

Inverse Analysis of Origin-Destination matrix for Microscopic Traffic Simulator

K. Abe¹, H. Fujii¹ and S. Yoshimura¹

Abstract: Microscopic traffic simulations are useful for solving various traffic-related problems, e.g. traffic jams and accidents, local and global environmental and energy problems, maintaining mobility in aging societies, and evacuation planning for natural as well as man-made disasters. The origin-destination (OD) matrix is often used as the input to represent traffic demands into traffic simulators. In this study, we propose an indirect method for estimating the OD matrix using a traffic simulator as an internal model. The proposed method is designed to output results that are consistent with the input of the simulator. The method consists of the following steps: (1) calculating link traffic volume from the OD matrix, and (2) updating the matrix. The estimated matrix is updated iteratively until it converges to a predefined tolerance level. Numerical experiments are then conducted using the proposed method on a grid network and on a representation of an actual road network. Finally, we discuss the characteristics of the proposed method and the non-negative constraint for the traffic volume.

Keywords: OD estimation, inverse problem, traffic simulation, Levenberg-Marquardt method, iterative method

1 Introduction

Microscopic traffic simulations are useful for solving various traffic-related problems, e.g. traffic jams and accidents, local and global environmental and energy problems, maintenance of mobility in aging societies, and evacuation planning for natural as well as man-made disasters. To use such microscopic traffic simulators, we need to input various types of traffic data. Data is typically input as an origin-destination (OD) matrix, which describes demands between origin-destination pairs in a traffic network, and is particularly necessary in microscopic simulators. Since the OD matrix cannot be observed directly, it has to be estimated in some way.

¹ The University of Tokyo, Bunkyo, Tokyo, Japan.

The approaches for OD matrix estimation can be roughly classified into two categories. The first is based on the population distribution. This approach is commonly used for traffic and civil planning using the four step model [McNally (2008)]. Since the population distribution is derived from traffic census data, we do not have to measure the actual traffic flow. However, the resolution of the census data is low, and the accuracy of the estimated results is not guaranteed.

The second approach is an inverse analysis of the link traffic volume data. Link traffic volume is the traffic volume counted at a fixed location. This approach is usually more accurate. Here, the OD matrix is optimized by minimizing the distance between the observed and estimated link traffic volume. This can be accomplished with a bi-level programming approach [Bera and Rao (2011)].

The inverse analysis approaches can be classified into two categories depending on the method used to solve the direct problem involved in the estimation process. Here, the direct problem refers to the assignment from the OD matrix to the link traffic volume. The solution of the direct problem can be obtained analytically, or approximated with the equilibrium assignment algorithm [Larsson and Patriksson (1992)]. Estimation of the OD matrix using this assignment algorithm was shown to be effective for large-scale road networks [Lundgren and Peterson (2008)]. For dynamic estimates, Off-line time-sliced OD estimation based on dynamic equilibrium, which is similar to the static OD estimation method, has been proposed [Barceló and Montero (2015)]. However, owing to the differences in network handling, the results of the equilibrium algorithm may not be compatible with the traffic simulator. The alternative is to use the traffic simulator to solve the direct problem. However, this latter method requires a high computational cost. Few studies have been conducted on the use of the traffic simulator for solving the direct problem, so the stability and robustness of the subsequent results are unclear.

In this paper, we newly propose an OD matrix estimation method using a microscopic traffic simulator. We then examine the accuracy and stability of our results for the inverse analysis.

2 Method

2.1 Outline of method

An outline of our method is shown in Figure 1.

The proposed method consists of the following steps: (1) calculating link traffic volume from the OD matrix, and (2) updating the optimal OD matrix solution. In step (1), a multi-agent based microscopic traffic simulator “ADVENTURE_Mates(MATES)” [Yoshimura (2006); Fujii, Yoshimura, and Seki (2010)] is used. In step (2), we use the Levenberg-Marquardt gradient method. After cal-

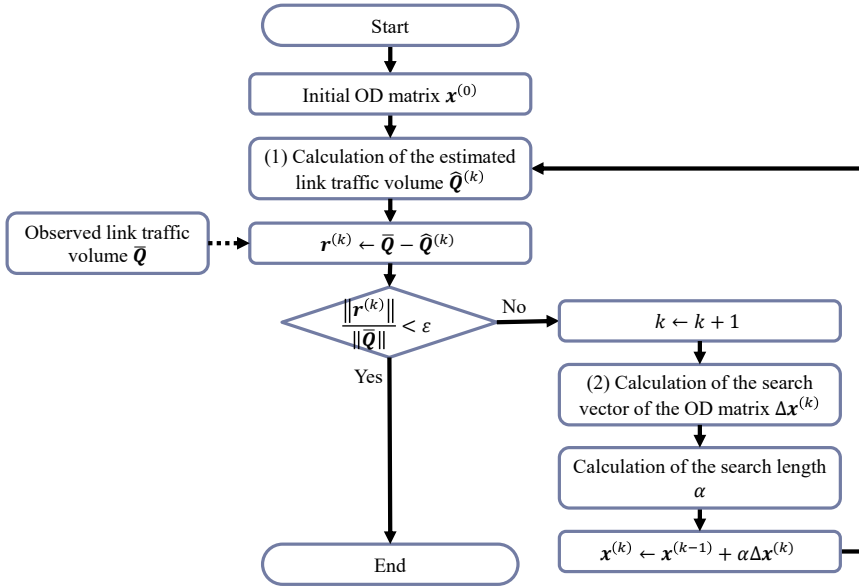


Figure 1: Flow diagram of the proposed method for estimating the OD matrix

calculating link traffic volume, the residuals of the link traffic volume are calculated. The optimal estimated OD matrix is updated using these residuals. These steps are iterated until the estimated OD matrix converges to a tolerance level.

2.2 Formulation

The OD matrix is estimated so as to minimize the residual norm of the link traffic volume between the estimated and the observed values. This process can be expressed as follows:

$$F = \|\mathbf{r}\|^2 \rightarrow \min. \quad (1)$$

where

$$\mathbf{r} = \bar{\mathbf{Q}} - \hat{\mathbf{Q}}(\mathbf{x}) \quad (2)$$

and The traffic simulation is used to evaluate $\hat{\mathbf{Q}}$ corresponds to the traffic simulation, where link traffic volume is calculated from the OD matrix \mathbf{x} .

We now define the dimensions of the vectors. N denotes the dimension of \mathbf{x} , i.e. the number of OD pairs, and M is the dimension of \mathbf{Q} , i.e. the number of observation points. In general, $N > M$. Therefore, in the OD matrix estimation, we have to estimate more variables from smaller datasets.

- \mathbf{x} : OD matrix (as a form of vector)
 \mathbf{r} : Residual (vector) of link traffic volume
 $\bar{\mathbf{Q}}$: Observed link traffic volume (vector)
 $\hat{\mathbf{Q}}(\mathbf{x})$: Estimated link traffic volume (vector)

In this paper, we use Euclid norms. If the variance and covariance of the observed link traffic volume are known and we use the Mahalanobis distance to calculate the norms, we can regularize the residuals, with consideration of the variance and covariance. However, this requires multiple observation of the traffic volume, which takes much observation cost. Therefore we assume that there is no variance and covariance in the observed link traffic volume data, and we employ Euclid norms.

For simplicity, we assume that the traffic flow is in a steady state and the relation between the OD matrix and the link traffic volume is linear. The assumption of such a linear relation is appropriate if traffic congestion does not occur. Here, the linear relation means that we can describe the relation in Equation 3 using a matrix \mathbf{J} .

$$\hat{\mathbf{Q}} = \mathbf{J}\mathbf{x} \quad (3)$$

In this study, we use the observed link traffic volume data which satisfies Equation 3. The assumption of a steady state means that both the OD matrix and the link traffic volume are stationary in time. This assumption is reasonable if an appropriate time period is used, e.g. a short time period where a drastic change in traffic flow does not occur.

\mathbf{J} in Equation 3 corresponds to the Jacobian matrix of F with respect to \mathbf{x} , which is an $M \times N$ matrix. Since this system has no explicit constitutive equation, we have to approximate \mathbf{J} . The value of \mathbf{J} is estimated as in Equation 4 using the results of the traffic simulator.

$$J_{i,j} = \begin{cases} 1 & \text{(when the link } i \text{ is included in the route of OD pair } j) \\ 0 & \text{(otherwise)} \end{cases} \quad (4)$$

Since a change in \mathbf{x} corresponds to that in the route choice, we re-evaluate \mathbf{J} at each iteration.

2.3 Modification of simulator

MATES is a multi-agent microscopic traffic simulator, which models individual driver behavior. This is useful for modeling individual vehicles on the microscopic

scale and for extrapolating traffic flow on the macroscopic scale in a form of the emergence phenomena. Since MATES uses OD matrix as input data and outputs link traffic volume, we use this simulator to calculate link traffic volume from OD matrix.

In order to model individual driver behavior, stochastic elements are necessary, so MATES simulates random numbers using some random seeds. In this study, we fix these random seeds to make MATES deterministic for simplicity.

In addition, the original version of MATES generates vehicles according to a Poisson distribution whose rate parameter corresponds to the OD traffic volume, i.e. the element of the OD matrix. However, when using a Poisson distribution, the number of generated vehicles in fixed interval of time is not always equal to that described in the OD matrix, which causes stochastic error in the OD estimation. Here, we modify the method to generate vehicles at fixed time intervals corresponding to the OD traffic volume. The number of generated vehicles thus is guaranteed to be exactly equal to the OD traffic volume.

2.4 Solution updating step

The OD matrix is updated using one of the gradient methods. This updating step is given in the following general form:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha \Delta \mathbf{x}^{(k)} \quad (5)$$

where

- α : Search length coefficient
- $\Delta \mathbf{x}$: Search direction vector
- $\bullet^{(k)}$: Variable for iteration step k

In this study, we use the Levenberg-Marquardt method (LMM) [Levenberg (1944); Marquardt (1963)]. The definition of $\Delta \mathbf{x}$ in the LMM is given as follows:

$$\Delta \mathbf{x} = -\mathbf{H}^{-1} \mathbf{G} \quad (6)$$

where

- \mathbf{G} : Gradient of F with respect to \mathbf{x} ; $\mathbf{G} = -\mathbf{J}^T \mathbf{r}$
- \mathbf{H} : Hessian matrix of F with respect to \mathbf{x} .

The steepest decent method (SDM) is commonly used in optimization problems, because it only requires a gradient matrix as input. However, the rate of convergence is slower than other methods. On the other hand, the LMM is based on the

Gauss-Newton method (GNM) and is applicable to nonlinear problems. In this method, the Hessian matrix is assumed to be the product of Jacobian matrices: $\mathbf{H} = \mathbf{J}^T \mathbf{J}$. However the GNM can only be applied when $N < M$. In the LMM, the Hessian matrix is regularized as in Equation 7

$$\mathbf{H} = \mathbf{J}^T \mathbf{J} + \lambda \mathbf{D} \quad (7)$$

where λ denotes a regularization parameter for LMM and \mathbf{D} denotes a diagonal matrix, e.g., the identity matrix. The addition of the diagonal matrix makes the Hessian matrix non-singular, so that its inverse can be calculated. Marquardt proposed initially setting λ to a large value and decreasing λ with successive iterations [Marquardt (1963)]. If λ becomes zero, Equation 7 is equivalent to the updating scheme for the GNM. If λ is sufficiently large, the search direction approaches that of the SDM. According to Marquardt's proposal, the initial iteration steps have the robustness of SDM, and the later iteration ones have the high rate of convergence of the GNM. In this study, evaluating F is computationally expensive. Thus, we decrease λ from an initial value $\lambda_0 (> 0)$ at constant rate $\gamma (0 < \gamma < 1)$ with successive iterations as follows:

$$\lambda = \lambda_0 \gamma^k \quad (8)$$

The convergence is judged comparing relative residual norm (RRN) $\|\mathbf{r}\|/\|\bar{\mathbf{Q}}\|$ and a tolerance ε . Since RRN is regularized by observed link traffic volume $\bar{\mathbf{Q}}$, it can be compared among different networks.

2.5 Introduction of non-negative constraints into the estimation method

Non-negative constraint of traffic volume is necessary for OD estimation. When the OD matrix satisfies the constraint, the link traffic volume then satisfies it. Thus it is sufficient to apply the constraint only to OD matrix. Since the original LMM has no constraint for variables, we have to introduce it into the LMM. We propose the following two kinds of methods to do so.

First, we describe the common assumption in both the methods. In this study, we make the non-negative constraint stricter, i.e. for all elements i , $x_i > \delta$, where δ is a small value and is larger than zero. The reason is that too small OD traffic volume causes some estimated link traffic volume to be zero. If it occurs, some elements of Jacobian matrix become zero, then the zero elements are propagated to some elements of $\Delta \mathbf{x}$. Consequently, the search direction gets limited.

Next, we describe the methods in sequence. The first, named method A, is designed to satisfy the constraint strictly. Here, the initial value of α is set to 1. If $x_i^k + \Delta x_i^k < \delta$ for any element i , α is modified to be so smaller as to satisfy

$x_i^k + \Delta x_i^k = \delta$ for all elements i . It does not modify the search direction from that given in the LMM. Nevertheless, it is expected that residuals stop decreasing within insufficient large values because α becomes smaller toward zero with successive iterations.

The second, named method B, is the method using heuristics. Here, the search length coefficient α is fixed to be 1 strictly. After the solution updating step, the updated x_i is forcibly modified to be δ individually when $x_i < \delta$ for any index i . Although it makes different search direction from that in the LMM, since α does not become zero, iteration is expected not to halt except the convergence.

We discuss these characteristics of the two methods in Section 4.

3 Numerical experiments

3.1 Outline of experiments

Before using the proposed method we describe previously, we have to determine the following two things; the optimal tolerance for convergence ε and which method to be used for the non-negative constraint. To determine them, it is necessary to use some evaluation index which can be discussed in terms of traffic engineering, e.g. the proportion of reproduced link traffic volume to observed one and the correlation coefficient between reproduced and observed link traffic volume. Here, we apply the proposed method to some cases considering observed link traffic volume with and without noise. Using these results, we determine these two things and discuss accuracy, stability and characteristics of each case or method.

3.2 Networks used for experiments

In the numerical experiments, we use two types of road networks, i.e. grids or road maps. The first map is a regular grid with simple topology, shown in Figure 2(a). In the figure, each line corresponds to a link. Numerical instabilities may arise when solving symmetric maps. Therefore, we slightly displace all grid nodes in a random manner. In this map, each link has 2 lanes, and all intersections have traffic signals.

The second map is an actual road network topology, i.e. a 3 km x 3 km area in Tokyo. The map is shown in Figure 2(b). This map includes one way links, so that the number of OD pairs N is less than $n(n-1)$, where n denotes the number of OD nodes.

The properties of these networks are shown in Table 1. In the table, detectors are the points at which link traffic volume is measured in a simulation. As a general rule in this study, detectors are located on all links, 5 m from the end point. Consequently, there are twice as many detectors as links.

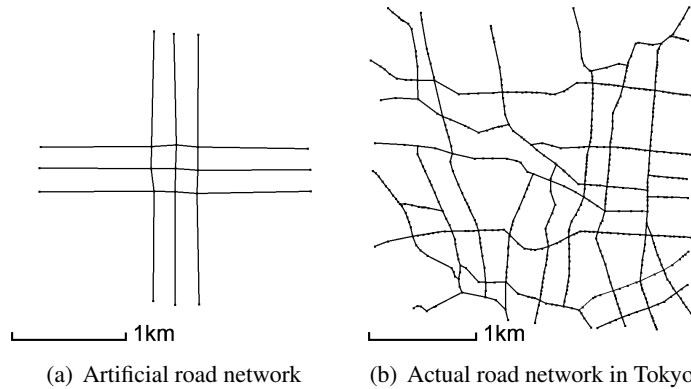


Figure 2: Networks used for experiments

Table 1: Network properties

Map	Number of OD nodes	Number of OD pairs N	Number of detectors M	OD traffic volume x_i which satisfies linearity Unit: veh/h
Regular grid	12	132	48	≤ 30
Tokyo 3 km	26	600	222	≤ 14

3.3 Initial conditions for experiments

The parameters used in the experiments are shown in Table 2.

Table 2: Initial conditions in the experiments

Symbol	Definition	Value
λ_0	Initial value of λ	10
γ	Reduction rate of λ	0.25

Next, we consider the observed link traffic volume. In this study, we use a simulated dataset in advance in all experiments. To generate this link traffic volume, we run the traffic simulator using OD matrices which satisfy the assumption of linearity described in Section 2.2. The results are given in Table 1. Since the data has been simulated under these conditions, the existence of the solution is guaranteed.

Using this data as a reference case, we can consider other cases where artificial noise has been added to the data in Table 1, to examine the accuracy and stability

of the proposal method. Noise is added to the networks as follows.

$$\bar{Q}_{NoiseIncluding} = \bar{Q} + \delta \quad (9)$$

Here, δ is a noise vector, where each element i follows a uniform distribution whose upper and lower limits are within $\pm 10\%$ of \bar{Q}_i . For the experiments with each map, 10 kinds of observed link traffic volume are prepared using 10 different random seeds.

4 Results and Discussions

4.1 Experiments on the regular grid

First of all, we apply the proposed method to the regular grid. As mentioned in Section 3.1, we apply both of the two methods of the non-negative constraint, i.e. methods A and B.

We apply the methods to the cases with and without noise. RRN transitions in these cases are shown in Figures 3 and 4. The x-axis denotes the number of iteration counts, while the y-axis is the RRN on the logarithmic scale. Red and green lines denote the methods A and B, respectively.

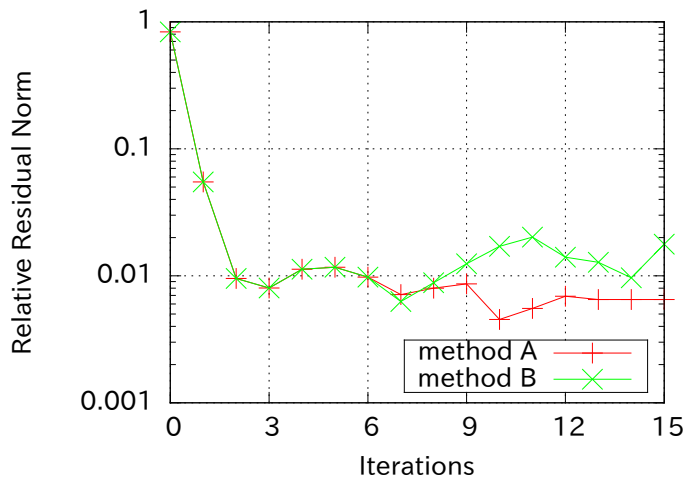


Figure 3: RRN transition in the regular grid without noise

These figures show that RRN transitions in both the methods A and B decrease in the same rate until early iteration steps, i.e. 6th iteration step without noise and 2nd iteration step with noise. The reason is that all elements of OD matrices are positive until those steps. Since the methods A and B consider both non-negative

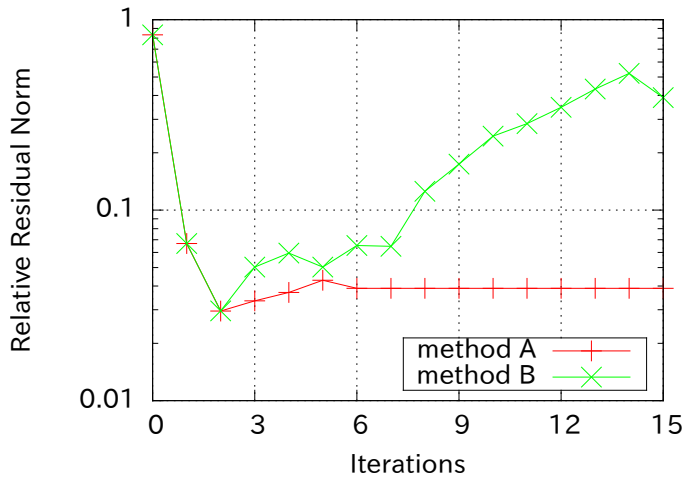


Figure 4: RRN transition in the regular grid with noise

constraints, there is no difference unless the constraints are applied to the solution updating step. On the other hand, RRNs in method B shows different transition between the cases with and without noise. The RRN in the case without noise oscillates, while the RRN in the case with noise almost continuously increases after the RRN becomes minimum. This difference is caused by the difference of the number of the elements to which the non-negative constraint is applied in OD matrices. If the non-negative constraint is applied for an element in OD matrices, its value approaches δ . Thus, the number of the elements of $x_i = 1$ denotes the number of the elements to which the constraint is applied. Here, we show the results in Figure 5.

This figure shows that there is a difference in the number of the elements of $x_i = 1$ between the cases with and without noise. In the case without noise, the number oscillates in small, i.e. from 0 to 3, while in the case with noise, the number increases almost continuously. As a result, the case with noise has more elements to which the constraint is applied than the case without noise. This suggests that the situation of the regular grid without noise has much simplicity and stability in iterative solution search.

To discuss the reasonable tolerance for convergence ε and the characteristics in estimated link traffic volume, we confirm the estimated link traffic volume using the minimum RRN in the following cases: method A without noise, method B without noise and method B with noise. These are considered with the difference of scale of minimum RRN among the cases. All the results are shown in Figure

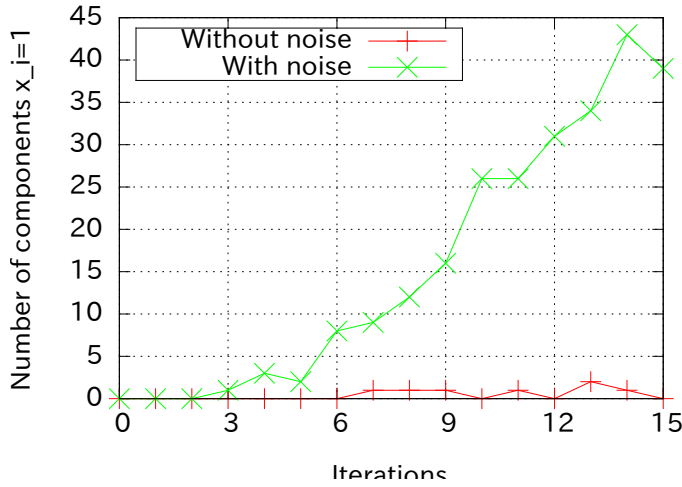
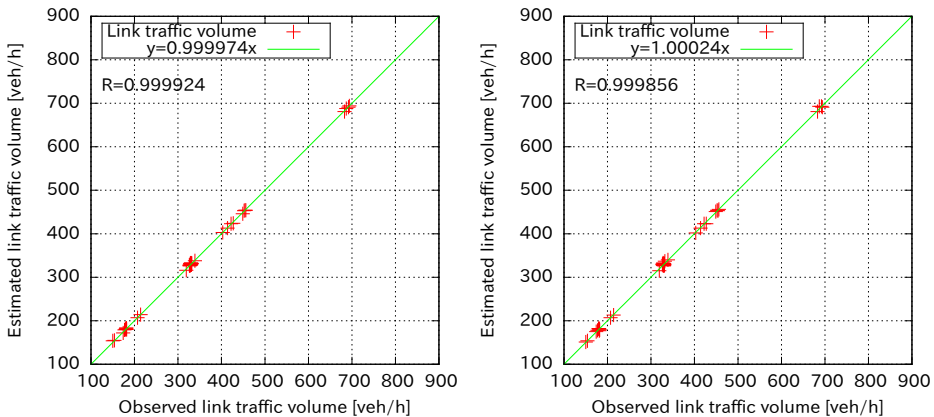


Figure 5: Transition of the number of the elements of $x_i = 1$ in the regular grid

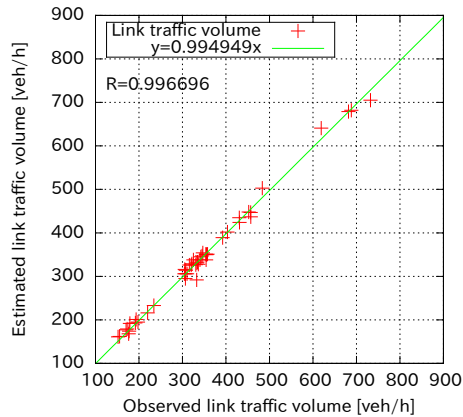
6. The x-axis denotes observed link traffic volume \bar{Q} , while the y-axis is the link traffic volume estimated from the OD matrix \hat{Q} . The approximate straight line is fixed at the origin point.

These figures of Estimated link traffic volume shows following two evaluation indexes: gradient of approximate straight line a and correlation coefficient R . The gradient of approximate straight line a shows the scale of estimated link traffic volume. If a equals to 1, it suggests that the scales of observed and the estimated link traffic volume are same. The scale of the estimated link traffic volume is higher than the observed one if a is greater than 1, otherwise the scale is lower. The correlation coefficient R shows the dispersion of accuracy of reproduction. If R equals to 1, the following relation is satisfied: $\hat{Q} = a\bar{Q}$. On the other hand, if some elements of link traffic volume are plotted away from approximate straight line, R becomes lower.

In terms of these evaluation indexes, the values of a in all of these cases equal to 1 in error by at most 0.006. The values of R in all cases also equal to 1 in error by at most 0.004. This high correlations arise from the characteristic that the LMM weights residuals equally. These results suggest the accuracy in these cases are the same or higher than 99% in proportion of the estimated link traffic volume to the observed one. Since these errors are sufficiently small to ignore and the highest RRN in these cases is 2.96×10^{-2} at that time, it is sufficient to set convergence coefficient ε to 0.03.



(a) Method A without noise; $RRN = 4.54 \times 10^{-3}$ at 10 iterations
 (b) Method B without noise; $RRN = 6.28 \times 10^{-3}$ at 7 iterations



(c) Method B with noise; $RRN = 2.96 \times 10^{-2}$ at 2 iterations

Figure 6: Estimated link traffic volume for the regular grid

4.2 Application to an actual road network

Next, we apply the proposed method to an actual road network, i.e. a 3 km x 3 km area in Tokyo. We show the RRN transitions in the cases with and without noise in Figures 7 and 8.

In this map, RRN transitions in both the two cases are alike. Until early several iteration steps, RRNs in method B are smaller than those in method A, while after that, RRNs in method B become larger than those in method A. The increase of RRNs in method B is explained by the same reason mentioned in Section 4.1, i.e.

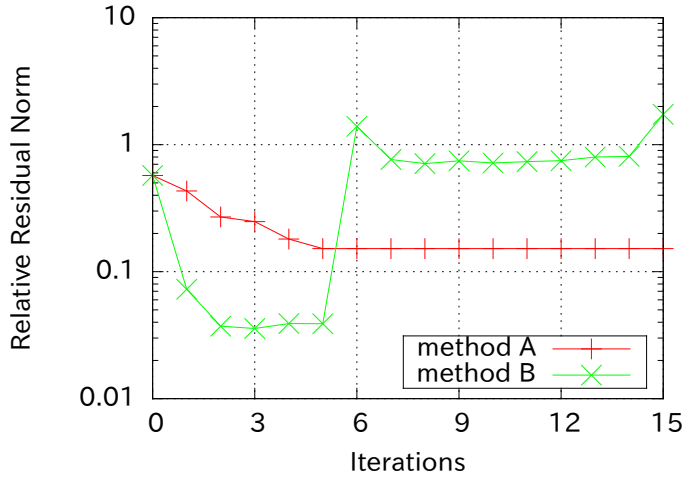


Figure 7: RRN transition for the 3 km x 3 km area in Tokyo without noise

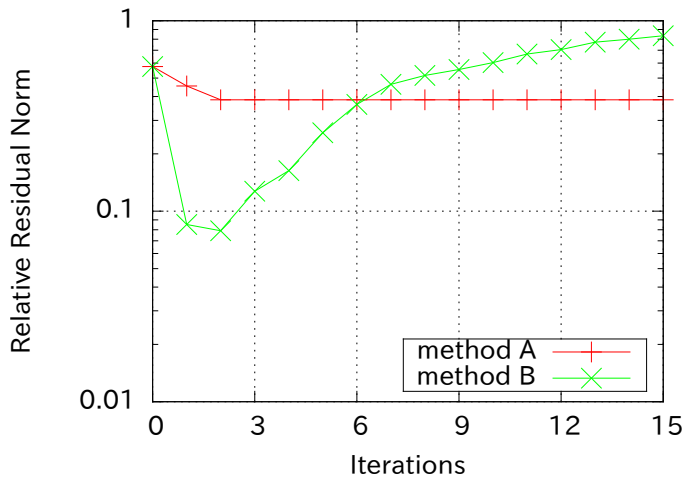


Figure 8: RRN transition for the 3 km x 3 km area in Tokyo with noise

the increase of the elements of $x_i = 1$ in OD matrices. Actually, in this map, the transitions of the number of the elements of $x_i = 1$ are similar to the green line in Figure 5 in both of the cases with and without noise. We show it in Figure 9.

On the other hand, these results are different from those in the regular grid, where RRNs are the same in method A and B at early several iteration steps. This suggests that in method A, the search length α approaches zero at early few iteration steps and therefore the RRNs decrease insufficiently. Actually, unlike in the regular grid,

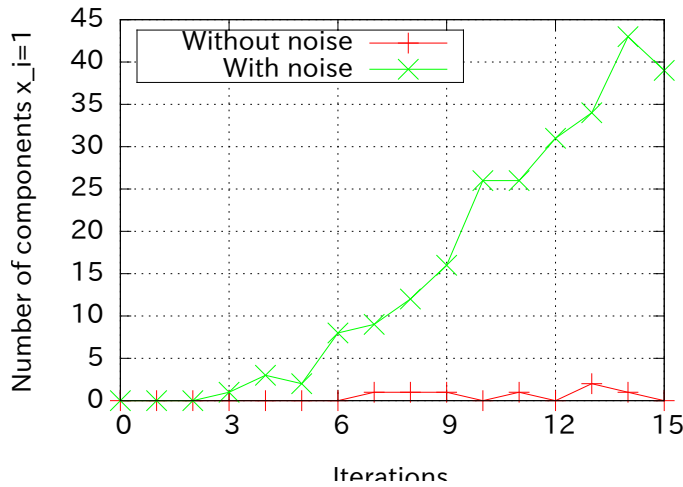


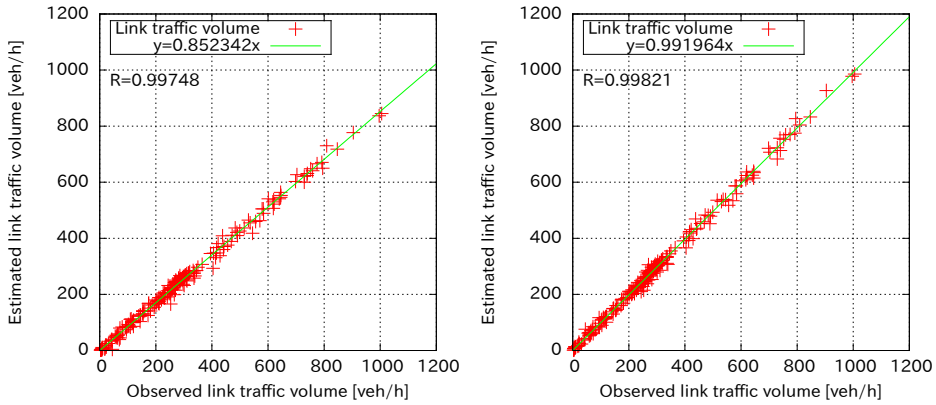
Figure 9: Transition of the number of the elements of $x_i = 1$ for the 3 km x 3 km area in Tokyo

the non-negative constraint is applied to OD matrices at only the first iteration in Tokyo map. This is caused by the large number of variables in OD matrices in Tokyo map.

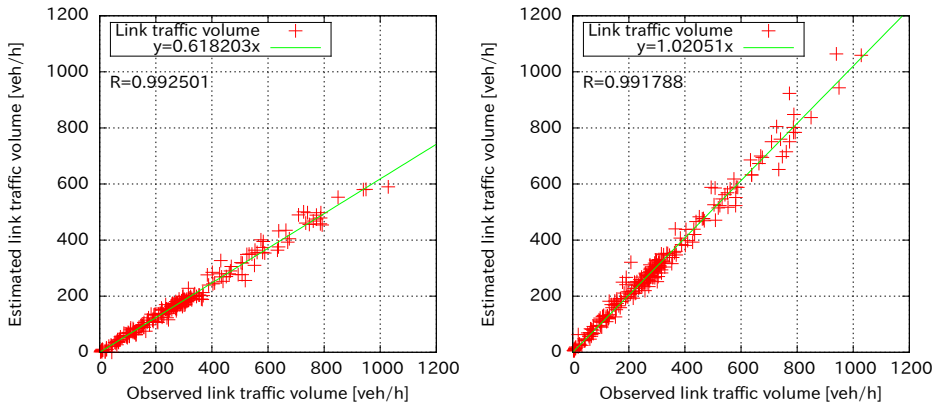
For method B, the iteration steps when RRNs become minimum are alike as follows: 3rd iteration steps in the case without noise and 2nd iteration steps in the case with noise. These values are almost the same as that in the regular grid with noise, whose value is 2, and these are all small number. Therefore the results suggest that even if not adding noise, solution search is more difficult in Tokyo map than in the regular grid, which is also caused by the large number of variables.

In terms of minimum RRNs, as similar as the regular grid, there is some difference between in the cases with and without noise. Here, we show the estimated link traffic volume when RRN is minimum in the following cases, considering the scale of minimum RRN: method A without noise, method B without noise, method A with noise and method B with noise. All the results are shown in Figure 10.

These figures show that the gradient of approximate straight line a is distant from 1 when method A is used, i.e. 0.852 without noise and 0.618 with noise. On the other hand, when method B is used, a equals to 1 in error by at most 0.022 as like as in the regular grid. The correlation coefficient R almost equals to 1 for all constraint methods and cases about noise. These results suggest method B is more applicable. Then, the accuracy is the same or higher than 97% in proportion of estimated link traffic volume to observed one. Even though this value is less than that in the regular



(a) Method A without noise; $RRN = 0.152 \times 10^{-2}$ at 5 iterations (b) Method B without noise; $RRN = 3.56 \times 10^{-2}$ at 3 iterations



(c) Method A with noise; $RRN = 0.384$ at 2 iterations (d) Method B with; $RRN = 7.90 \times 10^{-2}$ at 2 iterations

Figure 10: Estimated link traffic volume for the 3 km x 3 km area in Tokyo

grid, this is expected to be still sufficiently accurate. Consequently, it is sufficient to set the tolerance for convergence ϵ to 0.8, where only method B converges.

5 Conclusions

We newly proposed an OD estimation method using a traffic simulator, whose result is designed to be suitable for use of the simulator directly. In addition, we introduced two kinds of approaches of applying non-negative constraints to the proposed method. Through numerical experiments, we demonstrated the validity of the method. The estimated link traffic volume is strongly correlated with the ob-

served link traffic volume. This is due to the iterative process in the LMM. When we introduce the non-negative constraint into our method, the one using heuristics is found to be effective and gives smaller RRN. Then, the tolerance for convergence for the RRN is found to set 0.08, when the estimated link traffic volume has 97% or more accuracy to the observed one. To improve the accuracy, especially even if the number of observation points becomes smaller, it is necessary to improve the method of non-negative constraint in terms of heuristics. In future work, we plan to examine the accuracy and stability of the proposed method with fewer dataset and to consider the state of congestion.

Acknowledgement: This work was supported by JSPS KAKENHI Grant Number 15H01785.

References

Barceló, J.; Montero, L. (2015): A Robust Framework for the Estimation of Dynamic OD Trip Matrices for Reliable Traffic Management. *Transportation Research Procedia*, vol. 10, pp. 134–144.

Bera, S.; Rao, K. (2011): Estimation of Origin-Destination Matrix from Traffic Counts: the State of the Art. *European Transport - Trasporti Europei*, vol. 49, pp. 3–23.

Fujii, H.; Yoshimura, S.; Seki, K. (2010): Multi-agent Based Traffic Simulation at Merging Section Using Coordinative Behavior Model. *Computer Modeling in Engineering and Sciences*, vol. 63, no. 3, pp. 265–282.

Larsson, T.; Patriksson, M. (1992): Simplicial Decomposition with Disaggregated Representation for the Traffic Assignment Problem. *Transportation Science*, vol. 26, no. 1, pp. 4–17.

Levenberg, K. (1944): A Method for the Solution of Certain Non-Linear Problems in Least Squares. *The Quarterly of Applied Mathematics*, vol. 2, pp. 164–168.

Lundgren, J. T.; Peterson, A. (2008): A Heuristic for the Bilevel Origin-Destination-Matrix Estimation Problem. *Transportation Research Part B: Methodological*, vol. 42, pp. 339–354.

Marquardt, D. W. (1963): An Algorithm for Least-Squares Estimation of Non-linear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441.

McNally, M. G. (2008): The Four Step Model. *Center for Activity Systems Analysis*.

Yoshimura, S. (2006): MATES: Multi-Agent Based Traffic and Environment Simulator-Theory, Implementation and Practical Application. *Computer Modeling in Engineering and Sciences*, vol. 11, no. 1, pp. 17–25.