マルチソースデータに基づく東京 23 区における所得階層別の活動空間の社会的隔離の解明

孫　晨晨*・澁谷遊野・関本義秀*****

Unveiling Social Segregation of Activity Space among Different Income Groups in Tokyo 23 Wards Based on Multi-sources Data

Chenchen SUN*, Yuya SHIBUYA**, Yoshihide SEKIMOTO***

Abstract: With the world economy globalization and the advancement of the global industrial division system, the social stratification of global cities presents a trend of polarization. However, the existed literature on social segregation mostly focuses on residential segregation and lacks attention on segregation analysis based on activity space. Combining the Pseudo flow data with census data, this paper coupled socioeconomic attributes with activity trajectories to provide solutions to the common problem of lack of social attributes for trajectory data. Based on the coupled data sources, multiple metrics comparison of activity space social segregation in Tokyo 23 wards will be calculated to unveil the current segregation situation of activity space among different income groups and figure out the places facing more serious segregation problems.

Keywords: 社会隔離 (social segregation), 活動空間 (activity space), 東京 23 区 (Tokyo 23 wards)

1. Introduction

With the world economy globalization and advancement of the global industrial division system, the social stratification of global cities presents a trend of polarization (Sassen, 1991). As one of the most famous megacities in the world, Tokyo has attracted a large number of multinational corporations' headquarters to gather in the city center. The industries with high added value such as finance, trade, and technical service industries are booming, and the social elite is growing. While low-tech, low-wage migrant workers who provide all kinds of living services and labor-intensive production are also pouring into cities, which leads to the increasing polarization of society.

The economically and socially disadvantaged groups may face serious geographical isolation problems due to the uneven distribution of jobs, high housing prices, and the use threshold of service facilities. Spatial isolation will lead to a lack of communication and resource sharing among different classes. And vulnerable groups are concentrated in areas with a less friendly environment and scarce resources (Massey, 1987). For example, they may be exposed to a seriously polluted environment and have difficulty obtaining public services, thus further aggravating urban poverty and social injustice. In the past few years, the differences between different income groups have increased in Tokyo 23 wards (橋本健二, 2020). Meanwhile, rising commodity prices, as well as the sagging service industry hit by the epidemic which employs many low-income workers, could lead to further fragmentation of the activity space for different income groups. The possible increasing segregation in activity space could lead to a lack of communication and resource sharing among different classes, thus exacerbating the social inequality in Tokyo.

However, the existed literature on social segregation mostly focuses on residential segregation and lacks attention on segregation analysis based on activity space. Moreover, limited by the small sample of activity location trajectory data, city-scale analysis on activity space segregation has still been relatively unsearched.

*　学生会員　東京大学工学系研究科（The University of Tokyo）
〒135-0064　東京都江東区青海 2-2　E-mail：cc-sun@iis.u-tokyo.ac.jp
**　正会員　東京大学空間情報科学研究センター（The University of Tokyo）
***　正会員　東京大学空間情報科学研究センター（The University of Tokyo）

Therefore, this paper tries to unveil the current segregation situation of activity space among different income groups and figure out the places facing more serious segregation problems. Combining the People flow data with census data, this paper tries to couple socioeconomic attributes with activity trajectories to provide solutions to the common problem of lack of social attributes for trajectory data. Based on the coupled data sources, multiple metrics of activity space social segregation in Tokyo 23 wards will be calculated and compared to seek a further understanding of the segregation index differences and segregation status in Tokyo 23 wards.

2. Literature review

2.1 Rising concern about segregation of activity spaces

Research on social-spatial segregation originated from concern on western residential segregation (Herbert, 1967). However, people experience segregation not only in their residential places but also in other places where they undertake daily activities, such as the workplace and sites for non-work activities (Lee & Kwan, 2011; Wang, Li, & Chai, 2012). In the past few years, an increasing number of studies have been conducted to investigate activity-space-based segregation. Most of these studies focus on the differences of actual activities space among social groups (Järv et al., 2015; Wang, Li, & Chai, 2012). By comparing the spatial range and spatial-temporal modes of activities, some studies used geographic analysis tools and a series of indicators to measure the social-spatial segregation of different social strata and verified the effect of physical environments such as transportation facilities, and socio-economic attributes such as education level, race, income, car ownership and housing type on segregation. Other studies measured segregation according to the possibility of individuals contacting with other social groups in their activity space (Wong, 2011; Li & Wang, 2017).

2.2 Diversification trend of data sources

Small data, such as travel logs, are widely used in traditional social-spatial segregation research because they can provide activity trajectories and detailed socio-economic attributes. However, it takes a lot of workforce and resources to collect these small data with low update frequency. The development of mobile communication technology provides massive human-tracking data that record individuals' positions and times, including mobile phone, vehicle GPS, and social media data. They have advantages in the spatial-temporal analysis of activities for larger sample size and high update frequency. Järv et al.(2015) mapped individuals' activity spaces based on mobile phone data in Tallinn (Estonia) and relating these activity spaces to users' ethnic background. Huang and Wong (2016) combined the American Community Survey (ACS) data with Twitter data to analyze the activity patterns of Twitter users with different SES.

While existing studies have made significant advances in the analysis of activity-space-based segregation, several gaps in this area remain. Most researchers pay attention to the variation of segregation in different periods but rarely consider the time when measuring segregation, such as whether the time duration of different groups using the same activity space is same. Besides, although big data is being used more and more extensively in social segregation research, it lacks the respondents' socio-economic attributes obtained in traditional surveys for data collection methods and privacy protection limitations In the field of activity space analysis, there are few pieces of literature concerning the combination of large and small data to collect users' social attributes and location information at the same time. Moreover, there is a wide variety of segregation indicators used in different studies, making it difficult to compare them with each other.

Therefore, this study tries to address these limitations by using pseudo flow data which mimic actual people flow in the urban spaces based on simulation and census data, to export the dataset of individual activity location records with social class attributes for social segregation index comparison.

3 Methodology and data Sources

In this section, we describe the data and methods for the analysis. First, in subsection 3.1 we introduce the target city, Tokyo, and the data sources for analysis. Then we present four types of exposure index to compare in subsection 3.2.

3.1Study area and data sources

Tokyo 23 wards, covering an area of 627.53km2 with 9.65 million population, were selected as the study area (Figure 3.1). The basic analysis unit is 1km*1km grid with a radius of 500m. It is generally believed that the walking space within 500m radius is the area where people from different social classes can interact with each other. (reference) Thus, if different social groups are present in the same 1km grid at the same time, we consider this as alleviating social segregation, and conversely, if different social groups gather in different grids and are rarely exposed to other social groups within 1km, then the social segregation is severe.

As for the data sources, this paper mainly uses the Census data (2015) providing social attributes and Pseudo flow data for getting activity trajectories. The census data include information of respondents' age, gender, occupation, family type and housing attributes, which could represent respondents' social class to some extent. And the Pseudo flow data is a substitute microscopic representation of the actual people flow and provide everyone's spatial-temporal location in 1min interval (Kashiyama et al., 2022).

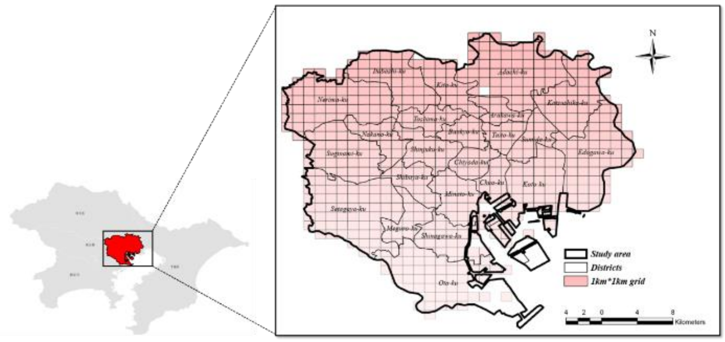


Figure3.1 study area

3.2 Combining census data with Pseudo flow data

Census data can provide socio-economic attributes of individuals to classify social classes but lacks information on individual activity trajectories, while Pseudo flow data can simulate the user's activity space but only contains limited socio-economic attributes such as age, gender, etc. Therefore, we need to combine the two types of data to obtain the social class classification and activity space distribution of users. Since census data do not contain information on residents' income, and the type of occupation is a determinant of income, this paper uses the type of occupation as an indicator to classify income groups.

Firstly, the 12 occupational types in census were divided into three categories, namely, Occupation Type A (Manager and technicians workers) including professional managers and professional technicians, Occupation Type B (General workers) including ordinary clerk workers, seller, service occupation workers, and occupation type C (Manual workers) including those engaged in agriculture, forestry and fishery, production engineering, construction and mining, transportation, cleaning and packaging, etc. Then, according to the proportion of these three occupation types in each census unit, four types of social areas were identified, namely, Clustering area of occupation group A, Clustering area of occupation group B, Clustering area of occupation group C and mixed residential area. The detailed division rules are as follows: Clustering area of occupation group A (Units with ratio of Occupation group A >=0.3 and ratio of Occupation group C<=0.1); Clustering area of occupation group B (Units with ratio of Occupation group B >=0.5 and ratio of Occupation group C <0.3 and ratio of Occupation group A <0.3); Clustering area of occupation group C (Units with ratio of Occupation group C >=0.3); Mixed area of different occupation groups (The rest units). Finally, the two types of data are overlaid according to the place of residence, and the social areas attributes of the census units in which the users live in the pseudo flow data are used as their social group attributes.

3.3 Segregation indexes

3.3.1 Traditional residence-based exposure index

Exposure refers to the degree of potential contact, or the possibility of interaction, between minority and majority group members within geographic areas of a city (Massey, 1988). Exposure indexes measure the extent to which two different social groups physically confront one another by sharing an activity space. The earliest well-known description of the exposure index is found in Massey's five-dimensional measure of social segregation proposed in 1988 (Massey, 1988). The indicators of the exposure dimension include interaction and isolation, which refer to the extent to which members of minority group A are exposed to members of majority group B and the extent to which minority members are exposed only to one other, respectively.

$$E_{A\times B} = \sum_{i=1}^{N} \frac{A_i}{A} * \frac{B_i}{T_i} \qquad (1)$$

$$E_{A\times A} = \sum_{i=1}^{N} \frac{A_i}{A} * \frac{A_i}{T_i} \qquad (2)$$

Where $E_{A\times B}$ is the interaction index and $E_{A\times A}$ is the isolation index. A represents the number of social group A members in the entire study area. A_i, B_i, and Ti are the numbers of group A members, group B members, and the total population of analysis unit i, respectively.

When only two social groups exist, the sum of these two indicators is 1. However, it should be noted that this is a global indicator that measures the exposure possibilities across the entire study area. In order to get local indicator, we calculated the proportion of group B residents in areal unit i as the residential exposure of group A to group B in areal unit i (i.e. $RE_{i,A\times B}$).

$$RE_{i,\ A\times B} = \frac{1}{A_i}\sum_{\alpha}^{A_i}\frac{B_{r\alpha}}{T_{r\alpha}} = \frac{B_i}{T_i} \qquad (3)$$

where A_i is the number of group A living in area unit i, $B_{r\alpha}$ and $T_{r\alpha}$ is the number of group B and all the groups living in the residence place of individual α. The indicator can be transferred to the ratio of group B in unit i finally.

3.3.2 Work place-based exposure index

The workplace-based exposure of analysis unit can be calculated as the average individual workplace exposure of residents living in this unit (Zhou, 2021). For example, if there is an individual α who belongs to group A, lives in areal unit i, and works in areal unit $j\alpha$. The individual workplace exposure to group B ($WE_{\alpha,A\times B}$) can be described as the ratio of group B in the workplace unit ($j\alpha$) of individual α.Then the workplace exposure of residence unit i can be calculated as follows:

$$WE_{i,\ A\times B} = \frac{1}{A_i}\sum_{\alpha}^{A_i}\frac{B_{j\alpha}}{T_{j\alpha}} \qquad (4)$$

$$WE_{i,\ A\times C} = \frac{1}{A_i}\sum_{\alpha}^{A_i}\frac{C_{j\alpha}}{T_{j\alpha}} \qquad (5)$$

$$WE_{i,\ C\times B} = \frac{1}{C_i}\sum_{\alpha}^{A_i}\frac{B_{j\alpha}}{T_{j\alpha}} \qquad (6)$$

Where the A_i and C_i are the number of group A and group C residents in analysis unit i respectively. And the $WE_{i,\ A\times B}$, $WE_{i,\ A\times C}$ represent the exposure possibilities to group B and C at their workplaces of group A living in unit i. The $B_{j\alpha}$ and $T_{j\alpha}$ are the number of group B and all the groups in workplace of individual α.

3.3.3 Activity places-based exposure index

To capture all relevant interaction spaces in evaluating the segregation experience of an individual, we adopt the concept of activity space. Many definitions of activity space have been proposed and one of them states that an activity space is "the subset of all locations within which an individual has direct contact as a result of his or her day-to-day activities" (Golledge and Stimson 1997, p. 279). In this study, we adopted this definition, and the activity space include space for work and nonwork activities. The activity places-based exposure of analysis unit can also be represented by the average individual activity exposure of residents living in this unit. And the individual activity exposure index should be the average value of exposure possibilities of each activity place. So, the calculation equation of $AE\_{i,A\times B}$ is as follows:

$$AE_{i,A\times B} = \frac{1}{A_i}\sum_{\alpha}^{A_i}\frac{1}{k}\sum_{\mu}^{k}\frac{B_\mu}{T_\mu} \qquad (7)$$

Where A_i is the number of group A residents in areal unit i, k is the number of activity places of individual α, B_μ and T_μ are the number of group B and all groups in the activity place of individual α.

3.3.4 Time considered activity places-based exposure index

Time duration spent in an activity place could also significantly influence the exposure possibilities and longer time staying in the same place together could highly improve the communication chances (Kwan, 2013; Kwan, 2018). Thus, time duration in the activity space is introduced in the calculation of exposure index as equation shows:

$$TAE_{\alpha,A\times B} = \frac{1}{t}\sum_{\mu}^{k}\frac{B_\mu}{T_\mu} * t_\mu \qquad (8)$$

$$TAE_{i,A\times B} = \frac{1}{A_i}\sum_{\alpha}^{A_i}AE_{\alpha,A\times B} \quad (9)$$

Where $TAE_{\alpha,A\times B}$ is the time-considered activity exposure index to group B of individual α who belongs to group A and $TAE_{i,A\times B}$ is the average individual exposure index of residents living in analysis unit i. B_μ and T_μ is the number of B group and all groups living in activity space μ of individual α, respectively. The t_μ is the time the individual α spent in activity space μ.

4 Results

4.1 Social area division

The trend of occupational polarization in global cities has been accelerating. There may be a serious social segregation between the social elite represented by managers of multinational companies, professionals,

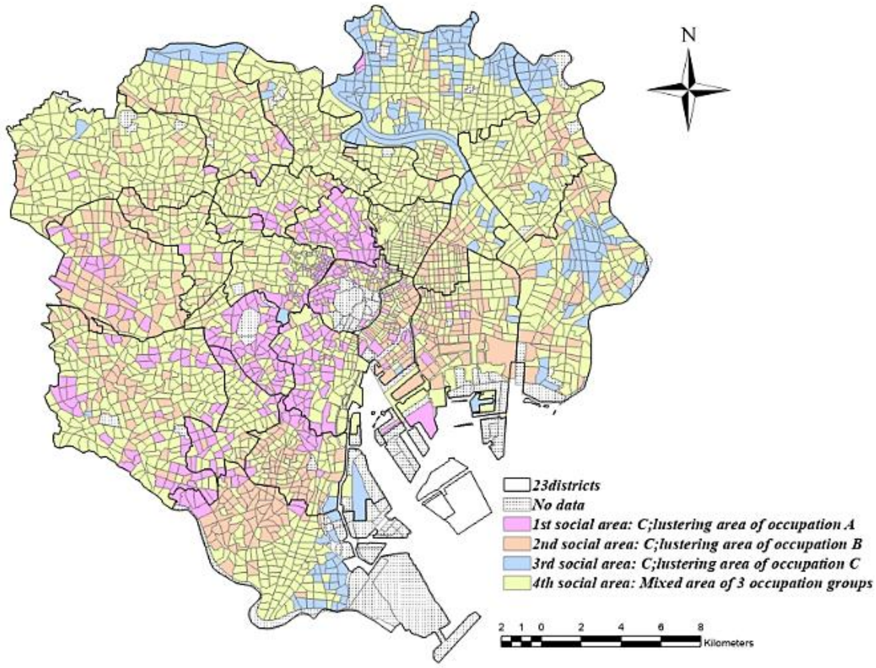


Figure 4.1 Spatial distribution of 4 types of social areas

technicians of consulting services and the social underclass of catering service, logistics and transportation workers who provide daily life services. Due to the lack of income attributes in the census, this paper finally used detailed occupation data to complete the social areas division. The statistics of number and proportion of employees from three occupations types in each block are shown in the following table：

Table 4.1 Statistics of employees from 3 occupation types in block units

	Occupation group A		Occupation group B		Occupation group C	
	Employees	Ratio/%	Employees	Ratio/%	Employees	Ratio/%
Max	2712	97.01	5535	100.00	1539	96.00
Min	0	0.00	1	2.98	0	0.00
Mean	288	26.91	605	54.81	202	18.28
SUM	917489	/	1926833	/	644036	/

Based on the ratio of each occupation type, the social areas division were conducted through threshold filtering. We got 362 units with clustering occupation group A, 599 units with clustering occupation group B, 213 units with clustering occupation C and 1871 mixed units. The spatial distribution of social areas is shown in the figure 4.1.

The degree of mixed residence of workers in Tokyo by occupation type is relatively high since over 60% of units in Tokyo are mixed social area with 3 occupation groups. Nearly 12% of the units belong to clustering area of Occupation A (managers and technicians), mainly in Bunkyo-ku, Shibuya-ku, Minato-ku, scattered in the southern part of Setagaya-ku, and the southern part of Meguro-ku. The 2nd social area, where Occupation B is concentrated, accounts for 20% of the total units, and is scattered in the northern part of Koto-ku, Chuo-ku, Shinagawa-ku, and the northern part of Ota-ku, while the 3rd social area, where Occupation C is concentrated, accounts for only 7% of the total, and is concentrated in Adachi-ku, the northern part of Itabashi-ku, the southern part of Ota-ku near Tokyo Bay and someplace near Shinozakimachi in Edogawa-ku.

4.2 Jobs-housing distribution of users from different social groups

The above subsection 4.1 described the social area division part. Then the Pseudo flow data with activity location were coupled with social area results based on users' residence location. Employment space can be an important part of activity space, so this paper first screens out users who have both employment and residence activity points, and calculates the Euclidean commuting distance between the employment and residence points. The number of users and average commuting distance were briefly counted by the social areas to which the residence belongs, as shown in Table 4.2.

Table 4.2 Basic statistics of commuting characteristics of users living in different social areas

	Social area 1	Social area2	Social area 3	Social area4
Number of users with job and residence	333809	726282	172976	2265966
Average commuting distance	6591	7630	8761	7757
Average commuting distance （Exclude JHS users）	7398	8442	9724	8639

From the table, it is easy to find that the commuting distance of residents from high income social class, i.e., social area1, which is dominated by managers and technicians, is shorter. The average commuting distance of residents from low-income social class, which is the social area3 dominated by manual workers, is 8761m, and the average commuting distance is as high as 9724m after excluding the JHS users who live and work in the same place. This may be related to the spatial distribution of the different social areas. Social area 1, which is closer to the city center, has access to enough jobs within a closer range, while social area 3, which is located at the edge of zone 23, is far from the employment center and requires a longer commute to obtain suitable jobs.

Moreover, we also visualized the spatial distribution of residential and employment kernel density of residents in each of the four types of social zones. In terms of residential density distribution, the overall distribution pattern can be seen in the previous description of spatial distribution of social areas in subsection 4.1, and there are some local high value points. In terms of employment density, the employment space of high-income social class has obvious clustering characteristics, and the high value of employment density is mainly located in Hibiya Station Tokyo Station, and the areas near Shinjuku, Shibuya, and Shinagawa Stations. In contrast, the employment space of the low-income social class is more evenly distributed, with higher employment densities around the residential area in addition to the high value of employment densities near Tokyo Station.

In summary, there are significant differences in commuting distances and spatial distribution characteristics of employment among residents from different social areas. Particular attention should be paid to the excessive commuting distance of residents in social area3 where the low-income social class is concentrated, which may worsen the employment problem of the low-income class.

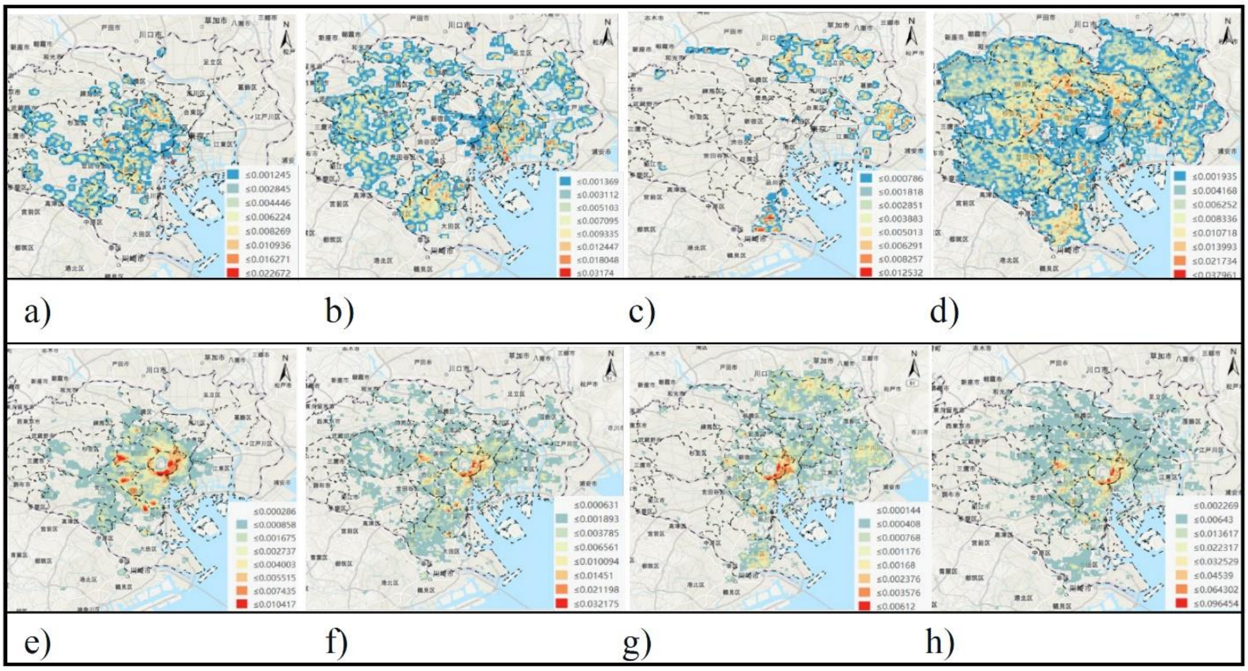


Figure 4.2 Jobs-housing spatial distribution of users living in different social areas

a), b), c), d) are the residential kernel density distributions of residents in 1st, 2nd, 3rd, and 4th social area respectively; e), f), g), h) are the distribution of employment kernel density distribution of residents in 1st, 2nd, 3rd, and 4th social area respectively

4.3 Exposure index aggregated in residence grid

In addition to comparing the jobs-housing characteristics of residents from different social areas, in order to further understand the social segregation of different income groups, this paper constructed four types of exposure indicators based on residential units, and the calculation method is detailed described in Subsection 3.3, and the final calculation results are as follows.

4.3.1 Global index of exposure

Using the 1km*1km grid as the basic analysis unit, the number of employees belong to three occupational groups in the census block cell is aggregated, and then the interaction and isolation indexes of occupational group A and occupational group C are calculated according to Equations 1, 2 (Subsection 3.3.1), respectively.

Table 4.3 Global exposure index results

Interaction index				Isolation index	
$E_{A\times B}$	0.555	$E_{A\times C}$	0.279	$E_{A\times A}$	0.166
$E_{C\times B}$	0.546	$E_{C\times A}$	0.216	$E_{C\times C}$	0.238

From the table, we can find that compared to the occupational group C (low-income social class), the occupational group A (high-income social class) has a lower isolation index and a higher interaction index. This means that Occupational Group A has a higher likelihood of interacting with other social classes and a lower level of social isolation. In the contrary, Occupational Group C is more exposed to groups in their own social class.

4.3.2 Comparison of four types of local exposure index

In the previous section, we measured the global exposure index based on residential space, and in this section, we will compare four types of local exposure index. In order to simplify the calculation, this paper only takes the 1st social area where the high-income social class gathers and the 3rd social area where the low-income social class gathers as the research objects, and calculates the probability that the residents of these two social areas have contact with other social groups in their residence, work, and activities space. The statistics of the calculation results are shown in Table 4.4.

Comparing the magnitudes of the four types of indexes, it can be found that the extreme differences of the indexes of each grid are decreasing as the space of the index measure expands from the residence to the work space and then to activity space. For example, the REi_AB range is [0.46, 0.67], the WEi_AB range is [0.50,0.60], and the AEi_AB range is [0.53,0.58]. Although the mean values of the indexes remained largely stable, the gap in social segregation between the grids narrowed significantly as the spatial scope of the measure was extended, especially the extreme difference between the exposure index from residence (RE) and the index from activity place (WE and AE) was significantly different. To some extent, this suggests that the daily activities of residents can facilitate contact between residents of some heavily residential segregated areas and other social class groups and reduce social isolation. Besides, the differences between activity space-based exposure result and time considered activity exposure index are not clear.

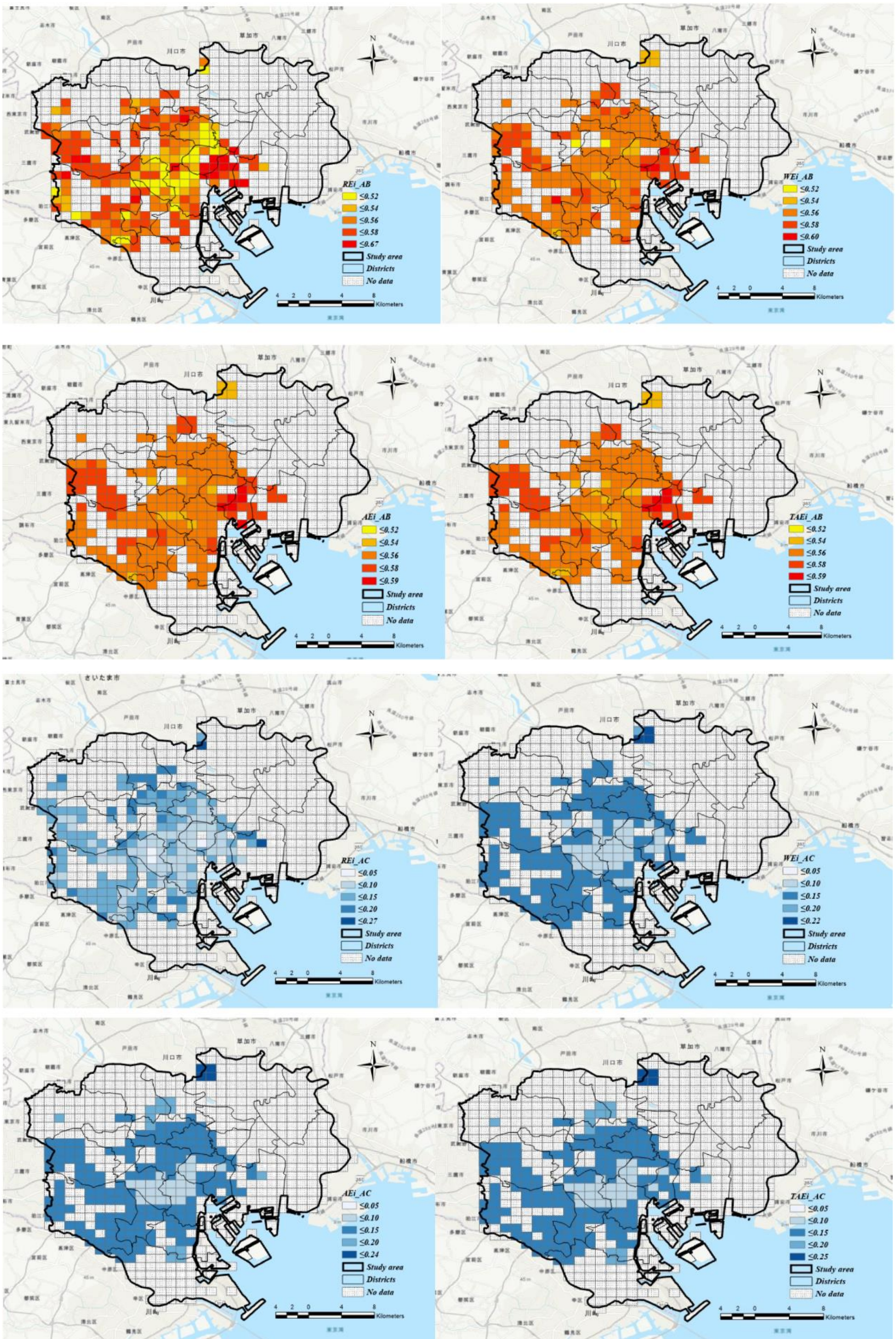
Table 4.4 Basic statistics of four types exposure index at the 1km*1km grid scale

Index	REi_AB	REi_AC	REi_CB	WEi_AB	WEi_AC	WEi_CB
Min	0.46	0	0.14	0.50	0.07	0.43
Max	0.67	0.27	0.65	0.60	0.22	0.58
ED	0.21	0.27	0.51	0.10	0.15	0.15
AVG	0.55	0.11	0.51	0.55	0.11	0.54

Index	AEi_AB	AEi_AC	AEi_CB	TAEi_AB	TAEi_AC	TAEi_CB
Min	0.53	0.09	0.45	0.52	0.08	0.45
Max	0.58	0.23	0.60	0.58	0.25	0.61
ED	0.05	0.14	0.15	0.06	0.13	0.16
AVG	0.55	0.12	0.54	0.55	0.12	0.54

Next, we visualize the four types of exposure metrics for each grid (Figure 4.3). Comparing the spatial distribution of the four indicators, it can be found that the residence-based exposure and workplace-based exposure are significantly different. The exposure possibilities to other social groups based on the workplace significantly increase. For example, in the REi_AB distribution map, Minato-ku, Bunkyo-ku, and Shibuya-ku have low local grid indicators, i.e., some degree of social segregation may exist, but in the WEi_AB and AEi_AB distribution maps, the indicators in these low-value areas improve significantly, i.e., compared to residential space segregation, the spatial segregation of activity space is lighter for residents living in these areas.

It is also noteworthy that the distribution of high value areas of exposure was generally consistent across indicators. For example, the occupation type A residents living in Chuo-ku, Suginami-ku, and Setagaya-ku have higher exposure possibilities to occupation group B and C than those residents in Minato-ku, Bunkyo-ku, and Shibuya-ku, in terms of residence exposure, workplace exposure, and activity space exposure. Furthermore, the



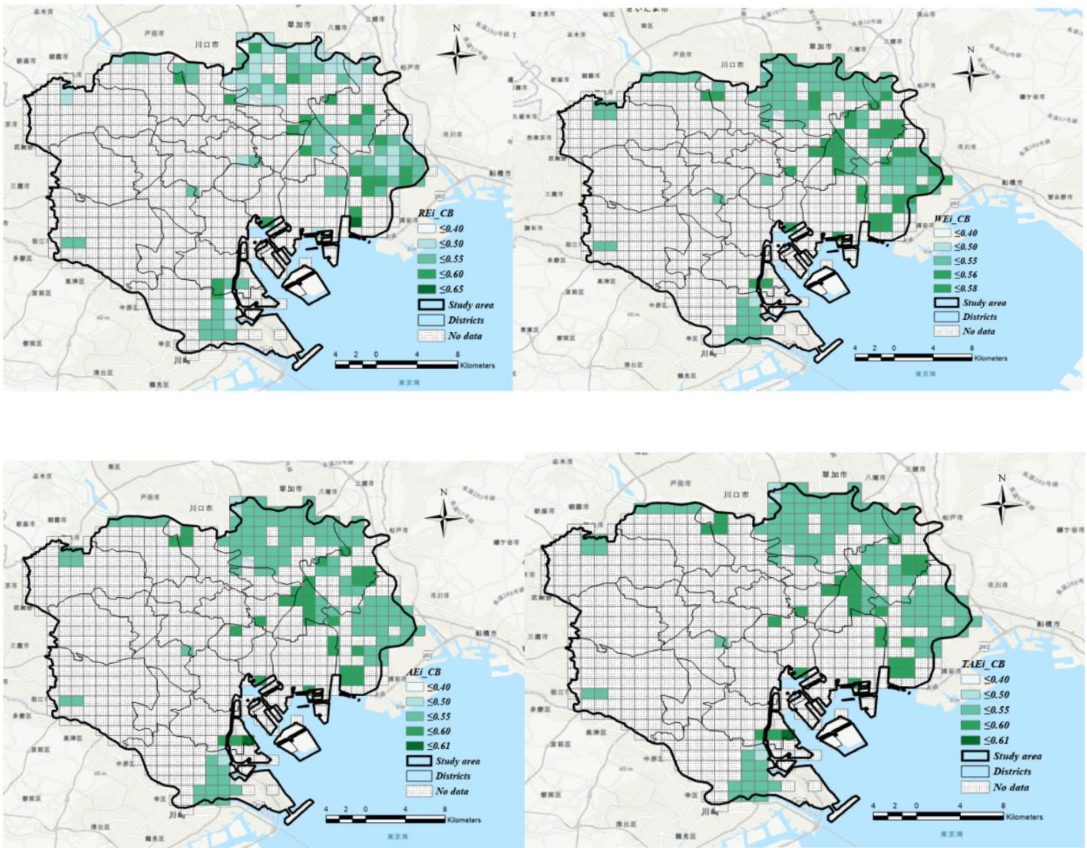


Figure 4.3 Spatial distribution of four types of exposure index at 1km*1km grid scale

spatial distributions of exposure in workplace and activity space are largely similar, with only local grid gaps. This may be because the pseudo flow data is a weekday data and most of the activity points are employment behavior points. Finally, activity space-based exposure and time-considered activity exposure indexes have no significant difference in spatial distribution and can even be said to be identical. It may be due to that we calculate the exposure possibilities using the occupation proportion of residents in activity grid, not the time duration proportion of actual visitors with different occupation of the activity grids.

5 Conclusions and discussion

This paper, taking the 23 wards of Tokyo as the study object, divided the social areas based on the occupation type from census and compared the commuting distance and jobs-housing spatial distribution of residents living in different social areas, and constructed four types of exposure indicators to measure the social segregation of residents in different social areas in terms of residential space, employment space and activity space by combining census data and pseudo flow data. The study results showed:1）There are significant differences in commuting distances and spatial distribution characteristics of employment among residents from different social areas. Particular attention should be paid to the excessive commuting distance of residents in social area3 where the low-income social class is concentrated. 2）Occupational Group A has a higher likelihood of interacting with other social classes and a lower level of social isolation. In the contrary, Occupational Group C is more exposed to groups in their own social class. 3）The results measured by residence-based exposure and workplace-based exposure index are significantly different. The exposure possibilities to other social groups based on the workplace are significantly higher than that based on residence in those areas with relatively low exposure chances.

By combining the two types of data, this paper makes up for the limited information on socioeconomic attributes in pseudo flow data and provides a reference for the application of people flow data in future studies of social segregation. Meanwhile, this paper constructs and compares four types of social exposure indicators, which solves the gap of existing studies that ignore the activity space and temporal dimension, and provides a system of social segregation indicators that can be directly compared with each other.

However, there are still some limitations. Firstly, this paper uses Pseudo flow data to perform segregation analysis. But it should be noted that the individual in the pseudo people flow is not identical to the actual person, even if the pseudo people flow matches various statistical distributions of the actual people flow. So, the main objective of this study focuses on the index comparison and methods exploration, not for the accurate actual segregation status evaluation. Besides, pseudo flow data is only a simulation of a working day and includes a limited number of activity points other than work. This can lead to the possibility that the activity space-based exposure measure may not reflect the spatial extent of activity during weekends and holidays. The second is about the unavoidable Ecological fallacy problem when combining census data and people flow data. Since the smallest analysis unit of census data is block, with an average area of 400*400m, the composition of people in the unit is still more mixed. So there is an error in taking the social area attribute of the block where the user lives as the social group attribute of the user directly. Finally, in the calculation of segregation index, in order to simplify the calculation and make a comparison between indexes, the proportion of occupations from census data of the grid where the workplace, activity, and residence are located, i.e., the proportion of occupations of the resident population, is directly taken as users' possible exposure to other social groups in that grid. This is not the proportion of occupations calculated based on actual users living, working or active in the grid at the same time.

References

橋本健二・浅川達人(2020)『格差社会と都市空間 ——東京圏の社会地図 1990-2010』, 鹿島出版会.

Golledge, R. G. (1997). *Spatial behavior: A geographic perspective*. Guilford Press.

Herbert, D. T. (1967). Social area analysis: A British study. *Urban Studies*, **4** (1), 41–60.

Huang, Q., & Wong, D. W. (2016). Activity patterns, socioeconomic status and urban spatial structure: what can social media data tell us?. *International Journal of Geographical Information Science*, **30** (9), 1873-1898.

Järv, O., Müürisepp, K., Ahas, R., Derudder, B., & Witlox, F. (2015). Ethnic differences in activity spaces as a characteristic of segregation: A study based on mobile phone usage in Tallinn, Estonia. *Urban Studies*, **52** (14), 2680–2698.

Kashiyama, T., Pang, Y., Sekimoto, Y., & Yabe, T. (2022). Pseudo-PFLOW: Development of nationwide synthetic open dataset for people movement based on limited travel survey and open statistical data. arXiv preprint arXiv:2205.00657.

Kwan, M.-P. (2013). Beyond space (as we knew it): Toward temporally integrated geographies of segregation, health, and accessibility: Space-time integration in geography and GIS Science. *Annals of the Association of American Geographers*, **103** (5),1078–1086.

Kwan, M.-P. (2018). The limits of the neighborhood effect: Contextual uncertainties in geographic, environmental health, and social science research. *Annals of the American Association of Geographers*, **108** (6), 1482–1490.

Lee, J. Y., & Kwan, M.-P. (2011). Visualization of socio-spatial isolation based on human activity patterns and social networks in space-time. *Tijdschrift voor Economische enSociale Geografie*, **102** (4), 468–485.

Li, F., & Wang, D. (2017). Measuring urban segregation based on individuals' daily activity patterns: A multidimensional approach. *Environment and Planning A: Economy and Space*, **49** (2), 467–486.

Massey, D. S., Condran, G. A., & Denton, N. A. (1987). The effect of residential segregation on black social and economic well-being. *Social Forces*, **66** (1), 29-56.

Massey, D. S., & Denton, N. A. (1988). The dimensions of residential segregation. *Social Forces*, **67** (2), 281-315.

Sassen, S. (1991). *The global city*. Princeton University Press.

Wang, D., Li, F., & Chai, Y. (2012). Activity spaces and socio-spatial segregation in Beijing. *Urban Geography*, **33** (2), 256–277.

Wong, D. W., & Shaw, S. L. (2011). Measuring segregation: An activity space approach. *Journal of Geographical Systems*, **13** (2), 127–145.

Zhou, X., Chen, Z., Yeh, A. G., & Yue, Y. (2021). Workplace segregation of rural migrants in urban China: A case study of Shenzhen using cellphone big data. *Environment and Planning B: Urban Analytics and City Science*, **48** (1), 25-42.