1

National Library
of Canada

Bibliothèque nationale
du Canada

CANADIAN THESES
ON MICROFICHE

THÈSES CANADIENNES
SUR MICROFICHE

44430

NAME OF AUTHOR/NOM DE L'AUTEUR LOUIS A.Y. SHALLAL

TITLE OF THESIS/TITRE DE LA THÈSE "Peak Hour Ratios Estimation, Application and Cost-Effectiveness
in the Urban Transportation Planning Process"

UNIVERSITY/UNIVERSITÉ Carleton University

DEGREE FOR WHICH THESIS WAS PRESENTED/
GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE Ph.D.

YEAR THIS DEGREE CONFERRED/ANNÉE D'OBTENTION DE CE DEGRÉ 1979

NAME OF SUPERVISOR/NOM DU DIRECTEUR DE THÈSE Ata M. Khan

Permission is hereby granted to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈQUE NATIONALE DU CANADA de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

L'auteur se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans l'autorisation écrite de l'auteur.

DATED/DATE June 5 / 1979 SIGNED/SIGNÉ Louis Shallal

PERMANENT ADDRESS/RÉSIDENCE FIXE 175 Bronson Apt. 1108

Ottawa, Ont. K1R6H4

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

PEAK HOUR RATIOS ESTIMATION, APPLICATION
AND COST-EFFECTIVENESS IN
THE URBAN TRANSPORTATION PLANNING PROCESS

by

• Louis A. Y. Shallal, P.Eng.
B.Sc. (Civil Engineering)
M.Sc. (Civil Engineering)

A thesis submitted to the Faculty of Graduate Studies
and Research in partial fulfillment of the requirement
for the degree of Doctor of Philosophy

Faculty of Engineering

Carleton University

Ottawa, Canada

April 1979

The undersigned recommend to the Faculty of Graduate Studies and Research acceptance of the thesis, "Peak Hour Ratios Estimation, Application and Cost-Effectiveness in the Urban Transportation Planning Process", submitted by Louis A. Y. Shallal, P.Eng., B.Sc. (Civil Engineering), M.Sc. (Civil Engineering) in partial fulfillment of the requirement for the degree of Doctor of Philosophy.

Ata M. Khan

Thesis Supervisor

John A. G. G. G.

Chairman
Department of Civil Engineering

ABSTRACT

In the course of most urban transportation planning studies, the transportation planner is faced with the task of determining which factors to use in order to convert 24 hour travel forecasts into peak hour forecasts. An extensive review of the literature reveals that the majority of planners opt for the highly simplified practice of applying uniform area-wide peak hour ratios to make this conversion.

The objective of this thesis is to develop an improved and cost-effective methodology for the estimation and application of peak hour ratios in the urban transportation planning process. Specifically, the research highlights both the importance of peak hour ratios in the urban transportation planning process and the ramifications of their improper application in terms of under or over-designing urban transportation facilities; reviews the existing practices of establishing peak hour ratios and identifies their deficiencies; develops and recommends improved approaches for peak hour ratios estimation and application; and investigates the cost-effectiveness of the proposed approach by employing Bayesian Statistical decision theory.

The research incorporates a modelling framework which focuses on the investigation of various determinants of peak hour ratios at different levels of analysis ranging from the disaggregate individual level to the aggregate district level. Actual data from an on-going urban transportation planning study are used in the analysis. Several statistical techniques such as factor analysis, category analysis and stepwise regression analysis are employed to relate peak hour ratios

to their determinants. Relationships are successfully derived for the important work travel component by mode. Zonal population, net residential density, total employment and a measure of travel time by mode are found to be the most significant variables for estimating zonal peak hour ratios for work travel.

The modelling framework culminates with recommendations for estimating peak hour ratios and applying them at different stages of the Urban Transportation Planning System Model. Specifically, improved approaches are recommended for applying peak hour ratios post the trip distribution and traffic assignment steps.

The cost-effectiveness of the recommended approaches is demonstrated using a decision theoretic framework. The framework incorporates a simple one stage design decision process. Four case studies are used and the design decisions based on decision theory are found to differ from the actual designs built in three of the four cases.

The revised methodology is practical, conceptually attractive and provides statistically significant results. Adopting this methodology will increase the reliability of peak hour volume estimates over the current practice. This in turn, will facilitate better transportation design decisions.

ACKNOWLEDGEMENTS

The research reported in this thesis was carried out under the supervision of a Committee composed of Professors A. M. Khan (Chairman), J. Braaksma, A. P. S. Selvadurai, B. Pagurek, A. Maslove, P. Tan and J. Salinas. The author wishes to express his sincere appreciation to the committee for guidance throughout the conduct of this thesis research.

The research was carried out while the author was employed as Head, Systems Planning Branch of the Regional Municipality of Ottawa-Carleton. Support for the research was gratefully received from the Transportation Department of the Regional Municipality in the form of access to information and data processing. The author wishes to thank Mr. M. Sheflin, Transportation Commissioner and Mr. J. Bonsall, Director Transportation Planning, for their help and cooperation. Valuable comments and encouragement were also gratefully received from the Study Team of the TRANS (Transportation Model Development Programme), a joint RMOC/NCC/CRO Project.

Financial support from the National Research Council of Canada operating grant number A7425 which enabled the author to acquire computer programming assistance is gratefully acknowledged.

The author wishes to express his sincere thanks to Mr. F. Williams and Mr. P. Beddall for their help in developing the computer programmes; to Mr. P. Ranniste, P. Mallette and R. Claude all of the Systems Planning Branch for their help in executing programmes and developing some of the input data. Special thanks to Mrs. L. Coleman for her efficient typing.

TABLE OF CONTENTS

	<u>Page</u>
ABSTRACT	i
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	iv
LIST OF FIGURES	ix
LIST OF TABLES	xiii
CHAPTER I.	
INTRODUCTION	
1.1 Introduction	1
1.2 Context of this Research	2
1.3 Objective and Scope of Investigation	5
1.4 Organization of Thesis	6
CHAPTER II.	
THE POSITION OF THE PEAK HOUR RATIOS IN THE URBAN TRANSPORTATION PLANNING PROCESS (UTPP)	
2.1 Introduction	9
2.2 Structure of the UTPP	10
2.3 Model Building in the UTPP	13
2.3.1 Early Models	14
2.3.2 Second Period Models	16
2.3.3 Third Period Models	19
2.4 Peak Hour Models vs. 24 Hour Models	20
CHAPTER III.	
EXISTING APPROACHES TO PEAK HOUR RATIOS ESTIMATION AND APPLICATION	
3.1 Introduction	23
3.2 Existing Approaches to Peak Hour Ratios Application	25
3.3 Use of Peak Hour Ratios in Second Generation Models	36
3.4 Deficiencies of Existing Approaches	39

	<u>Page</u>
CHAPTER IV	
A FRAMEWORK FOR MODELLING THE PEAK HOUR RATIOS	
4.1 Introduction	45
4.2 Hypothesis Formulation	49
4.3 Alternative Approaches for Peak Hour Ratios Estimation and Application	50
4.3.1 Zonal Trip Ends Approach	53
4.3.2 Zonal Trip Interchange Approach	56
4.3.3 District Level Peak Hour Ratios	58
4.3.4 Concentrated Work Trips Approach	58
4.3.5 Disaggregate Approach	60
4.4 Hypothesis Testing	61
CHAPTER V.	
DEVELOPMENT OF INPUT DATA	
5.1 Introduction	62
5.2 Sources of Data	63
5.3 Development of Independent Variables	64
5.4 Development of Dependent Variables	67
5.4.1 Zonal Level Peak Hour Ratios	67
5.4.2 District Level Peak Hour Ratios	76
5.4.3 Disaggregate Household and Individual Level	81
5.5 Input Data for Decision Analysis	82
5.6 Notes on the Definition of the Peak Hour	85
5.6.1 Hourly Variation of Travel	85
5.6.2 Methods of Defining the Peak Hour	88
5.6.3 Defining Peak Hour Trips-In-Motion	90
CHAPTER VI.	
STATISTICAL ESTIMATION OF PEAK HOUR RATIOS	
6.1 Introduction	94
6.2 Preliminary Analysis	96
6.2.1 Correlation Analysis	104
6.2.2 Scattergram Analysis	105

	<u>Page</u>
6.3 Canonical Correlation Analysis	108
6.4 Factor Analysis	110
6.4.1 Factor Analysis for the Dependent Variables	114
6.4.2 Factor Analysis for the Independent Variables	119
6.5 Step-Wise-Regression Analysis	132
6.5.1 Linear Regression	133
6.5.2 Non-Linear Regression	137
6.6 Category Analysis	144
6.6.1 Household Level Category Analysis	147
6.6.2 Individual Level Category Analysis	152
6.7 Analysis of the Peak Hour Ratios from Ground Counts	153

CHAPTER VII.

DISCUSSION OF PEAK HOUR RATIOS ESTIMATION EQUATIONS

7.1 Introduction	168
7.2 Data Limitations	169
7.3 Testing of Peak Hour Ratios Estimation Equations	170
7.4 Recommended Approach(es) for Peak Hour Ratios Estimation and Application	172
7.5 Other Considerations in Peak Hour Ratio Estimation and Application	186

CHAPTER VIII.

ECONOMIC BASIS FOR PEAK HOUR RATIOS ESTIMATION AND APPLICATION

8.1 Introduction	191
8.2 Link Peak Hour Volume Estimation	194
8.3 Economics of Error in Link-Volume Estimation	200

	<u>Page</u>
8.4 A Framework for Design Decisions	205
8.4.1 Problem Statement	206
8.4.2 Existing Approaches to Design Decisions and Their Deficiencies	207
8.4.3 Bayesian Decision Theory Approach	210
8.4.4 A Framework for the Design Decision Process	213
8.4.5 Expected Value of Information	220
8.5 Application of the Decision Theoretic Framework	220
8.5.1 Selecting the Options	223
8.5.2 Prior Probabilities	227
8.5.3 Conditional Probabilities	235
8.5.4 Cost Functions for Alternative Actions	237
8.5.5 Case Studies	246
8.6 Sensitivity Analysis	248
CHAPTER IX	
SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	
9.1 Introduction	255
9.2 A Framework For Future Research	256
9.3 Summary	261
9.4 Conclusions	264
9.5 Recommendations for Further Research	266
REFERENCES	270
APPENDIXES	
A. LIST OF NOTATIONS AND ABBREVIATIONS	275
B. LISTING OF VARIABLES WITHIN EACH DATA BANK	278
C. LISTING OF A PROGRAMME TO CALCULATE CAR AND TRANSIT ACCESSIBILITY INDEXES	283
D. BRIEF LISTING OF COMPUTER PROGRAMMES USED IN THE RESEARCH	286

	<u>Page</u>
E. LISTING OF THE TRIP-IN-MOTION PROGRAMME	289
F. FACTOR ANALYSIS	292
G. LISTING OF ZONAL LINEAR AND NON-LINEAR REGRESSIONS NOT IN THE MAIN TEXT	301
H. BAYESIAN DECISION THEORY	308
I. BAYESIAN DECISION THEORY PROGRAMME	319
J. LISTING OF A TYPICAL EXECUTION AND RESULTS OF THE BAYESIAN DECISION THEORY PROGRAMME	336

LIST OF FIGURES

<u>Figure</u>	<u>Description</u>	<u>Page</u>
2.1	The Urban Transportation/Land Use Planning Process	12
2.2	Models Classification	15
3.1	A Simplified Version of the Model Used in the Ottawa-Carleton Regional Transportation Study (O.C.R.T.S.)	24
3.2	Peak Hour Model Vs 24 Hour Model Application	26
3.3	Existing Alternative Approaches to the Use of Peak Hour Ratios in the UTPS Model	37
4.1	A Framework for Modelling the Peak Hour Ratios in the UTPP	47
4.2	Ottawa-Carleton Regional Transportation Study Area	48
4.3	Alternative Approaches for Peak Hour Ratios Estimation and Application in the UTPS Model	54
5.1	A Simplified Flowchart For the Development of Dependent Variables	68
5.2	The 300 Zones System	73
5.3	The 60 Districts System	79
5.4	Approaches to Peak Hour Definition	86
6.1	Statistical Estimation of Peak Hour Ratios	97
6.2	Typical Output from Scattergram Analysis	107

<u>Figure</u>	<u>Description</u>	<u>Page</u>
6.3	Rotated Factors From Factor Analysis for the Dependent Variables	118
6.4	Rotated Factors From Factor Analysis for the Independent Variables	124
6.5	Household Peak Hour Ratios Distribution Vs Vehicle Ownership	149
6.6	Household Peak Hour Ratios Distribution Vs Household Occupancy	150
6.7	Household Peak Hour Ratios Distribution Vs Household Income	151
6.8	Frequency Distribution of Total Individual and Household Peak Hour Ratios	154
6.9	Location of Stations for Ground Class- ification Counts	156
6.10	Geographic Areas Used in Ground Counts Analysis	159
6.11	Time of Occurrence of Highest 15 Minutes Volume on Screenlines by Distance from CBD	160
6.12	Variation in Vehicle Based Peak Hour From 1976 Ground Counts of 182 Stations	162
6.13	Variation in Person Based Peak Hour Ratios From 1976 Ground Counts of 169 Stations	163
6.14	Variation of Peak Hour Ratios by Facility Type	165
7.1	Simulated Vs Observed Origin Zonal Peak Hour Ratios Home-Work, Driver plus Passenger (HIDPHWRK)	171
7.2	Simulated Vs Observed Origin Zonal Peak Hour Ratios Home-Work, Transit (HITRANHWR)	173

<u>Figure</u>	<u>Description</u>	<u>Page</u>
7.3	Testing the Recommended Peak Hour Ratios	174
7.4	Summary of Alternative Approaches Recommended for Peak Hour Ratios Estimation and Application	185
7.5	Geographical Distribution of Cities Analyzed in Time of Day Study	188
7.6	Percentage of Total Travel in P.M. Peak Period (2-8 P.M.) Highest Hour	189
8.1	Interrelationships between the Various Steps of the Modelling Process Past Trip Distribution	196
8.2	Schematic Illustration of the Variation in Total System Cost by Different DHV Decisions	202
8.3	A Framework for the Design-Decision Process	214
8.4	Relationship between Marginal and Conditional Probabilities	215
8.5	A Simple Decision Tree	217
8.6	Preposterior Analysis	219
8.7	Expected Value of Sample Information	221
8.8	Application of the Decision Analysis Framework	222
8.9	Rectangular Prior Probability Distribution	228
8.10	Prior Probability Distribution Function Based on 1968/1978 Data	232
8.11	Prior Probability Distribution Function Subdivided by 5 Ranges of Demand States	233

<u>Figure</u>	<u>Description</u>	<u>Page</u>
• 8.12	Cost Functions Derived From Non-Linear Regressions	243
8.13	Cost in Millions of Dollars/Km Vs Peak Hour Volume	245
8.14	Sensitivity Analysis of the Variation in the Prior Probability Distribution Function	249
8.15	Sensitivity Analysis of the Variation in the Range of Demand States	251
8.16	Sensitivity Analysis of the Variation in Link Length	252
9.1	A Framework For Future Research	259
9.2	Statistical Estimation Sub-Framework	262
9.3	Sub-Framework for Investigating the Cost-Effectiveness of Researching the Element L	263

LIST OF TABLES

<u>Table</u>	<u>Description</u>	<u>Page</u>
3.1	An Example Illustrating the Impact of Variations in Peak Hour Ratios and Directional Split Factors on Facility Design	30
4.1	Determinants of the Peak Hour Ratios	51
5.1	Data Bank	65
5.2	Sources and/or Method of Development of Zonal Variables	66
5.3	Conversion Table: 300 Zones to 60 Districts	78
5.4	Summary of All Data Banks Developed in the Research	84
5.5	Approaches to Peak Hour Definition	89
5.6	Example of Trip-In-Motion Concept	91
5.7	Comparison of Peak Hour Ratios by Various Definitions of Peak Hour	92
6.1	Condensive Statistics for the 20 Dependent Variables	99
6.2	Typical Output from Pearson Correlation Analysis	106
6.3	Results of the Canonical Correlation Analysis	111
6.4	Varimax Matrix and Factor Score Matrix for an Illustrative Example of Factor Analysis	115
6.5	Varimax Matrix for the Dependent Variables	117

<u>Table</u>	<u>Description</u>	<u>Page</u>
6.6	Independent Variables Used as Input to Factor Analysis	120
6.7	Percent of Variance Explained by the Five Factors	119
6.8	Varimax Matrix for the Independent Variables	121
6.9	Zonal Level Factor Analysis	122
6.10	Zonal Level Correlation Analysis	128
6.11	Varimax Matrix for District Level Analysis	129
6.12	District Level Factor Analysis	130
6.13	District Level Correlation Analysis	131
6.14	Correlation Coefficients of Variables Used in Zonal Level Regressions	135
6.15	Correlation Coefficients of Variables Used in District Level Regressions	136
6.16	Some Results of Zonal Level Linear Regressions	138
6.17	Some Results of District Level Linear Regression	142
6.18	Some Results of Zonal and District Level Non-Linear Regressions	145
6.19	Typical Information Developed from Ground Counts	157
6.20	Screenline Peak Hour Ratios Data Bank	158

<u>Table</u>	<u>Description</u>	<u>Page</u>
6.21	Testing Variation of Peak Hour Ratios Distribution by Classifications	164
6.22	Results of Regression Analysis for Ground Counts Data	166
7.1	Recommended Peak Hour Ratios by Mode and Purpose for Non-Work Travel	177
7.2	Recommended Peak Hour Ratios Equation for Home-Work (Transit) HITRANHW	178
7.3	Recommended Peak Hour Ratios Equation for Home-Work (Driver and Passenger) HIDPHWRK	179
7.4	Recommended Peak Hour Ratios Equation for Home-Work (Transit) HJTRANHW	182
7.5	Recommended Peak Hour Ratios Equation for Home-Work (Driver Plus Passenger) HJDPHWRK	183
8.1	Estimates of Total Experimentation Cost by Option	226
8.2	D.H.V. Analysis Data For Urban and Semi-Urban Sections	230
8.3	Conditional Probability Measures $P(v t,e)$ by Experiment Alternatives	236
8.4	Cost Data for the Cost Curves	239
8.5	Results of Decision Analysis Application for 4 Case Studies	247
8.6	Summary of Decision Analysis Runs Under Various Input Assumptions	253

CHAPTER I

INTRODUCTION

1.1 Introduction

The conversion of daily travel forecasts into peak hour forecasts is a task which is generally experienced in the course of most practical urban transportation planning studies. This conversion is accomplished through the application of the so-called Peak Hour Ratios. These ratios are applied in order to arrive at an estimate of future peak hour volume at the corridor or facility level. In turn, this estimate is used in the functional planning and design stages.

The need for peak hour travel information is self-evident. It is indeed surprising that while tremendous efforts are expended in the development of 24 hour travel information, the conversion to peak hour travel in most practical studies is handled through the application of a simple area-wide ratio. This observation also holds for the treatment of the directional split factors.

The background for the above problem dates to the early periods in the development of structured transportation planning techniques. The main concern at that time was to develop tools that could be used for generating 24 hour travel forecasts. A succession of urban transportation planning studies refined these techniques by means of collecting finer disaggregated data on land use, the transportation system and the trip maker.

Significant changes in the planning process have taken place since that time. The most dramatic, of course, is the use of electronic computers throughout all phases of processing, storing, retrieval and analysis of data and for system simulation. The planning process is now not only iterative and continuous, but also variable in terms of emphasis on the macro or micro analysis levels.

Another dramatic shift in emphasis is from long range planning to short range planning complemented with concentration on specific rather than system-wide issues. The recent history of transportation planning has witnessed dramatic growth in the development and application of computerized modelling techniques. Nevertheless, the tradition of constructing 24 hour based transportation models continues to dominate the practical field and the application of area-wide peak hour ratios continues to be the common practice.

1.2 Context of This Research

The selection of a peak hour ratio for the conversion of estimates of daily travel into peak hour travel is an important and significant step in the urban transportation planning process. It has a direct bearing on the subsequent functional planning and design steps. The major issue is the determination of a suitable set of ratios, within the accuracy bounds of the planning process, that will minimize the probable under-designing or over-designing of transportation facilities.

The establishment of these peak hour ratios is generally an integral element of the systems planning process. However, as will be discussed later, there has been some tendencies in the past to establish these ratios as part of the functional planning and/or design processes. This points out the need for a proper definition and integration of these processes.

The application of area-wide peak hour ratios is probably adequate at the conceptual planning level. This is not the case, however, at the detailed planning level because of the errors that may be introduced in estimating the probable peak hour volume using a particular facility. The impact of such errors may be quite significant.

Admittedly, the estimation of a probable future peak hour volume on a given facility is a complex task which cannot be determined with certainty. Furthermore, the application of area-wide peak hour ratios is not the sole source of possible errors. In fact, the entire demand forecasting process is probabilistic in nature. This should not suggest that improvements to the various elements of this process are unwarranted. Quite the contrary, the task of the system planner is to refine as much as possible his tools to minimize the range of possible error.

The refinement of the existing approaches for estimating peak hour ratios has a direct and immediate application. These ratios have significant implications on many steps of the travel forecasting procedure. As will be described in Chapter II, their position with respect to other activities of the urban transportation planning process is central. The proper estimation and

application of peak hour ratios are particularly important in bridging the gap that presently exists between the systems planning and functional planning processes. Furthermore, from a purely economic point of view, it should be immediately evident that for a given daily travel demand across a particular transportation facility, the transport supply can vary significantly depending on the peak hour ratios used. As demonstrated in Chapter III, a small variation in the ratio may cause a large error in the design decision. Such an error may have serious economic implications.

Review of numerous working papers and transportation handbooks indicates that little contribution to this important topic has been made. The author believes that a more reliable method of estimating the peak hour ratios can be developed. Most of the data required for developing an improved method are already collected during the urban transportation planning process for reasons other than the peak hour ratios estimation. Additional data used specifically in the development of the peak hour ratios are either readily available or easily assembled.

The basic contentions of this research may therefore be stated as follows:

1. the application of area wide peak hour ratios would result in under-designing or over-designing transportation facilities in an urban region,
2. a more valid approach would be to establish the determinants of peak hour ratios and to develop relationships that can be used to convert the 24 hour travel to peak hour travel,

3. adopting the refined approach would result in a simulation of peak hour conditions at the base year that is more realistic than that obtained using area-wide peak hour ratios. Consequently better peak hour forecasts would be achieved, and
4. the refined approach would be cost-effective because the cost of experimentation and data collection would be lower than the cost of under-designing or over-designing the transportation facilities under consideration.

1.3 Objective and Scope of Investigation

The purpose of this research study is to develop an improved approach to the estimation and application of peak hour ratios in the urban transportation planning process and to demonstrate that this improved approach is cost-effective. In addition to the statistical modelling dimension, the scope of this research encompasses a philosophical dimension which attempts to bridge the gap between the systems planning and functional planning processes.

The major components of the research study are to:

1. outline and define the position of peak hour ratios in the modelling steps of the urban transportation planning process,
2. review the existing practices in estimating and applying the peak hour ratios,
3. identify the deficiencies and problems encountered in applying constant area-wide peak hour ratios,
4. identify the major determinants of peak hour ratios and generate a number of alternative approaches for peak hour ratios estimation and application,

5. employ statistical techniques and develop, using actual data and various new approaches, relationships that explain the variations in peak hour ratios in terms of socio-economic and/or transportation system variables,
6. recommend an improved approach for peak hour ratios estimation and application, and
7. employ Bayesian statistical decision theory to show the cost-effectiveness of the improved approach. This will be demonstrated by identifying the trade-offs between the cost of improved design information and the savings in the facility total cost.

The author believes that this investigation has merits from both academic and practical standpoints. The choice and the scope of this investigation were, in fact, dictated by the above two considerations.

1.4 Organization of Thesis

The position of the peak hour ratios in the urban transportation planning process is described in Chapter II. The structure of the urban transportation planning process is described with emphasis being placed on the model building aspect. The main considerations in deciding whether to construct 24 hour models or peak hour models are briefly discussed. Chapter III describes the use of peak hour ratios in conventional models. The deficiencies of the existing approaches for peak hour ratios estimation and application are presented in some detail, and numerous examples are cited from the literature.

Chapter IV is devoted to developing a framework for modelling the peak hour ratios. The various steps in the simulation process are defined and a brief description of the alternative approaches to modelling the peak hour ratios is presented. A general hypothesis is formulated which outlines the spectrum of the sets of determinants of peak hour ratios. A number of independent variables are identified within each set and the study area is defined. The development of input data for the study area is described in Chapter V. The sources of data are cited and the procedures used for developing the dependent and independent variables are outlined.

Chapter VI describes the simulation and data analysis steps. Statistical estimation techniques such as factor analysis, regression analysis and category analysis are employed to derive relationships between the peak hour ratios and their determinants. Ground counts from over 180 stations are analyzed in an attempt to relate observed peak hour ratios to station attributes. The simulation of peak hour ratios from a travel survey is attempted at different levels ranging from the individual at the micro scale to the district at the macro scale.

Chapter VII provides a discussion of the analysis results. Data problems and limitations are outlined and various approaches are recommended for peak hour ratios estimation and application. Those sets of determinants identified in Chapter IV, but not treated explicitly in the research are briefly discussed.

In Chapter VIII, the transport planner is given the decision-making role. He is presented with a set of options and he is asked to decide whether to experiment or not. This leads to a classical case of decision theory application. In order to resolve this question, the economic basis for peak hour ratios estimation and application are described and case studies are used to illustrate the application of Bayesian decision theory for arriving at the best course of action. The basic trade off is between the cost of experimentation and the cost of error in link volume estimation and thus in link design.

The final chapter briefly summarizes the research and presents conclusions and recommendations for further investigation. A framework based on the research structure described in this thesis is suggested for possible use in future research on other elements of the urban transportation planning system model.

CHAPTER II

THE POSITION OF PEAK HOUR RATIOS IN THE URBAN TRANSPORTATION PLANNING PROCESS

2.1 Introduction

The study of the overall urban transportation planning process (UTPP) is essential to the proper use and application of peak hour ratios. Of particular interest is the model building activity within which these ratios are introduced. This activity is commonly referred to as the Urban Transportation Planning System Model (UTPS).

This chapter outlines the structure of the urban transportation planning process and emphasizes the modelling aspect in order to illustrate the position of peak hour ratios and their interaction with the other elements of the process. In the next few paragraphs, a brief overview of this position is presented.

As stated earlier, most urban transportation planning studies develop travel demand models based on 24 hour conditions (1,2,3).^{*} Although some planners have attempted to build peak period models (4), their modelling efforts were mainly restricted to work travel since it constitutes over 50% of the total travel in the P.M. peak hour.

* References are provided at the end of the thesis.

The development of appropriate peak hour ratios requires extensive origin-destination surveys. Low sampling rates of such surveys; general under-reporting of non-work travel; and estimation errors in reported time of travel all contribute to make the task of determining peak hour ratios rather difficult.

Since one of the end products of the urban transportation planning system model (UTPS) is to develop peak hour travel forecasts which become a basis for designing a particular facility, it seems apparent that the peak hour ratios play an important and central role. Sooner or later in the process, data collected on a 24 hour basis or from other time periods must be manipulated to enable the planner to work with peak hour conditions. This manipulation may be performed early in the process (when origin-destination data are developed and the actual peak hour is determined) or when the 24 hour traffic assignment is converted to peak hour assignment.

The position of the peak hour ratios is in fact at the interface between systems planning and functional planning. Proper estimation and application of these ratios will therefore help to bridge the gap that presently exists between these two planning processes. This can also aid in matching the transport supply to the variable travel demand.

2.2 Structure of the Urban Transportation Planning Process

The urban transportation planning process has been defined as the process of guiding the changing transportation system environment in order to prepare a scheme to provide an urban area with a determined

level of mobility (5). The process is orderly, structured and goal oriented. Implicit in the definition of planning is the desire to guide the inevitable change in the face of uncertainty. Unlike pure science, there is always a certain level of error that cannot be avoided in the planning process.

While evidence in the past suggests that the urban transportation planning process has failed to keep up with its definition, particularly in the area of implementation, its value nevertheless should not be undermined. The process has served and continues to serve as an important tool in the hands of the decision maker.

A review of the theory of transportation planning indicates certain major facets of the transportation planning process. The land-use planning and transportation planning processes are highly interrelated. A composite structure of the transportation/land use planning process is shown in Figure 2.1.

In a typical urban transportation planning study, the process involves the following main activities:

1. goals and objectives formulation,
2. inventory of the existing land-use and transport systems,
3. development of land-use allocation and travel demand models,
4. formulation of alternative transport land-use plans,
5. forecasting, testing, and evaluation of the alternative plans,
6. selection and implementation of the preferred alternative,
7. development of staging priorities, and
8. monitoring the impacts and the continuous update of the study and the plan.

Each of the above activities may be further subdivided into numerous elements (6). However, only the third and fifth activities are of direct concern to this research. They involve the use of mathematical models which are described in the next section. It should be noted at this point that although peak hour ratios are generally applied within the above two activities, they can be indirectly implicit in some of the other activities. For example, in formulating the objectives, the decision-maker may be willing to accept a level of service higher or lower than a given datum. This may be reflected in increasing or decreasing the transport supply (by modifying the transport networks), or by means of policy changes affecting the peak hour travel demand. In effect, these peak hour ratios can be perceived as policy variables in the hands of the decision-maker. This point will be dealt with again as it relates to the application of decision theory.

2.3 Model Building in the Urban Transportation Planning Process

(Most urban transportation planning studies have relied heavily on mathematical models (5,6,7). These models serve three basic functions:

1. improve the planner's comprehension and understanding of the complex transportation system and its environment,
2. provide the planner with a tool to test the various land use and transportation concepts and their interaction, and
3. provide data and information that are essential for developing programmes, priorities and projects.

Transportation models are built to replicate in abstract terms the existing travel patterns and to provide a simple and effective tool for estimating future travel demand. Unlike the deterministic models of physics and science, transportation models are generally of a probabilistic nature, and thus an element of chance or error is always present. The accuracy and strength of such models are not necessarily enhanced by increasing or decreasing the number of variables introduced (8). This should not suggest that these models are useless. Quite the contrary, without these tools, travel demand forecasting would be a crude and frustrating exercise. For this reason, substantial advances in the model building activity have been made during the past two or three decades (9,10,11,12).

The history of transportation planning models is relatively brief. Their development dates from the late 1940's. Nevertheless, three distinct periods in the history of transportation model development may be identified. These are classified in a summary format in Figure 2.2.

2.3.1 Early Models

Transportation models built during the late 1940's and early 1950's were mainly of the factoring or trends type. Existing travel conditions would be examined and extrapolated to the design year. While such models have been abandoned for long range planning, they are still applied in practice for short range planning and in small urban areas experiencing little growth.

FIGURE 2.2

MODELS CLASSIFICATION

Because of the coarse nature of these models, peak hour travel forecasts are rarely developed. The major end products are usually forecasts of average annual daily traffic (AADT). Peak hour ratios, therefore are seldom necessary.

The major short-comings of these models are firstly their inability to take into account the impact of dramatic changes in land use and secondly, their inability to reflect changes in travel patterns resulting from changes in transportation supply and services (5).

2.3.2 Second Period Models

Transportation models development from the mid 1950's to the late 1960's was rather extensive. The early models were refined and structured through a series of transportation studies. From these efforts evolved the so-called Urban Transportation Planning System Models. These second period models (also referred to as second generation models) were adopted by most subsequent urban transportation planning studies. In addition to their structural appeal, the availability of computer packages specifically tailored for these models enhanced their continued use and improvement. Peak hour ratios are treated explicitly in these models. For this reason, a brief overview of their major components is presented below.

Four essential components of second generation models can be easily identified:

1. Trip Generation,
2. Trip Distribution,
3. Modal Split, and
4. Trip Assignment

Trip generation is the process of estimating the total trips produced in a given area or attracted to a given area within a certain time and for a stated trip purpose. Trip generation sub-models are usually based on socio-economic and land use data (2). Recent attempts have been made to incorporate some measure of transportation system service into the trip generation formulation (13).

Two techniques have been used extensively to formulate trip generation relationships. These are Regression Analysis and Category Analysis. In certain cases, when a traffic zone contains a specialized type of land use such as terminals, universities or shopping centres, a Special Generator Analysis may be performed.

Most trip generation relationships developed in practice apply to the 24 hour time period (1,2,3,13). The peak hour ratios may be applied to develop peak hour trip productions and attractions. When such is the case, the subsequent trip distribution, modal split and traffic assignment steps are all based on peak hour conditions.

Trip Distribution is the process of estimating the magnitude of travel between two geographic areas. For each origin and destination pair in a study area, the magnitude of travel is assumed to depend on the characteristics of the origin and destination zones and on some measure of the transportation system activity. The most widely used technique is by far the Gravity Model (14,15). Other trip distribution techniques include the Intervening Opportunity Distribution method; the Growth Factor method; the Detroit method; and the Fratar method. Reference (9) provides an excellent description of these techniques.

Peak hour ratios are usually applied to the 24 hour travel matrices produced by the trip distribution sub-model (1,2,16). A typical example of such applications is shown in Figure 3.1, Chapter III.

Modal Split is the process of estimating the magnitude of travel between various areas by different modes. These sub-models have received particular attention in recent years because of increased emphasis on public transit (11).

Techniques for estimating modal split generally fall into two main categories; regression analysis and diversion curves. In the first technique, the modal split is assumed to be a function of a number of socio-economic, transportation system activity and policy variables. In the second technique, data from travel surveys are stratified to derive some logical indices relating transportation system variables and socio-economic variables to transit usage.

The application of peak hour ratios to travel matrices produced by the modal split sub-model is rarely done in practice (1,3,9,11). As will be demonstrated later on, the introduction of modal peak hour ratios should aid in accounting for modal peaking differences and enhance the overall model results.

Trip assignment is the process of estimating the volume of travel demand on each link of the transportation network and for each mode. The most common practice is to develop separate trip tables by auto and transit and assign them to abstract simulated road and transit networks respectively (16). Several assignment techniques have been developed but the most common are the All-Or-Nothing; Multiple Path; Equilibrium; and Capacity Restraint.

By far, the all or nothing technique has been used most extensively in practice in spite of its limitations. This technique, in general, over-estimates travel on high speed routes since it does not recognize the capacity limitation of the individual links. Most traffic assignment results require manual re-adjustment in order to produce a reasonable flow on the transportation network.

Following the 24 hour travel matrix assignment to the transport network, peak hour ratios may be applied to convert the 24 hour estimates of link volumes into peak hour estimates (16). This is the procedure advocated by various transportation handbooks. As will be shown in Section 6.6, peak hour ratios may vary depending on the particular link type and its geographic location.

2.3.3 Third Period Models

Models developed in this third period are commonly referred to as third generation demand forecast models (17). They began appearing in the literature in the late 1960's and early 1970's. Most of these models have a behavioural basis. They may be grouped according to two major categories, namely the Econometric models and the Disaggregate Stochastic models (18). Econometric models treat travel as an economic commodity subject to the laws of supply and demand. In effect, these type of models introduce economic theories of consumer behaviour and combine them with statistical techniques (19). Because of the disaggregate behavioural basis of these models, they can be developed for peak hour conditions, thus eliminating the need for peak hour ratios application.

2.4 Peak Hour Models vs. 24 Hour Models

The determination of the time period upon which transportation models are based is an important and fundamental decision. It dictates to a large extent the scope and complexity of the modelling procedure. The choices available to the planner are few but distinct. The model may be based, in general, on one or more of the following time periods:

1. 24 hour,
2. a.m. peak hour or peak period,
3. p.m. peak hour or peak period, and
4. off peak hour or off peak period.

The main criterion in the selection of the time period is the objective for which the model is being developed. The other considerations include: the data base; the planning horizon; and the available time and staff resources. These considerations are discussed briefly in the next few paragraphs.

Origin-destination (O-D) travel surveys constitute the main data base of second generation type models. The choice of time period for the model depends on the time duration for which the data base was collected. If the O-D survey collected travel data for a 24 hour period based on a limited small sample (which is the common practice), peak hour models would be rather difficult to construct since the proportion of trips reported in the peak hour may be insufficient to derive statistically valid relationships. If, on the other hand, concentrated surveys for peak period travel are conducted, peak period and/or peak hour models may be developed. This would be easier than expanding the data base of the peak period survey to 24 hours and subsequently building 24 hour models.

For long range planning, it is probably sufficient to build 24 hour models since the ultimate objective of the exercise would be to determine the right-of-way for major corridors in the study area. For short-range planning, however, peak hour models (or 24 hour models with peak hour ratios calibration) are necessary since peak hour volumes are used as an input to functional planning and design.

The task of model development is a demanding task which requires substantial resources. These may impose serious limitations on the choice of the time period(s) of the model. Staff and time may not be available to develop both 24 hour and peak hour models.

Finally, it should be noted that for transit planning and for economic studies, 24 hour models are essential. For transit network design, 24 hour data are used to determine both peak and off peak services. Indications are that it is easier to build 24 hour models and, by using appropriate peak hour ratios, deduce the peak hour volumes than to first build peak hour models and subsequently infer the 24 hour travel demand.

Based on the author's experience and the working reports reviewed, it is believed that 24 hour models are usually necessary in the urban transportation planning process. Although peak hour models are also desirable, they suffer from a number of drawbacks and limitations. These include:

1. the vague and imprecise peak hour definition which may change because of the dynamic peak hour travel conditions,
2. direct simulation of peak hour conditions is difficult since the consistent travel patterns observed over the 24 hour period are not as evident or as consistent in the peak hour, (the short period

fluctuations in peak hour travel are difficult to consider because of the variable nature of peak hour demand), and

3. as mentioned earlier, in the common practice of carrying out small sample 24 hour O-D surveys, the number of observations available for developing peak hour models may be limiting and restrictive:

Regardless of the model's time basis, another important notion is that of seasonal variations. This fact introduces severe limitations on the model calibration exercise since the 24 hour counts used for calibration purposes are usually taken over a span of a few months. It is evident, for instance, that transit ridership peaks in winter and reaches its lowest level in the summer. Furthermore, counts used for the study area in this research (Ottawa-Carleton) were taken between the months of May and July. Inappropriate estimates of school travel and the significant reduction in work travel because of summer vacations may seriously contribute to errors in simulation.

The concepts of "Average Daily Traffic" (ADT) and "Design Hour Volumes" (DHV) are vague notions which are not explicitly treated in the urban transportation planning system model. They seem to have different meaning to different authors. These concepts need to be clearly defined and explicitly applied.

In summary, regardless of the type of model constructed, the system planner and the functional planner must be able to decide on the time frame of travel being simulated or forecasted. Failure to do so may result in confusion and eventually with improper design of transportation facilities.

CHAPTER III

EXISTING APPROACHES TO PEAK HOUR RATIOS

ESTIMATION AND APPLICATION

3.1 Introduction

During the last two or three decades, advances in the field of modelling urban travel demand has been overwhelming. Models based on home interview surveys, travel attitudes and utility theory, have been developed with varying degrees of success. Literature describing these models is abundant (5,17,19,20,21). The urban transportation planning system model (UTPS) evolved and spreaded to such an extent that most government agencies and authorities responsible for planning transportation systems have in one way or another adopted it. A simplified version of the model, as applied in the Ottawa-Carleton Regional Transportation Study (OCRTS) is shown in Figure 3.1. The four basic components of the model were developed in much the same way as in other studies (22,23,24).

In a typical urban transportation planning study, the trip generation phase develops 24 hour trips by purpose which are distributed over the aggregate zones of the study area. The modal split and travel assignment phases are performed for both peak hour and 24 hour conditions. Vehicle assignments to the road network are generally made for peak hour conditions and the peak hour volume is given to the functional planning engineer who develops from it the design hour volume (DHV) to be used for the facility design stage.

FIGURE 3.1

A SIMPLIFIED VERSION OF THE MODEL USED IN THE
 OTTAWA-CARLETON REGIONAL TRANSPORTATION STUDY (O.C.R.T.S.)

In order to convert the 24 hour total travel matrix into peak hour, a ratio or a set of peak hour ratios are applied as shown in Figure 3.2 (a). Most studies use an average area-wide ratio between 8 to 12% to make this conversion. Other studies apply constant peak hour ratios by trip purpose such as in the Ottawa-Carleton Study. The Toronto Area Regional Model Study (TARMS) further disaggregates these ratios by mode and purpose (2). However, most studies rarely endeavour to pursue this level of detail.

Recognizing that the eventual data needed in functional planning are peak hour volumes using a particular facility, a number of studies use the approach shown in Figure 3.2 (b). Trip productions and attractions are developed for the peak hour conditions directly and a peak hour travel matrix is developed and assigned to the appropriate network. This procedure, however, has a number of drawbacks and limitations which were discussed in section 2.4.

3.2 Existing Approaches to Peak Hour Ratios Application

The topic of peak hour ratios is perhaps one of the most neglected topics in the urban transportation planning process. Practical studies reviewed, overwhelmingly and consistently used constant area-wide ratios to convert 24 hour travel into peak hour travel. Transportation and traffic design and planning handbooks recognize and quote different ranges of variation in the peak hour ratios but they leave it to the transport planner or designer to "pick" the value to use.

FIGURE 3.2

PEAK HOUR MODEL
VS
24 HOUR MODEL APPLICATION

The meager treatment of the peak hour ratios in the literature is indeed surprising. Meyer (10)* confirms this by stating:

It is therefore, somewhat surprising that most urban transportation studies handle the peaking problem by an adjustment on the supply side. Typically, street and highway volumes are determined, analyzed, and forecasted on a 24 hour basis. Peaking is then accounted for by assuming that peak period volumes are a constant proportion of all 24 hour volumes at every kind of facility.

A report by the Trip Generation Project Committee of the Road and Transportation Association of Canada (RTAC) reviewed trip generation analysis procedures used in Canadian urban transportation planning studies. The report indicated that most of these studies developed trip generation relationships based on 24 hour periods (25).

An example of the use of a constant area-wide peak hour ratio may be found in a report on the Regional Municipality of York Transportation Study. The study was conducted by the Ministry of Transportation and Communications of Ontario (MTC) for the Region of York (3). The report indicates that to economize on time, the 24 hour work trip tables were factored to the p.m. peak hour by 0.215, which is the ratio of the p.m. peak hour to the 24 hour home based work trip generation for TARMS.

* Reference (10), page 116.

In theory, over an entire urban network, there is an infinite number of peak hour ratios to convert the 24 hour volumes on each individual node or link in the network into peak hour volumes. The peak hour ratio may vary depending on mode, trip purpose, geographic location, type of the facility, city size and a host of other variables.

The south-eastern Wisconsin Study assumed that the peak hour flows on the expressways and arterial highways were 10% of total average weekday flows and 8% on freeways (16). Oglesby and Hewes point out that the ratio between the 30th highest hourly volume and the average daily traffic (ADT) is far from constant for all roads and that actually recorded extremes have ranged from 8% to 38% (26)*.

D. M. Hill and others described a research project which contained a thorough analysis of the temporal distribution of vehicular travel in eight U.S. urban areas having a population ranging from 100,000 to 3.5 million people (21). Graphical models were developed during the analysis and five sub-models were used to aggregate hourly travel into similar time groups: "Wee" hours, morning, mid-day, afternoon and evenings. Area wide traffic counts data were used to determine total vehicle-miles of travel occurring on various classes of highways at varying distances and orientations from central core city, on directional and non-directional basis. The models developed in this study have area wide bases relating to city size and composition.

* Reference 26, page 185

The variables studied were:

1. the 24 hour vehicle-miles of travel divided by 10 times the hourly capacity,
2. employment compactness ratio defined as the central area employment divided by study area employment,
3. population compactness ratio defined as the central area population divided by study area population, and
4. density ratio defined as central city population density divided by study area population density.

The Cosmos Corporation in a report for the U.S. Department of Transportation* gives a simple example which is reproduced in Table 3.1 to show the importance of accuracy in the application of peak hour ratios. The table assumes an assigned total daily traffic of 125,000 vehicles. The variation in the number of lanes required assuming different values of the peak hour ratios and directional split factors ranges from 3 lanes to 5 lanes. This is based upon what might appear to be small variations in the peak hour ratios. If a 10% peak hour ratio and .55 directional split factor are assumed, 4 lanes are indicated in the heavy direction. However, just the change to a 12% peak hour ratio and .60 directional split factor would indicate 5 lanes as being required.

The Institute of Transportation Engineer's Handbook (27) states that the standard practice is to determine the design hour volume by multiplying the estimated or forecasted Average Daily Traffic (ADT) by the peak hour ratio. The handbook further states that for urban facilities the average ratio is about 11% with a range of 7% to 18%.

* Reference (16), page 83

TABLE 3.1

AN EXAMPLE ILLUSTRATING THE IMPACT OF VARIATIONS IN PEAK
HOUR RATIOS AND DIRECTIONAL SPLIT FACTORS ON FACILITY DESIGN*

<u>Total Daily Travel</u>	<u>Peak Hour Ratio</u>	<u>Directional Split Factor</u>	<u>Peak Hour Directional Volume</u>	<u>Number of Lanes @ 1800 Capacity (1 Direction)</u>	
				<u>Calculated</u>	<u>Rounded</u>
125,000	8	.55	5500	3.0	3
125,000	9	.55	6187	3.4	4
125,000	10	.55	6875	3.8	4
125,000	11	.55	7562	4.2	5
125,000	12	.55	8250	4.6	5
125,000	8	.60	6000	3.3	4
125,000	9	.60	6750	3.7	4
125,000	10	.60	7500	4.2	5
125,000	11	.60	8250	4.6	5
125,000	12	.60	9000	5.0	5

* Source: Reference(16), page 133.

An excellent treatment of the impact of traffic estimation errors is given in the National Cooperative Highway Research Program (NCHRP) report number 120. The report states that the traffic assignment output of the Strategic Transportation Planning Process (STPP), which is typically in terms of ADT, provides only the starting point for deriving design hour volume (DHV) data for design purposes. Errors in STPP output can be either offset or compounded by errors in the subsequent derivation of DHV data. Because the two sources of potential traffic estimation error are virtually inseparable the examination for sensitivity, although it explicitly concerns DHV data, also implicitly concerns STPP data (28).

Wohl (20) points out that no information can be more important to transport planners than reliable forecasts of the peak and off peak travel volumes on transport networks. Yet, he remarks, no reasonably complete and valid methodology has been proposed, much less developed and verified, that will permit the transport planners and analysts to properly differentiate travel by time of day, to determine realistically the duration and level of peaking and recognize its dependence on the transport system design and performance, and to account for shifts in trip making from one hour to another, from one mode to another, and from car pooling to driving alone in response to changes in transport systems designs.

The MTC's geometric design standards publication (29)* defines the design hour volume as that hourly traffic volume (including trucks) which the highway must accommodate in the design year. It determines to

* Reference (29), page 26.

a large degree the type of facility, pavement widths and various geometric features. The design hour volume for a given facility is obtained by plotting in descending order the traffic volumes for every hour in any year on that facility. The accepted hour for design within the Ministry of Transportation and Communications of Ontario is the 30th highest hour whether it be for an urban arterial or a rural freeway. This assumption permits 29 hours of operation during that one year which would be over the design capacity of that facility. The MTC publication further notes that the 30th highest hour expressed as a percentage of the AADT will be higher for the rural areas than for the urban. Consequently, it is not uncommon to find the design hour volume in an urban area as 7 to 15% of the AADT, whereas in the rural area this value might be 15 to 30%. However, the publication emphasizes that these are empirical relationships and that they will vary from place to place and highway to highway.

The peak hour travel demand on a given transportation facility varies considerably by direction. For design considerations, the distribution by direction is not as important on two lane roads as is the total two-way volume that governs capacity and design. However, it is important on multi-lane facilities where one direction is independent of the other when considering capacity. In general, directional split is more pronounced on rural transportation facilities than in the urban ones. In and near the central business district location, particularly in large cities, the traffic tends to equalize and in some cases approaches a 50-50 split (29).

The 1965 Highway Capacity Manual (30) states that:

On most two-way highways the actual average daily traffic has been found to be approximately the same in each direction. The volume during any specific hour may, however, be much heavier in one direction than in the other. A knowledge of the traffic load in each direction for the peak periods of traffic flow is essential because of the critical effect an unbalanced flow can have on the needed capacity as related to the design and operation of a highway.

The Manual recognizes the cyclical variation of travel with respect to time and points out that peaking characteristics within the peak hours should be considered. It does not, however, present any guidelines.

In its Policy on Design of Urban Highways and Arterial Streets (31), the American Association of State Highways Officials (AASHO)* states that although the annual average daily traffic AADT is essential in determining annual usage for expenditure justification, in design of structural elements of highway, and in estimating future traffic from which the design hour volume is derived, it has little direct application to the required geometric features of urban highways. Instead, a traffic volume representative of the twice daily rush hour period is used when designing urban roads. Because of the recurring morning and afternoon peak traffic flow in urban areas, the Policy states that there is little difference between the 30th highest hourly volume and the 200th highest hourly volume.

* Now named AASHTO: American Association of State Highways and Transportation Officials.

The AASHO Policy points out that traffic estimates used for the design of urban streets are usually expressed as ADT volumes derived from the urban transportation planning process. It further notes that recently, consideration has been given to the development of design hour volumes by making peak hour traffic assignments in lieu of AADT assignments.

The AASHO Policy recognizes the gap which presently exists between the system planning and functional planning processes by stating that:

... there occurs a transfer of information between planners responsible for determining future travel demand for urban regional transportation planning, and designers responsible for providing designs for specific facilities. Since the output of the planning process serves, almost without exception, as a basis for the development of design volumes, it is critical that an appreciation and awareness of the procedure and methods employed in each discipline be recognized.

Returning to the paper by Wohl (20), he states that the transport planner and analyst have fashioned numerous models dealing completely or in part with the travel forecasting problem. But, to his knowledge, none of the available models and techniques dealt realistically or structurally with the matter of peaking. More precisely, he points out that none of the models and techniques appropriately recognize that travel during the peak period is dependent both in concept and in actuality on the nature and extent of the transport system, on its performance and price characteristics during both peak and off peak periods, and on the attendant social economic conditions and preference patterns of travellers for peak and off peak travel choices.

Wohl concludes that without constructing demand models that simultaneously incorporate time of day preferences, the full range of service conditions, and modal possibilities, it seems evident that neither modal split nor the extent of peaking can be forecasted appropriately.

After describing the inadequacy of conventional models, Wohl proceeds to formulate what he calls a workable demand model capable of incorporating the most significant of the modal split and peaking aspects. But, Wohl gives no concrete operational application of his model. Instead, he admits that, in spite of his elaborate framework, it cannot be argued with certainty that the requisite effort in the last analysis will prove to have been worthwhile and to have improved significantly the decision-making process.

DeLeuw Cather and Company of Canada (32) advocate the carrying out of traffic assignments by inspection. They indicate that this makes it possible to assign peak hour ratios individually in accordance with the characteristics of particular areas, trip purposes and periods of the day or week. They further note that the peak hour will vary from point to point on the network and will vary differently on different days and in different seasons. They suggest that these variations can and should be taken into account. They propose that peak hour ratios be applied only where it is necessary to determine hourly volumes for the functional design of facilities in the immediate short range programme. They point out that peak hour ratios may be applied separately for separate trip purposes.

This permits more realistic peak hour estimates than assuming a single peak hour ratio for all trips or generating and distributing trips for peak periods of two or three hours and then factoring the peak hour volumes from that total.

The paper by Hill (21) mentioned earlier points out that the levels of congestion in major metropolitan areas are directly related to the fact that total vehicular travel is not uniform throughout the 24 hours of the day, while for the most part, the supply of transportation services is uniform. The cost of highway investment is directly related to this peaking phenomenon. Indeed, if it were not for this peaking in traffic demand, the required highway investment could be considerably less than it is now or than is anticipated to be in the future.

Consequently, urban planners acknowledge the need to develop a better understanding of urban travel behaviour and associated temporal characteristics so as to analyze and evaluate alternative levels of transportation investments. Thus, a case can be made for refining the current procedures of estimating and applying peak hour ratios in the urban areas. Furthermore, the net savings or benefits from such improvement would by far, exceed the cost of the extra data collected and analyzed.

3.3 Use of Peak Hour Ratios in Second Generation Models

There are basically four different approaches to the application of peak hour ratios in practice. They are presented below according to their order of complexity and may be traced in Figure 3.3.

FIGURE 3.3

EXISTING ALTERNATIVE APPROACHES TO THE USE
OF PEAK HOUR RATIOS IN THE URBAN TRANSPORTATION PLANNING SYSTEM MODEL

Approach (a)

The trip assignment output which is based on 24 hour conditions is given to the functional planning engineer who "picks" a suitable peak hour ratio to establish the design hour volume (DHV) required at the facility design stage.

Approach (b)

The matrix of 24 hour total travel is multiplied by a constant area-wide ratio (H^*) as follows:

$$t = T \cdot H \quad \dots \text{Eq. (3.1)}$$

where t = peak hour trip matrix

T = 24 hour trip matrix

H = area-wide peak hour ratio

The resultant peak hour travel matrix t is assigned to the peak hour network. The output from the peak hour assignment is then forwarded to the functional planning engineer.

Approach (c)

Different matrices of travel by trip purpose for the 24 hour conditions are converted into matrices of peak hour travel by trip purpose by means of applying a set of peak hour ratios by purpose as follows:

$$t^p = T^p \cdot H^p \quad \dots \text{Eq. (3.2)}$$

where the superscript p refers to trip purpose.

* The value commonly used is about 10%.

Proper directional splits are then applied and the matrices are added in order to establish the total peak hour travel which is then assigned as in approach (b).

Approach (d)

This approach is similar to c above except for one additional disaggregation. Matrices of 24 hour travel by purpose and mode (usually a set of matrices for auto travel and another set for transit) are developed. A set of peak hour ratios by mode and purpose $H^{m,P}$ are also developed in order to convert the 24 hour matrices into peak hour matrices, as follows:

$$t^{m,P} = T^{m,P} \cdot H^{m,P} \dots \text{Eq. (3.3)}$$

where the superscript m refers to travel mode.

This approach which appears rather simple and easy to implement is, surprisingly, rarely applied in practice. It is the author's contention that the application of such an approach will yield significantly better results than the other approaches described above.

3.4 Deficiencies of Existing Approaches

At the present state of the art in transportation planning, no reasonably complete methodology exists for the appropriate estimation and application of peak hour ratios within the urban transportation planning system model. The existing approaches all suffer from their basic assumption of constant area-wide peak hour ratios. Although some guidelines are given by various transportation handbooks, the choice of these peak hour ratios is left to the planner.

In order to outline the deficiencies of existing approaches, the flow diagram in Figure (3.3) is used and the following discussion refers to the blocks shown in that diagram.

The 24 hour trip table which is used as the main input matrix would normally be established using a gravity type trip distribution sub model. It is assumed throughout this research that such a trip table has indeed been satisfactorily developed from a given survey and compared to ground counts.*

In approach (a), the 24 hour trip table is assigned to a 24 hour road network to obtain daily estimates of travel. From these, the design hour volume is derived. It should be pointed out that the exercise of constructing a 24 hour network for the assignment of daily travel is rather difficult. In practice, an off peak network is usually employed. This procedure is not entirely correct since a high proportion of the 24 hour trips occur during the peak periods for which the off peak network would give a poor representation.

Some authors (20) have suggested to categorize the 24 hour travel into a number of time periods and to develop separate models for these time periods (21). The 24 hour traffic assignment is compared to 24 hour counts at certain strategically selected screenlines and cordons. It should be noted that full 24 hour classification counts in urban areas are rarely taken. Instead, eight or twelve hour classification counts are the common practice. This necessitates the use of conversion factors in order to

* In the decision-theory application presented in Chapter VIII, this matrix is assumed to be certain. A natural extension of the research would be to treat it as uncertain along the lines presented by Johnson (51).

bring these counts into daily totals. A certain amount of accuracy may be lost in making such conversions. Furthermore, classification counts do not always give accurate results, their accuracy being generally around 80 to 85%.

Factoring the 24 hour assignment into peak hour assignment for a given link or facility is generally done for the design year. This activity may be carried out by the transportation planner or, alternatively, by the facility designer who applies a highly subjective ratio to develop the design hour volume (DHV).

As the analysis of ground counts data discussed in Section 6.7 will show, the peak hour ratio for a given transportation facility depends on the type of facility. The ratio is lower for the higher order type of facilities such as freeways, and higher for the lower order of facilities such as local roads.

The use of the peak hour volume as the design hour volume is common in practice. However, before employing the assigned peak hour volume, considerable judgement is involved, particularly if the trip assignment technique used was the minimum path type. Such an assignment tends to underestimate traffic volumes on low order facilities because of the absorption of travel by the higher order, higher speed facilities. Other elements of judgement and subjective design considerations may also enter into the design decision process. These will be discussed in Chapter VIII.

Because of these subjective elements, it may be argued that a refined approach for peak hour ratios estimation and application may not prove more useful than the existing approach. In response to this argument, it should be pointed out that the purpose of the refined approach is to attempt to reduce the subjective input by placing in front of the transportation planner and designer the objective facts.

In approach (b), an area wide ratio is used to convert the 24 hour trip table into a total peak hour trip table. The most commonly used ratio is 10%. The primary reason for the application of such a ratio is probably because the percentage of the 24 hour volume occurring during the peak hour varies from about 7 or 8% to about 12 to 14% area-wide. Presumably, a value of 10% is a "happy medium". The simplicity of using the value of 10% particularly in converting the 24 hour volume directly to a peak hour volume may also have some influence on its selection.

An area wide ratio of 12% or more may result in over-designing certain segments of the transportation system. This will have serious economic implications particularly in terms of over-investment. However, some planners argue that because of budgetary and resources constraints, it is practically impossible to over-design a transportation system. Indeed, their point is that once the facility (say a supply of a six lane facility in lieu of a four lane demand) is open to traffic, the actual demand on that facility is higher than the predicted volume because of the diverted traffic from other facilities. Since most of the transportation systems in urban areas operate generally below level of service C, the over-designing or over-investment of a particular facility becomes rather difficult to pin point or define.

The use of a peak hour ratio of 8 or 9% may result in under-designing certain segments of the transportation system. This has been the case in many urban areas, but probably not as much due to the under-estimation of travel demand, as due to the financial and resources constraints imposed on transportation expenditures. The result of under-designing the system is a phenomenon which has serious implications on the society as a whole in terms of adverse effects. The literature is abundant with examples of the cost of congestion to the users, non users and society (16,18,26,30).

In approach (c), four or five trip purposes are used in a typical UTPS model study and thus a similar number of peak hour ratios by purpose are developed using origin-destination data. The approach is sensitive to the value of the ratio used for work trips since this type of travel constitutes over 50% of the total p.m. peak hour travel (and close to 90% of the a.m. peak hour travel). Although this approach produces significantly better results than the previous two approaches, it does not recognize the modal and locational characteristics of travel.

Approach (d) in which area wide ratios are applied to trip tables by purpose and mode should produce even better results. This disaggregation of peak hour ratios by mode is performed to reflect the hourly variation of travel by auto and transit. Since the peaking of transit travel is significantly more pronounced than the automobile travel, it is expected that the peak hour ratios for transit would be generally higher than those for the auto mode.

It is interesting to note that approach (d) is rarely employed in practice for a number of reasons. Firstly, because of the dominance of the auto mode particularly in the 1960's when urban transportation planning system models were evolving, the analysis was highly oriented towards automobile travel simulation and forecasting. The modal split step was relatively minor and was viewed as an adjustment applied to the total person trips in order to remove the minor transit component. Secondly, many studies, including the OCRTS, did not develop 24 hour transit trip tables and thus transit peak hour ratios were not necessary. Finally, in many cases and because of the very small origin-destination survey samples, non-work peak hour ratios by mode were difficult to establish.

In conclusion, approach (d) is by far superior to the previous three approaches. For immediate practical applications, it is the recommended approach. Nevertheless, this approach does not recognize the important locational aspects of travel, both at the origin and at the destination, nor the conditions of the transportation system. Thus, a refined framework for modelling peak hour ratios which incorporates these considerations is required.

CHAPTER IV

A FRAMEWORK FOR MODELLING THE PEAK HOUR RATIOS

4.1 Introduction

The premise that an improved approach for the estimation and application of peak hour ratios can be developed was the basic rationale for establishing the modelling framework discussed in this chapter. The framework was structured along the following eight activities:

1. definition of the peak hour ratios estimation problem,
2. preliminary investigation,
3. structuring of the new approach and formulation of hypotheses,
4. development of input data,
5. statistical analysis of data and simulation of peak hour ratios,
6. discussion and recommendation of the new approach for peak hour ratios estimation and application,
7. investigation of cost-effectiveness of the new approach using a Bayesian decision-theoretic model, and
8. conclusions and recommendations for further investigation.

The peak hour ratio estimation problem was defined in Chapter I. This problem definition became more significant when the author was faced with the task of revising original 1971 peak hour ratios by trip purpose used in the OCRTS study with a new set which would reflect 1975/1976 travel conditions. The author discovered that simple adjustments to these peak hour ratios caused significant differences to both the total peak hour

travel assigned to the transport network and to the individual link loading. In fact, the impacts of revising these peak hour ratios were so significant that the need for a number of new links in the transportation system was reconsidered.

The above led to some preliminary investigations of the peak hour ratios definition problem. The literature review described in the previous chapter and consultation with other professionals enforced the conclusion that the present treatment of peak hour ratios within the UTPS is meager and minimal.

The position of the peak hour ratios in the urban transportation planning process (UTPP) was described in Chapter II. The actual structure of the UTPP was examined to identify the interrelationship between the peak hour ratios and other elements of the process. Furthermore, and as part of this preliminary investigation, the basic considerations in the choice between 24 hour models and peak hour models were outlined.

Based on the above activities, a framework for modelling the peak hour ratios was formulated as shown in Figure 4.1. The framework provided an appropriate mechanism for guiding the research process and for illustrating the interaction among the various activities carried out in the research.

It was evident from the nature of this investigation that, to be feasible, it must utilize data that are at the disposal of the author (33). For this reason, the study area was defined as the Ottawa-Carleton Regional Transportation Study Area as shown in Figure 4.2.

The remainder of this chapter is devoted to the description of new alternative approaches for peak hour ratios estimation and application. A general hypothesis is formulated and the basic structure of each alternative approach is presented.

4.2 Hypothesis Formulation

The spectrum of peak hour ratios estimation and application may be described in terms of two extremes. At the simple end of the spectrum, constant peak hour ratios may be applied area-wide as discussed in Chapter III. At the other end, peak hour ratios may be applied at the individual trip maker level. Between these two extremes a range of possibilities emerges and these will be presented in the next section.

As a general hypothesis, it may be stated that the following six sets cover the spectrum of peak hour ratio determinants:

- (a) trip characteristics (e.g. trip mode and purpose),
- (b) origin and destination zone characteristics,
- (c) transport system variables (e.g. accessibility and facility type or classification),
- (d) regional or area-wide variables (e.g. city size, population and employment),
- (e) policy constraint variables, and
- (f) disaggregate socio-economic variables and trip maker characteristics.

Sets (d) and (e) are not treated explicitly in this research. They will be further discussed in Chapter VII. Set (d) requires data from different cities and the paper by Hill (21) provides the necessary background. For set (e), on the other hand, little

empirical work has been done. Furthermore, it can only be treated at the discussion level since the simulation of the impact of this variable set on peak hour ratios, though important, is extremely difficult.

One or more of the other four sets were used in generating different hypotheses depending on the alternative approach investigated. Table 4.1 shows the components of each set of peak hour ratios determinants. Appendix A provides a list of notations and abbreviations which may be used to explain some of the terms which appear in the table. It is quite evident that the list of specific zonal variables was developed based on data availability for the study area. Even though it was apparent that many of these variables would not enter into the peak hour ratios estimation equations, they were, nevertheless, developed to ensure that all available data were included and, also to permit the factor analysis technique to reduce the number of these independent variables into a manageable size.

4.3 Alternative Approaches for Peak Hour Ratios Estimation and Application

As indicated earlier, peak hour ratios may be estimated and applied at different levels of aggregation or disaggregation. These include:

1. area-wide or regional level,
2. large geographic area or district level,
3. zonal level,
4. household level, and
5. individual level.

TABLE 4.1
DETERMINANTS OF THE PEAK HOUR RATIOS

(a) Trip Characteristics:

Trip Mode: Auto, Transit

Trip Purpose: HBW, HBB, HBP&O, HBSCH. & NHB

(b) Origin and Destination Zone Characteristics

1. Labour Force
2. % Participation in Variable Work Hours
(i.e. Proportion of Jobs with Starting Times $t_1, t_2, \dots$)
3. Population Density
4. Residential Density
5. Total Employment
6. Total Population
7. Employment Density
8. Retail Employment
9. Other Employment
10. Zonal Income
11. Zonal Car Ownership
12. Aerial Distance from CBD
13. Road Accessibility Index
14. Transit Accessibility Index
15. Parking Availability
16. Trip Length From Zone i to Zone j
17. Geographic Location
18. Shopping Centres Gross Leasable Area
19. Labour Force Participation Ratio
20. Zonal Attractivity
21. Dwelling Units
22. Secondary School Enrollment
23. Post-Secondary School Enrollment

TABLE 4.1 (continued)

24. Federal Employment
25. Time by Car from CBD
26. Time by Transit from CBD
27. Net Zonal Acreage
28. Off Peak Travel Time from CBD
29. Difference of (Off Peak-Peak) travel time from CBD.

(c) Transport System Characteristics

1. Facility Type: Local, Collector, Arterial and Freeway
2. See (b) above (Variables 13 to 16, 25, 26, 28 and 29)

(d) Regional or Area-Wide Variables*

1. City Size
2. Vehicle-miles (24 HR)/Hourly Capacity
3. Employment Compactness Ratio: Central Area Employment/Total Area Employment
4. Population Compactness Ratio: Central Area Population/Total Area Population
5. Density Ratio: Central Area Population Density/Total Area Population Density

(e) Policy Constraint Variables

1. Discriminate Pricing Policies and Techniques
2. Restricted or Controlled Access During Peak Hour
3. Shortened Work Week
4. Communications in Lieu of Travel

(f) Disaggregate Socio-Economic Variables

1. Number of Vehicles/Household
2. Persons/Household
3. Household Type
4. Occupation
5. Income

* From U.S. research (21)

In this research, all the above levels were attempted with varying degrees of emphasis.

The regional or area-wide level involves the four existing approaches described in section 3.3. These are also shown in Figure 4.3 as approaches a, b, c and d. At both the district and the zonal levels, peak hour ratios may be estimated for the zonal or district trip ends and for the zonal or district trip interchanges. These are shown as the two new approaches in Figure 4.3 and they are described in the next four sub-sections.

The disaggregate level involves formulating an approach whereby peak hour ratios at the household or individual levels are related to trip makers characteristics and other socio-economic variables. This approach is described in sub-section 4.3.5.

4.3.1 Zonal Trip Ends Approach

In this approach the 24 hour zonal trip productions and attractions by purpose and mode are converted to peak hour trip productions and attractions by mode and purpose using the trip-end peak hour ratios as follows:

$$H_i = f(X_1, X_2, X_3, \dots) \dots \dots \text{Eq. (4.1)}$$

$$H_j = f(Y_1, Y_2, Y_3, \dots) \dots \dots \text{Eq. (4.2)}$$

where:

H_i = Peak hour ratio for origin zone i

H_j = Peak hour ratio for destination zone j

FIGURE 4.3

ALTERNATIVE APPROACHES FOR PEAK HOUR RATIOS ESTIMATION
AND APPLICATION IN THE URBAN TRANSPORTATION PLANNING SYSTEM MODEL

$x_1, x_2, x_3, \dots$ and

$y_1, y_2, y_3, \dots$ are zonal independent variables reflecting socio-economic, land use and/or transport system variables for the zone of origin and zone of destination respectively.

The above ratios are then applied as follows:

$$t_i^{m,p} = T_i^{m,p} \cdot H_i^{m,p} \quad \dots \text{Eq. (4.3)}$$

$$t_j^{m,p} = T_j^{m,p} \cdot H_j^{m,p} \quad \dots \text{Eq. (4.4)}$$

where:

$t_i^{m,p}$ & $t_j^{m,p}$ = total peak hour trips for purpose p and by mode m for origin zone i and destination zone j respectively.

and

$T_i^{m,p}, T_j^{m,p}$ = total 24 hour trips for purpose p and by mode m for origin zone i and destination zone j respectively.

Peak hour ratios at the origin zone, destination zone or both may be applied. It should be evident that a balancing procedure will be needed to produce the final peak hour trip tables by mode and purpose. These tables are converted into true origin-destination trip tables by means of applying appropriate directional split factors.

They are, subsequently, assigned to their respective transportation network.

Zonal end peak hour ratios can be developed using O-D survey data aggregated to the zonal level by means of statistical techniques such as factor analysis and regression analysis. In effect, this approach produces peak hour trip generation equations by purpose and mode. It may be argued that such equations may be derived directly at the beginning of the process from the origin-destination survey. However, if this is done, the advantages of the 24 hour trip distribution simulation would be lost. Furthermore, the limitations of peak hour models discussed in Section 2.4 would apply. Finally, in the zonal trip ends approach, control totals from the 24 hour trip tables would be available which can be used to judge the appropriateness of the results. Peak hour trip generation models on the other hand, cannot provide such control totals and their adequacy can only be judged at the end of the process when simulated peak hour travel is compared to peak hour traffic counts. It may be concluded therefore, that the zonal trip ends approach is superior to the existing approaches discussed in Section 3.3. However, it does not take into account inter-zonal variations of the peak hour ratios.

4.3.2 Zonal Trip Interchange Approach

This approach is somewhat equivalent to deriving a direct demand model of peak hour trip generation and distribution by purpose and mode. In this approach peak hour ratios are developed to convert every interchange cell of each 24 hour trip table by purpose and

mode to a corresponding peak hour cell interchange in each of the resulting trip tables. Just as in the case of zonal trip ends approach, a number of independent variables are related to the peak hour ratios (the dependent variables). The relationship may be established as follows:

$$H_{ij}^{m,p} = f(X_1, X_2, \dots, Y_1, Y_2, \dots, S_1, S_2, \dots) \dots \text{Eq. (4.5)}$$

where:

$H_{ij}^{m,p}$ = peak hour ratio for purpose p and mode m as applied to trips from zone i to zone j

$X_1, X_2, \dots$ = variables reflecting origin zone characteristics

$Y_1, Y_2, \dots$ = variables reflecting destination zone characteristics, and

$S_1, S_2, \dots$ = variables reflecting transport system characteristics

The above factors are then applied as follows:

$$t_{ij}^{m,p} = T_{ij}^{m,p} \cdot H_{ij}^{m,p} \dots \text{Eq. (4.6)}$$

where:

$t_{ij}^{m,p}$ = peak hour trips from zone i to zone j for purpose p and mode m

$T_{ij}^{m,p}$ = 24-hour trips from zone i to zone j for purpose p and mode m

It should be apparent from the above discussion that the zonal trip interchange approach is superior to all previous approaches since it takes into account:

1. variation of peak hour travel by mode,
2. variation of peak hour travel by purpose,
3. variation of peak hour travel by zone of origin and destination (thus accounting for geographic differences), and
4. inter-zonal variation of peak hour travel as reflected by the conditions of the transport system in the peak hour.

4.3.3 District Level Peak Hour Ratios

This approach is basically the same as the zonal trip end or zonal trip interchange approaches. However, because of data limitations and scatter, it was possible to group the zones into large districts and to establish different district level peak hour ratios by mode and purpose. This was attempted by converting the 300 zone system used in the analysis into 60 districts. All statistical analysis performed at the zonal level was also performed at the district level as described in Section 5.4.2.

4.3.4 Concentrated Work Trips Approach

In this approach the development of work related peak hour ratios by mode is emphasized. This approach improves upon the current practice* of applying the relationship:

$$t_{ij} = t_{ij}^w \cdot \phi \quad \text{..... Eq. (4.7)}$$

* In fact, this is the relationship which is currently employed in the UTPS computer package of UMTA (Urban Mass Transportation Administration) of the U.S. Dept. of Transportation. (6)

where:

- t_{ij} = total peak hour trips from zone i to zone j, all purposes,
 t_{ij}^w = work peak hour trips from zone i to zone j, and
 ϕ = factor to convert work peak hour travel matrix into total peak hour travel matrix.

The above relationship may be useful if total a.m.* peak hour trips are desired, but it may be error inviting if it is applied directly for p.m. peak hour travel since work trips constitute between 50% to 60% of total p.m. peak hour travel.

The alternative procedure advanced in this thesis is to develop peak hour ratios for work trips using either of the two new approaches (namely zonal trip ends or zonal trip interchange) and then to account for non-work trips as described below.

The average peak hour ratios $H^{m,nw}$ by mode for each of the non-work trip purposes (nw) may be used area-wide. The total peak hour travel may be obtained by multiplying these average ratios by their corresponding 24 hour trip tables and adding the product to the peak hour work trips.

Thus,

$$H_{ij}^{m,w} = f(X_1, X_2, Y_1, Y_2, S_1, S_2, \dots) \quad \dots \text{Eq. (4.8)}$$

* Close to 90% of the a.m. peak hour trips are work trips.

$$t_{ij}^{m,w} = H_{ij}^{m,w} \cdot T_{ij}^{m,w} \quad \dots \text{Eq. (4.9)}$$

$$t_{ij}^{m,nw} = H_{ij}^{m,nw} \cdot T_{ij}^{m,nw} \quad \dots \text{Eq. (4.10)}$$

and

$$t_{ij}^{m,p} = t_{ij}^{m,w} + t_{ij}^{m,nw} \quad \dots \text{Eq. (4.11)}$$

The reader should note that appropriate directional split factors and balancing procedures should be applied to each trip table by mode and purpose prior to developing the total true O-D peak hour trip tables which are assigned to the peak hour transportation networks.

4.3.5 Disaggregate Approach

Peak hour ratios in this approach, are defined on the basis of individuals and/or households. This approach utilizes the O-D survey records together with a totally different* set of independent variables reflecting socio-economic data of the individual trip maker or his household. The results are then aggregated in order to develop zonal data which may subsequently be used as input to the urban transportation planning system model.

Let $a, b, c, \dots$ be attributes of individual i in household j . Then

$$H_i^{m,p} = f(a, b, c, \dots) \quad \dots \text{Eq. (4.12)}$$

$$H_j^{m,p} = f(A, B, C, \dots) \quad \dots \text{Eq. (4.13)}$$

where

$A, B, C, \dots$ are attributes of household j .

* From those of the aggregate zonal level approach

Peak hour ratios may be derived for individual i and/or household j using category analysis or other techniques and then, in much the same way as in trip generation, resultant data are aggregated to the zonal level.

4.4 Hypotheses Testing

Having formulated the various approaches and hypotheses, the framework developed in Figure 4.1 was used for testing these hypotheses for the study area. As stated earlier, the choice of the RMOC/CRO* regions as the study area was dictated by the availability of data and resources for this research. In fact, for each set of the peak hour ratio determinants, only variables for which data were available or could be derived were included. This presented some difficulties because certain variables were not available until much later and others had to be developed throughout the course of the research.

The basic steps followed in testing the derived hypotheses were:

1. input data analysis and validation,
2. input data manipulation and development of dependent and independent variables,
3. statistical analysis and simulation, including canonical correlation analysis, factor analysis, category analysis and step-wise regression analysis,
4. screenline and cordon counts analysis, and
5. application of Bayesian statistical decision theory.

Chapters V through VIII elaborate on the above steps.

* The Regional Municipality of Ottawa-Carleton and the Communauté Régionale de l'Outaouais which make up the OCRTS Area.

CHAPTER V

DEVELOPMENT OF INPUT DATA

5.1 Introduction

In most urban transportation planning studies input data are obtained from different sources. Thus problems of incompatibility and inconsistency are usually encountered. The date of collection or availability of data varies because data are often collected over a long time span, extending sometimes, from one year to another. More importantly however, data are collected based on identified needs at a certain point in time. In addition, most of the data for the study area employed in this research were collected for different zone systems. It was therefore essential to convert all data to a consistent zone system.

The process of data development in this research was lengthy because of the following reasons:

1. the amount of input data (it was perhaps possible to use a small subset of the input data however, for practical considerations, the entire data set was used),
2. the necessity of performing lengthy data manipulation and modifications particularly when no computer programmes were available to perform such manipulation, and
3. the ambitious attempt to develop zonal trip interchange peak hour ratios.

It should be appreciated that the development of all input data could not be carried out solely by the author within the time frame of this research. Most of

the data developed were necessary for the Phase II* Transportation Model Development Study of the Regional Municipality of Ottawa-Carleton (23).

5.2 Sources of Data

The definition of variables to be used and the criteria for their selection were described in Section 4.2. Most of the input data were developed by various government agencies. Some data were developed by the author directly. The following agencies provided the land use and socio-economic data:

1. The Planning Department of the Regional Municipality of Ottawa-Carleton (RMOC).
2. The Planning Department of the Communauté Régionale de l'Outaouais (CRO).
3. The National Capital Commission (NCC).

The origin destination travel survey data were obtained from the following sources:

1. The Transportation Department of the Regional Municipality of Ottawa-Carleton.
2. The Ministry of Transportation and Communications of Ontario.

The basic travel data were obtained from the 1972 Home Interview Survey conducted by Canadian Facts Limited for the Regional Municipality of Ottawa-Carleton in 1972-1973. The 1971 census information was also used by the Planning Department to validate the socio-economic and land use data. Data used in developing the economic basis for the peak hour ratio estimation were obtained

* Also referred to as the TRANS programme

from the Programme Management Branch of the Regional Municipality of Ottawa-Carleton and from the City of Ottawa records for data applicable to the period prior to the creation of the Regional Government in Ottawa-Carleton.

5.3 Development of Independent Variables

Table 5.1 presents a data bank* listing of the zonal independent and dependent variables. Appendix A may be consulted for a listing of the notations and abbreviations used in describing some fields of this data bank. It may be noted that data for the independent variables were developed for the years 1972 and 1977. Data for the dependent variables, on the other hand, were developed for the year 1972 only using the Home Interview Survey as described in the next section. The matrix form of the data bank is reproduced in a variable label format in Appendix B.1. With the exception of fields 127 through 146, all the remaining variables in this data bank were labelled "Independent".

The source and/or method of development of each field of the data bank is indicated in Table 5.2. Later in the research, the following independent variables were also developed and used in the simulation (however, they were not entered into the data bank):

1. zonal off peak travel times by car from CBD,
2. difference in zonal travel times by car from CBD (peak minus off peak), and

* This data bank is actually a column matrix of 300 Zones or observations by X number of fields. In order to define a particular field, say 125, start with a row 120 and go horizontally to find under column 5 the variable "Car Index of Attractivity".

TABLE 5.1
DATA BANK

0	1	2	3	4	5	6	7	8	9
ZONE	POP. 1977	TOT. EMP.	RET. EMP.	D. U.	CARS	SEC. SCHOOL ENROL.	POST SEC. ENROL.	SHOP. CENTRES G.L.A.	LAB. FORCE PART. RAT.
GROSS ACREAGE	FEDERAL EMPL.	GROSS POP. DEN. PER/ACRE	POP. PER DWELLING	GROSS EMPL. DENSITY	CARS PER DWELLING	NON-FED NON-RET EMPL.			
	AERIAL DIST. FROM CBD	TIME BY TRANSIT FROM CBD	TIME BY CAR FROM CBD	PARKING AVAILAB- ILITY	NET ACREAGE	INCOME	LABOUR FORCE	FINISHING TIME FROM 1531-1545	FINISHING TIME FROM 1546-1600
FINISHING TIME FROM 1601-1615	FINISHING TIME FROM 1616-1630	FINISHING TIME FROM 1631-1645	FINISHING TIME FROM 1646-1700	FINISHING TIME FROM 1701-1715	FINISHING TIME FROM 1716-1730	FINISHING TIME FROM 1731-1745	FINISHING TIME FROM 1756-1800	FINISHING TIME FROM 1801-1815	FINISHING TIME FROM 1816-1830
	NET POP. DENSITY	NET RESID.	NET EMPL.	ZONAL ATTRACT.	CAR INDEX OF ATTRACT.	TRANSIT INDEX OF ATTRACT.	H(J) D + P HSRO	H(I) D + P PSRO	H(J) D + P HSCH
H(I) D + P HSCH	H(J) TRANS. HW	H(I) TRANS. HW	H(I) TRANS. HSRB	H(I) TRANS. HSRB	H(J) D + P HW	H(I) D + P HW	H(J) D + P HSRB	H(I) D + P HSRB	H(J) D + P HSCH
H(I) D + P HSCH	H(J) TRANS. HSRB	H(I) TRANS. HSRB	H(I) TRANS. HSRB	H(I) TRANS. HSRB	H(J) TRANS. HSRB	H(I) TRANS. HSRB	H(J) D + P HSRB	H(I) D + P HSRB	H(J) D + P HSRB
	POP. 1972	TOT. EMP.	RET. EMP.		CARS	SEC. SCHOOL ENROL.	POST SEC. SCH. DORN	SHOP CENT. G.L.A.	LAB. FORCE PART. RAT.
		TIME BY TRANSIT FROM CBD	TIME BY CAR FROM CBD	PARKING AVAIL.	NET ACREAGE	INCOME	LABOUR FORCE	FINISHING TIME FROM 1531-1545	FINISHING TIME FROM 1546-1600
FINISHING TIME FROM 1601-1615	FINISHING TIME FROM 1616-1630	FINISHING TIME FROM 1631-1645	FINISHING TIME FROM 1646-1700	FINISHING TIME FROM 1701-1715	FINISHING TIME FROM 1716-1730	FINISHING TIME FROM 1731-1745	FINISHING TIME FROM 1756-1800	FINISHING TIME FROM 1801-1815	FINISHING TIME FROM 1816-1830
	NET POP. DENSITY	NET RESIDENT	NET EMPL.	ZONAL ATTRACT.	CAR INDEX OF ATTRACT.	TRANSIT INDEX OF ATTRACT.			

TABLE 5.2

SOURCES AND METHOD OF DEVELOPMENT OF ZONAL VARIABLES*

<u>Field Number</u>	<u>Source and/or Method of Development</u>
001-009, 090-096, 201-209	Planning Departments/RMOC/CRO/NCC
101	Manually
102-103	1977 Road and Transit Networks respectively
302-303	1972 Road and Transit Networks respectively
104, 304	Manually
105, 305	Manually
106, 306	Reference (22) and Census Data
107, 307	Computer Programme
108-119, 308-319	Reference (34)
121-123, 321-323	Computer Programme
124, 324	References (22) and (23)
125-126, 325-326	Computer Programme Appendix C
127-146	Computer Programmes (See Section 5.4)

* Refer to Data Bank in Table 5.1 for definition of fields.

3. number of employees per zone who finish work between 4:25 p.m. to 5:25 p.m. The choice of this hour was dictated by the trip-in-motion concept which is described in Section 5.6.

5.4 Development of Dependent Variables

5.4.1 Zonal Level Peak Hour Ratios

A simplified flow chart describing the procedure followed in developing the dependent variables is shown in Figure 5.1. Appendix D.1 contains a listing of the titles of various computer programmes employed in the above procedure. No further information about these programmes are given but the interested reader should consult reference (35) for more details.

Some 25,000 survey records representing vehicular travel by the driver, passenger and transit modes of 2% of all households in the study area in 1972 were used. The following output data on the basis of the 300 zone system were desired:

$H_{ij}^{m,p}$ defined as zonal interchange peak hour ratios by mode m and purpose p,

$H_i^{m,p}$ defined as origin zonal trip-ends peak hour ratios by mode m and purpose p, and

$H_j^{m,p}$ defined as destination zonal trip-ends peak hour ratios by mode m and purpose p.

The first problem encountered in the development of these data was the need to convert all survey records into the 300 Zone System*. The second problem concerned

* The survey was coded initially on the basis of the Ontario Provincial Code System. The conversion to the 300 Zone System was done at MTC in Toronto.

FIGURE 5.1

A SIMPLIFIED FLOWCHART FOR THE DEVELOPMENT
OF DEPENDENT VARIABLES

the definition of the p.m. peak hour used for extracting all peak hour records from the total survey records. The investigation of this problem led to a separate examination of the peak hour definition concepts. Section 5.6 has been devoted to explore these concepts.

The 25,000 trip records represented travel in the 24 hour period starting from 3:00 a.m. of the survey day to 3:00 a.m. the next day. The author chose to adopt the trip-in-motion concept for defining the actual occurrence of the p.m. peak hour. A computer programme was written which defined both the a.m. and the p.m. peak hours using this concept. Another programme (FORTGCLC in Figure 5.1) extracted from the total 24 hour records, those records of trips which were in motion during the 4:25 p.m. to 5:25 p.m. time period anywhere in the study area. Appendix E contains a listing of the trip-in-motion programme used.

The above programmes produced some 3178 p.m. peak hour trip records. This represented about 13% of the 24 hour travel. It was quite obvious that the use of trip-in-motion concept and the trip selection method produced much higher peak hour ratios than those produced using conventional peak hour definition and trip selection.

The next step was to add expansion factors to the survey records. The calculation of expansion factors was done manually on the basis of unique individuals by zone. The method of deciding on the expansion factors and on later adjustment techniques are beyond the scope of this research.

A computer programme was then implemented to construct 20 origin-destination matrices by the five purposes and two modes for the 24 hour $T_{ij}^{m,P}$ and for the peak hour $t_{ij}^{m,P}$.

The driver and passenger (D+P) matrix for home-work (H-W) in the peak hour was divided by its corresponding 24 hour matrix and the resultant peak hour ratios were examined. The occurrences of peak hour ratios values of 0% or 100% were predominant. Few cells contained other values. The explanation was rather simple. Notwithstanding the expansion factors, the survey peak hour sample of 3178 trip records was being distributed among the 20 zonal matrices. A particular cell would likely have a value of 0% or 100% because of the few peak hour observations. Such being the case for the H-W, D+P travel which represent over 50% of peak hour travel, it was evident that the same problem will definitely be encountered for the other trip mode/trip purpose matrices.

Three remedial options were contemplated:

1. Increase the survey sample* to increase peak hour trip records. This option was immediately rejected since it would not have resolved significantly the above problem. Furthermore, it would have caused unacceptable delay of the research until the second survey was completed, coded and edited.
2. Convert the zonal matrices into district matrices. This was attempted but the resultant district peak hour ratio matrix H_{ij} , H-W, D+P) contained, in general, the values 0%, 100%, 33% (1 peak hour record

* Perhaps by adding the 1977 1% Telephone Survey which was being carried out during the research (62).

- out of a daily total of 3) and 66% (2 peak hour records out of a daily total of 3). This indicated that the sample is still too small to be used to simulate $H_{ij}^{m,p}$ at the district interchange level.
3. Abandon the $H_{ij}^{m,p}$ approach, or at least limit any further analysis to investigating H_{ij} for H-W, D+P at a larger geographic level than the 60 districts.

It was therefore decided to abandon this approach because of insufficient data.

The number of peak hour records was believed sufficient for the task of establishing zonal trip ends peak hour ratios by purpose and mode ($H_i^{m,p}$ and $H_j^{m,p}$). The procedure followed in their development is described below.

Using the 20 matrices of travel by mode and purpose in the 24 hour and in the peak hour ($T_{ij}^{m,p}$ and $t_{ij}^{m,p}$), a trip-ends summary programme was used to list the following data:

$$\sum_{i=1}^{300} P_{24i}^{m,p} = \sum_{i=1}^{300} T_{ij}^{m,p} \quad \dots \text{Eq. (5.1)}$$

$$\sum_{j=1}^{300} A_{24j}^{m,p} = \sum_{j=1}^{300} T_{ij}^{m,p} \quad \dots \text{Eq. (5.2)}$$

$$\sum_{i=1}^{300} P_{PKi}^{m,p} = \sum_{i=1}^{300} t_{ij}^{m,p} \quad \dots \text{Eq. (5.3)}$$

$$\sum_{j=1}^{300} P_{PKj}^{m,p} = \sum_{j=1}^{300} t_{ij}^{m,p} \quad \dots \text{Eq. (5.4)}$$

where:

- $P_{24i}^{m,p}$ = Total trip-ends (produced at the origin zone i) by mode m and purpose p in the 24 hour period.
- $A_{24j}^{m,p}$ = Total trip-ends (attracted to the destination zone j) by mode m and purpose p in the 24 hour period.
- $P_{pk_i}^{m,p}$ = Total trip-ends (produced at the origin zone i) by mode m and purpose p in the peak hour.
- $A_{pk_j}^{m,p}$ = Total trip-ends by mode m and purpose p in the peak hour attracted to the destination zone j.
- T_{ij} = Total 24 hour trips from zone i to zone j.
- t_{ij} = Peak hour trips from zone i to zone j.

The $H_i^{m,p}$ and $H_j^{m,p}$ in percentages, were then calculated as follows:

$$H_i^{m,p} = \left(\frac{\sum_{i=1}^{300} P_{pk_i}^{m,p}}{\sum_{i=1}^{300} P_{24i}^{m,p}} \right) \cdot 100 \dots \text{Eq. (5.5)}$$

$$H_j^{m,p} = \left(\frac{\sum_{j=1}^{300} A_{pk_j}^{m,p}}{\sum_{j=1}^{300} A_{24j}^{m,p}} \right) \cdot 100 \dots \text{Eq. (5.6)}$$

The zonal trip ends peak hour ratios were calculated and added, as field 127 through field 146, to the Data Bank shown in Table 5.1. The 300 zone system used in the data development is shown in Figure 5.2.

FIGURE 5.2

THE 300 ZONE SYSTEM (URBAN)

FIGURE 5.2 (Continued)

THE 300 ZONE SYSTEM (RURAL)

FIGURE 5.2 (Continued)

THE 300 ZONE SYSTEM (EXTERNAL)

Once these peak hour ratios were added to the data bank, the assembly of all zonal input data (dependent as well as independent variables) was complete. Analysis of these data and simulation of peak hour ratios commenced at the zonal trip-ends level. This analysis is described in the next chapter. The data bank was read into SPSS* and a system file was created containing the list of variables shown in Appendix B.1. Subsequent zonal level statistical analysis was performed using this system file.

5.4.2 District Level Peak Hour Ratios

The decision to proceed with the development of district trip-ends peak hour ratios $H_i^{m,P}$ and $H_j^{m,P}$ was based on the premise that the scatter of observations of the dependent variables at the zonal level can, perhaps, be reduced at the district level to produce better relationships with the independent variables. This scatter, on the one hand, indicates that the use of average values for peak hour ratios has some justification, but on the other hand, it points out to the variation in these ratios and the need to develop meaningful relationships to explain their variation. Admittedly, data manipulations and assumptions as well as the low sampling rate may be responsible for the scatter, but this cannot be concluded or verified without comparing it to other studies, none of which appear to have been carried out.

The premise of zone-to-district aggregation is that while the number of observations (and consequently, the degrees of freedom) are reduced, the intra-district (inter-zonal within a given district) variations would be eliminated

* SPSS: Statistical Package for The Social Sciences. Appendix D.2 contains a listing of the Statistical Procedures available in this package which were employed in the research.

and the new district (mean) peak hour ratios will show less scatter. Indeed, comparison of similar scattergrams at the zonal and district levels indicated that such was the case.

The second reason for developing district trip-ends peak hour ratios was to reduce the data processing costs and the labour required for deriving and adding new variables to the system file.

The third reason for developing district level peak hour ratios was to be able to visualize geographically their variation. This is difficult with 300 values but much simpler with 60 values. Geographic plots require less time to develop and may be easier to interpret.

The zonal level SPSS system file was used in the development of district trip-ends peak hour ratios. The file was first sorted using the district-zone equivalent list given in Table 5.3. Next, the file variables were assigned either a sum label (e.g. district population = sum of population of all zones within the district) or a mean label (e.g. district travel time from CBD = weighted average of all zonal travel times from CBD). The third and final step was to read this aggregated file in binary and create a new district level SPSS system file. After this was accomplished, the resultant file contained exactly the same structure as the zonal file. All subsequent analysis and simulation at the district level were performed using this file. The 60 districts system is shown in Figure 5.3.

TABLE 5.3

CONVERSION TABLE: 300 ZONES TO 60 DISTRICTS

<u>DISTRICT</u>	<u>ZONES</u>	<u>DISTRICT</u>	<u>ZONES</u>
1	1-14	31	196-201
2	15-17	32	202-204
3	18-22	33	205, 206
4	23-26	34	207
5	27-35	35	208
6	36-40	36	209
7	41-49	37	210
8	50-52	38	211
9	53-54	39	212, 213
10	55-63	40	214-216
11	64-72	41	217-219
12	73-83	42	220-224
13	84-97	43	225-233
14	98-106	44	234, 237, 238, 244-250, 251
15	107-113	45	235, 236, 240-243
16	114-119	46	239, 251, 253, 254
17	122-125	47	255-260
18	120, 121, 126-134	48	261-269
19	135-141	49	270-276
20	142-147	50	277-279
21	148-156	51	280
22	157, 158, 163, 164	52	281
23	159-162	53	282
24	165-172	54	283
25	173-177	55	284
26	178	56	285
27	179, 180, 183	57	290-296 (ONT.)
28	181, 182, 184, 185	58	297-299 (QUE.)
29	186, 187	59	286-289 (USA)
30	188-195	60	300

2

FIGURE 5.3

THE 60 DISTRICTS SYSTEM (URBAN)

FIGURE 5.3 (Continued)

THE 60 DISTRICTS SYSTEM (RURAL)

5.4.3 Disaggregate Household and Individual Level

The 24 hour trip records, in the form used for zonal level analysis of peak hour ratios, were also used for this disaggregate level data development. Similarly, the p.m. peak hour records developed using the trip in motion concept used at the zonal level analysis were also used.

A computer programme was written* to match identification fields of the 24 hour records and the p.m. peak hour records. When such a match was obtained an indicator was set on the 24 hour record which stated whether or not a p.m. peak hour record match was obtained for that particular record. This was done using either the value 0 to indicate no match or the value 1 to indicate a match.

Another programme* was written to calculate the peak hour ratios for each unique household. This was done by accumulating the total peak hour trips and dividing them by the total 24 hour trips accumulated for that particular household. This was initially done for total trips (i.e. overall peak hour ratio by household). Subsequently, at the household level, peak hour ratios were also developed for each of the 10 trip purpose/trip mode combinations.

Thus 11 peak hour ratios were developed at the household level and only one peak hour ratio was developed at the individual level. It was obvious that the number of trips by each unique individual will not be sufficient to estimate peak hour ratios by mode and purpose.

* In PL1. These two programmes are too long to include in the thesis.

The above 12 peak hour ratios were merged with the original trip records and the format of the trip file was extended. Additionally, the trip purpose and mode codes were redefined.

Three SPSS system files were then extracted from the total survey records. Each of these contained the variables described in Appendix B.2. The three files were defined as follows:

1. "HH72" which contained household records only,
2. "ID72" which contained individual records only, and
3. "PK72" which contained p.m. peak hour records only.

These files were used to carry out the category analysis described in Section 6.6.

5.5 Input Data for Decision Analysis

In this section, a brief outline of the basic input data used in the decision analysis exercise is presented. Section 8.5 will contain further discussion of these input data.

The following are the four basic input data sets for decision analysis:

1. cost of experimentation for each alternative option,
2. prior probability distribution,
3. conditional probability distribution, and
4. cost of alternative actions.

The cost of experimentation was developed for three of the five options used in the decision analysis as described in Chapter VIII. Experimentation costs for the other two options were developed subjectively based on the author's experience and judgement.

In order to develop the prior probability distribution function the following input data were assembled:

1. design data for more than 60 links of the RMOC transportation network,
2. ten years travel projection of the design hour volume by link obtained from the 1968 Road Needs Study Programme of the City of Ottawa, and
3. actual 1978 ground peak hour volume counts obtained from the RMOC Transportation Department.

The prior probabilities were derived using the above input data as described in Section 8.5.2.

The conditional probability distributions which are described in Section 8.5.3 were developed subjectively by the author in consultation with other professionals.

Finally, the cost functions for the alternative actions were developed using cost information for some 110 links of the RMOC road network. The input data assembled in order to derive these cost functions included:

1. number of lanes,
2. type of link (urban or semi-urban),
3. length of link in kilometers,
4. total operating, maintenance and capital costs in dollars per kilometer, and
5. design hour volume in vehicles per hour.

These data were placed into an SPSS system file defined as "COSTS". Appendix B.3 gives a listing of the variable labels contained in the file. These variables were used to develop the cost functions as described in Section 8.5.4.

Table 5.4 provides a summary of all the data banks developed in the research.

TABLE 5.4
SUMMARY OF ALL DATA BANKS DEVELOPED IN THE RESEARCH

Item	Name or Type of Data Bank	Number of Cases or Observations	Number of Variables or Fields	Type of File
1	Socio-Economic/Land Use Data Bank	300	500	AVJIO* File
2	Screenlines and Cordon Counts Data Bank	300	22	AVJIO File
3	Skim Trips (Transit and Road)	300	300	AVJIO File
4	Zonal-End Peak Hour Ratios	300	20	FORTRAN File
5	"SHALLAL" Zonal Level File	300	98	SPSS** File
6	"DISTR60" District Level File	60	98	SPSS** File
7	"COSTS" for Cost Functions/ Decision Analysis	110	14	SPSS** File
8	"HD19AGGR" for Disaggregate Level Analysis	24071	48	SPSS** File
9	"HH72" For Household Level Analysis	2166	48	SPSS** File
10	"ID72" For Individual Level Analysis	5490	48	SPSS** File
11	"PK72" For Peak Hour Records Only	3178	48	SPSS** File

* See MTC/STP Manual (35)

** SPSS, Statistical Package for the Social Sciences (61)

5.6 Notes on the Definition of the Peak Hour

The term peak hour in the transportation planning literature is one of the most commonly used and confused terms. Indeed part of the confusion relates to the fact that the term peak hour has been used synonymously with the expressions peak period, a.m. peak hour, p.m. peak hour, a.m. peak period and p.m. peak period. From a duration point of view the terms a.m. peak hour, p.m. peak hour and design hour volumes all refer to 60 minutes of person or vehicular travel. On the other hand the terms a.m. peak period, p.m. peak period or simply peak period may denote anywhere from one hour to three or sometimes three and a half hours in duration. The purpose of this discussion is to delineate on these definitions in an attempt to remove some of the ambiguity in their use. It should be noted that the erroneous use of the peak hour definition may cause significant consequences in most steps of the urban transportation planning process, particularly the assignment, trip distribution and modal split steps.

5.6.1 Hourly Variation of Travel

The shape of the hourly variation of travel for most urban centres is quite familiar to the transport planner, with the predominant two peaks occurring in the morning and in the afternoons because of the concentration of work trips as shown schematically in Figure 5.4. In most studies such a graph is plotted using 24 units on the horizontal axis representing each hour interval versus percentage, frequency or amount of vehicular or person travel on the vertical axis. There are basically two sources of information from which such

FIGURE 5.4

APPROACHES TO PEAK HOUR DEFINITION

a graph can be developed. These are:

1. actual ground counts, and
2. origin-destination surveys. 7

Although traffic counts can be obtained on a fairly detailed and accurate basis, the transportation planner rarely uses finer than 15 minutes interval data. With the development of computerized traffic control signals, it is now possible to calculate and monitor the performance of most intersections in a given urban area. Traffic volumes for each cycle can be obtained and stored for future manipulation. This may enable the transport planner to establish the actual occurrence of peak hour (defined as 60 minutes duration of highest traffic volume) for the entire study area or for individual geographic areas.

Such information can be extremely valuable since they provide an accurate and continuous source for use in the design of future facilities in a given location. The designer may also obtain the detailed peak hour travel variations from counts of an adjacent existing facility and use them in his design of the future link.

Hourly variation of travel developed using origin-destinations surveys may produce similar results to those obtained from actual ground counts. However, unlike ground counts which may be based on 100% coverage, O-D surveys are based on a limited sample. The advantage of using travel survey information is that they provide the means to distinguish among various trip purposes and traffic composition, the reasons behind the familiar shape of the hourly variation of travel curves.

5.6.2 Methods of Defining the Peak Hour

The peak hour from travel origin-destination surveys may be defined in terms of one of four methods or a combination thereof as shown in Table 5.5:

1. Trip Starting Time Method: This basically calls for stratification of all survey records using the trip starting time as reported by the traveller. Trip records may be stratified by 15 minute intervals. The four 15 minute intervals giving the highest frequency of records are considered to represent the peak hour for the study area. Obviously, the peak hour may be defined for various geographic areas using the same method. The main drawback of this technique is that it does not take the destination of the trip into consideration.
2. Trip Finishing Time Method: In this technique, trip finishing time as reported by the traveller may be used to stratify survey records. The four 15 minutes intervals giving the highest frequency of records would be taken as the peak hour. The method is therefore similar to the first but instead of using starting times, finishing times are used. The method suffers again from a similar shortcoming.
3. Mid Point Method: This is basically an attempt to reconcile the above two methods by taking the mid points of all trips accumulated using the starting times and those accumulated using the finishing trip times. Again, this method suffers from similar shortcoming as the previous two methods.

TABLE 5.5

APPROACHES TO PEAK HOUR DEFINITION

<u>Type</u>	<u>5 Minute Intervals</u>	<u>15 Minute Intervals</u>	<u>Approach</u>
1	X	✓	Starting times at origin are used and highest 4 @ 15 minutes intervals determine the peak hour.
2	X	✓	Arrival times at destinations are used and highest 4 @ 15 minutes intervals determine the peak hour
3	✓	✓	Starting times and arrival times at origins and destinations respectively are added and mid point established. The highest 12 @ 5 minutes intervals or the highest 4 @ 15 minutes intervals determine the peak hour
4	✓	X	Trip in Motion Concept is applied where both arrival times and starting times are used to calculate total trip length in minutes. The trip is added to each 6 minute intervals beginning with starting time until trip length is reached at the arrival time. The highest 12 @ 5 minutes intervals determine the peak hour. Note that it is possible with this concept to go to as much accuracy as deemed necessary.

4. Trip In Motion Method: Even though this technique is rather simple in concept, its applications have been rather limited. The technique examines the trip starting time, the trip finishing time and the trip duration and attempts to build an accounting procedure to check for every unit of intervals whether a given trip is in motion or not. An example of the concept is given in Table 5.6. The accuracy in this technique depends highly on the accuracy of the times provided by the traveller. Usually, respondents provide travel times to the closest 5 minutes. For this reason it is probably sufficiently accurate to accumulate over 5 minutes intervals and develop, using this concept, the highest 12 consecutive five minutes intervals giving peak occurrence of travel.

5.6.3 Defining Peak Hour Trips-In-Motion

Regardless of the method used to define the peak hour, the extraction of survey trips occurring during the peak hour involves examining those which have either their starting time, their finishing time or both occurring during the peak hour defined. Most urban transportation planning studies have extracted only those records which occur during the peak hour, not those that are in motion during the peak hour. The effect of extracting only the trips occurring during the peak hour, of course, would result with substantial underestimation of peak hour travel as shown in Table 5.7. Admittedly, the actual peak hour defined by any of the four methods described earlier may coincide.

TABLE 5.6

EXAMPLE OF TRIP IN MOTION CONCERT

Trip begins at 8:00 AM and ends at 8:30 AM

<u>Time</u>	<u>Trip Added</u>
6.9-7.9	No
7.0-8.0	No
7.1-8.1	Yes
7.2-8.2	Yes
7.3-8.3	Yes
7.4-8.4	Yes
7.5-8.5	Yes
7.6-8.6	Yes
7.7-8.7	Yes
7.8-8.8	Yes
7.9-8.9	Yes
8.0-9.0	Yes
8.1-9.1	Yes
8.2-9.2	Yes
8.3-9.3	Yes
8.4-9.4	Yes
8.5-9.5	No
8.6-9.6	No

TABLE 5.7
COMPARISON OF PK. HR. RATIOS BY
VARIOUS DEFINITIONS OF PEAK HOUR

<u>PK. Hr. Defined</u> <u>On the Basis Of</u>	<u>PK.HR.</u> <u>Records*</u>	<u>Total 24 HR</u> <u>Records</u>	<u>Peak Hour</u> <u>Ratio (%)</u>
Origin Time Hrs. (4-5) PM	2,296	24,071	9.5
Dest. Time Hrs. (5-6) PM	2,376	24,071	9.9
Trip in Motion (4:25-5:25) PM	3,178	24,071	13.2

* 1972 Ottawa-Carleton Home Interview Survey

Furthermore, although the use of the trip-in-motion concept may result with higher peak hour trips being extracted from the 24 hour trips, the actual distribution of these trips by purpose and mode may not cause the results to change significantly from those which may be developed on the basis of simple averaging or simple extraction. This does not suggest that transportation planners need not investigate accurately the trips occurring in the peak hour; quite the contrary, only with such an effort would it be possible to properly apply the definition of the peak hour in the urban transportation planning system model.

CHAPTER VI

STATISTICAL ESTIMATION OF PEAK HOUR RATIOS

6.1 Introduction

The modelling framework presented in Chapter IV described in broad terms the activities carried out in this research. This Chapter outlines the statistical techniques employed in the analysis of data and in the simulation of peak hour ratios. Discussion of the results of these simulation attempts is given in the next Chapter.

The statistical analysis was designed to cover an overall investigation of peak hour ratios determinants at the area-wide, district, zonal and disaggregate household and individual levels. Furthermore, ground counts from over 180 stations distributed throughout the area were analyzed in an attempt to relate peak hour ratios to individual station attributes.

The choice of a particular statistical technique was dictated by a number of factors including:

1. number of cases or observations,
2. number of dependent or independent variables,
3. availability of computer programmes and sub-routines,
4. suitability of the technique relative to other available techniques, and
5. continuity and usability of the results from one step to the next.

In brief, the statistical analysis performed involved the following six steps:

1. preliminary analysis which produced basic descriptive* statistics for the variables. (Correlation analysis and scattergram analysis were also performed to detect and visualize any possible relationship between the dependent and independent variables.),
2. canonical correlation analysis which attempted to develop two linear combinations from the two sets of dependent and independent variables in such a way that the correlation between the two linear combinations would be maximized,
3. factor analysis which attempted to reduce the large set of independent variables into a smaller set of factors or components that would be taken as source variables accounting for the observed interrelations in the data,
4. stepwise regression analysis which related the reduced set of independent variables from factor analysis to the dependent variables. (Both linear and non linear regressions were performed.),
5. category analysis which employed simple stratification of survey records in order to relate the disaggregated peak hour ratios to some attributes of the trip maker and/or his household, and finally,
6. both, stepwise regression and category analysis techniques were employed in the analysis of ground counts data.

* Statistics which describe the distribution, variability and central tendencies of the variables.

It is not the intention of this thesis to describe at length these statistical techniques. The interested reader is advised to consult an advanced text in statistics (36). Figure 6.1 summarizes in the form of a flowchart the sequence of statistical analysis activities carried out in this research. The reader will note that this flowchart basically represents the upper portion of the modelling framework presented in Figure 4.1.

6.2 Preliminary Analysis

The preliminary analysis step involved listing at various levels of analysis all dependent and independent variables in order to gain an overall understanding of the data. The zonal level system file, which contained all the variables was listed using an SPSS programme and the range of variation of each variable was examined.

Of particular interest was to note the missing data cases. Most of the independent variables contained over 250 cases. The missing cases in most instances were the external zones (from zone 285 to 300). There were however, certain independent variables such as secondary school enrollment, post secondary school enrollment and shopping centres gross leasable area which contained few observations. These variables which later in the research presented numerous problems at the factor analysis and regression steps, were therefore eliminated from further consideration.

The missing data problem for the dependent variables was also significant. Some trip purpose/trip mode combinations contained few observations only because of the small survey sample. Transit travel for purposes other than work in particular, contained significant amount of missing data. The driver and passenger modes on the other hand,

FIGURE 6.1

STATISTICAL ESTIMATION OF PEAK HOUR RATIOS

were well represented in the survey. The problem of missing data for the dependent variables in conjunction with the fact that some trip purpose/trip mode combinations did not contribute significantly to the total peak hour, indicated that all the dependent variables need not be studied. This conclusion was substantiated at a later stage in the research when weak relationships were derived for some of these dependent variables due to the few and scattered observations.

At both the zonal and district level analyses, 20 dependent variables (peak hour ratios by mode and purpose at the origin and destination ends) were developed as described in Chapter V. The statistical analysis step concentrated on simulation of peak hour ratios for those important trip mode/trip purpose combinations.

Condescriptive statistics were developed for all the independent and dependent variables. Of these statistics, the mean, the standard deviation and the number of missing observations were of particular interest. These are given in Table 6.1 for the 20 dependent variables.

An examination of the standard deviation and the mean for the dependent variables indicates that significant variations in the peak hour ratios are still present in the data in spite of the already stratified results by mode and purpose. In fact, the values of the standard deviation for each independent variable give a measure of the amount of variance or uncertainty still left in the data after the aggregation by mode and purpose. The statistical significance of these standard deviations and the fact that they are almost as large as the mean, if not larger, can lead into a lengthy discussion of the philosophy of their application and their eventual impacts.

TABLE 6.1

CONDESCRIPTIVE STATISTICS FOR
THE 20 DEPENDENT VARIABLES*

CORRECT CIST SYS FILE FOR MISS. V.									
FILE		SYSOPL		(CREATION DATE = 05/19/78)		05/22/78		PAGE	
VARIABLE		VAR127		HJ D AND P HSRO					
MEAN		70.518		STD ERROR		6.964		STD DEV	
VARIANCE		2730.634		KURTOSIS		11.710		SKENESS	
RANGE		244.000		MINIMUM		17.000		MAXIMUM	
SUM		3243.837						47.230	
								2.683	
								301.000	
VALID OBSERVATIONS		46		MISSING OBSERVATIONS		14			
VARIABLE		VAR128		HJ C AND P HSRO					
MEAN		58.875		STD ERROR		10.436		STD DEV	
VARIANCE		4900.777		KURTOSIS		5.444		SKENESS	
RANGE		354.500		MINIMUM		10.500		MAXIMUM	
SUM		3929.396						70.006	
								2.070	
								365.000	
VALID OBSERVATIONS		45		MISSING OBSERVATIONS		15			
VARIABLE		VAR129		HJ D AND P HSCH					
MEAN		109.767		STD ERROR		25.662		STD DEV	
VARIANCE		13622.461		KURTOSIS		5.341		SKENESS	
RANGE		489.000		MINIMUM		12.000		MAXIMUM	
SUM		2305.114						117.599	
								2.120	
								500.000	
VALID OBSERVATIONS		21		MISSING OBSERVATIONS		39			
VARIABLE		VAR130		HJ C AND P HSCH					
MEAN		63.125		STD ERROR		9.597		STD DEV	
VARIANCE		1842.172		KURTOSIS		0.172		SKENESS	
RANGE		143.000		MINIMUM		12.000		MAXIMUM	
SUM		1262.501						42.921	
								1.154	
								155.000	
VALID OBSERVATIONS		20		MISSING OBSERVATIONS		40			

* All variables are in thousands, e.g. the value 70.518 above represents an average peak hour ratio of 7.0518%

TABLE 6.1 (Continued)

CONDESCRIPTIVE STATISTICS FOR
THE 20 DEPENDENT VARIABLES

CORRECT CIST SYS FILE FOR MISS. V.									
FILE	SYSDFL	(CREATION DATE = 05/19/78)						05/23/78	PAGE 20
VARIABLE	VAR131	HJ	TRANS.	HW					
MEAN	295.167								
VARIANCE	9770.058				STD ERROR	16.474		STD DEV	98.844
RANGE	387.000				KURTOSIS	-0.643		SKEWNESS	0.120
SUM	9546.027				MINIMUM	75.000		MAXIMUM	462.000
VALID OBSERVATIONS	36				MISSING OBSERVATIONS				
									24
VARIABLE	VAR132	HJ	TRANS.	HW					
MEAN	202.219				STD ERROR	23.307		STD DEV	123.328
VARIANCE	15209.910				KURTOSIS	-0.231		SKEWNESS	0.707
RANGE	440.014				MINIMUM	27.200		MAXIMUM	467.214
SUM	5662.117								
VALID OBSERVATIONS	28				MISSING OBSERVATIONS				
									32
VARIABLE	VAR133	HJ	TRANS.	HSPB					
MEAN	160.815				STD ERROR	20.131		STD DEV	96.547
VARIANCE	9321.281				KURTOSIS	6.322		SKEWNESS	2.015
RANGE	466.909				MINIMUM	33.091		MAXIMUM	500.000
SUM	3698.745								
VALID OBSERVATIONS	23				MISSING OBSERVATIONS				
									37
VARIABLE	VAR134	HJ	TRANS.	HSPB					
MEAN	160.424				STD ERROR	19.089		STD DEV	68.825
VARIANCE	4716.824				KURTOSIS	0.021		SKEWNESS	0.513
RANGE	249.357				MINIMUM	54.143		MAXIMUM	303.500
SUM	2085.517								
VALID OBSERVATIONS	13				MISSING OBSERVATIONS				
									47

TABLE 6.1 (Continued)

CONDESCRIPTIVE STATISTICS FOR
THE 20 DEPENDENT VARIABLES

CORRECT DIST SYS FILE FOR MISS. V.									
FILE	SYSOFL	(CREATION DATE = 05/19/78)						05/21/78	PAGE 21
VARIABLE	VARI35	HJ D AND P HW							
MEAN	234.590			STD ERROR	17.090			STD DEV	117.166
VARIANCE	13727.855			KURTOSIS	-0.283			SKEWNESS	0.308
RANGE	492.500			MINIMUM	7.500			MAXIMUM	500.000
SUM	11025.742								
VALID OBSERVATIONS	47	MISSING OBSERVATIONS		13					
VARIABLE	VARI36	HI D AND P HW							
MEAN	188.924			STD ERROR	16.042			STD DEV	112.297
VARIANCE	12610.547			KURTOSIS	-0.043			SKEWNESS	0.705
RANGE	428.750			MINIMUM	29.750			MAXIMUM	457.500
SUM	9257.273								
VALID OBSERVATIONS	49	MISSING OBSERVATIONS		11					
VARIABLE	VARI37	HJ C AND P HSPB							
MEAN	103.380			STD ERROR	11.520			STD DEV	78.131
VARIANCE	6104.523			KURTOSIS	2.428			SKEWNESS	1.552
RANGE	352.500			MINIMUM	13.500			MAXIMUM	366.000
SUM	4755.461								
VALID OBSERVATIONS	46	MISSING OBSERVATIONS		14					
VARIABLE	VARI38	HI D AND P HSPB							
MEAN	107.814			STD ERROR	10.825			STD DEV	74.214
VARIANCE	5507.680			KURTOSIS	6.745			SKEWNESS	2.236
RANGE	413.333			MINIMUM	13.000			MAXIMUM	426.333
SUM	5067.270								
VALID OBSERVATIONS	47	MISSING OBSERVATIONS		13					

TABLE 6.1 (Continued)

CONDESCRIPTIVE STATISTICS FOR
THE 20 DEPENDENT VARIABLES

CORRECT CTSY SYS FILE FOR MISS. V.									
FILE	SYSOFL	(CREATION DATE = 05/19/78)						05/23/78	PAGE 22
VARIABLE VAR139 HJ D AND P NHB									
MEAN	127.144	STD ERROR	13.145	STD DEV				91.070	
VARIANCE	8293.824	KURTOSIS	6.449	SKEWNESS				2.355	
RANGE	467.800	MINIMUM	16.200	MAXIMUM				484.000	
SUM	6132.891								
VALID OBSERVATIONS -	48	MISSING OBSERVATIONS -	12						
VARIABLE VAR140 HI D AND P NFB									
MEAN	116.398	STD ERROR	12.240	STD DEV				84.804	
VARIANCE	7191.684	KURTOSIS	8.668	SKEWNESS				2.449	
RANGE	477.400	MINIMUM	22.600	MAXIMUM				500.000	
SUM	5597.125								
VALID OBSERVATIONS -	48	MISSING OBSERVATIONS -	12						
VARIABLE VAR141 HJ TRANS. NHD									
MEAN	154.100	STD ERROR	21.815	STD DEV				111.236	
VARIANCE	12373.488	KURTOSIS	1.720	SKEWNESS				1.436	
RANGE	410.125	MINIMUM	28.875	MAXIMUM				439.000	
SUM	4006.605								
VALID OBSERVATIONS -	26	MISSING OBSERVATIONS -	34						
VARIABLE VAR142 HI TRANS. NHD									
MEAN	144.560	STD ERROR	15.800	STD DEV				99.002	
VARIANCE	5831.441	KURTOSIS	2.785	SKEWNESS				1.598	
RANGE	392.286	MINIMUM	35.714	MAXIMUM				428.000	
SUM	3814.908								
VALID OBSERVATIONS -	25	MISSING OBSERVATIONS -	35						

TABLE 6.1 (Continued)

CONDESCRIPTIVE STATISTICS FOR
THE 20 DEPENDENT VARIABLES

CORRECT DTST SYS FILE FOR HTSS. V.											
FILE		SYSDFL		(CREATION DATE = 05/19/78)		05/23/78		PAGE 2			
VARIABLE		VARI43		HJ TRANS. HSR0							
MEAN		181.470				STD ERROR		25.375		120.867	
VARIANCE		14633.840				KURTOSIS		0.852		1.136	
RANGE		448.286				MINIMUM		51.714		500.000	
SUM		3810.866									
VALID OBSERVATIONS		21				MISSING OBSERVATIONS		39			
VARIABLE		VARI44		HI TRANS. HSR0							
MEAN		206.045				STD ERROR		38.950		140.436	
VARIANCE		10722.371				KURTOSIS		1.331		1.273	
RANGE		479.786				MINIMUM		63.714		543.500	
SUM		2678.583									
VALID OBSERVATIONS		13				MISSING OBSERVATIONS		47			
VARIABLE		VARI45		HJ TRANS. HSCH							
MEAN		107.563				STD ERROR		12.960		72.161	
VARIANCE		5237.211				KURTOSIS		1.684		1.032	
RANGE		323.222				MINIMUM		9.778		333.000	
SUM		3334.461									
VALID OBSERVATIONS		31				MISSING OBSERVATIONS		29			
VARIABLE		VARI46		HI TRANS. HSCH							
MEAN		86.509				STD ERROR		11.839		63.756	
VARIANCE		4064.826				KURTOSIS		0.684		1.030	
RANGE		250.500				MINIMUM		5.000		255.500	
SUM		2539.757									
VALID OBSERVATIONS		29				MISSING OBSERVATIONS		31			

on transport link design. Table 3.1 in Chapter III has already demonstrated these impacts in the form of a simple example. The author will return to this topic in Chapter VIII. It may be sufficient to state that since the frequency distribution of each of the 20 dependent variables showed significant variations, further analysis was necessary to establish the determinants of these variations.

To gain further insight into the nature of the data used, correlation analysis and scattergram analysis were performed.

6.2.1 Correlation Analysis

This type of analysis was performed for all the dependent and independent variables. The purpose was to establish at a glance the nature of correlation between the variables, as a guide for upcoming more thorough statistical analysis.

The Pearson type bivariate correlation analysis was used for this exercise. This type of correlation provides a single number which summarizes the relationship between two variables:

The following information was obtained from this analysis:

- a. the number of valid observations (zones),
- b. Pearson's correlation coefficients, and
- c. the level of significance of each correlation using a t-test with $n-2$ degrees of freedom.

A typical result of this analysis is shown in Table 6.2. Examination of the Pearson's correlation coefficients revealed that significant collinearity exists among the independent variables but no immediately apparent strong correlation exists between the dependent and independent variables.

6.2.2 Scattergram Analysis

This is basically a two-dimensional plot of data points for the values of the two variables being plotted. The vertical axis was chosen to represent the dependent variable while the horizontal axis represented the independent variable. A typical scattergram is given in Figure 6.2.

The scattergram plot serves mainly the following functions:

1. it provides a graphical representation of the association between two variables,
2. it outlines the range (upper and lower bounds) of the values of each variable,
3. it enables visual inspection of correlation patterns, linear or non-linear, and
4. it identifies extreme or unusual observations for possible explanation.

Literally, tens of scattergrams were plotted for various dependent variables against most of the independent variables. For each scattergram a number of statistics were developed including:

- Correlation Coefficient
- Significance Level
- Standard Error of Estimate
- Intercept of a Fitted Linear Curve

TABLE 6.2

TYPICAL OUTPUT FROM PEARSON CORRELATION ANALYSIS

RUK PEARSON CORRELATIONS, DEP AGAINST INDEP									
FILE SPALLAL (CREATION DATE = 05/05/70)									
PEARSON CORRELATION COEFFICIENTS									
	VAR201	VAR202	VAR203	VAR204	VAR205	VAR206	VAR207	VAR208	VAR209
VAR130	-0.2675 (173) S=0.231	-0.1562 (133) S=0.211	-0.0752 (117) S=0.201	-0.2042 (129) S=0.201	-0.2637 (123) S=0.201	-0.3067 (134) S=0.201	-0.2622 (131) S=0.201	0.3167 (27) S=0.201	0.0015 (127) S=0.201
VAR135	-0.1152 (152) S=0.177	-0.1127 (153) S=0.175	-0.1147 (146) S=0.175	-0.1137 (149) S=0.177	-0.1374 (140) S=0.173	-0.0443 (136) S=0.176	0.3083 (5) S=0.177	-0.0902 (27) S=0.177	-0.1655 (146) S=0.173
VAR140	-0.2474 (145) S=0.201	-0.1317 (113) S=0.201	-0.1115 (145) S=0.207	-0.2160 (141) S=0.201	-0.2154 (123) S=0.206	0.1453 (40) S=0.157	0.6602 (5) S=0.112	0.0421 (25) S=0.144	0.0131 (140) S=0.136
VAR141	-0.1350 (121) S=0.147	0.0741 (40) S=0.172	-0.1255 (37) S=0.170	-0.0131 (12) S=0.176	-0.0776 (12) S=0.174	-0.1374 (3) S=0.173	-1.0000 (2) S=0.177	-0.0440 (6) S=0.172	0.1627 (33) S=0.183
VAR142	0.1140 (34) S=0.163	0.0548 (45) S=0.163	-0.2330 (35) S=0.166	-0.1064 (32) S=0.173	0.1401 (31) S=0.216	-0.4340 (11) S=0.191	-0.7134 (3) S=0.147	-0.4974 (5) S=0.107	-0.0005 (33) S=0.100
VAR143	0.0745 (23) S=0.144	-0.2112 (24) S=0.172	-0.2256 (21) S=0.177	0.1177 (23) S=0.206	-0.1007 (21) S=0.132	-0.1125 (5) S=0.176	99.0000 (0) S=0.000	99.0000 (1) S=0.000	0.1645 (23) S=0.200
VAR144	-0.0644 (1) S=0.120	-0.2854 (23) S=0.120	-0.3157 (22) S=0.144	-0.1240 (1) S=0.106	-0.1141 (16) S=0.127	-0.3095 (7) S=0.170	1.0000 (2) S=0.000	1.0000 (2) S=0.000	-0.0367 (20) S=0.184
VAR145	-0.3157 (5) S=0.173	0.0363 (5) S=0.173	0.3174 (5) S=0.173	-0.2354 (5) S=0.173	-0.3600 (5) S=0.173	-0.1062 (17) S=0.173	99.0000 (1) S=0.000	0.1240 (6) S=0.177	0.3253 (58) S=0.173
VAR146	-0.2117 (41) S=0.157	0.0777 (44) S=0.174	0.1007 (10) S=0.241	-0.1641 (40) S=0.176	-0.1515 (3) S=0.179	0.1034 (20) S=0.173	-1.0000 (2) S=0.000	-0.0473 (3) S=0.166	0.2017 (40) S=0.106
PEARSON CORRELATION COEFFICIENTS									
	VAR201	VAR202	VAR203	VAR204	VAR205	VAR206	VAR207	VAR208	VAR209
VAR127	-0.1727 (144) S=0.227	-0.1562 (144) S=0.226	-0.1051 (171) S=0.116	-0.1340 (134) S=0.054	-0.0344 (123) S=0.346	0.2045 (5) S=0.112	-0.5657 (27) S=0.122	-0.0479 (20) S=0.120	-0.0216 (136) S=0.101
VAR128	-0.4400 (114) S=0.001	-0.1142 (145) S=0.001	0.1466 (134) S=0.144	-0.1463 (134) S=0.001	-0.4062 (123) S=0.001	-0.0041 (39) S=0.270	0.4145 (21) S=0.027	0.4257 (21) S=0.027	0.0963 (136) S=0.183
VAR129	0.0745 (27) S=0.144	-0.2112 (24) S=0.172	-0.2256 (21) S=0.177	-0.0730 (27) S=0.177	0.0533 (27) S=0.177	0.5313 (27) S=0.177	99.0000 (0) S=0.000	0.2055 (3) S=0.177	-0.3235 (27) S=0.177
VAR130	0.4175 (25) S=0.173	-0.1463 (25) S=0.173	-0.2665 (13) S=0.173	0.3244 (24) S=0.173	-0.2054 (19) S=0.173	-0.4851 (11) S=0.173	-0.7651 (3) S=0.173	-0.2844 (5) S=0.173	-0.2023 (24) S=0.173
VAR131	0.1754 (114) S=0.001	-0.2655 (112) S=0.117	-0.1125 (112) S=0.117	0.2164 (112) S=0.117	0.1747 (114) S=0.031	0.3865 (25) S=0.021	0.7471 (3) S=0.231	-0.3766 (14) S=0.072	-0.1448 (116) S=0.081
VAR132	-0.0524 (27) S=0.101	0.5100 (41) S=0.101	0.1472 (66) S=0.004	-0.5167 (64) S=0.001	-0.4706 (15) S=0.001	-0.3135 (13) S=0.127	1.0000 (2) S=0.000	0.0112 (6) S=0.075	-0.0111 (61) S=0.174
VAR133	-0.1271 (11) S=0.214	-0.0747 (11) S=0.112	-0.0820 (43) S=0.112	-0.0109 (41) S=0.102	-0.1465 (40) S=0.170	-0.1752 (13) S=0.280	99.0000 (0) S=0.000	-0.5403 (0) S=0.075	0.2496 (41) S=0.078
VAR134	-0.2667 (25) S=0.173	-0.1127 (10) S=0.173	-0.1747 (27) S=0.173	-0.1147 (15) S=0.173	-0.2165 (19) S=0.173	0.4146 (3) S=0.164	99.0000 (0) S=0.000	-0.5531 (21) S=0.223	0.1350 (24) S=0.232
VAR135	0.3142 (144) S=0.001	-0.4450 (147) S=0.001	-0.3113 (154) S=0.001	0.2757 (153) S=0.001	0.4148 (163) S=0.001	0.1235 (43) S=0.212	-0.9701 (4) S=0.313	-0.4022 (20) S=0.314	0.0761 (161) S=0.162
VAR136	-0.5547 (103) S=0.001	0.0511 (107) S=0.001	0.2055 (103) S=0.001	-0.5174 (111) S=0.001	-0.4701 (150) S=0.001	-0.0476 (42) S=0.373	0.9651 (27) S=0.174	0.2087 (21) S=0.147	0.1472 (157) S=0.038
VAR137	-0.0515 (114) S=0.001	-0.1651 (114) S=0.001	-0.1170 (125) S=0.001	-0.3123 (125) S=0.001	-0.3174 (125) S=0.001	-0.0834 (31) S=0.328	0.5574 (21) S=0.001	-0.7715 (21) S=0.075	0.7125 (126) S=0.008

(COEFFICIENT / CASES) / SIGNIFICANCE)

(A VALUE OF 99.0000 IS PRINTED IF A COEFFICIENT CANNOT BE COMPUTED)

FIGURE 6.2

TYPICAL OUTPUT FROM SCATTERGRAM ANALYSIS

- . Slope of the Fitted Line
- . Number of Plotted Values, and
- . Number of Missing Values

In summary, the preliminary analysis step revealed the following problem areas:

1. multicollinearity of independent variables,
2. missing data, and
3. input data inconsistency, particularly for the independent variables since data came from different sources and for different years.

6.3 Canonical Correlation Analysis

This type of analysis is basically similar to other types of multivariate linear statistical techniques in that its basic aim is to account for as much variance in certain variables as possible.

In order to explain the reason for using this type of analysis, it is probably useful to outline briefly how some multivariate statistics are used. Consider the common situation in which a few variables are to be analyzed against a dependent variable which is to be explained. The obvious choice here would be Multiple Regression. When the number of variables increases to a point where it becomes impossible to provide a meaningful interpretation of the findings without data reduction, the Factor Analysis technique may be employed. It should be noted that factor analysis does not make the distinction between dependent and independent variables. If it is desired to make such distinction, the data may be divided into two sets and factor analysis may be performed for each set separately. This was actually done in this research since the two sets were distinct.

However, it was desired to explore another strategy. This strategy, which is the basic element of canonical correlation analysis, is to derive a linear combination from each of the two sets of variables in such a way, that the correlation between the two linear combinations is maximized. Many such pairs of linear combinations may be derived. These are termed Canonical Variates (37,38).

These variates come in two sets, one for each of the subsets of variables entered into the analysis. These variates are composed of coefficients that reflect the importance of the original variables in the subset in forming the variates. The first canonical variate from the first set of variables and the first canonical variate from the second set of variables are chosen so as to maximally correlate with each other, and similarly for the second and all successive pairs of canonical variates. The amount of correlation between each corresponding pair of canonical variates is the canonical correlation between them, and its square represents the amount of variance in one canonical variate that is accounted for by the other canonical variate. Since a correlation coefficient, canonical or otherwise, between A and B is always equivalent to that between B and A, it makes no difference mathematically which subset of variables is considered to be the dependent variables and which the independent variables, the results will always be the same.

Canonical correlation analysis was attempted in this research to manipulate intercorrelation among variables to see if a particular type of patterning exists in the data. It was hoped that if such patterning did exist, some theoretical premises may be derived to explain it.

Table 6.3 shows the results of canonical correlation analysis. Two canonical correlations were produced with values over 0.5. The first which has a value of 0.84 indicates that the amount of variance shared by the two first canonical variates is 70 percent. The second canonical correlation with a value of 0.61 indicates that the second canonical variates share only 37% of their variance. Thus, only the first canonical variates were produced.

Examination of the loadings of the individual variables on the canonical variates shows that variables TMTRACBD and LABFORCE load with values higher than 0.5 from the first subset of variables with some 8 other variables loading between 0.2 and 0.5. From the second subset only HJDPHWRK loads with a value higher than 0.5. The interpretation of these results suggests that work peak hour ratios for driver and passenger at the destination end seems to intercorrelate with travel time by transit from the CBD and the zonal labour force. The relationships among the dependent and independent variables are almost all unaccounted for by the canonical variates developed. In fact, even the one pair of canonical variates developed does not seem to portray significant variables loadings. This evidence indicates that no significant canonical correlation can be developed for the two sets of variables used in this research.

6.4 Factor Analysis

The single most distinctive characteristic of factor analysis is its data reduction capability. Given an array of correlation coefficients for a set of variables, factor-analytic techniques enable the user to see whether

TABLE 6.3

RESULTS OF THE CANONICAL CORRELATION ANALYSIS

<u>Number</u>	<u>Eigenvalue</u>	<u>Canonical Correlation</u>	<u>Wilk's Lambda</u>	<u>Chi-Square</u>	<u>D.F.</u>	<u>Significance</u>
1	0.70151	0.83756	0.07978	284.45483	153	0.000
2	0.37135	0.60939	0.26728	148.43869	128	0.103
3	0.22955	0.47912	0.42517	96.21788	105	0.721
4	0.21111	0.45947	0.55185	66.87979	84	0.913
5	0.12674	0.35601	0.69952	40.20297	65	0.992
6	0.09476	0.30783	0.80105	24.95677	48	0.998
7	0.06606	0.25702	0.88489	13.75727	33	0.999
8	0.02936	0.17135	0.94748	6.06891	20	0.999
9	0.02386	0.15445	0.97614	2.71622	9	0.974

TABLE 6.3 (Continued)

COEFFICIENTS FOR CANONICAL VARIABLES
OF THE FIRST SET

	<u>CANVAR 1</u>
CARINDEX	0.27841
TRAINDEX	0.07832
POPULATN	0.38626
TOTEMPLY	0.41998
RETEMPLY	0.03119
DWLGUNIT	0.25827
TOTLCARS	0.10285
TMTRACBD	0.57856
TMCARCB	0.23063
PARKAVAL	0.07529
NACREAGE	0.07634
INCOMEZN	0.19342
LABFORCE	0.67125
NTPOPDEN	0.25582
NTRESDEN	0.41605
NTEMPDEN	0.11542
ZONATTR	0.10822

COEFFICIENTS FOR CANONICAL VARIABLES
OF THE SECOND SET

	<u>CANVAR 1</u>
HJDPHSRO	0.05059
HIDPHSRO	0.18487
HJTRANHW	0.04715
HJDPHWRK	0.52441
HIDPHWRK	0.46358
HJDPHSPB	0.05317
HIDPHSPB	0.06556
HJD&PNHB	0.11236
HID&PNHB	0.04081

some underlying pattern of relationship exists such that the data may be reduced to a smaller set of factors or components that may be taken as source variables accounting for the observed interrelations in the data (39).

Factor analysis may be used to explore or detect patterns of variation in the hope of possible reduction of data (this is the most common use); to confirm significant factors, or to develop new indices which may be used as new variables in later analysis.

There are basically three steps in factor analysis:

1. preparation of the correlation matrix,
2. extraction of the initial factors (the exploration of possible data reduction), and
3. the rotation to a terminal solution (i.e. the search for simple and interpretable factors).

Further discussion of the factor analysis technique is presented in Appendix F. For a more elaborate description, the reader is advised to consult an advanced text in Statistics (40).

In the next few paragraphs a simple example is presented to illustrate the use and interpretation of factor analysis results. The factor analysis procedure was performed for the following set of variables:

DWLGUNIT - Dwelling Units
 TOTLCARS - Cars
 LABFORPT - Labour Force Participation Ratio
 NACREAGE - Net Acreage
 INCOMEZN - Average Zonal Income
 NTPOPDEN - Net Population Density
 ZONATTRA - Zonal Attractivity
 TRAIINDEX - Index of Transit Accessibility

The table below gives the percent of common variance accounted for by unrotated factors. It clearly shows that a significant portion of variation in the data is explained by the first three factors. With the use of the factor analysis technique it may be possible to represent the original 8 variables by 3 factors.

<u>Factor</u>	<u>Pct. of Variance</u>	<u>Cum. Pct.</u>
1	42.9	42.9
2	38.9	81.8
3	8.7	90.5
4	7.1	97.6
5	2.4	100.0

The resultant varimax and factor score matrices are given in Table 6.4. The varimax rotated factor matrix (also referred to as the variables-factors matrix) gives some indication of the importance each variable plays with respect to the five factors. The reader can thus see that the DWLGUNIT and TOTLCAR variables significantly load factor 1; the NACREAGE and NTPOPDEN variables significantly load factor 2; and that TRAINDEX, and to a lesser degree, ZONATTRA load factor 3.

The factor score matrix shown in Table 6.4 provides the coefficients to apply on the standardized variables to obtain the factors.

6.4.1 Factor Analysis for the Dependent Variables

This analysis was attempted using the 20 dependent variables ($H_i^{m,P}$ and $H_j^{m,P}$) in order to establish whether the dependent variables themselves could be reduced to a smaller set. The value 0 for $H_i^{m,P}$ and $H_j^{m,P}$ was treated

TABLE 6.4

VARIMAX MATRIX AND FACTOR SCORE MATRIX FOR
AN ILLUSTRATIVE EXAMPLE OF FACTOR ANALYSIS

VARIMAX MATRIX

		F A C T O R				
		1	2	3	4	5
V						
A	DWLGUNIT	.915	.285	.050	-.006	-.004
R	TOTLCARS	.842	-.155	.085	.108	.169
	LABFORPT	-.055	.090	.156	.452	.022
I	NACREAGE	.397	-.567	-.149	-.074	-.288
A	INCOMEZN	.058	.015	-.053	.020	.436
	NTPOPDEN	.304	.752	.145	.180	-.108
B	ZONATTRA	.077	.081	.448	.144	-.095
L	TRAINDEX	-.060	.483	.518	.425	.112
E						

FACTOR SCORE MATRIX

		F A C T O R				
		1	2	3	4	5
V						
A	DWLGUNIT	.700	.343	-.186	-.012	-.116
R	TOTLCARS	.280	-.334	.280	-.107	.422
	LABFORPT	.019	-.068	-.017	.337	.014
I	NACREAGE	.130	.338	.006	.193	.415
A	INCOMEZN	.001	.004	-.076	.039	.331
	NTPOPDEN	-.024	.502	.104	.024	.264
B	ZONATTRA	.000	-.068	.304	-.005	.107
L	TRAINDEX	-.039	.061	.524	.378	.129
E						

as a missing value. R-type factoring* was used and 5 factors were sought. Six iterations were performed using the varimax method of rotation.

The factor analysis programme used in the research provided, as output, various condscriptive statistics and correlation matrices. It was interesting to note from the correlation coefficients produced by the programme that no significant multicollinearity among the dependent variables themselves seems to exist. The varimax rotated factor matrix did not show substantial loading of any one variable on any one factor, with the exception of factor 1 where the variables HITRANHW, HJTRANHW, HIDPHWRK and HJDPHWRK all loaded highly. This suggests that the home to work component is the only major component which can be represented by a single factor. It is quite possible to relate many variables to the other four factors but substantial number of variables will be involved, thus indicating that no single dimension represents any of these variables. The four factors will, in effect use most of the variables and therefore, little could be gained from using the factors instead of the variables.

Table 6.5 shows the varimax matrix developed while Figure 6.3 illustrates a typical plot of the rotated factors and the relative position of the variables. Consider for instance variables 9 and 5 which represent HJDPHWRK and HJTRANHW respectively. These two variables (H_j , Home-Work, Driver plus Passenger and Transit respectively) load basically Factor 1 and minimally Factor 2. Examination of the varimax matrix indicates this to be the case.

* See Appendix F for a brief overview of the factor analysis technique.

TABLE 6.5

VARIMAX MATRIX FOR THE DEPENDENT VARIABLES

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
HJDPHSRO	0.10142	0.09496	-0.55276	0.12249	0.12758
HIDPHSRO	-0.15612	0.41241	-0.08730	0.04654	0.17539
HJDPHSCH	0.03402	0.02819	-0.17584	0.66585	0.06807
HIDPHSCH	-0.02325	0.36460	-0.04441	0.22767	-0.15085
HJTRANHW	0.71007	-0.19973	-0.22714	-0.08645	0.09377
HITRANHW	-0.52012	0.33132	0.29522	-0.21896	0.08080
HJTRHSPB	0.28707	-0.01695	0.18609	0.39906	-0.12082
HITRHSBP	-0.47729	0.02239	-0.10170	-0.15161	0.21896
HJDPHWRK	0.68846	-0.15647	-0.25702	0.12723	-0.07823
HIDPHWRK	-0.53602	0.32035	0.11653	-0.24725	0.15522
HJDPHSPB	0.32239	0.02933	0.05363	0.01800	-0.00819
HIDPHSPB	-0.06873	0.00995	0.20162	-0.09979	0.52158
HJD&PNHB	0.13556	0.21117	0.04267	0.13220	0.23657
HID&PNHB	-0.15889	0.35106	0.10378	-0.00296	0.08131
HJTRANHB	0.05568	0.20647	0.38752	0.06193	0.18851
HITRANHB	-0.02271	0.66294	0.03258	-0.23914	-0.09914
HJTRHSRO	0.15116	-0.09742	0.09788	-0.03723	-0.20275
HITRHSRO	-0.01631	-0.01314	-0.04743	0.01184	0.32512
HJTRHSCH	0.22746	-0.00954	0.13271	0.01962	0.08210
HITRHSCH	-0.24601	0.07098	-0.01011	-0.26147	-0.22835

FIGURE 6.3

ROTATED FACTORS FROM FACTOR ANALYSIS
FOR THE DEPENDENT VARIABLES

6.4.2 Factor Analysis for the Independent Variables

Several attempts to factor analyze the independent variables were performed. These factor analyses included most of the independent variables available in the system file which are listed in Table 6.6. It was recognized from previous analyses performed that many of these variables were collinear. Indeed, the correlation matrix produced by the factor analysis programme confirmed that substantial collinearity existed among many "Independent" variables.

It is apparent from an examination of Table 6.7 below that the variation in the total set of data may be accounted for in terms of five factors. Furthermore, about 90% of the variation can be accounted for by the first three factors.

TABLE 6.7
PERCENT OF VARIANCE EXPLAINED BY THE FIVE FACTORS

<u>Factor</u>	<u>Pct. of Variance</u>	<u>Cum. Pct.</u>
1	44.1	44.1
2	28.9	73.0
3	16.2	89.2
4	7.1	96.3
5	3.7	100.0

The varimax rotated factor matrix given in Table 6.8 was closely examined. For each factor, the variables loading it heavily were identified and listed. The intention was to find out an interpretation of what each factor meant or represented in terms of its constituent variables. Table 6.9 shows for each factor those variables with significant loading. A "name" or a "label" was affixed to each factor. For instance, it was clear that

TABLE 6.6

INDEPENDENT VARIABLES USED AS
INPUT TO FACTOR ANALYSIS

<u>Variables</u>	<u>Labels</u>
POPULATN	Population
TOTEMPLY	Total Employment
RETEMPLY	Retail Employment
DWLGUNIT	Dwelling Unit
TOTLCARS	Cars
LABFORPT	Labour Force Part. Ratio
TMTRACBD	Time by Transit from CBD
TMCARCB	Time by Car from CBD
PARKAVAL	Parking Availability
NACREAGE	Net Acreage
INCOMEZ	Income
LABFORCE	Labour Force
NTPODEN	Net Population Density
NTRES	Net Resident. Density
NTEMPDEN	Net Empl. Density
ZONATTR	Zonal Attract.
CARINDEX	Car Index of Attract.
TRAINDEX	Transit Index of Attract.

TABLE 6.8

VARIMAX MATRIX FOR THE INDEPENDENT VARIABLES

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
POPULATN	-0.08275	0.96298	-0.11397	0.08258	-0.07025
TOTEMPLY	0.16412	-0.15193	0.93725	0.04947	0.06909
RETEMPLY	-0.00377	-0.07234	0.42706	-0.04509	-0.53954
DWLGUNIT	0.01906	0.90904	-0.06912	0.30871	-0.00866
TOTLCARS	-0.21912	0.80385	-0.14830	-0.12391	0.01064
LABFORPT	0.23661	-0.03502	0.14009	0.09120	0.32641
TMTRACBD	-0.97088	0.05134	-0.12446	-0.09137	-0.09481
TMCARCBD	-0.92082	0.10842	-0.20826	-0.21529	-0.04770
PARKAVAL	0.74880	0.04400	0.17312	0.34550	0.02816
NACREAGE	-0.67113	0.19259	0.03863	-0.18197	0.06815
INCOMEZN	-0.02523	0.13091	-0.29641	-0.12811	0.04445
LABFORCE	0.00022	0.96346	-0.09289	0.13245	0.07193
NTPOPDEN	0.42919	0.26450	0.00688	0.83888	0.04085
NTRESDEN	0.34371	0.15539	0.14354	0.85056	0.13514
NTEMPDEN	0.19594	-0.18961	0.71102	-0.06703	0.23889
ZONATTRA	0.11251	0.12409	0.94898	0.03248	-0.24901
CARINDEX	0.94336	-0.01791	0.01091	-0.00788	0.13218
TRAINDEX	0.83315	-0.03028	0.22465	0.17974	0.13659

TABLE 6.9
ZONAL LEVEL FACTOR ANALYSIS

FACTOR	V A R I A B L E		FACTOR LABEL
	LOADING	LABEL	
1	.97	TMTRACBD Time by Transit from CBD	Accessibility
	.92	TMCARCB Time by Car from CBD	
	.74	PARKAVAL Parking Availability	
	.67	NACREAGE Net Acreage	
	.94	CARINDEX Car Index of Attract.	
	.83	TRAINDEX Transit Index of Attract.	
2	.96	POPULATN Population	Origin Zone Demographic Characteristics
	.90	DWLGUNIT D.U.	
	.80	TOTLCARS Cars	
	.96	LABFORCE Labour Force	
3	.93	TOTEMPLY Total Employment	Destination Zone Demographic Characteristics
	.71	NTEMPDEN Net Employ. Density	
	.94	ZONATTRA Zonal Attractivity	
4	.83	NTPOPDEN Net Population Density	Origin Zonal Density
	.85	NTRESDEN Net Residential Density	
5	.53	RETEMPY Retail Employment	Same as Variable Label

the variables: POPULATN, DWLGUNIT, TOTLCARS and LABFORCE all reflect "Origin Zone Demographic Characteristics" and all load heavily on factor 2. In certain cases such as with the last factor, the naming of the factor is rather difficult since the constituent variables seem to be unique in interpretation. For this reason, another factor analysis was performed using seven factors instead of five and, although the amount of variance accounted for by the last two factors was small, they nevertheless did appear in the variables-factors matrix and certain variables uniquely loaded on these factors.

Returning to Table 6.8, it was interesting to note that each of the variables entered in Table 6.9 loaded heavily and almost uniquely on only one factor and not on two or more. In fact, if a variable did load significantly on two or more factors, it means that no simple labelling can be affixed to that variable but rather it has multi-dimensional explanation. This heavy loading on a single factor is achieved through the varimax factor rotation technique. Figure 6.4 shows a typical plot of the rotated factors and the final position of the independent variables.

At this stage two options were available:

- a. use the factor scores obtained from factor analysis as input to the stepwise regression. This was not attempted because, for application purposes, these factor scores are difficult to use since the meaning of some factors is difficult to define and quantify. Additionally, their use would mean that most of the "independent" variables which are constituents of these factors must be included and developed at the application and forecasting stages, and

FIGURE 6.4

ROTATED FACTORS FROM FACTOR ANALYSIS
FOR THE INDEPENDENT VARIABLES

- 2
- b. use only those variables in each factor that are orthogonal to each other and to all other variables in the other factors.

Option b above was selected. Collinearity between one variable and another within the same factor was examined by consulting the correlation matrix. If these variables were collinear, one variable was dropped and the other was retained. In the case of the "Origin Zone Demographic Characteristics" factor, for instance, population, dwelling units, cars and labour force were all found to be collinear. Thus, three of these were dropped and one was retained.

In deciding which independent variable to retain, the following criteria were used:

1. amount of loading of the variable on the factor,
 2. ease of developing and forecasting the variable,
 3. minimum correlation with other variables in the other factors, and
 4. method employed in the development of the variable.
- For instance, if the total number of cars per zone were developed based on zonal population and certain assumptions of average auto ownership, the population variable was retained and the number of cars per zone was excluded since this variable was derived from population.

Not only the collinearity of variables within the same factor was examined, but also collinearity with other variables within other factors. Such collinearity was not significant because it was minimized through the factor rotation technique which attempts to make them orthogonal.

In the case of a factor with two or more variables loading and no collinearity among those variables, all the variables had to be retained and used in the subsequent stepwise regression step.

Thus, factor analysis at the zonal level indicated that among all the original "independent" variables, only seven could be retained as orthogonal variables in the stepwise regression to explain the variation in the dependent variables.

At this point, an important note is in order. In selecting these seven independent variables, it was necessary to keep in mind the dependent variables against which they would be used. For instance, labour force may be a more logical variable to use with work peak hour ratios than population in spite of the collinearity between these two variables. For this reason, and as will be described in the next section, several regressions were performed with different independent variables.

The treatment of missing data in factor analysis posed another important and limiting consideration. There are two options in the factor analysis programme, namely:

- listwise deletion of missing data, and
- pairwise deletion of missing data.

Depending on the option selected, the final factors and even the ability to perform factor analysis can be affected. Some variables such as shopping centres gross leasable area (GLA) and school enrollment could not be entered into factor analysis because only few zones contained such activities. The inclusion of such variables with the listwise deletion option could have reduced the number of observations for factor analysis to only 6 or 7

from over 200. Since these variables had to be excluded from factor analysis, they were also excluded from step-wise regression.

With the above considerations resolved, the factor analysis exercise was repeated for both the zonal level and district level analyses. Tables 6.8 through 6.13 present the final results.

A comparison of the factor analysis results for the zonal and district levels of analysis (Tables 6.9 and 6.12 respectively) would indicate that while the first four factors developed were the same, the fifth factor however was not. It is interesting to note that the order of the variables themselves in each of the first four factors is quite similar for both levels of analysis. The fifth factor for both levels is not clear since the constituent variables seem to be unique in interpretation.

Furthermore, an examination of the list of collinear variables will indicate that collinearity is much more prevalent at the district level than at the zonal level (Tables 6.10 and 6.13). This is to be expected since the averaging effect of district data aggregation caused this increase in multicollinearity between the variables. The ramification of this increase in collinearity was that the regression step at the district level was much more restrictive than at the zonal level.

In summary, the factor analysis step led to the following conclusions or results:

1. It reduced the number of variables to be entered into each regression from over 20 variables to less than 7 variables.

TABLE 6.10
ZONAL LEVEL CORRELATION ANALYSIS

Variable	Collinear With
LABFORPT	NONE
INCOMEZN	NONE
TMTRACBD	TMCARCBD, PARKAVAL, NACREAGE, CARINDEX, TRAINDEX
TMCARCBD	TMTRACBD, PARKAVAL, NACREAGE, NTPOPDEN, NTRESDEN, CARINDEX, TRAINDEX
PARKAVAL	TMTRACBD, TMCARCBD, NACREAGE, NTPOPDEN, NTRESDEN, CARINDEX, TRAINDEX
NACREAGE	TMTRACBD, TMCARCBD, PARKAVAL, NACREAGE, CARINDEX, TRAINDEX
CARINDEX	TMTRACBD, TMCARCBD, PARKAVAL, NACREAGE, TRAINDEX
TRAINDEX	TMTRACBD, TMCARCBD, PARKAVAL, NACREAGE, NTPOPDEN, CARINDEX
POPULATN	DWLGUNIT, TOTLCARS, LABFORCE
DWLGUNIT	POPULATN, TOTLCARS, LABFORCE
TOTLCARS	POPULATN, DWLGUNIT, LABFORCE
LABFORCE	POPULATN, DWLGUNIT, TOTLCARS, LABFORPT
TOTEMPLY	NTEMPDEN, ZONATTRA
NTEMPDEN	TOTEMPLY, ZONATTRA
ZONATTRA	TOTEMPLY, RETEMPY, NTEMPDEN
NTPOPDEN	TMCARCBD, PARKAVAL, NTRESDEN, TRAINDEX
NTRESDEN	TMCARCBD, PARKAVAL, NTPOPDEN
RETEMPY	ZONATTRA

TABLE 6.11
VARIMAX MATRIX FOR DISTRICT LEVEL ANALYSIS

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
POPULATN	0.02045	0.97057	0.01462	0.17377	0.04197
TOTEMPLY	0.26228	0.01545	0.91797	0.10001	0.16419
RETEMPLY	0.14595	0.51552	0.48496	0.08079	0.23760
DWLGUNIT	0.10816	0.91218	0.05308	0.34905	0.04167
TOTLCARS	0.09315	0.91836	0.06775	0.04866	0.08593
LABFORPT	0.23621	0.14668	0.27945	0.12628	0.73607
TMTRACBD	0.94346	0.12093	0.19281	0.13457	0.15568
TMCARCBD	0.91620	0.06337	0.25594	0.24502	0.07833
PARKAVAL	0.72258	0.13612	0.29907	0.45152	0.04886
NACREAGE	0.73132	0.11988	0.07218	0.16076	0.15772
INCOMEZN	0.24308	0.79581	0.16920	0.12350	0.12909
LABFORCE	0.08492	0.96367	0.03285	0.20858	0.10038
NTPOPDEN	0.49471	0.21302	0.05783	0.82424	0.06529
NTRESDEN	0.42271	0.14832	0.12747	0.84350	0.11497
NTEMPDEN	0.18680	0.19109	0.86871	0.08947	0.16797
ZONATTRA	0.27675	0.42626	0.85299	0.08643	0.07140
CARINDEX	0.92347	0.21095	0.10781	0.05748	0.23231
TRAINDEX	0.83700	0.09395	0.30011	0.27834	0.17970

TABLE 6.12

DISTRICT LEVEL FACTOR ANALYSIS

FACTOR	V A R I A B L E		FACTOR LABEL
	LOADING	LABEL	
1	.94	TMTRACBD Time by Transit from CBD	Accessibility
	.91	TMCARCB D Time by Car from CBD	
	.72	PARKAVAL Parking Availability	
	.73	NACREAGE Net Acreage	
	.92	CARINDEX Car Index of Attract.	
	.84	TRAINDEX Transit Index of Attract.	
2	.97	POPULATN Population	Origin Zone Demographic Characteristic
	.91	DWLGUNIT D.U.	
	.91	TOTLCARS Cars	
	.96	LABFORCE Labour Force	
	.79	INCOMEZN Income	
3	.92	TOTEMPLY Total Employment	Destination Zone Demographic Characteristics
	.86	NTEMPDEN Net Empl. Density	
	.85	ZONATTR A Zonal Attractivity	
4	.82	NTPOPDEN Net Population Density	Origin Zonal Density
	.84	NTRESDEN Net Residential Density	
5	.73	LABFORPT L. F. Participation Ratio	Same as Variable Label

TABLE 6.13

DISTRICT LEVEL CORRELATION ANALYSIS

Variable	Collinear With
LABFORPT	NONE
TMTRACBD	TMCARCBD, PARKAVAL, NACREAGE, NTPOPDEN, NTRESDEN, ZONATTRA, CARINDEX, TRAINDEX
TMCARCBD	TOTEMPLY, TMTRACBD, PARKAVAL, NACREAGE, NTPOPDEN, NTRESDEN, ZONATTRA, CARINDEX, TRAINDEX
PARKAVAL	TOTEMPLY, TMTRACBD, TMCARCBD, NACREAGE, NTPOPDEN, NTRESDEN, ZONATTRA, CARINDEX, TRAINDEX
NACREAGE	TMTRACBD, TMCARCBD, PARKAVAL, CARINDEX, TRAINDEX
CARINDEX	NTPOPDEN, NTRESDEN, TMTRACBD, TMCARCBD, PARKAVAL, NACREAGE, TRAINDEX
TRAINDEX	TOTEMPLY, NTRESDEN, TMTRACBD, TMCARCBD, PARKAVAL, NACREAGE, NTPOPDEN, CARINDEX, ZONATTRA
POPULATN	DWLGUNIT, TOTLCARS, LABFORCE, INCOMEZN
DWLGUNIT	POPULATN, TOTLCARS, LABFORCE, INCOMEZN, NTPOPDEN
TOTLCARS	POPULATN, DWLGUNIT, TOTLCARS, INCOMEZN
LABFORCE	POPULATN, DWLGUNIT, TOTLCARS, INCOMEZN
INCOMEZN	POPULATN, RETEMPLY, DWLGUNIT, TOTLCARS, LABFORCE, ZONATTRA
TOTEMPLY	TMCARCBD, NTEMPDEN, ZONATTRA, PARKAVAL, TRAINDEX
NTEMPDEN	TOTEMPLY, ZONATTRA
ZONATTRA	DWLGUNIT, TOTEMPLY, RETEMPLY, NTEMPDEN, TMCARCBD, PARKAVAL, TRAINDEX
NTPOPDEN	DWLGUNIT, TMCARCBD, PARKAVAL, NTRESDEN, TRAINDEX, TMTRACBD, CARINDEX
NTRESDEN	TMTRACBD, TMCARCBD, PARKAVAL, NTPOPDEN, CARINDEX, TRAINDEX

2. It provided an interpretation of each set of variables belonging to a particular factor. This interpretation facilitated the substitution of one variable for another within the same factor.
3. It provided the correlation between the various independent variables. Subsequently, this was used to identify collinear variables which were eventually avoided in the regression steps.
4. It provided the option of using the factor scores instead of the independent variables in the regression step.

6.5 Stepwise Regression Analysis

Because of the wide scope of the research, an important consideration prior to performing regression analysis was to decide on those peak hour ratios by mode and purpose which may not be simulated. The 20 dependent variables shown in Table 6.1 were examined and 10 of these were selected for further detailed analysis. The criteria in the selection of these variables were:

1. the number of available observations,
2. the importance of the particular dependent variable in terms of its contribution, percentagewise, to the total peak hour volume,
3. the limited number of zones containing certain trip purpose activities such as shopping centres and post secondary schools, and
4. cost and time considerations.

Furthermore, as discussed in Subsection 4.3.4, it was recognized that emphasis should be placed on peak hour ratios for work travel by mode. This will become evident in terms of the number of regression analysis attempts for work travel which were performed. Even though regression

analyses were carried out for the other trip purposes, they were not repeated numerous times as in the case of work travel.

6.5.1 Linear Regression

Before presenting the results of the regression analysis performed, it is useful to briefly outline the basic concept in this type of analysis.

In simple bivariate regression, observations are made of the magnitude of a dependent variable Y and similar number of observations for an independent variable X and what is desired is to predict the values of the dependent variable from a linear function of the form:

$$Y' = A + BX \quad \text{..... Eq. (6.1)}$$

where:

Y' = the estimated value of the dependent variable,

A = the intercept of the regression line, and

B = the slope of the regression line

The residual is defined as the difference between the actual values of the dependent variable and the estimated ones. The residual is therefore a measure of the error in prediction and can be represented simply as $Y - Y'$. The basic regression strategy involves a selection of a regression line for which the sum of squared residuals (SS_{res}) is a minimum. This can be written as:

$$\sum (Y - Y')^2 = SS_{res} = \text{Minimum} \quad \text{..... Eq. (6.2)}$$

The above basic concept may be extended to situations involving two or more independent variables. For more elaborate treatment of the stepwise regression analysis technique, the interested reader may consult any advanced text in statistics (41).

In developing the regression equations, a simple process involving the following three steps were followed:

- Step 1: The relationship between the dependent variables and each of the independent variables was examined in order to detect any non-linearities. This was accomplished using scattergram analysis. Where non-linearities were suspected, the relationships were linearized by transforming the independent variables, the dependent variables or both.
- Step 2: The Pearson's correlation matrix was examined to detect collinearity between the independent variables as well as the association between the independent variables and the dependent variables. As mentioned earlier, the choice of the independent variables resulting from the factor analysis stage was highly constrained by the amount of collinearity in the data. The correlation coefficient matrices for the variables used in the zonal and district level analyses are shown in Tables 6.14 and 6.15 respectively.

TABLE 6.14

CORRELATION COEFFICIENTS OF VARIABLES USED IN ZONAL LEVEL REGRESSION

CARINDEX	POPULATN	TOTEMPLY	NTRESDEN	RETEMPLY	LABFORPT	INCOMEZN
1	1	1	1	1	1	1
POPULATN	1					
TOTEMPLY	-.13	1				
NTRESDEN	.34	-.19	1			
RETEMPLY	.14	-.25	.34	1		
LABFORPT	.22	1	.16	-.11	1	
INCOMEZN	.06		.25	.35	-.05	1
	.11		.25	-.05	1	
	-.27		.25	-.05	1	
	-.06		.25	-.05	1	
	-.15		.25	-.05	1	
	-.01		.25	-.05	1	
	1		.25	-.05	1	

TABLE 6.15

CORRELATION COEFFICIENTS OF VARIABLES
USED IN DISTRICT LEVEL REGRESSION

	<u>CARINDEX</u>	<u>POPULATN</u>	<u>TOTEMPLY</u>	<u>NTRESDEN</u>	<u>LABFORPT</u>
CARINDEX	1	.22	.40	.51	.45
POPULATN		1	.04	.28	.14
TOTEMPLY			1	.33	.43
NTRESDEN				1	.34
LABFORPT					1

Step 3: The parameters of each regression equation were developed and the following tests were performed:

- a. the magnitude of the coefficient of determination R ,
- b. sign and magnitude of the partial regression coefficients,
- c. the significance of the partial regression coefficients,
- d. the magnitude of the constant,
- e. the value of the standard error of estimate, and
- f. the significance of the overall regression equation.

Some results of the stepwise linear regression analysis at the zonal and district levels are presented in Tables 6.16 and 6.17 respectively. These will be discussed in Chapter VII.

6.5.2 Non-Linear Regression

As discussed in the previous subsection, some occasions may arise for which linear formulations are inadequate. It may be that:

1. the relationship based on a predefined notion or theory is expected to take a non-linear form or
2. the relationship is simply unknown and the examination of the scattergrams suggests non-linearity or
3. the combined effects of the independent variables are not additive.

One way to handle these types of non-linear situations is to transform the original variables in such a way that the resultant relationships among the transformed variables become linear. This was attempted in a number of situations in this research.

TABLE 6.16

SOME RESULTS OF ZONAL LEVEL
LINEAR REGRESSIONS

<u>Equation Number</u>	<u>Dependent* Variable</u>	<u>Equation</u>
1	HJTRANHW	$-0.088 (\text{NTEMPDEN}) + 0.006 (\text{POPULATN}) + 0.148 (\text{TMTRACBD}) + 338.0$
2	HJTRANHW	$-0.05 (\text{TRAINDEX}) - 0.01 (\text{ZONATTRA}) + 0.04 (\text{LABFORCE}) + 561.7$
3	HITRANHW	$0.05 (\text{TOTEMPLY}) - 0.026 (\text{POPULATN}) + 339.9$
4	HJTRHSPB	$-0.45 (\text{RETEMPY}) - 0.0068 (\text{POPULATN}) - 0.02 (\text{TRAINDEX}) + 461.9$
5	HITRHSBP	$-0.077 (\text{POPULATN}) - 0.062 (\text{TOTEMPLY}) + 0.93 (\text{NTRESDEN}) + 504.2$
6	HJDPHWRK	$0.034 (\text{POPULATN}) - 0.048 (\text{TOTEMPLY}) + 0.349 (\text{LABFORPT}) + 1.75$ $(\text{DIFFTIME}) - 6.33$
7	HJDPHWRK	$-0.08 (\text{TOTEMPLY}) + 0.02 (\text{POPULATN}) - 0.4 (\text{NTRESDEN}) + 0.3 (\text{LABFORPT})$ $+ 191.9$
8	HJDPHSPB	$0.008 (\text{POPULATN}) - 0.25 (\text{NTRESDEN}) + 0.12 (\text{NTEMPDEN}) - 0.01$ $(\text{ZONATTRA}) + 228.12$

* All values of the dependent variable are in thousands and should be divided by ten (10) to represent the peak hour ratios in percentages.

TABLE 6.16 (Continued)

SOME RESULTS OF ZONAL LEVEL
LINEAR REGRESSIONS

Equation Number	Multiple R	Standard Error	F- Regression	Degrees Of Freedom-Residual	F-Table ($\alpha=.05$) ($\alpha=.01$)	Significance At 95% At 99%
1	0.50	11.09	11.46	103	2.70 4.00	Yes Yes
2	0.47	11.23	10.20	106	2.70 4.00	Yes Yes
3	0.71	12.64	28.89	56	3.15 4.98	Yes Yes
4	0.23	13.02	0.69	36	2.84 4.39	No No
5	0.63	13.66	4.07	19	3.14 5.01	Yes No
6	0.66	13.53	36.40	192	2.37 3.32	Yes Yes
7	0.59	11.22	18.90	140	2.45 3.48	Yes Yes
8	0.44	9.90	6.35	105	2.47 3.55	Yes Yes

TABLE 6.16 (Continued)

SOME RESULTS OF ZONAL LEVEL
LINEAR REGRESSIONS

<u>Equation</u> <u>Number</u>	<u>Dependent*</u> <u>Variable</u>	<u>Equation</u>
9	HIDPHSPB	$-0.044 (\text{TOTLCARS}) + 0.02 (\text{INCOMEZN}) + 0.04 (\text{TOTEMPLY}) - 0.12 (\text{RETEMPLY}) + 2.5$
10	HJD&PNHB	$-0.22 (\text{CARINDEX}) - 0.89 (\text{POPULATN}) - 0.11 (\text{TOTEMPLY}) + 401.87$
11	HID&PNHB	$-0.015 (\text{POPULATN}) - 0.09 (\text{RETEMPLY}) - 0.013 (\text{CARINDEX}) + 320.1$

* All values of the dependent variable are in thousands and should be divided by ten (10) to represent the peak hour ratios in percentages.

TABLE 6.16 (Continued)

SOME RESULTS OF ZONAL LEVEL
LINEAR REGRESSIONS

Equation Number	Multiple R	Standard Error	F- Regression	Degrees Of Freedom-Residual	F-Table ($\alpha=0.05$) ($\alpha=0.01$)	Significance At 95% At 99%
9	0.49	10.62	8.46	105	2.45 3.52	Yes Yes
10	0.38	8.56	7.18	127	2.68 3.90	Yes Yes
11	0.43	7.75	8.98	116	2.68 3.98	Yes Yes

TABLE 6.17

SOME RESULTS OF DISTRICT LEVEL
LINEAR REGRESSIONS

<u>Equation Number</u>	<u>Dependent Variable</u>	<u>Equation</u>
1	HITRANHW	-0.007 (POPULATN) + 0.007 (TOTEMPLY) + 243.80
2	HJTRHSPB	-0.066 (TRAINDEX) + 431.2
3	HITRHSPB	0.088 (TRAINDEX) - 0.38 (NTRESDEN) -0.021 (RETEMPY) -180.14
4	HJDPHWRK	0.007 (POPULATN) -0.003 (ZONATTRA) + 178.20
5	HIDPHWRK	0.003 (ZONATTRA) -0.014 (LABFORCE) +0.46 (NTRESDEN) + 148.80
6	HJDPHSPB	-0.0034 (TOTEMPLY) -0.0069 (TOTLCARS) + 147.60
7	HIDPHSPB	-0.160 (NTRESDEN) -0.61 (TMTRACBD) -22.25 (PARKAVAL) -0.042 (NTEMPDEN) + 0.001 (ZONATTRA) -0.001 (POPULATN) + 327.60
8	HJD&PNHB	-0.023 (CARINDEX) -0.0019 (POPULATN) + 374.10
9	HID&PNHB	-0.023 (POPULATN) + 133.40

* All values of the dependent variable are in thousands and should be divided by ten (10) to represent the peak hour ratios in percentages.

TABLE 6.17 (Continued)

SOME RESULTS OF DISTRICT LEVEL
LINEAR REGRESSIONS

Equation Number	Multiple R	Standard Error	F- Regression	Degrees of Freedom-Residual	F-Table ($\alpha=.05$) ($\alpha=.01$)	Significance At 95% At 99%
1	0.720	8.46	12.90	24	3.40 5.61	Yes Yes
2	0.420	9.16	4.37	20	4.35 8.10	Yes No
3	0.750	4.35	3.44	8	4.07 7.59	No No
4	0.660	8.35	11.90	31	3.23 5.18	Yes Yes
5	0.690	7.33	12.60	41	2.84 3.83	Yes Yes
6	0.380	7.21	3.52	41	3.23 5.18	Yes No
7	0.523	4.60	1.75	28	2.45 3.53	No No
8	0.500	8.17	7.00	43	3.23 5.15	Yes Yes
9	0.420	4.79	8.53	41	4.08 7.31	Yes Yes

The scatterplots described in subsection 6.2.2 indicated wide distribution of observations and raised numerous possibilities for non-linear formulations. In few instances, the non-linear formulation results were considerably better (in terms of higher R value) than the linear regression results. Table 6.18 gives a few examples of non-linear regression equations developed together with their associated statistical tests. Again both district and zonal level non-linear regressions were attempted. Further results are summarized in Appendix G.

6.6 Category Analysis

The basic premise of this type of analysis is that individuals or households with different attributes would exhibit different peak hour travel making characteristics. Category analysis is simply a joint frequency distribution of data or observations according to two or more classificatory variables. The terms: Index Method, Cross Tabulation, and Classification Analysis are synonymous with the term Category Analysis.

The display of the distribution of observations by their position relative to two or more variables is the chief component of category analysis. These joint frequency distributions can be statistically analysed by certain tests of significance (such as the chi-square statistic) to determine whether or not the variables are statistically independent.

The value of the chi-square statistic obtained from a cross-tabulation may be compared to the chi-square value obtained from statistical tables. This later value depends on the degrees of freedom (the number of rows and columns in the table). If the chi-square value from the

TABLE 6.18
SOME RESULTS OF ZONAL AND DISTRICT
LEVEL NON-LINEAR REGRESSIONS

Equation Number	Dependent* Variable	Equation
1	HJTRANHW	$1388 / (\text{ZONATTRA} / 1000) + 2.0 (\text{LABFORCE}) + 0.03 (\text{TRAINDEX}) - 84.1$
2	HIDPHWRK	$1.64 (\text{TOTEMPLY}) + 181.0 / \text{POPULATN} - 0.28 \text{NTRESDEN} + 35.2$
3†	HIDPHWRK	$5.3 (\text{TOTEMPLY}) + 76.2$

† This equation is for zonal level analysis.

* All values of the dependent variable are in thousands and should be divided by ten (10) to represent the peak hour ratios in percentages.

TABLE 6.18 (Continued)
 SOME RESULTS OF ZONAL AND DISTRICT
 LEVEL NON-LINEAR REGRESSIONS

Equation Number	Multiple R	Standard Error	F- Regression	Degrees Of Freedom-Residual	F-Table ($\alpha=.05$) ($\alpha=.01$)	Significance At 95% At 99%
1	0.64	7.98	7.2	32	2.92 4.51	Yes Yes
2	0.74	6.83	16.6	41	2.84 4.31	Yes Yes
3	0.57	12.58	74.7	154	3.84 6.63	Yes Yes

cross-tabulation is greater than the chi-square from the statistical tables for a given level of significance, then it is assumed that a relationship does exist at that level.

Backgrounds for the development of peak hour ratios at both the household and individual level as well as the independent variables (attributes) were given in subsection 5.4.3. In the following two subsections a brief description of the category analysis performed is presented.

6.6.1 Household Level Category Analysis

The frequency distribution of the household peak hour ratios developed in this research, indicated wide variations by mode and purpose. Scattergrams of the inter-relationships between these peak hour ratios and certain household attributes were plotted in an attempt to detect any meaningful pattern. These attempts however were not successful.

The household peak hour ratios were recoded into some 13 categories and these were cross-tabulated against most of the household attributes available on the survey records*. The following household characteristics were tested:

- . Household Car Ownership
- . Household Occupancy (persons per household)
- . Household Type
- . Household Income
- . Number of People Employed per Household

* Appendix B.3

Figures 6.5, 6.6 and 6.7 present typical results of the above stratifications. The peak hour ratio codes used at both household and individual level analyses were as follows:

<u>Code</u>	<u>Range of Peak Hour Ratios</u>
1	Special Code*
2	.2 - .9%
3	1.0 - 9.9%
4	10.0 - 19.9%
5	20.0 - 29.9%
6	30.0 - 39.9%
7	40.0 - 49.9%
8	50.0 - 59.9%
9	60.0 - 69.9%
10	70.0 - 79.9%
11	80.0 - 89.9%
12	90.0 - 99.9%
13	100%**

Even though few comments and observations can be made for interpreting the results shown in these figures, it is evident that the differences among the various distributions are not statistically significant.

Figure 6.5 indicates that the household peak hour ratios tend to have a wider variation for households owning no cars than for households owning one, two, or three or more cars respectively. The mean peak hour ratio tends to decrease as the household's vehicle ownership increases. This simply reflects the fact that more

* This special code was assigned to survey record where a trip was reported in the 24 hour period but not in the p.m. peak hour.

** This code indicates that only p.m. peak hour trips were reported by a given household surveyed.

FIGURE 6.5
HOUSEHOLD PEAK HOUR RATIOS DISTRIBUTION
VS. VEHICLE OWNERSHIP

FIGURE 6.6

HOUSEHOLD PEAK HOUR RATIOS DISTRIBUTION
VS. HOUSEHOLD OCCUPANCY

Household Peak Hour Ratios

1.0% 9.9%

10.0% 19.9%

20.0% 59.9%

FIGURE 6.7
HOUSEHOLD PEAK HOUR RATIOS DISTRIBUTION
VS. HOUSEHOLD INCOME

non-peak trips are made by households owning more cars than those made by households owning no cars.

Figure 6.6 shows that as the household occupancy increases the peak hour ratios distribution tends to be wider and the values tend to decrease. A household with two or three persons tends to have a higher mean peak hour ratio than a household with five or six persons. This reflects the fact that tied peak hour trips, such as work trips, represent a significant portion of the overall daily travel of a household with few people. For a household with many people these tied peak hour trips represent lower percentages.

Figure 6.7 relates the peak hour ratio distribution to household income. The figure indicates that households with lower income have a higher mean peak hour ratio than households of higher income. This observation tends to substantiate the comments made on the previous two figures.

6.6.2 Individual Level Category Analysis

As indicated in subsection 5.4.3, only total individual peak hour ratios were developed since disaggregation by mode and purpose would not be meaningful at this micro level of analysis.

Stratification analyses of individual peak hour ratios similar to those at the household level were performed. Again, the frequency distribution of the individual peak hour ratios indicated wide variations. Attempts to explain these variations in terms of individual trip maker attributes were not successful.

The total individual peak hour ratios were recoded into 13 categories and these were cross tabulated against most of the individual attributes available on the survey records. The following individual characteristics were tested:

- . Sex
- . Age
- . Driver Licence Availability
- . Occupation
- . Income

Plots similar to those given for the household level analysis were developed. Figure 6.8 illustrates the frequency distribution of the individual level peak hour ratios compared to those of the household level.

As expected, the household level peak hour ratios tend to be more concentrated, more pronounced and show less variation than those at the individual level. An extrapolation of this observation from the household level to the aggregate zonal level should suggest even less variation. Indeed, transportation planners have proceeded one step further and selected a single value (area-wide average) in order to convert 24 hour travel to peak hour travel.

In conclusion, this thesis has defined peak hour ratios at various scales of analysis ranging from the macro scale area-wide level to the micro scale individual level.

6.7 Analysis of the Peak Hour Ratios from Ground Counts

Prior to performing the zonal, district and disaggregate level analysis, it was decided to examine the

FIGURE 6.8
FREQUENCY DISTRIBUTION OF TOTAL INDIVIDUAL
AND HOUSEHOLD PEAK HOUR RATIOS

1976 ground counts data to see if further insight into peak hour ratios can be gained using ground counts only. Figure 6.9 shows the location of some 186 stations for which classification count data were collected. Table 6.19 shows typical information developed for each station.

A small data bank containing summarized data from the 1976 ground counts programme was built. Table 6.20 shows the data which was stored in the 21 fields of the data bank. Figure 6.10 shows the 5 geographic areas used in the analysis. The class of the facility on which the station was located was divided into freeways, arterials, collectors and locals.

For each station entered into the data bank, the peak hour was identified and the two 15 minute intervals before and after the peak hour together with the four 15 minutes intervals representing the peak hour were developed. The peak hour identified was based on the four highest consecutive 15 minutes interval counts of total person trips crossing that station. The total p.m. peak hour two directional station volume was divided by the total two directional 24 hour station volume and the peak hour ratio was thus developed by station.

An interesting question was whether the peak hour throughout the study area was the same or whether it varied from one geographic area to another. Figure 6.11 shows the time of occurrence of the highest 15 minute volume on screenlines by distance from the CBD. A similar plot was also observed for the full 60 minute volume. It is apparent from these plots that the p.m. peak hour occurs much later in the suburbs than in the central city. This raises support for the controversial issue concerning the validity of comparing instantaneous assignments produced by simulation models with actual ground counts.

FIGURE 6.9

LOCATION OF STATIONS FOR GROUND
CLASSIFICATION COUNTS

TABLE 6.20
SCREENLINE PEAK HOUR RATIOS DATA BANK

<u>Field No.</u>	<u>Data Item</u>
1	Station Number
2	Facility Classification
3	Area Classification
4	Distance from CBD (Bank and Laurier)
5	Vehicle PM. Pk. HR. Inbound
6	Vehicle PM. PK. HR. Outbound
7	Vehicle 12 HR. Inbound
8	Vehicle 12 HR. Outbound
9	12-24 Hour Conversion Factor Inbound
10	12-24 Hour Conversion Factor Outbound
11	Occupancy PM. PK. HR. Inbound
12	Occupancy 12 HR. Inbound
13	Occupancy PM. PK. HR. Outbound
14	Occupancy 12 HR. Outbound
15	Street Direction One-Way/Two-Way
16	PM. PK. HR. Total Vehicles
17	Directional Split % Inbound Peak Hour
18	Directional Split % Outbound Peak Hour
19	24 Hour Inbound Vehicles
20	24 Hour Outbound Vehicles
21	% Peak Hour Ratio

FIGURE 6.10
GEOGRAPHIC AREAS USED IN
GROUND COUNTS ANALYSES

FIGURE 6.11

TIME OF OCCURRENCE OF HIGHEST 15 MINUTES
VOLUME ON SCREENLINES BY DISTANCE FROM CBD

Another important consideration was the question of the variation in the peak hour ratio itself. The frequency distribution of the peak hour ratio for vehicle trips is shown in Figure 6.12. As expected, the distribution shows considerable variation in the values of the peak hour ratio ranging from a minimum of 6.5% to a maximum of 16.5% with a mean of 10.12% and a variance of 2.89. This gives support to the argument that the application of an area wide peak hour ratio in the urban transportation planning system model will undoubtedly result with underdesigning and/or overdesigning a particular transportation facility. As a matter of fact, the exhibit shows that in about 24 to 35% of the time the peak hour ratio is outside the range of 8 to 12% suggested by some transportation handbooks reviewed earlier (27,29,31). The frequency distribution of the peak hour ratios for person trips is shown in Figure 6.13. Because of the influence of transit trips, the distribution shows even more variation in the values of the peak hour ratios than those shown in Figure 6.12, raising further support for the above argument.

Considerable stratification analyses and regression analyses were undertaken to study the cause of the variation in the peak hour ratios. Some fields in the data bank were treated as independent variables and were used to attempt to explain the variations in the peak hour ratios by means of stratification analysis and regression analysis. These attempts were unsuccessful except the two described below.

FIGURE 6.12

VARIATION IN VEHICLE BASED PEAK HOUR RATIOS FROM 1976 GROUND COUNTS OF 182 STATIONS

MEAN PK. HR. RATIO = 13.18
 STANDARD DEVIATION = 4.9
 VARIANCE = 24.2

FIGURE 6.13

VARIATION IN PERSON BASED PEAK HOUR RATIOS
 FROM 1976 GROUND COUNTS OF 169 STATIONS

It was postulated that a relationship exists between facility volume or type and the peak hour ratios. The category analysis technique was employed to plot the distributions shown in Figure 6.14. It is evident from these distributions that the higher the facility volume (higher order type of facility), the lower the peak hour ratio. The t-test was performed to examine whether there were any significant differences between the means of the distributions by each facility type. The results of this t-test are shown in Table 6.21 below.

TABLE 6.21
TESTING VARIATION OF PEAK HOUR RATIOS
DISTRIBUTION BY FACILITY CLASSIFICATION

<u>Hypothesis Tested</u>	<u>t-Statistic</u>	<u>Significance Level</u>
Freeways vs Arterial and Collectors	2.49	Significant at 95%
Freeway vs Locals	2.54	Significant at 95%
Arterial and Collectors vs Locals	0.47	Not Significant

Table 6.22 shows the results of the regression of peak hour ratios versus distance from the CBD for the four different facility types. It should be noted that the regression equations developed can only be applied in the Ottawa-Carleton area since the coefficients of the equations reflect the distribution of the stations in the study area. Although similar equations may be developed in other cities, the value of the coefficients will undoubtedly be different. It is evident from these regression equations that the peak hour ratio seems to increase as the hierarchy of facility type decreases (from freeways to collectors). They further suggest that peak hour ratios

FIGURE 6.14
VARIATION IN PEAK HOUR RATIOS BY FACILITY TYPE

TABLE 6.22
RESULTS OF REGRESSION ANALYSIS FOR GROUND COUNTS DATA

Facility Type	Regression Y=Peak Hour Ratio X=Distance From CBD	Degrees of Freedom	T Value	Correlation Co-Efficient		Standard Error
				Correlation	% Error	
Freeway 1	Y=1.45X	14	5.2355*	0.8136	60.38	56.87
Arterial 2	Y=1.58X	107	11.5496*	0.7449	67.89	69.13
Collectors 3	Y=2.20X	17	5.7588*	0.8131	61.69	61.76
Local 4	Y=5.07X	43	7.6873*	0.7608	66.68	69.34

* All the above equations were found statistically significant at 95%.

would increase with increasing distance from the CBD. Table 6.22 also provides the results of statistical tests of these regression equations.

The analysis of ground counts data was helpful since it demonstrated that indeed variations in the peak hour ratios do occur and that simple station location characteristics were insufficient to explain these variations.

CHAPTER VII

DISCUSSION OF PEAK HOUR RATIOS ESTIMATION EQUATIONS

7.1 Introduction

In Chapter IV, several formulations for peak hour ratios estimation and application were developed. These differed in terms of the analysis level (district, zonal or disaggregate) and the analysis technique (linear regression, non-linear regression and category analysis). The statistical analyses performed were briefly outlined in the previous chapter but the relative strengths or weaknesses of the derived relationships or their application were not discussed. This chapter is devoted to that end.

It was hypothesized in Section 4.2 that the following six sets cover the spectrum of peak hour ratio determinants:

- (a) Trip Characteristics.
- (b) Origin-Destination Zone Characteristics.
- (c) Transport System Variables.
- (d) Regional or Area-Wide Variables.
- (e) Policy Constraint Variables, and
- (f) Disaggregate Socio-Economic Variables.

Furthermore, attempts at estimating peak hour ratios using sets (a), (b), (c) and (f) were made as described in the previous chapter. Sets (d) and (e) were not explicitly treated in the research in terms of quantitative analysis, however, they will be further discussed in Section 7.5.

7.2 Data Limitations

Perhaps the most important factor influencing any research is the quality of input data. In this research, data presented numerous problems and limitations which made the possibility of obtaining good simulation rather doubtful. It is emphasized that for a research of this nature which used actual data instead of synthesized data, the likelihood of deriving strong mathematical formulations is rather low. This was anticipated at the inception of the project. However, data limitations should not be a deterrent to formulating or utilizing the recommended approach(es) to peak hour ratio estimation and application.

In order to avoid a prolonged discussion of data limitations, the following is a brief listing of the factors and circumstances that influenced the quality of input data:

1. data were obtained from different sources and were developed by different people,
2. data were collected for different years, different time periods, and different zone systems, and had to be converted to the same common base with considerable judgment,
3. the relatively small survey sample which constituted the main input to the research, particularly in deriving the dependent variables,
4. the relatively old age of data (5 to 6 year old information was used in the analysis),
5. the use of different computer programmes and packages to interface the data such as the interface between the AVJIO files of the Ministry of Transportation and Communications of Ontario's computer package and the SPSS package, (in this regard, the treatment of missing data posed a particular difficulty), and

6. at both the zonal and district level analysis, the relative importance of zones in terms of number of observations was not considered. For each zone the peak hour ratio was developed by dividing peak hour trips by the total 24 hour trips and a single value represented that zone regardless of the number of observations applicable to that zone. Since the sample is not homogeneous by zone this could have significantly affected the results of the analysis.

Although substantial effort was devoted to using reasonable input data, the above limitations strongly influenced the outcome of this research. Furthermore, time and cost limitations were also significant. The input data required substantial manipulations which were frequently costly. In addition, a research of this nature and this scope would have taken the author at least twice the time had it been conducted with the assistance of staff working with the author for the Regional Municipality of Ottawa-Carleton.

7.3 Testing of Peak Hour Ratios Estimation Equations

While it was intended that the equations developed in this research using the 1972 Home Interview Survey would be tested with respect to the 1977/1978 conditions, this could not be done due to both time and cost limitations mentioned earlier. This test is however, recommended for further research and investigation.

In order to ascertain the adequacy of some of the developed equations, scattergrams were prepared showing the observed vs the simulated peak hour ratios. Figure 7.1 shows a typical scattergram for home-work peak hour

FIGURE 7.1

SIMULATED VS. OBSERVED ORIGIN ZONAL PEAK HOUR RATIOS
HOME-WORK, DRIVER PLUS PASSENGER (HIDPHWRK)

ratios by driver and passenger, while Figure 7.2 shows a similar scattergram of simulated versus observed peak hour ratios for home-work by transit.

Another interesting test of the developed equations would be to apply the recommended peak hour ratios to the 24 hour travel matrix in order to establish the peak hour travel matrix and then assign this matrix to the appropriate transportation network. Subsequently, assignment results for a number of links using the existing area-wide approach may be compared to those using the recommended approach. These, in turn, may be compared to the actual ground counts. Figure 7.3 illustrates the above test in the form of a flow chart. As the figure indicates, this test in fact provides a measure of the reliability of the recommended approach versus the existing area-wide approach. This measure would then be used as input to the decision theory application described in Chapter VIII. Specifically, more objective conditional probabilities would be obtained than those subjectively derived in Section 8.5.3.

7.4 The Recommended Approach(es) for Peak Hour Ratios Estimation and Application

It is interesting to note that, when the research commenced, it was hoped that zonal interchange peak hour ratios could be developed to convert each cell of the 24 hour travel matrix into peak hour travel. Unfortunately, the zonal interchange approach was abandoned due to data limitations. As will be described in Section 8.5.1 however, such an approach, though costly in terms of data needs, could be cost-effective for certain conditions.

FIGURE 7.2

SIMULATED VS. OBSERVED ORIGIN ZONAL PEAK HOUR RATIOS

HOME-WORK, TRANSIT (HITRANHW)

FIGURE 7.3
TESTING THE RECOMMENDED PEAK HOUR RATIOS

As the research progressed, it became evident that the recommended approach to peak hour ratios estimation is highly a function of the way the peak hour ratios are applied and the stage of the urban transportation planning system model at which they are applied. In the discussion which follows, several approaches are recommended for peak hour ratios estimation and application depending on the particular situation under consideration.

Peak hour ratios for trips, by mode, and/or by purpose and mode may be developed for application at the trip generation stage. This may be done by simple stratification of peak and 24 hour trip records of a given home interview survey. These peak hour ratios can be applied at the production zone, attraction zone, or both in order to convert the 24 hour trip generation into peak hour. In this approach, the trip distribution is made for peak hour conditions only. The approach differs from a direct simulation of peak hour trip generation equation in that, first 24 hour trip generation relationships are established, then peak hour ratios are applied to the resultant total productions and attractions. This approach therefore, tends to limit the availability and the use of twenty four hour travel data.

Peak hour ratios may be applied at the trip distribution step to convert 24 hour travel matrices into peak hour. If the distribution step is made for person trips, average peak hour ratios for non-work purposes may be applied. Work trips on the other hand, may be converted by applying peak hour ratios developed through regression analysis as described in Section 6.5.

3

When the trip distribution step produces 24 hour travel matrices by mode and purpose as described in Section 8.2, the recommended peak hour ratios to be applied are shown in Table 7.1. These peak hour ratios may be applied to convert non-work 24 hour travel matrices into peak hour. These peak hour ratios are applied at both the production and the attraction ends, and a balancing procedure would be required. This balancing may be done using any of the available programmes which incorporate a balancing step such as the programme associated with the Fratar technique. Alternatively, peak hour ratios at the origin and destination may be added and divided by two to obtain an average value which can be applied at both the production and attraction ends. In any case, directional split factors would have to be applied to these matrices in order to convert them into true origin-destination peak hour matrices.

The use of mean peak hour ratios by mode and purpose for non-work travel is dictated by the fact that no appropriate and sufficiently strong simulation of peak hour ratios was possible in this research. Tables 6.17, 6.18, and 6.19 summarized the derived relationships. For work travel however, better simulation results were obtained for both transit and auto driver and passenger. Table 7.2 shows the recommended equation for origin zone peak hour ratios by transit for home-work travel, while Table 7.3 gives the recommended equation for origin zone peak hour ratios for home-work travel by driver and passenger.

TABLE 7.1
RECOMMENDED PEAK HOUR RATIOS BY MODE
AND PURPOSE FOR NON-WORK TRAVEL

Variable	Variable Label	Z O N A L			D I S T R I C T		
		Mean	Std.Dev.	Valid Obs.	Mean	Std.Dev.	Valid Obs.
HJDPHSRO	HJ D + P HSRO	6.5%	7.5	257	5.9%	5.0	55
HIDPHSRO	HI D + P HSRO	8.7%	11.5	257	7.2%	7.2	56
HJDPHSCH	HJ D + P HSCH	4.3%	10.1	142	5.3%	9.7	44
HIDPHSCH	HI D + P HSCH	3.5%	8.2	143	2.9%	4.2	44
HJTRHSPB	HJ TRANS HSPB	11.1%	16.9	124	11.6%	10.9	32
HITRHSPB	HI TRANS HSPB	8.6%	17.0	120	7.3%	9.2	29
HJDPHSPB	HJ D + P HSPB	8.4%	10.9	241	8.8%	8.1	54
HIDPHSPB	HI D + P HSPB	9.7%	13.1	247	9.2%	7.8	55
HJD&PNHB	HJ D + P NHB	10.1%	10.5	264	10.9%	9.5	56
HID&PNHB	HI D + P NHB	9.7%	10.9	264	10.0%	8.8	56
HJTRANHB	HJ TRANS NHB	9.6%	17.4	149	9.0%	11.4	45
HITRANHB	HI TRANS NHB	9.4%	16.2	154	8.1%	10.3	45
HJTRHSRO	HJ TRANS HSRO	9.3%	17.3	97	10.9%	12.9	35
HITRHSRO	HI TRANS HSRO	8.7%	16.9	102	7.3%	12.8	37
HJTRHSCH	HJ TRANS HSCH	6.6%	11.6	171	7.1%	7.8	47
HITRHSCH	HI TRANS HSCH	5.6%	12.3	174	5.1%	6.4	50

TABLE 7.2

RECOMMENDED PEAK HOUR RATIOS
EQUATION FOR HOME-WORK (TRANSIT)

The Model

$$\text{HITRANHW}^* = 0.09 (\text{TOTEMPLY}) - 0.16 (\text{POPULATN}) - 1.9 (\text{DIFFTIME}) + 180.5$$

Variable Labels

TOTEMPLY = Total Employment, F = 77.5

POPULATN = Population, F = 8.4

DIFFTIME = Difference of (Peak - Off Peak) Travel Time in
in Tenth of Minutes, F = 4.2

HITRANHW = Origin Zonal Peak Hour Ratios

Correlation Matrix

	TOTEMPLY	POPULATN	DIFFTIME
TOTEMPLY	1	-.37	-.34
POPULATN		1	.30
DIFFTIME			1

Statistics

Multiple R = 0.715, Standard Error = 14.66

F_{Regression} = 50.8

F-Table (@ $\alpha=.01$) = 3.8

Equation Significant at 99%

* In thousands, e.g. a value of 234 means that the peak hour ratio is 23.4%.

TABLE 7.3

RECOMMENDED PEAK HOUR RATIOS
EQUATION FOR HOME-WORK (DRIVER AND PASSENGER)

The Model

$$\text{HIDPHWRK}^* = 0.049 (\text{TOTEMPLY}) - 0.025 (\text{POPULATN}) + 242.6$$

Variable Labels

TOTEMPLY = Total Employment, F = 39.8

POPULATN = Population, F = 34.5

HITRANHW = Origin Zonal Peak Hour Ratios

Correlation Matrix

	TOTEMPLY	POPULATION
TOTEMPLY	1	-0.36
POPULATN		1

Statistics

Multiple R = 0.68 Standard Error = 11.14

F_{Regression} = 58.1

F_{Table} (@α=.01) = 4.7

Equation Significant at 99%

* In thousands, e.g. a value of 234 means that the peak hour ratio is 23.4%.

Even though equations with higher multiple correlation coefficients than those given* in the recommended equations for work travel were obtained, the most simple formulation of the relationships was selected. Total employment and population by zone appear in both equations, the first with a positive sign, while the second with a negative sign. This simply reflects the fact that, as the total employment in the origin zones increases, zonal peak hour ratios for work travel during the p.m. peak hour also increase. On the other hand, as the population of the origin zone increases, these zones will tend to be in the suburban areas where few p.m. peak hour work trips will be made and the peak hour ratios for work travel decrease.

The equation for transit shows an additional variable, namely, the difference between travel time from a given zone to the CBD in the peak and off peak conditions. This variable reflects the fact that transit service to the CBD during the peak hour is much faster than during the off-peak hour conditions. The equation suggests that as the difference in travel time variable increases, the peak hour ratios for transit work travel decrease. Stated in a different way, for zones that are further away from the CBD and for which the difference in travel time between the peak and off peak is large, the peak hour ratios for work travel tends to decrease. This is again a reflection of the fact that few p.m. peak hour trips will be generated in these suburban zones.

* Also those in Tables 6.17, 6.18 and 6.19

Finally, for both of the origin zonal equations, the value of the constant is large reflecting the variation which is not explained by the variables in the equation. In fact, the author believes that the value of the constant reflects the variation already accounted for by the disaggregation of peak hour ratios by mode and purpose.

Tables 7.4 and 7.5 present the recommended equations for destination zonal peak hour ratios for home-work, by transit and driver plus passenger (HJTRANHW and HJDPHWRK) respectively. Table 7.4 gives a simple formulation which directly relates HJTRANHW to the travel time by transit from the CBD. The interpretation of the equation stems from the fact that destination zones which are close to the CBD will tend to attract fewer p.m. peak hour work-home trips than those suburban zones which are further away from the CBD and which are typically residential zones. These residential zones tend to have higher work peak hour ratios because of the work to home trips destined to them during the p.m. peak hour.

In Table 7.5, HJDPHWRK are related to zonal population and net residential density. The destination zonal peak hour ratios for work by driver and passenger increase as the zonal population increases. This again is explained in terms of the work to home p.m. peak hour trips, whereby the higher the zonal population, the higher the labour force component returning from work. The zonal net residential density on the other hand, has a negative sign in the equation, which suggests that the peak hour ratios HJDPHWRK increase with decreasing net residential densities. This is in accord with the fact that these densities decrease with increasing distance from the CBD and thus support the explanation advanced above for the effect of variation of the population variable.

TABLE 7.4

RECOMMENDED PEAK HOUR RATIOS
EQUATION FOR HOME - WORK (TRANSIT)

The Model

$$HJTRANHW^* = 4.3 \text{ TMTRACBD} - 483$$

Variable Labels

TMTRACBD = Time by Transit From CBD in Tenth of Minutes.

HJTRANHW = Destination Zonal Peak Hour Ratios

Statistics

Multiple R = 0.93

Standard Error = 7.52

F_{Regression} = 26.0

F_{Table} (@ $\alpha=0.05$) = 7.7

Equation Significant at 95%

* In thousands, e.g. a value of 234 means that the peak hour ratio is 23.4%.

TABLE 7.5

RECOMMENDED PEAK HOUR RATIOS
EQUATION FOR HOME-WORK (DRIVER AND PASSENGER)

The Model

$$HJDPHWRK^* = 0.11 \text{ POPULATN} - 0.29 \text{ NTRESDEN} + 40.4$$

Variable Labels

POPULATN = Population, F = 41.4

NTRESDEN = Net Residential Density, F = 7.0

HJDPHWRK = Destination Zonal Peak Hour Ratios

Correlation Matrix

	POPULATN	NTRESDEN
POPULATN	1	0.25
NTRESDEN		1

Statistics

Multiple R = 0.96

Standard Error = 2.48

F_{Regression} = 21.2

F_{Table} (@α = 0.05) = 9.5

Equation Significant at 95%

* In thousands, e.g. a value of 234 means that the peak hour ratio is 23.4%.

There are also merits for recommending another approach to applying peak hour ratios. This is the approach in which peak hour ratios are estimated using stratification analysis or regression analysis from screenlines and cordon counts as described in Section 6.6. This approach may be referred to as the Post-Assignment approach. The 24 hour travel matrix by mode is assigned to the corresponding transportation network and the resulting 24 hour link volume may be converted into peak hour volume using the equations developed in Table 6.22 or in Figure 6.14. In this approach, the merits of both the statistical equations developed and the judgment of the functional planner and his knowledge of the conditions of the particular link being designed would be utilized in a most effective way.

In summary, the reader may appreciate that several approaches to peak hour ratios estimation and application may be recommended. The task of the transportation planner is to decide, within his overall modelling framework, that approach which seems most suitable. Admittedly, no simple approach can be optimum for any given situation, but what the research has delineated is the fact that peak hour ratios are not constant and that a single area-wide average should not be applied to convert 24 hour travel into peak hour travel. Figure 7.4 summarizes in mathematical terms the alternative approaches recommended for the estimation and application of peak hour ratios at different stages of the UTPS model. The following discussion relates to the equations presented in Figure 7.4.

Equation (7.1) illustrates how peak hour ratios can be applied past the 24 hour trip generation step. The basic concept is to apply different zonal trip-ends peak hour ratios to the 24 hour total productions and

FIGURE 7.4

SUMMARY OF ALTERNATIVE APPROACHES RECOMMENDED
FOR PEAK HOUR RATIOS ESTIMATION AND APPLICATION

attractions ($T_{Pi}^{m,P}$ and $T_{Aj}^{m,P}$ respectively) in order to produce peak hour productions and attractions ($t_{Pi}^{m,P}$ and $t_{Aj}^{m,P}$ respectively).

Equations (7.2) and (7.3) show how peak hour ratios are applied post the 24 hour trip distribution step. This is basically the approach presented in Section 4.3.1 and described by Equations (4.3) and (4.4). Equation (7.3) indicates that it may be advantageous to apply an average zonal peak hour ratio to the 24 hour travel matrix instead of using the two elements of Equation (7.2) and subsequently applying a trip table balancing procedure.

In Equation (7.4), peak hour ratios are applied by link l post the 24 hour trip assignment step. The ratios vary depending on the type of facility f .

Finally, it should be stated that the time horizon for the urban transportation study may dictate, to a large extent, the approach by which peak hour ratios are estimated and applied. The above recommended approaches should be employed for immediate and short range planning applications. For long range planning, the development of elaborate approaches for peak hour ratios estimation or application is neither recommended nor supported. In fact, the development of peak hour models for long range planning are not necessary since for corridor analysis, which is the ultimate objective of long range planning, 24 hour modelling of travel patterns would suffice.

7.5 Other Considerations in Peak Hour Ratio Estimation and Application

As mentioned earlier, two of the six sets of peak hour ratio determinants were not treated explicitly in

this research. These were the policy constraint variables set and the regional area wide variables set. They are discussed briefly in the following paragraphs.

The policy constraint set as illustrated in Table 4.1 includes a number of variables such as:

- . discriminant pricing policies and techniques,
- . restrictive or controlled access during peak hour,
- . shortened work week, and
- . communications in lieu of travel.

These type of variables were treated either implicitly by the introduction of another variable which directly or indirectly relates to them, or they were too difficult to quantify for the purposes of this research. For example, the first three policy constraint variables are somewhat reflected in variables 2, 15, and 29 or set b in Table 4.1. On the other hand, the fourth policy constraint variable (communications in lieu of travel) is simply difficult to quantify or to relate to any of the other zonal or transport variables used in the research. The report by Khan* on this subject prepared for the Canadian Transport Commission may be referred to by the interested reader.

The list of regional or area wide variables is given under item d in Table 4.1. These variables were found to describe the variation in peak hour ratios from one urban area to another through research conducted by Hill et al, (21) as mentioned in Section 3.2. Certain important aspects of this research are briefly discussed below.

*"Transportation and Telecommunications: A Study of Substitution, Stimulation and Their Implications" Canadian Transport Commission, Systems Analysis and Research Data Base Branch, Research Report 121, 1974.

FIGURE 7.5

GEOGRAPHICAL DISTRIBUTION OF CITIES
ANALYZED IN TIME OF DAY STUDY

Notes Adjustment for split workshift phenomenon in Louisville and Stockton; sum of top hour fit the trend. Remainder was assigned to second hour.

Primary Stratification based on 24 Hour Volume/Capacity Ratio at zero modal split

FIGURE 7.6

PERCENTAGE OF TOTAL TRAVEL IN P.M. PEAK PERIOD (2-8 P.M.) HIGHEST HOUR

Source: Reference (2)

The temporal distribution of travel was analyzed using data from 8 different urban centres shown in Figure 7.5. The basic design parameters studied were the peak hour ratios H and the directional split factors D . The research indicated that H values were in the range of 7-12% while the D values were in the range of 60-70% in the peak direction.

One of the most interesting models developed by Hill is that relating the non-directional peak hour ratios to total city population and the ratio of area-wide daily volume to capacity. This model is shown in Figure 7.6. Hill points out that the peak hour ratio generally increases with increasing city population reflecting diversity of travel purposes and patterns in larger areas, while increasing congestion (reflected by high 24 hour volume/capacity ratio) drives the percentage downward apparently forcing some travel to occur in other hours than the drivers might freely have chosen.

CHAPTER VIII

ECONOMIC BASIS FOR PEAK HOUR RATIOS ESTIMATION

8.1 Introduction

The theme of this research has been the establishment of an improved approach for peak hour ratios estimation and application in the UTPP. The premise is that if the peak hour volume is estimated and predicted within a reasonable level of accuracy, the facility would be designed with minimum under or over-investment. This premise may be delineated in the following simplistic form:

This chapter is devoted to exploring the above premise by:

1. outlining the interaction between the peak hour volumes and other elements of the UTPS model, particularly the post trip distribution steps,
2. describing the process of link peak hour volume estimation and design, and indicating the consequences of errors in the design decisions,
3. structuring and applying a decision theoretic framework to identify the trade offs between the cost of developing

the improved approach and the expected net gain from experimentation, and

4. testing the above framework using actual case studies.

In dealing with the application of decision theory, it is not the intention to suggest that decision theory be introduced in the design of each facility. In fact, the author is well aware of the usual skepticism expressed by decision makers even in the use of transportation models for future forecasts, let alone the application of decision theory. The basic intention is to demonstrate the cost effectiveness of the recommended approaches to peak hour ratios estimation and application.

It must be emphasized that design decisions are made based on a number of criteria and considerations, including:

1. the peak hour volume anticipated on the facility at the design year,
2. staging considerations,
3. subsidy levels, budgetary constraints and other fiscal considerations,
4. the individual attributes of the facility such as type of facility, mode of operation and traffic composition, and
5. subjective and qualitative judgment of the facility designer and the decision maker.

It is not within the scope of this research to elaborate on the above criteria and considerations in selecting the design hour volume. It will be assumed that the peak hour volume is the sole determinant of the design hour volume, an assumption which may not necessarily be true. However, in order to simplify the application of the statistical decision theory outlined in this chapter, the

design hour volume is defined in terms of the peak hour volume. It is interesting to note that this is not a departure from the standard practice suggested by various transportation handbooks (27, 30, 31).

As in any other area of prediction of future human activities, the prediction of future peak hour volume involves elements of uncertainty (42). There are a number of sources and reasons for this uncertainty including:

1. unexpected changes in traveller's choice of mode,
2. errors in estimation of future land use activities,
3. errors in data measurement,
4. errors due to simulation and modelling,
5. technological developments, environmental changes and political considerations, and
6. lack of information.

In this chapter, the problem of uncertainty will be treated explicitly to show how it affects the design decisions.

An important aspect of the urban transportation planning process is its continuous nature. The process is iterative and requires feedback information from one step to another. As indicated in Section 2.2, the position of peak hour ratios is central to this process. These ratios, like other elements of the process, interact and interplay both explicitly and implicitly to give the process its coherence. However, the sequential and iterative nature of the process does not lend itself to a final solution. This, in turn does not lead into an optimum link design, but only a design that satisfies a given state of nature of the process. The process is continuously and dynamically changing and the best a planner can do is to ensure that consistent, logical and reliable estimates are

used both for short term planning and for the design process. It is fundamental therefore, to understand the role of the peak hour ratios in the UTPP and in particular, their effect on the estimation of peak hour volume using a particular facility. In the section which follows the link peak hour volume estimation problem is discussed as it relates to the travel assignment in the system planning process and to the design hour volume (DHV) in the functional planning and design processes.

8.2 Link Peak Hour Volume Estimation

The estimation of peak hour volume on a particular link is generally made after the assignment step of the urban transportation planning system model (UTPS). The peak hour ratios, however, are generally applied at the trip distribution and/or modal split steps depending on the type of model used (43). Thus, it is not easily feasible to isolate the effects of peak hour ratios on the estimated link peak hour volume and the eventual design decision. The assignment technique employed may dictate, to a large extent, the link peak hour volume and in turn the design decision. However, this is also true of any step in the UTPS model. This research concentrated on peak hour ratios to show their direct relationship to the eventual design hour volume and the design decision.

It is not easily feasible to apply decision theory to demonstrate the cost effectiveness of developing peak hour travel matrices using the recommended peak hour ratios vs. developing peak hour travel matrices using area-wide ratios*. It was therefore decided to skip the traffic

* Comparison of these two sets of matrices may not produce concrete and explicitly meaningful results in terms of decision theory.

assignment step and to assume that the link peak hour volume developed using the recommended approach(es) will have a higher reliability than that using area-wide peak hour ratio. This assumption may be verified by following the testing procedure outlined in Figure 7.3.

An attempt is made in this section to elaborate on the role and inter-relationships among the main elements of the UTPS and in particular the trip distribution, modal split and traffic assignment components*. The intention is to point out that these elements are interdependent with the link design hour volume and that the latter will vary depending on the assumptions made concerning these elements and in particular, the peak hour ratios. Figure 8.1 will be used to illustrate these inter-relationships past the trip distribution process.

An appropriate formulation of the basic trip distribution sub-model may be given as:

$$T_{ij} = \frac{P_i \cdot A_j \left[M_{ij} \frac{TF_{ij}}{\sum_{x=i}^n A_x \cdot TF_{ix}} + (1-M_{ij}) \frac{HF_{ij}}{\sum_{x=i}^n A_x \cdot HF_{ix}} \right]}{\sum_{x=i}^n A_x \cdot \left[M_{ij} \frac{TF_{ij}}{\sum_{x=i}^n A_x \cdot TF_{ix}} + (1-M_{ij}) \frac{HF_{ij}}{\sum_{x=i}^n A_x \cdot HF_{ix}} \right]}$$

.....Eq. (8.1)

* In fact, in the above discussion, the author describes the actual process being employed in applying the revised peak hour ratios within the TRANS model. This model is being developed under the author's direction in co-operation with a study team from the Regional Municipality of Ottawa-Carleton, The Outaouais Regional Community and the National Capital Commission (23).

FIGURE 8.1.

INTERRELATIONSHIPS BETWEEN THE VARIOUS STEPS
OF THE MODELLING PROCESS PAST TRIP DISTRIBUTION

where:

n = number of traffic zones
 T_{ij} = number of trips from Zone i to Zone j
 P_i = number of trips originating in Zone i
 A_j = number of possible trip destinations in Zone j
 TF = Transit Travel Time Factor
 HF = Road Travel Time Factor
 M = Transit Probability Factor

In this formulation, the 24 hour person trips productions and attractions, preliminary estimate of transit probability and the highway and transit impedances are all used as inputs to the trip distribution procedure. Peak hour impedances are used to distribute the 24 hour work trip productions and attractions while off peak impedances are used to distribute all other trip purposes.

This trip distribution model produces 24 hour trip matrices for person (T_{ij}^P), transit ($T_{ij}^{T,P}$) and for auto driver and passenger ($T_{ij}^{DP,P}$) modes. An appropriate auto occupancy model may be used to convert the auto driver and passenger matrices to vehicular trip matrices $T_{ij}^{V,P}$ by purpose.

At this stage, the recommended peak hour ratios $H_i^{m,P}$ and the directional split factors $D^{m,P}$ may be applied to convert the 24 hour trip matrices by mode and purpose into peak hour matrices. For the work-trip matrices by mode, the recommended peak hour ratios equations (Tables 7.2 to 7.5) may be applied. For all other matrices, mean peak hour ratios by mode and purpose may be applied as given in Table 7.1.

With the application of appropriate directional split factors, the true origin-destination trip matrices by purpose and mode may be summed in order to obtain total peak hour travel matrices by mode t_{ij}^T and t_{ij}^V . These, in turn will be assigned to the appropriate network and the resultant assignment produces the link peak hour volume estimate by mode.

Thus,

$$t_{\ell}^m = \sum_m \sum_p \sum_r t_{ij}^{m,p,r} \quad \dots \text{Eq. (8.2)}$$

where:

t_{ℓ}^m = estimate of peak hour volume on link ℓ by mode m , and

$t_{ij}^{m,p,r}$ = peak hour trips from zone i to zone j for trip purpose p , mode m and all routes r using the link ℓ .

Depending on the method of assignment used and/or the judgment and experience of the transportation planner, the assigned link peak hour volume can be transformed into the appropriate design hour volume (DHV) used by the design engineer to decide on the type of facility to be provided, i.e.:

$$(\text{DHV})_{\ell} = t_{\ell}^m = \left(\sum_m \sum_p \sum_r t_{ij}^{m,p,r} \cdot f \right) \quad \dots \text{Eq. (8.3)}$$

where f = subjective adjustment factor to convert assigned peak hour volume into design hour volume.

While a lengthy description of the various traffic assignment techniques is not within the scope of this research, it should be noted that, regardless of the reliability and accuracy in the estimation of peak hour ratios, inappropriate assignment algorithms may result with significant over-estimate and/or under-estimate of the link peak hour volume. For this reason, there is an urgent need to develop efficient algorithms which produce reasonable flows in the assignment of traffic to a transport network (44).

In the past two decades, progress has been made in developing various assignment algorithms such as the all-or-nothing technique, multi-path technique, capacity restraint methods and recently, approaches based on network equilibrium theory. In fact, program UROAD of the UTPS computer package developed by UMTA (Urban Mass Transit Administration) contains most of the above capabilities. The equilibrium approach is probably the most desirable approach to traffic assignment (45,46). In fact, an interesting extension of this research would be to use the equilibrium approach in evaluating traffic flow cost impacts for each state of nature in the face of alternative decisions regarding peak hour flow capacity.

Trips assigned using the Wardrop's first principal* are referred to as equilibrium flows. Various authors such as Dafermos (45), Leblanc (46) and Nguyen (47) have developed algorithms that assign a fixed origin-destination trip matrix to a road network and result in equilibrium flows. However

* Wardrop's First Principal regarding the division of traffic flow among alternate routes states that for steady state traffic conditions, the journey times are equal on all the routes actually used for travel between a specified origin and a specified destination. The journey times on the routes used are less than the journey times that would be experienced by a vehicle which travels any other route between the specified origin and the specified destination.

most of their efforts have been theoretical and abstract. Recently, Florian and Nguyen (48) succeeded in validating the results of their equilibrium assignment method using the City of Winnipeg data. Further work by Ruiter (49) and by Tobin (50) using these equilibrium techniques have advanced the theory. It is likely that combined mode choice/ equilibrium traffic assignment models will be available within a few years. Such models will be extremely valuable and may greatly enhance the design-decision making process.

8.3 Economics of Error in Link Volume Estimation

Design policies and criteria for most transportation agencies are based on several standard references. Perhaps the most important of these are the Highway Capacity Manual (30) and the AASHO Policy Publications (31). Guidance is generally accepted from these references; however, many agencies develop their own specific design policies. These policies vary depending on a number of considerations including those indicated in Section 8.1. For example, one agency may decide, based on its own specific conditions, that a level of service C will be used in the design of urban freeways while another agency may consider the design for a level of service D acceptable. The design requirements may therefore differ. This point should be kept in mind throughout the following discussion of the economics of error in link volume estimation.

It should be immediately apparent that when the error in link volume estimation has a low sensitivity to changing the design decision, (any value between a and b in Figure 8.2), elaborate research on peak hour ratios may not be justified. On the other hand, when a high level of sensitivity is evident, this should suggest that more reliable estimates and additional research are required. Skillful judgment should be employed in deciding whether further information is necessary. This process however, will never be completely systematized to the point that the designer need but to "plug-in" the input requirements and automatically produce the design output.

The UTPS generally produces demand estimates in terms of average annual daily traffic (AADT) or in terms of assigned link peak hour volumes. The design process on the other hand requires actual design hour volume (DHV) which in conjunction with guidelines from the AASHO or Highway Capacity Manuals is employed to produce facility design data. It cannot be over-emphasized that presently a gap exists between systems planners who are thoroughly familiar with the urban transportation planning system model (UTPS) and functional planners who may not be aware of the basic assumptions inherent in the process. The following list summarizes the various aspects of the UTPS model that should be taken into consideration to bridge this gap:

1. errors in the urban transportation planning system model itself both in terms of land-use input, model formulation and output.

FIGURE 8.2

SCHEMATIC ILLUSTRATION OF THE VARIATION IN TOTAL SYSTEM COST BY DIFFERENT DHV DECISIONS

2. errors due to seasonal variations since the UTPS model generally produces demand estimate for an average daily traffic (Many of the standard manuals suggest that the peak hour volume from the UTPS may be employed as the design hour volume.),
3. weekly, daily and hourly variations of travel specific to the link under consideration,
4. the method of treatment of truck travel in the UTPS,
5. the procedure used in developing the directional split factors, and
6. the procedure used in applying the peak hour ratios to convert the daily traffic volume estimated by the UTPS model into peak hour volume, which is the central theme of this thesis.

As an example to demonstrate the sensitivity of the eventual design to the above factors, it can be easily shown that the application of one combination of the above design factors to an AADT of 60,000 would produce a DHV of about 3,400 vehicles per hour while the application of another combination may produce over 6,000 vehicles per hour. This would be the difference between a four lane freeway with low level of service and an eight lane freeway with an acceptable level of service.

Because the lane requirements for a given design hour volume is a step-like process (an increase of one lane accommodates an additional demand of say 1,200 vehicles per hour at a given L.O.S), the eventual design is sometimes insensitive to small variations or estimation errors of the design hour volume. The task of both the systems planner and the facility designer is to attempt to minimize the estimation errors which could result in under-designing or over-designing the facility.

In order to minimize the impact of these estimation errors, the following remedial actions should be considered:

- a. Traffic estimates from the UTPS model, when given to facility designers, should indicate the expected range of error of the estimates. The main advantage would be to warn the user of these demand estimates that errors do exist and should be taken into consideration in the design process.
- b. Estimation errors may be taken into consideration implicitly by resorting to various defence measures such as the skillful acquisition of right-of-way and staging.

Given that estimation errors will occur in the UTPS, it can be demonstrated that further research into the UTPS itself, more data acquisition, higher sampling rates for origin-destination surveys, more reliable land use data, and more accurate and practical modelling procedures, all should lead to better design decisions. The main task is to establish the trade-offs between the cost of research into the above topics and the expected net gain from this experimentation.

While the use of decision theory in this thesis assumes that the model reliability using the recommended approach for peak hour ratios estimation is higher than the reliability using the existing area-wide approach, the basic principles of applying the theory have wider scope which may cover the above research topics. Research by Johnson (51) for instance, demonstrated the application of the decision theoretic technique for information acquisition in transportation planning. Furthermore, the theory may be employed in evaluating traffic flow cost impacts for each state of nature (derived using equilibrium traffic

assignment technique) in the face of alternative decisions regarding peak hour flow capacity. Again, the trade-offs would be between the cost of the research and the associated net gain.

In deciding whether to experiment or not, the following factors should be considered:

1. the sensitivity of the design to the variation of the expected link volume (regardless of the approach or experiment, if the link volume is too low or too high to make a difference to the design decision, no experimentation will be necessary),
2. whether or not the design decision is based on considerations other than the anticipated link peak hour volume,
3. conflict of producing fine data for one element of the process versus coarse data for the rest of the process,
4. short term versus long term design considerations,
5. staging and right-of-way protection considerations, and
6. the overall transport supply at a given time relative to the overall travel demand (it may be that, because of resources limitations, the overall supply of transportation services is much lower than the anticipated demand which will make the question of over investment irrelevant).

8.4 A Framework for Design Decisions

An examination of the basic concepts in decision analysis reveals that the theory may be applied to illustrate the cost effectiveness of improving not only the peak hour ratios estimation, but also improving other elements of the UTPS model. The question which decision

theory addresses is simply that of gaining more information and in turn more reliable modelling procedures traded off against the cost of experimentation in the light of the eventual design decision. For example, it can be argued, as Johnson (51) did, that taking continuous counts for four months at a number of locations and at a high cost may prove to be more cost effective than making a design decision based on only limited low cost counts.

Design decisions are usually made by professionals other than those involved in the systems planning process. These decisions are highly influenced by planning decisions which dictate the overall transportation planning process. On a higher level, policy decisions, specifically those dealing with the amount of transport supply (or level of service), can and do influence the other two types of decisions. In fact, it has been the author's experience that policy decisions, influenced by limited resources and budgetary constraints, dictate to a great extent the other two types of decisions. It is for this reason that the design decision framework described in this section can only be simplistic in structure and in content. To make the framework more comprehensive, it will require the introduction of such complex aspects as priority planning, staging, and socio-environmental impacts.

8.4.1 Problem Statement

The problem can be stated as simply that of weighting the trade offs between the cost of researching peak hour ratios and the expected gain from the research as manifested by the link design decision. The objective is to minimize the under or over-designing of the link under consideration.

This basic definition points out:

- a. the need to explicitly treat the problem of uncertainty with respect to the peak hour volume which will actually use a given facility being designed,
- b. the need to understand the inter-dependency of the final design decision and the individual sequential steps of the urban transportation planning process,
- c. the fact that many considerations (other than peak hour volume) which cannot be treated explicitly enter into the design decision, and,
- d. the need for a simple systematic framework for delineating or responding to the problem posed by the above trade offs.

The above problem statement invites a simple and classical application of statistical decision theory termed Bayesian which was developed for practical problems by Raiffa and Schlaifer (52,53). A detailed description of the theory is presented in the next few sections.

8.4.2 Existing Approaches to Design Decisions and Their Deficiencies

In Section 8.1 a number of considerations entering the design decision process were noted. Because of the complexity and broad definition of these considerations, the design decision process has been and continues to be a highly qualitative and subjective process.

Several basic references such as the Highway Capacity Manual and the AASHO policy publications constitute at present, the main input to the design decision process. The integration of the output from the systems planning process into the design process is surprisingly neglected in the literature. This supports the argument repeatedly stated in this thesis that a gap exists between systems

planners and functional planners. One reason for this gap is probably because the systems planning process generally addresses intermediate and long term planning horizons while the functional planning and design process addresses short range and/or immediate applications.

In practice, it is still not uncommon to find transportation agencies making simple extrapolation of existing traffic volumes for a number of links to 5 and 10 years into the future and then use these projected volumes as input to the so-called Five Years Program.

Traffic counts on major arterials and along screenlines and cordons lines are still the most important data source for transportation planning in current urban and regional transportation studies. The main reasons for using traffic counts so extensively in the link design decision process are:

1. the ease by which counts can be obtained on a continuous basis,
2. the acquisition of these counts is or can be automated, and
3. unlike the origin-destination surveys which are based on a small sample, counts measure total quantities instead of partial samples.

Recently, with the introduction and wide spread availability of electronic computers, much more extensive use is being made of both traffic counts and origin-destination surveys.

One area which has not yet been extensively applied to transportation problems is the analysis of design decisions under uncertainty. The main reason for the interest in this topic is the fact that presently greater

recognition is given to the need to allocate the limited resources available more economically.

Applications of the Bayesian decision approach to real world problems are limited. Green (54) indicated application of the approach to problems arising in pricing, new product introduction, and marketing research. Hutchinson (55) formulated conceptual framework for pavement design decisions.

The decision analysis approach was also directed to the question of information acquisition and the evaluation of alternative transportation plans. Johnson (51) presented a rational procedure to allocate resources to information acquisition activities in the uncertain environment of transportation decision making. He formulated two models for selecting an optimum information acquisition activity to estimate the current state of travel demand. He presented a single link information acquisition model which is sensitive to the travel cost savings, investment cost of project alternatives, and to the specifications of the decision maker's prior probabilities over the true states of the system. He also formulated a network model with a global demand function and used observed flow counts on network links and origin destination survey results to revise the decision maker's prior probabilities over the state of the model parameters.

Manheim (56) developed a Bayesian decision model for guiding the engineering, planning or design process for the highway location problem. Khan (57) presented a framework for transport policy decisions which accounts for uncertainty of the possible impacts. He provided a systematic process for revising the prior knowledge of the public acceptance of a policy alternative through

market survey and demonstration studies. He also allowed a formal treatment of the decision maker's preference structure for consequences in conjunction with random nature of demand.

Decision theory has also been applied to an intersection problem. Burnes (58) used counts and delay data to revise prior knowledge about the conditions at the intersection in order to decide among a number of options to improve the intersection.

At present, no formal systematic decision theoretic framework to respond to the problem statement indicated earlier exists. In fact, the author is not convinced that all factors that enter into the design decision process can or should be explicitly treated by means of a decision theoretic framework. Most of the references cited above seem to address hypothetical situations with assumptions that may not be valid for real world problems. The application of statistical decision theory in this research was based on actual data obtained over the past 10 years from records and files of the Regional Municipality of Ottawa-Carleton.

The simple framework developed is prescriptive in nature. It provides a basis for selecting the upper level of investment in research or experimentation. The framework can also be extended to a two or more stage process. Additionally, the design decision problem itself can be expanded beyond the simple formulation described in this chapter.

8.4.3 Bayesian Decision Theory Approach

Statistical decision theory is usually termed Bayesian because it utilizes Bayes Theorem. The approach

has been developed and applied by Raiffa and Schlaifer (52;53).

The main reason for using the Bayesian statistical decision technique is to demonstrate how the added model reliability due to improvements in the estimation of peak hour ratios can be cost effective. The Bayesian approach does not explain how the design decision will be made but rather provides a structured framework to make the decision according to the decision maker's preferences and subjective judgment.

The basic elements used in the statistical decision theory approach are identified below. Further discussion of these elements and their measurements will be presented in later sections of this chapter.

1. Alternative Design Decisions (Actions); $a_1, a_2, a_3, \dots$

These are the basic choices available to the decision maker which indicate the type of facility to be designed. Examples of alternative actions include building a two lane, three lane or a four lane highway link.

2. States of Demand (Actual True Peak Hour Volume) $t_1, t_2, t_3, \dots$

This is the actual true peak hour volume which will use the facility after a particular action a_i is taken. This volume is treated as uncertain.

3. Alternative Approaches (Experiments): $e_0, e_1, e_2, e_3, \dots$

These are the various experiments that reflect different levels of investment in experimentation for the development of reliable tools for peak hour ratios estimation. The null approach which indicates no experimentation is e_0 .

4. Possible Peak Hour Volumes (Experimental Outcomes) $v_1, v_2, v_3, \dots$

For any experiment e , a number of possible outcomes will result. The experimental outcome v is the assigned peak hour volume using a given link under a particular experiment e .

5. Costs

There are basically two types of costs involved, cost of experimentation expressed as $C_s(e, v)$ and cost of action $C_l(a, t)$. The cost of experimentation is spread over the links being designed in the typical five year program. The cost of terminal action is the total system cost including capital, maintenance and operating costs for a particular design action. The total link costs for conducting experiment e , obtaining outcome v , taking action a and observing the true peak hour volume t is given as:

$$C(e, v, a, t) = C_s(e, v) + C_l(a, t) \quad \dots \text{Eq. (8.4)}$$

6. Prior and Conditional Probabilities

The prior probabilities are those probabilities of observing the peak hour volume t when no experimentation is made. The conditional probabilities $P(v|t, e)$ are those distributions over the possible outcome of experiment e of observing peak hour volume v given a true peak hour volume t .

The prior probability expresses the knowledge of the decision maker of the state of demand (i.e. true peak hour volume t prior to experimentation). The conditional probability distributions $P(v|t, e)$ establish the connection between the actual true peak hour volume and the experimental outcome. This probability is a measure of

the reliability of the proposed modelling approach or experiment.

8.4.4 A Framework for the Design Decision Process

The elements of the Bayesian decision theoretic approach described above may be systematically structured in order to develop a framework for making design decisions for any link in the transportation network. A simple illustration of the envisaged framework is shown in Figure 8.3.

Examination of the framework should make it immediately apparent that a rather simplistic structure is outlined, however, the framework may be extended in two directions. Firstly, a multi-stage process instead of a single stage process may be performed, and secondly, the input data, particularly the cost functions, may be further disaggregated and detailed for the various options or experiments.

In order to simplify the discussion of the framework, it may be described in terms of two basic options. The first is the null or the do-nothing option where an area-wide peak hour ratio is applied to convert 24 hour travel into peak hour travel. The second option specifies the experimentation alternative whereby an improved approach to peak hour ratios estimation is envisaged.

Application of the improved peak hour ratios estimation procedure would generate, once assigned to the transport network, various possible experimental outcomes. These are the possible assigned link peak hour volumes and they are represented by the V space in Figure 8.4. The probability distribution $P(v)$ of these experimental outcomes is

FIGURE 8.3

A FRAMEWORK FOR THE DESIGN DECISION PROCESS

T Space		Marginal Probty. of V_i
$t_1 \dots \dots \dots t_i$		
V_1	$P(t_1 \cap v_1)$ $P(t_1 v_1) P(v_1)$ $P(v_1 t_1) P(t_1)$	$P(t_1 \cap v_1)$ $P(t_1 v_1) P(v_1)$ $P(v_1 t_1) P(t_1)$
V_i	$P(t_i \cap v_i)$ $P(t_i v_i) P(v_i)$ $P(v_i t_i) P(t_i)$	$P(t_i \cap v_i)$ $P(t_i v_i) P(v_i)$ $P(v_i t_i) P(t_i)$
Marginal Probty. of t_1	$P(t_1) \dots \dots \dots P(t_i)$	Experiment e_j

FIGURE 8.4
RELATIONSHIP BETWEEN MARGINAL AND
CONDITIONAL PROBABILITIES

calculated using the prior probability distribution $P'(t)$, represented by the T space in Figure 8.4, and the conditional probability distribution $P(v|t,e)$, using the relationship shown in Figure 8.4 as follows:

$$P(v_i) = \sum_{t_i} P'(t_i) \cdot P(v_i|t_i,e) \quad \dots\dots\text{Eq. (8.5)}$$

The marginal probability $P(v)$ when used in conjunction with the conditional probability measure enables the designer to determine the posterior probability distribution of the actual or true peak hour volume which will use the link $P''(t|v)$. These posterior probability distributions are the decision maker's revised knowledge of the likelihood of the true peak hour volume t . Furthermore, these revised probabilities are used in turn to compute the expected cost of each option. Obviously, the option with the minimum expected cost C_e (min) is optimum.

The design decision for a given link in the network is simulated using the so called "preposterior analysis". The objective of the preposterior analysis in simulating the link design decision is to evaluate the various experimentations or options being tested.

A simple decision tree may be plotted as shown in Figure 8.5 in order to illustrate the application of preposterior analysis. The sequence of operations in this analysis may be briefly outlined as follows:

1. Develop $P'(t)$, the prior probability which expresses prior judgment about the relative likelihood of different states of nature (true peak hour volume t)
2. Determine for each experiment e , its conditional probability characteristics as given by $P(v|t,e)$
3. Compute the marginal probability measures

$$P(v|e) = \sum_t P'(t) \cdot P(v|t,e) \quad \dots\dots\text{Eq. (8.6)}$$

FIGURE 8.5

A SIMPLE DECISION TREE

4. Compute the posterior distribution

$$P''(t|v,e) = \frac{P'(t) \cdot P(v|t,e)}{\sum_t P'(t) \cdot P(v|t,e)} \quad \dots\dots \text{Eq. (8.7)}$$

5. Determine for each combination of e, v, a , and t its relative cost $C(e, v, a, t)$

6. Compute the expected cost for each act a

$$C^*(a, v, e) = \sum_t P''(t|v,e) \cdot C(e, v, a, t) \quad \dots\dots \text{Eq. (8.8)}$$

7. For each e, v combination, determine the optimal act

$$C^*(v, e) = \min_a C^*(a, v, e) \quad \dots\dots \text{Eq. (8.9)}$$

8. For each e , compute the expected cost

$$C^*(e) = \sum_v P(v|e) \cdot C^*(a, v, e) \quad \dots\dots \text{Eq. (8.10)}$$

9. The optimal experiment e^* is that e for which $C^*(e)$ is a minimum, i.e.:

$$C^*(e) = \min_e C^*(e) \quad \dots\dots \text{Eq. (8.11)}$$

The objective of the decision maker is to select the experimental option which has the lowest expected total cost in dollars as given by step 9 above. Figure 8.6 gives a flow chart which illustrates the sequence of operations outlined above. Details of the actual calculations are given in Appendices H, I, and J.

FIGURE 8.6

PREPOSTERIOR ANALYSIS

8.4.5 Expected Value of Information

The decrease in the total expected cost of an option is considered a measure of the effectiveness of the option. This reduction in terminal costs (i.e. the terminal costs with no information less the terminal costs with information) is the expected value of information. An effectiveness measure of the experiment which is closely related to the expected value of information is the expected net gain of peak hour volume information. The expected net gain of the peak hour volume information acquisition activity is the expected value of information less the cost of acquiring it. Figure 8.7 illustrates the type of calculation performed to develop the expected value of sample information. Detailed description of the actual equation derivation is given in Appendix H.

8.5 Application of the Decision Theoretic Framework

The purpose of this section is to show the application of the decision analysis framework by structuring the data needed for the design decision. The cost effectiveness of the improved approach for peak hour ratios estimation and application in the UTPP will be demonstrated along the activities indicated in Figure 8.8. A number of experimental options will be outlined and actual case studies will be used to show the actual decision analysis procedure.

The design decision problem can be simply stated as that of choosing the experimental option or modelling procedure that will have the lowest total expected cost to the decision maker. Preposterior analysis is employed

FIGURE 8.7

EXPECTED VALUE OF SAMPLE INFORMATION

FIGURE 8.8

APPLICATION OF THE DECISION ANALYSIS FRAMEWORK

to identify the optimal option or experiment to be performed as well as to compute the expected value of information.

A computer programme for design decision application was written (Appendix I) and was run interactively on a remote terminal. The input to the programme consists of the procedure, prior and conditional probabilities, cost of experimentation and cost of alternative actions. A single stage process is employed whereby the expected costs for each alternative action are computed together with the expected value of peak hour volume information resulting from improved peak hour ratios estimation as well as the expected cost of experimentation.

In the next few sub-sections the method of development of the various input data used in the decision analysis programme are described. It is emphasized that, although the objective criteria for the application of decision analysis are to minimize the cost to the decision maker, in practice however, other non-quantifiable factors may enter the design decision process and alter the decision. This should be borne in mind when comparing the results of the four case studies used in this design decision analysis with the decisions actually made in practice.

8.5.1 Selecting the Options

Five experiments or options (including the do-nothing "null" option) were used in the design decision analysis. The following is a brief description of each of these options.

The do-nothing option involves the application of an area-wide peak hour ratio (say 10%) to convert 24 hour travel matrix into peak hour travel matrix. This option represents,

more or less, the existing practice of applying peak hour ratios. Throughout the text this option is referred to as e_0 .

The next option, referred to as e_1 , represents an improved approach whereby peak hour ratios are estimated and applied by mode and purpose but still using area-wide average values. This option requires additional information and research over and above the do-nothing option. The activities and costs incurred in this option include writing a trip-in-motion programme, executing the programme, data processing and manpower costs. It also includes conversion of matrices by mode and purpose from 24 hour to peak hour. The total expected cost of this option can thus be developed.

The next option, referred to as e_2 , represents the application of the recommended approach to peak hour ratios estimation. Incurred in this option are all the activities described in the previous option. Additionally, the costs of all statistical analyses performed in this thesis are also included.

Experimental options e_3 and e_4 involve both, additional research and further data collection in the form of higher survey samples. Option e_3 for example, involves repeating the research described in this thesis using additional data obtained by increasing the survey sample from 2% to about 7 or 8%. Option e_4 on the other hand, involves a much larger sample, say up to 20% so that the concept of developing zonal interchange peak hour ratios can be applied. In addition to the cost of sampling, application of zonal interchange peak hour ratios in option e_4 involves the use of matrices instead of column fields and thus the cost of data processing and research procedure may be doubled.

Table 8.1 shows an estimate of the total experimentation cost for each of the five options. Note that no cost has been indicated for the do-nothing alternative since the overall steps in the urban transportation planning system model may be common to all five options.

Examination of the Five Year Programme for new links and/or improvements to existing links within the Regional Municipality of Ottawa-Carleton showed that about 20 links are included in the programme. Assuming that the cost of each of these options will be distributed equally among the 20 links in the five year programme, the resulting experimentation costs per link were calculated. These were rounded to the following values which were used as input to the computer programme:

- . cost per link for the e_0 option = \$0
- . cost per link for the e_1 option = \$100
- . cost per link for the e_2 option = \$400
- . cost per link for the e_3 option = \$2,000
- . cost per link for the e_4 option = \$6,000

The above distribution of experimentation costs represents an adequate approximation which may be used to describe the application of the Bayesian statistical decision technique to the design problem. Obviously, these experimentation costs are very small in comparison to the millions of dollars in operating, maintenance and construction costs incurred on a given facility. It will be demonstrated nevertheless, that in certain situations these experimentation costs are not justified.

TABLE 8.1
ESTIMATES OF TOTAL EXPERIMENTATION COST* BY OPTION

Experiment	Activity	Cost Per Activity	Cost Per Link \$	$C_s(e)**$
e_0	-	0	0	0
e_1	Trip in Motion Pg. Test Programme	\$ 600	75	100
	Data Processing	50		
	2 Man-Day Labour	200		
	Execute Programme	150		
		500		
		<u>\$ 1,500</u>		
e_2	e1 activities Develop matrices	\$ 1,500	375	400
		6,000		
		<u>\$ 7,500</u>		
e_3	e2 activities Increase Sample Size to 7%	\$ 7,500	2,375	2,000
		40,000		
		<u>\$ 47,500</u>		
e_4	e2 activities times 2 Increase Sample Size to 20%	\$ 15,000	5,750	6,000
		100,000		
		<u>\$115,000</u>		

* Assuming 20 link improvements in the 5 year programme
 ** Values used in the Computer programme, Appendix II

8.5.2 Prior Probabilities

The prior probability distributions over the possible demand states $P'(t)$, expresses the level of confidence of the decision maker with the particular modelling procedure prior to experimentation. For instance, suppose the possible demand states (actual peak hour volumes) observed on a given link have an equal probability of being 1000, 2000, or 3000 vehicles per hour, as shown in Figure 8.9, this will mean that the prior probability distribution is a rectangular distribution over the 3 possible demand states each with a probability of 0.333. However, in most cases, the prior knowledge exhibits central tendencies which take the form of a normal or triangular probability distribution.

As will be demonstrated later, the design decision is highly sensitive to the prior uncertainty of the demand state. In this research, the prior probability distribution was developed using traffic data for about 60 regional road links in the Regional Municipality of Ottawa-Carleton. The traffic data compiled, covered a period of 10 years. The methodology used in developing the prior probability distribution is described briefly in the following paragraphs.

On an annual basis, the Regional Municipality of Ottawa-Carleton compiles the so called Road Needs Study Sheets which are forwarded to the Ministry of Transportation and Communications of Ontario. These sheets are compiled for each link of the regional road system. Included on these sheets are five and ten year traffic projections. The 1968 Road Needs Study Sheets were obtained from the City of Ottawa records since the

FIGURE 8.9

RECTANGULAR PRIOR PROBABILITY DISTRIBUTION

Regional Government was not yet created. Over 60 links of the urban and semi-urban regional links were selected randomly as shown in Table 8.2. The existing 1968 design hour volume as well as the ten year projection to 1978 were noted*. In 1977 and 1978 traffic counts were taken over the entire regional road system. These 1978 traffic counts were compared to the originally projected 1978 volumes and the differences were noted and coded. The difference code column in the table indicates the range of differences between the projected 1978 DHV and the observed peak hour volumes. Code 1 indicates a range of $\pm 10\%$; Code 2 a range of $\pm 30\%$; Code 3 a range of $\pm 50\%$; and Code 4 a difference $\pm > 50\%$. Figure 8.10 shows the frequency distribution of the ratios of the projected and observed volumes.

Examination of the frequency distribution in Figure 8.10 indicates that the transportation planners tended to over-estimate future traffic volume projections for the period 1968 to 1978. In attempting to explain the reasons for the over-estimates, the following points are made:

1. the rate of growth in population of the Regional Municipality of Ottawa-Carleton did not conform to assumptions made in 1968,
2. the introduction of flexible work hours in 1974 for federal government employees, and
3. substantial increases in the transit level of service in recent years with corresponding increases in transit ridership (59).

* These were based on assigned peak hour volumes which were developed by the application of area-wide peak hour ratios by trip purpose.

TABLE 8.2

D.H.V. ANALYSIS DATA FOR URBAN SECTIONS

Link Number	Existing 1968 DHV	Projected 1978 DHV	Observed 1978 Volume	Difference Code	Projected ÷ 1978 DHV
11	1046	2200	1670	+2	1.32
20	1314	2980	2120	+3	1.40
21	996	2180	1420	+3	1.54
86	360	650	580	+2	1.12
136	400	800	700	+2	1.14
159	360	575	560	+1	1.03
172	340	550	550	+1	1.00
180	900	1450	1430	+1	1.01
173	250	400	450	+1	.89
218	800	1450	1110	+2	1.31
226	775	1330	1380	+1	.96
239	621	1100	930	+1	1.18
267	1010	1900	1580	+2	1.20
268	1010	1810	1550	+2	1.17
273	997	1670	1870	-2	.89
274	600	1000	980	+1	1.02
275	770	1290	1180	+1	1.09
286	430	940	620	+2	1.52
333	327	530	580	-1	.91
370	345	650	620	+1	1.05
497	700	1480	870	+4	1.70
498	1100	2100	1600	+3	1.31
589	870	1740	1200	+3	1.45
599	452	900	1230	-3	.73
600	440	880	930	-1	.95
330	620	1250	850	+3	1.47
88	534	750	800	-1	.94
90	540	870	670	+2	1.30
173	250	1250	460	+4	2.72

TABLE 8.2 (Continued)
D.H.V. ANALYSIS DATA FOR SEMI URBAN SECTIONS

Link Number	Existing 1968 DHV	Projected 1978 DHV	Observed 1978 Volume	Difference Code	Projected ÷ 1978 DHV
31	760	1470	1400	+1	1.05
43	660	1130	1350	-2	.84
92	500	860	900	-1	.96
108	1120	2330	1600	+3	1.46
110	830	1420	1100	+2	1.29
130	520	1040	620	+4	1.68
137	362	740	690	+1	1.07
204	489	980	1600	-4	.61
227	1080	2160	1460	+3	1.48
251	960	2000	1180	+4	1.69
252	1236	2724	1910	+3	1.43
253	1236	2724	1430	+4	1.90
255	700	1580	1140	+2	1.39
256	700	1580	1210	+2	1.31
263	895	1740	1580	+1	1.10
264	680	1320	1330	-1	.99
266	771	1450	1190	+2	1.22
285	430	940	810	+2	1.16
291	273	530	762	+3	.70
292	770	1500	980	+4	1.53
324	894	1705	1180	+2	1.44
325	1101	2160	1890	+2	1.14
342	380	830	270	+5	3.07
347	500	970	651	+4	1.49
348	610	1160	1100	+1	1.05
349	800	1520	1220	+1	1.25
350	390	740	600	+2	1.23
526	165	244	280	-1	.87
534	645	1650	1511	+1	1.09
536	550	1550	740	+5	2.09

FIGURE 8.10

PRIOR PROBABILITY DISTRIBUTION FUNCTION
BASED ON 1968/1978 DATA

A normal probability distribution curve was fitted to the data in Table 8.2 and the Chi-Square test was carried out. The observations were grouped into 7 categories as follows:

<u>Range Ratio</u>	<u>Number of Observations</u>
0.50-0.70	2
0.71-0.90	5
0.91-1.10	16
1.10-1.30	13
1.31-1.50	12
1.51-1.70	6
>1.70	4

The mean and the standard deviation for the above distribution were calculated as 1.22 and 0.294 respectively. The Chi-Square test compared the expected frequencies from the normal curve with the above observed frequencies. The value of Chi-Square obtained was 2.5 which, when compared to the value of 9.49 obtained from statistical tables (for $\alpha=.05$), indicated that the hypothesis of normal distribution of observations cannot be rejected.

Another problem which was encountered in developing this prior probability distribution was the difficulty of incorporating this continuous function into a discrete design decision process. For this reason, the area under the prior probability distribution curve was divided into five discrete segments. The following prior probability values for each

of the 5 segments were derived as shown in Figure 8.11:

$$P'(t_1) = 0.125$$

$$P'(t_2) = 0.200$$

$$P'(t_3) = 0.350$$

$$P'(t_4) = 0.200$$

$$P'(t_5) = 0.125$$

The above values were used as input to the design decision programme. These were later on varied as part of the sensitivity analysis described in Section 8.6 of this chapter.

8.5.3 Conditional Probabilities

These probabilities, referred to as $P(v|t,e)$, provide the relationship between the actual true peak-hour volume which will be observed on a given link and those measured by experimentation. These probabilities are therefore, characteristics of the option being considered and they reflect an assessment of the accuracy or reliability of the modelling procedure employed in that option.

The conditional probability distributions shown in Table 8.3 were derived for each experimental option subjectively after discussion with other professionals. These conditional probability measures can be developed objectively for some of the experimental options (see Figure 7.3) by means of comparing the assigned experimental peak hour volumes with the true observed peak hour volumes from ground counts. As indicated earlier, this process could not be performed due to the time constraints. Furthermore, for options e_3 and e_4 in particular, only subjective conditional probabilities can be developed since no empirical data exist to derive them objectively. Actually, assigning subjective

FIGURE 8.11
PRIOR PROBABILITY DISTRIBUTION FUNCTION
SUBDIVIDED BY 5 RANGES OF DEMAND STATES

TABLE 8.3

CONDITIONAL PROBABILITY MEASURES $P(v|t,e)$
BY EXPERIMENT ALTERNATIVES†

<u>Experiment Alternative</u>	<u>v</u>	<u>$P(v t,e)$</u>				
		<u>t_1</u>	<u>t_2</u>	<u>t_3</u>	<u>t_4</u>	<u>t_5</u>
e_1	v_1	.500	.200	.050	.100	.030
	v_2	.250	.500	.200	.100	.070
	v_3	.150	.200	.500	.150	.150
	v_4	.070	.075	.200	.500	.250
	v_5	.030	.025	.050	.150	.500
e_2	v_1	.600	.100	.050	.030	.040
	v_2	.300	.600	.150	.070	.060
	v_3	.070	.200	.600	.150	.100
	v_4	.020	.060	.150	.600	.200
	v_5	.010	.040	.050	.150	.600
e_3	v_1	.700	.125	.025	.075	.040
	v_2	.150	.700	.125	.075	.060
	v_3	.100	.125	.700	.100	.075
	v_4	.040	.035	.125	.700	.125
	v_5	.010	.015	.025	.100	.700
e_4	v_1	.800	.080	.030	.010	.030
	v_2	.075	.800	.070	.020	.040
	v_3	.055	.070	.800	.095	.055
	v_4	.040	.040	.080	.800	.075
	v_5	.030	.010	.020	.075	.800

† Subjectively Assigned

probabilities enables the decision maker to explore the implications of uncertainty on the results and conclusions. This permits the decision maker to see the basis for making a particular design decision.

8.5.4 Cost Functions for the Alternative Actions

The cost data used in this thesis were derived from files of the Programme Management Branch of the Regional Municipality of Ottawa-Carleton. All cost data were discounted to the base year conditions.

The input data to these cost functions were developed from the Economic Road Improvement Selector Programme (ERIS). This programme which was written by N. D. Lea and Associates for the Regional Municipality of Ottawa-Carleton, provides a method of determining road construction priorities on the basis of engineering economic analysis. Capital costs, road maintenance costs and operating costs developed to meet the Ottawa-Carleton conditions were included in this programme. The ERIS programme selects the optimum economic improvement for each link and ranks the priority of each recommended improvement within the road system being analysed.

The economic analysis for each link is based upon the vehicle operating, road maintenance and capital costs incurred each year and summarized over the forecast period. This period is usually twenty years. The breakdown of the cost variables used in the ERIS programme are briefly summarized below. The interested reader may consult the ERIS user's manual for detailed explanation of each variable (60).

1. Vehicle Operating Costs

- a. fuel costs consumed during normal driving stopping and idling conditions
- b. tire costs
- c. depreciation costs
- d. oil costs
- e. user costs
- f. accident costs
- g. fixed costs

2. Road Maintenance Costs

- a. annual maintenance
- b. periodic maintenance

3. Capital Costs

- a. right-of-way costs
- b. resurfacing costs
- c. construction costs

In order to construct the total system cost functions, the following data were obtained from the RMOC and placed in an SPSS system file (Table 8.4):

- . link number
- . number of lanes
- . type of link (urban or semi-urban)
- . total costs in thousands of dollars
- . length of link in miles*
- . design hour volume

Examination of the above data indicated the need to generate the following additional variables:

- . total costs in thousands of dollars per mile (or kilometers)

* These were converted later into kilometers

TABLE 8.4
COST DATA FOR THE COST CURVES

<u>Link Number</u>	<u>Number of Lanes</u>	<u>Type of Link</u>	<u>Cost in \$1000's</u>	<u>Length of Link-Miles *10</u>	<u>DHV</u>
11	2	2	1614	4	807
20	4	2	1531	1	1221
21	3	2	2197	4	903
22	4	2	2458	5	853
23	3	2	2570	3	959
24	3	2	1352	2	886
31	2	3	1109	4	481
42	4	2	2245	4	1021
43	2	3	1136	7	439
66	4	2	3162	7	854
67	4	2	2034	4	924
68	6	2	4245	8	1181
86	2	2	1674	8	428
87	4	2	285	1	296
88	2	2	530	2	522
90	2	2	723	1	608
92	2	3	1002	4	580
108	2	3	2482	4	1024
110	2	3	1569	8	947
130	2	3	1427	5	562
136	2	2	3230	3	871
137	2	3	6203	11	625
159	2	2	908	4	495
172	2	2	265	2	333
173	3	2	495	2	514
180	4	2	803	2	520
195	2	2	352	8	157
203	2	2	1875	10	363
204	2	3	1142	3	504
214	2	2	1232	3	820
215	4	2	1371	4	932
218	2	2	118	1	400
225	4	2	7714	9	1157
226	2	2	2352	3	697
227	2	3	12223	16	794
239	2	2	1321	7	340
244	2	3	1286	5	571
251	2	3	471	3	633
252	2	3	1235	6	680
253	2	3	880	1	1079
254	2	3	1229	2	785
255	2	3	942	2	946
256	2	3	771	2	813

TABLE 8.4 (Continued)

<u>Link Number</u>	<u>Number of Lanes</u>	<u>Type of Link</u>	<u>Cost in \$1000's</u>	<u>Length of Link-Miles *10</u>	<u>DHV</u>
260	4	2	810	3	440
263	2	3	2461	8	996
264	2	3	1321	6	611
266	2	3	3442	10	724
267	2	2	309	3	425
268	3	2	1810	6	653
273	4	2	21466	6	1291
274	2	2	2909	5	463
275	2	2	1583	3	664
285	2	3	607	6	501
286	2	2	722	4	566
291	2	3	5499	10	693
292	2	3	1726	9	730
324	2	3	4951	9	806
325	2	3	11464	4	1234
330	4	2	2311	6	773
333	2	2	1015	2	530
334	4	2	1501	2	1188
342	2	3	279	11	131
347	2	3	4448	14	437
348	2	3	1565	4	734
349	2	3	4462	7	864
350	2	3	1681	6	538
363	4	2	4503	4	1100
370	2	2	803	3	567
372	4	2	666	2	556
497	4	2	2209	6	736
498	4	2	3851	9	1045
257	2	3	3034	12	837
258	2	3	2686	4	917
280	2	3	3358	13	709
376	2	3	3	1	22
287	2	3	87	2	192
393	2	3	101	3	150
408	2	3	280	3	700
436	2	3	991	7	748
441	2	3	15	3	16
490	2	3	5054	14	679
491	2	3	1082	5	610
495	2	3	607	6	495
523	2	2	885	12	218
530	2	3	50	7	43
545	2	3	98	4	74
554	2	3	2434	12	517
555	2	3	632	4	411
567	2	3	66	5	45
571	2	3	2377	17	404

TABLE 8.4 (Continued)

<u>Link Number</u>	<u>Number of Lanes</u>	<u>Type of Link</u>	<u>Cost in \$1000's</u>	<u>Length of Link-Miles *10</u>	<u>DHV</u>
668	2	3	3109	13	762
669	2	3	1383	7	704
675	2	3	9	1	143
682	2	3	59	7	35
526	2	3	228	6	259
534	2	3	2148	10	974
536	2	3	2112	9	668
553	2	3	1583	7	517
589	4	2	2364	3	952
599	2	2	1156	5	385
600	2	2	1358	9	469
654	2	3	396	2	437
662	2	2	264	3	221
665	2	3	1407	9	751
672	4	2	1480	4	854
677	2	2	1941	13	317
689	2	3	547	4	216
690	2	3	376	3	194
800	3	2	1700	10	200
801	3	2	2400	10	325

- link length in miles (or kilometers) times the design hour volume (DHV)
- total costs in thousands of dollars per vehicle mile (or vehicle kilometer)

Regression analysis was performed on the above data in order to derive relationships between the total system costs in thousands of dollars per kilometer (C_a) and the design hour volume. Figure 8.12 shows a plot of the three equations developed by non-linear regression. These equations* are described as follows:

For a 2 lane urban link:

$$C_a = 0.266 (\text{DHV})^{1.5} - 766.9 \quad \dots\dots\text{Eq. (8.11)}$$

$$R = 0.87$$

$$F = 68.35$$

$$\text{Std. Error} = 1319.2$$

For a 3 lane urban link:

$$C_a = 0.0043 \left(\frac{\text{DHV}}{1000} \right)^3 + 1012.7 \quad \dots\dots\text{Eq. (8.12)}$$

$$R = 0.96$$

$$F = 65.85$$

$$\text{Std. Error} = 760.9$$

For a 4 lane urban link:

$$C_a = .0039 \left(\frac{\text{DHV}}{1000} \right)^3 + 685.8 \quad \dots\dots\text{Eq. (8.13)}$$

$$R = 0.86$$

$$F = 46.68$$

$$\text{Std. Error} = 2366.6$$

* All statistically significant at $\alpha = 0.05$

Because of the few observations available on file for links having a design hour volume less than 400 vehicles per hour, the expected overlap of the 3 lane and 4 lane total costs per kilometer functions did not materialize. Instead, the non-linear regression attempt forced the 4 lane curve to be below the 3 lane curve even at these low peak hour volume values. Instead of providing additional dummy data for further non-linear regressions attempts, it was decided to hand-fit the curves to the observed data. The curves were adjusted in order to make them reasonable and to remove inconsistencies produced by certain links in which the total cost was either too high or too low because of the specific conditions involved when that link was constructed. The resultant cost function curves are shown in Figure 8.13. These curves are characterized by a steep increase in the system costs as the peak hour volume increases due to increases in both operating and maintenance costs.

Since the road system in the RMOC does not contain many links with 6 lanes, it was not possible to derive the cost function for a 6 lane urban link. Furthermore, the limited number of observations did not permit the disaggregation of the cost functions by various categories such as divided or undivided roads, different pavement widths and different shoulder widths. This limited the choice among the alternative actions to only three. These were action a_1 requiring a two lane facility; action a_2 requiring a 3 lane facility and; action a_3 requiring a 4 lane facility.

The computer programme in Appendix I is presently limited to selecting a design decision from among the above three alternative actions, but it can easily be expanded to accommodate additional alternative actions if data needed to update the cost functions are made available.

8.5.5 Case Studies

Four links were selected from the regional road system to test the decision analysis framework. Data were obtained for the following four links:

- . Alta Vista Drive from Randall Street to Smyth Road
- . Fisher Avenue from Prince of Wales Drive to Dynes Road
- . Meadowlands Road from Withrow Street to Bayne Street
- . Merivale Road from Jamie Street to Viewmount Drive

All the above links represent major arterials in Ottawa-Carleton. The computer programme was executed interactively for each of these case studies. Appendix J shows a complete listing of one full programme execution. Table 8.5 summarizes the results of the application of the decision analysis programme for the above four case studies. The results indicate that in only one of the four case studies did the number of lanes actually built coincide with the number of lanes selected by the decision analysis programme. It is interesting to note that option e_0 is the optimal experimentation option for that decision.

Examination of the decision analysis paths for the four case studies show that in two of the above cases the do-nothing option was optimal while for the other two, option e_4 (the highest in terms of experimentation costs) was selected. This is not surprising since the options in the programme are numbered in order of increasing experimentation costs. Because of the relatively small values of experimentation costs relative to the total system costs for any alternative action, most of the results of application of the decision analysis programme will indicate either the do-nothing or the do-most options. However, as described in the following section, there are cases in which the experimental option selected is neither the do-nothing nor the do-most options.

TABLE 8.5

RESULTS OF DECISION ANALYSIS APPLICATION
FOR 4 CASE STUDIES

	<u>Built</u>	<u>Decision Analysis Minimum Path</u>
Alta Vista Drive (From Randall Street to Smyth Road)	a_1	$e_0-v_0-a_3$
Fisher Avenue (From Prince of Wales Drive to Dynes Road)	a_3	$e_0-v_0-a_3$
Meadowlands Road (From Withrow Street to Bayne Street)	a_2	$e_4-v_1-a_1$
Merivale Road (From Jamie Street to Viewmount Drive)	a_3	$e_4-v_1-a_2$

8.6 Sensitivity Analysis

The sensitivity of the optimal design decision to a number of basic changes to the input parameters was studied. The sensitivity analysis involved changing the distribution of prior probabilities $P'(t)$, changing the expected range of the actual peak hour volume t , and varying the length of link being designed. Additionally, a search process was made to identify, for each particular experimental option, when that option will be optimal.

The expected costs of alternative actions for different distributions of the prior probability $P'(t)$ are shown in Figure 8.14. This figure indicates that as the prior probability density function becomes more peaked, the expected value of information decreases. This conclusion agrees with a priori reasoning. The higher the peak of the prior probability density function, the better the initial model reliability will be and the value of additional information (in relative terms) decreases. This suggests that for the do-nothing option, if a lower prior uncertainty in the model results can be achieved by such means as obtaining more reliable counts or having more time allotted for the study, then the expected value of experimentation decreases.

It can also be observed that the shape of the prior probability density function can influence the choice of the alternative action. Flatter curves favour alternative action a_2 (within the ranges of t values tested in the computer programme).

FIGURE 8.14
SENSITIVITY ANALYSIS OF THE VARIATION IN
THE PRIOR PROBABILITY DISTRIBUTION FUNCTION

It should be evident that the total system cost curve favours the a_3 alternative action if there is a wide spread of t values for a given prior probability distribution. The reason is quite simply because the change in total system cost among each alternative action at a given t value is much greater for t values in the range where a_3 is optimal than for t values where a_1 or a_2 are optimal.

In the next sensitivity analysis performed, the prior probability density function was retained constant while the range of the possible demand state values t (actual peak hour volumes) was varied. As shown in Figure 8.15, when the range of possible demand states is narrow (i.e. it falls within the portion of the cost curves where one alternative action is always optimal), the net value of experimentation is negative since there will be no reason for experimentation. If the range of possible demand volume falls in the range where the cost curves overlap, the alternative action selected will probably be a_2 . The expected value of experimentation tends to vary depending on the spread in the values of the possible demand states.

It should be evident that as the length of a link being designed decreases, the importance of the cost of experimentation (particularly to the do-most option) becomes significant. Figure 8.16 indicates that if the optimum decision analysis path is identical for a number of cases, the minimum cost of the optimal action is proportional to the length of the link (assuming all other input parameters remain constant). This conclusion agrees with a priori reasoning. The figure also indicates that as the link length decreases, experiments involving low

FIGURE 8.15

SENSITIVITY ANALYSIS OF THE VARIATION IN THE
RANGE OF DEMAND STATES

FIGURE 8.16
SENSITIVITY ANALYSIS OF THE VARIATION
IN LINK LENGTH

TABLE 8.6

SUMMARY OF DECISION ANALYSIS RUNS UNDER VARIOUS INPUT ASSUMPTIONS

Length L (Km)	$P^*(t)$	t	Expt.	Alternative
0.0016	.125-.2-.35-.2-.125	400-550-700-850-1000	e ₀	a ₃
1.6090	.125-.2-.35-.2-.125	100-250-300-350-400		a ₁
1.6090	.125-.2-.35-.2-.125	900-1075-1250-1425-1600		a ₃
0.16090	.125-.2-.35-.2-.125	500-750-1020-1300-1500		a ₃
1.93080	.125-.2-.35-.2-.125	500-650-825-900-1000		a ₃
1.28720	.125-.2-.35-.2-.125	725-800-875-950-1025		a ₃
0.00804	.125-.2-.35-.2-.125	400-550-700-850-1000	e ₁	a ₂
0.169 < L < 0.2041	.125-.2-.35-.2-.125	400-550-700-850-1000	e ₂	a ₂
0.16090	.125-.2-.35-.2-.125	400-600-875-950-1080	e ₃	a ₂
0.2041	.125-.2-.35-.2-.125	400-550-700-850-1000	e ₄	a ₂
1.6090	.025-.25-.45-.25-.025	400-550-700-850-1000		a ₂
1.6090	.155-.22-.25-.22-.155	400-550-700-850-1000		a ₁
1.6090	.2-.2-.2-.2-.2	400-550-700-850-1000		a ₁
1.6090	.025-.1-.75-.1-.025	400-550-700-850-1000		a ₂
1.6090	.125-.2-.35-.2-.125	200-500-800-1100-1400		a ₂
1.6090	.125-.2-.35-.2-.125	600-650-700-750-800	a ₂	a ₂
1.2872	.125-.2-.35-.2-.125	400-600-875-950-1080	a ₂	a ₂
3.2180	.125-.2-.35-.2-.125	400-600-875-950-1080	a ₂	a ₂
1.2872	.125-.2-.35-.2-.125	250-562-875-1175-1400	a ₂	a ₂

sampling costs are favoured. This is obviously because of the above mentioned reason. The figure clearly indicates that if the length of the link being designed exceeds 0.2041 km, the do-most option e_4 is always selected.

A lengthy search process involving numerous attempts can also identify, given other input parameters, when the other experimental options e_1 , e_2 , and e_3 are selected. This is illustrated in Table 8.6 which summarizes all the computer runs of the Bayesian Decision Programme performed in this research. The table clearly shows that the do-nothing and the do-most experimental options are selected for numerous input parameter conditions.

The above sensitivity analysis demonstrated the use of the decision theory as a useful input into the design process. The author plans to expand the existing computer programme to cover other variations in the basic input parameters to make the programme more directly applicable to the design decisions within the Regional Municipality of Ottawa-Carleton.

CHAPTER IX

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

9.1 Introduction

In general, the results of this research achieved the objectives stated in Section 1.3. There are, however, several areas where further research could well lead to increased knowledge and improved peak hour ratios modelling and application. These areas will be delineated in the last section of this Chapter.

Initially the objective of this research was to develop, using the available data, an improved approach for peak hour ratios estimation and application. At the conclusion of the research however, the topic has led in two directions:

1. inward, in the sense that most levels of analysis ranging from the aggregate at the area-wide, district and zonal levels to the disaggregate household and individual level were attempted, and
2. outward, in the sense that the issue is not only the proper estimation of peak hour ratios, but their proper application in a given modelling procedure and at various stages of the UTPS model. This point in turn, raised numerous questions regarding the philosophy of transportation modelling, and indicated that not only are there numerous approaches for estimating peak hour ratios, but that there are equally several sets of applications which are interdependent with the other elements of the modelling process. Without understanding this interaction, there is little one can do with any results which the research produced regardless of their accuracy or statistical significance.

Simulating the urban travel demand is a difficult and a complex process that requires use of tools adequately tailored to fit the process. Wohl (20) was probably correct in stating that numerous time period models are necessary to understand the complex process of estimating urban travel demand. Indeed, a.m. peak hour models, p.m. peak hour models and 24 hour models may all be necessary, for certain conditions and situations, to fully understand and simulate the urban travel patterns. This viewpoint is obviously not shared by many transportation planners. Throughout the research process, numerous discussions between the author and other professionals demonstrated that a wide gap still exists between systems planners and functional planners and designers, and even among systems planners, a wide range of awareness and differing opinions exists on the significance of this gap.

Therefore, a procedure should be established to improve the estimation and application of those elements which help to bring the systems planning and functional planning processes closer together. The importance of coordinating and understanding the assumptions and limitations of these two processes cannot be overemphasized.

This research has demonstrated the need to bridge the above gap and, furthermore, as presented in the next section, a framework based on this thesis has been developed for guiding future research aimed at improving the various elements of the UTPS model.

9.2 A Framework for Future Research

The basic structure and the main activities in the UTPP have been described in Section 2.2. The complex and

continuous nature of the process was delineated and the model building aspects were presented at some length in Section 2.3.

Throughout the past two or three decades, the UTPP has attracted various schools of thought ranging from "map planning" to the so-called Fourth Generation Models. Although the process, in its historical development, appears to demand greater comprehensiveness, recently emphasis has shifted to the so-called Sketch Planning techniques. However, the process continues to be tailored to fit the environment within which it is employed.

In order to make the process adaptable to the changing environment, it is becoming increasingly complex and difficult to manage. New coordinative, administrative and research skills are needed to carry out the process effectively. This points out the need for new tools that can help the individuals participating in the process to better understand it and to make appropriate contributions.

It is therefore, the author's contention that a framework for future research into improving certain elements of the UTPP and in particular, the UTPS model is required. Furthermore, a preliminary framework for such a structure has been developed in the course of this research and is presented in the next few paragraphs.

The UTPP may be viewed as a goal oriented structure with numerous interrelated elements. Each element in turn, may be defined, estimated and applied in several ways within the overall framework. As mentioned earlier, the process is continuously changing to adapt to the changing needs of the users. This change, in many instances, necessitates revisions to the various elements of the

process. In order to guide these revisions, it is proposed that a framework for research be established which will not only assist in the systematic progress of the research, but will also aid in identifying the full range of options available to carry out the research.

The ultimate objective of the UTPS model is to provide demand estimates which are eventually used to design transportation facilities. The various elements of the UTPS model contribute to this objective in varying degrees since the eventual design decision may be more sensitive to one element than to another. For example, errors in the trip assignment step may produce more dramatic impacts on a given link being designed than, say, errors in the formulation of the friction factor curves.

Most elements of the UTPS model may be described in terms of one or more of the following:

1. Trip Purpose p
2. Trip Mode m
3. Geographic Base (Region-wide, District, Zone) i or j
4. Path or Route r
5. Link l
6. Time Period (Peak hour t or 24 hour T)

Furthermore, the basic unit of measurement is obviously the individual trip. Most elements may also be defined or applied at the disaggregate individual i or household j levels.

The proposed framework is presented in Figure 9.1. It calls for the following research activities:

1. definition of the researched element L ,
2. identification of the interaction between the element and the rest of the process.

FIGURE 9.1
A FRAMEWORK FOR FUTURE RESEARCH

3. assessment of the existing approaches to the definition, estimation and application of the element,
4. identification of problem areas and deficiencies,
5. hypothesis formulations,
6. definition of the range of possible options in terms of scale of analysis and resources limitations,
7. development of input data,
8. establishment of appropriate statistical analysis techniques,
9. simulation, validation and selection of the most promising revised approach, and
10. investigation of the cost-effectiveness of the revised approach using a decision theoretic model.

In formulating the hypotheses in step 5 above, the research element L may be described in terms of the following sets of determinants (or independent variables). Some of these are further described in Table 4.1.

- a. Trip Characteristics: These include trip purpose, mode, time period, route and geographic area.
- b. Trip Maker Characteristics: These include household and/or individual socio-economic attributes.
- c. Zonal Characteristics: These refer to origin and destination zonal variables, particularly land use data.
- d. Transportation System Characteristics: These reflect the transport supply and may be expressed in terms of networks attributes.
- e. Policy Constraint Variables: These relate to decisions exogenous to the UTPS model.
- f. Regionwide Characteristics: These reflect overall regional attributes and may be expressed in terms of city size and composition.

Each of the research activities describing the framework may, in turn, be subdivided into numerous interrelated tasks. Of particular importance, however, are those dealing with the statistical simulation (activities 8 and 9) and the investigation of cost-effectiveness of the revised approach (activity 10). Sub-frameworks were developed for these two tasks, as shown in Figures 9.2 and 9.3 respectively.

9.3 Summary

The objective of this research was to develop an improved and cost-effective approach for estimating and applying peak hour ratios in the UTPS model. In order to achieve this objective the following is a brief summary of the basic activities carried out in this research:

1. the central position of the peak hour ratios in the modelling steps of the urban transportation planning process was delineated and the importance of these ratios was demonstrated,
2. the existing practices for estimating and applying peak hour ratios were reviewed and their deficiencies were identified,
3. a number of alternative approaches for peak hour ratios application and estimation were generated (these ranged from the disaggregate household and individual levels to the aggregate zonal and district levels)
4. hypotheses and relationships between peak hour ratios and some socio-economic and transport system variables were derived using actual data obtained from the Regional Municipality of Ottawa-Carleton,
5. statistical techniques such as factor analysis, step-wise regression analysis and category analysis were employed to develop and test the revised approaches,

FIGURE 9.2

STATISTICAL ESTIMATION SUB-FRAMEWORK

FIGURE 9.3

SUB-FRAMEWORK FOR INVESTIGATING
THE COST-EFFECTIVENESS OF RESEARCHING THE ELEMENT L

6. various approaches were recommended for estimating and applying peak hour ratios at various stages of the UTPS model,
7. decision analysis was employed to show the cost-effectiveness of the improved approaches (this was achieved by identifying the trade offs between the cost of experimentation and the expected net gain from experimentation), and
8. a framework was developed which may be used as a preliminary structure for guiding future research into improving elements of the UTPS model.

9.4

Conclusions

Revised approaches for peak hour ratios estimation and application in the urban transportation planning process were developed in this research. These approaches were based on actual data obtained from the Regional Municipality of Ottawa-Carleton.

Even though it was hoped that stronger relationships could be obtained through regression analysis for many peak hour ratios by mode and purpose, unfortunately, it was not possible to develop statistically significant relationships for many of these combinations due to data limitations. Zonal peak hour ratios for work by the transit and auto driver and passenger modes were developed using regression analysis and these were recommended together with the mean peak hour ratios for the other trip purpose/mode combinations.

Because of time and cost limitations, further refinement of the relationships developed was not possible. However, it is believed that substantial improvement to the recommended approaches for peak hour ratios estimation

and application can be made. In the next section, suggestions in that direction are given.

The research has also pointed out that peak hour ratios may be estimated and applied at various stages of the urban transportation planning process. Aggregate zonal and district level relationships were successfully developed through linear regression techniques. Category analysis was employed to estimate peak hour ratios at the disaggregate household and individual levels. Finally, both category analysis and regression analysis were used to analyze peak hour ratios obtained from actual ground counts. On the basis of these analyses it can be deduced that improved approaches for peak hour ratios estimation and application could indeed be developed.

Bayesian statistical decision theory was employed to show the cost effectiveness of the revised approach(es) to peak hour ratios estimation and application. It was concluded that, in many situations, the resources expended in experimentation aimed at improving peak hour ratios estimation and application were justified. Actual case studies demonstrated the application of the decision analysis technique. Four road links within the Regional Municipality of Ottawa-Carleton were used to compare the output of the decision analysis and the actual decision made by the design engineer. The comparison indicated different lane requirements in three of the four cases studied.

Finally, a preliminary structure of a framework for future research was developed based on the actual procedure employed in this thesis. Examination of this framework concluded that it is logical, practical and may aid to structure future research aimed at improving various elements

of the UTPS model.

9.5

Recommendations for Further Research

This research, like many others, probably posed more questions than it actually answered. In order to avoid a lengthy discussion of the topics requiring further investigation and research, a brief outline of these topics in point format is presented.

1. In order to test the stability of the peak hour ratios relationships developed through regression analysis, 1977/1978 data may be used to verify the developed equations. This may be done using the 1977 Telephone Survey which, at the time of writing this thesis, was being validated.
2. Another neglected area of research in the transportation planning literature is that of directional split factors. It is suggested that this important topic be further researched separately or jointly with peak hour ratios.
3. In order to establish the level of reliability of the various approaches for peak hour ratios estimation and application described in Chapter 4, it is suggested that a comparative analysis be performed both at the link level and for the entire network. Such a comparative analysis will indicate at the link level, when model volumes are compared to counts, which approach is the most reliable. Traffic assignments of matrices produced through various applications of peak hour ratios approaches to the entire transport network can establish the range of possible variation in the resultant peak hour volumes as outlined in Figure 7.3.

4. It is recommended that research into 24 hour, p.m. peak hour and a.m. peak hour based models be supported, where feasible, since the cost of simulation is a small portion of the cost of the overall process while the benefits for individual link design or location analysis should undoubtedly off-set the cost of simulation. Furthermore, it is suggested that research into a.m. peak hour ratios (obviously with concentration on work travel) be performed. The proper application of different time period models is becoming increasingly important especially for short range planning and research into this topic should be supported.
5. It is recommended that further research be carried out to take into account the seasonal variation of travel in the UTPS model. The definition of the model output should be reconciled with the time basis of traffic counts used to validate the results.
6. Further research into the variation of peak hour ratios and directional split factors by facility type and by various geographic areas is suggested. This type of research can be performed using the data already developed in this thesis.
7. In the disaggregate household and individual level analysis, it is suggested that the Maximum Likelihood Technique be employed to investigate the peak hour ratios variation as a function of individual or household attributes. The investigation would take a behavioural form similar to that of the logit model in modal split analysis. The dependent variable would be the probability of making a peak hour trip given certain individual attributes in much the same way as the logit model establishes the probability of choosing a particular mode by various users.

8. Further investigation of the use of the equilibrium technique in traffic assignment in conjunction with the revised approaches to peak hour ratios application should be carried out in order to produce as reliable estimate of link peak hour volume as possible.
9. In the decision analysis exercise, the 24 hour matrix of travel is implicitly treated as certain. It is assumed that the 24 hour trip distribution was simulated as accurately as possible and compared to actual 24 hour ground counts. In actuality, this matrix should be treated as uncertain and further investigation of such an assumption is recommended.
10. Case studies used in the decision analysis exercise; all represented highway links. It is suggested to redefine the problem in terms of investment on transit.
11. The left side of the schematic curve in Figure 8.2 represents the cost of congestion resulting from underdesigning the facility while the right side reflects the cost of over investment resulting from overdesigning the facility. The point of minimum system cost is when the design hour volume approaches the actual peak hour volume. It is suggested that research into defining the upper and lower bounds of the optimum design hour volume be made. Within these bounds the actual total system cost curve and the design decision will not change.
12. Further research into applying the prior probability distribution function as a continuous function rather than as a discrete function is suggested.
13. Expansion of the decision analysis programme written in this research is recommended in order to allow for a number of additions such as the development of a multi-stage decision tree, (i.e. to allow for staging);

the analysis for separate components of the total system cost (i.e. capital, maintenance and operating costs); development of separate cost curves for each type of link (i.e. rural, urban and semi-urban) and extension of the alternative actions to cover more than a two, three or four lane facility.

14. Throughout the simulation of peak hour ratios by mode, auto driver and passenger were treated as a single mode. It is recommended that an investigation of treating auto occupancy as a separate component be carried out.
15. It is recommended that an expanded scope of peak hour ratios be investigated which addresses the modal split and the choice of peak hour travel as two related elements of the UTPS model.

REFERENCES

- (1) Ottawa-Carleton Regional Transportation Study (O.C.R.T.S.) Progress Report 1, RMOC, Planning Department, 1972.
- (2) Toronto Area Regional Model Study (T.A.R.M.S.), Trip Generation, Ministry of Transportation and Communications of Ontario (MTC), 1971.
- (3) Transportation Study, Short Term Transportation Forecast for the Regional Municipality of York, MTC, 1974.
- (4) Edens, H.J., Modal Split and Peak Period Model as Applied in the Ottawa-Hull Area Transportation Study (OHATS). DeLeuw Cather and Company of Canada Limited, 1966.
- (5) Ministry of Transportation and Communications of Ontario, Travel Demand Forecast Models, 1976.
- (6) Urban Transportation Planning General Information, U.S. Department of Transportation, Federal Highway Administration, 1972.
- (7) Moore, W.T., Ridel, F.J. and Rodriguez, C.G., An Introduction to Urban Development Models and Guidelines for Their Use in Urban Transportation Planning, U.S. Department of Transportation, Federal Highway Administration, 1975.
- (8) Alonso, W., "The Quality of Data and The Choice and Design of Predictive Models", Special Report 97, Highway Research Board, 1968.
- (9) Hangroves, B. and Rea, J.C., Techniques of Transportation Analysis, Pennsylvania State University, Course Notes.
- (10) Meyer, J.R., Editor, Systems Analysis and Simulation Models, Vol. 2 of Techniques of Transport Planning, The Brookings Institution, 1971.
- (11) Toronto Area Regional Model Study (T.A.R.M.S.), Modal Split, M.T.C., 1970.
- (12) Wohl, M., and Martin, B.V., Traffic System Analysis for Engineers and Planners, McGraw-Hill Book Company 1967.

- (13) Nakkash, T.Z., Activity-Accessibility Models of Trip Generation, Purdue University, 1969.
- (14) Bouchard, R.J. and Pyers, C.E., "Use of Gravity Model for Describing Urban Travel", Travel Patterns, Record 88, Highway Research Board, 1965.
- (15) U.S. Department of Commerce, Bureau of Public Roads, Calibrating and Testing a Gravity Model For any Size Urban Area, U.S. Government Printing Office, 1965.
- (16) Cosmos Corporation, Traffic Assignment, Methods, Applications, Products. For the office of Highway Planning, U.S. Department of Transportation, 1973.
- (17) Proceedings of a Conference held at Williamsburg, Virginia, Urban Travel Demand Forecasting, Highway Research Board, Special Report 143, 1972.
- (18) Charles River Associates, A Disaggregate Behaviour Model of Urban Travel Demand, 1972.
- (19) Golob, T.F., et al, "An Analysis of Consumer Preferences for a Public Transportation System", Transportation Research, Vol. 6, 1972.
- (20) Wohl, M., "A Methodology for Forecasting Peak and Off Peak Travel Volumes, Highway Research Board, Record 322, 1972 .
- (21) D. Hill, L.H. Themore, et al., An Analysis of Urban Travel by Time of Day, Peat, Marwick, Mitchell and Company, NTIS, PB247-776, 1972.
- (22) Shallal, L., 1972 Home Interview Survey, Preliminary Analysis, O.C.R.T.S., M.T.C., Internal Report, 1974.
- (23) Shallal, L., Prospectus, Phase II Transportation Model Development Study, O.C.R.T.S., RMOC, Internal Report, 1977.
- (24) Transportation Study, O.C.R.T.S., Travel Forecasts and Analysis, Report No. 2, RMOC, Planning Department, 1973.

- (25) Roads and Transportation Association of Canada (RTAC) A Review of Trip Generation Analysis Procedures Used in Canadian Urban Transportation Planning Studies, Technical Publication No. 4, 1974.
- (26) Oglesby and Hewes, Highway Engineering, John Wiley and Sons, Inc. 1963.
- (27) Institute of Traffic Engineers (ITE), Transportation and Traffic Engineering Handbook 1976.
- (28) Creighton, Hamburg Planning Consultants, Data Requirements for Metropolitan Transportation Planning, National Cooperative Highway Research Program, Report 120, 1971.
- (29) Ministry of Transportation and Communications of Ontario, Geometric Design Standards for Ontario Highways and Streets, 1965.
- (30) Highway Capacity Manual, Highway Research Board Special Report 87, National Research Council Publication 1328, 1965.
- (31) AASHO, American Association of State Highway Officials, A Policy on Design of Urban Highways and Arterial Streets, 1973.
- (32) DeLeuw Cather Co. of Canada Ltd., A New Procedure for Urban Transportation Planning, prepared for MTC of Ontario, 1972.
- (33) Canadian Facts Ltd., The RMOC/Outaouais Household Interview Origin-Destination Survey, 1972, The Social Survey Research Centre, 1973.
- (34) Bonsall, J., "Flexible Hours and Public Transit in Ottawa", presented at the Annual Conference of RTAC, Toronto, 1974.
- (35) Ministry of Transportation and Communications of Ontario, STPCP, Simplified Transportation Planning Computer Programs, User's Manual, 1977.
- (36) Overall, J.E., and Kleat, C., Applied Multivariate Analysis New York, McGraw-Hill, 1973.

44

OF / DE

- (37) Dempster, A.P., Elements of Continuous Multivariate Analysis, Reading, Mass., Addison-Wesley, 1969.
- (38) Van Der Geer, J., Introduction to Multivariate Analysis for the Social Sciences, San Fransisco, Freeman, 1971.
- (39) Cattell, R.F., Factor Analysis An Introduction to Essentials, Biometrics, Vol. 21, 1965.
- (40) Rummel, R.J., Understanding Factor Analysis, Evanston, Northwestern University Press, 1970.
- (41) Roo, R., Linear Statistical Inference and It's Application, New York, Wiley, 1973.
- (42) Lavalley, I.H., An Introduction to Probability, Decision and Inference, Holt, Rinehart and Winston, Inc., 1970.
- (43) U. S. Bureau of Public Roads, Modal Split, U. S. Department of Commerce, Office of Planning, 1968.
- (44) National Cooperative Highway Research Program, Comparative Analysis of Traffic Assignment Techniques with Actual Highway Use, Report 58, 1968.
- (45) Dafermos, S.C., "An Extended Traffic Assignment Model with Applications to Two-Way Traffic", Transportation Science, Vol. 5, 1971.
- (46) Leblanc, L.J., Mathematical Programming Algorithms for Large Scale Network Equilibrium and Network Design Problems, Ph.D Dissertation, Northwestern University, 1973.
- (47) Nguyen, S., "An Algorithm for the Traffic Assignment Problem", Transportation Science, Vol. 8, 1974.
- (48) Florian, M. and Nguyen S., "An Application and Validation of Equilibrium Trip Assignment Methods", Transportation Science, Vol. 10, 1976.
- (49) Ruiter, E.R., Network Equilibrium Capabilities for the UMTA Transportation Planning System, Cambridge Systematics, Inc., 1974.

- (50) Tobin, R.L., A Convergent Iterative Traffic Assignment Algorithm, GM Research Laboratories, 1975.
- (51) Johnson, W.F., Decision Theoretic Techniques for Information Aquisition in Transportation Planning, M.I.T. Research Report R-69-1, 1969.
- (52) Raiffa, H. and Schlaifer, R., Applied Statistical Decision Theory, The M.I.T. Press, 1968.
- (53) Raiffa, H., Decision Analysis: Introductory Lectures on Choices under Uncertainty, Addison-Wesley, Reading, Mass., 1968.
- (54) Green, P.E., "Bayesian Decision Theory in Pricing Strategy", Journal of Marketing, Vol. 27, 1963.
- (55) Hutchinson, B.G., The Evaluation of Pavement Structural Performance, Ph.D Thesis, University of Waterloo, 1965.
- (56) Manheim, M.L., "Search and Choice in Transport Systems Planning", Highway Research Board, 48th Annual Meeting, 1969.
- (57) Khan, A.M., Regional Transport System Plan Selection Problem: A Decision Theoretic Framework for Policy Decisions, Ph.D. Thesis, University of Waterloo, 1969.
- (58) Burnes, J.A., "Statistical Decision Theory Applied to a Traffic Problem", Traffic Quarterly, 1969..
- (59) Bonsall, J., Regional Transit: A Recent Success Story, Unpublished Report, 1976.
- (60) N. D. Lea's and Associates, Economic Road Improvement Selector Programme, ERI\$₂, Users Manual, 1972.
- (61) SPSS Inc., Statistical Package for the Social Sciences, Second Edition, McGraw-Hill, 1975.
- (62) Shallal, L., Comparative Analysis of 1963, 1972 and 1977 Travel O-D Surveys, RMOC, Internal Report, 1978.

APPENDIX A

LIST OF NOTATIONS AND ABBREVIATIONS

APPENDIX A

LIST OF NOTATIONS AND ABBREVIATIONS

<u>Notation</u>	<u>Description</u>
a	Alternative Actions
AADT	Average Annual Daily Traffic
A	Space of All Possible Alternative Actions
A_i	Attractivity of Zone i
CBD	Central Business District
D.F.	Degrees of Freedom
D.H.V.	Design Hour Volume
D + P	Auto Driver Plus Passenger
D.U.	Dwelling Units
e	Experiment or Option
E	Space of All Possible Experiments Or Options
f	Facility Type
G.L.A.	Gross Leasable Area
H	Peak Hour Ratios
HBB	Home Based Shopping and Personal Business Trips
HBSCH	Home Based School Trips
HBSRO	Home Based Social, Recreational and Other Trips
HBW	Home Based Work Trips
H.H.	Household
H_i	Peak Hour Ratios for the Origin Zone, Origin District, Individual or Household i
H_{ij}	Peak Hour Ratios for Travel from Origin Zone i to Destination Zone j
$H_i^{m,p}$	Peak Hour Ratios for Zone i by Mode m and Trip Purpose p.
H_j	Peak Hour Ratios for the Destination Zone or Destination District
L.O.S.	Level of Service

<u>Notation</u>	<u>Description</u>
HPA1	H_i/H_j , Transit (Appendix G)
HPA2	H_i/H_j , D+P (Appendix G)
HPAM1	Mean of H_i and H_j , i.e. $(H_i + H_j)/2$, Transit (Appendix G)
HPAM2	Mean of H_i and H_j , i.e. $(H_i + H_j)/2$, D+P (Appendix G)
Indiv.	Individual
ℓ	A Link in the Transportation Network
L	An Element of the UTPS Model
L.F.	Labour Force
m	Travel Mode
MTC	Ministry of Transportation and Communications of Ontario
OCRTS	Ottawa-Carleton Regional Transportation Study
NHB	Non-Home Based Trips
O-D	Origin-Destination
P	Trip Purpose
$\bar{P}$	Probability
Pct.	Percent
PK.HR.	Peak Hour
RMOC	Regional Municipality of Ottawa-Carleton
SPSS	Statistical Package for the Social Sciences
SS	Sum of Squares
t	Peak Hour Trips, also True Peak Hour Volume in Decision Analysis
T	24 Hour Trips, also the Space of All Possible True Peak Hour Volumes in Decision Analysis
Tr	Transit
UTPP	The Urban Transportation Planning Process
UTPS	The Urban Transportation Planning System Model
V	Space of all Possible Peak Hour Volume Outcomes From Experiment e
V(e)	Peak Hour Volume Outcomes From Experiment e
VARxxx	Variable Number xxx

APPENDIX B

LISTING OF VARIABLES
WITHIN EACH DATA BANK

- B.1 CONTENTS OF THE SPSS SYSTEM FILES "SHALLAL"
AND "DISTR66" DATA BANKS.
- B.2 CONTENTS OF THE SPSS SYSTEM FILES "HDISAGGR",
"HH72", "ID72" AND "PK72" DATA BANKS.
- B.3 CONTENTS OF THE SPSS SYSTEM FILE "COSTS"
DATA BANK.

APPENDIX B.1

CONTENTS OF THE SPSS SYSTEM FILES
"SHALLAL" AND "DISTR60" DATA BANKS

<u>SPSS VARIABLE NAME *</u>	<u>DESCRIPTION</u>	<u>VARIABLE LABEL</u>
VAR001	POPULATION 1977 BY ZONE	POPULATN(77)
VAR002	TOTAL EMPLOYMENT 1977	TOTEMPLY(77)
VAR003	RETAIL EMPLOYMENT 1977	RETEMPLY(77)
VAR004	DWELLING UNITS 1977	DWLGUNIT(77)
VAR005	CARS 1977	TOTLCARS(77)
VAR006	SECONDARY SCHOOL ENROLMENT 1977	SECSCHEN(77)
VAR007	POST SEC. SCHOOL DORM. 1977	PSSCHDOR(77)
VAR008	SHOPPING CENTRE G. L. A.	SHCTRGLA(77)
VAR009	LABOUR FORCE PART. RATIO 1977	LABFORPT(77)
VAR090	GROSS ACREAGE	GACREAGE(77)
VAR091	FEDERAL EMPLOYMENT 1977	FEDEMPLY(77)
VAR092	GROSS POP. DENSITY PERSON/ACRE 1977	GPDPRACR(77)
VAR093	POPULATION PER DWELLING 1977	POPPRDLG(77)
VAR094	GROSS EMPLOYMENT DENSITY 1977	GSEMPDEN(77)
VAR095	GROSS RESIDENTIAL DENSITY 1977	GSRESDEN(77)
VAR096	CARS PER DWELLING 1977	CARPRDLG(77)
VAR097	NON FEDERAL/NON RETAIL EMPL:1977	NFDNREMP(77)
VAR101	AERIAL DIST. FROM CBD 1977	ARDSTCBD(77)
VAR102	TIME BY TRANSIT FROM CBD 1977	TMTRACBD(77)
VAR103	TIME BY CAR FROM CBD	TMCARCBD(77)
VAR104	PARKING AVAILABILITY	PARKAVAL(77)
VAR105	NET ACREAGE 1977	NACREAGE(77)
VAR106	INCOME	INCOMEZN(77)
VAR107	LABOUR FORCE	LABFORCE(77)
VAR108	FINISHING TIME FROM 1531-1545	F1531X45(77)
VAR109	FINISHING TIME FROM 1546-1600	F46X1600(77)
VAR110	FINISHING TIME FROM 1601-1615	F1601X15(77)
VAR111	FINISHING TIME FROM 1616-1630	F1616X30(77)
VAR112	FINISHING TIME FROM 1631-1645	F1631X45(77)
VAR113	FINISHING TIME FROM 1646-1700	F46X1700(77)
VAR114	FINISHING TIME FROM 1701-1715	F1701X15(77)
VAR115	FINISHING TIME FROM 1716-1730	F1716X30(77)
VAR116	FINISHING TIME FROM 1731-1745	F1731X45(77)
VAR117	FINISHING TIME FROM 1756-1800	F46X1800(77)
VAR118	FINISHING TIME FROM 1801-1815	F1801X15(77)
VAR119	FINISHING TIME FROM 1816-1830	F1816X30(77)
VAR121	NET POPULATION DENSITY 1977	NTPOPDEN(77)
VAR122	NET RESIDENTIAL DENSITY 1977	NTRESDEN(77)
VAR123	NET EMPLOYMENT DENSITY 1977	NTEMPDEN(77)
VAR124	ZONAL ATTRACTION 1977	ZONATTRA(77)
VAR325	CAR INDEX OF ATTRACTIVITY 1977	CARINDEX(77)
VAR326	TRANSIT INDEX OF ATTRACTIVITY	TRAINDEX(77)

* These names were given arbitrarily to the SPSS Programme. They are listed in two sets: the 1977 data on this page and the 1972 data on the next page.

<u>SPSS VARIABLE NAME</u>	<u>DESCRIPTION</u>	<u>VARIABLE LABEL</u>
VAR125	CAR INDEX OF ATTRACT.	CARINDEX
VAR126	TRANSIT INDEX OF ATTRACT.	TRAINDEX
VAR127	HJ D AND P HSR0	HJDPHSRO
VAR128	HI D AND HSR0	HIDPHSRO
VAR129	HJ D AND P HSCH	HJDPHSCH
VAR130	HI D AND P HSCH	HIDPHSCH
VAR131	HJ TRANS. HW	HJTRANHW
VAR132	HI TRANS. HW	HITRANHW
VAR133	HJ TRANS. HSPB	HJTRHSPB
VAR134	HI TRANS. HSPB	HITRHSPB
VAR135	HJ D AND P HW	HJDPHWRK
VAR136	HI D AND P HW	HIDPHWRK
VAR137	HJ D AND P HSPB	HJDPHSPB
VAR138	HI D AND P HSPB	HIDPHSPB
VAR139	HJ D AND P NHB	HJD&PNHB
VAR140	HI D AND P NHB	HID&PNHB
VAR141	HJ TRANS. NHB	HJTRANHB
VAR142	HI TRANS. NHB	HITRANHB
VAR143	HJ TRANS. HSR0	HJTRHSRO
VAR144	HI TRANS. HSR0	HITRHSRO
VAR145	HJ TRANS. HSCH	HJTRHSCH
VAR146	HI TRANS. HSCH	HITRHSCH
VAR201	POPULATION 1972	POPULATN
VAR202	TOTAL EMPLOYMENT	TOTEMPLY
VAR203	TOTAL RETAIL EMPLOYMENT	RETEMPLY
VAR204	DWELLING UNITS	DWLGNIT
VAR205	CARS	TOTLCARS
VAR206	SEC. SCHOOL ENROLMENT	SECSCHEN
VAR207	POST SEC. SCHOOL DORM.	PSSCHDOR
VAR208	SHOPPING CENTRE G. L. A.	SHCTRGLA
VAR209	LABOUR FORCE PART. RATIO	LABFORPT
VAR302	TIME BY TRANSIT FROM CBD 1972	TMTRACBD
VAR303	TIME BY CAR FROM CBD	TMCARCB
VAR304	PARKING AVAILABILITY	PARKAVAL
VAR305	NET ACREAGE	NACREAGE
VAR306	INCOME	INCOMEZN
VAR307	LABOUR FORCE	LABFORCE
VAR308	FINISHING TIME 1531-1545	F1531X45
VAR309	FINISHING TIME 1546-1600	F46X1600
VAR310	FINISHING TIME 1601-1615	F1601X15
VAR311	FINISHING TIME 1616-1630	F1616X30
VAR312	FINISHING TIME 1631-1645	F1631X45
VAR313	FINISHING TIME 1646-1700	F46X1700
VAR314	FINISHING TIME 1701-1715	F1701X15
VAR315	FINISHING TIME 1716-1730	F1716X30
VAR316	FINISHING TIME 1731-1745	F1731X45
VAR317	FINISHING TIME 1746-1800	F46X1800
VAR318	FINISHING TIME 1801-1815	F1801X15
VAR319	FINISHING TIME 1816-1830	F1816X30
VAR321	NET POPULATION DENSITY	NTPOPDEN
VAR322	NET RESIDENTIAL DENSITY	NTRESDEN
VAR323	NET EMPLOYMENT DENSITY	NTEMPDEN
VAR324	ZONAL ATTRACT.	ZONATTRA

APPENDIX B.2

CONTENTS OF THE SPSS SYSTEM FILES
"HDISAGGR", "HH72", "ID72" AND "PK72" DATA BANKS

<u>VARIABLE LABEL</u>	<u>DESCRIPTION</u>
VAR001	ADDRESS ZONE
VAR002	H. H. NUMBER
VAR003	I. D. PERSON NUMBER
VAR004	LANGUAGE
VAR005	NUMBER OF VEHICLES PER H.H.
VAR006	PERSONS PER H.H.
VAR007	PERSONS OVER 5 PER H.H.
VAR008	H.H. TYPE
VAR009	SEX
VAR010	AGE
VAR011	EDUCATION
VAR012	DRIVERS LICENSE
VAR013	OCCUPATION
VAR014	FULL TIME EMPLOY. OCCUP.
VAR015	ORIGIN ZONE
VAR016	ORIGIN TIME HRS.
VAR017	ORIGIN TIME MINUTES
VAR018	ORIGIN AM PM CODE
VAR019	MIN. WALKED AT START
VAR020	FROM ESTABLISHMENT TYPE
VAR021	TO ESTABLISHMENT TYPE
VAR022	DESTINATION ZONE
VAR023	DEST. TIME HRS
VAR024	DEST. TIME MINUTES
VAR025	DEST. AM PM CODE
VAR026	MIN. WALKED AT END
VAR027	ANNUAL INCOME CODE
VAR028	H. H. INDICATOR FIRST 1
VAR029	INDIV. INDICATOR FIRST 1
VAR030	EXPANSION FACTORS
VAR031	PK. HR. FLAG
VAR032	MODE RECODED
VAR033	PURPOSE RECODED
VAR034	H TOT H.H.
VAR035	H TOT INDIV
VAR036	H H.H. HW DP
VAR037	H H.H. HSPB DP
VAR038	H H.H. HSRO DP
VAR039	H H.H. HSCH DP
VAR040	H H.H. NHB DP
VAR041	H H.H. HW TRANSIT
VAR042	H H.H. HSPB TRANSIT
VAR043	H H.H. HSRO TRANSIT
VAR044	H H.H. HSCH TRANSIT
VAR045	H H.H. NHB TRANSIT

APPENDIX B.3

CONTENTS OF THE SPSS SYSTEM FILE
"COSTS" DATA BANKVARIABLE LABELDESCRIPTION

VAR001	LINK NUMBER
VAR002	NUMBER OF LANES
VAR003	URBAN 2 OR SEMI URBAN 3
VAR004	TOTAL COSTS IN 1000 DOLLARS
VAR005	LENGTH IN MILES
VAR006	D.H.V.
VAR007	TOTAL COSTS IN THOUSANDS PER MILE
VAR008	LENGTH IN MILES X D.H.V.
VAR009	TOTAL COSTS IN THOUSANDS PER VEH. MILE
VAR006S	D.H.V. SQUARED
VAR006C	D.H.V. CUBED

APPENDIX C

LISTING OF A PROGRAMME TO CALCULATE
CAR AND TRANSIT ACCESSIBILITY INDEXES

APPENDIX C

LISTING OF A PROGRAMME TO CALCULATE
CAR AND TRANSIT ACCESSIBILITY INDEXES

This appendix briefly describes the procedure used in the development of zonal car and transit accessibility indexes. These two independent variables were described in Section 5.3.

The two accessibility indexes were derived as follows:

$$I_{ci} = \frac{A_i}{\sum_{j=1}^n A_j \cdot r_{ijc}} \quad \text{and}$$

$$I_{ti} = \frac{A_i}{\sum_{j=1}^n A_j \cdot r_{ijt}}$$

Where

I_{ci} and I_{ti} are accessibility indexes of Zone i by car and transit respectively,

A_i is the attractivity of Zone i measured in terms of total person trips generated by mode m and for trip purpose p ,

(i.e. $A_j = \sum_m \sum_p A_j^{m,p}$) and

r_{ijc} and r_{ijt} are travel impedances from Zone i to Zone j by car and transit developed from the RMOC road and transit networks respectively.

A listing of the computer programme employed is given on the next page.

APPENDIX C

LISTING OF A PROGRAMME TO CALCULATE
CAR AND TRANSIT ACCESSIBILITY INDEXES

```

...//FORTRAN EXEC FORTGCLG,REGION.GO=200K,TIME.GO=(5,0)
//FORT.SYSIN DD *
      INTEGER*4 OUTPUT(300,6),DATA(300)
      DIMENSION P(300,4)
      DATA C/1200*0./
      INTEGER*2 RECNO,RC7(300),RT7(300)
      REWIND 4
      WRITE(6,220)
220   FORMAT('1',*ZONE,OUTPUT1 OUTPUT2 OUTPUT3 OUTPUT4 OUTPUT5 OUTPUT6
      1   *)
C
C      GENERATE ATTRACTIVITY COEFFICIENTS
C
      REWIND 1
      DO 50 I=1,1000
      READ(1,20,END=60)RECNO,(OUTPUT(I,J),J=1,2)
      FORMAT(2X,A3,8(100X),194X,18,14(100X),*92X,18,14(100X),8X)
20   CONTINUE
50   CONTINUE
C
C      COMPUTE DENOMINATOR
C
60   REWIND 1
      REWIND 2
      DO 70 I=1,300
      READ(2,21,END=75)RECNO,(RC7(IN),IN=1,300)
      FORMAT(2X,A2,50(6A21))
21   WRITE(4,1)(RC7(IN),IN=1,300)
      FORMAT(100(316))
1   REWIND 4
      READ(4,1)(DATA(IN),IN=1,300)
      REWIND 4
      DO 65 J=1,300
      D(1,1)=D(1,1)+FLOAT(OUTPUT(I,1))*FLOAT(DATA(J))*0.08
      D(1,2)=D(1,2)+FLOAT(OUTPUT(I,1))*FLOAT(DATA(J))
65   CONTINUE
70   CONTINUE
75   REWIND 2
      REWIND 3
      DO 80 I=1,300
      READ(3,21,END=85)RECNO,(RT7(IN),IN=1,300)
      WRITE(4,1)(RT7(IN),IN=1,300)
      REWIND 4
      READ(4,1)(DATA(IN),IN=1,300)
      REWIND 4
      DO 76 J=1,300
      D(1,3)=D(1,3)+FLOAT(OUTPUT(I,2))*FLOAT(DATA(J))*1.25
      D(1,4)=D(1,4)+FLOAT(OUTPUT(I,2))*FLOAT(DATA(J))
76   CONTINUE
80   CONTINUE
85   DO 200 I=1,300
      IF(D(1,2).EQ.0.) GO TO 610
      OUTPUT(1,3)=(OUTPUT(1,1)*(10**4 ))/(D(1,2)/100000)
      GO TO 620
610   OUTPUT(1,3)=0
620   IF(D(1,4).EQ.0.) GO TO 630
      OUTPUT(1,4)=(OUTPUT(1,1)*(10**4 ))/(D(1,4)/100000)
      GO TO 640
630   OUTPUT(1,4)=0
640   IF(D(1,1).EQ.0.) GO TO 650
      OUTPUT(1,5)=(OUTPUT(1,2)*(10**4 ))/(D(1,1)/100000)
      GO TO 660
650   OUTPUT(1,5)=0
660   IF(D(1,3).EQ.0.) GO TO 670
      OUTPUT(1,6)=(OUTPUT(1,2)*(10**4 ))/(D(1,3)/100000)
      GO TO 680
670   OUTPUT(1,6)=0
680   CONTINUE
      WRITE(6,700)1,(OUTPUT(1,J),J=1,6)
      FORMAT(' ',14,6(18))
700   PUNCH 230,1,(OUTPUT(1,IN),IN=3,6)
      FORMAT(' ',14,4(18))
230   CONTINUE
200   WRITE(6,300)
300   FORMAT(' ',*** END OF JOB ***)
      STOP
      END
**
//GO.FT01F001 DD DSN=RMOC.PROJECT.DATA,BANK,DISP=(OLD,PASS),
// UNIT=(2400,,DEFF),
// VOL=SER=J00061,LAUFL=(2,,,IN),
// DCB=(RECFM=FB,LRECL=4005,BLKSIZE=4005)
//GO.FT02FJ01 DD DSN=RMCC.SPB0017,DISP=(OLD,PASS),UNIT=AF=FT01F001,
// VOL=SER=000075,LAUFL=(1,,,IN),
// DCB=(RECFM=FB,LRECL=604,BLKSIZE=604)
//GO.FT03FJ01 DD DSN=RMCC.SPB00174,DISP=(OLD,PASS),UNIT=AF=FT01F001,
// VOL=SER=000077,LAUFL=(1,,,IN),
// DCB=(RECFM=FB,LRECL=604,BLKSIZE=3624)
//GO.FT04F001 DD DSN=C&TFMP,DISP=(NEW,KFEP),UNIT=SYSCA,
// DCB=(RECFM=FB,LRECL=1800,BLKSIZE=1800),
// SPACE=(1200,(10,1))
//GO.SYSIN DD DUMMY

```


APPENDIX D.

BRIEF LISTING OF COMPUTER PROGRAMMES
USED IN THE RESEARCH

D.1. LISTING OF M.T.C./T.P.S. COMPUTER
PROGRAMMES USED IN THE RESEARCH.

D.2. LISTING OF SPSS PROCEDURES USED IN
THE RESEARCH

APPENDIX D.1

LISTING OF M.T.C./T.P.S.* COMPUTER PROGRAMMES
USED IN THE RESEARCH

<u>Program</u>	<u>Program Title</u>
TR02123	O/D Survey Record Manipulation
TR02124	Zone Record Conversion
TR02311	Build/Update Traffic Data Bank
TR02451	Transit Network Program
TR02452	Transit Path Finder
TR02453	Minimum Path Summary Program
TR02455	Trip Tables/Skim Trees Formatter
TR02464	Merge O/D Survey Record
TR02467	General Purpose Summary Program
TR02501	Print Trip Table
TR02504	Trip Table Balancing - Fratar
TR02505	Trip End Summary
TRQ2506	Road Network Build/Update
TR02509	Skim Trees Builder
TR02512	Modification of Trip Tables
TR02513	Trip Table Compressor
TR02515	Trip Table Manipulator
TR02533	Trip Table Builder
TR02534	Matrix Addition
TR02535	Matrix Printout
TR03301	Land Use Data Bank and Trip Generation
TR03302	Road Network Generation and Assignment
TR03305	Matrix Manipulation
TR03306	Traffic Surveys Analysis and O/D Surveys

* Ministry of Transportation and Communications of Ontario,
 Transportation Planning System.

APPENDIX D.2

LISTING OF SPSS* PROCEDURES USED IN THE RESEARCH

<u>Procedure</u>	<u>Procedure Description</u>
AGGREGATE	. Aggregates Sub-Population Statistics and unites them to a Raw Output File.
BREAKDOWN	. Describes Sub-Populations and Provides a One-Way Analysis of Variance and Test of Linearity.
CANCORR	. Performs Canonical Correlation and Produces Canonical Variables.
CONDESCRIPTIVE	. Provides Descriptive Statistics for Variables
CROSSTABS	. Crosstabulates Variables Optionally. Controlling for Other Variables.
FREQUENCIES	. Provides Frequency Tables and Descriptive Statistics.
REGRESSION	. Performs Simple or Multiple Regression with a Stepwise Option
SCATTERGRAM	. Produces Bivariates Plots
FACTOR	. Performs Five Different Types of Factor Analysis and Produces Factor Scores
PEARSON CORR.	. Calculates Pearson Product-Moment Correlations and significance levels.

* SPSS Statistical Package for the Social Sciences

APPENDIX E

LISTING OF THE TRIP-IN-MOTION PROGRAMME

APPENDIX E

LISTING OF THE TRIP-IN-MOTION PROGRAMME

```

FORTRAN IV G LEVEL . 21          MAIN          DATE = 78031

C LOUIS SHALLAL, TRIP IN MOTION PROJECT
C SHALLAL T1MP-1972 H.I. SURVEY DATA
0001  INTEGER TIME(300), TRIP(300), TOTAL(300), START, STUDY
0002  INTEGER OTIME, OAMP, OTIME, DAMPM

C CONTROL CARD - START TIME, STUDY DURATION, INTERVAL DU
C
0003  READ(5,10) START, STUDY, INTDUR
0004  10 FORMAT(21A,12)

C INITIALIZE TRIP VECTOR, TOTAL TRIPS VECTOR
C
0005  DO 12 K=1,300
0006  TRIP(K)=0
0007  TOTAL(K)=0
0008  12 CONTINUE
0009  MINUTE=0
0010  NINT=(60/INTDUR)*STUDY+1
0011  NSTART=START
0012  LINES=59
0013  NTOTAL=0
0014  INVAL=0
0015  NZERO=0
0016  NVALID=0

C INITIALIZE TIME VECTOR - 03:00 HRS TO 27:00 HRS
C
0017  J=0
0018  DO 20 I=1,NINT
0019  TIME(I) = 0
0020  MINUTE = INTDUR*(J)
0021  IF (MINUTE.NE.60) GO TO 15
0022  MINUTE = 0
0023  NSTART = NSTART+100
0024  J=0
0025  15 TIME(I) = NSTART + MINUTE
0026  J=J+1
0027  20 CONTINUE

C MAIN PROGRAM LOOP STARTS HERE
C
0028  50 READ(4,55,ERR=99,END=100) OTIME, OAMP, OTIME, DAMPM
0029  55 FORMAT(8A,X,14,11,10X,14,11,96X)
0030  IF(OTIME.EQ.0) GO TO 95
0031  IF(OTIME.EQ.0) GO TO 95

C CONVERT TIMES TO RANGE OF 03:00 HRS TO 27:00 HRS
C
0032  IF(OAMP.EQ.2) GO TO 56
0033  IF(OTIME.GT.1200) OTIME=OTIME+1200
0034  IF(OTIME.LT.0301) OTIME=OTIME+2400
0035  GO TO 57
0036  56 IF(OTIME.LE.1200) OTIME=OTIME+1200
0037  57 IF(DAMP.EQ.2) GO TO 58
0038  IF(OTIME.GT.1200) OTIME=OTIME+1200
0039  IF(OTIME.LT.0301) OTIME=OTIME+2400
0040  GO TO 59
0041  58 IF(OTIME.LE.1200) OTIME=OTIME+1200

```

```

C
C CHECK FOR INVALID TRIP TIMES.
0042 59 NDUR = DTIME-OTIME
0043 IF(NDUR.LT.0)GOTO 90
0044 IF(NDUR.GT.400)GOTO 90
0045 IF(OTIME.LT.300)GO TO 90
0046 IF(OTIME.GT.2700)GOTO 90
0047 IF(DTIME.LT.300)GO TO 90
0048 IF(DTIME.GT.2700)GOTO 90
0049 NVALID = NVALID + 1

C
C LOCATE ORIGIN TIME.
0050 DO 70 I=1,NINT
0051 IJ=1
0052 IF(OTIME.NE.TIME(I))GOTO 60
0053 TRIP(I)=TRIP(I)+1
0054 GO TO 80
0055 60 IF(OTIME.LT.TIME(I))GOTO 75
0056 70 CONTINUE
0057 GO TO 99
0058 75 TRIP(I-1)=TRIP(I-1)+1
0059 IJ=I-1

C
C LOCATE DESTINATION TIME. START AT ORIGIN TIME + 5 MIN
0060 IJ=IJ+1
0061 DO 85 J=IJ,NINT
0062 IF(DTIME.GT.TIME(J))GOTO 84
0063 TRIP(J)=TRIP(J)+1
0064 GO TO 99
0065 84 TRIP(J)=TRIP(J)+1
0066 85 CONTINUE
0067 GO TO 99

C
C REJECT INVALID RECORDS
0068 90 INVAL=INVAL+1
0069 LINES=LINES+1
0070 IF(LINES.NE.60)GOTO 93
0071 LINES=1
0072 WRITE(6,91)
0073 91 FORMAT('1',30X,'***** INVALID RECORDS *****')
0074 WRITE(6,92)
0075 92 FORMAT('0',8X,'ORIGIN AMPM DEST AMPM')
0076 93 IF(DTIME.GT.2400)OTIME=OTIME-2400
0077 IF(OTIME.GT.2400)DTIME=DTIME-2400
0078 WRITE(6,94)OTIME,DAMPM,DTIME,DAMPM
0079 94 FORMAT('0',10X,14.5X,11.6X,14.5X,11)
0080 GO TO 99

C
C COUNT NO-TRIP RECORDS
0081 95 NZERO = NZERO + 1

C
C COUNT ALL RECORDS. MAIN LOOP ENDS HERE.
0082 99 NTOTAL = NTOTAL + 1
0083 GO TO 50

C
C ACCUMULATE NO. OF TRIPS PER HOUR
0084 DO 110 N=1,12
0085 TOTAL(12)=TOTAL(12)+TRIP(N)
0086 110 CONTINUE
0087 DO 120 N=13,NINT
0088 I=N-12
0089 J=N-1
0090 TOTAL(N)=TOTAL(J)+TRIP(N)-TRIP(I)
0091 120 CONTINUE

C
C FIND TWO RUSH HOURS (6-11:00 AM AND 2:30-7:00 PM)
0092 MAX1=0
0093 MAX2=0
0094 DO 150 I=36,66
0095 MAX1=MAX0(MAX1,TOTAL(I))
0096 150 CONTINUE
0097 DO 160 I=138,192
0098 MAX2=MAX0(MAX2,TOTAL(I))
0099 160 CONTINUE

C
C PRINT RESULTS
0100 WRITE(6,200)
0101 200 FORMAT('1',10X,'TIME',10X,'TRIPS',10X,'TOTAL')
0102 DO 210 K=1,NINT
0103 IF(TIME(K).GT.2400)TIME(K)=TIME(K)-2400
0104 WRITE(6,205)TIME(K),TRIP(K),TOTAL(K)
0105 205 FORMAT('11X,14.10X,14.10X,15)
0106 210 CONTINUE
0107 WRITE(6,215)MAX1,MAX2
0108 215 FORMAT('0',10X,' MAX1 = ',15.11X,' MAX2 = ',15)
0109 WRITE(6,220)INVAL,NZERO,NVALID,NTOTAL
0110 220 FORMAT('0',10X,' INVALID = ',15.5X,' NO TRIP = ',15,
C5X,' VALID = ',15.5X,' TOTAL = ',15)
0111 STOP
0112 END

```

APPENDIX F
FACTOR ANALYSIS

APPENDIX F

FACTOR ANALYSIS*

F.1 Introduction

The single most distinctive characteristic of factor analysis is its data reduction capability. Given an array of correlation coefficients for a set of variables, the factor analytic technique indicates whether an underlying pattern of relationship exists such that the data may be reduced to a smaller set of factors or components that may be taken as source variables accounting for the observed interrelations in the data.

Factor analysis can be used to:

- a. explore or detect patterns or data variation in order to reduce data requirements,
- b. confirm significant factors, and
- c. develop new indicies which may be used as new variables in later analysis.

The first use is the most common.

F.2 Types of Factor Analysis

There are basically three steps in factor analysis:

1. preparation of the correlation matrix,
2. extraction of the initial factors (the exploration of possible data reduction),
3. the rotation to a terminal solution (the search for simple and interpretable factors).

* References (39, 40 and 61)

Major options at each stage or step may be summed up by three dichotomies:

R type vs Q type factor analysis in Step 1, defined vs inferred factors in Step 2, and orthogonal vs oblique in Step 3.

F.2.1 Preparation of Correlation Matrix

Granted that some type of correlation matrix is used as the basic input to factor analysis (in this thesis the actual raw data were used), one can look at the correlation between variables and perform the so-called R-Factor analysis. Alternatively one can look at the correlation matrix of units (in this thesis zones) and do the so-called Q-Factor analysis.

F.2.2 Extraction of Initial Factors

The second step in the analysis is to explore the data reduction possibilities by constructing a set of new variables on the basis of the interrelations shown in the data. In order to achieve this the new variables may be either defined as an exact mathematical transformation of the original data or inferences may be made about the structuring of variables and their source of variation. The former approach which uses defined factors is called Principal Component Analysis. Whether factors are exactly defined or are inferred, initial factors are usually extracted in such a way that one factor is independent from the other, that is all factors are orthogonal.

a. Defined Factors

The Principal Component Analysis is a method of transforming a given set of variables into a new set of composite variables or principal components that are orthogonal (i.e. unrelated) to each other. In this regard, it is desired to find out what is the best linear combination of variables that would account for more of the variance in the data as a whole than for any other linear combination of variables. The first component analysis will result with a certain residual. The second component analysis, which is orthogonal to the first, attempts to account for the residual variance remaining after the effect of the first component is removed from the data. Subsequent components are defined similarly until all the variance in the data is exhausted. Unless at least one variable is perfectly determined by the rest of the variables in the data, the principal component solution requires as many components as there are variables. Since the first few components will explain most of the variance they will be considerably less than the original number of variables and the analyst will tend to use these components as his new factors.

b. Inferred Factors

Classical factor analysis is based fundamentally on the premise that the observed correlation is mainly due to some underlying regularity in the data. It is assumed that the observed variable is influenced by various determinants, some of which are shared by other variables in the set while others are not shared by any other variable. The part of a variable that is influenced by the shared determinants is usually called "common".

the other part is called unique. Under this assumption the unique part of a variable does not contribute to relationships among variables. This implies that the observed correlation must be the result of the correlated variables sharing some of the common determinants. Thus, one concludes that those assumed common determinants will not only account for all or most of the observed relations in the data but will also be smaller in number than the variables. Therefore, if there is any correlation between two variables, say J and K, it is assumed to be due to the common factors. And if the common factors are assumed to be orthogonal (i.e. unrelated) an equation can be written showing the relation between these two variables in terms of a number of factors. If there is only a single common factor the correlation can be written in terms of that factor and if one were to control that particular factor the partial correlation between J and K would be 0.

Factor analysis therefore can be thought of as a technique by which a minimum number of hypothetical variables (factors) are specified in such a way that after controlling for these hypothetical variables, all the remaining (partial) correlation between the variables would become 0. The determination of communalities between variables remains one of the most difficult and ambiguous tasks in factor analysis. One of the main characteristics that differentiates one method of factoring from another is the procedure used for estimating the communalities. Whether a solution employs defined factors depends on whether the existence of unique variance is assumed. More practically one differs from the other depending on whether the correlation matrix is altered before performing

the factor analysis. All inferred factor techniques replace the main diagonals of the correlation matrix with communality estimates before factoring.

F.2.3 Rotation of Factors Into Terminal Factors

Regardless of whether factors are defined or inferred, the exact configuration of the factor structure is not unique, one factor solution can be transformed into another without violating the basic assumptions or the mathematical properties of a given solution. This is unfortunate because there is no unique and generally accepted best solution. On the other hand not all statistical factor solutions are equally meaningful in theoretical terms.

To achieve simpler and theoretically more meaningful factor patterns the analyst must choose between an orthogonal method or an oblique rotational method. The difference is that orthogonal factors are uncorrelated while oblique factors may be correlated. Table F.1 summarizes the types of factor analysis. Table F.2 summarizes the basic matrices of factor analysis that are produced depending on whether the solution is orthogonal or oblique.

F.3 Explanation of the Varimax Rotated Factor Matrix

A typical varimax rotated factor matrix is given in Table 6.11. Since it is an orthogonal factor matrix it stands for both a pattern and a structure matrix. That is, the coefficients in the table represent both regression weights and correlation coefficients. The loadings or number in a given row represent regression coefficients of factors used to describe a given variable.

TABLE F.1

TYPES OF FACTOR ANALYSIS

<u>Steps in Factor Analysis</u>	<u>Major Option</u>	<u>Key Reference to Literature</u>
1. Preparation of correlation matrix	a. Correlation between variables	a. R factoring
	b. Correlation between units	b. Q factoring
2. Extraction of initial factors	a. Defined factors	a. Principal-component solution
	b. Inferred factors	b. Classical - or common-factor solution
3. Rotation to terminal factors	a. Uncorrelated factors	a. Orthogonal factors or rotation
	b. Correlated factors	b. Oblique factors or rotation

TABLE F.2

BASIC MATRICES OF FACTOR ANALYSIS

	Orthogonal Solution	Oblique Solution
Basic Data	Same Correlation Matrix	
Initial Factors	Same Orthogonal-Factor Matrix	
Terminal Factors	Factor Matrix	a. Pattern Matrix b. Structure Matrix
Correlation Between Factors (No Matrix Output)		Factor Correlation Matrix
Factor-Estimate Matrix	Factor-Estimate Matrix	Factor Estimate Matrix

If each variable loads heavily on a single factor its factorial complexity is one. If on the other hand, a variable loads moderately on two or more factors its factorial complexity is 2 or more. The meaning of such a variable is no longer simple. It measures more than one theoretical dimension.

The importance of a given factor for a given variable can be exactly expressed in terms of the variance in the variable that can be accounted for by the factor. This is equivalent to the square of the coefficient in the varimax matrix. The square of the coefficient is a number which represents the percent of the total variance of that variable which is accounted for by that particular factor. Likewise, the variance of a variable accounted for by all factors is given by the sum of squares of the respective factor loadings, (for orthogonal solutions only). The total variance of a variable accounted for by the combination of all common factors is usually referred to as the "communality" of the variable. This value indicates the amount of the variance that is shared by at least one other variable in the set. The complement of communality represents the proportion of the unique variance of a variable which is not accounted for by the common factors or by any variable in the set. Note as mentioned earlier, that the coefficients in the varimax matrix also represent correlation coefficients between the factors and variables. Finally one may compare any two variables and determine the source of common variation or correlation.

The interested reader may consult references (39,40) for a further description of the factor analysis technique.

APPENDIX G

RESULTS OF ZONAL LINEAR AND
NON-LINEAR REGRESSIONS NOT IN THE MAIN TEXT

APPENDIX G

RESULTS OF ZONAL LINEAR AND NON-LINEAR REGRESSIONS
NOT IN THE MAIN TEXT

Equation Number	Dependent* Variable	Equation
1	HJTRANHW	$-0.120 (\text{NETMPDEN}) + 0.512 (\text{ZONATTRA}/1000)^2 - 1250/(\text{TMTRACBD}) + 446.0$
2	HJTRANHW	$2.6 (\text{LABFORCE})^{\frac{1}{2}} - 0.07 (\text{TRAINDEX}) + 63.8/(\text{ZONATTRA}) - 647.0/(\text{NTRESDEN}) + 577.9$
3	HJTRANHW	$-0.07 (\text{TRAINDEX}) + 1.2 (\text{POPULATN})^{\frac{1}{2}} - 287.9/(\text{NTRESDEN}) + 617.8$
4	HJTRANHW	$-0.04 (\text{TOTEMPLY}) + 0.012 (\text{POPULATN}) + 0.061 (\text{NTRESDEN}) + 357.7$
5	HITRANHW	$-0.03 (\text{POPULATN}) + 0.537 (\text{ZONATTRA}/1000)^2 + .03 (\text{NTEMPDEN}) - 0.19 (\text{NTRESDEN}) + 396.5$
6	HITRANHW	$7.03 (\text{TOTEMPLY})^{\frac{1}{2}} - 0.29 (\text{NTRESDEN}) - 5.5 (\text{RETEMPLY})^{\frac{1}{2}} + 190.6$
7	HITRANHW	$-0.037 (\text{POPULATN}) + 0.012 (\text{ZONATTRA}) + 0.029 (\text{NTEMPDEN}) + 0.731 (\text{TMTRACBD}) + 213.0$
8	HITRANHW	$0.062 (\text{TOTEMPLY}) - 0.026 (\text{POPULATN}) - 0.04 (\text{CARINDEX}) - 0.22 (\text{NTRESDEN}) - 0.098 (\text{RETEMPLY}) + 856.7$

* All values of the dependent variable are in thousands and should be divided by ten (10), to represent the peak hour ratios in percentages.

Equation Number	Multiple R	Standard Error	F- Regression	Degrees Of Freedom-Residual	F-Table ($\alpha=.05$) ($\alpha=.01$)		Significance At 95% At 99%	
							At 95%	At 99%
1	0.53	10.85	13.53	103	2.70	4.00	Yes	Yes
2	0.48	10.79	8.6	116	2.45	3.48	Yes	Yes
3	0.41	11.17	7.8	117	2.68	3.95	Yes	Yes
4	0.42	11.56	7.66	106	2.70	4.00	Yes	Yes
5	0.70	12.94	13.74	57	2.54	3.7	Yes	Yes
6	0.73	12.34	21.7	58	2.76	4.13	Yes	Yes
7	0.72	12.64	15.06	57	2.55	3.7	Yes	Yes
8	0.76	12.08	14.3	53	2.4	3.3	Yes	Yes

Equation Number	Dependent Variable	Equation
9	HITRANHW	$10 (\text{LABFORCE}) + 0.02 (\text{ZONATTRA}) - 0.05 (\text{TRAINDEX}) + 584.5$
10	HJDPHWRK	$0.11 (\text{POPULATN}) - 30364.08 / (\text{TMTRACBD}) - 0.535 (\text{ZONATTRA} / 1000)^2 + 0.38 (\text{INCOMEZN} / 1000)^2 - 0.057 (\text{NTEMPDEN}) + 355.41$
11	HJDPHWRK	$2.4 (\text{POPULATION}) - 0.05 (\text{TRAINDEX}) + 358.7$
12	HJDPHWRK	$4.2 (\text{LABFORCE}) - 0.05 (\text{TRAINDEX}) + 371.4$
13	HJDPHWRK	$0.028 (\text{POPULATN}) - .017 (\text{ZONATTRA}) - 30.16 (\text{PARKAVAL}) + 374.29$
14	HJDPHWRK	$-0.078 (\text{TOEMPLY}) + 0.019 (\text{POPULATN}) - 0.33 (\text{NTRESDEN}) + 310.7$
15	HJDPHWRK	$0.082 (\text{LABFORCE}) - 0.018 (\text{ZONATTRA}) - 0.46 (\text{NTRESDEN}) + 299.9$
16	HJDPHWRK	$-0.029 (\text{POPULATN}) + 0.411 (\text{ZONATTRA} / 1000)^2 + 0.026 (\text{NTEMPDEN}) + 284.1$
17	HJDPHWRK	$-0.035 (\text{POPULATN}) + 0.15 (\text{ZONATTRA}) + 0.15 (\text{NTRESDEN}) + 242.97$

* All values of the dependent variable are in thousands and should be divided by ten (10) to represent the peak hour ratios in percentages.

Equation Number	Multiple R	Standard Error	F- Regression	Degrees Of Freedom-Residual	F-Table		Significance	
					($\alpha=.05$)	($\alpha=.01$)	At 95%	At 99%
9	0.71	12.88	18.2	55	2.76	4.15	Yes	Yes
10	0.561	12.09	12.67	138	2.25	3.5	Yes	Yes
11	0.46	12.69	22.9	167	3.00	4.61	Yes	Yes
12	0.49	12.46	27.0	167	3.0	4.61	Yes	Yes
13	0.58	11.80	23.81	140	2.60	3.8	Yes	Yes
14	0.55	11.57	20.4	142	2.65	3.8	Yes	Yes
15	0.56	11.47	21.5	142	2.65	3.8	Yes	Yes
16	0.628	12.17	29.92	138	2.64	3.8	Yes	Yes
17	0.66	11.70	36.10	138	2.65	3.8	Yes	Yes

<u>Equation Number</u>	<u>Dependent* Variable</u>	<u>Equation</u>
18	HIDPHWRK	$0.049 (\text{TOTEMPLY}) - 0.025 (\text{POPULATN}) + 242.5$
19	HIDPHWRK	$-0.095 (\text{LABFORCE}) + 0.015 (\text{ZONATTRA}) + 0.24 (\text{NTRESDEN}) + 235.3$
20	HJDPHSPB	$-0.01 (\text{POPULATN}) - 0.23 (\text{NTRESDEN}) + 0.11 (\text{NTEMPDEN}) + 0.412$ $(\text{ZONATTRA}/1000)^2 + 207.08$
21	HJDPHSPB	$-0.054 (\text{TOTLCARS}) - 0.19 (\text{NTRESDEN}) - 0.071 (\text{RETEMPLY}) + 231.9$
22	HIDPHSPB	$-0.021 (\text{POPULATN}) + 0.660 (\text{INCOMEZN}/1000)^2 - 77.21/(\text{PARKAVAL})^2$ $+ 181.64$
23	HIDPHSPB	$0.012 (\text{POPULATN}) + 0.01 (\text{INCOMEZN}) - 0.475 (\text{TMTRACBD}) - 12.70$ $(\text{PARKAVAL}) + 268.64$
24	HPA1 **	$0.00096 (\text{TOTEMPLY}) - 0.00041 (\text{POPULATN}) + 1.99$
25	HPA2	$-0.0032 (\text{TOTEMPLY}) - 0.00036 (\text{POPULATN}) + 0.74$
26	HPAM1	$-0.0083 (\text{POPULATN}) + 0.0088 (\text{TOTEMPLY}) + 336.45$
27	HPAM2	$-0.0036 (\text{POPULATN}) + 0.0027 (\text{TOTEMPLY}) + 246.4$

* All values of the dependent variable are in thousands and should be divided by ten (10) to represent the peak hour ratios in percentages.

** See Appendix A

Equation Number	Multiple R	Standard Error	F- Regression	Degrees Of Freedom-Residual	F-Table		Significance	
					($\alpha=0.05$) ($\alpha=0.01$)		At 95%	At 99%
18	0.68	11.10	58.9	135	3.05	4.65	Yes	Yes
19	0.68	11.14	39.0	134	2.65	3.9	Yes	Yes
20	0.433	9.95	6.047	105	2.47	3.5	Yes	Yes
21	0.40	9.786	5.73	103	2.68	3.95	Yes	Yes
22	0.407	12.87	7.34	111	2.70	4.0	Yes	Yes
23	0.383	13.07	4.74	110	2.47	3.6	Yes	Yes
24	0.59	0.279	10.94	42	3.23	5.18	Yes	Yes
25	0.73	0.259	68.4	120	3.07	4.79	Yes	Yes
26	0.34	7.14	2.66	42	3.23	5.18	No	No
27	0.15	6.03	0.45	38	3.23	5.18	No	No

APPENDIX H

BAYESIAN DECISION THEORY

APPENDIX H

BAYESIAN DECISION THEORY

H.1 Basic Formulation

The particular principles of statistical decision theory termed Bayesian have been developed for practical problems by Raiffa and Schlaifer (52,53). The basic elements of the Bayesian decision theory approach are:

1. The set of alternative terminal actions:

$a_1, a_2, \dots, \in A$

2. The possible states of nature: The true peak hour volume

$t_1, t_2, \dots, \in T$

The states of nature listed have the possibility of occurrence after an act has been selected.

The decision maker believes that his choice of an act from the set A is dependent upon the actual state of nature which is not known with certainty.

3. The set of alternative experiments:

$e_0, e_1, \dots, \in E$

The decision maker may elect to perform one experiment from the set E , in order to get more information about the actual state of nature prior to the selection of an act.

4. The set of experimental outcomes:

$v_0, v_1, \dots, \in V$

For every experiment e , there is a space of possible outcomes of that experiment. As indicated by Raiffa and Schlaifer, the space of outcomes V is defined such that it encompasses any outcome of any e in E .

5. The costs $C(e, v, a, t)$, which represents the decision maker's preferences for all e, v, a, t combinations. In a typical course of action, the decision maker performs an experiment e , observes a result v , selects a particular act a , and then observes that a particular state of nature, t , occurs. The space of all possible combinations of (e, v, a, t) is $E \times V \times A \times T$. It is assumed that a single-valued cost function is defined over this space, denoted by $C(e, v, a, t)$, and that the costs $C(e, v)$ and $C(a, t)$ are additive.
6. The probability distributions $P'(t)$ and $P(v|t, e)$, which specify a joint probability measure $P(t, v|e)$, over the space $T \times V$ for each experiment. The $P'(t)$ represents the decision maker's judgement about the relative likelihood of values of t , and $P(v|t, e)$ characterizes each experiment e and is the probability that the outcome v will be observed if the experiment e is performed and t is the value of the state of nature variable. The probability measures are discussed further in Section H.2.

The basic objective of the decision maker is to pick the terminal act, a , which will minimize his cost. The objective of the Bayesian decision theory approach is to aid the decision maker to make a "rational" decision by explicitly considering all possible sequences of experiments, their outcomes, terminal acts and states of nature. Each such combination has a cost $C(e, v, a, t)$. In order to find the best course of action, the expected value of the costs of each combination must be computed. This requires the probability distribution over t .

FIGURE H.1
RELATIONSHIP OF PROBABILITY MEASURES

Source: Reference (55).

H.2 Probability Measures

The basic problem of concern is to assign, either directly or indirectly, a joint probability $P(t, v|e)$ to the joint distribution of t, v , over the space $T \times V$ for each experiment e . This means that the decision maker must define the reliability of each possible experimental outcome in predicting the true state of nature, for each experiment.

The joint probability measure determines four other probability measures. These are shown in Figure H.1, and are defined below:

1. The prior measure $P'(t)$ on the states of nature that the decision maker would assign to t prior to observing the outcome v of experiment e .
2. The conditional measure $P(v|t, e)$ on the space v (i.e. the probability that the outcome v will be observed if the experiment e is performed and t is the true value of state of nature variable).
3. The marginal measure $P(v|e)$ on the space v for all t (i.e. the probability of observing outcome v from experiment e).
4. The conditional measure $P''(t|v, e)$ on the space T (i.e. the likelihood of different states t , given v and e). This is the probability measure that the decision maker assigns to the space T posterior to knowing the outcome v of the experiment e .

The joint probability measure, $P(t, v|e)$ may be specified in three ways. These are:

1. Direct estimation of $P(t, v|e)$;
2. Estimation of $P'(t)$ and $P(v|t, e)$, then computing the probability measures 3 and 4 given above; and
3. Estimation of $P(v|e)$ and $P''(t|v, e)$, then computing the joint distribution and the measures 1 and 2 listed above.

The usual approach to specify $P(t, v|e)$ is as given in 2 and was used in this investigation. Given the probability measures $P'(t)$ and $P(v|t, e)$, the marginal measure $P(v|e)$ can be computed as:

$$P(v|e) = \sum_t P'(t) \cdot P(v|t, e) \quad \dots \text{Eq. (H.1)}$$

The posterior distribution, $P''(t|v, e)$ is computed by Bayes Theorem:

$$P''(t|v, e) = \frac{P'(t) \cdot P(v|t, e)}{\sum_t P'(t) \cdot P(v|t, e)} \quad \dots \text{Eq. (H.2)}$$

The assumption, termed Bayesian, is that each experiment e can be characterized by a conditional probability distribution $P(v|t, e)$, such that the relationship between the prior and posterior distribution is as given by Bayes Theorem.

H.3 Solving the Decision Problem

The basic data of the decision problem can be represented by a decision tree as shown in Figure H.2. Any sequence of (experiment, outcome, action, state of nature) can be represented by a path through this tree. The desirability of that sequence or course of action (e, v, a, t) is completely described by the cost $C(e, v, a, t)$.

The decision maker is interested in choosing an experiment (including the "null" experiment) and therefore he has to evaluate all the possible sequences or courses of action. The mode of analysis concerned with the evaluation of alternative courses of action for the purpose of determining the most valuable experiment is known as preposterior analysis.

FIGURE H.2

A SIMPLE DECISION TREE

H.3.1 Preposterior Analysis

The sequence of operations required in making a decision in preposterior analysis are given below:*

1. Express decision maker's prior judgement about the relative likelihood of different states of nature t , in the form of a probability distribution $P'(t)$.
2. Determine for each experiment, its probability characteristics as given by $P(v|t, e)$. Steps 1 and 2 generate the joint probability distribution $P(t, v|e)$.

3. Compute the marginal measures for each experiment:

$$P(v|e) = \sum_t P'(t) \cdot P(v|t, e) \quad \dots\dots \text{Eq. (H.3)}$$

Note: $P(v_0|e_0) = 1.0$

4. Compute the posterior distribution $P''(t|v, e)$ for each (experiment, outcome) combination:

$$P''(t|v, e) = \frac{P'(t) \cdot P(v|t, e)}{\sum_t P'(t) \cdot P(v|t, e)} \quad \dots\dots \text{Eq. (H.4)}$$

5. Determine for each combination of (experiment, outcome, action, state of nature) its relative cost to the decision maker i.e. $C(e, v, a, t)$.
6. Compute the expected cost for each act a , for each (experiment, outcome) combination:

$$C^*(a, v, e) = \sum_t P''(t|v, e) \cdot C(e, v, a, t) \quad \dots\dots \text{Eq. (H.5)}$$

Note: for prior branch, where no experiment is to be performed,

$$C^*(a, v_0, e_0) = \sum_t P'(t) \cdot C(e_0, v_0, a, t)$$

* It is assumed that the decision tree is already constructed.

7. For each (experiment, outcome) combination, determine the optimal act and its associated cost:

$$C^*(v, e) = \min_a C^*(a, v, e) \quad \dots\dots \text{Eq. (H.6)}$$

8. For each experiment, compute the expected cost:

$$C^*(e) = \sum_v P(v|e) \cdot C^*(v, e) \quad \dots\dots \text{Eq. (H.7)}$$

9. The optimal experiment e^* is that e for which $C^*(e)$ is a minimum. Therefore:

$$C^*(e^*) = \min_e C^*(e) \quad \dots\dots \text{Eq. (H.8)}$$

H.3.2 Value of Information

In the preposterior mode of analysis, if the optimal experiment turns out to be any experiment other than the "null" one, it is important to know the value of information acquired from sampling. In Section H.3.1, the preposterior mode of analysis, the relative value of various experiments was computed by assigning the cost $C(v, e)$ for every possible v of the experiment and then determining the expected value of these costs relative to the marginal probability measure $P(v|e)$. An assumption is usually made that the sampling and terminal costs are additive. Then, instead of finding the cost $C(e, v)$ for each possible experimental outcome, it is possible to compute the decrease in cost which would result for each v if the decision maker altered his prior choice of a terminal action after learning about it. A weighted average of these cost decreases can be calculated and is termed the value of information.

As indicated above, the additivity assumption allows one to write:

$$C(e, v, a, t) = C_s(e, v) + C_\ell(a, t) \quad \dots\dots\text{Eq. (H.9)}$$

The expected cost of an experiment which is defined by equation (H.7) is:

$$C^*(e) = \sum_v P(v|e) \cdot C^*(v, e) = \sum_v P(v|e) \cdot \min_a^m \sum_t P''(t|v, e) \cdot C(e, v, a, t) \quad \dots\dots\text{Eq. (H.10)}$$

From Equations (H.9) and (H.10):

$$\begin{aligned} C^*(e) &= \sum_v P(v|e) \cdot [\min_a \sum_t P''(t|v, e) \{C_s(e, v) + C_\ell(a, t)\}] \\ &= \sum_v P(v|e) \cdot [C_s(e, v) + \min_a^m \sum_t P''(t|v, e) \cdot C_\ell(a, t)] \quad \dots\dots\text{Eq. (H.11)} \end{aligned}$$

In other words:

$$C_s^*(e) = \sum_v P(v|e) \cdot [C_s(e, v)] \quad \dots\dots\text{Eq. (H.12)}$$

and

$$C_\ell^*(e) = \sum_v P(v|e) \cdot [\min_a^m \sum_t P''(t|v, e) \cdot C_\ell(a, t)] \quad \dots\dots\text{Eq. (H.13)}$$

Let a' be an act optimal under prior distribution of t :

$$\sum_t P'(t) \cdot C_\ell(a', t) \leq \sum_t P'(t) \cdot [C_\ell(a, t)] \quad \dots\dots\text{Eq. (H.14)}$$

and let a_v be defined as the act optimal under the posterior distribution of t determined by the outcome v of an experiment:

$$\sum_t P''(t|v, e) \cdot [C_\ell(a_v, t)] = \min_a^m \sum_t P''(t|v, e) \cdot C_\ell(a, t) \quad \dots\dots\text{Eq. (H.15)}$$

If instead of choosing the act a directly, the decision maker performs an experiment e , observes v , and then chooses a_v he decreases his terminal cost by:

$$W_{\ell}(e, v) = \sum_t P(t|v, e) \cdot C_{\ell}(a', t) - \sum_t P(t|v, e) \cdot C_{\ell}(a, t) \quad \dots\dots \text{Eq. (H.16)}$$

The term $W_{\ell}(e, v)$ is termed the conditional value of sample information. However, before v is observed the expected value of sample information is:

$$W_{\ell}^*(e) = \sum_v P(v|e) \cdot [W_{\ell}(e, v)] \quad \dots\dots \text{Eq. (H.17)}$$

The expected terminal cost of a particular experiment is the expected terminal cost of an immediate terminal action minus the expected cost of sample information. That is:

$$C^*(e) = C_{\ell}^*(e_0) - W_{\ell}^*(e) \quad \dots\dots \text{Eq. (H.18)}$$

The expected net gain of experimentation is, therefore, defined to be the expected value of sample information less the expected cost of obtaining it:

$$W^*(e) = W_{\ell}^*(e) - C_S^*(e) \quad \dots\dots \text{Eq. (H.19)}$$

where:

$$C_S^*(e) = \sum_v P(v|e) \cdot [C_S(e, v)]$$

APPENDIX I

BAYESIAN DECISION THEORY PROGRAMME

- I.1 VARIABLE LIST
- I.2 PROGRAMME LISTING
- I.3 PROGRAMME FLOWCHART

APPENDIX I.1

VARIABLE LIST

(I1) Input from file
 (I2) Input from user
 (C) Computed

In Common INFO

(I2) ALEN . length of link
 (C) CAT(3,5) . cost of sampling at t nodes for the null experiment
 (C) ISOLN . the optimum experiment index
 (C) ITEMPE(5) . the optimum V index for each experiment
 (C) ITEMPV(5,5) . the optimum action index for each experiment and V
 (I2) IVALA(5) . the indexes of the actions selected
 (I2) NAME(20) . the title of the run
 (I2) NUMA . the number of actions
 (I1) NUME . the number of experiments
 (I2) NUMT . the number of branches
 (C) PT(5) . the prior probabilities
 (C) PTV(5,5,5) . the posterior probabilities of a given V by experiment
 (C) PVE(5,5) . the marginal probabilities of V by experiment
 (I1) PVT(5,5,5) . the conditional probabilities of V given t by experiment
 (I1) SCOST(5) . the cost of sampling by experiment
 (C) SOLN . the optimum cost
 (C) TEMPA(5,5,3) . the costs at the action nodes
 (C) TEMPE(5) . the costs at the experiment nodes
 (C) TEMPV(5,5) . the optimum costs at the V nodes
 (I2) VT(5) . the traffic volume values at the t nodes
 (C) WTH(5) . the expected net gain of each experiment

In Main

IFLAG . used in OUTPT1 0 if complete output
is required 1 if only the summary table
is desired

In Subroutine COSTS

COST(3,13) . contains the cost curve data points
IND(2,5) . contains the computed indexes, from
volume values, for the COST matrix

In Subroutine OUTPT1

ITITLE(5,6) . contains the titles for each experiment
INAMA . contains the character 'A' used in
output
INAME . contains the character 'E' used in
output
INAMT . contains the character 'T' used in
output
INAMV . contains the character 'V' used in
output

INDEXES

IA . index value for action (up to 3 in the
programme)
IE . index value for experiment (up to 5)
IV . index value for V (up to 5)
IT . index value for t (up to 5)

APPENDIX I.2.

PROGRAMME LISTING

EDIT BAYESIANS
 EDIT HERE
 *IY 1-999

```

1.000 C
2.000 C PROGRAM NAME: BAYESIANS.
3.000 C
4.000 C THIS PROGRAM SOLVES A 1 STAGE DECISION TREE WITH PRIOR
5.000 C PROBABILITIES AND CONDITIONAL PROBABILITIES AS INPUT
6.000 C
7.000 C
8.000 C WRITTEN FOR : L. SHALLAL P.ENG
9.000 C BY : F. WILLIAMS ENG III
10.000 C CARLETON UNIVERSITY
11.000 C DATE : JULY 1978
12.000 C
13.000 C
14.000 C UNIT #4 IS FILE INPUT
15.000 C UNIT #5 IS TTY INPUT
16.000 C UNIT #6 IS TTY OUTPUT
17.000 C
18.000 C
19.000 C RESTRICTIONS
20.000 C
21.000 C MAXIMUM NUMBER OF EXPERIMENTS = 5
22.000 C MAXIMUM NUMBER OF "V" BRANCHES = 5
23.000 C MAXIMUM NUMBER OF ACTIONS = 3
24.000 C MAXIMUM NUMBER OF "T" BRANCHES = 5
25.000 C
26.000 C NOTE:
27.000 C COST VALUES FOR "T" NODES ONLY VALID FOR PEAK HOUR VOLUMES
28.000 C BETWEEN 200 AND 1400 CARS PER HOUR
29.000 C
30.000 C
31.000 C
32.000 C COMMON /INFO/ALEN,CAT(3,5),ISOLN,ITEMPE(5),ITEMPV(5,5),
33.000 C XIVALA(5),NAME(20),NUMA,NUME,NUMT,PT(5),PTV(5,5,5),
34.000 C XPVE(5,5),PVT(5,5,5),SCOST(5),SOLN,TEMPA(5,5,3),TEMPE(5),
35.000 C XIEMPV(5,5),VI(5),WTH(5)
36.000 C
37.000 C THE MAIN PROGRAM INPUTS THE VARIABLES
38.000 C THEN INITIALIZES THE PVT AND SCOST MATRICES
39.000 C THROUGH THE SUBROUTINE INITIAL
40.000 C IT THEN PROCEEDS TO SOLVE THE DECISION TREE
41.000 C BY CALLING THE SUBROUTINES:
42.000 C INITIAL INITIALIZE VARIABLES FROM FILE
43.000 C CALPRO CALCULATES PROBABILITIES
44.000 C COSTS CALCULATES THE FINAL COSTS
45.000 C SOLVE SOLVES THE TREE
46.000 C WORTH CALCULATES WORTH OF THE EXPERIMENT
47.000 C OUTPTI OUTPUTS THE RESULTS

```

```

48.000 C
49.000 C INPUT DATA FROM TERMINAL
50.000 C
51.000 WRITE(6,1)
52.000 1 FORMAT(' ', 'INPUT PROBLEM TITLE')
53.000 READ(5,2), (NAME(I), I=1,20)
54.000 2 FORMAT(20A4)
55.000 WRITE(6,3)
56.000 3 FORMAT(' ', 'INPUT LENGTH OF LINK')
57.000 READ(5,4) ALEN
58.000 4 FORMAT(1G)
59.000 WRITE(6,5)
60.000 5 FORMAT(' ', 'INPUT NUMBER OF T VALUES')
61.000 110 READ(5,4) NUMT
62.000 IF (NUMT.GT.0.AND.NUMT.LT.6) GOTO 100
63.000 WRITE(6,6)
64.000 6 FORMAT(' ', 'MAXIMUM VALUE OF 5', 'RETRY')
65.000 GOTO 110
66.000 100 CONTINUE
67.000 WRITE(6,7) NUMT
68.000 7 FORMAT(' ', 'INPUT ', 11, ' PRIOR PROBABILITY VALUES OF T')
69.000 DO 120 IT=1,NUMT
70.000 READ(5,4) PT(IT)
71.000 120 CONTINUE
72.000 WRITE(6,8) NUMT
73.000 8 FORMAT(' ', 'INPUT ', 11, ' PEAK HOUR VOLUME VALUES OF T')
74.000 DO 130 IT=1,NUMT
75.000 READ(5,4) VT(IT)
76.000 130 CONTINUE
77.000 WRITE(6,9)
78.000 9 FORMAT(' ', 'INPUT NUMBER OF ALTERNATIVE ACTIONS (A'S)')
79.000 READ(5,4) NUMA
80.000 WRITE(6,10) NUMA
81.000 10 FORMAT(' ', 'INPUT ', 11, ' A VALUES (INTEGER SUBSCRIPTS OF A)')
82.000 DO 140 IA=1,NUMA
83.000 READ(5,4) IVALA(IA)
84.000 140 CONTINUE
85.000 WRITE(6,99)
86.000 99 FORMAT('O', 'DO YOU WISH A COMPLETE LISTING 0=YES 1=NO')
87.000 READ(5,4) IFLAG
88.000 C
89.000 C CALL INITIALIZATION ROUTINE
90.000 C
91.000 C CALL INITAL
92.000 C
93.000 C CALL PROBABILITY CALCULATION ROUTINE
94.000 C
95.000 C CALL CALPRO
96.000 C
97.000 C CALL COST CALCULATION ROUTINE
98.000 C
99.000 C CALL COSTS
100.000 C
101.000 C CALL TREE SOLUTION ROUTINE
102.000 C
103.000 C CALL SOLVE
104.000 C
105.000 C CALL WORTH CALCULATION ROUTINE
106.000 C

```

```

107.000 CALL WORTH
108.000 C
109.000 C CALL FIRST OUTPUT ROUTINE
110.000 C
111.000 CALL OUTPT(IFLAG)
112.000 STOP
113.000 END
114.000 C
115.000 C SUBROUTINE INITIAL
116.000 C THIS SUBROUTINE INITIALIZES THE VARIABLE NUME, THE
117.000 C MATRIX PVT AND THE VECTOR SCOST FROM FORTRAN UNIT #4
118.000 C (USUALLY A FILE)
119.000 C
120.000 C SUBROUTINE INITIAL
121.000 C COMMON /INFO/ALN,CAT(3,5),ISOLN,ITEMPE(5),ITEMPV(5,5),
122.000 C XIVALA(5),NAME(20),NUMA,NUME,NUMT,PT(5),PIV(5,5,5),
123.000 C XPVE(5,5),PVT(5,5,5),SCOST(5),SOLN,TEMPA(5,5,3),TEMPE(5),
124.000 C XTEMPV(5,5),VT(5),WIH(5)
125.000 C
126.000 C READ VALUE FOR NUME
127.000 C
128.000 C READ(4,2) NUME
129.000 C
130.000 C READ VALUES FOR MATRIX PVT
131.000 C
132.000 C DO 10 IE=1,NUME
133.000 C DO 20 IV=1,NUMT
134.000 C READ(4,1) (PVT(IE,IV,II),II=1,NUMI)
135.000 1 FORMAT(5G)
136.000 20 CONTINUE
137.000 10 CONTINUE
138.000 C
139.000 C READ VALUES FOR VECTOR SCOST
140.000 C
141.000 C DO 30 IE=1,NUME
142.000 C READ(4,2) SCOST(IE)
143.000 2 FORMAT(1G)
144.000 3 CONTINUE
145.000 C RETURN
146.000 C END
147.000 C
148.000 C SUBROUTINE CALPRO
149.000 C THIS SUBROUTINE COMPUTES THE MARGINAL PROBABILITIES
150.000 C PVE AND THE POSTERIOR PROBABILITIES PIV
151.000 C
152.000 C SUBROUTINE CALPRO
153.000 C COMMON /INFO/ALN,CAT(3,5),ISOLN,ITEMPE(5),ITEMPV(5,5),
154.000 C XIVALA(5),NAME(20),NUMA,NUME,NUMT,PT(5),PIV(5,5,5),
155.000 C XPVE(5,5),PVT(5,5,5),SCOST(5),SOLN,TEMPA(5,5,3),TEMPE(5),
156.000 C XTEMPV(5,5),VT(5),WIH(5)
157.000 C DO 10 IE=1,NUME
158.000 C DO 20 IV=1,NUMT
159.000 C TEMP=0.0
160.000 C
161.000 C COMPUTE MARGINAL PROBABILITIES PVE
162.000 C DO 30 IT=1,NUMT
163.000 C TEMP=TEMP+PVT(IE,IV,IT)*PT(IT)
164.000 30 CONTINUE
165.000 C PVE(IE,IV)=TEMP

```

```

166.000 C
167.000 C
168.000 C
169.000 DO 40 II=1,NUM1
170.000 PIV(IE,IV,II)=(PT(II)*PVI(IE,IV,II))/PVE(IE,IV)
171.000 40 CONTINUE
172.000 20 CONTINUE
173.000 10 CONTINUE
174.000 RETURN
175.000 END
176.000 C
177.000 SUBROUTINE COSIS
178.000 THIS SUBROUTINE COMPUTES THE COST AT EACH "I" NODE FOR A
179.000 SPECIFIED ACTION FOR THE NULL EXPERIMENT. THE TABLE IS
180.000 VALID FOR "I" VOLUMES BETWEEN 200 & 1400. ANY VALUES
181.000 BELOW OR ABOVE THESE LIMITS WILL BE ASSIGNED THE MINIMUM
182.000 OR MAXIMUM COST RESPECTIVELY. LINEAR INTERPOLATION IS DONE
183.000 FOR ANY VOLUME VALUE WHICH DOES NOT FALL ON AN EXACT DATA
184.000 POINT
185.000 C
186.000 SUBROUTINE COSIS
187.000 COMMON /INFO/ALEN,CAT(3,5),ISOLN,ITEMPE(5),ITEMPV(5,5),
188.000 XIVALA(5),NAME(20),NUMA,NUME,NUMT,PT(5),PTV(5,5,5),
189.000 XPVE(5,5),PVI(5,5,5),SCOST(5),SOLN,TEMPA(5,5,3),TEMPE(5),
190.000 XTEMPV(5,5),VI(5),WIH(5)
191.000 DIMENSION COST(3,13),IND(2,5)
192.000 DATA COST/1.5E06,2.0E06,3.41E06,2.05E06,2.36E06,3.68E06,
193.000 X2.15E06,2.88E06,3.96E06,3.34E06,3.68E06,4.34E06,
194.000 X5.14E06,4.48E06,4.73E06,7.65E06,5.53E06,5.56E06,
195.000 X12.79E06,6.78E06,6.3E06,18.35E06,9.04E06,7.05E06,
196.000 X22.66E06,11.58E06,8.02E06,28.36E06,16.26E06,9.76E06,
197.000 X33.04E06,20.82E06,12.23E06,44.48E06,27.34E06,15.29E06,
198.000 X55.6E06,32.39E06,18.9E06/
199.000 C
200.000 TRAP VOLUME VALUES OUTSIDE LIMITS AND COMPUTE TABLE INDEXES
201.000 C
202.000 DO 10 IT=1,NUMT
203.000 IF (VT(IT).LT.200.0) VT(IT)=200.0
204.000 IF (VT(IT).GT.1400.0) VT(IT)=1400.0
205.000 IND(1,IT)=IFIX((VT(IT)-100.0)/100.0)
206.000 IND(2,IT)=IND(1,IT)+1
207.000 IF (IND(2,IT).GT.13) IND(2,IT)=13
208.000 10 CONTINUE
209.000 C
210.000 PERFORM LINEAR INTERPOLATION BETWEEN ADJACENT DATA POINTS
211.000 FOR EXACT COST
212.000 C
213.000 DO 20 IA=1,NUMA
214.000 DO 30 II=1,NUMT
215.000 TEMP=(COST(IA,IND(2,IT))-COST(IA,IND(1,IT)))/100.0
216.000 TEMP=TEMP*(VI(IT)-(100.0+IND(1,IT)*100.0))
217.000 CAT(IA,IT)=(COST(IA,IND(1,IT))+TEMP)*ALEN
218.000 30 CONTINUE
219.000 20 CONTINUE
220.000 RETURN
221.000 END
222.000 C
223.000 SUBROUTINE SOLVE
224.000 THIS SUBROUTINE COMPUTES THE SOLUTION TO THE 1 STAGE TREE

```



```

225.000 C
226.000 SUBROUTINE SOLVE
227.000 COMMON /INFO/ALN,CAT(3,5),ISOLN,IEMPE(5),IEMPV(5,5),
228.000 XIVALA(5),NAME(20),NUMA,NUME,NUMT,PI(5),PTV(5,5,5),
229.000 XPVE(5,5),PVI(5,5,5),SCOST(5),SOLN,IEMPA(5,5,3),TEMPE(5),
230.000 XTEMPV(5,5),VI(5),WIH(5)
231.000 C
232.000 C COMPUTE THE PROBABILISTIC COST AT EACH ACTION NODE
233.000 C
234.000 DO 10 IE=1,NUME
235.000 DO 20 IV=1,NUMT
236.000 IF (IE.EQ.1.AND.IV.GT.1) GO10 20
237.000 DO 30 IND=1,NUMA
238.000 IA=IVALA(IND)
239.000 IEMPA(IE,IV,IA)=0.0
240.000 DO 40 IT=1,NUMT
241.000 TEMPI=(PIV(IE,IV,IT)*(CAT(IA,IT)+SCOST(IE)))
242.000 IEMPA(IE,IV,IA)=TEMPI+IEMPA(IE,IV,IA)
243.000 40 CONTINUE
244.000 30 CONTINUE
245.000 20 CONTINUE
246.000 10 CONTINUE
247.000 C
248.000 C FIND MINIMUM ACTION COST FOR EACH "V" BRANCH AND IT'S
249.000 C CORRESPONDING INDEX
250.000 C
251.000 DO 50 IE=1,NUME
252.000 IEMPE(IE)=1
253.000 DO 60 IV=1,NUMT
254.000 IF (IE.EQ.1.AND.IV.GT.1) GO10 60
255.000 IEMPV(IE,IV)=IEMPA(IE,IV,IVALA(1))
256.000 ITEMPV(IE,IV)=1
257.000 DO 70 IND=2,NUMA
258.000 IA=IVALA(IND)
259.000 IF (IEMPV(IE,IV).LT.IEMPA(IE,IV,IA)) GO10 70
260.000 IEMPV(IE,IV)=IEMPA(IE,IV,IA)
261.000 ITEMPV(IE,IV)=IA
262.000 C
263.000 C FIND MINIMUM "V" COST FOR EACH EXPERIMENT AND STORE IT'S
264.000 C INDEX ONLY
265.000 C
266.000 70 CONTINUE
267.000 IF (IEMPV(IE,1).LT.IEMPV(IE,IV)) GO10 60
268.000 IEMPE(IE)=IV
269.000 60 CONTINUE
270.000 50 CONTINUE
271.000 C
272.000 C COMPUTE THE PROBABILISTIC COST AT EACH EXPERIMENT
273.000 C NODE
274.000 C
275.000 TEMPE(1)=IEMPV(1,1)
276.000 DO 80 IE=2,NUME
277.000 TEMPE(IE)=0.0
278.000 DO 90 IV=1,NUMT
279.000 TEMPE(IE)=TEMPE(IE)+((PVE(IE,IV)*IEMPV(IE,IV)))
280.000 90 CONTINUE
281.000 80 CONTINUE
282.000 C
283.000 C FIND THE MINIMUM EXPERIMENT COST AND IT'S CORRESPONDING
284.000 C INDEX

```

```

285.000 C
286.000 SOLN=TEMPE(1)
287.000 ISOLN=1
288.000 DO 100 IE=2,NUME
289.000 IF (SOLN.LT.TEMPE(IE)) GOTO 100
290.000 SOLN=TEMPE(IE)
291.000 ISOLN=IE
292.000 100 CONTINUE
293.000 RETURN
294.000 END
295.000 C
296.000 C SUBROUTINE WORTH
297.000 C THIS SUBROUTINE COMPUTES THE EXPECTED NET GAIN OF EACH
298.000 C EXPERIMENT RELATIVE TO THE NULL EXPERIMENT
299.000 C
300.000 SUBROUTINE WORTH
301.000 COMMON /INFO/ALEN,CAT(3,5),ISOLN,ITEMPE(5),ITEMPV(5,5),
302.000 XIVALA(5),NAME(20),NUMA,NUME,NUMT,PT(5),PTV(5,5,5),
303.000 XPVE(5,5),PVT(5,5,5),SCOST(5),SOLN,TEMPA(5,5,3),TEMPE(5),
304.000 XTEMPV(5,5),VT(5),WITH(5)
305.000 DO 10 IE=1,NUME
306.000 WITH(IE)=TEMPE(1)-TEMPE(IE)
307.000 10 CONTINUE
308.000 RETURN
309.000 END
310.000 C
311.000 C SUBROUTINE OUTPTI
312.000 C THIS SUBROUTINE OUTPUTS PROBABILITY, COST TABLES AND A
313.000 C SUMMARY IF IFLAG=0 OR JUST A SUMMARY TABLE IF IFLAG=1
314.000 C
315.000 SUBROUTINE OUTPTI(IFLAG)
316.000 COMMON /INFO/ALEN,CAT(3,5),ISOLN,ITEMPE(5),ITEMPV(5,5),
317.000 XIVALA(5),NAME(20),NUMA,NUME,NUMT,PT(5),PTV(5,5,5),
318.000 XPVE(5,5),PVT(5,5,5),SCOST(5),SOLN,TEMPA(5,5,3),TEMPE(5),
319.000 XTEMPV(5,5),VT(5),WITH(5)
320.000 DIMENSION ITITLE(5,6)
321.000 DATA ITITLE//AREA//,RAII//,ZONA//,E(2)//,ZONA//,
322.000 X'WIDE//,OS B//,L EN//,PLU//,L IN//,
323.000 X'RAI//,Y MO//,D RA//,S MO//,TERC//,
324.000 X'O DE-8//,TIOS//,RE D//,HANG//,
325.000 X'PUR//,ATA//,E RA//,
326.000 X'POSE//,TIOS//
327.000 DATA INAME//E//,INAMV//V//,INAMT//T//,INAMA//A//
328.000 IF (IFLAG.EQ.1) GOTO 900
329.000 WRITE(6,1)
330.000 1 FORMAT('O', 'PRIOR PROBABILITIES P(T)///',24('---'))
331.000 DO 10 IT=1,NUMT
332.000 WRITE(6,2) PT(IT)
333.000 2 FORMAT(5X,E13.7)
334.000 1 CONTINUE
335.000 WRITE(6,3)
336.000 3 FORMAT('O', 'CONDITIONAL PROBABILITIES OF V GIVEN T BY ',
337.000 X'EXPERIMENT, P(V/T,E)///',61('---'))
338.000 DO 20 IE=1,NUME
339.000 ICNT=IE-1
340.000 WRITE(6,4) ICNT
341.000 4 FORMAT('O', 'FOR EXPERIMENT E(//,11,')')
342.000 DO 30 IV=1,NUMT
343.000 IF (IE.EQ.1.AND.IV.GT.1) GOTO 30

```

```

344.000 WRITE(6,5) INAMV,IV,NUMT,(PVT(IE,IV,IV),IT=1,NUMT)
345.000 5 FORMAT(' ',A1,'(',11,')',1X,N(E13.7,2X))
346.000 30 CONTINUE
347.000 20 CONTINUE
348.000 WRITE(6,6)
349.000 6 FORMAT('0', 'COSTS FOR E(0) (IN $*S) A BY 1' // ',29(---)/)
350.000 DO 40 IA=1,NUMA
351.000 WRITE(6,5) INAMA,IA,NUMT,(CAT(IA,IT),IT=1,NUMT)
352.000 40 CONTINUE
353.000 DO 120 IE=2,NUME
354.000 ICNT=IE-1
355.000 WRITE(6,22) ICNT,SCOST(IE)
356.000 22 FORMAT(' ', 'COSTS FOR E(',11,') = E(0) + $',E13.7)
357.000 120 CONTINUE
358.000 WRITE(6,7)
359.000 7 FORMAT('0', 'COST VALUE OF A NODES (IN $*S)' // ',30(---)/)
360.000 DO 50 IE=1,NUME
361.000 ICNT=IE-1
362.000 WRITE(6,4) ICNT
363.000 DO 60 IV=1,NUMT
364.000 IF (IE.EQ.1.AND.IV.GT.1) GOTO 60
365.000 WRITE(6,5) INAMV,IV,NUMA,(TEMPA(IE,IV,IA),IA=1,NUMA)
366.000 60 CONTINUE
367.000 50 CONTINUE
368.000 WRITE(6,8)
369.000 8 FORMAT('0', 'COST VALUE OF V NODES (IN $*S)' // ',30(---)/)
370.000 DO 70 IE=1,NUME
371.000 ICNT=IE-1
372.000 NCNT=NUMT
373.000 IF (IE.EQ.1) NCNT=1
374.000 WRITE(6,5) INAME,ICNT,NCNT,(TEMPV(IE,IV),IV=1,NCNT)
375.000 70 CONTINUE
376.000 WRITE(6,9)
377.000 9 FORMAT('0', 'COST VALUE OF E NODES (IN $*S)' // ',30(---)/)
378.000 WRITE(6,21) NUME,(TEMPE(IE),IE=1,NUME)
379.000 21 FORMAT(6X,N(E13.7,2X))
380.000 WRITE(6,11)
381.000 11 FORMAT('0', 'POSTERIOR PROBABILITIES OF T GIVEN V BY' // ',50(---)/)
382.000 X'EXPERIMENT' // ',50(---)/)
383.000 DO 80 IE=1,NUME
384.000 ICNT=IE-1
385.000 WRITE(6,4) ICNT
386.000 DO 90 IV=1,NUMT
387.000 IF (IE.EQ.1.AND.IV.GT.1) GOTO 90
388.000 WRITE(6,5) INAMV,IV,NUMT,(PTV(IE,IV,IT),IT=1,NUMT)
389.000 90 CONTINUE
390.000 80 CONTINUE
391.000 WRITE(6,12)
392.000 12 FORMAT('0', 'MARGINAL PROBABILITIES OF V BY E' // ',32(---)/)
393.000 DO 100 IE=1,NUME
394.000 NCNT=NUMT
395.000 IF (IE.EQ.1) NCNT=1
396.000 ICNT=IE-1
397.000 WRITE(6,5) INAME,ICNT,NCNT,(PVE(IE,IV),IV=1,NCNT)
398.000 100 CONTINUE
399.000 900 CONTINUE
400.000 WRITE(6,13)
401.000 13 FORMAT(' ',5X,'SUMMARY TABLE')

```

```

402.000 WRITE(6,14),(NAME(I),I=1,20)
403.000 14 FORMAT('0',20A4)
404.000 WRITE(6,15) ALLEN
405.000 15 FORMAT(' ',LINK LENGTH = ',E13.7,' MILES')
406.000 WRITE(6,16)
407.000 16 FORMAT('0',28X,'COST OF SAMPLE',2X,'MINIMUM COST',2X,
408.000 X'EXPECTED NET GAIN OF E'/29X,52(' '))//)
409.000 DO 110 IE=1,NUME
410.000 ICNT=IE-1
411.000 WRITE(6,17) ICNT,(ITITLE(IE,I),I=1,6),SCOST(IE),TEMPE(IE),WITH
412.000 17 FORMAT(' ',E(' ',I1,' ')',6A4,2X,'$',E13.7,4X,'$',E13.7,4X,'$',
413.000 XE13.7)
414.000 110 CONTINUE
415.000 WRITE(6,18) SOLN
416.000 18 FORMAT('0',MINIMUM COST =$',E13.7)
417.000 ICNT=ISOLN-1
418.000 WRITE(6,19) ICNT,ITEMPE(ISOLN),ITEMPV(ISOLN,ITEMPE(ISOLN))
419.000 19 FORMAT('0',PATH IS E(' ',I1,' ') TO V(' ',I1,' ') TO A(' ',I1,' '))
420.000 WRITE(6,24)
421.000 24 FORMAT('0',OTHER PATHS ARE//)
422.000 DO 130 IE=1,NUME
423.000 IF (IE.EQ.ISOLN) GOTO 130
424.000 ICNT=IE-1
425.000 WRITE(6,23) ICNT,ITEMPE(IE),ITEMPV(IE,ITEMPE(IE))
426.000 23 FORMAT(9X,E(' ',I1,' ') TO V(' ',I1,' ') TO A(' ',I1,' '))
427.000 130 CONTINUE
428.000 RETURN
429.000 END
--EOF HIT AFTER 429.

```

APPENDIX I.3
PROGRAMME FLOWCHART

APPENDIX J

LISTING OF A TYPICAL EXECUTION AND
RESULTS OF THE DECISION THEORY PROGRAMME

APPENDIX J

LISTING OF A TYPICAL EXECUTION AND
RESULTS OF THE DECISION THEORY PROGRAMME

BT.
INPUT PROBLEM TITLE
?TEST OF PROGRAM EG MEADOWSLANDS DR.
INPUT LENGTH OF LINK
?.9
INPUT NUMBER OF T VALUES
?5
INPUT 5 PRIOR PROBABILITY VALUES OF T
?.125
?.2
?.35
?.2
?.125
INPUT 5 PEAK HOUR VOLUME VALUES OF T
?100
?250
?320
?400
?700
INPUT NUMBER OF ALTERNATIVE ACTIONS (A"S)
?3
INPUT 3 A VALUES (INTEGER SUBSCRIPTS OF A)
?1
?2
?3
DO YOU WISH A COMPLETE LISTING 0=YES 1=NO
?

PRIOR PROBABILITIES P(T)

.1250000E 00
 .2000000E 00
 .3500000E 00
 .2000000E 00
 .1250000E 00

CONDITIONAL PROBABILITIES OF V GIVEN T BY EXPERIMENT, P(V/T,E)

FOR EXPERIMENT E(0)

V(1)	.1000000E 01	.1000000E 01	.1000000E 01	.1000000E 01	.1000000E 01
FOR EXPERIMENT E(1)					
V(1)	.5000000E 00	.2000000E 00	.5000000E 01	.1000000E 00	.3000000E 01
V(2)	.2500000E 00	.5000000E 00	.2000000E 00	.1000000E 00	.6999999E 01
V(3)	.1500000E 00	.2000000E 00	.5000000E 00	.1500000E 00	.1500000E 00
V(4)	.6999999E 01	.7499999E 01	.2000000E 00	.5000000E 00	.2500000E 00
V(5)	.3000000E 01	.2500000E 01	.5000000E 01	.1500000E 00	.5000000E 00

FOR EXPERIMENT E(2)

V(1)	.6000000E 00	.1000000E 00	.5000000E 01	.3000000E 01	.4000000E 01
V(2)	.3000000E 00	.6000000E 00	.1500000E 00	.6999999E 01	.6000000E 01
V(3)	.6999999E 01	.2000000E 00	.6000000E 00	.1500000E 00	.1000000E 00
V(4)	.2000000E 01	.6000000E 01	.1500000E 00	.6000000E 00	.2000000E 00
V(5)	.1000000E 01	.4000000E 01	.5000000E 01	.1500000E 00	.6000000E 00

FOR EXPERIMENT E(3)

V(1)	.7000000E 00	.1250000E 00	.2500000E 01	.2500000E 01	.4000000E 01
V(2)	.1500000E 00	.7000000E 00	.1250000E 00	.7499999E 01	.6000000E 01
V(3)	.1000000E 00	.1250000E 00	.7000000E 00	.1000000E 00	.7499999E 01
V(4)	.4000000E 01	.3500000E 01	.1250000E 00	.7000000E 00	.1250000E 00
V(5)	.1000000E 01	.1500000E 01	.2500000E 01	.1000000E 00	.7000000E 00

FOR EXPERIMENT E(4)

V(1)	.8000000E 00	.7999998E 01	.3000000E 01	.1000000E 01	.3000000E 01
V(2)	.7499999E 01	.8000000E 00	.6999999E 01	.2000000E 01	.4000000E 01
V(3)	.5500000E 01	.6999999E 01	.8000000E 00	.9500003E 01	.5500000E 01
V(4)	.4000000E 01	.4000000E 01	.7999998E 01	.8000000E 00	.7499999E 01
V(5)	.3000000E 01	.1000000E 01	.2000000E 01	.7499999E 01	.8000000E 00

COSTS FOR E(0) (IN \$'S) A BY I

A(1) .1349999E 07 .1597499E 07 .1862999E 07 .1934999E 07 .6884999E 07
 A(2) .1817999E 07 .1970999E 07 .2217599E 07 .2591999E 07 .4976999E 07
 A(3) .3068999E 07 .3190499E 07 .3362399E 07 .3563999E 07 .5003999E 07
 COSTS FOR E(1) = E(0) + \$.1000000E 03
 COSTS FOR E(2) = E(0) + \$.4000000E 03
 COSTS FOR E(3) = E(0) + \$.2000000E 04
 COSTS FOR E(4) = E(0) + \$.6000000E 04
 COST VALUE OF A NODES (IN \$'S)

FOR EXPERIMENT E(0)

V(1) .2387922E 07 .2538132E 07 .3536801E 07

FOR EXPERIMENT E(1)

V(1) .1707203E 07 .2099415E 07 .3257972E 07
 V(2) .1975277E 07 .2193720E 07 .3327877E 07
 V(3) .2132421E 07 .2379164E 07 .3449049E 07
 V(4) .2554948E 07 .2735367E 07 .3657227E 07
 V(5) .4497066E 07 .3741598E 07 .4260925E 07

FOR EXPERIMENT E(2)

V(1) .1715681E 07 .2065297E 07 .3233039E 07
 V(2) .1809729E 07 .2137480E 07 .3291541E 07
 V(3) .2028796E 07 .2325431E 07 .3419643E 07
 V(4) .2475291E 07 .2736652E 07 .3657306E 07
 V(5) .4717629E 07 .3855303E 07 .4329980E 07

FOR EXPERIMENT E(3)

V(1) .1666483E 07 .2025610E 07 .3206272E 07
 V(2) .1829247E 07 .2149797E 07 .3301148E 07
 V(3) .1978733E 07 .2290772E 07 .3401133E 07
 V(4) .2262987E 07 .2653932E 07 .3606639E 07
 V(5) .5511697E 07 .4275163E 07 .4582568E 07

FOR EXPERIMENT E(4)

V(1) .1592466E 07 .1976516E 07 .3175347E 07
 V(2) .1761090E 07 .2086035E 07 .3263703E 07
 V(3) .1956680E 07 .2284454E 07 .3401122E 07
 V(4) .2125255E 07 .2612439E 07 .3581370E 07
 V(5) .5789353E 07 .4411969E 07 .4665776E 07

COST VALUE OF V NODES (IN \$'S)

E(0) .2387922E 07
 E(1) .1707203E 07 .1875277E 07 .2132421E 07 .2554948E 07 .3741598E 07
 E(2) .1715681E 07 .1809729E 07 .2028796E 07 .2475291E 07 .3855303E 07
 E(3) .1666483E 07 .1829247E 07 .1978733E 07 .2262987E 07 .4275168E 07
 E(4) .1592466E 07 .1761090E 07 .1956680E 07 .2125265E 07 .4411969E 07

COST VALUE OF E NODES (IN \$'S)

.2387922E 07 .2298309E 07 .2274708E 07 .2240918E 07 .2217959E 07

POSTERIOR PROBABILITIES OF I GIVEN V BY EXPERIMENT

FOR EXPERIMENT E(0)					
V(1)	.1250000E 00	.2000000E 00	.3500000E 00	.2000000E 00	.1250000E 00
FOR EXPERIMENT E(1)					
V(1)	.4347829E 00	.2782610E 00	.1217392E 00	.1391305E 00	.2608697E 01
V(2)	.1358696E 00	.4347826E 00	.3043479E 00	.8695656E 01	.3804349E 01
V(3)	.6637168E 01	.1415929E 00	.6194692E 00	.1061947E 00	.6637168E 01
V(4)	.3888890E 01	.6666666E 01	.3111112E 00	.4444444E 00	.1388889E 00
V(5)	.3157895E 01	.4210526E 01	.1473684E 00	.2526315E 00	.5263159E 00
FOR EXPERIMENT E(2)					
V(1)	.6072876E 00	.1619434E 00	.1417004E 00	.4858300E 01	.4048584E 01
V(2)	.1619872E 00	.5183587E 00	.2267820E 00	.6047519E 01	.3239742E 01
V(3)	.2904565E 01	.1327801E 00	.6970955E 00	.9958500E 01	.4149380E 01
V(4)	.1179245E 01	.5660379E 01	.2476416E 00	.5660377E 00	.1179246E 00
V(5)	.9487670E 02	.6072108E 01	.1328273E 00	.2277040E 00	.5692602E 00
FOR EXPERIMENT E(3)					
V(1)	.6666669E 00	.1904763E 00	.6666666E 01	.3809525E 01	.3809525E 01
V(2)	.8333331E 01	.6222224E 00	.1944445E 00	.6666666E 01	.3333333E 01
V(3)	.4008018E 01	.8016032E 01	.7855713E 00	.8412828E 01	.3006012E 01
V(4)	.2365464E 01	.3311650E 01	.2069781E 00	.6623301E 00	.7392073E 01
V(5)	.1037345E 01	.2489627E 01	.7261407E 01	.1659752E 00	.7261410E 00
FOR EXPERIMENT E(4)					
V(1)	.7561443E 00	.1209831E 00	.7939517E 01	.1512289E 01	.2835542E 01
V(2)	.4621075E 01	.7886634E 00	.1207641E 00	.1971657E 01	.2464573E 01
V(3)	.2104054E 01	.4284621E 01	.8569244E 00	.5814845E 01	.2104054E 01
V(4)	.2376709E 01	.3802734E 01	.1330957E 00	.7605469E 00	.4456329E 01
V(5)	.2935423E 01	.1565558E 01	.5479455E 01	.1174168E 00	.7827790E 00

MARGINAL PROBABILITIES OF V BY E

E(0)	.1000000E 01				
E(1)	.1437499E 00	.2299999E 00	.2824999E 00	.2249999E 00	.1187500E 00
E(2)	.1234999E 00	.2314998E 00	.3012499E 00	.2119999E 00	.1317499E 00
E(3)	.1312499E 00	.2249999E 00	.3118749E 00	.2113749E 00	.1205000E 00
E(4)	.1322498E 00	.2028748E 00	.3267499E 00	.2103749E 00	.1277499E 00

SUMMARY TABLE

TEST OF PROGRAM EG MEADOWSLANDS DR.
LINK LENGTH = .9000000E 00 MILES

COST OF SAMPLE MINIMUM COST EXPECTED NET GAIN OF E

E(0) AREAWIDE RATIO	\$.0000000E 00	\$.2387922E 07	\$.0000000E 00
E(1) RATIOS BY MODE & PURPOSE	\$.1000000E 03	\$.2298309E 07	\$.8961300E 05
E(2) ZONAL END RATIOS	\$.4000000E 03	\$.2274708E 07	\$.1132140E 06
E(3) E(2) PLUS MORE DATA	\$.2000000E 04	\$.2240918E 07	\$.1470040E 06
E(4) ZONAL INTERCHANGE RATIOS	\$.6000000E 04	\$.2217959E 07	\$.1699630E 06
MINIMUM COST =	\$.2217959E 07		

PATH IS E(4) TO V(1) TO A(1)
OTHER PATHS ARE

E(0) TO V(1) TO A(1)
E(1) TO V(1) TO A(1)
E(2) TO V(1) TO A(1)
E(3) TO V(1) TO A(1)

STOP 0

END

24 · 12 · 80

FIN