The Feasibility of Synthesizing an Origin-Destination Matrix for the Manitoba Road Network Using Routinely Collected Traffic Counts

Greg Mason and Associates
Winnipeg, Manitoba
May 1981

1.0 Executive Summary

a. Of the recent algorithms devised to synthesize origin-destination information from link flow data, the regression model based on the gravity model appears to be the most attractive both with respect to potential accuracy and computational viability.

b. A major problem in the application of this (or any other) synthesizing technique to Manitoba is that there is no benchmark trip interchange data to validate the synthesizing model. Validation would have to proceed by comparing the estimated link flows to the observed link flows. Several quite different 0-D matrices can generate similar link flow patterns, or alternatively, the same table may generate different link flows depending on the route assignment. The density of the provincial highway and road network could render simple all-or-nothing route assignments invalid.

c. The data available from the Department of Highways, while generally of high calibre, was not collected with a view to use in a synthesizing experiment. At each origin and destination important assumptions will be needed to convert the traffic counts into consistent link flows. It is unclear at this time what impacts these assumptions will have upon the final model outcome. These conversion assumptions will most likely be crucial in the Winnipeg region. Testing alternative conversion assumptions is a major cost in the synthesizing experiment.

d. Given the cost and the uncertainty associated with the state of the art it is hazardous to proceed with 0-D synthesis in the absence of benchmark trip interchange information developed by conventional survey procedures.

e. A research strategy of a small (n = 2000) mail out survey using a random sample (probability proportional to size) and the Autopac renewal canvass is a viable alternative to the O-D synthesis. The cost of this approach is competitive with the cost of 0-D synthesis, and after this information has been obtained, the 0-D synthesis can proceed with greater assurance of model results.

2.0 Introduction and Outline

This study reviews the recent developments in transportation planning with special emphasis on the synthesizing of origin-destination (0-D) tables from routinely collected link flow data. Further, this study evaluates the feasibility of applying these techniques to the data collected on the Manitoba highway network and compares the usefulness of this approach with a survey conventionally used to develop such information.

Origin-destination tables are an essential part of transportation planning. Coupled with a route assignment model they permit an analysis of future demand for specific links; allow an analysis of service implications when links are added or deleted; permit the analysis of cost both to the public provider and the private consumer as links are added or deleted, and as links are maintained at different levels. They also permit an analysis of the impacts of energy shortage on probable trip-making behaviour and flows of traffic along specific links.

This study focuses narrowly on the feasibility of synthesizing origin-destination tables from the link flow data at hand. There seems little point to developing an overall transportation planning model which incorporates a variety of policy initiatives unless the integrity of the synthesizing model has been assured.

The two approaches which are most attractive from a computational perspective are gravity-based regression models and entropy models. The former have the most support from the literature and have been validated against real world situations involving reasonably large matrices (40 x 40), while the latter represent recent, state-of-the-art approaches. The current off-the-shelf computer programs are also briefly described, although the literature available on most of these is relatively sketchy. Often they are sub-models in a larger study. The Manitoba data base contained in the Community Reports and the Highway statistics is evaluated with respect to its usefulness in developing an O-D synthesis. Finally, a research program is sketched to develop an O-D synthesizing model, and to develop benchmark interchange information crucial to the validation of the model.

3.0 A Review of Transportation Planning Methods

A majority of origin-destination synthesizing models are founded upon one of the two dominant paradigms currently in use. The first, which can be called the traditional transportation planning system (TTPS), is a sequential modelling process that feeds information recursively into subsequent elements. The second, commonly called network equilibrium models, and combined distribution-assignment models (CDA), views the transportation system as a simultaneous optimizing model.

3.1 Traditional Transportation Planning System

This sequential procedure has four steps as shown in Figure 3.1.1

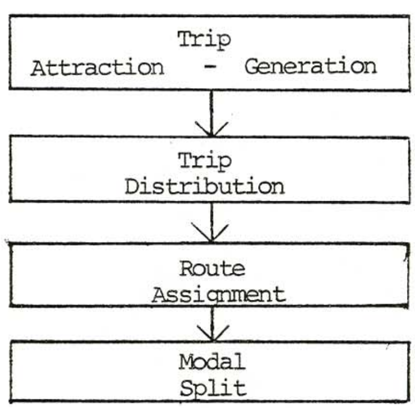


Figure 3.1.1

a. Trip Generation/Attraction

The first stage is usually a regression-based method which estimates the number of trips emanating from node i (0_i) and the number of trips destined for node j (D_j) as functions of socio-economic variables such as population, employment, recreational opportunity and the like. In some cases, proxies for 0i and Dj are used such as the population at i (P_i) and the employment at j (E_j). Earlier analytic trip generation/attraction models have largely been replaced by regression methods. For the purposes of 0-D synthesis, this step is of little interest.

b. Trip Distribution

This stage is the heart of the TTPS. A trip interchange matrix is prepared, usually from survey data, and related to various socio-economic variables at the nodes. Subsequent changes in the attributes of these nodes are then introduced to predict changes in the trip interchange matrix.

Overwhelmingly, the gravity model is used;

$$T_{ij} = \frac{0_i D_j}{c_{ij}^m} \qquad \text{...3.1.1}$$

where 0_i is the number of trips emerging from i, D_j the number of trips destined to j, c_{ij} an impedance function reflecting the cost of travel and m a parameter. An alternate expression is

$$T_{ij} = \frac{\alpha P_i E_j}{c_{ij}^m} \qquad \text{...3.1.2}$$

where P_i is the population at i and E_j the employment at j. In most cases c_{ij} is measured as time cost on simple linear distance between i and j.

The calibration of gravity models revolves around the need to satisfy the following constraints

$$\sum_i T_{ij} = D_j \qquad \text{...3.1.3a}$$

$$\sum_j T_{ij} = O_i \qquad \text{...3.1.3b}$$

$$\sum_{ij}\sum T_{ij} = T. \qquad \text{...3.1.3c}$$

The second of these (eqn. 3.1.3b) is always satisfied, but the first and third require a balancing procedure to adjust α and m.

Once α and m are found, subsequent values of P_i and E_j can be introduced to produce a new origin-destination matrix.

The gravity model has the advantage of good predictive power over smaller homogeneous networks where no major transport innovations have occurred, such as new freeways. It generally outperforms both the competing and intervening opportunities models. A drawback is that a doubling of P_i and E_j leads to a quadcoupling of T_{ij} which is clearly wrong. An oft voiced criticism is that gravity models have very little grounding in behavioural science and do not provide insight or incorporate information on why travellers make certain decisions.

c. Route Assignment

Another element from the TTPS used in O-D synthesis is route assignment which allocates internodal trips to certain routes or paths through the networks. The simplest assignment is the shortest path or the all-or-nothing assignment (AN). The probability of using a link "a" between i and j given by P^a_{ij} is either zero or one depending on whether it is part of the shortest path or not. This algorithm works quite well provided that the network does not have congestion, that the nodes are quite separated and that any but the shortest path imposes significant additional travel costs. A regional transport network is a likely candidate.

In urban transportation, where congestion effects can induce travellers to change routes, or even destinations to avoid time delays, the all-or-nothing assignment has given way to a variety of probability models. Perhaps the best known is Dial's multipath assignment which ranks routes according to distance and assigns travellers to links between i and j proportionally and probabilistically. Of course the shortest path has the greatest patronage. Other multipath methods incorporate threshold effects and learning curves in order to simulate more accurately typical traveller behaviours. Since these procedures are most important for congested urban areas, they will not be pursued here.

d. Modal Split

Modal split models can be incorporated as the fourth step as shown in Figure 3.1.1 or as a step prior to route assignment. The techniques of choice are the probabilistic choice models (probit and logit). The modal split is generally not included in 0-D synthesis at the outset but do come in when 0-D matrices are incorporated into the planning model.

3.2 The Synthesis of 0-D Matrices

The general approach (which is detailed in Section 4.0) involves the specification of a gravity model

$$T_{ij} = \frac{P_i E_j}{c_{ij}^m} \qquad \text{...3.2.1}$$

where P_i and E_j are as defined above, c_{ij} is usually linear distance or travel time and m an initial value. The trial 0-D matrix (T^*_{ij}) is then coupled with a route assignment which produces p_{ij}^a in the following equation:

$$V^*_a = T^*_{ij} p^a_{ij} \quad \text{for all } i,j \qquad \text{...3.1.2}$$

The trial V^*_a is compared to the actual link flows V_a and m is adjusted until there is a "match". This procedure is similar to the calibration procedure of a gravity model, and does not appear to necessitate the collection of trip interchange information from surveys. It is important to stress, however, that the individual steps such as the route assignment procedure, the validity of the gravity model itself and the accuracy of the link flow data must all be very high. It is quite possible to have substantially different T_{ij} generate similar values of link flow. In general, without corroborationg trip information, it is very hazardous to validate this procedure using secondary information (i.e., link flows).

3.2.3 Network Equilibrium Models

The assumption that the trip distribution phase and the route assignment phase of travelling behaviour are separate is wrong, especially for urban methods. Even for relatively dense regional networks or for modelling the travel behaviour around relatively large urban areas, congestion can often induce substitution in destinations. Network equilibrium models regard the distribution-assignment problem as a simultaneous optimizing model (mathematical program). An objective function, total system cost, or "entropy," is minimized or maximized subject to a set of constraints. These constraints conserve flows on various links in the network, ensure there is no doubling back, and that the trip interchange table produced actually conforms to the origin-destination information at hand. An underlying philosophy to the design of network equilibrium model is given by Wardrop's principle.

3.2.4 Wardrop's Principle

1. All routes actually used have equal cost (to the traveller) and any unused route must have a greater cost than any route actually travelled.

2. The total travel costs on the system are a minimum.

The first principle states that average travel costs on the routes chosen are equal, while the second requires that the marginal costs of 0-D pairs are equal. In actual practice, the first principle is taken as the most realistic. The second principle gives rise to a very simple linear program.

3.2.2 Mathematical (Linear) Program of Network Equilibrium

If the cost of travelling from i to j is given by c_{ij}, then the total cost of the system is given by

$$\sum_{i}\sum_{j} c_{ij}T_{ij}. \qquad \text{...3.2.1}$$

This expression is minimized with respect to

$$\sum_{j} T_{ij} = O_i \qquad \text{...3.2.2a}$$

$$\sum_{i} T_{ij} = D_j \qquad \text{...3.2.2b}$$

$$\Sigma\Sigma T_{ij} = T \qquad \text{...3.2.2c}$$

$$T_{ij} \geq 0, \quad c_{ij} \geq 0.$$

Given the costs (c_{ij} (perhaps simple linear distance, or time) and information on originating trips, destinating trips and total trips, this program will produce the 0-D matrix T^*_{ij} consistent with system wide cost minimizations. Needless to say, such a model suffers from the assumptions of linearity, no non linear cost effects, no congestion effects, and finally, it is doubtful that travellers arrange their personal travel decisions to reduce total system costs.

3.2.3 Combined Distribution-Assignment Models

Wardrop's two principles can be unified by link capacity functions which express link cost (C_a) as a function of link volume (V_a). The total system cost can now be expressed as

$$\sum_a C_a(V_a) = \sum_{ij}\sum c_{ij}(T_{ij}). \qquad ...3.2.3$$

If link flow constraints are added to reflect rational travel behaviour, etc., a non linear mathematical program can be specified.

$$\text{Min} \sum_a C_a(V_a) \qquad ...3.2.4$$

subject to

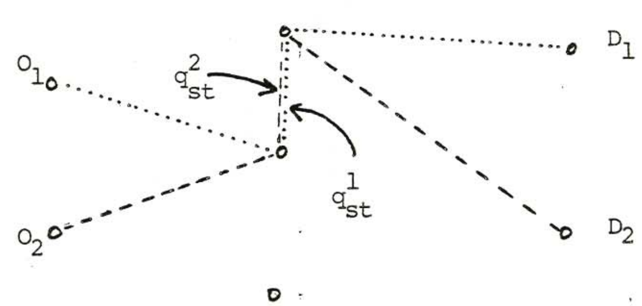
$$q_{st} = \sum_i q^i_{st} \quad \text{for all } s,t \qquad ...3.2.5$$

$$\sum_k q^i_{kj} - \sum_k q^i_{jk} = T_{ij} \quad \text{for all } i,j \qquad ...3.2.6$$

$$\sum_{st}\sum q^i_{st} \geq 0 \quad \text{for all } T \qquad ...3.2.7$$

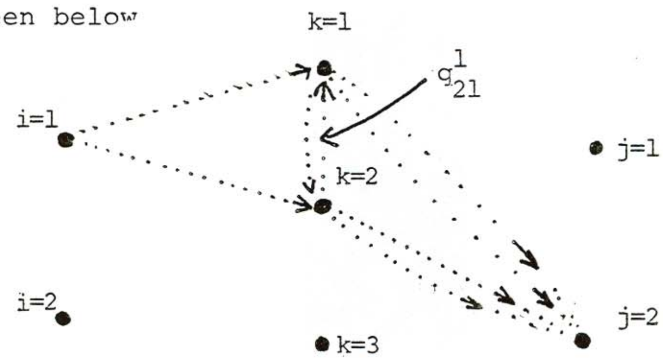
where q^i_{st} is the flow from node s to node t of a trip that originated at i.

Consider a simple network of two origins and two destinations. There are also three intermediate nodes.



The first constraint merely ensures that the flow on q_{st} is the sum of flows from s to t which originates at i. In this case the two flows marked with dots.

The second constraint (3.2.6) indicates that the net flow between points k and j which originate at i, for all values of k and j, is the trip interchange matrix. This can be seen below



The net sum of trips going from i to j is given by the solid flows from i to 1, 1 to 3, 3 to 2 and 2 to j, the dashed flow from i to 1, 1 to 2 and 2 to j less any flows in the reverse direction.

The final constraint simply ensures that each i contributes a positive flow over the network.

This model has been refined considerably, but the presentation above certainly gives the flavour of the approach. With respect to 0-D synthesis, these models have been used to replicate very small experimental networks. There ~~is~~ are apparently no experimental results obtainable from larger networks.

3.2.4 Entropy Models

A major advance in transportation modelling was introduced by Wilson in the form of entropy. As defined by information theorists, entropy is a measure of probability. As defined by transportation analysts, entropy is a measure of accessibility. Entropy maximization of a trip interchange table is a state where all trips between i and j are equally probable. Consider the fraction of trips per cell in an 0-D matrix T_{ij} given by T_{ij}/T. Maximum accessibility could be associated with a situation where each entry in the 0-D matrix produced a T_{ij}/T which was constant. Uneven distributions of trips over the cells might be associated with reduced accessibility From information theory the entropy of this 0-D matrix is defined as

$$H = - \sum_{ij} \frac{T_{ij}}{T} \log \frac{T_{ij}}{T}. \qquad \text{...3.2.8}$$

If T_{ij}/T is a constant, then over a total of n cells $T_{ij}/T = \frac{1}{n}$ and

$$H = -n \frac{1}{n} \log n = \log n.$$

If T_{ij}/T is very uneven, then H approaches 0. The constraints that operate on entropy maximizations are the usual constraints associated with trip interchange tables, namely

$$\sum_{j} T_{ij} = O_i \qquad \text{...3.2.9a}$$

$$\sum_{i} T_{ij} = D_j \qquad \text{...3.2.9b}$$

$$\sum_{ij} T_{ij} = T. \qquad \text{...3.2.9c}$$

As these constraints bind the value of H, travellers have less freedom over choosing their origin destination combinations.

It is possible to incorporate efficiency into the model. Let the average cost per journey be defined as

$$C = \frac{1}{T} \sum_{ij}\sum c_{ij} T_{ij}. \qquad \text{...3.2.10}$$

If this function is minimized subject to the above constraints (eqns. 3.2.9a - 3.2.9c) and an additional constraint regarding the minimal acceptable accessibility level, namely

$$H = -\sum_{ij}\sum \frac{T_{ij}}{T} \log \frac{T_{ij}}{T} \geq H_0, \qquad \text{...3.2.11}$$

then a trip interchange matrix T^*_{ij} will be located which minimizes average journey cost and achieves some minimum level of accessibility.

Entropy based modelling may appear highly abstract, and it is, but it does give rise to quite simple algorithms for synthesizing O-D networks from link flow data. These are outlined in part 4.

3.2.5 Combining the Entropy and the CDA Model

The state of the art in this work is reached when entropy concepts and link flow cost ideas are united in one model. The objective function appears as

$$\text{Min } \lambda \sum \int_0^{V_a} C_a(x)dx - \sum\sum T_{ij} \log T_{ij} \qquad \text{...3.2.12a}$$

$$\text{S.T.} \quad \sum_i T_{ij} = D_j \qquad \text{...3.2.12b}$$

$$\sum_j T_{ij} = O_i \qquad \text{...3.2.12c}$$

$$\sum_k q^i_{kj} - \sum_k q^i_{jk} = T_{ij}. \qquad \text{...3.2.12d}$$

The objective function minimizes total link cost and the negative of entropy (the same as maximizing entropy). The constraints are relatively straightforward adaptations from previous models.

O-D synthesis is also being attempted using this framework, but once again, the work is very experimental and preliminary (Nguyen, 1978).

3.2.6 Summary

Transportation modelling comprised either of the sequential transportation planning system (TTPS) or the network equilibrium models are the basis for current O-D synthesizing efforts. The first, with its reliance on gravity formulation, has produced a variety of regression based procedures, while the second, especially the entropy approach, has yielded some very interesting algorithms based upon estimating the most likely way a set of link flows could be generated from a trip interchange table.

4.0 SYNTHESIZING ORIGIN-DESTINATION MATRICES FROM LINK-FLOW DATA

Introduction

Several approaches have been developed to take routinely collected traffic information and synthesize an origin destination matrix. For the most part, these methods fall neatly into regression based methodologies which exploit the gravity model, and network equilibrium models which synthesize the O-D matrix in a mathematical program. This section outlines significant examples of each type of model, examines the experimental basis for evaluating their predictive power, and summarizes under which context a particular technique would prove most fruitful.

4.1 Regression Based Procedures

The general paradigm is quite simple. First a gravity model is specified using information on the origin and destinations such as population at i and employment at j. An initial value of m is chosen and a trial trip interchange matrix is given by

$$T^*_{ij} = \frac{P_i E_j}{c^m_{ij}}. \qquad \ldots 4.1.1$$

The trial trip information is then assigned to various routes according to some algorithm such as the all or nothing assignment or a multipath model. The route assignment information produces a matrix p^a_{ij} which is used in conjunction with T^*_{ij}

to produce estimated link volumes, V_a^* according to

$$V_a^* = T_{ij}P_{ij}^a. \qquad \ldots 4.1.2$$

Now, V_a^* is compared to the actual link flow data and the trip matrix T_{ij} recalculated using a different value of m until the difference between V_a^* and V_a is minimized.

Regression based methods set up the least squares criterion as the appropriate way of comparing V_a^* and V_a. These may be further classified into linear and non linear algorithms.

4.1.1 Linear Regression Methods

The method of Low (1972) follows the above paradigm very closely. The criterion for accepting a particular value of m is that it minimizes the expression $\Sigma(V_a - V_a^*)^2$ where V_a are the observed link volumes and V_a^* are computed from a gravity model and all or nothing route assignment. It is possible to consider several journey types or even modes in addition to work trips by car. In this case the calculation of V_a^* can be done according to a multiple regression

$$V_a^* = \alpha_0 + \alpha_1 \frac{P_i E_j}{c_{ij}^m} + \alpha_2 \frac{P_i P_j}{c_{ij}^m} \qquad \ldots 4.1.3$$

where the second term $\left(\frac{P_i P_j}{c_{ij}^m}\right)$ can be used to model recreational trips. The parameter m is common to both formulations, an obvious deficiency.

Another possible variant of this approach is the χ^2 test formed by

$$\chi^2 = \frac{\Sigma(V_a - V_a^*)^2}{V_a} \qquad \text{...4.1.4}$$

with $(n-1)^2$ degrees of freedom (n is the number of zones in the 0-D matrix and it is assumed the matrix T_{ij} is square). This test is a little more rigorous since it is possible for the model to fail to never reject the null hypothesis that V_a and V_a^* are the same. In this case the model will be abandoned. For the least squares approach merely finding a low sum of squares does not ensure statistical adequacy of the results.

Finally, it should be stressed that the criterion for success is whether the model reproduces observed link volumes. In a sense, this is at best an indirect test. It is conceivable that substantially different 0-D matrices could produce the same link volumes. This is probably true of denser regional networks and urban systems where an all or nothing assignment is invalid. For rural areas the methodology is quite robust.

More complicated approaches have been developed by Overgaard (1974) and Lamarre (1975). In particular, the form of the impedance function (given above as c_{ij}^m) has been amended. For example, Overgaard has a step function given by

$$c_{ij} = \begin{cases} 1.5t_{ij}^{2.8} & \text{for } t_{ij} \leq 1.5 \text{ hours} \\ t_{ij}^{1.8} & \text{for } t_{ij} \geq 1.5 \text{ hours} \end{cases} \qquad \text{...4.1.5}$$

Another wrinkle is to model the zone characteristics with regressions; in a sense this is the incorporation of trip generation/attraction functions into the gravity model. A typical equation is to model 0_i as

$$O_i = b_1E_i + b_2P_i + b_3S_iP_i. \qquad ...4.1.6$$

The Australian Commonwealth Bureau of Roads has constructed a model using central place theory. The impedance function is adjusted depending on whether the trip is to a low or high order centre. The basic model appears as

$$T^*_{ij} = \sum_r b_r f^r_{ij} \qquad ...4.1.7$$

where b_r is the parameter to be estimated, f^r_{ij} is the impedance between nodes i and j(i ≠ j), and r gives the order of the destination centre. Once T^*_{ij} has been given the link flows are calculated in the usual way using an all or nothing assignment. The criterion for adjusting b_r is the usual least squares procedure.

$$V^*_0 = \sum_{ij}\sum (\sum_r b_r f^r_{ij}) P^a_{ij} \qquad ...4.1.8$$

The values of f^r_{ij} is of the form $P_i c^{-k}_{ij}$ and is estimated for a total of six trip types (i.e., r goes from 1 to 6). This is a major weakness of the model in that the classification of trip purpose (that is trip to centres of various order) can be arbitrary and require survey data.

4.1.2 Non-Linear Methods

By making the impedance functions more complex, a non linear specification is produced. For example, Hoberg (1974) uses a function of the form

$$f(c_{ij}) = b_1 c_{ij} + b_2 c_{ij}^2 \qquad \ldots 4.1.9$$

where c_{ij} is the log of travel time. A trip interchange matrix is generated according to

$$T^*_{ij} = b_3 A_i O_i D_j f(c_{ij}) \qquad \ldots 4.1.10$$

where the factor A_i is a balancing factor to ensure row totals are consistent (see the discussion of the gravity model above). An all or nothing assignment is used to generate predicted values of link volumes. Now, however, the least squares function $\Sigma(V^*_a - V^*_a)^2$ is non-linear. Three parameters need to be calculated, namely b_1, b_2 and b_3 in what is clearly a non-linear least squares problem. The Marquardt algorithm can be used, but potential problems exist in the "tu~~r~~ning" of the results. Thus far, the published tests of the model concentrate on replicating simple matrices, although Lamarre (1975) reports that the non-linear least squares model performs better than the linear specification.

An important feature of this model is that is produces reasonably close results using only partial information. The addition of information from less important trips and links does not alter the trip interchanges between major areas, an important consideration for sensitivity analysis.

Robillard (1975) takes the model one step further by introducing balancing factors on both the rows and the columns. Formally the trial T_{ij} is calculated by

$$T_{ij} = R_i S_j f(c_{ij}) \qquad \ldots 4.1.11$$

where $R_i = A_i O_i$, and $S_j = B_j O_j$ (and where O_i and D_j are approximated by P_i and E_j). The trial link volumes are selected by proportional assignment and the problem is to minimize the function

$$[V_a - \sum_{ij}\sum R_i S_j f(c_{ij})]^2. \qquad \ldots 4.1.12$$

The impedance function $f(c_{ij})$ is non-linear, and the multiplicative nature of the balancing functions makes this a non-linear least squares problem.

Yet another non-linear model was developed by Wills (1976) who defines T^*_{ij} as a function of a set of k origin parameters at i (A_{ik}) and a set of k destination parameters at j (B_{jk}).

$$T^*_{ij} = b_0 \prod_k (A_{ik} B_{jk})^{b_k}. \qquad \ldots 4.1.13$$

The estimated link volumes are

$$V^*_a = \sum\sum T_{ij} P^a_{ij} \qquad \ldots 4.1.14$$

$$= \sum_{ij}\sum b_0 \prod_k (A_{ik} B_{jk})^{b_k} P^a_{ij}.$$

The usual minimand is used

$$\text{Min} \sum_a [V_a - V^*_a)^2. \qquad \ldots 4.1.15$$

The objective attributes may be P_i and E_j at which point this model collapses into the standard linear model. As other attributes of i and j are added such a recreational indices and the like, the number of parameters that need to be estimated $(b_0, \ldots, b_k)$ increases, necessitating the use of non-linear least squares.

This model has been tested in British Columbia for 37 centres and 3 parameters and has produced useful results. At this time a copy of the test results and report have not been received and detailed discussion is not possible.

4.2 Appraisal of Regression Models

The regression models all use the same paradigm. A trial trip interchange matrix is developed, usually from the gravity model. A route assignment, most often the all or nothing algorithm, is used to develop trial link volumes, which are then placed into the least squares framework. The complexity of the trip interchange model governs whether the model will be a linear or a non-linear least squares problem. The validation of the model is generally done by comparing and making as small as possible the difference between the estimated and actual link volumes.

On the one hand this validation procedure is simple and cost effective, but on the other hand, there is no assurance that the final value of T^*_{ij} corresponds to the true origin-destination matrix. A T^*_{ij} may produce link volume very close to what is observed, yet actually deviate significantly from the true T_{ij}. In a regional transport system, the all or nothing assignment will probably work well except for trips undertaken within the market shadow of Winnipeg. It is uncertain how one dominant node influences the calculation of T^*_{ij}, but most likely it will reinforce the tendency of regression models to underestimate small values in T_{ij} and overestimate larger interchanges.

Without expressly testing T^*_{ij} against the actual values, obtained by conventional methods, no assurance can be given as to the accuracy of these, or any 0-D synthesizing models for that matter. At this moment research is continuing by developing comprehensive T_{ij} values from well designed surveys and then experimenting with alternate models.

4.3 Mathematical Programming Models Based Upon Entropy and Combined Distribution-Assignment Methods

Both the entropy and network equilibrium paradigms have produced algorithms for synthesizing 0-D matrices. The former has produced a very simple procedure due to Williamsen (1978), and this method will be presented in some detail as it offers the greatest opportunity for application in the province. The latter methods are being developed to deal with congestion effects in urban networks and are not really applicable to the problem at hand.

4.3.1 Entropy Based Methods for O-D Synthesis

The essence of this procedure is very simple. A given set of link flows can be generated by a finite number of trip interchange matrices. Using some basic ideas from probability, the likelihood of a given trip interchange matrix generating a set of observed link volumes is computed. Clearly, that matrix T_{ij} which has the greatest likelihood of generating the observed flow is then selected as the final O-D matrix. An example serves to illustrate the procedure.

Consider a simple street intersection (this model is entirely applicable to regional networks) as shown in Figure 4.3.1.

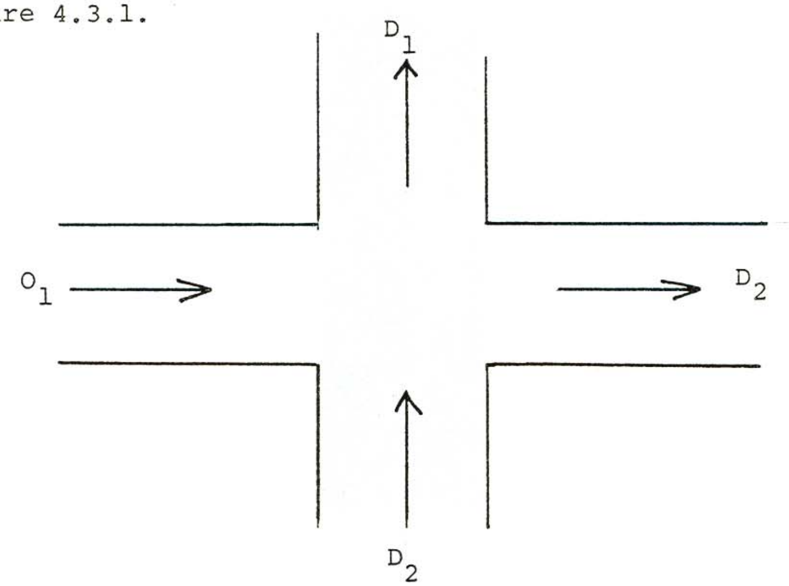


Figure 4.3.1

This situation may be summarized by a trip interchange table as shown below.

Destination

Origin	1	2	
1			2
2			4
Σ	3	3	

Figure 4.2.2

Notice that the only information at hand is the volume of traffic emerging from origins 1 and 2 (2 and 4 respectively) and the volume destined to 1 and 2 (3 in each case). The problem is to infer values for a, b, c and d (which is precisely the O-D synthesis).

Given these row and column totals there are only three possible matrices consistent with the flow data.

I

	1	2	
1	0	2	2
2	3	1	4
	3	3	

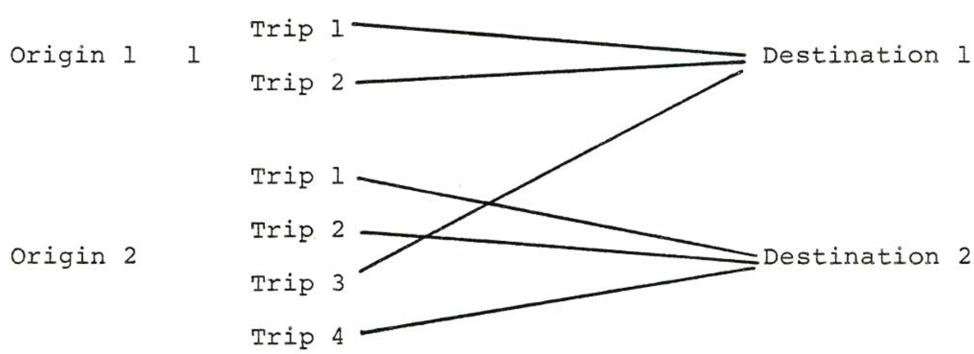
II

	1	2	
1	1	1	2
2	2	2	4
	3	3	

III

	1	2	
1	2	0	2
2	1	3	4
	3	3	

The important point is that each of these matrices can be generated in only a few ways. For example, matrix II can be generated by the following trip pattern



Clearly, the arrows may be arranged in other combinations, as long as three and only three arrows go to each destination. In general, the number of ways 1 trip can be selected from the two emerging from origin 1 is given by

$$2C_1 = \frac{2!}{1!1!} = 2$$

and the number of ways 2 trips can be selected from the 4 emerging from origin 2 is given by

$$4C_2 = \frac{4!}{2!2!} = 6.$$

Thus the number of ways matrix II can be determined is given by

$$2C_1 \cdot 4C_2 = 2 \cdot 6 = 12.$$

In the same way matrix I is generated by a total of $[2C_0 \cdot 4C_1] = 4$ ways and matrix III is generated by a total of $[2C_2 \cdot 4C_3] = 4$ ways. On the basis of this, one would select the trip table represented by matrix II as the most likely. In general, for the 2×2 matrix, the formula given by $[2C_a \cdot 4C_b]$ will give the total number of ways the flows can be reconciled with the table.

	1	2	
1	a		2
2		b	4
	3	3	

Further development is possible. Analysts typically have additional information on trip behaviour. For example, it may be that at this intersection left hand turns are less preferred. Thus, travellers coming from 0_1 prefer D_1 to D_2 by a ratio of 2:1 while travellers coming from 0_2 prefer D_2 to D_1 by 3:1. Now consider matrix 2

	1	2	
1	1	1	2
2	2	2	4
	3	3	

For a trip coming from 1 the chances of it going to D_1 and D_2 are 2/3 and 1/3 respectively. While for a trip coming from 0_2 the chances of it going to D_1 and D_2 are 1/4 and 3/4 respectively. Since there is 1 trip that goes to D_1 and D_2 and two trips from 0_2 going to D_1 and D_2 (total of 4 trips) each of the twelve ways of producing matrix II has a chance of

$$(\tfrac{2}{3})^1 (\tfrac{1}{3})^1 (\tfrac{1}{4})^2 (\tfrac{3}{4})^2 = 18/2304.$$

Similarly for matrix I the chance of each of the 4 ways occurring is

$$(\tfrac{2}{3})^0 (\tfrac{1}{3})^2 (\tfrac{1}{4})^2 (\tfrac{3}{4})^1 = 3/2304$$

and for matrix III the probability of each of the 4 ways is

$$(\tfrac{2}{3})^2 (\tfrac{1}{3})^0 (\tfrac{1}{4})^1 (\tfrac{3}{4})^3 = 108/2304$$

Now the relative frequency of each of the matrices occurring is

$4 \times 3 = 12$ (Matrix I)

$12 \times 18 = 216$ (Matrix II)

$4 \times 108 = 432$ (Matrix III)

On the basis of this information, Matrix III is now the most likely. The turning movement probabilities alter the most likely interchange matrix.

The extension of this procedure to more cells is straightforward. Formally, the problem appears as follows.

For any row i, the number of elementary links is given by

$$\binom{O_i}{T_{i1}}\binom{O_i - T_{i1}}{T_{i2}} \cdots\cdots \binom{O_i - T_{i1} - T_{i2} - \cdots - T_{i,n-1}}{T_{in}}$$

$$= \frac{O_i!}{\prod\limits_{i}^{n} T_{ij}!} . \qquad \ldots 4.3.1$$

For the entire matrix

$$w(T) = \frac{\prod\limits_{m} O_i!}{\prod\limits_{m}\prod\limits_{n} T_{ij}!} . \qquad \ldots 4.3.2$$

The point is to discover that matrix T^*_{ij} which results from the greatest number of equally probable link flows. To achieve this, the minimum of $\prod\limits_{m}\prod\limits_{n} T_{ij}!$ is needed, since $\prod O_i!$ is a constant for any matrix. To minimize the numerator, Williamsen has shown that this is closely approximated by minimizing $\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n} T_{ij}!$ Now the usual constraints are placed into the mathematical program, namely, $\sum\limits_{j} T_{ij} = O_i$, and $\sum\limits_{i} T_{ij} = D_j$ and $\sum\sum T_{ij} = T$.

For the problem where there is equal probability of elementary events (i.e., where, in our simple example, travellers are indifferent between left and right turns) the optimal value T^*_{ij} turns out to be

$$T^*_{ij} = \frac{O_i D_j}{T}.$$

At this time, such a model has been applied to situations where the elementary events (pairs of O-D nodes) are equally likely, and some modest success has been achieved with intercity airline 0-D matrices, and urban freeway flows. As yet no program has been made in the calculation of matrices when there is preference for one destination over others (i.e., turning movement preferences), a situation that clearly characterizes regional networks.

Note that this method falls generally into the camp of entropy models, by virtue of the minimization of the $\sum_i \sum_j \ln T_{ij}$ (which is a close approximation of maximizing entropy as conventionally defined in transportation).

Other methods closely related to this procedure start with an initial value of T_{ij}, say the results of the last survey, and use this year's traffic flow to amend this 0-D matrix. These procedures can be very useful once a benchmark T_{ij} has been established using conventional procedures.

4.3.3 Summary on Entropy Measures

Recently, these procedures have received the most academic attention. The integration of destination preference probabilities into the overall model framework remains unresearched. Also the viability of the simple procedures on more complex networks has not been undertaken. But, in general, this methodology has not been field tested.

4.4 Computer Algorithms for Synthesizing 0-D Matrices

A number of algorithms (computer programs) have been developed for synthesizing 0-D matrices. For the most part these programs are quite operational and "tuned" to a particular problem.

1. Casey-Hendrickson-Siddarthan Model. This program assumes:

a. the aggregate volume of traffic between two points is a function of socio-economic characteristics at the two nodes.

b. all intermodal volume estimates are consistent with the totals.

c. least squares is the appropriate objective function.

In essence this is a combined distribution assignment model based upon the work of Nguyen _et_ _al_. The quadratic programming procedure is used and is similar to the problem outlined in equations 3.2.12a to 3.2.12d above.

It is computationally complex and permits the user to specify travel impedances exogenously. It also requires detailed socio-economic data at each of the nodes and is primarily designed for use in large metropolitan areas where the links are high volume freeways and the nodes consist of employment, residential and shopping centres.

2. Frittabe

The usual gravity model is specified and the technique outlined for regression based procedure is employed. An initial trip interchange table is used for comparison in the calibration procedures. No documentation is presently available on this program.

3. Federal Highway Administration

The most comprehensive program has been developed for the U.S. government. It is designed to simulate an O-D matrix based on observed link flows and turning movements. Since it is designed for relatively small, but dense, urban streets networks, its relevance here is minimal.

4.6 Summary

The available computer programs are not relevant to Manitoba. They would require a substantial reprogramming which could only be accomplished in the light of reliable interchange information. For any transportation planning exercise in Manitoba, a computerized model should only be contemplated after the collection of data for and the computation of an O-D matrix has been thoroughly tested.

5.0 Application of 0-D Synthesis to Manitoba

Of the procedures that have been reviewed, the regression based techniques and the entropy model of Williamsen offer the most promise. They are computationally straightforward, can use socio-economic data (for the gravity model) which is available from the community reports, and do allow an exploration of various assumptions involved in developing link flows. The regional transport system probably does not have significant congestion effects and, therefore, an all-or-nothing route assignment is feasible and certainly simplifies matters. In the next section, the data at hand are evaluated, a research program to derive these 0-D matrices is outlined, and a brief description of the programming consideration is presented. Finally a validating survey is sketched which should be developed to complement or even replace the synthesizing exercise.

5.1 The Data

The regression based algorithms require socio-economic data such as population, employment, recreational indices and service characteristics at various nodes. Although probably not of the highest quality the information in the Manitoba Community Reports can be used. Where better information is obtainable, this basic data file could be augmented. As a start, a simple linear regression based model merely requires population and employment information and the community reports will suffice for this purpose.

In the entropy model, if the trips are uninfluenced by socio-economic factors, these data are not required. This is clearly a very narrow model and as such is unlikely to produce reliable trip interchange estimates. It is possible to use some type of probability model, based on socio-economic information to adjust the preferred destinations for travellers from a given origin. As mentioned in the previous section, this line of research has not been extensively reported in the literature reviewed, although it seems certain that it is being investigated. To develop such an exercie ~~is~~ experimental, although it could produce quite reasonable results.

More important is the quality of the link flow data available in the Traffic Map Statistics. These data can be divided into two groups. Approximately 60 points have information collected on a 24 hours basis from which an average annual daily traffic count can be developed. These data correspond to the provincial highway system. A small subset of these points have two way information which permits an estimate of traffice emerging from each origin i, and entering each destination j. In some cases, complete two-way information is available. For example, at Brandon North/South and East/West counts are available at the east, north and western boundaries. In more cases there is only a simple count for both flows of traffic at the boundaries. Presumably, one would average these flows for the estimate of traffic entering and exiting the point in question.

Most of the centres listed in the Community Reports have at least sufficient information to develop estimates on the entering and exiting flows. An important problem is the estimation of internodal flows V_a since these are the basis for validating the gravity model. Clearly, along any given link there is going to be leakage. This is especially so if the analysis is confined to the provincial highway system. At this point, the incorporation of the provincial road system appears too difficult, given the uncertainty as to the validity of these models. If it turns out that an acceptable 0-D matrix can be synthesized for the provincial highway system, then the incorporation of the road system can be undertaken with greater assurance.

An illustration of simplifying assumptions which will be needed is the area between Carman, Morden and Winkler. The information available is shown below in Figure 5.1. Notice that the information on 0_i and D_j (the entering and emerging traffic at Carman) must be obtained from counts which aggregate two directions of traffic, and in the case of north/south traffic between Elm Creek and Carman, do not provide accurate data for the northern perimeter of Carman.

Clearly some simplifying assumptions are needed such as how to derive and allocate flows, and imputing a flow to a boundary point on the basis of "downstream" counts. Also, leakage along a link can produce distortion. Although

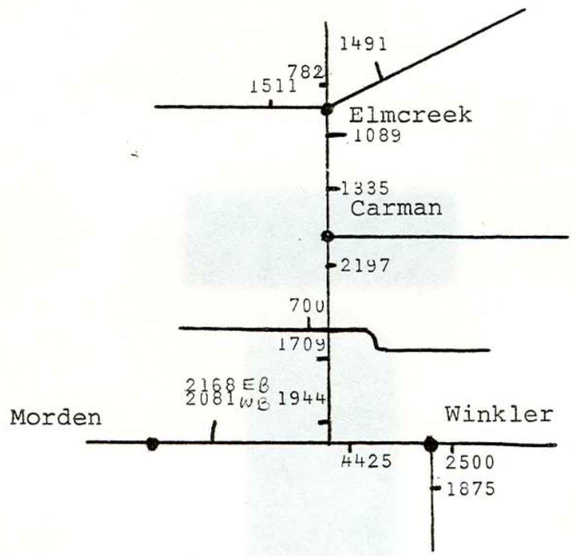


Figure 5.1

these assumptions need not be especially important individually, collectively they can produce a significant distortion in the estimated interchange matrix. Most of the validation done on existing regression type models has employed a very high quality link flow information on small networks.

An important feature of the road system of Manitoba is that the flat featureless plain has led to a relatively dense network. Compared to British Columbia, Manitoba has a large number of interchange nodes on the provincial highway system, with no significant population or employment mass. These dummy nodes can be included in the trip interchange matrix, or they can be submerged. The last approach is perhaps the best point at which to begin. Nonetheless, this could make the simple all-or-nothing route assignment problematic.

5.2. Validation

A final problem confronting both the regression and entropy approach is validation. The usual procedure is to derive a trip interchange matrix T^*_{ij} that, coupled with a route assignment assumption, produces a close match with observed link flows. Only limited matching of the computed T^*_{ij} with the actual T_{ij} has been reported in the literature and this has been with small well-defined matrices, most of which were synthetic. This is potentially a serious problem since there is no assurance that a distribution/assignment procedure which happens to replicate observed flows corresponds to a true interchange matrix. Given the existence of several dummy nodes, the need to disaggregate link flow data to produce originating and destinating flows (O_i, D_j) and the experimentation with different formulation of the gravity assumptions to produce interchange matrices for car and truck travel, there is a significant probability that different T^*_{ij} could produce quite similar link flows. Because of this, any O-D synthesis which cannot be verified against a benchmark interchange matrix must be treated with circumspection. This is not to say that the exercise is pointless, but to emphasize that demand forecasting on the basis of unverified O-D matrices is hazardous. There is much to learn about these procedures, and current research is concentrating upon testing these procedures against O-D matrices produced by conventional survey techniques.

Summary and Recommendation

The procedures which are currently under research and which promise the most likely candidates for application to Manitoba are the regression based models founded upon the gravity formulation coupled with a route assignment model, and entropy based models which calculate the most likely (entropy maximum) trip interchange.

The former have the advantage of proven success (although this is limited), while the latter represent state of the art procedures but remain quite experimental.

Computationally both procedures can be made as simple as desired, but they also have possibilities for development.

There are no "off-the-shelf" computer routines which can be adapted to the Manitoba road network at less cost than creating a model from scratch. Furthermore, given the nature and specificity of assumptions needed to devleop estimates of the 0_i and D_j there seems little chance of creating an algorithm to read highway statistics directly from tape or card without some considerable experimentation with and sensitivity analysis of these link flow aggregation assumptions.

The following steps constitute a research program which forms the basis for an 0-D synthesizing model.

(1) The most recent Manitoba Community reports should be consulted and the population, recreational facilities, health facilities, retail levels and governmental services recorded for the following urban areas.

Altona*
Arborg
Beausejour*
Benito
Birtle

Boissevain*
Brandon*
Carberry
Carman*
Crystal City

Dauphin*
Deloraine
Elkhorn
Emerson
Erickson

Ethelbert
Flin Flon*
Gilbert Plains
Gimli*
Gladstone

Glenboro
Grand Rapids
Grandview
Gretna
Hamiota

Haitney
Killarney
Lac du Bonnet*
MacGregor
Manitou

McCreary
Melita

Minnedosa
Morden*

Morris*
Neepawa
Niverville
Notre Dame de Lourdes
Pilot Mound

Pinewa
Pine Falls/Powerview
Plum Coulee
Portage la Prairie*
Rivers

Riverton*
Roblin
Rossburn
Russell
Ste. Anne

St. Claude
St. Pierre
Ste. Rose du Lac
Selkirk*
Shoal Lake

Somerset
Souris*
Steinbach
Swan River
Teulon

The Pas*
Thompson*

Virden*
Wawanesa

Winkler*
Winnipeg Beach
Winnipegosis
Winnipeg*

The socio-economic data set will record data in terms of numbers of people, numbers of jobs, etc., using these as gross indications of levels of economic activities and service levels. Subsequent testing will experiment with other ways of developing measures of soci-economic mass.

As a first step only those centres designated with a * will be included, and accordingly a simplified network will be constituted.

(2) For the most part, shortest path algorithms will produce acceptable route flows. A potential problem could arise if rural travellers turn out to use the provincial road system as well as the provincial highway systems. If there is extensive use of the road system, then link flows estimated by the model will seriously overestimate link flows actually observed on the highway system. The result is that synthesis will proceed resulting in an erroneous estimate of T_{ij}. This question can only be discovered experimentally, and most likely only with corroborating evidence from a travel survey.

(3) Initially the linear regression formulation should be pursued. Also, given the modal split provided by the highway statistics it should be possible to synthesize T_{ij} for cars and trucks.

It is also important to experiment with non-linear formulations. Without a benchmark 0-D matrix, validation can proceed only using link flows. Selection of a model is based upon the ability to replicate observed link flows.

Important aspects of this experiment will be

a. sensitivity analysis for different link flow aggregation procedures

b. the form of the impedance function

c. the specification of socio-economic data to be used in developing trip behaviour.

An important decision is whether a validating travel survey should be undertaken. Without this, there is no assurance that the synthesized 0-D matrix is accurate. It is strongly recommended that the synthesizing experiment not proceed unless there is a prior travel survey and unless the cost of 0-D synthesis is significantly lower than the cost of the survey.

Tentative Budget for 0-D Synthesis

This budget is designed to give a rough idea of how much a simple, linear regression O-D synthesizing experiment will cost.

1. Research Assistant: (to collect, compile and code socio-economic data and link flow data)
   ($10,000/annum @ 4 months = $3300)
2. Pprogrammer: (to set up alternative model for testing)
   ($10,000/annum @ 6 months = $5000)
3. Computer Time: (10 hours @ $300/hour: $3000)
4. Report Preparation: $1000

$12,300

(Note this budget does not include a stipend for the principal investigator.)

The approximate time to complete such an exercise would be about 6 months. The results would be:

a. A variety of estimated T^*_{ij} matrices based on various formulations of the gravity model.

b. A sensitivity analysis to demonstrate the effect of link flow disaggregation assumptions.

c. A validation analysis to determine how well each model replicated observed link flows.

5.3 A Survey Design for 0-D Matrices

Given the uncertainty associated with the 0-D synthesizing methods presently available, and the tentative results that would be available from the experiment outlined above, the development of a trip interchange matrix from survey procedures should be considered. Without this benchmark data it is difficult to construct a synthesizing program with any assurance.

A questionnaire, that was mailed out with the annual Autopac renewal notices or with driver re-registration could be constructed to produce a high quality trip interchange matrix. A random selection of about 2000 or even 1000 vehicle drivers could produce results with a relatively low standard error. Sampling on a probability proportional to size or a randomized cluster sample could also reduce the sample size yet also preserve the inferential basis of the sample.

The problem of no response can be handled in two ways.

a. Sample a large group (say 20,000) and hope for a 30% response rate. Randomly sample the nonresponses by telephone (about 200) and test whether the nonresponders are different from the responders. If not, then proceed with the anlysis on the sample at hand; if so, the main sample can be adjusted using the results from the telephone survey.

b. Provide an incentive to respond in the form of waiving the driver registration fee, or creating a lottery of about $1000 (spread over several prizes) to induce response.

The cost of the second alternative ranges between $10,000 (for a registration fee rebate (500 responses @ $20 per response) to $1000 for the lottery. Coding and analysis of the questionnaire would cost roughly about $5000. The second alternative seems the most cost effective and could produce a response rate as high as 75-80%.

Clearly, since a benchmark 0-D matrix is needed to verify the 0-D synthesizing model, and further, several benchmarks over a period of years are needed to establish the predictive powers of any models, it is useful to consider the role that surveys can play in developing a transportation planning system. After these synthetic matrices and the models used to develop them have been well established, greater reliance can be placed upon them to replace conventional trip interchange surveys. Until such time, however, it is risky to proceed with 0-D synthesis without having a reliable 0-D matrix to evaluate the model. The question in the final analysis is whether the 0-D synthesizing models as applied to Manitoba will be able to replace a conventional 0-D survey. At this time, given no such benchmark matrix exists, it would not be beneficial to commission experiments in 0-D synthesis without also undertaking a modest 0-D survey. Even so, if there was a divergence between the synthesized matrix and the survey based matrix, the latter would probably be most accurate. It is very difficult to forecast how well even uncomplicated regression models could synthesize 0-D information, but it is probably best to err on the side of pessimism. As a result, given the probable cost of synthesizing an 0-D matrix and the requirement that a survey based 0-D matrix is needed for verification, it might be wisest to thoroughly explore the conventional methods first before undertaking an experimental venture in 0-D synthesis.

Literature Consulted

1. Chan, V.K., R.G. Dowling, A.E. Willis, "Deriving Origin-Destination Information from Routinely Collected Traffic Counts" Institute for Transportation Studies, University of California, Berkeley, WP-80-1,1980

2. Dial, R.B., "A Probabilistic Multipath Traffic Assignment Model Which Obviates the Need for Path Enumeration" Transportation Research, vol. 5, 1971

3. Erlander, S., "Accessibility, Entropy and the Distribution and Assignment of Traffic" Transportation Research, vol. 11, 1977

4. Erlander, S., Nguyen, S. and N.F. Stewart, "On the Calibration of the Combined Trip Distribution-Assignment Model" Transportation Research, vol. 13B, 1977

5. Foster, R.E., J.D. Bensen, V. Stover, Consequences of Small Sample O-D Data Collection in the Transportation Planning Process, Federal Highway Administration, U.S. Government, Report #FJWA-RD-78-56, 1978

6. Hamburg, John and Associates, "Estimation of an Origin-Destination Trip Table Based on Observed Link Volumes and Turning Movements," vols. 1-3, Federal Highway Administration, U.S. Government, Report #FHWA RO-80/035

7. Hoberg, P., "Estimation of Parameters in Models for Traffic Prediction - A Nonlinear Regression Approach" Transportation Research, vol. 10, 1976

8. Lamarre, L., and M. Gaudry, "Estimating Origin-Destination Matrices from Traffic Counts: A Simple Linear Inter-City Model for Quebec" Centre de Recherche, University de Montreal, #105, 1978

9. Low, D., "A New Approach to Transportation Systems Modelling" Traffic Quarterly, vol 26.,1972

10. MacArthur, B., "Use of Origin-Destination Assignment Models in Vermont" Department of Civil Engineering, University of New Brunswick, 1973

11. Nguyen, S., "Estimating an O-D Matrix from Network Data: A Network Equilibrium Approach" Centre de Recherche, Universite de Montreal, Feb. 1977

12. Nguyen, S., "On the Estimation of an O-D Matrix by the Equilibrium Method Using Pseudo Delay Functions" Centre de Recherche, Universite de Montreal, 1978

13. Province of Manitoba, "1979 Traffic Statistics" Department of Highways and Transportation

14. Robillard, P., "Estimating the O-D Matrix from Observed Link Volumes" Transportation Research, vol. 9, 1975

15. Van Zuylen, H.J., "The Information Minimizing Method: Validity and Applicability to Transportation Planning" in New Developments in Modelling Travel Demands, Jensen, H. et al., Saxon House, 1978

16. Willumsen, L.G., "Estimation of an O-D Matrix from Traffic Counts - A Review" Institute of Transport Studies, University of Leeds, 1978

17. Wilson, A.G., "The Use of Entropy in Maximizing Models in the Theory of Trip Distribution, Modal Split and Route Split" Journal of Transport Economics and Policy, Jan. 1978