APPLICABILITY OF TRIP DISTRIBUTION MODELS

By Tomonori SUMI* and Masao KUWAHARA**

1. INTRODUCTION

The estimation of trip distribution is an important step of the well-known procedure to forecast travel demand. For this purpose, two types of methods are available. One is the present pattern method and the other is the application of mathematical models. Certain kinds of Gravity models and Opportunity models belong to the latter and they are frequently used.

It is practically quite impossible to specify demand functions or utility functions of almost all demand models of transportation field. So alternatively a priori forms of these functions are specified.

These functions above are also based on simplified assumptions regarding human behavior and urban activities, though they are more complicated¹⁾⁻³⁾. Then necessarily their validity must be empirically verified.

Up to this time, these models were examined from the view of the applicability in some papers4)~8) but they are limited in the accumulation of empirical results. For that reason, this report examines the applicability of the models by looking at various cases in a methodical way, and supplies a more confirmed stand point in actual usage. The applicability of the models is examined in the following aspects: The first is their reproducibility. The parameters included in the mathematical expressions of the models are calibrated using the data obtained from surveys. Subsequently trip distribution is again estimated using the obtained parameters. The reproducibility means the similarity of the estimated results to the observed distribution and it is a fundamental criterion to evaluate the performance of the models. The second is the manipulability. The concept of "manipulability" is complicated, however here, the effect of the sample size on the estimation is only examined because one of the main difficulties lies in the data collection. Those models that require less samples are more desirable. The third is the sensitivity. The sensitivity means the changing rate of the estimated results in accordance with the variation of the data handling such as the calibration procedure or zoning composition. The sensitivity is related in some degrees to the manipulability, however in this paper they are dealt separately. It is desirable that models should be insensitive. In the latter, the sensitivity resulting from erroneous parameters calibrated and that of variations between zones are discussed.

This report examines five trip-distribution models, which are frequently used, in the application of the same way to the data of six person-trip surveys in Japan and the Philippines, i.e. Hiroshima (1967, 1978), Maebashi (1978), Okayama (1972), Manila (1970) and Davao (1979).

2. MODEL AND DATA HANDLING

(1) Models

where.

The following five trip-distribution models are examined.

^{*} Member of JSCE, Associate Professor, Kyushu University

^{**} Graduate Student, University of Tokyo

 t_{ij} : estimated trips from zone i to zone j

 O_i ; trips produced in zone i

 D_j ; trips attracted in zone j

 C_{ij} ; intra- and inter-zonal travel time

 V_{j-1} ; total number of trips attracted in zones which are considered nearer than zone j from zone i

 V_j ; $V_{j-1}+D_j$

 $K, L, \alpha, \beta, \gamma$; parameters.

Models 1 and 2 do not satisfy the two constraints represented by Eqs. (8) and (9). On the other hand, model 3 satisfies the production-constraint represented by Eq. (8), while model 4 satisfies both constraints.

(production-constraint)

$$O_i = \sum_j t_{ij}$$
(8)

(attraction-constraint)

$$D_j = \sum_i t_{ij}$$
(9)

Therefore the number of trips calculated with models 1, 2 and 3 are modified by the Furness iterative procedure⁹⁾ in order to satisfy the above constraints.

(2) Data handling

The models are applied to the journey-to-work trips by road in the way mentioned below.

- Step 1: Journey-to-work trips by road, that is, trips from homes to workplaces by cars, buses, bicycles and motorcycles, are selected from original data of each person trip survey. The selected data are called as the base data in the following.
- Step 2: The mean travel time $\{C_{ij}\}$ and the observed trips $\{T_{ij}\}$ between the two of the zones are calculated from the base data. Intraand inter-zonal distributions are both taken into account.
- Step 3: The model's parameters are calibrated using $\{C_{ij}\}$ and $\{T_{ij}\}$. The values of the parameters which give the minimum of the residual sum of squares, are called as the optimal values. Note that zero elements in $\{T_{ij}\}$ matrix are not used to calibrate the parameters.
- Step 4: The estimated trip distribution $\{t_{ij'}\}$ is calculated using the optimal values of the parameters.
- Step 5: $\{t_{ij}'\}$ are modified by the Furness procedures so that $\{t_{ij}\}$ is obtained satisfying the production and attraction constraints, except for model 4.

(3) Statistics for evaluation

a) Reproducibility

The reproducibility is evaluated using the statistics shown as follows.

(Correlation Coefficient)

$$R = \frac{\sum \sum (t_{ij} - \bar{t})(T_{ij} - \bar{T})}{\sqrt{\sum \sum (t_{ij} - \bar{t})} \sqrt{\sum \sum (T_{ij} - \bar{T})}} \cdots \cdots (10)$$

(Chi-square)

$$\chi^2 = \sum_{i} \sum_{j} (T_{ij} - t_{ij})^2 / t_{ij} \qquad \cdots \qquad (11)$$

(Relative Error Index*)

where

 t_{ij} ; estimated trips from zone i to zone j

 T_{ij} ; observed trips from zone i to zone j

 \bar{t} ; average number of $\{t_{ij}\}$

 \bar{T} ; average number of $\{T_{ij}\}$

b) Effect of sample size

To examine the effect of the sample size, the following two questions are discussed.

- (1) Which model is least affected by the reduction of sample size?
- (2) How many samples does the best model selected in the above require?

Each model is applied to the data of various sample sizes obtained from sampling again from the base data randomly. The procedure of calculation is as follows.

- (1) The sampling rate (s) is selected in the range from 1% to 100%. Then the samples of reduced number are obtained from the base data by random sampling without replacement. The trip distribution $\{T_{ij}^{s}\}$ is determined from the sampled data.
- (2) The optimal parameters are obtained using $\{T_{ij}^{*}\}$ and $\{C_{ij}^{*}\}$ according to the procedures in the step 3 of 2.(2).
- (3) The estimated trip distribution $\{t_i _j^{s''}\}$ are calculated using the Eqs. (1) to (7).
- (4) $\{t_{ij}^{s''}\}$ is modified by the Furness procedure except model 4, so that $\{t_{ij}^{s'}\}$ satisfies the following marginal constants:

$$O_i^s = \sum_j t_{ij}^{s'} \qquad (13)$$

$$D_j^s = \sum_i t_{ij}^{s'} \qquad (14)$$

(5) $\{t_i^s\}$ is multiplied by the reciprocal of the sampling rate (s), so that the total amount of the estimated trip distribution coincides with that of the observed trip distribution, then $\{t_{ij}^s\}$ can be obtained.

$$t_{ij}^{s} = t_{ij}^{s'}/s$$
(15)

^{*} The Relative Error Index as defined as above is convenient to compare the errors produced in the application to various data of different sample sizes because it represents the absolute value of the dimensionless sum of square root of residuals.

The calculation mentioned above was made once for every sampling rate.

The statistics used to evaluate the effect of the sample size are Relative Error Index, (RE) Correlation Coefficient (R) and Chi-square (χ^2) defined as follows.

$$R = \frac{\sum \sum (t_{ij} - \bar{t}^s)(t_{ij} - \bar{t})}{\sqrt{\sum \sum (t_{ij} - \bar{t}^s)^2} \sqrt{\sum \sum (t_{ij} - \bar{t})^2}} \dots \dots \dots (16)$$

$$RE = \sqrt{\sum \sum (t_{ij} - t_{ij})^2} / \sum \sum t_{ij} \dots \dots \dots \dots (17)$$

$$\chi_s^2 = \sum \sum (t_{ij} - t_{ij}^s)^2 / t_{ij} \dots \dots \dots \dots \dots \dots (18)$$

c) Sensitivity

(1) Effect of calibration error

The calibration error of the parameters is simulated by giving a conscious deviation from the optimum values. Although models 1 and 2 include plural parameters, all parameters in the models 1 to 4 except α and depend on α when the Furness procedure is used to modify $\{t_{ij}\}^*$. Therefore the calibration error of parameter α in the models 1, 2, 3 and 4 is only taken account for. Parameter L of model 5 has the dimensions of (1/trips) and it is determined depending on the number of data. Therefore the calibration error

The final matrix $\{t_{ij}\}$ is not affected, even when arbitral columns or rows of $\{t_{ij}'\}$ are multiplied by any positive numbers. Therefore $\{t_{ij}\}$ is dependent only on $\{C_{ij}^{-i}\}$ in Models 1 to 4.

is expressed as the deviation of quantity "LT" where T is the total number of the observed trips.

The Chi-squares of Eq. (19) are used to test the sensitivity of parameters produced by the models statistically.

$$\tilde{\chi}_s^2 = \sum \sum (t_{ij} - \tilde{t}_{ij})^2 / t_{ij} \qquad \cdots (19)$$

where

 t_{ij} ; trips calculated using the deviated parameter

 $ilde{t}_{ij}$; trips calculated using the optimal parameter

(2) Effect of zoning composition

The sensitivity of the results to the difference in zoning composition is examined by comparing the estimated results obtained from the various zonings of particular areas.

3. RESULTS AND DISCUSSIONS

(1) Reproducibility

Table 1 shows the parameters obtained from the calibration. Table 2 shows the statistics obtained from the results. Models 3 and 4 yielded the almost best values of Correlation Coefficients, Relative Error Indices and Chi-squares for all surveys. Although Model 2 have two additional parameters in comparison with model 1, it gave almost the same values of the statistics as those given by model 1, and models 1 and 2 are the worst of the five models. Therefore it is supposed that the introduction of the marginal constraints given by Eqs. (8) and (9) in the models improves

Table 1 The optimal parameter values of each model in six cities.

Model No.	1		2				3	4	5	Population	Area	Number	
Parameter	k	α	k	α	β	r	α	α	L (×10-4)	(×10³)	(km²)	of Zones	
Hiroshima (1967)	0.000 201	0.650	4.21	0.418	0.294	0.195	2.82	2.89	9.64	750	430	111	
Hiroshima (1978)	0.007 35	1.01	2.06	1.11	0.429	0.475	3.04	2.92	6.70	1 080	720	40	
Maebashi	0.252	2.63	95.7	2.68	0.442	0.594	3.67	3.42	5.22	1 010	1 309	40	
Okayama	0.287	2.32	28.8	2.04	0.386	0.544	3.11	3.34	9.39	1 017	1 159	60	
Manila	0.020	1.27	2.08	1.16	0.425	0.538	2.69	2.70	4.52	3 890	594	51	
Davao	0.001 57	0.555	0.0758	0.650	0.586	0.691	0.876	1.17	5.41	360	137	21	

Table 2 The statistics of each model in six cities.

Statistical Indices Chi-square (χ²) (×104)					Re	lative	Erro	r (RE	E)	Correlation Coefficient (R)						
Model No.	1	2	3	4	5	χ2-05	1	2	3	4	5	1	2	3	4	5
Hiroshima (1967)	3.04	3.38	3.28	3.39	5.82	1.258	3.16	3.26	2.26	2.26	2,60	0.68	0.64	0.84	0.84	0.78
Hiroshima (1978)	0.689	0.634	0.404	0.404	0.593	0.169	4.01	3.92	2.72	2.70	3.75	0.74	0.76	0.89	0.89	0.77
Maebashi	0.680	0.650	0.580	0.530	2.44	0.169	3.06	2.99	2.47	2.41	3.32	0.93	0.93	0.94	0.94	0.89
Okayama	1.01	1.24	0.820	0.880	1.13	0.374	3.69	4.17	2.81	2.75	3.67	0.93	0.91	0.94	0.94	0.88
Manila	0.648	0.706	0.434	0.434	0.591	0.272	2.91	2.98	2.23	2.20	2.78	0.78	0.76	0.87	0.87	0.78
Davao .	0.198	0.189	0.175	0.172	0.184	0.049	6.20	6.13	5.92	5.90	6.03	0.78	0.78	0.80	0.80	0.79

^{*} The Furness procedure can be written as follows: $t_{ij} = r_i s_i t_{ij}', \quad r_i = \frac{O_i}{\sum_i s_i t_{ij}'}, \quad s_j = \frac{D_j}{\sum_i r_i t_{ij}'}$

the reproducibility more effectively than the additional parameters of model 2. Model 5 gave the intermediate values of the statistics.

(2) Effect of sample size

Figs. 1, 2 and 3 show the relationships between the sampling rates (s) and the goodness-of-fit statistics as an example in Okayama among the six surveys. The statistics given by models 3

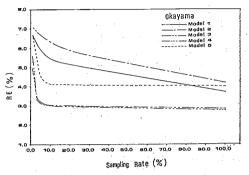


Fig. 1 Relationship between sampling rate and Relative Error Indices in Okayama.

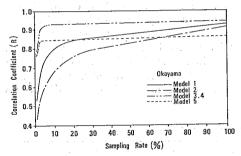


Fig. 2 Relationship between sampling rate and Correlation Coefficients in Okayama.

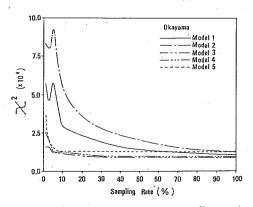


Fig. 3 Relationship between sampling rate and Chi-squares in Okayama.

and 4 are more desirable than those given by models 1 and 2 when the sampling rate is lower.

Fig. 4 shows the relationship between the sampling rate and the Chi-squares difined in Eq. (18). It indicates that the difference between the results derived from samples of considerably reduced size and those from the base data is not significant in the statistical sense. Especially models 3 and 4 are so robust that the difference of the results is not a serious problem even when the sample size is very small. Similar results were obtained for the other surveys. It is clear that models 3 and 4 are least affected by the sample size.

Fig. 5 illustrates the relationship between sample sizes and $\chi_1^2/\chi_{0.05}^2$ of Eq. (20) given by model 4 in application to six surveys.

$$\frac{\chi_s^2}{\chi_{0.05}^2} = \frac{\sum \sum (t_{ij} - t_{ij}^2)^2 t_{ij}}{\text{Chi-square value in contingency table with 0.05 significance level.}}$$

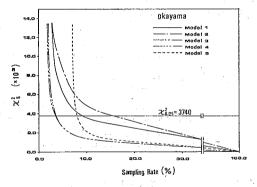


Fig. 4 Relationship between χ² and sampling rate in Okayama.

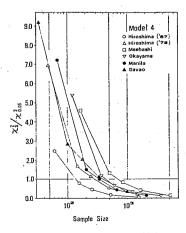


Fig. 5 Relationship between $\chi_s^2/\chi_{0.05}^2$ and the sample size in six cities for model 4.

When $\chi_s^2/\chi_{0.05}^2$ is less than 1.0, the significance level of the statistical fitness exceeds 5%. Even if the sample size is reduced to 500–600, the results derived by model 4 are virtually unaffected.

(3) Sensitivity

a) Effect of parameter deviation

Figs. 6 and 7 show the Chi-square values of each parameter for the selected range of the parameter values, defined by Eq. (18).

For all surveys and models, the ranges of parameters α and L which can satisfy the goodness-of-fit tests with 0.05 significance level, are

$$\alpha_{opt} \pm 0.8$$
 $L_{opt} \pm 2.0$

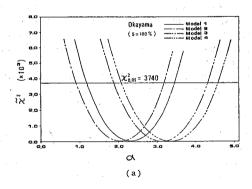
where.

 α_{opt} = the optimal value of α L_{opt} = the optimal value of L

It seems that the sensitivities of both parameters are small.

b) Effect of zoning composition

The models were applied to the base data and 10% sampled data of Hiroshima (1967) and Okayama (1972) under various zoning compositions. For example, the examined variation of



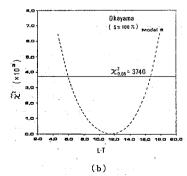
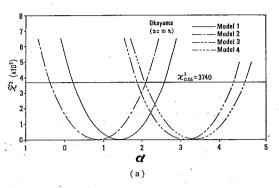


Fig. 6 $\bar{\chi}^2$ of each parameter for the selected range of the parameter values in Okayama, when the sampling rate is 190%.

the zoning in Hiroshima is shown in Fig. 8.

The calibrated values of the parameters are shown in Table 3. This table shows that the optimal values of the parameters for all models, especially models 3 and 4, are stable.



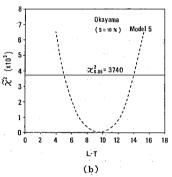


Fig. 7 $\tilde{\chi}^2$ of each parameter for the selected range of the parameter values in Okayama, when the sampling rate is 10%.

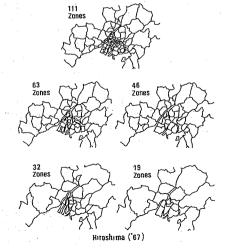


Fig. 8 Outline of the zoning composition in Hiroshima (1967).

Name The Number of Zones			Hiro	shima (1	.967)	Okayama					
		111	63	46	32	19	60	39	27 .	16	8
Model 1	T 1	0.650	0.689	1.290	1.270	1.660	2.320	3.510	3.800	4.080	4.220
Model 2	1 1	0.418	0.702	1.420	1.410	1.810	2.040	3.160	3.620	4.060	4.240
Model 3	α	2.82	2.99	2.91	3.25	3.38	3.11	3.11	3.22	3.80	3.33
Model 4		2.89	2.88	2.80	3.06	3.20	3.34	3.36	3.44	3.75	3.43
Model 5	LT	4.94	4.24	3.13	3,60	3.52	11.4	10.0	7.51	9.02	5.72

Table 3 The optimal parameter values of each model in Hiroshima (1967) and Okayama (1972) for different zoning composition.

4. SUMMARY AND CONCLUSIONS

In this report, five trip-distribution models were applied to journey-to-work trips by road obtained from the six person-trip survey data of five cities in order to examine the applicability of models.

The study is summarized as follows;

- (1) In order to examine the reproducibility, Correlation Coefficients, Relative Error Indices and Chi-squares were calculated. Models 3 and 4 gave almost the best values among the five models for all surveys.
- (2) In order to clarify the effects of the sample size on the estimated trip distribution, Correlation Coefficients, Relative Error Indices and Chi-squares between the estimated distribution derived from the base data and those derived from smaller samples were compared. The larger the sample size was, the better the goodness-of-fit statistics were obtained for all models. However, the difference between the estimated distribution derived from the base data and those derived from reduced samples of less than 20% is not significant in the statistical sense. Especially the estimated results derived from models 3 and 4 were least affected even when the sample size was reduced to 5% of the base data.
- (3) The effects of the parameter deviation from the optimal values on the estimated trip distribution were examined in six surveys. For all models, the variation of the estimated results in accordance with the deviation of the parameters was not significant within the following range of the parameter values:

$$\alpha_{opt} \pm 0.8$$
 $L_{opt} \pm 2.0$

This revealed that the sensitivities of the estimation to the deviation of parameters α and L were small. These sensitivities were also examined using smaller sample sizes, and it was also revealed that sensitivity did not depend on sample size.

(4) In order to examine the effect of zoning

composition, the models were applied to the survey data of Hiroshima (1967) and Okayama (1972) in various zoning systems. This revealed that the change of the calibrated values of the parameters were small unless the zoning composition was varied drastically.

Especially parameter α of models 3 and 4 was least affected of all.

We tried goodness-of-fit tests to examine the reproducibility (see Table 2). It revealed that none of the models satisfied our chosen level of the significance, even when a low significance level was selected.

Nevertheless, under consideration of a practical usefulness of the models, the models' characteristics, i.e. the statistically valid range of a parameter " α ", and the sensitivity of the parameter due to sample size and zoning composition, help planners to design better modelling schemes and to avoid excess efforts and expectations to the models, because the models themselves have a considerable limit of descripting human behavior.

With respect to journey-to-work trips by road we can conclude that models 3 and 4 are far superior to the others in actual usage. Additionally, sample size is possibly reduced to considerably small number when the models 3 and 4 are used.

5. ACKNOWLEDGEMENT

The authors are grateful to Prof. Yoshiji Matsumoto, Dept. of Civil Eng., Univ. of Tokyo for his advice on this matter, to Mr. Toru Tanioka, Japan Highway Public Corporation, for his assistance and to the members of the Laboratory of Transp., Dept. of Civil Eng., Univ. of Tokyo for their support of this research.

REFERENCES

 Wilson, A. G.: A statistical theory of spatial distribution models, Transportation Research, Vol. 1, pp. 253~269, 1967.

- Evans, S. P.: A relationship between the Gravity Model for trip distribution and the transportation problem in linear programming, Transportation Research Vol. 7, pp. 39~61, 1973.
- 3) Sasaki, T. and K. Kondo.: Present pattern method as an optimization problem, Proc. of JSCE, Note No. 727, pp. 139~142, April, 1978 (Japanese text).
- Heanue, K. E. and C. E. Pyers: A comparative evaluation of trip distribution procedures, Public Roads, Vol. 34, No. 2, pp. 43~51, June, 1966
- 5) Williams, I.: A comparison of some calibration techniques for doubly constrained models with an exponential cost function, Transportation Research Vol. 10, pp. 91~104, 1976.
- 6) Beardwood, J.: The space-averaging of deter-

- rent functions for use in Gravity Model distribution calculations, Transport and Road Research Laboratory Report LR 462, pp. 1~12, 1972.
- Kawakami, S.: Evaluation of trip forecasting methods from the point of work and school trips, Proc. of JSCE, No. 165, pp. 31~44, May, 1969 (Japanese text).
- 8) Beardwood, J. E. and H. R. Kirby: Zone definition and the Gravity Model; The separability, excludability, and compressibility Properties, Transportation Research Vol. 9, pp. 363~369, 1975.
- Evans, A. E.: Some properties of trip distribution methods, Transportation Research, Vol. 4, 1970.

(Received November 19, 1982)

土木学会論文報告集 第339号

分布交通量推定モデルの適用性

(角 知憲/桑原雅夫)

昭和58年11月

1. はじめに

交通需要予測の手順の1つである分布交通量の推定に は、現在パターン法とともに、重力モデルやオポチュニ ティモデルのあてはめが行われることが多い。これらの モデルは、人間の行動と都市のアクティビティについて のきわめて単純化された仮定に基づいて作成されたもの で,その論理的妥当性を比較することは不適当である. そこで, モデルの優劣は実用的な見地から論じるほかに なく、従来から、これらモデルの適用性の研究が行われ てきた. しかし、従来の研究は、特定の地域で行われた P. T. 調査のデータに、いくつかのモデルのあてはめを 行った単発的なものがほとんどで, モデルの適用性に関 する評価を一般化するためには、より多くの経験的蓄積 が必要である. そこで, 本報告では, 従来の研究の成果 を補強して、実務的により確かな情報を提供することを 目的とし、できるだけ多くの P.T. データを、統一的な 方法で取り扱って, 各種の分布交通量推定モデルにあて はめ、モデルの適用性を評価しようとするものである.

適用性は,①現象再現性,②サンブル数に対する依存性,③キャリブレーション誤差およびゾーニングの相違に対する感度,の3つの観点から評価する。あてはめに使用したデータは,広島(1967年,1978年),前橋(1978年),岡山(1972年),マニラ(1970年),ダバオ(1979年)の5都市6回のP.T.調査結果である。

2. モデルとデータの取り扱い

モデルは、よく使用される次の5つを取り上げた。 モデル 1 $t_{ij} = KO_i D_j C_{ij}^{-\alpha}$ (1) モデル 2 $t_{ij} = KO_i^{\alpha} D_j^{\alpha} C_{ij}^{-\alpha}$ (2) モデル 3 $t_{ij} = \frac{O_i D_j C_{ij}^{-\alpha}}{\sum_i D_i C_{ij}^{-\alpha}}$ (3)

モデル 4	$t_{ij} = A_i O_i B_j D_j C_{ij}^{-\alpha} \cdots (4)$
	$A_i=1/\sum_j B_j D_j C_{ij}^{-\alpha}$ (5)
	$B_j=1/\sum_i A_i O_i C_{ij}^{-\alpha} \cdots \qquad (6)$
モデル 5	$t_{ij} = O_i\{e^{-LV_{j-1}} - e^{-LV_j}\} \cdots (7)$

ここに、 t_{ij} : ゾーンij間の分布交通量、 O_i : ゾーンi の発生交通量、 D_j : ゾーンj の集中交通量、 C_{ij} : ゾーンi がらみてゾーンj より近いゾーンの集中交通量の和、 V_j : $V_{j-1}+D_j$ 、K、 α,β,γ : パラメーター、である。

これらのモデルは、P. T. データの中から 通勤目的の 道路交通のみを選び出したものに 適用した. このとき、 ソーン内々、内外交通量とも、交通抵抗を実測の平均所 要時間で与えた. また、バラメーターのキャリプレーションは実測 O.D. 交通量とその推定値の残差の平方和が 最小となるようキャリプレーションした.

3. 結果と考察

現象再現性については、モデル3および4が最良であり、モデル5がこれに次ぐ。

サンプル数と推定結果の関係については、どのモデルについても、サンプル数を相当減らしても結果に大差を生じなかった。特にモデル4では、サンプル数 1000以下でも、有効な推定結果を与えた。

キャリブレーション誤差とゾーニングの相違について、パラメーター値とゾーニングを意図的に変化させて推定結果を調べたところ、どのモデルも広い範囲で安定した結果を与えることがわかった。特に、モデル3および4が安定的である。

4. 結 論

再現性,サンプル数に対する推定結果の依存性の面では,モデル3および4(Voorhees 形および二重制約形グラビティモデル)が優れている。特に,モデル4は,著しくサンプル数が少ない場合にも使用できる可能性がある。

キャリプレーション誤差とゾーニングの相違に対する 感度は、どのモデルについても大きな問題にはならない が、やはりモデル3および4が優れている.