USE OF ON-OFF COUNTS FOR OD ESTIMATION
AN APPROACH TOWARDS MORE COST-EFFECTIVE BUS SURVEYS*

By Terdsak RONGVIRIYAPANICH**, Fumihiko NAKAMURA*** and Izumi OKURA****

1. Introduction

Bus operators, particularly in developed countries, have witnessed the declining number of patronage due to the rapid motorization in past decades. In Japan, despite of its advanced rail-based transit system, patronage of the bus system has been decreasing continually. Given high level of car ownership, more efficient operation is therefore required in order to make bus transit as much attractive as possible. It is clear that more up-to-date and in-depth data is important to the operation planning, which in turn is used to make the service meet passengers' needs. Conventional on-board survey, which is conducted in Tokyo Metropolitan area every five year to obtain the route OD matrix by employing two surveyors on each bus for a whole weekday, is rather labor and capital intensive. In addition, this estimation scheme cannot deal with day-to-day as well as seasonal variation, which is a nature of the OD matrices.

Recent attempts by some local bus operators, for instance those in Tokyo or Yokohama, to develop advanced management system such as Automatic Vehicle Location (AVL) system or automatic passenger counter system, so called BDCS, have helped broaden the database. This system offers us an abundant source of aggregate data of bus operation and passenger demand, which although cannot be used alone to estimate OD matrix, yet can capture the variation in travel demand. However, lack of capable data analysis framework has impeded the integration of such automatically collected information with the survey data. As a result, such data has been left unused though the installation cost of the system is considerably high. It is therefore of interest to address the following questions. First is "can data obtained automatically through the bus data collection system (BDCS) be utilized in bus service planning procedures?" Another question is "is it possible to replace the costly conventional OD survey by use of the BDCS data in conjunction with other readily available data?" These two questions are among motivations of this study.

OD matrix at the route level, as suggested by several researchers, is important information for bus service planning and management[1]. Its applications are for instance, patronage forecasting or predicting the effect of change in level of service[2]. Ben-Akiva et al.[3], among others, recommended the use of on-board survey in conjunction with ride-check data as an alternative method to estimate route-level OD matrix. It is revealed in their study that the simple expansion of on-board survey gives less accurate result than that obtained by the intervening opportunity method. However, comparisons of performances of the methods are still limited and mostly based on simulated data or test networks.

Recently on-off count data has become readily available due to the widespread use of automatic passenger counter system. Hence, it is of interest to evaluate performance of the methodologies to estimate OD matrices based on real data. In this study we use the real data obtained from two bus routes in Tokyo Metropolitan. Thanks to the availability of actual daily OD matrices of two years for the two bus routes, comparison of the models in terms of accuracy of the estimates for large-scale data is made possible.

In spite of obvious usefulness of on-off counts, which is readily available from the BDCS, as a basic data for estimating OD matrices, in practice the operator has still relied solely on OD matrices obtained by the conventional surveys. In this paper, we therefore attempt to show that by using the available information efficiently, the expenses for service planning can be significantly cut. Effects of the sample size on accuracy of the estimates are also examined. In addition, due to the simpler nature of route-level OD matrix coupling with an availability of BDCS data, we shall explore possibility in simplification of OD estimation problem at the route level.

The organization of this paper is as follows. In section 2 some past studies on OD estimation techniques are reviewed. Description of data and methodology used in the present study are outlined in section 3. We then show some empirical results of this study in section 4. Finally, in section 5 discussion and conclusion are made.

2. Theoretical Background

Estimation problems for road network OD and route-level bus OD are different in sense that the latter does not require any assumption on the route choice probability. Thus it could be estimated in more simple and accurate manner, as error in route choice probability becomes irrelevant. The nature of bus-route OD estimation problem is more or less comparable to that of expressway, in which inflow and outflow at each point along single route are known yet distribution of trips among them requires estimation.

For the case of road network, the most widely used methods to estimate OD matrices from traffic counts can be classified into two major families[4]. First, the entropy maximization method (ME), or sometimes known as the iterative proportional fitting method (IPF), is based on the concept of information minimization. The other family of OD estimator is known as the statistical inference method, including the maximum likelihood method (ML), the general least square method (GLS), the Markov chain process and the Bayesian method. These methods can be applied to the bus route-level OD estimation problem as well by using on-off counts instead of traffic counts. In addition, for bus route-level OD estimation, some other methodologies, as will be mentioned hereafter, can also be applied. The growth factor method, which is well known for its computational simplicity, is not recommended here given that BDCS data is available. However, its result will be used as a basis for evaluating performance of the other methods, which utilize BDCS data. It is also possible to estimate OD matrices when ride-check data alone is available, by using the intervening opportunity (IO) method. The formulation of estimation methodologies used in this study can be summarized as follows.

(1) ME estimators

This methodology is based on the concept of entropy maximization and information minimization, which was proposed by Willumsen[5]. The ME estimator for OD pair (i,j) is obtained by

$$\text{Max } S(\hat{t}_{ij} / t_{ij}^0) = -\sum_{ij} (\hat{t}_{ij} \log \hat{t}_{ij} / t_{ij}^0 - \hat{t}_{ij}) \tag{1}$$

$$\text{s.t.} \quad \sum_{j=1}^{J} \hat{t}_{ij} = t_{i\cdot} , \quad \sum_{i=1}^{I} \hat{t}_{ij} = t_{\cdot j} \tag{2}$$

The formal solution of the above problem is given by

$$\hat{t}_{ij} = a_i b_j t_{ij}^0 \tag{3}$$

where

$\hat{t}_{ij}$ = the number of passengers boarding at bus stop (origin) i and alighting at bus top (destination) j

t_{ij}^0 = the observed number of trips, which is originated at bus stop i and terminated at bus stop j of priori OD matrix

$t_{i\cdot}$ and $t_{\cdot j}$ = the passenger count constraints on bus stop i and bus stop j respectively

a_i and b_j = multiplying factors of row i and column j, determined iteratively until the row and column constraints as shown in equation 2 are satisfied

I and J = the total number of bus stops

(2) ML estimators

ML technique was presented by Landau et al.[6]. The technique is based on the concept of maximum likelihood of observing the experiment data conditional on the true trip matrix. Under the assumption of Poisson or Multinomial distribution of trips, the ML estimator is obtained by solving the following:

$$L = \prod_{i=1}^{I} \prod_{j=1}^{J} \left(t_{ij}^{t_{ij}^0} / t_{ij}^0! \right) \exp(-t_{ij}) \tag{4}$$

subject to equation 2. Applying Lagrangian for the above set of equations, we obtain the first-order condition as

$$\partial L^0 / \partial t_{ij} = (t_{ij}^0 / t_{ij}) - a_i - b_j = 0 \tag{5}$$

Thus the ML estimator is

$$\hat{t}_{ij} = t_{ij}^0 / (a_i + b_j) \tag{6}$$

(3) GLS estimators

The GLS method is based on the assumption that simple expansion gives unbiased estimators of the true OD (t_{ij}). The concept of GLS was first applied to OD matrix estimation problem by Mcneil[7]. The GLS estimator is determined by minimizing

$$\left(f \cdot t^0 - t \right)' V^{-1} \left(f \cdot t^0 - t \right) \tag{7}$$

subject to

$$Rt = r \tag{8}$$

where

f = the expansion factor, given by $t_{..}/t_{..}^0$

V = variance-covariance matrix of t^0 (in this study we assume that $V_{ij} = t_{ij}^0$)

r = row and column constraints and

R = constraint incidence matrix which contains only elements with value of 0 or 1

By solving equation (7) subject to equation (8), the GLS estimator can be expressed as follows;

$$\hat{t} = f \cdot t^0 + VR'(RVR')^{-1}(r - Rf \cdot t^0) \tag{9}$$

(4) IO estimators

Tsygalnitsky[8] proposed that IO method could be used for estimating OD matrix when the priori information on OD pattern is not available. Under the assumption of equal probability of alighting for all passengers on-board, whose travel distance already exceeded minimum required distance, we can estimate number of trips between particular OD pair (i,j). Here we assume that no passenger alights at the next bus stop from the stop he boards, unless the number of qualified passengers according to the above criteria is less than the observed number of alighting passengers.

The algorithm can be summarized as below:

For $j \geq i + 2$

$$V_{ij} = t_{i \cdot} - \sum_{k=i}^{j-1} \hat{t}_{ik} \tag{10}$$

and for $j < i + 2$; $V_{ij} = t_{ij} = 0$

At bus stop j ;

$$V_{\cdot j} = \sum_{i=1}^{j} V_{ij} \tag{11}$$

where

V_{ij} = the number of passengers who board at bus stop i and are qualified for alighting at bus stop j

$V_{\cdot j}$ = all qualified passengers for alighting at bus stop j

$t_{i \cdot}$ = total number of passengers boarding at bus stop i

As a result, IO estimator for OD pair (i,j) is given by

$$\hat{t}_{ij} = (t_{\cdot j} / V_{\cdot j}) \cdot V_{ij} \tag{12}$$

(5) Growth factor estimators

Under the assumption that OD pattern remains unchanged between the base and present year, the growth factor estimators can be simply obtained uniform scaling of the outdated OD as shown in equation (13).

$$\hat{t}_{ij} = (t_{..} / t_{..}^0) \cdot t_{ij}^0 \tag{13}$$

where

$t_{..}^0$ = the total number of trips in the base year OD matrix.

$t_{..}$ = the total number of trips in the present year OD matrix.

In summary, characteristics of the OD estimation techniques, mentioned above, can be summarized as shown below in table1.

Table1: Characteristics of the OD estimation techniques used in the present study

Method	Data required		Assumption
	OD survey	Ride check	
ME	*	**	Information minimization concept
ML	**	**	Sampling distribution of the base year OD is either multinomial or Poisson
GLS	**	**	Base year OD matrix is the unbiased mean of the true OD
IO	***	**	Equal probability of alighting for every passenger on- board
Growth factor	**	***	OD pattern remains unchanged

* necessity depends on the approach taken, ** necessary, and *** not required
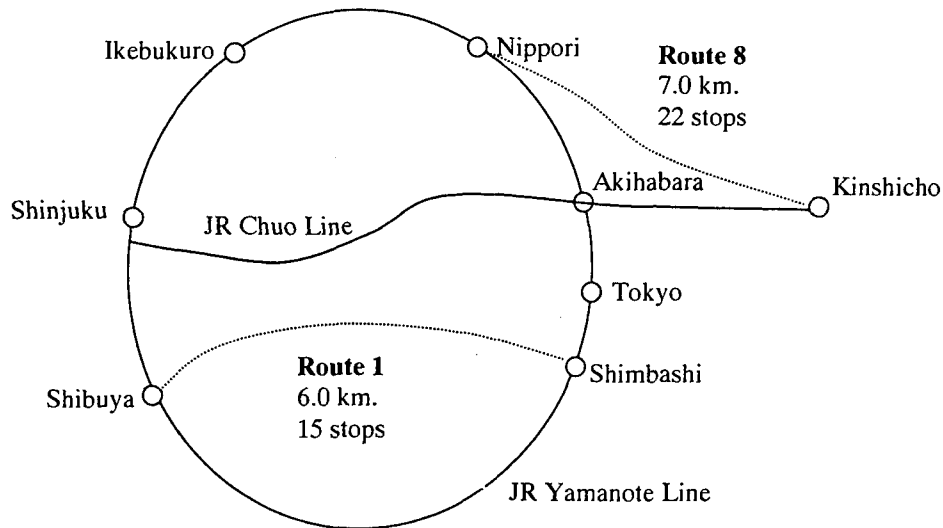
Fig1 Route map of route 1 and 8
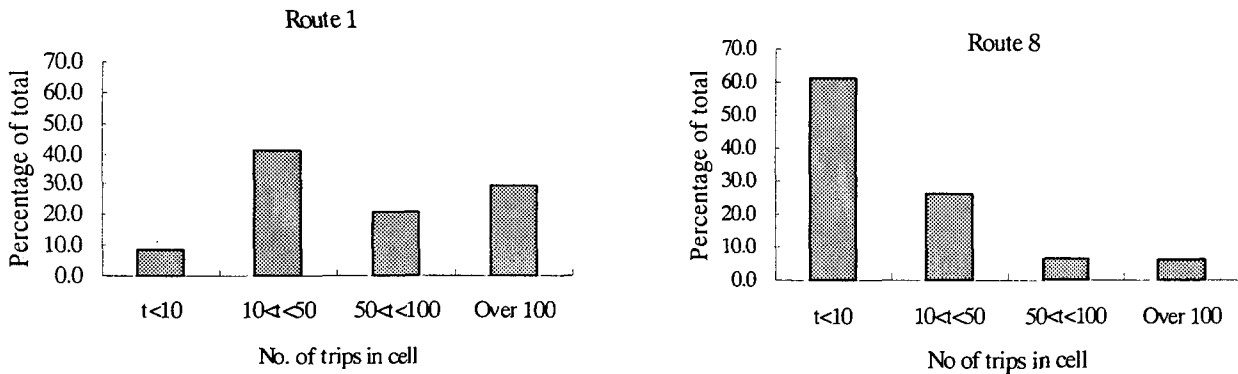
Fig.2 Classification of cells in the OD matrices in 1991 based on the number of trips

3. Data and methodology

(1) Description of data

Data, utilized in this study comprises of two major sources. First is one-day OD matrix, derived from the bus survey conducted every five year by the Tokyo metropolitan bus operator. The other data set is sampling ride-check, which is automatically collected by the BDCS system installed in 10 percent of the total bus fleet. It is a main purpose of this study that data to be used should be readily available so that operator can cut unnecessary cost of data collection for the service planning. The five-year bus survey provides us an actual daily OD matrix of that year. In this study we use the OD matrices in 1991 and 1996 of two bus routes in the Tokyo Metropolitan, namely route number 1 and 8 respectively. The daily OD matrices in 1991 are used as base year matrices, while the daily OD matrices in 1996 are used as the target matrices for estimations. The route 1 and 8 are comprised of 15 and 22 bus tops respectively. Distance between stops along each route is presented above in figure 1. Figure 2 shows the share of cells in the OD matrices in 1991 with different number of trips of route 1 and 8 respectively.

As shown above in figure 2, cells in the OD matrices are classified into four groups according to the number of trips. It is apparent that structure of the two matrices is considerably different. The OD matrix of route 1, which is shorter in distance, is dominated by cells with considerable number of trips, while that of route 8 mainly comprises of cells with few trips. This difference is attributed to the fact that route 1 runs through several major stops within Yamanote loop of Tokyo, while route 8 runs through comparatively less built-up area from Ueno to Kinshicho. We shall discuss later how difference in the structure of OD matrices affects the OD estimation. It is also important to compare the daily OD matrices surveyed in 1991 with those in 1996, as the extent of change in OD significantly influences the accuracy of OD estimates.
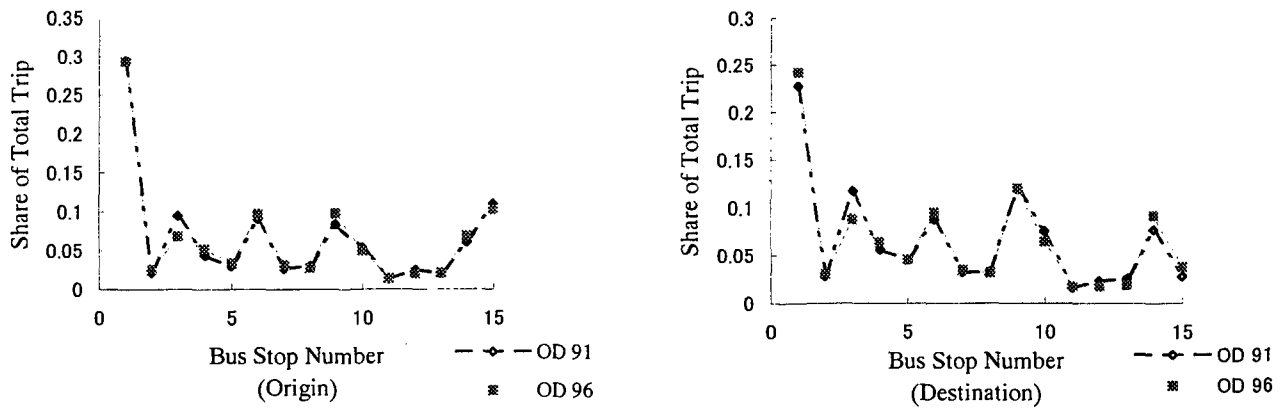
Fig.3 Share in proportion of total of originated trip and attracted trip by bus stops of route 1
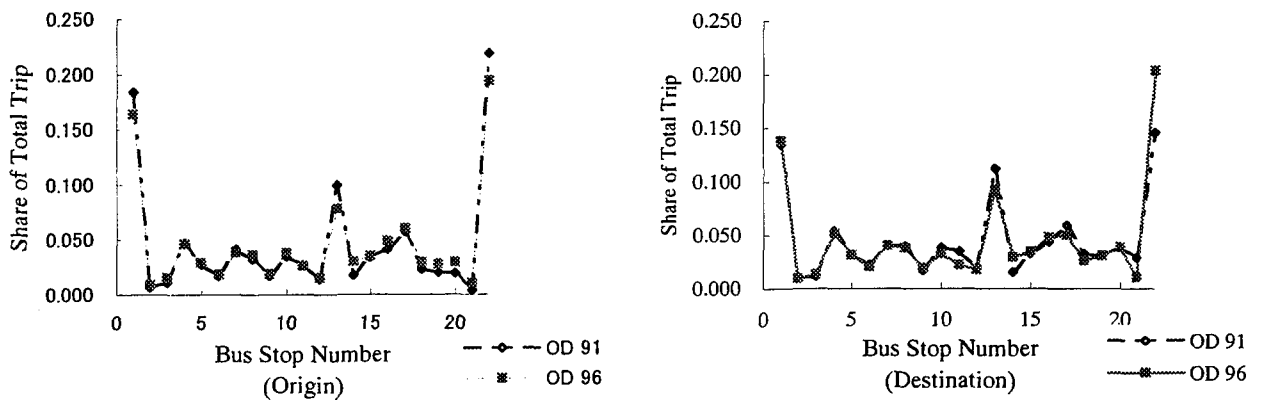

Fig.4 Share in proportion of total of originated trip and attracted trip by bus stops of route 8

Figures 3 and 4 present the pattern of OD matrices of route 1 and 8 in 1991 and 1996 in terms of share of each origin and destination in total trip respectively. It is clear that, except some change in share of stop no. 3 of route 1 and stop no. 13 of route 8, OD pattern of the both years is almost identical. The change in share of the stops may result from additional development nearby the stop or from change in the position for short turning of some buses. In addition to the above change, there was only minor change in total demand of the both routes between 1991 and 1996. The daily demand of route 1 decreased from 37,588 to 35,231 passengers, equivalent to 6 percent decrease in number of passengers. On the contrary, the patronage of route 8 increased by 3 percent from 12,202 to 12,626 passengers daily.

Sampling ride-check, which was obtained by the BDCS system, was collected in 1998 from the buses equipped with the system. Each data set contains ID number of the bus, running pattern, date and time of operation, ID number of the bus stops and the number of boarding as well as alighting passengers at each stop. Such data is classified according to the time of operation. Since collected data is more than the actual number of bus runs, the data is randomly picked out in order to match up it with the actual operation. As a result, the sampling data is assumed to represent a daily ride-check of the bus route.

(2) Methodology

In this study, three estimators namely the ME, ML and GLS estimators, obtained by using the base year OD matrix in 1991 in conjunction with constraints of the number of generating and attracting trips by each stop in 1996 are tested against the actual OD matrix in 1996. It should be noted here that in this study, we neglect variation of the OD so that the available OD matrices in 1996 can be treated as the actual OD matrices. Effects of sample size on the accuracy of the estimates are also examined by randomly sampling the base year OD at different sizes of 20, 25, 30, 40 and 50 percent of the original size respectively with 1000 repetitions. By using the Bootstrap technique, mean and variance of each estimate can be determined and compared with the actual value in terms of accuracy.

In order to compare the performance of different models in terms of accuracy of the OD estimates, we use the statistical measures shown below.

Root-mean square error:

$$RMSE = \sqrt{\frac{1}{N} \sum_{ij} (\hat{t}_{ij} - t_{ij})^2} \tag{14}$$

Root-mean square difference:

$$RMSD = \sqrt{(1/K)\sum_{ij}\left((\hat{t}_{ij} - t_{ij})/t_{..}\right)^2} \qquad (15)$$

Average weighted fractional error:

$$AWFE = \sqrt{\sum_{ij}(t_{ij}/t_{..})\left((\hat{t}_{ij} - t_{ij})/t_{ij}\right)^2} \qquad (16)$$

Weighted trace

$$WT = \sum_{ij}\frac{t_{ij}}{t_{..}}Var(t_{ij}) \qquad (17)$$

where
$\hat{t}_{ij}$ and t_{ij} = the estimated and actual number of passengers for particular OD pair.

$t_{..}$ = the total number of trips in the actual OD matrix.
N = the number of cells in the actual matrix
K = the number of nonempty cells in the estimated matrix.

IO estimators are determined by applying IO method to the sampling ride-check. As the available ride-check data is collected from only 10 percent of total buses, which are equipped with the automatic counters, some running patterns are missed out. Therefore, it is not surprising that its OD pattern is different from that of the daily OD survey. However, in order to assess the accuracy of estimate, IO estimate is compared with the actual OD matrix based on the trip length distribution.

4.   Results

Table 2: Statistical measures of the OD estimates of route 1 based on full data

Technique	RMSE	RMSD	AWFE
ME	35.43	10.4E-4	0.182
ML	35.40	10.4E-4	0.183
GLS	35.90	10.5E-4	0.184
Growth factor	96.40	28.3E-4	0.289

Table 3: Statistical measures of the OD estimates of route 8 based on full data

Technique	RMSE	RMSD	AWFE
ME	8.89	7.5E-4	0.369
ML	9.36	7.9E-4	0.394
GLS	9.00	7.6E-4	0.393
Growth factor	17.75	15.0E-4	0.517

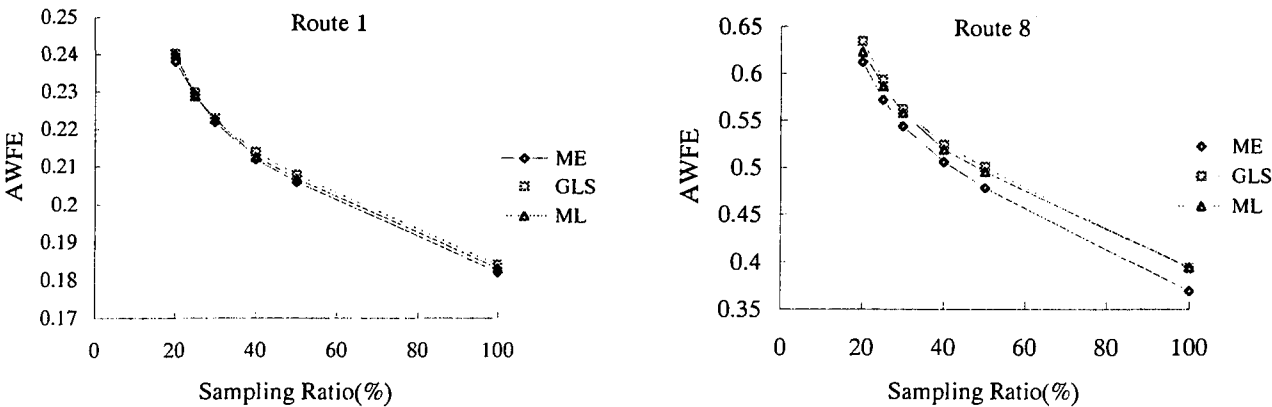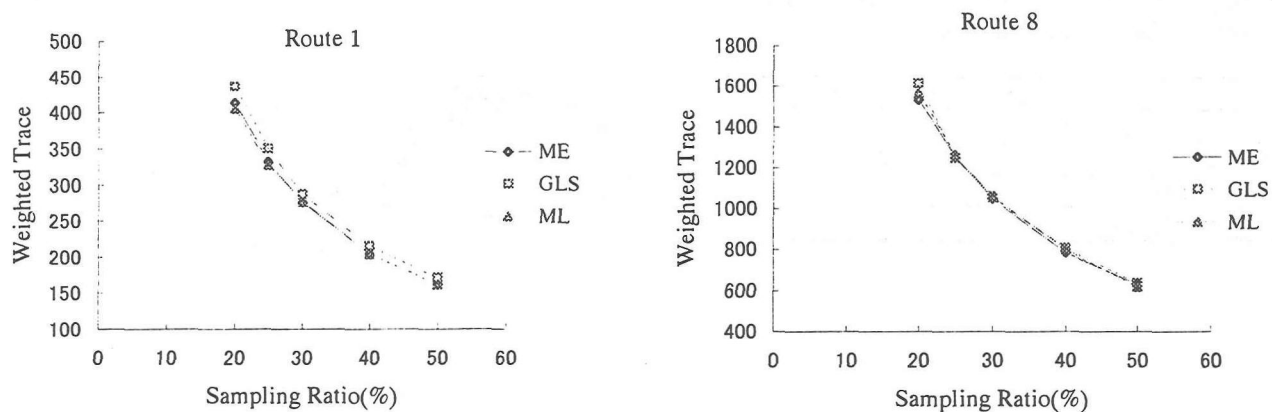**Fig.5** Change of AWFE of the estimates with size of sampling ratio

Fig.6 Change of weighted trace of the estimates with size of sampling ratio
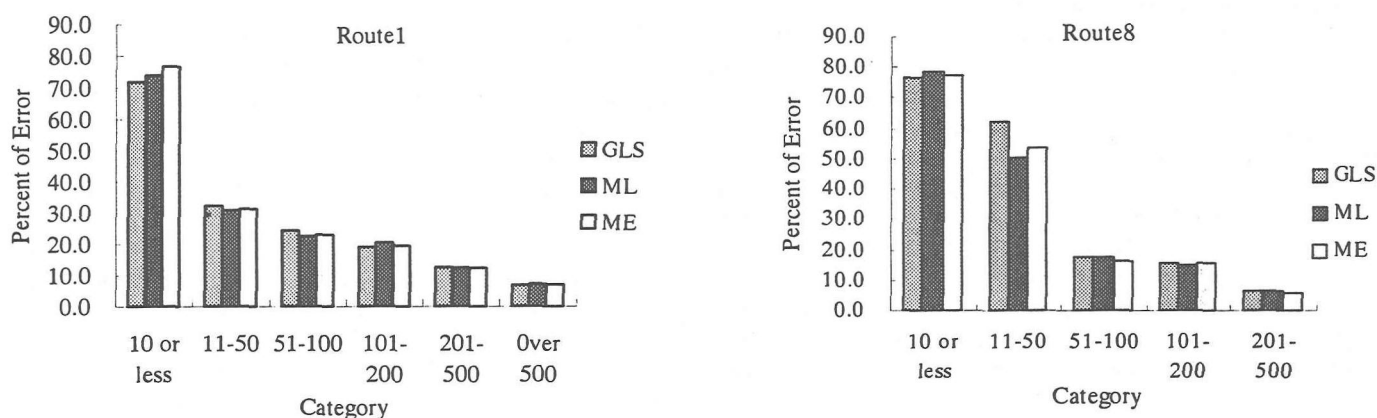

Fig.7 Percent of error of estimates classified by number of trips in cell based on full data

(1) Accuracy and effects of sampling ratio

Based on the estimation results, comparison of the performance of the models, as expressed by the statistical measures, can be summarized in table 2 and 3. As evidently found elsewhere for the case of test network, the ME and ML methods perform equally well even with the big sample data due to its similarity in mathematical formulation. Since it has been concluded in the study by Ben-akiva et al.[2] that the results obtained by the ME and ML methods are comparable when the sample size is small, we shall generalize here that one may choose either ME or ML technique to estimate the OD matrix. However, due to its computational simplicity, the ME technique appears to be more preferable to the ML method, if the assumption of deterministic constraints on on-off counts is taken. GLS method also gives comparable result with the above methods without requiring iterative calculations. However, the difficulty of this method lies in finding the inverse of variance-covariance matrix of the base year OD matrix, especially in the case of route 8, in which the ratio between the number of trips in biggest and smallest cell is large. Comparing the results obtained from the ME, ML and GLS techniques with that of growth factor method, benefits of using on-off counts data for OD estimation can be appreciated more clearly.

In figure 5, we show that by using the OD estimation techniques based on sufficient OD sample under constraints of on-off counts we can obtain the result within acceptable region of confidence. As shown by the statistics, sampling ratio of 50% or so, when used with the on-off counts already gives adequately desirable results. For the both routes, using the ME method with 50% sampling ratio gives better result than using the growth factor method. Comparing the results obtained from route 1 and 8, we may see that the effects of sample size on the accuracy are more visible for the case of route 8, as reducing size of the sample by half from 100% to 50% of the original size results in 10% and 30% more error of the estimates for route 1 and 8 respectively. We shall examine in detail below why this becomes the case.

Figure 6 presents the change of weighted trace of estimates with size of sampling ratio of priori OD matrix. We can see the consistent trend of growing rate of increase in variance of estimates as the size of sampling decreases. This means the effect of reducing the size of sampling ratio is two-fold. First, as previously seen from figure 5, it results in decrease in accuracy of the estimates. The other effect is it progressively decreases the degree of confidence of the estimates.

It has become clear from the above results that the performances of the three models are insignificantly different. This is attributed to the similarity in characteristics of the models from statistical viewpoint. All of the three models are based upon expansion of outdated OD with row and column constraints. The minor difference among them lies on the way in which row and column constraints are satisfied. Similarity in pattern between OD data in 1991 and 1996 may be another reason. Due to limitation on availability of data, we could not make concrete conclusion. Anyway there is a reason to believe that even if the difference between base year and present year OD pattern is more visible, the performances of the models are still similar. As we reduce sampling ratio to 20 percent but there was no sign of deviation in accuracy of the obtained results.

Table 4: Degree of correlation between the predicted and actual trip length distribution

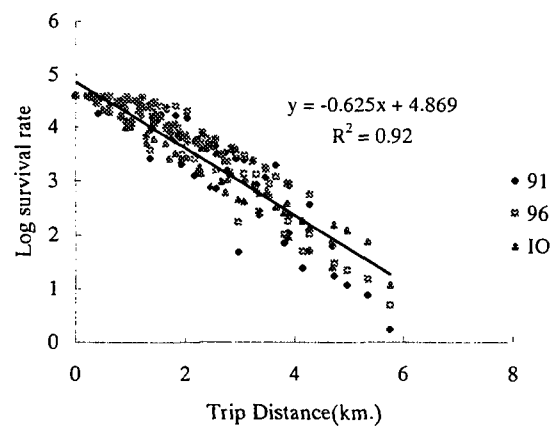| Case	Degree of correlation (R^2)
Route1/for Shimbashi	0.92
Route1/for Shibuya	0.50
Route8/for Kinshicho	0.91
Route8/for Nippori	0.78

bounding for Shimbashi station

:r 1

(2) Effects of OD flows and its pattern

We then attempt to examine the accuracy of estimates of cells with different number of trips. We classify OD pairs into groups according to the OD flows under the hypothesis that cells with large OD flows could be estimated with higher accuracy due to lower level of variance of the sampling counts. As shown in figure 7 above, although there is some minor variation, we may be able to generalize that the accuracy of estimates does not significantly change with the estimation method used. On the other hand, accuracy of estimates is dependent principally on the OD flows. This is a reason why the accuracy of estimates of route 8 is lower and rate of decrease in accuracy is faster than that of route 1, as the property of the ME and ML techniques is that empty cells of the base year matrix always results in zero estimates. In addition, reduction of the sampling size means coefficient of variation of sample is bounded to increase. Hence optimal sample size is in turn dependent on the pattern of OD matrix. In general, sufficiently large sample size would be necessary for a matrix, which is dominated by relatively small number of trips in each cell like the OD matrix of route 8 in this study, as shown in figure 1. This is attributed to the fact that reducing sampling size of the matrix with smaller OD flows has more visible effect. Moreover, by nature accuracy of estimates from a priori matrix with smaller OD flows is always less desirable. From figure 7, we propose that sampling ratio could be reduced as long as the OD flows are still large enough that majority of them do not fall into category, whose the accuracy is below desired level.

(3) Applicability of the estimation technique without outdated OD survey

Now let us investigate the applicability of the IO method under different OD pattern. Despite of its apparent drawback namely accuracy of the estimates intimately depends on the demand pattern, the IO method is useful in the sense that it requires only on-off counts for OD estimation. In addition, it is useful for estimating OD matrix for several periods of day, in which OD patterns typically vary from the pattern of average daily OD. Here, based on on-off counts, which are available from the BDCS system, we attempt to identify the applicability of the method. We assume that logarithm of survival rate should be linearly dependent on the trip distance. The detailed result of route 1 bounding for Shimbashi station is shown in figure 7. For the case of other route or directions, we present only the degree of correlation of the predicted trip length distribution with the actual value in 1996.

As shown in figure 8, trip length distribution in 1991 and 1996 is slightly different. Thus it is no wonder that the predicted trip length distribution, which is obtained from applying the IO method to the BDCS data in 1998, to some extent varies from the actual values in 1991 and 1996. However, as summarized in table 4, the result is satisfactory given that the degree of correlation between the predicted distribution and the actual one in 1996 for two cases is as high as 0.9. Meanwhile, degree of correlation of the worst case is as low as 0.5. It would be more desirable if we had the actual OD matrix in 1998 in order to make more solid conclusion. We may conclude in general here that results of IO method are highly dependent on the OD pattern and special care should be taken in applying it.

(4) Cost effectiveness evaluation

Finally, based on the above results we introduce criteria for determining applicability of the OD estimation techniques, which are investigated in the study. Table 5 illustrates major factors, which should be taken into account. Requisites in terms of data, labor and equipment cost and computational complexity as well as desired level of accuracy are expressed quantitatively in comparison among the studied methodologies.

Data, required for the estimation OD matrix, is either previous OD or on-off counts. Conventionally, OD survey is conducted periodically (for example, just one typical weekday every five year) for use as a representative OD matrix of particular period. However drawbacks of the approach based only on OD survey(s) are high cost of survey and low effectiveness of data utilization. Effectiveness of data utilization here means applicability of data to produce useful information, which cannot be grasped from the data itself. On the other hand, combining outdated OD survey with on-off counts brings about higher effectiveness of data utilization since OD matrices for different time period can be estimated with negligible additional cost. Furthermore this approach enables bus service to be designed to cope with day-to-day or seasonal variation in travel demand. One point to be noted on this approach is that it requires high initial investment for some equipment such as automatic counters and computer system, which may be depreciated in the long term. Concerning accuracy, it was proved that the OD

Table 5: Criteria for evaluating OD estimation techniques

Estimation technique	Requisites				Accuracy
	Data	Labor Cost	Equipment cost	Computational complexity	
ME	Outdated OD On-off counts	Moderate	High	Moderate	High
ML	Outdated OD On-off counts	Moderate	High	High	High
GLS	Outdated OD On-off counts	Moderate	High	High	High
IO	On-off counts	Low	High	Moderate	Low-moderate
Growth factor	Outdated OD Fare box count	High	Low	Low	Moderate

estimation based on outdated OD and on-off counts give acceptable results when compared to the actual value obtained from OD survey. In conclusion, BDCS is effective in that it provides on-off count data automatically in any time period, which can give higher effectiveness of data utilization with acceptable accuracy and affordable cost, if coupled with outdated OD matrix data.

Thus, decision to adopt any estimation technique should be reflected by differences in constraints and needs of each organization. For instance, an organization in developed country, which already installed equipment for automatic data collection, would consider ME or IO as alternatives for OD estimation depending on required accuracy and availability of outdated OD. On the other hand, their counterpart in developing country, in which budget to purchase new equipment is limited yet labor cost is comparatively low, would find growth factor method desirable and affordable. It should be noted that, in table 5, we assume that labor cost is accrued solely from conducting OD survey and equipment cost is only for purchasing automatic passenger counter. On-off counts can alternately be obtained through on-board ride checks, resulting in higher labor cost but lower equipment cost. Thus it is also feasible for an organization without BDCS system to use the OD estimation technique, which requires on-off counts for more accurate results.

5. Conclusion

This paper presents some empirical evidences on the performance of different OD estimation techniques based on on-off counts and/or the outdated OD matrices and its potential to replace conventional OD survey. The ME method is proved to be the most practical technique for the OD estimation problem under the assumption of deterministic count constraints due to its computational simplicity and desirable accuracy of the estimates, provided that priori OD matrix is available. Effects of sample size on the accuracy tend to be trivial given that most of the OD flows are large enough, say over 100. Although the results for the case of route 1 and 8 are considerably different in terms of accuracy, at this point it is clear that, given the availability of the on-off counts, expenses required for the service planning could be saved more or less depending on availability of priori OD matrix. If there exists a priori OD matrix, it is proved that based on the priori OD pattern, on-off counts can be used for OD estimation with acceptable accuracy. On the other hand, if that is not the case, conventional daily OD survey could be replaced by an OD sample without losing much accuracy as shown from the result of route 1 in this study, whose OD flows between pairs of stops are mostly large. Change in OD pattern between base year and present year OD matrices results in lower accuracy of OD estimates. However, as empirically found in this study, comparative performance of the three OD estimation techniques is independent on such change.

BDCS data also shows high potential for use in the route-level OD estimation as revealed in our finding that for some cases predicted trip length distribution, based on the IO method, is highly correlated with the actual one, with degree of correlation as high as 0.9. However, given that a priori OD matrix is available, there would be no reason to use the BDCS data alone for OD estimation, since it is not always the case that trip length distribution of trips originated from different stops in a route is identical. In other words, BDCS data should be used in conjunction with priori OD matrices to avert the necessity to conduct costly conventional on-board OD surveys.

Criteria for choosing OD estimation technique are proposed. Required data, costs, computational complexity and desired accuracy are the factors to be taken into consideration. Given differences in constraints in terms of budget or labor cost and in desired accuracy, there is no single best choice for every organization. Consequently, clear-cut policy is needed before any decision to adopt an OD estimation technique shall be made.

References

1) Fielding P. : Obtaining Bus Transit Service Consumption Information through Sampling Procedures, in New Survey Methods in Transportation; 2nd international conference, E.S. Ampt, A.S. Richardson and W. Brog, Utrecht, the Netherlands, pp. 227-239, 1983.
2) Ceder A. : Bus Frequency Determination Using Passenger Count Data, Transportation Research Vol. 18(A), pp. 439-453, 1984.
3) Ben-Akiva, M.E., Macke, P.P. and Hsu P.S. : Alternative Methods to Estimate Route-Level Trip Tables and Expand On-Board Surveys, Transportation Research Record 1037, TRB, pp. 1-11, 1985.

4) Cascetta E. and Nguyen S. : A Unified Framework for Estimating or Updating Origin/Destination Matrices from Traffic Counts, Transportation Research Vol. 22(B), pp. 437-455, 1988.
5) Willumsen L.G. : Estimation of an OD matrix from traffic counts: a review. Working Paper 99, Institute for Transport Studies, Leeds University, 1978.
6) Landau U., Hauer E. and Geva I. : Estimation of Cross-Cordon Origin-Destination Flows from Cordon Studies, Transportation Research Record 891, TRB, pp. 5-10, 1982.
7) Mcneil S. : Quadratic Matrix Entry Estimation Methods, Ph.D. thesis, Department of Civil Engineering, Carnegie-Mellon University, 1983.
8) Tsygalnitsky S. : Simplified Methods of Transportation Planning, thesis presented to MIT, 1977.

USE OF ON-OFF COUNTS FOR OD ESTIMATION
AN APPROACH TOWARDS MORE COST-EFFECTIVE BUS SURVEYS[*]

By Terdsak RONGVIRIYAPANICH[**], Fumihiko NAKAMURA[***] and Izumi OKURA[****]

Several OD estimation techniques are evaluated based on its accuracy under different sampling ratios by using real data from two bus routes operated in Tokyo metropolitan area. Given that prior OD matrix is available, it was found that the Maximum Entropy method could be applied to update OD matrix with desirable level of accuracy as long as most of OD flows in the OD sample are still large enough. Without priori OD matrix, the intervening opportunity method also offers reasonable estimation when applied to BDCS data. It is clear that more effective utilization of available data, for instance BDCS data or priori OD matrices can help bus operators save planning expenses. Some criteria were proposed for determining applicability of each OD estimation technique based on differences in constraints and needs of planning side.

バス乗降データを組み入れたバスＯＤ推計手法の改善手法に関する研究
より費用効果の高いバス交通計画調査に向けて[*]

ターサック　ロンビリヤパニチ[**]・　中村　文彦[***]・　大蔵　泉[****]

本研究は、各バス停での乗降人数を自動集計する装置（BDCS）を有する東京都交通局の２つのバス路線を対象として、同データを活用した何種類かのＯＤ推計手法の精度を比較評価したものである。現在活用不十分な BDCS データを組み入れてＯＤ表推計手法の向上が課題である。分析の結果、過去のＯＤ表が与件の場合は、過大推計の部分はあるが、エントロピー最大化法によるＯＤ表更新が適用可能とわかった。過去のＯＤ表が与件でない場合は、機会介在法が有用であることがわかった。また、BDCS や過去のＯＤ表夕等利用可能なデータ源をより効果的に活用することで、事業者の調査計画費用の節約にも貢献することが明らかになった。また、計画サイドの制約条件や必要性の違いに基づいて OD 推定手法の適用可能性を判断する基準を提案した。