1. INTRODUCTION

Various types of transport models so far have been developed and used for travel demand forecast in transport planning. As Ortuzar and Willumsen (1990) show, there are several features of transport models which must be taken into account when specifying an analytical approach: the decision-making context, accuracy required, availability of suitable data, state-of-the-art modeling, resources available for the study, data processing requirements, and levels of training and skills of the analysts. In practice, the choice of method often depends on the analyst's preference or his/her knowledge and experience. This may be because training costs are usually high and it is sometimes better to use an existing model that is well understood than to embark on acquiring and learning to use an advanced one. However, even if the same data set is used, the estimated results may vary among the travel demand forecast methods. This could render the analysis inadequate and/or may negatively affect decision-making.

This paper analyzes empirically to what extent the choice of travel demand analysis methods impact the estimated results of travel demand. The urban rail route demand forecast in the Tokyo Metropolitan Area will be used for the empirical analysis.

As far as the traffic assignment model is concerned, Lam and Lo (2004) show the comparison of link flows estimated by various traffic assignment models in a simple traffic network. Their case study covers the following models: the user equilibrium (UE), probit-based stochastic user equilibrium (SUE), C-logit SUE, and logit-base SUE models. Although they compare the methods from the model fitness viewpoint in a simple network, they do not discuss the
computation time in a real network. This study compares the methods from both the model fitness and the computation time. Our study covers a multinomial logit model, the structured probit model, and the all-or-nothing (AON) assignment in addition to the UE and SUE assignments. This is because we empirically analyze a real transport network which was actually used in the latest urban rail planning of the Tokyo Metropolitan Area.

The paper is organized as follows. Section 2 introduces the travel demand forecast for urban rail planning in Tokyo and presents the methods which will be compared in the empirical analysis. Section 3 shows the empirical analysis using the methods. Section 4 summarizes the results and discusses their implications on the transport demand analysis.

2. METHODS

2.1 Travel Demand Forecast in Urban Rail Planning: Tokyo’s Case

Transport network planning for the urban railway in the Tokyo Metropolitan Area started in 1925. A number of urban railway plans had been proposed after World War II. See Morich et al. (2001) for the detailed history of the urban rail planning in Tokyo.

It was in 1972 that a mathematical travel demand analysis was first introduced into the urban rail planning in Tokyo. In 1985, a four-step travel demand model was introduced into the rail demand analysis. Multinomial logit (MNL) models were used for the modal choice and the rail route choice models.

The latest urban rail development plan in the Tokyo Metropolitan Area was proposed by the Transport Policy Committee and commissioned by the Minister of Transport, Japan, in 2000. The four-step travel demand model was again used for the travel demand forecast. The MNL model was used for the modal choice analysis whereas a probit model was used for the route choice analysis. This is because it is necessary to consider the commonality of routes in the rail route choice analysis. A huge urban rail network with high density has been already developed in the Tokyo Metropolitan Area. Thus, to avoid the

---

1 The plan is shown in Report No.18 of the Transport Policy Committee. This report is a proposal to the Minister of Transport about rail investments going forward. Although the report does not have legal power, the transport authority and rail operators usually follow the proposals.
enormous amount of calculation time, the probit model with a structured error component was introduced (Yai et al., 1997). We call this model as a “structured probit model.” The coefficients are estimated with the simulation method by using the GHK recursive simulator (Train, 2003).

Although the structured probit model can reduce the computation time, it still takes much longer than the MNL model. Therefore, the MNL model is still popular in practical rail route choice analysis in the Tokyo Metropolitan Area. For example, Kato et al. (2003) simulates the urban rail traffic flows in the Tokyo Metropolitan Area using the MNL model.

2.2 Analytical Requirements for Urban Rail Route Demand Analysis in Mega Cities

This paper focuses on the three factors required in urban rail demand analysis, particularly in mega cities: in-vehicle congestion; stochastic route choice; and the route choice set. This paper examines the impacts of these three factors on the estimated results.

First, it is often observed that rail commuters suffer from serious in-vehicle congestion in the morning peak hours. As the in-vehicle congestion varies among rail routes, a rail user can choose a route by considering not only the travel time/cost but also the in-vehicle congestion. Then, we may analyze the equilibrium under which any route from an origin to a destination has the same expected utility level. Second, the urban rail network in mega cities is so complicated that rail users could find it difficult to understand the network well. If it is assumed that users have incomplete information, the stochastic approach would be more suitable. Third, the urban rail network in mega cities is so dense that it enables rail-users to choose between various routes. When using the probabilistic-choice-based technique for demand analysis, we should define the individual choice set. We usually generate a choice set under the practical generation rule. However, this rule may bias the estimation results.

2.3 Methods Compared

We select the following six methods as the rail route choice models: the MNL model, structured multinomial probit (SMNP) model, UE model, the logit-based SUE model, probit-based SUE model, and AON assignment model. We will
compare them empirically with the same dataset.

**MNL**
The MNL model is a discrete-choice type model, in which a consumer chooses an option discretely by maximizing his/her utility with the random factor (Ben-Akiva and Lerman, 1985). The conditional indirect utility function of a route from one origin to the other destination is formulated as follows:

\[
U_{ijr} = U_{ijr} + \varepsilon_{ijr} = \theta_c \cdot GC_{ijr} + \varepsilon_{ijr}
\]

where \( U_{ijr} \) is the universal component of the indirect utility when the \( r \)th route is chosen from zone \( i \) to \( j \), \( \varepsilon_{ijr} \) is the error component of the utility following the independently and identically distributed (iid) Gumbel, \( GC_{ijr} \) is the generalized cost including the travel time and travel cost of the \( r \)th route from zone \( i \) to \( j \), and \( \theta_c \) is an unknown parameter. When the total volume of traffic flow from zone \( i \) to \( j \) is given as \( Q_{ij} \), the expected volume of traffic flow \( q_{ijr} \) choosing the \( r \)th route from zone \( i \) to \( j \) is expressed as

\[
E[q_{ijr}] = Q_{ij} \cdot p_{ijr} = Q_{ij} \cdot \frac{\exp(\lambda V_{ijr})}{\sum_{r'=n}^{m} \exp(\lambda V_{ijr'})}
\]

where \( p_{ijr} \) is the probability of choosing the \( r \)th route from zone \( i \) to \( j \) and \( \lambda \) is a scale parameter corresponding to the Gumbel distribution with \( \lambda^2 = \pi^2 / 6 \sigma^2 \) (\( \sigma^2 \) is the variance of the Gumbel distribution).

**SMNP**
The SMNP model was originally proposed by Yai et al. (1997). This model is one of the probit models with a simplified structure of error components with which the commonality of the routes can be taken into account. The model is formulated as follows. First, the additive function of a conditional indirect utility function consisting of a systematic component and an error component \( \varepsilon_{ijr} \) is as shown in eq. (1). Second, assume that the error component is divided into the following two parts: the first part \( \eta_{ijr} \) is the systematic error component that is influenced by the length of the overlapping sections in the two routes;
and the second part $\eta_{i,r}$ is the white noise that follows the normal distribution with mean 0 and variance $\sigma_{0,i,r}^2$ that is independent among the routes. Third, assume that the variance of the systematic error is in proportion to the corresponding route’s length whereas the covariance of the systematic error is in proportion to the length of overlapping sections in the corresponding pair of routes. Then, the variance matrix of $\eta_{i,r}$ is shown as

$$\text{cov}(\eta_{i,r}^1, \eta_{i,s}^1) = d_{i,q} \sigma^2$$

where $d_{i,q}$ denotes the length of the overlapping sections of a pair of the $i$th route and the $q$th route from $i$ to $j$ and $\sigma^2$ represents the variance per unit length. Fourth, the variances of all the routes in the same O-D pair are assumed to be identical. Then, the variance matrix of the error component in the indirect utility function with respect to a specific route is shown as

$$\sum = \left( d_i \sigma^2 + \sigma_0^2 \right) \left( \begin{array}{cccc} 1 & \tilde{w}_{i,j} & \cdots & \tilde{w}_{i,s} \\ \tilde{w}_{i,j} & 1 & \cdots & \tilde{w}_{i,s} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{w}_{i,s} & \tilde{w}_{i,s} & \cdots & 1 \end{array} \right)$$

where $d_i$ represents the sum of the lengths of all the routes which connect $i$ and $j$; $\sigma_0^2$ represents the variance; $\xi$ represents the parameter with respect to the covariance; and $w_{i,q}$ is defined as follows:

$$w_{i,q} = \begin{cases} \frac{1}{d_{i,q}} & \text{if } q = r \\ \frac{d_{i,q}}{d_{i,q}} & \text{if } q \neq r \end{cases}$$

As $\left( d_i \sigma^2 + \sigma_0^2 \right)$ cannot be estimated independently, the unknown parameter with respect to the variance of the error component reduces only $\xi$. Since the probability of choosing a route cannot be derived in the closed form in the probit model, the expected likelihood can be maximized using the following simulated probability function to estimate the coefficients and parameters:

$$p_{i,r} = \int_{\mathcal{E}} f(\epsilon) \Phi(\epsilon) d\epsilon$$

where $B = \{ \epsilon: \text{ s.t. } V_{i,r} + e_{i,r} > V_{i,r} + e_{i,r}, \forall r \neq r \}$.

**UE**

The deterministic UE assignment is one of the traffic assignment techniques in which the transportation system is assumed to fall into stable equilibrium when all users maximize their individual utility including the congestion in choosing their routes (Sheffi, 1985). The conditions of the user equilibrium are
formulated as follows:

\[ q_{ij,r}^* (GC_{ij,r}^* - GC_{ij}^*) = 0 \]  

\[ GC_{ij,r}^* - GC_{ij}^* \geq 0 \]  

\[ q_{ij,r}^* \geq 0 \]

where \( q_{ij,r} \) is the volume of the traffic flow of the \( r \)th route from \( i \) to \( j \), \( GC_{ij,r}^* \) is the generalized cost of the \( r \)th route, and \( GC_{ij}^* \) is the generalized cost from \( i \) to \( j \) in equilibrium. The symbol “\(^*\)” denotes the equilibrium status. The relationship between (1) the traffic volume of a route and that of an O-D pair, (2) the traffic volume of a link and that of a route, and (3) the generalized cost of a route and that of a link are respectively shown as

\[ \sum_{r \in R_i} q_{ij,r} = Q_{ij} \]  

\[ X_i = \sum_{r \in R_i} q_{ij,r} \delta_{ij,r} \]  

\[ GC_i = \sum_{r \in R_i} GC_{ij,r} \delta_{ij,r} \]

where \( X_i \) is the traffic volume of link \( i \), \( \delta_{ij,r} \) is equal to 1 if a link \( i \) is on the route \( r \) else 0, and \( GC_i \) is the generalized cost of link \( i \). The generalized cost of the link is dependent upon the traffic flow of the link \( X_i \).

The abovementioned conditions of the user equilibrium are theoretically equal to the optimization problem,

\[ \min_{\omega} Z = \int_{0}^{X_i} GC_i(\omega) d\omega , \]

subject to eqs. (9) to (11).

This optimization problem is solved by the Frank-Wolfe method (Sheffi, 1985). The UE traffic pattern arises if each traveler has perfect knowledge about the network conditions and all travelers have identical perceptions of generalized cost.

**SUE**

In reality, travelers select routes to improve their perceived generalized costs rather than their actual generalized costs, which are not perfectly known to them. The route choice principle for this case with the perception variations is often called the stochastic user equilibrium (SUE). At the SUE, the perceived
generalized costs on all used routes are equal and less than or equal to the perceived generalized costs on any unused route for each O-D pair. The perceived generalized cost is formulated in the same manner as eq. (1) in which the perceived generalized cost is converted into utility with the error component. Equations (9) to (11) should also be satisfied at the SUE. When the error component is assumed to follow the Gumbel distribution, it is known as the logit-based SUE (Dial, 1971). The expected route flow of the logit-based SUE is expressed as eq. (2). The conditions of the SUE including eqs. (2), (9), (10), and (11) are theoretically equal to the following optimization problem:

$$\min Z = \sum_{i}^{\infty} \sum_{j}^{N} GC_{ij}(\omega)K_{ij} + \sum_{i}^{\infty} \sum_{k}^{N} Q_{ij} \sum_{k}^{N} \left[ \frac{q_{ij,k}^{*}}{Q_{ij}} \ln \frac{q_{ij,k}^{*}}{Q_{ij}} \right]$$

subject to eqs. (9) to (11).

When the error component is assumed to follow a normal distribution, it is known as the probit-based SUE (Daganzo and Sheffi, 1977). In this paper, we use the probit-based SUE with the structured error component which is the same as the SMNP.

### AON

The AON assignment is a traffic assignment method in which all the traffic flows from an origin to a destination are simply allocated to a route with the lowest cost among the alternative routes. This technique does not take into account the change in the generalized cost due to traffic congestion or in-vehicle congestion. The traffic volume assigned by this method is shown as

$$q_{ij,r} = Q_{ij} \quad \text{if} \quad GC_{ij,r} = \min \{GC_{ij,1}, GC_{ij,2}, \ldots \}$$

$$= 0 \quad \text{otherwise.}$$

### 3. EMPIRICAL ANALYSIS

#### 3.1 Data Used for the Empirical Analysis

The data used for the empirical analysis is the Tokyo Metropolitan Travel Census 2000 (Institution for Transport Policy Studies, 2000). This survey was conducted by a joint research team of the Ministry of Land, Infrastructure and Transport, Japan, and the Institute of Transport Policy Studies, Japan. The
survey comprises three surveys: the first is a paper-based interview survey on the home-to-work and the home-to-school travels of passengers with seasonal tickets; the second is a paper-based questionnaire survey on the travels of passengers without seasonal tickets; and the third is a paper-based interview survey on the service of rail operators. The passenger travel surveys were conducted in October 2000. The passenger travel data include the records of rail-use travels with an origin and destination, travel purpose, travel route, times of start, and termination with socio-demographic data. They cover the Kanto Region including eight prefectures: Tokyo, Kanagawa, Saitama, Chiba, Ibaraki, Gunma, Tochigi, and Yamanashi.

To estimate the unknown parameters in the models, a sample travel dataset was constructed with the following steps. First, we categorized the observed travels into the following four types of travels: home-to-work, home-to-school, private, and business. Second, we chose travelers with one or more alternative routes. This leads to the elimination of travelers having very short travel times. The purpose of this elimination is to analyze the rail route choice of travelers and gather the data of travelers with alternative routes. Finally, we obtained 1,390 samples for home-to-work travel, 1,191 samples for home-to-school travel, 644 samples for private travel, and 462 samples for business travel.

Then, we prepared the level-of-service data for the selected samples. For this, we constructed the rail network including the 1,877 zones, 4,850 nodes, and 9,796 links in the Tokyo Metropolitan Area. The network includes the links for the rail line, the transfer at the rail stations, and the access/egress travel to/from the rail stations; it also includes the nodes representing the origin/destination, the platforms, and the gates at the rail stations. When a rail line has more tracks in addition to the local service track, we set the additional links to the local-service-track link. We also set transfer links connecting all the pairs of all the platforms at the stations. We set three to four access/egress links from an origin/destination zone to the adjacent stations considering the road and bus service networks around the zones. Then, we constructed the dataset including the travel time, transfer time, access/egress travel time, travel cost, and in-vehicle congestion rate for each link. As far as the travel cost is concerned, we estimate the fare function by using the original rail fare table. As the rail operators set their own fare tables independently, the rail fare functions are also estimated for each rail operator². There are two reasons for

² There are more than 20 rail operators in Tokyo and most of them are private operators.
estimating the rail fare function. First, in Tokyo, all the rail operators use the distance-based fare system, in which the fare is in proportion to the travel distance. Second, the rail fare is not defined as a link-based fare but as an O-D-based fare. The rail fare function is estimated as a linear function with the initial fare and the distance-based fare. We allocate the rail fare to the links by using the estimated rail fare function. For example, the estimated initial fare is set for the link connecting the rail gate to the platform whereas the distance-based fare is set for the rail link.

3.2 Parameter Estimation of MNL and SMNP

For the simulation with the MNL and SMNP, we specify the generalized cost function of a route as follows:

\[
GC_{ijr}^{a} = C_{ijr}^{a} + \frac{1}{\theta_{C}} \sum_{k \in C} \theta_{k}^{a} X_{k,ijr}^{a}
\]

\[
= C_{ijr}^{a} + \theta_{C}^{a} / \theta_{C} \cdot T_{ijr,1}^{a} + \theta_{C}^{a} / \theta_{C} \cdot T_{ijr,2}^{a} + \theta_{C}^{a} / \theta_{C} \cdot T_{ijr,3}^{a} + \theta_{C}^{a} / \theta_{C} \cdot Cong_{ijr}
\]

(15)

where \( T_{ijr,1}^{a} \) is the access/egress travel time of the \( r \)th route from zone \( i \) to \( j \) for the travel purpose \( a \), \( T_{ijr,2}^{a} \) is the rail-ride travel time, \( T_{ijr,3}^{a} \) is the transfer time including the waiting time at the transfer station, \( C_{ijr}^{a} \) is the travel cost, \( Cong_{ijr} \) is the in-vehicle congestion, and \( \theta_{C}^{a} \) is the coefficient with respect to the corresponding variables. The in-vehicle congestion is defined as

\[
Cong_{ijr}^{a} = \sum_{l \in L_{ijr}} z_{il} \cdot T_{ijr,l}
\]

(16)

where \( z_{il} \) represents the congestion rate of a rail link \( l \) and \( T_{ijr,l} \) denotes the rail-ride travel time of the rail link \( l \). The congestion rate is defined as

\[
z_{il} = \sum_{r} \sum_{a} \sum_{u} q_{ijr,ul}^{a} / \text{cap}_{l}
\]

(17)

where \( q_{ijr,ul}^{a} \) represents the link traffic flow of link \( l \) of the \( r \)th route from zone \( i \) to \( j \) for the travel purpose \( a \) and \( \text{cap}_{l} \) is the traffic flow capacity of link \( l \).

Then, we estimate the coefficients of the MNL model and SMNP model with the empirical data shown earlier. The estimated results are shown in Table 1.

As the rail operators do not coordinate between setting a fare system, passengers have to pay the charge for each operator independently.
This shows that the models are well estimated from a statistical viewpoint. The results of the tests for most of the variables indicate that they are highly significant. The likelihood ratios are also high enough.

Then, we define the link performance functions for the UE and SUE assignment analyses. The link performance functions are expressed as follows:

### Table 1: Parameter estimation results: MNL and SMNP

<table>
<thead>
<tr>
<th>Variables</th>
<th>Units</th>
<th>MNL</th>
<th>SMNP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Home–to–work</td>
<td>Home–to–school</td>
</tr>
<tr>
<td>Rail–ride travel time</td>
<td>Minutes</td>
<td>–0.0797 (–8.8)</td>
<td>–0.0902 (–9.9)</td>
</tr>
<tr>
<td>Transfer time</td>
<td>Minutes</td>
<td>–0.1153 (–13.6)</td>
<td>–0.1335 (–13.6)</td>
</tr>
<tr>
<td>Access/egress travel time</td>
<td>Minutes</td>
<td>–0.1648 (–15.8)</td>
<td>–0.225 (–13.9)</td>
</tr>
<tr>
<td>Travel cost</td>
<td>Yen</td>
<td>–0.00347 (–5.2)</td>
<td>–0.0127 (–11.8)</td>
</tr>
<tr>
<td>In–vehicle congestion</td>
<td></td>
<td>–0.00582 (–2.2)</td>
<td>–0.00709 (–1.9)</td>
</tr>
<tr>
<td>Variance parameter</td>
<td></td>
<td>– (–)</td>
<td>– (–)</td>
</tr>
<tr>
<td>Number of observations</td>
<td></td>
<td>1,390</td>
<td>1,191</td>
</tr>
<tr>
<td>Likelihood ratio</td>
<td></td>
<td>0.228</td>
<td>0.402</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables</th>
<th>Units</th>
<th>SMNP</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Home–to–work</td>
<td>Home–to–school</td>
</tr>
<tr>
<td>Rail–ride travel time</td>
<td>Minutes</td>
<td>–0.1145 (–4.7)</td>
<td>–0.1637 (–5.9)</td>
</tr>
<tr>
<td>Transfer time</td>
<td>Minutes</td>
<td>–0.1422 (–6.4)</td>
<td>–0.201 (–7.9)</td>
</tr>
<tr>
<td>Access/egress travel time</td>
<td>Minutes</td>
<td>–0.1899 (–7.7)</td>
<td>–0.286 (–8.6)</td>
</tr>
<tr>
<td>Travel cost</td>
<td>Yen</td>
<td>–0.00319 (–3.8)</td>
<td>–0.0142 (–8.6)</td>
</tr>
<tr>
<td>In–vehicle congestion</td>
<td></td>
<td>–0.00908 (–2.0)</td>
<td>–0.0097 (–1.7)</td>
</tr>
<tr>
<td>Variance parameter</td>
<td></td>
<td>0.189</td>
<td>0.211</td>
</tr>
<tr>
<td>Number of observations</td>
<td></td>
<td>1,390</td>
<td>1,191</td>
</tr>
<tr>
<td>Likelihood ratio</td>
<td></td>
<td>0.232</td>
<td>0.418</td>
</tr>
</tbody>
</table>
Access/egress link:  \[ GC_i^n = c_i + \theta_x^n / \theta_T^n T_i \]  \hspace{1cm} (18)

Rail link:  \[ GC_i^n = c_i + \theta_x^n / \theta_T^n T_i + \theta_x^n / \theta_C Cong_{ijr} \]  \hspace{1cm} (19)

Transfer link:  \[ GC_i^n = c_i + \theta_x^n / \theta_T^n T_{ij} \]  \hspace{1cm} (20)

where \( T_{ij} \) is the travel time of link \( i \) \( (n = 1, 2, \text{ and } 3) \) and \( C_i \) is the travel cost of link \( i \). The estimated results shown in Table 1 are used for the coefficients in the link performance functions. When using the AON, we delete the in-vehicle congestion term from eq. (19) for the rail link performance function.

3.3 Traffic Flow Simulations with the Models

For the simulation with the MNL and SMNP models, we define the route choice set for each O-D pair. Although there are various ways to define it, in our analysis, we define the choice set in the following way. First, in a specific O-D pair, we search for a route with the lowest generalized cost. Second, we specify the stations by searching the shortest access/egress links between the origin/destination zones and the stations. Then, we search the rail route between the specific pairs of the specified origin stations and the specified destination stations. We list a maximum of ten routes. If the route list in the second step includes the route searched in the first step, it is eliminated from the list. Finally, the routes in the final route list are used for the choice set of the corresponding O-D pair.

Then, we simulate the traffic flows in the network with the abovementioned methods. In the simulations with the UE and SUE assignments, we calculate the traffic flows of the home-to-work travel and home-to-school travels using the multi-class assignment method. This is because the home-to-work travel flows intersect the home-to-school travel flows due to the in-vehicle congestion. We calculate the traffic flows of the private and business travels with the AON assignment under the assumption that they are free from in-vehicle congestion. This is because it is assumed that most of the private and business travels usually do not start in the peak hours of the morning.
Table 2: Comparison of iterations and computation time among the methods

<table>
<thead>
<tr>
<th></th>
<th>MNL</th>
<th>SMNP</th>
<th>UE</th>
<th>Logit-based SUE</th>
<th>Probit-based SUE</th>
<th>AON</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iterations</td>
<td>1</td>
<td>1</td>
<td>13</td>
<td>21</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td>Computation time (Minutes)</td>
<td>10</td>
<td>40</td>
<td>38</td>
<td>174</td>
<td>534</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 2 shows the comparison of the calculation time with the iteration times among the methods. First, the computation time required by the SMNP is about four times longer than that of the MNL, while the SMNP requires almost the same computation time as the UE. Second, the SUE requires three to nine hours with 21 iterations to complete the traffic assignment. Third, the AON can complete the simulation in the shortest computation time.

Figure 1 shows the comparison of the estimated link flows with the observed link flows among the six methods. Table 3 summarizes the results of the model fitness of the six methods.

These results show the following. First, Table 3 shows that the probit-based SUE has the best fitness followed by the logit-based SUE. However, the MNL has the worst fitness and the SMNP has the second worst fitness. This is probably because that the SUE and UE take into account the equilibrium unlike the MNL and SMNP. Second, surprisingly the AON assignment has a better fitness than the MNL and SMNP. This is probably because the definition of the choice set biases the simulation results. Third, the fitness of the UE is better than that of the AON. This indicates that it is worthwhile taking the in-vehicle congestion into account in the rail route choice. Fourth, the fitness of the SMNP is better than that of the MNL, whereas the fitness of the probit-based SUE is better than that of the logit-based SUE. This means that the estimation results can be improved by considering the route commonality. Fifth, Table 2 shows that the computation time with the probit-based SUE is three times longer than that with the logit-based SUE, whereas the computation time with the SMNP is four times longer than that with the MNL. This is because the simulation in probability calculations requires additional computation time. Sixth, the computation time with the probit-based SMNP reaches almost nine hours. This may be too long from a practical viewpoint. For example, the latest urban rail planning in the Tokyo Metropolitan Area in 2000 examined over 1000 cases in six months (Morichi et al., 2001). If we would use the probit-based SMNP with a single computer, it would take more than five months only for
predicting the future traffic demand, which is not realistic.

Figure 1: Comparison of estimated vs. observed link flows among the assignment methods
4. CONCLUSIONS

This paper compares six traffic assignment methods with the same empirical dataset of the route choice. The multinomial logit (MNL), structured multinomial probit (SMNP), user equilibrium (UE) assignment, logit-based stochastic user equilibrium (logit-based SUE) assignment, probit-based stochastic user equilibrium (probit-based SUE) assignment, and all-or-nothing (AON) assignment are applied to the comparative empirical analysis. The revealed preference data of urban rail route choice in the Tokyo Metropolitan Area thus obtained is used for the empirical case analysis. The empirical case analysis shows that the probit-based SUE simulates link traffic flows with the highest accuracy. However, the probit-based SUE requires almost nine hours of computation time. It also shows that the route commonality and in-vehicle congestion significantly influence the accuracy of the simulation. We thus conclude that in the Tokyo urban rail choice analysis, it is critical to take these factors into consideration.

In the latest urban rail planning in the Tokyo Metropolitan Area, the planning process requires the computation time of traffic demand analysis to be less than an hour. Note that the traffic demand analysis includes not only the rail route choice analysis but also the modal choice. This makes it difficult for transport demand analysts and urban rail planners to select an analysis method in the urban route choice analysis; this is because the computation time using an assignment method with high accuracy is longer than the required time.

There are three potential solutions to the abovementioned problems for future rail planning in Tokyo. The first way is simply to increase the computation time. If more time is given for the travel demand analysis in the planning process, we could increase the computation time. However, on the whole, this makes the planning process longer. The second way is to make the required accuracy

<table>
<thead>
<tr>
<th></th>
<th>MNL</th>
<th>SMNP</th>
<th>UE</th>
<th>Logit-based SUE</th>
<th>Probit-based SUE</th>
<th>AON</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation</td>
<td>0.955</td>
<td>0.949</td>
<td>0.974</td>
<td>0.974</td>
<td>0.975</td>
<td>0.968</td>
</tr>
<tr>
<td>RMS errors</td>
<td>53483.4</td>
<td>57385.9</td>
<td>42909.7</td>
<td>41325.3</td>
<td>41543.7</td>
<td>46852.2</td>
</tr>
</tbody>
</table>
lower than that obtained with the probit-based SUE. As the SMNP was used in the latest planning process in Tokyo, decreasing the accuracy would be acceptable. The simplification of the zoning and/or network can also contribute to the decrease in computation time although the accuracy may be reduced. However, this modeling would not be state of the art. The third way is by introducing a new method. For example, the C-logit model (Cascetta et al., 1997) may be an alternative to the SMNP. As the choice probability in the C-logit is the same as that of the MNL (except the commonality factor in the utility function), the computation time with the C-logit SUE is expected to be almost the same as that with the logit-based SUE. However, the logit-based SUE still requires almost three hours for a unit computation, which is much longer than the required computation time. Thus, there is still a need for developing a more efficient assignment method in the future.

Acknowledgement

We are grateful to Mr. Shio Hayasaki (Creative Research and Planning Co.) for his support of the data analysis.

References

Dial, R. B. (1971) A probabilistic multi-path traffic assignment model which obviates the need for path enumeration, Transportation Research, 5, 83-111.
in the Tokyo Metropolitan Area. (in Japanese)