Estimating Traffic Conditions of the Radial-ring Expressway Network by Fusing Probe and Detector Data into Traffic Simulations

Azusa Goto [1*], Kazunori Ooshima [1], Kosuke Yamada [2], Ryota Horiguchi [3], Shin Sakaki [1], Hidenori Yoshida [1]

[1] Intelligent Transport Systems Division, National Institute for Land and Infrastructure Management, Asahi 1, Tsukuba, Japan

[2] Urban Planning and Management Division, Pasific Consultants Co, Ltd., 3-22 Kandanishiki-cho, Chiyoda-ku, Tokyo, Japan

[3] i-Transport Lab. Co., Ltd., 9th floor, 3-10 Kandaogawa-cho, Chiyoda-ku, Tokyo, Japan

[*] goto-a92uj@mlit.go.jp

Abstract: This paper introduces a framework to estimate road network traffic conditions by fusing probe and detector data into traffic simulations. In this traffic simulation, the trajectories of some vehicles are controlled by the observed probe vehicles, then the speed reductions of these vehicles propagate according to the kinematic wave theory. In addition, the origin-destination matrix is calibrated using a mathematical model so that the errors in the estimated link traffic volumes from the detector data can be minimized. By applying this framework, the traffic conditions on the Tokyo Metropolitan Expressway Network were estimated. By comparing the estimation results with the observation, it was confirmed that the estimation showed a good accuracy in terms of link traffic volumes, and traffic congestion at the typical bottlenecks could be reasonably represented to some extent.

1. Introduction

The Tokyo Metropolitan Expressway Network in Japan has developed for a long time since the 1960s. First, radial roads were constructed to make the surrounding areas accessible to the metropolitan centres, currently the ring roads that connect these radial roads have been mostly completed [1]. Since various routes will become available for road users after the completion of this radial-ring expressway network, it will be more important to utilize ITS technologies, such as dynamic route guidance or congestion charges, to realize a more efficient road network operation.

In order to implement such operational measures, it is essential to understand traffic conditions in a network, which means not only traffic flows and speeds of individual sections but also traffic flow distributions from each origin to destination. It will allow road network operators to recognise where congestion happens, where vehicles within the congestion come and go, and where and how they can alleviate congestion by encouraging vehicles to detour.

However, existing detectors (or sensors) give us limited information about traffic flows and speeds at specific sites. Since these detectors are not so densely located on some sections, e.g., 1 or less in 5 km, this information is not enough to understand traffic flow distribution conditions over a whole network. On the other hand, we are recently able to obtain some probe data by vehicle-infrastructure cooperative system so-called "ETC 2.0 system" in Japan. Although the number of probe vehicles is still limited, this data contains a series of time and position (longitude and latitude) information of probe vehicles for every 200 m, which tells us where and when they reduced their speeds. It is expected that we will able to estimate traffic conditions in more detail by fully utilizing these different kinds of data.

Therefore, the objective of this research is to develop a framework to estimate road network traffic conditions by fusing probe and detector data into mesoscopic traffic simulations. In the simulation, vehicles are generated and attracted according to the time-varying origin and destination matrix (hereafter referred to as "OD matrix"), which cannot be directly observed and therefore is estimated by calibrating the initial OD matrix using the observed link traffic counts from detector data. Then positions of the vehicles are estimated by incorporating the time-space trajectories of probe vehicles into the fundamental diagram of traffic flow based on the variational theory developed by Daganzo [2]. This can give us information about traffic conditions where detectors are not installed, as well as traffic flow distributions, which are not limited for probe vehicles.

So far, the framework has been developed for performing post evaluations of an operational measure by comparing estimated traffic conditions before and after implementation, thus all data was collected and inputted offline. However, it is also applicable to real-time traffic monitoring if all the data is connected online in the future. This will contribute to realizing effective road network management by implementing some operational measures based on dynamic traffic conditions.

The following parts of this paper consist of a literature review and the position of this research (section 2), an explanation of the methodology (section 3), a case study to verify how much accurately traffic conditions are estimated by the proposed methodology (section 4) and finally the conclusion and future work (section 5).

2. Literature Review

Previous studies ([3] ~ [8]) have already developed the frameworks for estimating and/or monitoring road network traffic conditions by inputting observed data into traffic simulations. In the traffic simulations, the most

important tasks are estimation of demand (i.e., OD matrix) and calibration of supply parameters (i.e., fundamental diagram).

In most of the existing frameworks, the OD matrix is estimated or calibrated according to the observed traffic volumes by some detectors and/or sensors based on typical traffic conditions. For example, Dynasmart-X [3], DynaMIT [4], and Aimsun Live [5] select the OD matrix based on the database that contains the relationship between OD matrices and observed traffic volumes by detectors and/or sensors in the past on the highway and motorway network. HEROINE [6] and RISE [7], which were developed by the urban expressway companies in Japan, utilize the onramp traffic volumes directly for calibrating OD matrix between the onramps and off-ramps. The calibration is based on the historical pattern of OD matrix considering the day of the week and other factors with the temporary changes of the observed onramp traffic volumes.

In order to apply these frameworks ([3] ~ [7]), it is necessary to obtain a historical pattern of OD matrices. However, it needs a large amount of data and thus is very difficult, since we cannot observe ODs of all vehicles in general. Besides, the Tokyo Metropolitan Expressway Network, which is our subject to study, has not been completed yet; some of the sections are still under construction. Thus, the historical patterns of OD matrices may not be applicable due to the openings of new sections in the future.

Without using the historical patterns of OD matrices, for another example, since road administrators have recently been able to obtain the OD matrix of vehicles that use the electric toll collection (ETC) system in Japan, Hirai, et al. [8] utilized this data for estimating the OD matrix of all vehicles (including the ones which pay tolls in cash, not through the ETC system) on the entire inter-urban expressway network. This OD matrix of ETC users was called "ETC-OD matrix". Because users of the ETC system accounted for about 90% of all expressway users, the ETC-OD matrix could be a good initial input for estimating the OD matrix of all users. Since this approach does not need historical pattern of OD matrices but can fully utilize the data available for road administrators, it was also adopted in our study.

On the other hand, the calibration of supply parameters had been conventionally achieved though the analysis on the bottleneck capacities on expressways by using historical detector data, as also incorporated in the above study [8]. However, it also requires plenty of data, since supply parameters vary by road geometry (e.g., number of lanes, lane width, gradient) as well as day of the week, weather condition, ambient lighting condition, type of road users (e.g., long-distance traffic or commuter traffic), etc. Instead of calibrating the parameters to these influencing factors, many of the above-mentioned frameworks ([3] ~ [7]) use observed speed data by detectors and/or sensors. The use of actually observed data is advantageous also in grasping the congestion caused by not only fixed bottlenecks but also some temporary incidents (i.e., traffic crashes or falling objects on roads) that cannot be calibrated in advance.

In our study, we focused on the use of probe data, because the detectors and/or sensors are sometimes located with large spacing (for example, 5~10 km) especially on
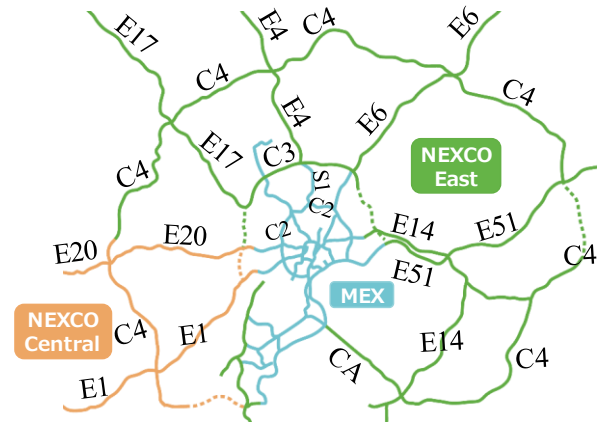

Fig. 1. *Expressway network in the Tokyo Metropolitan area with route numbers (dotted lines means under construction as for 2017)*

intercity expressways in Japan, only detector data is not precise enough for detecting congestions. From this viewpoint, probe data had a great advantage because it contains continuous speed information with denser spacing as far as probe vehicles are driving on the roads.

The utilization of probe data is one of the most substantial challenges in traffic estimation. For example, Work et al. [9], Herring et al. [10] showed the applicability of probe data for estimating travel time (or travel speed). Choi and Chung [11] developed a data fusion technique to combine probe and detector data for more reliable and accurate estimation. While most of the existing studies including above used only travel time (or travel speed) information of probe vehicles, Mehran et al. [12] proposed a methodology to utilize individual vehicle trajectories (time-space) based on Daganzo's variational theory [2] for estimating trajectories of other vehicles at the corridor level. This approach has the great extensibility to utilize what probe data contains, not only time-space but also vehicle type for example in the future research. Therefore, our study was aimed to incorporate this approach for calibrating the supply parameters into network simulations.

3. Methodology

3.1. Subject network

The subject road network was a whole expressway network in the Tokyo metropolitan area as illustrated in Fig. 1. It has 1599.7 km for both directions covering the area of about 1.1×10^4 km^2, consisting of 2214 links in Digital Road Map (DRM) [13]. As shown in the figure, the network is outlined by three ring roads (C2, C3, and C4) and several radial roads with 29 junctions. This allows road users to choose several routes, especially when passing through the metropolitan area from and to outside the outer ring road (C4).

The three different colours in Fig. 1 stand for the companies that are in charge of construction, operation, and management for these expressway sections: Metropolitan Expressway Co., Ltd., (MEX) for the central part, Central Nippon Expressway Co. Ltd., (NEXCO Central) for the west part, and East Nippon Expressway Co., Ltd., (NEXCO East) for east and north part.

3.2. Overview of traffic estimation process

In this research, traffic conditions were estimated using a traffic simulator with different kinds of observed data. Fig. 2 shows the flowchart of the whole estimation process.

The estimation process consisted of two steps. At the first step, we calibrated the OD matrix, which was one of the most important inputs for traffic simulations. The OD matrix in this study must specify the traffic demands between the onramps and off-ramps on the subject expressway network during each time interval. In order to obtain the OD matrix, we prepared an initial OD matrix based on the ETC-OD matrix, and calibrated it through the iterative process of traffic simulation and minimization of the errors in link traffic volumes from the detector data.

By inputting the calibrated OD matrix at the first step, we carried out the traffic simulation with ETC 2.0 probe data at the second step. Actually, the iteration at the first step also included traffic simulations as shown in Fig. 2, but these were done without the probe data.

3.3. Data

As stated, three kinds of data were utilized in this framework: detector data, ETC- OD matrix and probe data. The following subsections introduce them in detail.

3.3.1 Detector data: In total, 408 detectors were installed in the subject network; that is, on average for every 300-600 m on the expressways under MEX and at least one between every on/off-ramps (about 5~10 km) on the expressways under NEXCO Central and East. With these detectors, we obtained traffic volume and average spot speeds aggregated once every 5 minutes for two categories of vehicle types (i.e., passenger cars and heavy vehicles).

Several types of detectors were used at different sites, for example, inductive loop detectors, ultrasonic detectors, and video image processor detectors. From them, we excluded detectors with errors and missing records, and the detectors that seemed inaccurate when comparing traffic counts with upstream and downstream detectors in the data cleansing process.

3.3.2 ETC-OD matrix: As briefly mentioned in section 2, the ETC-OD matrix was the OD matrix of ETC users which specified the traffic volumes between the onramps and off-ramps during each time intervals. This matrix was calculated in each of the three expressway companies (MEX, NEXCO East, and NEXCO Central) by aggregating the records of times when vehicles passed through the tollgates located at onramps, off-ramps, or boundary junctions between different expressways with the ETC on-board units (including ETC 2.0 on-board units that are explained in 3.3.3).

In this research, we first combined these matrices into one matrix in order to make the ETC-OD matrix in the whole subject network with the aggregation intervals of 15 minutes for each type of vehicles (passenger cars or heavy vehicles). Here, several assumptions were made about diverging and merging volumes at boundary junctions where expressways under different companies intersect.

Then, we approximately converted it into the OD matrix of all vehicles by multiplying the elements by the
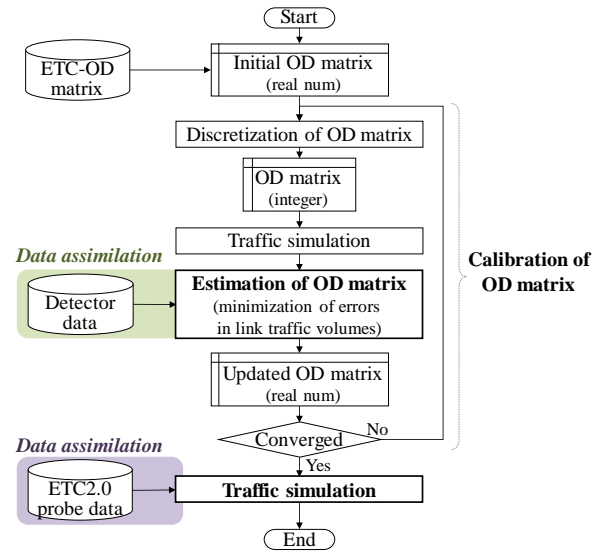

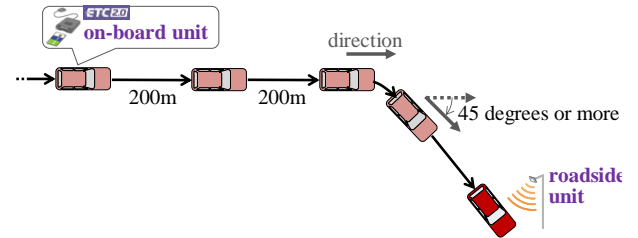**Fig. 2.** *Flowchart of the whole estimation process*

Fig. 3 *schematic image of travel record in probe data by ETC2.0 system*

inverse of the average percentage of ETC users in the expressway companies according to their destination off-ramps. The ETC system in Japan has been in operation nationwide since 2001, and the percentage of ETC users is currently almost 90% among all expressway users.

The OD matrix obtained in the above procedure was used for the initial OD matrix in Fig. 2. Since this procedure was done with several simplifications, the initial OD matrix would not be accurate enough for estimating dynamic traffic conditions. That was why the optimization was further carried out as shown in Fig. 2.

3.3.3 Probe data: Probe data was obtained from the ETC 2.0 system. The ETC 2.0 system is a vehicle-infrastructure cooperative system using the DSRC (Dedicated Short Range Communications) between ETC 2.0 on-board units and roadside units. The system has been deployed nationwide by the Ministry of Land, Infrastructure, Transport and Tourism since 2011. While the ETC system was only for electric toll collection, the ETC 2.0 system had additional functions of dynamic route guidance, safe driving assistance, and probe data collection.

In the probe data collection of the ETC 2.0 system, time and position (longitude and latitude) information of each probe vehicle is recorded and accumulated in its on-board unit every 200 m or when its direction changed 45 degrees or more as illustrated in Fig. 3. Then, the accumulated data is uploaded when probe vehicles passed by roadside units, which are generally set every 10~15 km on intercity expressways and 4 km on urban expressways on
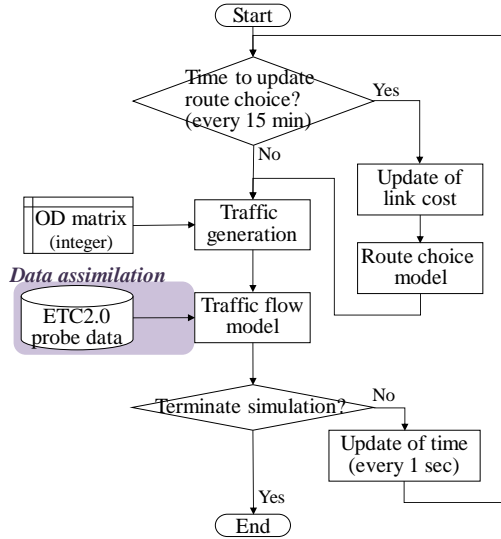
Fig. 4. *Flowchart of traffic simulation*

average. The roadside units send the data to the central server, and finally the central server connects the data of each probe vehicle downloaded by different roadside units for every day in order to make continuous trajectories of probe vehicles. For more information, please refer to the report [15].

Although the use of the ETC 2.0 system is gradually increasing year by year, it is still limited in total traffic volume. On average, we can collect probe data from about 2~3% of all vehicles traversing the subject expressways. For example, probe data was collected from 40 vehicles, which accounted for 3.3% of the total 1,221 vehicles on one section of the E1 expressway (bound for Tokyo, upstream of Yokohama-Machida IC) during the evening peak from 16:00 to 16:15 on Sunday, 5th June, 2016.

3.4. Traffic simulation

In the main part of the estimation process, we simulated traffic with the existing simulator, namely SOUND (Simulation On Urban road Network with Dynamic route choice; [16], [17]), which had been developed by i-Transport Lab. Co., Ltd., by adding a special function for incorporating probe data. Fig. 4 shows the general flowchart. SOUND is a mesoscopic traffic simulator dealing with discrete vehicles with 1-second update intervals. In this simulation, vehicles are generated according to the OD matrix, and their movements are simulated by the traffic flow model and route choice model in the following way.

3.4.1 Traffic generation: Traffic was randomly generated at every time step (i.e. 1 second) from onramps according to a 15-minute OD matrix. Here, the OD matrix was updated during the calibration step as explained in subsection 3.5 in detail.

3.4.2 Traffic flow model: SOUND estimates the number of vehicles and the positions of individual vehicles on each link at every time step according to the fundamental diagram. Here, a simple fundamental diagram is assumed as shown in Fig. 5. It is defined by link capacity q_c and jam density d_j. In a free-flow regime (blue line in Fig. 5), forward wave speed, which is the angle from the horizontal axis, decreases
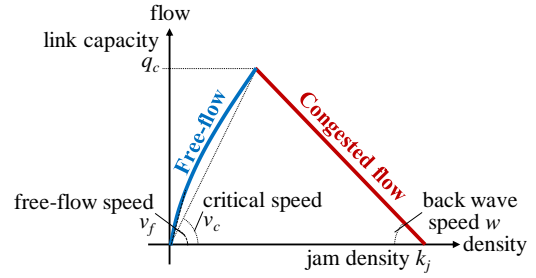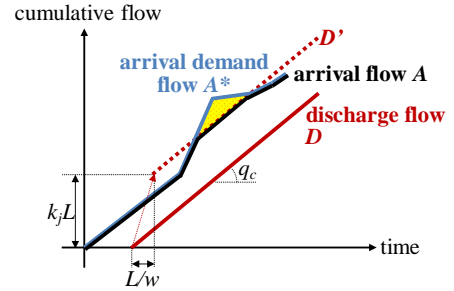

Fig. 5. *Fundamental diagram*

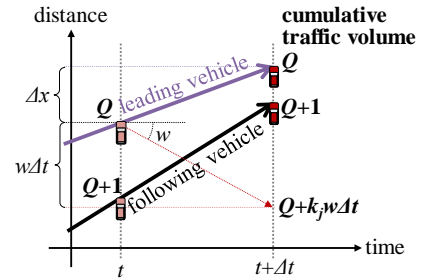**Fig. 6.** *Estimation of cumulative arrival flow*

Fig. 7. *Impact of a probe vehicle trajectory*

from free-flow speed v_f to critical speed v_c as density increases. In a congested flow regime (red line in Fig. 5), backward wave speed w is calculated as follows.

$$w = \frac{q_c}{k_j - {q_c}/{v_c}} \quad (1)$$

In SOUND, expressways are segmented into links on which vehicles are conserved (no entry and exit exist). Then the number of vehicles on each link is first calculated as follows:

In principle, each vehicle is put into the list of free-flow states when entering into a link, and transferred into the list of ready-to-discharge states after spending the link travel time by free-flow speed, which is expressed by L/v_f where L denotes link length, while keeping the first-in-first-out. Then, vehicles in the list of ready-to-discharge states can move to the downstream link according to the cumulative arrival flow curve, which is obtained by applying Newell's simplified kinematic wave theory [18].

Fig. 6 illustrates an example of the cumulative flow diagram of a link. The cumulative arrival-demand flow curve $A*$ is determined by the cumulative number of vehicles in the list of ready-to-discharge states of an upstream link (blue solid line). On the other hand, for its downstream link, the cumulative discharge flow curve D at its downstream end is drawn with the slope of its link

4

Table 1 parameters of the links
(a)MEX

Parameters	6-lane	4-lane	IC and JCT
q_c [pcu/h/lane]	2200	2200	2000
k_j [pcu/km/lane]	80		
v_f [km/h]	100	80	60
v_c [km/h]	60	54	36

(b)NEXCO East and Central

Parameters	4-/ 6-lane	2-lane	IC and JCT
q_c [pcu/h/lane]	2400	2000	2000
k_j [pcu/km/lane]	80		
v_f [km/h]	100	80	60
v_c [km/h]	70	54	36

Table 2 Coefficients of the route choice model

Vehicle type	Passenger cars		Heavy vehicles	
Weekday/end	Weekday	Weekend	Weekday	Weekend
α	0.240	0.242	0.257	0.341
β	0.330	0.311	0.219	0.181
γ	-	-	-	-
θ	0.064	0.064	0.055	0.055

Note: "-" means not available

objects on roads) that cannot be calibrated in advance. This is the great advantage of incorporating probe data, especially for further development of a real-time traffic monitoring system.

3.4.4 Route choice model: SOUND assumes that road users have perfect information and choose their routes dynamically based on updated travel time and tolls of their alternative routes every fifteen minutes. Thus, user route choice probability is formulated by the logit model as follows.

$$p_j = \frac{\exp(-\theta c_j)}{\sum_{k=1}^{J} \exp(-\theta c_k)} \quad (2)$$

$$c_j = \alpha \times time_j + \beta \times \frac{distance_j}{v_{fj}} + \gamma \times toll_j \quad (3)$$

Where, p_j: probability to choose route j, J: total number of available routes, c_j: cost of route j, $time_j$: expected travel time of route j based on link travel time of the current time step (every fifteen minutes), $distance_j$: distance of route j, v_{fj}: free-flow speed of route j, $toll_j$: toll of route j, θ: route choice sensitivity parameter and α, β and γ: coefficients.

Coefficients α, β and γ were set as listed in Table 2. These coefficients were estimated based on the observed trajectories of the ETC 2.0 probe vehicles (for dependent variables) and route travel times from the detector data (for independent variables) by the maximum likelihood method. The data used for this estimation was observed during June of 2016 and 2017, which contains the trajectories of 1,034,542 vehicles in total (i.e., 380,370 vehicles in 2016 and 654,172 vehicles in 2017).

In the above estimation, the impact of tolls represented by the coefficient γ could not be estimated. This was because tolls are determined by only origins and destinations regardless of routes for almost of OD-pairs in the subject network after April 2016 until today; in other words there was no difference in tolls when drivers chose their routes. Here, it should be noted that there was an exception in the toll policy that tolls on the routes using MEX's sections costed higher than other routes for several OD-pairs, in order to reduce number of vehicles penetrating the centre of the metropolitan area. However, the impact of this exception could not be reflected in the model due to the lack of the detailed toll data. In addition, please note that the above-mentioned toll policy has been applied only for ETC users (about 90% of all vehicles) in reality, but the model was applied for all users in the simulation for simplicity.

3.5. Calibration of OD matrix using detector data
The OD matrix was calibrated using detector data as

capacity q_c (red solid line). Because of backward wave propagation, the acceptable cumulative flow curve D' at the upstream end of this link can be obtained by shifting the curve of D with L/w to the right and k_jL to the top (red dotted line). Finally, the cumulative arrival flow curve A is the minimum of $A*$ and D (black solid line).

After that, with the given boundary of the cumulative number of vehicles at the downstream end of each link, the positions of individual vehicles inside the link are calculated one by one from the downstream to upstream by applying Daganzo's variational theory [2]. Since not only the downstream end of links but also the observed vehicle trajectories of probe vehicles can be regarded as boundaries, we could incorporate probe data in this simulation. Fig. 7 illustrates how to estimate the position of the following vehicle based on the leading vehicle trajectory. As shown in Fig. 7, when the cumulative traffic volume of the leading vehicle is Q at time t, that of the vehicle at $w\Delta t$ downstream of the leading vehicle at time $t + \Delta t$ is calculated as $Q + k_j w \Delta t$. Assuming the vehicles are linearly distributed between them, SOUND estimates the position of the following vehicle of which the cumulative traffic volume is $Q + 1$ at time $t + \Delta t$.

The parameters for determining a fundamental diagram are traffic capacity q_c, jam density k_j, free-flow speed v_f and critical speed v_c. Usually they need to be calibrated especially at bottlenecks based on empirical analysis, so that congestion phenomena can be reasonably represented. However, since this approach requires a large amount of empirical data, these parameters were uniformly set according to Table 1 in this research. Instead of calibrating these parameters, probe data was incorporated for representing congestion.

3.4.3 Incorporation of probe data: Specifically, in this simulation, we controlled the speeds of some vehicles so that they could precisely track the observed trajectories of probe vehicles. Then, the speed reduction of a probe vehicle influenced the following vehicles as explained in 3.4.2. As a result, this triggered congestion in the simulation.

If the number of probe vehicles was not so small, most of the traffic congestion was captured by this approach. Furthermore, although it is outside the focus of this paper, by using this approach, it becomes possible to grasp the congestion caused by not only fixed bottlenecks but also some temporary incidents (i.e., traffic crashes or falling

shown in Fig. 2. That means the OD matrix was iteratively revised in order to minimize the sum of errors in estimated link traffic volumes from the observed volumes in detector data.

For this calibration, we used the mathematical model, which had been proposed by Kobayashi et al. [19] and applied by Hanabusa et al. [14]. The advantage of this approach was that the number of iterations could be drastically reduced. The previous application to the entire Japan expressway network [14] needed only three to five simulation runs to converge the results. This was important since it took time to run traffic simulation.

This model minimizes the sum of errors in estimating link traffic volume E, which is expressed by the following equation.

$$E = \sum_i \sum_t \left(q_{i,t} - \widehat{q_{i,t}} \right)^2 \quad (4)$$

Where, $q_{i,t}$: the estimated traffic volume of link i in time step t in the simulation, $\widehat{q_{i,t}}$: the observed traffic volume of link i in time step t from detectors.

In the above equation, although $q_{i,t}$ cannot be generally obtained without running the simulation, it is mathematically expressed with the relation to OD traffic volumes as follows.

$$q_{i,t} = \sum_\omega \sum_\tau p_{i,t}^{\omega,\tau} Q^{\omega,\tau} \quad (5)$$

$$p_{i,t}^{\omega,\tau} = \frac{q_{i,t}^{\omega,\tau}}{Q^{\omega,\tau}} \quad (6)$$

Where, $Q^{\omega,\tau}$: traffic volume of OD pair ω that departed in time step τ, $p_{i,t}^{\omega,\tau}$: probability that a vehicle of OD pair ω that departs in time step τ passes link i in time step t, $q_{i,t}^{\omega,\tau}$: the estimated traffic volume of OD pair ω that departed in time step τ passed link i in time step t in the simulation.

Assuming a binary choice for the logit model (c.f., equation (2)), $p_{i,t}^{\omega,\tau}$ is also expressed by the following equation.

$$p_{i,t}^{\omega,\tau} = \frac{1}{1 + exp\left(-\theta \Delta c_{i,t}^{\omega,\tau} \right)} \quad (7)$$

Where, $\Delta c_{i,t}^{\omega,\tau}$: difference in costs for a vehicle of OD pair ω that departed in time step τ passing link i in time step t from passing aother link under the alternative routes.

By substituting equation (7) into (5), $q_{i,t}$ is expressed as a function of $Q^{\omega,\tau}$ (subject for this calibration), θ (given parameter from Table 2) and $\Delta c_{i,t}^{\omega,\tau}$ as follows.

$$q_{i,t} = \sum_\omega \sum_\tau \frac{Q^{\omega,\tau}}{1 + exp\left(-\theta \Delta c_{i,t}^{\omega,\tau} \right)} \quad (8)$$

Generally, it is difficult to obtain $\Delta c_{i,t}^{\omega,\tau}$. In this model, $\Delta c_{i,t}^{\omega,\tau}$ is estimated by solving the simultaneous equations (6) and (7), assuming that $\Delta c_{i,t}^{\omega,\tau}$ is calculated by the simulation result and constant while optimizing $Q^{\omega,\tau}$.

Table 3 percentage of ETC users (passenger cars)

Month	MEX	NEXCO East	NEXCO Central
Jun. 2016	94.7	89.6	91.8
Jun. 2017	95.5	88.9	92.2

$$\Delta c_{i,t}^{\omega,\tau} = -\frac{1}{\theta} \ln \left(\frac{Q^{\omega,\tau}}{q_{i,t}^{\omega,\tau}} - 1 \right) \quad (9)$$

As a result, equation (4) becomes differentiable at $Q^{\omega,\tau}$ with the following constraints.

$$Q^{\omega,\tau} \geq 0 \quad (10)$$

$$\sum_\tau Q^{\omega,\tau} = Q^\omega \quad (11)$$

Where, Q^ω: total traffic volume of OD pair ω.

3.6. Output

In order to evaluate operational measures, it is necessary to see the impacts with several indices at different scale, because some measures may have a positive impact on a certain section but a negative impact on a whole network, and vice versa. For example, link speeds and link traffic volumes are necessary to see congested conditions of individual road sections; on the other hand, we also need to calculate the total delay in a whole network. Because the proposed framework in this study could simulate movements of all vehicles on a network, these indices could be calculated at various scales.

Furthermore, as mentioned in chapter 1, the purpose of traffic estimations in this research was to understand traffic flow distributions that make it possible to think about how to operate road networks more efficiently by diverging some traffic into alternative routes. For this purpose, the proposed framework could calculate route-sharing rates for any route of an arbitrary OD pair by dividing the number of subject OD traffic volumes that used the subject route by the total OD traffic volume. This is one of the most important outputs of the proposed framework, since it cannot be directly obtained by the observed data in most of the routes and OD pairs.

4. Validation

This section introduces a validation to see whether the simulated results can reasonably represent actual traffic conditions or not.

4.1. Outline of the estimated scenarios

We estimated traffic conditions on Monday, 6th June in 2016 and Monday, 5th June in 2017 using the observed probe and detector data on these days.

The average percentages of ETC users were as listed in Table 3 during the study periods. We used this values to convert the ETC-OD matrix into the initial OD matrix as explained in 3.3.2.

As mentioned already, tolls were determined by only origins and destinations regardless of routes. That resulted in the route choice model used in this study insensitive to tolls as shown in 3.4.4. Therefore, we did not input tolls for all links in the simulation.

We used a personal computer with Intel Core i7-5820K 3.30 Ghz CPU, SSD 128 GB. In order to get the results converged, we needed to repeat the process of calibration of OD matrix two times. That means, in total, we run the traffic simulation three times (two times during the calibration of the OD matrix and one time with probe data incorporation). It took approximately 10 hours in total: 2.5 hours for running the traffic simulation and 1.1 hours for estimating the OD matrix by the iteration with the mathematical model respectively. Although the calculation time could be shortened by applying the mathematical model for the estimation of the OD matrix as mentioned in subsection 3.5, it was still long. That must be improved in the future work.

4.2. Estimation of link speeds and traffic volumes

Link speeds and link traffic volumes were the most fundamental outputs of the traffic estimations, which could be used for other indices such as the total delay. Therefore, in order to verify the simulated results, these estimated values were compared with the ones observed by detectors.

4.2.1 Link speed: Fig. 8 and Fig. 9 are examples of the speed contours which show 5-minute link speeds on sections of the E1 expressway and the E4 and S1 expressways where congestion often occurred. The speeds in the upper figures *(a)* were observed by detectors and the ones in the lower figures *(b)* were estimated by the simulator.

We found that congestion, where the link speeds were lower (indicated by reddish colour in the speed contours), could be represented in the simulation mostly at the same location and time with the observed data. For example, in Fig. 8, the congestion that happened from the mainline tollgate and the JCT for the C2 expressway in the morning and evening hours were indicated in both observation and estimation. However, the link speeds were estimated much lower than the observation especially after the beginning of the congestion from the JCT for the C2 expressway in the morning. This happened because sometimes there was a difference in travel speeds by lane, especially at merging and diverging sections; in other words, sometimes one lane was congested while the other was not congested depending on which direction these lanes were connected. Because the differences in traffic conditions by lane could not be considered in the traffic flow model applied in the simulation, if one probe vehicle was driving on a congested lane, the speed reduction was reflected on the cross-section even though the other lane were not congested. Such kinds of overestimations of speed reductions were also found on other sections, typically around merging and diverging sections.

On the other hand, the beginning of the congestion in the evening in Fig. 8 delayed about 10 minutes in the simulation from the observation. For another example, in Fig. 9, the link speeds in the morning from the JCT for the C2 expressway were overestimated. One of the reasons of these examples was because speed reductions in this simulation were triggered by those of probe vehicles only. Therefore congestion did not occur in the simulation until a probe vehicle reached the actual congested section, and speed reductions were less reflected if there were few probe vehicles.
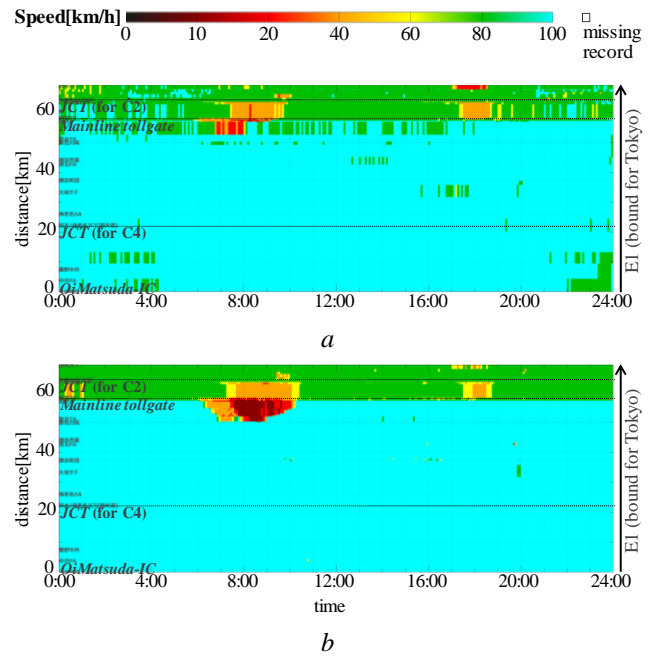

Fig. 8. *Comparison of speed contours: the E1 expressway on 5th June, 2017 (a) observed by detectors and (b) estimated by simulator*
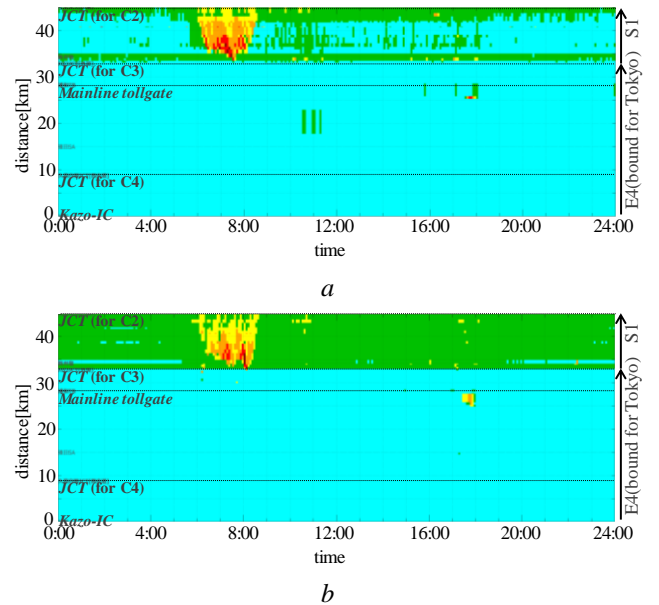

Fig. 9. *Comparison of speed contours: the E4 and S1 expressways on 6th June, 2016 (a) observed by detectors and (b) estimated by simulator*

In addition, in Fig. 9, we found that the estimated link speeds during uncongested periods did not agree with the observed speeds on the S1 expressway (the upper part of the figure). This was because the free-flow speeds were simply inputted as listed in Table 1, and not calibrated considering the geometry and other influencing factors of each link.

As a whole network, Fig. 10 shows the distributions of the estimated link speeds according to the observed link speeds for every 10 km/h in all links. As shown in Fig. 10, link speeds were overestimated in many links, especially when the observed speeds were low. This implied that
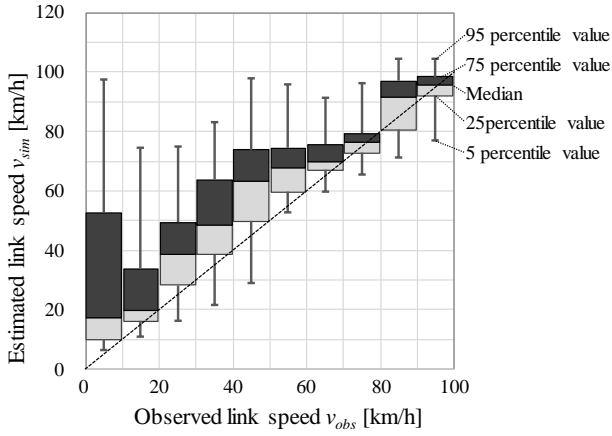
7

Fig. 10. *distributions of estimated link speeds by range of observed ones on 5th June, 2017*

Fig. 11. *comparison of estimated and observed link traffic volumes on 5th June, 2017*

congestion and speed reductions were not sufficiently represented in the simulation because of the reasons mentioned above.

4.2.2 Link traffic volume: Fig. 11 shows the scatter diagram of the observed and estimated link volumes. We found that traffic volumes were well estimated in general although there were still some links with over- and underestimations.

Fig. 12 shows the correlation coefficients between the estimated and observed link traffic volumes by time of day of the two scenarios in 2016 and 2017. The result showed that the estimated link traffic volumes were highly correlated with the observed volumes in all the time units, but actually there were some fluctuations. Because link speeds were over- and underestimated on some sections as shown in 4.2.1, it must affect the estimation of link traffic volumes, too though both the traffic flow model and the route choice model. Analysis in more detail (e.g., difference in accuracy by traffic flow level, impact of the number of probe vehicles) needs to be conducted to identify the reasons in order to improve the accuracy of the estimated results.

5. Conclusion

In this research, we developed the framework to estimate traffic conditions by fusing probe and detector data into traffic simulations, and applied it on the Tokyo Metropolitan Expressway Network. By comparing the estimation results with the detector data, it was confirmed that the estimation showed a good accuracy in terms of link traffic volumes. Traffic congestion at the typical bottlenecks could be reasonably represented to some extent.

However, the link speeds were likely to be underestimated especially at merging and diverging bottlenecks, but also overestimated in many cases. Against this issue, more analysis in detail in each bottleneck needs to be conducted in order to identify the causes of low accuracy considering the impact of the number of prove vehicles and make necessary improvement for each section. Besides, the impact of such errors in link speeds on the accuracy of the estimated traffic volumes needs to be investigated.

Furthermore, in the current framework, the impact of tolls cannot be taken into account because it was not included in the route choice model. We would like to
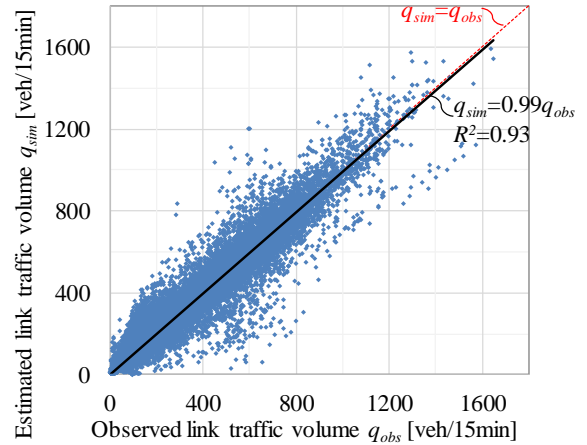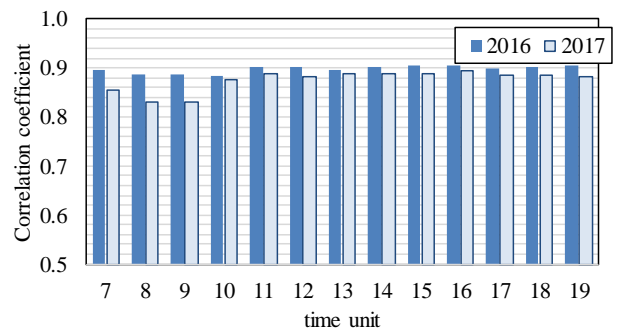

Fig. 12. *Correlation coefficient between observed and estimated link traffic volumes by time of day*

improve this by obtaining the detailed toll data and analysing the impact of tolls in detail for each OD pairs, so that the framework will be applicable to evaluate the impact of changing tolls.

After these validation and improvements, we would like to conduct case studies to evaluate the operational measures.

References

[1] 'Kanto Regional Development Bureau, Ministry of Land, Infrastructure, Transport and Tourism: The history of three-ring expressways plannning', http://www.ktr.mlit.go.jp/honkyoku/road/3kanjo/history/chronologic.htm, accessed 18 January, 2018 (in Japanese)

[2] Daganzo, C. F., Menendez, M.: 'A variational formulation of kinematic waves: bottleneck properties and examples', Transportation Research Part B, 2005, 39, (2), pp.187-196

[3] Mahmassani, H. S.: 'Dynamic network traffic assignment and simulation methodology for advanced system management applications', Networks and Spatial

Economics, 2001, 1, (3)-(4), pp. 267-292

[4] Ben-Akiva, M., H. N. Koutsopoulos, C. Antoniou, et al.: 'Traffic simulation with DynaMIT' in J. Barcelo (Ed.): 'Fundamentals of traffic simulation' (Springer, 2010), pp. 363-398

[5] 'Aimsun Live', https://www.aimsun.com/aimsun-live/, accessed 18 January, 2018

[6] Saita, K., Kurauchi, F., Okushima, M., et al.: 'Establishment of HEROINE (Hanshin Expressway Real-time Observation-based & Integrated Network Evaluator)', Proceedings of the 9th ITS World Congress, Chicago, USA, Oct 2002, pp.1-8

[7] Munakata, K., Tamura, Y., Warita, H. et al.: 'A case study about the traffic prediction under accidents using dynamic traffic simulation on Tokyo metropolitan expressway', Proceedings of the 16th ITS World Congress, Stockholm, Sweden, Sep 2009, pp. 1-12

[8] Hirai, S., Xing, J., Horiguchi, R., et al.: 'Development of a network traffic simulator for the entire inter-urban expressway network in Japan', Transportation Research Procedia, 2015, 6, pp.285-296.

[9] Work, D. B., Blandin, S., Tossavainen, O., et al.: 'A distributed highway velocity model for traffic state reconstruction', Applied Mathematics Research eXpress (AMRX), 2010, pp. 1-35

[10] Herring, R., Hofleitner, A., Abbeel, P., et al.: 'Estimating arterial traffic conditions using sparse probe data', Proceedings of the 13th International IEEE Conference on Intelligent Transport Systems, Madeira Island, Portugal, Sep 2010, pp. 1-8

[11] Choi, K., Chung, Y.: 'A data fusion algorithm for estimating link travel time', ITS Journal: Intelligent Transportation Systems Journal, 2002, 7, pp. 235-260

[12] Mehran, B., Kuwahara, M., Naznin, F. : 'Implementing kinematic wave theory to reconstruct vehicle trajectories from fixed and probe sensor data', Procedia Social and Behavioral Science, 2011, 17, pp. 247-268

[13] 'Japan Digital Road Map Association: Digital Road Map', http://www.drm.jp/english/drm/e_index.htm, accessed 13 July, 2018

[14] Hanabusa, H., Kobayashi, M., Koide, K. et al.: 'Development of the nowcast traffic simulation system using floating car data', Proceedings of the 3rd International Conference on Models and Technologies for Intelligent Transport Systems, Dresden, Germany, Dec 2013, pp.1-10

[15] U.S. Department of Transportation ITS Joint Program Office: 'Assessment report of US-Japan-Europe collaborative research on probe data, International probe data work group phase 2', (https://ntl.bts.gov/lib/60000/60100/60131/FHWA-JPO-16-356.pdf, 2016), pp. 82-85

[16] Yoshii, T., Kuwahara, M.: 'SOUND: a traffic simulation model for oversaturated traffic flow on urban expressways', Proceedings of the 7th World Conference on Transportation Research, Sydney, Australia, Jul 1995, pp. 1-15

[17] 'SOUND Ver.5, a traffic simulator for wide area road networks', http://www.i-transportlab.jp/en/products/sound.html., accessed 15 January 2018

[18] Newell, G. F.: 'A simplified theory of kinematic waves in highway traffic, Part II: queueing at freeway bottlenecks', Transportation Research Part B, 1993, 27B, (4), pp. 289-303

[19] Kobayashi, M., Horiguchi, R., Hanabusa, H., et al.: 'Estimation of time-dependent O-D matrix and route choice model parameter for a large scale network simulation', Proceedings of 32nd Conference of Japan Society of Traffic Engineers, Tokyo, Japan, Sep 2012, pp. 1-4 (in Japanese)