SESSION C3: TAXI AND NETWORK MODELLING

SESSION C4: PUBLIC TRANSPORT

SESSION C5: SAFETY AND TRAFFIC ACCIDENTS

LEMBAR PENGESAHAN

Judul	:	Public Transport Deman Estimation by Calibrating the Combined Trip Distribution-Mode Choise (TDMC) Model From Passenger Counts
Penulis	:	Rahayu Sulistyorini
NIP	:	19741004 2000032002
Instansi	:	Fakultas Teknik, Universitas Lampung
Publikasi	:	Prosiding Internasional
	:	ISBN: 978-988-98847-7-2
	:	Bulan Desember dan Tahun 2009
Penerbit	:	Hongkong Society for Transportation Studies

Bandar Lampung, 31 Mei 2012

Mengetahui,
a.n. Dekan Fakultas Teknik
Universitas Lampung
Pembantu Dekan I

Dr. Eng. Helmy Fitriawan, S.T, M.Sc.
NIP 19750928 2001121002

Penulis,

Dr. Rahayu Sulistyorini, S.T, M.T.
NIP. 19741004 2000032002

Menyetujui:

Ketua Lembaga Penelitian
Universitas Lampung

Dr. Eng. Admi Syarif.
NIP. 196701031992031003

PUBLIC TRANSPORT DEMAND ESTIMATION BY CALIBRATING THE COMBINED TRIP DISTRIBUTION-MODE CHOICE (TDMC) MODEL FROM PASSENGER COUNTS

Ofyar Z. TAMIN [a] and Rahayu SULISTYORINI [b]

[a] Jl.Ganesha 10 Bandung
Department of Civil Engineering
Institute Technology Bandung (ITB)
[b] Department of Civil Engineering
Institute Technology Bandung (ITB)
Jl.Ganesha 10 Bandung

ABSTRACT

The conventional method to estimate the O-D matrices require very large surveys and very expensive. The need for inexpensive methods, which require low-cost data, generally called as 'unconventional method'. The development of techniques for calibrating the trip distribution models from traffic volumes to obtain the O-D matrices is well advanced. However, the previous research still in a burden condition of "All or Nothing" which is not realistic for some congested road networks in urban area. So, the main objective and contribution of this research is the estimation of origin-destination matrices by calibrating the combined gravity with multinomial logit under equilibrium assignment. The estimation methods, namely: Non-Linear-Least-Squares (NLLS) will be used to estimate the parameters of transport demand models. The combined model and its calibration method have been implemented. The model was able to obtain the calibrated parameters which can then be used for forecasting purposes. The advantageous and the applicability of the model are given.

1. BACKGROUND

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 urbanization and increase rate of vehicles, rapid growth of population, improvement of income level, inefficient public transport system, etc. 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 restraint, 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 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 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 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 inexpensively in terms of time and manpower, easier in terms of organization 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 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 development of techniques for calibrating the trip distribution models from traffic counts to obtain the O-D matrices is well advanced (see Tamin, 1988; Tamin and Willumsen, 1988; Tamin, 1992).

Therefore, positive results on this development will be further developed and extended to enable the transport planner to estimate the demand for public transport for short, medium or long term planning. The main idea is by combining a Trip Distribution and Mode Choice (TDMC) model and calibrating it using low-cost traffic (passenger) count information (see Tamin, 1988). However, the previous research still in a burden condition of "All or Nothing" which is not realistic for some congested road networks in urban area. This method is not realistic for some congested road network in urban area because it never consider to the traffic jam effect and various perception in considering of route selection. Referring to the previous research, the main objective of the research development is develop and demonstrate the application of a combined trip distribution-mode choice model estimated from traffic count under equilibrium condition.

2. THE MAIN PRINCIPLE OF THE PROBLEM

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 approach assumes that either trip distribution or mode 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 by 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 minimized.

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

[illegible] = the observed O-D matrix from origin *i* to destination *d*.

[illegible] = the total trips of each mode m generated by origin *i*.

[illegible] = 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 traveling from origin i to destination d by mode m.
β = the unknown estimated parameter to be calibrated.
$\hat{V}_l^m, V_l^m$ = 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.

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^2 cells. [N^2-N] cells if intra zone 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 traveling 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} ($0 \le p_{id}^{lm} \le 1$).

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:

$$V_l^m = \sum_i \sum_d T_{id}^m \, p_{id}^{lm} \qquad (1)$$

Given all the p_{id}^{lm} and all the observed traffic counts (V_l^m), then there will be N^2 unknown T_{id}^m to be estimated from a set of L simultaneous linear equations (1) where L is the total number of traffic (passenger) counts. In principle, N^2 independent and consistent traffic counts are required in order to determine uniquely the O-D matrix [T_{id}^m]. (N^2 - 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 modeling the trip making behavior.

3. GRAVITY-MULTINOMIAL LOGIT MODEL

The Gravity (GR) model is developed by analogy with Newton's law of gravitation. Newton asserted that the force of attraction, F_{id}, between two bodies is proportional to the product of their masses, m_i and m_d, divided by the square of the distance between them (d_{id}^2). In geography, 'force' is associated with the numbers of movements or trips between two regions; 'mass' is replaced by a variable such as population size and measures a region's capacity either to generate or to attract trips; and distance is either measured in physical terms or replaced by a more relevant variable such as travel cost or time. The analogous transport gravity model is:

$$T_{id} = k\frac{O_i O_d}{d_{id}^2} \quad k \text{ is a constant} \tag{2}$$

This model has some sensible properties. It says that the number of trips from zone i to zone d is directly proportional to each of O_i and D_d and inversely proportional to the square of the distance between them. Hence, if a particular O_i and a particular D_d are each doubled, then the number of trips between these zones would quadrate according to equation (1), when one would be expected that they would only double.

Therefore, the following constraint equations on T_{id} should always be required, such constraints are not satisfied by equation (1):

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

where O_i and D_d directly represent the total number of trips originating and terminating at i and d respectively. These constraint equations can be satisfied if sets of constants A_i and B_d associated with production zones and attraction zones respectively are introduced. They are sometimes called 'balancing factors'. Suppose now there are M modes traveling between zones, the modified gravity model (Doubly Constrained Gravity Model) can then be expressed as:

$$T_{id} = \sum_m \left(O_i^m . D_d^m . A_i^m . B_d^m . f_{id}^m\right) \quad \text{where} \tag{3}$$

A_i^m and B_d^m = the balancing factors expressed as:

$$A_i^m = \left[\sum_d \left(B_d^m . D_d^m . f_{id}^m\right)\right]^{-1} \text{ and } B_d^m = \left[\sum_i \left(A_i^m . O_i^m . f_{id}^m\right)\right]^{-1} \tag{4}$$

The equations for A_i^m and B_d^m are solved iteratively, and it can be easily checked that they ensure that T_{id} given in equation (3) satisfies the constraint equation (2). This process is repeated until the values of A_i^m and B_d^m converge to certain unique values. So far, there is no reason to think that distance plays the same role in transport. Hence, a general function of travel distance or generalized cost, normally called as 'deterrence function', is introduced.

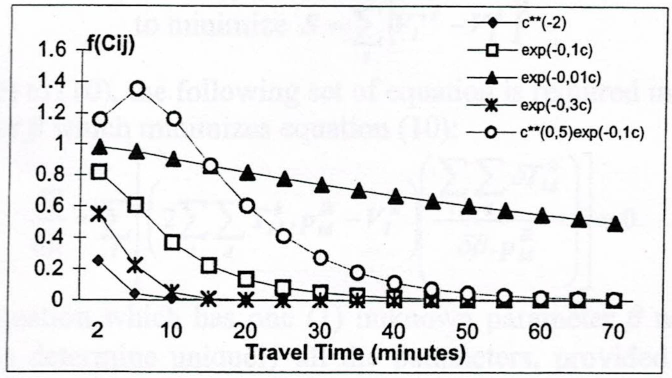

Figure 1 Types of deterrence function

There are three types of deterrence functions being used in this study which are also shown in figure 1, namely:

- $f(C_{id}) = C_{id}^{-\alpha}$ ('Power' function) (5)
- $f(C_{id}) = e^{-\beta C_{id}}$ ('Exponential' function) (6)
- $f(C_{id}) = C_{id}^{\alpha}.e^{-\beta C_{id}}$ ('Tanner' function) (7)

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 mode 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 mode split stage also treats those trips which use more than one mode to complete the whole trip. The most general and simplest mode choice model (Multi-Nomial Logit Model) was used in this study. It can be expressed as:

$$T_{id}^{k} = T_{id}.\frac{\exp\left(-\beta.C_{id}^{k}\right)}{\sum_{m}\exp\left(-\beta.C_{id}^{m}\right)} \tag{8}$$

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:

$$V_l^k = \sum_{d}\sum_{i}\left[O_i^k.D_d^k.A_i^k.B_d^k.f_{id}^k.p_{id}^{lk}\frac{\exp\left(-\beta.C_{id}^{k}\right)}{\sum_{m}\exp\left(-\beta.C_{id}^{m}\right)}\right] \tag{9}$$

Equation (9) is a system of L simultaneous equations with only one (1) unknown parameter (β) need to be estimated, assuming that O_i^k, D_d^k, p_{id}^{lk}, C_{id}^k are known for all i, d, k, and l. For public transport which has fixed routes, it is not difficult to determine p_{id}^{lk}. The following variables O_i^k and D_d^k can be obtained by estimating the trip generation and attraction factors for each mode k using certain trip generation procedures.

Hence, having known all variables O_i^k, D_d^k, p_{id}^{lk}, C_{id}^k, the only problem now is how to estimate the unknown parameter so that the model reproduces the estimated traffic (passenger) counts as close as possible to the observed ones. The main idea of Non-Linear-Least-Squares Estimation Method (NLLS) is to estimate the unknown parameter which minimizes the sum of the squared differences between the estimated and observed traffic counts. The problem now is:

$$\text{to minimize } \boldsymbol{S} = \sum_{l}\left[V_l^{+k} - V_l^k\right]^2 \tag{10}$$

Having substituted (9) to (10), the following set of equation is required in order to find a set of unknown parameter $\boldsymbol{\beta}$ which minimizes equation (10):

$$\frac{\delta S}{\delta \beta} = \sum_{l}\left[\left(2\sum_{i}\sum_{d} T_{id}^{k}.p_{id}^{lk} - V_l^k\right)\left(\frac{\sum_{i}\sum_{d}\delta T_{id}^{k}}{\delta\beta . p_{id}^{lk}}\right)\right] = 0 \tag{11}$$

Equation (11) 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 $L \geq 1$. Newton's method is then be used to solve equation (11). This method can also be used to solve for sets of parameters β_m for M types of modes of transportation from simultaneous equations built for that purpose. Let β_0 be an approximate solution and suppose β_0+h is an exact solution, then $f(\beta_0+h)=0$, and to the first order in h:

$$f(\beta_0) + h\ \delta f/\delta\beta|_{\beta 0} = 0 \tag{12}$$

Equation (12) is a linear equation for h and can be solved easily to obtain value h_0 of h, then β_0+h_0 is a better solution. This process is repeated until h_0 is very small, indicating convergence.

4. METHODOLOGY

The methodology of this research can be seen on figure 2.

5. APPLICATION MODEL

The model has been tested in artificial network consisting of four zones and 42 links representing the road network. Therefore, there will be (5x5=25) number of p_{id}^{lk} for each link and (42x25=1050) number of p_{id}^{lk} for total number of available links within the study area. There are two mode, bus and car. The data as input for estimation process are: *Oi, Dd*, and network system.

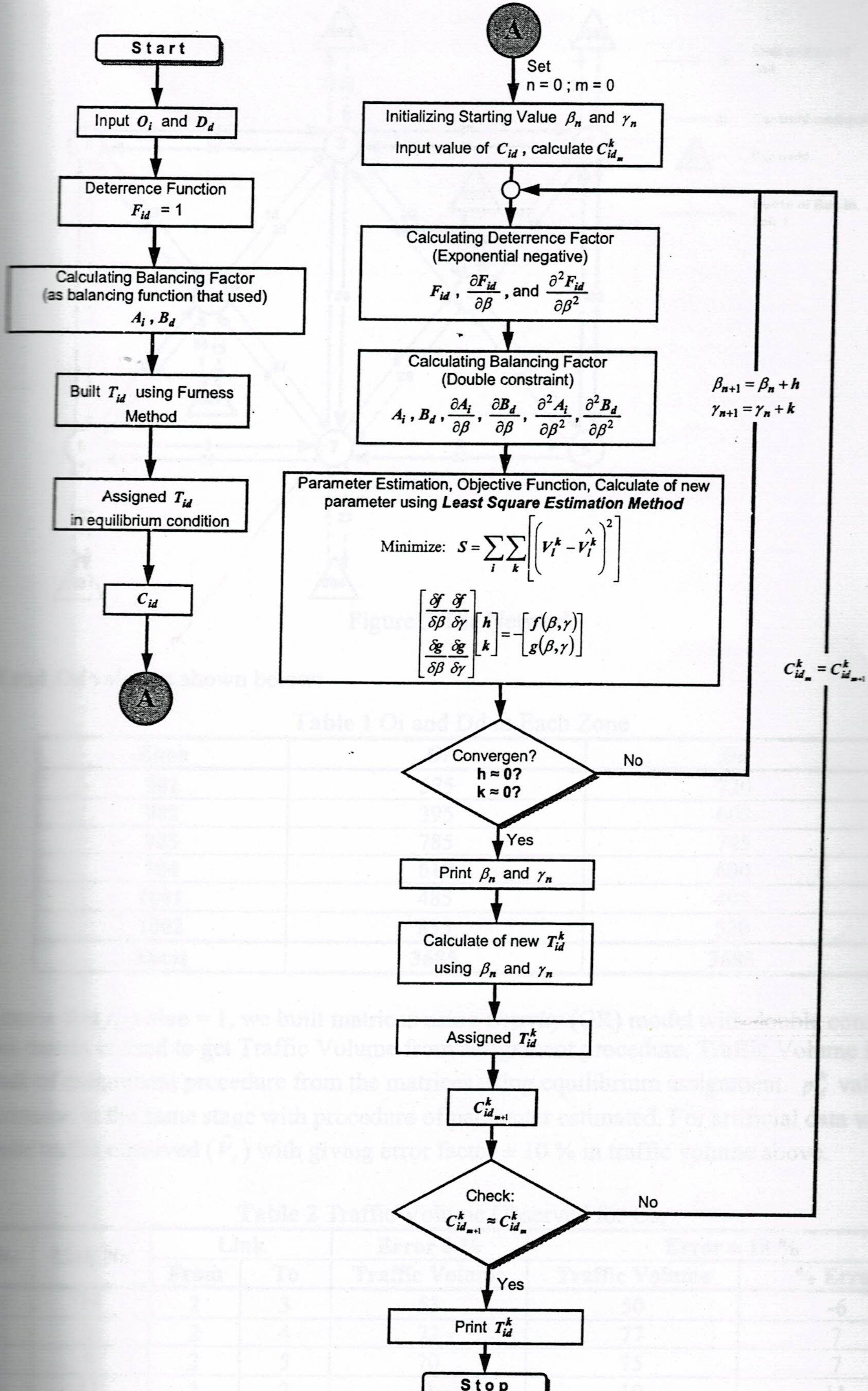

Figure 2 Research Methodology

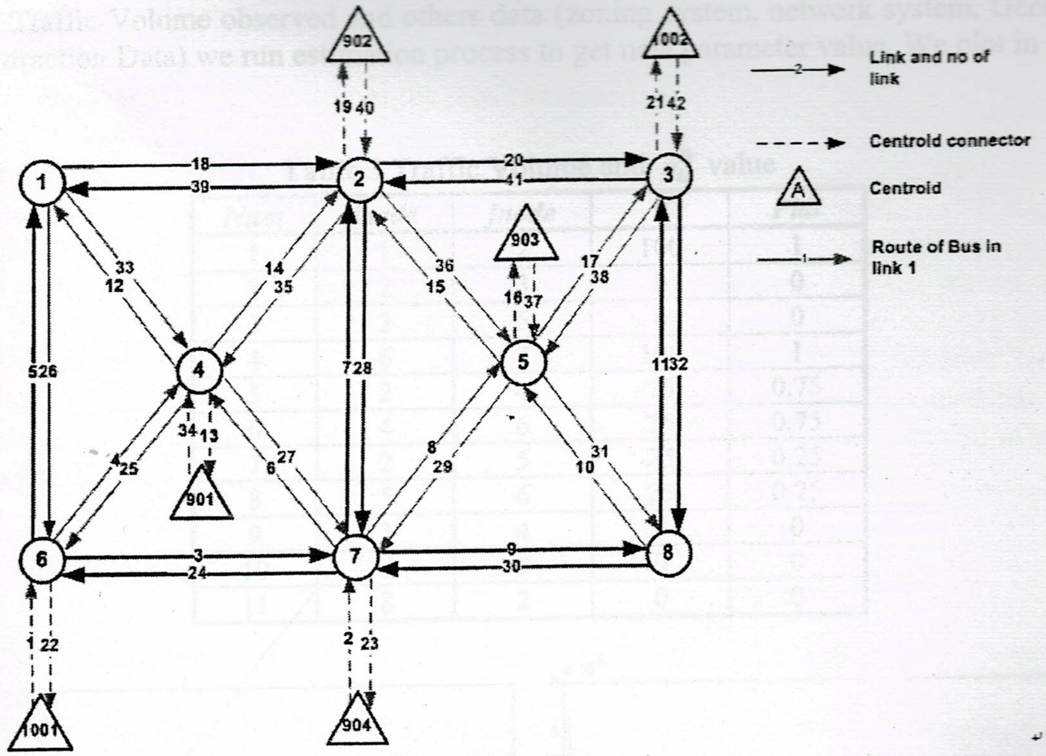

Figure 3 The Network

Oi and *Dd* value as shown below:

Table 1 Oi and Dd in Each Zone

Zona	*Oi*	*Dd*
901	575	720
902	395	605
903	785	745
904	610	600
1001	485	495
1002	835	520
Total	**3685**	**3685**

Assume that f_{id} value = 1, we built matrices using *Gravity* (GR) model with double constraint. This matrix is used to get Traffic Volume from assignment procedure. Traffic Volume is the result of assignment procedure from the matrices using equilibrium assignment. p_{id}^{lk} value is determine in the same stage with procedure of parameter estimated. For artificial data we create traffic observed ($\hat{V}_l$) with giving error factor ± 10 % in traffic volume above.

Table 2 Traffic Volume Observed for Car

No	Link No	Link		Error 0 %	Error ± 10 %	
		From	To	Traffic Volume	Traffic Volume	% Error
1	13	2	3	53	50	-6
2	2	2	4	72	77	7
3	9	2	5	70	75	7
4	11	2	7	9	10	11
5	14	3	2	110	103	-6
6	20	3	5	325	318	-2
7	1	4	2	3	3	0

Using Traffic Volume observed and others data (zoning system, network system, Generation and Attraction Data) we run estimation process to get new parameter value. We plot in graphs like:

Table 3 Traffic Volume and p_{id}^{lk} value

Num	*inode*	*jnode*	*V*	*Pidl*
1	1	2	100	1
2	2	3	0	0
3	3	6	0	0
4	6	7	100	1
5	2	4	75	0.75
6	4	6	75	0.75
7	2	5	25	0.25
8	5	6	25	0.25
9	3	4	0	0
10	8	3	0	0
11	8	2	0	0

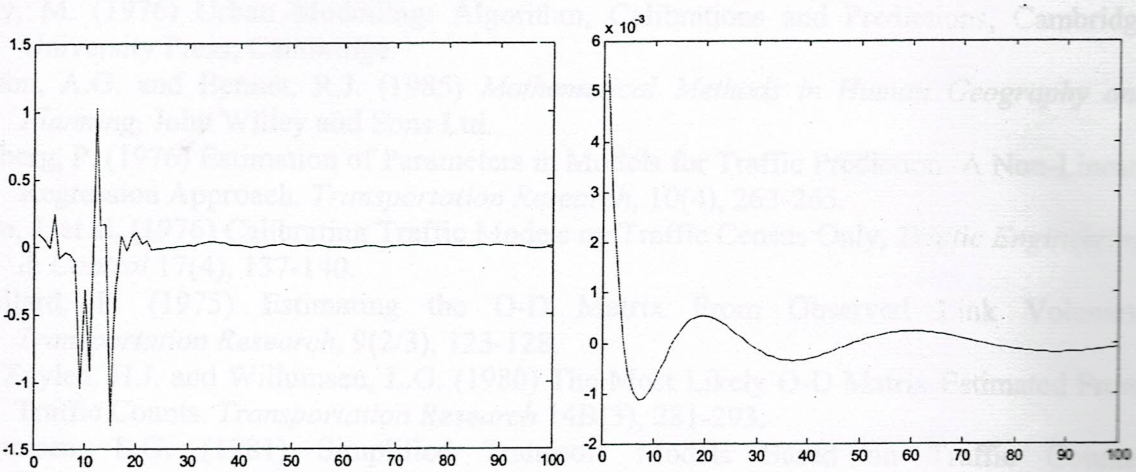

Figure 4 Convergence of *h* and *k* value

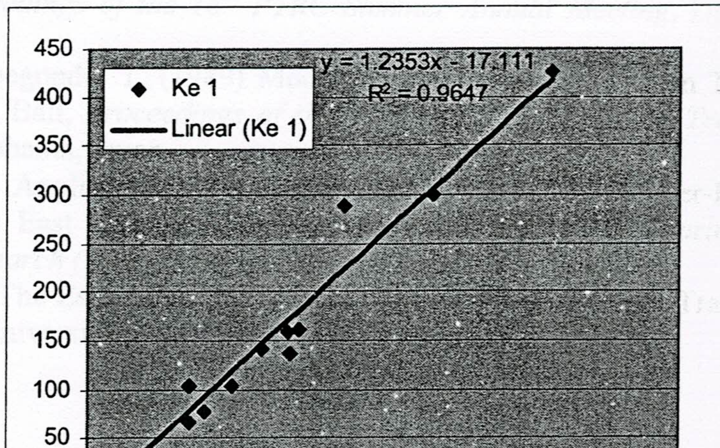

Figure 5 Statistical value of *V* observed vs *V* model

6. CONCLUSION

Some conclusions can be drawn from the result obtained:

- The number of observed 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.
- It is found that by having the information of passenger flows using Angkot, we can obtain the O-D matrices for private and Angkot.
- It is also shown that the TDMC model with negative exponential deterrence function produced estimation parameter (β and γ) for NLLS and ML estimation methods. This result is very important in terms of time and money for estimating the demand of public transport and also for forecasting purposes.

REFERENCES

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

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

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 & Control* 17(4), 137-140.

Robillard, P. (1975) Estimating the O-D Matrix From Observed Link Volumes, *Transportation Research*, 9(2/3), 123-128.

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

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. (1992) Application of Transport Demand Models for Inter-Regional Vehicle Movements in East Java (Indonesia), *Proceedings of the 6th World Conference on Transport Research (WCTR)*, Lyon, France.

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