INSTITUTE OF POLICY AND PLANNING SCIENCES

Discussion Paper Series

No. 1102

Modeling with GIS: OD Commuting Times by Car and Public Transit in Tokyo

by
Mizuki Kawabata, Akiko Takahashi

December, 2004

UNIVERSITY OF TSUKUBA
Tsukuba, Ibaraki 305-8573
JAPAN

Modeling with GIS: OD Commuting Times by Car and Public Transit in Tokyo

Mizuki Kawabata
Department of Risk Engineering
Graduate School of Systems and Information Engineering
University of Tsukuba
Tsukuba, Ibaraki, JAPAN
Email: *mizuki@risk.tsukuba.ac.jp*

and

Akiko Takahashi
Center for Spatial Information Science
University of Tokyo
Meguro-ku, Tokyo, JAPAN
Email: *akuri@csis.u-tokyo.ac.jp*

Abstract

This paper proposes a GIS modeling methodology for estimating OD commuting times by car and public transit in the Tokyo metropolitan region. Using GIS and spatial data on roads, railroads, and stations, we first compute OD travel times between centroids of basic planning zones, a spatial unit used in the 1998 Tokyo Metropolitan Region Person-Trip Survey (PT Survey). Next, the GIS-modeled OD times are adjusted through regression modeling using actual average OD commuting times reported in the PT Survey. The work is an initial attempt to develop a modeling method that can readily be applicable to other major urban areas. Limitations and directions of future refinements are discussed.

1. INTRODUCTION

Origin to destination (OD) travel times are frequently used for urban and transportation modeling. In Japan, OD travel times by travel mode for major metropolitan areas are available from the Person Trip Surveys (PT Surveys). Since the PT Survey data are sample-based, however, travel times for a complete set of OD pairs are usually not available, especially for small spatial units. When travel times in a given metropolitan area are needed for a certain mode for a particular purpose, researchers often estimate travel times in various ways.

In his study on the extent to which commuting is wasteful due to inefficient balance of jobs and housing in the Tokyo metropolitan area, Suzuki (1992), for example, creates a railway network database and estimates OD travel times by constructing a model where travel times for shortest-distance paths are used as the dependent variable. Hino et al. (1999) use GIS to calculate average travel times on existing bus network and new transit system routes to evaluate public transit routes in the Sapporo Northern Area.

In order to examine the effect of railways on in-region travel time, Koshizuka et al. (2003) construct railway network and calculate travel times between 100 points that are uniformly distributed across the Tokyo metropolitan region. Building transit network with GIS, Yoshinaga (2003) computes travel times on railroad links, which he uses to examine changes in time-distance areas from stations in the Tokyo metropolitan area.

Metropolitan-wide OD commuting times by travel mode for OD zones that are widely used and relatively small in size such as cities would be useful for a variety of purposes, yet a systematic method to estimate such OD commuting times has not been proposed explicitly. Recent Geographic Information Systems (GIS) provide a wide range of tools that allow us to manage and analyze spatial information with ease. Using GIS and spatial data on roads, railroads, and stations, in this paper we propose a modeling methodology that systematically estimates OD commuting times by car and public transit in the Tokyo metropolitan region. The OD zone employed in this study is the basic planning zone, a PT Survey zone widely used in urban and transportation planning. Our goal is to develop a GIS modeling method that can readily be applicable to other major urban areas.

The remainder of the paper starts with a description of the study area and data in section two. In section three the modeling methodology for estimating OD commuting times by car and public transit is presented, and in section four limitations and directions of future refinements are discussed.

2. STUDY AREA AND DATA

The study area is the Tokyo metropolitan region as defined for the 1998 PT Survey. Covering Tokyo, Kanagawa, and Chiba Prefectures and the southern portion of Ibaraki Prefecture, the region comprises about 15,000 square kilometers and accommodates

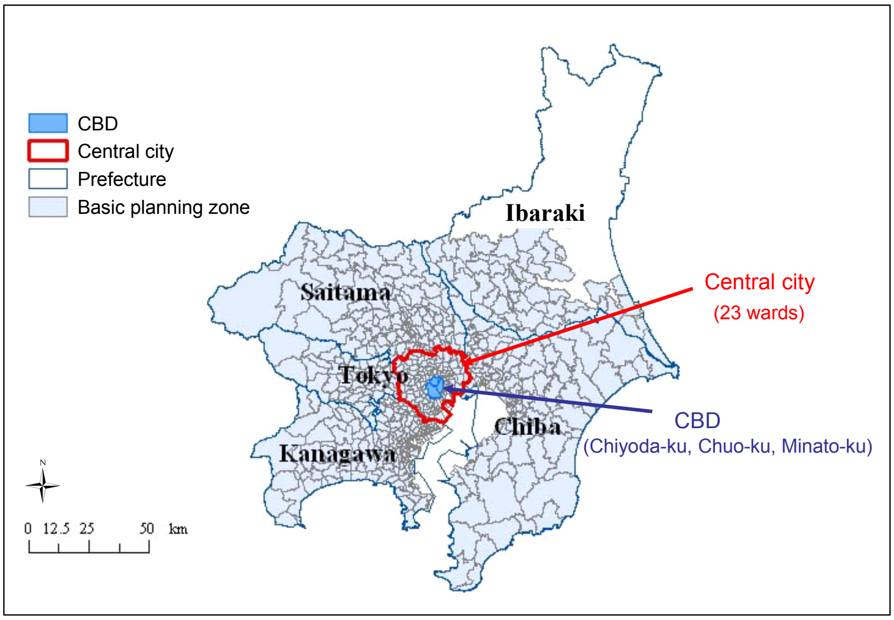

Figure 1. Geographic boundaries of study area

about 35 million people and 19 million workers (persons in the labor force). The geographic unit used in this study is the basic planning zone that is defined especially for the Tokyo PT Survey. The Tokyo metropolitan region has 595 basic planning zones. The geographic boundaries of the study area are shown in Figure 1.

The modeling of OD commuting times uses the following data: spatial data on roads, railroads, and stations extracted from 2001 GISMAP 25000V of Hokkaido-Chizu Co, Ltd.; and OD average commuting times including waiting times by car and public transit during morning peak hours (7:00-9:30am) compiled from the 1998 PT Survey.

3. MODELING OD COMMUTING TIMES

Because the PT Survey data are sample-based, a considerable number of OD pairs do not have reported commuting times. Out of 354,025 OD pairs (=595×595 zones), 286,521 (81%) pairs for car users and 252,520 (71%) pairs for public transit users have no reported commuting times. In this study we estimate OD commuting times for 353,430 inter-zone OD travels (=595×594 zones) between centroids of basic planning zones. The estimation is made through two modeling parts: GIS modeling and regression modeling. ArcView 3.2 and ArcGIS 8.3 Desktop are used as GIS software. In the following two subsections, the steps to model OD commuting times by car and public transit, respectively, are described.

3.1 Modeling OD Commuting Times by Car

We estimate inter-zone OD commuting times by car as shown in Figure 2.

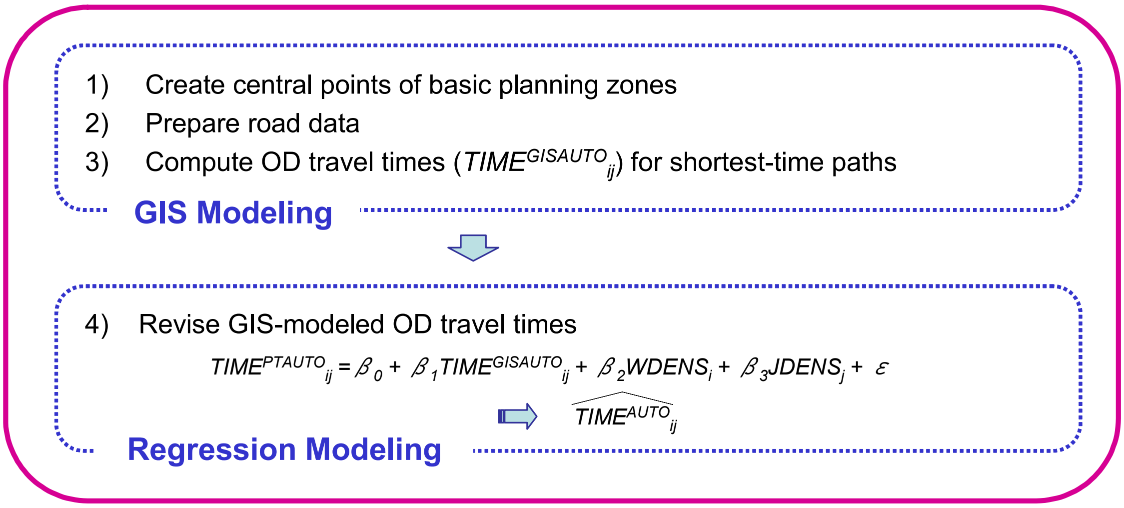

Figure 2. Estimation of OD commuting times by car

<u>GIS Modeling Part</u>

The first step of the GIS-modeling part is to create 595 central points of basic planning zones. The second step is to prepare road data as follows. 1) Firstly, road links are

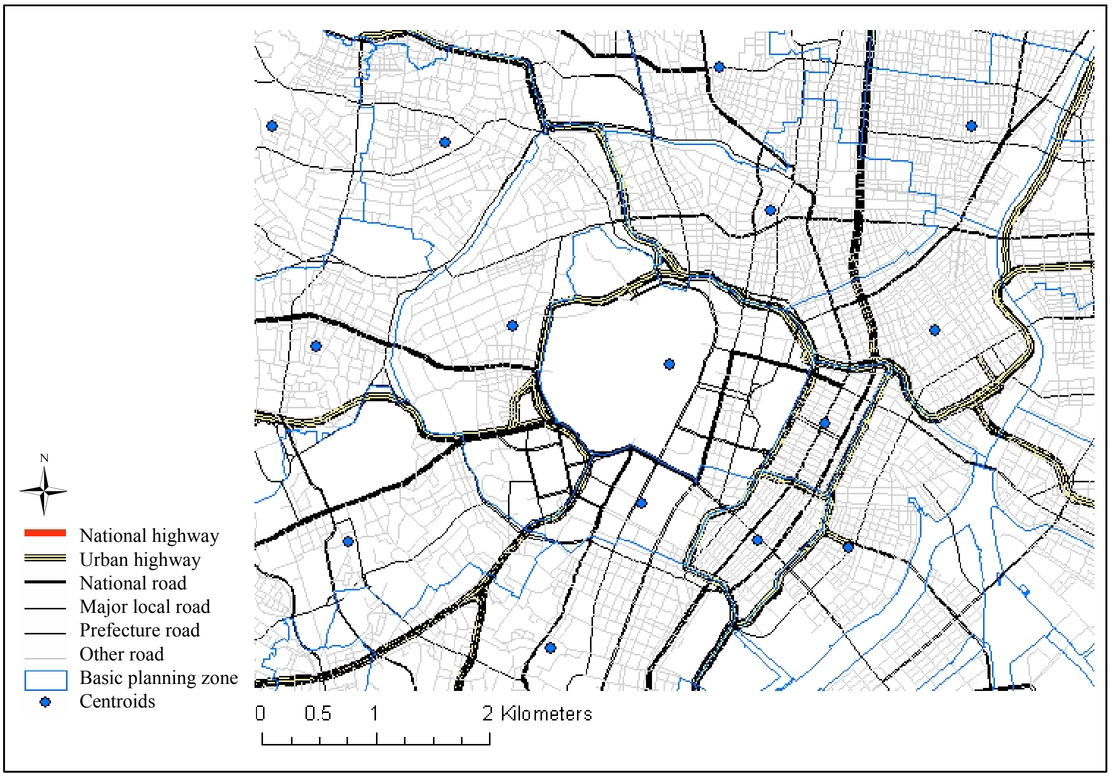

Figure 3. Roads and centroids of basic planning zones

extracted from GISMAP 25000V. The central points and extracted roads are shown in Figure 3. 2) Secondly, the length of each link is calculated. 3) Thirdly, average speed is assigned to each road link based on road type as indicated in Table 1. We elected to use the average speed during weekday peak time by road type for the Kanto-Rinkai block, obtained from the 1999 Road Transportation Census. The Kanto-Rinkai block comprises four out of five Tokyo study area prefectures: Tokyo, Kanagawa, Saitama, and Chiba. 4) Fourthly, for each road link travel time is calculated by dividing length by speed.

The third step is to compute OD travel times ($TIME^{GISAUTO}_{ij}$) using the road data. This step is done by finding a minimum travel-time path for each inter-zone travel and then calculating total travel time for each path.

Table 1. Average speed by road type

Road type code[*1)]	Road type[*1)]	Weekday average speed during peak time[*2)]
1	National highway	72.5 km/h
2	Urban highway	35.2 km/h
3	National road	26.5 km/h
4	Major local road	24.4 km/h
5	Prefecture road	25.6 km/h
6	Other road	19.4 km/h

Source: [*1)] 2001 GISMAP25000V. [*2)] 1999 Road Transportation Census.

Regression Modeling Part

Next, we proceed to the regression-modeling part in which GIS-modeled OD commuting times are adjusted using the actual commuting times reported in the PT Survey. Table 2 shows the basic statistics of the GIS-modeled OD travel times from zone i to zone j ($TIME^{GISAUTO}_{ij}$) and PT Survey's average OD commuting times from zone i to zone j ($TIME^{PTAUTO}_{ij}$). For the GIS-modeled OD travel times, statistics for both all inter-zone travels (353,430 OD pairs) and those inter-zone travels where the PT data are available (66,911 OD pairs) are presented for the comparative purpose.

Table 2. Descriptive statistics of OD commuting times by car

Variables	Number of observations	Mean	Std. Dev.	Min.	Max.
GIS-modeled OD travel time ($TIME^{GISAUTO}_{ij}$)	353,430	101.0	51.6	2.2	415.4
	66,911*	53.9*	33.1*	2.2*	357.1*
PT Survey OD commuting time ($TIME^{PTAUTO}_{ij}$)	66,911*	67.5*	41.3*	2.0*	960.0*

Note: *Statistics are for inter-zone OD travels where the PT Survey data are available.

That the mean GIS-modeled travel time for inter-zone travels where the PT Survey data are available (53.9 minutes) is shorter than that the mean PT Survey travel time (67.5 minutes) is likely related to the fact that the former does not include waiting times while the latter does.[1]

Next, using Wang's (2003) approach, we estimate the following regression:

$$TIME^{PTAUTO}_{ij} = \beta_0 + \beta_1 TIME^{GISAUTO}_{ij} + \beta_2 WDENS_i + \beta_3 JDENS_j + \varepsilon, \qquad (1)$$

where $TIME^{PTAUTO}_{ij}$ is the average commuting time from zone i to zone j obtained from the PT Survey, and $TIME^{GISAUTO}_{ij}$ is the GIS-modeled travel time from zone i to zone j. The density of workers (the number of workers per square kilometers) in origin zone i is given by $WDENS_i$, and the density of jobs (the number of jobs per square kilometers) in destination zone j is represented by $JDENS_j$. The error term is indicated by ε.

The numbers of workers and jobs are extracted from the 2000 Population Census of Japan. The census data by city are converted to data by basic planning zone and then the densities are calculated. The density variables are included to capture congestion in residence and workplace zones. We would expect that increases in the densities imply rise in congestion, which in turn increases commuting times. Congestion in the other zones along the commuting paths is not included for the sake of simplicity. The regression results are presented in Table 3.

All estimated coefficients are significant at the one percent level. As expected, both

[1] The median travel time for inter-zone travels where the PT Survey data are available from the GIS modeling and from the PT Survey is 47.5 and 60.0 minutes, respectively.

the density variables have positive effects on the average commuting times of the PT Survey. The R^2 value of 0.45 suggests fairly good fit; 45% of variation in PT Survey's average commuting times is explained by the fitted line (measured by the regression sum of squares). The regression results are then used to estimate OD commuting times ($\hat{TIME}^{AUTO}_{ij}$) for all inter-zone travels. Table 4 presents the basic statistics of the estimated OD commuting times in comparison with the actual OD average commuting times reported in the PT Survey.

Table 3. Regression results of OD commuting times

Variable	Auto		Public transit	
	Coefficient	t statistics	Coefficient	t statistics
Constant	13.254***	39.89	28.430***	113.12
GIS-modeled OD travel time ($TIME^{GISAUTO}_{ij}$)	0.877***	233.10	0.686***	300.50
Density of workers in origin zone ($WDENS_i$)	0.001***	31.91	0.0003***	10.68
Density of jobs in destination zone ($JDENS_j$)	0.0002***	21.35	0.00003***	6.42
Number of observations	66,911		100,968	
R^2	0.45		0.49	

Note: * Significant at the 0.10 level. ** Significant at the 0.05 level. *** Significant at the 0.01 level.

Table 4. Descriptive statistics of adjusted OD commuting times by car

Variables	Number of observations	Mean	Std. Dev.	Min.	Max.
Estimated OD commuting time ($\hat{TIME}^{AUTO}_{ij}$)	353,430	107.6	43.5	17.2	378.1
	66,911*	67.5*	27.9*	17.2*	327.0*
PT Survey OD commuting time ($TIME^{PTAUTO}_{ij}$)	66,911*	67.5*	41.3*	2.0*	960.0*

Note: *Statistics are for inter-zone travels where the PT data are available.

3.2 Modeling OD Commuting Times by Public Transit

Next, we estimate OD commuting times by public transit based on the following assumptions: first preference for means of commuting is railway; local trains that stop at every station are used; Shinkansen (rapid rail), cable cars, and ropeways are not used; and there is no waiting or transfer time. Each OD commuting time comprises three components: (a) travel time from origin zone's centroid to its nearest station on road ($TIME_{(1)}$); (b) travel time from the station nearest the origin zone's centroid to a station closest to the centroid of a destination zone on railroads ($TIME_{(2)}$); and (c) travel time from the station closest to the destination zone's centroid to the destination zone's centroid on roads ($TIME_{(3)}$). Figure 4 presents the flow diagram for the estimation of OD commuting times by public transit.

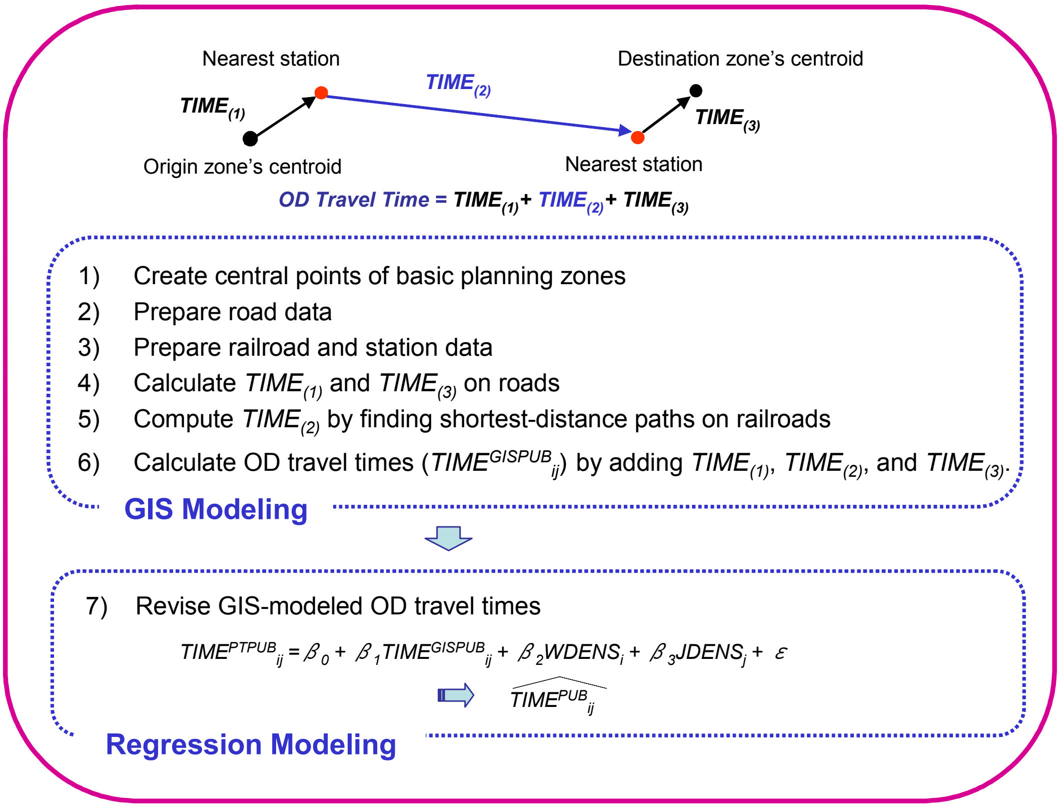

Figure 4. Estimation of OD commuting times by public transit

<u>GIS Modeling Part</u>

The first and second steps, the creation of central points of basic planning zones and the preparation of road data, respectively, are the same as the ones described in the previous subsection. The third step is to prepare railroad and station data as follows. (1) First, railroad links and station points are extracted from GISMAP 25000V (Figure 5). (2) Second, errors in links (such as dangled nodes) and redundant station points (such as multiple existences of the same stations) are corrected. (3) Third, the length of each link is calculated. (4) Fourth, average train speed is assigned to each railroad link based on train types as shown in Table 5. The speed is calculated for Tokyo-bound trains arriving at terminal stations between 8:30am and 9:00am using data in the Myline Tokyo Timetable (Mae, 2004). (5) Fifth, travel time is calculated for each railroad link by dividing length by speed. (6) Sixth, station points are snapped onto the railroad links.

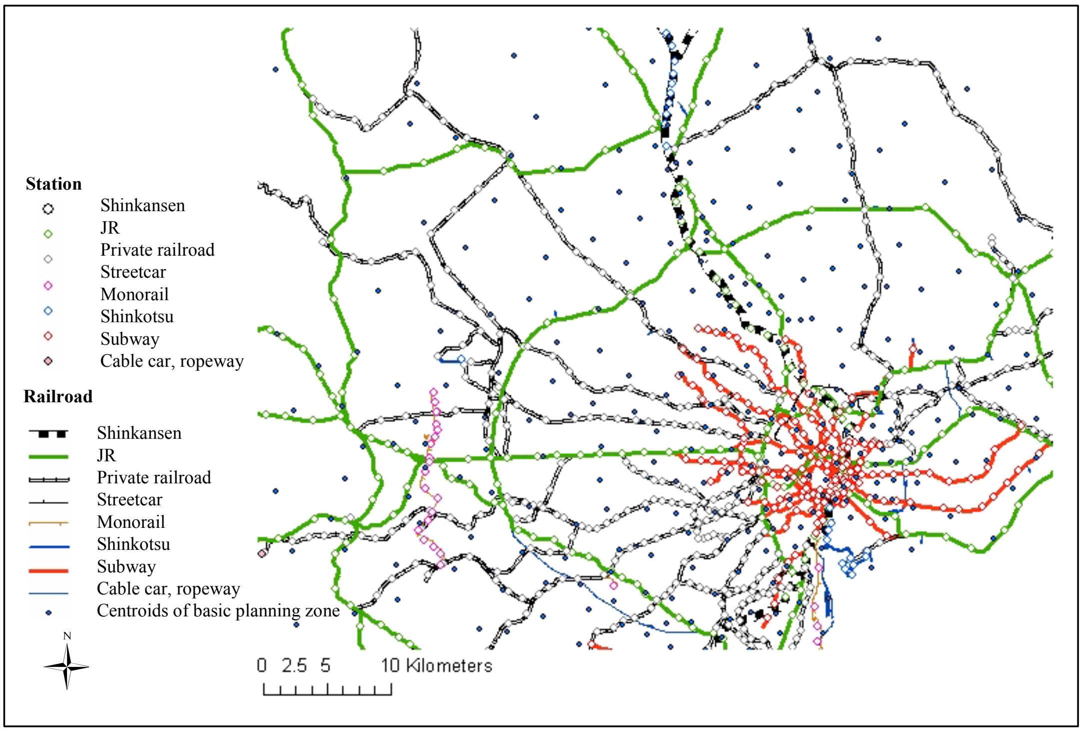

Figure 5. Railroads, stations, and centroids of basic planning zones

Table 5. Average train speed by train type

Train type code*1)	Train type*1)	Average train speed*2)
2	JR	45 km/h
3	Private railroad	35 km/h
4	Streetcar	15 km/h
5	Monorail	30 km/h
6	Shinkotsu	25 km/h
7	Subway	30 km/h

Source: *1) 2001 GISMAP25000V. *2) Myline Tokyo Timetable (Mae, 2004).
Note: Train types 1 (shinkansen) and 8 (cable car, ropeway, freight train, and sidetrack) are excluded.

The fourth step of the GIS modeling part is to calculate travel times between zones' central points and their nearest stations on roads ($TIME_{(1)}$ and $TIME_{(3)}$). The process for this step has four phases. (1) First, a station nearest each zone's central point is identified. (2) Second, the shortest-distance path between zones' central point and its nearest station on roads is found. (3) Third, total distance for each path is calculated. (4) Fourth, travel time for each path is calculated based on the assumption that people walk if walking time is equal to or less than 15 minutes; otherwise, people take buses. Walking speed is set at 80 meters per minute, the official speed determined by the Real Estate Fair Council in Japan. People therefore are assumed to walk if the distance is 1,200 meters

(=80 meters per minute × 15 minutes) or less. If the distance exceeds 1,200 meters, people take buses. The speed of bus is set at 20 kilometers an hour, about the average speed of long and short-distance buses, which is calculated with data in the Association of Tokyo Bus (2003).

The fifth step of the GIS modeling part is to compute travel time from the station nearest the origin zone's central point to a station closest to destination zones' central point for each OD pair ($TIME_{(2)}$). Using the railroad and station data prepared in the third step, the minimum travel-time path for each inter-zone OD pair is identified and travel time for each path is calculated. Finally, in the sixth step, total travel time for each OD pair is calculated by adding up $TIME_{(1)}$,, $TIME_{(2)}$, and $TIME_{(3)}$.

Regression Modeling Part

Table 6 shows the descriptive statistics of the OD travel times from the GIS modeling part ($TIME^{GISPUB}_{ij}$) and those from the PT Survey ($TIME^{PTPUB}_{ij}$). For the GIS-modeled OD travel times, statistics for both all inter-zone commuting (353,430 OD pairs) and inter-zone commuting where the PT Survey data are available (100,968 OD pairs) are presented.

Table 6. Descriptive statistics of OD commuting times by public transit

Variables	Number of observations	Mean	Std. Dev.	Min.	Max.
GIS-modeled OD travel time ($TIME^{GISPUB}_{ij}$)	353,430	119.9	55.6	7.6	405.0
	100,968*	75.0*	32.6*	7.6*	314.6*
PT Survey OD commuting time ($TIME^{PTPUB}_{ij}$)	100,968*	81.4*	31.5*	5.0*	870.0*

Note: *Statistics are for inter-zone OD travels where the PT data are available.

The mean commuting time for inter-zone travels where the PT Survey data are available from the GIS modeling (75.0 minutes) is shorter than that from the PT data (81.4 minutes). As in the case of car times, this result is likely related to the fact that the GIS-modeled travel times do not take into account waiting as well as transfer times while PT Survey's commuting times do.[2]

Using the same approach in the previous subsection, we estimate the following regression:

$$TIME^{PTPUB}_{ij} = \beta_0 + \beta_1 TIME^{GISPUB}_{ij} + \beta_2 WDENS_i + \beta_3 JDENS_j + \varepsilon, \qquad (2)$$

where $TIME^{PTPUB}_{ij}$ is the average commuting time from origin zone i to destination zone j of the PT Survey, and $TIME^{GISPUB}_{ij}$ is the GIS-modeled commuting time from origin zone

[2] The median travel time for inter-zone travels where the PT Survey data are available from the GIS modeling is 71.3 minutes, and the median travel time from the PT Survey is 78.1 minutes.

i to destination zone *j*. The regression results are shown in Table 3.

The GIS-modeled OD travel time ($TIME^{GISPUB}_{ij}$) has a strong and significant effect on the average OD times of the PT Survey ($TIME^{PTPUB}_{ij}$). The density variables also have positive and significant effects, but the magnitude of the effects is much lesser than that for auto users. This result is not surprising given that congestion is less likely to affect train times than to impact car times. The estimated model shows moderately good fit as indicated by the R^2 value of 0.49. The results are then used to estimate OD commuting times ($\hat{TIME}^{PUB}_{ij}$) for all inter-zone travels. The basic statistics of the estimated OD commuting times in comparison with the actual OD average commuting times reported in the PT Survey are shown in Table 7.

Table 7. Descriptive statistics of adjusted OD commuting times by public transit

Variables	Number of observations	Mean	Std. Dev.	Min.	Max.
Estimated OD commuting time	353,430	111.7	37.8	34.2	306.2
($\hat{TIME}^{PUB}_{ij}$)	100,968*	81.4*	22.1*	34.2*	244.2*
PT Survey OD commuting time ($TIME^{PTPUB}_{ij}$)	100,968*	81.4*	31.5*	5.0*	870.0*

Note: *Statistics are for inter-zone travels where the PT data are available.

3. Conclusions

In this paper we have proposed a GIS modeling method to systematically estimate OD commuting times by car and public transit in the Tokyo metropolitan region. The work is our initial attempt to utilize GIS to develop a modeling method that can readily be applied to other major urban areas. The proposed method is useful especially when estimating travel times for a large number of OD pairs in a given metropolitan area.

The proposed method, however, has some limitations and further refinements are under way. One limitation is that the GIS modeling part of the estimation process of OD travel times does not take into account waiting and transfer times. In the case of public transit, rapid and express trains are also not considered. Realistic public transit travel times can be obtained from such software as Eki Spurt by Val Laboratory Co. Kotoh (1995), for instance, uses Eki Spurt to get travel times between major city stations. This manual approach, however, is practically infeasible for a large number of OD pairs like this study's 353,430 OD pairs. In the case of cars, alternative data that have more detailed information on roads can improve the travel time estimation. Road data based on the Japan Digital Road Map Database developed by the Japan Digital Road Map Association, for example, provide in-depth information on road links including travel speed and road traffic regulations, but such data are considerably expensive. We are currently researching the practical ways to improve the modeling method that provide better estimation results.

Another limitation is that the gap between the estimated and actual travel times between centroids and their nearest stations ($TIME_{(1)}$ and $TIME_{(3)}$) for large zones may be significant. This problem is likely to happen in large zones around the metropolitan peripheries where public transit is not well developed. A way to deal with this problem is to use additional data that allow us to estimate the locations of workers and jobs within a zone. Residential Maps by Zenrin Co., Ltd., for example, include residential location information, and Townpage data by the NTT directory Services Co. have information on job locations. With these data, it is possible to investigate the spatial distributions of workers and jobs within a zone, which can then be used to calculate travel times between these locations and nearest stations.

Although these limitations exist, the OD commuting times estimated in this study would be useful for various purposes, especially for macro-scale analyses. Zone-based accessibility to jobs in a metropolitan area, for instance, can be calculated using the estimated OD commuting times as impedance. The refinement and utilization of the estimation of OD commuting times are of interest in our future research.

Acknowledgement

This study was supported by a research grant from the Center for Spatial Information Science at the University of Tokyo. We are grateful to Atsu Okabe, Akinori Morimoto, Tetsuo Morita, and Daisuke Shibatani for their data assistance.

References

Association of Tokyo Bus. (2003) *Route Guide to Tokyo Buses, No. 12*, Tokyo: Association of Tokyo Bus.

Hino, S., Haraguchi, M. and Satou, K. (1999) A calculation of user’s average travel time by GIS and an evaluation of public transportation routes: a case study of Sapporo northern area, Hokkaido, *Journal of the City Planning Institute of Japan*, **34**, 973-978.

Koshizuka, T., Ohara, H. and Nakagawa, M. (2003) A calculation of movable space expansion by railway in Kantoh region, *Journal of the City Planning Institute of Japan*, **38**(3), 151-156.

Kotoh, H. (1995) Visualization of multi-city structure using time-distance network, *Journal of the City Planning Institute of Japan*, 30, 553-558.

Mae, K. ed. (2004) *Myline Tokyo Timetable*, Tokyo: Transportation Co., Ltd.

Suzuki, T. (1992) An empirical study on the optimal commuting assignment problem in the Tokyo metropolitan area, *Journal of the City Planning Institute of Japan*, **27**, 337-342.

Wang, F. (2003) Job proximity and accessibility for workers of various wage groups. *Urban Geography*, **24**(3), 253-271.

Yoshinaga, T. (2003) *Study on the Transition of Time Range in the Tokyo Metropolitan Area*. Graduation Thesis, College of Policy and Planning Sciences, University of Tsukuba.