Statistical Journal of the IAOS 29 (2013) 223–232
DOI 10.3233/SJI-130778
IOS Press

Can mobile phone network data be used to estimate small area population? A comparison from Japan

Naoki Makita[a,*], Masakazu Kimura[b], Masayuki Terada[c], Motonari Kobayashi[c] and Yuki Oyabu[c]
[a]*National Statistics Center, Shinjuku-ku, Tokyo, Japan*
[b]*Department of Economics, Aoyama Gakuin University, Shibuya-ku, Tokyo, Japan*
[c]*NTT DOCOMO R&D Center, Yokosuka City, Kanagawa, Japan*

Abstract. For a long time, only national statistics offices have had the capability of producing small area statistics on population, such as, in Japan, the Grid Square Statistics of Population Census (GPC). With the proliferation of mobile phones in daily life and the significant development of ICT, there emerges a new kind of statistics called the Mobile Spatial Statistics (MSS), which estimates small area population derived from the operational data of a mobile phone network. There are several differences between the two statistics, particularly relating to measuring de jure versus de facto population, and research was conducted to examine the validity of MSS by comparing the two statistics. Deviation rates were calculated to compare the 2010 GPC population with the MSS population before dawn on the 2010 Census day and concluded that, to a certain degree, MSS would be plausible for use in densely populated areas. We also found that MSS highlighted the busy blocks of central Tokyo early in the morning. The research proved the utility of MSS, and at the same time, it confirmed the ongoing usefulness of GPC as the benchmark of population statistics.

Keywords: Small area statistics, grid square statistics, mobile spatial statistics, mobile phone network, gridded population, geostatistics

1. Introduction

Small area statistics on population is essential information and widely used in various fields: from city planning, to regional developments, disaster management, and area marketing. Today, with the development of information and communication technology, the utilization of and demands for small area statistics have been booming.

For a long time, only national statistics offices have been capable of producing such statistics. In Japan, the Grid Square Statistics of Population Census (GPC) is disseminated by the Statistics Bureau of Japan (SBJ). Within this context, a new kind of statistics called the Mobile Spatial Statistics (MSS) is being developed by NTT DOCOMO, a major mobile phone operator in Japan. MSS provides hourly estimates of small area population, derived from the operational data of the mobile phone network.

GPC and MSS are set apart by several key differences, for instance, the characteristics of population. GPC depicts *de jure population*, that is, an individual's location is recorded to their place of residence on Census day. On the other hand, MSS estimates *de facto population*, i.e. the location is recorded according to where the person is present at the time of reference. For example, suppose there is a man who lives in area X and goes to his office in area Y. GPC attributes him to X only, but MSS attributes him to X when he is at home, and to Y when at work. He may go shopping

---

[*] Corresponding author: Naoki Makita, National Statistics Center, 19-1 Wakamatsu-cho, Shinjuku-ku, Tokyo, Japan. Tel.: +81 3 5273 1282; Fax: +81 3 3203 8189; E-mail: naoki@nstac.go.jp.

somewhere neither X nor Y, and would be attributed to that area at that time.

Although in this way the GPC and MSS populations will not always be aligned, it seems reasonable to expect that the two population measurements will be somewhat close at midnight or before dawn on the Census day, as most people are at home during those times.

Another key difference between GPC and MSS is their respective data sources. While GPC is aggregated from the Population Census, which enumerates all of the Japanese population, MSS is based on estimates from DOCOMO's mobile phone users, which covers over 40% of the Japanese population. GPC can provide the population for any given small area, as the whole of Japan is exclusively and exhaustively partitioned by Census Enumeration Districts (CEDs). On the other hand, there are limitations on applying MSS to a small area with sparse population because the sample of the mobile phone users in those areas would be too small to accurately inform estimation. In addition, although the coverage area of base stations (cell towers) encompasses all the 1700+ municipality offices in Japan, it does not cover the whole of Japan, where more than half of the land area is forest and mountainous zone. MSS could be, nonetheless, viable for a small area with large enough population.

Despite these limitations, MSS remains an attractive alternative to GPC, as it incorporates several features that GPC doesn't have. First, MSS has advantages over GPC in terms of periodicity and timeliness. Population Census of Japan is conducted every five years, and it takes around two years to produce the GPC results. In contrast, the periodicity of MSS is as short as one hour thanks to the mobile phone network which operates around the clock. MSS requires only a few weeks to finalize the result to ensure consistency and confidentiality.

Second, MSS would provide cost savings through efficiency of data collection, compared to GPC; whereas a statistical agency bears huge costs to mobilize enumerators to collect census questionnaires and to capture the data, MSS utilises operational data already routinely collected by DOCOMO from the mobile phone network.

To enhance access to statistical information and to meet growing needs of small area statistics, it should be worthwhile to explore possibilities and limitations of MSS. For these reasons, the National Statistics Center (NSTAC), an incorporated administrative agency attached to SBJ, and mobile telecommunications provider DOCOMO conducted a joint research project to examine the validity of MSS by comparing the two statistical methods in more detail. This paper outlines the methodology used to produce GPC and MSS, and provides a comparison between them, based on the calculation of 'deviation rates'. This research concludes that MSS would be plausible for densely populated areas.

1.1. Related studies

Several studies have been published previously that, like MSS, aim to discover population dynamics by using operational data from mobile phone networks.

Recently, some researchers have inferred car traffic flows [1], origin-destination flows [2,3] and road usages [4,5] based on the location data extracted from mobile phone networks. However, such studies have mainly focused on monitoring people's flow or traffic conditions, rather than the geographic distribution of population.

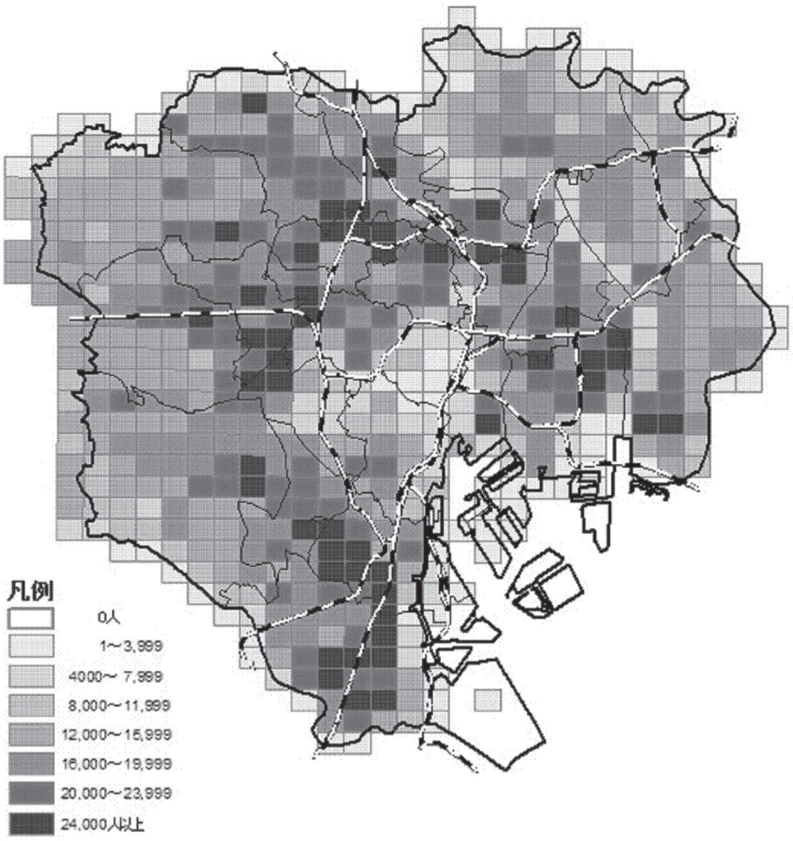
City API [6] developed by the urban API project [7] demonstrated that the geographic distribution of mobile phone users could be monitored by using mobile phone network data, but this method only considered the distribution of mobile phone user, not that of the population as a whole. The relationship between census data and their estimation was not evaluated.

Becker et al. [8] estimated laborshed numbers (catchment populations who commute to a city of interest) in the United States from mobile phone network data, and drew a scatter plot which showed a correspondence between the resulting numbers and Census data. However, this study assumed that the estimated numbers are simply proportional to population, i.e. the demographic variation of the usage ratio of mobile phone is not considered. Their study focused on the neighboring area of a single city in the U.S., but the nation-wide applicability of their research is unclear.

2. Grid Square Statistics of Population Census

2.1. Background on Japanese Grid Square Statistics

Grid Square Statistics (GSS) present statistical data by small areas of almost the same size and shape, demarcated by latitude and longitude at fixed intervals (grid squares). It is easy to measure and compare data for areas across regions and in time-series without being affected by the change of administrative boundaries or city blocks due to community redevelopment, etc.



Source: URL: http://www.stat.go.jp/data/mesh/h22_w.htm (Japanese language)

Fig. 1. Grid Square Statistics of 2010 Population Census, Tokyo area by Basic Grid Square, Statistics Bureau of Japan.

In Japan in the middle of the twentieth century, there arose the statistical need to collect and compare a variety of data by a common small area, for planning civil engineering and public services. To respond to the needs, the then Administrative Management Agency released Announcement 143: Standard Grid Square and Grid Square Code Used for the Statistics in 1973 to direct the Japanese standard method in compiling the common small area statistics.

The Announcement stipulates Primary Area Partition, Secondary Area Partition, Basic Grid Square (Third Area Partition) and Half Grid Square. A Primary Area Partition, almost 6,400 km^2 (80 km by 80 km), is an area of Japanese territory demarcated by intervals of 40 minutes latitude and 1 degree longitude. Secondary Area Partition is defined by dividing a Primary Area Partition into 8 × 8 parts, each of which is almost 100 km^2 (10 km by 10 km). In the same way, dividing a Secondary Area Partition into 10 × 10 parts, Basic Grid Squares (almost 1 km^2) are given, and dividing one of these into 2 × 2 parts, gives Half Grid Squares (almost 500 m by 500 m).

Complying with the Announcement, statistical agencies have compiled several GSSs: SBJ produces GSS of Population Census (GPC) and GSS of Economic Census; the Ministry of Economy, Trade and Industry produces GSS of Census of Manufactures; and the National Institute for Agro-Environmental Sciences compiles GSS of Census of Agriculture and Forestry.

2.2. *Grid Square Statistics of Population Census*

The Statistical Bureau of Japan (SBJ) conducts five-yearly Population Censuses, and disseminates statistical results on population including small area statistics such as GPC (Fig. 1). The latest Population Census in Japan was taken as of 1st October 2010, and the 2010 GPC was released in November 2012. The 2010 GPC covers all of Japan by Basic Grid Square and by Half Grid Square. NSTAC, under the direction of SBJ, processes various kinds of data from censuses and surveys, and tabulates statistics. GPC is one of the products of NSTAC.

GPC is compiled by assigning the Census populations of CEDs to grid squares, and aggregating by grid square. Statistical disclosure control is applied by suppressing the data for a grid square where the population is so small that disclosing the number might violate the privacy of respondents.

It is not straightforward to compile GPC population from CED populations, even though a geographical information system streamlines the process. There are several ways of determining the correspondence between CEDs and grid squares: If a CED is compact enough to be fully contained in a single grid square (as is usually the case in highly populated areas), the CED population belongs to the population of the grid square. If a CED is too broad to fill in a single grid square but is overlaid with multiple grid squares, the population is divided in certain ways, for instance: in case of the CED in a region where the population is homogeneously dispersed, the CED population is shared by the proportion of the CED area that fell within each grid square; in case of the CED in a region where population is sparse, the CED population is distributed based on consideration of factors such as the geographical distribution of dwellings in the CED [9].

GPC can illustrate population distribution as well as population composition by various socio-demographic characteristics derived from the questionnaire of the Population Census, such as not only sex and age group, but also labor force status by industrial and occupational classification, education level, etc.

3. Mobile spatial statistics

3.1. Overview of mobile spatial statistics

Japanese mobile telecommunications company NTT DOCOMO has been developing MSS since 2008, as a method of estimating small area population from the operational data of their mobile phone network. DOCOMO has formed partnerships with some municipalities and universities in order to explore the possible use of MSS in fields such as urban planning (analysis of the catchment area of a city center [10]), disaster prevention planning (projection of stranded commuters [11]), and regional revitalization (estimation of the number of visitors [12]). MSS is currently still in research phase, and MSS data are not yet publically available.

MSS is able to illustrate population distribution, population transition for every an hour, as well as population composition by sex, age group and residential area of people (Fig. 2.1). The demographic disaggregation available in MSS is useful, if not so comprehensive as that of GPC. Figure 2.2 is the population distribution maps of Japan in the predawn (2.2a) and the afternoon (2.2b) on a weekday, demonstrating the difference in population dynamics between night and day.

Although the geographic unit of MSS population can be either grid squares or administrative areas, in this paper we focus on MSS population for grid squares to simplify the discussion.

3.2. Operational data of mobile phone network

MSS is compiled from the operational data of DOCOMO's nationwide mobile phone network. The number of DOCOMO users is approximately 60 million. With the mobile phone penetration rate reaching almost 100 percent in Japan (whose population is 128 million), more than 40% of the Japanese population carry DOCOMO mobile phones. Although it is not mandatory under Japanese regulation to obtain users' consent for aggregating personal information to produce statistics, in compiling MSS, DOCOMO accepts opt-out requests from customers who don't wish their mobile phone data to be used for that purpose.

The operational data consist of two categories: location data and attribute data.

3.2.1. Location data

Location data are a set of information detailing where DOCOMO's mobile phones are located. The data is obtained through base stations in the mobile phone network [13], not via global navigation satellite systems such as GPS. The location data is indispensable for the mobile phone network, as it enables any mobile phones to be paged anytime, anywhere.

Mobile phones connect to the mobile phone network through the base stations which facilitate wireless communications in their coverage areas, each of which is called a "cell". Mobile phones routinely contact nearby base stations, and the mobile phone network is able to monitor the location of each mobile phone with regard to cell.

The distribution of base stations, as well as the granularity of cells, depends on population density. The denser the population in an area is, the more base stations are installed there, the smaller each of the cells is, and thus the more accurately a phone in the area can be located.

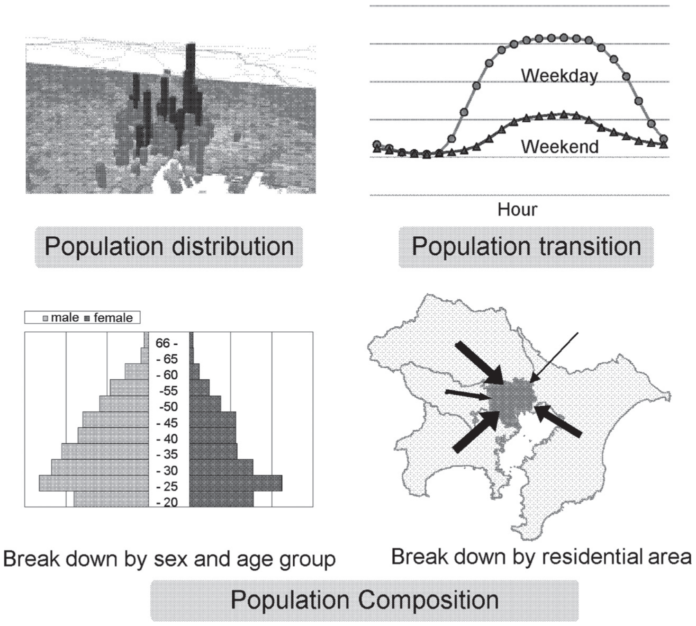
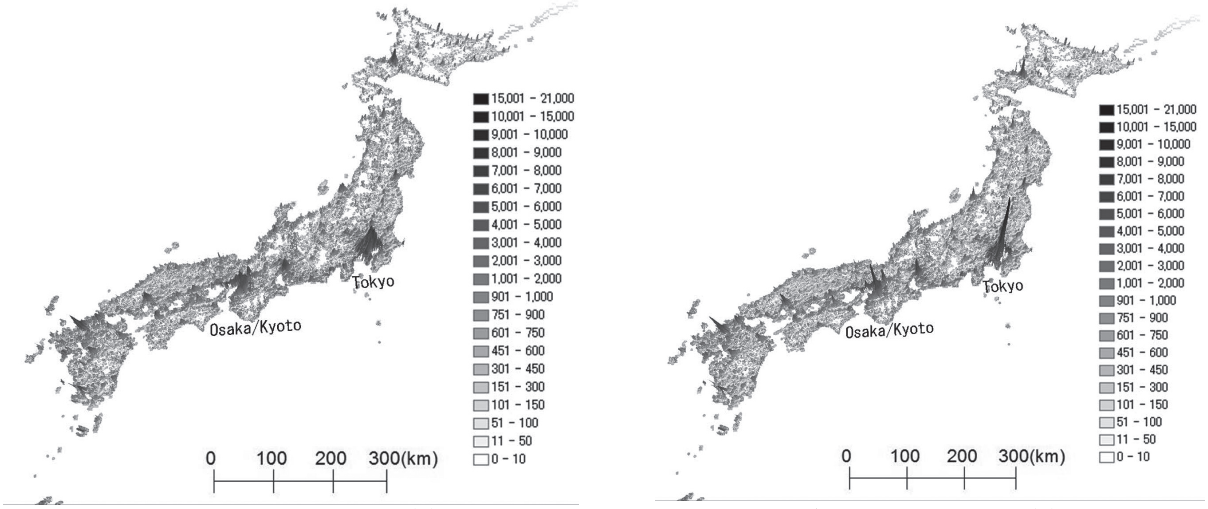


Fig. 2.1. Variation of Mobile Spatial Statistics presentation.



(a) Japan, 4 AM on a weekday

(b) Japan, 2 PM on a weekday

Fig. 2.2. Mobile Spatial Statistics, NTT DOCOMO.

3.2.2. *Attribute data*

Attribute data are a set of demographic information relating to mobile phone users. The mobile phone network holds the attribute data (such as user name and residential address, sex and date of birth) for service provision control and billing purposes.

Population estimates by MSS covers the population aged 15 to 79 only. Those aged 14 and under and 80 and over, are excluded from MSS, because mobile phone penetration rates for these age groups are low, and thus the sample size is too small for the MSS data to provide accurate estimates. Note that the population aged 15–79 makes up around 80 percent of the total population according to the 2010 Population Census.

3.3. *Compilation process of MSS*

Both the location data and the attribute data contain sensitive information about the mobile phone user, and ample care to protect their privacy must be taken when

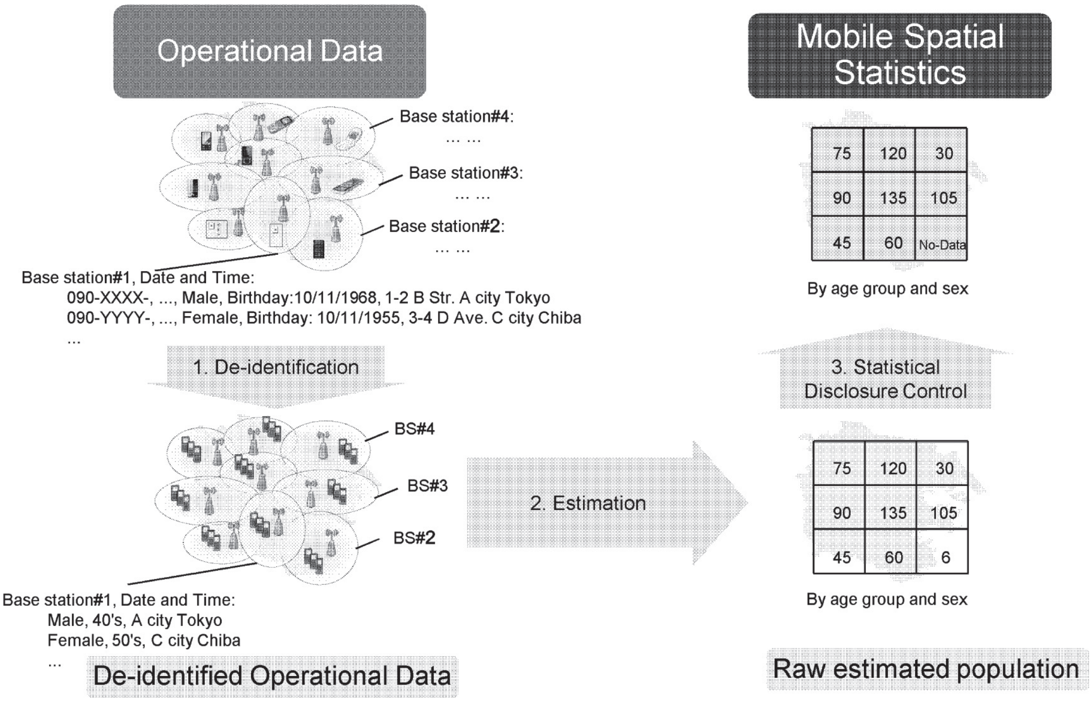


Fig. 3. Three-step process to compile Mobile Spatial Statistics.

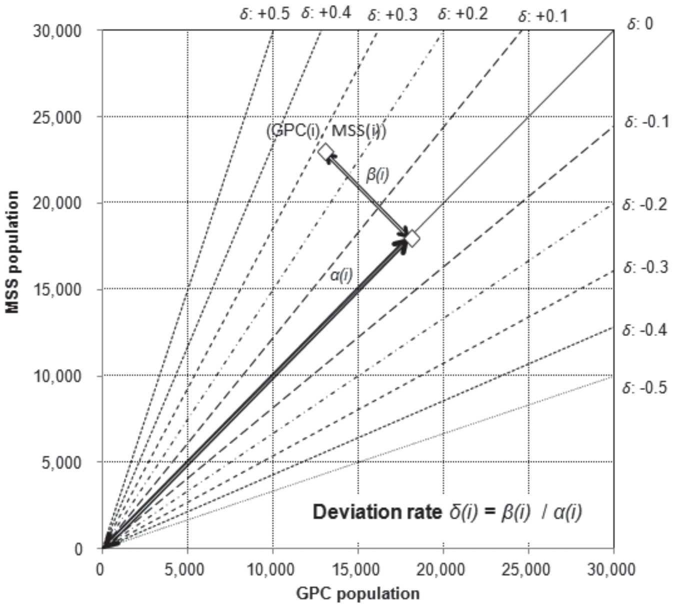


Fig. 4. Concept of the deviation rate.

handling the operational data. To that end, DOCOMO has established a three-step process (Fig. 3) for ensuring privacy in compiling MSS, incorporating the recommendations of an external expert study [14] covering social, legal and technical aspects of MSS.

3.3.1. De-identification

As the first step, unnecessary information for the population estimate is deleted from the operational data. In compiling MSS, it is sufficient to have sex, age group and city code of residence of mobile phone users, so before performing the estimation process, information such as dates of birth and residential addresses are coded in groups, and information that could be used for personal identification is removed, such as names, telephone numbers, and identifiers of mobile phone, so that no one can follow a mobile phone nor a particular user over time.

3.3.2. *Estimation*

In the second step, the raw population estimate for each grid square is compiled from the de-identified operational data. The estimation step includes several sub processes, such as to remove the bias due to the fluctuation of the interval time of routine contacts between mobile phones and base stations; to estimate the number of mobile phones on an hourly basis by sex, age group and the residential area of users for each cell; to extrapolate the whole population for each cell including those other than DOCOMO users, reflecting the DOCOMO's share in the mobile phone market by sex, age group and location (prefecture); and to convert the extrapolated population for cells to those for grid squares. The details of these sub processes are described in [15].

3.3.3. *Statistical disclosure control*

To ensure the privacy of DOCOMO users, statistical disclosure control is applied to the raw estimated population compiled in the above sub process, in the same way as has been described for GPC.

4. Evaluation of deviation rate between MSS and GPC

With regard to MSS data, which seems to have immense possibilities for statistics users, DOCOMO has formed partnerships with municipalities and universities in order to research the application of MSS. However, this research had done little to examine the validity of MSS data, and a joint research project was launched by DOCOMO and NSTAC to compare MSS with GPC.

For the comparison, we introduce a 'deviation rate' which measures the extent of the difference of the two statistics in a given grid square, and evaluated the accuracy of MSS by examining the deviation rates between MSS in early morning on the 2010 Census day and 2010 GPC.

The accuracy of MSS largely varies depending on the size of grid square and the population density. This is primarily, because these characteristics (size and population density) are reflected in the cell size (coverage area) of each base station, with smaller cells providing more accurate location data, and hence more accurate MSS populations. We show and discuss these relationships and examine the overall pattern of deviation rates in Section 4.2. In Section 4.3, we closely investigate the Tokyo area where the deviation rates are relatively high, which illustrates the inherent difference between the de facto population and the de jure population in definition.

4.1. *Definition of deviation rate*

Assume $MSS(i)$ is the MSS population in a grid square i, and $GPC(i)$ the GPC population between the ages 15–79 in i (since MSS population excludes those aged under 15 and over 79). $\mu(i)$ denotes the average of the two populations, i.e. $\mu(i) = \{MSS(i) + GPC(i)\}/2$.

The deviation rate $\delta(i)$ in the grid square i is defined as the difference between the MSS estimated population $MSS(i)$ and the average of the two estimation methods, $\mu(i)$, relative to $\mu(i)$: $\delta(i) = \{MSS(i) - \mu(i)\}/\mu(i)$.

By definition, $\delta(i)$ ranges between -1 and 1, which can be seen by substituting in the original $\mu(i) = \{MSS(i) + GPC(i)\}\ /2$ to give the following equation: $\delta(i) = \{MSS(i) - GPC(i)\}/\{MSS(i) + GPC(i)\}$.

The closer $\delta(i)$ is to zero, the closer the MSS and GPC population estimates in i are to each other. The nearer $\delta(i)$ is to 1 or -1, the further they are from one another. If the deviation rate $\delta(i)$ is positive, this signifies that the MSS population estimate exceeds the GPC estimate, and if negative, vice versa.

Figure 4 illustrates the concept of the deviation rate $\delta(i)$ on a Cartesian plane with axes X = GPC and Y = MSS. The 45 degree slope on the plane represents points at which the MSS population estimate agrees with GPC estimate. $\delta(i)$ is calculated as the ratio of $\beta(i)$ (the perpendicular distance from the point ($GPC(i)$, $MSS(i)$) to the 45 degree slope) to $\alpha(i)$ (the distance between (0,0) and projection of ($GPC(i), MSS(i)$) on to the 45 degree slope.

4.2. *Results of the deviation rate*

Deviation rates were calculated by inputting the pairs of MSS and GPC populations for the whole of Japan. Based on the assumption that MSS and GPC before dawn on the Census day were closely aligned, the MSS population at 4 AM on Friday 1st October 2010

Table 1
Proportion of the grid squares for which the deviation rates fall within a certain absolute distance from zero

Grid square type (approx. on a side)		Proportion (%) of grid squares for which the deviation rates fall within		Number of grid squares inputted for the calculation
		± 0.2	± 0.1	
Primary Area Partition	(80 km)	83.1	57.4	136
Secondary Area Partition	(10 km)	64.6	46.3	3,130
Basic Grid Square	(1 km)	56.9	33.5	69,531
Half Grid Square	(500 m)	50.1	27.8	137,926

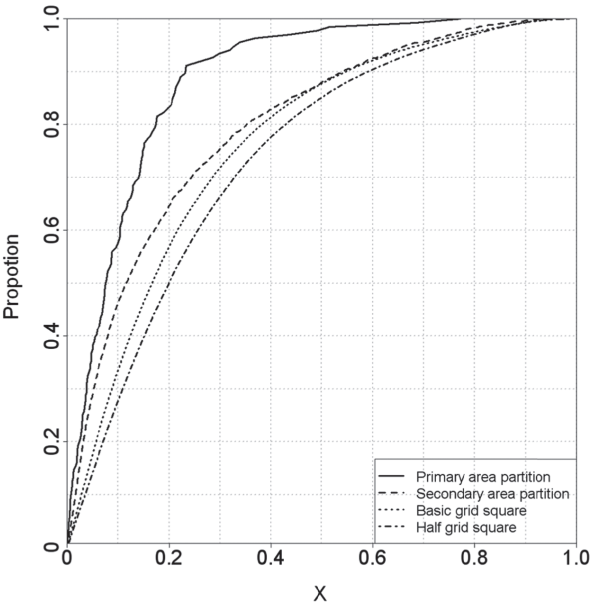


Fig. 5. Chart of the proportions of grid squares for which the absolute values of the deviation rates are less than a given value. Vertical axis: Proportion of grid squares for which the absolute values of the deviation rates is less than a given value X.

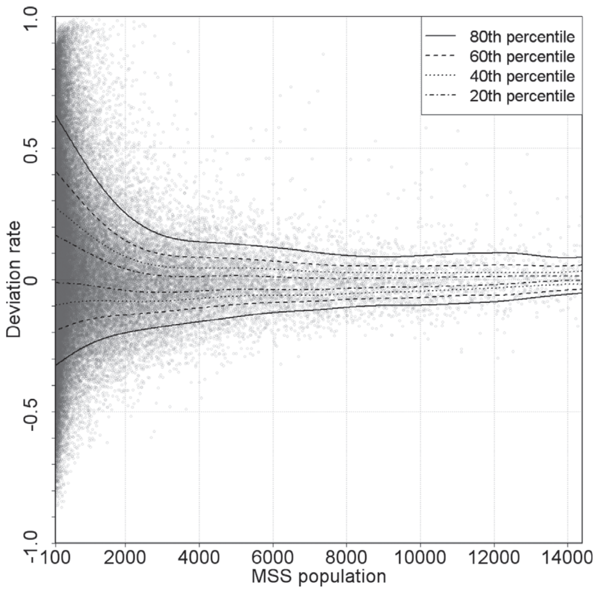


Fig. 6. Scatter plot of Basic Grid Squares' deviation rate along the MSS population.

(the Census day) is chosen for comparison against the 2010 GPC population so as to minimize the difference between de facto population and de jure population. In calculation, Basic and Half Grid Squares for which the MSS population estimate is less than 100 are removed from the input data to avoid unnecessary and extreme fluctuation of deviation rates.

Table 1 shows that the larger the grid square type is, the higher the proportion of individual grid squares for which the deviation rates fall within a certain absolute distance from zero, implying that for larger grid squares, *MSS(i)* and *GPC(i)* tend to be closer each other than for smaller grid squares. We found that 83.1% of Primary Area Partitions have deviation rates between a threshold of ± 0.2, followed by 64.6% of Secondary Area Partitions, 56.9% of Basic Grid Squares and 50.1% of Half Grid Squares. The same pattern is observed when considering deviation rates between ± 0.1. This pattern seems reasonable, because the larger the size of grid square is, the larger the sample of mobile phones in the grid square to be used for estimating the MSS population is, and the more stable the estimation is.

Figure 5 shows the relationship between the proportion of grid squares of a given grid square type (vertical axis) where the absolute values of the deviation rates are less than a given value X (horizontal axis). It is Primary Area Partition which approaches 100% the fastest, followed by that of Secondary Area Partition, Basic Grid Square and Half Grid Square. This indicates that larger grid sizes are likely to experience the least fluctuation in the deviation rate.

Figure 6 is a scatter plot of Basic Grid Squares with the deviation rate as vertical axis and the MSS population as horizontal axis, overlaid with two-sided percentile curves of the deviation rate distributions. These percentile curves indicate that the fluctuation in deviation rates tends to decrease as the MSS population in

Fig. 7. Choropleth map of the deviation rate by Basic Grid Square, Tokyo area.

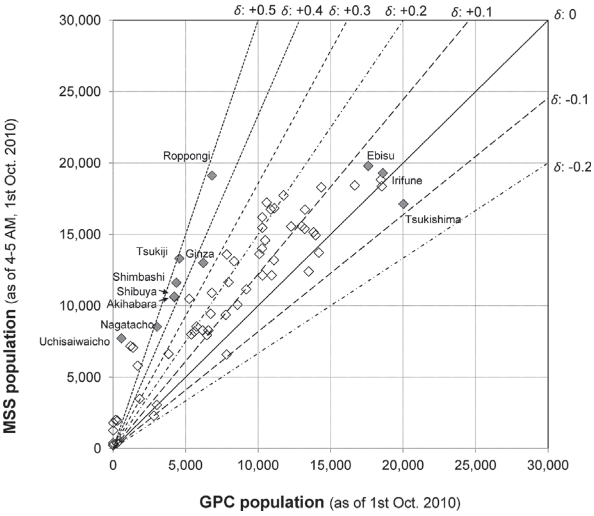


Fig. 8. Scatter plot of MSS and GPC populations by Basic Grid Square, central Tokyo.

the grid square increases, up to a population of 4,000. The higher the population density in a grid square is, the more stable the deviation rate becomes. This pattern may reflect the fact that more base stations for mobile phones are installed as the population density increases, leading to more accurate estimates. For MSS populations above 4,000, the percentile curves level off, implying that at this level of population, the maximum impact has been reached, and that further increases in population density do not continue to improve accuracy.

4.3. *Distribution of deviation rate in Tokyo area*

Figure 7 is a choropleth map of the deviation rate of Basic Grid Squares in Tokyo area, the most crowded area in Japan. For the most part the deviation rate is between ± 0.1, although there are some areas where the deviation rates exceed this range, mostly beyond +0.1 (darker gray).

There are two reasons that the deviation rates become unstable in some grid squares even in such a densely populated area. First, there are some city blocks where few people usually reside, and so have low GPC populations, but which MSS reflects as being busy with workers, hotel guests and nightlife. This makes the deviation rates noticeably higher even in the predawn. Figure 8 is a scatter plot of Basic Grid Squares in central Tokyo (Chiyoda, Chuo and Minato Cities), with GPC population as horizontal axis and MSS population as vertical axis. While the grid squares of residential neighborhoods, such as Ebisu, Irifune and Tsukishima, are close to the 45 degree slope (zero deviation), those of downtown districts are plotted in the zone of high deviation. These include Roppongi, Ginza and Shibuya (entertainment districts); Shimbashi, Akihabara and Uchisaiwaicho (business quarters); Nagatacho (government complexes); and Tsukiji (popular early morning fish market).

Another reason is that there are geographically unique districts which make it technically difficult to compile small area statistics, such as those next to city parks, riverside, bayside and other vacant spaces. It is often troublesome for both GPC and MSS to properly assign populations of CEDs/cells to grid squares in such districts.

5. Concluding remarks

With the proliferation of mobile phones in daily life and the significant development of ICT, MSS emerges as a means of harnessing these new technologies for statistical production and use. Acknowledging the differences between MSS and GPC, we compared them by inputting the MSS population before dawn on the 2010 Census day and the 2010 GSS population, to calculate the deviation rate. Though there are limitations in applying MSS to a small area with sparse population, we found that, to a certain degree, the use of MSS would be plausible in densely populated area. We also found that MSS was impacted by particular patterns of behavior, as seen in the busy blocks of central Tokyo early in the morning. The research proved the utility of

MSS, and at the same time, it confirmed that GPC remains useful as the benchmark of population statistics.

In today's society, enormous amounts of operational data, as well as administrative data, are generated day by day or even hour by hour not just by mobile phone networks, but in various areas of society. In the same way as MSS, new kinds of statistics could be created in a timely, low cost manner. Such statistics may enrich our social life with more information, though evaluating their credibility often proves difficult. In the age of big data, it would be worthwhile for national statistics offices, being committed to retaining trust in official statistics, to investigate possibilities and limitations of those new statistics and to shed light on the pros and cons of the statistics [16].

Acknowledgements

The authors would like to thank colleagues who provided us with many comments and suggestions, in particular Mr. Nobuyuki Sakashita and Mr. Masayoshi Takahashi of NSTAC.

The views and opinions expressed in this document are those of the authors, and not necessarily those of the organizations to which the authors belong. The authors take responsibility for any mistakes or errors.

References

[1] F. Calabrese, M. Colonna, P. Lovisolo, D. Parata and C. Ratti, Real-time urban monitoring using cell phones: A case study in Rome, *IEEE Transactions on Intelligent Transportation Systems* **12**(1) (2011), 141–151.

[2] N. Caceres, J.P. Wideberg and F.G. Benitez, Deriving origin-destination data from a mobile phone network, *IET Intelligent Transport Systems* **1**(1) (2007), 15–26, doi: 10.1049/iet-its.

[3] F. Calabrese, G. Di Lorenzo, L. Liu and C. Ratti, Estimating Origin-Destination Flows Using Mobile Phone Location Data, *IEEE Pervasive Computing* **10**(4) (2011), 36–44, doi: 10.1109/MPRV.2011.41.

[4] K. Sohn and D. Kim, Dynamic origin-destination flow estimation using cellular communication system, *IEEE Transactions on Vehicular Technology* **57**(5) (2008), 2703–2713.

[5] P. Wang, T. Hunter, A.M. Bayen, K. Schechtner and M.C. Gonzalez, Understanding road usage patterns in urban areas, Scientific reports, 2:1001, (2012) doi: 10.1038/srep01001.

[6] W. Loibl and J. Peters-Anders, Mobile phone data as source to discover spatial activity and motion patterns, in: *GI_Forum 2012: Geovizualisation*, T. Jekel, A. Car, J. Strobl and G. Griesbner, eds., society and learning, Herbert Wichmann Verlag Berlin, (2012), 524–533, ISBN: 978-3-87907-521-8.

[7] urbanAPI: Interactive Analysis, Simulation and Visualisation Tools for Urban Agile Policy Implementation, available from: http://www.urbanapi.eu.

[8] R.A. Becker, R. Cáceres, K. Hanson, J.M. Loh, S. Urbanek, A. Varshavsky and C. Volinsky, A tale of one city: using cellular network data for urban planning, *Pervasive Computing, IEEE* **10**(4) (2011), 18–26, doi: 10.1109/MPRV.2011.44.

[9] Chiiki messhu toukei no sakusei ('Compilation of grid square statistics' in Japanese language), Statistics Bureau, Ministry of Internal Affairs and Communications, Japan, available from: http://www.stat.go.jp/data/mesh/pdf/gaiyo2.pdf.

[10] T. Odawara and H. Kawakami, Using Mobile Spatial Statistics in Field of Urban Planning, *NTT DOCOMO Technical Journal* **14**(3) (2013), 31–36.

[11] T. Suzuki, M. Yamashita and M. Terada, Using Mobile Spatial Statistics in Field of Disaster Prevention Planning, *NTT DOCOMO Technical Journal* **14**(3) (2013), 37–45.

[12] T. Nagata, S. Aoyagi and H. Kawakami, Using Mobile Spatial Statistics for Regional Revitalization, *NTT DOCOMO Technical Journal* **14**(3) (2013), 46–50.

[13] Functional stage 2 description of location services (LCS), 3rd Generation Partnership Project (3GPP), Technical Specification 23.271.

[14] Shakai sangyou no hatten ni kiyosuru mobairu kuukan toukei rikatsuyou no arikata ni kansuru houkokusho ('A recommendation to the utilization of Mobile Spatial Statistics for the contribution to social and industrial development' in Japanese language), Ad hoc panel on Mobile Spatial Statistics for the contribution to social and industrial development, Mobile Society Research Institute (2010), available from http://www.moba-ken.jp/research/research2010/r10-01.html.

[15] M. Terada, T. Nagata and M. Kobayashi, Population estimation technology for mobile spatial statistics, *NTT DOCOMO Technical Journal* **14**(3) (2013), 10–15.

[16] C. Capps and T. Wright, Toward a Vision: Official Statistics and Big Data, *Amstat News* **434** (2013), 9–13.