_{i=1}^{n}\sum_{j\neq i}^{n} T_{ij} \times \frac{1}{2}\sum_{r=1}^{R}\sum_{r'=1}^{R} \frac{\partial^2}{\partial x_r \partial x_{r'}} P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\bar{\mathbf{x}}_{ij}} \operatorname{cov}\left(x_{ijr}, x_{ijr'}\right)
\end{aligned}
\tag{5.16}
$$

where $P(m|\mathbf{x}_{ij})$ could either be the MNL model in the following form or the NL model:

$$
P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\bar{\mathbf{x}}_{ij}} = \frac{e^{\boldsymbol{\beta}'\bar{\mathbf{x}}_{ij}}}{\sum_{j'=1}^{n} e^{\boldsymbol{\beta}'\bar{\mathbf{x}}_{ij'}}}.
$$

As shown in Equation (5.16), the area-wide total error for the number of trips choosing a given mode requires the knowledge of the covariance between each pair of the variables included in a MNL mode choice model. However, this information is not likely to be available to transportation professionals. The trips allocating to zone pair i and j may also be a contributing factor in the discrepancy of mode choice estimate. For simplification purpose, such a variation is discarded in the study. In addition, to quantify the discrepancy of total number of trips selecting a given transportation mode, the probability for a mode choice model aggregated at the TAZ level, i.e., $P(m|\mathbf{x}_{ij})$, has to be precisely estimated. Since the true $P(m|\mathbf{x}_{ij})$ is unknown to the modelers, it may be replaced with a calibrated MNL or NL model. Equation (5.16) thus becomes:

$$
Err_m \cong \sum_{i=1}^{n}\sum_{j\neq i}^{n} T_{ij} \times \frac{1}{2}\sum_{r=1}^{R}\sum_{r'=1}^{R} \frac{\partial^2}{\partial x_r \partial x_{r'}} \hat{P}\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\bar{\mathbf{x}}_{ij}} \operatorname{cov}\left(x_{ijr}, x_{ijr'}\right) \tag{5.17}
$$

See the appendix II for the deviations of the second derivative of $P\left(i\middle|\mathbf{x}_g\right)$ with respect to x_r and $x_{r'}$. The complete vector of choice-relevant attributes for every individual in a given TAZ is generally unknown to the transportation professionals. As shown in Equation (5.17), the area-wide total error for the number of trips choosing a given transportation mode requires the knowledge of each covariance between the attributes included in the mode choice model. The research team is currently collecting the necessary information to investigate the aggregation error exhibited in both MNL and NL models.

In addition to aggregation error, modal split models are calibrated with variables collected from limited number of samples and thus subjected to the errors introduced in the sampling and estimation process. In the following section, the errors introduced by the parameters and variables of the modal split models are described.

Errors Subjected to Both Aggregation and Predictive Variability

The previous section assumes that perfect model parameters and variables are used, i.e., no prediction error exists. This section relaxes part of the assumption and incorporates statistical variability of model parameters into the derivation of modal split errors. No error in variables is assumed for simplicity. However, this simplification will not pose a serious concern if the error in variables is assumed additive along with the aggregation error. For notational convenience, parameters and variables are expressed in vector for the majority of derivation steps.

$$\begin{aligned} P(m \mid \mathbf{x},\boldsymbol{\beta}) \cong\ & P(m \mid \bar{\mathbf{x}},\bar{\boldsymbol{\beta}}) + (\mathbf{x}-\bar{\mathbf{x}})^T \frac{\partial P}{\partial \mathbf{x}}\Big|_{\mathbf{x}=\bar{\mathbf{x}}} + (\boldsymbol{\beta}-\bar{\boldsymbol{\beta}})^T \frac{\partial P}{\partial \boldsymbol{\beta}}\Big|_{\boldsymbol{\beta}=\bar{\boldsymbol{\beta}}} + \frac{1}{2}(\mathbf{x}-\bar{\mathbf{x}})^T \frac{\partial^2 P}{\partial \mathbf{x}\partial \mathbf{x}'}(\mathbf{x}-\bar{\mathbf{x}}) \\ & + \frac{1}{2}(\boldsymbol{\beta}-\bar{\boldsymbol{\beta}})^T \frac{\partial^2 P}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}'}(\boldsymbol{\beta}-\bar{\boldsymbol{\beta}}) + (\mathbf{x}-\bar{\mathbf{x}})^T \frac{\partial^2 P}{\partial \mathbf{x}\partial \boldsymbol{\beta}'}(\boldsymbol{\beta}-\bar{\boldsymbol{\beta}}) \end{aligned} \tag{5.18}$$

Application of the expection operator, $E(\ )$, on both sides of the above equation yields:

$$\begin{aligned} & E\left[P(m \mid \mathbf{x},\boldsymbol{\beta})\right] \\ & \cong P(m \mid \bar{\mathbf{x}},\bar{\boldsymbol{\beta}}) + \frac{1}{2}E[(\mathbf{x}-\bar{\mathbf{x}})^T \frac{\partial^2 P}{\partial \mathbf{x}\partial \mathbf{x}'}(\mathbf{x}-\bar{\mathbf{x}})] + \frac{1}{2}E[(\boldsymbol{\beta}-\bar{\boldsymbol{\beta}})^T \frac{\partial^2 P}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}'}(\boldsymbol{\beta}-\bar{\boldsymbol{\beta}})] \\ & \qquad + E[(\mathbf{x}-\bar{\mathbf{x}})^T \frac{\partial^2 P}{\partial \mathbf{x}\partial \boldsymbol{\beta}'}(\boldsymbol{\beta}-\bar{\boldsymbol{\beta}})] \\ & = P(m \mid \bar{\mathbf{x}},\bar{\boldsymbol{\beta}}) + \frac{1}{2}E\left\{tr[(\mathbf{x}-\bar{\mathbf{x}})^T \frac{\partial^2 P}{\partial \mathbf{x}\partial \mathbf{x}'}(\mathbf{x}-\bar{\mathbf{x}})]\right\} + \frac{1}{2}E\left\{tr[(\boldsymbol{\beta}-\bar{\boldsymbol{\beta}})^T \frac{\partial^2 P}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}'}(\boldsymbol{\beta}-\bar{\boldsymbol{\beta}})]\right\} \\ & \qquad + E\left\{tr[(\mathbf{x}-\bar{\mathbf{x}})^T \frac{\partial^2 P}{\partial \mathbf{x}\partial \boldsymbol{\beta}'}(\boldsymbol{\beta}-\bar{\boldsymbol{\beta}})]\right\} \\ & = P(m \mid \bar{\mathbf{x}},\bar{\boldsymbol{\beta}}) + \frac{1}{2}E\left\{tr[\frac{\partial^2 P}{\partial \mathbf{x}\partial \mathbf{x}'}(\mathbf{x}-\bar{\mathbf{x}})(\mathbf{x}-\bar{\mathbf{x}})^T]\right\} + \frac{1}{2}E\left\{tr[\frac{\partial^2 P}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}'}(\boldsymbol{\beta}-\bar{\boldsymbol{\beta}})(\boldsymbol{\beta}-\bar{\boldsymbol{\beta}})^T]\right\} \\ & \qquad + E\left\{tr[\frac{\partial^2 P}{\partial \mathbf{x}\partial \boldsymbol{\beta}'}(\boldsymbol{\beta}-\bar{\boldsymbol{\beta}})(\mathbf{x}-\bar{\mathbf{x}})^T]\right\} \\ & = P(m \mid \bar{\mathbf{x}},\bar{\boldsymbol{\beta}}) + \frac{1}{2}tr\left\{\frac{\partial^2 P}{\partial \mathbf{x}\partial \mathbf{x}'}E[(\mathbf{x}-\bar{\mathbf{x}})(\mathbf{x}-\bar{\mathbf{x}})^T]\right\} + \frac{1}{2}tr\left\{\frac{\partial^2 P}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}'}E[(\boldsymbol{\beta}-\bar{\boldsymbol{\beta}})(\boldsymbol{\beta}-\bar{\boldsymbol{\beta}})^T]\right\} \\ & \qquad + tr\left\{\frac{\partial^2 P}{\partial \mathbf{x}\partial \boldsymbol{\beta}'}E[(\boldsymbol{\beta}-\bar{\boldsymbol{\beta}})(\mathbf{x}-\bar{\mathbf{x}})^T]\right\} \\ & = P(m \mid \bar{\mathbf{x}},\bar{\boldsymbol{\beta}}) + \frac{1}{2}tr[\frac{\partial^2 P}{\partial \mathbf{x}\partial \mathbf{x}'}\mathrm{cov}(\mathbf{x})] + \frac{1}{2}tr[\frac{\partial^2 P}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}'}\mathrm{cov}(\boldsymbol{\beta})] + tr[\frac{\partial^2 P}{\partial \mathbf{x}\partial \boldsymbol{\beta}'}\mathrm{cov}(\boldsymbol{\beta},\mathbf{x})] \end{aligned}$$

where $tr(\ )$ is the trace operator. Assume that the interaction between the variance in parameters and the variance in variables is small and can be ignored, i.e, $\mathrm{cov}(\beta, \mathbf{x}) = \mathbf{0}$. Therefore,

$$E\left[P(m \mid \mathbf{x},\boldsymbol{\beta})\right] \cong P(m \mid \bar{\mathbf{x}},\bar{\boldsymbol{\beta}}) + \frac{1}{2}tr[\frac{\partial^2 P}{\partial \mathbf{x}\partial \mathbf{x}'}\mathrm{cov}(\mathbf{x})] + \frac{1}{2}tr[\frac{\partial^2 P}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}'}\mathrm{cov}(\boldsymbol{\beta})] \tag{5.19}$$

and the error in modal split is:

$$
\begin{aligned}
Err_m &= \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \left\{ T_{ij} \left(E\left[P(m \mid \mathbf{x}, \boldsymbol{\beta})\right] - P(m \mid \overline{\mathbf{x}}, \overline{\boldsymbol{\beta}}) \right)_{ij} \right\} \\
&= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \left\{ T_{ij} \left(tr[\frac{\partial^2 P}{\partial \mathbf{x} \partial \mathbf{x}'} \text{cov}(\mathbf{x})] + tr[\frac{\partial^2 P}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} \text{cov}(\boldsymbol{\beta})] \right)_{ij} \right\}
\end{aligned}
\qquad (5.20)
$$

This error term can be applied to derive the accuracy standard for transit mode share once their attributes are collected and covariance matrices of variables and parameters are obtained through estimation of the model parameters.

<u>Variability of Modal Split Trips</u>

This section presents the theoretical background of variability of the total number of trips estimated for a given mode. This information can be used to establish confidence intervals at a user-specified level of significance and an alternative form of accuracy standards to the one presented in the previous section. From Equation (5.11),

$$
\begin{aligned}
\text{var}\left(N_T(m)\right) &\cong \text{var}\left(\sum_{i=1}^{n} \sum_{j\neq i}^{n} T_{ij} \times P\left(m \middle| \mathbf{x}_{ij}\right) \right) \\
&= \sum_{i=1}^{n} \sum_{j\neq i}^{n} T_{ij}^{\,2} \times \text{var}\left(P\left(m \middle| \mathbf{x}_{ij}\right)\right) \\
&\quad + \sum_{i=1}^{n} \sum_{j\neq i}^{n} \sum_{i'=1}^{n} \sum_{\substack{j'\neq i' \\ j'\neq j}}^{n} T_{ij} \times T_{i'j'} \times \text{cov}\left(P\left(m \middle| \mathbf{x}_{ij}\right), P\left(m \middle| \mathbf{x}_{i'j'}\right)\right)
\end{aligned}
\qquad (5.21)
$$

As shown in Equation (5.21), the variance of the aggregate share prediction is composed of the sum of variance terms for the choice probabilities for the trips between every zone pair and the covariance of choice probabilities for every pair of trips for a given purpose. The variance and covariance terms in Equation (5.21) may be approximated as follows (Koppelman 1975a):

$$
\text{var}\left(P\left(m \middle| \mathbf{x}_{ij}\right)\right) \cong \left[\frac{\partial P\left(m \middle| \mathbf{x}_{ij}\right)}{\partial \mathbf{x}_{ij}{}' \boldsymbol{\beta}} \right]^2 \text{var}\left(\mathbf{x}_{ij}{}' \boldsymbol{\beta}\right) \qquad (5.22a)
$$

$$
\text{cov}\left(P\left(m \middle| \mathbf{x}_{ij}\right), P\left(m \middle| \mathbf{x}_{i'j'}\right)\right) \cong \left[\frac{\partial P\left(m \middle| \mathbf{x}_{ij}\right)}{\partial \mathbf{x}_{ij}{}' \boldsymbol{\beta}} \right] \left[\frac{\partial P\left(m \middle| \mathbf{x}_{i'j'}\right)}{\partial \mathbf{x}_{i'j'}{}' \boldsymbol{\beta}} \right] \text{cov}\left(\mathbf{x}_{ij}{}' \boldsymbol{\beta}, \mathbf{x}_{i'j'}{}' \boldsymbol{\beta}\right) \qquad (5.22b)
$$

Further assume that the errors introduced to the model variables in the various stages of the model formulation and prediction process by a random additive disturbance term as follows:

$$\mathbf{x}_{ij} = \overline{\mathbf{x}}_{ij} + \boldsymbol{\delta}_{ij}$$

where

$\tilde{\mathbf{x}}_{ij}$ = a vector of observed or predicted variables for a given zone pair;

$\overline{\mathbf{x}}_{ij}$ = a vector of true variables; and

$\boldsymbol{\delta}_{ij}$ = a vector of random disturbance terms.

The variance of the utility functions in Equation (5.21) may thus be disaggregated into the variance contributed by errors in parameters ($\boldsymbol{\beta}$) and the errors in choice variables ($\boldsymbol{\delta}_{ij}$) as follows, given that no interacting variance between the model parameters ($\boldsymbol{\beta}$) and the choice variables ($\mathbf{x}_{ij}$) (Koppelman 1975a):

$$\begin{aligned}\operatorname{var}\left(\mathbf{x}_{ij}{}'\boldsymbol{\beta}\right) &= \overline{\mathbf{x}}_{ij}{}' \operatorname{var}\left(\boldsymbol{\beta}\right)\overline{\mathbf{x}}_{ij} + \boldsymbol{\beta}\operatorname{var}\left(\boldsymbol{\delta}_{ij}\right)\boldsymbol{\beta}' + \mathbf{1}'\operatorname{var}\left(\boldsymbol{\beta}\right)\operatorname{var}\left(\boldsymbol{\delta}_{ij}\right)\mathbf{1} \\ &= \overline{\mathbf{x}}_{ij}{}'\mathbf{A}\overline{\mathbf{x}}_{ij} + \boldsymbol{\beta}\mathbf{Y}_{ij}\boldsymbol{\beta}' + \mathbf{1}'\mathbf{A}\mathbf{Y}_{ij}\mathbf{1}\end{aligned} \tag{5.23}$$

where

A = variance-covariance matrix for errors in parameters;

$\mathbf{Y}_{ij}$ = variance-covariance matrix for errors in variables for trips between TAZs i and j; and

1 = sum vector unit value for each of its elements.

Similarly, the covariance of utility values between pairs of trips may be disaggregated as follows (Koppelman 1975a):

$$\begin{aligned}\operatorname{cov}\left(\mathbf{x}_{ij}{}'\boldsymbol{\beta}, \mathbf{x}_{i'j'}{}'\boldsymbol{\beta}\right) &= \overline{\mathbf{x}}_{ij}{}'\operatorname{cov}\left(\boldsymbol{\beta},\boldsymbol{\beta}\right)\overline{\mathbf{x}}_{ij} + \boldsymbol{\beta}\operatorname{cov}\left(\boldsymbol{\delta}_{ij},\boldsymbol{\delta}_{i'j'}\right)\boldsymbol{\beta}' + \mathbf{1}'\operatorname{cov}\left(\boldsymbol{\beta},\boldsymbol{\beta}\right)\operatorname{cov}\left(\boldsymbol{\delta}_{ij},\boldsymbol{\delta}_{i'j'}\right)\mathbf{1} \\ &= \overline{\mathbf{x}}_{ij}{}'\operatorname{var}\left(\boldsymbol{\beta}\right)\overline{\mathbf{x}}_{ij} + \boldsymbol{\beta}\operatorname{cov}\left(\boldsymbol{\delta}_{ij},\boldsymbol{\delta}_{i'j'}\right)\boldsymbol{\beta}' + \mathbf{1}'\operatorname{var}\left(\boldsymbol{\beta}\right)\operatorname{cov}\left(\boldsymbol{\delta}_{ij},\boldsymbol{\delta}_{i'j'}\right)\mathbf{1} \\ &= \overline{\mathbf{x}}_{ij}{}'\mathbf{A}\overline{\mathbf{x}}_{ij} + \boldsymbol{\beta}\mathbf{Y}_{ij,i'j'}\boldsymbol{\beta}' + \mathbf{1}\mathbf{A}\mathbf{Y}_{ij,i'j'}\mathbf{1}\end{aligned} \tag{5.24}$$

where $\mathbf{Y}_{ij,i'j'}$ is the covariance matrix for measurement errors in variables between zone pairs i-j and i'-j'. If the interaction between the variance in parameters and the variance in variables is small and is thus ignored, Equations (5.22) becomes:

$$\operatorname{var}\left(P\left(m\middle|\mathbf{x}_{ij}\right)\right) \cong \left[\frac{\partial P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\overline{\mathbf{x}}_{ij}}}{\partial \mathbf{x}_{ij}{}'\boldsymbol{\beta}}\right]^2 \left(\overline{\mathbf{x}}_{ij}{}'\mathbf{A}\overline{\mathbf{x}}_{ij} + \boldsymbol{\beta}\mathbf{Y}_{ij}\boldsymbol{\beta}'\right) \tag{5.25a}$$

and

$$\operatorname{cov}\left(P\left(m\middle|\mathbf{x}_{ij}\right),P\left(m\middle|\mathbf{x}_{i'j'}\right)\right)\cong\left[\frac{\partial P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\overline{\mathbf{x}}_{ij}}}{\partial\mathbf{x}_{ij}{}'\boldsymbol{\beta}}\right]\left[\frac{\partial P\left(m\middle|\mathbf{x}_{i'j'}\right)\Big|_{\overline{\mathbf{x}}_{i'j'}}}{\partial\mathbf{x}_{i'j'}{}'\boldsymbol{\beta}}\right]\left(\overline{\mathbf{x}}_{ij}{}'\mathbf{A}\overline{\mathbf{x}}_{ij}+\boldsymbol{\beta}\mathbf{Y}_{ij,i'j'}\boldsymbol{\beta}'\right)\tag{5.25b}$$

Equation (5.21) becomes

$$\begin{aligned}\operatorname{var}\left(N_T\left(m\right)\right)\cong&\sum_{i=1}^{n}\sum_{j\neq i}^{n}T_{ij}{}^2\times\left[\frac{\partial P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\overline{\mathbf{x}}_{ij}}}{\partial\mathbf{x}_{ij}{}'\boldsymbol{\beta}}\right]^2\left(\overline{\mathbf{x}}_{ij}{}'\mathbf{A}\overline{\mathbf{x}}_{ij}+\boldsymbol{\beta}\mathbf{Y}_{ij}\boldsymbol{\beta}'\right)\\&+\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{i'=1}^{n}\sum_{\substack{j'\neq i'\\ j'\neq j}}^{n}T_{ij}\times T_{i'j'}\times\left[\frac{\partial P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\overline{\mathbf{x}}_{ij}}}{\partial\mathbf{x}_{ij}{}'\boldsymbol{\beta}}\right]\left[\frac{\partial P\left(m\middle|\mathbf{x}_{i'j'}\right)\Big|_{\overline{\mathbf{x}}_{i'j'}}}{\partial\mathbf{x}_{i'j'}{}'\boldsymbol{\beta}}\right]\left(\overline{\mathbf{x}}_{ij}{}'\mathbf{A}\overline{\mathbf{x}}_{ij}+\boldsymbol{\beta}\mathbf{Y}_{ij,i'j'}\boldsymbol{\beta}'\right)\end{aligned}\tag{5.26}$$

As addressed in (Koppelman 1975a), Equation (5.26) may be further simplified to the following equation since the trips with the same trip purposes are usually attributed with similar derivatives, equal variance among trips, and equal covariance among pairs of trips and also most of the parameter estimates in $Y_{ij,i'j'}$ are the same as those in Y_{ij}:

$$\begin{aligned}\operatorname{var}\left(N_T\left(m\right)\right)\cong&\sum_{i=1}^{n}\sum_{j\neq i}^{n}T_{ij}{}^2\times\left[\frac{\partial P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\overline{\mathbf{x}}_{ij}}}{\partial\mathbf{x}_{ij}{}'\boldsymbol{\beta}}\right]^2\left(\overline{\mathbf{x}}_{ij}{}'\mathbf{A}\overline{\mathbf{x}}_{ij}+\boldsymbol{\beta}\mathbf{Y}_{ij}\boldsymbol{\beta}'\right)\\&+\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{i'=1}^{n}\sum_{\substack{j'\neq i'\\ j'\neq j}}^{n}T_{ij}\times T_{i'j'}\times\left[\frac{\partial P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\overline{\mathbf{x}}_{ij}}}{\partial\mathbf{x}_{ij}{}'\boldsymbol{\beta}}\right]^2\left(\overline{\mathbf{x}}_{ij}{}'\mathbf{A}\overline{\mathbf{x}}_{ij}+\boldsymbol{\beta}\mathbf{Y}_{ij}\boldsymbol{\beta}'\right)\end{aligned}\tag{5.27}$$

Denote $T_{\bullet\bullet}$ as the total row and column sum of person trips for every zone pair in study area T by excluding the intrazonal trips on the diagonal of the trip table for a given purpose. Equation (5.27) becomes:

$$\begin{aligned}\operatorname{var}\left(N_T\left(m\right)\right)\cong&\sum_{i=1}^{n}\sum_{j\neq i}^{n}T_{ij}{}^2\times\left[\frac{\partial P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\overline{\mathbf{x}}_{ij}}}{\partial\mathbf{x}_{ij}{}'\boldsymbol{\beta}}\right]^2\left(\overline{\mathbf{x}}_{ij}{}'\mathbf{A}\overline{\mathbf{x}}_{ij}+\boldsymbol{\beta}\mathbf{Y}_{ij}\boldsymbol{\beta}'\right)\\&+\sum_{i=1}^{n}\sum_{j\neq i}^{n}T_{ij}\times\left(T_{\bullet\bullet}-T_{ij}\right)\times\left[\frac{\partial P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\overline{\mathbf{x}}_{ij}}}{\partial\mathbf{x}_{ij}{}'\boldsymbol{\beta}}\right]^2\left(\overline{\mathbf{x}}_{ij}{}'\mathbf{A}\overline{\mathbf{x}}_{ij}+\boldsymbol{\beta}\mathbf{Y}_{ij}\boldsymbol{\beta}'\right)\\=&\,T_{\bullet\bullet}\times\sum_{i=1}^{n}\sum_{j\neq i}^{n}T_{ij}\times\left[\frac{\partial P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\overline{\mathbf{x}}_{ij}}}{\partial\mathbf{x}_{ij}{}'\boldsymbol{\beta}}\right]^2\left(\overline{\mathbf{x}}_{ij}{}'\mathbf{A}\overline{\mathbf{x}}_{ij}+\boldsymbol{\beta}\mathbf{Y}_{ij}\boldsymbol{\beta}'\right)\end{aligned}\tag{5.28}$$

As shown in Equation (5.28), the errors in parameters and errors in variables contribute to the variance of total trips for a given mode. Again, since the true $P(m|\mathbf{x}_{ij})$ and other parameters in Equation (5.28) are unknown to the modelers, they may be replaced with the variables estimated from a calibrated MNL model. Equation (5.28) thus becomes

$$\operatorname{var}\left(N_T(m)\right) \cong T_{\bullet\bullet} \times \sum_{i=1}^{n}\sum_{j\neq i}^{n} T_{ij} \times \left[\frac{\partial \hat{P}\left(m\left|\mathbf{x}_{ij}\right.\right)\Big|_{\bar{\mathbf{x}}_{ij}}}{\partial \mathbf{x}_{ij}{}'\boldsymbol{\beta}}\right]^2 \left(\bar{\mathbf{x}}_{ij}{}'\hat{\mathbf{A}}\bar{\mathbf{x}}_{ij} + \hat{\boldsymbol{\beta}}\hat{\mathbf{Y}}_{ij}\hat{\boldsymbol{\beta}}'\right) \tag{5.29}$$

where

$\hat{P}\left(m\left|\mathbf{x}_{ij}\right.\right)\Big|_{\bar{\mathbf{x}}_{ij}} = \dfrac{e^{\bar{\mathbf{x}}_{ij}{}'\hat{\boldsymbol{\beta}}}}{\sum_{j'=1}^{n} e^{\bar{\mathbf{x}}_{ij'}{}'\hat{\boldsymbol{\beta}}}}$. if MNL model is chosen; or takes on the NL model form

$\hat{\mathbf{A}} = \operatorname{var}\left(\hat{\boldsymbol{\beta}}\right)$; and

$\hat{\mathbf{Y}}_{ij} = \operatorname{var}\left(\hat{\boldsymbol{\delta}}_{ij}\right)$.

Equation (5.29) reveals the following characteristics in quantifying the variance for the total number of trips allocating to mode m:

- The variance is positively related to the total number of trips for a given purpose estimated in study area T.
- The variance is positively related to the number of trips interchanging between a given zone pair.
- The variance is positively related to the estimated probability for trips between a given zone pair that are allocated to mode m.
- The variance is positively related to the sum of $\bar{\mathbf{x}}_{ij}{}'\hat{\mathbf{A}}\bar{\mathbf{x}}_{ij}$ and $\hat{\boldsymbol{\beta}}\hat{\mathbf{Y}}_{ij}\hat{\boldsymbol{\beta}}'$, which in terms may increase or reduce the variance in aggregate prediction.

Since $\hat{\mathbf{A}}$, i.e., $\operatorname{var}\left(\hat{\boldsymbol{\beta}}\right)$, contains the maximum likelihood estimates calibrated with the records collected at the individual level, the Cramer-Rao Theorem may be applied to quantify the lower limits associated with the variance-covariance matrix as follows (Ben-Akiva and Lerman 1985):

$$\operatorname{var}\left(\hat{\boldsymbol{\beta}}\right) \geq \mathbf{B}^{-1} \tag{5.30}$$

and

$$\mathbf{B}=\begin{bmatrix} -E\left[\frac{\partial^2 L}{\partial\beta_1^2}\right] & -E\left[\frac{\partial^2 L}{\partial\beta_1\partial\beta_2}\right] & \cdots & -E\left[\frac{\partial^2 L}{\partial\beta_1\partial\beta_k}\right] \\ -E\left[\frac{\partial^2 L}{\partial\beta_2\partial\beta_1}\right] & -E\left[\frac{\partial^2 L}{\partial\beta_2^2}\right] & \cdots & -E\left[\frac{\partial^2 L}{\partial\beta_2\partial\beta_k}\right] \\ \vdots & \vdots & \ddots & \vdots \\ -E\left[\frac{\partial^2 L}{\partial\beta_k\partial\beta_1}\right] & -E\left[\frac{\partial^2 L}{\partial\beta_k\partial\beta_2}\right] & \cdots & -E\left[\frac{\partial^2 L}{\partial\beta_k^2}\right] \end{bmatrix}; \tag{5.31}$$

where

N = sample size;

Y_{in} = 1 if observation n chose alternative i; 0 otherwise;

C_n = choice set for observation n; and

$$L = \sum_{n=1}^{N}\sum_{i\in C_n} y_{in}\left(\mathbf{x}_{in}'\hat{\boldsymbol{\beta}} - \ln\sum_{j\in C_n} e^{\mathbf{x}_{in}'\hat{\boldsymbol{\beta}}}\right). \text{ if the MNL model is chosen} \tag{5.32}$$

Again, since the true model parameters are unknown, Equation (5.31) may be approximated as follows:

$$\hat{\mathbf{B}}=\begin{bmatrix} -E\left[\frac{\partial^2 L}{\partial\hat\beta_1^2}\right] & -E\left[\frac{\partial^2 L}{\partial\hat\beta_1\partial\hat\beta_2}\right] & \cdots & -E\left[\frac{\partial^2 L}{\partial\hat\beta_1\partial\hat\beta_k}\right] \\ -E\left[\frac{\partial^2 L}{\partial\hat\beta_2\partial\hat\beta_1}\right] & -E\left[\frac{\partial^2 L}{\partial\hat\beta_2^2}\right] & \cdots & -E\left[\frac{\partial^2 L}{\partial\hat\beta_2\partial\hat\beta_k}\right] \\ \vdots & \vdots & \ddots & \vdots \\ -E\left[\frac{\partial^2 L}{\partial\hat\beta_k\partial\hat\beta_1}\right] & -E\left[\frac{\partial^2 L}{\partial\hat\beta_k\partial\hat\beta_2}\right] & \cdots & -E\left[\frac{\partial^2 L}{\partial\hat\beta_k^2}\right] \end{bmatrix} \tag{5.33}$$

where

$$\frac{\partial^2 L}{\partial\hat\beta_k\partial\hat\beta_l} = -\sum_{n=1}^{N}\sum_{i\in C_n}\hat{P}_n(i)\left[x_{ink}-\sum_{j\in C_n}x_{jnk}\hat{P}_n(j)\right]\left[x_{inl}-\sum_{j\in C_n}x_{jnl}\hat{P}_n(j)\right]; \text{ and} \tag{5.34}$$

$$\hat{P}_n(i) = \frac{e^{\mathbf{x}_{in}'\hat{\boldsymbol{\beta}}}}{\sum_{j\in C_n}^{n} e^{\mathbf{x}_{jn}'\hat{\boldsymbol{\beta}}}}. \text{ if the MNL model is chosen.}$$

Summary

As illustrated in Equation (5.16), errors in the aggregate share estimates in modal split are different from errors in the estimation of model parameters ($\boldsymbol{\beta}$) or in the predictions of model variables ($\mathbf{x}_{ij}$) since aggregation errors are deterministic and do not introduce random errors when the true probability of selecting a given mode, $P\left(m\middle|\mathbf{x}_{ij}\right)$, is given. Since $P\left(m\middle|\mathbf{x}_{ij}\right)$ is unknown to the modelers, area-wide total aggregation error may be approximately estimated when $P\left(m\middle|\mathbf{x}_{ij}\right)\Big|_{\bar{\mathbf{x}}_{ij}}$ and the covariance term, $\mathrm{cov}\left(x_{ijr}, x_{ijr'}\right)$, for every zone pair are available. Equation (5.29) may be considered as the variance of Err_m in Equations (5.12) and (5.16) since no variance is associated with the $E\left(P\left(m\middle|\mathbf{x}_{ij}\right)\right)$ term. Consequently, Equation (5.16) may be applied to estimate the expected discrepancy in number of trips choosing mode m. The Cramer-Rao Theorem may then be implemented to estimate the lower limit of $\mathrm{var}\left(\hat{\boldsymbol{\beta}}\right)$ and subsequently incorporated to estimate the variance of $N_T(m)$, given that the covariance term, i.e., $\hat{\mathbf{Y}}_{ij}$ or $\mathrm{var}\left(\hat{\boldsymbol{\delta}}_{ij}\right)$, are known for every zone pair.

While the modal split model and the mean values for the associated model variables may be precisely estimated, the information for the covariance terms between the model variables cannot be obtained without devoting significant effort in data collection.

5.4 Statistical Variability of Transit Ridership

The travel demand model must be calibrated and validated against the observation measures before they can be used to provide meaningful results. The observation measures to be used prefer a data source that is different from the calibration data set, if and as available. Otherwise, the calibration data can be used to determine the accuracy of the model estimates. The variability of these observed values have been generally poorly identified or ignored. Without explicit and rigorous statistical recognition of variation in the transit ridership data, the process of validation may lead to erroneous corrections to models and deteriorate the performance of model.

In almost all model validation efforts, only mean measures of interest regardless their variability are taken into account. The validation process on the basis of mean measures constrains the comparison between point model forecasts and point observed values collected in the field. Such practice may underestimate the overall error between the modeled and observed transit trips and produce biased estimates. Furthermore, overlooking data variation may prevent alternative demand estimations that are statistically insignificant from being revealed. This section addresses the statistical quantification of the variability in the observed transit patronages.

The validation process of travel demand model conventionally focuses only on the highway volumes. The various FSUTMS standard models simulate peak season trip productions and attractions from zonal distributions of residential, employment, and socio-economic input data. FSUTMS traffic assignment volumes represent Peak Season Weekday Average Daily Traffic (PSWADT) projections for the roads represented in the modeled highway network. The peak season is the 13 consecutive weeks of the year during which the highest weekday volumes occur. When the forecasted link-flow estimates are evaluated for reasonableness, the ground counts must be converted from Average Daily Traffic (ADT) to PSWADT before comparing with the model trip assignments.

Although a guideline for converting traffic counts to the counterparts for model validation has been well established for the highway side, no procedure for the transit side has been documented. The observed transit ridership used for validating the transit model differs from agency to agency. For example, the Tampa regional model uses weekday average transit ridership of a single month (March in this case) as the peak season figures for validation. To be consistent with the procedure used for traffic volumes on highway networks, three consecutive months with highest weekday ridership were selected for variation computation.

The following three studies areas in Florida with transit systems were selected for the investigation of transit ridership variation: Jacksonville, Miami, and Tampa. The transit ridership data were estimated with the farebox system by individual transit agencies. This counting technique has been reported to be useful in determining system-wide ridership (McKnight *et al.*, 2000). The estimated daily patronages of each route were averaged through a period of one month for weekdays and weekends, respectively. Even though the information is normally verified by comparing the results to revenue, the accuracy tests should be further performed with other methods. For instance, the comparison between farebox estimates and recording by manual checkers can be conducted. Further research into the ridership variation is recommended to test the accuracy with the farebox method.

The distribution of transit ridership data by route was examined before computing the errors. Some ridership data lie apparently outside the overall pattern of the distribution for a specific route. The data contain outliers that may not be observations from the same population. The outliers in the weekday average ridership may point to special events, holidays, or an error in data entries. The outliers stemming from these abnormal conditions should not be included in the analysis of data variation. In addition, the sample mean and sample standard deviation are sensitive to the influence of outliers. The t distribution used here is established on the basis of these two statistical measures. Consequently, the transit routes were eliminated from statistical analysis with a clear conscience.

The monthly weekday average riderships over a 12 month period were used in this study. To draw a steady result to reflect the current operation circumstance, the data were selected not only for the transit ridership of the latest period of time, but also the reliability of data itself. Consequently, the time coverage for each study area is different from one to another. After screening route data with missing records or with extraordinary condition, the Jacksonville system consists of 43 routes, the Miami system consists of 87 routes, and the Tampa system consists of 48routes. The transit ridership data for three study areas are given in Tables 5.5-5.7.

By comparing aggregated values of consecutive three months of all possible combination, the peak three months for each route and system-wide were selected and shown in shared cells of the tables. The three-month variation was then computed by route and system-wide individually. Since the assumption that the population distribution is normal cannot be sufficiently verified with only a few observations, the t instead of z distribution was used to approximate normal distribution. The errors resulting from the observed transit ridership variability is computed as follows:

$$d = \frac{t_{\alpha/2} \times s / \overline{x}}{\sqrt{n}} \times 100 = \frac{t_{\alpha/2} \times C_v}{\sqrt{n}} \times 100 \tag{5.35}$$

where

n = sample size of weekday average ridership;
s = sample standard deviation;
$\overline{x}$ = mean of weekday average riderships;
$t_{\alpha/2}$ = t statistic at significance level α with $n - 1$ degree of freedom;
C_v = coefficient of variation; and
d = error in percent of the weekday average ridership of peak-three months.

A sample size of 3 corresponds to the three highest weekday ridership. Three confidence levels (CLs) were selected in this study to estimate the respective errors. Given confidence level ($1 - \alpha$) and the degree of freedom ($n - 1$), the t can be obtained from the t distribution table. The t statistics of each level of confidence are 1.321 for 68% confidence level, 2.403 for 85%, and 4.303 for 95%. Based on the transit ridership data obtained, the sample means, the coefficients of variation, and the errors estimated at the three levels of confidence are summarized in Tables 5.8-5.11. The coefficients of variation for each route were plotted against the corresponding weekday average ridership to analyze the relationship. As expected, the relationship between the two variables reveals a decreasing function as shown in Figures 5.1-5.3 for the three urban areas.

Table 5.5 Weekday Average Transit Ridership of Jacksonville for Years 2003-2004

Route	Oct-03	Nov-03	Dec-03	Jan-04	Feb-04	Mar-04	Apr-04	May-04	Jun-04	Jul-04	Aug-04	Sep-04
B6	410	397	375	397	362	385	396	393	395	412	442	426
B7	1,285	1,229	1,159	1,208	1,159	1,166	1,145	1,163	1,181	1,174	1,224	1,154
B9	500	480	428	479	477	492	477	437	449	445	493	465
E2	1,218	1,166	1,146	1,217	1,180	1,171	1,163	1,126	1,180	1,218	1,204	1,121
E5	439	398	410	403	399	396	425	404	455	434	412	348
F1	805	754	662	695	700	740	729	701	704	710	788	734
H2	213	221	193	192	177	174	181	194	220	209	207	189
I6	1,570	1,564	1,434	1,513	1,570	1,609	1,600	1,569	1,657	1,655	1,712	1,616
J1	1,425	1,312	1,210	1,343	1,302	1,330	1,293	1,261	1,319	1,295	1,301	1,268
K1	1,395	1,358	1,258	1,319	1,332	1,301	1,317	1,284	1,327	1,289	1,349	1,322
K2	1,710	1,595	1,472	1,568	1,578	1,640	1,622	1,488	1,547	1,536	1,776	1,739
L7	1,403	1,435	1,352	1,422	1,423	1,425	1,448	1,402	1,449	1,463	1,464	1,386
L8	2,106	2,071	2,025	2,036	2,072	2,094	2,007	1,880	1,914	1,916	1,995	1,861
M4	1,079	1,075	1,072	1,117	1,068	1,098	1,089	1,118	1,126	1,169	1,216	1,137
M5	1,377	1,366	1,224	1,311	1,324	1,349	1,395	1,280	1,327	1,329	1,388	1,328
WS 38	30	29	27	22	28	32	26	20	26	24	22	19
N6	877	817	721	770	779	786	765	753	746	738	822	776
O1	357	355	343	354	361	374	379	373	410	377	383	374
P2	571	553	524	588	597	610	633	580	568	566	621	578
P3	512	493	469	479	440	445	457	465	496	465	472	454
P4	1,520	1,466	1,431	1,498	1,453	1,467	1,469	1,419	1,397	1,406	1,513	1,420
P7	1,578	1,413	1,328	1,417	1,402	1,373	1,362	1,359	1,416	1,422	1,457	1,438
Q3	322	315	289	308	317	317	337	319	305	311	305	286
Q4	208	183	184	195	157	156	142	159	175	154	155	152
R1	1,834	1,867	1,739	1,856	1,830	1,892	1,968	1,758	1,807	1,787	1,723	1,678
R4	222	206	197	202	194	201	212	200	206	214	207	215
R5	1,163	1,155	1,029	1,122	1,209	1,183	1,204	1,087	1,165	1,162	1,148	1,175
S1	460	427	382	428	406	438	433	423	468	484	446	431
NS14	422	407	382	416	426	424	427	400	433	423	474	472
NS20	477	473	418	430	430	433	416	398	404	431	444	406
NS21	249	228	228	245	248	240	232	208	187	196	223	207

Table 5.5 Weekday Average Transit Ridership of Jacksonville for Years 2003-2004 (Continued)

Route	Oct-03	Nov-03	Dec-03	Jan-04	Feb-04	Mar-04	Apr-04	May-04	Jun-04	Jul-04	Aug-04	Sep-04
NS22	144	142	146	142	148	161	154	140	140	148	157	135
X2	31	25	27	32	34	37	35	33	29	28	29	26
X4	69	58	52	62	60	58	55	60	63	64	67	58
AR20	443	434	451	432	427	437	436	427	453	452	431	435
SS 3	510	499	462	499	486	495	491	461	477	485	499	468
SS 5	377	369	337	369	362	350	361	339	328	321	369	346
SS35	43	40	36	40	47	50	48	47	45	43	48	49
TR 6	618	659	625	622	632	651	684	714	735	717	770	687
TR 7	296	305	337	377	267	267	276	288	511	460	276	316
IL 9	880	807	742	778	794	848	843	822	868	877	935	874
FADJ	85	83	71	87	81	82	84	87	100	94	114	100
APT	246	313	213	212	156	224	260	264	278	292	263	305
TOTAL	31,479	30,542	28,610	30,202	29,894	30,401	30,476	29,303	30,486	30,395	31,344	29,974

Table 5.6 Weekday Average Transit Ridership of Miami for Years 2002-2003

Route	Jul-02	Aug-02	Sep-02	Oct-02	Nov-02	Dec-02	Jan-03	Feb-03	Mar-03	Apr-03	May-03	Jun-03
A	362	413	361	339	458	466	410	451	542	393	357	323
B	1,427	1,472	1,408	1,397	1,456	1,402	1,369	1,480	1,635	1,587	1,499	1,388
C	3,202	3,255	3,433	3,434	3,289	3,118	3,457	3,762	3,867	3,464	3,414	3,477
E	928	929	898	994	929	838	923	1,148	1,175	1,126	1,176	1,070
G	2,940	3,003	3,185	3,205	2,864	2,447	2,685	2,958	3,019	2,953	2,944	2,796
H	4,140	4,284	4,081	4,463	4,353	4,538	4,827	4,912	4,792	4,313	4,402	4,065
J	4,457	4,533	4,658	4,776	4,826	4,507	4,500	4,497	4,486	4,268	4,118	3,902
K	4,073	4,080	4,374	4,518	4,315	4,222	4,302	4,669	4,541	4,154	4,157	3,963
L	10,057	10,090	10,028	10,191	10,231	9,816	10,039	10,511	10,324	10,127	9,876	9,391
M	1,860	1,907	1,818	1,957	1,851	1,799	1,818	2,050	1,909	1,799	1,790	1,717
R	329	320	334	361	350	306	327	339	455	525	415	419
S	12,745	12,679	12,214	12,068	12,062	11,811	12,203	12,265	12,700	11,777	11,412	11,442
T	1,990	2,048	2,057	2,128	2,092	2,160	2,173	2,367	2,383	2,304	2,115	2,108
V	259	270	297	282	278	285	325	413	403	382	386	344
W	359	373	413	380	400	395	310	353	260	226	196	258
1	1,404	1,308	1,409	1,456	1,532	1,317	1,462	1,584	2,400	1,439	1,400	1,380
2	3,428	3,651	3,983	3,951	3,881	3,495	3,540	3,910	3,747	3,488	3,678	3,218
3/16	12,457	12,496	13,050	13,323	12,835	12,731	12,565	12,670	12,761	12,594	12,076	11,916
6	277	246	266	298	277	248	226	262	325	239	309	249
7	3,489	3,568	3,324	3,611	3,664	3,474	3,679	3,805	3,761	3,754	3,652	3,743
8	7,482	7,480	7,440	7,698	7,498	6,706	6,853	7,276	7,118	7,034	7,286	7,527
9	4,593	4,648	5,244	5,217	4,892	4,600	4,610	5,255	5,150	5,082	4,863	4,670
10	2,316	2,436	2,477	2,602	2,543	2,388	2,461	2,553	2,548	2,318	2,374	2,304
11	12,009	12,206	12,698	12,536	12,256	10,938	11,147	12,280	12,088	11,561	11,741	11,397
12	3,107	3,137	3,241	3,258	3,192	2,846	2,987	3,316	3,251	3,073	3,109	2,937
17	4,789	4,891	5,690	5,803	5,646	4,945	5,365	5,798	5,346	4,911	5,089	4,479
21	2,210	2,306	2,562	2,525	2,370	2,159	2,226	2,355	2,352	2,370	2,329	2,192
22	3,714	3,732	4,091	4,222	3,852	3,511	3,654	3,726	3,769	3,598	3,872	3,578
24	4,016	3,741	4,003	3,938	4,178	3,959	4,077	4,537	4,106	3,966	3,970	3,748
27	7,349	8,114	8,731	8,612	8,864	8,203	8,662	9,301	8,710	8,597	8,807	8,237

Table 5.6 Weekday Average Transit Ridership of Miami for Years 2002-2003 (Continued)

Route	Jul-02	Aug-02	Sep-02	Oct-02	Nov-02	Dec-02	Jan-03	Feb-03	Mar-03	Apr-03	May-03	Jun-03
28	600	658	777	746	784	629	700	789	805	785	923	789
29	288	278	300	292	259	225	279	288	362	296	253	227
31	1,005	1,049	1,108	1,162	1,208	1,230	1,136	1,202	1,058	1,030	986	921
32	3,350	3,367	3,848	4,064	3,715	3,288	3,574	4,030	3,854	3,727	3,586	3,595
33	1,665	1,676	1,933	2,010	2,113	2,069	1,996	2,221	2,227	2,147	2,169	1,980
35/70	1,802	1,813	2,103	2,154	2,116	1,904	2,240	2,413	2,376	2,467	2,889	2,678
36	3,255	3,191	3,289	3,385	3,215	2,824	3,071	3,216	3,218	3,118	3,165	2,985
37	3,520	3,551	3,542	3,631	3,509	3,445	3,384	3,569	3,536	3,262	3,454	3,369
38	3,437	3,623	3,529	3,653	3,549	3,605	3,807	3,559	3,538	3,935	3,652	3,721
40	1,996	1,994	2,069	2,247	2,148	2,038	2,104	2,305	2,173	2,023	2,099	2,054
42	964	970	974	1,039	995	853	965	1,140	1,199	1,203	1,309	1,288
48	461	508	483	575	424	395	372	377	315	273	301	315
51 FLAGLER MAX	1,615	1,620	1,727	1,448	2,273	2,356	2,534	2,691	2,908	3,062	3,172	2,699
52	1,247	1,234	1,356	1,423	1,075	921	1,072	1,130	1,053	1,065	1,366	1,334
54	3,106	3,230	3,296	3,340	3,068	2,682	2,856	3,055	2,958	2,675	2,657	2,681
56	1,175	1,168	1,066	1,262	1,463	1,471	1,575	1,530	1,614	1,467	1,106	1,015
57/72	1,695	1,666	1,730	1,788	1,754	1,658	1,687	1,901	1,898	1,565	1,340	1,871
62	4,648	4,481	4,970	5,094	5,143	4,727	4,932	5,599	5,229	4,929	5,309	4,572
65EX	203	115	108	242	310	231	144	353	418	405	352	319
71	1,180	1,163	1,324	1,369	1,331	1,214	1,348	1,366	1,329	1,342	1,287	1,286
73	1,985	2,000	2,122	2,164	2,114	1,797	1,942	2,203	2,029	1,995	2,026	1,874
75	3,218	3,197	4,325	4,308	4,114	3,293	3,729	4,115	3,923	3,638	3,831	3,252
77	8,997	9,324	9,831	9,977	9,870	8,904	9,369	9,986	10,012	9,895	9,952	9,413
83	3,418	3,477	4,222	4,317	4,265	3,749	4,208	4,344	4,151	4,037	4,076	3,637
87	1,294	1,348	1,410	1,474	1,560	1,528	1,634	1,748	1,743	1,683	1,790	1,623
88	2,512	2,494	2,587	2,636	2,736	2,571	2,601	2,676	2,646	2,512	2,637	2,584
91	1,040	1,095	1,276	1,276	1,279	1,069	1,226	1,445	1,421	1,336	1,404	1,196
BISC MAX 1	1,823	1,842	1,970	2,030	2,314	2,255	2,232	2,290	2,267	2,189	2,137	2,095
95EX	1,488	1,508	1,565	1,659	1,646	1,669	1,754	1,745	1,744	1,679	1,689	1,561
27 MAX	597	569	741	759	727	613	649	646	667	653	594	573

Table 5.6 Weekday Average Transit Ridership of Miami for Years 2002-2003 (Continued)

Route	Jul-02	Aug-02	Sep-02	Oct-02	Nov-02	Dec-02	Jan-03	Feb-03	Mar-03	Apr-03	May-03	Jun-03
104	1,380	1,462	1,642	1,771	1,560	1,311	1,451	1,611	1,538	1,205	1,335	1,165
TR-36 ST	52	75	52	52	41	48	40	49	38	27	33	45
TR-MIA	416	224	223	212	307	492	544	460	474	344	452	570
134- RIVERSIDE SHUTTLE	37	35	40	38	39	37	39	39	44	46	36	36
137-WEST DADE CONN	1,265	1,207	1,165	1,207	1,251	1,131	1,199	1,226	1,586	1,127	1,021	942
231-BUSWAY LOCAL	760	887	950	1,042	750	725	949	974	845	568	580	551
236-A RPORT OWL	336	315	295	293	294	351	335	371	368	349	299	236
237 DOUGLAS BRIDGE	38	53	62	37	38	28	39	42	43	46	32	33
238 EAST/WEST CONN	479	495	493	547	485	447	409	404	384	487	552	567
240-BIRD MAX	425	485	530	525	525	530	490	499	337	387	503	465
241-NORTH DADE CONN	307	305	353	341	340	340	301	298	372	357	296	270
242-DORAL CONN	430	479	456	545	433	395	369	400	250	328	358	422
243-SEAPORT CONN	212	228	164	184	219	192	180	169	83	232	160	232
245-OKEECHOBEE CONN	198	228	315	257	240	225	136	131	143	214	172	285
246-NIGHT OWL	494	441	323	393	518	437	373	336	627	566	346	316
248-BRICKELL KEY SHUTTLE	278	367	333	392	367	384	333	439	345	749	645	774
252-CORAL REEF	922	915	920	821	855	835	954	897	987	844	953	959
267-LUDLUM MAX	367	414	409	374	401	416	404	469	358	354	367	395
287-SAGA BAY	272	225	255	249	239	264	344	320	270	325	278	261
KAT-KILLIAN	1,017	1,154	1,257	1,032	948	1,306	1,055	1,086	756	1,164	1,084	1,307
KAT-SUNSET	811	813	858	729	847	927	629	654	788	904	697	815
KAT-KENDALL	585	612	582	634	576	627	662	690	543	492	545	531
GREEN HILLS SHUTTLE	5	8	9	8	10	7	29	22	13	19	13	12
KINGS CREEK SHUTTLE	3	10	9	7	9	8	9	10	9	12	10	10
SIERRA LAKES SHUTTLE	5	8	10	11	5	7	7	11	10	12	11	10
ROBERT SHARPE SHUTTLE	6	8	9	8	14	7	8	8	9	10	9	9
TOTAL	196,452	199,278	209,063	212,776	209,622	196,320	203,590	216,350	213,814	204,983	205,110	196,963

Table 5.7 Weekday Average Transit Ridership of Tampa for Years 2003-2004

Route	Oct-03	Nov-03	Dec-03	Jan-04	Feb-04	Mar-04	Apr-04	May-04	Jun-04	Jul-04	Aug-04	Sep-04
1	2,656	2,642	2,534	2,640	2,440	2,452	2,266	2,313	2,273	2,243	2,419	2,523
2	3,223	3,266	3,086	3,358	3,331	3,326	3,348	3,415	3,402	3,342	3,432	3,493
4	441	432	394	403	393	405	379	387	387	372	393	388
5	1,213	1,216	987	1,150	1,192	1,158	1,135	1,121	1,134	1,160	1,336	1,354
6	1,912	2,046	2,013	2,284	2,190	2,064	2,127	2,169	2,049	2,100	2,325	2,282
7	1,958	1,944	1,827	2,110	1,886	1,828	1,898	1,906	1,836	1,774	1,937	1,843
8	1,110	1,098	1,042	1,136	1,024	975	899	854	915	846	898	904
9	1,463	1,435	1,330	1,392	1,335	1,327	1,314	1,403	1,382	1,358	1,600	1,494
10	782	800	688	764	802	791	434	379	387	373	407	394
11	581	575	488	555	541	525	617	600	607	567	597	615
12	2,151	2,090	1,966	2,203	2,192	2,106	2,120	2,139	2,124	2,133	2,306	2,104
14	509	500	439	487	478	456	443	449	439	457	525	514
15	735	710	666	739	788	715	728	741	803	799	855	864
16	447	457	444	468	459	434	434	424	420	432	489	498
17	484	477	429	445	434	424	428	441	417	388	402	413
18	1,315	1,222	1,098	1,210	1,148	1,129	1,175	1,175	1,191	1,166	1,220	1,203
19	1,090	1,156	1,062	1,241	1,065	976	960	976	959	999	1,033	1,017
30	1,261	1,257	1,186	1,239	1,212	1,208	1,313	1,369	1,330	1,387	1,432	1,444
31	283	297	291	287	292	288	254	240	240	246	232	238
32	911	925	861	940	948	908	907	932	884	878	977	983
33	353	355	335	354	351	360	374	394	381	376	435	423
34	1,444	1,425	1,367	1,496	1,514	1,490	1,513	1,629	1,547	1,435	1,538	1,511
36	1,130	1,132	1,071	1,117	1,106	1,100	1,073	1,057	1,003	997	1,047	1,057
37	449	490	417	466	500	500	514	501	450	419	571	599
39	1,077	1,086	1,000	1,057	1,021	1,054	1,006	936	906	892	927	884
41	337	298	280	289	297	240	264	254	204	222	345	354
46	107	115	114	118	103	98	164	190	185	189	226	244
70	60	57	77	69	56	45	36	39	52	54	51	48
71	66	61	78	77	54	48	54	48	66	60	50	48
72	55	52	62	66	38	37	32	28	65	68	39	36

Table 5.7 Weekday Average Transit Ridership of Tampa for Years 2003-2004 (Continued)

Route	Oct-03	Nov-03	Dec-03	Jan-04	Feb-04	Mar-04	Apr-04	May-04	Jun-04	Jul-04	Aug-04	Sep-04
73	32	31	40	34	14	26	19	24	31	37	24	31
81	24	24	23	24	23	29	54	55	54	55	58	55
83	543	529	499	513	521	495	435	446	467	492	522	561
84	124	122	130	132	122	133	120	131	146	147	139	135
88	101	93	93	84	91	95	83	89	87	72	71	69
96	581	492	436	435	519	526	557	436	527	581	504	540
20X	45	52	43	46	48	51	50	53	55	51	58	47
22X	67	67	64	74	69	63	67	67	65	59	64	57
23X	45	44	37	41	38	39	38	43	38	37	38	33
25X	72	62	56	79	81	82	74	74	78	71	88	71
26X	28	27	21	23	23	22	26	26	22	21	24	21
27X	74	71	55	78	76	68	68	80	76	74	82	64
28X	73	62	56	63	61	58	54	67	69	67	61	56
50X	29	31	27	33	30	34	30	34	33	29	27	29
200X	65	65	61	63	52	60	60	54	50	47	48	49
57LX	156	174	154	171	164	153	160	166	149	141	173	168
58LX	107	95	98	105	107	103	91	98	98	94	102	106
TOTAL	31,768	31,657	29,523	32,155	31,224	30,504	30,192	30,453	30,084	29,807	32,130	31,864

According to the updated transit validation standards for FDOT, the acceptable error range for total area transit trips is ±1% as given in Chapter 3. Table 5.8 shows that none of the error percentages stemming from ridership variability of any study areas meet the accuracy standard, irregardless of the levels of confidence selected. On the other hand, the error percentages by route were compared against the allowable errors defined on the basis of daily riders of each route. Tables 5.9-5.11 show that all the error percentages by route are far below the accuracy standard, with one exception: Route 1 of the Miami system fails to meet the standard at the 95% confidence level. To demonstrate the discrepancy between the derived error percentages and FDOT accuracy standard, the error percentages were categorized on the specified range of daily ridership as in Table 5.12. It is noted that the derived error percentages are rounded to the integer.

Table 5.8 System-Wide Errors Percentage Resulting from Transit Ridership Variability

Study Area	Weekday Average Ridership	Cv	Error (%)		
			68% CL	85% CL	95% CL
Jacksonville	30,742	0.0395	1.30	2.36	4.23
Miami	211,716	0.0282	2.15	3.91	7.00
Tampa	31,639	0.0170	2.02	3.67	6.57

Because the weekday transit ridership data are averaged within a period of one month before being used for computing error percentages statistically, the estimated values reflect only the variation of weekday average transit ridership by month. Small sample sizes could be a problem as they contribute to a larger error due to the relative larger values of $t_{\alpha/2}$ and $s/\sqrt{n}$. Further examination of variability for weekday ridership is necessary to determine the true error percentage arising from the daily variation of transit ridership.

Table 5.9 Route Errors Percentage Resulting from Transit Ridership Variability in Jacksonville

Route	Weekday Average Ridership	Cv	Error (%)		
			68% CL	85% CL	95% CL
B6	427	0.0352	2.68	4.88	8.74
B7	1,224	0.0516	3.93	7.15	12.81
B9	483	0.0169	1.29	2.34	4.19
E2	1,201	0.0160	1.22	2.22	3.98
E5	434	0.0496	3.78	6.88	12.32
F1	744	0.0537	4.10	7.45	13.34
H2	212	0.0330	2.52	4.58	8.20
I6	1,675	0.0193	1.47	2.68	4.80
J1	1,325	0.0158	1.21	2.19	3.93
K1	1,337	0.0530	4.04	7.35	13.17
K2	1,684	0.0767	5.85	10.65	19.07
L7	1,459	0.0057	0.44	0.80	1.43
L8	2,067	0.0197	1.50	2.73	4.88
M4	1,174	0.0338	2.58	4.70	8.41
M5	1,356	0.0266	2.03	3.68	6.60
WS 38	29	0.0533	4.06	7.39	13.24
N6	805	0.0978	7.46	13.56	24.28
O1	390	0.0451	3.44	6.25	11.20
P2	613	0.0297	2.27	4.12	7.38
P3	491	0.0439	3.34	6.08	10.90
P4	1,473	0.0156	1.19	2.17	3.88
P7	1,440	0.0883	6.73	12.25	21.94
Q3	324	0.0340	2.59	4.71	8.44
Q4	192	0.0738	5.63	10.25	18.35
R1	1,897	0.0364	2.78	5.06	9.05
R4	212	0.0206	1.57	2.85	5.11
R5	1,199	0.0115	0.88	1.60	2.86
S1	466	0.0409	3.12	5.68	10.17
NS14	456	0.0633	4.83	8.78	15.73
NS20	456	0.0723	5.51	10.03	17.96
NS21	244	0.0165	1.26	2.29	4.11
NS22	154	0.0422	3.22	5.85	10.47
X2	35	0.0432	3.30	6.00	10.74
X4	65	0.0322	2.46	4.47	8.00
AR20	445	0.0279	2.13	3.87	6.93
SS 3	493	0.0135	1.03	1.87	3.35
SS 5	361	0.0586	4.47	8.13	14.57
SS35	48	0.0316	2.41	4.38	7.85
TR 6	741	0.0364	2.78	5.05	9.04
TR 7	420	0.2784	21.23	38.63	69.17
IL 9	895	0.0384	2.93	5.33	9.54
FADJ	103	0.1000	7.62	13.87	24.84
APT	287	0.0750	5.72	10.41	18.63

Table 5.10 Route Errors Percentage Resulting from Transit Ridership Variability in Miami

Route	Weekday Average Ridership	Cv	Error (%)		
			68% CL	85% CL	95% CL
A	468	0.1445	11.02	20.04	35.89
B	1,574	0.0438	3.34	6.08	10.89
C	3,698	0.0565	4.31	7.84	14.05
E	1,159	0.0247	1.88	3.42	6.13
G	3,131	0.0355	2.71	4.93	8.83
H	4,844	0.0127	0.97	1.77	3.17
J	4,753	0.0181	1.38	2.52	4.51
K	4,504	0.0414	3.15	5.74	10.27
L	10,321	0.0186	1.42	2.58	4.62
M	1,926	0.0607	4.63	8.42	15.08
R	465	0.1197	9.13	16.61	29.75
S	12,546	0.0231	1.76	3.20	5.73
T	2,351	0.0178	1.35	2.46	4.41
V	399	0.0396	3.02	5.50	9.84
W	398	0.0418	3.19	5.80	10.39
1	1,815	0.2809	21.43	38.98	69.79
2	3,938	0.0132	1.01	1.84	3.29
3/16	13,069	0.0187	1.43	2.60	4.65
6	291	0.1572	11.99	21.81	39.05
7	3,773	0.0073	0.56	1.02	1.82
8	7,545	0.0179	1.37	2.49	4.46
9	5,162	0.0169	1.29	2.34	4.19
10	2,541	0.0246	1.88	3.41	6.11
11	12,497	0.0179	1.36	2.48	4.45
12	3,230	0.0106	0.81	1.47	2.64
17	5,713	0.0142	1.08	1.97	3.52
21	2,486	0.0410	3.13	5.69	10.18
22	4,055	0.0463	3.53	6.42	11.49
24	4,240	0.0608	4.63	8.43	15.09
27	8,891	0.0400	3.05	5.55	9.94
28	838	0.0890	6.79	12.35	22.12
29	315	0.1288	9.82	17.87	32.00
31	1,200	0.0289	2.21	4.01	7.18
32	3,876	0.0454	3.47	6.31	11.29
33	2,198	0.0203	1.55	2.81	5.04
35/70	2,678	0.0788	6.01	10.93	19.57
36	3,296	0.0259	1.97	3.59	6.42
37	3,575	0.0137	1.05	1.90	3.40
38	3,769	0.0391	2.99	5.43	9.73
40	2,194	0.0466	3.55	6.46	11.56
42	1,267	0.0443	3.38	6.15	11.01
48	522	0.0911	6.95	12.64	22.64
51 FLAGLER MAX	3,047	0.0435	3.32	6.04	10.81
52	1,338	0.0716	5.46	9.94	17.80
54	3,289	0.0168	1.28	2.34	4.18
56	1,573	0.0267	2.04	3.71	6.64
57/72	1,829	0.0671	5.12	9.31	16.67
62	5,253	0.0636	4.85	8.83	15.80
65EX	392	0.0877	6.69	12.17	21.80
71	1,348	0.0137	1.05	1.90	3.41
73	2,133	0.0126	0.96	1.75	3.13

Table 5.10 Route Errors Percentage Resulting from Transit Ridership Variability in Miami (Continued)

Route	Weekday Average Ridership	Cv	Error (%)		
			68% CL	85% CL	95% CL
75	4,249	0.0276	2.10	3.83	6.85
77	9,964	0.0062	0.47	0.86	1.53
83	4,268	0.0111	0.85	1.55	2.77
87	1,739	0.0308	2.35	4.28	7.66
88	2,653	0.0286	2.18	3.97	7.11
91	1,401	0.0409	3.12	5.67	10.16
BISC MAX 1	2,267	0.0187	1.42	2.59	4.64
95EX	1,748	0.0032	0.24	0.44	0.78
27 MAX	742	0.0216	1.65	3.00	5.37
104	1,658	0.0642	4.89	8.90	15.94
TR-36 ST	60	0.2226	16.97	30.88	55.29
TR-MIA	499	0.0850	6.48	11.79	21.12
134- RIVERSIDE SHUTTLE	43	0.0839	6.40	11.63	20.83
137-WEST DADE CONN	1,337	0.1616	12.33	22.42	40.15
231-BUSWAY LOCAL	960	0.0812	6.20	11.27	20.18
236-A RPORT OWL	363	0.0329	2.51	4.56	8.17
237 DOUGLAS BRIDGE	51	0.2377	18.13	32.98	59.06
238 EAST/WEST CONN	535	0.0794	6.06	11.02	19.73
240-BIRD MAX	527	0.0055	0.42	0.76	1.36
241-NORTH DADE CONN	345	0.0210	1.60	2.91	5.21
242-DORAL CONN	493	0.0936	7.14	12.99	23.26
243-SEAPORT CONN	208	0.1999	15.24	27.73	49.65
245-OKEECHOBEE CONN	271	0.1453	11.08	20.16	36.09
246-NIGHT OWL	513	0.2881	21.97	39.97	71.58
248-BRICKELL KEY SHUTTLE	723	0.0947	7.22	13.13	23.52
252-CORAL REEF	946	0.0481	3.67	6.68	11.96
267-LUDLUM MAX	430	0.0805	6.14	11.17	20.00
287-SAGA BAY	311	0.1213	9.25	16.82	30.13
KAT-KILLIAN	1,185	0.0953	7.27	13.23	23.68
KAT-SUNSET	834	0.1194	9.11	16.56	29.66
KAT-KENDALL	660	0.0478	3.65	6.64	11.89
GREEN HILLS SHUTTLE	21	0.3760	28.67	52.16	93.40
KINGS CREEK SHUTTLE	11	0.1083	8.26	15.02	26.89
SIERRA LAKES SHUTTLE	11	0.0909	6.93	12.61	22.58
ROBERT SHARPE SHUTTLE	10	0.3111	23.73	43.16	77.28

Table 5.11 Route Errors Percentage Resulting from Transit Ridership Variability in Tampa

Route	Weekday Average Ridership	Cv	Error (%)		
			68% CL	85% CL	95% CL
1	2,611	0.0256	1.95	3.55	6.36
2	3,422	0.0223	1.70	3.09	5.53
4	422	0.0594	4.53	8.25	14.77
5	1,284	0.0834	6.36	11.58	20.73
6	2,236	0.0535	4.08	7.42	13.29
7	1,960	0.0725	5.53	10.06	18.01
8	1,092	0.0435	3.32	6.04	10.82
9	1,484	0.0819	6.25	11.37	20.36
10	786	0.0251	1.92	3.49	6.24
11	608	0.0144	1.10	2.00	3.58
12	2,188	0.0470	3.58	6.52	11.67
14	499	0.0734	5.60	10.18	18.23
15	840	0.0420	3.21	5.83	10.44
16	473	0.0761	5.80	10.56	18.91
17	464	0.0646	4.93	8.96	16.05
18	1,212	0.0902	6.88	12.52	22.42
19	1,153	0.0773	5.90	10.73	19.22
30	1,421	0.0212	1.62	2.94	5.27
31	292	0.0173	1.32	2.40	4.29
32	946	0.0623	4.75	8.65	15.49
33	411	0.0768	5.86	10.66	19.08
34	1,563	0.0382	2.92	5.30	9.50
36	1,111	0.0313	2.38	4.34	7.77
37	530	0.1832	13.97	25.42	45.52
39	1,054	0.0450	3.43	6.24	11.18
41	307	0.2392	18.25	33.19	59.44
46	220	0.1294	9.87	17.95	32.14
70	68	0.1459	11.13	20.24	36.25
71	72	0.1354	10.33	18.79	33.64
72	60	0.1217	9.28	16.88	30.22
73	35	0.1269	9.68	17.60	31.52
81	56	0.0265	2.02	3.68	6.58
83	525	0.0660	5.04	9.16	16.41
84	144	0.0332	2.53	4.61	8.25
88	96	0.0479	3.65	6.64	11.89
96	542	0.0712	5.43	9.87	17.68
20X	55	0.0615	4.69	8.53	15.27
22X	69	0.0726	5.54	10.07	18.03
23X	42	0.0933	7.11	12.94	23.17
25X	81	0.0170	1.30	2.36	4.22
26X	26	0.1473	11.24	20.44	36.60
27X	77	0.0528	4.03	7.33	13.12
28X	68	0.0141	1.07	1.95	3.50
50X	33	0.0699	5.33	9.70	17.38
200X	64	0.0370	2.82	5.13	9.19
57LX	166	0.0643	4.91	8.93	15.98
58LX	105	0.0163	1.24	2.26	4.05

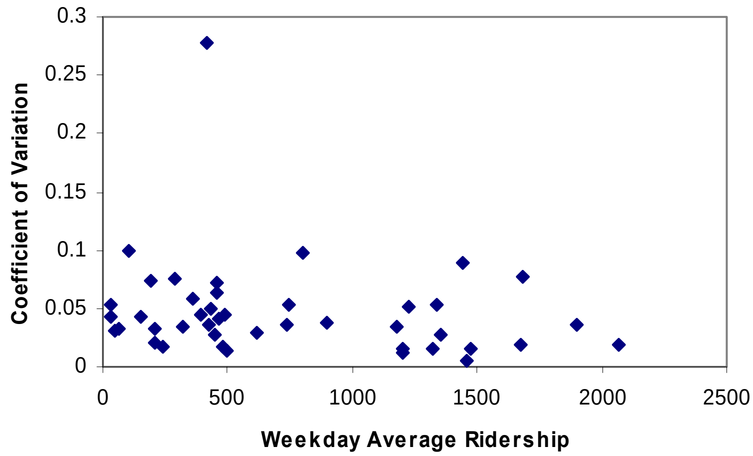

Figure 5.1 Relationship between Cv and Weekday Average Ridership for Jacksonville

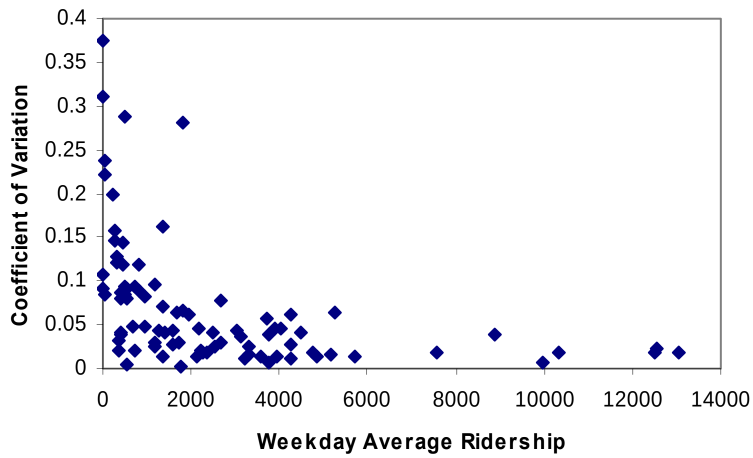

Figure 5.2 Relationship between *C*v and Weekday Average Ridership for Miami

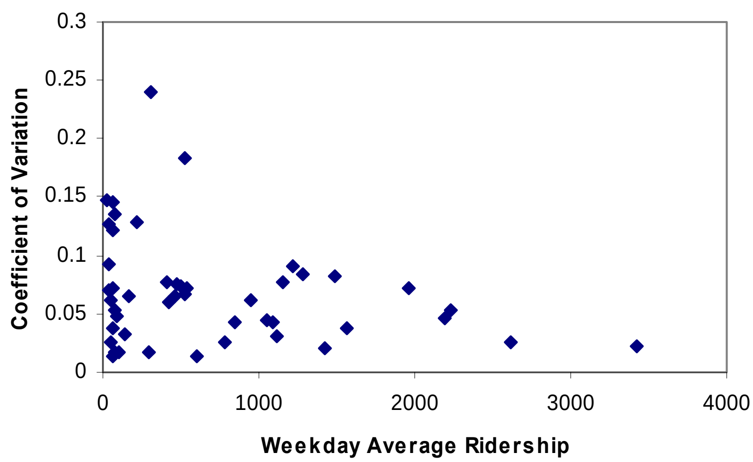

Figure 5.3 Relationship between *C*v and Weekday Average Ridership for Tampa

Table 5.12 Comparison between the FDOT Accuracy Standard and Transit Ridership Variation

Daily Ridership	FDOT Accuracy Standard	Error		
		68% CL	85% CL	95% CL
Jacksonville				
< 1,000 Passengers/Day	-100% – 150%	±22%	±39%	±70%
1,000 – 2,000 Passengers/Day	± 90%	±7%	±13%	±22%
2,000 – 5,000 Passengers/Day	± 70%	±2%	±3%	±5%
5,000 – 10,000 Passengers/Day	± 45%	–	–	–
10,000 – 20,000 Passengers/Day	± 35%	–	–	–
> 20,000 Passengers/Day	± 30%	–	–	–
Miami				
< 1,000 Passengers/Day	-100% – 150%	±29%	±53%	±94%
1,000 – 2,000 Passengers/Day	± 90%	±22%	±39%	±41%
2,000 – 5,000 Passengers/Day	± 70%	±7%	±11%	*
5,000 – 10,000 Passengers/Day	± 45%	±5%	±9%	±20%
10,000 – 20,000 Passengers/Day	± 35%	±2%	±4%	±16%
> 20,000 Passengers/Day	± 30%	–	–	±6%
				–
Tampa				
< 1,000 Passengers/Day	-100% – 150%	±19%	±34%	±60%
1,000 – 2,000 Passengers/Day	± 90%	±7%	±13%	±23%
2,000 – 5,000 Passengers/Day	± 70%	±5%	±8%	±14%
5,000 – 10,000 Passengers/Day	± 45%	–	–	–
10,000 – 20,000 Passengers/Day	± 35%	–	–	–
> 20,000 Passengers/Day	± 30%	–	–	–

* The value reflects a result excluding the outlier.

CHAPTER SIX

CONCLUSIONS AND FUTURE WORKS

6.1 Conclusions

The recommended modifications of the current accuracy standards in FSUTMS are summarized as follows:

1. In highway networks, sampling and model misspecification errors are considered in deriving accuracy standards. For sampling errors, except several facility types with limited sample size (less than 50), almost all error limits are below 10% (average is 5%) at the 95% level of confidence. One-way and toll facilities were slightly higher than others due to the higher variability of traffic counts observed on those facilities.

2. Allowable error limits to account for model misspecification errors are derived based on the perception error associated with travelers in choosing routes. The overall standards are derived by adding 5% sampling error on the model misspecification error, assuming the sampling error is independent of the model misspecification error. As a comparison, the FHWA standards are about 10% higher than the proposed ones for ADTs equal to 15,000 vpd or higher, and are substantially lower for ADTs below 15,000 vpd. The proposed accuracy standards meet the Michigan DOT standards quite closely, except when the average assigned volume falls between 2,500 vpd and 15,000 vpd. The current practices by both FHWA and Michigan DOT seem to underestimate the assignment error by 10-15 percent within the abovementioned ADT range compared to the proposed standards. The RMSE standards by link volume groups are revised as follows:

Proposed Modified Accuracy Standards for Highway Networks in FSUTMS

Control Statistic	**Original Accuracy Standard**	**Modified Accuracy Standard**
Root Mean Square Error (%) (Area)	35% - 50%	See the following worksheet
Root Mean Square Error (%) (Link volume group)	25% ≥ (AADT ≥ 50,000 vpd) 30-100% (AADT < 50,000 vpd) >100% (AADT < 3,000 vpd)	150% (ADT<1,000) 100% (ADT 1,000–2,500) 65% (ADT 2,500-5,000) 45% (ADT 5,000–10,000) 35% (ADT 10,000–15,000) 25% (ADT 15,000 – 25,000) 15% (ADT 25,000 – 50,000) 10% (ADT>50,000)

3. The current area-wide RMSE standard of 35%-50% seems arbitrary. This standard should be a composite measure of the actual (or estimated) distribution of the roadway ADT groups and the corresponding allowable error in each volume group. Posting accuracy standards by a coarse classification of area types will not be reasonable if the actual distribution of roadway ADTs are significantly different from those designated area types. Therefore, the following

worksheet is designed and proposed for planners in different jurisdictional areas to develop their own standards.

Proposed Accuracy Standards Worksheet on the Area-wide Allowable Error

ADT	Mean ADT (vpd) (1)	Distribution of Roadway ADT (2)	Proposed Allowable Error (3)	Col. (1) × Col. (2) (4)	Col. (3) × Col. (4) (5)
< 1,000	500		1.5		
1,000 – 2,500	1750		1.0		
2,500 – 5,000	3750		0.65		
5,000 – 10,000	7500	User	0.45		
10,000 – 15,000	12500	Inputs	0.35		
15,000 – 25,000	20000		0.25		
25,000 – 50,000	37500		0.15		
> 50,000	75000		0.1		
Sum	-	1.0	-	(6)	(7)
Area-wide Allowable Error = (7) ÷ (6) =					

4. This study recommends eliminating the facility-type and size-of-facility specific accuracy standards, since they have created some contradictions to the standards by volume-group. Current practices show that different standards were applied onto different roadway facilities even though they carried similar traffic volumes. While this may seem to contradict the initial intent of this study to accommodate the two-digit codes for facility type, it is suggested that the extra categories be eliminated to simplify and standardize the structure of the current accuracy standards for highway networks.

5. Based on the current accuracy standards, the required sample sizes are computed for each functional class area-wide at 68%, 85%, and 95% level of confidence. The results show that the required sample size increase significantly as the level of confidence increases. However, the required sample size remains at an affordable range of 10 and 60. Except for one-way facility, which may require a larger sample such as 50, the use of 30 count stations is a rule of thumb for general sampling purposes.

6. For transit networks, the accuracy standards were derived to account for statistical variability, aggregation errors, and prediction errors. Errors in modal share subjected to aggregation and prediction variability are derived for both multinomial logit and nested logit models based on a second-order Taylor expansion theory. Variability of the modal share is also derived to establish confidence intervals at a user-specified level of significance as an alternative form of accuracy standards. Due to the limitation of data sources in both quantity and quality in supporting estimation of the multinomial and/or nested logit models, however, the accuracy standards to account for aggregation and prediction errors could not be developed and thus will be reserved for future studies. On the other hand, errors due to statistical variability are significantly lower than the current accuracy standards in terms of daily ridership, which is expected since no other sources of errors are considered.

6.2 Future Works

The accuracy standards to account for aggregation and prediction errors for transit network could not be developed due to the limitation of data sources in both quantity and quality in supporting estimation of the multinomial and/or nested logit models. This will be reserved for future studies.

REFERENCES

Abdel-Aty, M. and H. Abdelwahab (2001). *Calibration of Nested-Logit Mode-Choice Models for Florida*. Final Report, Prepared for the Florida Department of Transportation, Tallahassee, Florida.

Akçelik, R. (1991) Travel Time Functions for Transport Planning Purpose: Davidson's Function, Its Time-Dependent Form and an Alternative Travel Time function, *Australian Road Research*, Vol. 21, No. 3, pp. 49-59.

Alpern, B. and D. M. Levinsohn (1984). Calibration of a Modal Split Model Using 1980 Census Data, *Compendium of Technical Papers*, ITE, pp. 8-16

Barton-Aschman Associates, Inc. and Cambridge Systematics, Inc. (1997) *Model Validation and Reasonableness Checking Manual*, Prepared for Federal Highway Administration.

Ben-Akiva, M. and S. R. Lerman (1985). Discrete Choice Analysis: Theory and Application to Travel Demand, The MIT Press, 1985.

Bureau of Transportation Statistics (2001). National Household (Personal) Transportation Survey.

Bureau of Transportation Statistics (2001). American Travel Survey.

Caldwell, L. C. III and M. J. Demetsky (1980). Transferability of Trip Generation Models, *Transportation Research Record* 751, pp. 56-62.

Carr Smith Corradino (2000). *Southeast Florida Regional Travel Characteristic Study*. Executive Summary Report, Prepared for District 4, Florida Department of Transportation.

Chung, W. L. and M. J. Henry (1999). Calibration of Travel Demand Models for Small Jurisdictions - Our Standards Are Way Too Low.

COMSIS (1981). *Develop Standard Distribution and Assignment Models*, Task C Report, Urban Transportation Model Update – Phase II.

COMSIS Corporation (1981). *Develop Standardized Modal Split Models*, Task E Report, Urban Transportation Planning Model Update – Phase II, Prepared for Florida Department of Transportation.

Cynecki, M. J., S. Khasnabis, and M. A. Flak (1982). Multimodal Logit Travel-Demand Model for Small and Medium-Sized Urban Areas, *Transportation Research Record* 848, pp. 28-36.

Daganzo, C. F. (1977). Some Statistical Problems in Connection With Traffic Assignment, *Transportation Research*, Vol. 11, pp. 385-389.

Daganzo, C. F. and Y. Sheffi (1977). On Stochastic Models of Traffic Assignment, *Transportation Science*, Vol. 11, No. 3, pp. 253-274.

Daly, A. J. and J. de D. Ortúzar (1990). Forecasting and Data Aggregation: Theory and Practice. *Traffic Engineering & Control*, 31(12), pp. 632-643.

Daly, A. J. and S. Zachary (1975). Calibration of Modal Split Models, Planning and Transport Research and Computation Co. Ltd., pp. 236-245.

Design Manual for Roads and Bridges. Chapter 4, Traffic Model Development, Vol. 12, Sec. 2, May 1996. (U.K.) Note: The name GEH comes from the initials of the person who developed the test.

Easa, S. M. (1991) Traffic Assignment in Practice: Overview and Guidelines for Users, *Journal of Transportation Engineering*, Vol. 117, No. 6, pp. 602-623.

Evans, A. W. (1971) The Calibration of Trip Distribution Models with Exponential or Similar Cost Functions, *Transportation Research*, Vol. 5, pp. 15-38.

Florida Department of Transportation (1997). *FSUTMS HEVAL Model*, Final Technical Report.

Florida Department of Transportation (1997). *FSUTMS Mode Choice Models (Mode)*. Prepared by Systems Planning Office, Tallahassee, Florida.

Florida Department of Transportation (1997). *FSUTMS Transit Network Model*, Final Technical Report No. 11.

Federal Highway Administration (1977). *Computer Programs for Urban Transportation Planning: PLANPAC/BACK-PAC General Information*, U. S. Department of Transportation, Stock No. 050-001-00125-0, Washington, DC.

Fleet, C. and P. DeCorla-Souza (1990). Increasing the Capacity of Urban Highways - The Role of Freeways, presented at the 69th Annual Meeting of the TRB, January 7-11.

Florian, M, R. Chapleau, S. Nguyen, C. Achim, L. James-Lefebvre, S. Galarneau, J. Lefebvre, and C. Fisk (1979). Validation and Application of an Equilibrium-Based Two-Mode Urban Transportation Planning Method (EMME), *Transportation Research Record* 728, pp. 14-23.

Florian, M. and S. Nguyen (1976). An Application and Validation of Equilibrium Trip Assignment Methods, *Transportation Science*, Vol. 10, No. 4, pp. 374-390.

Ganek, J. and R. Saulino (1976). A disaggregate modal-split for work trips involving three mode choices, *Transportation Research Record* 610, pp. 25-29.

Goode, L. R. and C. Heimbach (1983). Evaluation of the Transferability of Trip Generation from One Urban Area to Another, *Transportation Research Record* 931, pp. 120-125.

Horowitz, J. L. (1984) Testing disaggregate travel demand models by comparing predicted and observed market shares, *Transportation Research Record* 976, pp. 1-7.

Hoang, L. T. and V. P. Poteat (1980). Estimating Vehicle Miles of Travel by Using Randoming Sampling Techniques, *Transportation Research Record* 779, pp.6-10.

Hyman, G. M. (1969). The Calibration of Trip Distribution Models, *Environment and Planning* 1, pp. 105-112.

Ismart, D. (1990) *Calibration and Adjustment of System Planning Models*, Prepared for Federal Highway Administration.

JHK & Associates and Dowling Associates (1992). *Travel Forecasting Guidelines*, Prepared for State of California Department of Transportation.

Kirby, H. R. (1974) Theoretical Requirements for Calibrating Gravity Model, *Transportation Research*, Vol. 8, pp. 97-104.

Koppelman, F. S. (1975a). *Travel Prediction with Models of Individualistic Choice Behavior*. Ph.D. Dissertation, Department of Civil Engineering, Massachusetts Institute of Technology, Camgridge, Massachusetts.

Koppelman, F. S. (1975b). “Develop Alternative Aggregate Models for Testing Purposes and Select Procedure for Use with TRANS.” Task 4, Final Report, *Development of An Aggregate Model of Urbanized Area Travel Behavior*. Center for Transportation Studies, Massachusetts Institute of Technology, Camgridge, Massachusetts.

Koppelman, F. S. (1976). “Methodology for Analyzing Errors in Prediction with Disaggregate Choice Models.” *Transportation Research Record 592*, Transportation Research Board, National Research Council, Washington, D.C., 17-23.

Kumar, A. and D. Levinson (1993). Specifying, Estimating, and Validating a New Trip Generation Model: Case study in Montgomery County, *Transportation Research Record* 1413, pp. 107-113.

Lan, C-J and P. Hu. (2000) A Mixed Generalized Linear Model for Estimating Household Trip Production, *Transportation Research Record*, 1718, pp. 61-67.

Landau, U. (1981). “Reducing the Aggregation Bias: A Heuristic Approach.” *Transportation Research A*, 15(5), 367-375.

Levinson, D.M. and A. Kumar (1994). Multimodal Trip Distribution: Structure and Application, *Transportation Research Record* 1466, pp. 124-131.

Levinson, H. S. and A. Roark (1975). Guide for Traffic Volume Counting Manual. Federal Highway Administration, U.S. Department of Transportation, preliminary.

Li, S. C. and A. Lapkowski (1999). Improving the Accuracy of Demand Forecast Models, *Florida Transportation Modeling Newsletter*, pp. 3-4.

Martin, W. A. and N. A. McGuckin (1998). *Travel Estimation Techniques for Urban Planning*. NCHRP Report 365, Transportation Research Board, Washington, DC.

McFadden, D. and F. Reid (1975). "Aggregate Travel Demand Forecasting from Disaggregate Behavioral Models." *Transportation Research Record 534*, Transportation Research Board, National Research Council, Washington, D.C., 24-37.

McFadden, D., A. P. Talvitie, S. Cosslett, I. Hasan, M. Johnson, F. A. Reid, and K. Train (1977). *Demand Model Estimation and Validation*. Final Report, Urban Travel Demand Forecasting Project, Phase 1, Vol. V, Institute of Transportation Studies, University of California, Berkeley, California.

Morrall, J. F. (1971) *Work Trip Distribution and Modal Split in the Metropolitan Toronto Region*, Waterloo University, Canada.

Murchland (1969). Road Network Traffic Distribution in Equilibrium, Paper presented at the *Conference on Mathematical Methods in the Economics Sciences*, Mathematisches Forschungsinstitut, Oberwolfach.

Nguyen, S. (1974) An Algorithm for the Traffic Assignment Problem, *Transportation Science* 8, pp. 203-216.

Parody, T. E. (1977). "Analysis of Predictive Qualities of Disaggregate Modal-Choice Models." *Transportation Research Record 637*, Transportation Research Board, National Research Council, Washington, D.C., 51-57.

Parsons Brinckerhoff Quade & Douglas, Inc. and Kittelson & Associates, Inc. (1995) *Travel Demand Model Development and Application Guidelines*, Prepared for Oregon Department of Transportation Planning Section, Transportation Planning Analysis Unit.

Reid, F. A. (1977). *Systematic and Efficient Methods for Minimizing Error in Aggregate Predictions from Disaggregate Models*. Institute of Transportation Studies, University of California, Berkeley, California.

Robbins, J. (1978) Mathematical Modeling – the Error of Our Ways, *Traffic Engineering & Control*, Vol. 19, No. 1, pp. 32-35.

San Diego Association of Governments (1995) *Regional Transportation Models*.

Schimpeler-Corradino Associates (1984). *Modal Split Model Refinement and Calibration Standards*, Task H Report, Urban Transportation Planning Model Update – Phase III, Prepared for Florida Department of Transportation.

Sheffi, Y. (1985) Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods, Prentice-Hall.

Spiess, H. (1990) Conical Volume-Delay Functions, *Transportation Science*, Vol. 24, No. 2, pp. 153-158.

Stover, V. G., J. D. Benson, and J. Buechler (1976). A Sensitivity Evaluation of the Traffic Assignment Process, *Transportation Research Record* 610, pp. 50-52.

Tekiele, L. (1993) *Urban Model Calibration Standards*, Prepared for Michigan Department of Transportation.

Texas Transportation Institute (1999). *Metropolitan Transportation Planning Issues: A Primer/ Anthology For Small and Medium MPOs*, Travel Model Improvement Program Report.

TMIP (1997). *Model Validation and Reasonableness Checking Manual*, Submitted to Travel Model Improvement Program, FHWA by Barton-Aschman Associates and Cambridge Systematics, Inc.

Train, K. (1978) Validation test of disaggregate mode choice model, *Transportation Research*, Vol. 12, 3, pp. 167-174.

Travel Model Improvement Program (1997). *Model Validation and Reasonableness Checking Manual*, Prepared for FHWA by Barton-Aschman Associates and Cambridge Systematics.

Tretvik, T. and S. Widlert (1998). The stability of parameter estimation in household-based structured logit models for travel-to-work decisions, Chapter 25, *Travel Behavior Research: Updating the State of Play*, pp. 457-472.

U.S. Department of Transportation (2000). Census Transportation Planning Package.

UMTA (1978). Characteristics of Urban Transportation Demand – A Handbook of Transportation Planners.

Van Zijpp, N. J., B. G. and M. G. H. Bell. (1998). Improved Algorithms for Calibrating Gravity models. *Transportation Networks: Recent Methodological Advances*, pp. 99-113.

Volusia County MPO (2000). *2020 Long Range Transportation Plan – Refinement*, November.

Williams G. J. and R. F. Fisher (1979). Georgia's Evaluation of Federal Highway Administration Procedures For Estimating Urban Vehicle Miles of Travel, pp.15-22.

Willis, C. O. Jr. and J. W. Lee (1980). Time Stability of Attitudes Toward Transit Use in the Orlando, Florida, Urbanized Area, *Transportation Research Record* 761, pp. 56-58.

Wilson, F. R., B. G. Bisson, and K. B. Kobia (1986). Factors That Determine Mode Choice in the Transportation of General Freight, *Transportation Research Record* 1061, pp. 26-31.

Zhao, Y. and K. M. Kockelman (2001). The Propagation of Uncertainty Through Travel Demand Models: An Exploratory Analysis. Submitted to Annuals of Regional Science.

APPENDICES

Appendix I. Hypothesis Testing on the Weekly ADT

Table I.1 Freeway and Expressway Weekly and Annual Morning-ADT *t*-test and *F*-test

Station ID	Weekly Stat		Annual Stat			*t*-test	*F*-test
	Mean	Std	Mean	Std	Size		
100110	64350	23345	64323	23361	510	0.0030	0.9986
100123	67089	19242	66918	19299	359	0.0232	0.9941
100194	44233	15971	43900	16023	710	0.0547	0.9935
100224	24505	7227	24731	7576	694	-0.0786	0.9100
109922	24162	4708	23998	4685	113	0.0898	1.0096
109926	61790	20381	57660	21520	453	0.5042	0.8969
150183	20630	7145	20690	7113	703	-0.0222	1.0091
489924	17927	5768	17801	5738	534	0.0577	1.0105
550304	25486	10186	25541	9696	549	-0.0149	1.1036
709919	14962	3525	14891	4359	605	0.0429	0.6539
720121	34227	10354	34030	10329	251	0.0498	1.0048
720171	46351	26555	47019	24883	251	-0.0699	1.1389
720216	29226	12209	29577	12061	409	-0.0763	1.0247
729914	33464	13631	34302	13386	359	-0.1640	1.0369
729923	23160	4602	24968	4666	396	-1.0164	0.9728
730292	25349	6494	26778	6720	632	-0.5597	0.9339
750130	56951	15533	55313	15469	578	0.2785	1.0083
750196	74588	26291	75071	25774	696	-0.0493	1.0405
750204	14147	3178	14619	3209	370	-0.3855	0.9804
770343	52717	18145	50033	17715	245	0.3950	1.0491
790133	23029	6648	22408	6569	529	0.2485	1.0242
799906	23801	7604	20726	9250	525	0.8753	0.6758
860163	104280	35517	104110	34418	588	0.0130	1.0649
860331	122660	39224	122340	40322	595	0.0209	0.9463
870108	48375	16395	47921	16463	435	0.0724	0.9918
870187	75879	28910	76027	29239	201	-0.0132	0.9776
930198	60417	16735	61131	17031	609	-0.1103	0.9655
930217	29018	6055	28832	6090	653	0.0804	0.9887
940260	20341	5359	20696	5169	648	-0.1807	1.0749
940334	17010	3791	17734	3769	603	-0.5053	1.0115

Table I.2 Divided Arterial Weekly and Annual Morning-ADT *t*-test and *F*-test

Station ID	Weekly Stat		Annual Stat			*t*-test	F-test
	Mean	Std	Mean	Std	Size		
30094	10660	2901	10421	2902	553	0.2165	0.9990
80283	8705	1831	8723	1831	648	-0.0259	0.9992
80294	3487	1399	3497	1404	591	-0.0187	0.9927
100162	11782	4675	11813	4725	472	-0.0172	0.9789
100321	32439	14011	32837	14105	657	-0.0743	0.9867
100339	38106	15557	36551	14961	155	0.2686	1.0813
110177	18059	6655	17966	6687	593	0.0366	0.9902
130180	9590	3080	9702	3062	660	-0.0959	1.0117
140199	25419	7220	25287	7341	216	0.0468	0.9674
150066	12340	3173	12336	3129	614	0.0034	1.0287
150086	16665	7614	16600	7595	502	0.0225	1.0050
150295	23791	8434	23520	8457	553	0.0843	0.9947
150302	22914	10479	23319	10497	676	-0.1016	0.9966
160128	13742	3927	13801	3947	442	-0.0392	0.9900
169927	6489	2436	6344	2495	398	0.1531	0.9529
170181	21749	8689	21930	8601	642	-0.0554	1.0204
260185	11664	4086	11645	4137	626	0.0121	0.9758
260323	21637	8963	20823	8695	692	0.2464	1.0626
290286	3808	1489	3844	1474	613	-0.0648	1.0213
360249	8177	3021	8161	3057	544	0.0138	0.9768
360264	6284	1334	6333	1337	561	-0.0958	0.9952
460166	5245	1291	5829	1850	563	-0.8330	0.4869
460308	7878	2765	7834	2732	718	0.0426	1.0243
480282	17089	6704	16976	6546	655	0.0454	1.0490
530117	8273	3002	8569	2974	305	-0.2604	1.0189
550151	16777	7331	16451	7354	507	0.1165	0.9938
550207	4990	2296	4804	2260	263	0.2138	1.0326
550209	8497	3600	8380	3529	529	0.0870	1.0409
550226	15027	6352	14758	6292	567	0.1124	1.0189
570167	24663	10319	24200	10389	588	0.1172	0.9866
570250	15056	6554	14851	6538	373	0.0822	1.0047
570293	27909	8752	28580	8597	577	-0.2052	1.0362
580261	18089	5826	18672	5909	698	-0.2598	0.9722
700113	24583	8217	24568	8348	446	0.0047	0.9689
700114	7286	2728	7348	2661	573	-0.0609	1.0508
710189	14213	6102	14528	6673	134	-0.1222	0.8362
710233	10667	4244	10662	4220	494	0.0031	1.0117

(**Table I.2** continued)

Station ID	Weekly Stat		Annual Stat			*t*-test	*F*-test
	Mean	Std	Mean	Std	Size		
720062	17860	5839	17878	5883	471	-0.0080	0.9850
720161	19164	7906	18801	7953	653	0.1201	0.9881
720172	27818	12717	27585	12782	605	0.0480	0.9899
740182	6063	2381	6103	2361	585	-0.0445	1.0168
750038	22349	8303	22145	8375	631	0.0641	0.9829
750154	17171	3740	17345	3753	562	-0.1219	0.9932
750175	19995	8577	19341	8993	538	0.1913	0.9097
770102	16480	5600	16305	5577	531	0.0825	1.0084
770197	22158	8732	22150	8535	502	0.0025	1.0469
780311	25384	9893	25398	9789	717	-0.0038	1.0214
780329	15245	4891	15093	5172	553	0.0773	0.8942
799929	6191	1622	5979	1599	301	0.3458	1.0297
860150	34188	11900	33838	11846	689	0.0778	1.0091
860176	13652	3205	13498	3178	664	0.1275	1.0168
860214	33962	14149	33406	14168	628	0.1033	0.9973
860222	27392	9666	27426	9750	670	-0.0092	0.9829
860298	30801	10864	30420	10913	688	0.0919	0.9910
860306	8473	1954	8565	2020	570	-0.1196	0.9363
870031	44917	12311	44237	12312	626	0.1453	0.9998
870096	10859	4564	11037	4508	481	-0.1037	1.0251
870178	57225	24092	56608	24101	663	0.0674	0.9993
870188	37236	16936	37670	16999	615	-0.0672	0.9926
870193	37290	19340	37223	19205	699	0.0092	1.0141
870258	13597	3062	13279	3090	596	0.2708	0.9821
870266	13801	5962	13959	5958	656	-0.0698	1.0014
879930	25802	7714	23357	7797	599	0.8250	0.9788
880314	14936	4945	14895	4968	698	0.0217	0.9904
890332	16880	5177	17242	5273	580	-0.1806	0.9640
900164	11728	2934	11820	2933	663	-0.0826	1.0011
900165	14269	3423	14245	3462	633	0.0182	0.9781
920265	18594	3431	18268	3394	533	0.2525	1.0219
930010	12456	3920	12250	4012	481	0.1349	0.9545
930101	27985	11618	28397	11603	632	-0.0934	1.0026

Table I.3 Undivided Arterial Weekly and Annual Morning-ADT *t*-test and *F*-test

Station ID	Weekly Stat		Annual Stat			*t*-test	*F*-test
	Mean	Std	Mean	Std	Size		
110246	5248	1975	5449	2004	596	-0.2630	0.9718
460315	4935	1771	4862	1740	658	0.1113	1.0364
550201	7505	3207	7560	3237	325	-0.0446	0.9820
550300	6893	3128	6720	3125	616	0.1452	1.0022
559908	7697	3106	7635	3147	397	0.0510	0.9742
890259	2787	797	2830	813	606	-0.1382	0.9619

Table I.4 Collectors Weekly and Annual Morning-ADT *t*-test and *F*-test

Station ID	Weekly Stat		Annual Stat			*t*-test	*F*-test
	Mean	Std	Mean	Std	Size		
160275	5093	1623	5038	1566	637	0.0922	1.0748
550206	2949	1243	2949	1273	691	0.0004	0.9543
550208	2290	774	2241	778	656	0.1674	0.9904
550212	2970	875	3028	868	674	-0.1771	1.0174
550213	1704	678	1719	686	632	-0.0587	0.9778
860215	6173	1709	6276	1729	671	-0.1561	0.9773
880326	6788	2642	6791	2655	680	-0.0036	0.9905
930087	15410	5724	15351	5838	712	0.0266	0.9613

Table I.5 Toll Facility Weekly and Annual Morning-ADT *t*-test and *F*-test

Station ID	Weekly Stat		Annual Stat			*t*-test	*F*-test
	Mean	Std	Mean	Std	Size		
970403	35513	14097	35442	14225	150	0.0129	0.9821
970410	27978	10094	27274	10023	404	0.1842	1.0142
970413	24756	7889	25170	8002	337	-0.1355	0.9720
970416	21280	6379	21022	6334	639	0.1072	1.0143
970417	15248	3696	14717	3683	666	0.3794	1.0071
970430	13383	4347	13469	4355	594	-0.0519	0.9960
979913	12692	2980	12548	3359	225	0.1120	0.7869
979933	26714	12009	26818	11957	652	-0.0229	1.0087

Table I.6 Freeway and Expressway Weekly and Annual Midday-ADT *t*-test and *F*-test

Station ID	Weekly Stat		Annual Stat			*t*-test	*F*-test
	Mean	Std	Mean	Std	Size		
100110	59670	6598	62327	6503	510	-1.0734	1.0292
100123	67758	6904	66918	6939	359	0.3172	0.9900
100194	46991	4992	47889	5279	710	-0.4480	0.8942
100224	22863	3956	22439	4417	694	0.2529	0.8019
109922	30195	2365	28452	2127	113	2.0913	1.2366
109926	49259	6406	52705	6489	453	-1.3946	0.9747
150183	19223	3063	19250	3182	703	-0.0223	0.9265
489924	25047	2914	26429	2959	534	-1.2278	0.9693
550304	24162	4324	26406	4298	549	-1.3726	1.0120
709919	16000	5767	15852	4639	605	0.0837	1.5454
720121	31996	5439	29996	5460	251	0.9560	0.9925
720171	47084	6299	46551	9250	251	0.1513	0.4637
720216	23167	2520	23256	2536	409	-0.0921	0.9873
729914	29125	4041	29667	3727	359	-0.3805	1.1756
729923	32453	4433	32748	4543	396	-0.1704	0.9522
730292	29288	5721	26970	6236	632	0.9788	0.8418
750130	69190	6161	70742	6051	578	-0.6745	1.0369
750196	78636	7565	78816	7490	696	-0.0633	1.0201
750204	26465	3953	24751	3730	370	1.2033	1.1233
770343	60216	5148	58372	6266	245	0.7707	0.6748
790133	19349	3145	19500	3510	529	-0.1132	0.8027
799906	26870	7123	24081	7495	525	0.9786	0.9033
860163	98592	11529	96107	11437	588	0.5714	1.0162
860331	113570	13264	112590	14031	595	0.1838	0.8937
870108	44489	6764	45915	6774	435	-0.5526	0.9972
870187	61623	5677	61607	5723	201	0.0073	0.9840
930198	72126	8180	73772	8069	609	-0.5365	1.0275
930217	40095	3601	39887	4093	653	0.1339	0.7743
940260	19970	3712	22230	3834	648	-1.5517	0.9372
940334	21830	3141	21701	3549	603	0.0957	0.7835

Table I.7 Divided Arterial Weekly and Annual Midday-ADT *t*-test and *F*-test

Station ID	Weekly Stat		Annual Stat			*t*-test	*F*-test
	Mean	Std	Mean	Std	Size		
30094	17281	2709	16253	2671	553	1.0118	1.0284
80283	9087	1028	9331	1010	648	-0.6349	1.0358
80294	4831	684	4668	613	591	0.6986	1.2431
100162	15487	3916	16702	4149	472	-0.7697	0.8910
100321	30203	4050	28890	4044	657	0.8545	1.0032
100339	25262	2570	27208	2668	155	-1.8901	0.9280
110177	16644	2981	16938	2985	593	-0.2591	0.9974
130180	10383	1210	10812	1317	660	-0.8577	0.8443
140199	32033	3387	31971	3472	216	0.0465	0.9519
150066	10567	1357	10751	1393	614	-0.3475	0.9490
150086	14904	2097	14106	2177	502	0.9635	0.9279
150295	22158	2606	22295	2966	553	-0.1216	0.7723
150302	18135	2941	17981	3042	676	0.1333	0.9349
160128	11895	1849	12495	2120	442	-0.7441	0.7610
169927	6164	933	6313	992	398	-0.3938	0.8847
170181	18861	2816	18198	3076	642	0.5676	0.8379
260185	12645	1634	12575	1667	626	0.1105	0.9599
260323	18088	3371	18656	3404	692	-0.4393	0.9806
290286	4578	693	4562	699	613	0.0602	0.9808
360249	11885	1506	11742	1525	544	0.2465	0.9754
360264	10931	1591	10826	1621	561	0.1704	0.9633
460166	4733	1690	5867	1820	563	-1.6396	0.8621
460308	11800	2047	11802	2085	718	-0.0025	0.9637
480282	16116	2660	15838	2698	655	0.2712	0.9720
530117	8891	1956	8894	1923	305	-0.0034	1.0344
550151	21297	4518	21441	4607	507	-0.0821	0.9615
550207	6533	1272	6853	1245	263	-0.6710	1.0434
550209	8481	1056	8331	1056	529	0.3754	0.9998
550226	14072	2419	13826	2383	567	0.2714	1.0299
570167	14860	1934	16430	1991	588	-2.0750	0.9441
570250	14988	2259	14724	2271	373	0.3047	0.9891
570293	28856	4498	25781	4383	577	1.8445	1.0529
580261	27624	4577	24497	4635	698	1.7763	0.9753
700113	25783	2822	25697	2983	446	0.0757	0.8952
700114	7202	948	7755	1020	573	-1.4272	0.8653
710189	15817	1670	13827	1594	134	3.2133	1.0975
710233	8457	1446	8472	1418	494	-0.0287	1.0399

(**Table I.7** continued)

Station ID	Weekly Stat		Annual Stat			*t*-test	*F*-test
	Mean	Std	Mean	Std	Size		
720062	18933	2675	18728	2803	471	0.1922	0.9112
720161	25761	4592	25294	4550	653	0.2701	1.0188
720172	18570	2349	18602	2402	605	-0.0351	0.9564
740182	7103	1218	7108	1378	585	-0.0082	0.7808
750038	18509	1910	18686	1776	631	-0.2621	1.1574
750154	26092	3269	26292	3248	562	-0.1619	1.0127
750175	20205	4188	18772	5006	538	0.7538	0.7001
770102	20689	3126	20472	3091	531	0.1845	1.0229
770197	22829	3332	22661	3340	502	0.1322	0.9955
780311	19157	1845	19181	1828	717	-0.0346	1.0192
780329	12585	3073	12173	3392	553	0.3197	0.8205
799929	7467	1018	6676	1085	301	1.9077	0.8802
860150	31344	4363	31764	4588	689	-0.2411	0.9042
860176	13450	1515	14587	1549	664	-1.9320	0.9557
860214	24632	3599	25506	3719	628	-0.6186	0.9367
860222	29953	4121	30630	4111	670	-0.4335	1.0048
860298	27293	2953	26839	2985	688	0.4004	0.9786
860306	7228	2184	8787	2505	570	-1.6392	0.7603
870031	44484	5430	45185	5658	626	-0.3261	0.9209
870096	12207	2024	12317	2037	481	-0.1418	0.9872
870178	44136	6323	44177	6232	663	-0.0173	1.0295
870188	28094	4203	28253	4184	615	-0.1000	1.0091
870193	27602	2946	27667	2914	699	-0.0587	1.0217
870258	13686	1233	13706	1267	596	-0.0415	0.9471
870266	10950	1231	11060	1224	656	-0.2366	1.0118
879930	30618	4108	28127	4536	599	1.4459	0.8205
880314	14302	2223	13816	2214	698	0.5778	1.0075
890332	27543	3410	29850	3380	580	-1.7947	1.0176
900164	12001	2488	11878	2436	663	0.1329	1.0437
900165	18104	1712	17439	1618	633	1.0812	1.1205
920265	29912	2108	28955	2515	533	1.0019	0.7024
930010	15882	2646	15306	2630	481	0.5752	1.0121
930101	22329	2716	22232	2680	632	0.0952	1.0269

Table I.8 Undivided Arterial Weekly and Annual Midday-ADT *t*-test and *F*-test

Station ID	Weekly Stat		Annual Stat			*t*-test	*F*-test
	Mean	Std	Mean	Std	Size		
110246	4865	609	4826	590	596	0.1717	1.0660
460315	6211	1775	6229	1774	658	-0.0263	1.0008
550201	6366	1284	6347	1273	325	0.0405	1.0170
550300	4767	804	4724	801	616	0.1412	1.0071
559908	6434	1281	6396	1259	397	0.0777	1.0359
890259	3620	753	3338	779	606	0.9548	0.9336

Table I.9 Collectors Weekly and Annual Midday-ADT *t*-test and *F*-test

Station ID	Weekly Stat		Annual Stat			*t*-test	*F*-test
	Mean	Std	Mean	Std	Size		
160275	4685	899	4865	915	637	-0.5173	0.9655
550206	2226	662	2218	670	691	0.0315	0.9781
550208	5135	795	5191	769	656	-0.1917	1.0693
550212	6599	903	6616	945	674	-0.0477	0.9131
550213	1475	271	1466	273	632	0.0887	0.9832
860215	7788	1315	7198	1314	671	1.1809	1.0008
880326	5032	399	4817	417	680	1.3632	0.9178
930087	10508	3081	10579	3132	712	-0.0597	0.9676

Table I.10 Toll Facility Weekly and Annual Midday-ADT *t*-test and *F*-test

Station ID	Weekly Stat		Annual Stat			*t*-test	*F*-test
	Mean	Std	Mean	Std	Size		
970403	35591	2855	34273	2983	150	1.1444	0.9162
970410	27918	3297	28532	3715	404	-0.4342	0.7877
970413	26648	4100	28395	4044	337	-1.1310	1.0276
970416	20405	5074	20502	4599	639	-0.0554	1.2169
970417	14392	4282	15329	4139	666	-0.5956	1.0701
970430	12947	1528	12808	1565	594	0.2337	0.9539
979913	14549	4388	14827	4913	225	-0.1478	0.7978
979933	26484	2460	26330	2437	652	0.1663	1.0191

Table I.11 Freeway and Expressway Weekly and Annual Afternoon-ADT *t*-test and *F*-test

Station ID	**Weekly Stat**		**Annual Stat**			***t*-test**	***F*-test**
	Mean	**Std**	**Mean**	**Std**	**Size**		
100110	58398	8553	59067	8725	510	-0.2015	0.9611
100123	66559	12111	66918	12088	359	-0.0778	1.0038
100194	47318	9940	47535	9966	710	-0.0573	0.9947
100224	20543	3604	20534	3555	694	0.0067	1.0276
109922	35037	5720	35055	6228	113	-0.0075	0.8435
109926	50273	6480	50138	6607	453	0.0537	0.9619
150183	21289	2644	21247	2896	703	0.0382	0.8337
489924	31123	7700	30856	7672	534	0.0915	1.0073
550304	23057	4721	25844	4652	549	-1.5747	1.0296
709919	17161	4448	16387	4514	605	0.4511	0.9712
720121	27541	4281	28552	5429	251	-0.4881	0.6218
720171	42984	12899	42100	11479	251	0.2004	1.2627
720216	23187	3255	22043	3183	409	0.9427	1.0462
729914	27971	5643	27857	5397	359	0.0553	1.0934
729923	35720	5944	35594	6025	396	0.0549	0.9733
730292	24165	4482	24811	5057	632	-0.3364	0.7855
750130	72408	5391	72013	6093	578	0.1707	0.7829
750196	72629	6796	73543	6634	696	-0.3626	1.0494
750204	28848	7688	29462	7735	370	-0.2081	0.9877
770343	58845	9839	57728	10095	245	0.2888	0.9500
790133	15081	2655	16322	2665	529	-1.2239	0.9921
799906	29942	6309	25215	8777	525	1.4194	0.5167
860163	94971	14102	94365	13730	588	0.1161	1.0549
860331	115690	15845	115280	16717	595	0.0645	0.8984
870108	45921	6833	45593	6855	435	0.1256	0.9936
870187	47648	2325	47096	2395	201	0.6000	0.9426
930198	76819	12891	77894	12954	609	-0.2183	0.9903
930217	44968	8949	45271	8833	653	-0.0903	1.0265
940260	20182	3705	20245	3545	648	-0.0467	1.0922
940334	21538	3506	21534	3566	603	0.0030	0.9666

Table I.12 Divided Arterial Weekly and Annual Afternoon-ADT *t*-test and *F*-test

Station ID	Weekly Stat		Annual Stat			*t*-test	*F*-test
	Mean	Std	Mean	Std	Size		
30094	19618	4119	19468	4172	553	0.0946	0.9749
80283	9800	1447	9748	1467	648	0.0931	0.9735
80294	5230	1247	5381	1236	591	-0.3228	1.0185
100162	18762	5693	18749	5611	472	0.0061	1.0292
100321	25289	3530	25472	3427	657	-0.1405	1.0610
100339	26540	2697	25957	3189	155	0.4757	0.7156
110177	18173	2674	18275	2956	593	-0.0908	0.8185
130180	11801	1769	11335	1771	660	0.6925	0.9977
140199	32949	6238	33853	6208	216	-0.3791	1.0099
150066	10999	1618	11170	1783	614	-0.2526	0.8235
150086	13524	2646	13677	2729	502	-0.1474	0.9402
150295	20703	3757	21510	3775	553	-0.5621	0.9903
150302	15591	3024	15667	2992	676	-0.0668	1.0215
160128	11186	1593	10935	1639	442	0.4022	0.9448
169927	6325	1332	6199	1382	398	0.2399	0.9298
170181	16469	2232	16370	2361	642	0.1104	0.8935
260185	12853	2795	12902	2784	626	-0.0463	1.0078
260323	17940	3102	16776	3094	692	0.9905	1.0056
290286	4873	1041	4922	1028	613	-0.1244	1.0249
360249	14186	4150	14528	4133	544	-0.2175	1.0083
360264	12386	2340	12478	2325	561	-0.1040	1.0130
460166	5097	2087	5826	2164	563	-0.8857	0.9303
460308	14154	3123	14096	3102	718	0.0492	1.0137
480282	14950	2833	14714	2977	655	0.2087	0.9055
530117	8892	1905	8855	1845	305	0.0523	1.0659
550151	19535	4380	20448	4384	507	-0.5473	0.9980
550207	7540	2300	7384	2428	263	0.1674	0.8977
550209	8006	1525	7978	1546	529	0.0467	0.9719
550226	12516	2636	12602	2725	567	-0.0830	0.9357
570167	11890	1618	12088	1570	588	-0.3317	1.0622
570250	16726	3346	16787	3349	373	-0.0477	0.9979
570293	25433	3761	24151	3877	577	0.8698	0.9408
580261	27456	4994	27624	4977	698	-0.0889	1.0068
700113	23565	3841	23506	3883	446	0.0399	0.9786
700114	7425	1252	7350	1245	573	0.1572	1.0110
710189	12893	2237	12421	2266	134	0.5375	0.9748
920265	31934	2935	32314	3099	533	-0.3225	0.8970
930010	16105	3223	15906	3184	481	0.1642	1.0249
930101	18342	1701	17882	1772	632	0.6834	0.9221

(**Table II.12** continued)

Station ID	Weekly Stat		Annual Stat			*t*-test	*F*-test
	Mean	Std	Mean	Std	Size		
720062	19256	2440	19178	2538	471	0.0807	0.9237
720161	26026	7822	26587	7818	653	-0.1888	1.0009
720172	15647	1576	15803	1575	605	-0.2606	1.0022
740182	7567	1556	7695	1586	585	-0.2129	0.9627
750038	15130	1581	15184	1570	631	-0.0905	1.0138
750154	31081	3684	30887	3709	562	0.1376	0.9867
750175	21242	5772	19032	5903	538	0.9845	0.9561
770102	23097	4191	23071	4243	531	0.0161	0.9756
770197	24699	4242	24623	4202	502	0.0475	1.0195
780311	16120	2244	15987	2231	717	0.1570	1.0113
780329	11285	1257	10690	1475	553	1.0619	0.7260
799929	6458	1010	6455	1076	301	0.0068	0.8823
860150	29940	3594	29458	3617	689	0.3508	0.9870
860176	16590	1918	15893	2002	664	0.9166	0.9179
860214	20273	3482	20165	3211	628	0.0884	1.1760
860222	31262	4232	31937	4207	670	-0.4223	1.0121
860298	25902	4090	25452	4104	688	0.2887	0.9932
860306	7786	1153	7846	1159	570	-0.1343	0.9892
870031	42922	3961	43834	4010	626	-0.5985	0.9759
870096	13249	4047	13040	3905	481	0.1405	1.0741
870178	42528	2668	41423	2711	663	1.0729	0.9681
870188	23111	1929	23061	1861	615	0.0707	1.0744
870193	16172	1720	15589	1653	699	0.9282	1.0830
870258	13955	1521	13977	1505	596	-0.0384	1.0214
870266	10218	1421	10242	1413	656	-0.0447	1.0115
879930	30265	5027	28275	5078	599	1.0310	0.9803
880314	30265	5027	28275	5078	698	1.0319	0.9803
890332	12822	2447	12997	2505	580	-0.1838	0.9545
900164	31888	7445	32454	7428	663	-0.2005	1.0046
900165	19421	3684	19477	3712	633	-0.0397	0.9847
920265	31934	2935	32314	3099	533	-0.3225	0.8970
930010	16105	3223	15906	3184	481	0.1642	1.0249
930101	18342	1701	17882	1772	632	0.6834	0.9221

Table I.13 Undivided Arterial Weekly and Annual Afternoon-ADT *t*-test and *F*-test

Station ID	Weekly Stat		Annual Stat			*t*-test	*F*-test
	Mean	Std	Mean	Std	Size		
110246	18342	1701	17882	1772	596	0.6832	0.9221
460315	6669	2230	6650	2243	658	0.0225	0.9887
550201	5427	431	5324	435	325	0.6145	0.9806
550300	3851	454	3778	458	616	0.4161	0.9819
559908	5318	408	5348	417	397	-0.1914	0.9574
890259	3682	841	3682	844	606	0.0000	0.9931

Table I.14 Collectors Weekly and Annual Afternoon-ADT *t*-test and *F*-test

Station ID	Weekly Stat		Annual Stat			*t*-test	*F*-test
	Mean	Std	Mean	Std	Size		
160275	5003	1086	4808	1091	637	0.4709	0.9907
550206	1894	180	1901	180	691	-0.0906	0.9978
550208	6765	2069	6713	2065	656	0.0655	1.0037
550212	8889	2025	8856	2030	674	0.0425	0.9948
550213	1391	200	1318	195	632	0.9853	1.0455
860215	7523	1049	7574	1146	671	-0.1165	0.8371
880326	3901	691	3895	699	680	0.0226	0.9762
930087	8080	2248	8164	2265	712	-0.0983	0.9858

Table I.15 Toll Facility Weekly and Annual Afternoon-ADT *t*-test and *F*-test

Station ID	Weekly Stat		Annual Stat			*t*-test	*F*-test
	Mean	Std	Mean	Std	Size		
970403	39865	10163	39147	9748	150	0.1902	1.0871
970410	29446	6928	29462	7107	404	-0.0059	0.9504
970413	28782	6887	28798	7058	337	-0.0059	0.9522
970416	17821	4527	17580	4487	639	0.1413	1.0180
970417	13758	4146	13815	4135	666	-0.0363	1.0052
970430	14192	2950	14276	2950	594	-0.0749	0.9997
979913	13724	3952	14407	4079	225	-0.4366	0.9384
979933	30354	5501	30043	5495	652	0.1489	1.0023

Appendix II. Derivatives for the MNL and NL models

The Multinomial Logit (MNL) model

The first derivative of $P\left(i\middle|\mathbf{x}_g\right)$ with respect to x_r is:

$$\frac{\partial}{\partial x_r}P\left(i\middle|\mathbf{x}_g\right)=\frac{\beta_r e^{\boldsymbol{\beta}'\mathbf{x}_{ig}}\sum_{l=1,l\neq i}^{C_g}e^{\boldsymbol{\beta}'\mathbf{x}_{lg}}}{\left(\sum_{C_g}e^{\boldsymbol{\beta}'\mathbf{x}_{lg}}\right)^2}\text{, if } x_r \text{ is the characteristic for mode } i; \quad \text{(II.1)}$$

or

$$\frac{\partial}{\partial x_r}P\left(i\middle|\mathbf{x}_g\right)=\frac{-\beta_r e^{\boldsymbol{\beta}'\left(\mathbf{x}_{ig}+\mathbf{x}_{lg}\right)}}{\left(\sum_{C_g}e^{\boldsymbol{\beta}'\mathbf{x}_{lg}}\right)^2}\text{, if } x_r \text{ is the characteristic for mode } l,\ l\neq i. \quad \text{(II.2)}$$

Likewise, if $x_{r'}$ is the characteristic for mode i, the second derivative of $P\left(i\middle|\mathbf{x}_g\right)$ with respect to x_r and $x_{r'}$ is:

$$\frac{\partial}{\partial x_r\partial x_{r'}}P\left(i\middle|\mathbf{x}_g\right)=\frac{\left(\beta_r\beta_{r'}e^{\boldsymbol{\beta}'\mathbf{x}_{ig}}\sum_{l=1,l\neq i}^{C_g}e^{\boldsymbol{\beta}'\mathbf{x}_{lg}}\right)\left(\sum_{C_g}e^{\boldsymbol{\beta}'\mathbf{x}_{lg}}\right)\left(\sum_{C_g}e^{\boldsymbol{\beta}'\mathbf{x}_{lg}}-2\right)}{\left(\sum_{C_g}e^{\boldsymbol{\beta}'\mathbf{x}_{lg}}\right)^4}. \quad \text{(II.3)}$$

Otherwise, assuming $x_{r'}$ is the characteristic for mode k, Equation (II.3) becomes

$$\frac{\partial}{\partial x_r\partial x_{r'}}P\left(i\middle|\mathbf{x}_g\right)=\frac{\beta_r\beta_{r'}^2e^{\boldsymbol{\beta}'(\mathbf{x}_{ig}+\mathbf{x}_{kg})}-2\beta_{r'}e^{\boldsymbol{\beta}'(\mathbf{x}_{ig}+\mathbf{x}_{kg})}\sum_{C_g}e^{\boldsymbol{\beta}'\mathbf{x}_{lg}}}{\left(\sum_{C_g}e^{\boldsymbol{\beta}'\mathbf{x}_{lg}}\right)^4}. \quad \text{(II.4)}$$

Similarly, if $x_{r'}$ represents one of the characteristics for mode i or l, the second derivative of $P\left(i\middle|\mathbf{x}_g\right)$ with respect to x_r and $x_{r'}$ is:

$$\frac{\partial}{\partial x_r\partial x_{r'}}P\left(i\middle|\mathbf{x}_g\right)=\frac{\beta_r\beta_{r'}e^{\boldsymbol{\beta}'\left(\mathbf{x}_{ig}+\mathbf{x}_{kg}\right)}\sum_{C_g}e^{\boldsymbol{\beta}'\mathbf{x}_{lg}}\left(1-\sum_{C_g}e^{\boldsymbol{\beta}'\mathbf{x}_{lg}}\right)}{\left(\sum_{C_g}e^{\boldsymbol{\beta}'\mathbf{x}_{lg}}\right)^4}; \quad \text{(II.5)}$$

Otherwise,

$$\frac{\partial}{\partial x_r \partial x_{r'}} P\left(i \middle| \mathbf{x}_g\right) = \frac{-\beta_r \beta_{r'} e^{\boldsymbol{\beta}'\left(\mathbf{x}_{ig}+\mathbf{x}_{kg}\right)} \sum_{C_g} e^{\boldsymbol{\beta}'\mathbf{x}_{lg}}}{\left(\sum_{C_g} e^{\boldsymbol{\beta}'\mathbf{x}_{lg}}\right)^4}. \tag{II.6}$$

<u>The Nested Logit (NL) model</u>

$$\begin{aligned}
\frac{\partial P_{L|W|TR}}{\partial \mathrm{x}} &= \frac{\partial}{\partial \mathrm{x}}\left(\frac{e^{U_{LW}}}{e^{U_{LW}} + e^{U_{EW}} + e^{U_{MW}} + e^{U_{TW}}}\right) \\
&= \frac{e^{U_{LW}} \cdot \beta_{\mathrm{LW}}}{e^{U_{LW}} + e^{U_{EW}} + e^{U_{MW}} + e^{U_{TW}}} - \frac{e^{U_{LW}} \cdot (e^{U_{LW}} \cdot \beta_{\mathrm{LW}} + e^{U_{EW}} \cdot \beta_{\mathrm{EW}} + e^{U_{MW}} \cdot \beta_{\mathrm{MW}} + e^{U_{TW}} \cdot \beta_{\mathrm{TW}})}{(e^{U_{LW}} + e^{U_{EW}} + e^{U_{MW}} + e^{U_{TW}})^2} \\
&= P_{L|W|TR}\left(\beta_{\mathrm{LW}} - \frac{e^{U_{LW}} \cdot \beta_{\mathrm{LW}} + e^{U_{EW}} \cdot \beta_{\mathrm{EW}} + e^{U_{MW}} \cdot \beta_{\mathrm{MW}} + e^{U_{TW}} \cdot \beta_{\mathrm{TW}}}{e^{U_{LW}} + e^{U_{EW}} + e^{U_{MW}} + e^{U_{TW}}}\right) \\
&= P_{L|W|TR}\left[\frac{e^{U_{EW}}(\beta_{\mathrm{LW}} - \beta_{\mathrm{EW}}) + e^{U_{MW}}(\beta_{\mathrm{LW}} - \beta_{\mathrm{MW}}) + e^{U_{TW}}(\beta_{\mathrm{LW}} - \beta_{\mathrm{TW}})}{e^{U_{LW}} + e^{U_{EW}} + e^{U_{MW}} + e^{U_{TW}}}\right]
\end{aligned} \tag{II.7}$$

$$\begin{aligned}
\frac{\partial P_{W|TR}}{\partial \mathrm{x}} &= \frac{\partial}{\partial \mathrm{x}}\left(\frac{e^{\tau_{WK} I_{WK}}}{e^{\tau_{WK} I_{WK}} + e^{\tau_{DV} I_{DV}}}\right) \\
&= \frac{\tau_{WK} e^{\tau_{WK} I_{WK}} \frac{\partial I_{WK}}{\partial \mathrm{x}}}{e^{\tau_{WK} I_{WK}} + e^{\tau_{DV} I_{DV}}} - \frac{e^{\tau_{WK} I_{WK}} (\tau_{WK} e^{\tau_{WK} I_{WK}} \frac{\partial I_{WK}}{\partial \mathrm{x}} + \tau_{DV} e^{\tau_{DV} I_{DV}} \frac{\partial I_{DV}}{\partial \mathrm{x}})}{(e^{\tau_{WK} I_{WK}} + e^{\tau_{DV} I_{DV}})^2} \\
&= P_{W|TR}\left(\tau_{WK} \frac{\partial I_{WK}}{\partial \mathrm{x}} - \frac{\tau_{WK} e^{\tau_{WK} I_{WK}} \frac{\partial I_{WK}}{\partial \mathrm{x}} + \tau_{DV} e^{\tau_{DV} I_{DV}} \frac{\partial I_{DV}}{\partial \mathrm{x}}}{e^{\tau_{WK} I_{WK}} + e^{\tau_{DV} I_{DV}}}\right) \\
&= P_{W|TR} \cdot e^{\tau_{DV} I_{DV}} \left(\frac{\tau_{WK} \frac{\partial I_{WK}}{\partial \mathrm{x}} - \tau_{DV} \frac{\partial I_{DV}}{\partial \mathrm{x}}}{e^{\tau_{WK} I_{WK}} + e^{\tau_{DV} I_{DV}}}\right)
\end{aligned} \tag{II.8}$$

$$\begin{aligned}
\frac{\partial P_{W|TR}}{\partial \mathrm{x}} &= \frac{\partial}{\partial \mathrm{x}}\left(\frac{e^{U_{WK}}}{e^{U_{WK}} + e^{U_{DV}}}\right) = \frac{e^{U_{WK}} \beta_{\mathrm{WK}}}{e^{U_{WK}} + e^{U_{DV}}} - \frac{e^{U_{WK}} (e^{U_{WK}} \beta_{\mathrm{WK}} + e^{U_{DV}} \beta_{\mathrm{DV}})}{(e^{U_{WK}} + e^{U_{DV}})^2} \\
&= P_{W|TR}\left(\beta_{\mathrm{WK}} - \frac{e^{U_{WK}} \beta_{\mathrm{WK}} + e^{U_{DV}} \beta_{\mathrm{DV}}}{e^{U_{WK}} + e^{U_{DV}}}\right) \\
&= P_{W|TR} \cdot e^{U_{DV}} \frac{\beta_{\mathrm{WK}} - \beta_{\mathrm{DV}}}{e^{U_{WK}} + e^{U_{DV}}}
\end{aligned} \tag{II.9}$$

$$\begin{aligned}
\frac{\partial P_{TR}}{\partial \mathrm{x}} &= \frac{\partial}{\partial \mathrm{x}}\left(\frac{e^{U_{TR}}}{e^{U_{TR}} + e^{U_{HY}}}\right) \\
&= \frac{e^{U_{TR}} \beta_{\mathrm{TR}}}{e^{U_{TR}} + e^{U_{HY}}} - \frac{e^{U_{TR}} (e^{U_{TR}} \beta_{\mathrm{TR}} + e^{U_{HY}} \beta_{\mathrm{HY}})}{(e^{U_{TR}} + e^{U_{HY}})^2} = P_{TR} \cdot P_{HY} \cdot (\beta_{\mathrm{TR}} - \beta_{\mathrm{HY}})
\end{aligned} \tag{II.10}$$

$$
\begin{aligned}
\frac{\partial P_{LW}}{\partial \mathrm{x}} &= \frac{\partial P_{L|W|TR}}{\partial \mathrm{x}} \cdot P_{W|TR} \cdot P_{TR} + P_{L|W|TR} \cdot \frac{\partial P_{W|TR}}{\partial \mathrm{x}} \cdot P_{TR} + P_{L|W|TR} \cdot P_{W|TR} \cdot \frac{\partial P_{TR}}{\partial \mathrm{x}} \\
&= P_{L|W|TR}\left[\frac{e^{U_{EW}}(\beta_{\mathrm{LW}}-\beta_{\mathrm{EW}}) + e^{U_{MW}}(\beta_{\mathrm{LW}}-\beta_{\mathrm{MW}}) + e^{U_{TW}}(\beta_{\mathrm{LW}}-\beta_{\mathrm{TW}})}{e^{U_{LW}} + e^{U_{EW}} + e^{U_{MW}} + e^{U_{TW}}}\right] \cdot P_{W|TR} \cdot P_{TR} + \\
&\qquad P_{L|W|TR} \cdot P_{W|TR} \cdot e^{U_{DV}} \frac{\beta_{\mathrm{WK}}-\beta_{\mathrm{DV}}}{e^{U_{WK}} + e^{U_{DV}}} \cdot P_{TR} + P_{L|W|TR} \cdot P_{W|TR} \cdot \frac{\partial P_{TR}}{\partial \mathrm{x}} \\
&= P_{LW}\left[\frac{e^{U_{EW}}(\beta_{\mathrm{LW}}-\beta_{\mathrm{EW}}) + e^{U_{MW}}(\beta_{\mathrm{LW}}-\beta_{\mathrm{MW}}) + e^{U_{TW}}(\beta_{\mathrm{LW}}-\beta_{\mathrm{TW}})}{e^{U_{LW}} + e^{U_{EW}} + e^{U_{MW}} + e^{U_{TW}}}\right] + \\
&\qquad P_{LW} \cdot e^{U_{DV}} \frac{\beta_{\mathrm{WK}}-\beta_{\mathrm{DV}}}{e^{U_{WK}} + e^{U_{DV}}} + P_{L|W|TR} \cdot P_{W|TR} \cdot P_{TR} \cdot P_{HY} \cdot (\beta_{\mathrm{TR}}-\beta_{\mathrm{HY}}) \qquad \text{(II.11)} \\
&= P_{LW}\left[P_{EW|W|TR}(\beta_{\mathrm{LW}}-\beta_{\mathrm{EW}}) + P_{MW|W|TR}(\beta_{\mathrm{LW}}-\beta_{\mathrm{MW}}) + P_{TW|W|TR}(\beta_{\mathrm{LW}}-\beta_{\mathrm{TW}})\right] \\
&\qquad P_{LW} \cdot P_{D|TR}(\beta_{\mathrm{WK}}-\beta_{\mathrm{DV}}) + P_{LW} \cdot P_{HY} \cdot (\beta_{\mathrm{TR}}-\beta_{\mathrm{HY}}) \\
&= P_{LW}[P_{EW|W|TR}(\beta_{\mathrm{LW}}-\beta_{\mathrm{EW}}) + P_{MW|W|TR}(\beta_{\mathrm{LW}}-\beta_{\mathrm{MW}}) + P_{TW|W|TR}(\beta_{\mathrm{LW}}-\beta_{\mathrm{TW}}) + \\
&\qquad P_{D|TR}(\beta_{\mathrm{WK}}-\beta_{\mathrm{DV}}) + P_{HY} \cdot (\beta_{\mathrm{TR}}-\beta_{\mathrm{HY}})]
\end{aligned}
$$

The second derivatives of the NL model are omitted here due to its complexity.