Calibration the Gravity Model for Areas with Limited Data by Maximum Likelihood Method —Case Study in Nairobi (KENYA)—

By

Josphat K. Z. Mwatelah* and Yasunori Iida**

(Received June 22, 1994)

Abstract

This research was concerned with the task of calibrating a model that can be used to map out travel demand in metropolitan regions in developing countries where the cost of data collection is not negligible. The Maximum Likelihood (ML) method is found to be an appropriate approach using traffic count data at road sections, and population and employment data for the zones. It calibrates the gravity model for OD matrix estimation without any constraints. The proposed model is then applied to Nairobi (Kenya) and it is found that the feasibility of the model is sufficient for estimating the OD travel demand. Assuming that the calibrated parameters remain stable, the model can be used to predict future travel demand if the land use patterns are known.

1. Introduction

One of the major problems facing metropolitan regions in the world is vehicular traffic congestion. During the 1984 National Academy of Engineering Panel which examined the question of technology and methods of development in transportation, Boyce (1990) reported that the panel identified urban traffic congestion as the leading problem which US cities were facing. Ortuzar and Willumsen (1990) mentioned that vehicular traffic congestion has become more widespread and severe in both industrialized and less industrialized metropolitan regions.

The reasons for vehicular traffic congestion differ from one nation to an other. This variation is especially marked when we consider developed and developing countries. It is recorded that the number of passenger cars on the road in the industrialized countries doubled between 1970 and 1986, while its growth rate reached more than double in the developing countries (Ortuzar and Willumsen, 1990). Dimitriou and Banjo (1990) reported that the absolute value of trips per head popula-

* Lecturer, Department of Civil Engineering, Jomo Kenyatta University of Agriculture & Technology (Kenya).

** Professor, Department of Transportation Engineering, Kyoto University.

tion in developing countries is not higher than in developed countries, but it is high enough because the available road capacity is very limited in developing countries.

A large proportion of the transportation of goods and passengers in most developing countries is road-based. That is attributed to the weak economies in these countries. Better transportation network facilities are necessary for economic development since they stimulate economic growth and facilitate inter-regional trade. The road-based transportation system is the cheapest in terms of investment, and is convenient as long as environmental qualities are preserved.

It can be assumed that the number of vehicular-kilometers increases in any country as the economy considerably expands. Nairobi is not an exception to those cities whose countries are faced with the problem of traffic congestion while experiencing economic growth (Mwatelah, et al. 1992). Gichaga (1989) states that congestion in Kenya continues to be prevalent in most major towns. The Central Bureau of Statistics (1990) show that Nairobi (the major economic region) holds about 60% of the total National vehicle population. As shown in Table 1, Kenya in general has experienced a rapid growth of vehicle population. Table 2 shows the registration of new vehicles in Nairobi and Kenya which correspondingly confirm the high proportion of vehicle population in the Nairobi area.

Table 1 Number of Registered Vehicles in Kenya and Nairobi (1984-99)

Year	1984	1985	1986	1987	1988	1989
Kenya	266,610	280,218	280,435	316,403	336,609	336,609
Nairobi	138,469	138,273	152,531	162,668	194,421	194,978

Source : Central Bureau of Statistics & Ministry of Transport and Communication (1990)

Table 2 New Vehicle Registration in Kenya and Nairobi (1984-99)

Year	1984	1985	1986	1987	1988	1989
Kenya	15,694	13,663	16,955	18,727	19,524	20,206
Nairobi	8,151	6,462	9,272	10,261	11,559	10,284

Source : Central Bureau of Statistics & Ministry of Transport and Communication (1990)

However, the road network has not correspondingly expanded especially in the large towns due to the lack of adequate funds, and road maintenance has not been carried out in an appropriate time interval. This trend, however, is bound to change as reported in the current National Development Plan. It is reported that within the period of July 1977 to July 1992 the total classified road network recorded a growth of 12,700 km, of which 4,400 km were bitumen (National Development Plan 1994-96). It is important to investigate and determine the road links which are mostly affected by vehicular traffic congestion, so that we can recommend the level of improvement required for such links in the process of alleviating vehicular traffic congestion.

In order to make a proper plan for a transportation network in a metropolitan region, an estimate of a present Origin Destination (OD) matrix is needed. The i-j element of the matrix, T_{ij}, shows the level of interaction between zone i and zone j in the region. The level of interaction is measured by the number of passenger trips per day or the number of vehicle trips per day. The OD matrix is either descriptive of the present condition or predictive of the future travel demand. The present OD matrix and corresponding level of service can be adopted to evaluate the present traffic condition and used for short term transport planning, while the predictive situation can be used for both medium and long term planning horizons in order to avert an unfavorable travel demand situation.

Due to limited financial and technical resources, developing countries can not obtain sufficient transportation data that is available in sophisticated transportation models in order to produce OD matrices. The so called conventional OD estimation methods use the data obtained through very expensive data collection methods. In fact, some methods like road side interviews have the disadvantage of disrupting traffic flow and causing delays to road-users that may prove unpopular. In addition, these methods incorporate sampling strategies, and hence are not completely free from sampling bias (O'Neill, 1988).

In the traditional four-step sequential transportation models, each sub-model produces independent calculations and the outputs of one sub-model are used in the preceding sub-model calculations. If an error occurs at one stage, it will be transferred to the next stage. Therefore, the conventional method may produce unreliable results. Another disadvantage is that there is no consideration of the level of service (LOS) offered by the road network during the estimation of OD distribution, hence making this approach unrealistic. There is therefore a need for determining a method(s) that can produce a reliable OD matrix at a reasonable cost.

The main objective of this study is to develop an OD matrix estimation method which is available in regions with limited data and cost constraint on data collection. Such a method is thought useful for transportation facility planning in developing countries faced with rapid population growth in their urban areas and resulting increase to travel demand. Such a model(s) is seen as one that can be promptly utilized by transportation planners to alleviate both immediate and future transportation crisis.

2. OD Estimation Model

Apart from the traditional methods and the classic transportation models, the method that has attracted a lot of interest is the estimation of OD matrices using traffic counts which are easily and cheaply observed at road sections. The other merit of this approach is that the cost of data collection can be shared over several planning projects, e. g., road intersection design, road traffic analysis, and road monitoring. In this chapter, the previous studies on OD matrix estimation using traffic counts are reviewed and then an OD estimation model is formulated assuming the gravity form.

2.1 Previous Studies

(1) General Modelling Techniques

A study area of a transportation planning is divided into zones ; a centroid node represents a zone. A network consists of nodes and links. N and K denote a set of nodes ($n \in N$) and a set of directed links ($k \in K$), respectively. The network is represented by N and K, which is written $G=(N, K)$. A centroid is a node and included in the set of N. Each link is associated with a positive link function of traffic flow at a particular time. The total traffic flow in a link, v_k, is composed by link flows of zonal pair i, j. This relation is written as,

$$v_k = \sum_i \sum_j v_{ijk} \tag{2.1}$$

where v_{ijk} which denotes the number of trips of OD pair i, j who uses that link. If the left hand side of (2.1) is regarded as the existing information and the right hand side the expected flow on link k, this is a matrix estimation which describes a feasible region or constraint space over which a functional form of distribution assumption is optimized.

There are several types of OD matrix estimation methods, but the basic approaches are classified into two categories ;

(i) Consistency is sought between the observed and estimated traffic counts and consistency is also considered between a prior matrix and the matrix being updated, thus it is called Matrix Estimation Method (MEC).

(ii) Consistency between observed and estimated traffic counts is achieved through calibrating parameters, hence it is called Parameter Calibration Method (PCM).

Gur et al. (1980) refers to these estimation method as statistical and optimization techniques, respectively. The former method applies regression analysis in the estimation of parameters and the latter uses optimization methods to obtain the most likely trip table consistent with the traffic counts. The merit of the application of these two variations depends on the reliability of the available information. In each case, estimation of the OD matrix is through a trip distribution model, usually the gravity type model.

(2) O-D Matrix Estimation by Optimization Methods

In the OD matrix estimation methods, an existing OD matrix called the prior matrix is compared with the estimated matrix. The methods take advantage of mathematical programming techniques by which estimation of trip tables can be conducted directly from known information as opposed to calibrating parameters of travel demand models (Van Zulyen and Willumsen 1980). These methods share with parameter calibration models the need to combine trip distribution assumption. The main difference among those models is that some are derived as mathematical programming optimization techniques while some are developed from the principles of econometric analysis.

The well known expression for the maximization problem is generally formulated as,

$$\text{min. } F(\boldsymbol{T},\ \boldsymbol{T}^*) \tag{2.2}$$

subject to,

$$v_k=\sum_i\sum_j T_{ij}P_{ijk}$$

$$0\leqq P_{ijk}\leqq 1,$$

where $\boldsymbol{T}=\{T_{ij}\}$ is the prior O-D matrix, $\boldsymbol{T}^*$ is the targeted O-D matrices, v_k is the traffic volume on link k, and P_{ijk} is the proportion of trips on link k originating from zone i to zone j. The objective function denotes an index which shows the difference between $\boldsymbol{T}$ and $\boldsymbol{T}^*$.

Additional information such as mean trip length may be appended to the constraints above. Further, if non-proportional traffic assignment is assumed, other constraints that express proportions as implicit functions of travel demand are necessary because of under-specification of the model. Traffic counts provide insufficient information for determining a unique trip matrix, because the number of knowns (traffic volumes) is usually smaller than the number of unknowns (interchange trips). Moreover, conservation of traffic flow conditions tend to reduce the number of independent traffic counts, and subsequently the number of independent equations. Under-specification is not a problem in parameter calibration methods since the number of parameters to be calibrated are less than the number of zones plus two.

This modelling technique may not be appropriate for urban regions, where the existing matrix may not be reliable enough to give reasonable comparisons between the targeted and existing matrices during the computation. Such a condition is prevalent in urban regions in developing countries, where comprehensive studies have not been done to produce reliable OD matrices.

(3) Parameter Calibration Methods

Models involved in this category first assume the functional form for OD distribution and then calibrate the model parameters so that the calculated link flows are as consistent with the observed ones as possible. The typical functional form of OD distribution is the gravity model which is generally formulated as,

$$T_{ij}=f(U_i,\ V_j,\ t_{ij},\ \boldsymbol{\theta}) \tag{2.3}$$

where T_{ij} is the total number of trips generated between zones i and j, U_i and V_j are the trip generation and attraction factors in the respective zones. $\boldsymbol{\theta}$ denotes a parameter vector of the model.

There are many functional forms of the gravity model and the choice of the function depends on the availability of explanatory variables. Once the functional form of the model is determined, the values of the parameters are calibrated by regression analysis or maximum likeihood methods. These are further grouped into linear and non-linear methods depending on whether the model employ linear (Low, 1972 ; Bendetsen 1974 ; Holm et al. 1976 ; Gaudry and Lamare, 1979) or non-linear (Robillard, 1975 ; Hodberg, 1976) regression analysis.

Robillard's work shed a lot of light on parameter calibration models, and it has received significant recognition (O'Neill, 1987). Assumptions made in his model are:

(i) In the study area every origin communicates with every destination ;

(ii) Observed traffic volumes do not have to conserve flow ;

(iii) Observations do not have to be made on every link ; and

(iv) The error term between the observed and estimated flow is independently and identically distributed with the mean which is equal to zero.

The basic idea of the most likely OD matrix estimation is to obtain consistency or goodness of fit between the estimated and observed traffic flows (Holm et al., 1976 ; Iida et al., 1986). However, due to the fact that travel patterns vary between countries and cities, the use of parameter estimated from one city in the process of transportation planning is not recommended. In fact, it seems that it is not known what model formulation is the most useful for practical use (Iida et al., 1986). The reason is that the parameters may vary over long planning period even for the urban area. This point makes it imperative that each planning region is treated independently.

All papers on the linear methods, except that by Holm et al., assume the All-or-Nothing traffic assignment method. However, it is not easy to give appropriate level of service (e. g. travel time) for the parameter calibration of the OD estimation model. Thus, it is necessary to develop a model in which both OD estimation and traffic assignment are combined. In the following section of this paper, we present the basic model for OD estimation with given travel time and then combine it with Incremental Assigment.

2.2 Model Formulation

(1) O-D Estimation for Given Travel Time

Because of the easiness of determining population and employment data, in addition to easy calculation of the travel time along links, the functional from of the gravity model is determined as Eqn. (2.4). The model suggests that the number of generated trips between any two zones decrease as a function of the travel time. Here we consider the randomness factor. The drivers' perception of travel costs and their intention to minimize the composite measure result in traffic leaving a zone and entering in a link in a stochastic manner. Likewise, it can be assumed that arrival in the zones will also be stochastic. Therefore, we assume the following gravity type model with random component for OD distribution model.

$$T_{ij}=\alpha U_i V_j t_{ij}^{-\gamma}+\varepsilon_{ij} \tag{2.4}$$

where T_{ij} is the number of trips between zone pair i and j, and U_i and V_j are the population plus employment in zone i and zone j, respectively. t_{ij} is the given travel time of zone pair i and j. α and γ are the traffic generation parameter and the deterrence parameter, respectively. ε_{ij} denotes the error term.

The gravity model predicts the traffic volume in link k as;

$$v_k=\sum_i \sum_j T_{ij}P_{ijk} \tag{2.5}$$

where P_{ijk} is the proportion of traffic in link k generated from zone i to zone j. The value of P_{ijk} is assumed to be predetermined by any route assignment model using given travel time. For example, the Dial's method can be applied.

Introduction an auxiliary variable for the link traffic volum, let

$$v_k = \alpha X_k \tag{2.6}$$

where X_k is an auxiliary variable for the traffic volume in link k. Eqn. (2.6) can be treated as stochastic in that the traffic generation parameter α is only considered a stochastic variation. At the same time traffic assigned on the link is deterministic.

Suppose that the observed link thaffic counts, v_k^0, follows the Normal Distribution with the mean v_k and the variance in proportion to the mean value. This condition is,

$$v_k^0 = v_k + \varepsilon_k, \qquad \varepsilon_k \sim N[0,\ (v_k^0)^{\omega}\sigma^2] \tag{2.7}$$

where ε_k is the error term in link k, σ is the standard deviation independent of k and ω is a real constant weighing the differences between observed and estimated traffic flow. $\omega=0$ means that the standard deviation of estimated traffic is independent of its magnitude, $\omega=1$ corresponds to the case in which the mean and variance are proportional and $\omega=2$ is the case in which the standard deviation of traffic is proportional to the magnitude of traffic.

The classical theory of maximum likelihood is based on a situation in which H sets of observations $\boldsymbol{q}=(q_1, \cdots, q_h, \cdots, q_H)$ are independently and identically distributed, whose joint density function is given by ;

$$L(\boldsymbol{q}\ ;\ \boldsymbol{\theta}) = \prod P(q_h) \tag{2.8}$$

where $P(q_h)$ is the joint probability density function of the h-th set of observations. $\boldsymbol{\theta}$ denotes a parameter vector of the probabilistic function. $L(\boldsymbol{q}\ ;\ \boldsymbol{\theta})$ is interpreted as the Likelihood Function (LF) and the Maximum Likelihood Estimator (MLE) is found by maximizing the function with respect to $\boldsymbol{\theta}$ (Harvey, 1989).

The joint probability function of traffic flow between zone pairs i, j through link k is given by the following expression ;

$$\begin{aligned} L &= \prod_k P_k \\ &= \prod_k [2\pi(v_k)^{\omega}\sigma^2]^{-1/2} \exp[-(v_k^0 - v_k)^2/(2\sigma^2(v_k)^{\omega})] \end{aligned} \tag{2.9}$$

substituting for v_k in Eqn. (2.9) from Eqn. (2.6), we obtain the following expression,

$$L=\prod_k[2\pi(aX_k)^{\omega}\sigma^2]^{-1/2}\exp[-(v_k{}^0-v_k)^2/(2\sigma^2(aX_k)^{\omega})] \tag{2.10}$$

For maximization, we obtain the likelihood function by taking the logarithm of Eqn. (2.10) as;

$$lnL=-\frac{1}{2}\sum_k[ln(2\pi(aX_k)^{\omega}\sigma^2)]-\sum_k[(v_k{}^0-v_k)^2/2\sigma^2(aX_k)^{\omega}] \tag{2.11}$$

It seems analytically difficult to obtain all parameter values which minimize Eqn. (2.11). Let us derive the closed form of the parameters α and σ for given values of ω and γ. Partially differentiating Eqn. (2.11) with respect to α and σ, and then equating to zero,

$$\sigma^2=1/2\sum_k(v_k{}^0-v_k)^2/(v_k)^{\omega} \tag{2.12}$$

$$\alpha=\sum_k v_k{}^0(X_k)^{1-\omega}/\sum_k(X_k)^{2-\omega} \tag{2.13}$$

The solution to the maximum likelihood problem requires the solution of com-

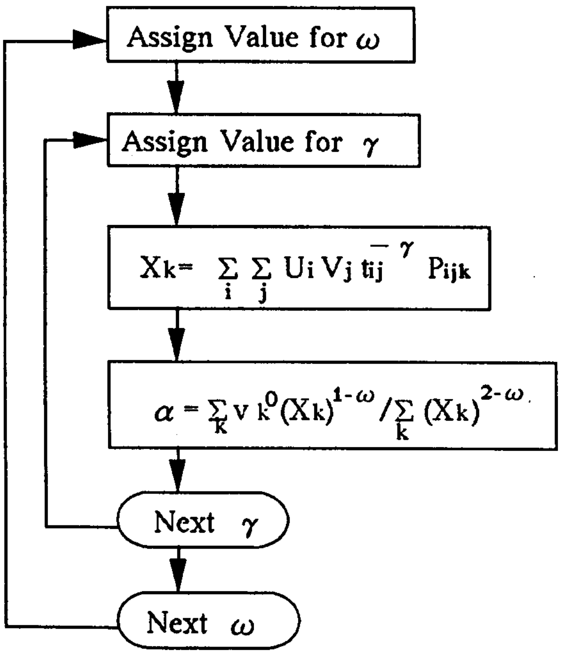

Figure 1 Algorithm for Estimating Parameter α

plex functions. The numerical methods involve approximations of the solutions and comparative methods until a stable solution is found. The O-D matrix is estimated through the iterative method at specific values of ω and γ. A combination of the values of ω and γ is given and then an OD matrix is calculated using the parameters above. Examining all possible combinations of ω and γ, we choose the OD matrix which minimizes the difference between observed link traffic counts and estimated ones. The computational procedure is depicted in Figure 1.

(2) O-D Estimation Model Combined with Traffic Assignment

The above mentioned OD estimation process assumed that the value of P_{ijk} was predetermined by any route assignment model using given travel time. In this section, we relax this condition and intend to determine the value of P_{ijk} within the model. It means that the OD estimation model is combined with Traffic Assignment model which calculates travel time and route. Let us explain the link travel time function and the traffic assignment before combining it with the OD estimation model.

At any particular instance, the travel time is determined on the basis of the amount to traffic loaded on to the road link. The speed-flow relationship is usually used to calculate the travel time at a certain traffic volume,

$$S_k=f(v_k) \tag{2.14}$$

where, v_k denotes traffic volume on link k, and $f(\ )$ is an increasing function with respect to traffic volume, taking into consideration the congestion effect.

The purpose of the traffic assignment is to provide an estimate of the amount of traffic which uses a part of the road network under given conditions. If drivers have complete information about the traffic conditions on a network and they are free to choose routes between OD pairs, the User Equilibrium (UE) conditions (Wardrop's first equilibrium principle) will be achieved. It is well known that the UE conditions are equivalent to solving a mathematical optimization program. However, it needs a lot of computational effort if we are to solve the UE assignment in the combined OD estimation and assignment model. Therefore, the following approximation method is considered instead of solving the UE assignment.

Initially drivers are free to choose routes in the network as in the UE condition. The majority of them attempt to minimize travel cost, taking into account traffic that results from other drivers' decisions. As traffic builds up to link capacity in the shortest route between zone pairs i and j, drivers diverge for the next shortest route and the process then repeats itself. Here the shortest path in the road network is found using Dijkstra method (Dijkstra, 1959). When all or part of the traffic between

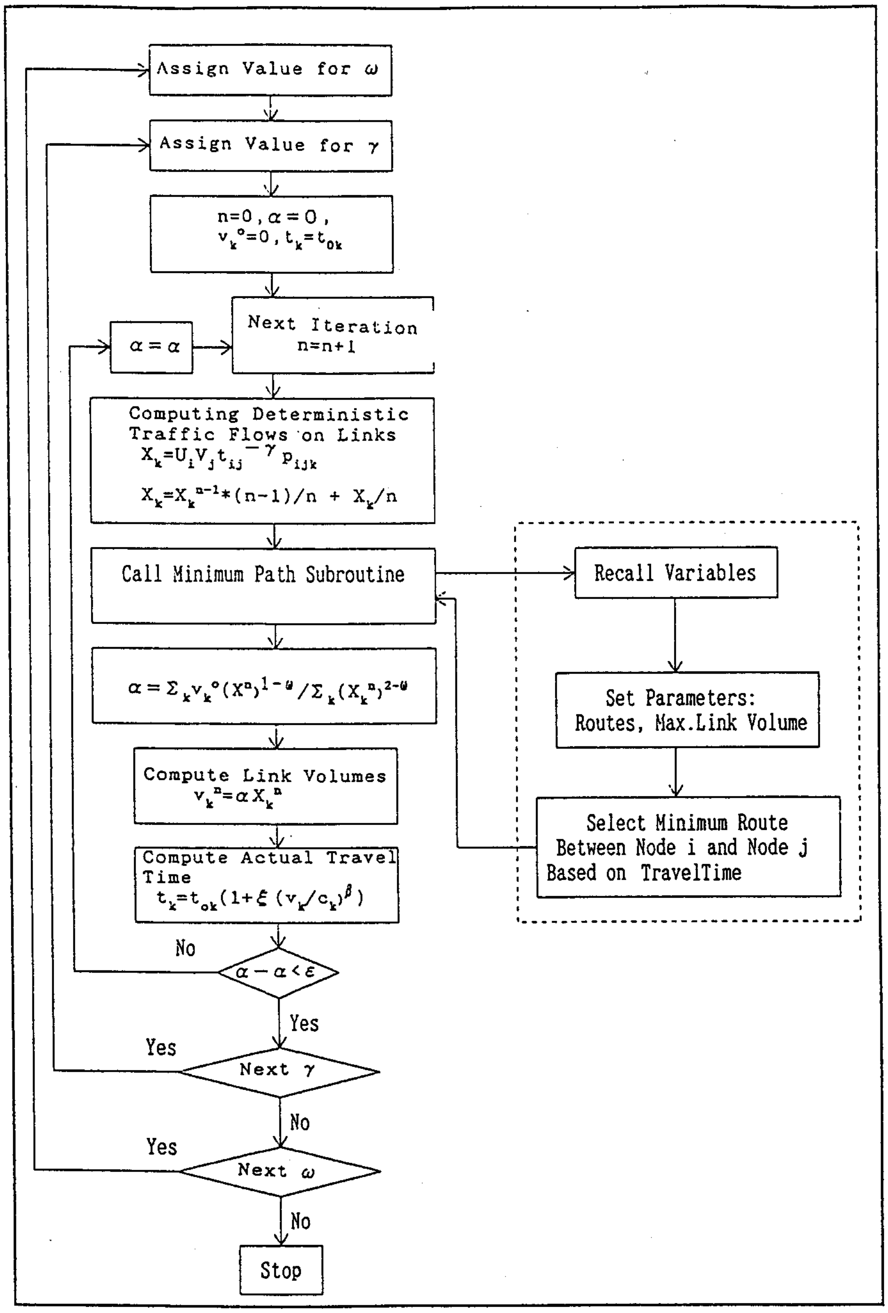

Figure 2 OD Matrix Estimation Combined with Incremental Assignment

zone pairs i and j are assigned to the routes in such a way, the averaged link traffic volume at an iteration is determined as follows,

$$v_k=((n-1)/n)v_k^{n-1}+v_k^n/n \tag{2.15}$$

where v_k^n is the traffic volume assigned along the minimum path on a link at the n-th iteration.

The above mentioned iterative assignment procedure is combined with the OD estimation process. The concept is that the link traffic volume and the link travel time are calculated and the value of P_{ijk} is up-dated in each inner iteration. Figure 2 depicts the flowchart of the combined model.

The calculation steps are explained as follows.

Step 0 : Set values for ω, assume zero traffic flow and calculate t for every link and assign traffic on every link.

Step 1 : Set $n=1$. Determine the fastest routes between zone pairs in the network.

Step 2 : Using the calculated travel times, compute the trips between zone pairs using the gravity model, without the traffic generation factor. This deterministic portion of traffic is assigned to the fastest routes between zones.

Step 3 : Estimate the trip generation factor so that the difference between observed and predicted traffic counts are minimized.

Step 4 : Calculate traffic volumes in the network by multiplying the deterministic portion of the traffic with the trip generation factor.

Step 5 : Calculate new travel times in the network.

Step 6 : If the value of the traffic generation is stabilized, stop ; otherwise go to step 2.

3. Case Study in Nairobi

3.1 Transport and Traffic Problems in Nairobi

The urbanization trend in developing countries has been disproportionate to the available resources. The idea that urban areas offer better job opportunities than rural areas lures most of the population into these regions, resulting in a high rise of travel demand, and putting a lot of pressure on transportation facilities.

In 1983 the Kenya government introduced the District Development committee (DDC) whose aim was to provide a mandate to the local communities which would initiate appropriate projects to offer employment in the respective areas. Another step taken by the government was to upgrade small towns to urban areas in order to give the employees buying power that would be equivalent to that of those people

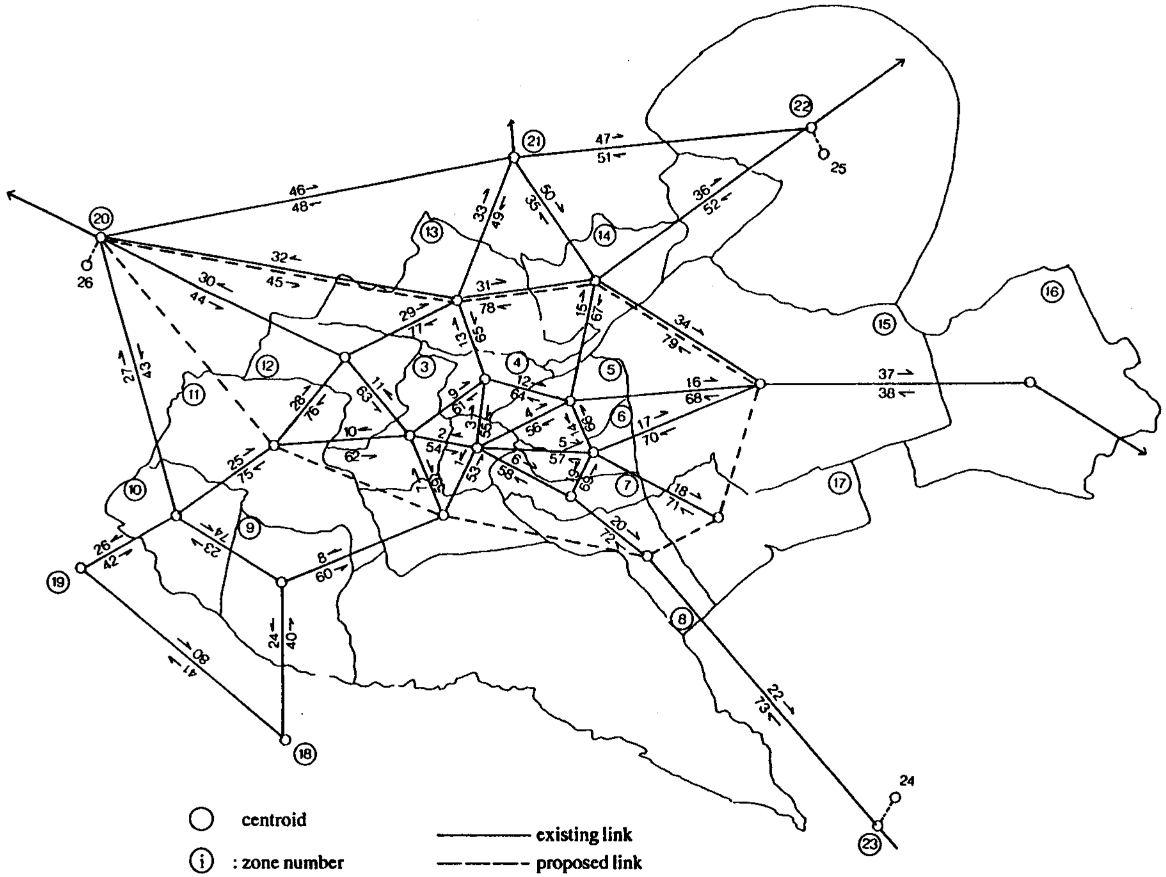

Figure 3 Zoning and Road Network Configuration in Nairobi

working in the larger towns. Despite this strategy, the urban population growth nationally was recorded at 4.8% and Nairobi had a 5% growth rate (Central Bureau of Statistics 1990). The existence of transportation problems in Nairobi has been discussed in another paper (Mwatelah et al. 1992). Other reports that discuss the same issues in Nairobi are, for example: Nairobi Growth Strategy (1973), Transurb Consult (1986), and JICA report (1988). Although the report by JICA actually produced an OD matrix, it is reported that it may not be reliable.

3.2 Zoning and Data Collection Methods

The Nairobi Metropolitan Region was divided into 17 zones having generally similar socio-economic characteristics as the existing development plan from the Nairobi City council. The surrounding zones, i.e. Athi Rever, Kiserian, Ngong, Kikuyu, Limuru and Kiambu, were included, making the region 23 zones. The geographical centres for each zone were chosen as centroids. They are connected to their neighboring centroids, where main roads exists. Figure 3 shows the zoning and the road network configuration.

The explanatory variables data and traffic count data used to describe the present OD matrix for the region were adopted as shown in the JICA report for the Ministry of Public Works of the Government of Kenya 1988 (JICA, 1988).

Table 3 shows the population and employment data for every zone. These were exogenous explanatory variables used for traffic distribution through the gravity model. This is due to the understanding that within 24 hours, all trips originating from a zone will likely be attracted back into it, because they arise from the employment and population in the zones.

Nairobi is an economic core for the Eastern and Central African region, apart

Table 3 Night Population and Employment in Nairobi (1986)

Zone	Pop.+Emp.	Zone	Pop.+Emp.	Zone	Pop.+Emp.
1	249,106	9	9,454	17	39,025
2	110,457	10	8,248	18	14,000
3	66,888	11	148,833	19	14,000
4	66,394	12	13,050	20	126,000
5	246,094	13	12,559	21	63,000
6	174,179	14	55,128	22	166,172
7	85,267	15	190,172	23	66,000
8	23,154	16	4,479	Total	1,951,659

Source: JICA Report (1988)

Table 4 Link Capacity, Free Flow Travel Time and Observed Link Traffic

Link	C_k	t_{k0}	v_k^0	Link	C_k	t_{k0}	v_k^0	Link	C_k	t_{k0}	v_k^0
1	6000	5	—	28	6000	10	—	55	6000	5	—
2	18000	5	5944	29	6000	7	—	56	12000	6	—
3	12000	5	—	30	6000	12	156	57	12000	7	—
4	12000	6	—	31	6000	8	—	58	27000	5	13497
5	12000	7	—	32	6000	20	1456	59	12000	5	—
6	27000	5	9365	33	6000	8	—	60	6000	4	6167
7	12000	5	—	34	6000	15	—	61	6000	6	—
8	6000	4	5061	35	6000	11	—	62	6000	6	4161
9	6000	6	—	36	18000	15	—	63	18000	5	87
10	6000	6	4200	37	6000	6	—	64	6000	6	—
11	18000	5	414	38	6000	6	—	65	6000	5	2844
12	6000	6	—	39	12000	5	—	66	6000	5	—
13	6000	5	1935	40	6000	10	1570	67	18000	8	5005
14	6000	5	—	41	6000	15	—	68	6000	11	—
15	18000	8	4597	42	6000	5	2354	69	12000	5	11233
16	6000	11	—	43	6000	15	—	70	6000	8	2486
17	6000	8	2827	44	6000	12	163	71	6000	8	4077
18	6000	8	3094	45	6000	20	1285	72	18000	8	7811
19	6000	5	8896	46	6000	30	—	73	6000	10	1272
20	12000	3	7179	47	6000	20	—	74	6000	7	653
21	12000	5	—	48	6000	30	—	75	6000	5	—
22	6000	10	1293	49	6000	8	—	76	6000	10	—
23	6000	7	610	50	6000	11	—	77	6000	7	—
24	6000	10	1293	51	6000	20	—	78	6000	8	—
25	6000	5	—	52	18000	15	—	79	6000	15	—
26	6000	5	1666	53	6000	5	—	80	6000	15	—
27	6000	15	—	54	18000	5	6031				

from being the capital city of Kenya, where two international roads traverse (Mwatelah et al. 1992). So the traffic passing through the Nairobi Metropolitan Region is not so small. In order to remove the effects of this traffic from the calculation, the traffic attracted into districts that are not involved in the study area were aggregated as per the direction of approach into or out of the study area. This traffic data was obtained from the O-D matrix in the JICA report. Three dummy centroids connected by 6 dummy links were created at points where the through roads enter the study area (see Figure 3). The traffic was then assigned once onto the network in the study area, and the assigned trips subtracted from the observed traffic counts. Table 4 summarizes the link capacities, zero flow travel time and the observed traffic flows.

The link performance function was taken as the US-BPR function,

$$t_k = t_{k0}\{1 + \xi[v_k/c_k]^{\beta}\} \tag{3.1}$$

where t_k is the travel time with a certain traffic volume on the link k, t_{k0} is the zero flow travel time, v_k is the actual traffic assigned on the link, c_k is the link capacity and ξ and β are coefficients characteristic to the region. The initial travel time was calculated from the average distances between centroids divided by the average speeds of the respective road link. A lane on a highway has an average of 9,000 vehicles per day and that of a regional road is 6,000 vehicles capacity. So in this study the mean capacity for each link taken was 7,500.

3.3 Parameter Values

Two sets of parameter values of the link travel time function were compared in the calculation. The first set is that of the original US BPR function whose parameters are $\xi=0.15$ and $\beta=4.00$. This case is recommended for areas having low density population and wide roads. The second set is the modified US BPR function whose values are $\xi=2.62$ and $\beta=5.00$, which are recommended for regions having high density population and narrow roads. The range of the parameter γ of the deterrence function is chosen from 4.00 to 2.25. This range is regarded as having practical meaning.

Weighting factor $\omega=0$ has no practical meaning and it is ignored. Considering the figures on population, employment and number of registered vehicles, it is assumed that at least one vehicle produces 1.5 inter-zonal trips. Therefore, $\omega=1$ is acceptable for estimating an OD matrix. $\omega=2$ was found to produce an unreasonably large travel demand, hence this parameter was ignored.

3.4 Calculations

Figures 4 shows the graph of the weighted variance against deterrence coefficient γ. The set of the parameter values of $\xi=0.15$ and $\beta=4.0$ are used. Figure 5 shows the case for the set of parameters of $\xi=2.62$ and $\beta=5.0$. In both figures, it is found that the optimum values of the deterrence coefficient is $\gamma=2.75$ which minimizes the weighted variance.

The set of $\xi=2.62$ and $\beta=5.0$ shows higher correlation coefficient between the observed and estimated link traffic volume. Figure 6 shows a plot of the estimated and observed traffic counts of this parameter set. The value of correlation coefficient is 0.8673 and it shows a high correlation. It confirms the fact that these results are

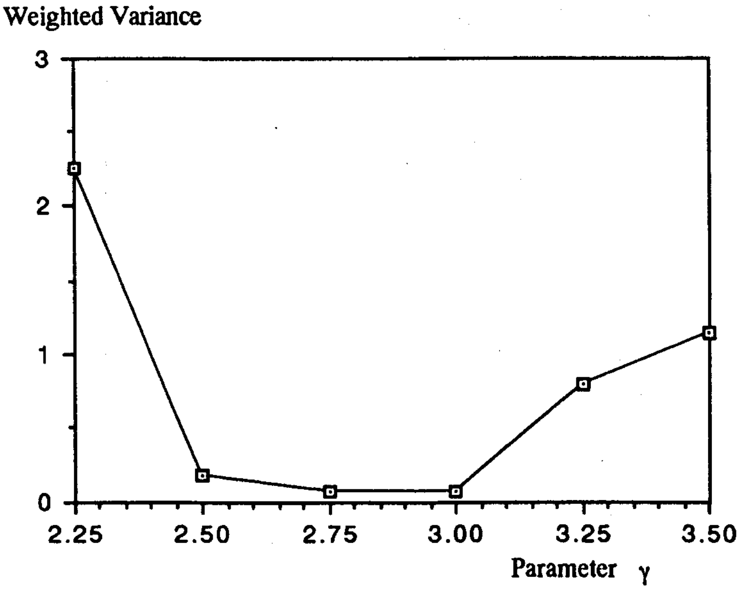

Figure 4 Weighted Variance of Link Traffic

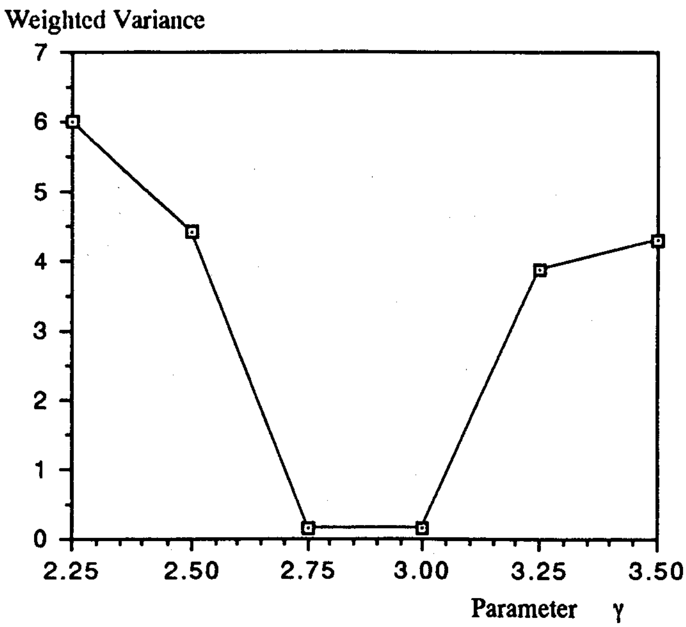

Figure 5 Weighted Variance of Link Traffic

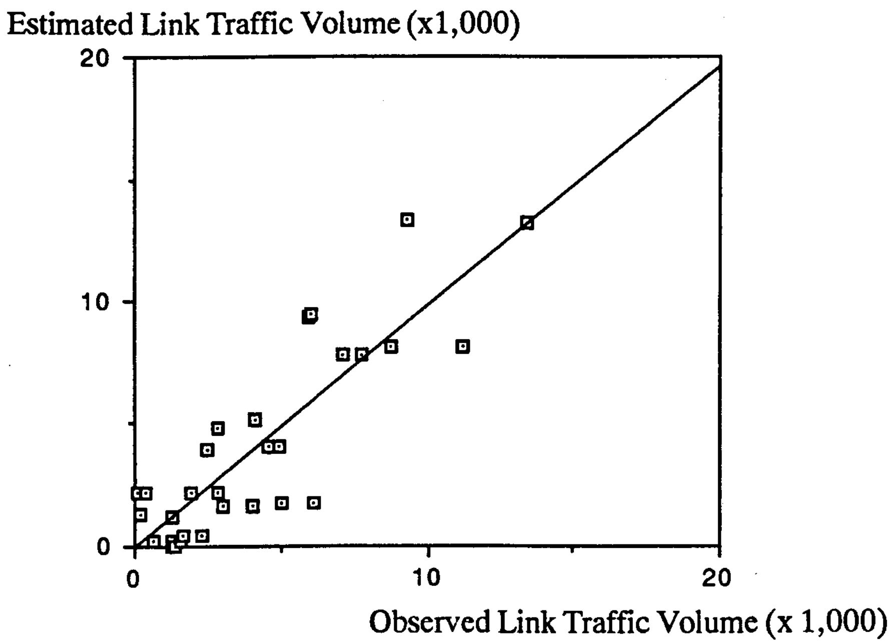

Figure 6 Correlation between Observed and Estimated Link Traffic Volume

reliable.

The traffic generation parameter was estimated as $\alpha=5.572\times10^{-5}$. The estimated OD matrix is shown in Table 5. The total amount of travel demand in the study area is estimated as 208,626 vehicle trips per day. This value is 2.5 times larger than that presented in the JICA report (1988), even though it includes intra-zonal trips. Since the method used in the JICA report was not clear to us, the direct comparison of the figures between these two cases is not easy. However, the results obtained by the proposed model seem more useful because its methodology is clear and the estimated values are credible.

4. Conclusion

The method has successfully estimated an OD matrix for the Nairobi Metropolitan Region. The gravity model was consecutively calibrated through estimating the matrix. Therefore, the method can be used to describe the travel demand by deriving the OD matrix table. The assumption of stochastic traffic flow is suited to a region

Table 5 Estimated OD Matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	0	3979	5867	7098	8512	6389	8911	713	120	32	1286	230	244	323	799	9	365	36	29	303	270	169	257
2	3976	0	4495	462	573	417	328	51	1277	69	713	136	33	50	126	2	44	61	42	154	57	41	41
3	5798	4497	0	1368	683	379	324	45	79	25	1626	582	55	57	97	1	36	15	17	194	65	36	32
4	7101	457	1340	0	3273	411	357	49	18	8	321	56	546	175	206	2	38	7	7	77	200	64	33
5	8512	573	591	3200	0	5391	713	111	28	12	376	51	154	2456	3138	23	215	14	13	131	262	409	90
6	6518	422	386	415	5391	0	7026	538	20	8	253	34	35	214	2964	20	1240	10	9	89	67	115	187
7	8090	328	326	357	713	6990	0	4085	14	5	173	25	25	48	437	4	548	7	5	56	44	36	242
8	712	51	46	49	111	536	4001	0	2	1	31	4	4	9	71	1	602	1	1	11	8	8	144
9	120	1277	79	18	28	20	14	2	0	21	84	5	2	3	7	0	2	13	8	13	3	3	2
10	32	69	25	8	12	8	5	1	21	0	816	3	1	2	4	0	1	3	77	34	3	2	1
11	1279	712	1625	323	381	250	173	31	84	816	0	192	43	65	99	1	31	24	206	276	75	54	32
12	228	136	528	57	52	34	25	4	5	3	192	0	43	23	12	0	4	2	3	99	27	10	4
13	245	33	54	546	156	34	25	4	2	1	43	43	0	127	24	0	4	1	1	27	145	21	4
14	323	49	57	175	2456	214	48	9	3	2	65	23	127	0	341	3	17	2	2	45	265	298	10
15	722	117	91	203	3094	2438	385	64	7	4	94	12	24	341	0	344	123	5	4	46	86	153	57
16	8	2	1	2	23	18	4	1	0	0	1	0	0	3	344	0	1	0	0	1	1	2	1
17	365	44	36	38	215	1240	548	602	2	1	31	4	4	17	138	1	0	1	1	13	9	14	80
18	36	61	15	7	14	10	7	1	13	3	24	2	1	2	5	0	1	0	6	7	2	2	2
19	29	42	17	7	13	9	5	1	8	77	206	3	1	2	5	0	1	6	0	26	3	3	1
20	302	154	194	78	132	89	56	11	13	34	276	99	27	45	46	1	13	7	26	0	51	40	15
21	270	56	65	200	262	67	44	8	3	3	75	25	145	265	86	1	9	2	2	51	0	154	10
22	169	41	36	64	408	115	36	8	3	2	54	10	21	297	152	2	14	2	3	40	154	0	12
23	257	41	32	33	90	186	242	144	2	1	32	4	4	10	61	1	80	2	1	15	10	12	0

with less traffic data since it has no constraint in the iterative process when the matrix is estimated. The calibrated gravity model can be used to simulate travel demand in either short term or long term plans by assuming that the parameters of the gravity model do not change. Such an approach may be applicable to metroptlitan regions in developing countries, since the data required for the model can be cheaply acquired. There is need to further test this model in a similar study area in order to confirm the results presented here.

Acknowledgements

The authors are indebted to the Japan International Cooperation Agency (JICA) for providing financial support and giving us access to useful data. Many thanks to Prof. H. Nakagawa of Kyoto University for his useful advice, Prof. H. Tsukaguchi of Ritsumekan University and Prof. J. Takayama of Kanazawa University for their untiring encouragement. This research would not look the way it is today without the assistance of Dr. T. Uchida, Mr. N. Uno and Mr. K. Iida in the Department of Transportation Engineering Kyoto University. Finally, the gratitude is also expressed to Prof. Y. Asakura of Ehime University for his suggestions on this paper.

References

1) Boyce D. E. (1990): Research Data and Technology Development, Environment & Planning A, Vol. 22, pp. 1275-1280.
2) Central Bureau of Statistics (1990): Statistical Abstract, Government Printers.
3) Dijkstra E. W. (1959): A Note on Two Problems in Connection with Graphs, Numerische Mathematik 1, pp. 269-271.
4) Dimitriou H. T. & Banjo G. A. (1990): Transport Planning for Third World Cities, Routledge.
5) Gichaga F. J. (1989): The Engineer in Road Building; The Kenyan Perspective (Inaugural Lecture at the University of Nairobi).
6) Republic of Kenya (1994): National Development Plan 1994-'96, Government Press.
7) Harvey A. C. (1989): Forecasting Structural Time Series Models and the Kalman Filter, Cambridge University Press.
8) Hogberg P. (1975): Estimation of Parameters in Models for Traffic Prediction: Non-linear Regression Approach, Transportation Research, Vol. 10, pp. 263-265.
9) Holm J. et al. (1976): Calibrating Traffic Models on Traffic Census Results Only, Laboratory of Road Data Processing, Danish Road Directorate.
10) Iida Y. and Takayama J. (1986): O-D Matrix Estimation Models From Traffic Flows, Proceedings of the Japan Society of Civil Engineering; Infrastructure Planning, pp. 97-104.
11) JICA (1988): The Nairobi South By-pass, Construction Feasibility Study Project; A Report for the Ministry of Public Works of the Government of Kenya.
12) Mwatelah, et al. (1992): Urban Transportation Planning Issues in Kenya (Case Study-Nairobi), Proceedings of Infrastructure Planning JSCE, Vol. 15, pp. 219-224.
13) Nairobi Metropolitan Growth Strategy Study Group (1973): The Nairobi Growth Strategy, Nairobi City Council.
14) O'Neill W. A. (1988): Origin Destination Trip Tabke Using Traffic Counts, Ph. D. Dissertation, State University of New York at Buffalo.
15) Ortuzar and Willumsen (1990): Modelling Transport, John Wiley.
16) Otto H. K. et al. (1971): Infrastructure Problems in the Cities of the Developing Countries, World Bank.
17) Robillard P. (1975): Estimating O-D Matrix from Observed Link Volumes, Transportation Research, Vol. 9, pp. 123-128.

18) Transurb Consult (1986) : Study of Urban Transport Needs of Nairobi.
19) Van Zuylen H. J. and Willumsen L. G. (1980) : The Most Likely Trip Matrix Estimation from Traffic Counts, Transportation Research, Vol. 14B, pp. 281-293.
20) Yang H. (1991) : Estimating Origin-Destination Matrices from Traffic Counts in Congested Networks, Ph. D. Dissertation, Kyoto University, Japan.