# THE IMPACT OF LOCATION AND NUMBER OF TRAFFIC COUNTS IN THE ACCURACY OF O-D MATRICES ESTIMATED FROM TRAFFIC COUNTS: A CASE STUDY IN BANDUNG (INDONESIA)

**Ofyar Z. TAMIN**
Professor
Department of Civil Engineering
Institute of Technology Bandung
Jalan Ganesha 10
Bandung 40132, Indonesia
Fax: +62-22-2512395
E-mail: ofyar@trans.si.itb.ac.id

**Titi L. SOEDIRDJO**
Senior Lecturer
Department of Civil Engineering
Institute of Technology Bandung
Jalan Ganesha 10
Bandung 40132, Indonesia
Fax: +62-22-2512395
E-mail: office@trans.si.itb.ac.id

**Rudi S. SUYONO**
Lecturer
Department of Civil Engineering
University of Tanjung Pura
Pontianak, Indonesia
E-mail: office@trans.si.itb.ac.id

**Abstract**: To estimate an Origin-Destination (OD) matrix from traffic counts, traffic count is the major input which obviously affects the accuracy of the estimated matrix. Any process regarding the traffic counts have to be carefully studied. Theoretically, the more traffic counts we have, the better will be the estimated OD matrix. Therefore, the objective is to obtain the best locations and the optimum number of traffic counts. The model considered 3 (three) major factors i.e. (a) proportion factor of the trip interchanges for each link, (b) independence and inconsistency conditions, and (c) physical link conditions. The model has been tested in Bandung consisting of 145 zones and 2485 links (arterial, collector, and local roads). The first stage has been able to select 969 links out of 2485 links (1516 links unselected) and 646 links of them has been reselected in the second stage (another 323 unselected). The best links obtained in the second stage is ranked using the link condition criteria and finally obtain the best locations together with their ranks. The study has also found that the optimum number of traffic count data is 90 links (around 3.6% of total links).

**Key Words:** OD matrix, traffic counts, location, impact, optimisation

## 1. INTRODUCTION

### 1.1 Background

Many problems in transport planning and management tasks require an **Origin-Destination (OD) matrix** as a fundamental input to represent the travel pattern. An OD matrix gives a very good indication of travel demand, and therefore, it plays a very important role in various types of transport studies, transport planning, and managemen tasks. The **conventional method** to estimate the OD matrices require very large surveys such as: **roadside** or **home interviews**; which are very expensive involving considerable resources in terms of manpower, time, subject to large errors, and time disruptive to trip makers. Moreover, the realibility of the resulting OD matrices seems to have relatively short life.

The need for inexpensive methods, which require low-cost data, less time and less manpower generally called as **unconventional method**, seems particularly attractive especially for developing countries due to **time and money constraints**. This become even more valuable for problems which require **quick-response** treatment such as urban transport problems. Therefore, unconventional method for estimating OD matrices is an effective and efficient method since the main input required is the traffic count information which is easily available and relatively inexpensive to obtain.

There are several reasons why traffic counts are so attractive as a data base to estimate the OD matrices (**Tamin, 2000**):

a. **Low-cost,** this type of data are relatively inexpensive to obtain since they require less manpower, less time, easier in terms of organization and management, and moreover, automatic traffic counters have been well developed.
b. **Availability,** traffic counts are routinely colloected by many authorities due to their multiple uses in many transport planning tasks; all of these make them easily available.
c. **Non-disruptive,** traffic counts can be obtained without generating any delay or disruption to vehicles and trip makers.

### 1.2 The Analysis Process

It is mentioned that the unconventional method uses traffic count information as the main input for estimating the OD matrices. Because of that, any process regarding the traffic counts should be clearly and deeply understood in order to obtain the best estimates of OD matrices; especially those which are related to data collection process e.g. number of traffic counts and their best locations. The data collection process is very important since it is the first action in the whole process of OD matrix estimation.

#### 1.2.1 The best location

Some basic analysis used in finding the best location are as follows:

**a. Proportion of trip interchanges on a particular link**

The total volume of flow in a particular link **l** ($\hat{V}_l$) is the summation of the contributions of all trip interchanges between zones within the study area to that link. Mathematically, it can be expressed as follows:

$$\mathbf{V_l = \sum_i \sum_d T_{id} . p^l_{id}} \qquad (1)$$

**Tamin (2000)** stated that the most important stage for the estimation of OD matrix from traffic counts is to identify the paths followed by the trips from each origin **i** to each destination **d**. In other words, the proportion of trip interchanges between zone **i** and zone **d** have to be uniquely identified for all those links involved. In this case, the variable $\mathbf{p^l_{id}}$ is used to define the proportion of trip interchanges from origin **i** to destination **d** travelling through link **l**. Thus, in other words, the flow on each link of a network is a result of:

- trip interchanges from origin **i** to destination **d** or combination of different type of interchanges, travelling between zones within the study area ($\mathbf{T_{id}}$);
- the proportion of trips from origin **i** to destination **d** whose trips use link **l** defined by $\mathbf{p^l_{id}}$ ($\mathbf{0 \le p^l_{id} \le 1}$).

Therefore, in finding the best location, the traffic counts having many information of the trip interchanges should be chosen. This information can be identified by analysing the total number and value of $\mathbf{p}^{l}_{id}$ in each link. This information will then be taken as the main criteria in determining choosing the best location of traffic counts.

**b. Inter-link relationship**

- **Inter-dependence** **Figure 1** shows that flows on link **5–6** are the summation of flows on link **1–5** and on link **2–5**, then there is no additional information can be extracted by counting flows on link **5–6** because of the **flow continuity condition**, $\bar{V}_{56} = \bar{V}_{15} + \bar{V}_{25}$.

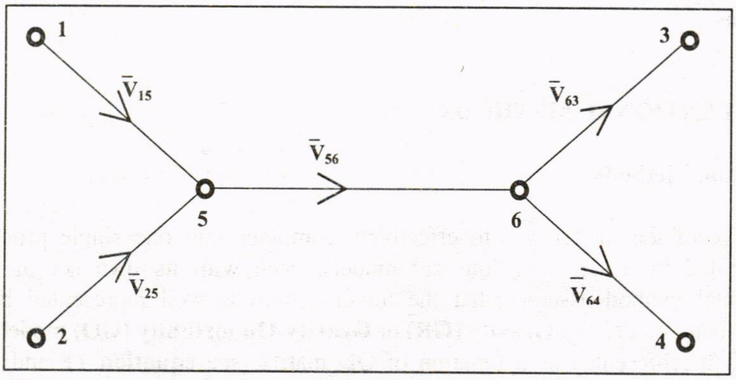


**Figure 1**: A simple network with link counts (Source: **Tamin, 2000**)

In principle, we need counts for only **3 (three) independent counts** in order to find the flows of all links in **figure 1**. Therefore, from an economic point of view, some efforts are needed in choosing the appropriate links to be counted (**Tamin, 2000**).

- **Inconsistency** In practice, the problem of inconsistency in link counts may arise when the flow continuity conditions are not satisfied by the observed volumes. In the case of **figure 1**, it may well happen that the observed flows are such that:

$$\bar{V}_{56} \neq \bar{V}_{15} + \bar{V}_{25} \tag{2}$$

or

$$\bar{V}_{15} + \bar{V}_{25} \neq \bar{V}_{63} + \bar{V}_{64} \tag{3}$$

This inconsistency in counts may arise due to human or counting errors and counting at different times or dates. As a result of all this, no solution for the OD matrix can be estimated that reproduces all these inconsistent traffic counts. One possible way to remove this problem is by choosing only independent links for counted (**Tamin, 2000**).

### 1.2.2 Optimum number of traffic counts

**Equation (1)** is the fundamental equation developed for estimating the OD matrices from traffic counts information. In this model, the parameters $\mathbf{p}^{l}_{id}$ are estimated using traffic assgnment technique. Given all the $\mathbf{p}^{l}_{id}$ and all the observed traffic counts ($\hat{V}_l$), then there will be $\mathbf{N}^2$ buah $\mathbf{T}_{id}$'s to be estimated from a set of **L** simultaneous linear equations **(1)** where **L** is the total number of traffic counts. **In principle,** $\mathbf{N}^2$ independent and consistent traffic counts are required in order to determine uniquely the OD matrix $[T_{id}]$; $[N^2-N]$ if intrazonal 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 **(Tamin, 1988).**

One possible way to overcome this problem is to restrict the number of possible solutions by modelling the trip making behaviour such as using: **Gravity (GR)** or **Gravity-Opportunity (GO)** model **(Tamin, 1988)**. In this case, hypothetically, it can be said that the more traffic count we use in estimating the OD matrices, the higher the accuracy of the estimated OD matrices. However, it requires more cost and more processing time. Therefore, the objective of this research is to study the effect of number of traffic count used on the accuracy of the estimated OD matrices as well as to obtain the optimum number of traffic counts and their best locations.

## 2. UNCONVENTIONAL METHODS

### 2.1 Estimation Methods

The objective of the model was to effectively combines into one single process what is usually handled in a series of four sub-models, each with its own set of errors. The unconventional method assumes that the travel pattern is well represented by a certain transport demand model e.g. **Gravity (GR)** or **Gravity-Opportunity (GO)** model. The traffic flows are then represented as a function of OD matrix (see **equation 1**), and; hence as a function of a transport demand model together with its parameters. **Nguyen (1982)** and **Tamin (1988)** provide a very good overview on **the state-of-the-art** in this research domain related to the OD matrix estimation from traffic counts.

Consider that there are trips travelling between zones and it is assumed that the trip interchanges can be represented using a transport demand model. Hence, the total number of trips $T_{id}$ with origin in **i** and destination in **d** can be expressed as:

$$T_{id} = O_i . D_d . A_i . B_d . f_{id} \quad \text{where:} \tag{4}$$

$A_i$ and $B_d$ = the **balancing factors** expressed as:

$$A_i = \frac{1}{\sum_d (B_d . D_d . f_{id})} \quad \text{and} \quad B_d = \frac{1}{\sum_i (A_i . O_i . f_{id})} \tag{5}$$

$f_{id}^k$ = the **deterrence function** (negative exponential $\exp(-\beta . C_{id}^k)$)

By subtituting **equation (4)** to **equation (1)**, then **the fundamental equation** for the estimation of a transport demand model from traffic counts can be expressed as:

$$V_l = \sum_i \sum_d O_i . D_d . A_i . B_d . f_{id} . (p_{id}^l) \tag{6}$$

This fundamental equation is very important and usually used in many literatures either to estimate the OD matrices or to calibrate the transport demand model from traffic counts **(Tamin, 1988).**

## 2.2 Calibration Methods

In estimating the OD matrices from traffic counts, model calibration is a key element in solving the problem. The objective is to calibrate the transport demand model from traffic counts to obtain the calibrated parameters which will be used to reproduced the OD matrix. In this case, the number of traffic counts should at least as many as the number of unknown parameters of the proposed transport model.

In practice it seems impossible, if ever, to obtain the traffic count information totally free from errors. There will always be some errors in the traffic counts themselves, in the network definition or zoning representation, or, perhaps, in the assignment techniques used in determining the routes taken by the drivers. Therefore, more traffic counts are needed in order to compensate for these types of errors.

**Tamin (1988)** has developed **2 (two)** main groups of estimation methods that can be used to calibrate the parameters of proposed transport demand model from traffic counts. They are: **Least-Squares (LS)** and **Maximum-Likelihood (ML)** estimation methods. The main ideas behind these estimation methods is that we try to calibrate the unknown parameters of the postulated model so that to minimize the deviations or differences between the traffic flows estimated by the calibrated model and the observed flows. This can be carried out by using the likelihood measure between the estimated and observed flows, e.g. **maximum-likelihood** or **least-squares**.

# 3. METHODOLOGY

## 3.1 The rank analysis for the best location of traffic counts

The rank analysis to obtain the best location of traffic counts will be grouped into **3 (three)** stages of selection. Each stage of analysis will determine the best locations of traffic counts based on its corresponding criteria. Model **Gravity-Opportunity (GO)** will be used to represent the transport demand model. They are as follows:

**a. 1st stage selection: based on proportion of trips on a links ($p_{id}^l$)**

The main parameter used in the **1st stage** of selection is the value of trip proportions for each link ($p_{id}^l$) within the study area. The value of $p_{id}^l$ for each link is actually the proportion of trips travelling from zone <u>i</u> to zone <u>d</u> whose trips use link <u>l</u> as part of their best routes. Hence, the value of $p_{id}^l$ will be within the range of **0–1**. The larger the value of $p_{id}^l$ represents the larger proportion of trips from zone <u>i</u> to zone <u>d</u> which use link **l**.

The total volume in a particular link <u>l</u> is the summation of the contributions of all trips interchanges travelling between zones within the study area. As in **equation (1)**, each link may have several different values of $p_{id}^l$ which depends on how attractive the links chosen by the drivers as part of their best route. The total number of $p_{id}^l$ for each link depends on the total number of zones within the study area. Therefore, if there are **N** zones, there will be $N^2$ values of $p_{id}^l$.

In the **1st stage**, the selection of traffic count location will use the $p_{id}^l$ as the main parameter which can represent the characteristics of trip interchanges for each link. Hence, to obtain the relationship between the trip proportion and total trip interchanges, the following weighted mean parameter will be used:

$$\hat{\mu} = \frac{k.\sum_{id} T_{id}.P_{id}^{l}}{N^2.\sum_{id} T_{id}} \quad (7)$$

where:

$\hat{\mu}$ = weighted mean  k = total number of $p_{id}^{l}>0$ for each link

N = number of zones within the study area

In this stage, the links which have the values of $\hat{\mu}=0$, or in other words, links which were not chosen by the drivers, will be dropped out in this stage. Whilst, the other links will be ranked depending on the values of its weighted mean. All ranked links will be further used in the next stage of selection.

**b. 2nd stage selection: based on inter-link relationship**

In the 2nd **stage**, the link selection will be carried out based on the inter-link relationship. In this stage, the selection process will consider the **inter-dependence** and **inconsistency** conditions of links. This 2nd stage of selection process will be carried out for all links which have passed the 1st stage.

Based on the **inter-dependence principle**, the total flow of a particular link is a summation of flows from all links connecting to that particular link because of the flow continuity condition. In practice, the problem of inconsistency in link counts may arise when the observed volumes do not satisfy the flow continuity conditions. This inconsistency in counts may arise due to human or counting errors and counting at different times or dates. As a result of all this, no solution for the OD matrix can be estimated that reproduces all these inconsistent traffic counts.

One possible way to remove the problem is by choosing only independent links for counted. However, based on the inter-dependence principle and from an economic point of view, some efforts are needed in choosing the appropriate links to be counted (see explanation in **figure** 1). Therefore, these independent links have to be re-evaluated again and only the appropriate independent links will be chosen. Hence, all selected links will be further used in the final stage of selection.

**c. 3rd stage selection: based on link condition**

This stage is the final process in selecting the best location of traffic counts for estimating the OD matrices. Some criteria related to the physical condition of links will be considered in this stage. They are as follows:

- **Link rank criteria based on 1st and 2nd stage of selection process** The links which have passed 1st and 2nd stage will be assumed as the best links for counted. The criteria will be used since the selected links have the high value of $\hat{\mu}$; meaning that the links have many trip interchanges information and also a part of their best routes. Moreever, the selected links have considered the inconsitency condition and fulfilled the inter-dependence principle.The weighting process in this criteria will be based on the rank list of all links which have passed the 1st and 2nd stage of selection.

- **Congestion condition criteria (degree of saturation)** This criteria has an assumption that the higher the level of congestion in a particular link, the lower the probability of the drivers to choose that particular link to be part of their best route to reach their destination. In this case, the level of congestion will be determined by the value of **Degree of Saturation (DS)** parameter. Therefore, the weighting process in this criteria will be based on the value of DS parameter of all links which have passed the 1st and 2nd stage of selection.

- **Side friction condition criteria** This criteria has an assumption that the higher the side friction in a particular link, the lower the probability of the drivers to choose that particular link to be part of their best routes to reach their destinations. Therefore, the weighting process in this criteria will be based on the **Side Friction (SF)** parameter of all links which have passed the 1st and 2nd stage of selection. The **SF** parameter will determined by **IHCM (1997)**.

This stage considers the influence of condition of a particular link to the rank. The higher the influence, the lower the rank of that link. A simple **multi-criteria** analysis will be used in the selection process of the 3rd stage. Firstly, the **Quantification of Criteria ($Q_C$)** process have to be determined based on each criteria as given in **table 1**.

**Table 1**: Quantification of criteria for analysing the ranks of traffic count location

| No | Criteria I | | Criteria II | | Criteria III | |
|---|---|---|---|---|---|---|
| | Rank lists (The best of) | $Q_C$ | DS | $Q_C$ | Class of SF | $Q_C$ |
| 1. | 0% – 10% | 1 | 0.0 – 0.1 | 1 | Very Low | 1 |
| 2. | 10% – 20% | 2 | 0.1 – 0.2 | 2 | Low | 2 |
| 3. | 20% – 30% | 3 | 0.2 – 0.3 | 3 | Medium | 3 |
| 4. | 30% – 40% | 4 | 0.3 – 0.4 | 4 | High | 4 |
| 5. | 40% – 50% | 5 | 0.4 – 0.5 | 5 | Very High | 5 |
| 6. | 50% – 60% | 6 | 0.5 – 0.6 | 6 | - | - |
| 7. | 60% – 70% | 7 | 0.6 – 0.7 | 7 | - | - |
| 8. | 70% – 80% | 8 | 0.7 – 0.8 | 8 | - | - |
| 9. | 80% – 90% | 9 | 0.8 – 0.9 | 9 | - | - |
| 10. | 90% –100% | 10 | 0.9 – 0.10 | 10 | - | - |

**Note:**
1. Criteria I : based on selection stage I and II
2. Criteria II : Degree of Saturation (DS)
3. Criteria III : Side Friction (SF)

Some scenarios will be adopted in the ranking process since the weighting of each criteria will be assumed to be different in different conditions. **Table 2** shows the values of the weightings ($W_C$) of each criteria for each scenario.

**Table 2**: Weighting for each scenario

| No | Criteria | Weighting ($W_C$) | | |
|---|---|---|---|---|
| | | Scenario I | Scenario II | Scenario III |
| 1. | Criteria I | 1 | 3 | 5 |
| 2. | Criteria II | 3 | 5 | 1 |
| 3. | Criteria III | 5 | 1 | 3 |

Therefore, in order to obtain the optimal result, different weightings (**$W_C$**) to each criteria will be determined based on each scenario. The values of **$W_C$** is taken within the range of **1–5**. Finally, the link ranks will be determined based on the value of $W_C$. The smallest value of $W_C$ will be the highest rank and the largest value will be the lowest rank.

## 3.2 The Determination of Optimum Number of Traffic Count

The optimum number of traffic count can be determined by analysing the relationship between number of traffic counts and the level of accuracy of the estimated OD matrix based on that counts. The level of accuracy of the estimated matrices will be determined by comparing those matrices with the initial matrix (matrix estimated by using all selected traffic counts). This process will be conducted in **2 (two)** conditions representing the sensitivity between the number of traffic counts and the link rank, namely: **sorted condition** and **random condition**.

### a. Condition I (sorted)

In condition I, several combination of traffic counts will be created based on the rank list of selected links which have passed the selection process. Each combination of traffic counts will then be used to estimate the OD matrices. The combination of traffic counts created varies from **2**, **4**, …, **646** links (total number of selected links). The **$R^2$ statistical test** will be used to compare the estimated matrices with the initial one. By analysing the relationship between the number of traffic counts and the level of accuracy of the estimated matrices, the optimum number of traffic count could be determined. The $R^2$ statistical test is calculated using equation (**8**).

$$\mathbf{R}^2 = \frac{\sum_i \left(\hat{Y}_i - \overline{Y}_i\right)^2}{\sum_i \left(Y_i - \overline{Y}_i\right)^2} \qquad \textbf{(8)}$$

where: $Y_i$ and $\hat{Y}_i$ = estimated and observed values

$\overline{Y}$ = average value of $\hat{Y}_i$

### b. Condition II (random)

Similar with condition I, several combination of traffic counts will be created not based on the rank list of selected links but based on random selection. Each combination of traffic counts will then be used to estimate the OD matrices. The combination of traffic counts created randomly varies from **2**, **4**, …, **646** links. The $R^2$ statistical test will be used to compare the estimated matrices with the initial one. By analysing the relationship between the number of traffic counts and the level of accuracy of the estimated matrices, the optimum number of traffic count could be determined.

## 4. STUDY AREA

The study area comprises the Kotamadya Bandung area and some parts of Kabupaten Bandung area surrounding the Kotamadya Bandung. The zone boundaries were **administrative boundaries** of **kelurahans**, **kecamatans**, and **sub-districts** (wilayah). The study area comprises of **139 zones**, **100 zones** are in Kotamadya Bandung and **39 zones** are in the Kabupaten Bandung, and another **6 external zones** representing the other area outside the study area.

The road network system in Indonesia can be grouped according to their hierarchy of the function, status, and class. Hierarchy of the function can divide the road as arterial, collector, and local roads. By status, the road can be divided into **national, provincial**, and **kabupaten** roads. By class, the road can be divided into class I, II, and III. In this paper, the road network is classified by road function such as: **arterial, collector**, and **local** roads and consists of 2485 links representing the modelled road network.

## 5. THE RESULT OF ANALYSIS

### 5.1 The Determination of The Best Location of Traffic Counts

#### a. 1st stage selection: based on proportion of trips on a links ($p_{id}^{l}$)

As mentioned above that in the 1st stage of selection process, the value of $\mathbf{p_{id}^{l}}$ will be determined by using trip assignment technique. In this research, equilibrium assignment will be used to consider the congestion effect in route choice selection. The study are has been divided into **145 zones** and consists of **2495 links** representing the road network. Therefore, there will be **(145x145=21025)** number of $\mathbf{p_{id}^{l}}$ for each link and **(2485x21025)** number of $\mathbf{p_{id}^{l}}$ for total number of available links within the study area.

**Table 3**: Selection result of the 1st stage (only 20 links having the highest ranks are shown)

| Ranks | Link number | Node From | Node To | Flows pcu/hour | Total $P_{id}^{l}$ >0.0 | Weighted mean | Street name |
|---|---|---|---|---|---|---|---|
| 1. | 331 | 1044 | 601 | 1131 | 3921 | 0.0039 | Otto Iskandar D. |
| 2. | 311 | 1037 | 660 | 1049 | 3946 | 0.0036 | Merdeka |
| 3. | 1324 | 201 | 102 | 855 | 3238 | 0.0024 | Elang |
| 4. | 313 | 1038 | 21504 | 561 | 3927 | 0.0019 | Lembong |
| 5. | 1458 | 432 | 604 | 1076 | 1884 | 0.0018 | Otto Iskandar D. |
| 6. | 891 | 660 | 65802 | 811 | 2275 | 0.0016 | Perintis |
| 7. | 1967 | 695 | 404 | 904 | 2002 | 0.0016 | Cipaganti |
| 8. | 1100 | 437 | 44001 | 646 | 2666 | 0.0015 | Otto Iskandar D. |
| 9. | 927 | 207 | 1044 | 1029 | 1661 | 0.0015 | Otto Iskandar D. |
| 10. | 889 | 42603 | 1037 | 938 | 1798 | 0.0015 | Merdeka |
| 11. | 2142 | 657 | 42502 | 551 | 3019 | 0.0014 | Wastu Kencana |
| 12. | 1448 | 210 | 207 | 985 | 1661 | 0.0014 | Suniaraja |
| 13. | 256 | 1019 | 405 | 841 | 1922 | 0.0014 | Dr. Setia Budi |
| 14. | 342 | 1048 | 32304 | 679 | 2338 | 0.0014 | Sukarno Hatta |
| 15. | 939 | 101 | 201 | 304 | 5128 | 0.0014 | Rajawali Barat |
| 16. | 1457 | 20803 | 432 | 918 | 1640 | 0.0013 | Otto Iskandar D. |
| 17. | 1545 | 32302 | 220 | 491 | 3044 | 0.0013 | Sukarno Hatta |
| 18. | 1024 | 326 | 221 | 517 | 2826 | 0.0013 | Sukarno Hatta |
| 19. | 1178 | 633 | 1089 | 1106 | 1285 | 0.0012 | Kiara Condong |
| 20. | 469 | 1087 | 633 | 1087 | 1304 | 0.0012 | Kiara Condong |

The value of weighted mean ($\hat{\mu}$) for each link will be calculated based on the values of $p_{id}^{l}$ and the trip interchanges ($T_{id}$), see **equation (7)**. All links which having the value of **$\hat{\mu}$=0** will be dropped out automatically in this 1st stage of selection process. In the 1st stage, **969** links have passed the process out of **2485** links (**1516** links dropped out). All links which have passed the 1st stage will be further analysed in the next stage by analysing their inter-link relationship. Some of the selected links in the 1st are shown in **table 3** (only.20 links having the highest rank are shown).

**b. 2nd stage selection: based on inter-link relationship**

In the **2nd stage** of selection process, the links are selected based on the inter-link relationship so that the selected link will consider the **inter-dependence** and **inconsistency** conditions. The links which passed the 1st stage will be reevaluated based on their inter-link relationship with other selected links. This process can be carried out manually (by analysing the road netowrk map) or by using a computer algorithm. In this stage, **969** links which have passed the 1st stage will be reevaluated; **646** links are reselected out of **969** links (another **323** links dropped out). Some of the selected links in the 2nd are shown in **table 4** (only 20 links having the highest rank are shown).

**Table 4**: Selection results of the 2nd stage (only 20 links having the highest ranks are shown)

| No | Link number | Node | | Street name | Ranks after selection stage II | Flows pcu/hour |
|---|---|---|---|---|---|---|
| | | From | To | | | |
| 1 | 331 | 1044 | 601 | Otto Iskandar D. | 1 | 1131 |
| 2 | 311 | 1037 | 660 | Merdeka | 2 | 1049 |
| 3 | 1324 | 201 | 102 | Elang | 3 | 855 |
| 4 | 313 | 1038 | 21504 | Lembong | 4 | 561 |
| 5 | 1458 | 432 | 604 | Otto Iskandar D. | 5 | 1076 |
| 6 | 891 | 660 | 65802 | Perintis | 6 | 811 |
| 7 | 1967 | 695 | 404 | Cipaganti | 7 | 904 |
| 8 | 1100 | 437 | 44001 | Otto Iskandar D. | 8 | 646 |
| 9 | 927 | 207 | 1044 | Otto Iskandar D. | 9 | 1029 |
| 10 | 889 | 42603 | 1037 | Merdeka | 10 | 938 |
| 11 | 2142 | 657 | 42502 | Wastu Kencana | 11 | 551 |
| 12 | 1448 | 210 | 207 | Suniaraja | 12 | 985 |
| 13 | 256 | 1019 | 405 | Dr. Setia Budi | 13 | 841 |
| 14 | 342 | 1048 | 32304 | Sukarno Hatta | 14 | 679 |
| 15 | 939 | 101 | 201 | Rajawali Barat | 15 | 304 |
| 16 | 1457 | 20803 | 432 | Otto Iskandar D. | 16 | 918 |
| 17 | 1545 | 32302 | 220 | Sukarno Hatta | 17 | 491 |
| 18 | 1024 | 326 | 221 | Sukarno Hatta | 18 | 517 |
| 19 | 1178 | 633 | 1089 | Kiara Condong | 19 | 1106 |
| 20 | 469 | 1087 | 633 | Kiara Condong | 20 | 1087 |

**c. 3rd stage selection**

The **3rd stage** will the final stage of the whole process in finding the best location for the links for counted. This final stage will consider **3 (three)** criteria which are related to the physical condition of each link. 646 links which have passed the 2nd stage will be re-ranked in this stage to obtain the final rank of the selected link for counted. Some of the selected links in the 3rd are shown in **table 5** (only 20 links having the highest rank are shown).

**Table 5**: Selection result for stage III (only 20 links having the highest ranks are shown)

| No. | Link no. | Node | | Street name | Flows pcu/hr | Rank in stage I | Rank in stage II | Selection stage III | |
|---|---|---|---|---|---|---|---|---|---|
| | | From | To | | | | | Total $W_C$ | Ranks in stage III |
| 1 | 2142 | 657 | 42502 | Wastu Kencana | 551 | 11 | 11 | 36 | 1 |
| 2 | 939 | 101 | 201 | Rajawali Barat | 304 | 15 | 15 | 36 | 2 |
| 3 | 2146 | 65801 | 42403 | Wastu Kencana | 292 | 51 | 50 | 36 | 3 |
| 4 | 2138 | 42401 | 657 | Wastu Kencana | 276 | 62 | 60 | 36 | 4 |
| 5 | 313 | 1038 | 21504 | Lembong | 561 | 4 | 4 | 45 | 5 |
| 6 | 891 | 660 | 65802 | Perintis | 811 | 6 | 6 | 45 | 6 |
| 7 | 233 | 1304 | 302 | Cimindi | 260 | 37 | 37 | 45 | 7 |
| 8 | 309 | 1036 | 42601 | Merdeka | 406 | 49 | 48 | 45 | 8 |
| 9 | 1959 | 403 | 402 | Sukajadi | 638 | 50 | 49 | 45 | 9 |
| 10 | 268 | 1023 | 409 | Ir. H. Juanda | 629 | 52 | 51 | 45 | 10 |
| 11 | 2088 | 681 | 7251 | Taman Sari | 463 | 60 | 58 | 45 | 11 |
| 12 | 2473 | 93502 | 936 | Sumatera | 314 | 65 | 63 | 45 | 12 |
| 13 | 1349 | 21602 | 104 | Veteran | 360 | 94 | 85 | 45 | 13 |
| 14 | 2077 | 42501 | 417 | Wastu Kencana | 351 | 104 | 92 | 45 | 14 |
| 15 | 2430 | 69201 | 69303 | Cipaganti | 349 | 111 | 98 | 45 | 15 |
| 16 | 1500 | 21501 | 93503 | Sumatera | 241 | 139 | 121 | 45 | 16 |
| 17 | 311 | 1037 | 660 | Merdeka | 1049 | 2 | 2 | 54 | 17 |
| 18 | 1324 | 201 | 102 | Elang | 855 | 3 | 3 | 54 | 18 |
| 19 | 1967 | 695 | 404 | Cipaganti | 904 | 7 | 7 | 54 | 19 |
| 20 | 889 | 42603 | 1037 | Merdeka | 938 | 10 | 10 | 54 | 20 |

It can be seen from **table 5**, that the rank list of the selected link resulted from the 2nd stage is reevaluated again in the final stage based on the link conditions. For example, link 2142 which was previously at rank 11 in the 1st stage and the 2nd stage has moved at rank 1 after the 3rd stage. Similar to link 2146, it was previously at rank 51 in the 1st stage, moved to rank 50 at the 2nd stage, finally moved to rank 3 at the final stage. However, different effect happened to link 311. Link 311 was previously at rank 2 in the 1st stage and the 2nd stage has moved at rank 17 after the 3rd stage. It can be concluded here that the road physical condition have a great influence in determining the rank of the selected links.

### 5.2 The Determination of the Optimum Number of Traffic Counts

As mentioned above, the determination of the optimum number of traffic count will be conducted under **2 (two)** conditions representing the sensitivity between the number of traffic counts and the link rank to the acuuracy of the estimated OD matrices, namely: **sorted condition** and **random condition**. In condition I, several combination of traffic counts will be created based on the rank list of selected links which have passed the selection process. Each combination of traffic counts will then be used to estimate the OD matrices. In condition II, several combination of traffic counts will be created not based on the rank list of selected links but based on random selection.

In this research, the initial OD matrix was created by calibrating the **Gravity-Opportunity (GO)** model from traffic counts by using all selected links (**646 links**). **Anindito (2000)** reports that the best of values of parameters (ε and μ) are ε=0,4 and μ=1,0 for the GO model. The other unknown parameters (α and β) of the GO model were then calibrated using 646 selected traffic counts by using **Non-Linear-Least-Squares (NLLS)** estimation method. The initial OD matrix to be used for comparison purposes will then be created using the GO model together with the values of its calibrated parameters.

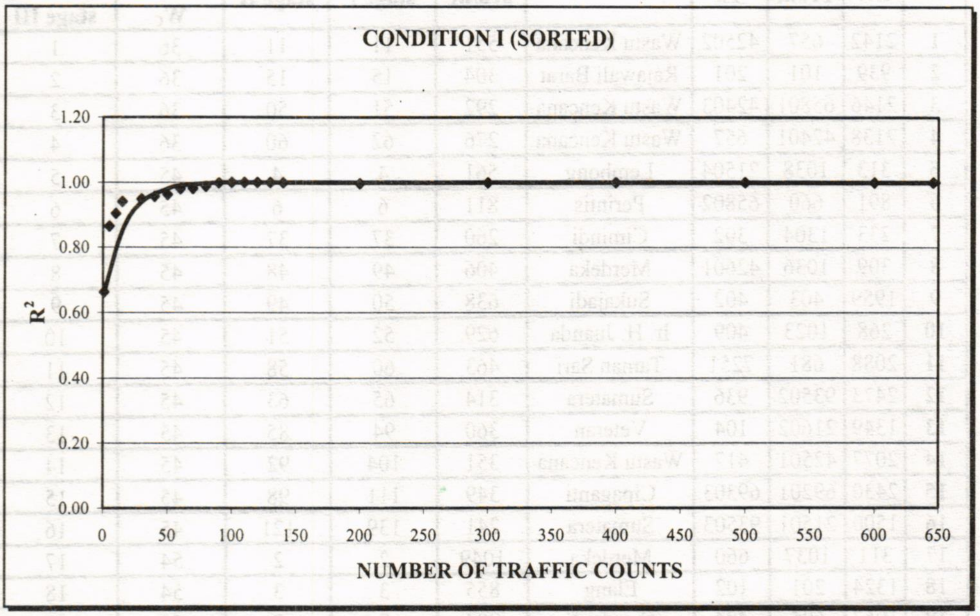


**Figure 2**: Number of traffic counts and the value of $R^2$ (sorted condition)

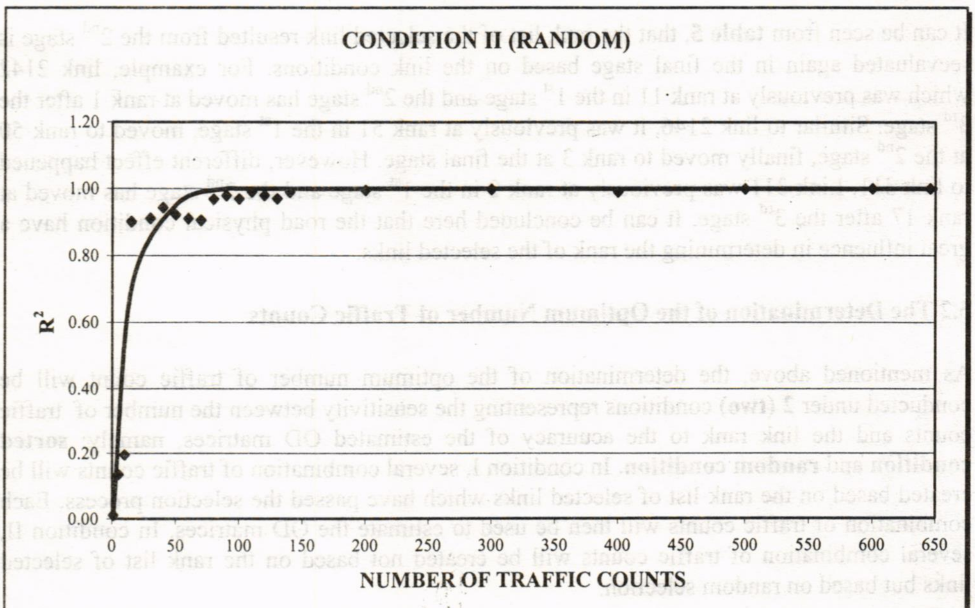


**Figure 3**: Number of traffic counts and the value of $R^2$ (random condition)

**Figures 2–3** show the relationship between the level of accuracy of the estimated matrices compared to the initial one and the number of selected traffic counts under **condition I (sorted)** and **condition II (random)**.

It can be seen from **figures 2–3** that the use of 90 links either for sorted and random conditions has reproduced the relatively high accuracy of estimated OD matrices compared to the initial one (in terms of $\mathbf{R^2}$). The use of 90 links has relatively the same accuracy with the use of 646 links for both conditions (sorted and random). It can be concluded here that the optimum number of traffic counts is **90 links (14% of 646 selected links** or **3.6% of 2485 available links)**. It can also be seen from those figures that the better accuracy of the estimated OD matrices are obtained under sorted condition rather than under random condition.

### 5.3 Testing Using Unselected Link Counts

The process to select the best location of traffic counts and to obtain the optimum number of traffic counts for Bandung have been mentioned. It is found that the optimum number of traffic counts is only 90 links of the selected links. To validate this result, another testing will be carried out to see the effect in the accuracy of the estimated OD matrices using traffic counts taken from the unselected links.

To do the validation, 20 groups of traffic counts were created taken randomly from the unselected links (each group contains of 90 link counts). Each group of traffic counts was then be used to estimated the OD matrices and then compared them with the initial one to see their accuracies. **Figure 4** shows the value of $R^2$ of the estimated OD matrices compared to the initial one for each group of combination of traffic counts (each group contains of 90 link counts taken from the unselected link).

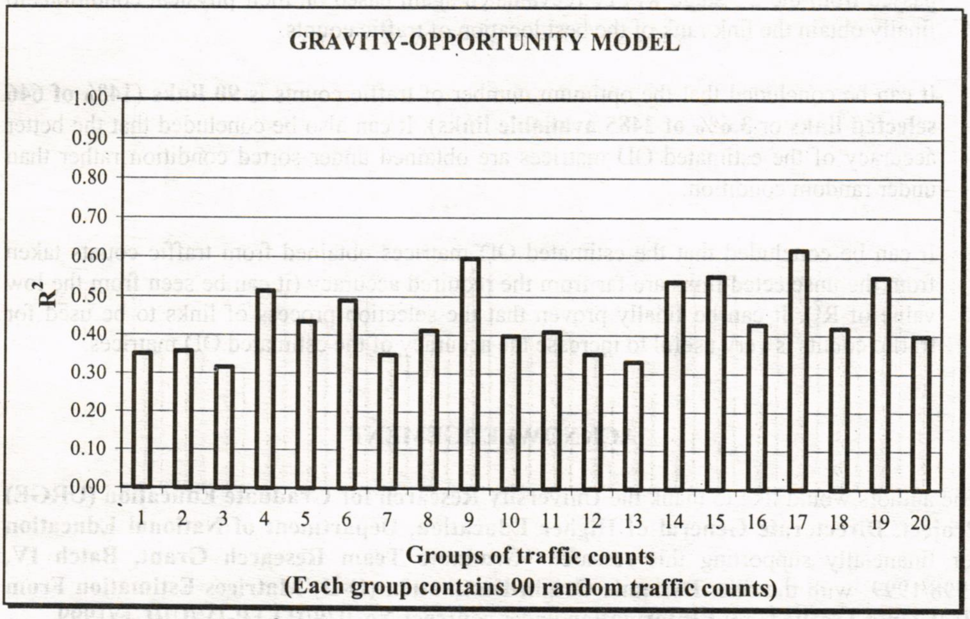


**Figure 4**: The value of $R^2$ for each group of traffic counts (taken from the unselected links)

It can be seen from **figure 4** that the values of $R^2$ vary from **0.30–0.60**; whilst if the selected links are chosen as the links for counted, the value of $\mathbf{R^2 \approx 1.00}$. It can be concluded that the estimated OD matrices obtained from traffic counts taken from the unselected links are far from the required accuracy (it can be seen from the low value of $\mathbf{R^2}$). It can be finally proven that the selection process of links to be used for traffic counts is very useful to increase the accuracy of the estimated OD matrices.

## 5. CONCLUSIONS

Several conclusions can be written as follows:

- The objective of this research is to identify the best location of links for counted and the optimum number of traffic counts required to estimate the high accuracy OD matrices.

- The selection process consider **3 (three)** main factors: **(a)** trip proportion of trip interchanges which will use a particular link, **(b)** the inter-link relationship which consider the interdependence and the inconsistency condition, and **(c)** the physical link condition of each selected link.

- The selection procedures have been tested in Kotamadya Bandung which consists of **145 zones** (**139 internal zones** and **6 external zones**) and considers **2485 links** to be selected for the best location of traffic counts (the road netowrk consists of arterial, collector, and parts of local roads).

- The **1st stage** of selection procedure can select **969 links** out of **2485 links (1516 links** dropped out for the next stage). The **2nd stage** of selection can select **646 links** out of **969 links** (another **323 links** dropped out). In the **3rd stage**, the **646 links** which have been passed from the 2nd stage will be reevaluated again based on their physical conditions to finally obtain the link rank of the best location of traffic counts.

- It can be concluded that the optimum number of traffic counts is **90 links (14% of 646 selected links** or **3.6% of 2485 available links)**. It can also be concluded that the better accuracy of the estimated OD matrices are obtained under sorted condition rather than under random condition.

- It can be concluded that the estimated OD matrices obtained from traffic counts taken from the unselected links are far from the required accuracy (it can be seen from the low value of $R^2$). It can be finally proven that the selection process of links to be used for traffic counts is very useful to increase the accuracy of the estimated OD matrices.

## ACKNOWLEDGEMENT

The authors would like to thank the **University Research for Graduate Education (URGE) Project, Directorate General of Higher Education, Department of National Education** for financially supporting this research **'Graduate Team Research Grant, Batch IV, 1998/1999'** with the title: **Dynamic Origin-Destination (O-D) Matrices Estimation From Real Time Traffic Count Information** under **contract No. 029/HTTP-IV/URGE/1999**.

## REFERENCES

**a) Books and Books chapters**

Tamin, O.Z. (2000) **Transport Planning and Modelling,** 2nd Edition. ITB Press, Bandung (in Indonesian).

Wilson, A.G. (1970) **Entropy in Urban and Region Modeling**. Pion Ltd, London.

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

**b) Journal papers**

Willumsen, L.G. (1981) Simplified Transport Model Based on Traffic Counts, **Transportation, Vol 10**, Elsevier Scientific Publishing Company, Amsterdam, 257–278.

**c) Papers presented to conferences**

**d) Other documents**

Anindito, W. (2000) The Impact of Intersection Delay and Route Choice Model in the Accuracy of the Estimated OD Matrices From Traffic Counts, **MSc Thesis** in Highway Engineering and System, ITB (in Indonesian).

Patunrangi, J. (2000) The Impact of Zoning System and Road Netowork Definition in The Accuracy of the Estimated OD Matrices From Traffic Counts, **MSc Thesis** in Transportation Engineering, ITB (in Indonesian).

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

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