AN ENTROPY MAXIMISING MODEL FOR ESTIMATING TRIP MATRICES FROM TRAFFIC COUNTS

Luis G.Willumsen

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

THE UNIVERSITY OF LEEDS
Department of Civil Engineering

September 1981

ABSTRACT

The main objective of this research is to develop and test a technique for estimating trip matrices from traffic counts. After discussing conventional methods for obtaining trip matrices an analysis is made of the problem of estimating them from traffic counts: it is found that in general the problem is underspecified in the sense that there will be more than one trip matrix which, when loaded onto a network, may reproduce a set of observed counts. A review is made of some models put forward to estimate a trip table from volume counts, the majority of them based on a travel demand model.

A new model is then developed by the author within an entropy maximising formalism. The model may be interpreted as producing the most likely trip matrix consistent with the information contained in the counts and a prior trip matrix if available. This model does not require counts on all links in the network, can make efficient use of outdated trip matrices and other information, and is fairly modest in computer requirements.

The model is then tested against real data collected by the Transport and Road Research Laboratory in the central area of Reading. Considerable temporal variability was found in the sampled trip matrices. The matrices estimated by the model are not very close to the observed ones but their errors are in general within the daily variations of the sampled matrices. A number of tests on the sensitivity of the model to errors and availability of traffic counts and route choice models used are also reported.

A technique has been developed to rank links according to their potential contribution to the improvement of an estimated trip matrix. This scheme may be used to select new counting sites.

The availability of a reasonable prior estimate of the trip matrix considerably improves the accuracy of the origin-destination matrix generated by the model.

Some suggestions for extensions and further research are presented towards the end of this work.

ACKNOWLEDGEMENTS

It is a pleasure to warmly thank the following.

- Professor A.D.May and Dr. D.Van Vliet for their valuable help and advice.
- Dennis Robertson and his staff at the Urban Networks Division of the Transport and Road Research Laboratory for access to the Reading data used in this research.
- The Science Research Council for financial support.
- Mrs. Saskia Fry for her very able secretarial assistance.
- Many other colleagues and friends both within and outside Leeds University for constructive discussion and criticism.

The errors and inadequacies of the work are, of course, the author's own.

CONTENTS

LOCATION OF FIGURES

LOCATION OF TABLES

CHAPTER 1
INTRODUCTION

1.1 OBJECTIVES

The analysis of transport problems and the design of measures to reduce them very often requires an estimation of the number of trips taking place between different points in the study area. This pattern of trips is usually condensed in the form of an origin-destination trip matrix (O-D matrix) representing a measure of the demand for travel in the area.

Conventional methods for estimating trip matrices are fairly expensive, involving considerable resources in terms of manpower, time, disruption to trip makers and data processing. Moreover, their reliability is somewhat suspect and trip matrices thus estimated seem to have a relatively short life.

It is perhaps not surprising therefore that many measures, the design and appraisal of which would benefit from an O-D matrix, are adopted without having recourse to one. A typical example is a local Traffic Management Scheme. This type of project is likely to include measures like introducing one-way schemes, banning certain turning movements at some junctions, banning certain classes of vehicles from some streets, etc. These modifications to the network may result in important changes in the routes followed through the area and these will in turn produce new flow levels, travel times (delays), fuel consumption and expenditure levels for different groups of users. The simple observation that average speeds have increased is not enough to justify a Traffic Management Scheme. Higher speeds might be achieved at the cost of longer routes thus increasing indicators like travel time and expenditure in the system. The only practical way to assess these changes is to follow their effects using an O-D matrix. This is the case whether a model is used to design a Traffic Management Scheme or simply a before-and-after assessment is sought for a specific scheme.

The main objective of this research is to develop and test a technique for estimating trip matrices from traffic counts. These may be either vehicle matrices or person trip matrices as the same concepts also apply to public transport and pedestrians. In principle, at least the same ideas could be applied to freight, telecommunications traffic, and to other countable units on a network.

The possibility of devising ways of estimating a trip matrix from traffic counts seems particularly attractive. For a start, traffic counts are relatively inexpensive to obtain. They are usually collected regularly by local authorities for several purposes, for example intersection design, accident analysis, maintenance planning and simply monitoring flow levels. Secondly, the automatic collection of vehicle counts is well advanced and there are several computer packages which provide efficient pre-processing. Thirdly, counting vehicles, passengers or pedestrians is much simpler than carrying out surveys requiring interviews and formfilling. Finally, most counting operations can be performed without interrupting traffic and causing delays to users.

1.2 OUTLINE OF THIS RESEARCH

This report is organised as follows. Following this chapter, Chapter 2 analyses the problem of estimating a trip matrix, beginning with some basic definitions and notation. The chapter then analyses conventional methods for obtaining a trip matrix, distinguishing between 'direct' and 'indirect' techniques. The next section in Chapter 2 discusses the problem of estimating a trip matrix from traffic counts. It is noted that in general the problem is 'underspecified' in the sense that there will be more than one trip matrix which, when loaded onto the network, reproduces the observed link flows. The use of an appropriate route choice model turns out to be an important element in the whole process. Finally, a quick analysis is made of the errors involved in any estimation process and it is concluded that because of the sparsity of trip matrices, errors are likely to be relatively large, whatever the method used to obtain them.

Chapter 3 contains a review of the methods already proposed to estimate a trip matrix from traffic counts. The chapter considers approaches which use other data in addition to traffic counts, for example population, and are based on a gravity model or a direct demand model. Another group of approaches makes use of network data only. It is noted that most proposed techniques have only been tested against their ability to reproduce the observed flows. Some comments are made about the likely areas of applications of each type of approach.

Chapter 4 describes a model developed by the author and based on an entropy maximising formalism. This model estimates the most likely trip matrix consistent with the observed counts and the hypothesised route choice. The chapter begins with a general presentation of the ideas behind entropy maximisation as a model building tool. Then the basic model developed by the author is presented and its main characteristics are discussed. A simple example and a small simulation program are used to illustrate the ideas behind the model. A formally similar model proposed by Van Zuylen is also presented and the main differences outlined.

Chapter 5 discusses solution methods for the Entropy Maximising Matrix Estimation (ME2) model. First some problems generated by errors in the counts and in the route choice proportions assumed are discussed. One of these problems is that at certain places in the network link flow continuity conditions (i.e. total flow 'into' a node should equal total flow 'out of' it) may not be met. A maximum likelihood solution is proposed to this problem and illustrated with an example. A second section in the chapter sets the ME2 model in the context of three related mathematical programmes, the properties of which may be used to devise solution algorithms. Two algorithms are discussed, one based in the Newton-Raphson method and the other on multi-proportional adjustments. This second technique proved to be more robust and to require less computer resources. The third section in this chapter attempts to identify the potential value of additional traffic counts. An indicator of this value is developed using the same theoretical framework and its use as a ranking criterion for deploying resources for counting traffic is outlined.

Some basic tests on the performance of the model with artificial data is the main theme of Chapter 6. First the problem of comparing trip matrices is outlined and a series of error (or similarity) indicators is selected from work by other researchers. The tests with artificial data proved to be very encouraging and also helped in improving the computer programs.

One of the reasons why it is so unusual for models to be tested against an observed trip matrix is that observed trip matrices are expensive to obtain, in particular if reliability is important. This researcher was fortunate in that the Transport and Road Research Laboratory (TRRL) made available a suitable data base (collected in the central area of Reading in October 1976) for these tests. Chapter 7 describes the main characteristics of this data collection exercise which resulted in a set of trip matrices, traffic counts and observed route choice proportions for four consecutive afternoon peak periods. The analysis of these data showed that the daily variations of the sampled trip matrices was much higher than expected. The daily variations at the link flow level were, nevertheless, within normal ranges.

Chapter 8 is devoted to a series of validation tests of the ME2 model using the Reading data. The tests were performed using different route choice models, complete and incomplete sets of couts and counts with artificial errors in them. On the whole it was found that although the estimated matrices were not very close to the sampled ones, they were roughly within the range of their daily variations. In order to test the practical implications of using a matrix estimated using ME2 rather than an observed one, both were loaded onto a modified network and the resulting flow patterns compared.

Chapter 9 reports on attempts which were made to incorporate an 'equilibrium assignment' route choice model which should be better adapted to the congestion levels prevailing in central Reading. Although a minor improvement in the accuracy of the model was obtained the results were disappointing as this approach requires considerably more computer time.

Chapter 10 describes an improved approach which tries to overcome some of the difficulties encountered in Chapter 9. This new scheme tackles the problem from the point of view of path flow estimation so that the trip matrix results from an aggregation of path flows. Results with this approach were better than with the previous method but it is recognised that both schemes are essentially heuristic and that there is scope for further, and more theoretically robust, improvements.

All the tests to this stage were performed without assuming any prior knowledge of the trip matrix itself, despite the fact that the ME2 model can incorporate this very efficiently. Chapter 11 explores some of the issues involved in the use of prior information. Tests were made to ascertain the best size for the 'seed' to be used on zero cells when a prior trip matrix is very sparse. Tests using the observed matrix of one day as a prior matrix for another day or an aggregate of the four days, were fairly successful. These results suggest that the ME2 model could be used to advantage to update trip matrices or to estimate a trip matrix using a combination of conventional survey and traffic counts.

Finally, Chapter 12 summarises the research findings and conclusions, with emphasis on their implications for further theoretical and practical work.

CHAPTER 2
THE ESTIMATION OF A TRIP MATRIX

A discussion of the main issues involved in the estimation of a trip matrix is the main theme of this chapter. Section 2.1 includes a definition of the most important elements of the problem and introduces the notation used throughout this work. Section 2.2 is devoted to an appraisal of conventional techniques to estimate a trip matrix either through direct observation or by means of a travel demand (distribution) model. Section 2.3 presents the characteristics of the problem of estimating a trip matrix from traffic counts. The role of different kinds of route choice models in this estimation problem is discussed in Section 2.4. Finally, Section 2.5 is devoted to the analysis of the sources of error in conventional techniques and some evidence about their relative accuracy is presented.

2.1 DEFINITIONS AND NOTATION

2.1.1 Basic elements

Throughout this work use is made of many familiar concepts employed in conventional (aggregate) transport modelling. No attempt is made to redefine them but the concepts which are most relevant to this dissertation are described below together with their notation. As some additional definitions and variables will be introduced later, a more extensive account of the notation used appears in Appendix A1 in alphabetical order.

For the purpose of this work a study area will be divided into zones of more or less homogeneous character. Each zone has a centroid associated with it and all trips are assumed to originate and terminate at them. The study area can also be considered as divided into two sub-areas: an internal one representing the area of interest itself, and an external one representing the rest-of-the-world in so far as it affects the internal transport system. Internal and external zones belong to these two sub-areas.

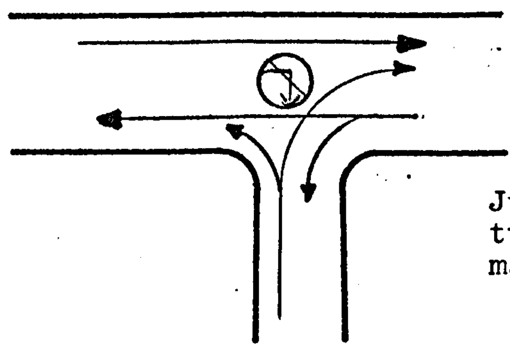
The road network is represented by a set of N <u>nodes</u> and a set of L <u>links</u>. Nodes are usually associated with points of interest in the network such as junctions or parts of them, and they are consecutively labelled from 1 to N. A link, or arc, is represented by an ordered pair of nodes, for example (d,f) if there is a link from node d to node f. Links are always one way and for some purposes it is useful to associate consecutive numbers to them $\ell = 1,2,..,L$. A link may be used to represent a particular (and homogenous) stretch of road or simply a turning movement. Figure 2.1 depicts two ways in which an intersection may be modelled as a combination of links and nodes.

Several properties can be associated with links and the most relevant to this work are:

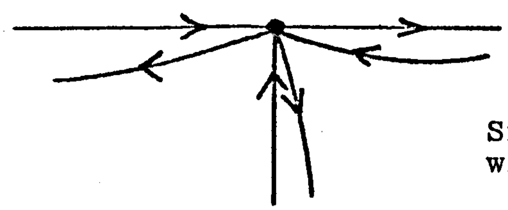
<u>distance</u>, $d_{\ell m}$, usually measured in metres
<u>speed</u>, $s_{\ell m}$, measured in km/h
<u>travel time</u>, $\mu_{\ell m}$, usually measured in minutes
<u>travel cost</u>, $c_{\ell m}$, usually a weighted combination of travel time and distance
<u>flow or traffic volume</u>, $V_{\ell m}$, measured in pcu or vehicles per hour
a <u>cost-flow</u> relationship, $c_{\ell m}(V_{\ell m})$, a function of the amount of traffic using link ℓm relating travel cost on the link to link volumes.

A special type of arc, a centroid connector, is used to link centroids with real nodes in the network. The cost associated with their use is deemed to represent the average cost of travelling over the local streets from the origin before joining the main street system. This cost is normally considered to be independent of the traffic flow.

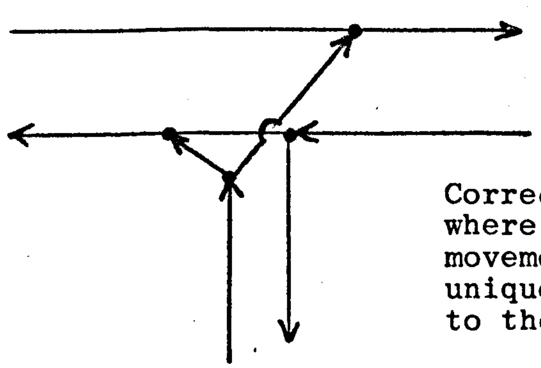
A trip between i and j will use a particular sequence of links called a <u>path</u> or <u>route</u> and T_{ijr} will be the trips from i to j which use route r. The cost of travelling along this route is the sum of the costs of the individual links used and will be represented by C_{ijr}. The variable $\delta_{ijr}^{\ell m}$ can be used to identify links used by



Junction with right-turning banned in main street



Simple representation without banning turn



Correct representation where all turning movements have a unique link associated to them

Figure 2.1: Node and link representation of junctions

route r between origin i and destination j. It is defined as

$$\delta_{ijr}^{\ell m} = \begin{cases} 0 & \text{if link } \ell m \text{ is not used by route } r \text{ between } i \text{ and } j \\ 1 & \text{if link } \ell m \text{ is used by route } r \end{cases}$$

The route cost is then

$$C_{ijr} = \sum_{\ell m} C_{\ell m} \, \delta_{ijr}^{\ell m} \; .$$

The relationship between these variables and traffic assignment methods is described in Section 2.4.

2.1.2 Trip matrices

The number of trips going from origin i to destination j is represented by T_{ij} and the complete set of trips covering all the centroids in the area constitute the trip matrix $[T_{ij}]$. Thus a trip matrix is simply a representation of the tripvolumes moving between pairs of zones. It depends on two forms of aggregation, spatial and temporal. Spatial aggregation involves the grouping of areas into discrete spatial units or zones. In large scale modelling exercises the number of zones can be in the thousands, whereas for small scale traffic management schemes 25 to 50 zones may be enough. But even in this latter case the number of cells is quite large (2500 for the 50 x 50 case) and most of them are likely to contain zeroes or small numbers. Trip matrices are therefore fairly sparse matrices.

Temporal aggregation is concerned with the time interval during which trips between zones are considered. The choice of this time slice or interval has a major impact on the O-D matrix and its usefulness in the analysis of particular problems. Detailed analysis of a system of saturated traffic signals require time slices of the order of 15 minutes to be able to follow the build-up and decrease of queue lengths and travel times. On the other hand most traffic management problems require discrimination at the hourly traffic level only and many problems involving new road construction could be handled with O-D matrices based on 16 or 24 hours. Indeed, certain analyses such as inter-city demand studies, use weekly, monthly or even yearly trip matrices.

Trip matrices based on a small time slice present some particular problems. For a start, it may be difficult to allocate unequivocally trips to matrices. For example, consider the difficulty of allocating a trip starting at 9.10 a.m. and ending at 9.19 a.m. (in the study area) to either the 9.00 to 9.15 or the 9.15 to 9.30 a.m. O-D matrix. It is desirable to use time slices larger than the average trip length in the area. A second problem is sparsity. The smaller the time slice, the greater will be the number of cells containing zero trips in the trip matrix. It will be seen later that this sparsity is also associated with high levels of error in the matrix.

The objectives of a particular study requiring O-D information will help to define the spatial aggregation and the time interval for the trip matrix of interest. But even then, another consideration has to be borne in mind.

Trip matrices are subject to hourly, daily, weekly and seasonal variations over time in the same way as traffic counts or trip rates. One can think of a distribution of O-D matrices over, say, a year and it will depend again on the objectives and resources of the study which matrix can be said to be the 'trip matrix of interest'. Some of the alternatives to be considered are:

- the average trip matrix (or mean)
- the most likely trip matrix (mode)
- the critical trip matrix, perhaps the one that produces maximum congestion
- the trip matrix generating <u>average delay</u>
- the trip matrix corresponding to the 29th/30th most congested hour.

The problems and costs of data collection for estimating an O-D matrix are such that most of these issues are usually overlooked in a study. It appears that in practice almost any O-D matrix would do provided it has been obtained through an accepted technique and it has no obvious inconsistencies. This is borne by the fact that this matrix is later referred to as the O-D matrix for the area rather than the estimated mean/mode/critical/average delay/30th hour trip matrices.

2.2 CONVENTIONAL METHODS FOR ESTIMATING A TRIP MATRIX

There are many transport problems whose resolution requires the identification of the spatial travel pattern over the area of interest. The most common representation of this spatial travel pattern is by means of trip matrices and their use ranges from large scale modelling of nationwide transport systems to before-and-after studies of local traffic management schemes. It is not surprising then that a variety of methods for estimating an O-D matrix have been put forward and used. This section provides a brief overview of these conventional techniques, attempting to identify their strengths and limitations. For a more detailed and technical review of the methods the reader is referred to publications such as Road Research Laboratory (1965), and US Federal Highway Administration (1975), or Hutchinson (1974). Three groups of techniques are considered here: first, direct methods such as road side interviews which use direct measurement of trip matrices; second, indirect or synthetic methods like a distribution model, which use other data to infer a trip matrix; and finally, hybrid methods which attempt to combine both approaches.

Most of these techniques have been developed in order to obtain more than just one type of information about trips in a study area. Home interviews, for example, are used to provide information about trip rates and purposes, trip length distributions, modal choice and so on, in addition to an estimated trip matrix. More attention will be paid in this review to methods specifically designed to estimate trip matrices.

2.2.1 Direct methods

2.2.1.1 Home/workplace interviews. This method is employed in conventional transport planning studies for large towns, major conurbations and regions. It is a fairly expensive technique, usually involving large numbers of staff carrying out interviews in a selected number of households and/or workplaces. The data thus collected covers a wide range of issues and variables, but information about the origin and destination of trips is given a good deal of attention. Because of the large cost of collecting and processing home interview data only a sample of all households is surveyed. Sampling rates normally range between 1 to 10 per cent but may be as high as 20 per cent*. As information of different sorts and for a variety of purposes is collected the sampling rates are a compromise between these objectives and survey costs. A recent and fairly practical discussion of the design of home interview sampling rates can be found in Smith (1979).

A less expensive technique is to use self-response or postcard methods.

In this case the individual fills in a (much simpler) questionnaire without an interviewer being present and the form is either collected later or returned (pre-paid) through the post. Response rates are often poor and it is much more difficult to keep quality control of the replies.

2.2.1.2 Roadside interviews. This technique requires motorists to be stopped and questioned regarding their origin and destination and other trip data. These interviews usually take place on the road at cordon or screenline points. Practical considerations restrict the

---

* The sampling rate here refers to the ratio between the number of units or elements surveyed and the total (maximum) number which could have been surveyed using the method in question. A sample rate of 10 per cent in home interviews represents one interview in every 10 households. The same rate in roadside interviews is one interview for every 10 vehicles passing the survey point in the survey direction on the interviewing day(s).

number of sites at which these interviews can be held. The number of vehicles to be stopped and the length of the interview are also limited by considerations of safety and inconvenience to road users. Sampling rates between 0.1 and 0.4 seem to be typical in urban areas. In-vehicle surveys and sometimes interviews at major interchanges are the analogous methods for public transport users. Self response questionnaires may also be used.

2.2.1.3 Flagging methods. These include a variety of techniques based on unequivocal identification of the entry and exit (and sometimes intermediate) points of randomly selected vehicles to the study area. The methods require observers located at each of these key points and some type of flag to identify vehicles. Sometimes coloured and numbered stickers may be used as codes for each entry point. These are attached on entry and recorded at intermediate and exit points. More often the registration number is used instead of a physical tag and this has the advantage of minimising drivers' inconvenience. Computer based and manual techniques exist to match observations at different sites to produce a trip matrix; see for example Clarke and Davies (1970), and Dawson (1979).

A variation of the flagging method has been put forward by Bebee (1959) and used in roundabout studies. It relies on asking drivers at one entry point of a small study area to switch their headlights on for a fixed period of time. Observers at key points then record the number of vehicles with their lights on for given intervals of time. The process is then repeated for different entry points on successive days. The method can only be used during daylight and for small study areas.

2.2.1.4 Vehicle following methods. This method requires observers to follow vehicles through the study area recording its passage through key points in the network. It has successfully been used by its proposer in the Westminster area of London, using taxicabs as 'followers'; see Wright and Orrom (1976), and Wright (1978). This method seems more appropriate for route choice than origin-destination studies and probably is only advantageous in large and busy central areas.

2.2.1.5 Aerial photography. This method is based on time-lapse aerial photography of a study area from a stationary helicopter hovering (ideally) at a fixed altitude; see for example Garner and Mountain (1978) and (1979). The data collection stage is quite fast and inexpensive compared with alternative methods but this is achieved at the expense of processing effort. The method still requires following individual vehicles frame by frame through the study area using a human observer and digitizing equipment. It has the advantage that quality control can be carried out fairly easily and that the photographs can be used for other purposes, for example parking surveys. This method will certainly become more attractive when automatic identification of vehicles in the frames can be achieved. In principle a sampling ratio of 100 per cent for the survey period could be obtained with this method, but practical reasons restrict sampling ratios to values similar to roadside interviews.

All the methods described are fairly expensive in either manpower or processing effort or both. In addition to this, interview methods require the co-operation of the tripmaker and, at least in the case of roadside interviews, imply disruption and delays. Interview methods also suffer from reporting errors of which the omission of intermediate stops in a journey is a common one and may be a significant hindrance in O-D studies. Registration number surveys are attractive because of their non-disruptive nature but they require both large numbers of observers and processing effort and are sensitive to mis -recording. A recent variation on the registration number method involves recording number plates at intermediate points providing data for route choice and checking purposes (Robertson, 1979, and Cathcart and Fearon, 1980). This variant provides better information at an increased cost.

All methods imply sampling and as such they can only provide an estimation of the trip matrix for the survey period. Even if a 100 per cent sample ratio is achieved for a given time period (with aerial photography for example), the problem still remains of estimating how close this matrix is to the trip matrix of interest for the study objectives.

2.2.2 Indirect and hybrid methods

Indirect methods are used to estimate an O-D matrix by means of a trip-distribution model of some kind. It is beyond the scope of this work to discuss the wide range of models and practices used to achieve this in any great detail. Only those points most relevant to this research will be touched upon here.

The main alternatives in this group of techniques are:

- to calibrate and use a conventional distribution model as part of a larger modelling study; the gravity model is a frequent choice here,
- to use a trip distribution model with parameters 'borrowed' from other studies,
- to calibrate an 'ad hoc' gravity model using partial matrices techniques.

Finally, hybrid methods combine a direct technique with one of the possibilities mentioned above.

2.2.2.1 Conventional distribution models. Conventional transport models are usually structured as a sequence of sub-models, namely

- trip generation, where the number of trips originated and attracted to each zone are modelled,
- trip distribution, where the number of trips between each origin and each destination are synthesised,
- modal split modelling choice of mode of travel,
- traffic assignment when route choice is modelled.

The sub-model of interest here is the distribution model and a usual choice is to use a gravity model at this stage. This choice immediately restricts the functional form of the model. Most recent studies in the UK have used a form following Wilson (1970)

$$T_{ij} = O_i D_j A_i B_j f(c_{ij}) \qquad (2.1)$$

where O_i and D_j are total number of trips generated and attracted to zones i and j, A_i and B_j are balancing factors calculated as

$$A_i = \frac{1}{\sum_j D_j B_j f(c_{ij})} \quad ,$$

$$B_j = \frac{1}{\sum_i A_i O_i f(c_{ij})} \quad ,$$

and $f(c_{ij})$ is a deterrence function or measure of separation with at least one parameter for calibration. A textbook choice is $f(c_{ij}) = e^{-\beta c_{ij}}$ with one parameter β.

The parameters of $f(c_{ij})$ are usually calibrated so that the model produces a trip length distribution which is as close as possible to the one obtained from survey data. In principle at least, the parameters of $f(c_{ij})$ are specific to each study area.

The use of this approach requires:

- data on trip ends O_i and D_j
- data on trip length distribution
- the assumption that the gravity model reasonably represents tripmaking behaviour in the area.

It should be noted that this model will produce a non-zero number of trips in each cell albeit some of them will contain a very small number and may indeed be rounded down to zero in some computer packages.

2.2.2.2 Cordoning sub-matrices. It is possible to make use of a larger trip matrix produced by a conventional method for a wider area and to reduce it to make it applicable to a smaller, cordoned area. The area of interest is 'cordoned' and the original O-D matrix is assigned to the network. The point where these assigned trips cross the cordon are considered to be the generators/ attractors of the reduced matrix. The zones internal to the cordoned area are usually kept.

Many traffic assignment packages, for example the TRADVV suite in Leeds, have facilities for cordoning a trip matrix in this way. The value of the resulting sub-matrix depends on the quality of the large O-D matrix, the assignment model used and the cordoning policy.

2.2.2.3 Distribution models with 'borrowed' parameters. This constitutes one of the practices advocated in the United States of America for small urban areas (Federal Highway Administration, 1977). Conventional distribution models require a good deal of data collection and a short cut is to use parameters calibrated for other areas. Synthetic trip end rates and rules of the thumb for borrowing values for the parameter β or equivalent are provided, for example, in Grecco et al (1976) and Soslau et al (1978). These methods are now becoming more elaborated (Pigman and Deen, 1979), perhaps losing some of their advantages of simplicity.

2.2.2.4 Partial matrices techniques. Partial matrices techniques have been developed mainly in the UK in order to synthesise a trip matrix using incomplete data. The techniques attempt to calibrate a variation of the gravity model in Equation (2.1) using only partial information about trip ends and trip lengths, obtained through a less expensive survey method, usually roadside interviews (Neffendorf and Wootton, 1974).

The technique is attractive in its survey cost saving potential, but apparently many questions still remain regarding, among other issues, the errors involved and the choice of good survey patterns; see Kirby (1979) and Day and Hawkins (1979).

2.2.2.5 Hybrid methods. Combinations of direct and indirect methods have been suggested in order to reduce the sparsity of the trip matrix obtained from direct observations. During the West Yorkshire Transportation Study the trip matrices obtained from roadside and 'in vehicle' interviews were complemented by the use of a distribution model, WYTCONSULT (1977). Due to the location of the screen lines for roadside interviews, many O-D pairs were not observable while many other cells were also empty probably due to underreporting of journeys. A simple gravity model calibrated from household interview data was then used to 'in fill' these empty cells. WYTCONSULT claim that a better 'base year' O-D matrix is obtained in this way, in particular for short term analysis of problems and alternative solutions.

This approach may be considered suitable where the study is large enough to provide sufficient data to calibrate the distribution model used to 'in fill' cells. A related approach is to apply correction factors to an O-D matrix so that when assigned it reproduces the observed flows crossing a screen line. For example, traffic counts on the River Thames crossings have been used in the GLTS studies to correct the corresponding trip matrices.

2.2.3 Comments

Despite the variety of methods for obtaining a trip matrix the state of the art is far from satisfactory. Dial (1973) gave the following description of a typical exercise in collecting information for an origin-destination matrix.

> "... after three months of interviewing, a truckload of interviews is entered into an archaic data processing chain. Months of keypunching and verifying move into months of edit checking. Zone numbers are related to addresses. More checking follows more fixing. A year

> later a factoring process begins and is followed by other accuracy checks and general wholesale handwriting on why census numbers and survey numbers do not match, and on and on......"

Data processing has improved since 1973 but on the whole Dial's is a good description of the resources, time and errors involved in the use of most conventional methods. It is not surprising then that many traffic management schemes, for example, are designed without recourse to an O-D matrix. The idea of devising techniques for synthesising trip matrix information from more easily available and less expensive data seems to be a promising proposition.

Local authorities and the TRRL among others have permanent programmes for regular counting of certain links. These counts are used to help programming roadworks and maintenance, redesigning junctions and monitoring traffic levels in general. In addition to this it is part of any Urban Traffic Control installation to monitor traffic levels, often on-line, in the area covered by the scheme.

These sources, plus ad hoc traffic counts where required, are likely to produce a wide and inexpensive data base which has not been fully exploited so far. The possibility of using these relatively inexpensive data to synthesise origin-destination trip matrices appears particularly attractive.

2.3 THE PROBLEM OF ESTIMATING A TRIP MATRIX FROM TRAFFIC COUNTS

In recent years a number of models have been proposed to estimate an origin-destination trip matrix using traffic counts on road links. These models are reviewed in Chapter 3, but this section discusses the general problem of estimating a trip matrix from counts.

2.3.1 Statement of the problem

Consider a study area which has been divided into M zones, each one with its corresponding centroid. The road network has been coded into N nodes and L one-way links. Traffic counts are available for $L_c \leq L$ of these links. The level of resolution of the zoning system and road network will depend on the objectives of the study and the modelling technique adopted to meet them.

It is assumed that the trips of the O-D matrix $[T_{ij}]$ use the links of the network and that the observed flow levels are the result of the passage of these trips over the counted links. A key issue in the estimation of a trip matrix from counts is the identification of the origin-destination pairs whose trips use a particular (counted) link. Throughout this work the variable $p_{ij}^{\ell m}$ will be used to this end. The variable is defined as the proportion of trips between origin i and destination j which use link ℓm. In general

$$0 \leq p_{ij}^{\ell m} \leq 1$$

where the extreme values occur when the link is not used by any trips from i to j and when all those trips travel over the link ℓm. The 'fundamental equation' in the estimation of a trip matrix from traffic counts can then be written as

$$V_{\ell m} = \sum_i \sum_j p_{ij}^{\ell m} T_{ij} \quad . \tag{2.2}$$

The value of the variable $p_{ij}^{\ell m}$ is clearly related to route choice and hence the values estimated by the modeller must depend on the type of assignment model used to represent route choice in the study area. This issue is discussed further in the next section. For the time being it is assumed the $[p_{ij}^{\ell m}]$ can be obtained independently from the O-D matrix estimation process.

The problem can now be described as follows.

Given a study area with M zones, a set of traffic counts $[V_{\ell m}]$ on (some of) the links in the network and a set of (assumed) values $[p_{ij}^{\ell m}]$ representing route choice between each O-D pair, estimate the underlying trip matrix from a set of linear equations

$$\sum_{ij} p_{ij}^{\ell m} T_{ij} = V_{\ell m} \quad . \tag{2.3}$$

This is a problem with M^2 unknowns T_{ij} and as many linear equations (2.3) as observed traffic counts $V_{\ell m}$ are available, say L_c links with counts. In practically all study areas M^2 will be much greater than the total number of links in the network. * Under these conditions the problem is underspecified and in general there will be more than one trip matrix satisfying the L_c linear equations (2.3).

A possible way of explaining this lack of uniqueness of the solution to the problem in physical terms is as follows. Once two or more different 'strands' of trips, e.g. two different T_{ij} trips or routes, come together on the same link there is no way that a count can distinguish between them. One can envisage the theoretical situation in which all the drivers on a certain link decide to get together and re-allocate their destinations amongst themselves, thus producing exactly the same counts on the rest of the network but a different trip matrix.

For practical reasons it is desirable to estimate a single O-D matrix from a given set of counts. A mechanism must be found then to reduce the number of unknowns of the problem so that it becomes fully specified. Assumptions about tripmaking behaviour will play that role by restricting the independence of the variable T_{ij} . The methods reviewed in Chapter 3 differ basically in the assumptions adopted to achieve this.

It is opportune, at this stage, to analyse some of the problems that may be associated to the problem as stated.

---

* Although one can set up artificial examples with more links than O-D pairs (see Section 2.3.5) this does not tend to happen in practice. For example Van Vliet (1978) reports on 5 real networks coded for assignment purposes and the ratio of links to O-D pairs range for them between 0.015 and 0.5.

2.3.2 Independent and inconsistent counts

There are two additional considerations to bear in mind regarding the solution to the problem represented by Equations (2.2). The first one is that not all the L_c equations necessarily add new information, i.e. some of them may be redundant. The second is that it is possible that some of the Equations (2.2) may be incompatible and then the problem has no solution. These considerations will be described in physical terms and then supported by standard results from linear algebra.

2.3.2.1 Independence. Not all traffic counts contain the same amount of 'information'. For example in Figure 2.2 traffic link 5-6 is made up of the sum of traffic on links 1-5 and 2-5. Counting traffic on link 5-6 is then redundant and only two counts there can be said to be independent.

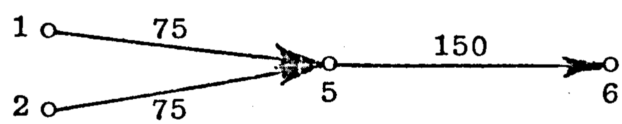


Figure 2.2: Dependent counts

Wherever a flow continuity equation of the type 'flows into' a node equals 'flows out of' the node linear dependence of equations will tend to occur. In this case it will always be possible to describe one link flow as a linear combination of the rest. Note that a centroid connector attached to node 5 will remove the dependency in Figure 2.2.

2.3.2.2 Inconsistency. Counting errors and asynchronous counting are likely to lead to inconsistencies in the flows. In other words, the expected continuity relationships will not be met.

If the flow on link 5-6 in Figure 2.2 were to be 160 instead of 150 the corresponding equations would be inconsistent and no trip matrix could satisfy these equations. Some of the ways of reducing this problem are to introduce an error term in Equations (2.3) or to remove the inconsistencies beforehand.

2.3.3 Solving a system of linear equations on flows

Some concepts of linear algebra may shed additional light onto the above problems, see for example Goult (1978). It is possible to discuss the problem in terms of two matrices and two vectors.

The vectors are column vectors containing the observed traffic volumes for the counted links $V_{\ell m}$. It can be represented as

$$\begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_{\ell c} \end{bmatrix} \quad \text{or} \quad \underline{V} \tag{2.4}$$

It is convenient to represent the trip matrix as a vector of M^2 unknowns $\underline{T}$.

$$\begin{bmatrix} T_{11} \\ \vdots \\ T_{ij} \\ \vdots \\ T_{mm} \end{bmatrix} \tag{2.5}$$

The first matrix is the matrix of coefficients $[p_{ij}^{\ell m}]$ or matrix $\underline{\underline{P}}$

$$\begin{bmatrix} p_{11}^{\ell 1} & p_{12}^{\ell 1} \cdots\cdots p_{1m}^{\ell 1} \\ \vdots & \\ p_{m1}^{\ell c} & p_{m2}^{\ell c} \cdots\cdots p_{mm}^{\ell c} \end{bmatrix} \tag{2.6}$$

?

This matrix has L_c (counted links) rows and M^2 (O-D pairs) columns.

The second matrix is the extended matrix [P,V] or $\underline{\underline{Pv}}$

$$\left[\begin{array}{c|c} p_{11}^{\ell 1} \cdots\cdots\cdots\cdots p_{1m}^{\ell 1} & V_{\ell 1} \\ \vdots & \\ p_{m1}^{\ell c} \cdots\cdots\cdots\cdots p_{mm}^{\ell c} & V_{\ell c} \end{array}\right] \tag{2.7}$$

This is obtained by adding the column vector $\underline{V}$ to the matrix $\underline{\underline{P}}$ and it has L_c rows and M^2+1 columns.

The system of linear equations is now represented by

$$\underline{\underline{P}}\ \underline{T} = \underline{V} \tag{2.8}$$

and the solution, according to the theory of linear equations (Goult, 1978) depends on the ranks of the matrices $\underline{\underline{P}}$ and $\underline{\underline{Pv}}$.*

Three situations are possible.

(a) If P is an $M^2 \cdot L_c$ matrix of rank M^2 the fundamental equations $\underline{\underline{P}}\underline{T} = \underline{V}$ are soluble for all constants $V_{\ell m}$ and this solution is unique. It has already been mentioned that this condition is unlikely to occur in practice for this type of problem.

(b) If $\text{rank}(\underline{\underline{P}}) < M^2$, a solution only exists if $\text{rank}(\underline{\underline{P}}) = \text{rank}(\underline{\underline{Pv}})$ and in this case infinitely many distinct solutions will exist. The rank of the matrix $(\underline{\underline{P}})$ may be interpreted as the maximum number of linearly independent equations in (2.8). If $\text{rank}(\underline{\underline{P}}) < L_c$ (number of links with counts) some rows in $\underline{\underline{P}}$ may be generated as linear combination of other rows. The fact that $\text{rank}(\underline{\underline{P}}) = \text{rank}(\underline{\underline{Pv}})$ implies that these linear combinations will also be applicable to the extended matrix $(\underline{\underline{Pv}})$ and the equations will be consistent. The solution though will not be unique.

(c) $\text{Rank}(\underline{\underline{P}}) < M^2$ and $\text{rank}(\underline{\underline{P}}) \neq \text{rank}(\underline{\underline{Pv}})$. In this case the system of equations is inconsistent and has no solution. A row in $\underline{\underline{P}}$ may be expressed as a linear combination of other rows but these operations do not apply equally to the extended matrix. One of the sources of this inconsistency is the violation of the link-flow

---

* The rank of a matrix is defined as the maximum number of linearly independent column vectors (or row vectors) in it and it is thus related to the independence of the original linear equations.

continuity conditions. A second source resides in the relationship between $[p_{ij}^{\ell m}]$ and $[V_{\ell m}]$, when the observed counts and the implied traffic assignment model are incompatible. A simple case of this would be a link with non-zero observed flow but where all $p_{ij}^{\ell m} = 0$ because the implied assignment model does not load any trips onto it. In practice though, the inconsistencies may be of a more subtle nature which can only be detected by matrix operations. As this second source of inconsistency is related to the interplay of flows on assignment it will be referred to in the future as the 'path flow' continuity conditions in contrast with the first and more obvious source of 'link flow' continuity equations.

In the future the number of independent traffic counts in a network will be identified by L_c' and in general $L_c' \leq L_c \leq L$.

2.3.4 Intrazonal trips

Trips which have their origin and destination in the same zone cause some problems in conventional models, traffic assignment and in the estimation of trip matrices from counts. The problems originate in the fact that most of the trip length of an intrazonal trip takes place outside the coded real links. This makes it difficult, for example, to calculate proper trip costs for a distribution model or to allocate them to the network in the assignment stage.

The extent of the problem depends on the level of resolution of the zoning system and network representation. One of the considerations in choosing this level is usually the type of trips which can safely be disregarded, at least in terms of network loadings. For example trips from an external zone to itself will not appear in the study area and can then be disregarded.

Any study attempting to estimate a trip matrix from traffic counts should consider the difficulties in assigning intrazonal trips onto the network and choose a level of resolution such that intrazonal trips can safely be disregarded. However, it does not imply much loss of generality to assume, as we do from now on, that all $T_{ij} = 0$ for $i = j$ for a suitable level of detail of the network and zoning representation. This new condition reduces the number of unknowns from M^2 to M^2-M but does not affect any of the considerations discussed in this chapter.

2.3.5 An example

It is possible to illustrate some of the issues in this section by means of the simple network in Figure 2.3..

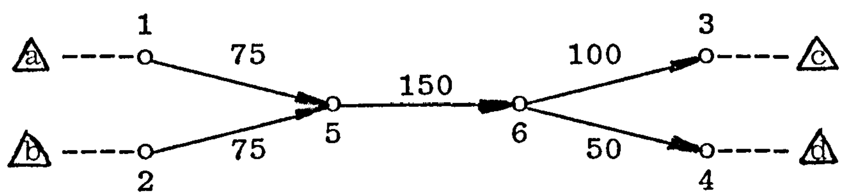


Figure 2.3: A simple network with traffic counts

This network has 2 origins (a and b), two destinations (c and d), 4 centroid connectors (a1, b2, c3, d4) and 5 links (1-5, 2-5, 3-6, 4-6, 5-6). The counts on the 5 links are depicted in the figure, for example $V_{25} = 75$ trips. In this example the 5 linear equations are

$$T_{13} + T_{14} - 75 = 0 \tag{2.9}$$

$$T_{23} + T_{24} - 75 = 0 \tag{2.10}$$

$$T_{13} + T_{23} - 100 = 0 \tag{2.11}$$

$$T_{14} + T_{24} - 50 = 0 \tag{2.12}$$

and $$T_{13} + T_{14} + T_{23} + T_{24} - 150 = 0 \quad . \tag{2.13}$$

There are 5 equations for 4 unknowns; these equations are compatible but only three of them are independent. For example Equation (2.13) results from combining Equation (2.9) and Equation (2.10) and Equation (2.12) from Eq(2.9) + Eq(2.10) - Eq(2.11) . If the matrices $\underline{\underline{P}}$ and $\underline{\underline{Pv}}$ of this system of linear equations are formed their rank will be 3.

If the condition $T_{ij} \geq 0$ is imposed on the O-D matrix and only integer values are accepted for each cell then there are in this case 51 different trip matrices which satisfy Equations (2.9 to 2.13). These 51 matrices can be quite different and the ranges for each cell are depicted below.

cell	from	to
a - c	25	100
a - d	0	50
b - c	25	75
b - d	0	50

These 51 matrices are discussed further in Section 4.3. In general the number of (integer) matrices which solve a given underspecified problem such as this one is very large depending on the number of unknowns, the number of independent counts and their absolute values. How to choose among these matrices is a central theme of this research.

2.4 THE ROLE OF TRAFFIC ASSIGNMENT

It was shown in the previous section that the variable $p_{ij}^{\ell m}$ plays a keyrole in the estimation of a trip matrix from counts and that this variable is related to traffic assignment. This section will briefly describe current techniques for modelling drivers' route choices discussing their relevance to this research. The two main sources of difficulty in modelling route choice seem to be the differences between the behaviour of different drivers and the effect of congestion on link costs. It is known that different drivers perceive travel costs differently; some may give more importance to time than to distance, or vice versa, others may be more concerned with reducing risks, and so on. Drivers pursuing different objectives in their route choice or perceiving route attributes differently will tend to generate a number of distinct routes between an origin and a destination. On the other hand, the fact that the cost on a given link depends on the volume of traffic using it has been the starting point of a family of capacity restraint assignment models. The actual model to choose for a given study area will depend on the relative importance of these two factors.

2.4.1 All-or-nothing assignment

This is the simplest and fastest (in CPU time) assignment model. It assumes that all drivers try to minimise the same type of cost (travel time, travel distance or a combination of both) and that they all perceive these attributes in the same way. With a unique definition of costs there are a number of computer programs to calculate the minimum cost paths, one for each origin-destination pair. For a discussion see Van Vliet (1978). All-or-nothing assignment models load all the traffic in each O-D pair to its single minimum cost path.

This is a very efficient and easy to use assignment model and as reported by Lai and Van Vliet (1979) it is the loading method most commonly used by British Local Authorities. But there is considerable empirical evidence that this model cannot explain the route choices of all drivers, see for example Wright and Orrom (1976) and Outram and Thompson (1978).

Van Vliet (1976) suggested the use of the dimensionless parameter E_1 to represent the effect of the variation in drivers' perceptions. The parameter is defined as

$$E_1 = \frac{\sum_{ijr} \hat{T}_{ijr}\, \hat{c}_{ijr} - \sum_{ij} T_{ij}\, c_{ij}^{min}}{\sum_{ij} T_{ij}\, c_{ij}^{min}} \qquad (2.14)$$

where $\hat{T}_{ijr}$ = observed number of trips between i and j via path r

$\hat{c}_{ijr}$ = observed route cost between i and j via r

c_{ij}^{min} = minimum cost between i and j

and a single definition of 'cost' has been set by the modeller.

The parameter E_1 measures then the degree of dispersion of routes which can be associated with variations in drivers' perceptions. Van Vliet (1976) suggests that a value of E_1 around 10 per cent may be taken as representative.

2.4.2 Stochastic assignment

Stochastic assignment models attempt to represent the effect of varying drivers' perceptions on route choice. The most commonly used stochastic assignment model in Britain has been developed by Burrell (1968). This method identifies a number of distinct routes for each O-D pair by allowing link costs to vary in some random manner about their pre-determined means. The number of routes to be calculated and the degree of 'noise' around the mean link costs are controllable by the user.

A second group of stochastic assignment models relies on 'splitting functions'. Every time a group of trips reaches a node from which there are at least two alternative paths to reach their final destination, the traffic is split among these paths according to some function of their respective costs. A number of different functional forms have been put forward for the splitting function by, among others, McLaughlin (1966), Murchland (1968), Outram (1972), Dial (1972) and Robertson (1977). The most widely available model of this type is Dial's which uses the splitting function f_i defined as

$$f_i = e^{-\theta \Delta c_i} \tag{2.15}$$

where Δc_i is the extra cost incurred in travelling through the next node instead of the minimum cost path from the current node, and

θ is a parameter controlling dispersion of routes.

The model also uses a rule to limit trips to 'reasonable routes' so that travellers are continually choosing links that are taking them away from the origin. This rule also has certain operational advantages from a programming point of view.

It should also be noted at this stage that Burrell's routes are subject to statistical sampling errors whereas the Dial-type methods are 'deterministic' in the sense that no random numbers are used in their derivation. The 'splitting' function is a good substitute to route enumeration thus saving computer time and core but at the same time making it more difficult to calculate $[p_{ij}^{\ell m}]$ s.

Ratcliffe (1972) calculated travel costs on reported routes for work journeys in Newcastle and compared them with the corresponding (objective) minimum costs. He found that all-or-nothing assignment would have allocated correctly about half of the trips with approximately 75 per cent of the journeys choosing routes less than 10 per cent above the minimum.

2.4.3 Congestion effects

All-or-nothing and stochastic assignment methods tend to overlook the interaction between link volume and link cost (speed). Capacity restrained assignment tries to restrict the volumes assigned to any link by its capacity and to relate the cost on a link to the assigned flow via cost-flow functions $c_{\ell m}(x)$. Various methods for capacity restrained assignment have been proposed. For example 'repeated assignments' adjust link costs according to the most recent assignment and then reassign all trips. Under certain conditions these adjustments may lead to oscillations in costs and volumes. Incremental loading and multipath iterative assignment methods have also been developed to cope with this problem.

Incremental assignment requires the O-D matrix to be loaded in fractions and link costs are adjusted after each assignment.

Iterative loading follows a modified approach. In each iteration the full O-D matrix is loaded but the resulting volumes are linearly combined with the flows from the previous iteration.

$$V_{\ell m}^{(n)} = \lambda \times V_{\ell m}^{(a)} + (1-\lambda)V_{\ell m}^{(n-1)} \tag{2.16}$$

where λ is a constant between 0 and 1

$V_{\ell m}^{(n-1)}$ is volume resulting from last iteration

$V_{\ell m}^{(n)}$ is volume resulting from current iteration

$V_{\ell m}^{(a)}$ is volume resulting from all-or-nothing assignment with link costs from last iteration.

The new link costs are then calculated using the flows $V_{\ell m}^{(n)}$ after each iteration.

There are different ways of choosing λ at each iteration but perhaps the most popular is due to Smock (1962) who suggested that λ should be made equal to the reciprocal of the number of iterations. It can be seen that this results in a small value for λ after some iterations, ensuring small volume and cost changes.

The most important advances in capacity restraint assignment have resulted from a more rigorous application of the principles under which equilibrium can be said to have been reached and these have led to the development of equilibrium assignment methods.

2.4.4 Equilibrium assignment

The principles defining traffic equilibrium conditions on a network were formulated by Wardrop (1952). His first principle states: "Traffic on a network distributes itself in such a way that the travel costs on all routes used from any origin to any destination are equal while all unused routes have equal or greater costs".

This is equivalent to saying that traffic distributes itself in such a way that no driver can reduce his travel cost by switching to another route.

This principle implicitly assumes that all drivers perceive costs in the same manner and that the only source of route dispersion lies in the congestion effects.

Wardrop's second principle states that the distribution of traffic must be such that "the total travel cost on all routes in the system is minimum"(the marginal travel costs on all paths between an O-D pair are equal).

Wardrop's first principle is usually referred to as user optimiser or selfish equilibrium whereas the second results in a system optimum equilibrium. Both principles can be embodied in a mathematical programming framework but only the first one will be briefly described here.

Beckman et al (1956) and Potts and Oliver (1972), among others, have shown that finding link costs and volumes according to Wardrop's first principle in a network is equivalent to the following mathematical program.

$$\text{Minimise} \quad Z = \sum_{\ell m} \int_0^{V_{\ell m}} c_{\ell m}(x)dx \qquad (2.17)$$

subject to:

$$V_{\ell m} = \sum_r \sum_{ij} T_{ijr}\delta_{ijr}^{\ell m} \quad , \qquad (2.18)$$

$$T_{ij} = \sum_r T_{ijr} \quad , \qquad (2.19)$$

$$\text{and} \quad T_{ijr} \geq 0 \quad , \qquad (2.20)$$

where T_{ijr} is trips between i and j on route r

$$\delta_{ijr}^{\ell m} \text{ is} \begin{cases} 1 & \text{if link } \ell m \text{ is used by route } r \text{ between } i \text{ and } j \\ 0 & \text{otherwise} \end{cases}$$

It can be shown that the objective function Z is convex if the costs $c_{\ell m}(x)$ are monotonically increasing functions with flow. Under these conditions it is possible to find a solution to this mathematical program. The solution is <u>unique</u> in the link <u>costs</u> if the first derivatives of the cost functions are greater or equal to zero and unique in the link <u>costs and volumes</u> if these derivatives are greater than zero. However, in general the solution is not unique in the path flows T_{ijr} .

The great advantage of the mathematical programming framework (Van Vliet, 1979a) is that it enables a systematic treatment of equilibrium conditions and the development of algorithms known to converge to the correct solution. The algorithms used today are quite efficient (see Nguyen (1974) and LeBlanc et al (1975)) even for fairly large networks. They basically employ a similar structure to iterative assignment methods in which an all-or-nothing assignment is used to find 'steepest descent' direction and λ is chosen so as to minimise at each iteration the objective function Z. Variations on this basic algorithm can be found for example in Dow and Van Vliet (1979).

Research is currently active in further improving traffic assignment models. A better approach would define equilibrium as a state in which no driver can reduce his perceived (rather than objective) or expected travel cost by switching route. A kind of stochastic user optimiser equilibrium model will then result and different ideas has been put forward in this direction, see Van Vliet (1979b), Daganzo and Sheffi (1977), Daganzo (1977 and 1980), and Fisk (1980).

2.4.5 Proportional assignment

Robillard (1975) has introduced a distinction which is useful from the point of view of estimating a trip matrix from traffic counts by classifying assignment models into 'proportional' and non-proportional' ones. A proportional assignment model is one in which

(i) the total assigned flow on a link equals the sum of the assigned flows obtained when the method is applied to each O-D pair separately, and

(ii) if all the entries of the O-D matrix are multiplied by a constant factor γ all the assigned flows on each link will increase by that factor.

Models not meeting these two conditions are said to be non-proportional assignment models. It is clear that equilibrium assignment methods do not meet this condition as congestion effects will ensure that costs and paths change when demand changes. On the other hand, all-or-nothing and stochastic assignment models with fixed link costs do meet these requisites and can be classified as proportional assignment techniques. The advantage of proportional assignment models is that the variable $p_{ij}^{\ell m}$ can be calculated independently from the estimation of the O-D matrix. These 'proportions' depend only on the network and the parameters of the assignment model (for example θ in Equation 2.15). For non-proportional assignment techniques the O-D matrix estimation and the route choice problem must be solved jointly and it will be seen later that this involves additional complications.

To help in the choice of the assignment model to be used for a given set of conditions Van Vliet (1976) has suggested the following non-dimensional parameter to characterise the route dispersion resulting from congestion effects:

$$E_2 = \frac{(\text{total travel cost at equilibrium}) - (\text{total free flow cost})}{(\text{total travel cost at equilibrium})} \tag{2.21}$$

Van Vliet suggests that wherever $E_1 >> E_2$ (where E_1 has been defined by Equation (2.14)), a stochastic assignment model should be chosen as this means that route dispersion due to variations in users' perceptions of costs is more important than the dispersion due to congestion effects. If $E_1 << E_2$ the choice should be equilibrium assignment methods. In the other cases practical considerations like level of detail available and familiarity with a model will probably decide.

2.5 ERRORS AND ACCURACY CONSIDERATIONS

Any origin-destination matrix, whether obtained from questionnaire surveys, large-scale transport modelling exercises or from less expensive data such as traffic counts, will only be an approximate representation of the actual O-D matrix. Assessing the accuracy an origin-destination matrix is a very important issue which has received, however, very little attention. Origin-destination matrices are often used as if they were error free despite the fact that, if questioned, their users will immediately accept that they contain an unknown but probably large level of 'noise'.

One way of approaching this problem is to ascertain how accurate conventional methods are and to ask for a technique to be at least as accurate at a cheaper cost. A more rigorous approach would be to review the potential applications for synthesised O-D matrices and to determine in each case the costs associated with each error level and the effort involved in reducing them by choosing a better technique or larger sample. A minimum combined cost solution could then be adopted. It will be seen that research into errors in trip matrices is so scarce that this rigorous approach is still some years away.

2.5.1 Stability of the trip matrix of interest

In using any O-D matrix there is an implicit assumption of stability and permanence. The trip matrix for the central area of Reading between 16.10 and 18.10 on Tuesday, October 19, 1976, is of little historical interest; its value resides in the assumption that it can be used to design a better traffic management scheme to cope with similar demands at about that time on future days.

There is little evidence to support such an assumption. From our own experience we know that we tend to repeat journeys at about the same time; but from the same experience we also know that we introduce a wide range of variations, in time and space, to our movements, and this experience is supported by the known daily variations, for example, of traffic counts.

Choosing a representative trip matrix for design purposes is similar to choosing a representative or critical design volume for a road construction project. The main problem is that due to the high costs of data collection only one estimated trip matrix may usually be surveyed. This problem is usually tackled by surveying a trip matrix during a representative day (often Tuesdays or Thursdays) and including a critical period, normally a peak hour. Ideally, one would like to know the level of error contained in this estimated matrix and to have some idea of the degree of daily/seasonal variation to be expected.

During this research one of the few pieces of evidence about the daily variation of trip matrices has been produced and this will be reported in Chapter 7. Mean while, evidence available elsewhere about the accuracy of trip matrices obtained by conventional methods will be discussed.

2.5.2 Accuracy of O-D matrices estimated by conventional methods

2.5.2.1 Sources of error. The direct and indirect methods reviewed in Section 2.2 are subject to a range of sources of error:

(i) daily/seasonal variations and survey period expansion errors
(ii) data collection errors
(iii) data processing errors
(iv) sampling errors.

In addition to these, an indirect method will be subject to

(v) calibration errors
(vi) misspecification errors.

Survey period expansion errors occur when correction factors are applied to convert the original survey data to an O-D matrix for a different time slice or period (for example to expand a 16 hour survey to 24 hours). These errors are mainly caused by the time variations of the trip matrix.

Data collection errors occur during the survey period due, for example, to misreporting of trips, misidentification of vehicles, incomplete questionnaires and even errors while writing down answers or number plates. Again good quality control can decrease but not eliminate these errors and the task is more difficult than at the processing stage.

Data processing errors occur while transferring and compiling raw survey data. Mispunching of cards, miscoding, double counting, misallocation of addresses to zones, missing records and even programming errors are some of the main sources. A good quality control system may help to reduce these errors but again there is almost no evidence regarding their relative size.

Sampling errors take place when the surveys cannot cover all the trips during the survey period. This may be due to the location of the survey stations in roadside interview and flagging methods which makes it impossible to sample from certain trips, for example trips with both origin and destination inside a cordoned area. Another source is the fact that practical considerations have required the adoption of sampling rates of less than 1; these are termed sampling fraction errors.

Calibration errors occur in indirect methods when a 'wrong' value of a model parameter is obtained in the calibration, either due to errors in the procedure itself or due to shortages in the data.

Finally, misspecification errors are due to the fact that real tripmaking behaviour does not exactly conform to the assumed functional form of the travel demand model, for example a gravity model.

The only type of error which has a standard theoretical treatment is the error due to the sampling fraction or sample size errors.

2.5.2.2 <u>Sample size errors in roadside interviews</u>. The theoretical treatment of sampling ratio errors depends to an extent on the survey method used. In this sub-section a brief discussion of the sampling ratio errors for roadside interviews will be given and some of the ideas related to other methods and errors are discussed in subsequent sub-sections.

The sampling population in a roadside interview survey may be represented by a trip matrix including <u>all</u> vehicles passing through a survey point during the interviewing period. If T_{ij} is the number of vehicles having origin i and destination j than an unbiased estimate of T_{ij}, is given by

$$T^*_{ij} = \frac{N}{n} t_{ij} \quad , \tag{2.22}$$

where t_{ij} is the sampled number of vehicles having origin i and destination j

n is the number of vehicles selected for interview (at random)

N is the total number of vehicles passing through the survey point.

From sampling theory (Cochran, 1960) the unbiased estimate of the variance of T_{ij} is

$$v(T^*_{ij}) = \frac{N(N-n)}{n-1} p(1-p) \tag{2.23}$$

where $p = t_{ij}/n$, the proportion of trips in the sample having a particular O-D pair.

Errors due to sampling ratios may be expressed as a coefficient of variation of the estimate of the number of vehicles having a particular O-D pair.

$$C_v(T^*_{ij}) = \frac{\sqrt{v(T^*_{ij})}}{T_{ij}} \quad . \tag{2.24}$$

By substituting (2.23) into (2.24) and replacing r (sampling ratio) for n/N one obtains

$$C_v(T^*_{ij}) = \sqrt{\frac{1-r}{r-1/N} \left(\frac{1}{T_{ij}} - \frac{1}{N}\right)} \quad , \tag{2.25}$$

As the total number of vehicles passing through a survey point N is much larger than the total number from a particular O-D pair T_{ij} and the sampling rate r is also large relative to $1/N$, Equation (2.25) may be simplified to produce the approximate formula

$$C_v(T^*_{ij}) \approx \sqrt{\left(\frac{1-r}{r}\right)\frac{1}{T_{ij}}} \quad . \tag{2.26}$$

Figure 2.4, adapted from Hajek (1977), depicts the curves (on log-log paper) for several sampling ratios corresponding to Equation (2.26). These curves are straight lines with a slope of $-\frac{1}{2}$. Assuming that the errors are normally distributed these curves can be interpreted as follows: the probability is about two-thirds that the error obtained by repeated sampling would be equal to or less than the C_v error (%). For example, assuming a sampling rate of 0.25 and a trip interchange volume T_{ij} of 20 trips, two-thirds of all sampling errors may be expected to be within 40 per cent (between 12 and 28 trips).

It can be seen from Figure 2.4 that for the usually small values encountered in the cells of a trip matrix the theoretical errors are quite large. These theoretical results have been supported by empirical findings in other investigations compiled by Hajek (1977). An earlier study by Sosslau and Rokke (1960) had pragmatically fitted similar curves finding a slope of -0.4884 (≈ -0.5) confirming again these theoretical results.

2.5.2.3 Other survey methods. Similar results have been obtained by Dobson (1965) for home interviews using data from two real surveys. Vaughan (1972) has studied the sample size problem for home interviews and produced some recommendations regarding stratified sampling, but these are applicable to trip rates rather than interchanges. More recently Smith (1979) discussed the design of samples for home interviews and produced some more practical recommendations. In relation to sampling for a trip matrix he concludes that the sample size requirements are such that for ordinary volumes of the order of 20-30 trips per cell "even the large surveys conducted in the past had no hope of reproducing interchange volumes at the zonal level within a reasonable degree of accuracy". This is again consistent with the curves in Figure 2.4.

Sampling size errors are likely to follow a similar pattern in other direct methods provided the sample can be said to be a random one. This is not easy to achieve with the vehicle following method for example.

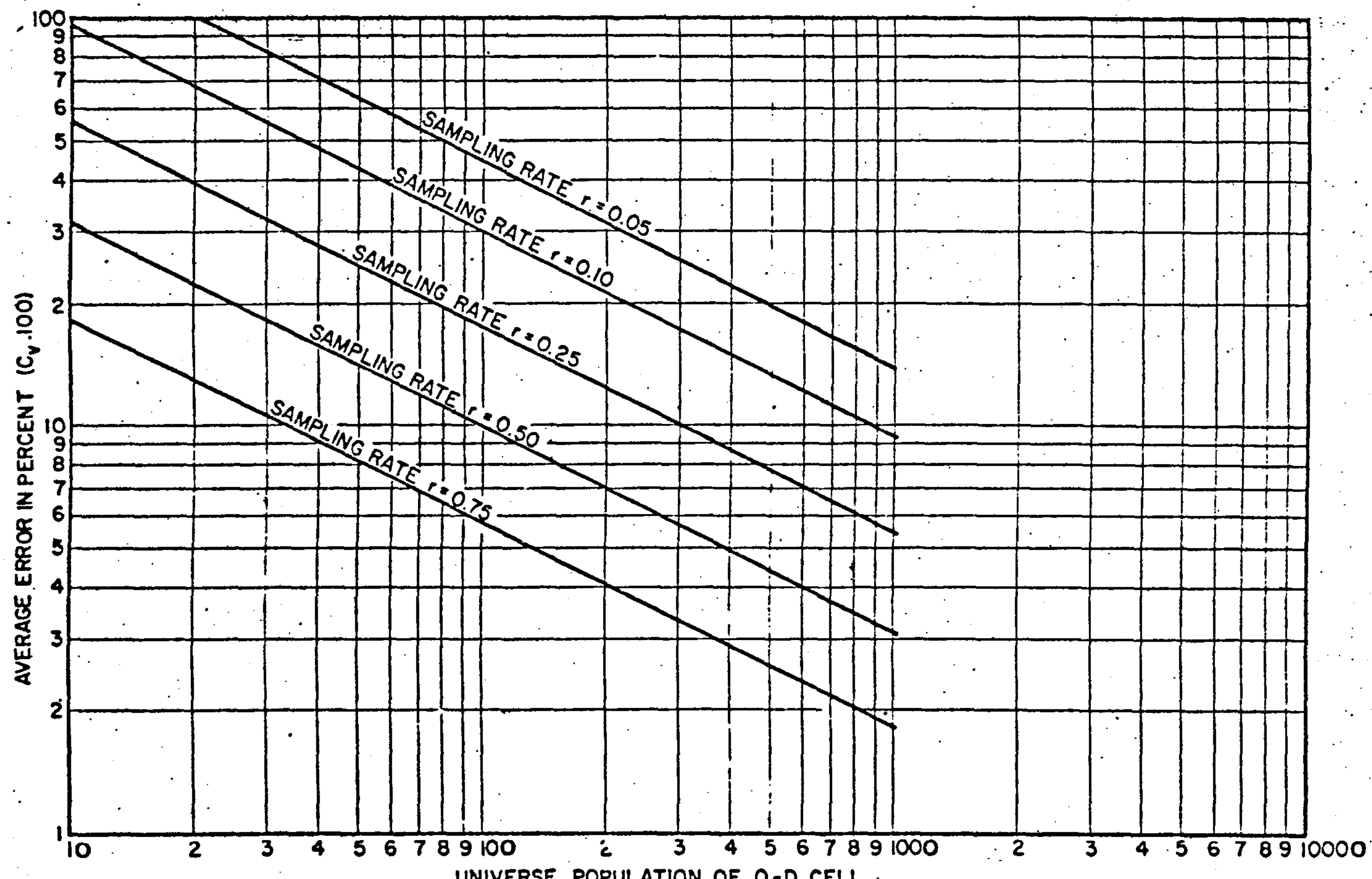


Figure 2.4: Theoretical sampling errors for O-D surveys

2.5.2.4 Total errors. It is possible to study the way in which different errors are compounded to form a total error. For example Hajek (1977) identifies Equation (2.27) for the total error for roadside interviews

$$e_t = [b^2 T_{ij}^2 + (T_{ij}+b)^2 e_b^2 + b^2 e_c^2]^{\frac{1}{2}} \quad , \qquad (2.27)$$

where e_t is the total error
b is the survey period expansion factor (> 1.0)
e_b is the error associated to b
e_c is the error associated to the daily variations of T_{ij}

Hajek suggests typical values for e_b and e_c to be of the order of 8 per cent and 13 per cent respectively.

This formula suggests that for many applications (with b around 1.35 ∿ 1.5) the total error so defined will be between 1.5 and 2 times the theoretical error.

Note that data collection and processing errors have not been included in this total error measure. Lack of information about these and other error components have generated some efforts in calculating total errors directly.

The US Federal Highway Administration (1975) has produced a table depicted in Figure 2.5 to estimate a rather loosely defined total error for home interviews. The error measures used are the root mean square (RMS) error defined as

$$\text{RMS} = \sqrt{\frac{(T^*_{ij} - T_{ij})^2}{n}} \quad , \qquad (2.28)$$

where n is the number of observations, and the percent RMS defined as

$$\%\text{RMS} = \frac{100 * \text{RMS}}{(\text{mean of observations})} \quad . \qquad (2.29)$$

The percent RMS has similar properties to the coefficient of variation C_v . The curves correspond to the equation

$$\%\text{RMS} = \frac{1624}{T_{ij}^{\;0.4884}\sqrt{100r}} \quad , \qquad (2.30)$$

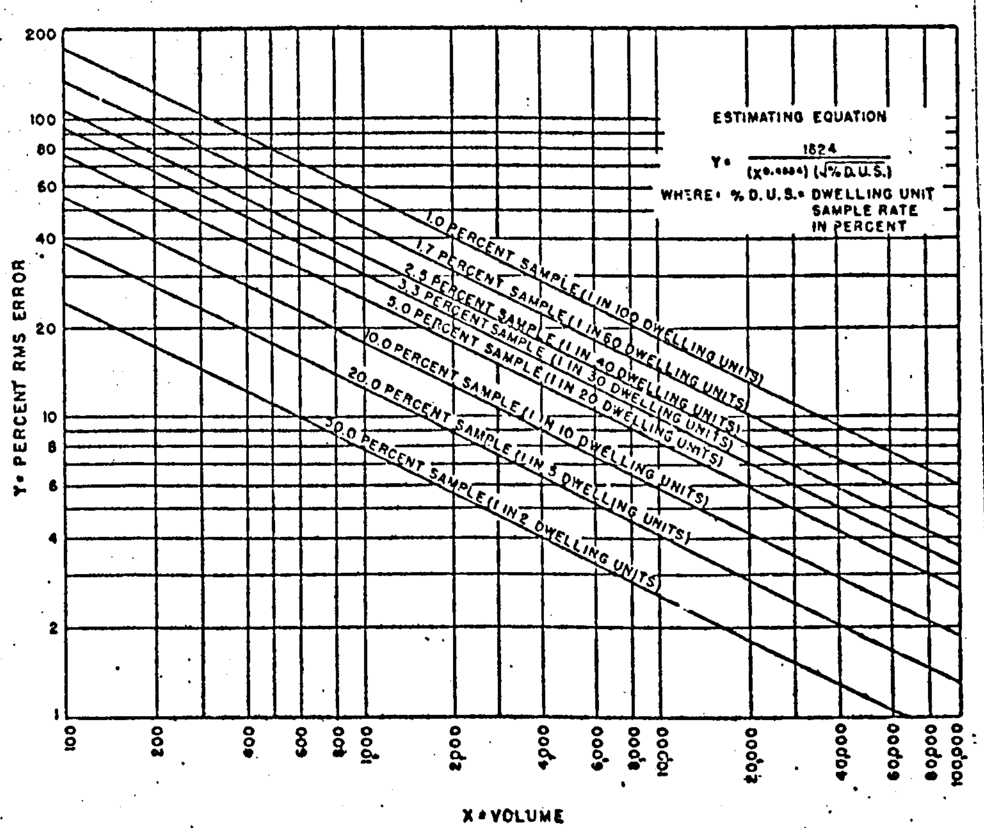


Figure 2.5: Total error for home interview O-D surveys

and they seem to bear out the statement above about the ratio of sampling ratio to total error.

Synthetic indirect methods are likely to suffer from these errors plus calibration and misspecification errors. A recent study by Sikdar and Hutchinson (1980) and Sikdar et al (1980) used data from 28 study areas in Canada to calibrate and test doubly constrained gravity models. The researchers found that the performance of the models was equivalent to a randomly introduced error in the observations of about 75 to 100 per cent and concluded that the "continued use of these models for estimating spatial interaction patterns ..... cannot be justified."

On the other hand, in many cases one is not interested in the accuracy of the trip matrix itself but in the errors of some aggregate measure of T_{ij}'s such as screen line movements, assigned flows, total delays and so on. The relative errors at these levels are likely to be much smaller than at T_{ij} level. For example, Stover et al (1976) studied the impact of errors in origin-destination data on traffic assignment accuracy using real data and stochastic simulation. They concluded that "while (expanded) origin-destination trip tables are subject to substantial error in terms of the resulting zonal trip ends and interzonal interchange volumes, these trip tables have generally given reasonable assignment results." They explained this improvement in accuracy at this level by the power of capacity restrained assignment to produce 'reasonable' flow levels for a given network thus masking errors in previous modelling stages.

2.5.3 Errors in traffic counts

The errors associated with traffic counts are likely to be of the types (i) to (iii) described in Section 2.5.2.1, that is daily/seasonal variation, data collection and data processing errors. Note that normally there is no sampling ratio involved in counting traffic. Time variation and data processing errors are likely to be common to most traffic counting techniques.

Cleveland (1965) suggests that pneumatic traffic counters have a data collection error of about 5 per cent due to the fact that this type of counter records axles rather than vehicles.

Induction loop counters detect vehicles rather than axles but may require careful tuning, and may undercount vehicles when two of them pass over the loop at almost the same time. Bellamy (1979) has studied these problems and concluded that errors of the order of 3 to 10 per cent may be encountered.

The impact of daily and seasonal variations has been studied, among others, by Bellamy (1978) and Phillips (1979a and 1979b). To summarise some of these results it can be said that it is possible to estimate the average annual daily traffic at any point in an urban road to within 10 per cent from a single 4 hour count on any day of the year. Expansion and correction factors have been developed to give this accuracy.

Advances in automatic vehicle classification and identification of turning movements are currently underway and these will certainly help to reduce costs and increase the accuracy of traffic volume data; see for example the report by the Institution of Civil Engineers (1978).

2.6 SUMMARY

This research attempts to develop ways of synthesising O-D information from traffic counts and the main characteristics of that problem were described in Section 2.3. The problem is in general underspecified, in the sense that the traffic counts are unlikely to uniquely determine an O-D matrix. Section 2.4 discussed the role of traffic assignment models and the way in which they affect the trip matrix estimation problem. The question of the accuracy attainable by conventional methods used to estimate a trip matrix was addressed in Section 2.5. It was found that due to the sparsity of this matrix theoretical errors were quite high and that other errors compounded this effect.

CHAPTER 3
REVIEW OF PROPOSED METHODS

3.1 OVERVIEW

It has been shown in the previous chapter that one of the key issues in estimating an origin-destination matrix from traffic counts is the fact that the problem is generally underspecified. The traffic counts alone do not provide enough information to uniquely determine an O-D matrix. One can say that the degree of underspecification or degree of freedom in choosing the O-D matrix is the difference between the number of unknowns T_{ij} $(M^2 - M)$ and the number of independent traffic counts available.

Thus, the task of estimating a trip matrix from counts can be seen as one of injecting additional information to the system of L_c', Equations (3.1):

$$\sum_{ij} T_{ij} p_{ij}^{\ell m} = V_{\ell m} \qquad (3.1)$$

for $\ell m \in L_c'$ (independent and counted links)

so that the problem becomes fully specified. This extra information may take the form of a theory of tripmaking behaviour which introduces restrictions to the values of variables T_{ij} can take; for example, one may require $[T_{ij}]$ to be linked to a gravity or other travel demand model. For example one may add to Equations (3.1) the following equations

$$T_{ij} = R_i S_j / d_{ij}^2 \qquad (\text{all } T_{ij}) \qquad (3.2)$$

where R_i and S_j are auxiliary variables restricting the freedom of $[T_{ij}]$ and d_{ij} is the distance between zone i and zone j.

It is almost impossible to avoid overspecifying the problem with this approach in the sense that there may be no solution to the system (3.1) plus (3.2), some of these equations being inconsistent * . The introduction of an error term ε in the traffic counts increases the realism of the problem as it acknowledges the uncertainty surrounding count data and solves the overspecifying problem. The new system is

$$\sum_{ij} T_{ij} p_{ij}^{\ell m} = V_{\ell m} + \varepsilon_{\ell m} \quad (\ell m \in L_c) \qquad (3.3)$$

and

$$T_{ij} = R_i S_j / d_{ij}^2 \quad (\text{all } T_{ij}) \qquad (3.2)$$

where $\varepsilon_{\ell m}$ is an error term for each count.

This problem is again underspecified in $R_i S_j$ and $\varepsilon_{\ell m}$ but a specific solution can be obtained by seeking to minimise some aggregate measure of the error terms. A frequent choice is to minimise $\sum_{L_c} \varepsilon_{\ell m}^2$ for which standard software is available.

The different behavioural assumptions used to estimate a trip matrix from counts have been used to group together similar techniques. The models proposed by a number of researchers have been classified into three groups.

(i) Techniques using some form of gravity model as an underlying principle for tripmaking behaviour.

(ii) Techniques using other assumptions about tripmaking behaviour and embodying them into a (direct) travel demand model.

The techniques in groups (i) and (ii) usually require zonal data in addition to traffic counts.

(iii) Techniques using only <u>network data</u> (costs, counts) and some general principle like entropy maximising to select an O-D matrix.

---

* Of course (3.1)+(3.2) is no longer a linear system but it has in effect only 2M-1 unknowns $R_i S_j$ as one of them is not independent .

Other classification schemes are possible and this one is far from perfect in the sense that it does not allocate models to groups in an unequivocal way. But at least this scheme usefully enables the discussion of some common issues and ideas to proceed in a more or less orderly fashion. The most important overlaps and ambiguities will be acknowledged in the last section of this chapter.

One of the problems hindering a review of methods to estimate trip matrices from counts is that relatively few publications have appeared in technical journals. During this research the author established contact and exchanged ideas with several researchers working in this area, some of them at graduate level. This chapter deals mainly with published papers and only makes reference to unpublished material and comments when these make a major contribution to the subject which would otherwise be lost. The review relies heavily on an earlier publication by the author (Willumsen, 1978a), which has already served as a base to more recent reviews elsewhere (Chan et al, 1980; Hauer and Shin, 1980).

3.2 APPROACHES BASED ON A GRAVITY MODEL

3.2.1 Background

The gravity model was one of the first mathematical models used to synthesise inter-zonal trips in an area (Casey, 1955). Because of its simplicity and a certain intuitive appeal it has attracted social scientists and engineers alike (Carruthers, 1956). It is not surprising then that the first ideas on estimating an O-D matrix from traffic counts were based on the gravity model. In fact, these first approaches saw the problem as one of calibrating a gravity model using inexpensive and widely available data: volume counts. These models were expected to constitute simpler substitutes for more expensive and data intensive conventional models and in this vein they were considered by the OECD group studying the scope for simplifying traffic models (OECD, 1974).

All these models assume that most of the tripmaking behaviour in the area of interest can be explained in terms of three types of factors:

- trip generation or origin factors
- trip attraction or destination factors
- separation or travel cost factors.

The models use, where possible, easily available planning data such as population and employment to produce the trip end factors; E_j^q and G_i^q will be used throughout to represent these trip end factors for particular journey purposes. For example E_j^2 could be the value of the trip attraction variable corresponding to journey purpose 2 (shopping) at destination j (perhaps employment in retail activities). The general function of cost $f(c_{ij})$ may be used to represent the separation factor in more advanced models.

In almost every study a proportion of trips have their origin or destination (or both) outside the area of interests. These trips are sometimes called external (or through) trips and most of their length occurs outside the study area. A gravity model using measures of separation internal to the study area cannot be expected to model well this type of movements *. The most common way of tackling this problem in practical applications is to carry out roadside interviews at the boundaries of the study area. These can be used to build an O-D matrix for the external trips. These trips are then loaded onto the network and the flows thus obtained are subtracted from the counts. In this way the modified traffic counts reflect only the movements internal to the study area.

An alternative approach would be to consider the external zones as part of the study area. These new 'pseudo internal' zones require special trip end and separation factors; see for example the model in Section 3.4.2.

* This not necessarily holds true for a negative exponential deterrence function.

The models in this section are rooted in the traditional four stage conventional, transport modelling approach although they tend to collapse trip generation, distribution and assignment into one estimation exercise. It could be possible to extend the approach to include a modal split element but this has not yet been attempted succesfully and in any case it would be beyond the scope of this research.

3.2.2 D.E.Low's model

3.2.2.1 Description. The first model based on traffic counts to be reported in the literature was probably that put forward by Low (1972). His objective was to "effectively combine into one process what is usually handled in a series of three or four sub models, each with its own set of errors.".He attempts to develop a model whose output is traffic volumes but estimates a trip matrix during an intermediate state. Low sees his model as applicable, in principle, to a large range of study areas but especially to small ones (around 50,000 inhabitants). His approach can be described as follows.

(a) Assignment of external trips. Current external trips as obtained from external cordon roadside interviews are assigned to the existing network to produce estimates of 'current external volumes' throughout the network.

(b) Estimation of 'current internal volumes'. The so called 'current internal volumes' are then obtained from the actual counted flows less the assigned 'current external volumes'. These current internal volumes are the result of trips "wholly explainable in terms of area characteristics".

(c) Internal volume forecasting model.

(i) Inter-zonal trip opportunity matrices of the form $G_i^q E_j^q$ are developed. For the journey to work population and employment are used. Each element of the trip opportunity matrix is then multiplied by a friction factor of the type C_{ij}^{-k} . In general time is suggested as measure of cost. The product $G_i^q E_j^q C_{ij}^{-k}$ is called then an 'inter zonal trip probability factor' and the full set of factors is the trip probability matrix. Several trip probability factors, say for different journey purposes and person types, may be so developed:

$$X_{qij} = G_i^q E_j^q C_{ij}^{-k} \quad , \qquad (3.4)$$

where X_{qij} is the trip probability factor between i and j for person type or journey purpose.

(ii) Trip probability matrices are then assigned separately to the current network (in practice an all-or-nothing assignment) just as if they were trips. For a link ℓm an auxiliary variable can be defined as

$$X_q^{\ell m} = \sum_{ij} p_{ij}^{\ell m} X_{qij} \quad , \qquad (3.5)$$

where $p_{ij}^{\ell m} = \begin{cases} 1 & \text{if the least cost route from } i \text{ to } j \text{ passes through link } \ell m \\ 0 & \text{otherwise .} \end{cases}$

(iii) Multiple regression techniques are used to develop equations of the following form

$$V_{\ell m}^* = b_0 + b_1 X_1^{\ell m} + b_2 X_2^{\ell m} + \ldots + b_1 X_q^{\ell m} \qquad (3.6)$$

where $V_{\ell m}^*$ is the internal traffic volume on link ℓm and b_q are the constants to be obtained. Note that b_0 here represents intrazonal or local traffic.

(iv) Low suggests that separate equations can be developed for different types of roads or areas or that additional parameters defining the characteristics of the link may be included.

3.2.2.2 Scope and applications. It is suggested that the model could be used for forecasting purposes as well as replicating the present pattern of trips. To this end the value of the socio-economic parameters should be obtained for the design year as well as the alternative networks that would be tested.

Low also suggests that more sophisticated versions of his model can be developed, especially in relation to the way in which the trip probability matrix is obtained.

The model was used during the Monongalia County Transportation Study in West Virginia in 1970-71. Conventional techniques were used in parallel to this model. The basic information used was:

- base year population and employment by zone (52 zones)
- base year road traffic assignment network
- base year traffic counts
- base year external trips table.

A very simple form of the deterrence or friction function was used but Low claims reasonable results.

Two inter-zonal trip probability factors were used

$$X_{1ij} = P_i E_j C_{ij}^{-2} \quad , \quad \text{and} \quad X_{2ij} = P_i P_j C_{ij}^{-2} \quad ,$$

where P_i , P_j = population in zone i and zone j
E_j = employment in zone j
C_{ij} = travel time between i and j .

Twenty-three counting points were used and the best regression line obtained was

$$V_{\ell m} = 730 + 0.005 \, X_1^{\ell m} \, , \tag{3.7}$$

with coefficient of correlation $R = 0.974$ for a comparison of observed against modelled flows.

3.2.2.1 Further developments. Smith and McFarlane (1978) have recently applied Low's model to the county of Fond du Lac in Central Wisconsin. The primary purpose of the study was to evaluate Low's model, now called Internal Volume Forecasting model (IVF), as a replacement for the conventional urban travel demand model in small and medium-sized urban areas. As this application involved some improvements in the model and also a critique of its theoretical basis a summary of the report will be given here.

Using the 1970 network and data and a technique analogous to the one used in the Monongalia study it was found that the best regression equation was

$$V_{\ell m} = 263 + 0.043\ X_2 \quad , \qquad (3.8)$$

where $X_{2ij} = P_i P_j C_{ij}^{-2}$, and C_{ij} = travel time.

Having selected the trip probability factor to use the next parameter examined was the travel time exponent. Two alternative exponents, -1.5 and -2.5, were tested on either side of the standard -2.0 exponent. It was found the $n = -2.5$ gave the best explanatory power. No other exponents were tested.

Finally, tests were made with multipath assignment techniques to see if it was possible to improve the fit of the model. The model used was the standard UMTA multipath assignment program UROAD. Only marginal improvements were obtained with this approach.

A conventional model was not applied to the area so it was not possible to compare their accuracies. Nevertheless, the level of accuracy of the IVF model in reproducing base year link volumes was "certainly within the limits of conventional models". This is to some extent surprising as only population and zone-to-zone travel times are included in the model. The reported correlation coefficient was $R = 0.87$ and the Root Mean Square Error (RMSE) values were 209 in absolute terms and 53 per cent in percentage terms at link flow levels.

In order to evaluate the forecasting ability of the model empirically the model was applied to 'forecast' 1960 traffic volumes using the 1960 Fond du Lac network and census data. The results of this exercise showed that overall the "absolute error of the 1960 estimate of assigned volumes was less than that of the 1970 estimates".

The researchers, though favourably impressed by the accuracy of the IVF model, were critical of its theoretical basis. Smith and McFarlane pointed out three basic mis-specification errors in the model.

(a) Changes in the tripmaking propensity of the study area population over time cannot be considered in the model. For instance, the impact of car ownership is not reflected in the model.

(b) The unconstrained nature of the model. If population is doubled, traffic flows will be multiplied by a factor of four instead of just doubled as would be expected. This criticism applies mainly to the model used in the forecasting mode. This objection is easily overcome by the inclusion of a 'normalising factor' equal to $(\sum_j E_j^q)^{-1}$ in Equation (3.4).

(c) The different measures of the trip probability factors are likely to be co-linear in which case only one probability factor should be included in the equation.

The researchers considered several possible improvements to the model most of which appear in other models reviewed in this section. They finally concluded that the inclusion of these improvements would probably make the IVF model far too similar to conventional techniques and its original attractiveness would be lost.

3.2.3 Overgaard's model

3.2.3.1 Description. Overgaard's model, though probably developed independently, can be considered as an improved version of Low's approach. As it includes a proxy measure for trip making propensity it goes some way to answer some of the criticisms of Smith and McFarlane. This model has been applied to Silkeborg, a Danish town of 44,000 inhabitants and it is described in some detail in OECD (1974) and Bendtsen (1974). External trips are obtained in the same way as in Low's method. The main change is regarding trip generation.

Car ownership levels were known for Silkeborg so that trip generation equations were stated in terms of trips per car per day (it has to be remembered these methods are only concerned with car movements). Furthermore, as a proxy for socio-economic level, the type of dwelling (single-family house or apartment) was also introduced in the trip generation term.

The general trip generation (and attraction) term was then expressed in the form

$$G_i = b_1E_i + b_2P_i + b_3(P_ir_i) \qquad (3.9)$$

where G_i = generation/attraction 'force' in zone i
E_i, P_i = employment and population in zone i
r_i = percentage of population in 'one-family houses'.

The model then uses a more or less conventional gravity form.

3.2.3.2 Application. When this model was applied in Silkeborg, it was found that

b_1 = 1.75 (trips per workplace)
b_2 = 0.7 (trips per inhabitant)
b_3 = 0.008 (trips per inhabitant in 'one-family houses')

gave the best results. Also, an exponent of -1.8 for travel time was calibrated for trips with $t > 90$ secs. The accuracy of the model in terms of the comparison of observed and predicted link flows was measured by an RMS (percentage) error. This was found to be around 20 per cent and better for higher flows.

3.2.4 Hogberg's model

3.2.4.1 Description. A model based on a more general type of gravity model has been proposed by Hogberg (1976). This version of the gravity model allows up to three journey purposes and a more flexible deterrence function.

Hogberg assumes that the joint generation distribution model is of the form

$$T_{ij} = b_1G_i^1A_i^1E_j^1f(c_{ij}) + b_2G_i^1A_i^2E_j^2f(c_{ij}) + b_3G_i^2A_i^2E_j^2f(c_{ij}) \qquad (3.10)$$

where

G_i^1 and G_i^2 are different types of trip generation parameters

E_j^1 and E_j^2 are different types of trip attraction parameters

A_i^1 and A_i^2 are balancing factors of the form

$$A_i^1 = (\textstyle\sum_j E_j^1 f(c_{ij}))^{-1} \tag{3.11}$$

and $f(c_{ij}) = C_{ij}^k \exp(b_4(\log_e C_{ij})^2)$.

Using the indicator variable $p_{ij}^{\ell m}$ the modelled flow at each link becomes as usual

$$V_{\ell m} = \sum_{ij} p_{ij}^{\ell m} T_{ij} \quad . \tag{3.12}$$

For a subset of links we can then compare the observed with the modelled flows and minimise the square of the difference

$$(\hat{V}_{\ell m} - \sum_{ij} p_{ij}^{\ell m} T_{ij})^2 \ .$$

3.2.4.2 Application. Hogberg carried out a desk study of this method on a synthetic network with 16 nodes and 44 one-way links using population and employment as origin and destination specific parameters. He assumed the three elements in the distribution model stood for home-work, home-home and work-work trips. Hogberg then used an algorithm for non-linear regression and from a synthetic O-D matrix he sampled flows from half of the links. After introducing an artificial error component he used this sample to obtain a minimum squared difference between observed and modelled traffic. He then compared the flows predicted for the rest of the 22 links and found the model "acceptable".

Hogberg (1975) has carried out some theoretical work on the 'contribution' of each link to the accuracy of the model. He found that after the inclusion of a few links in the calculations there was little gain in the accuracy of the model. This is not entirely surprising. The artificial data was created using a composite gravity model. The model to be estimated then was perfectly specified and the only source of error was 'noise' in the traffic counts. At least 5 counts are required to calibrate the 5 parameters of the model and if these 5 had no noise they would uniquely determine the original model. The number of counts above 5, required to obtain a good estimation of the model, would depend on the amount of noise introduced into the 'observations'.

In a personal communication to this author, Hogberg has confirmed that practical applications are being carried out in Sweden with this model but they have not yet been reported in English.

3.2.5 Commonwealth Bureau of Roads' model

3.2.5.1 Description. This approach was developed for the Commonwealth Bureau of Roads (CBR) in Australia by Symons, Wilson and Paterson (1976). The objective of the exercise was "to devise a rigorous methodology for disaggregating the Australian National System of Urban Centres into distinct zones of economic activity; then to use this zoning to construct a model of traffic generation for the National Highway System". This approach is only applicable to inter-city travel but includes an interesting innovation at the trip generation stage.

A relationship is assumed between traffic generation and the urban hierarchy which is determined by the provision of services to lesser urban centres. Each urban centre is assigned a level in the hierarchy from Central Place Theory considerations and it is assumed that trips will be generated from a low order centre to a higher order centre only. The trip generation rates per head of population are expected to vary from level to level. Each of the Australian urban centres was then assigned to one of seven broad ranks using population and employment type statistics. As explanatory variables both population and a recreational attractiveness index were used. The model could then be expressed as a group of sub models.

Trip purpose	Functional form (F_r)
Seeking level 7 services	$P_i\ C_{ij}^{-k_1}$
Seeking level 6 services	$P_i\ C_{ij}^{-k_2}$
Seeking level 5 services	$P_i\ C_{ij}^{-k_3}$
Seeking level 4 services	$P_i\ C_{ij}^{-k_4}$
Recreation	$P_i\ R_j\ C_{ij}^{-k_5}$
Inter-capital	$P_i\ P_j\ C_{ij}^{-k_6}$

where

P_i = population urban centre i
R_i = recreational index urban centre i
C_{ij}^{-k} = deterrence function
F_r = functional form for journey purpose r .

The general model is then

$$T_{ij} = \sum_r F_r b_r \tag{3.13}$$

and $$V_{\ell m} = \sum_{ij} T_{ij} p_{ij}^{\ell m} \tag{3.12}$$

that is

$$V_{\ell m} = \sum_{ij} \sum_r b_r F_r p_{ij}^{\ell m} \tag{3.14}$$

$$V_{\ell m} = \sum_r b_r \sum_{ij} F_r p_{ij}^{\ell m} , \tag{3.15}$$

where Equations (3.15) can be used to determine the calibration parameters b_r for each journey purpose. Of course an implicit assumption in this model is that route choice does not depend on journey purpose and that proportional assignment is sufficiently realistic.

In order to simplify calculations only trips between State capitals and trips between a centre and the nodes in its market area were considered. It was felt that this set of trips would capture the bulk of inter-urban traffic.

3.2.5.2 Application. In the Australian case assignment was performed manually and multiple linear regression techniques were used to calibrate the model. In addition different exponents k for the deterrence function for each journey purpose were tested and those giving best fit were adopted. It was found that the smallest exponent was that of trips to level 5 suggesting that perhaps the most essential and irreplaceable goods and services were supplied by these centres. The researchers also found a high exponent for inter-capital trips. This was explained as a result of the heavy competition from other modes (air and rail services).

Symons et al concluded that the model was to be of great interest to State Highway Authorities for it indicates how much information may be extracted from census data and road counts. They were satisfied that the results were encouraging and statistically sound and envisaged no conceptual difficulties in transforming the model into a 'robust predictive tool'. The researchers mentioned a number of desirable improvements in that direction mainly in terms of automating processes within the model, for example traffic assignment.

3.2.6 Danish Road Directorate approach

3.2.6.1 Description. The Danish Road Directorate (DRD) has acquired considerable practical experience with this type of model. At least two full scale applications in rural areas, Jensen and Nielsen (1973) and Holm et al (1976), have been reported in English. The method has been improved since its first conception but only the version described by Holm et al (1976) will be discussed here.

In general the method is again seen as a way of directly calibrating a trip generation distribution assignment model from traffic counts. Probably the main point of interest is the use of an iterative assignment technique instead of the simpler all-or-nothing via minimum cost path apprach of previous researchers.

In general 'external trips' are not obtained independently but as a result of the use of the model. In effect, the 'external zones' are coded in the same way as the internal ones and treated in the same way. This is mainly possible because the model has been used in rather large and relatively self-contained rural areas. The basic trip generation distribution model is

$$T_{ij} \doteq b\, G_i E_j C_{ij}^{-k} \qquad (3.16)$$

where b is the trip generation factor.

The problem is then to determine a value for b consistent with the iterative assignment and which minimises the difference between observed and modelled link flows.

An iterative assignment has been chosen, hopefully converging towards a Wardrop's equilibrium, and consequently the $p_{ij}^{\ell m}$ values change at each iteration. The Danish Road Directorate chose Smock's (1968) iterative algorithm for assignment and developed a method for calculating b as part of the iterative process.

3.2.6.2 The iterative process. The 'cost-flow' relationship for a link ℓm is assumed to be

$$C_{\ell m} = C_{\ell m}(V_{\ell m}) .$$

Then all or part of the traffic between each pair of zones is assigned for routes in such a way that, for each link, the flow at iteration n is

$$V_{\ell m}^{n} = (1 - \tfrac{1}{n})\, V_{\ell m}^{n-1} + \tfrac{1}{n}\, V'_{\ell m} \qquad (3.17)$$

where $V'_{\ell m}$ is the traffic volume resulting from all-or-nothing assignments to the link in iteration n, using the cost obtained in iteration $(n-1)$.

Two auxiliary variables are used, X_{ij} and $Y_{\ell m}$, defined as

$$\left.\begin{aligned} X_{ij} &= G_i E_j C_{ij}^{-k} \\ \text{and} \quad Y_{\ell m} &= \sum_{ij} G_i E_j C_{ij}^{-k} \cdot p_{ij}^{\ell m} = \sum_{ij} X_{ij} p_{ij}^{\ell m} \end{aligned}\right\} \qquad (3.18)$$

where $p_{ij}^{\ell m}$ is the proportion of traffic T_{ij} which uses link ℓm.

At iteration n in the estimation process

$$V_{\ell m}^{n} = b\, Y_{\ell m}^{n} \qquad (3.19)$$

where

$$Y_{\ell m}^{n} = (1 - \frac{1}{n})\, Y_{\ell m}^{n-1} + \frac{1}{n} Y'_{\ell m} \quad . \qquad (3.20)$$

At this stage it would be possible to calculate, after each iteration, the value of b using linear regression, and one might accept that convergence has been achieved when the estimations of b do not differ significantly in two consecutive iterations.

The Danish Road Directorate, however, preferred a maximum likelihood method. It is assumed that the observed flows $\hat{V}_{\ell m}$ are mutually independent, normally distributed variables with mean $V_{\ell m}^{n}$. that is

$$\hat{V}_{\ell m} = V_{\ell m}^{n} + \varepsilon_{\ell m} \qquad (3.21)$$

where $\varepsilon_{\ell m}$ is a normally distributed variable with mean zero and variance $[(V_{\ell m}^{n})^{\gamma} \sigma^2]$ *.

---

* $\gamma = 0$ implies that the standard deviation is independent from the mean

$\gamma = 1$ requires mean and variance to be in proportion

$\gamma = 2$ means that average flow and standard deviation are proportional.

Using the maximum likelihood method it is possible to find the best estimator for b as

$$b^* = \frac{\sum_{\ell m} V_{\ell m}(Y^r_{\ell m})^{1-\gamma}}{\sum_{\ell m} (Y^n_{\ell m})^{2-\gamma}} \tag{3.22}$$

and also

$$\sigma^* = \frac{1}{n-1} \sum_{\ell m} \frac{(V_{\ell m} - V^n_{\ell m})^2}{(V^n_{\ell m})^\gamma} \quad . \tag{3.23}$$

The derivation of these relationships is shown in Willumsen (1978a).

The algorithm developed by the DRD can be summarised as follows.

Step 1 = calculate free flow travel times

Step 2 = determine minimum cost routes for all O-D pairs

Step 3 = assign, following Smock's algorithm, $X'_{ij} = O_i O_j C_{ij}^{-k}$ factors to the network

Step 4 = estimate b* with the maximum likelihood formulae

Step 5 = calculate traffic volumes on all links as $b^* Y_{\ell m}$

Step 6 = calculate new travel times

Step 7 = if b* has not converged proceed to Step 2, otherwise stop

This algorithm is tested with different values of the exponent γ and k in a manner similar to Smith and McFarlane (1978). It is not difficult to see that the convergence of this algorithm as defined requires $Y^n_{\ell m}$ to change little from one iteration to the next. This will happen by force after some iterations as a large n in (3.20) ensures a small change in $Y^n_{\ell m}$ without guaranteeing Wardrop's equilibrium.

3.2.6.3 <u>Application</u>. The model has been applied, in slightly different versions, in the Aarhus and in the South Zealand rural areas in Denmark. In this last case the data used consisted of a road network, with seven types of roads (and speed-flow relationships), parish populations and about 40 counting points. The network included 73 zones and 334 links. It was found that $\gamma = 1$ and $k = 3.25$ gave the best fit.

The researchers found that the model tended to overestimate large flows and underestimate small ones. This is certainly not very desirable from a planning view point. They thought the source of this error was the fact that intra-zonal trips were not considered by the model.

The researchers estimated the percentage error in flows on the links was around 17 per cent and that this did not deteriorate much when the model was calibrated with only half the counts.

The researchers also prepared confidence intervals for the Equations (3.19) and (3.20). They found that the 95 per cent confidence interval was, for example, ± 5200 pcu/day at 10,000 pcu/day and about ± 3000 pcu/day at 5,000 pcu/day levels. They concluded that this accuracy was no worse than that obtained through the use of more expensive traditional models.

The Danish Road Directorate has also extended the approach to include a more general form of the gravity model similar to Hogberg's. This model has not been reported in English.

3.2.7 Robillard's model

Pierre Robillard was one of the pioneers of a more rigorous analysis of the information that could be extracted from traffic counts. In Robillard and Trahan (1973) and Robillard (1975) two methods to estimate an O-D matrix from counts are put forward. The first one is a composite method requiring first an estimation of the total trip ends O_i and D_j and then the use of a gravity model on similar lines to Low's. The second method is more general and of interest here.

Robillard assumes a generalised gravity model of the type used in partial matrix techniques; see for example Kirby (1979). This takes the form

$$V_{\ell m} = \sum_{ij} R_i S_j f(c_{ij}) p_{ij}^{\ell m} + \varepsilon_{\ell m} \tag{3.24}$$

where $\varepsilon_{\ell m}$ is an error term and R_i and S_j are the parameters to be estimated. $f(c_{ij})$ is a known cost function or known 'cost bins' if empirical separation ranges are used.

The parameters R_i and S_j are not unique. It is always possible to multiply the R_i by a constant k and the S_j by k^{-1} without affecting Equations (3.24). The estimation of the parameters R_i and S_j can be tackled as a non-linear regression problem. It is always possible to solve this problem so that the R_i and S_j are positive.

The number of independent links with traffic counts available should be greater than the number of the variables R_i and S_j . In terms of Hogberg's previous example it means that at least 33 of the 44 one-way links should be sampled. Robillard indicates that the method suggested by Lawtcn and Sylvester (1971) can be used to solve this model in order to minimise the difference between observed and synthesized flows.

$$\text{Min} \sum_{\ell m} (\hat{V}_{\ell m} - \sum_{ij} R_i S_j f(c_{ij}) p_{ij}^{\ell m})^2 \tag{3.25}$$

for all counted links.

Of course other algorithms are also possible, but it is claimed that this approach should reduce computer time. It is interesting to note that this method does not require any information regarding population or other trip generating parameters. These are all embodied in the R 's and S 's . One can compare the standard doubly constrained gravity model

$$T_{ij} = G_i A_i E_j B_j (f(c_{ij})) \tag{3.26}$$

with

$$T_{ij} = R_i S_j f(c_{ij}) \tag{3.27}$$

to find the equivalence

$$R_i \equiv G_i A_i$$

$$S_j \equiv E_j B_j \quad .$$

Robillard's approach has been included here because of its link with the gravity model. As it only uses network data it could have been included under Section 3.4 as well.

3.2.8 Comments

Approaches based on some form of the gravity model have certainly an intuitive appeal and some of the models, in particular those resulting in linear regression calibration, may be quite simple to use.

The main criticisms leveled at these models seem to be

- extreme simplicity
- lack of policy sensitiveness of the variables
- lack (in some) of balancing factors resulting in counter-intuitive behaviour in forecasting mode
- too simple traffic assignment.

These models are indeed simple and can only respond to changes through the cost or distance component. The deterrence function can be made quite flexible as in Hogberg's model and one may even conceive the use of generalised costs as measure of separation.

These generalised cost should have parameters (value of time and so on) externally determined as in many conventional studies. It is not out of the question to attempt to combine such an elaborate model with similar ones for other modes of transport. The author is aware of at least one such attempt (unpublished) by foreign consultants in a developing country. As mentioned before, balancing factors and better assignment model can also be incorporated.

One of the most attractive markets for this type of approach is in less developed countries. The poor data bases available and the relative simplicity of the networks make this approach particularly suited to these countries. Having this type of application in mind, it may be of interest to devise the best model which could be calibrated from traffic counts using only linear regression techniques. This model would be of the form

$$X^q_{ij} = \frac{G^q_i E^q_j}{\sum_j E^q_j} C^{-k}_{ij} \quad , \tag{3.28}$$

where q stands again for journey purpose,

$$Y^q_{\ell m} = \sum_{ij} X^q_{ij} \cdot p^{\ell m}_{ij} \tag{3.29}$$

and $$V_{\ell m} = b_0 + b_1 Y^1_{\ell m} + b_2 Y^2_{\ell m} + b_3 Y^3_{\ell m} + \ldots \quad . \tag{3.30}$$

The model would be calibrated using linear regression for 3 or 4 values of the exponent k and best fit would be chosen. With classified counts the model could be calibrated for car, buses and freight.

A summary of the models in this section is presented in Table 3.1, which also includes the model just described.

3.3 APPROACHES BASED ON DIRECT DEMAND MODELS

3.3.1 Background

The few models in this section are rooted in the tradition of direct demand modelling in transport. Early criticism of the sequential (4-stage) structure of conventional transport modelling resulted in an alternative approach which uses a single estimated equation to relate travel demand by mode directly to modal attributes. *

Probably the first direct demand model is the SARC

* A variation on this is the quasi-direct approach which accepts a form of separability between mode split and total travel demand problems.

Table 3.1: Comparison of gravity model based approaches

Model	Local traffic intercept b_0	Number of journey purposes	Balancing factor	Deterrence function $f(c_{ij})$	Assignment proportional/ iterative	Calibration technique	Applications
Low	Yes	3	No	c_{ij}^{-k}	prop.	linear regr.	Urban
Overgaard	No	2	No	c_{ij}^{k}	prop.	mixed	Urban
Hogberg	No	3	$(\sum_j E_j f(c_{ij}))^{-1}$	$c_{ij}^{k}\exp(b_4(\log c_{ij})^2)$	prop.	non-linear regression	---
Commonwealth Bureau of roads	No	6 based on rank values	No	c_{ij}^{-kq}	prop.	linear regr.	inter-urban
Danish Road Directorate	No	1	No	c_{ij}^{-k}	iterat.	maximum likelihood	inter-urban
Robillard	No	1	implicit	any	prop.	non-linear regression	---
Best linear model	Yes	3	$(\sum_j E_j)^{-1}$	c_{ij}^{-k}	prop.	linear regr.	---

(Kraft, 1963) whose implicit form can be written as

$$T_{ijs} = b_0^s \prod_u A_{iju}^{b_u^s} \prod_w \prod_v C_{ijvw}^{d_{swv}} \tag{3.31}$$

where the following sub(super)-scripts refer to the variables:

w and s	- mode of travel
u	- activity or socio-economic variable
v	- level of service characteristic
A_{iju}	is a function of the u^{th} activity and socio-economic variable, for example population
C_{ijvw}	is the element of cost associated to 'level of service' variable v on mode w between origin i and destination j

b_0^s , b_u^s and d_{swv} are the parameters for calibration.

This is a multiplicative model with potentially very many parameters. Generalised friction factors are obtained by an (exponentially) weighted product of modal attributes. The elasticity of demand to changes in the characteristics of the mode itself can be obtained from making $w = s$ whereas the demand elasticities to changes in other modes or cross elasticities are obtained from making $w \neq s$. The abstract mode models suggested by Quandt and Baumol (1966), and a variation on the SARC model put forward by Domencich et al (1968) are some of the best known direct demand models.

These models have not been used much in the United Kingdom but because of their non-sequential nature they lend themselves to calibration from observations on link volumes. It must be said though, that potential users would be much more familiar with conventional (sequential) approaches.

3.3.2 Wills's approach

Michael Wills has produced a very complete investigation of the use of direct demand models in the estimation of trip matrices from traffic counts (Wills, 1978). He considered several variations on the multiplicative model in (3.31) of which the gravity model is a special case.

He also used a general deterrence function of the type proposed by Tanner (1961) and similar to the one used by Hogberg (1975).

Wills also studied the potential of simpler linear models but concentrated his efforts on the multiplicative model. Wills used a quasi-Newton algorithm to minimise the sum of the square of the differences between observed and modelled link volumes,

$$\sum_{\ell m} (\hat{V}_{\ell m} - V_{\ell m})^2 = \sum_{\ell m} (\hat{V}_{\ell m} - \sum_{ij} T_{ij} p_{ij}^{\ell m}) \tag{3.32}$$

where T_{ij} is replaced by Equation (3.31). The best functional form for the direct demand model and the parameters that result in a best fit can be obtained.

The main attraction of Wills's approach is the wide range of variables which can be introduced in this model. He tried, for example, three groups of variables for an inter urban system:

- socio-economic variables like population and employment
- accessibility variables; he used a centrality index assuming that towns in the periphery will have a different propensity to generate trips than those near the 'centre'
- separation, time or distance, variables.

The models were tested by Wills on two transport systems. The first test was performed on a coarse version of the Canadian highway network in its entirety consisting of 107 zones and 300 links. The second was on the road network in British Columbia including 76 centroids and only 3 parameters. As traffic counts for several years were available it was possible to compare the performance of the model in the forecasting mode.

In both cases the comparisons were made against observed flows and not against 'real' O-D matrices. In both cases the predictions of the model were found 'satisfactory', for example, in the British Columbia case an R^2 of 0.82 was reported. Chan et al (1980) report that a similar model has been put forward by Carey et al (1979) at Carnegie Mellon University.

3.3.3 Lamarre's model

Following some of the suggestions of Wills (1978) Louis Lamarre developed and tested a simple linear model incorporating calibratable parameters affecting socio-economic and travel time indicators, see Lamarre (1977) and Gaudry and Lamarre (1978). Lamarre developed two alternative formulations. The first one is an 'O-D based' formulation:

$$T_{ij} = b_0 + b_1(G_i+G_j) + b_2C_{ij} \tag{3.33}$$

and

$$V_{\ell m} = b_0 \sum_{ij} p_{ij}^{\ell m} + b_1 \sum_{ij} (G_i + E_j)p_{ij}^{\ell m} + b_2 \sum_{ij} p_{ij}^{\ell m}C_{ij} \tag{3.34}$$

where G_i and E_j are indicators of the generation and attraction 'power' of each zone and C_{ij} is the travel time between zones.

His second model is a 'link-based' formulation where

$$V_{\ell m} = b_0 + b_1 \sum_{ij} (G_i + E_j)p_{ij}^{\ell m} + b_2 \sum_{ij} p_{ij}^{\ell m}C_{ij} . \tag{3.35}$$

Lamarre tested his models on an inter-city network and used independent variables including population, employment, total income and car ownership.

His results were not as good as those obtained from the application of Wills's models to the same network. This is not entirely surprising as the models were developed for the sake of computational efficiency rather than any accepted theory about travel behaviour.

3.3.4 Comments

There is some attraction in embedding a direct demand model in the process of estimating a trip matrix from traffic counts, in particular because of the generality of the demand function. Interesting issues like the best functional form for the model can then be discussed and the potential user has the choice of employing the planning and socio-economic data already available.

The drawback of this approach seem to be two-fold. Firstly, as Lamarre's work has shown, it is almost certain that a non-linear model will have to be fitted and this implies the use of less familiar and more expensive software. Secondly, as Wills himself acknowledges, potential users are much more familiar with sequential than with direct demand models. A model making use of familiar concepts like trip generation parameters such as those described in Section 3.2 are more likely to gain acceptance in practice. But the final test must surely be which model best reproduces the underlying O-D matrix rather than the observed traffic counts and so far all the models have only undergone this second test.

3.4 MODELS BASED ON NETWORK DATA ONLY

3.4.1 Background

In many cases an O-D matrix will be required for areas which cannot be said to be self contained, in other words local sub-areas like a city centre. In these cases, the observations will be made over only part of the real trip length and any indicator of separation (distance) is likely to be misleading. The models described in this section tackle this problem making use of network data only rather than assuming an explicit demand model.

The first group of models follow a heuristic approach without much analysis of the theoretical aspects behind the solution methods adopted. The second group approaches the problem from the point of view of equilibrium assignment and develops a mathematical programming framework consistent with this tradition. Finally, models based on entropy maximising (or information minimising) principles are described. The model put forward by the author is a member of this family and is discussed in greater detail in Chapter 4.

3.4.2 Heuristic approaches

3.4.2.1 NEMROD model. The Institut de Recherche des Transports in Paris has developed the model NEMROD to estimate trip matrices at different times of the day. The model, as described by S.Debaille (1977), attempts to answer the question:"Given a good estimation of the O-D matrix for a particular time of the day, say 8-9 am, how could traffic counts on links at other times be used to estimate the corresponding trip matrices."

A simple 'distribution' model is used of the form

$$T_{ij}^{t} = R_{i}^{t} S_{j}^{t} B_{ij} \qquad (3.36)$$

where T_{ij}^{t} is the number of trips between i and j during time slice t

R_{i}^{t} and S_{j}^{t} are generation and attraction,

and B_{ij} is an interchange indicator.

The interchange indicators are meant to represent the level of 'connectivity' or association between two zones and they are assumed to be symmetric and independent of the time of day:

$$B_{ij} = B_{ji} \quad .$$

These interchange indicators are calculated from the original O-D matrix $[T_{ij}^{*}]$. Several ways of doing this were tested, for example

$$B_{ij} = (T_{ij}^{*} + T_{ji}^{*}) / 2 \qquad (3.37)$$

but it was found that the results were very similar.

The parameters R_i^t and S_j^t are then to be calibrated by minimising the error measure

$$(\hat{V}^t_{\ell m} - \sum_{ij} R_i^t S_j^t B_{ij} \cdot p_{ij}^{\ell m})^2 \quad . \tag{3.38}$$

The model and variations on it were tested using data collected using time lapse aerial photography over the city of Roanne in France, see Figure 3.1. A rather coarse zoning system was used but some sensitivity analysis was carried out to test whether aggregating the zones even further would improve the model.

The investigators were satisfied with the performance of the model which produced a percentage RMS error of around 15 per cent.

3.4.2.2 Linear programming approaches. Given that in general the problem of estimating an O-D matrix is underspecified it is tempting to consider Equations (3.1) as the constraints in a linear programming framework. This has been suggested by some authors, for example Gur et al (1978) who put forward the model

$$\text{Min } Z = \sum_{ij} (W_{ij}+X_{ij}) + \sum_{\ell m} K_{\ell m}(Y_{\ell m}+Z_{\ell m}) \tag{3.39}$$

subject to

$$\sum_{ij} T_{ij} p_{ij}^{\ell m} - Y_{\ell m} + Z_{\ell m} = V_{\ell m} \qquad \text{some } \ell m$$

$$T_{ij} - W_{ij} + X_{ij} = t_{ij} \qquad \text{all } ij$$

$$T_{ij}, W_{ij}, K_{ij} \geq 0 \qquad \text{all } ij$$

$$Y_{\ell m}, Z_{\ell m} \geq 0 \qquad \text{some } \ell m$$

where $[t_{ij}]$ is a 'target' O-D matrix, for example an outdated trip table

W_{ij} and X_{ij} are the positive and negative difference between T_{ij} and t_{ij}

$Y_{\ell m}$ and $Z_{\ell m}$ are the positive and negative difference between modelled and observed flows

$K_{\ell m}$ are (arbitrary) weightings for link errors.

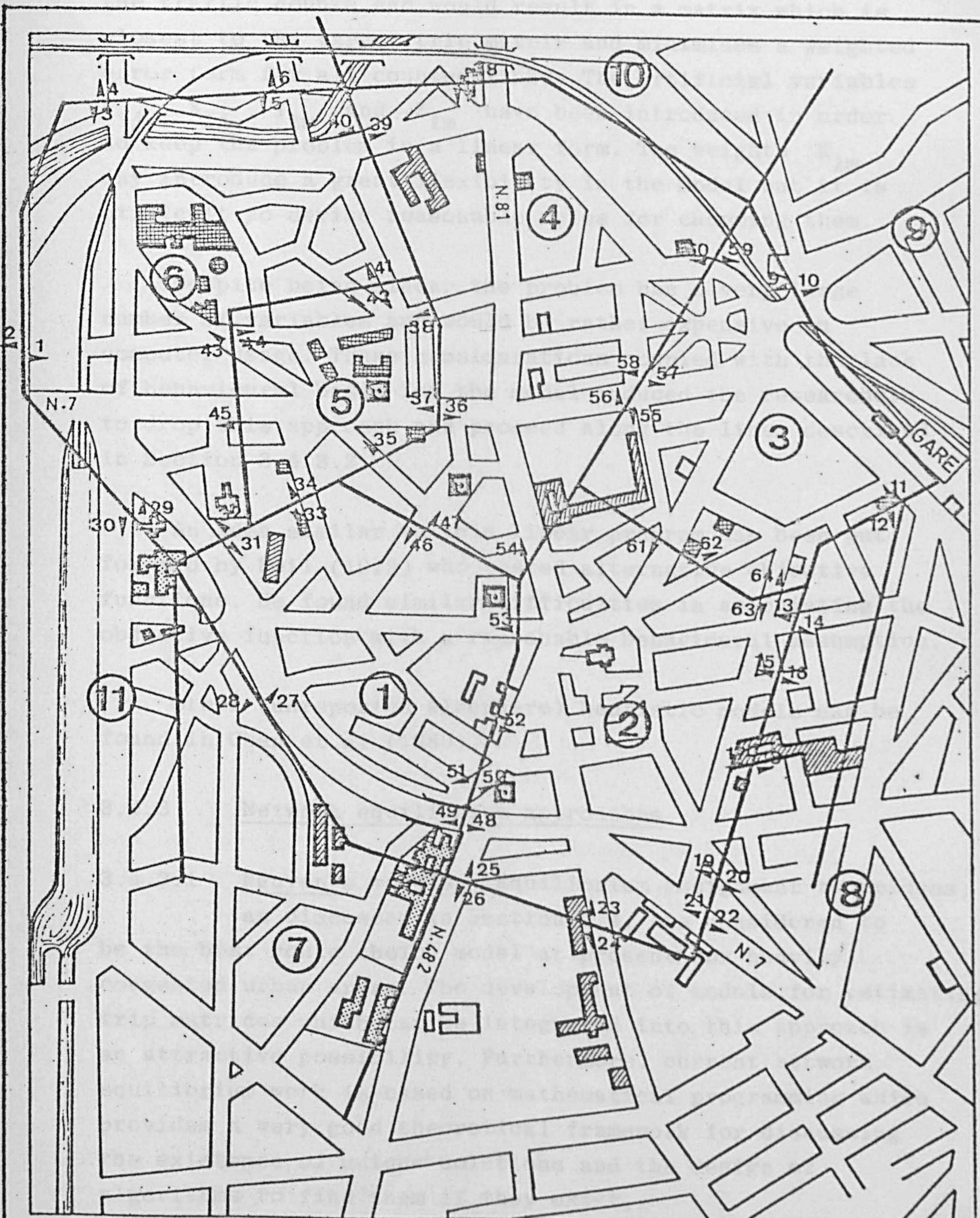


Figure 3.1: <u>Study area in Roanne, France</u>

This is a linear program which allows for errors in the traffic counts and would result in a matrix which is closest to the target trip matrix and minimises a weighted error term for all counted links. The artificial variables W_{ij} , X_{ij} , $Y_{\ell m}$ and $Z_{\ell m}$ have been introduced in order to keep the problem in a linear form. The weights $K_{\ell m}$ may introduce a great flexibility in the model but it is difficult to devise reasonable rules for choosing them.

Despite being linear the problem has a very large number of variables and would be rather expensive in computer usage. These considerations coupled with the lack of behavioural basis for the model induced the researchers to drop this approach and proceed along the lines described in Section 3.4.3.2.

An idea similar to this linear program has been put forward by Beil (1979) who tested alternative objective functions. He found similar difficulties in associating the objective function with a reasonable behavioural assumption.

Other (unreported elsewhere) heuristic models can be found in Chan et al (1980).

3.4.3 Network equilibrium approaches

3.4.3.1 Nguyen's models. Equilibrium assignment techniques, as discussed in Section 2.4, are considered to be the best route choice model at present for heavily congested urban areas. The development of models for estimating trip matrices which can be integrated into this approach is an attractive possibility. Furthermore, current network equilibrium work is based on mathematical programming which provides a very good theoretical framework for discussing the existence of unique solutions and the design of algorithms to find them if they exist.

Nguyen (1977) has put forward two mathematical programs whose solutions are O-D matrices which satisfy equilibrium assignment conditions and are consistent with the observed flows. These will be described in some detail here.

(a) Nguyen's model, Case I

The first model proposed by Nguyen requires all links in the network to be counted. The model is described by Nguyen (1977) as follows.

Let $[\hat{V}_{\ell m}]$ be the observed link flows and $[\hat{C}_{ij}]$ the observed costs of travelling between i and j by any used route. For each O-D pair only one $\hat{C}_{ij}$ will exist since $[\hat{V}_{\ell m}]$ is assumed to be in Wardrop's equilibrium. The cost-flow relationship in any link is $C_{\ell m}(\hat{V}_{\ell m})$. If $\hat{T}_{ij}$ is the trip matrix generating the $[\hat{V}_{\ell m}]$ then the equilibrium state is expressed by

$$\sum_{ij} \hat{C}_{ij}\hat{T}_{ij} = \sum_{\ell m} C_{\ell m}(\hat{V}_{\ell m})\hat{V}_{\ell m} \quad . \tag{3.40}$$

Two necessary conditions for a trip matrix $[T_{ij}]$ to be identical to $[\hat{T}_{ij}]$ are

$$\sum_{ij} \hat{C}_{ij}T_{ij} = \sum_{\ell m} C_{\ell m}(\hat{V}_{\ell m})\hat{V}_{\ell m} \tag{3.41}$$

and

$$C_{ij} = \hat{C}_{ij} \quad \text{for all O-D pairs} \tag{3.42}$$

where C_{ij} is the 'modelled' cost on all used routes between i and j when $[T_{ij}]$ is assigned onto the network according to Wardrop's first principle.

Nguyen goes on then to prove that the T_{ij} matrix satisfying (3.41) and (3.42) can be obtained by solving the programme P1 .

$$\text{P1} \quad \text{Min } F_1(V_{\ell m}) = \sum_{\ell m} \int_0^{V_{\ell m}} C_{\ell m}(x)dx \tag{3.43}$$

subject to

$$C_1 \left\{ \begin{array}{llll} T_{ij} - \sum_{p} T_{ijr} & = & 0 & \text{for all } ij \quad (3.44) \\ T_{ijr} & \geq & 0 & \text{for all } r \text{ , } ij \quad (3.45) \\ T_{ij} & \geq & 0 & \text{for all } ij \quad (3.46) \\ V_{\ell m} & = & \sum_{ijr} \delta_{ijr}^{\ell m} T_{ijr} & \quad (3.47) \end{array} \right.$$

$$\text{and} \quad C_2 \quad \sum_{\ell m} C_{\ell m}(\hat{V}_{\ell m})\hat{V}_{\ell m} - \sum_{ij} \hat{C}_{ij}T_{ij} = 0 \tag{3.48}$$

where T_{ijr} is flow on route r connecting i with j, and $\delta_{ij}^{\ell m} = 1$ if T_{ijr} uses link ℓm, $= 0$ otherwise.

Using Kuhn and Tucker conditions Nguyen (1977) established the equivalence between P1 and Equations (3.41) and (3.42). Nguyen then describes an algorithm to solve P1 .

The problem P1 without constraint C_2 has the same structure as a network equilibrium problem. Accordingly, the following algorithm is suggested by Nguyen.

Step 1: Select an initial feasible $[T_{ij}]$, for example $T_{ij} = K / \sum_{ij} \hat{C}_{ij}$ where $K = \sum_{\ell m} C_{\ell m}(\hat{V}_{\ell m})\hat{V}_{\ell m}$. Determine an initial flow pattern $V_{\ell m}$ using T_{ij}.

Step 2: Determine the shortest route between each O-D pair and let C_{ij} be the travel cost on this route.

Step 3: Find the O-D pair rs for which $C_{rs} / \hat{C}_{rs} = \min_{i,j} C_{ij}/\hat{C}_{ij}$ and load $K/\hat{C}_{ij}$ onto the shortest route from r to s. Let $[\hat{V}_{\ell m}]$ be the resulting traffic pattern.

Step 4: Test convergence.
If $\sum_{\ell m} C_{\ell m}(V_{\ell m})(\hat{V}_{\ell m}-V_{\ell m}) / F(V_{\ell m}) < \varepsilon$ (3.49)
the solution is ε optimal and STOP .

Step 5: Find an optimum combination between previous flow pattern and the new one. Determine λ minimising

$$F(V_{\ell m} + \lambda(\hat{V}_{\ell m}-V_{\ell m})) \tag{3.50}$$

subject to $0 \leq \lambda \leq 1$.

Step 6: Revise the trip matrix and flows as follows

$$T_{ij} = T_{ij} - \lambda T_{ij} \quad \text{for all } ij \neq rs \tag{3.51}$$

$$T_{rs} = T_{rs} + \lambda\left(\frac{k}{\hat{C}_{rs}} - T_{rs}\right) \tag{3.52}$$

Return to Step 2.

(b) Nguyen's approach, Case II

The Case I method is only suitable where it is possible to obtain traffic counts for all links. This would restrict its application to small networks. In addition, the algorithm is likely to be slow for a large number of O-D pairs. Nguyen then proposed a second method which reduces somewhat the data requirements. What is needed now is an estimation of the tòtal travel costs which could be obtained by a floating car survey for example.

He starts by proving that the solution fo the following problem P2 also is a solution to the initial model.

$$\text{P2} \quad \text{Min } F_2(V_{\ell m}, T_{ij}) = \sum_{\ell m} \int_0^{V_{\ell m}} C_{\ell m}(x)dx - \sum_{ij} \hat{C}_{ij} T_{ij} \qquad (3.53)$$

subject to

$$C_1' \begin{cases} T_{ij} - \sum_r T_{ijr} = 0 & \text{for al } ij & (3.54) \\ T_{ijr} \geq 0 & \text{for all } r, ij & (3.55) \\ T_{ij} \geq 0 & \text{for all } ij & (3.56) \\ V_{\ell m} = \sum_{ijr} \delta_{ijr}^{\ell m} T_{ijr} & & (3.57) \end{cases}$$

It can be noted that here only the observed costs $\hat{C}_{ij}$ are required and not the traffic counts.

It should also be noted that P2 has the same form as a formulation of traffic assignment problem with elastic demand. By analogy Nguyen considers that the inverse of the demand function for an O-D pair ij is constant and equal to $\hat{C}_{ij}$. He then suggests that an algorithm such as the one stated in Nguyen (1976) could be used to solve the problem.

Nguyen tested both methods on a synthetic network with 4 centroids and 18 one-way links as in Figure 3.2. This is a system with 4 unknowns and 18 equations. The errors obtained were of about 16 per cent for the T_{ij} cells, and about 3 per cent on link flows. Both methods gave an answer in this case with approximately the same precision.

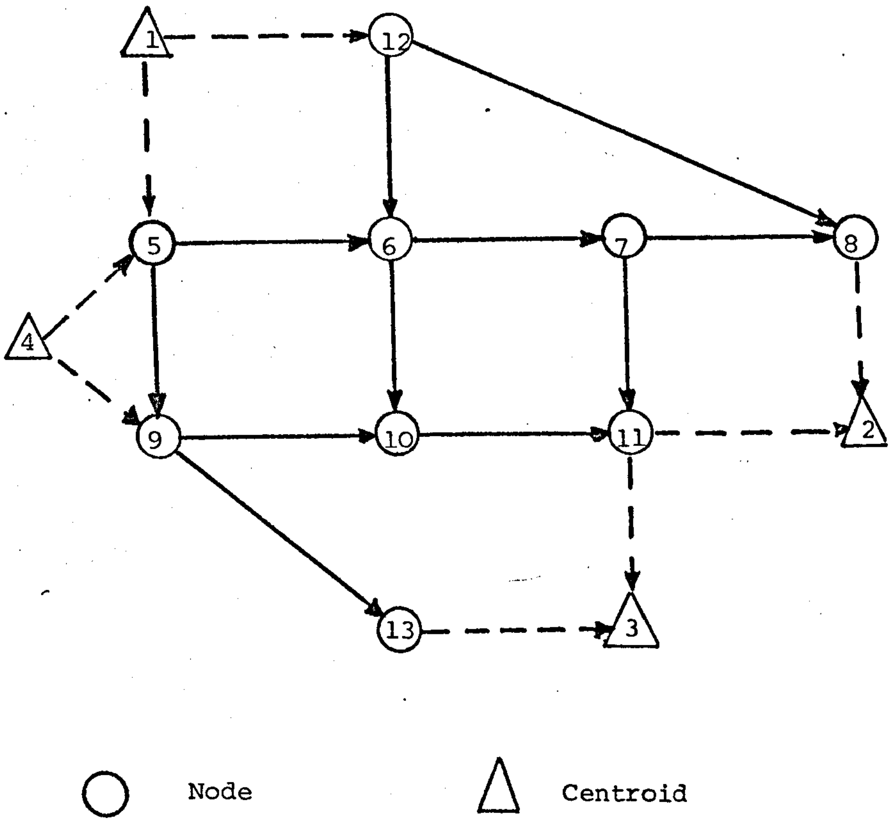


Figure 3.2: Nguyen's test problem network

As Nguyen himself pointed out the solution to this problem is in general not unique and there are likely to by multiple optima to P2 . This poses the problem, yet again, of how to choose among them.

An interesting alternative is to choose the matrix which maximises an entropy measure of the trip matrix. This then becomes the following.

(c) Jörnsten and Nguyen's model, Case III

The rationale behind choosing the trip matrix which maximises the 'entropy' of the system is discussed in greater detail in Chapter 4. Here the modified network equilibrium problem will be presented.

Jörnsten and Nguyen (1979) modify P2 which now becomes

$$\text{P3} \quad \text{Max } F_3(T_{ij}) = -\sum_{ij} T_{ij} \log T_{ij} \qquad (3.58)$$

subject to

$$C_3 \left\{ \begin{array}{ll} \sum_{ij} T_{ij} - \hat{T} = 0 & (3.59) \\ \sum_{ij} \hat{C}_{ij} . T_{ij} - \sum_{\ell m} \hat{C}_{\ell m} . \hat{V}_{\ell m} = 0 & (3.60) \\ T_{ij} \geq 0 & \end{array} \right.$$

$$C_4 \left\{ \begin{array}{lll} T_{ij} - \sum_{r} T_{ijr} = 0 & \text{all } i , j & (3.61) \\ \sum_{ijr} T_{ijr} \delta_{ijr}^{\ell m} - \hat{V}_{\ell m} = 0 & \text{all } \ell m & (3.62) \\ T_{ijr} \geq 0 & & \end{array} \right.$$

The first group of constraints, C_3 , corresponds to constraints which involve only T_{ij} and observed travel costs. The second group, C_4 , involves network constraints and in particular Equations (3.62) which make the problem difficult to tackle.

Jörnsten and Nguyen put forward a decomposition-relaxation approach. The problem P3 is decomposed into two. The master problem involves P3 and the non-network constraints C_3 which results in a classical form of the gravity model. (Wilson, 1970). The second problem is a feasibility problem which involves the relaxed constraints C_4 . The general scheme for the solution algorithm is as follows.

(i) Solve the master problem P3 and obtain a trip matrix $[T_{ij}]$.

(ii) Check the feasibility of the relaxed constraints with respect to the current $[T_{ij}]$. If the constraints are feasible terminate, otherwise add a cut (constraint) to the master problem and repeat the procedure.

The feasibility problem in (ii) is the most difficult part. This problem is tackled by solving the auxiliary equilibrium assignment problem

$$\text{Min} \sum_{\ell m} \int_0^{V_{\ell m}} C_{\ell m}(x)dx \qquad (3.63)$$

subject to

$$\sum_r T_{ijr} - T_{ij} = 0 \qquad \text{all} \quad ij \tag{3.64}$$

$$T_{ijr} \geq 0 \qquad \text{all} \quad ijr \tag{3.65}$$

If the resulting total costs are equal to the observed total costs the constraints C_4 are met and then terminate. If this is not the case, the following constraint should be added to the master problem

$$\sum_{ij} C_{ij} T_{ij} \leq \sum_{\ell m} C_{\ell m}(V_{\ell m}) . \hat{V}_{\ell m} \tag{3.66}$$

So far this third approach has only been tested with fairly small networks using artificial data.

3.4.3.2 The FHWA model. The Federal Highway Administration of the US Department of Transportation asked consultants to develop a practical model for estimating an O-D matrix from traffic counts for use in small central areas. After discussing some alternative approaches the work centered around improving Nguyen's approach (Case II).

The main problem to be tackled was again how to choose among the feasible trip matrices but now a different approach to that of Jörnsten and Nguyen (1979) was followed.

The consultants (Turnkist and Gur, 1979), formulated the problem as one of

$$\text{P4} \quad \text{Min } F_4 = \sum_{\ell m} \int_0^{V_{\ell m}} C_{\ell m}(x) dx - \sum_{ij} \hat{C}_{ij} T_{ij} \tag{3.53}$$

subject to

$$T_{ij} - \sum_r T_{ijr} = 0 \qquad \text{all} \quad ij \tag{3.54}$$

$$\sum_{ijr} T_{ijr} \delta_{ij}^{\ell m} - \hat{V}_{\ell m} = 0 \qquad \text{all} \quad \ell m \tag{3.67}$$

$$T_{ijr} \geq 0 \qquad \text{all} \quad ijr$$

$$T_{ij} \geq 0 \qquad \text{all} \quad ij$$

The main difference with Nguyen (Case II) is that in Equation (3.67) the observed rather than the modelled flows appear. The non-uniqueness of the solution to P4 was tackled by creating a 'target trip matrix'. A modified algorithm was developed which found an O-D matrix solving P4 and laying closest to this target O-D matrix.

The algorithm in summary is as follows.

(i) Find the target O-D matrix $[T^*_{ij}]$.
Set $[T_{ij}] \equiv [T^*_{ij}]$, let $[C_{ij}]$ be free flow costs.

(ii) Load the estimated trip matrix $[T_{ij}]$ to the network using $[C_{ij}]$ and obtain modelled link flows $[V_{\ell m}]$.

(iii) Find the corresponding new link costs and build new minimum cost trees.

(iv) Find a correction trip table $[T^c_{ij}]$ which is closer to a solution and load it to the network using new trees obtained in (iii) and obtain a set of correction flows $[V^c_{\ell m}]$.

(v) Find a new set of trips $[T_{ij}]$ and modelled flows $[V_{\ell m}]$ as the linear combination of the previous ones and the correction trip tables and flows $[T^c_{ij}]$, $[V^c_{ij}]$.

(vi) Check for convergence. If not satisfied, go to step (iii).

The consultants (see Turnkist and Gur, 1979, and Gur et al, 1978) tested several heuristic methods for estimating the correction trip table $[T^c_{ij}]$. They found that the best approach was to make

$$T^c_{ij} = T_{ij} + 2(\hat{C}_{ij} - C_{ij}) / (C_{ij} - C^*_{ij}) \qquad (3.68)$$

where C^*_{ij} are the costs under free flow conditions.

Turnkist and Gur also made use of simplified pseudo-delay functions. Sang Nguyen (1978) had found that in the use of his method (Case II) the solution replicated the observed link flows provided that the cost-flow relationship

- is an increasing function of the link volume
- takes the value of the observed cost at the observed link volume.

These conditions give some room to select suitable functions and Turnkist and Gur used piecewise-linear pseudo-delay functions of the type in Figure 3.3.

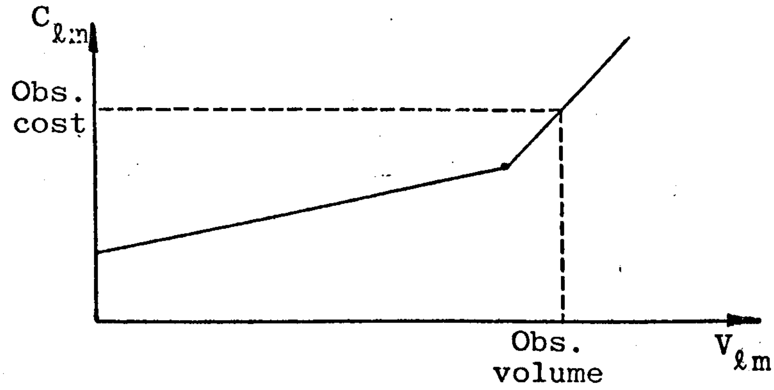


Figure 3.3: Typical linear pseudo-delay functions

In order to choose a sensible 'target' trip matrix a special distribution model for small areas (SMALD) was developed by Kurth et al (1979). This model takes the form

$$T^*_{ij} = \frac{R_i S_j}{(R_i+S_j)} f(c_{ij}) \quad . \qquad (3.69)$$

The parameters R_i and S_j are related to the entry and exit flows at particular links and the domains associated with the respective link type. The idea of a domain here is to identify the part of a region served by a particular type of link or point. The characteristic forms of domains are depicted in Figure 3.4.

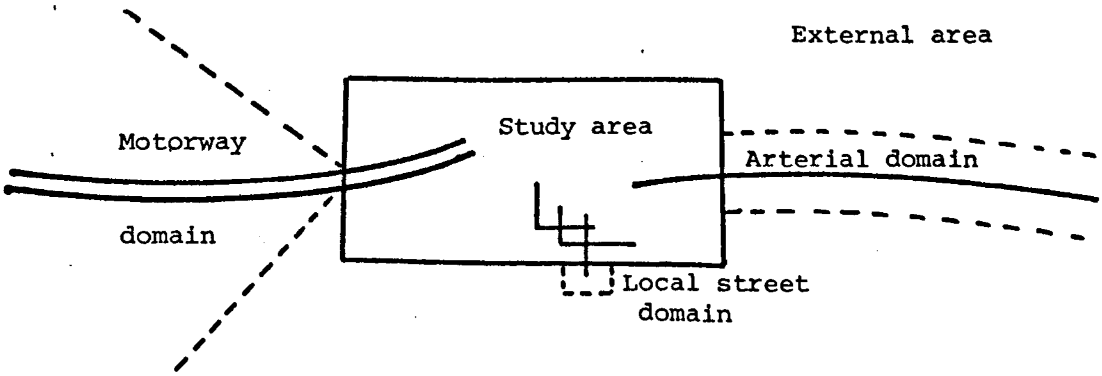


Figure 3.4: Domains in SMALD

The deterrence functions $f(c_{ij})$ are modified Bessel functions which again depend on the domain in question. Typical deterrence functions are shown in Figure 3.5.

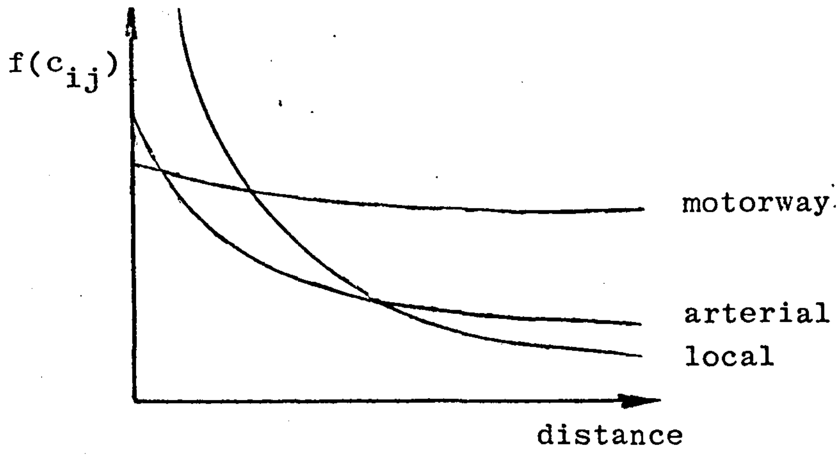


Figure 3.5: Typical deterrence functions

The authors accept that there are limitations to the theory behind SMALD. The models, both SMALD and O-D estimation, have been tested against observed link volumes in Hudson County, New Jersey. The network had 58 zones and 369 links. The observed flows were reproduced with a relative RMS error of between 14 and 18 per cent.

Finally and within the same framework, LeBlanc and Farhangian (1980) have explored alternative algorithms for estimating a matrix closest to the target trip table and solving Nguyen's problem. The problem was tackled in two stages. First Nguyen's problem P2 (Equations (3.53)-(3.57) is solved and the value F_2^* is found for the objective function. Then, in order to choose a trip table the following auxiliary problem is set up.

$$\text{P5} \quad \text{Min } F_5 = \sum_{ij} (T^*_{ij} - T_{ij})^2 \qquad (3.70)$$

subject to

$$T_{ijr} \text{ and } T_{ij} \geq 0 \qquad \text{all } i,j,r \qquad (3.71)$$

$$\sum_{ij} T_{ijr}.p^{\ell m}_{ijr} - V_{\ell m} = 0 \qquad \text{all } \ell m \qquad (3.72)$$

$$\sum_{ij} \int_0^{V_{\ell m}} C_{\ell m}(x)dx - \sum_{ij} \hat{C}_{ij}T_{ij} \leq F_2^* \qquad (3.73)$$

The new restriction (3.73) results from solving the original problem. The inequality sign is due to the approximate character of the original solution. LeBlanc and Farhangian discussed alternative solution methods to these problems and tested them with a network with 24 centroids and 76 links on which a known trip table had been loaded using equilibrium assignment techniques and the resulting link flows were used as 'observed' volumes. The solutions tended to reproduce the 'observed' volumes but the resulting trip matrices were found to be significantly different from the original one. The solution methods developed were not considered by the authors to be suitable for very large networks.

3.4.4 Entropy maximising approaches

3.4.4.1 Background. The concept of entropy and its related measure of information has found several applications in transport, urban and regional systems, see for example Wilson (1970). Its best known application is in the derivation of a fully constrained gravity model as the most likely arrangement of trips consistent with trip end and total cost constraints.

Wilson and MacGill (1977) have observed that the entropy maximising method is particularly suited to modelling systems with large numbers of components with apparent disorganized complexity. The approach supplies the minimum of external information (structure) so that a problem can be solved.

It is possible to argue that when one is assuming an underlying transport demand model and seeks to calibrate it from traffic counts one is probably not fully using the information content of the observed flows. This is to some extent confirmed by a theoretical study by Hogberg (1975). He studied the contribution of each extra link count on the accuracy of his non-linear model (for his 16 nodes, 44 one-way links network). He observed that after including the 5 most important one-way links "the gain in precision is very small" using artificial data and a gravity model with precisely 5 parameters for calibration.

Only an outline of two similar models put forward by the author and by Henk Van Zuylen will be given here for completeness. The whole of Chapter 4 is devoted to entropy maximising methods and these two models.

3.4.4.2 Van Zuylen's model. In many cases some other information is available in addition to traffic counts, for example an old O-D matrix. It would be advantageous to be able to use this information as well. To this end, it is possible to use an information-minimising framework, closely and unambiguously related to the entropy maximising approach.

Van Zuylen (1978) put forward a model based on Brillouin's information measure for estimating an O-D matrix from counts where an 'a priori' trip matrix (t_{ij}) is available.

The model requires minimising.

$$I = \sum_{\ell m} \sum_{ij} T_{ij} p_{ij}^{\ell m} \log \left(\frac{T_{ij}}{V_{\ell m} t_{ij}} \right) \tag{3.74}$$

subject to the usual

$$V_{\ell m} - \sum_{ij} T_{ij} p_{ij}^{\ell m} = 0 \qquad \text{some } \ell m \text{ (counted)} \tag{3.75}$$

The solution to this problem results in a multi-proportional model of the form

$$T_{ij} = t_{ij} \prod_{\ell m} X_{\ell m}^{p_{ij}^{\ell m}/g_{ij}} \tag{3.76}$$

where

$$g_{ij} = \sum_{\ell m} p_{ij}^{\ell m} \quad . \tag{3.77}$$

The solution to this model is discussed in detail later.

3.4.4.3 Willumsen's model. The author has generated a model for estimating an O-D matrix from traffic counts under proportional assignment conditions from an entropy maximising approach. The derivation parallels the derivation of a gravity model but replaces the usual trip end and cost constraints by constraints associated with the traffic counts; see Willumsen (1978a). The problem can be stated as

$$\text{Max } S = -\sum_{ij} (T_{ij} \log_e T_{ij} - T_{ij}) \qquad (3.78)$$

subject to

$$V_{\ell m} - \sum_{ij} T_{ij} p_{ij}^{\ell m} = 0 \quad . \qquad (3.75)$$

Using Langrangian methods the formal solution to this program is obtained as

$$T_{ij} = \prod_{\ell m} X_{\ell m}^{p_{ij}^{\ell m}} \qquad (3.79)$$

where

$$X_{\ell m} = e^{-\lambda_{\ell m}} \qquad (3.80)$$

and $\lambda_{\ell m}$ is the Langrangian multiplier associated with the count on link ℓm .

This model can also be extended to incorporate prior information about the trip matrix and the solution becomes simply

$$T_{ij} = t_{ij} \prod_{\ell m} X_{\ell m}^{p_{ij}^{\ell m}} \quad . \qquad (3.81)$$

Some properties of this model are worth stating at this stage.

(a) The model requires the traffic counts to be consistent although not necessarily independent.

(b) The model generates the most likely O-D matrix consistent with the information contained in the counts. It does not require counts on <u>all</u> links in the network. Of course a more complete set of counts is likely to improve the accuracy of the estimated trip matrix.

(c) In the limit the solution <u>always</u> reproduces the observed flows.

The multiproportional model in Equation (3.81) can be solved using an extension of the Furness method for balancing row and columns in a distribution model.

3.4.5 Comments

Models based on network data only offer a very interesting alternative for local areas where it may be more difficult to assume any particular travel demand model to be realistic. Among these models, those based on equilibrium assignment and entropy maximising considerations possess a more advanced theoretical framework. The final test for any such model must be in its ability to replicate the underlying O-D matrix from traffic counts only and this will be attempted from Chapters 6 onwards.

3.5 SIMPLE SYSTEMS

There are a number of cases in which an O-D matrix is to be estimated from traffic counts (or equivalent) without some of the complexities of the problems described so far. In these cases, some of the models discussed above can be used but at the same time the problem can be made simpler due to the structure of the system of interest.

Applications to this type of problem will be discussed here. All the problems have in common their relatively small size and the fact that traffic assignment does not play a role in them. There is always only one route for joining origin and destination.

3.5.1 Estimating turning movements at junctions

In this case traffic counts are available in and out of a junction (perhaps only automatic counters were used) and the proportions or flows turning left, right and going straight ahead are wanted for each stream. In essence, a mini trip matrix is required, see Figure 3.6.

Jeffreys and Norman (1977), and Marshall (1979) have suggested some heuristic methods for estimating turning flows. But perhaps more interestingly, Mekky (1979) proposed independently an entropy maximising solution, which is equivalent to Van Zuylen's and Willumsen's models (these

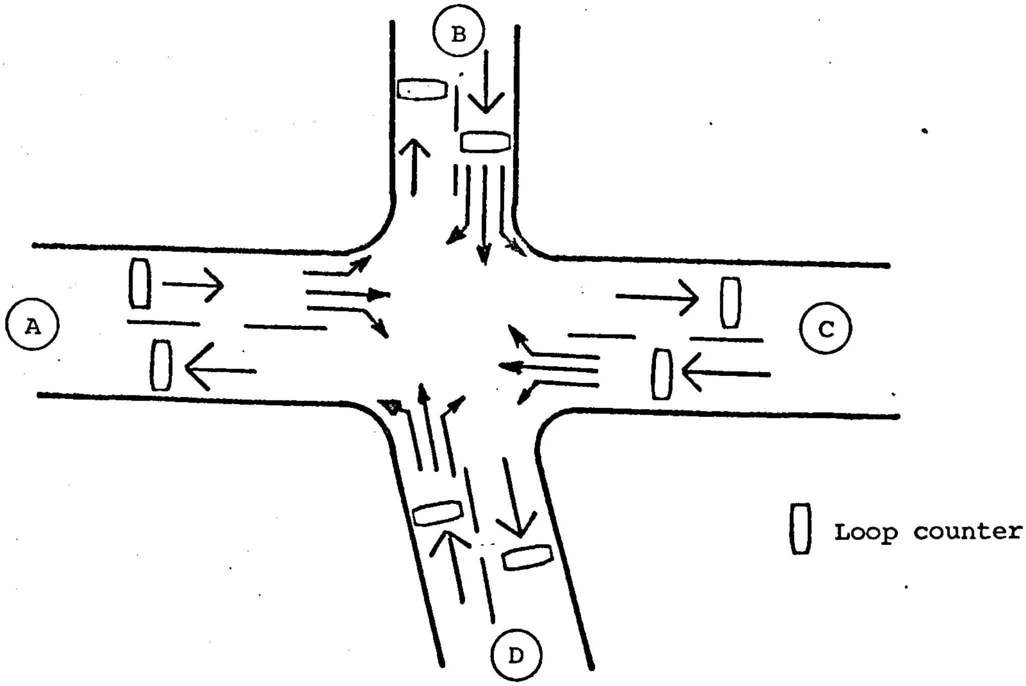


Figure 3.6: A junction with turning movements

two coincide in this case) as shown later by Van Zuylen (1979). In this case the models result in a bi-proportional problem

$$T_{ij} = t_{ij}\, a_i\, b_j \tag{3.82}$$

where a_i and b_j are to be found using Furness type iterations from

$$\textstyle\sum_i T_{ij} = b_j \tag{3.83}$$
$$\textstyle\sum_j T_{ij} = a_i \tag{3.84}$$

This type of model had already been discussed by Potts and Oliver (1972) although without mentioning this particular application.

As reported by Van Zuylen (1979) this model produces quite good results, in particular if a good prior estimation t_{ij} is available, perhaps from an old full set of counts.

3.5.2 Flows on motorways

A second type of simple system is a stretch of motorway where sliproads (or ramps) provide the only entry and exit points. The problem here is to estimate the trip matrix with counts only on the sliproads and on the motorway itself. (Only the counts at the beginning and end of the area of interest on the motorway are relevant here.)

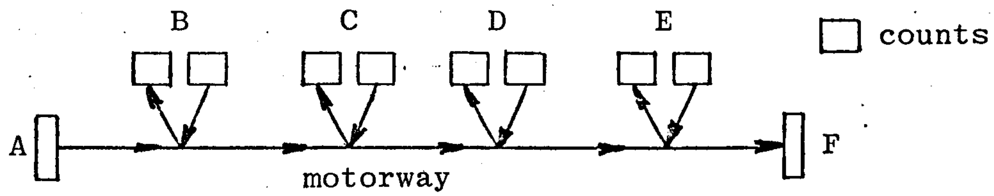


Figure 3.7: Traffic counts on a motorway

In this case the O-D matrix has a triangular shape as no trips are possible, say from D to C or B . This reduced the degrees of freedom as for example flow EF is uniquely determined by motorway bound traffic at E .

Hauer and Shin (1980) have applied the entropy maximising model to this problem taking advantage of this particular structure as shown in Figure 3.8.

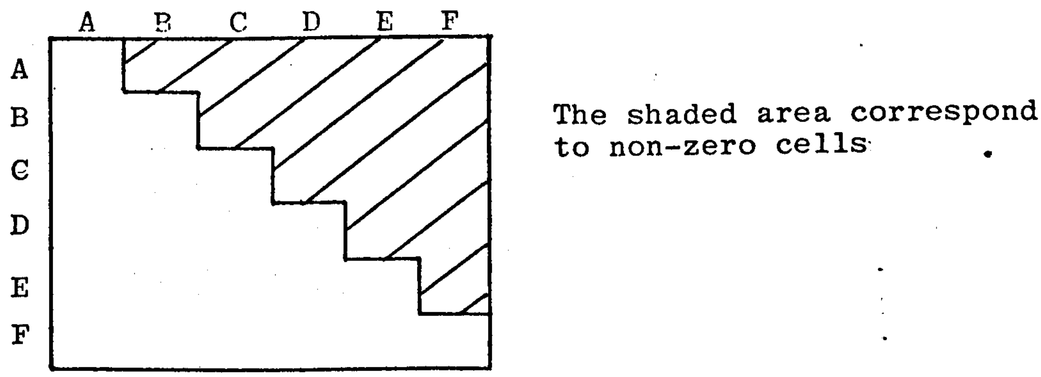


Figure 3.8: Structure of a motorway O-D matrix

Of course in a problem with this structure one knows all the flows in and out of the section and it seems reasonable to assume that the fractions leaving at each exit is independent of entry points. This assumption results here consistent with the entropy maximising solution to this problem. The authors applied the model to the Queen Elizabeth Way in Toronto where an independent O-D survey had been carried out. They found encouraging results as reproduced in Figure 3.9.

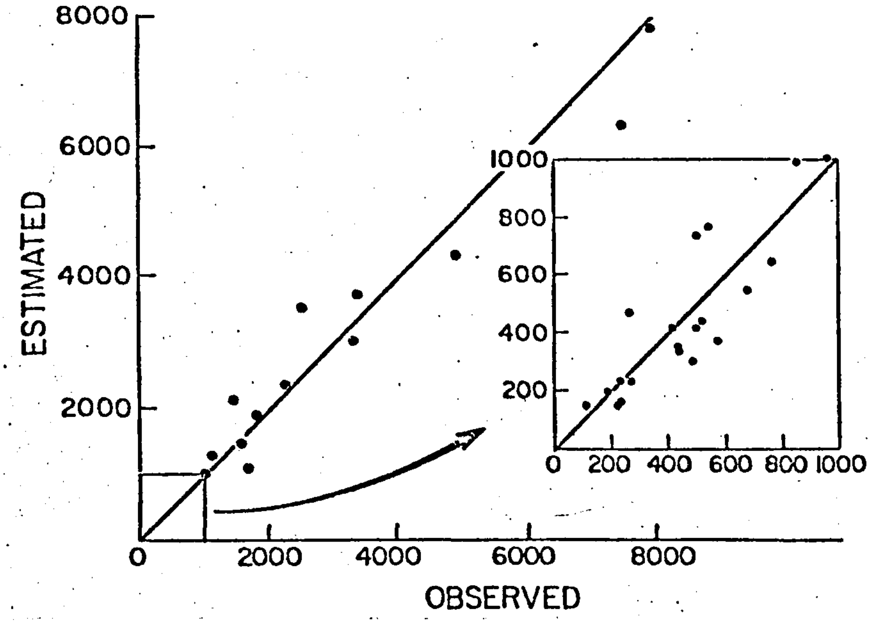


Figure 3.9: Observed vs estimated trips on Queen Elizabeth Way

Chan et al (1980) report the model SYNODM developed by Wang and May (1973) and improved by Eldor (1976) to estimate trip matrices on motorways. The model uses 'technological coefficients' similar to those of the French NEMROD model. Eldor used data from the San Francisco Bay area and Santa Monica to conclude that the derived trip tables were "accurate enough for the design of ramp control" systems.

3.5.3 Flows on rail service line

Another relatively simple system is to be found in the passenger flow pattern on a rail service. Entry and exit points are fixed and counts can be easily obtained of the passengers boarding or leaving the train at each station (at turnstiles for example).

Hauer and Shin (1980) used the entropy maximising model to synthesise trip tables from two commuter lines in Canada and found fairly good results.

A different model to tackle the same problem has been developed by Jensen (1980) in Denmark. In this case passenger counts were not available at stations but on the trains in between stations; these were of course directional counts. Jensen's approach is to study possible distributions for what is equivalent to the technological coefficients of Wang and May. The selection of a particular probability function is given by the nature of the data but she found that the Poisson distribution gave a good fit provided some anomalies in the data were removed (passengers travelling in groups).

One of the advantages of Jensen's approach (or any other based on inexpensive counts) is that the demand at different times of the day can easily be obtained from the corresponding volume counts.

3.5.4 Comments

The work reported in this section is encouraging as it shows that models based on traffic counts can find a wide range of applications. Other public transport services like buses could also benefit from these approaches. Some of the additional problems encountered in these svstems (such as the difficulty of obtaining reliable automatic counts) can be compensated by the simpler structure of the system.

3.6 SYNTHESIS

A number of models for estimating trip matrices from traffic counts has been reviewed. The large majority of these models have been tested against their ability to reproduce the observed counts. It is argued in this work that due to the underspecified nature of the problem this is not a sufficiently exacting test. Little can be said at this stage then on the relative accuracy of each model.

On the other hand, one can speculate on other aspects of the models, for example

- the conditions under which the assumptions are likely to be acceptable,
- the amount of extra data (in addition to traffic counts) required,

- internal consistency,
- their eventual use for forecasting purposes and not just short term estimation of a trip matrix;
- their flexibility for incorporating information already available about trip making behaviour, and
- their general parsimony and transparency.

(a) Approaches based on a travel demand model

This group includes the models reviewed in Sections 3.2 and 3.3. It must be said that most planners are familiar with the idea that travel demand may follow a 'gravity type' pattern. This assumption is more likely to be valid in large free standing areas, where trip cost or trip length is an important factor in the formation of a travel demand pattern.

Of this group, very general approaches like the ones suggested by Hogberg and Wills, are in general more theoretically sound. On the other hand, they tend to require more data and greater effort in calibration. The 'best linear' model goes some way at getting the most out of simplicity and standard software.

The extra data used in the model will depend on the type of data readily available or obtainable with reasonable effort. Robillard's model, for example, passes most specification tests, requires no extra data, but implies a fairly large calibration effort. (The minimum number of parameters to determine by minimum squares techniques is twice the number of zones plus one.)

It is not by chance that most applications of this approach have been in inter-urban areas. The method seems ideally suitable for problems where each town represents a single zone.

Applications to free standing towns will usually imply a more detailed zoning system and an independent roadside interview to determine 'external trips'. Simple congestion effects can be incorporated as shown in the Danish Road Directorate model.

This type of approach is probably not suitable for local areas where probably no demand model relying on a measure of separation or friction will perform well. This would also apply to simpler systems like motorways.

On the other hand, the fact that these models use external information like population and employment make them more suited for planning purposes. The values of these planning variables in the future could be used to run strategic planning models which could be used to select options for a more detailed analysis later. The fact that traffic counts are used in calibration means that the models can be updated and improved as more information becomes available.

(b) Approaches based on network data only

These models seem most suited to small local areas and 'simple systems'. In general they try to make the best use of the data available in the form of network characteristics and counts.

The models based on equilibrium assignment constitute an interesting alternative. Their main drawbacks seem to be the requirement of a full set of counts or at least a travel time survey good enough to estimate times between all O-D pairs. A second problem may be the computational effort required and the soundness of the method used to estimate the 'target' trip matrix.

Models based on entropy maximising principles come from a different theoretical tradition which emphasises the need to 'inject' the minimum of external information into the problem in order to obtain the maximum out of the data available. This scheme provides a good basis for incorporating information about the O-D matrix which is not in the form of counts, for example an old trip matrix or one synthesised from a study of a larger area. It will be shown later how congestion effects can be incorporated into the estimation problem but it is fair to say, at this stage, that this inclusion has not the rigorous treatment found in equilibrium assignment.

The fact that an incomplete set of counts can be used with entropy maximising (and travel demand model based) approaches is a valuable characteristic. It is important as well to be able to improve an estimate as more information (counts) become available without having to undertake another complete (travel time) survey.

The ultimate model would probably be one which

(a) does not require a full set of counts,

(b) can incorporate a wide variety of external information but can also estimate an O-D matrix without it,

(c) can cope with inconsistent traffic counts and errors in the counts themselves,

(d) would assume route choice to be close to equilibrium assignment but could also accept departures from it,

(e) could incorporate information as it becomes available without requiring a full new survey,

(f) would use robust and efficient software,

and most importantly,

(g) would estimate an O-D matrix with a satisfactory degree of accuracy.

The next chapter will explore the extent to which an entropy maximising model developed by the author can meet these requirements.

CHAPTER 4
ENTROPY MAXIMISING APPROACH

This chapter introduces the entropy maximising approach and uses it to develop a model for the estimation of a trip matrix from traffic counts. The chapter is organised as follows. Section 4.1 is devoted to a brief review of the concept of entropy, its relation with the notion of information and the model building method which is based on these ideas. Section 4.2 applies the entropy maximising framework to the problem of estimating an O-D matrix from the information contained in traffic counts. The resulting model is applied to a very simple problem in Section 4.3. Finally Section 4.4 discusses the main properties of the model.

4.1 THE ENTROPY MAXIMISING FORMALISM

The concept of entropy has its origin in physics but its use has been extended to several other fields. This section describes firstly the use of entropy in the context of physical systems and its relation to 'disorder' and information. It is the use of entropy as a measure of information which has generated a large number of applications in other areas and some of them are touched upon later in the section.

Whatever the use, entropy is related to the wellknown functional form $x \log x$. Different theoretical frameworks sometimes modify this functional form slightly and the most relevant variations are identified together with their corresponding interpretations. The section then discusses the use of entropy maximising as a methodology for model building and uses the now conventional derivation of a gravity model to illustrate its application to transport problems.

4.1.1 Entropy in physical systems

The concept of entropy originated in physics, in particular from the analysis of closed physical systems. Consider a physical system with N indistinguishable particles or elements. Each particle can take one of M states i where each state is identified by certain parameters such as kinetic energy or speed, and let m_i be the number of particles in each state i. A micro-state then is a description of the system in which the state of each individual particle is specified. In the absence of additional information one can assume all states to be equally likely.

The number of ways or selections of micro-states, in which a particular system can be achieved is given by the combinational formula

$$W(N) = \frac{N!}{\prod_i m_i!} \tag{4.1}$$

It is not unreasonable to assume that the most likely arrangement for the elements in this closed physical system will be the one that can be achieved in the greatest number of ways. Some configurations will be more likely than others in the sense that they could be obtained in a greater number of ways. For example, an arrangement requiring one particle to be in state 1, the rest in state 2 and none in any of the M-2 remaining states can almost certainly be said to be a very unlikely one. On the other hand, a configuration in which particles can take a wider range of micro-states can be considered more likely.

One may now try to find properties associated with the most likely arrangement of elementary units. The search for the properties may begin with maximising $W(N)$ or a monotonic function of $W(N)$ such as $\log(W(N))$.

$$\log(W(N)) = \log N! - \sum_i \log m_i! \tag{4.2}$$

The first term is a constant which does not affect the optimum point. The short version of Stirling's approximation $\log X! \simeq X \log X - X$ can be used to obtain

$$\log(W(N)) \simeq - \sum_i m_i \log m_i - \sum_i m_i + \log N!$$

which leads to the associated measure

$$S_1 = - \sum_i m_i \log m_i - \sum_i m_i \tag{4.3}$$

The configuration which maximises this entropy measure can be said to be the most likely arrangement of micro-states.

The same idea can be represented in probabilistic terms where $p_i = m_i/N$ represents the probability of a particle i being in state i . From (4.2) then

$$\log(W(N)) = - \sum_i \log\{m_i!/N!\}$$

and using Stirling's approximation and $\sum_i p_i = 1$ leads to

$$S_2 = - \sum_i p_i \log p_i \qquad (4.4)$$

a wellknown entropy measure. S_2 is also called Boltzman's H function (Nathanson, 1978) and its value is maximum when all p_i are equal to the number of possible micro-states M divided by the number of particles N . In other words, the most likely configuration results in a uniform distribution of probabilities which can be associated with a state of maximum disorder.

4.1.2 Entropy and information

Information theory has been linked since its conception to the idea of entropy. It can be seen, intuitively at least, that a state of maximum disorder is also one containing a minimum of information. The potential information content of a message grows as the sequence of symbols departs from a purely random (high disorder) sequence.

Shanon (1948) considered a series of events $i = 1,2,..,N$, and a probability of each event occurring p_i . The maximum uncertainty (or minimum information) occurs when each event is equally probable. Shanon suggested the use of S_2 as measure for information and called it entropy. In this context S_2 has the advantage that it can be axiomatically derived so that it satisfies certain 'natural' conditions for measuring information rate. In particular, S_2 is non-negative and additive for independent events.

A variation of this idea considers information as a relative quantity, which can only be measured by comparing one state of the system with another and computing the information difference or 'gain'. The argument has interesting philosophical implications (Batty and March, 1976b) as it relates to the way in which reality is perceived relative to some expectation or 'a priori' idea about the form of the world.

To measure information gain consider a set of prior probabilities q_i and a set of unknown posterior probabilities p_i. For this relative case the combinational combinatorial formula analogous to (4.1) is

$$W[p:q] = \frac{N!\ \pi\ q_i^{m_i}}{\pi\ m_i!} \qquad (4.5)$$

which is essentially the multinomial probability function for the posterior distribution. Taking natural logarithms and following steps analogous to those leading from (4.1) to (4.2) (including the use of Stirling's approximation for $X!$) one gets

$$\log W[p:q] = \sum_i^N [m_i \log q_i - (m_i \log m_i - m_i)] + N \log N - N$$

which, as $N = \sum_i m_i$, is equivalent to

$$\log W[p:q] = \sum_i [m_i \log q_i - m_i \log m_i + m_i + m_i \log N - m_i]$$

$$\log W[p:q] = \sum_i [m_i (\log q_i - \log \frac{m_i}{N}]$$

Dividing by $-N$

$$-\log W[p:q]/N = -\sum_i [\frac{m_i}{N} \log q_i - \frac{m_i}{N} \log \frac{m_i}{N}] \qquad (4.6)$$

Using the fact that $p_i = \frac{m_i}{N}$ the measure of information gain becomes

$$S_3 = \sum_i p_i \log \frac{p_i}{q_i} \qquad (4.7)$$

This function takes its minimum value of zero when all $p_i = q_i$, that is when there is no difference in the two configurations and hence there is no information gain.

If the distribution of prior probabilities is uniform (i.e. $q_i = 1/M$ for all i) Equation (4.7) becomes

$$S = \sum_i p_i [\log p_i + \log M] \tag{4.8}$$

The minimisation of (4.8) subject to the normalising constraint $\sum_i p_i = 1$ is equivalent to the maximisation of S_2 under the same constraint. There are reasons to prefer S_3 to S_2; see Batty and March (1976b) for example. For a start, S_2 can always be derived from S_3 but not vice versa. But more importantly, S_3 makes possible the use of prior information in a consistent and efficient manner.

By setting $q_i = m_i^o / N^o$ in Equation (4.6) and multiplying by N we obtain

$$-\log W[m_i : m_i^o] = -\sum_i [m_i \log \frac{m_i^o}{N^o} - m_i \log \frac{m_i}{N}]$$

$$-\log W[m_i : m_i^o] = \sum_i m_i [\log m_i / m_i^o - \log \frac{N}{N^o}]$$

If $N = N^o$ then one obtains the measure of information gain

$$S_4 = \sum_i m_i [\log m_i / m_i^o - 1] \tag{4.9}$$

equivalent to S_1.

March and Batty (1976) propose the use of a generalised information gain measure related to Reyni's α entropy. Reyni's (1965) formula, which March and Batty derive from Bayes's theorem, becomes then

$$S_5 = \frac{1}{\alpha - 1} \log \left[\sum_i p_i \left(\frac{p_i}{q_i} \right)^{\alpha - 1} \right] \tag{4.10}$$

It can be shown, using l'Hopital's rule, that S_5 becomes S_3 when α is made equal to one. Alternative values for α would result in other entropy-like measures of information gain.

These five measures are not the only ones. Other authors have proposed a number of related measures and put forward complementary interpretations for entropy. For example Walsh and Webber (1977) discuss alternative information statistics as measures of the uncertainty present in distributions. Tribus (1969) suggested the use of S_3 as an approximation to the chi-square statistic. In the same vein, Erlander (1980) has shown that S_2 is related to the variance of a population and suggested its use as a measure of dispersion. For the same reason Sheppard (1976) has suggested the use of S_2 as a descriptive statistic for spatial patterns. However, it is beyond the scope of this work to discuss these and other proposals in any more detail here.

4.1.3 The entropy maximising formalism

A number of researchers have used the measures S_1 to S_5 as tools for model building and this type of application is of particular interest to this work. The method of entropy maximisation was first formally presented by Shannon (1948) and then extended by Jaynes (1957) who introduced a Lagrangian optimisation procedure. Some time later Wilson (1970) adapted this technique to trip-distribution and other related problems initiating a prolific and important effort by transport analysts and geographers to elaborate, extend and refine the approach. This general strategy for model building has received the name of 'the' entropy maximising framework. As it stands now, it involves the following steps.

(i) An identification of micro-states and their prior probabilities q_i . If nothing is known about them the assumption that all micro-states or 'elementary events' are equiprobable may be satisfactory. An entropy measure is then identified in terms of meso-states, that is aggregations of micro-states at suitable resolution levels.

(ii) An identification of the macro-state constraints for the system. Any real constraint on the states the system can take and any theoretical property which the system is known to have must be incorporated in the constraints. Normally these constraints take a linear form.

(iii) A technique is required to optimise the desired measure of entropy (or information) subject to the constraints identified in (ii).

The measures S_1 to S_5 all have first and second derivatives and the second one has the right sign for optimisation purposes. The functions S_1 and S_2 are strictly convex and S_3 , S_4 and S_5 are strictly concave; the existence of a unique solution to the problem depends on the characteristics of the constraints. Either:

- the constraints do not permit any feasible solutions to the optimisation problem; or
- the constraints are feasible and there is a unique solution to the problem.

The popularity of the entropy maximising method can be said to rely as much on the interesting interpretations, that can be given to the solutions found as on the favourable properties of the entropy measures for optimisation purposes.

A maximum entropy solution can be described as the most probable or unprejudiced meso-state compatible with current knowledge about a system embodied in the constraints (Wilson, 1974). Recent work leading to further applications or refinement of the formalism is reported in Beckman (1974), Wilson (1975), Slater (1978), Erlander (1977), Fisch (1977), Sheppard (1976) and in a very useful review by MacGill and Wilson (1979).

An important development has been the study of the link between geometric programming and entropy maximising techniques analysed by Nijkamp and Paelink (1974) and Dinkel et al (1977). Finally, Coelho and Wilson (1977) have discussed the ways in which entropy maximising submodels can be embedded within overall mathematical programming frameworks.

4.1.4 An example - the gravity model

As an illustration, the now conventional way of deriving a double-constrained gravity model from entropy maximising will be outlined here following Wilson (1970).

First three levels of description of the trips in a transport system are chosen. At the finest level of detail (micro) one could consider each individual tripmaker, his origin and destination (plus time or other travel characteristics). This is the micro-level of description. At the coarsest (macro) level only the total number of trips generated and attracted to each zone (Q_i and D_j) plus the total expenditure on transport C are recorded. Finally, at a medium (meso) level of detail only the total number of trips made for each O-D pair T_{ij} are of interest.

In the absence of other information it is assumed that each micro-state or elemental event is equally likely. The number of micro-states associated with some meso-state $[T_{ij}]$ is given by

$$W([T_{ij}]) = \begin{Bmatrix}\text{number of ways of} \\ \text{selecting } T_{11} \text{ from } T\end{Bmatrix} \times \begin{Bmatrix}\text{number of ways of} \\ \text{selecting } T_{12} \text{ from } T-T_{11}\end{Bmatrix} \times \ldots$$

where T is the total number of trips in the system. The number of ways of selecting T_{11} for example from T is

$$\frac{T!}{(T-T_{11})!T_{11}!}$$

then

$$W([T_{ij}]) = \frac{T!}{(T-T_{11})!T_{11}!} \times \frac{(T-T_{11})!}{(T-T_{11}-T_{12})!T_{12}!} \times \ldots$$

$$W([T_{ij}]) = \frac{T!}{\prod_{ij} T_{ij}!}$$

The matrix $[T_{ij}]$ which maximises $W([T_{ij}])$ is sought and for convenience one maximises the monotonic function

$$\log(W([T_{ij}])) \tag{4.11}$$

subject to the constraints representing the knowledge about the macro-state of the system. These are usually taken to be

$$\sum_j T_{ij} = O_i \tag{4.12}$$

$$\sum_i T_{ij} = D_j \tag{4.13}$$

$$\sum_{ij} T_{ij}c_{ij} = C \tag{4.14}$$

The equivalence between maximising (4.11) and the measure S_1 has already been shown. By replacing T_{ij} for m_i one obtaines the problem

$$\text{Max } S_1 = -\sum_{ij} T_{ij}(\log T_{ij} - 1) \tag{4.15}$$

subject to (4.12), (4.13) and (4.14) and of course to $T_{ij} \geq 0$. In addition to the non-negativity condition the problem has $2N+1$ constraints. The application of the Lagrangian method leads to the Lagrangian

$$\mathcal{L} = S_1 + \sum_i \lambda_i^1 (O_i - \sum_j T_{ij}) + \sum_j \lambda_j^2 (D_j - \sum_i T_{ij}) + \beta(C - \sum_{ij} T_{ij}c_{ij}) \tag{4.16}$$

where λ_i^1, λ_i^2 and β are the multipliers associated to the constraints (4.12) to (4.14). The extreme value of (4.16) is obtained by taking first derivatives with respect to the variables in (4.16).

$$\frac{\partial \mathcal{L}}{\partial T_{ij}} = -\log T_{ij} - \lambda_i^1 - \lambda_j^2 - \beta c_{ij} = 0 \tag{4.17}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_i^1} = O_i - \sum_j T_{ij} = 0 \tag{4.18}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_j^1} = D_j - \sum_i T_{ij} = 0 \tag{4.19}$$

$$\frac{\partial \mathcal{L}}{\partial \beta} = C - \sum_{ij} T_{ij}c_{ij} = 0 \tag{4.20}$$

Equations (4.18) to (4.20) are, as expected, repetition of the constraints, The solution is obtained from (4.17) in terms of the multipliers

$$T_{ij} = e^{-\lambda_i^1 - \lambda_j^2 - \beta c_{ij}} \tag{4.21}$$

The variables λ_i^1 and λ_j^2 can be replaced by

$$A_i O_i = e^{-\lambda_i^1}$$

$$B_j D_j = e^{-\lambda_j^2}$$

and then (4.21) becomes

$$T_{ij} = A_i B_j O_i D_j e^{-\beta c_{ij}} \tag{4.22}$$

where substitution in Equations (4.12) and (4.13) shows that

$$A_i = \frac{1}{\sum_j B_j D_j e^{-\beta c_{ij}}} \tag{4.23}$$

$$B_j = \frac{1}{\sum_i A_i O_i e^{-\beta c_{ij}}} \tag{4.24}$$

are the usual balancing factors of a doubly constrained gravity model. In practice, the value of the total expenditure C is not available directly from survey data and the Lagrance multiplier β must be calibrated from trip length distributions, see Hyman (1969).

For a given value of β the value of each cell T_{ij} must be found by solving (4.22) with (4.23) and (4.24). One practical way of doing this is to use Furness iterations. This method implies solving (4.22) and (4.24) for an initial set of values A_i (say 1) and then using the B_j thus obtained to solve (4.22) and (4.23). The new A_i values are then used to solve (4.22) and (4.24) and the process is repeated until the values of A_i and B_j do not change much in consecutive iterations. The algorithm is then said to have converged to the unique solution of the problem. Because of the nature of the repeated corrections to match row and column totals this algorithm is also called a 'bi-proportional' adjustment procedure.

It can also be seen that β must be such that a feasible trip matrix may exist. Too low a total expenditure may prevent any matrix from meeting trip end constraints, for example. A second condition is the consistency of the trip end constraints, that is

$$\sum_j D_j \text{ must equal } \sum_i O_i \ .$$

As Potts and Oliver (1972) have shown singly constrained gravity models can be obtained by simply omitting one set of trip end constraints, There are also alternative derivations of a gravity model using other measures of entropy and information; see for example Evans (1976) or Erlander (1980).

4.2 AN ENTROPY MAXIMISING MODEL WITH TRAFFIC COUNT CONSTRAINTS

4.2.1 The basic model

Early in this research the author (Willumsen, 1978a,b) developed a model based on entropy maximisation considerations to estimate an origin-destination matrix from traffic counts. This derivation parallels Wilson's derivation for a gravity model as shown in Section 4.14. Following Wilson, the same equiprobable micro-states and the same meso-states $[T_{ij}]$ are chosen. But in this case, the trip end and total cost constraints are replaced by a set of (linear) constraints representing the counted flows. These reflect the current knowledge or macro-state information about the system. The problem then presented is

$$\text{Max } S_1 = -\sum_{ij} T_{ij}(\log T_{ij} - 1) \qquad (4.15)$$

subject to

$$\sum_{ij} T_{ij} p_{ij}^{\ell m} - V_{\ell m} = 0 \quad \text{for some } \ell m \text{ (counted links)} \qquad (4.25)$$

and $T_{ij} \geq 0$.

The formal solution is obtained by forming the Lagrangian

$$\mathcal{L} = \sum_{ij} T_{ij}(\log T_{ij} - 1) + \sum_{\ell m} \lambda_{\ell m}(V_{\ell m} - \sum_{ij} T_{ij} p_{ij}^{\ell m})$$

and differentiating it with respect to T_{ij} :

$$\frac{\partial \mathcal{L}}{\partial T_{ij}} = -\log T_{ij} - \sum_{\ell m} \lambda_{\ell m}\, p_{ij}^{\ell m} = 0 \quad .$$

Thus

$$T_{ij} = e^{-\sum_{\ell m} \lambda_{\ell m} p_{ij}^{\ell m}}$$

By a change of variable

$$e^{-\lambda_{\ell m}} = X_{\ell m} \tag{4.26}$$

one obtains

$$T_{ij} = \prod_{\ell m} X_{\ell m}^{p_{ij}^{\ell m}} \tag{4.27}$$

Equations (4.25) and (4.27) belong to a class of models termed by Murchland (1978) multi-proportional problems and discussed in detail in Section 5.2. In outline, these problems can be tackled iteratively by taking each link flow equation (4.25) at a time and modifying the O-D matrix so that the modelled flows reproduce the observed ones until some convergence criterion is satisfied. The new variables $X_{\ell m}$ play a similar role to the balacing factors A_i and B_j . They may be considered balancing factors associated with the link counts instead of the usual trip end counts. In the gravity model each O-D pair appears in only one row and one column sum. Here, it may appear in several links depending on the characteristics of the network and the traffic counts available. Each O-D pair will then be adjusted proportionally once for every count to which it makes a contribution; hence the multi-proportional nature of the problem.

4.2.2 The model with prior information

This model can also be extended to make use of prior information about the trip matrix along the lines put forward by Batty and March (1976b) and Nathanson (1978). Only the objective function (4.15) changes by using the measure S_4 instead of S_1 as discussed in Section 4.1.2. The problem then becomes

$$\text{Max } S_4 = - \sum_{i} T_{ij}(\log T_{ij}/t_{ij} - 1) \qquad (4.28)$$

subject to the same constraints (4.25). The formal solution is again obtained by forming a Lagrangian and differentiating it with respect to T_{ij}. The equivalent change of variable (4.26) leads to the new form

$$T_{ij} = t_{ij} \prod_{\ell m} X_{\ell m}^{p_{ij}^{\ell m}} \qquad (4.29)$$

which is again a version of the multi-proportional problem.

When no prior information is available a reasonable assumption is to make all t_{ij} equal, for example to 1, in which case the model (4.29) reverts to (4.27)*.

In this context the prior matrix $[t_{ij}]$ can be obtained from an old trip matrix, a small sample survey, or by cordoning a matrix from a larger transport study. In such cases entropy plays the role of a measure of 'distance' between $[t_{ij}]$ and $[T_{ij}]$; thus the model (4.29) can be interpreted as the O-D matrix consistent with the observed counts and lying closest to the prior trip matrix.

As the models in (4.27) and (4.29) were developed from entropy maximising considerations it has been found

---

* Incidentally, the application of the same approach to the gravity model may also be of use. Maximising (4.28) subject to the usual constraints (4.12), (4.13) and (4.14) leads to

$$T_{ij} = t_{ij}A_iB_jO_iD_je^{-\beta c_{ij}}$$

The prior t_{ij} could be an efficient use of the sampled O-D matrix obtained through home interviews whose value is usually reduced to estimating β. This model also lends itself to interesting dynamic extensions.

convenient to call them the ME2 models (Maximum Entropy Matrix Estimation) with and without prior information.

In both models the variable $X_{\ell m}$ may have a similar interpretation due to its relationship to the Lagrange multiplier $\lambda_{\ell m}$. In the ME2 model without prior information (4.27) the variable $X_{\ell m}$ can be associated with the contribution of the observed count on link ℓm to the formation of the trip matrix. In the ME2 model with prior information $X_{\ell m}$ is associated with the contribution to the modification of the matrix (or to the information gained) from observing the particular volume. In both cases this contribution is weighted by the exponent $p_{ij}^{\ell m}$ representing the proportion of trips of each O-D pair which uses that link.

4.2.3 Van Zuylen's model

Quite independently from this author and almost at the same time, Henk Van Zuylen of the Verkeersakademie, Tilburgh, Holland, had been working on an approach similar to the one used in the entropy maximising model ME2. Van Zuylen (1978a, b) uses Brillouin's information measure, one of the variants discussed by Walsh and Webber (1977). The information I_B contained in a set of N observations where the state k has been observed n_k times is defined by Brillouin (1956) as

$$I_B = \log N! \prod_k \left[q_k^{n_k} / n_k! \right]$$

where q_k is the prior probability of observing state k. If the observations are counts on a particular link ℓm, it is possible to define state ij as the state in which the vehicle observed has been travelling between origin i and destination j.

So

$$n_{ij}^{\ell m} = T_{ij}\, p_{ij}^{\ell m} \quad .$$

It is possible to express the 'a priori' probability of observing state ij on link ℓm as a function of 'a priori' information about the O-D matrix as

$$q_{ij}^{\ell m} = \frac{t_{ij}p_{ij}^{\ell m}}{\sum_{ij} t_{ij}p_{ij}^{\ell m}}$$

where t_{ij} is the 'a priori' number of trips between i and j provided, for example, by an old O-D matrix. The information contained in $V_{\ell m}$ counts on a link is then

$$I_B^{\ell m} = -\log V_{\ell m}! \prod_{ij} \left\{\frac{t_{ij}p_{ij}^{\ell m}}{Z^{\ell m}}\right\}^{T_{ij}p_{ij}^{\ell m}} \Big/ \prod_{ij} (T_{ij}p_{ij}^{\ell m})!$$

where

$$Z^{\ell m} = \sum_{ij} t_{ij}p_{ij}^{\ell m} \quad .$$

Using Stirling's approximation, $\log X! \simeq X\log X - X$, it is possible to obtain

$$I_B^{\ell m} \simeq \sum_{ij} T_{ij}p_{ij}^{\ell m} \log\frac{T_{ij}Z^{\ell m}}{V_{\ell m}t_{ij}} \quad .$$

Summing over all the links in the network with counts:

$$S_6 = \sum_{\ell m} \sum_{ij} T_{ij}p_{ij}^{\ell m} \log\frac{T_{ij}Z^{\ell m}}{V_{\ell m}t_{ij}} \tag{4.30}$$

is the total information contained in the observed flows. The problem of finding an O-D matrix consistent with the observations and adding a minimum of extra information to them is equivalent to minimising S_6 , subject to the flow constraints (4.25). This is why Van Zuylen calls his model an information minimising approach. The formal solution to this problem can be obtained by differentiation of the Langrangian:

$$\mathcal{L} = \sum_{\ell m} \sum_{ij} T_{ij}p_{ij}^{\ell m} \log\frac{T_{ij}Z^{\ell m}}{V_{\ell m}t_{ij}} + \sum_{\ell m} \lambda_{\ell m}(\sum_{ij} T_{ij}p_{ij}^{\ell m} - V_{\ell m}) \tag{4.31}$$

where again $\lambda_{\ell m}$ are the Lagrange multipliers corresponding to the counted links. By taking the partial derivate

$\frac{\partial \mathcal{L}}{\partial T_{ij}}$ and equating it to zero one obtains

$$\left(\frac{T_{ij}}{t_{ij}}\right)^{\Sigma_{\ell m} p_{ij}^{\ell m}} = \underset{\ell m}{\pi} \left(\frac{V_{\ell m}}{Z_{\ell m}}\right)^{p_{ij}^{\ell m}} \cdot e^{-\Sigma p_{ij}^{\ell m}(1+\lambda_{\ell m})} \qquad (4.32)$$

and by making $g_{ij} = \sum_{\ell m} p_{ij}^{\ell m}$

and $$X_{\ell m} = \frac{V_{\ell m}}{Z_{\ell m}} e^{-(1+\lambda_{\ell m})}$$

the expression (4.32) is transformed into

$$T_{ij} = t_{ij} \underset{\ell m}{\pi} X_{\ell m}^{p_{ij}^{\ell m}/g_{ij}} \qquad (4.33)$$

This model is very similar in form to the one developed by this writer and has again the form of a multi-proportional problem. Some of its properties are further discussed in Section 4.4.

4.3 AN EXAMPLE OF THE USE OF THE ME2 MODEL

The example presented in this section has been devised to illustrate the nature of the solutions obtained by means of the ME2 model. The use of a simple system allows one to identify <u>all</u> the trip matrices that satisfy a set of link count constraints. A simulation program is then used to generate trip matrices satisfying these constraints and the number of times each of these matrices is generated is recorded. The ME2 solution should also be the most often generated trip matrix.

4.3.1 The simple system

Consider the network depicted in Figure 4.1 (see also Section 2.3). Let us assume that counts are only available for three links (1,5), (5,6) and (6,4) and that these are 100, 150 and 75 trips respectively. There are three independent counts and there is no prior information about the matrix. This problem with four unknowns and three independent equations has then one 'degree of freedom'.

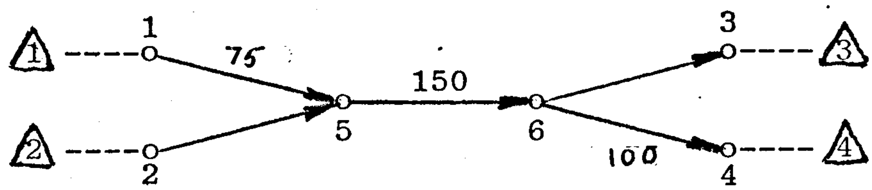


Figure 4.1: <u>A simple transport network</u>

The entropy maximising model (4.27) $T_{ij} = \underset{\ell m}{\pi} X_{\ell m}^{p_{ij}^{\ell m}}$ was applied using the counted link equations:

$$\left.\begin{aligned} T_{13} + T_{14} - 100 &= 0 \\ T_{13} + T_{14} + T_{23} + T_{24} - 150 &= 0 \\ T_{14} + T_{24} - 75 &= 0 \end{aligned}\right\} \quad (4.34)$$

These equations permit expressing all the variables in terms of one, say T_{13} , and so S becomes a function of one variable which may be easily maximised to give $T_{13} = 50$ and the rest 25, 50, 25. To facilitate comparison with larger problems the preferred algorithm for multi-proportional problems, to be described in Section 5.2, was used and the iterations were stopped when the modelled flows were within 2 per cent of the observed ones. The algorithm took 0.09 seconds of CPU time in the Leeds University Amdahl V7 machine and the results were

		Destination	
		3	4
Origin	1	50.5	24.6
	2	49.8	24.6

This is a good approximation to the integer solution 50, 50, 25, 25 which satisfies the constraints (4.34) exactly. Of course other O-D matrices may also satisfy these constraints but they would be 'less likely' as defined before.

4.3.2 A simulation program

In order to have a better feeling for the relative likelihood of the solution found compared with other feasible O-D matrices a simple simulation program was written to replicate the way in which trip matrices can be generated for this system. It can be seen from (4.34) that each matrix must have a total of 150 trips and these are generated sequentially by the simulation model. To generate each trip the program chooses one origin-destination pair at random, allocates one trip to it and increments the simulated flows on links used by that O-D pair by one. As soon as the simulated flow on one link reaches the level of the observed flow that link is said to be 'saturated' and all the O-D pairs using that link are 'closed' so that they cannot receive any new trips. The corresponding flow chart is depicted in Figure 4.2.

As this problem has one degree of freedom, fixing one cell value uniquely, determines the other three O-D pairs. With the assumed flow levels, the number of trips between 2 and 4 can take any value between 0 and 50. As the simulation model uses whole trips (integers), there are 51 distinct trip matrices which are consistent with the link counts.

Each trial results in a particular O-D matrix and the number of times each O-D matrix is generated is recorded in block (10) of the flow chart in Figure 4.2. After a predetermined number of trials these numbers are printed (block (12)).

The test for feasibility in block (9) is not required for this simple system but in larger networks it is possible for a matrix to have all its cells 'closed' and some links still 'unsaturated'. This would result in an unfeasible and unsuccessful trial. The problem has no solution if it is not possible to generate a matrix which satisfies blocks (8) and (9). This was described before as a problem generated by an inconsistency between traffic counts and the assumed assignment model.

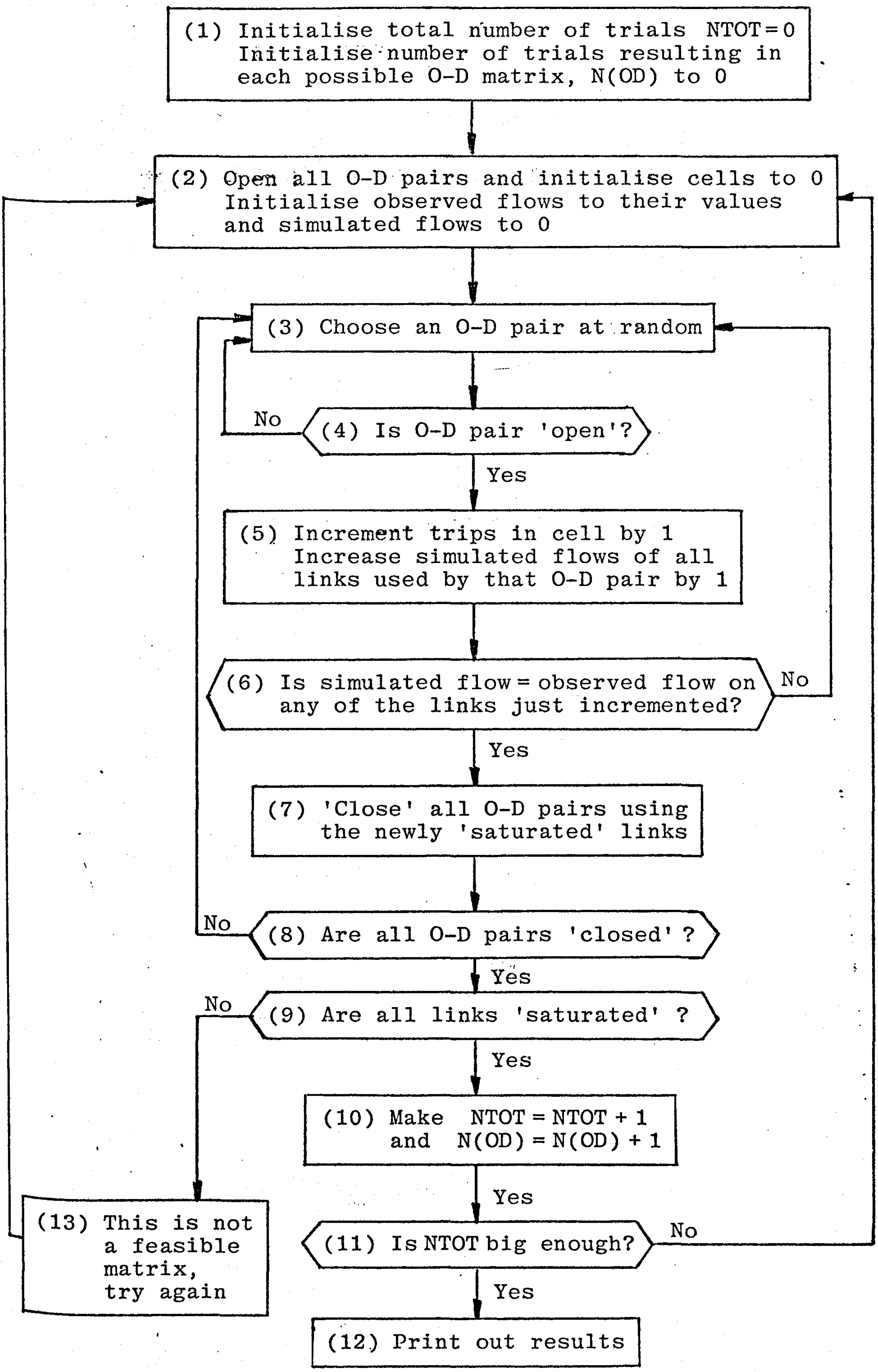


Figure 4.2: Flow-chart of simulation program for simple system

4.3.3 Results

In order to ascertain the number of trials needed to accurately identify the most likely trip matrix tests were made with 21, 51, 100 and 1000 trials and the results are depicted in Table 4.1 together with the entropy value of each matrix. The trip matrix with maximum entropy is underlined. It can be seen that about 100 trials are needed to identify this matrix but even then there is a risk of error as the frequency distribution has some discontinuities. The 1000 trials case has a much smoother frequency distribution. The cpu time is likely to be proportional to the number of trips times the number of links times the number of trials. The total number of trials will depend directly on the number of possible O-D matrices. Even for this simple system the simulation model takes much longer than the ME2 model solution reported in Section 4.3.1.

A result which perhaps confirms initial expectations is that in this case although all 51 matrices are possible 30 of them were not generated at all in 1000 trials. These matrices are therefore very unlikely matrices in the sense that they have a very low probability of being generated.

On the other hand, the most likely matrix occurs approximately once every eight trials (12.5 per cent) and two other matrices are generated in more than 10 per cent of the trials. Even for this simple system the optimum does not appear to be very 'sharp'. However, if one considers all the matrices whose cell values are within $\pm$ 2 trips from the optimum (5 matrices) these are found to be generated more than 50 per cent of the time (51.7 per cent).

These results are only illustrative and should not be directly generalised to larger networks and other degrees of freedom.

Table 4.1: Results from simulation program

Matrix 1-3	1-4	2-3	2-4	Entropy S_4	Trials	Trials	Trials	Trials
25	50	75	0	-449.88	0	0	0	0
26	49	74	1	-443.91	0	0	0	0
27	48	73	2	-439.40	0	0	0	0
28	47	72	3	-435.48	0	0	0	0
29	46	71	4	-431.97	0	0	0	0
30	45	70	5	-428.78	0	0	0	0
31	44	69	6	-425.86	0	0	0	0
32	43	68	7	-423.18	0	0	0	0
33	42	67	8	-420.17	0	0	0	0
34	41	66	9	-418.45	0	0	0	0
35	40	65	10	-416.35	0	0	0	0
36	39	64	11	-414.43	0	0	0	0
37	38	63	12	-412.67	0	0	0	0
38	37	62	13	-411.06	0	0	0	0
39	36	61	14	-409.60	0	0	0	0
40	35	60	15	-408.27	0	0	0	2
41	34	59	16	-407.09	0	0	0	2
42	33	58	17	-406.04	0	3	4	15
43	32	57	18	-405.12	0	0	1	19
44	31	56	19	-404.32	0	1	2	36
45	30	55	20	-403.65	4	6	8	56
46	29	54	21	-403.11	0	3	5	69
47	28	53	22	-402.69	3	5	10	98
48	27	52	23	-402.39	2	4	8	107
49	26	51	24	-402.21	2	6	12	105
50	25	50	25	-402.15	2	6	15	125
51	24	49	26	-402.21	1	5	9	95
52	23	48	27	-402.39	1	3	6	85
53	22	47	28	-402.69	2	3	7	62
54	21	46	29	-403.11	2	3	6	47
55	20	45	30	-403.65	1	1	2	31
56	19	44	31	-404.32	0	0	2	22
57	18	43	32	-405.12	0	0	1	11
58	17	42	33	-406.04	0	0	0	8
59	16	41	34	-407.09	1	2	2	4
60	15	40	35	-408.27	0	0	0	1
61	14	39	36	-409.60	0	0	0	0
62	13	38	37	-411.06	0	0	0	0
63	12	37	38	-412.67	0	0	0	0
64	11	36	39	-414.43	0	0	0	0
65	10	35	40	-416.35	0	0	0	0
66	9	34	41	-418.45	0	0	0	0
67	8	33	42	-420.17	0	0	0	0
68	7	32	43	-423.18	0	0	0	0
69	6	31	44	-425.86	0	0	0	0
70	5	30	45	-428.78	0	0	0	0
71	4	29	46	-431.97	0	0	0	0
72	3	28	47	-435.48	0	0	0	0
73	2	27	48	-439.40	0	0	0	0
74	1	26	49	-443.91	0	0	0	0
75	0	25	50	-449.88	0	0	0	0
CPU TIME FOR SIMULATION (SEC)					0.02	0.05	0.13	1.38
FOR A TOTAL NUMBER OF TRIALS					21	51	100	1000

NOTE: For 1000 trials σ^2 is approximately 100 implying an error of about 2 per cent and hence the optimum would be generated 12.5 per cent ± 2 per cent of the time. This is not enough precision to properly distinguish an optimum and more computer time would have been required to obtain the 16,000 trials to reduce errors to 0.5 per cent.

4.4 SOME PROPERTIES OF THE MODELS

4.4.1 Differences between Willumsen's and Van Zuylen's models

Although the basic assumptions are similar the two models have some interesting formal differences as was recently explained in a joint paper (Van Zuylen and Willumsen, 1980).

The main difference between the models in Equations (4.29) and (4.33) lies in the exponents, $p_{ij}^{\ell m}/\sum_{\ell m} p_{ij}^{\ell m}$ for Van Zuylen's model and simply $p_{ij}^{\ell m}$ in Willumsen's. These have the form of 'weights' to be associated to observations on link ℓm. The similarity is not surprising as the close links between entropy maximising and information minimising have long been recognised. In fact, it is possible to show that the basic difference between the models resides in what is considered to be the unit of observation or the relevant meso-state: Van Zuylen's model uses a counted vehicle and Willumsen's a trip.

One way of showing this is to derive Van Zuylen's model by means of maximum entropy considerations.

Each trip from i to j is counted $g_{ij} = \sum_{\ell m} p_{ij}^{\ell m}$ times on average. Then in general

$$\tilde{T} = \sum_{ij} g_{ij} T_{ij} \quad \text{and} \quad \tilde{t} = \sum_{ij} g_{ij} t_{ij}$$

vehicle counts are assigned to origin-destination pairs. Following Wilson (1970) again the number of possible ways of doing this may be used to determine the most likely trip matrix. The 'a priori' probability that a counted vehicle actually moved from i to j is

$$q_{ij} = g_{ij} / \sum_{ij} g_{ij} t_{ij} = g_{ij} / \tilde{t} \quad .$$

The problem is then to maximise

$$\log\left[\frac{\tilde{T}!}{\tilde{t}} \prod_{ij} \frac{(g_{ij}t_{ij})^{g_{ij}T_{ij}}}{(g_{ij}T_{ij})!}\right]$$

or the resulting new entropy measure

$$S_7 = \log\frac{\tilde{T}!}{\tilde{t}} - \sum_{ij} g_{ij}T_{ij} \log(T_{ij}/t_{ij} - 1) \quad . \qquad (4.35)$$

A further maximisation of S_7 subject to (4.25) using Lagrangian multipliers gives again Equation (4.33).

It can be argued whether the assignment of traffic counts to origin-destination pairs or the determination of the most likely O-D trip matrix consistent with the observed flows best represents the real problem. Both S_4 and S_7 give a measure of the difference between a trip matrix $[T_{ij}]$ and an 'a priori' matrix $[t_{ij}]$. As Murchland (1978) suggests, there is little or no theoretical argument to prefer any of such measures. Practical issues like computational problems and ease of use are likely to be more important.

Both models result in a multi-proportional problem whose treatment will be discussed in Section 5.2. A comparison of the performances of both models with artificial data is reported in Section 6.3. Most of the properties of the ME2 model discussed in the rest of this work are also applicable to Van Zuylen's model.

4.4.2 Data requirements

The first data requirement for the ME2 model is, as for any other model for this purpose, information on the structure of the network and on link costs. These data combined with a reasonably good route choice model should enable one to estimate a satisfactory set of $[p_{ij}^{\ell m}]$.

The second requirement is, of course, traffic counts and here the model is particularly attractive. The model does not require a full set of counts to estimate an O-D matrix and it can do so without having to impose a particular travel demand model on the trip matrix. The ME2 model reduces the number of unknowns to one per counted link thus making the problem fully specified. This reduction of variables implies that the model would give no information on origin-destination pairs which do not use any of the observed links. These cells will not be considered in any solution method and they will retain their 'a priori' values.

As mentioned in Section 2.3 the observed volumes have to be consistent. The first consistency, corresponding to link flow continuity conditions, can be ensured by methods discussed in Section 5.1 The second consistency, corresponding to path flow continuity conditions, will depend basically on the match between assumed and real route choices. It does not seem right to modify counts to become consistent with an imperfect traffic assignment model. A better approach is to use the best route choice model for the conditions prevailing in the area of interest. These issues are discussed in greater detail in Chapters 8 to 10.

If the flows are consistent in these two senses the ME2 model without prior information will always converge to a solution which will reproduce the observed link volumes.

4.4.3 Use of prior information

The convergence of the entropy maximising model with prior information will depend on the characteristics of the matrix $[t_{ij}]$, which will in term depend on the way in which it was obtained.

As mentioned before there are several ways of obtaining a prior trip matrix $[t_{ij}]$. One is to use an old trip matrix so that the problem becomes one of updating a trip table. A second source could be a small sample survey. Alternatively, a prior matrix could be cordoned from a larger study or even obtained using a simple demand model, perhaps a gravity model. Any observed trip matrix, and in particular if estimated from small sampling fractions, is likely to contain a large number of zeroes. These may be 'true' zeroes if they represent impossible or highly unlikely journeys, but are much more likely to be zeroes 'by chance' due to the limitations in the sampling framework. Even for quite large samples a cell with no observations cannot be said to be significantly different from a cell with one or two observed trips. However, from Equation (4.29)

$$T_{ij} = t_{ij} \prod_{\ell m} X_{\ell m}^{p_{ij}^{\ell m}}$$

it is clear that if the observed t_{ij} is zero so must be the updated or expanded T_{ij} . This is the source of a third type of inconsistency, one between the prior trip matrix and the observed link flows.

In effect, it is possible to observe a positive flow on one link but to have all the O-D pairs using that link ($p_{ij}^{\ell m} > 0$) assigned a prior value zero ($t_{ij} = 0$). In this case the modelled flows will always be zero and the equations will have no feasible solution. This problem does not occur in the ME2 model without prior information as all t_{ij} are implicitly made equal to one.

A pragmatic solution to this problem is to 'seed' the zero cells in the prior trip matrix $[t_{ij}]$ with suitable small values (say 0.5) so that all potential trip interchanges will be possible in the updated matrix. Under this arrangement the 'seeded' cells will be modified by the link balancing factors $X_{\ell m}$, some of them 'growing' to become full trips in the matrix and others returning to zero depending on the counts affecting them.

This procedure should produce a more realistic and less distorted trip matrix. The use of a method like this is likely to be a better way of expanding survey data than simply applying grossing up factors depending on the sampling rates only.

The application of this technique to update an old trip matrix in Harrogate is described in Sections 9.5 and 11.2.

4.4.4 Use of other information

Due to the structure of the ME2 model in Equation (4.29) it is possible to introduce other types of independent information by making use of 'pseudo links'. Two of these possibilities are described below.

(i) Information regarding total number of trips in an area. This requires the creation of a 'pseudo link' $\ell m'$ with all $p_{ij}^{\ell m'} = 1$. The 'observed volume' of this link is then made equal to the total number of trips. Similar arrangements are possible for information about the total number of trips from two or more particular zones or between pairs of zones.

(ii) Trip length distribution, perhaps obtained through a small scale survey. If the total number of trips is also known it is possible to create 'pseudo links', each one corresponding to a particular trip length range. The $p_{ij}^{\ell m}$ values corresponding to these links will be one for all O-D pairs whose trip length falls within the range represented by the link and zero for the rest. There is of course the risk of the trip length distribution being inconsistent with the observed link flows.

An approximate solution which does not carry this risk is to use the prior trip matrix $[t_{ij}]$ to introduce in the estimation problem a particular trip length distribution. The solution found in this way can be said to be the O-D matrix which while satisfying the link counts lies closest to the surveyed one.

Other pieces of information, even of the form "there are 50 per cent more trips between i and k than between j and m " can be introduced through the prior matrix.

These lines for using additional information were not pursued in this work which concentrated on the problem where network data is the only information available.

CHAPTER 5
SOLUTION METHODS

An entropy maximising model (ME2) has been put forward to estimate an origin-destination matrix from traffic counts. The model has been presented in the form of a mathematical program:

$$\text{Maximise } S = -\sum_{ij} T_{ij}(\log T_{ij}/t_{ij} - 1) \qquad (5.1)$$

subject to

$$\sum_{ij} T_{ij}p_{ij}^{\ell m} - V_{\ell m} = o \qquad \text{for } L_c \text{ counted links } \ell m \; (L_c \leq L) \qquad (5.2)$$

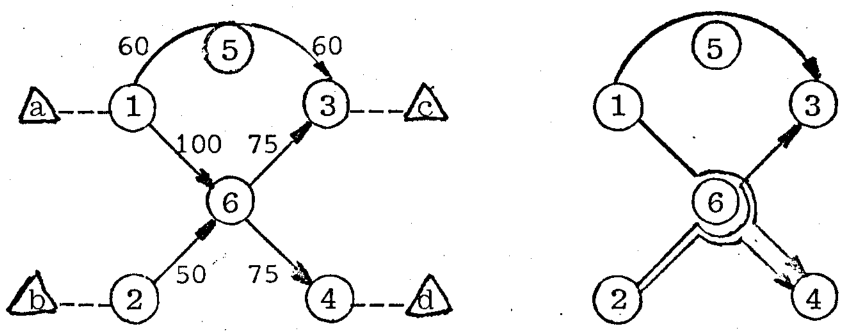
and $T_{ij} \geq 0$ (5.3)

This chapter addresses the main issues related to the solution of this mathematical program. First, how can one remove at least some of the possible inconsistencies in the set of constraints (5.2). Second, what algorithms can be used to solve this matrix estimation problem and what are their relative advantages? Third, if resources are not available to obtain traffic counts in all the links in a network, how can one choose which links to count first?

5.1 TREATMENT OF INCONSISTENT TRAFFIC COUNTS

5.1.1 The sources of inconsistency in traffic counts

Two sources of inconsistency in traffic counts were identified in Section 2.3. The first one is simply the fact that errors in the counts may lead to situations in which the 'total flow into' a node does not equal the 'total flow out of' the same node, thus not meeting link flow continuity conditions. The second source was due to a mismatch between the assumed traffic assignment model in Equation (5.2) and the observed flows. This type of inconsistency occurs whenever path flow continuity is not met. Consistency at the link flow level is a necessary but not sufficient condition for consistency at path flow level. Consistency at path flow level is, however, a sufficient condition for link flow consistency.



(a) Network and flows (b) Assumed route choices

Figure 5.1: An example of path flow inconsistencies with counts

The example in Figure 5.1 may help to illustrate the above mentioned problems.

It can be seen that the link flow continuity conditions are met. (They would not have been met if the flow on link (2,6), for example, had been 55 instead of 50.) However, the assumed assignment depicted in Figure 5.1b is incompatible with the flows shown in Figure 5.1a. No feasible trip matrix can reproduce the count of 75 at link (6,3) because the only path using it, b-c, is limited to a maximum of 50 by link (2,6).

The set of linear equations corresponding to this example is given by

link (1,5)	$T_{ac} \cdot 1$	$= 60$	(5.4)
link (5,3)	$T_{ac} \cdot 1$	$= 60$	(5.5)
link (1,6)	$T_{ad} \cdot 1$	$= 100$	(5.6)
link (2,6)	$T_{bc} \cdot 1 + T_{bd} \cdot 1$	$= 50$	(5.7)
link (6,3)	$T_{bc} \cdot 1$	$= 75$	(5.8)
link (6,4)	$T_{ad} \cdot 1 + T_{bd} \cdot 1$	$= 75$	(5.9)

Clearly Equations (5.7) and (5.8) are incompatible with the condition of non-negativity as are Equations (5.6 and (5.9), making it impossible to solve this set of equations. In simple problems like these inconsistencies can be ascertained by inspection but in more complex networks they can only be identified by means of row and column operations on the constraints. For large systems these operations are likely to be expensive in terms of computer requirements.

In this simplistic example it is not difficult to see that the problem originates in the assumed single route between a and c . If two paths were allowed, one via node 5 and the other via node 6, the inconsistency could be removed. Furthermore, the value of the resulting variable p_{ac}^{63} cannot be arbitrarily chosen; in effect, a feasible solution requires

$$\frac{25}{(25+60)} = 0.294 \leq p_{ac}^{63} \leq 0.556 = \frac{75}{(75+60)} \quad .$$

The fact that the <u>path flow</u> continuity conditions are not met seems to reflect errors in assignment whereas the <u>link flow</u> discontinuities are a reflection solely of errors in the traffic counts. It seems reasonable then to develop a technique for removing the <u>link flow</u> inconsistencies in the counts to ensure link flow continuity conditions are met. On the other hand, a reasonable approach to deal with the lack of consistency at the <u>path flow</u> level seems to be the adoption of a better route choice model.

5.1.2 Identification of flow continuity conditions

In order to remove inconsistencies at the link flow level only the following set of linear equations must be considered.

$$\sum_{\ell} V_{\ell k} - \sum_{m} V_{km} = 0 \quad \text{for all nodes } k \quad . \tag{5.10}$$

But not all these equations are a source of inconsistency at this level and there are two reasons for this. Firstly, where a centroid connector is attached to a node it is always possible to remove any inconsistency by assigning an appropriate flow to the centroid connector itself. Accordingly, equations based on these nodes need not be considered. Secondly, it is likely that some links will not have counts, thus reducing the opportunities for inconsistencies to arise, as the non-counted link flow can take any (non-negative) value. But in this case, equations corresponding to nodes with non-counted links should not be automatically dropped as a linear combination for some of them may constitute a proper flow continuity relationship. This possibility is illustrated by the network in Figure 5.2 where counted flows on link ℓm are represented by $\hat{V}_{\ell m}$ and non-counted ones by $V_{\ell m}$.

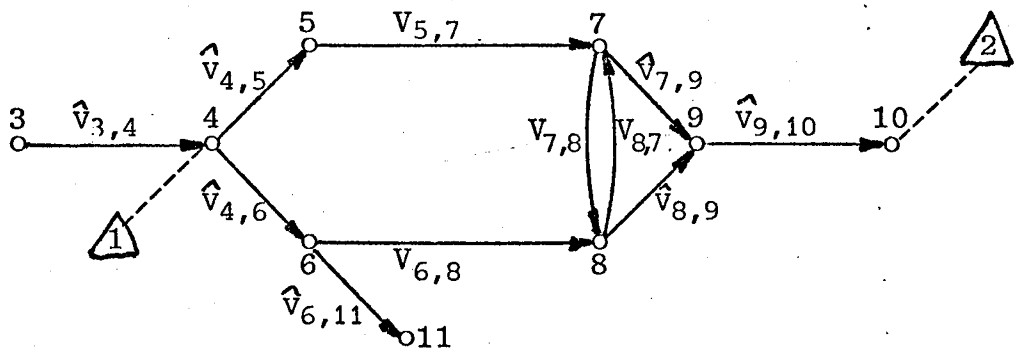


Figure 5.2: A network segment

It can be seen that node 9 generates the only direct flow continuity condition:

$$\hat{V}_{7,9} + \hat{V}_{8,9} - \hat{V}_{9,10} = 0 \qquad (5.11)$$

where the convention adopted is that 'flows into' the node have a positive sign and 'flows out of' a negative one. Nodes 4 and 10 do not generate suitable conditions as the flows on the centroid connectors can take any values. Consider now the equations corresponding to nodes with non-counted links. Using the same convention these can be written as

node 5	$\hat{V}_{4,5} - V_{5,7}$	$= 0$	(5.12a)
node 6	$\hat{V}_{4,6} - V_{6,8} - \hat{V}_{6,11}$	$= 0$	(5.12b)
node 7	$V_{5,7} + V_{8,7} - V_{7,8} - \hat{V}_{7,9}$	$= 0$	(5.12c)
node 8	$V_{6,8} + V_{7,8} - V_{8,7} - \hat{V}_{8,9}$	$= 0$	(5.12d)

It can be seen that each non-counted link appears twice, once with each sign. It is a simple matter to perform additions (linear combinations) of these equations to generate a single equation including only counted links. The sum of (5.12a) + (5.12b) + (5.12c) + (5.12d) results in a flow continuity condition over observed volumes only

$$\hat{V}_{4,5} + \hat{V}_{4,6} - \hat{V}_{7,9} - \hat{V}_{8,9} - \hat{V}_{6,11} = 0$$

It is interesting to note that had a centroid connector been attached to, say, node 8 thus removing Equation (5.12d), this combination would have not been possible.

These considerations suggest the following procedure to identify the set of relevant flow continuity conditions in a full network.

(a) Take each node in turn and classify it in one of three groups as follows

- nodes with centroid connectors attached to them
- nodes with counts on all the attached links (and no centroid connectors)
- nodes with either some or all non-counted attached links (and no centroid connectors).

(b) Set up flow continuity conditions for the second group.

(c) Set up equations of the form (5.10) for each of the nodes in the third group by replacing non-counted flows by unknowns $V_{\ell m}$ and seek to form proper flow continuity conditions by eliminating the unknowns $V_{\ell m}$ by linear combination of these equations.

Due to the nature of the problem, (c) can be achieved by the following steps.

(i) Set up a list of equations obtained from (c)

(ii) Take the next (first) equation from the list and consider it as a candidate for a flow continuity condition.

(iii) Eliminate each of the unknowns $V_{\ell m}$ by adding the complementary equation containing the same variable (with opposite sign). Remove this complementary equation from the list. If this complementary equation does not exist remove the current equation and proceed to (ii).

(iv) Eliminate any new (added) unknowns by the same process and repeat until no unknowns remain.

(v) If the set of un-used equations has not been exhausted proceed to (ii), otherwise add the extra set of flow continuity conditions to those obtained in (b) above.

The resulting set of flow continuity conditions is now better written as

$$\sum_{\ell m} \hat{V}_{\ell m} \delta^{k}_{\ell m} = 0 \qquad (5.13)$$

where $\delta^{k}_{\ell m}$ is an indicator of the sign associated with volume $\hat{V}_{\ell m}$ in equation k and is defined as

$$\delta^{k}_{\ell m} = \begin{cases} +1 & \text{if flow } \hat{V}_{\ell m} \text{ appears with positive sign} \\ 0 & \text{if flow } \hat{V}_{\ell m} \text{ does not enter in equation } k \\ -1 & \text{if flow } \hat{V}_{\ell m} \text{ appears with negative sign} \end{cases}$$

Note that some k correspond to node equations and others to linear combinations. In any case, it can be seen that for each counted link ℓm there will be at most two non-zero indicators $\delta^{k}_{\ell m}$ and that these will have opposite signs. This is obvious for equations based on real node as a volume can only be 'out of' one node and 'into' another. The method used to obtain the extra continuity conditions through the addition of node equations with some non-counted links ensures that this will also be the case for them.

Once the set of flow continuity conditions has been identified the remaining problem is to eliminate any inconsistencies existing in the observed flows.

5.1.3 A maximum likelihood solution

It is convenient to accept the existence of an explicit error term for each observed volume $\hat{V}_{\ell m}$

$$V_{\ell m} = \hat{V}_{\ell m} + \varepsilon_{\ell m} \quad \text{all counted links } \ell m \qquad (5.14)$$

where $V_{\ell m}$ is the 'true' volume
$\hat{V}_{\ell m}$ is the observed flow
and $\varepsilon_{\ell m}$ is a (small) error term.

It is then possible to replace the variables $V_{\ell m}$ for their expressions in (5.14) and solve the set (5.13) so that some measure of the total error is minimised.

A frequent choice is to minimise the sum of the square of the errors

$$Z = \sum_{\ell m} (V_{\ell m} - \hat{V}_{\ell m})^2 = \sum_{\ell m} \varepsilon^2_{\ell m} \tag{5.15}$$

Software is available for solving this problem, for example the NAG routine F04AUF (Numerical Algorithms Group, 1978). This method may yield negative corrected flows but is unlikely so for links with relatively large volumes.

A perhaps simpler approach can be adopted following ideas put forward by Hamerslag and Huisman (1978). This approach is based on maximum likelihood and results in a set of equations of simple resolution.

Without much loss of generality it is possible to assume that the observed flows on link ℓm are Poisson distributed with mean $V_{\ell m}$. In this case, the probability of observing $\hat{V}_{\ell m}$ vehicles on link ℓm is given by

$$P(\hat{V}_{\ell m}/V_{\ell m}) = \frac{e^{-V_{\ell m}} \, V_{\ell m}^{\hat{V}_{\ell m}}}{\hat{V}_{\ell m}!}$$

Then the likelihood (Hogg and Craig, 1970) of observing a set of flows $[\hat{V}_{\ell m}]$ is given by the function

$$L = \prod_{\ell m} e^{-V_{\ell m}} \, V_{\ell m}^{\hat{V}_{\ell m}} / \hat{V}_{\ell m}!$$

The best estimates of $[V_{\ell m}]$ will be those maximising this likelihood function subject to the constraints (5.13). For practical reasons the equivalent maximisation of the monotonic function

$$\log L = \sum_{\ell m} (-V_{\ell m} + \hat{V}_{\ell m} \log V_{\ell m} - \log \hat{V}_{\ell m}!)$$

is preferred.

A Langrangian can be formed by adding the constraints (5.16) and the Lagrange multipliers λ_k.

$$\mathcal{L} = \sum_{\ell m} (-V_{\ell m} + \hat{V}_{\ell m} \log V_{\ell m} - \log \hat{V}_{\ell m}!) + \sum_k \lambda_k \cdot \sum_{\ell m} \delta^k_{\ell m} V_{\ell m} .$$

Taking the partial derivatives and equating them to zero

$$\frac{\partial \mathcal{L}}{\partial V_{\ell m}} = -1 + \frac{\hat{V}_{\ell m}}{V_{\ell m}} + \sum_k \lambda_k \delta^k_{\ell m} = 0$$

which can be rearranged as

$$\frac{\hat{V}_{\ell m}}{V_{\ell m}} = 1 - \sum_k \lambda_k \delta^k_{\ell m}$$

and hence leads to

$$V_{\ell m} = \frac{\hat{V}_{\ell m}}{1 + \sum_k \lambda_k \delta^k_{\ell m}} . \tag{5.16}$$

Where $\sum_k \lambda_k \delta^k_{\ell m}$ has no more than two non-zero elements k and they cannot have the same sign.

Solving the K equations (5.16), where K is at most equal to the number of counted links, requires finding the values of $[\lambda]$ by substituting the Equations (5.16) into the Equations (5.13). However, this leads to a non-linear problem which is somewhat difficult to solve.

It can be seen from (5.16) that the Lagrange multipliers λ_k play the role of correction factors for the observed $\hat{V}_{\ell m}$. These corrections can be expected to be small, close to zero, as $\hat{V}_{\ell m}/\bar{V}_{\ell m}$ should not be too different from one. It is possible to take advantage of the fact and the approximation $(1+x)^{-1} \approx 1-x$ for x close to zero, so that Equations (5.16) become

$$V_{\ell m} = \hat{V}_{\ell m}(1 - \sum_k \lambda_k \delta^k_{\ell m}) . \tag{5.17}$$

Substituting into (5.13) leads to K equations

$$\sum_{\ell m} \delta^k_{\ell m} \hat{V}_{\ell m}(1 - \sum_k \lambda_k \delta^k_{\ell m}) = 0 \tag{5.18}$$

This is now a set of K linear equations with K unknowns λ_k, where K is the number of flow continuity relationships obtained in 5.1.2.

For normal networks this number of unknowns is not too large and a standard method can be used for its resolution.

The proposed method for obtaining consistent counts may be summarised as follows.

(i) Identify a set of flow continuity conditions as described in Section 5.1.2, resulting in a set of Equations (5.13).

(ii) Set up a series of K linear Equations (5.18)
$$\sum_{\ell m} \delta^k_{\ell m} \hat{V}_{\ell m}(1 - \sum_k \lambda_k \delta^k_{\ell m}) = 0$$
and solve it for values of λ_k.

(iii) Modify the observed links by making
$$V_{\ell m} = \hat{V}_{\ell m}(1 - \sum_k \lambda_k \delta^k_{\ell m}) \qquad (5.17)$$

While the author has not been able to prove yet that the set of Equations (5.18) can always be solved, a number of practical examples have failed to produce a case where these equations were inconsistent and impossible to solve by this method. Branston (forthcoming) has shown how this technique can be extended to consider links with more than one count available.

5.1.4 An example

A computer program, METWO1, has been written by the author to perform the task described in (i) to (iii) above. The program reads a description of the network and a set of observed counts and produces as output a new set of (link flow) consistent counts. Of course if the set of counts is already consistent on input their consistency is checked and the volumes remain unmodified.

The application of this technique to the set of inconsistent counts in Figure 5.3 results in the set of equations

$$100(1-\lambda_5) + 110(1-\lambda_5) - 200(1+\lambda_5-\lambda_6) = 0$$

and $$200(1+\lambda_5-\lambda_6) - 50(1+\lambda_6) - 140(1+\lambda_6) = 0$$

the solution of which is

$$\lambda_5 = 0.0492$$

and $$\lambda_6 = 0.0509$$

leading to

$$\left.\begin{aligned}\hat{V}_{15} &= 95.08\\ \hat{V}_{25} &= 104.59\end{aligned}\right\} 199.67$$

$$\hat{V}_{56} = 199.66$$

$$\left.\begin{aligned}\hat{V}_{63} &= 52.55\\ \hat{V}_{64} &= 147.13\end{aligned}\right\} 199.68$$

which is a fairly good approximation to perfect consistency at link flow level.

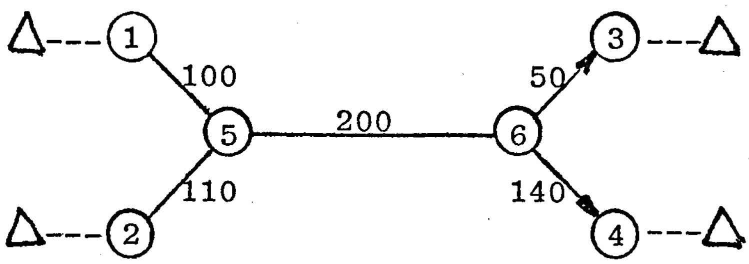


Figure 5.3: A set of inconsistent counts

5.2 SOLVING THE ENTROPY MAXIMISING PROBLEM

This section discusses two alternative methods for solving the entropy maximising problem presented in Chapter 4.

We begin by presenting the original problem in the framework of Rockafellar's linear, convexly separable, mathematical programming theory to facilitate a discussion of solution algorithms in the following sub-sections.

5.2.1 The mathematical programmes

The problem of estimating a trip matrix from traffic counts was presented in Section 4.2 as one of maximising

$$S = -\sum_{ij} T_{ij}(\log T_{ij}/t_{ij} - 1) \qquad (5.1)$$

subject to

$$\sum_{ij} T_{ij} p_{ij}^{\ell m} - \hat{V}_{\ell m} = 0 \qquad \ell m \in L_c \qquad (5.2)$$

and $T_{ij} \geq 0$ all i,j (5.3)

The problem can also be rewritten

$$\text{Minimise } S' = \sum_{ij} T_{ij}(\log T_{ij}/t_{ij} - 1) \qquad (5.19)$$

subject to the same constraints. It was mentioned in 4.1.3 that S' is the sum of a number of strictly convex functions, each of one variable T_{ij} and so itself strictly convex. A strictly convex function defined over a closed bounded convex set has a unique minimum so that provided the set defined by (5.2) (5.3) is non-empty then S' will have a unique minimum. This unique minimum may either be on the boundary of the feasible region with at least one $T_{ij}=0$ or it may be an interior point with all T_{ij} strictly positive. If it is an interior point, then the argument in section 4.2 using Lagrange multipliers $\lambda_{\ell m}$ shows that the optimum $\overline{T_{ij}}$ must be of the form

$$T_{ij} = t_{ij} \prod_{\ell m} X_{\ell m}^{p_{ij}^{\ell m}} \qquad (5.20)$$

where $X_{\ell m} = e^{-\lambda_{\ell m}}$ (5.21)

That the T_{ij} have this form also follows from the properties of mathematical programmes of this type as discussed by Murchland (1977). Murchland uses Rockafellar's (1970) linear, convexly separable, programming theory to present the programme in the form of primal, dual and existence problems. This treatment facilitates the simultaneous discussions of the existence of a solution and the algorithms to find it. Murchland's analysis is now outlined using the specific problem in hand and its notation.

Three Rockafellar problems can be written as follows.

Primal problem

$$\text{Min} \sum_{ij} T_{ij} \log T_{ij}/t_{ij} - T_{ij} + t_{ij} \tag{5.22}$$

subject to

$$\sum_{ij} T_{ij} p_{ij}^{\ell m} - \hat{V}_{\ell m} = 0 \tag{5.2}$$

and

$$T_{ij} > 0 \tag{5.3a}$$

over variables T_{ij} with $p_{ij}^{\ell m}$, t_{ij} and $\hat{V}_{\ell m}$ given.

The (Rockafellar) primal problem corresponds closely to the minimising problem (5.19) subject to (5.2) and (5.3). The differences are that the T_{ij} are restricted in (5.3.a) to be strictly positive, that is $T_{ij}>0$, and that in (5.22), an additional term $\sum_{ij} t_{ij}$ is incorporated to the objective function only for convenience. The term $\sum_{ij} t_{ij}$ is equal to t, the total number of prior trips, a constant which could be dropped from the objective function without affecting the properties nor the solution of the problem.

Dual problem

$$\text{Min} \left[\sum_{ij} t_{ij}(e^{z_{ij}}-1) + \sum_{\ell m} \hat{V}_{\ell m} \lambda_{\ell m} \right] \tag{5.23}$$

subject to

$$z_{ij} = - \sum_{\ell m} \lambda_{\ell m} p_{ij}^{\ell m} \tag{5.24}$$

over variables $\lambda_{\ell m}$ and z_{ij} (which, unlike the previous problem, have no restrictions on their signs) with t_{ij} , $\hat{V}_{\ell m}$ and $p_{ij}^{\ell m}$ given.

<u>Existence problem</u>

Find a solution to

$$\sum_{ij} T_{ij} p_{ij}^{\ell m} - \hat{V}_{\ell m} = 0 \qquad (5.2)$$

$$z_{ij} = - \sum_{\ell m} \lambda_{\ell m} p_{ij}^{\ell m} \qquad (5.24)$$

$$T_{ij} = t_{ij}\, e^{z_{ij}} \qquad (5.25)$$

with variables T_{ij} , $\lambda_{\ell m}$ and z_{ij} .

Rockafellar's theory shows that a necessary and sufficient condition for the primal problem to have a unique solution is that the set of T_{ij} defined by the constraints (5.2) and (5.3a) is non-empty, i.e. that the constraints are not inconsistent. Moreover, the dual and existence problem have a solution if and only if the primal problem does, and the same set of T_{ij}, $\lambda_{\ell m}$ and z_{ij} satisfies all of them. Hence, we may solve any of the problems by finding solutions to any other.

Another result from Rockafellar's theory is that if the primal problem above has a solution then the trip matrix problem defined at the begining of this section by equations (5.1), (5.2) and (5.3) has the same solution. In other words, relaxing the constraints from T_{ij} strictly positive (5.3a) to T_{ij} greater or equal zero (5.3), does not lead to a different solution. However, it is possible for the trip matrix problem to have a solution where the primal problem above does not. This arises if there is no feasible solution to the constraints (5.2) and (5.3a) but there is to the equations (5.2) and (5.3), or in other words if the only feasible solution to (5.2) and (5.3) have some T_{ij} equal to zero. In this case the tripmatrix problem still has a solution but this solution has not necessarily the form

$$T_{ij} = t_{ij} \prod_{\ell m} X_{\ell m}^{p_{ij}^{\ell m}} \qquad (5.20)$$

and $$X_{\ell m} = e^{-\lambda_{\ell m}} \qquad (5.21)$$

In what follows we ignore the possibility of boundary solutions of the type mentioned above (some $T_{ij}=0$) and concentrate in the case where there is a feasible solution to (5.2) and (5.3) or (5.3a) with all T_{ij} strictly positive

Two algorithms can be devised almost directly from this treatment of the problem and these will be discussed below.

5.2.2 Newton-Raphson's method

One possible way of searching for a solution to the three mathematical programs is to attempt to solve the existence problem (5.2), (5.24) and (5.25) directly. This approach has been put forward (in another context) by Erlander (1977) and by Eriksson (1978). This method is now described in outline following Eriksson (1978) which contains a detailed analysis of the problem.

First substitute Equations (5.24) and (5.25) into the original link flow Equations (5.2) to obtain

$$\Lambda = \sum_{ij} t_{ij} p_{ij}^{\ell m} \, e^{-\sum_{\ell m} \lambda_{\ell m} p_{ij}^{\ell m}} - \hat{V}_{\ell m} = 0 \tag{5.26}$$

for L_c (counted) links ℓm

These correspond to a set of non-linear equations in $\lambda_{\ell m}$ which can be tackled by Newton-Raphson's iterative method. This approach can be seen as a method of successive approximations to the solution to (5.26) in which the direction of search is given by the first derivatives of those equations with respect to the unknowns $\lambda_{\ell m}$. These derivatives are incorporated in the related Jacobian matrix

$$\frac{\partial \Lambda}{\partial \lambda_{pq}} = \sum_{ij} t_{ij} p_{ij}^{\ell m} \left[e^{-\sum_{\ell m} \lambda_{\ell m} p_{ij}^{\ell m}} \right] p_{ij}^{pq}$$

for L_c links pq

After a series of approximations suitable values for $\lambda_{\ell m}$ can be found. These can be substituted into Equations (5.24) and (5.25) to obtain the estimated trip matrix $[T_{ij}]$ which constitutes the original problem.

Eriksson has produced two routines in FORTRAN to implement this method and these were adapted by the author to work within the general framework of the models described in Chapters 7 to 10.

This method requires considerable computer core memory for real size problems since the whole of the array $[p_{ij}^{\ell m}]$ must be made available at all times. It is possible to 'compress' this array by including only the non-zero values and using pointers indicating the O-D pair to which they relate. This was done in the author's program METWO3, but even then core requirements restrict this approach to small networks.

Often in practice the Equations (5.2) turn out to be inconsistent at the path flow level and no feasible solution with non-negative T_{ij} can be found. This method seems to fail to provide in these cases a reasonable approximation.

5.2.3 The multi-proportional problem

A second approach to solving the entropy maximising programme follows from the dual problem. This algorithm, which can be seen as an extension of the well known bi-proportional problem, has a long history. Its evolution and relationship with the three Rockafellar problems will be presented below.

5.2.3.1 The bi-proportional problem. Consider the following problem. Given a matrix of numbers, none of them negative, and some desired row and column sums, adjust the numbers, as little as possible, so that the target sums are achieved. This is a common problem in traffic forecasting, be it vehicle (Fratar, 1954) or telephone (Bear, 1976) and also in forecasting inter-industry input-output matrices (Bacharach, 1970).

The problem is a bi-proportional problem if the numbers are adjusted by multiplying by two factors, one for the row and one for the column in which they appear. The most commonly used algorithm to find the modified numbers proceeds by finding the multiplicative factors which balance each (row or column) sum in turn.

According to Murchland (1978) the earliest application of this approach was made by Kruithof (1937) for forecasting future telephone traffic in Luxembourg. The balancing factor operations used in solving a doubly-constrained gravity model correspond to the Kruithof algorithm.

The bi-proportional problem can be extended if, instead of requiring just row and column sums to be met, a more general set of target sums are to be enforced. This problem is treated next.

5.2.3.2 The multi-proportional extension. The multi-proportional problem has been extensively studied by Murchland (1977, 1978) who describes its simplest version as:

> "Suppose a set of positive numbers are given, and that certain subsets of these numbers are intended to sum to given totals: -- If the numbers do not sum as intended, a multi-proportional problem arises if we allow each number to be adjusted by multiplication by a proportionality factor for each sum in which it appears."

The desired sums can be written as

$$\sum_i q_i \delta_i^k = S^k \qquad (5.27)$$

where S^k is the k^{th} sum, q_i are the numbers, and $\delta_i^k = 0$ or 1 indicating whether or not q_i appears in sum k .

If the original numbers are represented by q_i^o the desired numbers are obtained from

$$q_i = q_i^o \prod_k X_k^{\delta_i^k} \qquad (5.28)$$

where the X_k are the proportionality or correction factors, one for each sum k .

For example, the original numbers

$$q_1^o = 1 \ , \ q_2^o = 2 \ , \ q_3^0 = 3 \ , \ q_4^o = 4 \ , \ q_5^o = 5 \ , \ q_6^o = 6$$

are required to be modified to meet the following sums.

$$q_1 + q_3 = 5 \quad . \tag{5.29a}$$
$$q_2 + q_4 = 7 \tag{5.29b}$$
$$q_1 + q_4 + q_5 = 9 \tag{5.29c}$$
$$q_5 + q_6 = 10 \tag{5.29d}$$

In this case there are four proportionality factors, one for each sum and the resulting adjusted values are

$$q_1 = 0.992$$
$$q_2 = 2.816$$
$$q_3 = 4.008$$
$$q_4 = 4.184$$
$$q_5 = 3.824$$
$$q_6 = 6.176 \quad .$$

It is possible to generalise the multi-proportional problem by allowing proportions of the numbers q_i to enter in the sums, resulting in

$$\sum_i \gamma_i^k X_i = S^k \tag{5.30}$$

where γ_i^k can take any value between 0 and 1 and replaces δ_i^k .

The solution retains the form of (5.28) but γ_i^k replaces δ_i^k .

Murchland (1977) has suggested an extension of Kruithof's algorithm to solve this problem. This involves taking each target sum in turn and recalculating the proportionality factor X_k so that the target is met. The first iteration in the example above would result in

$$X_1 = 1.25 \quad q_1 = 1.25 \quad q_3 = 3.75 \tag{5.31a}$$
$$X_2 = 1.17 \quad q_2 = 2.33 \quad q_4 = 4.67 \tag{5.31b}$$
$$X_3 = 0.82 \quad q_1 = 1.03 \quad q_4 = 3.85 \quad q_5 = 4.12 \tag{5.31c}$$
$$X_4 = 0.99 \quad q_5 = 4.07 \quad q_6 = 5.93 \tag{5.31d}$$

Successive iterations will modify the factors and numbers further until the solution quoted above is reached.

The extended Kruithof algorithm of repeated balancing of the target sums can be seen as a dual problem method as it deals with the correction factors X_k . In effect, the dual objective function (5.23) can be written in a general notation as

$$\text{Minimise} \left[\sum_i q_i^o (e^{z_i} - 1) + \sum_k S^k \lambda^k \right] \tag{5.32}$$

subject to $z_i = - \sum_k \lambda^k \gamma_i^k$.

Taking the first derivative of (5.32) with respect to λ^k and making it equal to zero results in

$$\frac{\partial}{\partial \lambda^k} = - \sum_i \gamma_i^k q_i^o e^{z_i} + S^k = 0$$

$$- \sum_i \gamma_i^k q_i^o \prod_k X_k^{\gamma_i^k} + S^k = 0$$

Kruithof's algorithm is just a rearrangement of this. The algorithm can be seen as one which minimises the dual objective function (5.32) with respect to each variable in turn and each balancing operation produces a definite decrease in this dual objective. Being a dual method, this algorithm would only produce $[T_{ij}]$ which exactly satisfy the primal constraints in the optimum.

It can be seen that substituting T_{ij} for q_i , t_{ij} for q_i^o , $p_{ij}^{\ell m}$ for γ_i^k and $\hat{V}_{\ell m}$ for S^k results in the formal solution of the entropy maximising problem with proportionality factors $X_{\ell m}$.

5.2.3.4 <u>A practical algorithm</u>. The extended Kruithof's algorithm can now be presented in full. The algorithm requires setting-up an ordered list of counted links $\ell m = 1, 2, \ldots, L_c$, where ℓm is interpreted below as a single label in this list. The algorithm processes one link at a time. In each step the O-D matrix (as estimated up to that point) is modified so that the modelled flow equals the observed volume on that particular link. This can be seen as a process in which one optimises for one variable $X_{\ell m}$ at a time.

More formally the algorithm entails the following.

(i) Obtain, using a suitable assignment method, the values for $[p_{ij}^{\ell m}]$ and set the number of iterations $n=1$.

(ii) Set $X_{\ell m}=1$ for all links and set $T_{ij}=t_{ij}$ for all O-D pairs.

(iii) For each link ℓm in the list in turn:

(a) calculate the modelled (estimated) flow

$$V_{\ell m}^{(n)} = \sum_{ij} T_{ij} p_{ij}^{\ell m}$$

(b) replace $X_{\ell m}$ by $X_{\ell m} \cdot Y_{\ell m}$ where $Y_{\ell m}$ is to be obtained by solving

$$\hat{V}_{\ell m} = \sum_{ij} p_{ij}^{\ell m} T_{ij} \cdot Y_{\ell m}^{p_{ij}^{\ell m}} \qquad (5.33)$$

and

(c) make $T_{ij} = T_{ij} \cdot Y_{\ell m}^{p_{ij}^{\ell m}}$ for all O-D pairs.

(iv) If all the modelled flows $V_{\ell m}^{(n)}$ are suitably near to the observed ones (say within about 5 per cent) stop, otherwise increment n by 1, and return to step (iii).

For all-or-nothing assignemnt ($p_{ij}^{\ell m}$ equal 0 or 1 only) Equation (5.33) in (iii) has a simple direct solution

$$Y_{\ell m} = \frac{\hat{V}_{\ell m}}{\sum_{ij} p_{ij}^{\ell m} \cdot T_{ij}} = \frac{\hat{V}_{\ell m}}{V_{\ell m}^{(n)}} \qquad (5.34)$$

i.e. $Y_{\ell m}$ is simply the ratio of observed against modelled link volume.

For other assignment techniques where $0 \le p_{ij}^{\ell m} \le 1$ a Newton-Raphson method may be used to solve (5.33) as follows.

(a) Start with a first estimation of $Y_{\ell m}$ as $Y_{\ell m} = \hat{V}_{\ell m}/V_{\ell m}$.

(b) Calculate a correction term ΔY

$$\Delta Y = \frac{\sum_{ij} p_{ij}^{\ell m} T_{ij} Y_{\ell m}^{p_{ij}^{\ell m}} - \hat{V}_{\ell m}}{\sum_{ij} p_{ij}^{\ell m} T_{ij} Y_{\ell m}^{(p_{ij}^{\ell m}-1)}} \tag{5.35}$$

and make $Y_{\ell m} = Y_{\ell m} - \Delta Y$ (5.36)

(c) If ΔY is greater than an acceptably small error ε repeat step (b), otherwise accept the current $Y_{\ell m}$ as the solution to

$$\sum_{ij} p_{ij}^{\ell m} T_{ij} Y_{\ell m}^{p_{ij}^{\ell m}} - \hat{V}_{\ell m} = 0 \tag{5.33'}$$

It has been found that in practice 2 or 3 iterations are sufficient to find suitable values for $Y_{\ell m}$. Moreover, whenever modelled and observed flows are not too different their ratio provides a good approximation to $Y_{\ell m}$ in all cases. The Newton-Raphson method therefore is only required in the first few iterations of Kruithof's algorithm and although it slows down calculations it keeps the number of Kruithof's iterations at a reasonable level. The use of Equation (5.34), for $0 \le p_{ij}^{\ell m} \le 1$ instead of (5.35) and (5.36) increases considerably the number of iterations required to solve the original problem as it underestimates $Y_{\ell m}$, i.e. the correction to the $X_{\ell m}$ factors, in each case.

Of course convergence is generally achieved more rapidly for all-or-nothing assignment as the solution to (5.33) is obtained directly with (5.34).

This algorithm is quite simple to program as it consists of a process of repeated corrections to current values. The algorithm need not keep the whole of the array $[p_{ij}^{\ell m}]$ in core as only the elements corresponding to the link ℓm being processed are required. An external disk file containing the variables $p_{ij}^{\ell m}$ ordered by (observed) link can be used. The program METWO2 has been developed by the author to estimate a trip matrix using this extended Kruithof's algorithm.

Although Murchland (1977) did not prove the ultimate convergence of this algorithm a very recent paper by Lamond and Stewart (1981) has shown it to be a special case of an efficient general balancing method studied by Bregman (1967). According to Bregman the algorithm converges to the unique solution provided the constraints (5.2) are consistent. But even if the path flow continuity conditions are not met this algorithm can still be used. In effect, the algorithm can be programmed to carry on minimising the (dual) objective function until further reductions are no longer possible due to the inconsistencies between the assumed assignment and observed traffic counts. The trip matrix thus can be said to be the one that is nearest to meeting the (inconsistent) path flow continuity conditions.

5.2.4 Comparison of solution methods

The two algorithms suggested in Sections 5.2.2 and 5.2.3 have been programmed for the Leeds University computers (ICL1906A and Amdahl V7) in FORTRAN. The O-D estimation problem for the simple network presented in Section 4.3 was correctly solved by both methods; the direct maximisation (Eriksson's) algorithm took 0.1 CPU seconds and the multi-proportional algorithm took 0.3 CPU seconds on the Amdahl V7. The direct maximisation method seems to be faster for small and (path flow) consistent problems but has the following limitations in medium to large problems.

(a) It requires the whole of the array $[p_{ij}^{\ell m}]$ to be in core and uses additional working space proportional to the size of this array. This limits the size of the problem to be handled to, say, 20 zones and 50 links.

(b) The method fails to provide an estimate of the O-D matrix if the constraints (5.2) are inconsistent.

(c) The program, as it stands, seems to be sensitive to linearly dependent constraints (5.2).

The multi-proportional algorithm on the other hand is robust to linearly dependent constraints (5.2) and requires less core memory, enabling larger problems to be handled. But perhaps its most important characteristic is that it produces an estimated trip matrix even if path flow continuity conditions are not fully met. It will be seen in later chapters that this feature is particularly useful when O-D estimation and assignment are handled iteratively.

5.3 THE VALUE OF INDIVIDUAL TRAFFIC COUNTS

5.3.1 Choosing counting sites

The use of the entropy maximising model to estimate an O-D matrix from link flows requires a more or less up to date set of traffic counts * . In principle at least, the greater the number of counts available the less underspecified the problem and the better the estimated O-D matrix should be. It is rare in practice, though, to have a full set of counts available; budget and practical considerations usually restrict the number of counts that can be carried out.

Faced with a partial set of traffic counts the transport analyst will often have to tackle the following problem. Given a limited budget to collect some extra traffic counts, what is the best choice of counting sites in terms of improving the estimate of the O-D matrix? In other words, one would like to know how valuable is each prospective count in terms of obtaining a good estimated trip matrix.

---

* Although not all counts need be taken on the same date or even month. Standard techniques can be used to 'convert' a count on a given day to another day, or perhaps to an annual average daily traffic count.

This section discusses some theoretical considerations in the choice of counting sites.

5.3.2 The information content of a traffic count

The only piece of work identified by the author in this area is restricted to approaches based on the gravity model. Hogberg (1975) studied the value of traffic counts in calibrating a gravity model using artificial data. He observed that links with expected high flow and those used by O-D pairs not yet counted on other links were particularly valuable. He then developed a more rigorous selection approach which is only applicable to estimations based on a simple travel demand model on the assumption that there is no specification error. He concluded, for example, that once enough links have been included to enable calibration of the model (at least as many links as parameters in it), the gain from adding extra counts diminished rapidly.

Given the entropy-maximising/information-minimising framework of this research it seems natural to attempt to ascertain the value of an extra traffic ccunt in terms of the expected information gain resulting from its inclusion. It is assumed at this stage that a basic but incomplete set of counts and perhaps an outdated O-D matrix are available. These enable a prior estimation of the O-D matrix $[t_{ij}]$ to be made. What is required is a measure of the information likely to be added by each new link counted without having to reproduce the whole estimation process using the ME2 model. To obtain such an approximation consider first the modified objective function discussed in Section 5.2.1

$$S'' = \sum_{ij} T_{ij} \log T_{ij}/t_{ij} - T_{ij} + t_{ij} \qquad (5.22)$$

The Taylor's series for $\log x$ for $x \geq \frac{1}{2}$ is

$$\log x = \left(\frac{x-1}{x}\right) + \frac{1}{2}\left(\frac{x-1}{x}\right)^2 + \frac{1}{3}\left(\frac{x-1}{x}\right)^3 + \ldots$$

and for x close to 1, $\log x$ can be approximated by the first two terms.

By assuming that $[T_{ij}]$ is sufficiently close to $[t_{ij}]$ so that $x = T_{ij}/t_{ij}$ is ~ 1 for all ij, S'' can be approximated by

$$S'' \approx \sum_{ij} \left\{ T_{ij}\left(\frac{T_{ij}/t_{ij} - 1}{T_{ij}/t_{ij}}\right) + \frac{T_{ij}}{2}\left(\frac{T_{ij}/t_{ij} - 1}{T_{ij}/t_{ij}}\right)^2 - T_{ij} + t_{ij} \right\}$$

$$= \sum_{ij} \left\{ T_{ij}\left(\frac{t_{ij}}{T_{ij}}\ \frac{(T_{ij}-t_{ij})}{t_{ij}}\right) + \frac{T_{ij}}{2}\left(\frac{t_{ij}}{T_{ij}}\ \frac{(T_{ij}-t_{ij})}{t_{ij}}\right)^2 - T_{ij} + t_{ij}\right.$$

$$= \sum_{ij} \frac{1}{2} \frac{(T_{ij}-t_{ij})^2}{T_{ij}} \tag{5.37}$$

In this equation the approximated objective function has the form of an error-like measure between prior and estimated trip matrices.

The introduction of a new count $\hat{V}_{\ell m}$ will modify the prior trip matrix $[t_{ij}]$ into $[T_{ij}]$ and the first order effect of such change can be represented by

$$T_{ij} = t_{ij}\, Y_{\ell m}^{p_{ij}^{\ell m}} \qquad \text{all } i,j$$

The contribution of this new count $\hat{V}_{\ell m}$ can then be estimated from (5.37) as

$$I_{\ell m} = \frac{1}{2} \sum_{ij} \frac{\left(t_{ij} Y_{\ell m}^{p_{ij}^{\ell m}} - t_{ij}\right)^2}{t_{ij} Y_{\ell m}^{p_{ij}^{\ell m}}}$$

$$I_{\ell m} = \frac{1}{2} \sum_{ij} t_{ij} \frac{\left(Y_{\ell m}^{p_{ij}^{\ell m}} - 1\right)^2}{Y_{\ell m}^{p_{ij}^{\ell m}}} \tag{5.38}$$

In the case of all-or-nothing assignment this can be simplified further as

$$Y_{\ell m} = \frac{\hat{v}_{\ell m}}{v_{\ell m}}$$

where

$$v_{\ell m} = \sum_{ij} t_{ij} p_{ij}^{\ell m} \tag{5.39}$$

is the 'prior' count. The contribution to the objective function S" becomes

$$I_{\ell m} = \frac{1}{2} \sum_{ij} p_{ij}^{\ell m} t_{ij} \; \frac{(\hat{v}_{\ell m}/v_{\ell m} - 1)^2}{\hat{v}_{\ell m}/v_{\ell m}}$$

$$= \frac{1}{2} \sum_{ij} p_{ij}^{\ell m} t_{ij} \; \frac{(\hat{v}_{\ell m} - v_{\ell m})^2}{\hat{v}_{\ell m} \cdot v_{\ell m}}$$

Substituting from (5.39) for $v_{\ell m}$ we obtain

$$I_{\ell m} = \frac{1}{2} \frac{(v_{\ell m} - \hat{v}_{\ell m})^2}{\hat{v}_{\ell m}} \tag{5.40}$$

The most valuable new link to count is the one with maximum $I_{\ell m}$ in Equation (5.40). This has a form close to χ^2, an error measure between prior and 'observed' count. Equation (5.40) provides a simple measure for the value of each new count and it is suggested that it should be used for other types of assignment (non all-or-nothing) models as well.

Of course the value of $\hat{V}_{\ell m}$ can only be guessed beforehand, perhaps from local experience. As a result of this analysis the following practical selection method can be put forward.

(i) Use all the information available (outdated trip matrix and/or existing counts) to obtain a first estimation of modelled flows $v_{\ell m}$ for all non-counted links.

(ii) Produce some rough estimates (from experience, road characteristics or quick 6 minute counts) of the expected volumes $\hat{V}_{\ell m}$.

(iii) Estimate $I_{\ell m} = 0.5(v_{\ell m} - \hat{V}_{\ell m})^2 / \hat{V}_{\ell m}$ for each uncounted link and rank them according to these values.

(iv) Select as many links for future counts as required from this list according to their ranking but exclude any link which is linearly dependent on other counts already available.

This recommended practice supports the idea that if a modelled flow seems to be far out of one's expectation that is a fairly important link to count.

The importance of second order effects will be reduced if linearly dependent links are explicitly removed from the ranking.

Large differences between modelled and 'expected' flows will result in a high ranking for the corresponding link.

However, the measure in (5.40) was obtained under the assumption of small differences hence it can only be considered as a useful approximation.

CHAPTER 6
BASIC TESTS OF THE ENTROPY MAXIMISING MODEL

6.1 COMPARING TRIP MATRICES

The entropy maximising model (ME2) for estimating an O-D matrix from traffic counts will always tend to reproduce the observed link counts to any desired level of accuracy provided there is a feasible solution and enough Kruithof iterations are performed. Thus, it does not seem fair to test the model on the basis of its ability to reproduce the observed counts.

A much better test is to compare the estimated trip matrix with the 'real' trip matrix, in other words the trip matrix which is responsible for the generation of the observed link volumes. Only one of the models outlined in Chapter 3 has been tested in this way with <u>real</u> data: the French NEMROD model.

The main reason for this is that 'real' O-D matrices are very expensive to obtain. However, Hogberg ard Nguyen have carried out tests with artificial data analogous to those reported for the ME2 model below.

This section discusses the best statistics to use when comparing trip matrices to ascertain how close the estimated matrix is to the original 'real' trip matrix.

6.1.1 <u>Goodness of fit statistics</u>

Dissatisfaction with some of the results obtained with the use of conventional, and in particular distribution, models has led to a surge in research on the most suitable statistics for comparing trip matrices. Some recent studies are particularly interesting in that they compare the performance of different statistics; see for example Black and Salter (1975) and Smith and Hutchinson (1981).

The goodness of fit statistics most often advocated are reviewed below. In the formulae n is the number of cells in the matrix (M^2-M), T_{ij} are the estimated values for these cells, that is trips from origin i to destination j and $\hat{T}_{ij}$ are the corresponding observed values against which the estimated ones are to be compared.

The problem is to find a suitable measure of how 'close' estimated values are to the 'real' or observed ones.

(i) Chi-square

The chi-square goodness of fit test indicates whether or not a set of data may be regarded as consistent with a model. This is achieved by calculating a test statistic

$$X^2 = \sum_{ij} \frac{(\hat{T}_{ij}-T_{ij})^2}{T_{ij}} \qquad (6.1)$$

If the model is correct X^2 can be assumed to have a χ_v^2 distribution where v are degrees of freedom given by $v = n - k - 1$ where k is the number of estimated parameters in the model.

For n large compared with the calibrated parameters χ_v^2 is approximately a normal distribution with mean v and variance $2v$. The null hypothesis, that is that the model estimates agree with the observations is rejected when $X^2 \geq \chi_v^2$.

However, the sampling distribution of X^2 is only approximately χ^2 distributed, the approximation being good only if the estimated values are above a certain minimum value m. Minimum cell values of 5 or 6 are usually recommended.

Due to the sparseness of O-D matrices these conditions are not normally met and methods must be found to either group or ignore low value cells. A systematic way of doing this has been proposed by Pitfield (1978). Leese (1977) has suggested a modification to the chi-square test which requires a direct calculation of the moments of X^2

(without making any assumption about the size of $\hat{T}_{ij}$) and then to use them to determine the distribution of X^2 without having to approximate it to χ^2 . However, this is not a simple task and for certain conditions ($\hat{T}_{ij} = 0$ for example) it is not possible to calculate the moments of X^2 .

(ii) Coefficient of determination R^2

The coefficient of determination is defined here as

$$R^2 = 1 - \frac{\sum_{ij}(\hat{T}_{ij}-T_{ij})^2}{\sum_{ij}(\hat{T}_{ij}-\bar{T}\ )^2} \qquad (6.2)$$

where $\bar{T}$ is the mean cell value $\bar{T} = \frac{1}{n}\sum_{ij} T_{ij}$. This is probably the most common measure of goodness of fit but also one of the least meaningful. It originates as a measure of the degree of linear, dependence between two random variables but in this context T_{ij} cannot be considered to be one as it results from a deterministic model. R^2 also gives a very high weight to large absolute errors and accordingly a high R^2 can be obtained with a matrix with small errors for large cells but very poor correspondence at lower cell values. A partial compensation for this effect can be achieved if the R^2 value is calculated over the square root of cell values as

$$SR^2 = 1 - \frac{\sum_{ij}(\hat{T}_{ij}^{\frac{1}{2}}-T_{ij}^{\frac{1}{2}})^2}{\sum_{ij}(\hat{T}_{ij}^{\frac{1}{2}}-\tilde{T}\ )^2} \quad . \quad \tilde{T} = \frac{1}{n}\sum_{ij} T_{ij}^{\frac{1}{2}} \qquad (6.3)$$

Used in this context both (6.2) and (6.3) may turn out to be negative for large differences between modelled and observed cell values.

(iii) Root mean square error

This is a measure of error based on Euclidean distance and usually defined as

$$\text{RMS or RMSE} = \left[\sum_{ij} (\hat{T}_{ij}-T_{ij})^2 / n \right]^{\frac{1}{2}} \quad . \qquad (6.4)$$

Other authors use the related standard deviation of the residuals defined as

$$\hat{\sigma} = \left[\sum_{ij} (\hat{T}_{ij} - T_{ij})^2 / (n-1) \right]^{\frac{1}{2}} \tag{6.5}$$

In both cases if a normal distribution of the errors can be assumed a probability can be readily assigned to the error ranges of the model. For large n (6.4) and (6.5) will be virtually identical and one could equally call RMS the standard deviation.

The percentage root mean square error (and $\% \hat{\sigma}$) can be calculated for different ranges of cell values and can be defined as

$$\% \, \text{RMS} = \frac{\text{RMS}}{\bar{T}} * 100 \tag{6.6}$$

and in similar terms it is possible to define a coefficient of variation

$$\hat{C}_v = \frac{\hat{\sigma}}{\bar{T}} \tag{6.7}$$

These error measures have already been discussed in Section 2.5.

(iv) Mean absolute error

A related measure of error is the mean absolute error defined as

$$\text{MAE} = \sum_{ij} |\hat{T}_{ij} - T_{ij}| / n \quad . \tag{6.8}$$

This measure is less sensitive to large absolute errors than the RMS.

(v) Likelihood ratio, or log likelihood difference

This measure has been advocated by Wilson (1976) in this context and it is also frequently used as a goodness of fit test in disaggregate modelling.

It is defined here as

$$\Delta L = \sum_{ij} \hat{T}_{ij} \left[\log \frac{\hat{T}_{ij}}{T} - \log \frac{T_{ij}}{T} \right] \qquad (6.9)$$

where T is the total number of observed trips. This is in fact the log likelihood difference version of this statistic. The advantages of ΔL are that it requires a minimum of assumptions and that it is naturally related to normal calibration procedures. Of two models with different ΔL values it can be said that the one with the higher ΔL value will be a more plausible model of the data.

(vi) Phi statistic

This is a measure related to the information gain or relative information measure as discussed in Section 4.1 and used by Thomas (1977), Batty and March (1976) and Morphet (1975) who is responsible for the name used here.

The statistic is defined as

$$\text{Phi} = \sum_{ij} \hat{T}_{ij} \left| \log \hat{T}_{ij}/T_{ij} \right| \qquad (6.10)$$

The main reason for using the absolute value of $\log \hat{T}_{ij}/T_{ij}$ is to allow the analysis of the contribution of all parts of the matrix to the total error measure. For example, the contribution of under and overestimated cells can both be handled in this way.

The larger the value of Phi the poorer the fit of the model.

6.1.2 Normalisation of error measures

Some of the error measures above cannot be used to compare model fits at different sites or at different times of the day as their values depend on local conditions like matrix size, $\bar{T}$ and so on. The following measures can be used to this end:

R^2 %RMS

SR^2 $\hat{C}_v$

the normalised Phi:

$$NPhi = \sum_{ij} \frac{\hat{T}_{ij}}{T} \left| \log \hat{T}_{ij}/T_{ij} \right| = \frac{Phi}{T} \qquad (6.11)$$

and the normalised mean absolute error:

$$NMAE = \sum_{ij} \left| \frac{\hat{T}_{ij}-T_{ij}}{T} \right| \qquad (6.12)$$

6.1.3 Comparison of statistics

A recent study by Smith and Hutchinson (1981) compared the performance of six of these statistics, X^2 , ΔL , Phi , R^2 , $\hat{\sigma}$ and MAE . The method used was to introduce different levels of errors to a set of eight Canadian observed O-D matrices * and to observe the behaviour of these six statistics. Their results can be summarised as follows.

- R^2 is a poor statistic in the sense that fairly high percentage errors still produced 'reasonable' R^2 , for example when the allowable percentage error was 100 per cent in one case an R^2 of 0.7 was still found.

- X^2 increased sharply for percentage errors beyond 75 per cent but tended to level off for errors greater than 100 per cent.

- ΔL had a similar behaviour to X^2 and both statistics were not considered suitable unless used to compare two different models applied to the same data base.

- $\hat{\sigma}$, Phi and MAE had a reasonable behaviour along all the error spectrum and the normalised versions Nphi and NMAE were recommended for future studies.

These normalised statistics have been used to compare the goodness of fit of double constrained gravity models. Sikdar and Hutchinson (1980) report on the calibration and testing of a double constrained gravity model (and

* Obtained from the 1971 census and representing an 11 per cent sample of all work trips.

variations on it) to an almost 100 per cent sample of trips in Edmonton, Canada. The comparison with the observed data led them to conclude that the model "... produced trip tables which had goodness of fit characteristics equivalent to a trip table produced by introducing random errors of about 75 per cent into the observed trip table."

6.2 TESTS WITH AN ARTIFICIAL DATA BASE

6.2.1 Objectives

The use of an artificial data base in model testing has advantages and disadvantages. On the positive side, an artificial data base gives the analyst full control over data errors and a variety of different cases can be studied. On the other hand, artificial data tests cannot tell how likely the model is to perform well in practice (unless likelihood is so qualified as to mean rather little).

The author performed several initial tests with the ME2 model with the following objectives in mind.

(i) To compare the performance of the model under different conditions, in particular its ability to estimate original O-D matrices of different types.

(ii) To assess the influence of the number of iterations on the goodness of fit between original and estimated trip matrices.

(iii) To study ways of improving the convergence of the algorithm.

(iv) To compare the performance of the ME2 model with the related model put forward by H.Van Zuylen.

(v) To debug the computer programs.

6.2.2 Data base

The chosen artificial data base consisted of the following elements.

(i) An artificial network with 15 centroids (zones) and 36 two-way links as depicted in Figure 6.1. Each link was assigned a distance and travel time.

(ii) Four different original or 'real' O-D matrices numbered I to IV and generated as follows.

Matrix I: Each cell in the matrix (with the exception of those in the main diagonal) contains 50 trips. This does not correspond to any practical situation but it was expected that the ME2 would perform well with it.

Matrix II: In order to generate this matrix trip generations and attractions were arbitrarily chosen for each centroid and a singly constrained gravity model was run with them and the calculated costs from the inter-zonal basic network.

Matrix III: Again a singly constrained gravity model was used with the trip ends obtained as in Matrix II. However, the lengths of all 13 links entering the central area were increased and the resulting new costs were used in the gravity model. For example link (1,3) was increased from 120 to 1200 metres. If the Matrix II can be deemed to represent an O-D matrix in a free standing town one can think of Matrix III as one obtained in a central area where only local counts are available. In both cases the average number of trips per cell was about 50.

Matrix IV: This trip matrix was generated by filling each cell with a random number between 1 and 100. This case may be deemed to represent a situation in which destination choice reflects almost no perceivable regularity.

The four trip matrices of (ii) were loaded (all-or-nothing) onto the network in Figure 6.1 and the resulting flows were used as observations or traffic counts. Both the entropy maximising model and the variant suggested by Van Zuylen (1978) were run on the Leeds University 1906A computer until all modelled link flows were within 2 per cent of the 'observed' volumes. In both cases no information other than the network characteristics and the

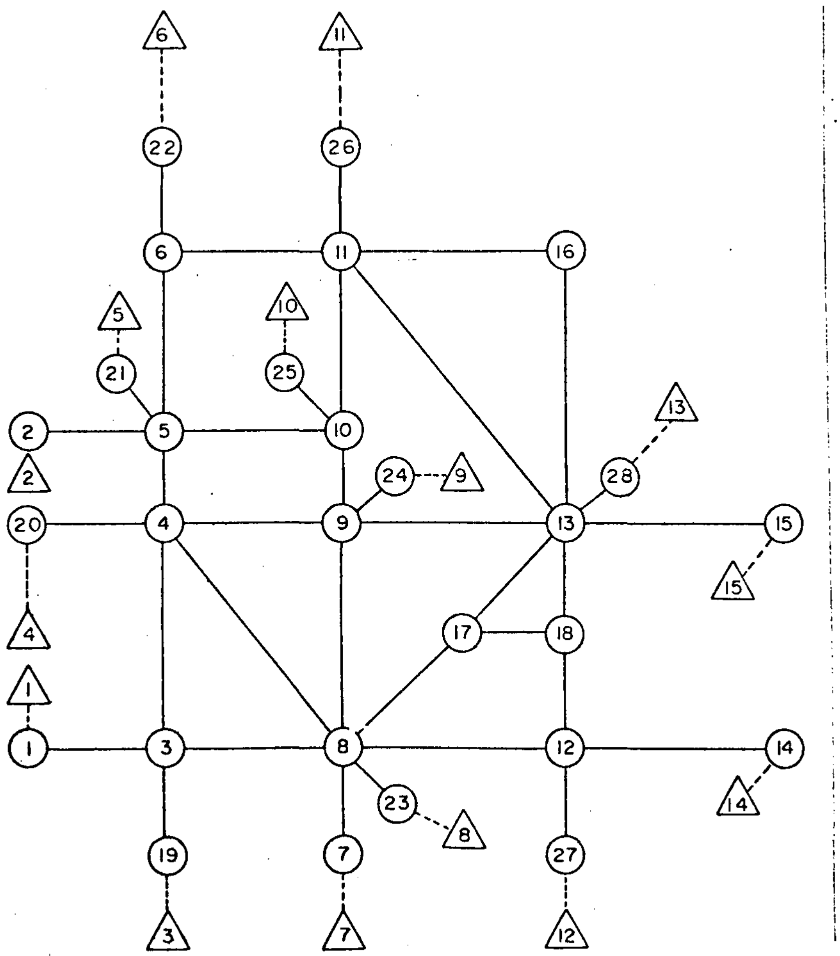


Figure 6.1: Test network

link volumes were used to estimate the trip matrices, i.e. all 'a priori' t_{ij} were set equal to 1 . For simplicity only the RMSE and %RMSE were used in the comparisons.

6.3 RESULTS

6.3.1 Relative accuracy

The results of running both models are summarised in Table 6.1. As expected, the models predicted trip matrix I rather well. A very low RMSE was obtained but could have been reduced even further by increasing the number of iterations to replicate link flows better.

The results for matrices II and III were also very good as can be seen from the scatter diagrams in Figures 6.2 and 6.3 for matrix III. The error levels are about the same for the two matrices but the author's model seems to be predicting better (RMSE of 1.7 against 16-20). The reason behind this agreement probably resides in the common framework (entropy maximising) of the gravity model and the author's model. It should be noted that link flows are counted in effect in and out of each zone. Accordingly the trip ends are 'counted' in both the gravity and ME2 model. Additionally, all links are counted and this information must be equivalent to the information on 'total expenditure' in the gravity model. *

Both the entropy formulation of the gravity model and of the ME2 model have the same objective function. The link flow constraints

$$\sum_{ij} T_{ij}\, p_{ij}^{\ell m} - V_{\ell m} = 0 \qquad (6.13)$$

multiplied by a unit cost which may depend on the link will give

$$\sum_{ij} T_{ij}\, c_{ij} - C = 0 \quad . \qquad (6.14)$$

---

* Both doubly and singly constrained models. This assumes that only network costs are considered in the gravity model.

Table 6.1: Results of tests with synthetic data
Comparison of estimated against 'real' trip matrices

	Van Zuylen's model						Willumsen's model (ME2)					
	Running time (secs)*	Number of iterations	Average RMSE **	Relative RMSE(%) 'real' trips ranges 0-4	5-10	10-∞	Running time (secs)	Number of iterations	Average RMSE	Relative RMSE(%) 'real' trips ranges 0-4	5-10	10-∞
Matrix I (50)	37	14	1.2	0.0	0.0	2.4	16	11	1.1	0.0	0.0	2.1
Matrix II (gravity)	36	15	21.3	68.0	44.3	25.0	22	16	1.7	18.2	4.4	2.0
Matrix III (part grav.)	34	13	16.3	59.5	55.0	18.9	21	15	1.7	20.1	0.0	2.0
Matrix IV (random)	44	18	25.8	1888.1	365.7	43.5	16	11	25.5	1870.0	362.0	43.1

* In ICL 1906A at Leeds University including INPUT/OUTPUT

** Average RMSE calculated as $\text{RMSE} = \sqrt{\frac{\Sigma_k (E_k - O_k)^2}{n}}$

where n = number of cells (210)
E_k = estimated number of trips for the k th origin-destination pair
O_k = 'observed' or 'real' number of trips for the k th origin-destination pair

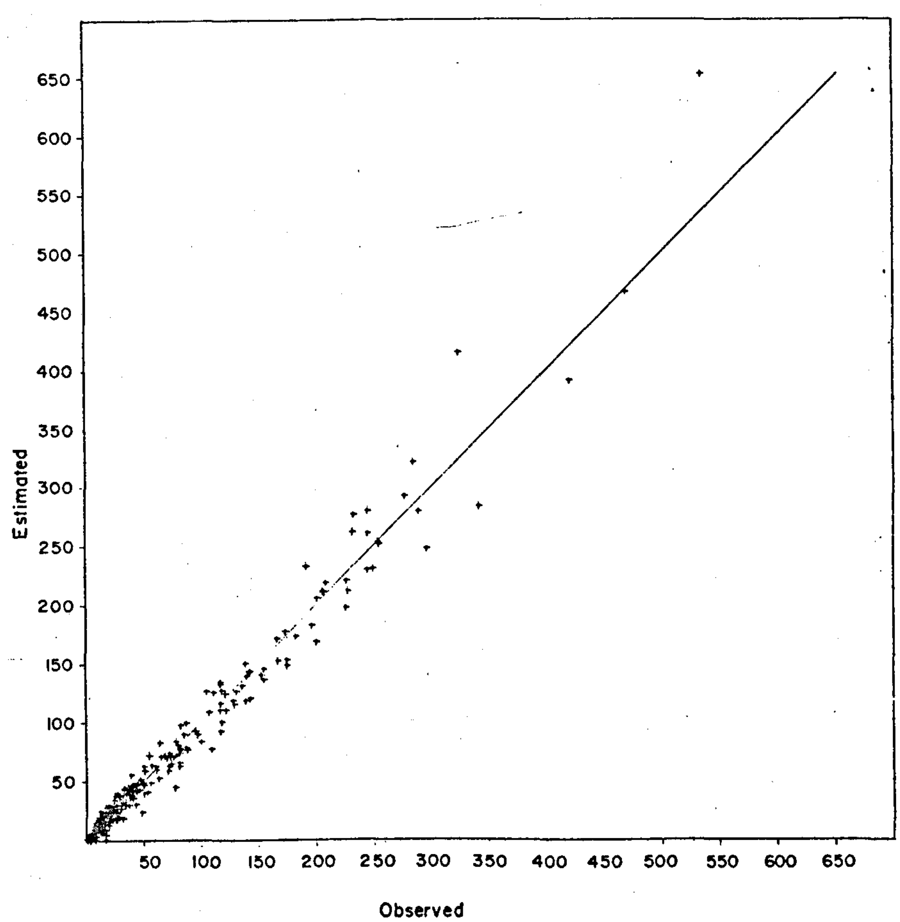


Figure 6.2: <u>Van Zuylen's model. 'Observed' vs estimated trip matrix for matrix III</u>

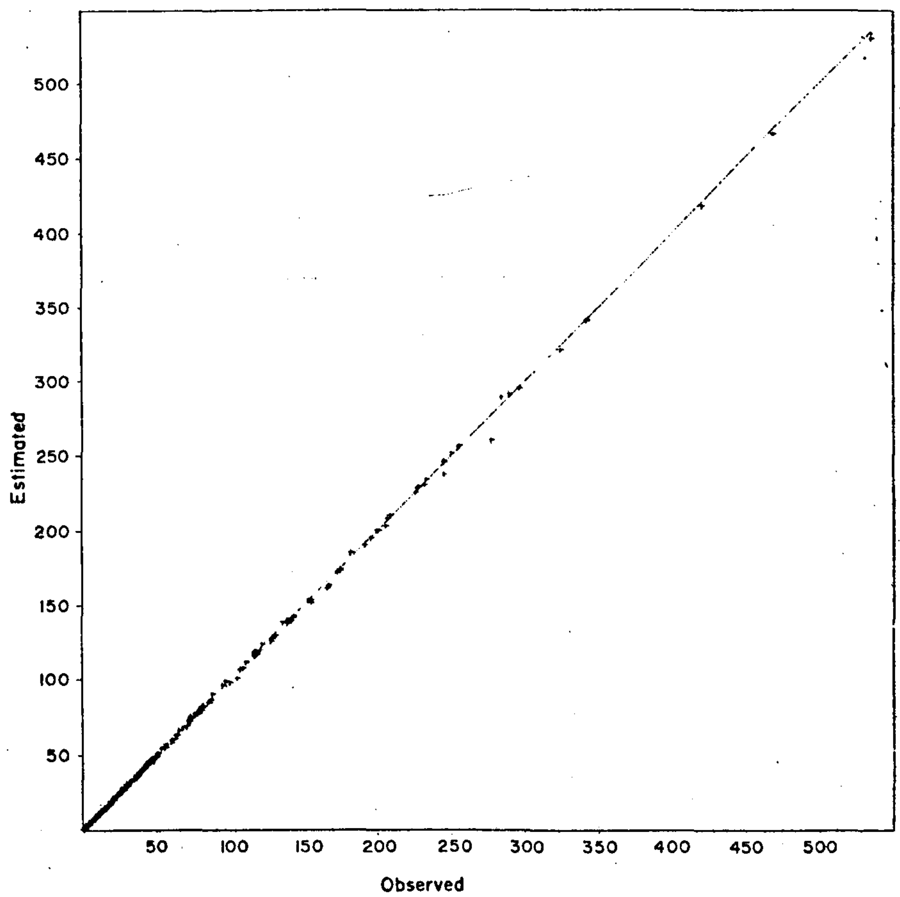


Figure 6.3: Willumsen's ME2 model. 'Observed' vs estimated trip matrix for matrix III

Thus the ME2 model can replicate a gravity model if counts are available on all links (including centroid connectors). Of course, the ME2 model can also pick up some of the structure of less systematic matrices.

Results in case IV were not so good, in particular for cells with few trips in the 'oberved' matrix. Both models performed equally poorly here and a good estimation in this case would benefit from some prior information.

6.3.2 Use of resources

Van Zuylen's model took on average 2.52 seconds per iteration and the entropy maximising model only 1.42 seconds. The difference is mainly due to the non-integer exponent in the first model and should disappear with other types of proportional assignment. Both models required about the same number of iterations and yielded similar Root Mean Squared Error figures for matrices I and IV. The O-D matrices estimated by the entropy maximising model were closer to the original ones for cases II and III. The only other extra requirement of Van Zuylen's model is a REAL array of the same size as the trip matrix and some minor initial calculations to obtain $\sum_{\ell m} p_{ij}^{\ell m}$ for each O-D pair to be stored in that array.

Tests were also made with the O-D matrix in case III to compare the results of the entropy maximising model under different criteria for convergence. These results are displayed in Table 6.2 and Figure 6.4.

Table 6.2: Comparison of estimated against real O-D matrix for case III entropy maximising model ME2

Maximum error in the modelled link flows (%)	Running time (secs)	Number of iterations	Average RMSE	Relative RMSE(%) range		
				0 - 4	5-10	10-∞
15	10	5	14.2	20.1	10.8	16.5
10	13	8	8.3	20.1	7.3	9.6
5	17	12	3.4	20.1	0.0	4.0
2	21	15	1.7	20.1	0.0	2.0
1	22	18	0.9	20.1	0.0	1.0
0.5	25	22	0.5	20.1	0.0	0.6

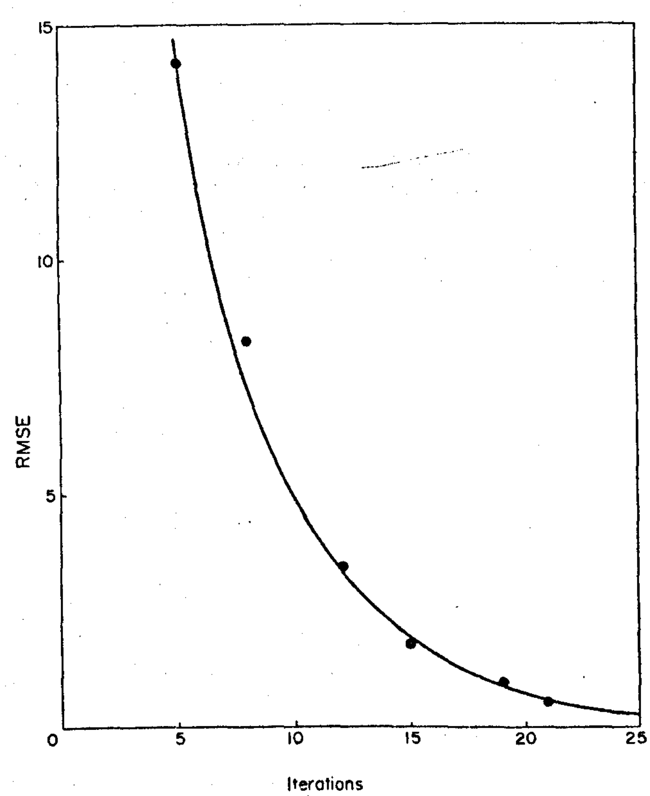


Figure 6.4: RMSE against number of iterations for ME2, entropy maximising model, case III

As expected more stringent convergence criteria produce better results but these improvements take place mainly for cells with 5 or more trips. It appears desirable to aim at replicating the observed flows to at least within 5 per cent to keep errors and running time at a reasonable level.

Tests were also carried out to see if it was possible to accelerate the convergence of Kruithof's algorithm by dealing with the links in particular orders, for example ranking them by volume or by the number of O-D pairs using them. It was found that there was little or no difference in convergence rates and estimated trip matrices. However, Willis and Chan (1980) working with a modified version of the program supplied to them by the author claim that sorting links may influence convergence and in some cases reduce the goodness of fit of the trip matrix. This is in contradiction with what the author has found.

6.4 CONCLUSIONS

A number of alternative indicators of the closeness between 'real' or observed and estimated trip matrices have been discussed. The most promising indicators seem to be the Root Mean Squared Error (percentage) and the statistic Phi (normalised) based on information theory considerations.

Tests were carried out on the entropy maximising model ME2 with artificial data and four different trip matrices. The results were promising for three of them and disappointing for the matrix generated using random numbers.

It must be stressed at this stage that these tests take no account of a number of possible error sources. For example, the following errors are likely to be encountered in real situations but <u>were not</u> included or simulated in these tests:

- network coding errors
- errors and inconsistencies in the counts
- errors in link costs (measured and perceived) and in the assumed assignment model (all or nothing)
- inconsistency between counts and assignment and between counts and prior trip matrix
- daily/weekly and seasonal variations
- sampling errors.

Accordingly, these results only reflect the maximum accuracy one may expect from the model in well conditioned cases. The tests with these synthetic matrices were nevertheless quite valuable for studying the solution algorithm and for debugging the computer programs to be used.

While further synthetic tests incorporating some or all of the above effects would no doubt have been of interest, it was felt that tests with real life data would be of greater value. These are described in the following chapters.

CHAPTER 7
THE READING DATA BASE

The validation of a model for estimating a trip matrix from traffic counts requires a real data base, ideally one consisting of

- a set of traffic counts on a network
- network data (distance, speed-flow curves)
- an independently estimated trip matrix of reasonable accuracy, obtained for example through aerial photography.

This type of data base, and in particular its third element, is difficult to come by and is one of the reasons why most O-D estimating models have been tested with artificial data or on their ability to reproduce observed counts. These two tests are not considered sufficient for the validation of the ME2 model as by definition it tends to reproduce the traffic counts.

This research used a data base collected by the Transport and Road Research Laboratory in the central area of Reading. This is one of the very few data sets available to this end (Chan et al (1980). This chapter describes the characteristics of these data.

7.1 DATA COLLECTION AND ANALYSIS

7.1.1 The Reading survey

The Urban Networks Division of the Transport and Road Research Laboratory has developed a technique based on registration number surveys to obtain comprehensive information about trips in an area. The technique was first applied in Reading in 1976 (Leonard and Tough, 1979) and has subsequently been used in Bedford (Cathcart and Fearon, 1980).

The technique involves locating a large number of observers at key points in the area of interest to record registration numbers and times of vehicles passing their location. These locations are such that a computer program can later on match registration numbers and 'follow' vehicles through the area. In this way not only the origin (entrance point) and destination (exit point) of the vehicle is obtained but also its path through the area.

A major survey using this method was carried out in 1976 by TRRL in order to collect data for a variety of purposes, one of them being the validation of their Traffic Management Simulation model CONTRAM (Leonard et al, 1978). The survey covered the central area of Reading, approximately a 2 x 2 km square as depicted in Figure 7.1. The area contains 5 roundabouts, 19 signal controlled junctions, 6 signal controlled pedestrian crossings, a short stretch of urban motorway and a number of priority junctions. A total of 82 observation points were established, each one with 2 observers one calling out the numberplate (last three digits and year letter) and the other recording it and the time to the nearest 5 seconds. The observers were so located that it was possible to follow the path of any vehicle through the area as a unique chain of observations. The location of the observers is depicted in Figure 7.2.

These data were collected over four consecutive evening peak periods in October 1976 * in such a way that a useful time interval of two hours is available for each day (16.10 to 18.10 approximately for Monday 18th to Thursday 21st October 1976). In order to reduce data collection requirements only vehicles whose numberplates ended in the digit 4 (before the year letter) were recorded. Buses and lorries were included in the observations but were not identified separately.

In addition, automatic pneumatic tube counters were placed at 20 key links in the network. These provide 15 minute counts and they were operated over a much longer

---

* Data was also collected for the morning peak but these have not been used in this research.

Figure 7.1: Reading survey area

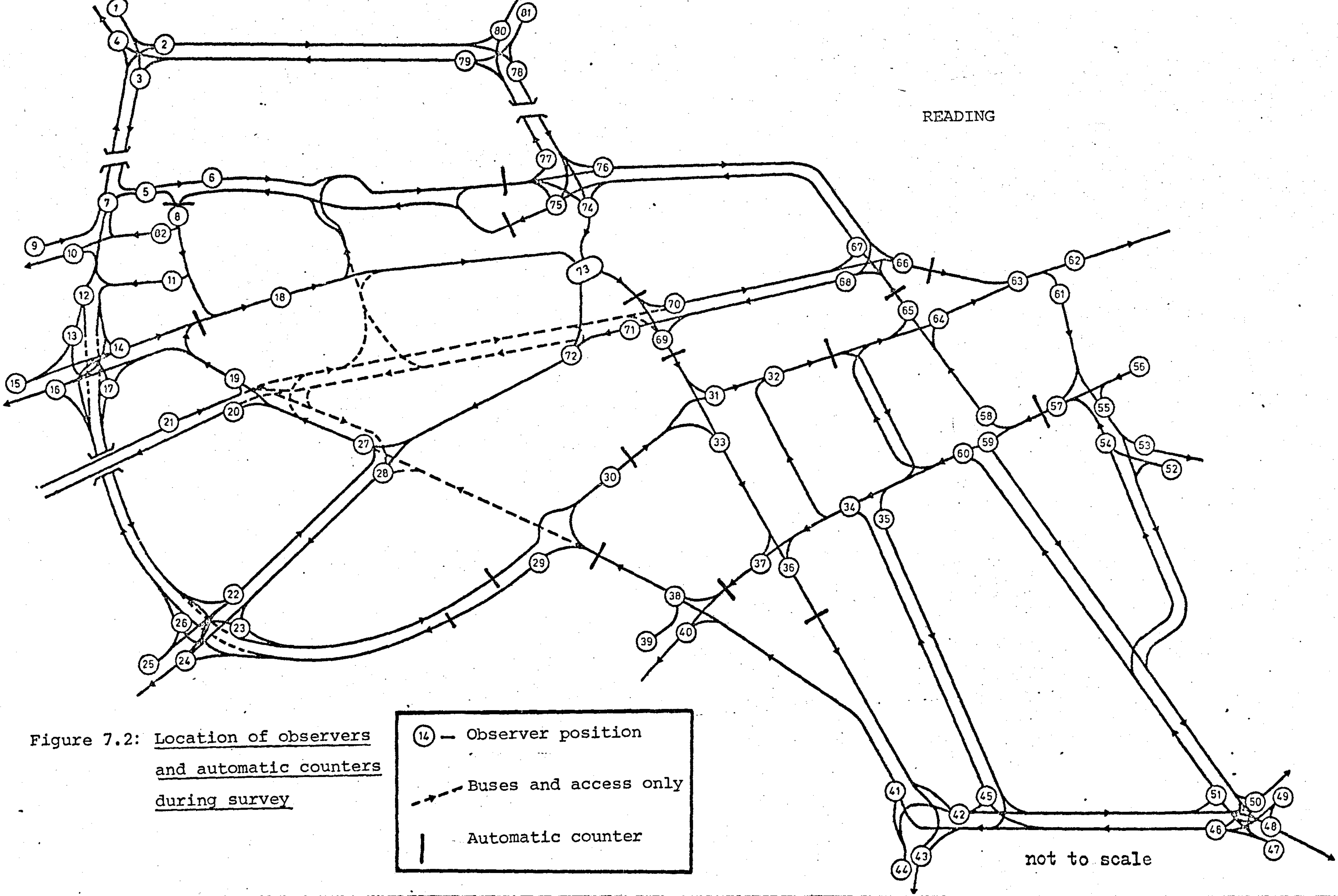


Figure 7.2: Location of observers and automatic counters during survey

period (5 weeks). Their location is depicted in Figure 7.2. Some checks on the reliability of the data were made using information from floating cars but no figures have been made available.

7.1.2 Coding and error trapping

The road network of the area of interest was then coded in a format compatible with the input to the program CONTRAM as depicted in Figure 7.3. The network description and coding conventions are of course those required by the CONTRAM program.

Registration number and observation time data were matched and analysed using this (CONTRAM) coded network. Several error trapping routines were used to eliminate spurious matches due to recording errors. The result of these was a series of successful 'chains' of observations, some 'chains' with gaps or missing links and a set of loose observations which could not be paired.

A second set of routines were then used to allow 'patches' in the chains; these were produced and flagged using special characters in the data base. A detail of these appears in Appendix A2.

Any loose observations remaining after these matches were taken out of the sample leaving the 'processed' observations only. The final results of this process are 4 files, one for each day, containing a series of records, one for each distinct vehicle observed and matched. Each record starts with the numberplate followed by 3 fields for each observation in the chain: special character (interpolation flag), observer number and observation time (in seconds after 16.10 hours). These files were made available to the author for this research, an example of part of one is shown in Appendix A3.

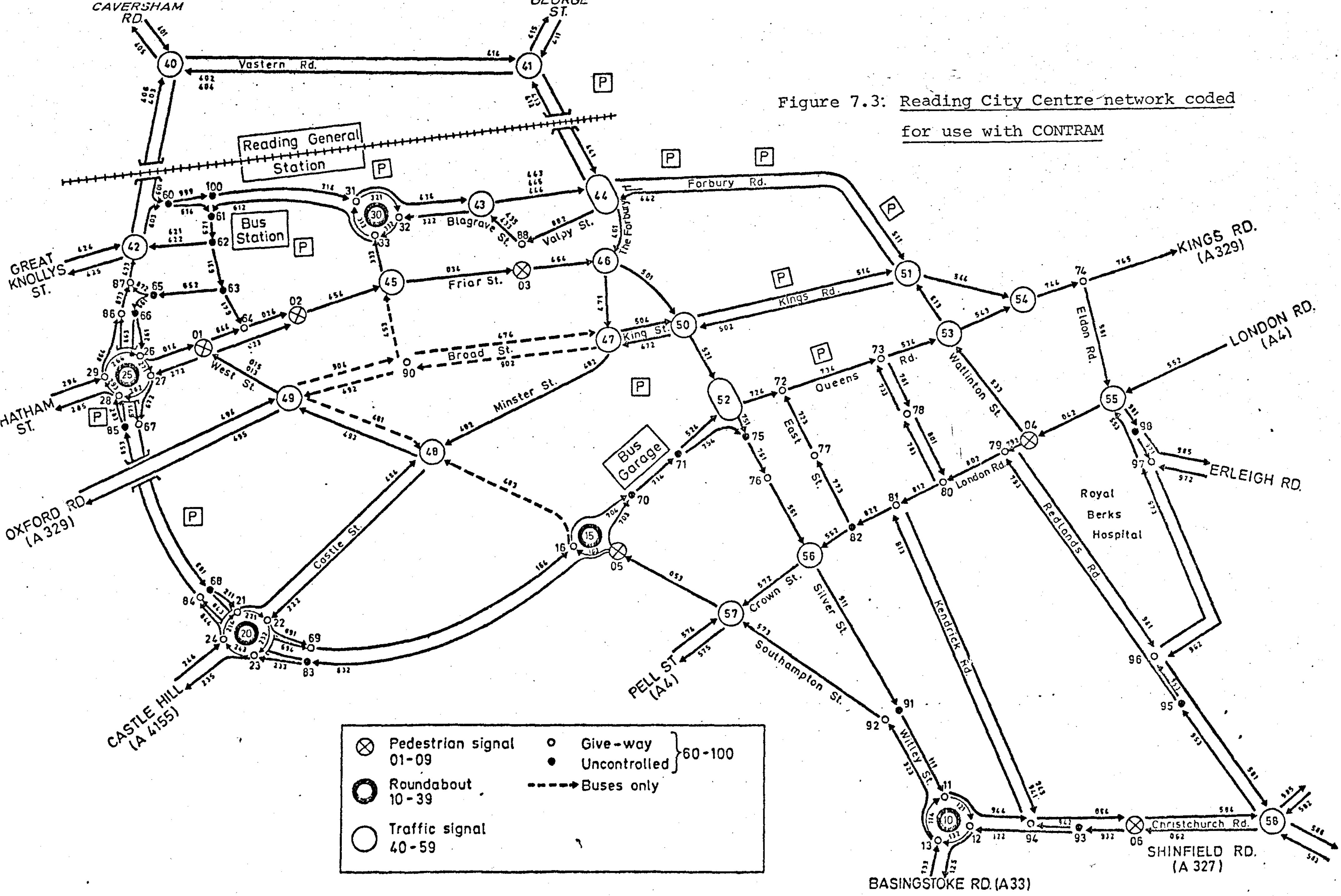


Figure 7.3: Reading City Centre network coded for use with CONTRAM

7.2 NETWORK CODING IN LEEDS

7.2.1 The TRADVV Suite

It was considered desirable to perform most of this research using a general purpose suite of programs for network modelling the TRADVV Suite, although it could have been argued that the use of a software package with more detailed route choice models specifically designed for central areas would be more appropriate; see for example CONTRAM (Leonard et al, 1978) or SATURN (Bolland et al, 1978).

There are two basic objections to this approach:

(i) models like CONTRAM and SATURN were just being developed when this research started, and

(ii) the use of a specific simulation package would make it almost impossible for other researchers to attempt to reproduce these tests with their own assignment programs.

The TRADVV suite of programs developed mainly by Dr. Dirck van Vliet at the Institute for Transport Studies, Leeds University, and based on some of his previous work at the Greater London Council, was selected for the tests of the ME2 model. The suite provides the usual facilities for network coding, tree-building, matrix manipulation, and implements a variety of assignment models including stochastic, capacity restraint and equilibrium assignment. Documentation for the suite is available at the Institute for Transport Studies.

All the programs in this work have been written to be used with the TRADVV Suite * but they can readily be adapted to other suites.

---

* With the exception of two programs for use with SATURN.

7.2.2 Zones and centroids

Since the CONTRAM network coding reflected in Figure 7.3 was not appropriate for TRADVV it was necessary to produce a new description of the network in the more general TRADVV format. The first important decision was the definition of zones or, since in this case no zonal data was used, simply centroids and centroid connectors. One possibility was to define 82 centroids, one for each observation point but, since most of these stations were used to make observations in one direction only, it was possible to group them so that two or at the most three observation stations were associated with one centroid. The result of this choice was a set of 39 centroids as depicted in Figure 7.4. An attempt was made to locate and link these centroids to the road network where natural origin/destination points occurred, for example parking lots and major entrance/exit points to the area.

7.2.3 Link definitions

The definition of the network followed the basic structure of the TRRL coded network presented in Figure 7.3 although the author also visited and inspected the area some time after the survey. All the junctions used by CONTRAM are included as nodes or groups of nodes and all the major stretches of road are also included. However, turning movements are coded only where they are required to describe the only possible movements at a junction; see for example junction 44 in Figure 7.3 and nodes 44 and 89 in Figure 7.4.

Link lengths were obtained from Ordnance Survey 1:1250 maps of the area with the exception of a few links constructed since the map was updated which were then estimated from CONTRAM data. Information on road widths was obtained from the same sources and these values plus an idea of the character of the road in question were used to select speed-flow relationships. These were obtained

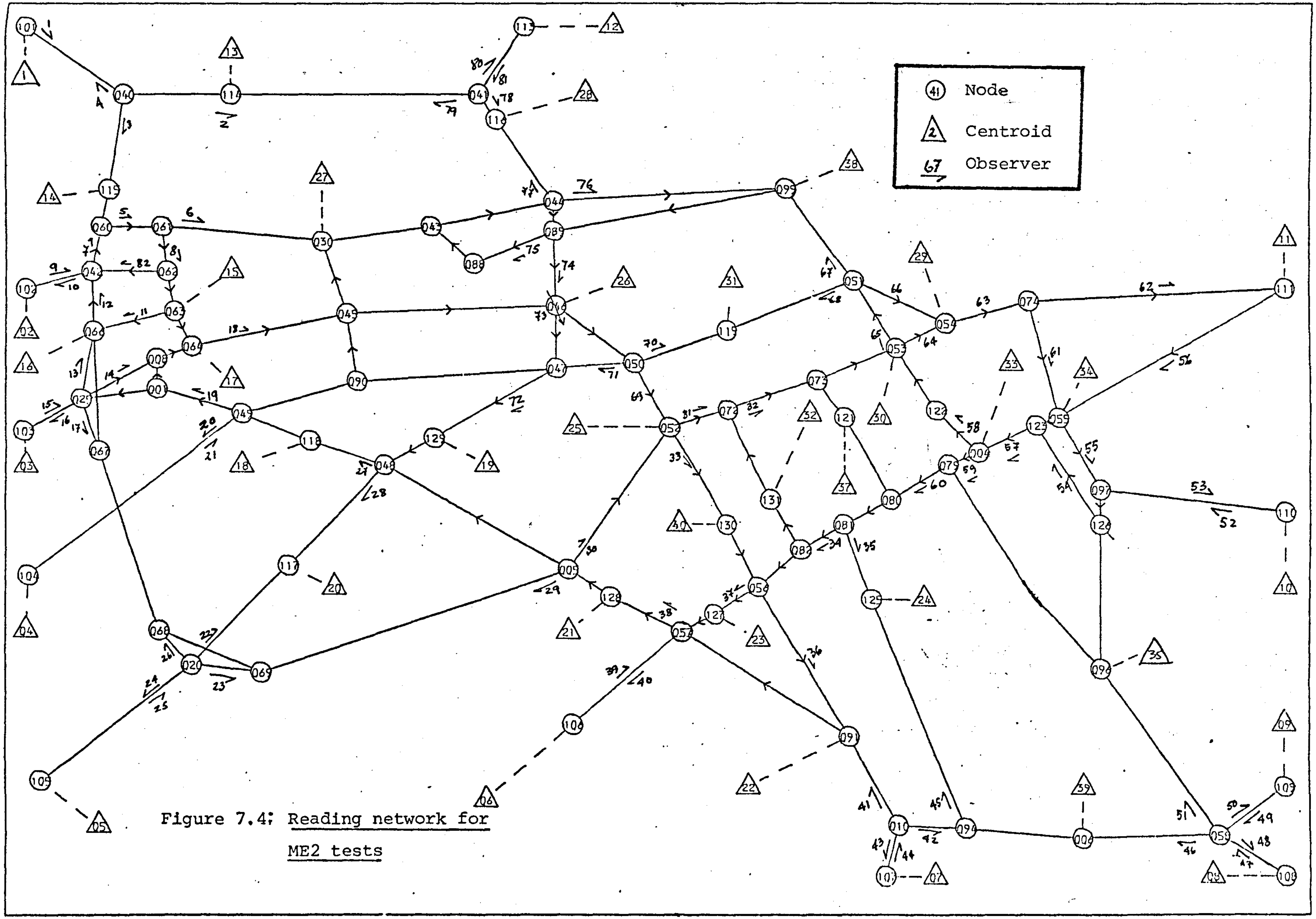


Figure 7.4: Reading network for ME2 tests

from Document 504 (WYTCONSULT, 1975) 'Speed flow relationships on road networks', which offers advice which is more detailed than, but consistent with, DoE (1971) Advice Note 1A.

The central area of Reading has 7 sections of road which can be classified as bus only-access only links. These links were coded for completeness but their distance was entered as a very large number to prevent cars from using them.

The full listing of the network coding appears in Appendix A4.

7.3 TRIP PROCESSING

The available Reading data recorded the movement of vehicles, not trips with clearly defined origins and destinations. The process of mapping vehicle movements into trips is not trivial. For example, one chain of observations might record a vehicle entering the area from Oxford Road, visiting the British Rail Station and then leaving the area again via Oxford Road. These movements are better considered as two trips, one from zone 4 (where Oxford Road enters the area) to zone 27 (British Rail Station) and another from zone 27 to zone 4. Furthermore, the author observed a small number of cars actually using the access/bus only links as through routes in violation of the traffic orders. Some of these violations also appear in the Reading data base.

The problem of mapping vehicle movements into trips was tackled in the following steps.

(i) The program RDGO detected all observed vehicles using access/bus only links and tagged them with a character 'B'. These vehicles represented 3.6 per cent of the data and they were later on effectively removed from the observations. An alternative treatment of these vehicles would have been to consider them as following a fixed route (buses) and keep track of them independently. This approach was followed on the CONTRAM validation work (Leonard and Tough, 1979). In the author's experience private cars also used, illegally, these bus only links for through journeys; it was then considered more appropriate to remove these vehicles altogether from the sample.

(ii) The program RDG1 had three basic inputs:

- a discription of the network (TRADVV format)
- the Reading observations from RDG0
- a table describing the correspondence between a pair of observations and the corresponding (TRADVV format) nodes visited.

The program then processed the data and replaced the 'chain' of observation points by a 'string' of (TRADVV format) nodes visited by each vehicle. RDG1 also checked whether the resulting 'string' of nodes was feasible as the above mentioned 'correspondence table' could not be made sufficiently general for all combinations of observation pairs. Unfeasible node strings were later on corrected by hand using the original Reading data.

(iii) The program RDG2 read the (corrected) node string file output by RDG1 and:

- checked 'looping' trips, that is vehicles visiting a node more than once; these strings were flagged for later analysis by hand
- translated the path of vehicles into $\hat{V}_{\ell m}$ link flows
- translated the path of vehicles into a route choice file with the corresponding values of the variable $[\hat{p}_{ij}^{\ell m}]$.

The looping trips were analysed individually by hand and in general they were found to correspond to visits to major trip origin/destinations such as the British Rail station, the Post Office, the University and a major hospital. Accordingly, these journeys were split into 2 trips. A handful of journeys apparently visited 3 of these places and were consequently split into 3 trips. These corrections were edited into the 'node string' file and the program RDG2 run again.

To summarise, all legal observed vehicular journeys (chains) were interpreted as single trips with the exception of 'looping' journeys which were split into two or three trips depending on the number of repeated visits to nodes.

7.4 ANALYSIS OF THE READING DATA

7.4.1 Sample size

It was possible to ascertain the effective sampling ratio of the observations by comparing the resulting link flows in 7.3 (iii) with 16 independent counts described in Section 7.1.1. The results of this comparison are presented in Table 7.1.

It can be seen that on average the sample implies a sampling rate of nearly 6.9 per cent after error trapping. This is probably an underestimation as the traffic counts were made available to the author only in terms of passenger car units (pcu's).

In order to avoid introducing extra errors into the sample it was decided to perform all tests using sampled flows and the observed O-D matrix generating them. No expansion factors have been used and accordingly the models always estimate the trip matrix corresponding to this 7 per cent sample. However, in order to maintain a correct treatment of delays wherever speed-flow relationships were used the link flows were expanded to represent 100 per cent volumes.

One may consider the trip matrix for this area to be made up of 10 sub-populations, each containing only those vehicles whose numberplates end in a particular digit. The observations collected by TRRL in Reading can be said to be of one sub-population on four consecutive days. Because of inevitable errors in the observations (missed vehicles, missrecording, edited-out chains, etc.) the effective sample is about 7 per cent of the total population and approximately 70 per cent of the sub-population of

Table 7.1: Automatic counts and sampled flows in Reading
(Observations in October 1976 between 16.10 and 18.10 hours)

Counter site	Monday 18			Tuesday 19			Wednesday 20			Thursday 21		
	Count (pcu)	Sample (vehs)	Per cent	Count (pcu)	Sample (vehs)	Per cent	Count (pcu)	Sample (vehs)	Per cent	Count (pcu)	Sample (vehs)	Per cent
1	4961	320	6.4	4905	321	6.5	4966	340	6.9	5178	358	6.9
2	--	218	--	3010	212	7.0	3139	223	7.1	3450	219	6.3
3	3937	256	6.5	4054	204	5.0	4152	304	7.3	4258	261	6.1
4	1672	129	7.7	1408	149	10.5	1106	147	13.3	1173	146	12.4
5	--	74	--	1049	87	8.3	1068	81	7.6	1225	99	8.1
6	2682	192	7.1	2877	175	6.1	2777	172	6.2	2931	197	6.7
7	2860	189	6.6	2790	218	7.8	3069	213	6.9	3092	202	6.5
8	1738	103	6.0	1877	103	5.5	1876	103	5.5	2018	121	6.0
9	2258	143	6.3	2466	172	7.0	2443	174	7.1	2011	160	8.0
10	2477	161	6.5	2693	156	5.8	2566	168	6.6	2644	146	5.5
11	3408	262	7.7	3647	249	6.8	3782	261	6.9	3995	244	6.1
12	1615	117	7.2	1693	111	6.6	1649	112	6.8	1735	101	5.8
13	--	225	--	3101	227	7.3	3083	250	8.1	3108	239	7.7
14	1539	90	5.9	1606	101	6.3	--	105	--	1617	99	6.1
15	--	184	--	2214	168	7.6	2287	185	8.1	2295	150	6.5
16	3527	242	6.9	3544	240	6.8	3512	264	7.5	3768	264	7.0
Total	32674	2204	6.8	42934	2893	6.7	41475	2997	7.2	44498	3006	6.8

-- missing values

Overall 6.87 per cent

interest. In the rest of this work these 7 per cent sample matrices will be referred to as 'observed' or 'sampled' matrices in contrast with the 'real' or 'full' 100 per cent trip matrices.

7.4.2 Stability of trip matrices

The four sampled O-D matrices obtained from the program RDG2 were compared against each other to give some idea of the daily variations in trip making behaviour. These comparisons were carried out with the same programs used later to compare estimated and observed trip matrices. It will be of interest to contrast in later chapters the performance of O-D matrix estimation programs with the daily variations in the observed trip matrices themselves.

The results for the daily variations of the observed trip matrices are summarised in Table 7.2. The indicators used were defined and discussed in Section 6.1.1. Following Smith and Hutchinson (1980) results the statistics Mean Absolute Error (MAE and %MAE), Root Mean Squared Error (RMSE and %RMSE) and Phi (Phi and NPhi) were adopted. As some other studies have used R^2 and SR^2 ($\sqrt{T_{ij}}$) these were also included.

It can be seen that the daily variations at the level of trip matrices are quite important. The Mean Absolute Error is around 0.9 trips, that is about 70 per cent of the average cell value. The coefficients of determination (R^2) for both cells and their square roots are also indicative of important daily variations. Looking at the Phi measure of similarity, it can be seen that cells containing 5 or less trips contribute nearly half of the total (Phi) differences between trip matrices. These cells also contribute nearly half of all the trips in the matrix.

Table 7.2: Comparison of observed trip matrices for different days in October 1976, Reading

Indicator	Dates: 18-19	18-20	18-21	19-20	19-21	20-21
Mean Absolute Difference MAE	0.9	0.9	0.9	0.8	0.9	0.9
Relative MAE %	76	72	75	68	85	70
Coefficient of Determination R^2	0.56	0.62	0.56	0.68	0.61	0.66
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.47	0.51	0.49	0.54	0.49	0.50
RMSE	1.9	1.8	1.9	1.7	1.8	1.7
%RMSE	160	147	154	135	148	134
Phi measure: Range (trips per cell)						
0 - 2	124	133	139	146	158	144
3 - 5	302	286	324	228	234	275
Over 5	413	384	454	388	499	401
All	839	803	917	762	891	820
NPhi (Phi/T)	0.46	0.44	0.51	0.42	0.49	0.43

7.4.3 Coarse zones

One would expect some reduction in the daily variations if a smaller number of coarser zones were used. Figure 7.5 shows how the 39 original zones were grouped into 20 coarser zones on a geographical basis while still retaining the main features of the original zoning system.

The daily variations of the observed trip matrices at this new condensed level are summarised in Table 7.3.

It can be seen that there is some improvement in the indicators of similarity (or error), in particular in relative terms that is %MAE, R^2, SR^2 and RMSE(%). However, the reductions in variability are not dramatic, reinforcing the idea that daily variations at the O-D matrix level seem to be fairly important.

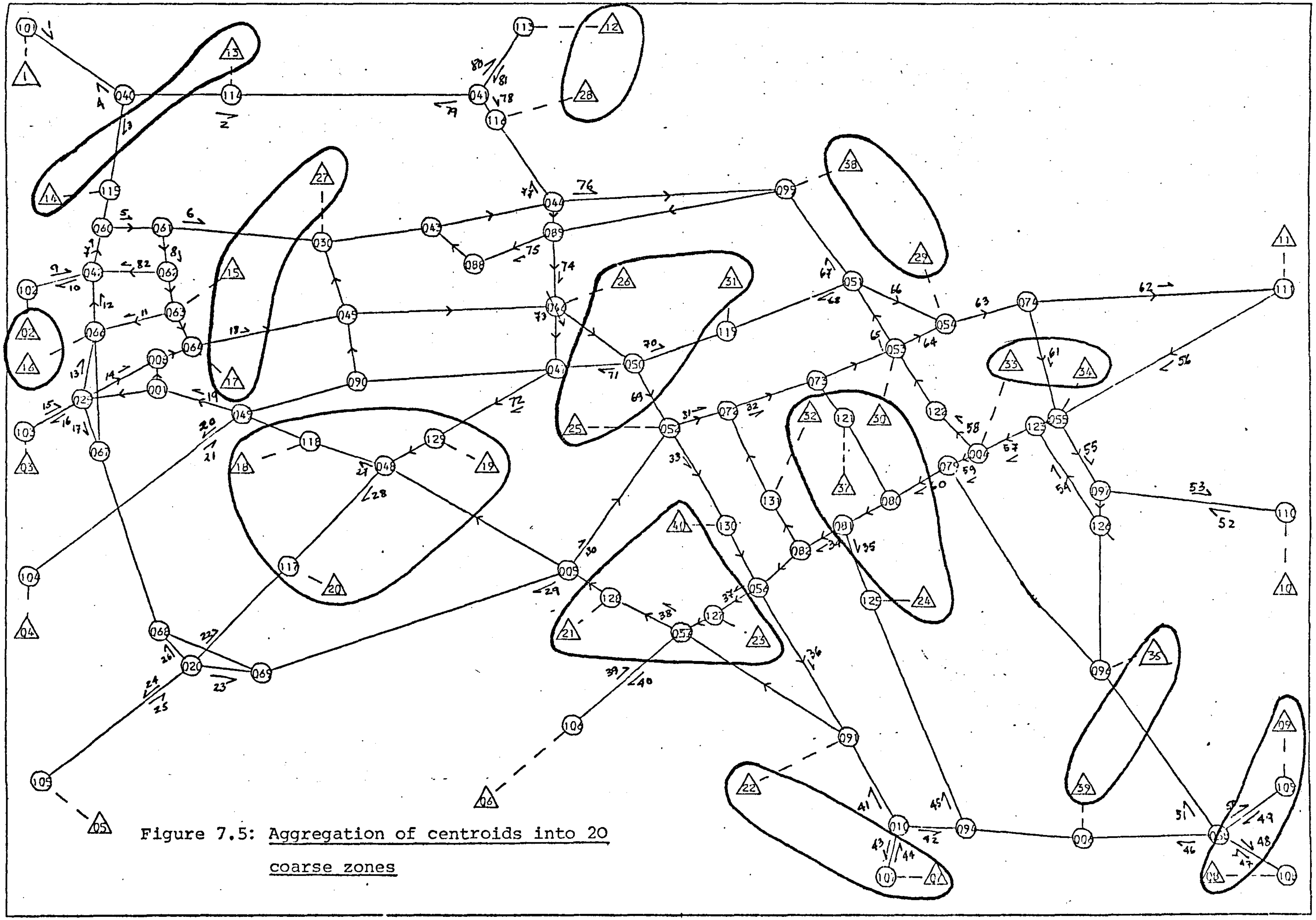


Figure 7.5: Aggregation of centroids into 20 coarse zones

Table 7.3: Comparison of observed trip matrices for different days (grouped), October 1976, Reading

Indicator	Dates:					
	18-19	18-20	18-21	19-20	19-21	20-21
Mean Absolute Difference MAE	2.4	2.3	2.4	2.0	2.2	2.1
Relative MAE %	54	51	52	44	48	44
Coefficient of Determination R^2	0.58	0.63	0.60	0.75	0.69	0.73
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.58	0.59	0.59	0.70	0.66	0.66
RMSE	4.0	3.6	3.7	3.1	3.5	3.1
%RMSE	88	80	82	68	76	65
Phi measure: Range (trips per cell)						
0 - 2	66	58	67	85	81	59
3 - 5	166	138	171	149	173	193
Over 5	638	563	589	474	558	511
All	871	760	828	708	812	763
NPhi (Phi/T)	0.480	0.42	0.46	0.39	0.44	0.41

7.4.4 Variations at link flow level

A similar analysis of daily variations can be carried out for the link flows. Due to the increased aggregation we would expect, again, a reduction in the variability indices. Table 7.4 summarises the results obtained at this level of aggregation using the same indices. It can be observed that the daily variations in link flows are much reduced and are indeed comparable to those found in other studies on the stability of traffic counts, see for example McShane and Crowley (1976).

One must conclude then that the apparent stability of link volumes is likely to mask much greater daily and other temporal variations at the O-D level.

Table 7.4: Comparison of observed link volumes for different days, October 1976, Reading

Indicator	Dates:					
	18-19	18-20	18-21	19-20	19-21	20-21
Mean Absolute Difference MAE	10.7	10.7	11.8	12.0	13.5	11.8
Relative MAE %	11	11	12	13	14	12
Coefficient of Determination R^2	0.91	0.96	0.95	0.94	0.94	0.96
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.96	0.97	0.96	0.97	0.95	0.96
RMSE	19.9	17.0	17.4	19.1	18.5	16.0
%RMSE	21	18	18	20	19	16
Phi measure:						
All	1635	1591	1841	1799	2113	1944
NPhi (Phi/T)	0.109	0.11	0.12	0.12	0.14	0.12

7.4.5 Results with artificial errors

In order to have a better understanding of the magnitude of the daily variations (and estimation errors later) of trip matrices, tests were conducted using artificial error components. These tests consisted of introducing increasing levels of errors in the sampled trip matrix of one day and then comparing the resulting matrix with the original using the same indicators employed in the daily variations tests. In this way, one could assess the level of (artificial) error required to achieve a new matrix as different from the original one as the matrix of that day is different from that of any other day. The matrix observed for Tuesday 19 October was arbitrarily chosen for these tests.

Three probability distributions were used to generate these artificial errors.

(a) A rectangular distribution with mean equal to the sampled cell value t_{ij} and error ranges from $\pm$ 25 per cent to $\pm$ 250 per cent. Wherever the application of this method resulted in a negative cell value this was truncated to zero.

(b) A log-normal distribution. This distribution is non-negative, continuous and allows for a constant variance-to-mean ratio of β , see Aitchinson and Brown (1969).

(c) A normal distribution of errors with mean t_{ij} and $\sigma = a + b\,t_{ij}$, also truncated if necessary for non-negativity.

Table 7.5: The impact of different artificial errors introduced onto the observed 19 October matrix

Indicator	Rectangular distribution range $\pm 100\%$	$\pm 150\%$	$\pm 250\%$	Log-normal distribution ' $\beta = 1$	$\beta = 2$	$\beta = 4$	Normal distribution "
Mean Absolute Difference MAE	0.6	0.9	1.2	0.4	0.6	0.8	1.0
Relative MAE %	50	69	101	35	47	60	81
Coefficient of Determination R^2	0.75	0.64	0.52	0.86	0.76	0.62	0.72
Coefficient of Determination for $(T_{ij})^{½}$ SR^2	0.79	0.64	0.54	0.92	0.87	0.79	0.48
RMSE	1.7	2.4	3.7	1.1	1.6	2.2	1.7
%RMSE	140	196	298	90	126	175	139
Phi measure: Range (trips per cell)							
0 - 2	141	145	192	312	410	493	204
3 - 5	219	195	196	172	240	329	163
Over 5	446	501	553	220	311	483	306
All	806	841	941	704	961	1261	673

' Log-normal with $Var(x) = \beta\, t_{ij}$

" $\sigma = 1.5 + 0.3t_{ij}$

Table 7.5 summarises the results of these tests for those parameters which produced error structures closest to those in the sampled matrices (Tables 7.2 and 7.3). It can be seen that different error distributions (noise) have different effects on the error measures chosen. For example, the observed daily variations as measured by MAE and R^2 are similar to the errors generated by a rectangular distribution in the range $\pm$ 250 per cent. On the other hand, the RMSE measure suggests the range should be only 100 per cent. In effect none of the artificial error structures fully reproduces the pattern of errors encountered in the sampled matrices. The best approximations seem to be the log-normal distribution with $\beta = 4$ and the normal distribution with $\sigma = 1.5 + 0.3t_{ij}$.

7.5 COMMENTS

This chapter has been devoted to the analysis of a unique data base collected in Reading by the TRRL. A description was made of the debugging process applied to the data and of the way in which the network for the ME2 model was coded.

The Reading data is exceptional because it offers the possibility of comparing observed O-D matrices collected for four consecutive days. No analysis of the daily variations of trip matrices is known to the author, surely because of the high cost of collecting reliable trip matrices for different days. This analysis was carried out for the Reading data. It was found that the daily variations at the trip matrix level were fairly large, about seven times bigger (in relative terms) than the daily variations at the link flow level.

There are some obvious difficulties in the interpretation of these findings. One possible interpretation is that trip patterns on different days of the week are not very stable but it can equally be argued that those patterns change not only across weekdays but also across Mondays, Tuesdays, etc. Furthermore, the observed matrices probably show large variations due to their relatively small size and it is possible that daily variations at the 100 per cent of the population level would show lower differences.

In any case, these findings cast some doubts on the current practice of estimating an O-D matrix from observations on single (but often different) days, for example roadside interviews. This evidence suggest that either a more expensive and extended O-D survey should be carried out or that some method must be developed to relate estimated trip matrices to 'representative' conditions.

On the other hand, one is often interested mainly in forecasting link flow levels and these appear fairly stable, even for 7 per cent samples, despite the variations at the trip matrix level. In this case what is needed is an estimated trip matrix which captures enough of the trip pattern to produce reliable link volume forecasts.

The next four chapters will attempt to show that the ME2 model provides means for estimating trip matrices from link volumes and a framework to use traffic counts to update them.

CHAPTER 8
TESTS OF THE MODEL WITH READING DATA

8.1 INTRODUCTION

The previous chapter described the main characteristics of the Reading data base and the form in which it was coded and validated. The study area was divided into 39 zones and the road network contained 80 nodes and 159 (real) one-way links. Each centroid was linked to the network through one (two-way) centroid connector. Each of the observations was followed through the network generating a trip matrix, a set of traffic counts and a set of observed route choice proportions $[p_{ij}^{\ell m}]$.

Two sets of traffic counts can be used from these data. The first one is the set of 159 real one-way links. These are all link flow consistent counts; in effect, it is possible to write 41 flow continuity conditions at every node except those with centroid connectors. The number of independent counts therefore is only 118 (159-41). The second set of traffic counts is an extended set including the 159 real link counts plus 78 (2 x 39) centroid connector counts. However, when centroid connector counts are included 39 additional link flow continuity conditions can also be established making the nett set of independent counts to be 157 (159 - 41 + 78 - 39). These two sets of traffic counts (identified by the words 'real links only' and 'trip end counted' or 'real links plus centroid connectors' or 'extended link counts') are used throughout the tests in the rest of this work. It can be seen that in the best of cases 157 independent pieces of information are used to estimate 1482 unknown trip matrix cells (39 x 39 - 39) making the ratio of unknowns to independent observations 9.4.

Section 8.2 reports on tests of the ME2 model using observed route choice proportions. As the traffic counts used are error free with respect to the sampled trip matrices these tests also remove the possible errors in the assumed route choice model The next section tests the model using the shortest free-flow cost routes as the source for the route choice proportions. This is a crude approximation and one would expect better route choice models to produce results somewhere between these two extreme tests.

These two sets of tests were carried out for each of the four observed days, 18 to 21 October, plus aggregate data from all four days. It was expected that these aggregate data would represent average conditions better. These data sets are referred to either by their dates of by the phrase "4-day data set".

The three computer programs used, METWO1, METWO2 and METWO4 are described in Appendix 5.

Section 8.4 reports on tests using Burrell's route choice model while the rest of the chapter describes sensitivity tests for the ME2 model. In these sections only some of the data sets are used, generally Tuesday 19 October and the 4-day data set.

Through out Chapters 8 to 10 no prior information about the trip matrix has been assumed; that is all t_{ij} were made equal at the start of the iterations * .

In all cases iterations continued until one of two conditions was met:

(a) all modelled flows were within 5 per cent of the observed link volumes, or

(b) it was not possible to reduce the total absolute error in the link flows in two consecutive iterations.

In most cased the first condition was reached first.

---

* The average number of trips per cell was used and estimated from

$$t = \frac{\sum_{\ell m} \hat{V}_{\ell m}}{\sum_{\ell m} \sum_{ij} p_{ij}^{\ell m}} \qquad (8.1)$$

where $\hat{V}_{\ell m}$ represents only the observed links. All prior t_{ij} were made equal to t . It was found that a faster convergence was achieved in this way (1 to 3 iterations less than when making $t_{ij} \doteq 1$).

8.2 TESTS WITH OBSERVED ROUTE CHOICE

The route choice proportions $[p_{ij}^{\ell m}]$ obtained from the Reading data represent, in effect, perfect knowledge about choice of route in the area. The METWO2 program was run for the 5 sets of observations; Table 8.1 presents the results when only real link counts were considered, whereas Table 8.2 contains performance indicators when both real links and centroid connectors are included. Finally, Table 8.3 presents results of these tests for a reduced (condensed) set of 20 zones following the same criteria as in the previous chapter for their aggregation.

The computer times presented are CPU times on the Leeds University Amdahl V7 and do not include output operations.

The following observations can be made at this stage.

(i) The estimated trip matrices are not very close to the observed ones even when perfect knowledge about route choice is available. The observed link flows plus the entropy maximising principle do not provide enough information to reconstruct the original O-D matrix. These results reflect the best that entropy maximising can perform given the information available. One would expect the application of the model under less ideal conditions (i.e. errors in the counts, errors in route choice) to produce poorer results.

(ii) Results for the data set including centroid connector counts are much better than those obtained from counts on real links only, in particular in respect of Relative MAE and the coefficients of determination.

(iii) The model performs very similarly for all observed days and it appears to reproduce the 4 days trip matrix marginally better than a single day. This would suggest that the entropy maximising framework is better suited to model average conditions since some of the daily variations would cancel out.

Table 8.1: Observed vs estimated trip matrix using observed route choice, real link only counted, 39 zones

Indicator	October 18	October 19	October 20	October 21	October 4 days	Typical daily variations * (October) 18-19	Typical daily variations * (October) 19-20
Mean Absolute Difference MAE	1.2	1.2	1.3	1.2	3.5	0.9	0.8
Relative MAE %	99	101	101	100	71	76	68
Coefficient of Determination R^2	0.60	0.62	0.54	0.53	0.50	0.56	0.68
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.55	0.55	0.54	0.55	0.52	0.54	0.66
RMSE	1.8	1.8	2.0	1.9	7.2	1.9	1.7
%RMSE	143	149	156	153	144	160	135
Phi measure Range:							
0 - 2	273	311	289	277	498	124	146
3 - 5	226	256	293	297	694	302	228
Over 5	518	533	603	576	5384	413	388
All	1116	1101	1186	1150	6567	839	762
NPhi (Phi/T)	0.62	0.60	0.63	0.62	0.89	0.46	0.42
Iterations	15	11	9	9	15	--	--
CPU time (secs)	2.76	2.17	1.90	1.84	5.93	--	--

* Monday 18 vs Tuesday 19 October, and Tuesday 19 vs Wednesday 20 October

Table 8.2: Observed vs estimated trip matrix using observed route choice, centroids counted, 39 zones

Indicator	October 18	October 19	October 20	October 21	October 4 days
Mean Absolut Difference MAE	0.8	0.8	0.9	0.8	3.0
Relative MAE %	67	69	68	67	61
Coefficient of Determination R^2	0.64	0.66	0.64	0.60	0.70
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.66	0.66	0.67	0.67	0.63
RMSE	1.6	1.7	1.7	1.7	5.6
%RMSE	130	134	134	138	112
Phi measure Range:					
0 - 2	337	394	365	358	467
3 - 5	394	299	370	356	545
Over 5	469	526	505	578	3339
All	1201	1219	1241	1231	4351
NPhi (Phi/T)	0.66	0.66	0.66	0.67	0.60
Iterations	25	26	24	27	33
CPU time (secs)	4.83	4.79	4.37	4.83	10.32

Table 8.3: Observed route choice condensed 20 zones

Indicator	Centroids counted 19 Oct.	Centroids counted 4 days	Real links only counted 19 Oct.	Real links only counted 4 days
Mean Absolute Difference MAE	2.2	7.6	3.3	9.6
Relative MAE %	49	41	73	52
Coefficient of Determination R^2	0.68	0.74	0.60	0.54
Coefficient of Determination for $(T_{ij})^{½}$ SR^2	0.71	0.72	0.57	0.55
RMSE	3.5	11.0	4.2	14.6
%RMSE	76	59	93	79
Phi measure Range:				
0 - 2	115	76	198	83
3 - 5	151	140	164	133
Over 5	602	2594	474	4087
All	868	2811	836	4301
NPhi (Phi/T)	0.47	0.38	0.46	0.58
Iterations	26	33	11	15
CPU time (secs)	4.79	10.32	2.17	5.93

(iv) When both modelled and observed matrices are condensed into 20 (bigger) zones the two become closer in terms of all the indicators of model performance. This can be interpreted as the model missing out some minor variations which tend to compensate when aggregated into bigger zones.

(v) Compared with the daily variations of the observed trip matrix the performance of the ME2 model looks quite reasonable. The estimated trip matrices are in general closer to the sampled one of one day than are those sampled on another day. In practical terms, if one is trying to obtain a trip matrix representing the conditions of a particular day it seems more accurate to use one estimated by the ME2 model (under these ideal conditions) than to use a matrix from a 7 per cent sample from a different day of the same week.

(vi) Improvements in the performance of the ME2 model over the results reported in this section can only come in practice from two main sources.

- More traffic counts in order to reduce the ratio of unknowns to observations from 9.8 to, say, 5.0 by counting turning movements. In the current network only a few turning movements are counted. This approach would certainly increase the cost of applying the model. Section 8.5 below reports on a sensitivity analysis of the potential value of adding extra counts.
- More and better information about trip making patterns. For example an outdated trip matrix $[t_{ij}]$ could introduce very valuable information which would improve the accuracy of the model. Alternatively, if there are reasons to believe that a gravity model of some sort might be applicable one might use it to generate a prior trip matrix $[t_{ij}]$.

8.3 TESTS WITH ALL-OR-NOTHING ASSIGNMENT

A second group of tests was run by obtaining the route choice proportions $p_{ij}^{\ell m}$ from the shortest routes between each O-D pair. The TRADVV program T1 was used to build the set of minimum (free flow) cost trees for the area.

The program METWO4 * was developed by the author to read those trees and to output a file containing the route choice proportions. The program METWO4 can also be used to add and combine two sets of route choice proportions. The METWO4 program required on average 26 CPU seconds (Amdahl V7) to process a single set of trees for the Reading network.

The program METWO2 was then used to estimate a trip matrix from the observed counts only for each of the five data sets. Table 8.4 shows the results of tests run with only real links being considered, while Table 8.5 shows results with centroid connector counts available. Finally, Table 8.6 presents results for two data sets when the trip matrices are condensed into 20 zones.

---

* See Appendix A5 for documentation on the programs developed during this research.

Table 8.4: Observed vs estimated trip matrix obtained using all-or-nothing route choice, real links only counted, 39 zones

Indicator	October 18	19	20	21	4 days
Mean Absolute Difference MAE	1.1	1.1	1.2	1.2	4.0
Relative MAE %	92	92	95	95	80
Coefficient of Determination R^2	0.43	0.47	0.43	0.42	0.49
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.36	0.38	0.36	0.35	0.43
RMSE	2.0	2.1	2.2	2.1	7.5
%RMSE	170	172	175	170	150
Phi measure Range:					
0 - 2	478	493	451	475	510
3 - 5	561	397	467	563	667
Over 5	694	715	740	649	4729
All	1733	1604	1658	1687	5907
NPhi (Phi/T)	0.96	0.88	0.88	0.91	0.80
Iterations	12	13	13	18	13
CPU time (secs)	1.73	1.82	1.83	2.27	1.80

Table 8.5: Observed vs estimated trip matrix obtained using all-or-nothing route choice, real links and centroids counted, 39 zones

Indicator	October 18	19	20	21	4 days
Mean Absolute Difference MAE	1.1	1.1	1.1	1.1	3.7
Relative MAE %	89	86	86	88	75
Coefficient of Determination R^2	0.45	0.50	0.54	0.47	0.54
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.37	0.42	0.41	0.39	0.46
RMSE	2.1	2.1	2.0	2.1	7.3
%RMSE	169	169	160	166	146
Phi measure Range:					
0 - 2	476	512	470	463	473
3 - 5	567	426	521	533	734
Over 5	598	671	701	629	4934
All	1642	1609	1692	1625	6142
NPhi (Phi/T)	0.91	0.88	0-90	0.87	0.84
Iterations	32	24	23	24	25
CPU time (secs)	4.24	3.56	3.28	3.44	3.53

Table 8.6: Observed vs estimated trip matrix obtained using all-or-nothing route choice, condensed 20 zones

Indicator	No counts 19 Oct.	counts 19 Oct.	No counts 4 days	counts 4 days
Mean Absolute Difference MAE	2.7	2.5	10.0	9.5
Relative MAE %	59	55	54	51
Coefficient of Determination R^2	0.57	0.59	0.58	0.61
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.50	0.52	0.52	0.54
RMSE	4.1	4.1	14.7	14.2
%RMSE	90	89	79	77
Phi measure				
Range:				
0 - 2	117	105	82	73
3 - 5	164	171	144	142
Over 5	706	729	3585	3967
All	987	1004	3810	4183
NPhi (Phi/T)	0.54	0.55	0.52	0.57
Iterations	13	24	13	25
CPU time (secs)	1.82	3.67	1.80	3.53

The use of (free flow) minimum cost trees to estimate the route choice proportions is a very simple, and perhaps over simplistic, approach. The results of these tests are indeed less good than when observed route choice proportions were used. The level of congestion in the study area is illustrated by the fact that, for example, 54 out of 159 links ($\sim$ 34 per cent) are loaded to more than half of their normal capacity. Seen in this perspective it is perhaps surprising how well the ME2 model performs with all-or-nothing route choice.

Part of the explanation for this may be found in the comprehensive system of one-way streets in Reading which certainly reduced the number of potentially attractive routes between any two points. There is no reason to believe, though, that these traffic management measures were any more restrictive than those operative in similar British towns.

The following additional comments can be made.

(i) Again the trip matrices estimated using trip end counts were closer to the observed ones than those obtained with counts on real links only.

(ii) The model performs similarly for all observed days; some indicators suggest that the performance for 18 or 19 October are better (%MAE and R^2), others that the other days produce improved results (%RMSE and Phi/T). The 4-day results are again better than those for any single day and the difference seems to be more significant than when observed route choice proportions were used. This finding reinforces the argument that the ME2 model appears to produce better average conditions.

(iii) Looking at the contribution of the total Phi error from different cell values it appears that cells with more than 5 trips account for only 1/3 of the total error. Compared with the 50 per cent contribution of these cells when observed route proportions are used it appears that most of the increase in error originates in low (< 5) value cells.

(iv) The ME2 model requires less computer time per iteration when operating with all-or-nothing route choice. This is due to the fact that all-or-nothing route proportions are integer values (0 or 1). The number of iterations required is very similar for observed and all-or-nothing route proportions.

(v) The ME2 model with all-or-nothing route proportions estimates trip matrices only slightly outside the range of daily variations of sampled trip matrices (for our Reading data base); see for example the %MAE and %RMSE indicators included in Table 8.1 for comparative purposes * .

(vi) One would expect that the use of a more realistic route choice model would produce results somewhere in between those obtained with all-or-nothing and those using observed route proportions.

Bearing in mind the rather homogeneous results obtained with data from different days of the week and in order to reduce the total number of tests to perform and report on all the rest of this research centred around the Tuesday 19 and 4 days data bases.

8.4 TESTS USING BURRELL's ROUTE CHOICE MODEL

Given the well known weaknesses of all-or-nothing route choice, it was considered of interest to test the performance of the model using a stochastic assignment model. Burrell's model, the most frequently used stochastic assignment technique in Britain (Lai and Van Vliet, 1979), was chosen for these tests. Burrell's approach is to

---

* Table 8.9 on page 213 presents a summary of the performances of the ME2 model with observed, all-or-nothing and stochastic route choice.

introduce a rectangularly distributed error on each link cost and to build one or more trees per origin with different sets of randomised costs. The greater the spread of the rectangular distribution, the greater the spread in routes generated.

The TRADVV program T1 was used to generate three trees for each origin using Burrell's method and two different cost spreads, 10 and 30 per cent (10 per cent is probably the most frequently used in practice). METWO4 was then used to obtain the route choice proportions $[p_{ij}^{\ell m}]$ as an average of these trees in each case. The METWO2 model was then applied using these proportions to estimate trip matrices for 19 October and the 4-day data base and for cost spreads of 10 and 30 per cent, see Tables 8.7 and 8.8. Table 8.9 summarises the results for different route choice proportions.

The application of Burrell's model produced in each case sets of $[p_{ij}^{\ell m}]$ different from those obtained from the all-or-nothing route choice model. However, because of the limited number of trees used the route choice proportions could only take one of the following values: 0, 1/3, 2/3 and 1. The following comments can be made regarding the performance of the METWO2 model under these conditions.

(i) The results depicted in Tables 8.7 and 8.8 are compared to but no better than those obtained using all-or-nothing route choice models. In some cases a set of indicators shows one route choice model to generate a better estimated matrix but the situation is reversed in other tests.

(ii) There seems to be little difference in the performance of the ME2 model with 10 per cent and 30 per cent spread in Burrell's trees. According to some indicators 30 per cent spread route choice produces inferior results to the use of 10 per cent spread only; but the situation is reversed for other indices.

Table 8.7: Observed vs estimated trip matrix obtained using average of 3 Burrell trees, 10 per cent spread route choice, 19 October data

Indicator	10% spread: No centroids counted: 39 zones	10% spread: No centroids counted: condensed	10% spread: Centroids counted: 39 zones	10% spread: Centroids counted: condensed	30% spread 39 zones: no centroids counted	30% spread 39 zones: centroids counted
Mean Absolute Difference MAE	1.3	3.2	1.1	2.6	1.2	1.1
Relative MAE %	101	70	85	57	95	87
Coefficient of Determination R^2	0.29	0.36	0.51	0.48	0.40	0.46
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.31	0.36	0.43	0.53	0.34	0.39
RMSE	2.7	5.7	2.1	4.1	2.3	2.3
%RMSE	221	125	167	89	187	183
Phi measure Range:						
0 - 2	518	122	512	109	528	531
3 - 5	411	179	430	166	388	382
Over 5	838	858	694	750	869	602
All	1767	1160	1630	1025	1785	1514
NPhi (Phi/T)	0.96	0.63	0.87	0.65	0.97	0.83
Iterations	15	15	20	20	13	26
CPU time (secs)	6.63	6.63	9.55	9.55	6.69	12.58

Table 8.8: Observed vs estimated trip matrix obtained using average Burrell trees, 4 days flow

Indicator	10% spread: No centroids counted: 39 zones	10% spread: No centroids counted: condensed	10% spread: Centroids counted: 39 zones	10% spread: Centroids counted: condensed	30% spread 39 zones: no centroids counted	30% spread 39 zones: centroids counted
Mean Absolute Difference MAE	4.3	11.8	3.7	9.7	4.1	3.9
Relative MAE %	87	64	74	53	83	78
Coefficient of Determination R^2	0.34	0.40	0.55	0.59	0.45	0.50
Coefficient of Determination for $(T_{ij})^{½}$ SR^2	0.36	0.40	0.47	0.54	0.39	0.42
RMSE	9.4	19.8	7.2	14.6	8.1	8.0
%RMSE	189	107	144	79	162	160
Phi measure Range:						
0 - 2	511	81	462	68	542	504
3 - 5	694	145	724	144	750	755
over 5	5135	4149	4801	3835	5608	4674
All	6340	4376	5987	4047	6899	5933
NPhi (Phi/T)	0.86	0.59	0.81	0.55	0.93	0.80
Iterations	15	15	16	16	12	27
CPU time (secs)	6.8	6.8	7.62	7.62	6.55	12.79

Table 8.9: Summary of results of ME2 model with different route choice models. 19 October and centroid connectors counted data base.

Indicator	Observed routes	All-or-nothing	Burrell's 10 %	Burrell's 30 %	19-21 October observed variations
Mean Absolute Difference MAE	0.8	1.1	1.1	1.1	0.9
Relative MAE%	69	86	85	87	75
Coefficient of Determination R^2	0.66	0.50	0.51	0.46	0.61
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.66	0.42	0.43	0.39	0.49
RMSE	1.7	2.1	2.1	2.3	1.8
%RMSE	134	169	167	183	148
Phi measure Range:					
0 - 2	394	512	512	531	158
3 - 5	299	426	430	382	234
Over 5	526	671	694	602	499
All	1219	1609	1630	1514	891
NPhi (Phi/T)	0.66	0.88	0.87	0.83	0.49
Iterations	26	24	20	26	--
CPU time (secs)	4.79	3.56	9.55	12.58	--

(iii) These tests could also be considered representative of a sensitivity analysis of errors in route cost evaluation.

(iv) Although Burrell's approach (and presumably those of other stochastic assignment models) can reproduce some of the spread of trips over the network, it does not seem to capture it with enough consistency and accuracy to improve significantly the estimated trip matrix.

(v) The greatest contribution to the total error originates, for all the route choice proportions used, from cells containing 5 trips or less. This is certainly related to the sparseness of the sampled or observed matrices since the average cell value is 1.25 trips.

(vi) The use of the ME2 model with Burrell's route choie proportions requires considerably more computer time than all-or-nothing but it was still within reasonable limits. The quality of the results, however, do not seem to warrant these extra resources.

8.5 SENSITIVITY ANALYSIS

8.5.1 A strategy for sensitivity tests

It is important to have an idea of how the results obtained so far would be affected by errors and limitations in the input data. The following questions seem to be relevant in this context.

(a) How would the results be affected by errors in the in the traffic counts themselves?

Although the sampled Reading trip matrices may indeed contain errors, the corresponding traffic counts are completely consistent with them. In essence, these traffic counts are error free in relation to the matrices.

(b) How would the performance of the ME2 model be affected by a limited set of traffic counts?

It is rather unlikely that practical applications of the ME2 model will use a full set of counts.

(c) For an incomplete set of counts, how would results be affected if new counts were added to the data set?

It is also interesting to test the rules put forward in Section 5.3 for choosing counting links.

(d) How would results be affected by errors in link cost estimation and in the route choice model assumed?

This question has already been investigated with the use of all-or-nothing and Burrell's randomised cost route choice models. Furthermore, the next chapter will explore the scope for using capacity restrained route choice models.

Some of these questions might be attempted using an analytical treatment to trace, under suitable assumptions, input errors through the errors in the model output. This proved to be a non-trivial affair as the author found the relationships to be analytically intractable. Furthermore, a direct comparison with the results already obtained was highly desirable. These two considerations suggested that a simulation approach would yield interesting results of practical importance. This approach makes use of the models already developed but uses them with input data subject to known and controllable error levels.

A second issue is to plan the tests so that an efficient use is made of computer resources. It has already been shown that the model results are roughly comparable for any of the sampled days, are slightly better for the 4-day data base and are generally better for the coarse (20 centroids) zoning system. It was decided then to concentrate efforts on a particular data base so that a greater variety of tests could be performed. The data base chosen is Tuesday 19 October, fine zoning system (39 centroids), all links counted (centroid connectors and real links), no prior information.

8.5.2 Errors in the traffic counts

It is well known that it is almost impossible to have error free traffic counts. It has already been shown in Chapter 7 that the counts vary from one day to the next. One source of errors then would be the fact that counts taken on different occasions will often be used in the estimation process. Other sources are simply counting and data transcription errors. There is some evidence to support the commonly made assumption that link flows are Poisson distributed. In this case the standard deviation σ is related to the mean μ by

$$\sigma = \sqrt{\mu} \tag{8.2}$$

For large values of μ the Poisson distribution can be approximated by the normal distribution with equal mean and variance.

In our tests a normal distribution was used to randomly simulate 'errors' on the observed link flows at three different levels for the ratio $\sigma/\sqrt{\mu}$: 0.5, 1.0 and 2.0. (The error-free values correspond, of course to $\sigma/\sqrt{\mu} = 0$.) After these errors were introduced the program METWO1 was used to remove any link flow inconsistencies thus generated. The program generates flow continuity conditions (80 when trip ends are counted, otherwise only 41) and removed any inconsistencies in 3 to 4 iterations taking less than 1 second of CPU time.

The resulting flows were used to estimate the 19 October O-D matrix using both observed and all-or-nothing route choice proportions. The results for real links and centroid connectors counted are depicted in Table 8.10.

It can be seen that for values of $\sigma/\sqrt{\mu}$ up to 1.0 the loss in accuracy of the estimated trip matrix does not seem to be very large.

This can be partially explained by the use of the METWO1 program to remove link flow discontinuities. In effect, while removing discontinuities the program also generates an improved estimation of the 'true' flow levels. Of course, with a reduced set of counts there will be fewer opportunities to use this facility.

It was also observed that in a few cases the number of iterations and computer times rose considerably. This seem to be due to possible inconsistencies between route choice model and traffic counts. The ME2 model with observed route choice seem to be more sensitive to this effect.

Table 8.10: Observed vs estimated trip matrix obtained using counts with errors

Indicator:	Observed route choice				All-or-nothing route choice			
σ/μ:	0	0.5	1.0	2.0	0	0.5	1.0	2.0
Mean Absolute Difference MAE	0.8	0.9	1.0	1.3	1.1	1.1	1.1	1.2
Relative MAE %	69	72	81	101	86	87	91	95
Coefficient of Determination R^2	0.66	0.61	0.53	0.35	0.50	0.49	0.46	0.40
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.66	0.61	0.50	0.30	0.42	0.41	0.37	0.31
RMSE	1.7	1.8	2.0	2.6	2.1	2.1	2.3	2.5
%RMSE	134	142	159	207	169	173	183	199
Phi measure Range:								
0 - 2	394	437	466	410	512	516	509	508
3 - 5	299	345	368	387	426	446	475	460
Over 5	526	523	617	834	671	697	768	852
All	1219	1304	1451	1631	1609	1659	1752	1820
NPhi (Phi/T)	0.66	0.71	0.79	0.88	0.88	0.90	0.95	0.99
Iterations	26	57	51	70	24	23	46	19
CPU time (secs)	4.79	9.53	10.44	16.70	3.56	3.17	6.47	2.74

8.5.3 Tests with incomplete set of counts

In essence almost any set of counts is incomplete. For example, in our Reading data base we could have coded every turning movement and obtained more than twice the number of counts. Nevertheless, it is interesting to see how the results of the model deteriorate as fewer counts are available. In order to do this, a modified version of the METWO2 program was written in which a randomly selected sub-set of link counts was used. The percentage selected ranged from 10 per cent to 100 per cent.

Tables 8.11a and b show the results under these conditions for both observed and all-or-nothing route choice proportions. The following comments can be made.

(i) Results with 10 per cent and 25 per cent of the counts seem to be poor for both route choice proportions. The improvement obtained by counting 75 per cent instead of only 50 per cent of the links is less substantial than that obtained by increasing the sample from 25 per cent to 50 per cent for both cases.

(ii) For all-or-nothing route choice models a larger number of counts seem to be required. Counting on 75 per cent of the links seem to produce in this case results comparable to the application of the observed route choice proportions data over only 50 per cent of the links.

(iii) As expected CPU time increases with the number of counted links.

Table 8.11a: Observed vs estimated trip matrix obtained using incomplete set of counts, 19 October

Indicator	Observed routes					All-or-nothing				
	10%	25%	50%	75%	100%	10%	25%	50%	75%	100%
Mean Absolute Difference MAE	1.4	1.3	1.0	0.9	0.8	1.6	1.4	1.2	1.1	1.1
Relative MAE %	115	105	83	74	69	130	115	97	87	86
Coefficient of Determination R^2	0.07	0.30	0.59	0.62	0.66	0.06	0.14	0.34	0.53	0.50
Coefficient of Determination for $(T_{ij})^{½}$ SR^2	0.17	0.41	0.55	0.62	0.66	0.06	0.14	0.32	0.42	0.42
RMSE	2.7	2.4	1.8	1.7	1.7	2.8	2.6	2.3	2.0	2.1
%RMSE	220	192	147	140	134	222	213	188	160	169
Phi measure Range:										
0 - 2	262	321	368	386	394	234	363	483	506	512
3 - 5	447	291	295	297	299	424	483	410	407	426
Over 5	1782	1073	576	561	526	1589	1442	937	621	671
All	2489	1685	1239	1245	1220	2248	2288	1830	1534	1609
NPhi (Phi/T)	1.36	0.92	0.68	0.68	0.67	1.23	1.24	1.00	0.84	0.88
Iterations	17	20	29	43	27	8	15	17	17	24
CPU time (secs)	0.34	1.03	2.87	5.61	4.78	1.06	1.81	2.21	2.46	3.30

Table 8.11b: Observed vs estimated trip matrix obtained using incomplete set of counts, 4-day flows

Indicator	Observed routes					All-or-nothing				
	10%	25%	50%	75%	100%	10%	25%	50%	75%	100%
Mean Absolute Difference MAE	4.8	4.3	3.4	3.2	3.0	5.9	5.1	4.1	3.5	3.7
Relative MAE %	97	87	66	64	61	119	103	83	75	75
Coefficient of Determination R^2	0.01	0.17	0.59	0.66	0.70	0.07	0.16	0.43	0.55	0.54
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.06	0.33	0.56	0.60	0.63	0.06	0.17	0.40	0.48	0.46
RMSE	10.6	9.4	6.1	5.9	5.6	9.8	9.4	7.8	7.1	7.3
%RMSE	212	188	130	118	112	197	188	157	142	146
Phi measure Range:										
0 - 2	429	521	452	453	467	702	543	491	499	476
3 - 5	760	621	617	571	545	415	490	644	689	740
Over 5	14172	7973	4058	3566	3339	7364	7346	5431	4444	4983
All	15361	9115	5127	4590	4351	8482	8379	6567	5632	6198
NPhi (Phi/T)	2.08	1.23	0.69	0.62	0.60	1.15	1.13	0.89	0.76	0.84
Iterations	12	17	28	33	33	8	15	23	53	28
CPU time (secs)	0.44	1.89	5.84	8.73	10.32	1.02	1.89	3.01	6.76	3.97

8.5.4 Selected link sampling

Sampling links to be counted at random does not seem to be the best of strategies. A better approach would be to select those links most likely to produce an improved trip matrix. A simple method for making this choice was put forward in Section 5.3. The technique involves calculating for each candidate link the indicator

$$\frac{1}{2}\,\frac{(v_{\ell m} - V_{\ell m})^2}{V_{\ell m}} \qquad (8.3)$$

where $v_{\ell m}$ is the modelled flow and $V_{\ell m}$ is a (prior) estimate of the count for the candidate link.

A modified version of the METWO2 program was used to choose new links to be included. Its use in this context represents in effect ideal conditions as the value taken for the prior estimate in the calculation of (8.3) is in fact the <u>true</u> value. In practice this estimation will be subject to inevitable errors making the process less reliable than it appears here.

In each case, after deciding which link counts to incorporate next the program estimated the trip matrix using no prior information, that is all t_{ij} with the same initial value. Table 8.12 presents the results using all-or-nothing route choice proportions.

The following comments can be made.

(i) It can be seen that the use of the indicator (8.3) in the selection of links to counts appears to produce very good results. The selection of 24 new counts to pass from a sampling ratio of 0.25 to 0.35 immediately reduces the error measures to levels comparable to those obtained with 100 per cent of the counts. The exceptional results obtained for a 50 per cent selective counting can only be considered as achieved by chance as they are better than those obtained with 100 per cent of the links.

Table 8.12: Observed vs estimated trip matrix obtained using selected link counts, all-or-nothing route choice

Indicator	Percentage of links used 25%	35%	50%	75%	100%
Mean Absolute Difference MAE	1.4	1.1	1.0	1.1	1.1
Relative MAE %	114	89	84	86	86
Coefficient of Determination R^2	0.14	0.48	0.53	0.49	0.50
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.14	0.39	0.44	0.42	0.42
RMSE	2.6	2.2	2.0	2.1	2.1
%RMSE	214	173	158	169	169
Phi measure Range:					
0 - 2	363	506	488	505	512
3 - 5	483	428	418	427	426
Over 5	1442	710	693	673	671
All	2287	1645	1599	1604	1609
NPhi (Phi/T)	1.27	0.89	0.85	0.85	0.86
Iterations	15	18	19	30	24
CPU time (secs)	1.91	2.25	2.46	3.98	3.56

(ii) An inspection of the disaggregated Phi indicators shows that most of the improvement in accuracy is generated for cells with more than 5 trips. This is a desirable result.

(iii) Even if the method cannot be expected to perform so well in practice (with poorer estimates of the expected count) it seems worthwhile devoting resources to its application. It seems to be the case that much of the structure of a trip matrix is somehow reflected in a relatively small number of key link volumes whose indentification would be worth the effort.

It is possible to use the matrix estimated, say, with 25 per cent of the counts as a prior estimate when running the ME2 model with more counts. This was tried and it was found that it produced neither an improvement in the resulting matrix nor a reduction in the number of iterations. The only 'fresh' information available is, in effect, contained in the latest counts which have been chosen precisely because they add new information modifying the prior estimate most. Only a prior trip matrix estimated by independent means (outdated matrix, other large study) would introduce fresh information to the process.

8.6 LOADING AN ESTIMATED TRIP MATRIX

A very practical question is to ask whether the use of an estimated rather than an observed trip matrix will result in very different forecasts of the impact of network changes. Origin-destination matrices are seldom sought for their own sake; it is the prediction of the likely results of modifying a network which requires them. The question above can often be quantified in terms of the resulting network loadings. If the network is modified and an estimated (ME2) matrix is loaded onto it, will it produce flows very different from those obtained from loading an observed trip matrix?

It is clearly difficult to answer this question in absolute terms as the results would depend on the particular characteristics of the network itself (unless estimated and observed matrices are equivalent). But one can at least try to compare the results on plausible changes to the Reading network.

In this case a second, modified version of the Reading data was produced incorporating the planned extension of the inner ring road from Southampton Street roundabout to (a future) Forbury Road roundabout. This change is likely to have important effects in the route choice pattern as it essentially replaces part of an elaborate one-way system south-east of the city centre. The proposed improvement is depicted in Figure 8.1. Taking advantage of these new fast links a couple of minor roads could revert to two-way operation. The full coding of this modified Reading network, labelled Network B, appears in Appendix A6.

The observed tripmatrices for 19 October and 4 days data base were loaded, all-or-nothing, to this new network and the resulting flows recorded. Different estimates of these matrices were also loaded and the resulting flows compared with the previous loadings obtained from the observed matrices. The results of these tests are summarised in Table 8.13.

It can be seen that the resulting performance measures at the flow level are well within the range of daily variations of the same (c.f. Table 7.4). The link loadings are in fact remarkably similar even for matrices which were not considered particularly good, such as the 19 October, all-or-nothing, real links counted matrix.

The estimated trip tables seem to preserve enough of the characteristics of the underlying observed matrices to produce similar indicators when loaded to a modified network.

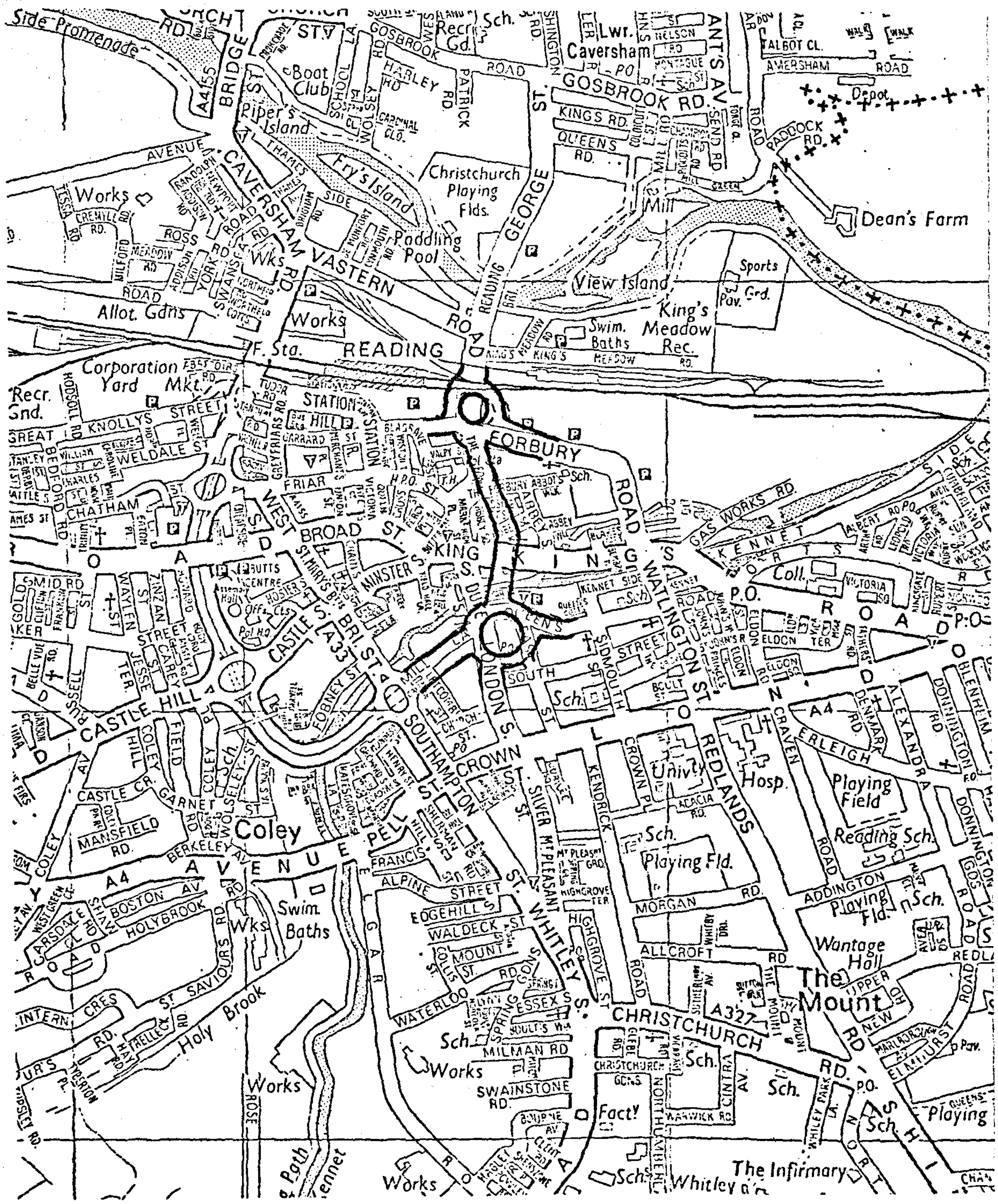


Figure 8.1: The modified Reading network (B)

Table 8.13: Comparison of all-or-nothing loaded flows on Network B from observed and estimated matrices

Indicator	Estimated matrix, Observed routes: 19 Oct. trip end c	Estimated matrix, Observed routes: 4 days trip end c	All-or-nothing route choice: 19 Oct. real links o	All-or-nothing route choice: 19 Oct. trip end c	All-or-nothing route choice: 4 days real links o	All-or-nothing route choice: 4 days trip end c
Mean Absolute Difference MAE	8.1	22.6	10.6	9.8	42.4	36.9
Relative MAE %	7	6	11	11	11	10
Coefficient of Determination R^2	0.98	0.99	0.96	0.97	0.96	0.97
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.98	0.99	0.97	0.97	0.97	0.97
RMSE	11.6	33.4	15.7	14.4	63.5	56.3
%RMSE	12	9	17	15	17	15
Phi measure Range:						
0 - 10	14	0	19	18	0	0
11 - 50	194	65	179	260	81	87
Over 50	1119	3521	1494	1363	6639	6146
All	1326	3586	1692	1640	6720	6233
NPhi (Phi/T)	0.087	0.058	0.111	0.107	0.109	0.101

8.7 CONCLUSIONS

It is possible to draw the following conclusions.

(i) The trip matrices estimated using the ME2 model are not very close to the observed ones, even when perfect knowledge about route choice and error free counts are made available. This is not entirely surprising as the ratio of unknowns to observations is around 9.4 in the best of cases.

(ii) Under favourable conditions (observed route choice proportions and/or trip end counts) the estimated trip matrices are within the range of daily variations of the sampled trip matrices obtained in the Reading survey.

(iii) The ME2 model seems to perform better with the aggregated 4-day data base. The entropy maximising framework tends to assign a low probability to matrices having extreme cell values and these extreme values are less likely to occur in larger samples or under 'average' conditions.

(iv) Errors in the assumed route choice model, exemplified by the use of a naive all-or-nothing route choice model, deteriorates the performance of the model but to a lesser extent than one might expect. It is possible that the one-way system in Reading enables all-or-nothing route choice to capture most of the paths used by vehicles.

(v) Stochastic route choice models, at least as exemplified by Burrell's, do not seem to improve the estimated trip matrix enough to compensate the extra cost and work involved.

(vi) In this particular application the ME2 model estimations benefited from the use of counts on at least 50 per cent of the links for good route choice models and at least 75 per cent of the links for all-or-nothing. Of course this cannot be considered a general recommendation as it depends on the size and level of detail of the network. However, it is worth stressing that no prior information about the trip matrix was used and that one would expect better results when an independent estimate of the trip table is employed.

(vii) There is a very good case for devoting resources to the problem of selecting additional links to count. Local experience and the use of the principles outlined in Section 5.3 and demonstrated in 8.5.4 can produce very cost-effective results.

(viii) Even when the estimated trip matrices are not very close to the sampled ones they seem to pick up enough of their characteristics to warrant their use for some forecasting tasks. This was illustrated by the similar link flows obtained by loading sampled and estimated trip matrices onto a modified network in Reading. Of course, one must include the rider that this applies only to conditions similar to those tested in Reading.

(ix) The ME2 model does not seem to be very demanding in terms of CPU time for networks of the size of Reading (39 zones, 80 nodes, 159 one-way links).

CHAPTER 9
EXTENSION FOR CAPACITY RESTRAINT ASSIGNMENT

9.1 INTRODUCTION

The ME2 model developed and tested in the preceding chapters is based on the assumption that it is possible to obtain the route choice proportions $[p_{ij}^{\ell m}]$ independently from the O-D estimation process. This is only possible for situations in which proportional assignment techniques (Robillard, 1975) are considered to be sufficiently realistic.

Wherever congestion plays an important role in route choice this assumption becomes questionable. The route choice proportions and the trip matrix become interdependent under these conditions. This chapter considers the theoretical and practical problems involved in adapting the ME2 model to congested conditions. Section 9.2 reviews capacity constraint assignment problems and the currently preferred technique in this area: equilibrium assignment. Both the mathematical properties of this method and its implications for the O-D estimation problem are outlined. The next section discusses the problem of linking equilibrium assignment and O-D estimation from traffic counts and proposes a heuristic method.

Section 9.4 reports on tests, within a more conventional equilibrium assignment framework, using Reading data. Finally Section 9.5 draws some conclusions regarding the use of the ME2 model in this context.

9.2 THEORETICAL CONSIDERATIONS

The role of route choice models in the estimation of trip matrices from traffic counts has already been discussed in Section 2.4 of this work. In the case of congested networks equilibrium assignment methods have a number of advantages over alternative techniques; in particular economic use of computer resources, guarantee of convergence under certain conditions and stability of the convergence process towards the optimum solution. The main properties of equilibrium assignment have already been discussed in Subsection 2.4.4 and accordingly only those directly relevant to our problem are detailed below.

The problem is to find, for a fixed trip matrix $[T_{ij}]$, a set of link flows and corresponding route costs in a network that satisfy Wardrop's First Principle (Wardrop, 1952) that all routes used for an O-D pair have equal costs and that no unused route has less cost. This 'user-optimised' requirement implies that all drivers perceive route and link costs in the same manner. As there is in general more than one route between an O-D pair we introduce the indicator r to identify each. Thus T_{ijr} stands for the number of trips between i and j via route r . A solution satisfies Wardrop's principle if

$$\left.\begin{aligned} c_{ijr} &= c^*_{ij} && \text{for all } T_{ijr} > 0 \\ c_{ijr} &\geq c^*_{ij} && \text{for all } T_{ijr} = 0 \end{aligned}\right\} \qquad (9.1)$$

The second requirement is the introduction of the capacity restraint element by relating the cost of travel on a link to the flow on that link by

$$c_{\ell m} = c_{\ell m}(V_{\ell m})$$

where in general $c_{\ell m}$ is a non-decreasing function of $V_{\ell m}$. The problem of finding a feasible solution which satisfies Wardrop's First Principle is equivalent to solving (Beckman et al, 1956):

$$\text{Minimise} \quad Z = \sum_{\ell m} \int_0^{V_{\ell m}} c_{\ell m}(x)dx \qquad (9.2)$$

subject to

$$\sum_{ij} T_{ijr} \cdot \delta^{\ell m}_{ijr} - V_{\ell m} = 0 \qquad (9.3)$$

$$\sum_{r} T_{ijr} - T_{ij} = 0 \qquad (9.4)$$

$$T_{ijr} \geq 0 \qquad (9.5)$$

where $\delta^{\ell m}_{ijr}$ is one if route r between i and j uses link ℓm and zero otherwise.

The proof that minimising Z subject to (9.3 - 9.5) produces a Wardrop equilibrium solution is well known, see for example Beckman et al (1956), Potts and Oliver (1972), Evans (1976) and Van Vliet (1979c). The reason is that the necessary and sufficient (Kuhn-Tucker) conditions for a solution to (9.3 - 9.5) to be minimum of (9.2) results in flows $V_{\ell m}$ and costs c_{ij} which satisfy Wardrop's conditions (9.1). This solution is unique in the flows $[V_{\ell m}]$ and costs $[c_{ijr}]$ and $[c_{ij}]$ but in general not unique in the path flows $[T_{ijr}]$.

A relatively simple algorithm for minimising (9.2) based on the well known iterative assignment method is as follows (Van Vliet and Dow, 1979).

<u>Iteration 1</u>

(i) Set all link costs to some predetermined value, usually free flow costs.

(ii) Build minimum cost trees and assign all T_{ij} to them (all-or-nothing) to produce a set of 'auxiliary' link flows $[F_{\ell m}^{(0)}]$. Set the current 'main flows' $V_{\ell m}^{(n)} = F_{\ell m}^{(0)}$ where $n = 1$.

<u>Iteration n+1</u>

(iii) Change the link costs according to the new flow levels $V_{\ell m}^{(n)}$

$$c_{\ell m}^{(n)} = c_{\ell m}(V_{\ell m}^{(n)}) \; .$$

(iv) Build minimum cost trees using $c_{\ell m}^{(n)}$ and assign all T_{ij} to them (all-or-nothing) to produce a set of auxiliary flows $[F_{\ell m}^{(n)}]$.

(v) Generate an 'improved' set of 'main flows' $V_{\ell m}^{(n+1)}$ as a linear combination of the old and auxiliary flows

$$V_{\ell m}^{(n+1)} = (1-\lambda)V_{\ell m}^{(n)} + \lambda F_{\ell m}^{(n)} \qquad (9.6)$$

where $0 \leq \lambda \leq 1$ and choose λ so as to minimise Z .

(vi) Increment n by 1 and return to step (iii) unless equilibrium (or other related) conditions have been reached.

The key element of this algorithm is the choice of λ . It should be noted that the algorithm does not make direct use of the path flows $[T_{ijr}]$ and that the choice of λ

only refers to combination of link flows and not of path flows. It may be argued that by implication this method essentially combines path flows (as generated by the auxiliary flows) and that the final solution has a well-defined set of path flows associated with it. The only problem is of course that these path flows are not unique, other combinations might have led to the same solution. Under equilibrium conditions the path flows are essentially undefined and this will pose a major difficulty in trying to extend ME2, which is precisely based on identifiable route choice proportions $[p_{ij}^{\ell m}]$, to these conditions.

9.3 AN ITERATIVE METHOD

The problem of estimating an O-D matric from link flows under equilibrium assignment conditions can be described as follows:

> One would like to use a set of route choice proportions $[p_{ij}^{\ell m}]$ with the observed link flows to estimate a trip matrix $[T_{ij}]$ (using ME2) such that when loaded to equilibrium it reproduces both the observed flows <u>and</u> the original route choice proportions $[p_{ij}^{\ell m}]$ used to estimate it.

This description suggests the following heuristic algorithm for finding self-consistent trip matrices.

(i) Assign (using equilibrium methods) a base-year (prior) matrix $[t_{ij}]$ to obtain a first estimate of $[p_{ij}^{\ell m}]^{(1)}$.
Set the cycle (or iteration) counter c to 1.

(ii) Estimate $[\hat{T}_{ij}]^c$ using $[p_{ij}^{\ell m}]^c$ and the observed flows $[\hat{V}_{\ell m}]$.

(iii) Assign $[T_{ij}]^c$ (equilibrium) to obtain new $[p_{ij}^{\ell m}]^c$.

(iv) Increment c by 1. Return to step (ii) unless the changes in $[p_{ij}^{\ell m}]$ or $[T_{ij}]^c$ have been sufficiently small.

There are two main difficulties with this scheme. The first is that convergence is not guaranteed. The second is the identification of the route choice proportions $[p_{ij}^{\ell m}]$ which, as mentioned before, are in general not unique under equilibrium conditions. An ad hoc solution to this second problem is to derive the route choice proportions $[p_{ij}^{\ell m}]$ from the trees and flow combination parameters λ in steps (iv) and (v) in the equilibrium assignment algorithm described in the previous section. It must be stressed that this is only an ad hoc device chosen to explore ways of extending the ME2 model to equilibrium assignment conditions.

A more rigorous approach to the problem is suggested in Section 12.2 for further research.

9.4 TESTS WITH READING DATA

It was decided to use the Reading data base to test this heuristic algorithm, again using assignment software from the TRADVV suite. These programs represent the state of the art in practical equilibrium assignment software provided one accepts that the influence of flow levels on link costs can be represented through the use of standard speed flow relationships.

It is of interest to first test how well the TRADVV equilibrium assignment programs reproduce the observed flows in the Reading area. In order to do this the observed O-D matrix for Tuesday 19 October was loaded all-or-nothing and to equilibrium and the resulting flows (adequately scaled) were compared with the observed ones. The results of these tests are summarised in Table 9.1. It can be seen that equilibrium assignment produces link flows which are closer to the observed ones than all-or-nothing assignment. These results are comparable to and even slightly better that those obtained by other researchers, see for example Van Vliet and Dow (1979). It is also of interest to note that all-or-nothing assignment also produces fairly reasonable flow levels suggesting that in this case it is not a bad approximation.

Table 9.1: Observed vs loaded flows (from 19 October O-D matrix)

Indicator	Loading technique: All-or-nothing	Loading technique: Equilibrium assignment
Mean Absolute Difference MAE	12.8	9.8
Relative MAE %	13	10
Coefficient of Determination R^2	0.92	0.96
Coefficient of Determination for $(V_{\ell m})^{\frac{1}{2}}$ SR^2	0.94	0.96
RMSE	18.8	16.0
% RMSE	20	16
Phi	1914	1740

The flow chart in Figure 9.1 describes the way in which these tests were conducted. The process starts in blocks 1 and 2 with the estimation of the first trip matrix using all-or-nothing $[p_{ij}^{\ell m}]^0$. This estimated matrix is then loaded to equilibrium through a series of N iterations (blocks 4-5 and 6) with an improved estimation of the equilibrium route choice proportions in block 5. at each iteration. After a sufficient number of iterations (3 to 5 in practice) the loading cycle is stopped and a new trip matrix is estimated using the last combined route choice proportions (block 7). If more cycles are required this new matrix is used to initiate a new set of loading iterations, otherwise results are printed out (blocks 8-9).

In the case of the tests with Reading data it was possible to monitor at each step how close an intermediate matrix estimated at that level would be to the observed one. This is denoted in the segmented block 10. Of course, this is not possible when one is estimating an unknown trip matrix.

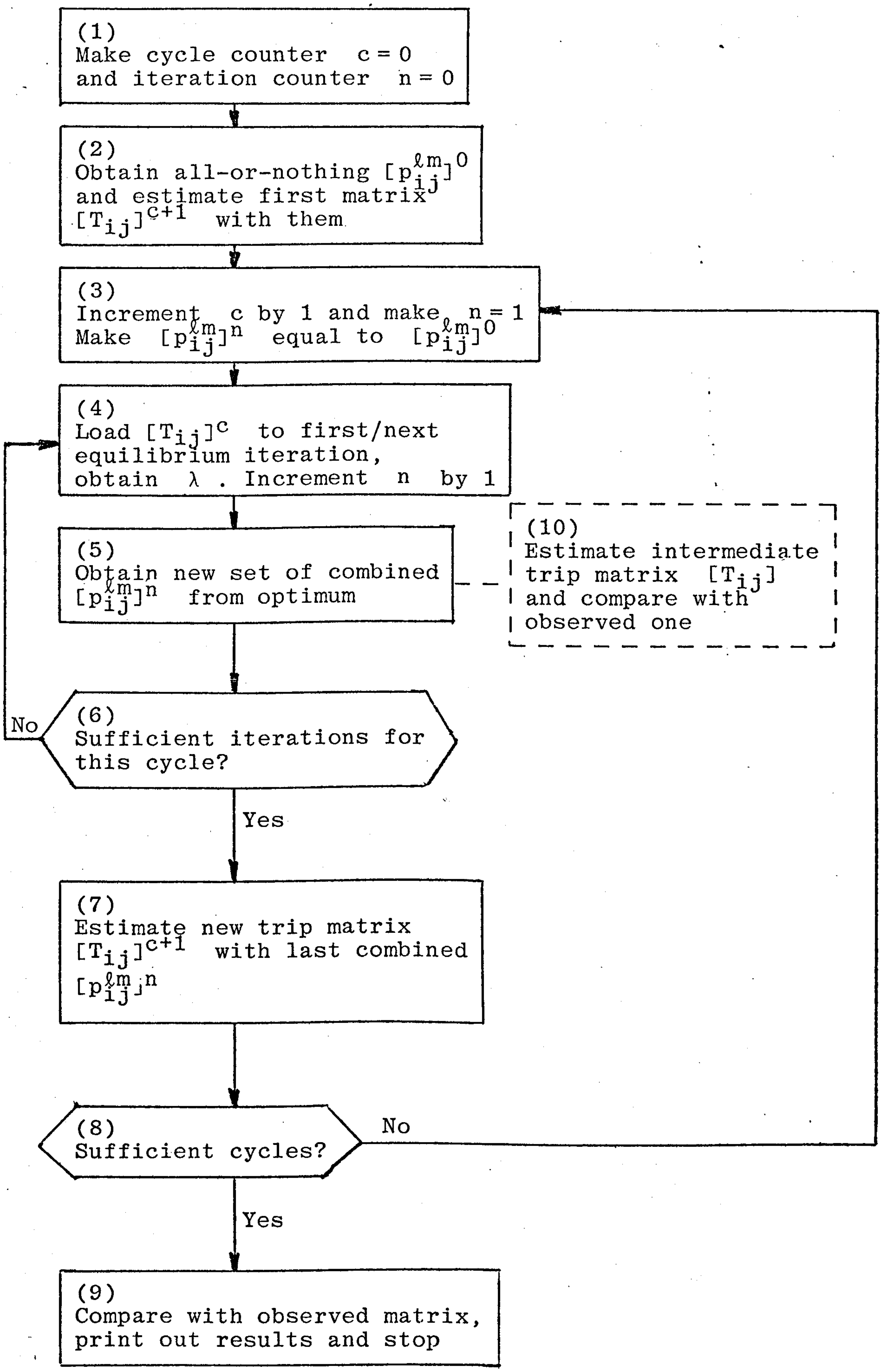


Figure 9.1: <u>Flow-chart for iterative O-D estimation method tests</u>

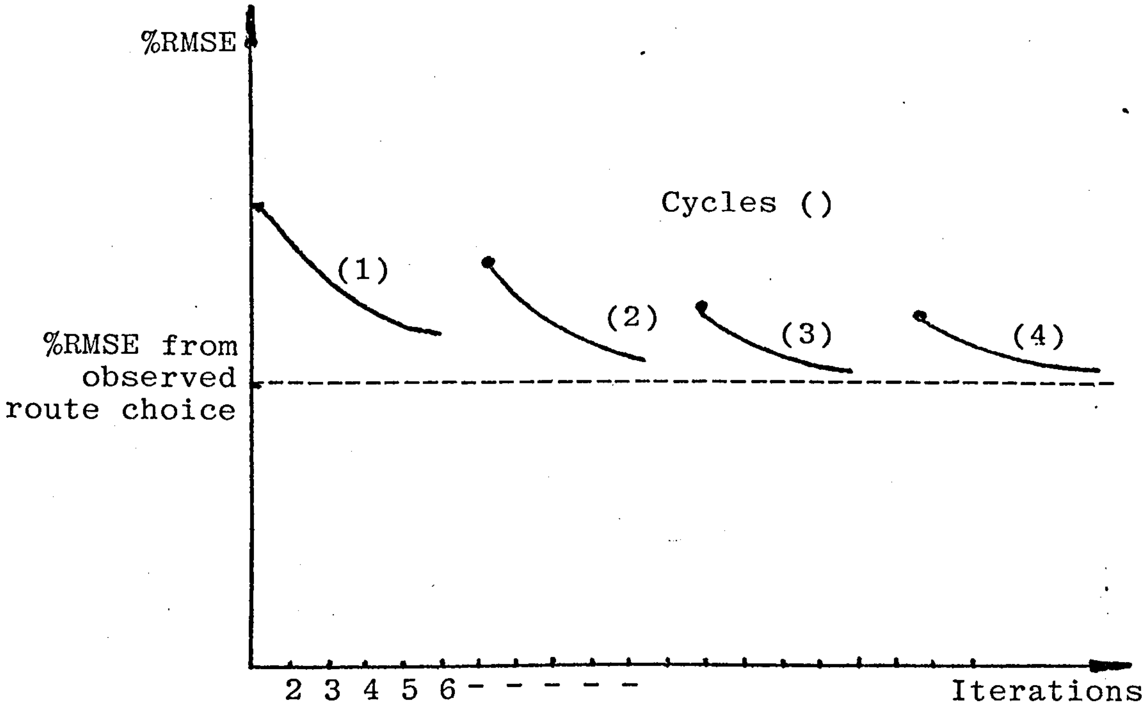


Figure 9.2: Idealised evolution of iterative O-D estimation method

It is interesting to speculate on the evolution one would like to see for a particular error measure, say %RMSE throughout these cycles. A desirable pattern is depicted in Figure 9.2.

In other words, one would expect the accuracy of the estimated matrix to improve as the number of iterations increases within each cycle and also for the matrix estimated at the end of each cycle. If the matrix estimated at the end of a cycle is not closer to the observed one than that estimated at the end of the previous cycle the whole process can be said to have converged.

The tests were carried out using Tuesday 19 October data base with the extended set of counts (trip end counts included). The two hour flow levels (16.10 to 18.10) were appropriately scaled when used to update costs with the cost-flow relationships. The results of these tests are summarised in Table 9.2. In this table the accumulated CPU times include not only ME2 iterations but also tree building, extracting route choice proportions and loading sequences for the matrices.

Table 9.2: Tests with iterative technique for O-D estimation

Indicator	1st cycle (iteration)					2nd cycle			3rd cycle		
	All-or-nothing	2 $p_{ij}^{\ell m}$ combined	3 $p_{ij}^{\ell m}$ combined	4 $p_{ij}^{\ell m}$ combined	5 $p_{ij}^{\ell m}$ combined	2 $p_{ij}^{\ell m}$ combined	3 $p_{ij}^{\ell m}$ combined	4 $p_{ij}^{\ell m}$ combined	2 $p_{ij}^{\ell m}$ combined	3 $p_{ij}^{\ell m}$ combined	4 $p_{ij}^{\ell m}$ combined
Mean Absolute Difference MAE	1.1	1.1	1.1	1.1	1.1	1.0	1.0	1.1	1.0	1.0	1.0
Relative MAE %	86	88	85	86	86	84	84	88	84	85	84
Coefficient of Determination R^2	0.50	0.50	0.52	0.51	0.52	0.54	0.53	0.50	0.53	0.52	0.53
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.42	0.43	0.43	0.44	0.44	0.44	0.45	0.40	0.45	0.45	0.45
RMSE	2.1	2.1	2.0	2.0	2.0	2.0	2.0	2.1	2.0	2.0	2.0
%RMSE	169	167	163	164	162	162	160	179	160	161	160
Phi measure Range:											
0 - 2	512	511	537	509	531	520	509	531	508	530	512
3 - 5	426	423	407	408	414	420	400	461	399	418	401
Over 5	671	685	633	652	626	654	629	709	633	627	630
All	1609	1619	1577	1569	1572	1594	1538	1702	1539	1574	1543
NPhi (Phi/T)	0.88	0.88	0.86	0.86	0.86	0.87	0.84	0.92	0.84	0.86	0.84
Iterations	24	21	24	20	29	20	21	28	22	24	26
CPU time (secs)	3.76	5.8	6.52	5.63	7.34	10.41	10.30	14.76	9.82	12.02	11.47
Accumulated CPU time incl. $p_{ij}^{\ell m}$ production	24.12	52.02	84.59	199.91	157.19	199.74	241.60	291.36	331.17	374.69	416.35

The evolution of the indicator %RMSE throughout the sequence is depicted in Figure 9.3 presented graphically in Figure 9.2.

The following comments can be made at this stage.

(i) The iterative technique seems to produce an improved estimate of the sampled trip matrix albeit at a high cost in CPU time. This improvement is, however, not very large.

(ii) The greatest improvement seems to be produced during the second cycle. Additional iterations did not improve the estimated matrix much.

(iii) The technique is not completely 'well behaved', in the sense that after certain iterations the goodness of fit worsens, albeit only marginally. The expected evolution of the indicator %RMSE is not reproduced in Figure 9.3.

(iv) Although not apparent in Table 9.2, it was observed that on occasions (for example 1st cycle, 5 $p_{ij}^{\ell m}$ and 2nd cycle, 4 $p_{ij}^{\ell m}$) full convergence of the ME2 model was not achieved because the linear constraints were inconsistent. Iterations were stopped when it was not possible to reduce the errors in the observed counts rather than when all of them were less than 5 per cent. Apparently this method tends to generate certain route choice proportions which are not fully compatible with the observed link flows.

(v) On the whole it cannot be said that this iterative method is entirely satisfactory. Further research would be necessary to improve it or develop alternative approaches. Some suggestions in this direction are given in the next chapter.

9.5 USE OF THE MODEL WITH SATURN

SATURN is a simulation-assignment model for the evaluation of traffic management schemes developed at the Institute for Transport Studies, University of Leeds (Bolland et al, 1977, 1979). The SATURN model treats junctions in great detail providing a better representation of the way in which equilibrium might be achieved in an urban area.

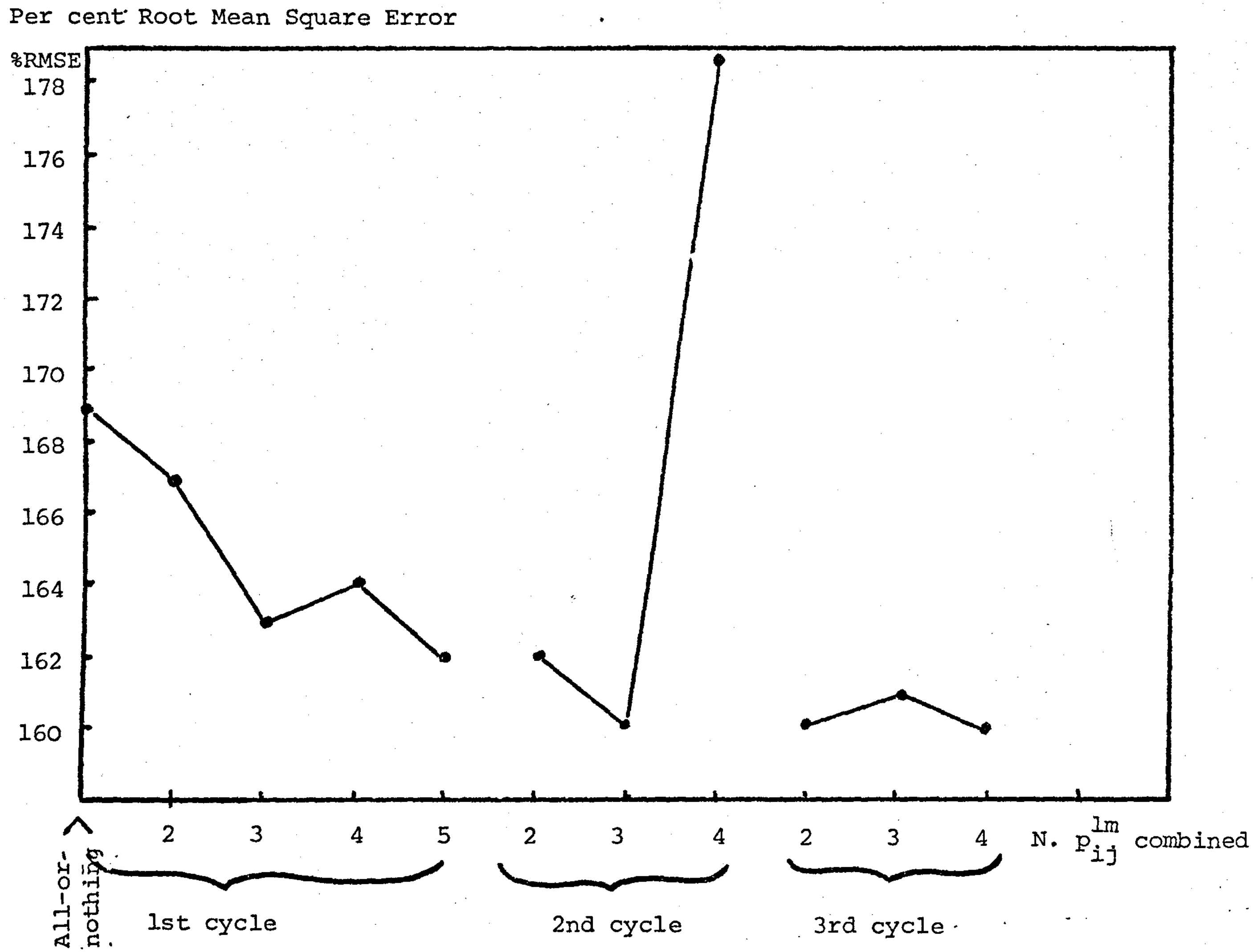


Figure 9.3: <u>Tests with iterative O-D estimation method, 19 October</u>

The algorithm described in this chapter to estimate a trip matrix from counts under congested conditions has been adapted for use with SATURN. The (adapted) ME2 model now forms part of the SATURN computer package and it has been used to update trip matrices using information contained in traffic counts (Hall, Van Vliet and Willumsen, 1980).

It was found that the use of the trip matrix updating program improved the capability of SATURN to replicate the observed flow patterns. However, as these results are part of a research project in which the author played only an auxiliary role no further report will be given here.

CHAPTER 10
EXTENSIONS FOR PATH FLOW ESTIMATION

10.1 INTRODUCTION

Extending the ME2 model to equilibrium assignment was not entirely succesful as reported in Chapter 9. One apparent difficulty with the approach seemed to be the generation of route choice proportions $[p_{ij}^{\ell m}]$ which, while consistent with capacity restraint assignment, occasionally prevent the system of linear Equations (10.1)

$$\sum_{ij} T_{ij} p_{ij}^{\ell m} - V_{\ell m} = 0 \qquad (10.1)$$

to have at least a feasible solution. It should be noted that under equilibrium assignment conditions path flows and the corresponding $[p_{ij}^{\ell m}]$ are in general not unique. One would like to be able to identify the $[p_{ij}^{\ell m}]$ which generate a consistent set of Equations (10.1). The method put forward in Chapter 9 was based in the optimum values of the parameter λ and although this produces optimum flow combinations it does not seem to generate, in general, entirely suitable route choice proportions.

10.2 A MODEL WITH PATH FLOW ESTIMATION

Consider first the problem of identifying the most likely path flows in a network with a fixed matrix loaded to equilibrium onto it. Link flows and costs are unique and known, but not path flows. This problem may be of some practical interest if one would like to identify the O-D pairs most likely to be affected by, for example, banning car traffic on a particular link. One may assume that R_{ij} feasible routes or paths have been identified for each O-D pair. The problem can be presented in the entropy maximising formalism as follows

$$\text{Maximise} \quad -\sum_{ijr} T_{ijr}(\log T_{ijr}/t_{ijr} - 1) \qquad (10.2)$$

subject to

$$T_{ij} - \sum_r T_{ijr} = 0 \qquad (10.3)$$

and $$\sum_{ijr} T_{ijr}\delta_{ijr}^{\ell m} - V_{\ell m} = 0 \qquad (10.4)$$

$$T_{ijr} \geq 0 \qquad (10.5)$$

where T_{ijr} are the trips from i to j using route r, t_{ijr} are the corresponding 'prior' estimates and $\delta_{ijr}^{\ell m}$ is 1 if T_{ijr} uses link ℓm and zero otherwise. The $V_{\ell m}$ are now not the observed flows but the (equilibrium) modelled flows.

This is a concave mathematical programme with linear constraints (10.3-10.5). Its solution can be found using Lagrangian methods to be

$$T_{ijr} = t_{ijr} X_{ij} \prod_{\ell m} X_{\ell m}^{\delta_{ijr}^{\ell m}} \qquad (10.6)$$

where

$$X_{ij} = e^{-\phi_{ij}}$$

$$X_{\ell m} = e^{\gamma_{\ell m}}$$

and ϕ_{ij} is the Lagrangian multiplier associated with Equation (10.3) and $\gamma_{\ell m}$ the multiplier associated with Equation (10.4).

This model can be solved using the same algorithms put forward for ME2. A sensible 'a priori' estimate for t_{ijr} is

$$t_{ijr} = T_{ij} / R_{ij} \qquad (10.7)$$

as the O-D matrix $[T_{ij}]$ is given. There are, of course, some practical problems, such as efficiently identifying the complete set of minimum cost routes. However, this type of approach may provide some basis for improving the performance of the ME2 model

Consider now the O-D matrix estimation problem in which at some stage a set of R feasible routes has been identified. If instead of fixing the set of route choice proportions from an assignment, one assumes that each of these routes is 'a priori' equally likely to be used, one may restate the problem as

$$\text{Maximise} \quad -\sum_{ijr} T_{ijr}(\log T_{ijr}/t_{ijr} - 1) \tag{10.2}$$

subject to

$$\sum_{ijr} T_{ijr}\delta_{ijr}^{\ell m} - \hat{V}_{\ell m} = 0 \tag{10.4'}$$

$$\text{and} \quad T_{ijr} \geq 0 \tag{10.5}$$

Equations (10.3) do not add any information as $[T_{ij}]$ is unknown in this problem. It is always possible, of course, to reconstruct the O-D matrix $[T_{ij}]$ by aggregating the path flow matrices $[T_{ijr}]$.

Again the solution to this new programme is

$$T_{ijr} = t_{ijr} \prod_{\ell m} X_{\ell m}^{\delta_{ijr}^{\ell m}} \tag{10.8}$$

$$\text{and} \quad T_{ij} = \sum_r T_{ijr} \quad .$$

The prior path flows may be calculated from the prior trip matrix as $t_{ijr} = t_{ij}/R_{ij}$.

There is no reason to restrict this approach to equilibrium assignment. The set of feasible routes might have been identified using another multiple route choice model, for example Burrell's. The performance of this type of model with the Reading data base is explored below.

10.3 INDEPENDENT MULTIPLE ROUTES USING BURRELL's MODEL

Burrell's stochastic route choice model was used in Section 8.4. In those tests three randomised cost trees were used and the route choice proportions $[p_{ij}^{\ell m}]$ calculated as an average of the three routes generated. In this new version of the model, each Burrell route is considered a priori equally likely to be chosen and only the observed link flows will introduce information to modify these estimates.

The Reading data base for Tuesday 19 October with the extended set of counts (trip ends counted) was used for these tests. Two to five Burrell's trees with a 10 per cent spread were used and the results are summarised in Table 10.1 which also depicts results with all-or-nothing assignment for comparison. We note the following.

(i) The use of independent multiple routing seems to improve the performance of the ME2 model slightly. This improvement is greatest when only two routes are used but decreases as the number of routes is increased thereafter.

(ii) Of course the technique increases computer time both in terms of the ME2 model and in terms of converting trees into $\delta_{ijr}^{\ell m}$ variables.

(iii) The advantage of the technique seems to be that it allows variability in destination and route choice. This variability is reduced or controlled by the information in the observed link counts. For example, the method allows the increase of the flow over one path for an O-D pair without necessarily increasing the flow over the other paths.

(iv) The method introduces more degrees of freedom and then it should not, in principle, give worse results than using the same paths in non-optimal fixed proportions.

10.4 ESTIMATING PATH FLOWS WITH EQUILIBRIUM ASSIGNMENT

The basic idea of using independent multiple routeing for O-D estimation in the context of equilibrium assignment is to avoid having to specify rigid route choice proportions. Here we shall use equilibrium assignment to identify feasible routes under congested conditions.

Table 10.1: Tests with path flow estimation using Burrell's trees with 10 per cent spread. 19 October data

Indicator	2 trees	3 trees	4 trees	5 trees	All-or-nothing
Mean Absolute Difference MAE	1.0	1.0	1.0	1.1	1.1
Relative MAE %	83	84	84	85	86
Coefficient of Determination R^2	0.53	0.53	0.53	0.52	0.50
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.45	0.44	0.44	0.43	0.42
RMSE	2.0	2.0	2.0	2.0	2.1
%RMSE	159	160	161	163	169
Phi measure Range:					
0 - 2	506	503	503	507	512
3 - 5	388	390	396	395	426
Over 5	640	651	656	668	671
All	1534	1544	1555	1569	1609
NPhi (Phi/T)	0.83	0.84	0.85	0.86	0.88
Iterations	13	13	18	13	24
CPU time (secs)	3.08	4.58	5.96	7.71	3.76

The general heuristic scheme for estimating a trip matrix from an equilibrium assignment described in Section 9.3 requires only minor modification. It can now be stated as follows.

(i) Assign, using equilibrium assignment methods, a base year matrix $[t_{ij}]$ to the network and save the corresponding routes (trees). Set the cycle counter $n = 1$.

(ii) Estimate a trip matrix $[T_{ij}]^{(n)}$ using independent routes $[\delta_{ijr}^{\ell m}]$ and observed flows $[\hat{V}_{\ell m}]$.

(iii) Assign $[T_{ij}]^n$ to equilibrium saving the routes (trees) used in the process.

(iv) Increment n by 1 and return to step (ii) unless the change in routes $[\delta_{ijr}^{\ell m}]$ or estimated matrices have been sufficiently small.

Practical considerations (computer memory and CPU time) restrict the number of independent routes to be used to a maximum of about five. In tests with the Reading data it was possible to monitor the improvement (or otherwise) in the estimated matrix induced by the inclusion of an extra route per O-D pair.

Table 10.2 summarises the results obtained with this approach for the Reading data for Tuesday 19 October with the extended (trip ends counted) set of observed links. Figure 10.1 presents results for %RMSE in a graphical form.

The following comments can be made.

(i) Multiple routeing with equilibrium assignment seems to produce a better estimation of the O-D matrix than the iterative model tested in Chapter 9 (cfr Figure 9.3). In fact, these results are the best obtained for the model for single days with <u>modelled</u> routes, but are still some way away from those obtained with <u>observed</u> routes.

(ii) The model seem to perform reasonably well with 3 or 4 independent routes but no great improvement is achieved beyond the third cycle.

(iii) This approach requires considerable computer time but not significantly more than the iterative model reported in Chapter 9.

Table 10.2: Tests with path flow estimation and equilibrium assignment. Tuesday 19 October, trip ends counted.

Routes:	First cycle				Second cycle			Third cycle			Fourth cycle		
Indicator	1	2	3	4	2	3	4	2	3	4	2	3	4
Mean Absolute Difference MAE	1.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Relative MAE %	86	84	84	84	84	83	83	83	84	83	83	83	83
Coefficient of Determination R^2	0.50	0.58	0.58	0.58	0.58	0.59	0.59	0.58	0.57	0.58	0.58	0.58	0.58
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.42	0.45	0.45	0.45	0.45	0.46	0.47	0.46	0.45	0.46	0.46	0.46	0.46
RMSE	2.1	1.9	1.9	1.9	1.9	1.8	1.8	1.9	1.9	1.9	1.9	1.9	1.9
%RMSE	169	151	151	151	151	148	148	151	152	150	151	150	150
Phi measure Range:													
0 - 2	512	494	496	499	492	490	491	489	491	493	491	488	499
3 - 5	426	376	374	373	374	369	360	373	376	369	373	375	361
Over 5	671	664	663	661	666	639	632	656	664	649	654	646	640
All	1609	1534	1533	1535	1532	1498	1483	1517	1531	1511	1518	1508	1500
NPhi (Phi/T)	0.88	0.84	0.84	0.85	0.84	0.82	0.81	0.83	0.84	0.82	0.83	0.82	0.81
Iterations	24	14	20	24	13	13	12	12	12	23	12	12	21
CPU time (secs)	3.76	3.30	7.24	9.91	3.16	4.47	5.40	3.02	4.20	9.97	3.01	4.04	8.98
Accumulated CPU time (secs) including assignment	33.71	63.66	102.28	143.30	178.98	205.93	231.83	284.84	319.37	357.62	387.43	423.49	462.91

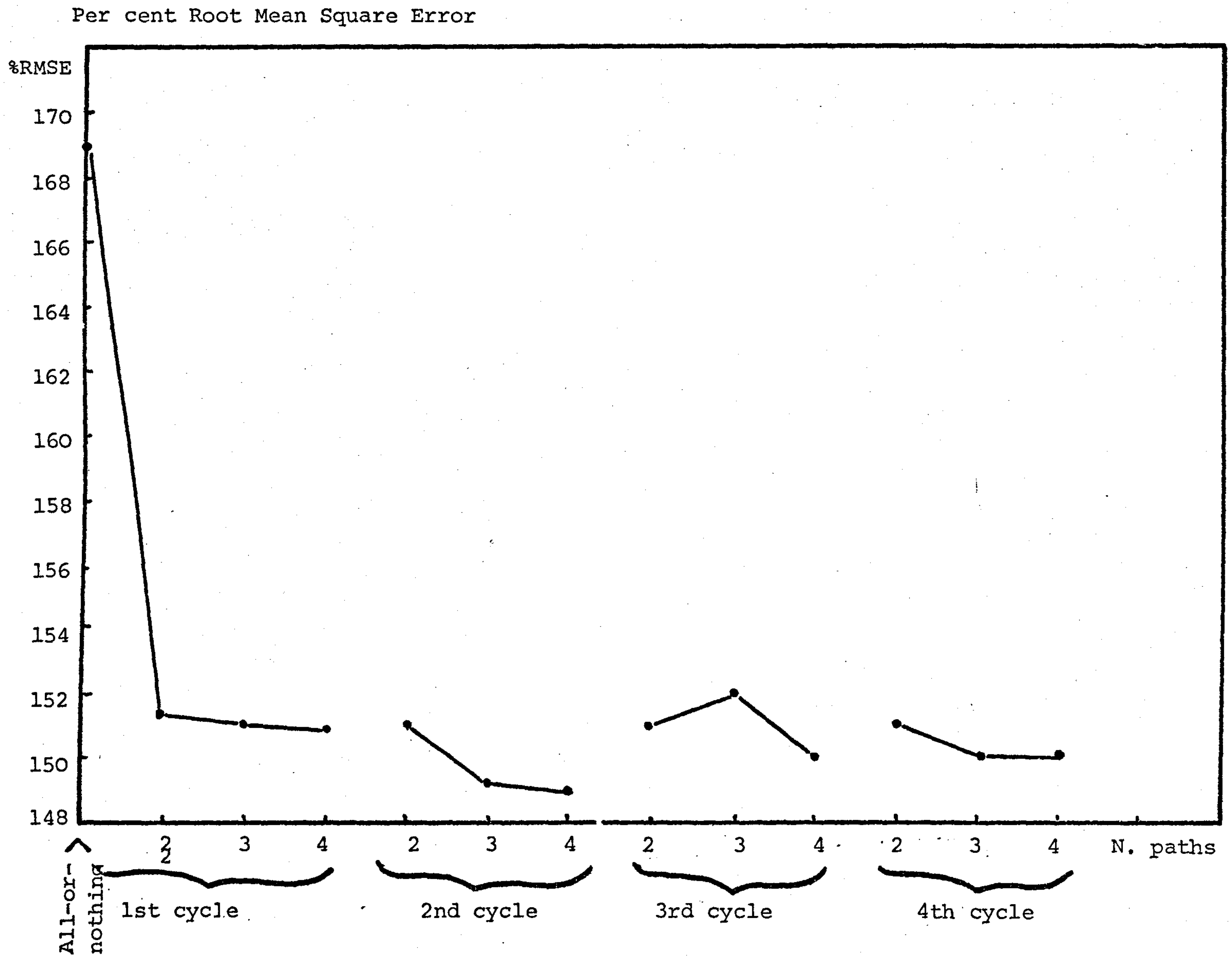


Figure 10.1: Tests with path flow estimation and equilibrium assignment, 19 October, extended set of counts (trip ends incl.)

Further tests were also made using the same data but without trip end counts with a maximum of 3 independent routes and 3 cycles. The results are summarised in Table 10.3 and Figure 10.2 for %RMSE.

It can be seen that similar results were obtained. The matrix estimated with this method was closer to the observed (sampled) one than that obtained, for a single day, by any non-capacity restrained method, even with trip end counts. This approach therefore seems to produce an improvement in the estimated O-D at least equivalent to the one generated by using trip end counts and a simpler route choice model.

These conclusions are not entirely independent of the particular network and data used. Further tests under different conditions should be performed to warrant a generalisation of these conclusions.

On the whole, the extension of the ME2 model to estimate path flows seems to offer an attractive technique for estimating trip matrices from counts under congested conditions.

Table 10.3: Tests with independent multiple routeing and equilibrium assignment. Tuesday 19 October, 159 links counted (real links only)

Routes:	First cycle			Second cycle		Third cycle	
Indicator	1	2	3	2	3	2	3
Mean Absolute Difference MAE	1.1	1.1	1.1	1.1	1.1	1.1	1.1
Relative MAE %	92	89	89	89	89	89	89
Coefficient of Determination R^2	0.47	0.52	0.52	0.52	0.53	0.52	0.52
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.38	0.40	0.40	0.40	0.41	0.40	0.41
RMSE	2.1	2.0	2.0	2.0	2.0	2.0	2.0
%RMSE	172	159	160	159	158	159	159
Phi measure Range:							
0 - 2	493	478	474	483	477	482	472
3 - 5	397	373	376	372	377	374	366
Over 5	715	728	740	728	733	729	731
All	1604	1580	1590	1583	1587	1585	1586
NPhi (Phi/T)	0.88	0.86	0.86	0.86	0.86	0.86	0.86
Iterations	13	13	23	13	16	19	22
CPU time (secs)	1.82	3.09	6.86	3.04	5.00	4.25	6.88
Accumulated CPU time (secs)	34.91	71.60	99.76	132.61	169.05	200.68	229.43

Per cent Root Mean Square Error.

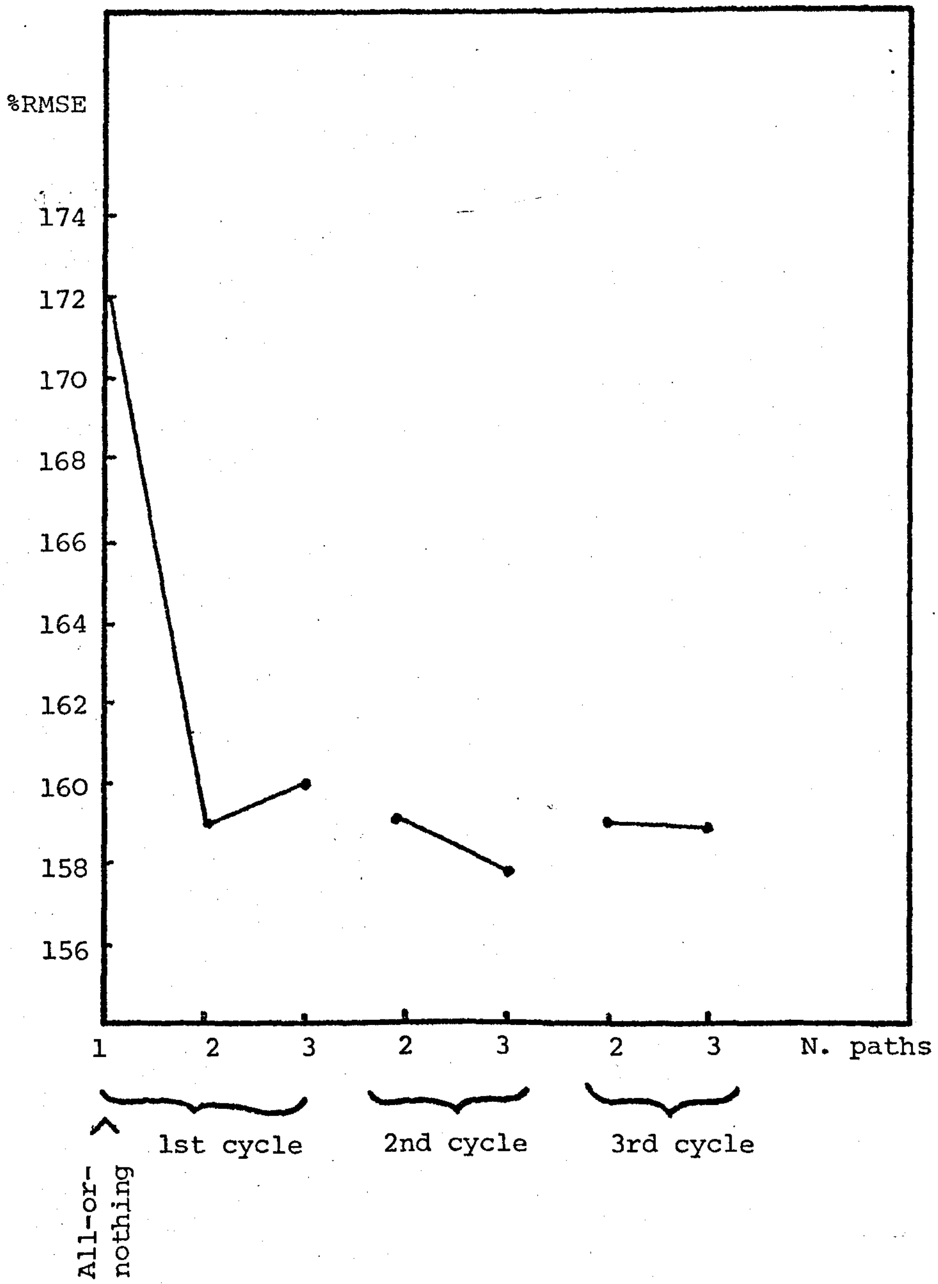


Figure 10.2: Tests with path flow estimation and equilibrium assignment, 19 October, real links only counted

CHAPTER 11
USE OF PRIOR INFORMATION

11.1 INTRODUCTION

All the tests reported so far have made no use of any prior information available on the trip matrix. One of the most interesting features of the ME2 model is precisely its capability of making use of prior information such as:

- an outdated trip matrix
- a trip matrix obtained from a study covering a larger area
- a model, perhaps of the gravity type, thought to be realistic for the area of interest
- a small sample survey.

The possibility of making use of such information will undoubtedly make the ME2 model more attractive for practical applications as it will almost certainly increase the accuracy of the estimated matrices. In the case of the Reading data there was no real prior information which could be used to test these aspects of the model. However, a number of tests have been carried out using the sampled matrix from one day as a prior estimate for other days.

11.2 TREATMENT OF ZEROES IN THE PRIOR TRIP MATRIX

The ME2 model in its general form can be related as

$$T_{ij} = t_{ij} \prod_{\ell m} X_{\ell m}^{p_{ij}^{\ell m}} \tag{11.1}$$

Solving the model results in a trip matrix $[T_{ij}]$ which is closest to the prior matrix $[t_{ij}]$ and reproduces the observed counts $[\hat{V}_{\ell m}]$. It is clear from (11.1) that if $t_{ij}=0$ so should T_{ij}, not an entirely satisfactory result unless the O-D pair ij is a structural zero, that is an impossible trip.

Observed matrices, and in particular those obtained from small sampling fractions, are very likely to contain a large number of zero cells. These may be 'true' or structural zeroes as defined above, but are much more likely to be zeroes 'by chance' due to limitations in the sampling framework. A cell with no observations in a small sample is almost certainly not significantly different from another cell where 1 trip was observed. The zeroes by chance will be preserved, together with structural ones, by the ME2 model and this may lead to unsatisfactory estimated matrices.

A pragmatic solution to this problem is to 'seed' the zero cells in the prior matrix $[t_{ij}]$ with a suitable small value so that all potential trip interchanges are possible in the updated matrix. Under this scheme the 'seeded' cells will be modified by action of the link balancing factors $X_{\ell m}$, some of them 'growing' to full trips in the matrix and others returning to zero, depending on the observed counts affecting them.

An issue of interest is to explore the magnitude of the 'seed' to allocate to empty cells. This was explored using the sample trip matrix from one day (19 October) as a prior estimate of the matrix for another day (21 October). Different seed values between 0.0 and 2.0 were tested using the extended set of counts (real links plus trip end counts) and all-or-nothing assignment. The results are summarised in Table 11.1 for the most important indicators. This table also shows the corresponding performance without prior information for comparison. The following comments can be made.

(i) The use of a prior matrix produces important improvements in the estimated matrix. For the MAE, R^2 and SR^2 indicators these results are better than those obtained in Chapter 10 using path flows.

(ii) Depending on the indicator used, seed values between 0.50 and 0.75 seem to produce the best results.

(iii) The extreme of using no seed seem to produce worse results than using too large a seed value (2.0)

(iv) For simplicity a seed value of 0.5 seems quite satisfactory

Table 11.1: 21 October matrix estimated using 19 October as prior matrix, trip ends counted, all-or-nothing route choice, different seed values

Indicator	Seed 0.0	0.1	0.25	0.50	0.75	1.0	2.0	No prior matrix
Mean Absolute Difference MAE	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.1
Relative MAE %	82	78	77	77	77	78	79	88
Coefficient of Determination R^2	0.56	0.61	0.63	0.64	0.64	0.63	0.62	0.47
Coefficient of Determination for $(T_{ij})^{½}$ SR^2	0.46	0.48	0.50	0.50	0.49	0.48	0.46	0.39
RMSE	2.3	2.1	2.0	2.0	2.0	2.0	2.0	2.1
%RMSE	177	162	157	155	154	154	156	166
Iterations	31	32	27	29	28	29	20	24
CPU time (secs)	4.20	4.38	4.05	4.07	4.05	4.16	2.92	3.44

A few comments should be made at this stage on the use of the model to update trip matrix in combination with SATURN as reported by Hall, Van Vliet and Willumsen (1980). An outdated trip matrix for Harrogate in North Yorkshire (24 zones) was available from roadside interviews. This matrix was used as a prior estimation in combination with a set of 63 recent traffic counts, most of them turning movements at key junctions. A 'seed' of 0.45 was used for cells with zero observations. After the updating process it was found that the average errors at flow levels with the use of SATURN had been reduced by about 60 per cent.

11.3 TRADE-OFF BETWEEN A PRIOR MATRIX AND COUNTED LINKS

One possible method for obtaining a reliable trip matrix for an area is to devote some resources to a conventional, small sample, survey (perhaps roadside interviews) and some resources to obtaining counts at selected links. It is possible to explore the advantages of this approach by using a Reading sampled matrix as a prior matrix and a subset of counts with the O-D estimation model. Thus the sampled trip matrix for 19 October was used as a prior estimate in combination with a random selection of counts (25 per cent to 100 per cent) to estimate the 4 days trip matrix. A seed of 0.5 for empty cells and all-or-nothing route choice were used.

The results of these tests are presented in Table 11.2 from which the following comments can be made.

(i) The use of a prior trip matrix appears even more valuable with the 4-day trip matrix data base as large improvements were obtained. One reason is that the 19 October matrix is about 1/4 of the matrix to be estimated and accordingly the traffic counts are only required to supply 3/4 of the total information. Another explanation is that the 4-day trip matrix has already been found to be a more favourable data base for the ME2 model because of its larger sample size.

(ii) The use of a prior matrix in this case seems very valuable as even with 25 per cent of the links counted the model performs better than with 100 per cent of the counts and no prior matrix.

(iii) The use of a prior matrix in this case produces better results with 75 per cent of the links counted and all-or-nothing route choice, than the performance of the model with 100 per cent of links counted, no prior matrix and observed route choice proportions.

(iv) In retrospect, it would have been better to use one day to estimate the other three, but it can be argued that the test performed is representative of the case in which a small sample roadside survey is used in combination with traffic counts to estimate a trip matrix. What these results have shown is that this approach may well prove to provide the best allocation of resources to estimate a trip matrix.

(v) The ME2 model in this updating mode provides what looks like a valuable tool in transport planning and management. The possibility of accumulating knowledge and being able to improve our estimation of a trip matrix as more information becomes available seems to be of great value. Its use may result in better trip matrices or a reduction in the cost of obtaining and updating them, or both.

Table 11.2: Estimated 4 days matrix using 19 October matrix as prior, seed = 0.5, all-or-nothing route choice, sampled links

Indicator	Sampled links 25 %	50 %	75 %	100%	100 % no prior	100 % no prior observed route choice
Mean Absolute Difference MAE	3.2	3.1	3.0	2.9	3.7	3.0
Relative MAE %	64	63	61	59	75	61
Coefficient of Determination R^2	0.62	0.68	0.72	0.73	0.54	0.70
Coefficient of Determination for $(T_{ij})^{\frac{1}{2}}$ SR^2	0.58	0.60	0.63	0.64	0.46	0.63
RMSE	6.5	6.5	6.3	6.3	7.3	5.6
%RMSE	130	129	127	125	146	112
Iterations	18	35	56	28	25	33
CPU time (secs)	2.32	4.40	7.63	4.06	3.53	10.32

CHAPTER 12
CONCLUSIONS

This, the last chapter of this dissertation, is organised as follows. The first section contains a synopsis of the main findings of this research. The corresponding section numbers appear in parentheses in the left hand margin. The second section puts forward some areas for further research around the problem of estimating a trip matrix from counts. Finally, the last section sets out a set of conclusions which the author feels can be drawn from this research.

12.1 SYNOPSIS

The main objective of this research was to develop and validate a model for the estimation of trip matrices from traffic counts. We began by identifying the basic
(2.1.1) elements of the problem, namely a set of centroids, nodes, links (some with volume counts) and an origin-destination matrix. Links may be defined as representing stretches of road and/or particular turning movements.

(2.1.2) In discussing trip matrices it was noted that they involved two types of aggregation: spatial aggregation in terms of grouping origins and destinations into zones, and temporal aggregation in terms of putting together journeys taking place at different times. Decisions about these two levels of aggregation have to be taken to define the 'trip matrix of interest' for the analysis of a particular problem. It was noted that in practice most trip matrices are fairly sparse.

(2.2) Three main groups of conventional techniques for estimating trip matrices were identified. The first group, direct methods, includes home and roadside interviews, flagging and vehicle following methods and aerial photography. Indirect methods include conventional trip distribution models, distribution models with 'borrowed' parameters, partial matrix techniques and cordoning sub-matrices.

The third group, hybrid methods, combine direct observations with a modelling effort to produce less sparse matrices.
(2.2.3) It was recognised that these techniques involve considerable resources and time and that the results are often of dubious accuracy. This, plus the more or less general availability of traffic counts, make the development of a method for estimating trip matrices from them particularly attractive.

(2.3.1) The problem of estimating a trip matrix was then described in some detail. Route choice variables $p_{ij}^{\ell m}$ link the trip matrix with the volume counts via a system of L_c linear equations

$$\sum_{ij} T_{ij} p_{ij}^{\ell m} - \hat{V}_{\ell m} = 0 \tag{12.1}$$

for each counted link. In most practical cases the number of these equations turns out to be far less than the number of unknowns T_{ij} . In these cases the problem can be said to be underspecified in the sense that more than one non-negative matrix may satisfy Equation (12.1). Some
(2.3.2) other potential problems with Equations (12.1) must be recognised at this stage. Firstly, some of these equations
(2.3.2.1) may be linearly dependent and thus add no 'new information'. It is also possible for the Equations (12.1) to be
(2.3.2.2) inconsistent, that is to have no feasible solution. These
(2.3.3) problems were discussed in terms of linear algebra.

The above discussion underlines the importance of the
(2.4) estimation of the route choice proportion $[p_{ij}^{\ell m}]$ and its relationship with traffic assignment. The most important route choice models were outlined starting with minimum
(2.4.1) cost all-or-nothing assignment, continuing with stochastic
(2.4.2) assignment and the introduction of congestion effects. The
(2.4.3) preferred model in this latter case is based on equilibrium
(2.4.4) assignment techniques. From the point of view of trip matrix estimation route choice models were classified as
(2.4.5) either proportional or non-proportional; the second group including all capacity restrained models and the first most of the remaining techniques. The importance of this distinction lies in the fact that for a proportional

assignment model the $[p_{ij}^{\ell m}]$ can be determined independently from the O-D estimation process. This has to be done jointly with the O-D estimation technique in the case of non-proportional assignment.

(2.5) The general problem of the accuracy of an O-D matrix
(2.5.1) was then discussed. The different sources of error affecting
(2.5.2) the accuracy of conventional methods were identified. The only type of error with a more or less developed theoretical treatment seem to be those due to sampling. It was found that given the sparsity of observed trip matrices whatever the survey method, sampling errors were likely to be quite large (for example of the order of ± 55 per cent for a sampling rate of 0.25 and an expected cell value of 10).
(2.5.3) Traffic counts are, of course, not error free and counting errors of the order of 3 to 10 per cent have been reported in the literature.

(3.1) A complete review of the methods proposed for estimating trip matrices from counts was also undertaken. Three families of techniques were identified. The first and
(3.2) largest group included approaches based on a gravity model; they centre on the calibration of a gravity model from link flow observations. This calibration may involve the use of multiple linear regression or a more general least squares approach.

(3.3) A second group uses more general direct demand models at the cost of greater calibration effort. Models in both the first and second groups made use of information like population and employment in addition to traffic counts.
(3.4) The third group includes those models which make use of network data only. Models in this group are of particular interest for estimating trip matrices in non-free standing areas as they require a minimum of assumptions regarding trip making behaviour. The model developed by the author belongs to this group.

(3.5) A special case of the third group occurs in very simple networks such as sections of a motorway, or a public transport line. In this case route choice presents no problem and the model structure and solution methods are greatly simplified.

(3.6) The great majority of these models have <u>not</u> been tested against observed O-D matrices, thus restricting the analysis of their relative value to theoretical considerations. The approaches based on some travel demand model seem appropriate to situations where the underlying model may be considered valid, for example in free standing towns or other self-contained study areas. Approaches based on network data only, seem more attractive in small, local areas, and wherever traffic counts are used to <u>update</u> trip matrices.

(4.1) A detailed presentation of the model put forward by the author began with a discussion of the entropy
(4.1.1) maximising formalism. This discussion covered the origin of
(4.1.2) the concept in physics, its use as a measure of information and information gain and its link with measures of error
(4.1.3) and dispersion. The ways in which entropy maximising may be used for model development was demonstrated using the
(4.1.4) classical example of the derivation of the gravity model.
(4.2.1) The basic model put forward by the author results from maximising an entropy function subject to link flow constraints. When the entropy function includes prior
(4.2.2) information maximising

$$S = -\sum_{ij} T_{ij}(\log T_{ij}/t_{ij} - 1) \qquad (12.2)$$

subject to (12.1) and non-negativity constraints for T_{ij} results in:

$$T_{ij} = t_{ij} \, \pi \, X_{\ell m}^{p_{ij}^{\ell m}} \qquad (12.3)$$

(4.2.3) A similar model put forward by Van Zuylen was also discussed and contrasted with the author's. A simple
(4.3) example using Monte Carlo simulation was used to demonstrate the sense in which the author's model can be said to produce the most likely trip matrix consistent with the observed link flows. Some of the properties of the author's
(4.4) Maximum Entropy Matrix Estimation (ME2) model were discussed. The model does not require a full set of counts; it seems to be able to use prior information in an efficient and consistent manner; and, provided the constraints (12.1) and the prior matrix $[t_{ij}]$ are consistent, it will generate a matrix that when loaded onto the network will reproduce the observed counts.

After deriving the model it is important to discuss methods for solving it. The first step in this task is to try to eliminate, as far as possible, the two identified
(5.1.1) sources of inconsistency in the constraints: inconsistency at the link flow and at the path flow level. To avoid the second source an appropriate route choice model should be
(5.1.2) used. Inconsistencies at link flow level are easily encountered wherever traffic counts contain errors and/or
(5.1.3) have been taken on different occasions. A maximum likelihood technique was put forward to solve link flow inconsistencies and embodied in the METWO1 computer program.

(5.2.1) In discussing the solution of the ME2 model it was found convenient to present the problem in three equivalent and related ways: a primal convex programme, its dual programme and an existence problem. It was shown that, provided the constraints define a feasible solution space, these three problems had a solution and that two algorithms to find it could be devised. The first algorithm involved a Newton-Raphson method to solve the system of non-linear equations resulting from substituting (12.3) into (12.1). It was noted that this method has limitations
(5.2.3) for large scale problems. A second algorithm involved an extension of the well known bi-proportional adjustment technique. This multi-proportional adjustment algorithm

has been incorporated into the METWO2 program and has been used throughout this research.

In many practical applications of the ME2 model budget consideration will limit the number of links to be
(5.3) counted. It is possible to obtain a first approximation to the information added by an extra traffic count within the same entropy maximising formalism. It was found that the indicator

$$I_{\ell m} = \frac{1}{2}\,\frac{(v_{\ell m} - \hat{V}_{\ell m})^2}{\hat{V}_{\ell m}} \tag{12.4}$$

(where $v_{\ell m}$ is the flow estimated without the count and $\hat{V}_{\ell m}$ is the expected value of that flow) gave a good estimation of the expected value of a new count at link ℓm.

(6.1) As a preparation for tests on the accuracy of the ME2 model a discussion was undertaken on goodness-of-fit statistics for trip matrices and a set of error measures was selected.

(6.2) Preliminary tests on the ME2 model were performed with an artificial data base in order to acquire a better understanding of its properties and to debug the software. It was found that the model was very good at reproducing from counts a trip matrix generated using a gravity model but performed less well when an aritificial matrix was generated using random numbers.

(7.1) This research was greatly facilitated by the existence of a comprehensive data base gathered by the TRRL in Reading (1976) and made available to the author. The
(7.1.1) characteristics of the data collection exercise and the error correction performed by TRRL were described. The data base is such that an observed trip matrix, a set of 159 traffic counts and a set of oberved route choice proportions can be obtained for four consecutive afternoon
(7.2) peaks (16.10 - 18.10). The process of coding the network into 39 zones, 80 nodes and 159 one-way links, further error
(7.3) trapping and data extraction by the author was also described.

(7.4) An analysis of the Reading data showed that the sample variations from day to day at the trip matrix level were much larger than the variations at link flow level. These variations are somewhat reduced by using a coarser zoning system (20 instead of the full 39 centroids).

The main effort of this research has been the testing of the ME2 model with the Reading data under a variety of route choice models and assumptions about the availability
(8.2) of data. The first set of tests was performed using the observed route choices and it was found that:

- the estimated trip matrices, although not very close to the observed ones, were well within the range of daily variations of the samples,
- trip end information proved quite valuable,
- the ME2 model performed much better with a coarser zoning system and with the data base representing an aggregate of the 4 observed days; sample size seems to be the main reason for this.

(8.3) A second set of tests using a very simplistic route choice model, all-or-nothing over free flow costs, gave similar results with the difference that the estimated matrices were now slightly outside the daily variations of the sampled one. One would expect a more realistic route choice model to produce results in between these two extremes.

(8.4) Tests were also performed using routes generated by a Burrell stochastic route choice, but the results were no better than those obtained using all-or-nothing route choice.

(8.5) A comprehensive set of sensitivity tests were then performed to study the performance of the model with an incomplete set of counts and counting errors. Incomplete sets of counts were chosen both at random and using the
(8.5.2) selective sampling criterion developed in Chapter 5. The
It was found that although random sampling of
(8.5.3) counts requires a large number of counts to produce

reasonable results, selective sampling is much more efficient enabling sample sizes of 30 - 40 per cent to be used in this
(8.5.4) case. Artificial errors were introduced into the counts using normally distributed noise with $\sigma/\sqrt{\mu}$ at different levels. It was found that for $\sigma < \sqrt{\mu}$ the performance of the ME2 model did not deteriorate beyond the daily variations of the observed matrices.

(8.5.5) For many practical applications it will be more important to know whether the use of an estimated rather than an observed matrix will result in significantly different forecasts. Both sampled and estimated trip matrices were loaded onto a modified network and it was found that the resulting flow patterns were closer than the observed daily variations.

(8.6) On the whole it was found that the model was reasonably robust, did not require much computer resources and produced results which, although not very close to the sampled matrices, are quite valuable considering that the ratio of unknowns to observations was at best around 9.4 to 1 .

The ME2 model is based on the assumption that the route choice proportions can be identified independently from the O-D estimation process. However, it is of interest
(9.1) to explore to what extent the model can be extended to include congestion effects which invalidate this assumption.
(9.2) Consideration was given to the equilibrium assignment framework noting that one of the properties of an equilibrium solution was that, unlike link flows and costs, path flows in general are not uniquely determined, making any identification of route choice proportions $[p_{ij}^{\ell m}]$ rather
(9.3) arbitrary. However, the iterative method used to solve the equilibrium assignment problem was also used to (heuristically) generate the $[p_{ij}^{\ell m}]$.

(9.4) Tests with this approach produced only a marginal improvement at a considerable computational cost. One reason for this disappointing result seems to be that some of the calculated $[p_{ij}^{\ell m}]$ are not fully compatible with the observed link flows.

(10.1) An alternative approach was then followed in which the path flows were estimated directly and the trip
(10.2) matrix aggregated over path flows. This approach may be used with capacity restrained or any other multi-route
(10.3) assignment model. The model was tested with Burrell's trees and a marginal improvement was obtained. Tests within
(10.4) an equilibrium assignment sequence produced better and more consistent results than the fixed route proportions method used in Chapter 9. The technique seems attractive but, as with any iterative method, increases the amount of computer time required.

(11.1) Throughout all the previous tests no prior information about the trip matrix was assumed. A separate series of tests was carried out in which the sampled trip matrix of one day was used as a prior estimate of the matrix for another day. As part of this procedure it was found convenient to 'seed' those cells with no observations due to limitations in the sampling framework. It was found that the use of a prior matrix greatly improved the accuracy of the ME2 model and that a seed value of around 0.5 seemed to produce the best overall results.

(11.3) In order to consider the trade-off between a prior trip matrix and counted links a one-day matrix (19 October) was used as a prior estimate of the 4-day matrix. It was found that this prior trip matrix combined with 25 per cent of the links counted produced better results than using 100 per cent of the links and no prior matrix.

These findings make it attractive to combine in the future direct (survey) and indirect (traffic counts) methods to estimate reliable trip matrices.

12.2 SUGGESTIONS FOR FUTHER RESEARCH

During this research several areas for further examination have been identified. The ones the author considers most promising are outlined below.

12.2.1 Further tests with a larger data base

One of the limitations of the Reading data base has been its limited sample size which probably explains some of the relatively large daily variations of the sampled matrices. Given that the ME2 model performed much better with the 4-day data base (larger sample size) one would expect better results with data collected with a higher sampling fraction or over a longer time period. In any case it is desirable to validate the model with other data bases to reduce the possibility of good results by chance.

12.2.2 Optimum combination of survey methods

A larger data base would also facilitate a more realistic analysis of the best deployment of resources between conventional surveys and traffic counts. For example samples of the observed vehicles can be taken as prior matrices for use with random or indicator-selected link counts. The relative costs of obtaining each of these types of information could then be incorporated into the problem of choosing the best allocation. The results from this type of study will be of considerable interest to practitioners.

12.2.3 Modelling a prior trip matrix

The possibility exists for using a model, say of the gravity type, to estimate the prior trip matrix $[t_{ij}]$. It is of particular interest in this context to check whether some of the ideas incorporated in the US Federal Highway Administration model (see Section 3.4.3.2) are of value. It should also be of interest to compare the performance of both models using a large data base.

To be of practical value the modelling of the prior matrix should use different information from that embodied in the traffic counts, but at the same time this new information should not involve a major data collection effort. It can be seen that this work is closely related to the one in Section 12.2.2.

12.2.4 Development of a better method for combining O-D estimation and equilibrium assignment

The two heuristic methods proposed in Chapters 9 and 10 to estimate trip matrices under capacity constrained conditions are not entirely satisfactory. At least from a theoretical point of view one can envisage a more rigorous treatment of the problem.

One attractive possibility is to develop a treatment analogous to the combined distribution assignment problem. It has been shown by, among others, Evans (1976) and Florian et al (1975) that the combined problem is equivalent to the following mathematical programme

$$\text{Minimise} \sum_{\ell m} \int_0^{V_{\ell m}} C(x)dx + \frac{1}{\beta} \sum_{ij} T_{ij}(\log T_{ij} - 1) \qquad (12.5)$$

subject to

$$\sum_{ijr} T_{ijr}\, \delta_{ijr}^{\ell m} - V_{\ell m} = 0 \qquad (12.6)$$

$$\left.\begin{array}{l} \sum_i T_{ij} - D_j = 0 \\ \sum_j T_{ij} - O_i = 0 \\ T_{ij} - \sum_r T_{ijr} = 0 \end{array}\right\} \qquad (12.7)$$

$$T_{ijr} \geq 0$$

where the distribution model is a doubly constrained gravity model. The solution to this problem is a matrix generated by a gravity model which when loaded onto the network to equilibrium results in the path costs being consistent with the costs used in the gravity model.

In the O-D estimation problem there is no information (in principle) on trip ends so constraints (12.7) should be dropped resulting in an unconstrained gravity model - equilibrium assignment problem. But the problem is complicated because some of the link flows are not variables $V_{\ell m}$ but observations $\hat{V}_{\ell m}$ thus introducing additional constraints. These can be said to be 'bundle' constraints as opposed to trip end ones and they introduce complexities requiring additional treatment. However, one suspects that this or a related approach will result in a more satisfactory theoretical treatment of the O-D estimation problem under equilibrium assignment conditions and an incomplete set of counts.

12.2.5 Better treatment of counting errors

During this research Monte Carlo simulation was used to study the impact of errors in the counts over the performance of the model. In this case after the link flow inconsistencies had been removed using the program METWO1, the model estimated the matrix which reproduced the corrected counts, in essence assuming they were error free. This is probably quite reasonable if those counts are the only information available.

However, there may be cases in which one would like to trade-off modifications to the prior trip matrix against error levels in the counts - in qualitative terms to obtain a matrix which is not too different from the prior matrix at the cost of not quite reproducing the traffic counts.

One possibility is to replace each link flow equation by two inequalities recognising the expected range of the 'true' value of the traffic count. An alternative approach, which seems easier to integrate to the ME2 model, may be outlined as follows.

Consider the entropy function as a measure of separation or error between the prior and the 'true' or current value of a variable. In the case of link flows the true value is represented by $V_{\ell m}$ and the observed value by $\hat{V}_{\ell m}$. The following mathematical programme may be built.

$$\text{Minimise} \sum_{ij} T_{ij}(\log T_{ij}/t_{ij} - 1) + \alpha \sum_{\ell m} V_{\ell m}(\log V_{\ell m}/\hat{V}_{\ell m} - 1) \qquad (12.8)$$

subject to

$$\sum_{ij} T_{ij}\, p_{ij}^{\ell m} - V_{\ell m} = 0 \qquad (12.9)$$

and

$$T_{ij} \geq 0$$

where α represents the relative weight attached to errors in the counts compared to changes to the prior matrix. The solution to this problem becomes

$$T_{ij} = t_{ij} \,\pi\, X_{\ell m}^{p_{ij}^{\ell m}} \qquad (12.3)$$

and

$$V_{\ell m} = \hat{V}_{\ell m}\, X_{\ell m}^{-V_{\ell m}/\alpha} \qquad (12.10)$$

This now looks like a problem in which the multi-proportional adjustment terms $X_{\ell m}$ are used to modify the prior trip matrix and the observed flows and the relative magnitude of these corrections is governed by the weight α. For a very large α only minor modifications to $\hat{V}_{\ell m}$ are allowed and the ME2 model can be seen as one in which $\alpha = \infty$. One may also envisage an option in which different weights are attached to different counts, for example according to their expected reliability.

Of course a good deal more work is required in order to produce a theoretically sound model along these lines, but again, the approach does seem promising.

12.2.6 Use of classified counts

It is not difficult to extend the use of the ME2 model to cases in which in at least some links the available counts distinguish between different kinds of vehicles. A third dimension, vehicle type, can be added to the trip matrix by making T_{ijk} to represent trips between i and j by vehicle k. A different 'link' ℓm should be associated to each vehicle type for classified counts and the new route choice proportions $[p_{ijk}^{\ell m}]$ calculated accordingly. For unclassified counts only, one link would be needed.

Probably the main problem in this extension would be to find an efficient way of handling these issues. This task would be simpler if one could assume that all vehicle types follow the same rules in route choice but this is unlikely to be realistic in the case of buses.

12.2.7 Application to public transport

The use of the ME2 model in public transport systems offers interesting possibilities. The simplest case is perhaps its use in fixed track systems where access is well controlled at stations. Further to the work already reported by Hauer and Shin (1980) it is interesting to use additional information to provide a better prior estimate. Probably the most promising source of additional data is, for variable fare systems, the distribution of ticket values sold. This information may act as a proxy for trip length distribution incorporated in the prior matrix.

In the case of buses much depends on the ticketing system and the type of information recorded in the way-bills.

12.2.8 Simplified algorithms for large networks

The algorithm used in the METWO2 program (multi-proportional adjustments) is fairly efficient provided the route choice proportions $[p_{ij}^{\ell m}]$ are made available in a suitable order. The extraction of the $[p_{ij}^{\ell m}]$ from trees and their sorting in the appropriate order is a limiting factor when dealing with fairly large networks (say more than 100 zones, 400 links). It would be desirable to find an approximate solution method which could make use of route choice trees directly. This would result in considerable savings in computer memory and should produce important savings in computer time.

12.2.9 Extensions for inequality constraints

In many practical cases some of the information on flow levels is better represented by inequalities, for example 'the flow on link a is at least V_a ', or the 'maximum volume on link b is V_b (its capacity)'. Inequality constraints could also be of interest when using traffic counts obtained some time in the past together with current counts.

The solution on entropy maximising problems with inequality constraints has been studied by Jefferson and Scott (1979) and Macgill (1979) among others. More recently Lamond and Stewart (1981) have shown how the general balancing method studied by Bregman (1967) can be applied to these problems. This method requires only minor modifications to the extended Kruithof's algorithm used in METWO2.

12.3 CONCLUSIONS

A detailed account of the findings of this research may be found in each chapter and, in synopsis, in Section 12.1. The author would like to conclude this presentation with some more personal comments summing up the main conclusions.

This research has reviewed the general problem of estimating a trip matrix from traffic counts and a number of methods proposed to carry out this task. The author has developed a new model based on an entropy maximising formalism which can be interpreted as producing the O-D matrix which is closest to its prior estimate (if any) and consistent with the observed counts.

This (ME2) model has some valuable properties; it does not require a full set of counts; it makes efficient use of prior and/or other information, the matrix generated reproduces the observed counts when loaded onto the network; and it is fairly modest in computational requirements

The model assumes that route choice proportions are identifiable using a suitable model and this generates some new questions, in particular under equilibrium assignment conditions, some of which have only been dealt with in a heuristic manner. There is scope for further refining the approach in this area.

A number of tests using data collected in Reading have shown the model to be reasonably robust and reliable. The model did not generate matrices very close to the observed ones but in general their errors were within the range of daily variations of the sampled matrices. The model performed very well where a better prior estimate of the trip matrix was provided.

It is possible to speculate how would ME2 or a model of this type fit in three of the main themes of current transport planning practice.

(i) 'Making the best possible use of the facilities already available'

One of the most common approaches in this area is the implementation of Comprehensive Traffic Management Schemes. As already pointed out in Chapter 1 only the simplest of these schemes can be designed and assessed without recourse to an O-D matrix. The ME2 model provides an ideal tool to this purpose and it has already been incorporated to one of the models (SATURN) developed to this end.

The main attractions of using ME2 in this context are not only its modest resource requirements but also that trip matrices can easily be obtained for different times of the day facilitating design. This is only possible with conventional models at a much higher cost.

The more extensive use of trip matrices facilitated by ME2 will also enable planners to identify better which groups of users benefit and which lose as a result of a measure, at least in geographical terms.

(ii) 'The need for continuous or recurrent planning'

It is more or less generally recognised that the era of large scale one-off expensive Transportation Studies is now over. A new style of planning based on a recurrent updating of designs and forecasts is called for. However, this new style has been slow to emerge in practice; a possible reason for this is a lack of efficient technical tools to assist this task.

In addition to the progress already made in relation to developing data banks and monitoring techniques a family of simplified models, less 'data greedy' than conventional techniques, is required. Models like ME2 or similar (perhaps based on a travel demand model) using traffic counts and other readily available data as major inputs, should play an important role in supporting a continuous planning effort (see Willumsen, 1981). The ability to make use of outdated information is not only a useful device but seems particularly appropriate in a scheme of recurrent updating of plans and forecasts.

(iii) 'Transport planning in developing countries'

This theme has been gaining importance both in the developing world and among consultants hoping to work there. The direct transfer of conventional techniques from developed to developing countries has proved to fall well short of initial expectations, see Willumsen (1977) and Stopher (1980) for example.

There are several reasons for this failure, in particular the fast rate of change in developing countries, lack of technical resources, poor quality of data and data collection and a different set of transport related problems. Transport planning techniques requiring large amounts of data and considerable technical resources are not appropriate to this environment. What is needed are techniques well adapted to the problems in hand which, while making fewer demands on technical resources and data, can be used more often to adapt plans to a changing environment.

Models like ME2 can only play a supporting role in an effort to develop a more appropriate planning style for developing countries (Willumsen, 1980), but the ideas of using generally available data, including outdated information, seems particularly applicable to this problem.

To conclude, the proposed model seems to be a promising technique to be used in the estimation or updating of trip matrices in several areas. The model seems well adapted to some of the main themes in transport planning today. While there are still some loose ends and several extensions to explore these are likely to be solved in time in the same way that a better understanding has been gained over the years on the specification, calibration and use of other transport models.

REFERENCES

AITCHINSON, J, and BROWN, J.A.C. (1969) The Log-Normal Distribution. Cambridge University Press.

BACHARACH, M.O.L. (1970) Biproportional Matrices and Input Output Change. Cambridge University Press.

BATTY, M. (1976) Entropy in spatial aggregation. Geographical Analysis, 8, 1-21.

BATTY, M. and MARCH, L. (1976a) The method of residues in urban modelling. Environment and Planning A, 8, 189-214.

BATTY, M. and MARCH, L. (1976b) Dynamic urban models based on information minimising. Geographical Paper 48, Department of Geography, University of Reading.

BEAR, D. (1976) Principles of Telecommunication-Traffic Engineering. P.Peregrinus.

BEBEE, M.A. (1959) Novel traffic survey method utilises vehicle lights. Bulletin of the Highway Research Board 224.

BECKMAN, M., McGUIRE, C.B. and WINSTEN, C. (1956) Studies in the Economics of Transportation. Yale University Press, New Haven Connecticut.

BECKMAN, M.J. (1974) Entropy, gravitv and utility in transportation modelling. In: G.Menges (ed.) Information, Inference and Decision. D.Reidel Publishing Co.

BEIL, D. (1979) Matrixmodelle zur Ermittlung von Teilströmen aus Querschnittmessungen. Universität (TH) Karlsruhe.

BELLAMY, P. (1978) Seasonal variations in traffic flows. TRRL Supplementary Report SR437. TRRL, Crowthorne.

BELLAMY, P. (1979) Undercounting of vehicles with single loop detector systems. TRRL Supplementary Report SR473. TRRL, Crowthorne.

BENDTSEN, P.H. (1974) Simplification of traffic models. Traffic Engineering & Control, 15, 817-819.

BLACK, J.A. and SALTER, R.J. (1975) A statistical evaluation of the accuracy of a family of gravity models. Proceedings of the Institution of Civil Engineers, 59, 1-20

BOLLAND, J., HALL, M., VAN VLIET, D. and WILLUMSEN, L.G. (1977) A model for the simulation of traffic management schemes. PTRC Summer Annual Meeting, Warwick University, June 1977.

BOLLAND, J., HALL, M., VAN VLIET, D. and WILLUMSEN, L.G. (1979) SATURN - simulation and assignment of traffic in urban road networks. Proceedings of the International Symposium on Traffic Control Systems, University of California, Berkeley, August 1979.

BREGMAN, L. (1967) The relaxation method of finding the common point of convex sets and its applications to the solution of problems in convex programming. USSR Computational Mathematical Physics, 7, 200-217.

BRILLOUIN, L. (1956) Science and Information Theory. Academic Press, New York.

BURRELL, J.E. (1968) Multiple route assignment and its applications to capacity restraint. Proceedings of the 4th International Symposium on the Theory of Traffic Flow, Karlsruhe 1968.

CAREY, M., HENDERSON, C. and SIDDHARTHAN, K. (1979) A Decomposition Method for Direct Estimation of Origin-Destination Trip Matrices. Carnegie Mellon University.

CARRUTHERS, G.A.P. (1956) A historical review of the gravity and potential concepts of human interaction. Journal of the American Institute of Planners, 22, 94-102.

CASEY, H.J. (1955) Applications to traffic engineering of the law of retail gravitation. Traffic Quaterly, XIX(1), ENO Foundation.

CATHCART, R. and FEARON, E.J. (1980) A database for CTM (Comprehensive Traffic Management) model validation. Traffic Engineering & Control, 21(7).

CHAN, V.K., DOWLING, R.G. and WILLIS, A.E. (1980) Deriving origin-destination information from routinely collected traffic counts. Working Paper UCB-ITS-WP-80-1, Institute of Transportation Studies, University of California, USA.

CLARKE, R.H. and DAVIES, D.H. (1970) Origin and destination survey by registration number method. Traffic Engineering & Control, 11(12).

COCHRAN, W.G. (1960) Sampling Techniques. John Wiley and Sons, New York.

COELHO, J.D. and WILSON, A.G. (1977) An equivalence theorem to integrate entropy-maxising submodels within overall mathematical programming frameworks. Geographical Analysis, IX, 160-173.

DAGANZO, C.F. (1977) On the traffif assignment problem with flow dependent costs. Transportation Research, 11, 433-441.

DAGANZO, C.F. (1980) An equilibrium algorithm for the spatial aggregation problem of traffic assignment. Transportation Research, 1B, 221-228.

DAGANZO, C.F. and SHEFFI, Y. (1977) On stochastic models of traffic assignment. Transportation Science, 11, 253-274.

DAWSON, J.A.L. (1979) Comprehensive traffic management in York - the monitoring and modelling. Traffic Engineering & Control, 20(11).

DAY, M.J.L. and HAWKINS, A.F. (1979) Partial matrices, empirical deterrence functions and ill-defined results. Traffic Engineering & Control, 20(8/9), 429-433.

DEBAILLE, S. (1977) Un modèle de réconstitution de matrices origine-destination en milieu urbain. Institut de Recherche des Transports, Raport de Recherche IRT 27. Arceuil, France.

DEPARTMENT OF THE ENVIRONMENT (1971) Advice Note 1A. Department of the Environment, URPA Division, London.

DEGROOT, M.H. (1970) Optimal Statistical Decisions. McGraw-Hill, London.

DIAL, R.B. (1972) A probabilistic multipath assignment model which obviates path enumeration. Transportation Research, 5, 83-111.

DIAL, R.B. (1973) Demand forecasting for new options and technology. Resource Paper, Highway Research Board Special Report 143 on Urban Travel Demand Forecasting, 77-85.

DINKEL, J.J., KOCHENBERGER, G.A. and WONG, S-N (1977) Entropy maximisation and geometric programming. Environment and Planning A, 9, 419-427.

DOBSON, G.G. (1965) Accuracy of home questionnaire surveys. Traffic Engineering & Control, 6(12), 723-726.

DOMENCICH, T., KRAFT, G. and VALETTE, J. (1968) Estimation of urban passenger travel behaviour: an economic demand model. Highway Research Record 238, 64-78.

DOW, P.D.C. and VAN VLIET, D. (1979) Capacity restrained assignment. Traffic Engineering & Control, 20(6), 296-305.

ELDOR, M. (1976) Integrated Real-Time Freeway On-Range Control Strategies. Institute of Transportation Studies, University of California, Berkeley, California.

ERIKSSON, J. (1978) Solution of large sparse maximum entropy problems with linear equality constraints. Report LITH-MAT-R-1978-02. Department of Mathematics, Linköping University, Linköping, Sweden.

ERLANDER, S. (1977a) Entropy in linear programs - an approach to planning. Report LITH-MAT-R-1977-03. Department of Mathematics, Linköping University, Linköping, Sweden.

ERLANDER, S. (1977b) Accessibility, entropy and the distribution and assignment of traffic. Transportation Research A, 11(3), 149-153.

ERLANDER, S. (1980) Optimal spatial interaction and the gravity model. Springer-Verlag Lecture Notes in Economics and Mathematical Systems, Berlin.

EVANS, S.P. (1976) Derivation and analysis of some models combining trip distribution and assignment. Transportation Research, 10, 37-57.

FEDERAL HIGHWAY ADMINISTRATION (1975) Urban Origin-Destination Surveys. US Department of Transportation, Washington DC.

FEDERAL HIGHWAY ADMINISTRATION (1977) Transportation Planning in Small Urban Areas: The Appropriate Level of Effort. US Department of Transportation, Washington DC.

FISCH, O. (1977) On the utility of entropy maximisation. Geographical Analysis, 9, 79-84.

FISK, C. (1980) Some developments in equilibrium traffic assignment. Transportation Research, 14B, 243-255.

FLORIAN, M., NGUYEN, S. and FERLAND, J. (1975) On the combined distribution assignment of traffic. Transportation Science, 9, 43-53

FOSTER, BENSON and STOVER (1976) Consequences of small sample origin-destination. Federal Highway Administration FHWA-RD-76-43. US Department of Transportation, Washington DC.

FRATAR, T.J. (1954) Forecasting distribution on interzonal vehicular trips by successive approximation. Highway Research Board Proceedings, 33, 376-385.

GARNER, J.B. and MOUNTAIN, L.J. (1978) Traffic data collection - an alternative method? Traffic Engineering & Control, 19(10).

GARNER, J.D. and MOUNTAIN, L.J. (eds) (1979) Proceedings of the Traffic Data Collection Conference at Leeds. Leeds University, Extra-Mural Studies.

GAUDRY, M. and LAMARRE, L. (1978) Estimating origin-destination matrices from traffic counts: a simple linear intercity model for Quebec. Centre de Recherche sur les Transports Publication 105. University of Montreal, Canada.

GOULT, R.J. (1978) Applied Linear Algebra. Halstead Press, John Wiley and Sons, Chichester.

GRECCO, W.L. and others (1976) Transportation planning for small urban areas. National Cooperative Highway Research Programme Report 167.

GUR, Y. and TURNQUIST, M.A. (1979) Estimation of trip tables from observed link volumes. Transportation Research Record 730. Transportation Research Board, Washington.

GUR, Y., TURNQUIST, M.A., SCHNEIDER, M. and LEBLANC, L.J. (1978) Determining an O-D trip table based on observed volumes. Proceedings Annual Meeting of Operations Research Society of America. New York.

HAJEK, J.J. (1977) Optimal Sample Size of Roadside Interview Origin-Destination Surveys. Ontario Ministry of Transportation and Communications, Canada.

HALL, M.D., VAN VLIET, D. and WILLUMSEN, L.G. (1980) SATURN - a simulation assignment model for the evaluation of traffic management schemes. Traffic Engineering & Control, 21(4), 168-176.

HAMERSLAG, R. and HUISMAN, H.C. (1978) Binaire Calibratie. Verkeerskunde, 166-168.

HAUER, E. and SHIN, S.B.T. (1980) Estimation of origin-destination matrices from observed flows: simple systems. University of Toronto-York University Joint Program in Transportation Research Report 65. Toronto, Canada.

HOGBERG, P. (1975) Planning of traffic counts. University of Goteborg, Department of Statistics Research Report 1975:3.

HOGBERG, P. (1976) Estimation of parameters in models for traffic prediction: a non-linear regression approach. Transportation Research, 10, 263-265.

HOGG, R.V. and CRAIG, A.T. (1970) Introduction to Mathematical Statistics. Third Edition. Collier Macmillan, London.

HOLM, J., JENSEN, T., NIELSEN, S., CHRISTENSEN, A., JOHNSON, B. and RONBY, G. (1976) Calibrating traffic models on traffic census results only. Traffic Engineering & Control, 17.

HUTCHINSON, B.G. (1974) Principles of Urban Transport Systems Planning. McGraw-Hill.

HYMAN, G. (1969) The calibration of trip distribution models. Environment and Planning, 1, 105-112.

INSTITUTION OF CIVIL ENGINEERS (1978) Traffic Data Collection. Institution of Civil Engineers, London.

JAYNES, E.T. (1957) Information theory and statistical mechanics. Physical Review, 106, 620-630; and 108, 171-190.

JEFFERSON, T.R. and SCOTT, C.H. (1979) The analysis of entropy models with equality and inequality constraints. Transportation Research, 13B, 123-132.

JEFFREYS, M. and NORMAN, M. (1977) On finding realistic flows at road junctions. Traffic Engineering & Control, 18(1), 19-25.

JENSEN, S.M. (1980) Analysis of intercity travels by railways in Denmark. Institute of Mathematical Statistics and Operations Research, Research Report 3/1980. Technical University of Denmark, Lyngby.

JENSEN, T. and NIELSEN, S. (1973) Calibrating gravity models and estimating their parameters using traffic volume counts. 5th Conference of Universities Transport Study Group, London, January 1973.

JÖRNSTEN, K.O. and NGUYEN, S. (1979) On the estimation of a trip matrix from network data. Working Paper LITH-MAT-R-79-36. Department of Mathematics, Linköping University, Sweden.

KIRBY, H.R. (1979) Partial matrix techniques. Traffic Engineering & Control, 20(8/9), 422-428.

KRAFT, G. (1963) Demand for intercity passenger travel in the Washington-Boston corridor. North-east Corridor Project Part V. Systems Analysis and Research Corp., Boston.

KRUITHOF, J. (1937) Calculation of telephone traffic. De Ingenieur, 52(8), E15-E25. English translation by UK Post Office Research Department Library (No. 2663) Dollis Hill, London.

KULLBACK, S. (1959) Information Theory and Statistics. John Wiley and Sons, New York.

KURTH, D., SCHNEITHER, M. and GUR, Y. (1979) Small-area trip distribution model. Transportation Research Record 728. Transportation Research Board, Washington.

LAI, F.K. and VAN VLIET, D. (1979) Current trip assignment techniques in the U.K. Traffic Engineering & Control, 20(7), 348-352.

LAMARRE, L. (1977) Une méthode linéaire simple d'estimation de la matrice des origines et des destinations à partir de comptages sur les liens d'un réseau: une application au reséau routier du Québec. Centre de Recherche sur les Transports, Université de Montréal, Publication 76.

LAMOND, B. and STEWART, N.F. (1981) Bregman's balancing method. Transportation Research, 15B(4), 239-248.

LAWTON, W.H. and SYLVESTER, E.A. (1971) Elimination of linear parameters in non-linear regression. Technometrics, 13(3), 461-467.

LEBLANC, L.J. and FARHANGIAN, K. (1980) Selection of a trip table which reproduces observed link flows. Technical Report OREM 80007. Department of Operations Research and Engineering Management. Southern Methodist University, Texas.

LEBLANC, L.J., MORLOK, E.K. and PIERSKALLA, W.P. (1975) An efficient approach to solving the road network equilibrium traffic assignment problem. Transportation Research, 9, 309-318.

LEESE, M.N. (1977) A statistical test for trip distribution models. 9th Conference of Universities Transport Study Group, Universities of Glasgow and Strathclyde, January 1977.

LEONARD, D.R. and TOUGH, J.B. (1979) Validation work on CONTRAM - a model for use in the design of traffic management schemes. PTRC Summer Annual Meeting, University of Warwick, July 1979. Seminar J, 135-153.

LEONARD, D.R., TOUGH, J.B. and BAGULEY, P.C. (1978) CONTRAM: A traffic assignment model for predicting flows and queues during peak periods. TRRL Report LR841. TRRL, Crowthorne.

LOW, D. (1972) A new approach to transportation systems modelling. Traffic Quarterly, 26, 391-404.

MacGILL, S.H. (1979) Convergence and related properties for a modified biproportional problem. Enviroment and Planning A, 11, 499-506.

MacGILL, S.M. and WILSON, A.G. (1979) Equivalences and similarities between some alternative urban and regional models. Sistemi Urbani, 1, 9-40.

McLAUGHLIN, W.A. (1966) Multipath system traffic assignment. Report R108. Department of Highways, Ontario, Canada.

McSHANE, W.R. and CROWLEY, K.W. (1976) Regularity of some detector-observed arterial traffic volume characteristics. Transportation Research Record, 596, 33-37.

MARCH, L. and BATTY, M. (1976) Generalised measures of information, Bayes's likelihood ratios and Jayne's formalism. Environment and Planning B, 2(99-105.

MARSHALL, M.L. (1979) Labour saving methods for counting traffic movements at three-and four-arm junctions. Traffic Engineering & Control, 20(4), 159-162.

MEKKY, A. (1979) On estimating turning flows at road junctions. Traffic Engineering & Control, 20(10), 486-487.

MORPHET, R. (1975) A note on the calculation and calibration of doubly constrained trip distribution models. Transportation, 4, 43-53.

MURCHLAND, J.D. (1968) Method of diverted assignment over adjacent routes used in London Underground Programs. Report LBS-TNT-102.1. London School of Business Studies.

MURCHLAND, J.D. (1977) The multiproportional problem. University College London, Transport Studies Group Note JDM-263. (unpublished)

MURCHLAND, J.D. (1978) Applications, history and properties of bi- and multi-proportional models. Seminar at the London School of Economics, London. (unpublsihed)

NATHANSON, M. (1978) Information minimisation, marker transition and dynamic modelling. Environment and Planning A, 10, 879-888.

NGUYEN, S. (1974) An algorithm for the traffic assignment problem. Transportation Science, 8, 203-216.

NGUYEN, S. (1976) Procedures for equilibrium traffic assignment with elastic demand. Centre de Recherche sur les Transport, University of Montreal, Publication 39.

NGUYEN, S. (1977) Estimating an O-D matrix from network data: a network equilibrium approach. Centre de Recherche sur les Transport, University of Montreal, Publication 87.

NGUYEN, S. (1978) On the estimation of an O-D matrix by the equilibrium method using pseudo delay functions. Centre de Recherche sure les Transports, University of Montreal, Publication 81.

NUMERICAL ALGORITHM GROUP (1978) NAG FORTRAN Library Manual Mark 7, Volume 4. Numerical Algorithm Group, Oxford.

NIJKAMP, P. and PAELINCK, J.H.P. (1974) A dual interpretation and generalization of entropy maximisation models in regional science. Papers of the Regional Science Association, 33, 13-31.

OECD ROAD RESEARCH GROUP (1974) Urban Traffic Models: Possibilities for Simplification. OECD, Paris.

OPENSHAW, S. (1976) An empirical study of some spatial interaction models. Environment and Planning A, 9, 169-184.

OUTRAM, V.E. and THOMPSON, E. (1978) Drivers perceived cost in route choice. Proceedings PTRC Summer Annual Meeting, University of Warwick, July 1978. PTRC Education and Research Services Ltd., London.

PHILLIPS, G. (1979a) Accuracy of annual traffic flow estimates from short period counts. TRRL Supplementary Report SR514. TRRL, Crowthorne.

PHILLIPS, G. (1979b) Accuracy of annual traffic flow estimates from automatic counts. TRRL Supplementary Report SR515. TRRL, Crowthorne.

PIGMAN, J.G. and DEEN, R.C. (1979) Simulation of travel patterns for small urban areas. Transportation Research Record 730.

PITFIELD, D.E. (1978) Algorithm 6: the χ^2 test for predicted trip matrices. Environment and Planning A, 10, 1201-1206.

POTTS, R.B. and OLIVER, R.M. (1972) Flows in Transportation Networks. Academic Press, New York.

QUANDT, R.E. and BAUMOL, W.J. (1966) The abstract mode model: theory and measurements. Journal of Regional Science, 6, 13-26.

RATCLIFFE, E.P. (1972) A comparison of drivers route choice criteria and those used in current assignment processes. Traffic Engineering & Control, 13, 526-529.

REYNI, A. (1965) On the foundations of information theory. Review of the International Statistical Institute, 33, 1-14.

ROAD RESEARCH LABORATORY (1965) Research on Road Traffic. HMSO, London.

ROBERTSON, D.I. (1977) A deterministic method of assigning traffic to multiple routes of known cost. TRRL Report LR757. TRRL, Crowthorne.

ROBERTSON, D.I. (1979) Traffic models and optimum strategies of control: a review. Proceedings of the International Symposium on Traffic Control Systems, University of California, Berkeley, 1, 262-288.

ROBILLARD, P. (1975) Estimating an O-D matrix from observerd link volumes. Transportation Research, 9, 123-128.

ROBILLARD, P. and TRAHAN, M. (1973) Estimating the O-D matrix and network characteristics from observed link volumes. Proceedings of International Conference of Transportation Research. First conference, 736-740. Also: Publication 128, Département d'Informatique, Université de Montréal.

SHANNON, C.E. (1948) The mathematical theory of communication. Bell Systems Technical Journal, 27, 379-423, 623-656.

SHEPPARD, E.S. (1976) Entropy, theory construction and spatial analysis. Environment and Planning A, 8, 741-752.

SIKDAR, P.K. and HUTCHINSON, B.G. (1980) Empirical studies of work trip distribution models. To be published in Transportation Research.

SIKDAR, P.K., SMITH, D.P. and HUTCHINSON, B.G. (1980) An empirical assessment of trip distribution models. Proceedings of World Conference on Transport Research, London, April 1980.

SLATER, P.B. (1978) Gravity-model estimates and exact maximum entropy solutions: their differences. Environment and Planning A, 10, 1131-1136.

SMITH, D.P. and HUTCHINSON, B.G. (1981) Goodness of fit statistics for trip distribution models. Transportation Research, 15A(4), 295-303.

SMITH, M.E. (1979) Design of small-sample home-interview travel surveys. Transportation Research Record 701. Transportation Research Board, Washington.

SMITH, R. and McFARLANE, W. (1978) Examination of a simplified travel demand model. Transportation Engineering Journal of ASCE, 104 (TE1).

SMOCK, R. (1962) An iterative assignment approach to capacity restraint on arterial networks. Highway Research Board Bulletin 347, 60-66.

SOSSLAU, A.B. and BROKKE, G.E. (1960) Appraisal of O-D survey sample size. Public Roads, 31(5), 114-119.

SOSSLAU, A.B., HASSAM, A.B. and CARTER, M.M. (1978) Manual techniques and transferable parameters for urban transportation planning. Transportation Research Record 673. Transportation Research Board, Washington.

STEWART, N.F. (1979) Notes on the mathematical structure of equilibrium models. Department of Mathematics LITH-MAT-R-79-8. University of Linköping, Sweden.

STOPHER, P.R. (1980) Transfering urban transport planning methods to developing countries. Journal of Institution of Highway Engineers, October, 8-14.

STOVER, V.G., BENSON, J.D. and BUECHLER, J. (1976) Sensitivity evaluation of the traffic assignment process. Transportation Research Record 610, 50-52. Transportation Research Board, Washington.

SYMONS, J., WILSON, R. and PATERSON, J. (1976) A model of inter-city motor travel estimated by link volumes. Australian Road Research Board Proceedings, 8, Session 20.

TANNER, J.C. (1961) Factors affecting the amount of travel. Road Research Laboratory Technical Paper 51. HMSO, London.

THOMAS, R.W. (1977) An interpretation of the journey-to-work in Merseyside using entropy-maximising methods. Environment and Planning A, 9, 817-834.

TRIBUS, M. (1969) Rational Descriptions, Decisions and Designs. Pergamon Press, New York.

TURNQUIST, M. and GUR, Y. (1979) Estimation of trip tables from observed link volumes. Transportation Research Record 730, 1-6. TRB 'Issues in transportation planning for small and medium-sized communities'.

VAN VLIET, D. (1976) Road assignment, parts I, II and III. Transportation Research, 10, 137-143.

VAN VLIET, D. (1978) Improved shortest path algorithm for transport networks. Transportation Research, 12, 7-20.

VAN VLIET, D. (1979a) The application of mathematical programming to road assignment. Bulletin of the Institute of Mathematics and its Applications, 15(11/12), 287-290.

VAN VLIET, D. (1979b) Four unsolved problems in road assignment. Proceedings PTRC Summer Annual Meeting, University of Warwick, July 1979. PTRC Education and Research Services Ltd., London.

VAN VLIET, D. (1979c) The common man's guide to the theory of equilibrium assignment. Technical Note 13. Leeds University, Institute for Transport Studies.

VAN ZUYLEN, H.J. (1978a) A method to estimate a trip matrix from traffic volume counts. Proceedings PTRC Summer Annual Meeting, University of Warwick, July 1978. PTRC Education and Research Services Ltd., London.

VAN ZUYLEN, H.J. (1978b) The information minimising method: validity and applicability to transport planning. In: G.R.M.Jansen et al (eds) New Developments in Modelling Travel Demand and Urban Systems. Saxon House, Farnborough.

VAN ZUYLEN, H.J. (1979) The estimation of turning flows on a junction. Traffic Engineering & Control, 20(11), 539-541.

VAN ZUYLEN, H.J. and WILLUMSEN, L.G. (1980) The most likely trip matrix estimated from traffic counts. Transportation Research, 14B, 281-293.

WALSH, J. and WEBBER, M. (1977) Information theory, some concepts and measures. Environment and Planning A, 9, 395-417.

WANG, J. and MAY, A.D. (1973) Analysis of freeway on-ramp control strategies. Institute of Transportation and Traffic Engineering, University of California, Berkeley.

WARDROP, J.G. (1952) Some theoretical aspects of road traffic research. Proceedings Institute of Civil Engineers, Part II, 1, 325-378.

WILLIS, A.E. and CHAN, V.K. (1980) Deriving origin-destination information from routinely collected traffic counts. Working Paper UCB-ITS-WP-80-10, Institute of Transportation Studies, University of California, Berkeley.

WILLS, M.J. (1977) Non linear estimation of origin-destination flows and gravity model parameters from traffic counts on links. Department of Geography, University of British Columbia, mimeographed notes.

WILLS, M.J. (1978) Linear and non-linear estimators of the O-D matrix. Ph.D. Thesis. Department of Geography, University of British Columbia.

WILLUMSEN, L.G. (1977) Transport planning techniques for developing countries: the experience of some Latin American countries. Proceedings PTRC Summer Annual Meeting, University of Warwick, June, 1977. Seminar E. PTRC Education and Research Services Ltd., London.

WILLUMSEN, L.G. (1978a) Estimation of an O-D matrix from traffic counts: a review. Institute for Transport Studies Working Paper 99. University of Leeds.

WILLUMSEN, L.G. (1978b) O-D matrices from network data: a comparison of alternative methods for their estimation. Proceedings PTRC Summer Annual Meeting, University of Warwick, July 1978. PTRC Education and Research Services Ltd., London.

WILLUMSEN, L.G. (1979) Estimating the most likely O-D matrix from traffic counts. 11th Annual Conference of Universities Transport Study Group, University of Southampton, January 1979.

WILLUMSEN, L.G. (1980) Appropriate transport planning techniques for developing countries. Proceedings World Conference on Transport Research, London, 1980.

WILLUMSEN, L.G. (1981) Simplified transport models based on traffic counts. To be published in Transportation.

WILSON, A.G. (1970) Entropy in Urban and Regional Modelling. Pion Press, London.

WILSON, A.G. (1974) Urban and Regional Models in Geography and Planning. John Wiley, London.

WILSON, A.G. (1975) Some new forms of spatial interaction model: a review. Transportation Research, 9, 167-179.

WILSON, A.G. and MacGILL, S.M. (1977) A system analytical framework for comprehensive urban and regional model building. School of Geography Working Paper 209. University of Leeds.

WILSON, S.R. (1976) Statistical notes on the evaluation of calibrated gravity models. Transportation Research, 10, 343-345.

WRIGHT, C.C. (1978) Westminster route choice survey; evaluation of study method. Transport Research Note 25. City Engineer's Department, Westminster City Hall, London.

WRIGHT, C.C. and ORROM, H. (1976) The Westminster route choice survey: a new technique for traffic studies. Traffic Engineering & Control, 17, 348.

WYTCONSULT (1975) Speed flow relationships on road networks. Document 504. West Yorkshire Transportation Studies, Wakefield.

WYTCONSULT (1977) The structure and calibration of the Leeds and Bradford transport models. Document 313, Vols I-III. West Yorkshire Transportation Studies, Wakefield.

APPENDIX A1

NOTATION

A_i, B_j	Balancing factors used in a gravity model
C_{ij}	Cost of travel between origin i and destination j
C_{ijr}	Cost of travel between i and j via route r
$C_{\ell m}(V_{\ell m})$	Cost-flow function for link ℓm assumed to be a non-decreasing function
D_j	Total number of trips attracted to destination j
E_j	Variable associated to the attraction of zone j, for example employment
$f(c_{ij})$	A deterrence function relating the number of trips between i and j to the cost C_{ij} of travelling between them
G_i	Variable associated to the trip generation power of zone i, for example population
L	Total number of links
L_c	Total number of links counted
$L_{c'}$	Total number of <u>independent</u> links counted
M	Total number of zones in study area
N	Total number of nodes
O_i	Total number of trips originated in zone i
$p_{ij}^{\ell m}$	Proportion of trips between i and j using link ℓm
t_{ij}	Prior trips between i and j
T_{ij}	Trips between i and j
T_{ijr}	Trips between i and j via route r
$V_{\ell m}$	Volume on link ℓm
$\hat{V}_{\ell m}$	Observed volume on link ℓm
$X_{\ell m}$	Link balancing factor corresponding to the count on ℓm
$\delta_{ijr}^{\ell m}$	Indicator, one if route r between i and j uses link ℓm, zero otherwise

CONVENTIONS

[ ] A variable in square brackets is taken to represent the whole set, for example $[T_{ij}]$ represents the whole trip matrix

^ Used on a variable to indicate an observed values, for example $\hat{V}_{\ell m}$

$\pi_{\ell m}$ Used to indicate the product over all variables with a subscript ℓm, for example

$$\pi_{\ell m} X_{\ell m}^{p_{ij}^{\ell m}} = X_1^{p_{ij}^1} \cdot X_2^{p_{ij}^2} \cdot X_3^{p_{ij}^3} \cdot \ldots$$

APPENDIX A2

<u>Special characters identifying patches in the observations in Reading</u>

The special characters preceding certain observations have the following meanings.

*	The observation point was interpolated
?	An adjustment was made to the observation time
S	The observation was obtained by 'matching' to an observation differing by a single digit in the number plate
Y	Matched to an observation differing by the year letter
T	Matched to an observation with the first two digits transposed
Z	Matched by S, Y or T and observation time adjusted

All observation times are measured in seconds after 16.10.

APPENDIX A3

Example of basic data from Reading made available for testing the ME2 model

```
number  num. of       observer n.
plate   observations  time, secs after 16.10

534L    4   476234   466276   416330   406408
564D    2   6362407  626259
894K    3   816240   796309    46402
844D    5   6362467  616270   576288   596303   486415
694L    6   576246   596261   606270   546300   376348   406372
514H    7   256252   266258   146330*  186368   756420   286480   246546
BUS AS DETECTED BY PAIR=      5  *****************************************
524L    3   446256   426270   506372
244G    4   476258   466282   416339   406402
284G    2   216258   696534
BUS AS DETECTED BY PAIR=      1  *****************************************
274N    4   476256   516306   546390   576420
134D    3   446259   416276   406336
224R    4   815279   786305   706552   666579
234     3   756282    86372   116426
814K    3   396282   386294   506316
514N    5   6862857  716373Y  726444*  286522   246600
764P    5   6862857  716395   726485   286540   246582
894F    5   766288   746390   756414   266480   246534
BUS AS DETECTED BY PAIR=      5  *****************************************
774N    6   546288   576294   586311   656378S  676456   756498
564L    5   5963007  606314   546330   376366   406384
654     2   616300   576354
864J    3    16300    26312   806469
404L    2   116306   166348
994D    2   496318   486325
574J    3    86318   826336    76516
574D    3   616318   556366   556372
794B    3   446323   426336   486395
474H    3Y  576324   586336   656384
764L    3   646328*  656377   626390
994N    8   576330*  596345*  606355   546366   376372   386396   296414   246474
 84M    6   116330   306426   556456   506498   426570   486641
864N    2   756336    86456
684G    5Y  746336   756354*  706359*  666373Y  636396
974L    3   756351    86468   826480
584C    3   496366   466411   456462
064H    3    26372   786414   766618
114G    4Y   16372    26378   786420   766624
```

APPENDIX A4

Network coding for Reading area, TRADVV format, Unmodified network A

LIST OF THE LINK CARDS READ

ANODE	BNODE	CODE	TIME	DIST	INDEX
1	8	1	1	1	0
1	25	1	3	10	12
4	122	1	4	18	10
4	79	1	2	4	6
5	48	1	999	19	0
5	52	1	7	34	9
5	69	0	5	38	1
5	53	0	7	43	9
8	64	1	2	4	9
10	94	1	2	10	9
10	91	1	4	33	3
10	107	0	3	15	9
20	117	1	3	15	7
20	68	0	3	13	10
20	69	0	3	13	10
20	105	0	3	15	9
25	8	1	3	9	10
25	67	0	4	18	10
25	66	0	3	12	10
25	103	0	3	10	10
30	61	0	4	25	7
30	43	0	3	12	12
40	114	1	3	10	12
40	115	1	4	20	7
40	101	0	3	15	2
41	116	1	2	4	12
41	113	0	3	10	10
41	114	1	5	39	4
42	60	1	3	9	4
42	102	0	3	10	10
43	44	1	3	13	9
44	99	1	5	31	7
44	89	1	1	1	0
44	116	1	4	20	5
45	46	1	5	25	9
45	30	1	4	15	10
46	47	1	4	13	12
46	50	1	3	12	9
47	129	1	3	15	9
47	90	0	999	15	0
47	50	0	3	8	13

48	113	1	2	11	9
48	117	1	3	15	7
49	1	1	4	14	10
49	118	1	999	11	0
49	30	0	999	20	0
49	104	0	3	10	10
50	119	1	4	23	9
50	52	1	4	16	12
51	54	1	5	24	9
51	99	1	5	31	7
51	112	1	4	23	9
52	72	1	3	12	9
52	130	1	3	15	9
53	54	1	4	15	9
53	51	1	4	15	9
54	74	1	2	8	6
55	123	1	1	1	0
55	97	1	2	8	9
56	127	1	3	14	9
56	91	1	7	30	9
57	128	1	3	13	9
57	106	0	3	15	7
58	96	0	8	51	14
58	108	0	2	10	7
58	109	0	2	10	7
60	61	1	3	10	9
60	115	0	3	15	9
61	62	1	2	6	9
62	42	1	3	12	8
62	63	1	2	7	9
63	64	1	3	10	11
63	66	1	3	12	9
64	45	1	6	25	9
66	42	1	4	12	9
66	67	0	5	30	1
67	68	0	3	21	1
68	69	0	4	27	1
72	73	1	4	19	9
73	53	1	4	19	9
73	121	0	3	13	8
74	111	1	3	15	9
74	55	1	3	18	5
79	80	1	4	18	6
79	96	0	8	54	6
80	81	1	2	11	6
81	82	1	2	16	6
81	125	0	4	30	5
82	131	1	4	17	8
82	56	1	2	15	6
88	43	1	3	8	7
89	88	1	3	14	7
89	46	1	5	21	14
90	45	1	999	63	0
91	57	1	7	50	5
91	10	1	5	33	4
94	6	0	3	17	9
94	10	1	3	10	10
96	126	0	9	59	14
97	110	0	3	10	7
97	126	1	1	1	0
99	51	1	6	31	7
99	89	1	5	31	7

111	55	1	3	10	6
114	40	1	3	10	5
114	41	1	5	39	12
115	40	1	4	20	14
116	41	1	2	4	5
116	44	1	4	20	7
117	20	1	3	15	7
117	48	1	3	15	7
118	48	1	299	11	0
118	49	1	2	11	9
119	51	1	4	23	9
119	50	1	4	23	13
121	80	0	3	13	8
122	53	1	2	9	10
123	4	1	4	16	9
125	94	0	6	43	5
126	123	1	2	8	10
127	57	1	2	6	6
128	5	1	2	5	6
129	48	1	3	14	10
130	56	1	3	16	3
131	72	1	4	17	8
C 1	101	0	1	0	0
C 2	102	0	1	0	0
C 3	103	0	1	0	0
C 4	104	0	1	0	0
C 5	105	0	1	0	0
C 6	106	0	1	0	0
C 7	107	0	1	0	0
C 8	108	0	1	0	0
C 9	109	0	1	0	0
C 10	110	0	1	0	0
C 11	111	0	1	0	0
C 12	113	0	1	0	0
C 13	114	0	1	0	0
C 14	115	0	1	0	0
C 15	63	0	1	0	0
C 16	66	0	1	0	0
C 17	64	0	1	0	0
C 18	118	0	1	0	0
C 19	129	0	1	0	0
C 20	117	0	1	0	0
C 21	128	0	1	0	0
C 22	91	0	1	0	0
C 23	127	0	1	0	0
C 24	125	0	1	0	0
C 25	52	0	1	0	0
C 26	46	0	1	0	0
C 27	30	0	1	0	0
C 28	116	0	1	0	0
C 29	54	0	1	0	0
C 30	53	0	1	0	0
C 31	119	0	1	0	0
C 32	131	0	1	0	0
C 33	4	0	1	0	0
C 34	55	0	1	0	0
C 35	96	0	1	0	0
C 37	121	0	1	0	0
C 38	99	0	1	0	0
C 39	6	0	1	0	0
C 40	130	0	1	0	0

READ IN SPEED-FLOW PARAMETERS

LISTING OF THE SPEED FLOW PARAMETERS

INDEX	S0 KM/HR	S1 KM/HR	F1 PCU/HR	F2 PCU/HR
1	80.0	66.0	3400.0	4800.0
2	65.0	56.0	2800.0	4400.0
3	50.0	30.0	1200.0	2200.0
4	45.0	30.0	2200.0	4400.0
5	45.0	25.0	500.0	1000.0
6	40.0	27.0	1850.0	3700.0
7	35.0	25.0	500.0	1000.0
8	25.0	15.0	250.0	700.0
9	35.0	23.0	1400.0	2800.0
10	25.0	10.0	600.0	1000.0
11	15.0	8.0	320.0	680.0
12	20.0	5.0	330.0	750.0
13	16.0	5.0	210.0	420.0
14	35.0	10.0	400.0	700.0

APPENDIX A5

Documentation for the ME2 O-D estimation programs

1. General

This document contains basic information describing the use of the ME2 programs for the estimation of O-D matrices from traffic counts.

The technique is implemented through three programs written in FORTRAN IV and currently implemented in the Leeds University Amdahl 470 machine operating under VMS. The programs have been written to be used in conjunction with the TRADVV suite of traffic assignment models and the programs make extensive use of the TRADVV utilities, in particular for network description, matrix handling, tree building and different assignment algorithms. Of course, it should not be too difficult to modify the programs to interface with other packages. This task is simplified by the fact that these suite specific operations are handled by subroutines the equivalent of which is almost certainly provided by any good assignment package.

This description assumes familiarity with the TRADVV suite. For more information about it consult DVV DOCUMENT.

The Leeds implementation is written for interactive use (the normal use in the Amdahl machine at Leeds University). This interactive use relates only to the determination of the basic parameters controlling the algorithms and this INPUT/OUTPUT is restricted to a few subroutines. For batch use it is quite simple to replace these routines for either a &NAMELIST or card reader input in some suitable format.

With the exception of these interactive intput/output routines the FORTRAN code is considered to be machine-dependent.

The functions of the three programs in this suite are as follows.

METWO1 This program checks a set of observed traffic counts to ensure that flow continuity conditions are maintained at all nodes in the network. If these conditions are not met the program modifies the flows to ensure the satisfaction of these conditions.

METWO2 This program accepts as input a set of counts and if available, a prior trip matrix. In addition to this it uses either a mark 6 file (see below) or simply a tree file to estimate a trip matrix.

METWO4 This program performs tree and mark 6 file manipulation.

2. Outline of the technique

The ME2 suite uses an entropy maximising formalism to estimate the most likely trip matrix consistent with the information contained in a set of counts (on links, turning movements, entrances and exits of motorway/rapid transit system). For each count a link must be defined in the network description. Usually many more links than traffic counts will be coded to represent a network.

In addition to the counts and network description the model requires an array containing 'the proportions of trips from each O-D pair using each counted link' or Pija factors. These are usually calculated from an assignment package. In this suite the Pija factors are extracted from TREEs built by the TRADVV progra, T1 and are stored in a TRADVV file mark 6. (A mark is a filetype identified in the TRADVV suite. All TRADVV files are kept in binary unformatted form for efficiency and the mark identifier enables the correct interpretation of these files. The mark 6 identifier is reserved for this type of file used by the ME2 programs.) A mark 6 TRADVV file contains the usual network file information (see TRADVV DOCUMENT) plus the Pija factors corresponding to a particular assignment.

The algorithm in outline consists of taking each counted link in turn and comparing the estimated with the observed flow. If the estimated flow is different from the observed one all the O-D pairs using that particular link are modified (weighted by Pija) so that both flows are equal. The process is repeated for all counted links until all links are processed and this is considered here to be one iteration. Successive iterations process again the whole of the counted link list until convergence is achieved.

3. METWO1 program

This program corrects observed link flows so that flow continuity conditions are met at all nodes. The ME2 model can estimate an O-D matrix without recourse to METWO1 but its use goes a long way at ensuring true convergence is achievable in METWO2.

Input

The program requires

a TRADVV network file with the structural characteristics of the network

a set of traffic counts, either via cards or via a TRADVV matrix file

Output

The program produces a corrected version of the traffic counts in the form of a TRADVV matrix file. It also lists them out on a printer

Channels

5	Terminal
6	Line printer
8	Network file
9	Observed flows file (optional)
10	Corrected flow file suitable for input to METWO2
1	Observed flows on cards (optional)

Parameters

None, the program will prompt the user for a run title and for the channel to be used to input the flows.

4. METWO2 program

This program estimates a trip matrix from traffic counts. The following inputs are required.

Input

channel	description
5	terminal
6	line printer
7	prior trip matrix, if any (optional) in TRADVV format
8	tree file (optional) from TRADVV.T1
9	observed flow file, in TRADVV matrix format
10	TRADVV mark 6 file (optional)

Output

10	TRADVV mark 6 file, if not provided originally
12	estimated trip matrix in TRADVV matrix format

Parameters

The program will prompt the user for a run title and:

OLDPJ	True if an old mark 6 file is to be used, otherwise false
ITERMX	Maximum number of iterations
IPRINT	Print out intermediate results every IPRINT iterations
EPSILN	Relative error at link flow level admissable for terminating the iterations. The program will terminate iterations when ITERMX has been reached or if all modelled link flows are within EPSILN observed flows. Recommended value is 0.05.
VZ	True if the Van Zuylen variant of the model is to be used.

5. METWO4 program

This program manipulates TRADVV mark 6 files either creating them from a tree file or combining them before being used by METWO2. The basic manipulations are:

reading a tree file and creating a mark 6 file

or

reading a tree file and combining the resulting Pija factors with an old mark 6 file. This combination will take the Pija factors, multiply them by

LAMBDA (between 0 and 1) and add to the old Pija factors multiplied by 1-LAMBDA. This option is useful when using the programs in an equilibrium assignment framework.

In the near future an option allowing up to 5 sets of Pija factors to be used in parallel without combining them will be implemented.

Input

channel	description
3	tree file, if any
5	terminal
6	line printer
7	first mark 6 file
8	second mark 6 file
9	combined mark 6 file
10	scratch file, unformatted

Parameters

The program will prompt the user for a run title and the following parameters.

CREAND	True if a tree file is to be read and a mark 6 file to be produced
CRONLY	True if only one mark 6 file is to be created
COMB	True if the created mark 6 file is to be combined with an old one
LAMBDA	Parameter between 0 and 1 for combing the mark 6 files.

APPENDIX A6

<u>Network coding for Reading area, TRADVV format, Modified network B</u>

LIST OF THE LINK CARDS READ

ANODE	BNODE	CODE	TIME	DIST	INDEX
1	8	1	0	1	0
1	25	1	3	10	8
4	122	1	4	18	11
4	79	1	2	4	2
5	48	1	999	19	0
5	52	0	7	34	1
5	69	0	5	38	1
6	58	0	7	43	3
8	64	1	2	4	2
10	94	1	2	10	2
10	91	1	4	38	2
10	107	0	3	15	3
20	117	1	3	15	2
20	58	0	3	13	10
20	69	0	3	13	10
20	105	0	3	15	3
25	8	1	3	9	7
25	67	0	4	18	10
25	66	0	3	12	10
25	103	0	3	10	3
30	61	0	4	25	10
30	43	0	3	12	10
40	114	1	3	10	6
40	115	1	4	20	4
40	101	0	3	15	2
41	116	1	2	4	5
41	113	0	3	10	10
41	114	1	5	39	6
42	60	0	3	9	4
42	102	0	3	10	10
43	44	1	3	13	5
44	99	1	5	31	4
44	89	1	1	1	0
44	116	1	4	20	5
45	46	1	5	25	10
45	30	1	4	15	5
46	47	1	4	13	12
46	50	1	3	12	12
47	129	1	3	15	9
47	90	0	999	15	0
47	50	0	3	8	10

48	118	1	2	11	9
48	117	1	3	15	2
49	1	1	4	14	4
49	118	1	999	11	0
49	90	0	999	20	0
49	104	0	3	10	10
50	119	1	4	23	2
50	52	1	4	16	10
51	54	1	5	24	5
51	99	1	5	31	9
51	119	1	4	23	2
52	72	1	3	12	4
52	130	1	3	15	3
52	44	0	10	65	1
53	54	1	4	15	5
53	51	1	4	15	7
54	74	1	2	8	4
55	123	1	1	1	0
55	97	1	2	8	9
56	127	1	3	14	4
56	91	1	7	30	4
57	128	1	3	13	4
57	106	0	3	15	7
58	96	0	8	51	7
58	108	0	2	10	7
58	109	0	2	10	7
60	61	1	3	10	4
60	115	0	3	15	4
61	62	1	2	6	2
62	42	1	3	12	7
62	63	1	2	7	9
63	64	1	3	10	11
63	66	1	3	12	9
64	45	1	6	25	7
65	42	0	4	12	4
66	67	0	5	30	1
67	68	0	3	21	1
68	69	0	4	27	1
72	73	1	4	19	6
73	53	1	4	19	7
73	121	0	3	13	8
74	111	1	3	15	4
74	55	1	3	18	7
79	80	1	4	18	6
79	96	0	8	54	6
80	81	1	2	11	6
81	82	1	2	16	6
81	125	0	4	30	6
82	131	1	4	17	8
82	56	1	2	15	7
88	43	1	3	8	7
89	88	1	3	14	6
89	46	1	5	21	6
90	45	1	999	63	0
91	57	1	7	50	5
91	10	1	5	38	3
94	6	0	3	17	5
94	10	1	3	10	3
96	126	0	9	59	6
97	110	0	3	10	7
97	126	1	0	1	0
99	51	1	6	31	10

	99	99	1	5	31	7
	111	55	1	3	10	7
	114	40	-1	3	10	5
	114	41	1	5	39	5
	115	40	1	4	20	5
	116	41	1	2	4	5
	116	44	1	4	20	5
	117	20	1	3	15	3
	117	48	1	3	15	3
	118	48	1	999	11	0
	118	49	1	2	11	10
	119	51	1	4	23	10
	119	50	1	4	23	10
	121	80	0	3	13	9
	122	53	1	2	9	11
	123	4	1	4	16	7
	125	94	0	6	43	6
	126	123	1	2	8	10
	127	57	1	2	6	5
	128	5	1	2	5	5
	129	48	1	3	14	10
	130	56	1	3	16	3
	131	72	1	4	17	8
C	1	101	0	1	0	0
C	2	102	0	1	0	0
C	3	103	0	1	0	0
C	4	104	0	1	0	0
C	5	105	0	1	0	0
C	6	106	0	1	0	0
C	7	107	0	1	0	0
C	8	108	0	1	0	0
C	9	109	0	1	0	0
C	10	110	0	1	0	0
C	11	111	0	1	0	0
C	12	113	0	1	0	0
C	13	114	0	1	0	0
C	14	115	0	1	0	0
C	15	63	0	1	0	0
C	16	66	0	1	0	0
C	17	64	0	1	0	0
C	18	118	0	1	0	0
C	19	129	0	1	0	0
C	20	117	0	1	0	0
C	21	128	0	1	0	0
C	22	91	0	1	0	0
C	23	127	0	1	0	0
C	24	125	0	1	0	0
C	25	52	0	1	0	0
C	26	46	0	1	0	0
C	27	30	0	1	0	0
C	28	116	0	1	0	0
C	29	54	0	1	0	0
C	30	53	0	1	0	0
C	31	119	0	1	0	0
C	32	131	0	1	0	0
C	33	4	0	1	0	0
C	34	55	0	1	0	0
C	35	36	0	1	0	0
C	37	121	0	1	0	0
C	38	99	0	1	0	0
C	39	6	0	1	0	0
C	40	130	0	1	0	0

READ IN SPEED-FLOW PARAMETERS

LISTING OF THE SPEED FLOW PARAMETERS

INDEX	S0 KM/HR	S1 KM/HR	F1 PCU/HR	F2 PCU/HR
1	80.0	66.0	3400.0	4800.0
2	65.0	55.0	2800.0	4400.0
3	50.0	30.0	1200.0	2200.0
4	45.0	30.0	2200.0	4400.0
5	45.0	25.0	500.0	1000.0
6	40.0	27.0	1850.0	3700.0
7	35.0	25.0	500.0	1000.0
8	25.0	15.0	250.0	700.0
9	35.0	23.0	1400.0	2800.0
10	25.0	10.0	600.0	1000.0
11	16.0	3.0	320.0	680.0
12	20.0	5.0	380.0	750.0
13	16.0	5.0	210.0	420.0
14	35.0	10.0	400.0	700.0