PUBLIC TRANSPORT DEMAND ESTIMATION BY CALIBRATING A COMBINED TRIP DISTRIBUTION-MODAL CHOICE FROM PASSENGER COUNTS: A CASE STUDY IN BANDUNG (INDONESIA)

Ofyar Z TAMIN
Senior Lecturer and Researcher
Department of Civil Engineering
Institute of Technology Bandung
Jalan Ganesha 10 Bandung, 40132
INDONESIA
Phone/Fax: +62-22-2502350
E-mail: ofyar@trans.si.itb.ac.id

abstract: The estimation of public transport demand is an expensive and time consuming; the need of low-cost method is therefore obvious. The objective is to develop the model for forecasting the public transport demand using low-cost passenger counts. The paper will report on the Gravity (GR) combined with the Multi-Nomial Logit (MNL) model which can be calibrated from traffic (passenger) counts and other simple zonal planning data. Non-Linear-Least-Squares (NLLS) and Maximum-Likelihood (ML) estimation methods were used to calibrate the parameter of the model. The model has been tested using the 1988 Public Transport Survey in Bandung (Indonesia). The model was found to provide a reasonably good fit and the calibrated parameter can then be used for forecasting purposes.

1. INTRODUCTION

Urban traffic congestion is one of the most important and critical problems always found in most of large cities in developing countries. This may due to high urbanisation and increase rate of vehicles, rapid growth of population, improvement of income level, inefficient public transport system, etc. In Indonesia, these already occur in large cities of more than 2 million population such as Jakarta, Surabaya, Medan, etc. followed by several other towns by the end of year 2000.

Delay, congestion, air pollution and vibration are among the problems. In order to alleviate these problems, various measures and actions have to be planned and implemented such as road-network extension, transport management schemes, traffic restraints, public transport policies, etc. It is therefore necessary to understand the cause of the problems which are usually due to travel pattern. The notion of Origin-Destination (O-D) matrix has been adopted by transport planners as an important tool to represent the existing travel pattern.

Unfortunately, '**conventional methods**' for estimating O-D matrices rely much on extensive surveys and interviews for example: three large transport planning studies carried out in Jakarta: Jakarta Metropolitan Area Transportation Study (**JMATS, 1975**), Arterial Road System Development Study (**ARSDS, 1987**) and Transport Network Planning Regulation (**TNPR, 1992**). All of these make them very expensive in terms of time and manpower and also time disruptive to trip makers. In developing countries, the need for inexpensive methods generally called '**unconventional methods**' which require low-cost data, less time and less manpower is therefore obvious due to time and money constraints. This becomes even more valuable for problems which require 'quick-response' treatment.

Traffic (passenger) counts, the embodiment and the reflection of the O-D matrix, provide direct information about the sum of all O-D pairs which use those links. Some reasons why traffic counts are so attractive as a data base are: firstly, they are routinely collected by many authorities due to their multiple uses in many transport planning tasks. All of these make them easily available. Secondly, they can be obtained relatively inexpensive in terms of time and manpower, easier in terms of organisation and management and also without disrupting the trip makers. Finally, the automatic collection of traffic counts is well advanced and its accuracy is satisfactorily.

A flexible model approach allows monitoring a long term plan in order to check its short term performance at regular intervals using easily-available data. If found necessary, changes to the plan may be evaluated and eventually implemented. For this reason, the approach is deemed appropriate for long term planning and project evaluation even in the case of rapid changing in land-use, socio-economic, population etc. usually occurs in most of developing countries.

A key element of the approach is a system to update the forecasting model (in particular its trip distribution and modal choice elements) using low-cost and/or easily-available information. Traffic (passenger) counts have been widely accepted as an easily-available and inexpensive information to obtain which makes them particularly attractive to be used in developing countries for planning purposes.

The estimation of public transport demand, particularly important for planning purposes in developing countries, is an expensive and time consuming undertaking. The need of low-cost method to estimate the public transport demand is therefore obvious. The development of techniques for calibrating the trip distribution models from traffic counts to obtain the O-D matrices is well advanced. Therefore, positive results on this development will be further developed and extended to enable the planner to estimate the demand for public transport for short, medium or long term planning, by combining a Trip Distribution and Modal Choice (TDMC) model and calibrating it using low-cost traffic (passenger) count information.

The paper will report on a family of aggregate model combined with a family of modal choice logit model which can be calibrated from traffic (passenger) counts and other simple zonal planning data. The TDMC model examined was the Gravity (GR) model combined with the Multi-Nomial Logit (MNL) model. Non-Linear-Least-Squares (NLLS) and Maximum Likelihood (ML) estimation methods were used to calibrate the parameters of the combined model. The combined TDMC model and the calibration methods have been implemented into a micro-computer package capable of dealing with the study area consisting of up to 300 zones, 3000 links and 6000 nodes.

2. PREVIOUS RESEARCH

One can interpret link flows (or traffic counts) as resulting from a combination of two elements: an O-D matrix and the route choice pattern selected by drivers on the network.

These two elements may be linearly related to traffic counts, see equation (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 idea of combining 'traditional' data sources (home or roadside interviews) with low cost data like traffic counts is not entirely new (see **Van Zuylen and Willumsen**, **1980**; **Willumsen**, **1981ab**; **Tamin**, **1988ab**; **Tamin** ***etal*****, 1990**). The models can be used to combine, for example, roadside interview data with traffic counts and this can be achieved with or without an explicit travel demand model (trip distribution model) (see **Tamin 1988b**; **Willumsen**, **1981a**). The models has also been tested for several cases for example for estimating the freight transport demand (see **Tamin**, **1985**; **Tamin and Willumsen**, **1988**; **Tamin and Soegondo**, **1989**).

For the purpose of public transport demand estimation, this idea can be extended to the development of a practical estimation approach to calibrate the combined trip distribution and modal choice (TDMC) model using traffic (passenger) counts and other simple zonal planning data. This approach assumes that either trip distribution or modal choice model is represented by certain model forms. As usual, the traffic (passenger) counts are expressed as a function of the TDMC model. In this case, the TDMC model is represented a function of a model form and relevant parameters. The parameters of the postulated model are then estimated, so that the errors between the estimated and observed traffic (passenger) counts are minimised.

Estimation methods of this kind have been proposed by **Low (1972)** and **Holm** ***etal*** **(1976)** using linear-least-squares, **Robillard (1975)**, **Högberg (1976)** and **Tamin (1985)** using non-linear-least-squares estimation method. Other relevant method namely maximum-likelihood estimation method has been proposed by **Tamin (1988b)**. There has been, however, little validation of these methods and there is a scope for exploring more advanced and flexible model forms.

3. DEFINITIONS

For the simplification purposes, we define the following terms as follows:

$[T_{id}]$	=	the observed O-D matrix from origin i to destination d.
O_i^m	=	the total trips of each mode m generated by origin i.
D_d^m	=	the total trips of each mode m attracted by destination d.
A_i^m, B_d^m	=	the balancing factors for each mode m for origin i and destination d.
C_{id}^m	=	the trip cost of travelling from origin i to destination d by mode m.
β	=	the unknown estimated parameter to be calibrated.
V_l^{+m}, V_l^m	=	the estimated and observed traffic counts for mode m, respectively.
p_{id}^{lm}	=	the trip assignment proportion for trips by mode m from origin i to destination d which use link l.
L	=	the total number of links observed.

M	=	the total number of modes.
N	=	the total number of origins or destinations.

We use the notational conventional that $\sum_{m}$ means the summation begins at m=1 and continues over the entire range of the subscript.

4. TRIP DISTRIBUTION-MODAL CHOICE (TDMC) MODEL BASED ON TRAFFIC COUNTS

4.1 The Main Principle of the Problem

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 **N²** cells. **[N²-N]** cells if intra-zonal trips can be disregarded. The most important stage for this combined transport model based on traffic (passenger) counts is to identify the paths followed by the trips from each origin to each destination.

The variable p_{id}^{lm} is used to define the proportion of trips by mode m from zone i to zone d travelling through link l. Thus, the flow on each link is a result of:

- trips by mode m from zone i to zone d (T_{id}^{m}), and
- the proportion of trips by mode m from zone i to zone d whose trips use link l which is defined by p_{id}^{lm} ($0 \leq p_{id}^{lm} \leq 1$).

The flow (V_l^m) in a particular link l is the summation of the contributions of all trips by mode m between zones to that link. Mathematically, it can be expressed as follows:

$$V_l^m = \sum_i \sum_d T_{id}^m p_{id}^{lm} \tag{1}$$

Given all the p_{id}^{lm} and all the observed traffic counts (V_l^m), then there will be **N²** unknown T_{id}^m's to be estimated from a set of **L** simultaneous linear equations (1) where **L** is the total number of traffic (passenger) counts.

In principle, **N²** independent and consistent traffic counts are required in order to determine uniquely the O-D matrix [T_{id}^m]. (**N²-N**) if intra-zonal trips can be disregarded. In practice, the number of observed traffic (passenger) counts is much less than the number of unknowns T_{id}^m'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.

4.2 Some Trip Assignment Techniques

The trip assignment aims at the determination of the routes taken through the network and hence the total number of trips using each particular link. There are two main groups of trip assignments as classified by **Robillard (1975)**, namely: **proportional** and **non-proportional** trip assignment. There are several characteristics of the study area that should help to define the best assignment technique. These include information on how the drivers perceive their travel costs, level of congestion and the availability of alternative routes and their corresponding costs.

In **proportional** assignment, it is assumed that proportion of drivers choosing each route will depend on assumed drivers and route characteristics but not on flow levels. The most common example is **all-or-nothing** assignment. It is assumed that all drivers wish to minimise their perceived travel costs and all the drivers perceive these costs in the same way, 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 congestion effect. However, it is the fastest and easiest assignment and is useful for simple networks where there are only few alternatives O-D paths. The value of p^l_{id} for this assignment is defined as follows:

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

In **non-proportional** assignment, it is said that, under congested conditions, the cost of travelling on a link depends on the flow on that link and the cost-flow relationship. Several techniques have been developed to explicitly consider these effects and they are usually called **capacity-restrained** assignment.

4.3 Trip Distribution-Modal Choice Model

4.3.1 Gravity Model (GR) as a Trip Distribution Model

This model is developed based on the analogy with Newton's law of gravitational. Suppose now there are **M** modes travelling between zones, the modified gravity model (**Doubly-Constrained Gravity Model**) can then be expressed as:

$$T_{id} = \sum_m \left(O^m_i . D^m_d . A^m_i . B^m_d . f^m_{id}\right) \tag{2}$$

where: A^m_i and B^m_d = the balancing factors expressed as:

$$A^m_i = \left[\sum_d \left(B^m_d . D^m_d . f^m_{id}\right)\right]^{-1} \text{ and } B^m_d = \left[\sum_i \left(A^m_i . O^m_i . f^m_{id}\right)\right]^{-1} \tag{3}$$

f^m_{id} = the deterrence function = $\exp\left(-\beta . C^m_{id}\right)$

4.3.2 Multi-Nomial Logit Model (MNL) as a Modal Choice Model

A modal split model will be required if public transport proposals are being considered, or if it is thought likely that proposals for the highway or parking system will lead to a significant transfers of trips between modes, hence altering modal split.

The purpose of this stage is the modelling of the choice between modes of conveyance for each trip, usually car and one or more public transport modes in the case of passengers. For freight movement the choice is also often between lorry and train. The modal split stage also treats those trips which use more than one mode to complete the whole trip.

The most general and simplest modal choice model (Multi-Nomial Logit Model) was used in this study. It can be expressed as:

$$T_{id}^{k} = T_{id} \cdot \frac{\exp\left(-\beta . C_{id}^{k}\right)}{\sum_{m} \exp\left(-\beta . C_{id}^{m}\right)} \tag{4}$$

4.3.3 Fundamental Basis

By substituting equations (2)-(4) to equation (1), then **the fundamental equation** for the estimation of a combined transport demand model from traffic counts is:

$$V_{l}^{k} = \sum_{d} \sum_{i} \left[O_{i}^{k} . D_{d}^{k} . A_{i}^{k} . B_{d}^{k} \cdot f_{id}^{k} \cdot p_{id}^{lk} \frac{\exp\left(-\beta . C_{id}^{k}\right)}{\sum_{m} \exp\left(-\beta . C_{id}^{m}\right)} \right] \tag{5}$$

Equation (**5**) is a system of **L** simultaneous equations with only one (**1**) unknown parameter β need to be estimated. The problem now is how to estimate the unknown parameter β so that the model reproduces the estimated traffic (passenger) counts as close as possible to the observed ones.

4.4 Non-Linear-Least-Squares Estimation Method (NLLS)

The main idea of this method is to estimate the unknown parameter which minimises the sum of the squared differences between the estimated and observed traffic (passenger) counts. The problem now is:

to minimise $$S = \sum_{l} \left[V_{l}^{+k} - V_{l}^{k} \right]^2 \tag{6}$$

Having substituted (**5**) to (**6**), the following set of equation is required in order to find a set of unknown parameter β which minimises equation (**6**):

$$\frac{\delta S}{\delta \beta} = \sum_{l} \left[\left(2 \sum_{i} \sum_{d} T_{id}^{k} \cdot p_{id}^{lk} - V_{l}^{k} \right) \left(\frac{\sum_{i} \sum_{d} \delta T_{id}^{k}}{\delta \beta . p_{id}^{lk}} \right) \right] = 0 \tag{7}$$

Equation (7) is a equation which has one (1) unknown parameter β need to be estimated. Then it is possible to determine uniquely all the parameters, provided that **L**>**1**. Newton's method combined with the Gauss-Jordan Matrix Elimination technique can then be used to solve equation (7) (see **Batty, 1976; Wilson and Bennet, 1985**).

4.5 Maximum-Likelihood Estimation Method (ML)

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

$$\mathbf{L}\left(\mathbf{H}|\mathbf{R}\right) = c.\mathbf{P}\left(\mathbf{R}|\mathbf{H}\right) \tag{8}$$

The arbitrary constant of proportionality '**c**' enables us to use the same definition of likelihood for discrete and continuous variables. Though it is a constant in any one application, involving many different hypotheses but the same data and probability model, it is, of course, not necessarily the same constant in another application.

The word '**Support**' as suggested by **Edward (1972)** has been used to refer to the natural logarithm of the likelihood or the likelihood ratio, principally in order to convert the multiplicative properties into the additive ones. The value of Support in favour of the one hypothesis than of the other will be in the range of **0** up to indefinite value (∞).

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. At **S=5** it is about 150. It is obvious now that the greater the value of **S**, the greater will be the likelihood in favour of the one hypothesis than of the other.

Suppose now we have available a traffic (passenger) count data survey involving a total number V_T^m passengers (of mode m) counted in **L** independent links. If V_l^m denotes the observed independent traffic (passenger) count obtained for each particular link l, then we must have:

$$\sum_l V_l^m = V_T^m \tag{9}$$

Now, let assume that p_l^m be the probability of having an estimated independent traffic (passenger) count (of mode m) for each particular link l expressed as:

$$p_l^m = \frac{V_l^{+m}}{V_T^m} \tag{10}$$

Following the 'normal distribution' and assuming that there be '**L**' links counted and also the probability of having an independent link count in the 'l^{th}' is p_l^m, then the probability of obtaining 'V_1^m' in the first link, 'V_2^m' in the second link, and in general, in the l^{th} link is:

$$\mathbf{P} = \frac{V_T^m\,!}{V_1!.V_2!.V_3!\ldots V_L}\prod_l \mathbf{p}_{\mathrm{l}}^{\mathrm{V}_\mathrm{l}^{\mathrm{m}}} \tag{11}$$

Then by subtituing equation (**11**) to equation (**8**), the '**Likelihood Function**'can be expressed as:

$$\mathbf{L} = \prod_l \mathbf{p}_{\mathrm{l}}^{\mathrm{V}_\mathrm{l}^{\mathrm{m}}} \tag{12}$$

Where the term of $\dfrac{V_T^m\,!}{V_1!.V_2!.V_3!\ldots V_L!}$ has been absorbed into the arbitrary constant **c**.

Finally, the framework of the Maximum-Likelihood estimation method is that the choice of the hypothesis **H** maximising equation (**12**) subject to particular constraint, will yield a distribution of V_l^m giving the best possible fit to the survey data (V_l^{+m}). The objective function for this framework is expressed as:

to maximise $$\mathbf{L} = \prod_l \mathbf{p}_{\mathrm{l}}^{\mathrm{V}_\mathrm{l}^{\mathrm{m}}} \tag{13}$$

subject to $$\sum_l V_l^{+m} - V_T^m = 0 \tag{14}$$

The purpose of equation (**14**) is to constrain that the total of estimated passenger flows should be equal to the total of observed passenger counts for each particular links obtained in the survey. Taking the natural logarithm term of equation (**13**) and using the following '**Lagrangian Multiplier**' method, equations (13)-(**14**) can then be written into a single equation. That is:

to maximise $$\mathbf{L}_1 = \sum_l V_l^m \ln p_l^m + \ln c - \theta\left(\sum_l V_l^m - V_T^m\right) \tag{15}$$

By subtituting equation (**1**) to equation (**15**), finally, the objective function can then be expressed:

to maximise $$\mathbf{L}_1 = \sum_l \left[V_l^m.\ln\left(\sum_i\sum_d T_{id}^m \cdot p_{id}^{lm}\right) - \theta.\sum_i\sum_d T_{id}^m \cdot p_{id}^{lm} \right] + \theta.V_l^m - V_l^m.\ln V_T^m + \ln c \tag{16}$$

with respect to parameters β and θ.

In order to determine uniquely parameter β of the GR model together with an additional parameter θ which maximises equation (**16**), the following two sets of equations are then required. They are as follows:

$$\frac{\delta \mathbf{L}_1}{\delta\beta} = f\beta = \sum_l \left[V_l^m \frac{\sum_i\sum_d \dfrac{\delta T_{id}^m}{\delta\beta.\, p_{id}^{lm}}}{\sum_i\sum_d T_{id}^m \cdot p_{id}^{lm}} \right] - \theta.\sum_i\sum_d \frac{\delta T_{id}^m \cdot p_{id}^{lm}}{\delta\beta.\, p_{id}^{lm}} = 0 \tag{17a}$$

$$\frac{\delta \mathbf{L}_1}{\delta\theta} = f\theta = -\theta\left[\sum_i\sum_d T_{id}^m \cdot p_{id}^{lm} - V_T^m\right] = 0 \tag{17b}$$

Equation (**17ab**) is in effect a system of 2 (two) simultaneous equations which has 2 (two) parameters β and θ need to be estimated. Again, the Newton's method combined with the Gauss-Jordan Matrix Elimination technique can then be used to solve for equation (**17**).

5. MOTORS PACKAGE PROGRAM

Micro-computers as scaled down mainframe computers offer reasonable capacity, speed and reliability at a very low cost. These characteristics make micro-computers very attractive and are the most important tools for transport engineers and planners. Willumsen (1985) mentioned that there are at least three comprehensive commercially marketed transport planning suites in Britain available on both 8 bit and 16 bit machines, they are: **MICROTRIPS**, **MOTORS** and **MINITRAMP**. Computer programs have been written in a such a way that it is fully integrated and interactive with **MOTORS** suite. MOTORS is a transport planning suite designed by professional transport planners and micro-computer specialists (**Steer Davies and Gleave, 1984**).

The work was carried out using the IBM compatible micro-computer with 640K RAM, two 360Kbyte 5½' floppy disk drives, 30M byte Winchester disk, a monitor and a printer. The MS-DOS 6.0 operating system was used in combination with the MOTORS suite plus word processing software.

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 are the resulting models 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) statistical tests are used, as suggested in **Wilson (1967)** and **Tamin (1988b)**:

<u>Root Mean Square Error (RMSE) and Relative %RMSE</u>

a. $$\text{RMSE} = \sqrt{\left[\sum_i \sum_d \frac{\left(T_{id} - T_{id}^{+}\right)^2}{\text{N}(\text{N}-1)}\right]} \quad \text{for } i \neq d \qquad (18)$$

b. $$\%\text{RMSE} = \text{RMSE}/T_1 \text{ x } 100\% \qquad (19)$$

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

$$T_1 = \frac{\left(\sum_i \sum_d T_{id}\right)}{\left[\text{N}(\text{N}-1)\right]} \quad \text{for } i \neq d \qquad (20)$$

7. TESTS WITH BANDUNG PUBLIC TRANSPORT DATA SURVEY

7.1 Bandung Data Set

Bandung as the capital of West Java Province is now facing some transport problems especially congestion due to inefficient public transport system. There are two groups of public transport available in Bandung: Damri operated under government agency and Angkot (Angkutan Kota) under private agency. This study stresses in analysing the demand of Angkot since Angkot has more routes and serves more area compared to Damri. Furthermore, it seems that Angkot gives more contribution to congestion in Bandung especially nearby intersection. Therefore, there are only two modes being considered in this study: mode by private vehicle and mode by Angkot.

The Regional Government of Bandung together with BUTP (Bandung Urban Transport Planning) and some related agencies have been carried out several extensive transport surveys and studies in order to overcome the problems of which some of them was used in this study. A public transport data survey of Bandung carried out in 1988 was used to test the proposed model and estimation method. The Bandung data set comprises the following:

- the 1988 O-D matrices for private transport (home-based) and public transport (Angkot) compiled from roadside interview collected in 1988;
- the 1988 Bandung network description together with information of 85 Angkot routes; and
- 65 traffic (passenger) counts for Angkot taken throughout Bandung in 1988.

The study area was coded into 15 internal zones and 10 external zones representing the remaining areas of the city. The network of this study area has 264 nodes and 734 one-way links connecting pairs of nodes and zone centroids. In addition to the information obtained from the network definition, some information of the observed O-D matrix $[\boldsymbol{T_{id}^m}]$ with its corresponding trip ends ($\boldsymbol{O_i^m}$ and $\boldsymbol{D_d^m}$) for each mode m were also obtained. The units adopted in (**5**) are given as follows:

V_l^m = passenger flows counted for each particular link $\boldsymbol{l}$ for mode $\boldsymbol{m}$ (Angkot) in persons/day

$\boldsymbol{O_i^m}, \boldsymbol{D_d^m}$ = trip generation and attraction factors for each mode m (private and Angkot), respectively in persons/day.

7.2 Results

In assessing the value of a new transport model, someone is of course interested in the accuracy of the estimated travel behaviour. In this case, the 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 a 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.

Due to limited time, the model chosen to represent the trip making behaviour was the gravity (GR) model and the Multi-Nomial Logit for the modal choice model. It is expected the better the model we use, the better the acuracy of the estimated O-D matrices. Other factors that also affect the level of accuracy of the estimated O-D matrix are the traffic counts themselves and the estimation method. 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 it to use this information to estimate the model parameter. A number of tests have been carried out using Non-Linear Least Squares (NLLS) and Maximum-Likelihood estimation methods and using 65 independently observed traffic (passenger) counts. **Table 1** shows the Goodness of Fit (GoF) statistical test of the estimated matrices compared to observed one using 65 observed traffic (passenger) counts.

Table 1: GoF Statistical Test using 65 Observed Traffic (Passenger) Counts

	PRIVATE				BUS			
MODEL	**TDMC Model**				**TDMC Model**			
Method	**NLLS**	**ML**	**TDMC[1]**	**Furness**	**NLLS**	**ML**	**TDMC[1]**	**Furness**
RMSE	94.231	94.125	94.087	96.802	60.334	60.122	60.087	64.802
%RMSE	90.853	90.811	90.700	92.370	58.953	58.866	58.700	62.370

Note: TDMC[1] obtained from the full O-D Survey

It is found that by having information of Angkot' passenger counts, we can obtain the O-D matrices for private and Angkot. It is also shown that the TDMC model with NLLS and ML estimation methods produced better result compared to the Furness model. The most encouraging results is the use of traffic (passenger) counts to estimate the O-D matrix is only marginally worse compared to the one obtained by the full O-D survey. This result is very important in terms of time and money for estimating the demand of public transport and also for forecasting purposes.

8. CONCLUSION

The estimation method used in this study are Non-Linear-Least-Squares (NLLS) and Maximum-Likelihood (ML) together with the TDMC model. In all cases the models were calibrated using observed traffic (passenger) counts and the selected methods were then tested using Bandung data (urban public transport movement). The study area was coded into 15 internal zones and 10 external zones and the network has 264 nodes and 734 one-way links connecting pairs of nodes and zone centroids.

The resulting equation was then solved by Newton's method and Gauss-Jordan Matrix Elimination. 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 (passenger) counts required are at least as many as the number of parameters. The more link flows you have, the faster the estimation method will converge and also the more accurate the estimated O-D matrix we have. From several research and application, it can be concluded that the optimal number of traffic counts required is between 25-30% of total number of links in the network.
- The calibrated model can then be used to forecast and also to evaluate the effect of ticket fare to the public transport demand.

- It is found that by having the information of traffic flows using Angkot, we can obtain the O-D matrices for private and Angkot.
- Furthermore, the results are very encouraging since the estimated O-D matrices obtained using traffic count information are only marginally worse than those obtained by full O-D survey.
- The level of accuracy of the estimated O-D matrices depends on some following factors:
  a. the transport demand model itself in representing the trip making behaviour within the study area,
  b. the estimation method used to calibrate the model from traffic counts,
  c. errors in traffic (passenger) counts,
  d. 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.

REFERENCES

Batty, M. (1976) **Urban Modelling: Algorithm, Calibrations and Predictions**. Cambridge University Press, Cambridge.

Bina Marga (1987) **Arterial Road System Development Study**, Final Report.

Dephubdat (1975) **Jakarta Metropolitan Area Transportation Study**, Final Report.

Dephubdat (1992) **Transport Network Planning Regulation**, Final Report.

Edwards, A.W.F. (1972) **Likelihood**. Cambridge University Press, Cambridge.

Högberg, P. (1976) Estimation of Parameters in Models for Traffic Prediction: A Non-Linear-Regression Approach. **Transportation Research 10(4)**, 263-265.

Holm, J. *etal* (1976) Calibrating Traffic Models on Traffic Census Only. **Traffic Engineering and Control 17(4)**, 137-140.

Low, D.E. (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 Volumes. **Transportation Research 9(2/3)**, 123-128.

Steer, Davies and Gleave (1984) **MOTORS Transportation Suite**, User Manual, London.

Tamin, O.Z. (1985) Estimation of Matrices for Freight Movement From Traffic Counts Using a Non-Linear Regression Approach. **MSc(Eng) Thesis**, University of London.

Tamin, O.Z. (1988a) Transport Demand Model Estimation From Traffic Counts. **Proceedings of the 20th Universities Transport Studies Group Annual Conference**, University College London, London.

Tamin, O.Z. (1988b) The Estimation of Transport Demand Models From Traffic Counts, **PhD Dissertation**, University of London.

Tamin, O.Z. and Willumsen, L.G. (1988) Freight Demand Model Estimation From Traffic Counts. **Proceedings of the 16th PTRC Summer Annual Meeting**, University of Bath, England.

Tamin, O.Z. and Soegondo, T. (1989) Modelling Freight Demand From Traffic Counts: A Case Study in Bali. **Proceedings of the 5th World Conference on Transport Research (WCTR)**, Yokohama, Japan.

Tamin, O.Z. *etal* (1989). Simplified Methods Based on Traffic Counts for Estimating O-D Matrices for Developing Countries. **Proceedings of SEATAC Urban Transport Seminar**, Jakarta.

Van Zuylen, H.J. and Willumsen, L.G. (1980) The Most Likely O-D Matrix Estimated From Traffic Counts. **Transportation Research 14B(3)**, 281-293.

Willumsen, L.G. (1981a) An Entropy Maximising Model for Estimating Trip Matrices From Traffic Counts, **PhD Dissertation**, University of Leeds.

Willumsen, L.G. (1981b) Simplified Transport Models Based on Traffic Counts. **Transportation 10(3)**, 257-278.

Willumsen, L.G. (1985) Microcomputer Software for the Application of O.R. Techniques to Traffic and Transport, Proceedings of the Conference of Operational Research Models on Microcomputers, Lisbon.

Wilson, A.G. (1967) A Statistical Theory of a Spatial Distribution Models. **Transportation Research 1**, 253-269.

Wilson, A.G. and Bennet, R.J. (1985) **Mathematical Methods in Human Geography and Planning**. John Willey and Sons Ltd.