Characteristics of ETC2.0 probe data compared to road traffic census data

Shohei Yasuda[1], Takamasa Iryo[1], Yuta Koyama[1], Katsuya Sakai[1]

[1]Department of Civil Engineering, Graduate School of Engineering, Kobe University, Japan

1. Introduction

Probe vehicle data have been examined in many studies. One of the reasons for this is that this type of data contains considerable information, including detailed spatial-temporal vehicle trajectories, speed, and acceleration. To obtain comprehensive traffic information throughout Japan, the national government had to spend a considerable amount of resources to conduct road traffic censuses. Recently, the probe-vehicle system called electronic toll collection (ETC) 2.0 has covered the entire country, and data collection costs are reduced considerably compared to the traditional survey methods such as questionnaire.

ETC2.0 has multiple functions such as toll collections on expressways, information provisions to drivers, and collections of trajectory data (with privacy protection measures provided). Various types of data that are relevant to these services are transmitted via dedicated short range communications (DSRC). Antennas of DSRC are installed throughout Japan (approx. 1,600 units), mostly along with expressways and major roads. No mobile network is used. Data stored by on-board equipment are transmitted to the servers when vehicles pass by antennas only.

Since the system is rather new (data has been collected since 2011), studies utilising the ETC2.0 probe data are not many and consequently characteristics of the date have yet been fully known. The present study especially focuses on the comparisons between probe data and road traffic census data to understand how far the probe data are different from the traditional survey data and how we can correct the differences.

2. Data and Methodology

The nationwide probe data of ETC2.0 taken in October-December 2015 and data of the road traffic census conducted in October-December 2010 are used. Because the census data in 2015 has yet been available, we had to use the data of the census in 2010 instead of the latest one. Thus, we need to be aware that this study compares these different sets of data from the same seasons but from different years. In addition, this study employs supplement datasets such as the number of vehicles equipped by ETC2.0 on-board units and as well as the taxable incomes in each region.

Trajectories taken by ETC2.0 are in the form of temporal-spatial coordination, i.e. longitude, latitude, and time. We define the term of *dot* as a combination of them. Each trajectory of a vehicle consists of the series of the dots. Dots within 500 meters after vehicle ignition and 500 meters before the engine is turned off are erased for privacy protection. The dot data are first stored in ETC2.0 on-board units. Then, when the vehicle pass by the DSRC antennas, all stored data are transmitted to the server This implies that, if a vehicle does not pass by any antenna for a certain distance, data of old dots are overwritten by new dots and lost owing to a limitation of the storage in an on-board equipment.

We divide each trajectory of a vehicle in a single day into trips according to the criteria described below. A trajectory is divided into two trips when the vehicle speed between two dots is less than 20 km/h, and the time gap between them is 30 minutes or more for standard vehicles (15 minutes or more for large vehicles). Then, the divided trajectories are matched onto the digital road map (DRM) by the algorithm of Asakura et al.[2)] with minor modifications.

The cumulative number of ETC2.0 setups is acquired from an ETC2.0 general information portal site[2)]. The setup rate of ETC2.0 is then calculated by dividing the cumulative setup number by the number of car registrations in each region. This is regarded as ETC2.0 penetration of each region. Data of taxable income in 2010 are acquired from an official statistical data provided by the cabinet office in the government of Japan[3)].

3. Results

3.1 Total number of trips and sample rate of ETC2.0

Table 1 shows the total number of vehicle trips taken on weekdays based on ETC2.0 probe and road traffic census data, as well as sample rates of ETC2.0. The sample rates were approximately 0.1% for both standard and large vehicles.

Table 1: Total number of trip on week day and sample-rate

	Standard-size vehicle (week day)	Large-size vehicle (week day)
Census	142,334,684 trip	12,746,035 trip
ETC2.0	189,851.57 trip	13,710.43 trip
Sample-rate	0.133%	0.108%

3.2 Number of trip generations in each region

Figs. 1 and 2 show the number of trip generations of standard and large vehicles. The value of ETC2.0 was divided by the average sample rate of ETC2.0. From Fig. 1, in the Kanto, Chubu, and Kinki regions including major metropolitan areas such as Tokyo, Nagoya, and Osaka, the estimated numbers of trip generations of standard vehicles acquired by ETC2.0 were higher than that acquired by census data. This was probably because the penetration ratios of ETC2.0 on-board equipment in these major cities were higher than that in other cities. However, Fig. 2 shows that trip generation of large vehicles acquired by ETC2.0 in these areas was lower than that acquired by census. On the other hand, those acquired by ETC2.0 in the Tohoku, Hokuriku, Chugoku, and Shikoku regions were higher than those acquired by census. This was probably due to the different years of the data (i.e. 2010 for the census and 2015 for ETC2.0) and because trip generations of large vehicles are particularly sensitive to the economic circumstances. These results may have also been influenced by the Great East Japan Earthquake of 2011, but this cannot be confirmed by the data sets used in this study.

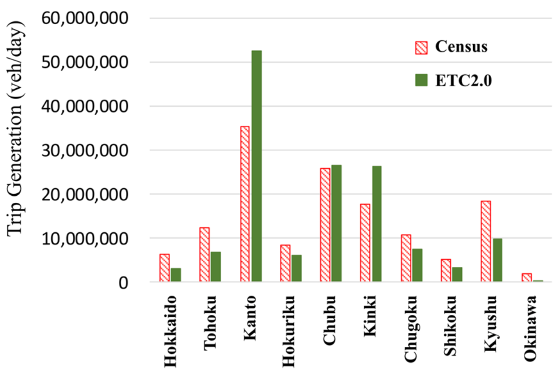

Fig. 2: Trip generation of each region (Standard vehicles)

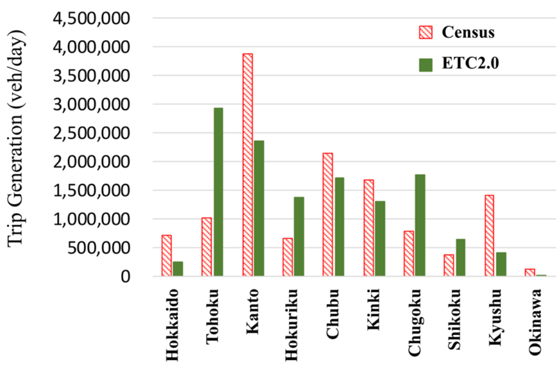

Fig. 2: Trip generation of each region (Large vehicles)

3.3 Relationship between ETC2.0 setup ratio and sample ratio of trip generation in each region

Fig. 3 shows the relationship between the ETC2.0 setup ratios and the sample ratios of trip generations. Each dot represents the relationship between the two rates according to prefecture. The intercept of the regression is fixed to zero so that if the sample rate is zero, the trip generation is also zero. The slope of the regression is estimated as 0.07. This result suggests that even if all vehicles are equipped by ETC2.0, the sample rate of trip generation will be only approximately 7%, and the remaining will be lost.

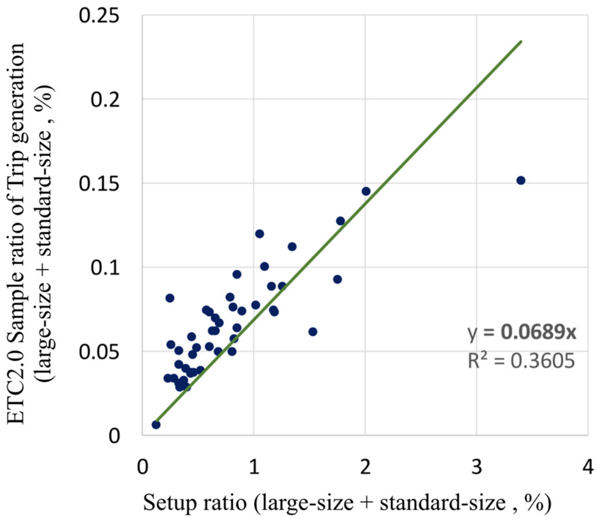

Fig. 3: The relationship between setup ratio and ETC2.0 sample ratio

3.4 Other analysis

One analysis we conducted is that of the relationship between the ETC2.0 sample ratio and income (Fig. 4). The results suggest that a positive correlation exists between the two. We also analysed the sample ratio of the traffic volume of vehicles departing from a certain prefecture to any city. Fig. 5 shows the result for vehicles departing from Tokyo. The results show that ETC2.0 contained more data about longer trips than do census data. This is probably because the antennas are mostly on highways.

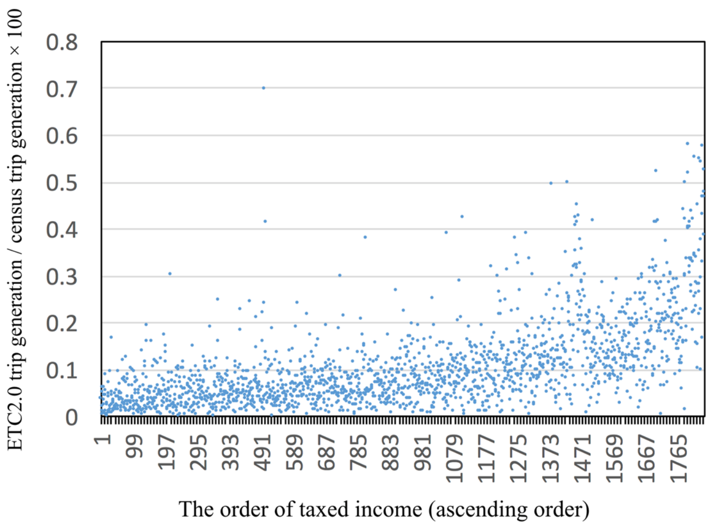

Fig. 4: Relationship between sample ratio (against the census data) of generation trips obtained by ETC2.0 and the order of taxed income for all cities / towns / villages in Japan

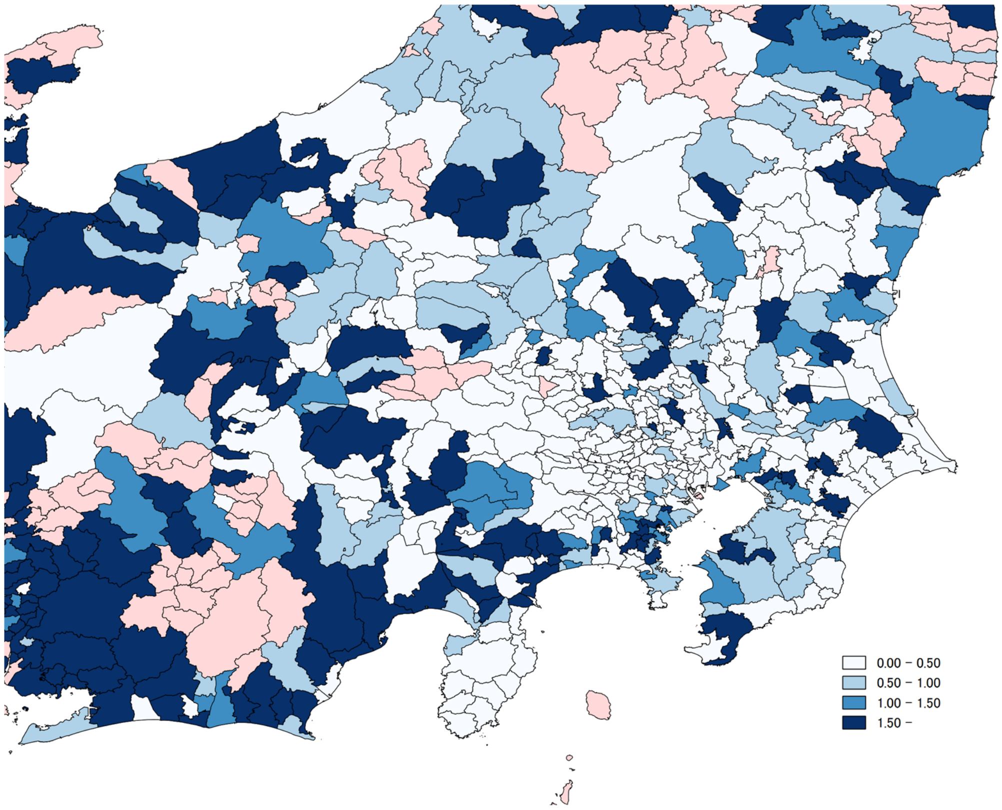

Fig. 5: Sample ratio (against the census data) of the number of trips obtained by ETC2.0 from Tokyo metropolitan area to cities / towns / villages. Numbers in the legend indicates the sample ratio in %. Pink colour implies that there is no sample in either ETC2.0 or the census data.

4. Conclusion

This study analysed the characteristics of ETC2.0 by comparing them with road traffic census data. The results show that there are biases in the sample ratio of ETC2.0. For example, the long-distance trips and those using expressways tend to be acquired more by the ETC2.0 system than the short-distance trips remaining on arterial roads. While this bias can be mitigated by installing many antennas on arterials, while it would not be a realistic solution because installing them are costly. The bias of the penetration ratio also causes the biases in the sample ratio in different regions. Methodologies that correct such biases are demanded to utilise ETC2.0 data for the replacement of the traditional road census data.

Acknowledgements

This study was conducted as a part of the technical researches and developments for quality improvement of roads, which was financially supported by Ministry of Land, Infrastructure, Transport and Tourism in Japan.

References

1) General incorporation foundation ITS technology enhancement association, ETC portal site, http://www.go-etc.jp/fukyu/index.html.
2) Asakura, Y. et al, Monitoring travel behaviour using PHS based location data, Journal of JSCE, 653/IV-48, pp. 95-104, 2000.
3) Government of Japan cabinet office, economic and population related data, http://www5.cao.go.jp/keizai-shimon/kaigi/special/future/keizai-jinkou_data.html.