Transportation **16**: 3–26, 1989.

Transport demand model estimation from traffic counts

O.Z. TAMIN[1] & L.G. WILLUMSEN[2,*]

[1]*Department of Civil Engineering, Institute of Technology Bandung, Bandung, Indonesia;* [2]*Transport Studies Group, University College London, and Steer Davies & Gleave Ltd., London, UK (*requests for offprints)*

Key words: demand modelling, traffic counts, model estimation, gravity model

Abstract. Many problems in transport planning and management tasks require an origin-destination (O-D) matrix to represent the travel pattern. However, O-D matrices obtained through a large scale survey such as home or roadside interviews, tend to be costly, labour intensive and time disruptive to trip makers. Therefore, the use of low cost and easily available data is particularly attractive.

The need of low-cost methods to estimate current and future O-D matrices is even more valuable in developing countries because of the rapid changes in population, economic activity and land use. Models of transport demand have been used for many years to synthesize O-D matrices in study areas. A typical example of this is the gravity model; its functional form, plus the appropriate values for the parameters involved, is employed to produce acceptable matrices representing trip making behaviour for many trip purposes and time periods.

The work reported in this paper has combined the advantages of acceptable travel demand models with the low cost and availability of traffic counts. Three types of demand models have been used: gravity (GR), opportunity (OP) and gravity-opportunity (GO) models. Three estimation methods have been developed to calibrate these models from traffic counts, namely: non-linear-least-squares (NLLS), weighted-non-linear-least-squares (WNLLS) and maximum-likelihood (ML).

The 1978 Ripon (urban vehicle movement) survey was used to test these methods. They were found to perform satisfactorily since each calibrated model reproduced the observed O-D matrix fairly closely. The tests were carried out using two assignment techniques, all-or-nothing and the stochastic method due to Burrell, in determining the routes taken through the network.

1. Introduction

Travel has become an integral part of our daily (urban) life. This activity generates its good share of problems to any community, including traffic congestion, delay, air pollution and visual intrusion. Any measure to alleviate these problems, presupposes an understanding of the underlying travel pattern. The concept of an "O-D matrix" has been adopted by transport planners to represent the most important features of this travel pattern. When an O-D matrix is assigned onto the network, a flow pattern will be produced. From examining this, one can identify the problems and some kind of solution may be devised. Therefore, an O-D matrix plays a very important role in various transport studies.

Unfortunately, "conventional methods" for estimating O-D matrices rely much on extensive surveys which make them fairly expensive in terms of time, manpower and also disruption to trip makers. All of these have led researchers to investigate alternative, less expensive methods for estimating O-D matrices. A popular option in this respect is the use of traffic counts because of their availability, low-cost and non-disruptive nature. Traffic counts are routinely collected by many authorities due to their multiple uses in traffic and transport planning tasks. Therefore, there is always a basic set of traffic counts available in almost any city. Secondly, traffic counts are inexpensive in terms of time, manpower, organizational and management requirements. Furthermore, they can be collected without disrupting travellers. Finally, the automatic collection of traffic counts is well advanced and good accuracies can now be expected from such methods. This makes them particularly useful in studying hourly, daily and seasonal variations in demand together with their corresponding O-D matrices.

2. Previous research

One can interpret traffic counts as resulting from a combination of two elements: an O-D matrix and the route choice pattern selected by drivers to travel on the network. These two elements may be linearly related to traffic counts, see eq. (4.1) below, but under normal circumstances there will never be enough traffic counts to identify a single O-D matrix as the only possible source of the observed flows. Traffic counts alone are not enough to estimate O-D matrices, something else is needed.

The methodology adopted to provide the additional information required to estimate a unique O-D matrix may be used to classify research work in this area into two main groups: unstructured and structured methods. Under the first group we include all methods attempting to estimate an O-D matrix without imposing any external structure or model. Researchers working with this approach employ general principles like entropy maximization (or information minimization) to estimate a matrix with the minimum of assumptions. Workers in the second group prefer to benefit from the wealth of experience in the calibration and use transport demand models and therefore impose such structure to their matrices. They seek to calibrate these models from low cost data like traffic counts instead of using more expensive surveys. For an early review of these approaches see Willumsen (1978, 1981).

Unstructured methods

In this case the additional information required to estimate a matrix is provided by a general principle reflecting the probability of observing a particular matrix given our current state of knowledge about the conditions that this matrix should meet. This principle often use the concept of "information" or "entropy" for defining a most likely trip pattern and the knowledge about the properties of the matrix must meet is embodied in the form of restrictions to an optimization programme. Pioneers of this approach were Van Zuylen and Wilumsen (1980). A number of related models have been developed to improve estimation techniques allowing for errors in the traffic counts or introducing a variation in the estimation method.

Transport demand model estimation

This approach assumes that the trip making behaviour in the study area is well represented by a certain general model form, e.g. a gravity model. As usual, the link flows are expressed as a function of the O-D matrix but in this case the matrix is a function of a model form and relevant parameters, for example the exponent beta in a gravity model. The parameters of the postulated model are then estimated, so that the errors between the estimated and observed link flows are minimized.

Estimation methods of this kind have been proposed by Low (1972) with a simple gravity model and linear-least-squares and by Robillard (1975) and Högberg (1976) using non-linear-least-squares estimation techniques. Other relevant methods have been proposed by Holm et al. (1976) and Tamin (1985). There has been, however, little validation of these techniques and there is scope for exploring more advanced and flexible model forms.

3. Definitions

For the rest of the paper we define the following terms:

$[T_{id}]$ = the *observed* O-D matrix from origin i to destination d
$[Z_{ij}]$ = the *ordered* O-D matrix from origin i to the j^{th} destination in ascending order of distance away from i
O_i^k = the total trips of each trip purpose k generated by origin i
D_d^k = the total trips of each trip purpose k attracted by destination d
A_i^k, B_d^k = the balancing factors for each trip purpose k for origin i and destination d
C_{id} = the trip cost of travelling from origin i to destination d
α_k, β_k = the unknown estimated parameters to be calibrated
V_l^+, V_l = the estimated and observed link flows, respectively
p_{id}^l = the trip assignment proportion for trips from origin i to destination d which use link l

L = the total number of links observed
K = the total number of trip purposes or commodity types
N = the total number of origins or destinations
f^k_{id}, F^k_{ij} = the observed and ordered proportionality factors relating origin i and destination j, respectively for each trip purpose k
δ^i_{jd}, δ^{i-1}_{jd} = the transformations relating $[T_{id}]$ and $[Z^{ij}]$
U^k_{ip} = the opportunity function relating i and the p^{th} destination away from i
ε, μ = transformation parameters

We use the notational convention that Σ_k means the summation begins at $k = 1$ and continues over the entire range of the subscript.

4. Transport demand model estimation from traffic counts

The basic principles

Consider a study area which is divided into N zones, each of which is represented by a centroid. All of these zones are inter-connected by a road network which consists of series of links and nodes. Furthermore, the O-D matrix for this study area consists of N cells. $[N - N]$ cells if intra-zonal trips can be disregarded. The most important stage for the estimation of a transport demand model from traffic counts is to identify the paths followed by the trips from each origin to each destination. The variable p^l_{id} is used to define the proportion of trips from zone i to zone d travelling through link l. Thus, the flow on each link is a result of:

- trips from zone i to zone d $(= T_{id})$, and
- the proportion of trips from zone i to zone d whose trips use link l which is defined by $p^l_{id}(0 \le p^l_{id} \le 1)$.

The flow (V_l) in a particular link l is the sum of the contributions of all trips between zones that use that link. Mathematically, it can be expressed as follows:

$$V_l = \sum_i \sum_d T_{id} \cdot p^l_{id} \tag{4.1}$$

The variable p^l_{id} can be estimated using various trip assignment techniques ranging from a simple all-or-nothing to more advanced equilibrium methods. Assuming a good estimation of the p^l_{id} and given a set of observed traffic counts (V_l), then there will be N^2 unknown T_{id}'s to be estimated from a set of L simultaneous linear equations (4.1) where L is the total number of traffic counts.

In principle, N^2 independent and consistent traffic counts are required in order to determine uniquely the O-D matrix $[T_{id}]$, $(N^2 - N)$ if intra-zonal trips can be disregarded. In practice, the number of observed traffic counts is much less than the number of unknowns T_{id}'s. Therefore it is impossible to determine uniquely the solution. In general, it can be said that there will be more than one O-D matrix which will satisfy the traffic counts. One possible way to overcome this problem is to restrict the number of possible solutions by modelling the trip making behaviour in the study area. But first, we must discuss the alternative route choice or traffic assignment models which can be used to estimate the p^l_{id}.

Some trip assignment techniques

One of the main objectives of trip assignment methods is the identification of the routes taken through the network and from them the total number of trips using each particular link. Robillard's (1975) classification of traffic assignment techniques is relevant to this problem: "proportional" and "non-proportional" methods.

If the proportion of drivers from an O-D pair using a particular link is independent from flow levels the method is said to be "proportional". The most common example is "all-or-nothing" assignment. It is assumed that all drivers wish to minimize their perceived travel costs, all the drivers perceive these costs in the same way and that these costs do not depend on flow level; hence drivers from one zone to another will use the same minimum cost route. This assignment is considered to be unrealistic for many networks because it does not consider variations in the perception of costs nor the effects of congestion. However, it is the fastest and simplest assignment technique and is useful for sparse networks where there are few alternatives O-D paths. The value of p^l_{id} for this assignment is defined as follows:

$$p^l_{id} = \begin{cases} 1 & \text{if trips from origins } i \text{ to destinations } d \text{ use link } l \\ 0 & \text{otherwise or } i = d \end{cases}$$

Most "stochastic" assignment techniques, like Burrell's and Dial's, are also "proportional" methods. They have been developed to take into account variations in individual perceptions of cost. They often ignore congestion effects but they produce a more realistic spread of routes than all-or-nothing.

Burrell (1968) proposed the multi-flow model where average link costs are defined, coupled with a form for the distribution of individual perceived costs about the mean. A different set of "perceived link costs" can then be generated by taking random samples of these link cost distributions. Random numbers, for example based on a rectangular distribution, are used to repeatedly select costs for each link. The model then find and load the quickest routes minimizing the sum of their perceived travel costs. Therefore, one can build several sets of minimum perceived cost trees, one for each set of sampled perceived costs. The O-D matrix is then divided into M fractions and each fraction is assigned to a set of minimum cost routes. This is repeated M times.

Under congested conditions, the cost of travelling on a link depends on the flow on that link and a cost-flow relationship. Several techniques have been developed to explicitly consider these effects and they are usually called "capacity-restrained" methods. Some of them are: repeated all-or-nothing, iterative loading, incremental loading and equilibrium assignment.

Wardrop proposed in 1952 the basic idea behind rigorous equilibrium assignment techniques in the form of a "principle". This may be written as: "Under equilibrium conditions, traffic arranges itself in such a way that all routes actually used by trips between a given origin and destination have equal and minimum costs and all unused routes have greater (or equal) costs". In other words, under equilibrium conditions no driver can switch to another route and thereby lower his travel cost since all routes have either same cost as his own or greater.

The most appropriate assignment technique in each case would depend on the characteristics of the study area. The level of congestion, the availability of alternative routes and their corresponding costs and some idea of drivers variability in the perceptions of costs should help in choosing the best assignment technique to estimate the values of p^l_{id}. The matrix estimation problem is simpler when a "proportional" assignment model is considered to be sufficiently realistic. The use of "non-proportional" assignment methods requires an iterative estimation process where the assumed p^l_{id} are used to estimate a matrix which in turn is used to improve the estimation of the p^l_{id} values. This issue is discussed in greater detail in Willumsen (1984). In this paper we will concentrate on p^l_{id} based on proportional assignment methods.

Gravity model (GR) as a transport demand model

This model was originally developed by analogy with Newton's law of gravitational force F_{id} between two masses m_i and m_d separated by a distance d_{id}, see Wilson (1967):

$$F_{id} = \mu \cdot m_i \cdot m_d/(d_{id})^2, \quad \mu \text{ is a constant} \tag{4.2}$$

The analogous transport gravity model is then,

$$T_{id} = k \cdot O_i \cdot D_d/(d_{id})^2, \quad k \text{ is a constant} \tag{4.3}$$

This model has some sensible properties. It proposes that the number of trips from zone i to d is directly proportional to each trips emanating from i (O_i) and those attracted to d (D_d) and inversely proportional to the square of the distance between them. But eq. (4.3) has at least one obvious deficiency. If a particular O_i and a particular D_d are each doubled, then the number of trips between these zones would quadruple accordingly, when it would be expected that they would only double. Therefore, it is often advisable to constrain the model so that:

$$\sum_d T_{id} = O_i \quad \text{and} \quad \sum_i T_{id} = D_d \tag{4.4}$$

These constraints can be satisfied if a set of variables A_i and B_d, associated with production and attraction totals respectively, are introduced. They are sometimes called "balancing factors" and must be calculated as part of the estimation process to ensure the constraints (4.4) are met. Furthermore, there are no good reasons to use just distance and a power of 2 in the denominator. Therefore, a more general deterrence function and generalized costs are often used instead of distance squared in the model. Assuming that there are K trip purposes travelling between zones, then the modified gravity model can be expressed as:

$$T_{id} = \sum_k [O_i^k \cdot D_d^k \cdot A_i^k \cdot B_d^k \cdot f_{id}^k] \tag{4.5}$$

where:

– A_i^k and B_d^k = the balancing factors expressed as:

$$- \; A_i^k = \left[\sum_d (B_d^k \cdot D_d^k \cdot f_{id}^k)\right]^{-1} \tag{4.6}$$

$$- \; B_d^k = \left[\sum_i (A_i^k \cdot O_i^k \cdot f_{id}^k)\right]^{-1} \tag{4.7}$$

– f_{id}^k = the deterrence function, often $\exp(-\beta_k \cdot C_{id})$

Eq. (4.5) expresses the doubly-constrained gravity model (DCGR).

Gravity-opportunity model (GO) as a transport demand model

A different framework to modelling trip making behaviour is incorporated in the opportunity model. Consider first a single origin and its destinations, embedded in a two dimensional surface allowing the full effects of contiguity to operate between all destinations; the destinations are then ordered by distance (or generalized cost) from the origin. Select one destination and induce change in its attributes, for example an increase in the shopping area there. Such a change is likely to lead to a different, and lesser, effect on trips from the origin to those destinations nearer to it, compared with those which are more distant.

The magnitude of this difference is a consequence of the intervening-opportunity effect of that destination, and it is an intuitively clear phenomenon that the gravity model fails to recognize. Whereas the gravity model is deficient in intervening-opportunity effects, the opportunity model constructed is equally deficient in omitting the trip impedance. It seems logical that an ideal model should contain both these distinct effects.

Wills (1986) developed a flexible gravity-opportunity (GO) model for trip distribution in which standard forms of the gravity and opportunity model are obtained as special cases. The choice between gravity or opportunity approach is decided empirically and statistically by restrictions on parameters which control the global functional form of the trip distribution mechanism.

$$\text{Given that } \delta^i_{jd} = \begin{cases} 1 & \text{if destination } d \text{ is the } j^{\text{th}} \text{ position in ascending} \\ & \text{order of distance away from } i \\ 0 & \text{otherwise} \end{cases}$$

then the ordered O-D matrix can be obtained by the following transformation:

$$Z_{ij} = \sum_d [\delta^i_{jd} \cdot T_{id}] \tag{4.8}$$

While the ordering transformation δ^i_{jd} produces an ordered O-D matrix, its inverse δ^{i-1}_{jd} allows the observed O-D matrix to be recovered by:

$$T_{id} = \sum_j [\delta^{i-1}_{jd} \cdot Z_{ij}] \tag{4.9}$$

It should be noted that this part of transformations is applicable to any variable based on the O-D matrix, notably the cost matrix and the proportionality factor, in addition to the O-D matrix.

Transformations
In order to provide a monotonic scaling of variables in such a manner as to generate families of specific functional forms, the Box-Cox transformation is used. The theoretical basis, and motivation for, these transformations are extensively covered in Box-Cox (1964). The direct Box-Cox transformation of a variable y is defined as:

$$y^{(\varepsilon)} = \begin{cases} (y^{\varepsilon} - 1)/\varepsilon & \varepsilon \neq 0 \\ \ln y & \varepsilon = 0 \end{cases} \tag{4.10}$$

and the inverse Box-Cox transformation as:

$$y^{(1/\varepsilon)} = \begin{cases} (y\varepsilon + 1)^{1/\varepsilon} & \varepsilon \neq 0 \\ \exp y & \varepsilon = 0 \end{cases} \tag{4.11}$$

These transformations may be combined into a new function which we introduce as a convex combination in μ:

$$y^{(\varepsilon,\mu)} = \mu y^{(\varepsilon)} + (1 - \mu) y^{(1/\varepsilon)}, \quad 0 \leq \mu \leq 1 \tag{4.12}$$

The proposed model
Assuming that there are K trip purposes travelling between zones, the proposed model is:

$$T_{id} = \sum_k [O_i^k \cdot D_d^k \cdot A_i^k \cdot B_d^k \cdot f_{id}^k] \tag{4.13}$$

where:

- A_i^k, B_d^k defined as eqs. (4.6) – (4.7), respectively;

- $$f_{id}^k = \sum_j [\delta_{jd}^{i-1} \cdot F_{ij}^k] \tag{4.14}$$

The opportunity function (U_{ip}^k), see eq. (4.16), is incorporated into a general proportionality factor (F_{ij}^k) which is defined by the difference in functions of the cumulative opportunities from i and j^{th} destination away from i, and from i to the $(j-1)^{\text{th}}$ destination away from i, and can be defined as:

- $$F_{ij}^k = \left(\sum_p^j U_{ip}^k\right)^{(\varepsilon,\mu)} - \left(\sum_p^{j-1} U_{ip}^k\right)^{(\varepsilon,\mu)} \tag{4.15}$$

– $U_{ip}^k = \exp[(1-\varepsilon) \cdot \alpha_k \cdot \ln D_{pk}^i - \beta_k \cdot C_{ip}]$ (4.16)

– $D_{jk}^i = \sum_d [\delta_{jd}^i \cdot D_d^k]$ (4.17)

– the (ε, μ) transformtion is defined by eqs. (4.10)–(4.12).

The general proportionality factor is subjected to a convex combination of direct and inverse Box-Cox transformations. From this general form several special cases may be derived by setting ε and μ to a particular value, see Fig. 1. Three specific models are easily identified: the gravity (GR), the logarithmic-opportunity (LO) and the exponential-opportunity (EO).

Fundamental equation

By substituting eq. (4.5) to eq. (4.1), then "the fundamental equation" for the estimation of a transport demand model from traffic counts is:

$$V_l = \sum_k \sum_i \sum_d [O_i^k \cdot D_d^k \cdot A_i^k \cdot B_d^k \cdot f_{id}^k \cdot p_{id}^l] \tag{4.18}$$

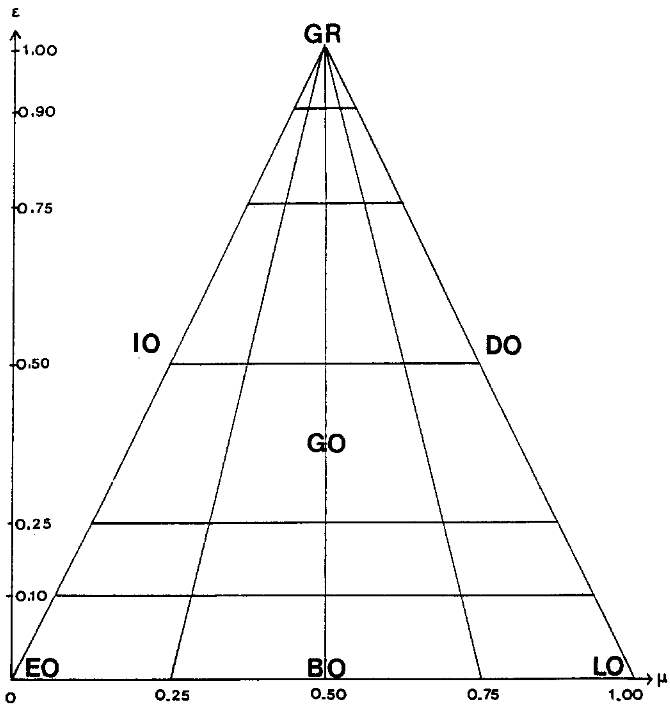

Fig. 1. Diagrammatic structure of the proportionality factor and its special cases (Source: Wills 1986).

It is important to note that V_l's are not available for each trip purpose but they are available for all trip purposes taken together. Assuming that we use the GO model, eq. (4.18) constitutes a system of L simultaneous equations with $2K$ unknown parameters for estimation. The problem now is how to estimate the unknown parameters so that the model reproduces the estimated link flows as close as possible to the observed ones.

5. Methods of estimation

Non-linear-least-squares estimation method (NLLS)

The main idea of this method is to estimate the unknown parameters which minimizes the sum of the squared differences between the estimated and observed link flows. This method gives more importance to the reduction of differences at high flows than at low flow levels. To compensate for this one could use a *weighted-non-linear-least-squares* (WNLLS) method, where weights (in our case equal to the reciprocal of the observed flows) are associated to each difference and the resulting sum of squares is minimized.

If we choose the gravity-opportunity (GO) model to represent the trip making behaviour the problem becomes:

$$\text{to minimize } S = \sum_l [(V_l^+ - V_l)/V_l^*] \tag{5.1}$$

where: $V_l^* = 1$ for NLLS or $V_l^* = V_l$ for WNLLS

Having substituted eq. (4.1) to eq. (5.1), the following two sets of equations are required in order to find a set of unknown parameters α_k and β_k which minimizes eq. (5.1), they are: for $k = 1, \ldots, K$.

$$\begin{aligned} \delta S/\delta\alpha_k &= \sum_l \left[2\left(\sum_i \sum_d T_{id} \cdot p_{id}^l - V_l\right)\left(\sum_i \sum_d \delta T_{id}/\delta\alpha_k \cdot p_{id}^l\right)/V_l^* \right] = 0 \\ \delta S/\delta\beta_k &= \sum_l \left[\left(\sum_i \sum_d T_{id} \cdot p_{id}^l - V_l\right)\left(\sum_i \sum_d \delta T_{id}/\delta\beta_k \cdot p_{id}^l\right)/V_l^* \right] = 0 \end{aligned} \tag{5.2}$$

Eq. (5.2) is a system of $2K$ simultaneous equations which has $2K$ unknown parameters α_k and β_k for estimation. It is possible to determine uniquely all the parameters provided that $L \geq 2K$ and preferably much greater. Newton's method could then be used to solve eq. (5.2).

Maximum-likelihood estimation method

A different approach to model estimation results from maximum likelihood techniques. Suppose we have a probability model, a set of statistical hypotheses and data, which form the foundation of statistical inference. Let $P(R/H)$ be the probability of obtaining results R given the hypothesis H, according to the probability model. This probability is defined for any member of the set of possible results given any one hypothesis. It may be regarded as a function of both R and H, but is usually used as a function of R alone, for some specific H. When considered as a function of R it defines a statistical distribution, either discrete or continuous. As such, if we sum or integrate over all possible results R we will obtain unity, by one of the axioms of probability.

The likelihood $L(H/R)$, of the hypothesis H given data R, and a specific model, is proportional to $P(R/H)$, the constant of proportionality being arbitrary. Whereas with the probability R is the variable and H is constant, with the likelihood H is the variable for constant R. This distinction is fundamental.

$$L(H/R) = cP(R/H) \tag{5.3}$$

The arbitrary constant of proportionality enables us to use the same definition of likelihood for discrete and continuous variables. The word support is suggested by Edwards (1972) to refer to natural logarithm of the likelihood, principally in order to change the multiplicative properties into additives ones. By taking logarithm, the arbitrary multiplicative constant becomes an arbitrary additive one. Support for one hypothesis against another ranges from zero to an infinitely large amount.

Denoting support by S, at the value $S = 2$ the likelihood in favour of the one hypothesis is about 7.4 times the likelihood in favour of the other, and at $S = 3$ the factor is about 20. It is obvious now that the greater the value of S, the greater will be the likelihood in favour of the one hypothesis than the other.

The framework of the maximum-likelihood estimation method

Suppose we have available a traffic count data survey involving a total number of V_T vehicles counted in L independent links. If V_l denotes the observed independent traffic count in each particular link l, then we must have:

$$\sum_l V_l = V_T \tag{5.4}$$

Now, let p_l be the probability of having an estimated independent link flow for each particular link l:

$$p_l = V_l^+ / V_T \tag{5.5}$$

Following the multinomial distribution, if there are "L" links and the probability of a trial falling in the "l^{th}" link is "p_l", then the probability of obtaining "V_l" out of "V_T" in the 1st link, "V_2" in the 2nd link, and in general, "V_l" in the l^{th} link is:

$$P = \frac{V_T!}{V_1! \cdot V_2! \cdot V_3! \cdots V_L!} \prod_l p_l^{V_l} \tag{5.6}$$

Then by substituting eq. (5.6) into eq. (5.3), the following equation, often referred to as the "Likelihood-Function", is built.

$$L = c \prod_l p_l^{V_l} \tag{5.7}$$

where the term of $V_T!/V_1! \cdot V_2! \cdot V_3! \cdots V_L!$ has been absorbed into the arbitrary constant c.

The idea of the maximum-likelihood method is that the choice of the hypothesis H maximizing eq. (5.7), will yield a distribution of V_l giving the best possible fit to the survey data (V_l). The objective function now is:

$$\text{to maximize } L = c \prod_l p_l^{V_l} \tag{5.8}$$

$$\text{subject to the total flow constraint: } \sum_l V_l^+ - V_T = 0 \tag{5.9}$$

Taking the log term of eq. (5.8) and using the following Lagrangian form, eq. (5.8) and eq. (5.9) can then be written into a single equation, that is:

$$\text{to maximize } L_1 = \sum_l V_l \ln p_l + \ln c - \theta\left(\sum_l V_l^+ - V_T\right) \tag{5.10}$$

By substituting eq. (4.1) to eq. (5.10), the objective function now is:

$$\begin{aligned} \text{to max. } L_1 = \sum_l \left[V_l \cdot \ln\left(\sum_i \sum_d T_{id} \cdot p_{id}^l\right) - \theta \sum_i \sum_d T_{id} \cdot p_{id}^l \right] \\ + \theta V_T - V_T \ln V_T + \ln c \end{aligned} \tag{5.11}$$

w.r.t. α_k, β_k and θ (for $k = 1, \ldots, K$)

In order to find a set of parameters which minimizes eq. (5.11), then the following three sets of equations are required (for $k = 1, \ldots, K$):

$$\delta L_1/\delta\alpha_k = \sum_l \left\{ V_l \frac{\sum_i \sum_d (\delta T_{id}/\delta\alpha_k \cdot p^l_{id})}{\sum_i \sum_d (T_{id} \cdot p^l_{id})} - \theta \sum_i \sum_d \delta T_{id}/\delta\alpha_k \cdot p^l_{id} \right\} = 0$$

$$\delta L_1/\delta\beta_k = \sum_l \left\{ V_l \frac{\sum_i \sum_d (\delta T_{id}/\delta\beta_k \cdot p^l_{id})}{\sum_i \sum_d (T_{id} \cdot p^l_{id})} - \theta \sum_i \sum_d \delta T_{id}/\delta\beta_k \cdot p^l_{id} \right\} = 0$$

$$\delta L_1/\delta_\theta = -\theta \left\{ \sum_i \sum_d T_{id} \cdot p^l_{id} - V_T \right\} = 0 \qquad (5.12)$$

Eq. (5.12) is a system of $2K + 1$ simultaneous equations which has $2K + 1$ unknown parameters α_k, β_k and θ need to be estimated. Newton's method can then be used to solve eq. (5.12). Hamerslag and Immers (1988) have used a ML estimator with equations similar to that of eq. (5.12) for the same problem. There are, however, little differences: the authors estimate α, β, θ and they use the ML estimators for discrete values of deterrence functions and groups of balancing factors.

6. Some basic statistical tests

The development of good model estimation techniques to ensure that the fitted parameters result in flows as close as possible to the observations is not enough to show the value of a new technique. We also need an indication of how accurate the resulting models are and this requires a comparison between estimated and an independently observed O-D matrix using appropriate statistical indicators of this fit. In this study, the following goodness-of-fit (GOF) statistic has been used:

– *Root mean square error RMSE*

$$\text{RMSE} = \left[\sum_i \sum_d \frac{(T^+_{id} - T_{id})^2}{N(N-1)} \right]^{0.5} \quad \text{for } i \neq d \qquad (6.1)$$

where: T^+_{id}, T_{id} = observed and estimated O-D matrix, respectively.

7. Tests with Ripon data (urban vehicle movement)

Ripon data set

In order to validate the model estimation technique a real data set of urban vehicle movements corresponding to Ripon was used. Ripon is a busy market town in North Yorkshire lying north of Harrogate on the A61. There are three main groups of movements using the roads in this town. First, there is the local traffic from the surrounding area as people come into Ripon to work or to make use of the various retail, educational and social facilities. Second, Ripon is a popular tourist site with the cathedral attracting many visitors each summer. Finally, some traffic passes through Ripon on its way to join the A1 further north. North Yorkshire County Council (NYCC) conducted roadside O-D surveys in 1978 and 1985 which, in conjunction with separate traffic counts collected by Steer, Davies and Gleave Ltd. (SDG), form the basis of this data set. The Ripon data set comprises the following:

- The 1985 O-D matrix, a 24 hour average annual daily traffic (AADT) matrix, compiled from roadside inteview data collected by NYCC in May 1985.
- The 1985 network description.
- 63 Traffic counts taken throughout Ripon by SDG in November 1985 and by NYCC at the interview sites in 1978 and 1985.

The study area was coded into 19 internal zones and 7 external zones and a network with 82 nodes and 188 one-way links connecting pairs of nodes and zone centroids. In addition to the network definition, we also have 63 observed traffic counts (V_l) and trip generation/attraction factors (O_i^k and D_d^k) for each zone. The units adopted in eq. (4.18) are given as follows:

V_l = link flow counted for each particular link l in vehicles/day.

O_i^k, D_d^k = trip generation and attraction factor, respectively in vehicles/day.

Computer programs

There are several computer packages for transport modelling which include the four conventional stages of trip generation, trip distribution, modal choice and trip assignment. The suite MOTORS, developed and marketed by Steer, Davies and Gleave Ltd. (1984) was used as a basis for our new model fitting software. Our computer programs have been written in such a way that they are fully integrated and interactive with MOTORS suite. The work was carried out using the Apricot-XEN micro-computer with 895K RAM, one 720 Kbyte 3.5′ floppy disk drive, 20M byte hard disk, and running under MS-DOS 3.1.

Results

In assessing the value of a new transport model one is of course interested in the accuracy of the estimated travel behaviour. In our case, our new approach relies on traffic counts as basic data input and offers a choice of demand models to represent trip making behaviour. This flexibility enhances the value of the model but it is also important to have some feeling on how the choice of model form may affect the accuracy of the resulting O-D matrix in at least one real case. We can use the Ripon O-D matrix to obtain an indication of the most suitable model in this case. We used the observed O-D matrix to estimate demand models directly using the NLLS estimation method. Three models were fitted: gravity (GR), opportunity (OP) and gravity-opportunity (GO) model.

In the GO model, the choice between gravity or opportunity is decided empirically and statistically by restriction on parameters, ε and μ, which control the global functional form of the trip distribution mechanism. By setting $\varepsilon = 1$, the GO model will behave as the GR model since the opportunity part is omitted from the opportunity function (4.16) and similarly, by omitting the cost part of eq. (4.16), the OP model is then created.

By using various values of ε and μ, we can then plot the contour of S. In order to show the contours more clearly, the triangular shape is then replaced by the standard rectangular one in Fig. 2a. It can be seen from Fig. 2b that the minimum value of S is obtained at points $\varepsilon = 0.5$ and $\mu = 1$. This means that for the Ripon case it is the GO model the one that gets closer to reproduce the observed O-D matrix thus offering some improvement over the gravity model.

The GOF statistic to compare the estimated O-D matrix with the observed one, using NLLS method and the real trip matrix information is given in Fig. 3. It can be seen that the GO model performs better than the GR and OP models since the value of RMSE of GO is less than those for GR and for OP. The GOF statistic of the estimated O-D matrix produced by the conventional Furness trip distribution is also given in Fig. 3. It can be seen that GR and GO performs better than Furness since the values of RMSE for GO and GR are less than for Furness.

It was not unexpected that the performance of GO was better than that

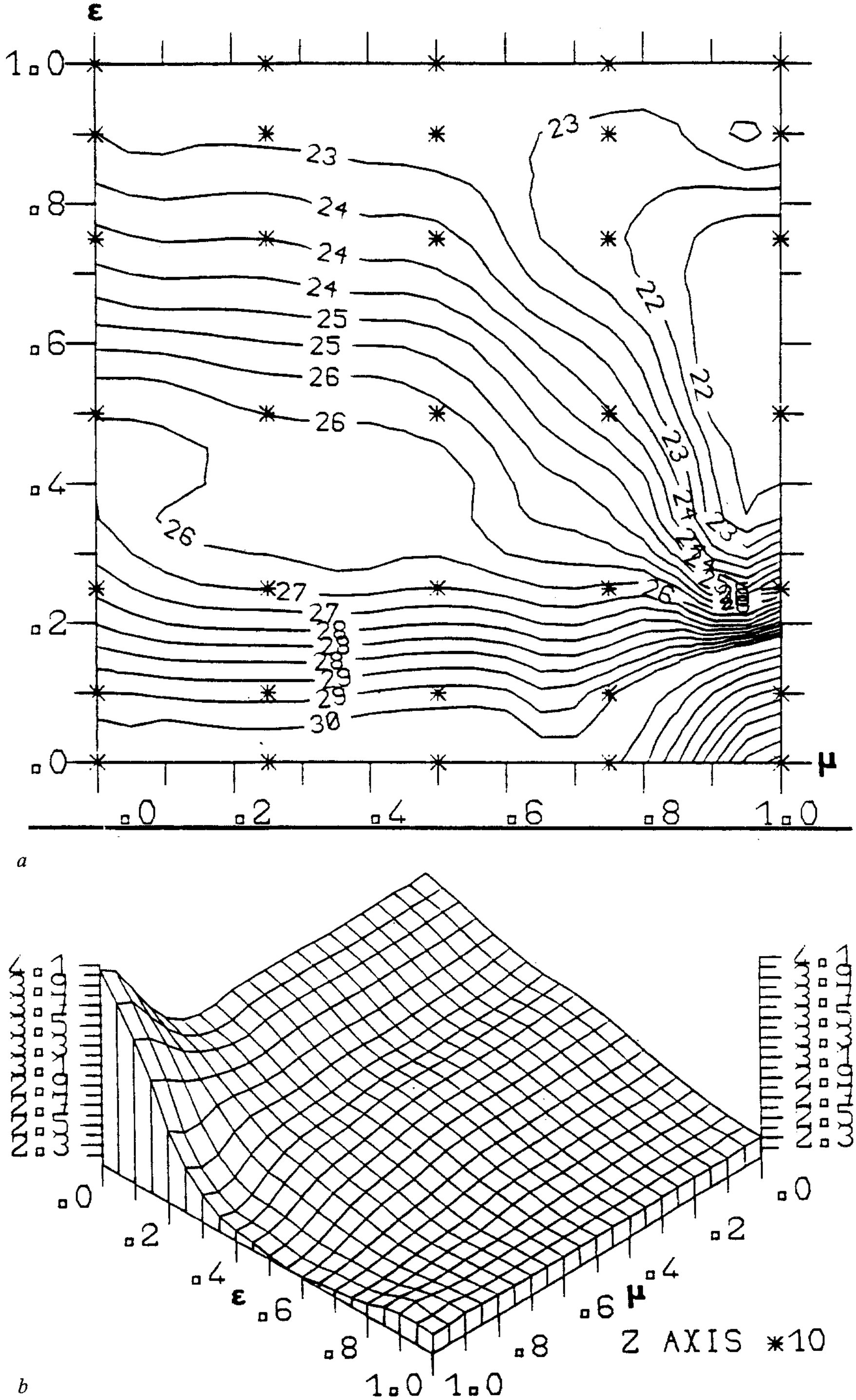

Fig. 2. The contour and the isometric projection of S for the GO model using various values of ε and μ.

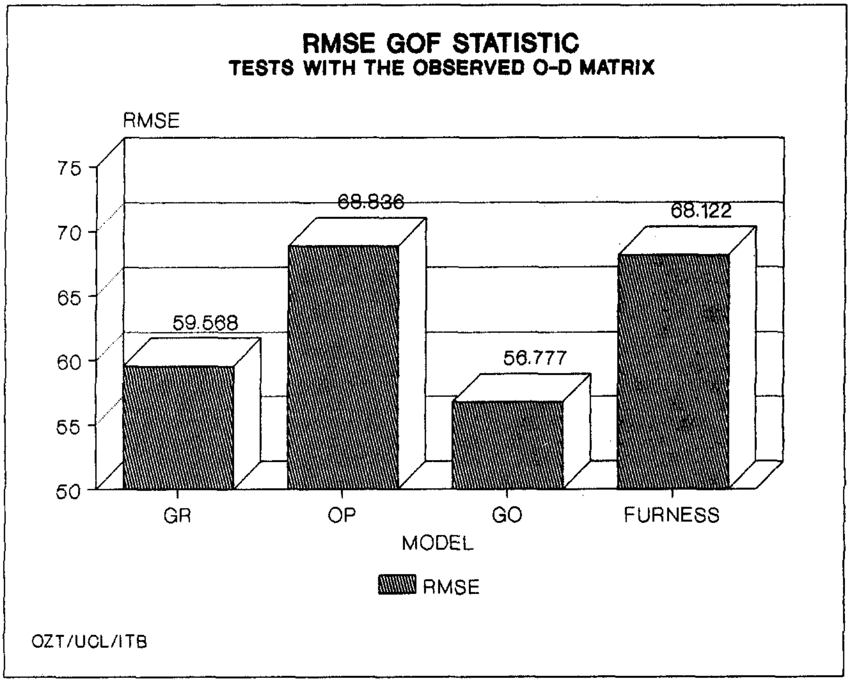

Fig. 3. GOF statistic of estimated O-D matrix compared with the observed one, using the NLLS method and direct O-D matrix information.

of GR since GO have more parameters than GR. The question is, later on, whether this richness in parameters can be exploited using traffic counts only. It was found that the OP model is not good enough to represent the trip making behaviour in Ripon since it is worse than Furness.

Other factors that also affect the level of accuracy of the estimated O-D matrix are the traffic counts themselves, the estimation method and the trip assignment technique. It is clear that the traffic counts, the reflection of the O-D matrix on the network, are never free from error and the problem now is to use this information to estimate the model parameters.

There are two important key issues that should be considered carefully: the estimation method and the trip assignment technique in determining the routes taken through the network. A number of tests have been carried out using three types of estimation methods (NLLS, WNLLS, ML), in association with two different trip assignment techniques (all-or-nothing and Burrell (10%, 30%)), and using 63 independently observed traffic counts (10% means that the variance of the link cost is 10% of the mean value). Three values of the objective functions (S, S_1, L_1) in association with three different trip assignment techniques are given in Table 1.

With all-or-nothing and Burrell (10%) assignments, the GO model

Table 1. The value of S(NLLS), S_1(WNLLS) and L_1(ML) for each trip assignment and model.

Objective function	Trip assignment								
	All-or-nothing			Burrell (10%)			Burrell (30%)		
	GR	OP(1)	GO(1)	GR	OP(1)	GO(1)	GR	OP(1)	GO(1)
$S(\times 10^6)$	48.33	78.40	49.04	55.12	83.77	56.06	99.74	107.32	102.27
$S_1(\times 10^3)$	31.39	31.12	29.19	35.75	34.60	33.51	45.02	43.31	43.27
$L_1(\times 10^4)$	108.21	107.98	108.25	108.07	107.85	108.07	106.17	106.37	106.15

Note: (1) $\rightarrow \varepsilon = 0.5$; $\mu = 1.0$

produced the closest fit to the observed traffic counts of the three models and with all three estimation methods. However, with the Burrell (30%), the best fit is obtained using the GO model for WNLLS, the OP model for ML and the GR model for NLLS. The best overall fits were always obtained using all-or-nothing assignment therefore the value of Burrell's assignment is in doubt in the case of our Ripon data base.

A close fit between estimated and observed traffic counts is not enough to show the value of a model of this kind. The acid test is to compare estimated with observed O-D matrices. Results on these tests, in terms of GOF statistic are given in Figs. 4, 5 and 6 to compare the estimated O-D

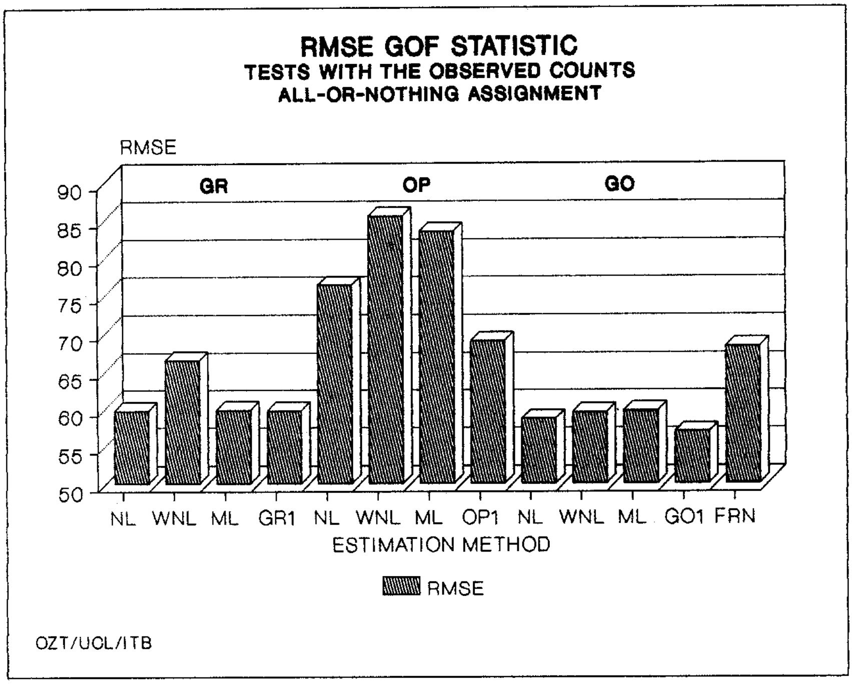

Fig. 4. GOF statistic of estimated O-D matrix compared with the observed one, using 63 observed traffic counts and all-or-nothing assignment.

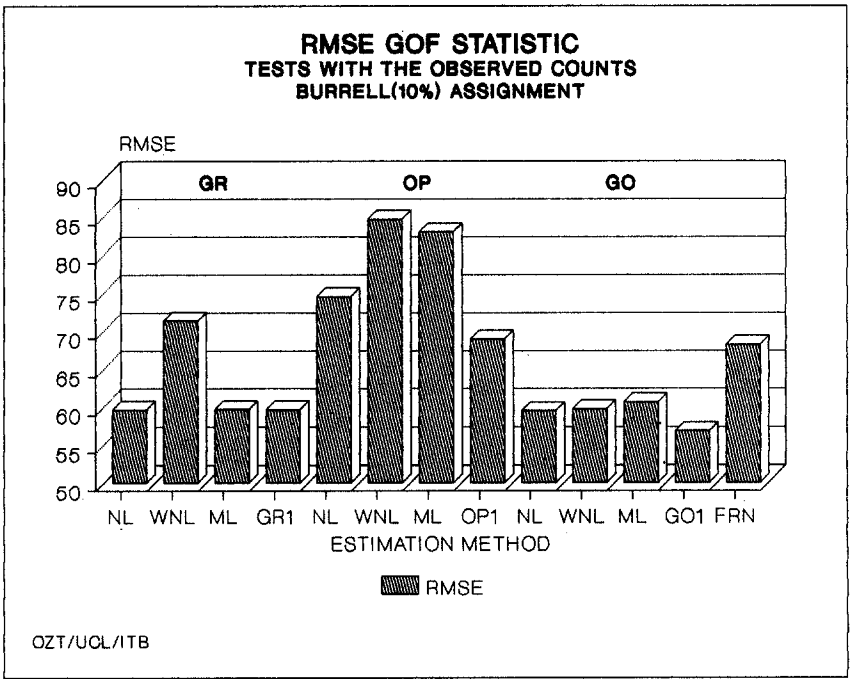

Fig. 5. GOF statistic of estimated O-D matrix compared with the observed one, using 63 observed traffic counts and Burrell (10%) assignment.

matrices, using 63 observed traffic counts with all-or-nothing, Burrell (10%) and Burrell (30%) route choice models respectively.

The RMSE GOF statistic in Fig. 4 reveals that, with all-or-nothing assignment, the NLLS and the ML methods produced the best estimated O-D matrix for the GR model whilst the WNLLS performed best for the GO model. Although the GO model reproduces the traffic counts best it does not estimate the O-D matrix most accurately as shown in Fig. 4.

The RMSE GOF statistic in Fig. 5 shows that, using Burrell (10%) assignment, the GR model produced the best fit for each estimation method. With Burrell (10%), the NLLS and ML methods obtained a slightly improved estimated O-D matrix compared with all-or-nothing assignment.

However, using the Burrell (30%), the RMSE GOF statistic in Fig. 6 shows that, each model produced worse fit for each estimation methods compared with all-or-nothing and Burrell (10%) assignments. However, it is found that the model which gives the best overall fit at the O-D matrix level is the GR model with the NLLS method and Burrell (10%) assignment. However, it can be seen from Figs. 4, 5, 6 that there is not much difference between NLLS and ML, therefore, ML is rather "robust" in that the RMSE is not very sensitive to the method of assignment.

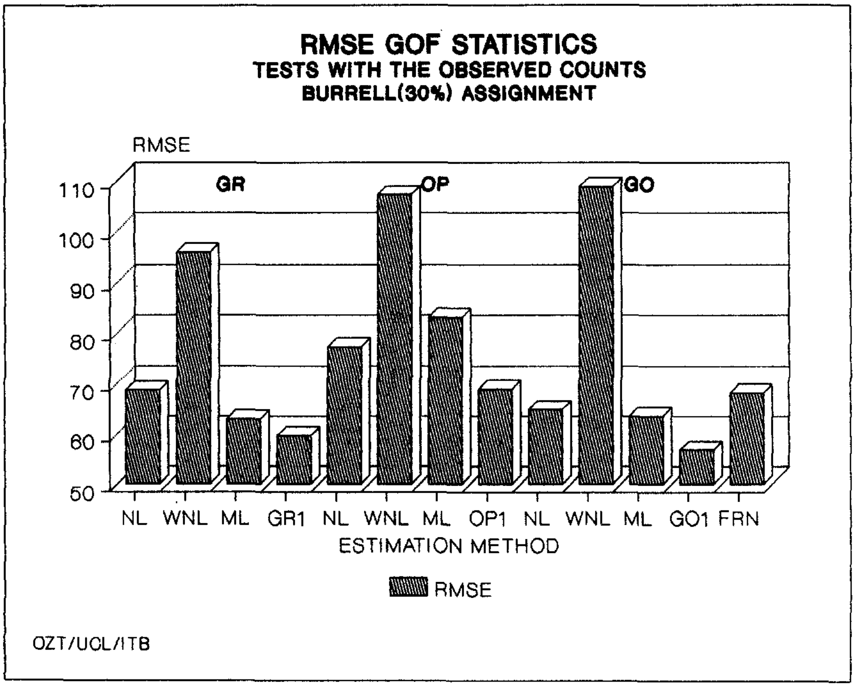

Fig. 6. GOF statistic of estimated O-D matrix compared with the observed one, using 63 observed traffic counts and Burrell (30%) assignment.

8. Conclusion

Three types of model have been used in this study: gravity (GR), opportunity (OP), and gravity-opportunity (GO) model. Three estimation methods have also been developed: least-squares (NLLS and WNLLS) and maximum-likelihood (ML) estimation methods. In all cases the models were calibrated using observed traffic counts and this requires assumptions about the most appropriate route choice model for the study area. The selected methods were then tested using Ripon data (real urban vehicle movement). The study area was divided into 19 internal zones and 7 external zones and the network has 82 nodes and 188 one-way links connecting pairs of nodes and zone centroids. The resulting equations of each method were then solved by Newton's method and Gauss-Jordan Matrix Elimination. Two trip assignment techniques were used: all-or-nothing and "stochastic" Burrell's assignment with dispersion parameters of 10% and 30%. All programs were written to be fully integrated and interactive with the MOTORS transport planning suite.

Some conclusions can be drawn from the result obtained:

- The number of observed traffic counts required are at least as many as the number of parameters. The more traffic counts you have, the faster the estimation method will converge and also the more accurate the estimated O-D matrix we have.
- It is found, not unexpectedly, that the GO and OP model are more time consuming than the GR model since they use more complicated algebra and procedures which require longer time to solve. Also, the use of all-or-nothing assignment produces faster cpu times by a factor of 3 than with stochastic Burrell, either 10% or 30%.
- A major problem found in using Newton's method is in finding the best starting point for each unknown parameter, otherwise, the method will fail to converge.
- It was found that when calibrated directly from O-D matrix data the GO model produces the best fit in Ripon. Both GO and GR models perform better than Furness.
- It is clear that, see Table 1, the choice of Burrell assignment, either 10% or 30%, gives no better fit to the traffic counts compared with the all-or-nothing. Moreover, the GO model is found as the best model in matching the observed traffic counts for all-or-nothing and Burrell (10%) assignment.
- Although the GO model performs as the best model in matching the traffic counts, it cannot be guaranteed that it will also produce the best O-D matrix, see Figs. 4, 5, 6. It was found that gravity model (GR) gives this best fit and the best estimation method in this case is non-linear-least-squares coupled with Burrell (10%) assignment. However, it is concluded that ML is rather "robust" in that its RMSE is not very sensitive to the method of assignment.
- The adoption of a more realistic assignment technique to include congestion effects is suggested in order to obtain the more accurate estimated O-D matrix. Capacity-restrained and equilibrium assignments are proposed for future work.
- The level of accuracy of the estimated O-D matrix reproduced by the model whose parameters calibrated from traffic counts depends on some following factors:
  - the transport demand model itself in representing the trip making behaviour within the study area,
  - the estimation methods used to calibrate the model from traffic counts,
  - the trip assignment techniques in determining the routes taken through the network,
  - errors in traffic counts,

- finally, the level of resolution of the zoning system and the network definition. This information should be specified carefully in order to obtain the required level of the accuracy. It is expected that the more detailed network representation and the more disaggregated the zoning system, the better the level of accuracy. Some tests by using higher level or resolution of network and zoning definition is suggested for future work.

Acknowledgements

This research was undertaken at the Transport Studies Group, University College London and has been supported by the Institute of Technology Bandung (ITB), Indonesia.

References

Box GEP & Cox DR (1964) An analysis of transformations. Journal of Royal Statistical Society 26B: 211–243

Burrell JE (1968) Multiple route assignment and its application to capacity restraint. Proceeding of the Fourth International Symposium on the Theory of Traffic Flow, Karlsruhe 1968

Edwards AWF (1972) Likelihood. Cambridge University Press, Cambridge

Hamerslag R & Immers LH (1988) The estimation of trip matrices: Shortcomings and possibilities for improvement. Faculty of Engineering, TU Delft

Högberg P (1976) Estimation of parameters in models for traffic prediction: a non-linear-regression approach. Transportation Research 10(4): 263–265

Holm J et al. (1976) Calibrating traffic models on traffic census only. Traffic Engineering and Control 17(4): 137–140

Low DE (1972) A new approach to transportation systems modelling. Traffic quarterly 26(3): 391–404

Robillard P (1975) Estimating the O-D matrix from observed link volume. Transportation Research 9(2/3): 123–128

Smith DP & Hutchinson BG (1981) Goodness of statistics for trip distribution models. Transportation Research 15A(4B): 295–303

Steer Davies & Gleave Ltd (1984) MOTORS Transportation Suite: User Manual, London

Tamin OZ (1985) Estimation of matrices for freight movement from traffic counts using a non-linear regression approach. Special study for the degree of MSc of the University of London, Imperial College and University College London (unpublished)

Tamin OZ (1988) The estimation of transport demand models from traffic counts. PhD thesis, Transport Studies Group, University College London, University of London

Van Zuylen HJ & Willumsen LG (1980) The most likely trip matrix estimated from traffic counts. Transportation Research 14B(3): 281–293

Wills MJ (1986) A flexible gravity-opportunities model for trip distribution. Transportation Research 20B(2): 89–111

Willumsen LG (1978) Estimation of an O-D matrix from traffic counts—a review. Working Paper no. 99, Institute for Transport Studies, University of Leeds

Willumsen LG (1981) Simplified transport models based on traffic counts. Transportation 10(3): 257–278

Willumsen LG (1984) Estimating time-dependent trip matrices from traffic counts. Proceedings of the Ninth International Symposium on Transportation and Traffic Theory, Delft (pp 397–411)

Wilson AG (1967) A statistical theory of a spatial distribution models. Transportation Research 1: 253–269