Cities 95 (2019) 102384

Contents lists available at ScienceDirect

Cities

journal homepage: www.elsevier.com/locate/cities

Revealing the relationship of human convergence–divergence patterns and land use: A case study on Shenzhen City, China

Xiping Yang[a,b], Zhixiang Fang[c,*], Ling Yin[d], Junyi Li[a,b], Shiwei Lu[e], Zhiyuan Zhao[f]

[a] School of Geography and Tourism, Shaanxi Normal University, Xi'an 710119, China
[b] Shaanxi Key Laboratory of Tourism Informatics, Xi'an 710119, China
[c] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China
[d] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, 1068 Xueyuan Road, Shenzhen 518005, China
[e] School of Architecture and Urban Planning, Huazhong University of Science and Technology, Wuhan 430074, China
[f] National & Local Joint Engineering Research Center of Geo-spatial Information Technology, Fuzhou University, Fuzhou 350002, China

ARTICLE INFO

Keywords:
Human mobility
Land use
Mobile phone data
Human convergence-divergence

ABSTRACT

Understanding the relationship between human mobility and land use has been a longstanding topic in multiple disciplines, including transport geography and urban planning. Recently, urban collective mobility patterns have become a hot research direction and has been explored at an unprecedented space–time scale due to the emerging big human tracking datasets (e.g., mobile phone data). However, only a few studies have comprehensively quantified the effects of land use on human mobility patterns while considering the influence of the scale of spatial analysis units. This study attempts to reinforce this knowledge by investigating urban human convergence–divergence patterns and their relationship with land use distribution characteristics at three popular types of spatial analysis units of human mobility studies (voronoi polygons, grid cells, and traffic analysis zones) using mobile phone data. A case study on Shenzhen, China is implemented, and results indicate that eight distinct convergence–divergence patterns could be extracted to describe urban collective mobility patterns despite the use of different types of spatial analysis unit. Moreover, the scale of spatial analysis units exerts a few effects on the quantification of the influence of land use distribution on human convergence–divergence patterns, but some common characteristics could be summarized from these discrepant results. The findings can help policy makers understand urban human mobility and can serve as a guide for urban management and planning.

1. Introduction

Investigating the interactions between human mobility and urban land use characteristics has been a classic research topic since the emergence of cities and has attracted considerable attention from the fields of human behavior, urban studies, and transport geography; such research attention has contributed to its potentially tremendous implications in urban spatial planning, traffic forecast, and optimization (Chen, Chen, & Barry, 2009; Gan, Yang, Feng, & Timmermans, 2018; Goodchild, Klinkenberg, & Janelle, 1993; Lee & Holme, 2015; Liu, Wang, Xiao, & Gao, 2012; Næss & Jensen, 2002; Newman & Kenworthy, 1996). Collective human mobility patterns are strongly associated with the spatial distribution of urban land use, which is the primary inherent motivation for many people to move in the city. Such movements are highly time dependent (e.g., commuting is a typical human activity during rush hours and traverses between work-related and residential lands).

Researchers worldwide have long made substantial efforts to understand complex human travel behavior and their interaction with the urban environment from individual and collective perspectives. For example, a body of literature has maximized travel survey data to examine the characteristics of individual travel behavior and its influence factors, such as sociodemographic (Kim, Woosnam, Marcouiller, Aleshinloye, & Choi, 2015; Kwan & Ren, 2008; Ta, Kwan, Chai, & Liu, 2015), built environment (Hong, Shen, & Zhang, 2014; Wang & Zhou, 2016), and land use characteristics (Litman & Steele, 2012). Although this type of dataset is able to present individual travel behavior at a micro level due to its abundant records of space–time activities and socioeconomic attributes of interviewees, two distinct drawbacks still exist. First, the collection of this dataset is costly, laborious, and time consuming, hence it is difficulty to capture travel behavior in time when changes occur (e.g., changes in land use or built environment of a

* Corresponding author.
E-mail address: zxfang@whu.edu.cn (Z. Fang).

https://doi.org/10.1016/j.cities.2019.06.015
Received 20 December 2018; Received in revised form 6 May 2019; Accepted 7 June 2019
0264-2751/

place). Second, the small sample size and low spatial–temporal resolution hinder the comprehensive understanding of human mobility from a macroscopic perspective (e.g., human spatiotemporal dynamics of a whole city) (Xu et al., 2015; Yuan & Raubal, 2012).

Fortunately, with the emergence of mobile intelligence that urges the ubiquitous use of location-aware devices and social applications, massive human tracking datasets (such as mobile phone call data, social media data, and smart card data) have been collected as a by-product of these devices, thereby increasing the possibility of monitoring the space–time movements of people at a large scale (Liu et al., 2015; Toch, Lerner, Ben-Zion, & Ben-Gal, 2018; Yue, Lan, Yeh, & Li, 2014). These emerging datasets stimulate researchers to review the conventional research questions about human mobility at an unprecedented space–time scale, with contents including but not limited to human mobility prediction models (González, Hidalgo, & Barabási, 2008; Song, Qu, Blumm, & Barabási, 2010), human activity space (Xu et al., 2016; Yuan & Raubal, 2016), socioeconomic characteristics (Lenormand et al., 2015; Luo, Cao, Mulligan, & Li, 2016; Xu, Belyi, Bojic, & Ratti, 2018), and urban spatial structures (Jiang, Joseph Ferreira, & Gonzalez, 2012; Galpern, Ladle, Alaniz Uribe, Sandalack, & Doyle-Baker, 2018; Lee, You, Eom, Song, & Min, 2018; Louail et al., 2015; Zhang, Liu, Tang, Cheng, & Wang, 2019). These datasets have also increased the attention to the study of collective human mobility patterns from different perspectives (Gao, 2015; García-Palomares, Salas-Olmedo, Moya-Gómez, Condeço-Melhorado, & Gutiérrez, 2018; Shaw, Tsou, & Ye, 2016; Sui & Shaw, 2018). Deep insights into human spatiotemporal mobility patterns and their interaction with the urban environment could benefit domains from urban planning and transportation to public health.

The human mobility patterns are closely related to land use characteristics. On the basis of this fundamental knowledge, some attempts have been made to infer or detect urban land use distribution from human spatiotemporal dynamics, which is extracted by using big geospatial datasets (Pei et al., 2014; Ríos & Muñoz, 2017; Toole, Ulm, Gonz, & Bauer, 2012). The general procedure is to construct a human dynamic time series for each spatial unit and cluster spatial unit with similar variation curves and then assign a certain land use to the spatial units of each group. However, the spatial unit is usually composed of different land uses, and its function changes over time (Tu et al., 2017); thus, further steps are needed to examine the land use characteristics of each group. Another main thread of this field is to investigate the relationship between land use and human movements to identify the effects of land use on human mobility (Gan et al., 2018; Gong, Lin, & Duan, 2017; K. Kim, 2018; Tu et al., 2018; Z. Yang et al., 2018). However, these studies mainly focus on quantifying the effects of land use on the ridership of urban public transport (e.g., subway or taxi) and thus cover only some specific areas (e.g., areas near subway stations) or people who take public transport. Moreover, the variations in the number of people for a place over time are not considered as a dynamic process in these analyses.

Along this thread of work, this study aims to reinforce knowledge on the relationship between land use distribution and human mobility patterns using mobile phone location data. For a spatial analysis unit, the convergence or divergence status respectively indicates whether the number of people is increasing or decreasing over time. The function of a spatial unit changes over time due to mixed land use distribution (Tu et al., 2017); this effect may lead to an alternative occurrence between human convergence and divergence during the day. Therefore, the main goal of this study is to investigate urban human convergence–divergence patterns and their underlying land use characteristics by addressing the following questions: What are the alternative patterns of convergence–divergence that occur in the city? What differences exist in land use distribution among these convergence–divergence patterns? We also examine whether common characteristics could be derived from different types of spatial analysis units. On the one hand, addressing these questions would give insights into the urban spatial–temporal dynamics of human convergence–divergence as well as their relationship with land use distribution. On the other hand, it would consider the effects of the modifiable area unit problem (MAUP) to seek some common characteristics from different types of spatial analysis units.

In this work, we establish a time series indicator to describe the human convergence–divergence alternative sequence. Clustering analysis was utilized to classify the spatial analysis units with similar convergence–divergence alternative patterns into groups. For each group, we sketch the outline of the dynamic curve to present the main human convergence–divergence processes. Next, multinomial logistic regression (MLR) is employed to reveal the land use characteristics for each human convergence–divergence pattern. Using mobile phone data in Shenzhen, China as a case study, we implement the above steps on three popular types of spatial analysis units (voronoi polygons, grid cells, and traffic analysis zones (TAZs)), which are frequently used to study human mobility patterns, to check whether consistent conclusions can be derived from different types of spatial units.

The rest of the paper is organized as follows. The study area and mobile phone data are introduced in Section 2. The methodology used in this study, including the extraction of human convergence–divergence patterns and a brief statement on MLR, is described in Section 3. Specific results and discussions are presented in Section 4. The conclusion of this study and future work are provided in Section 5.

2. Study area and dataset

2.1. Study area

The study area of this research is Shenzhen, which is located in the southeastern part of China and is adjacent to Hong Kong. Since the reform and opening-up policy in 1979, Shenzhen has experienced rapid development and has developed into a famous financial and technological center in China. Currently, the total area of Shenzhen is approximately 2000 km^2 and is divided into 10 administrative districts, which can be further classified into downtowns, suburbs, and rural areas on the basis of the extent of their economic development (Fig. 1). Therefore, the Futian, Luohu, and Nanshan districts are concentrated with various commercial, financial, and high-tech companies, whereas the other districts are mainly occupied by industrial parks and factories. Shenzhen's rapid economic development, openness, and inclusiveness have attracted a mass of immigrant workers from other provinces who seek job opportunities. As a result, Shenzhen has become one of the most populated cities in China, with a population of more than 15 million. In this study, we utilized the following kinds of spatial analysis units to check the effects of land use characteristics on human convergence–divergence dynamics: voronoi polygons, regular grid cells, and TAZs. The voronoi polygons are produced on the basis of cell phone towers, and the spatial scale for grid cells is 500 m × 500 m. TAZs are often used for traffic management and forecast. The mean area of voronoi polygons, grid cells, and TAZs are 0.32, 0.25, and 1.86 km^2, respectively.

2.2. Data description

The mobile phone location data used in this study were acquired from a main mobile phone operator, which accounts for more than 60% of the mobile phone market in Shenzhen. The dataset covers the footprints of approximately 16 million mobile phone users within a typical workday in 2012. Unlike the data comprising call detail records that sample individual locations only during communication events (such as phone calls and text messages) (Fang, Yang, Xu, Shaw, & Yin, 2017), the mobile phone data used in this study actively capture every mobile phone record (at the cell phone tower level) with a regular interval of approximately 1 h. These data were originally collected for troubleshooting by the mobile operator. Therefore, each record only contains

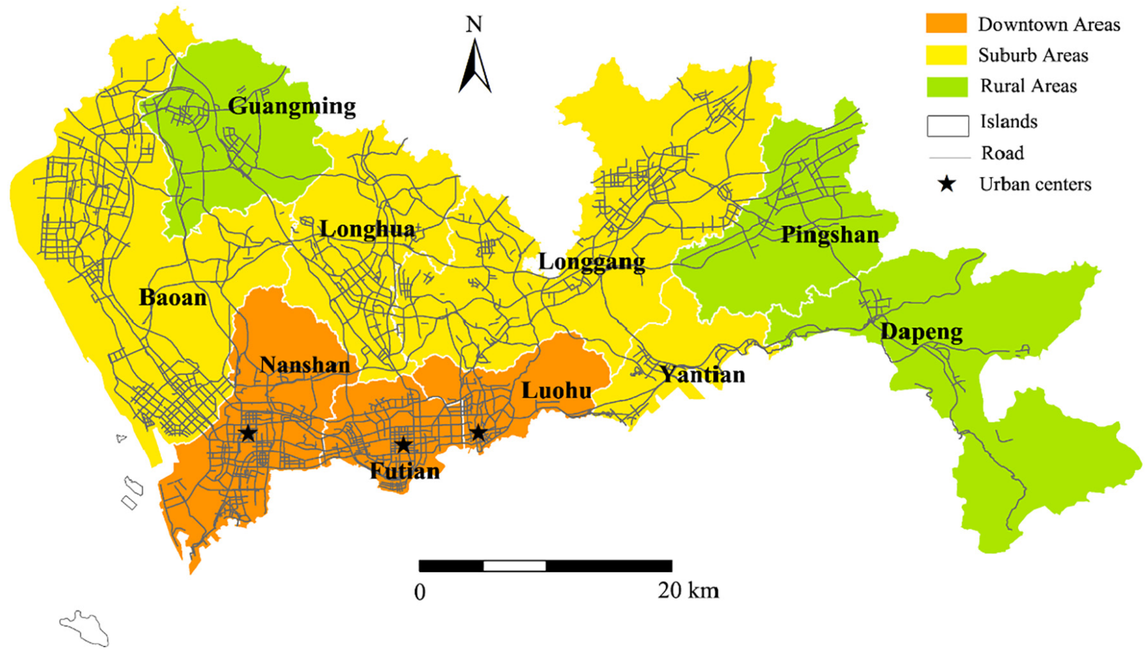


Fig. 1. Study area of Shenzhen.

Table 1
Instance of an individual's cell phone records in a day.

User ID	Record time	Time window	Longitude	Latitude
5u3d2c9 ******	00:20:33	00:00–01:00	113.***	22.***
5u3d2c9 ******	01:22:45	01:00–02:00	113.***	22.***
5u3d2c9 ******	02:26:50	02:00–03:00	113.***	22.***
5u3d2c9 ******	…		…	…
5u3d2c9 ******	23:30:52	23:00–24:00	113.***	22.***

four fields: user ID, recording time, and the longitude and latitude of the corresponding cell phone tower that the phone is connecting to. For privacy protection, the ID was encrypted before it was used for this study. Table 1 shows an instance of a cell phone user's records during a day. In summary, more than 5900 cell phone towers were extracted from the dataset, and each tower was labeled using a unique number.

Another dataset used in this study was Shenzhen's land use data, which were generated in the same year as the mobile phone location data. The dataset included 53 land use types in detail; we aggregated these land use types into six categories according to the urban land use and planning standards of land development. The six categories were commercial (wholesale, retail, accommodation, catering, and financial land, etc.), industrial (industrial parks, factories, and warehouses, etc.), residential, public (schools, hospitals, scenic spots, and public parks, etc.), transport and special lands (military, jail, and funeral, etc.). The special lands are little related with activities of urban residents, thus the other five land categories were used to analyze the relationship between land use and human convergence–divergence dynamics.

3. Methodology

The methodological workflow of this study is presented in two subsections. Section 3.1 describes how to extract human convergence–divergence patterns from mobile phone location data. Section 3.2 briefly explains the principle of multinomial logistic regression, which is employed to quantify the effects of land use characteristics on human convergence–divergence patterns.

3.1. Extracting human convergence–divergence patterns

In our previous studies, we extracted cell phone tower-based movements by checking the transformation of an individual's location between two adjacent windows and generated 23 movement matrices in an entire day (Fang et al., 2017; Yang et al., 2016). In this study, three popular divisions, namely, voronoi polygons, regular grid cells, and TAZs, are considered as spatial analysis units, which have been widely used in urban transport geographic studies and human mobility analyses. Therefore, we should further transform the tower-based movements into movements on the basis of the spatial scales of the three analysis units.

For voronoi polygons, which are generated on the basis of the original cell phone tower; thus tower-based movements can be directly used to extract human convergence–divergence patterns for voronoi polygons. Regular grid cells and TAZs need further aggregation to transform tower-based movements into grid-based and TAZ-based movements. Inspired from areal interpolation method in existing study (Yin et al., 2015), this study allocated the tower-based flows according to the proportion of overlapping areas between voronoi polygons and grid cells or TAZs. Fig. 2 depicts an example of aggregating tower-based OD flow into TAZ-based movements, and the green areas are overlapping areas between voronoi polygons and TAZs. The specific formula can be described as follow:

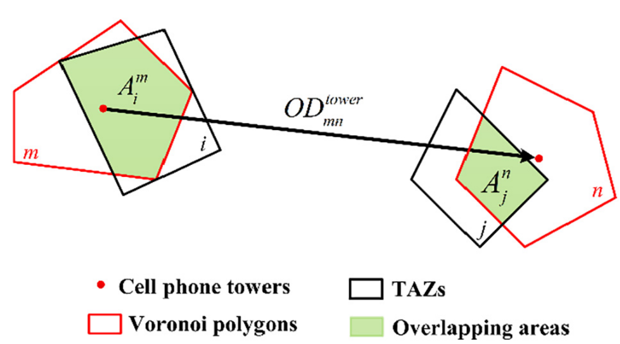


Fig. 2. An example of aggregating tower-based OD flow into TAZ-based movements.

$$OD_{ij}^{TAZ} = \sum_{m=1}^{N}\sum_{n=1}^{N}\frac{A_i^m}{A^m}\cdot\frac{A_j^n}{A^n}\cdot OD_{mn}^{tower} \quad (1)$$

where A_i^m represents the overlapping area between voronoi polygon m and TAZ i, A_j^n represents the overlapping area between voronoi polygon n and TAZ j. A_m and A_n represent the area of voronoi polygon m and n respectively, OD_{mn}^{tower} represents the tower-based flows from voronoi polygon m to voronoi polygon n, N is the number of cell phone towers. In this manner, we could generate human flows between grid cells or TAZs for each time slot, which would be used to extract the human convergence-divergence patterns.

The movement matrix can be denoted as $(i,j,C_{ij},T_t)^U$, where U represents the spatial analysis units, including the voronoi polygon, grid cell, and TAZ. C_{ij} represents the number of people from spatial unit i to spatial unit j, and T_t represents the time slot. On the basis of the movement matrix, we could calculate the inflow for each spatial unit by summing up the numbers of people arriving at the spatial unit from other spatial units during time slot T_t; similarly, the outflow of the spatial unit could be generated by summing up the flows departing from this spatial unit to other spatial units. The net flow is defined as the difference between the inflow and the outflow of the spatial unit; a positive net flow indicates that the number of people in a spatial unit increases during a time slot, whereas a negative net flow indicates otherwise. Therefore, net flow is considered as an indicator of the convergence (positive net flow) or divergence (negative net flow) status of the spatial unit during time slot T_t (Yang et al., 2016). According to Fang et al. (2017), the cumulative net flow of a spatial unit in time slot T_s can be calculated as

$$n_s = \sum_{t=1}^{s} netflow_t \quad 1 \le t \le 23, \quad (2)$$

where n_s represents the change in the number of people during the time period from T_1 to T_s. Thus, we can generate a time series of cumulative net flow during the day, denoted as $n = \{n_1, n_2, \cdots, n_{23}\}$, which is used to indicate the human convergence–divergence dynamic of the day.

For each spatial unit, we normalize the time series of cumulative net flow to compare the differences in the human convergence-divergence of the spatial units (e.g., comparing when the convergence or divergence reaches a maximum or minimum during the day). Normalization is calculated as

$$N_t = \frac{n_t}{max(|n|)}, \quad (3)$$

where $|n|$ represents the absolute value of each element in time series$n$ and $max()$ is the maximum function. In this manner, a normalized cumulative net flow time series $N = \{N_1, N_2, \cdots, N_{23}\}$ can be created for each spatial unit. The normalized time series is convenient for comparing the convergence–divergence patterns, and it is appropriate for the succeeding cluster analysis because the effect of magnitude is eliminated.

The objective of cluster analysis is to group spatial units with similar human convergence–divergence pattern into the same class. This study employs the self-organizing map (SOM) approach to cluster the voronoi polygons, grid cells, and TAZs according to their time series features. SOM is an unsupervised neural network-based approach that represents multidimensional features into a two-dimensional topological space, which makes cluster analysis feasible to execute for time series or multi-attribute data that can be characterized by vectors. This approach has been widely applied to geospatial data to mine hidden knowledge, including spatial distribution features and human mobility patterns (Andrienko et al., 2010; Sagl, Delmelle, & Delmelle, 2014). Therefore, this study utilizes the SOM approach for the above normalized time series to cluster the spatial units into different classes according to their time variation features. Each cluster indicates a distinct human convergence–divergence pattern.

3.2. Multinomial logistic regression

The human mobility patterns that occur in an area depend on the land function of the place. For example, residential and industrial areas would produce opposite human mobility patterns. This study employs MLR to address the relationship between land use characteristics and human convergence–divergence patterns. This approach, which has been widely applied to transport geographical analysis, is an extension of binary logistic regression to handle situations in which the dependent variables are more than two discrete categories (Jun, Choi, Jeong, Kwon, & Kim, 2015; Kim, 2018). Generally, the approach needs to select a class as the reference category and then compares the probability of other categories with the probability of the reference category. Given that the number of dependent variables is M, the first variable ($Y = 0$) is set as the reference group. The probability of each category can be calculated as

$$P(Y = k) = \frac{exp(Y_k)}{1 + \sum_{j=1}^{M\text{-}1} exp(Y_j)}, \quad (4)$$

and the model for each non-reference category can be expressed as

$$Y_k = ln\frac{P(Y = k)}{P(Y = 0)} = \beta_{0k} + \beta_{1k}x_1 + \cdots + \beta_{mk}x_m,\ k = 1,\ 2,\ \cdots M\text{-}1, \quad (5)$$

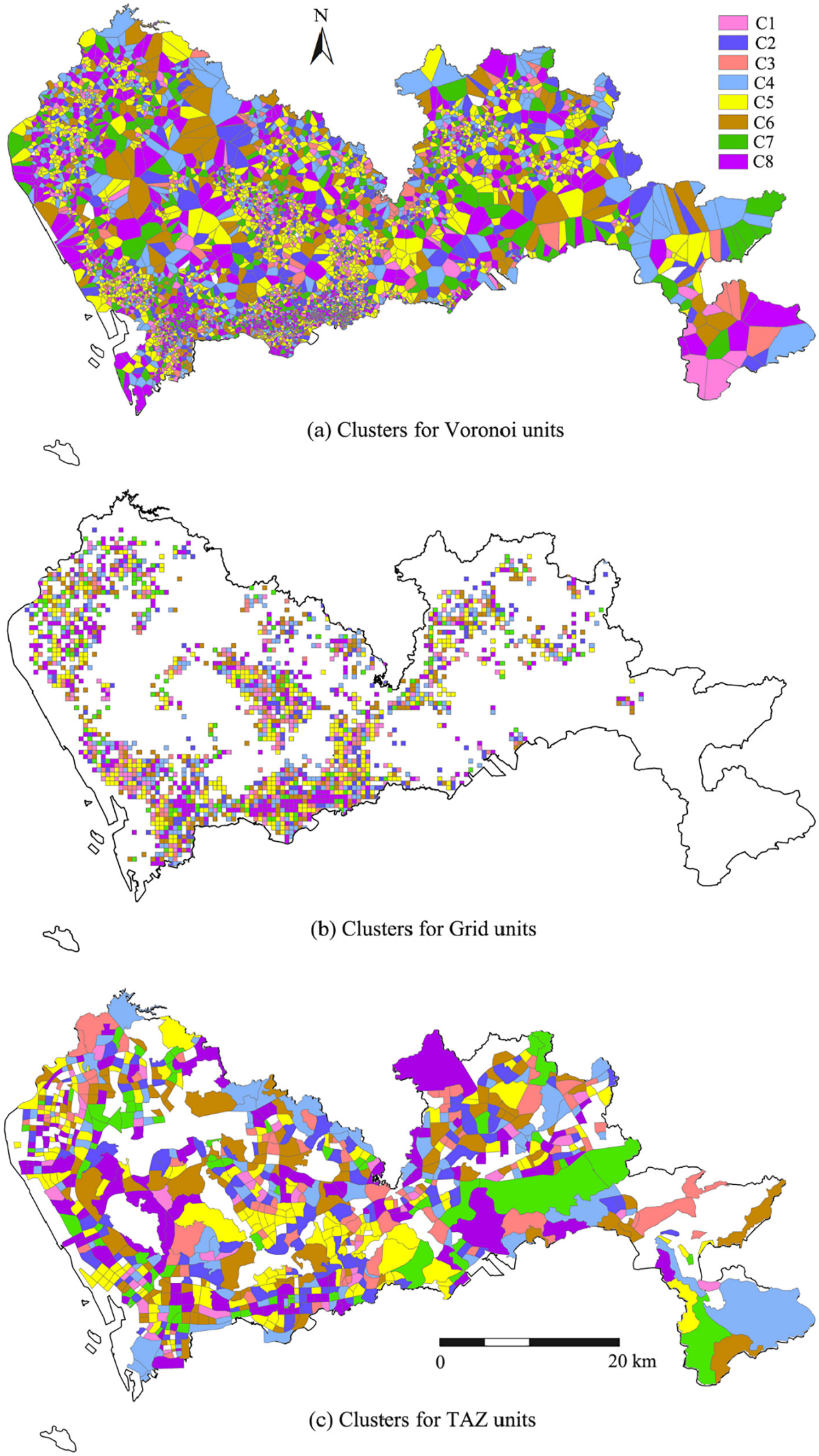
where M represents the total number of categories; k represents the non-reference categories; and β_i and x_i represent the coefficient and independent variables, respectively. The coefficient is estimated using maximum likelihood estimation. The coefficient β_i indicates the extent of the influence of independent variable x_i on the dependent variable. A positive value means that the independent variable increases the probability of the outcome, whereas a negative coefficient means otherwise.

The MLR model can be used to predict the probability of discrete dependent variables by some independent variables and to quantify the differences in the influence of independent variables on different dependent variables. The main objective of this study is to explore the relationship between human convergence–divergence patterns and land use characteristics. Therefore, the dependent variables of MLR in this study are the clusters that extracted human convergence–divergence patterns in the last section, and the independent variables represent the land use characteristics in the spatial analysis units, namely, the percentages of commercial, industrial, residential, public, and transport lands. Moreover, the land use mix of each spatial analysis unit is considered as an independent variable. We employ entropy to quantify the extent of the mix of different land uses, denoted as E_i. The calculation formula of land use mix can be found in the work of Tu et al. (2018). In addition to land use characteristics, we add the distance to nearest urban centers as an important independent variable because some studies have found a relationship between human mobility patterns and distance to urban central business districts (Gan et al., 2018). We utilize the Euclidean metric to calculate the distance from the center of each spatial analysis unit to its nearest urban center, denoted as D_i.

4. Results and discussion

4.1. Human convergence–divergence patterns

On the basis of the temporal characteristics of human convergence and divergence, we group the voronoi polygons, grid cells, and TAZs into eight distinct clusters, which are denoted as $C1$, $C2$, $C3$, $C4$, $C5$, $C6$, $C7$, and $C8$. The number of the clusters is determined according to our previous research (Yang et al., 2016). Fig. 3 shows the spatial distribution of the eight distinct human convergence–divergence patterns for the three spatial units. Some differences can be observed in the spatial distributions of the clusters among the different types of spatial



(a) Clusters for Voronoi units

(b) Clusters for Grid units

(c) Clusters for TAZ units

Fig. 3. Spatial distribution of each convergence–divergence dynamic for three spatial units.

analysis units because of the typical MAUP. This study focuses on checking whether some common patterns can be identified from the three different types of spatial analysis units.

Fig. 4 illustrates the centroids of eight clusters for the three spatial analysis units to show the temporal dynamics of human convergence and divergence during the day. According to the definition in the work of Fang et al. (2017), a successive convergence or divergence status for several time slots can be modeled as a human convergence process (HCP) or human divergence process (HDP). We apply the two conceptual models to the dynamic curve of each cluster to sketch the

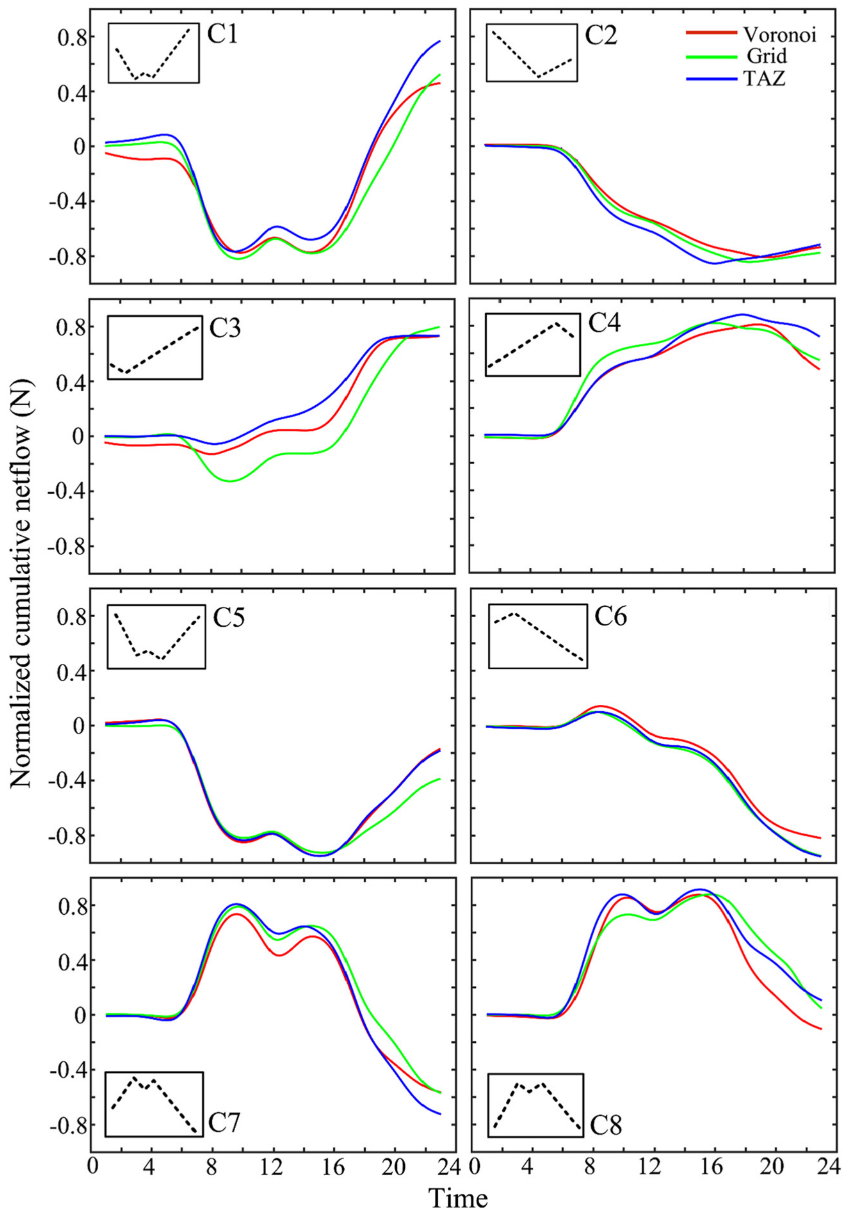


Fig. 4. Human convergence–divergence patterns for the three spatial analysis units.

outline of the human convergence-divergence process, which is illustrated by the black dashed line in the small inset. When the slope of the dashed line is greater than zero, it is an HCP; otherwise, it is an HDP. Although three different kinds of spatial analysis units are used to extract human convergence–divergence patterns, eight common patterns are uncovered from the three spatial units to describe human convergence-divergence during the day (Fig. 4). Moreover, some common characteristics or distinct differences in the human convergence–divergence process could be singled out from the eight patterns. A brief conclusion is provided as follows.

(1) There are no significant convergence-divergence appearing in all eight clusters from midnight (00:00) to early morning (around 05:00) when most people are asleep. Thus, the whole city is like a calm sea lying in a status of peace during those periods. After 05:00, obvious HCPs or HDPs arise in each cluster, indicating that citizens in Shenzhen have started their commuting activities in the early morning. The pressure of high housing prices and rent in Shenzhen (especially in the Futian, Luohu, and Nanshan districts) has caused many people working in urban centers to but opt to live in urban suburban areas to relieve their economic burden. Moreover, these people have to commute long distances for work, making them early birds.
(2) For clusters *C*1 and *C*5, a common feature is that they go through an analogous W-shaped dynamic process, forming a convergence–divergence alternate sequence: HDP–HCP–HDP–HCP. Specifically, an HDP occurring from 05:00 to 10:00 and an HCP after 15:00 indicate that these spatial units may be associated with urban residential land, where numerous residents disperse from these areas to workplaces in the morning and return in the afternoon. In addition, an HCP–HDP dynamic process occurs at noon (from 10:00 to 14:00) possibly because of lunch activities, leading to a small fluctuation appearing around 12:00. However, a significant difference between *C*1 and *C*5 is mainly reflected in the last HCP (occurring after 15:00 until midnight). The normalized cumulative net flow of *C*1 achieves approximately 0.5 at midnight, and the corresponding value of *C*5 only returns to the original level. On the contrary, clusters *C*7 and *C*8exhibit a reverse W-shaped dynamic process, namely, an M-shaped process, showing an HCP–HDP–HCP–HDP alternative sequence. Similarly, the main

difference between the two patterns lies in the last HDP after 15:00. Intuitively, industrial land may be dominant in the spatial analysis units of both clusters, and further investigation into their land use characteristics is still needed to determine the specific reason.

(3) As for clusters $C2$ and $C3$, an analogous V-shaped dynamic process is outlined, thereby generating an HDP–HCP alternative sequence from early morning to midnight. The only difference is that the fluctuation of $C3$ is smaller than that of $C2$. Moreover, the turning point between HDP and HCP in $C3$ happens around 09:00 in the morning, whereas the corresponding turning point in $C2$ occurs around 16:00. Conversely, the spatial analysis units in $C4$ and $C6$ first attract people from other places and reveal an HCP, followed by an HDP, which forms an inverted V-shaped dynamic process. The difference also lies in the time of the turning point; one happens at 08:00 in the morning, whereas the other takes place at about 17:00.

(4) From the above observation of human convergence–divergence dynamics, mutual complementary or opposite dynamic processes are observed among these clusters. For example, clusters $C1$ and $C7$ seem to experience an exactly opposite dynamic process. We speculate that the reason for this phenomenon is strongly related to the land use characteristics in the spatial analysis units and is the primary intrinsic motivation of human movement in cities.

4.2. Difference in land use characteristics among convergence–divergence patterns

For a spatial analysis unit, land use distribution determines its primary functional property, which further determines the period of occurrence of human convergence and divergence and further results in different human mobility patterns. Therefore, time-dependent human convergence–divergence patterns are strongly associated with the land use characteristics within the analysis unit. To drive the intrinsic motivation of each human convergence-divergence patterns, we use the MLR approach in quantifying the influence of land use on human convergence–divergence patterns and compare the main differences in the land use characteristics among these patterns. By comparing with the reference category, the approach can distinguish the differences in land use characteristics between other patterns and referential patterns by analyzing the logistic regression coefficients.

In this study, cluster $C1$ is set as the basis of the reference category. We utilize the boxplot to visualize the characteristics of land use distribution in this cluster for each type of spatial analysis unit (Fig. 5). Regardless of the type of spatial analysis unit, the percentages of commercial and public lands in most spatial units are less than 0.05 and 0.1, respectively. Meanwhile, the percentage of residential land in this cluster is significantly dominant relative to the percentage of industrial land, which may be the reason that the cumulative net flow being negative (W-shaped dynamic outline) during the daytime in this pattern (Fig. 4). In other words, the sum outflow in these spatial analysis units continues to be greater than the sum inflow during daytime. In addition, the entropy values in most of the analysis units are more than 0.7, which indicates that land use is highly mixed in this cluster. For the location of the spatial analysis units in this cluster, more than half of the places are located within 20 km of nearby urban centers. These characteristics may be the primary factors for generating a W-shaped dynamic convergence–divergence process in these areas. The next section investigates the land use characteristics of other clusters by comparing them with those of this cluster on the basis of the results of the logistic regression. In addition, the outliers of boxplot may be some units with extreme value (e.g., land use distribution, mixed entropy or distance) relative to most units. It can be seen that the number of outliers is decreasing from voronoi polygons, to Grid cells, to TAZs, which may be caused by spatial aggregation.

We utilize the statistical software Stata to execute multinomial logistic regression analysis. Tables 2, 3, and 4 show the regression results

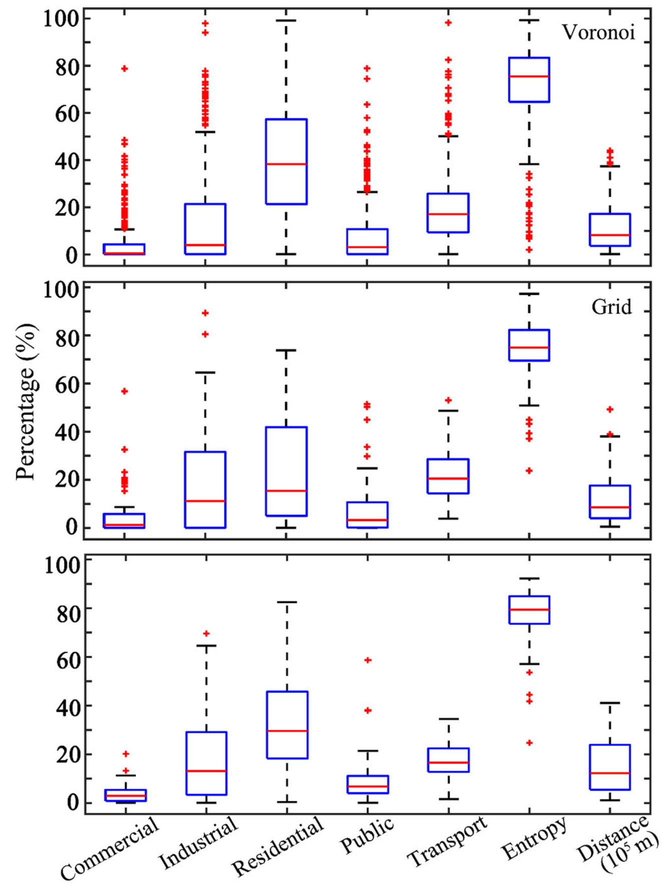


Fig. 5. The boxplot shows the distribution of each independent variable in cluster $C1$ for the voronoi polygons, Grid cells, and TAZs.

of the explanatory land use variables of each cluster for the three spatial analysis units. The results of the likelihood-ratio test (Prob > chi-square) for all three spatial analysis units indicate that the three regression models are significant. The pseudo R^2 demonstrates that the goodness of fit for TAZs is the best and that for the voronoi is the worst, which may indicate that the spatial aggregation may be a positive effect on the result of the multinomial logistic regression. In general, the coefficients are used to explain quantitatively the effects of independent variables on categorical dependent variables. According to Eq. (5), the coefficient measures the change of the natural logarithm of the odds ratio when the corresponding independent variable x_i has a unit increase. In this study, a coefficient greater than zero means that a unit increase in land use would enhance the probability of the convergence–divergence pattern belonging to other clusters versus the referential cluster $C1$. Conversely, a value less than zero would decrease the probability when a unit increase occurs in the independent land use variable. For the convenience of comparative analysis, we place the clusters with similar convergence–divergence patterns together to distinguish their land use characteristics according to the logistic regression coefficients.

For convergence-divergence pattern with a W-shape, such as clusters $C1$ and $C5$, all the coefficients of cluster$C5$ fail the test of significance for TAZs. In addition, commercial and public lands have a significant negative effect on cluster $C5$ at the spatial scale of voronoi polygons, whereas the mixed entropy has a negative effect at the level of grid cells. This finding demonstrates that the variations in the spatial scales of the analysis units make a significant difference in the regression results of cluster $C5$. The reason may be that land use distributions for both clusters are highly similar, making most of the coefficients statistically insignificant. However, transport land has a negative effect

Table 2
Results of multinomial logistic regression of convergence–divergence patterns for Voronoi polygons, with cluster *C*1 as the reference category.

Number of observation = 5860; The likelihood ratio chi-square = 1806.17; Prob > chi-square = 0.000; Pesudo R^2 = 0.078.

Variables	The logistic regression coefficient of each cluster						
	*C*2	*C*3	*C*4	*C*5	*C*6	*C*7	*C*8
	β_i	β_i	β_i	β_i	β_i	β_i	β_i
Commercial	0.045	1.904**	3.903***	−2.615***	0.134	1.142	3.711***
Industrial	−0.454	0.139	0.958**	−0.389	−0.092	0.592*	1.686***
Residential	−2.596***	−1.037***	−2.304***	−0.141	−4.440***	−5.074***	−4.000***
Public	0.753	0.294	1.120**	−1.056**	0.331	1.393***	0.958*
Transport	0.644	0.851	1.416***	−1.997***	0.410*	0.502	1.040**
Mixed entropy	−1.042**	−1.065**	−2.102***	0.356	−1.465***	−1.342***	−1.877***
Distance to center	−0.004	−0.005	0.016**	0.012**	−0.002	−0.002	−0.030***
Constant	1.572***	0.744*	1.695***	1.373***	2.071***	1.911***	2.827***

* Represents the significance of the regression coefficient at the 0.1 level.
** Represents the significance of the regression coefficient at the 0.05 level.
*** Represents the significance of the regression coefficient at the 0.01 level.

on cluster*C*5 for voronoi polygons and grid cells, thereby indicating that compared with the case of cluster *C*1, a unit increase in transport land of the spatial analysis unit may lead to a decrease in the probability of appearing in the convergence–divergence pattern of cluster *C*5.

As for clusters *C*2 and *C*3, which show a V-shaped convergence–divergence dynamics, residential land has a significant negative effect on both clusters for all three spatial analysis units. A unit increase in residential land would lead to a decrease in the probability of appearing in pattern *C*2 and *C*3 relative to the referential cluster *C*1. Compared with cluster *C*2, commercial land has a significant positive effect on cluster *C*3 for the spatial units of the Voronoi polygons and grid cells (Tables 3 and 4), indicating that the convergence-divergence process of cluster *C*3 may likely occur in areas with more commercial land relative to cluster *C*2. Commercial land has an appeal to urban citizens and may thus explain the prolonged HCP occurrence during the daytime in cluster *C*3 (Fig. 4). In addition, the significant negative coefficients of mixed entropy in cluster *C*3 for voronoi polygons and grid cells illustrate that the mix of land use in the spatial units of cluster *C*3 is less than that in the spatial units of clusters *C*1 and *C*2.

For the inverted V-shaped convergence–divergence process of clusters *C*4 and *C*6, a common characteristic is that residential land has a significant negative effect on the occurrence of both clusters for all three spatial units, whereas transport land has a positive effect on them. That is, compared with cluster *C*1, clusters *C*4 and *C*6 tend to appear in areas with less residential land and more transport land. However, a distinct difference between both clusters is embodied in commercial land, which only has a significant positive effect on cluster *C*4. This result also explains why intense HCP happens throughout work hours in this cluster. The mix of land use has a significant negative effect on both clusters versus cluster *C*1 at the spatial scale of voronoi polygons and grid cells. From the absolute value of coefficients, the spatial units of cluster *C*4 are less mixed than those of cluster *C*6.

For clusters *C*7 and *C*8, the regression coefficients of land use also show some similar characteristics. Specifically, industrial and public lands may have a positive effect on both clusters, thereby indicating that compared with those of cluster *C*1, the convergence–divergence processes of clusters *C*7 and *C*8 are more likely to occur at areas with more industrial and public lands. The significant negative effect of residential land further demonstrates that the spatial units of clusters *C*7 and *C*8 may be dominated by work-related land. Therefore, a similar M-shaped convergence–divergence process appears in these places (Fig. 4). Furthermore, the mix of land use in both clusters is significantly less than that of cluster *C*1. However, some differences are observed between the two clusters, as reflected by commercial and transport land. Compared with that of cluster *C*7, the human convergence-divergence process of cluster *C*8 is more inclined to arise in areas with more commercial and transport lands.

In addition, the effects of distance from spatial units to nearest urban centers on human convergence–divergence patterns has no significant and unified conclusion, which is inconsistent with Gan et al.

Table 3
Results of multinomial logistic regression of convergence–divergence patterns for grid cells, with cluster *C*1 as the reference category.

Number of observation = 2056; The likelihood ratio chi-square = 956.06; Prob > chi-square = 0.000; Pesudo R^2 = 0.109.

Variables	The logistic regression coefficient of each cluster						
	*C*2	*C*3	*C*4	*C*5	*C*6	*C*7	*C*8
	β_i	β_i	β_i	β_i	β_i	β_i	β_i
Commercial	−1.922	4.005**	4.076**	−2.697	1.144	1.397	4.470**
Industrial	−1.150*	−1.148**	0.393	−1.161**	−1.196**	0.043	1.034
Residential	−2.905***	−3.518***	−6.020***	−0.185	−7.533***	−9.224***	−9.576***
Public	2.902	2.051	3.499	−0.933	2.644*	4.468***	4.640***
Transport	0.288	1.985	0.962*	−2.431**	1.257	1.246	2.267*
Mixed entropy	−1.523	−4.124***	−3.237***	−2.245**	−2.204**	−3.252***	−3.397***
Distance to center	−0.009	0.023**	−0.018	0.001	0.008	−0.019	−0.004
constant	2.668***	3.503***	3.670***	3.786***	3.892***	3.827***	4.442***

* Represents the significance of the regression coefficient at the 0.1 level.
** Represents the significance of the regression coefficient at the 0.05 level.
*** Represents the significance of the regression coefficient at the 0.01 level.

Table 4
Results of multinomial logistic regression of convergence–divergence patterns for TAZs, with cluster *C*1 as the reference category.

Number of observation = 855; The likelihood ratio chi-square = 385.45; Prob > chi-square = 0.000; Pesudo R^2 = 0.130.

Variables	The logistic regression coefficient of each cluster						
	*C*2	*C*3	*C*4	*C*5	*C*6	*C*7	*C*8
	β_i	β_i	β_i	β_i	β_i	β_i	β_i
Commercial	−6.121	−5.905	9.155*	−0.676	−0.508	5.200	10.505**
Industrial	0.474	0.391	0.870	0.745	0.106	3.281**	2.931**
Residential	−3.557***	−2.902**	−9.491***	−0.243	−9.457***	−10.695***	−12.378***
Public	0.289	−0.221	4.909	1.848	4.615	5.711	3.763*
Transport	4.267**	3.829	4.141**	1.049	3.457**	5.599**	5.398***
Mixed entropy	−0.059	−2.094	−2.471	−2.354	−1.031	−4.515**	−3.929**
Distance to center	−0.030	0.003	−0.016	−0.034*	−0.028	−0.058**	−0.044*
constant	1.761	2.409*	3.596***	2.333**	2.341**	4.356***	4.938***

* Represents the significance of the regression coefficient at the 0.1 level.
** Represents the significance of the regression coefficient at the 0.05 level.
*** Represents the significance of the regression coefficient at the 0.01 level.

(2018); they found that the distance to urban centers is a significant factor that affects the ridership profiles of metro stations. The reason may be that the locations of urban metro stations are fixed, and unlike mobile phone data, smart card data can only capture the dynamics of these metro stations and cannot cover the whole area of the city.

In general, the scale of a spatial analysis unit makes a difference in the results of multinomial logistic regression, which is known as MAUP in spatial analysis. Compared with those of the spatial units of voronoi polygons and grid cells, the regression coefficients in some independent variables for TAZs are statistically insignificant. This result illustrates that large-scale spatial analysis units may aggregate some subtle differences in land use characteristics. Although differences exist in regression coefficients for the three types of spatial analysis units, some uniform characteristics can be summarized from these three tables. For example, a unit increase in commercial land would lead to an increase in the probability of appearing in convergence-divergence patterns *C*3 (for voronoi ploygons and grid cells, not significant for TAZs), *C*4 (for voronoi polygons, grid cells and TAZs) and *C*8 (for voronoi polygons, grid cells and TAZs). This indicates that the convergence-divergence pattern *C*3, *C*4 and *C*8 may likely occur in areas with more commercial land relative to the referential cluster *C*1. Therefore, we used up-arrow to label the three convergence-divergence patterns at the row of commercial land (Table 5). Based on the above discussion, we could draw a brief conclusion about the effects of change in land use on human convergence–divergence patterns. We employ up-arrow and down-arrow to denote the increasing and decreasing of probability in appearing corresponding human convergence–divergence patterns when there is an increase in the percentage of certain land use (Table 5). These conclusions are on the premise of cluster *C*1 being the referential category; and the blank indicates that the effect of land use is not significant.

Table 5
Effects of change in land use on human convergence-divergence patterns. The up-arrow and down-arrow denote the increasing and decreasing of probability in appearing corresponding human convergence–divergence patterns when there is an increase in the percentage of land use, and the blank indicates that the variation of land use has no significant effect on corresponding convergence–divergence patterns.

Land use	*C*2	*C*3	*C*4	*C*5	*C*6	*C*7	*C*8
Commercial		↑	↑				↑
Industrial						↑	↑
Residential	↓	↓	↓		↓	↓	↓
Public						↑	↑
Transport			↑	↓	↑		↑
Entropy		↓	↓		↓	↓	↓
Distance							

5. Conclusion

With the rapid development of information and communication technologies, the collection of massive human sensing datasets has become easy, providing valuable resources for the study of human mobility patterns and their relationship with the urban environment from a comprehensive space–time perspective. In this study, we aim to reveal the human convergence–divergence patterns of a city and their latent land use characteristics. Taking Shenzhen, China as a case study, we implemented the proposed methodological workflow on three popular types of spatial scales (cell phone tower-based voronoi polygons, grid cells, and TAZs). Although the spatial analysis units are used differently, some uniform conclusions can still be drawn from the experimental results.

First, eight common convergence–divergence dynamic patterns could be extracted from the three kinds of spatial analysis units via clustering analysis. By sketching the outline of each pattern using models of HCP and HDP, distinct characteristics were observed in the convergence-divergence processes among these patterns. These characteristics were mainly manifested in the alternating numbers and orders of HCP and HDP and the duration of HCP and HDP. In terms of the shape of convergence-divergence processes, four types of process were generalized from the eight patterns, namely W-shape (*C*1 and *C*5), inverted W-shape (*C*7and *C*8), V-shape (*C*2and *C*3), inverted V-shape (*C*4and *C*6). Among these shape, the first two may be mainly distributed in urban residential and industrial land respectively, and are strong related with people's daily home-work routine. For pattern *C*2, *C*4 and *C*6, they tend to appear in areas with less residential land and more transport land. *C*3 is likely to more related with urban commercial land.

As for land use, the scales of the spatial analysis units exerted some influence on the results of logistic regression. Nevertheless, some common effects of land use on the probability of the occurrence of human convergence–divergence patterns could be summarized from the three regression results. For example, a unit increase in commercial land for each spatial analysis unit would enhance the probability of occurrence of human convergence–divergence patterns *C*3, *C*4 and *C*8. The percentage of industrial and public land have a significant positive effect on convergence-divergence pattern *C*7 and *C*8. For residential land, its increase would decrease the probability of appearing in patterns*C*2,*C*3,*C*4,*C*6,*C*7 and *C*8. The transport land has a significant negative influence on pattern *C*5 but is positive related with pattern *C*4, C6 and C8. Finally, an increase in mixed of land use would generate a significant negative influence on patterns C3,C4,C6,C7 and C8.

These findings shed light on the relationship between human convergence–divergence patterns and land use distribution. Moreover, the results could help policy makers understand human mobility and eventually guide them in urban management and planning. In fact, convergence-divergence patterns represent the dynamic change of difference between inflow and outflow over time, which further indicates the spatiotemporal travel demand of urban residents, thus these findings could help optimize regional traffic planning to meet the dynamic travel demand of people. In addition, one could have an elementary knowledge about the characteristic of human convergence–divergence if the land use planning of a place (such as planning a new district or rebuilding urban village) is available, which could help adjust the land use planning.

However, some limitations can still be improved in future studies. First, this study only analyzed one weekday of human dynamics; further studies should focus on multiple workdays and compare the differences in human convergence–divergence patterns between weekdays and weekends. Second, apart from land use distribution, other impact factors (e.g., socioeconomic characteristics and accessibility of a place) should be considered to gain a better understanding of human convergence–divergence. Another important direction for future work is to predict human mobility patterns on a fine space–time scale in the context of an urban environment.

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgements

This research was jointly supported by the National Natural Science Foundation of China (NO.41801373, 41771473, 41771441), China Postdoctoral Science Foundation (No. 2019T120876, 2017M623112, 2018M632860, 2019M652244), Fundamental Research Funds for the Central Universities (Grants GK201803049, 2042017kf0235), and National key R & D plan (2017YFC1405302, 2017YFB0503802), and the Open Fund of Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Land and Resources (KF-2018-03-006).

References

Andrienko, G., Andrienko, N., Bak, P., Bremm, S., Keim, D., von Landesberger, T., ... Schreck, T. (2010). A framework for using self-organising maps to analyse spatio-temporal patterns, exemplified by analysis of mobile phone usage. *Journal of Location Based Services, 4*, 200–221.

Chen, C., Chen, J., & Barry, J. (2009). Diurnal pattern of transit ridership: A case study of the New York City subway system. *Journal of Transport Geography, 17*, 176–186.

Fang, Z., Yang, X., Xu, Y., Shaw, S.-L., & Yin, L. (2017). Spatiotemporal model for assessing the stability of urban human convergence and divergence patterns. *International Journal of Geographical Information Science, 31*, 2119–2141.

Galpern, P., Ladle, A., Alaniz Uribe, F., Sandalack, B., & Doyle-Baker, P. (2018). Assessing urban connectivity using volunteered mobile phone GPS locations. *Applied Geography, 93*, 37–46.

Gan, Z., Yang, M., Feng, T., & Timmermans, H. (2018). Understanding urban mobility patterns from a spatiotemporal perspective: Daily ridership profiles of metro stations. *Transportation,* 1–22.

Gao, S. (2015). Spatio-temporal analytics for exploring human mobility patterns and urban dynamics in the mobile age. *Spatial Cognition & Computation, 15*, 86–114.

García-Palomares, J. C., Salas-Olmedo, M. H., Moya-Gómez, B., Condeço-Melhorado, A., & Gutiérrez, J. (2018). City dynamics through Twitter: Relationships between land use and spatiotemporal demographics. *Cities, 72*, 310–319.

Gong, Y., Lin, Y., & Duan, Z. (2017). Exploring the spatiotemporal structure of dynamic urban space using metro smart card records. *Computers, Environment and Urban Systems, 64*, 169–183.

González, M. C., Hidalgo, C. A., & Barabási, A.-L. (2008). Understanding individual human mobility patterns. *nature, 453*, 779.

Goodchild, M. F., Klinkenberg, B., & Janelle, D. G. (1993). A factorial model of aggregate spatio-temporal behavior: application to the diurnal cycle. *Geographical Analysis, 25*, 277–294.

Hong, J., Shen, Q., & Zhang, L. (2014). How do built-environment factors affect travel behavior? A spatial analysis at different geographic scales. *Transportation, 41*, 419–440.

Jiang, S., Joseph Ferreira, J., & Gonzalez, M. C. (2012). Discovering urban spatial-temporal structure from human activity patterns. *Proceedings of the ACM SIGKDD international workshop on urban computing* (pp. 95–102). Beijing, China: ACM.

Jun, M.-J., Choi, K., Jeong, J.-E., Kwon, K.-H., & Kim, H.-J. (2015). Land use characteristics of subway catchment areas and their influence on subway ridership in Seoul. *Journal of Transport Geography, 48*, 30–40.

Kim, H., Woosnam, K. M., Marcouiller, D. W., Aleshinloye, K. D., & Choi, Y. (2015). Residential mobility, urban preference, and human settlement: A South Korean case study. *Habitat International, 49*, 497–507.

Kim, K. (2018). Exploring the difference between ridership patterns of subway and taxi: Case study in Seoul. *Journal of Transport Geography, 66*, 213–223.

Kwan, M.-P., & Ren, F. (2008). Analysis of human space-time behavior: Geovisualization and geocomputational approaches. *Understanding dynamics of geographic domains* (pp. 93–113). New York: CRC Press.

Lee, K.-S., You, S. Y., Eom, J. K., Song, J., & Min, J. H. (2018). Urban spatiotemporal analysis using mobile phone data: Case study of medium- and large-sized Korean cities. *Habitat International, 73*, 6–15.

Lee, M., & Holme, P. (2015). Relating land use and human intra-city mobility. *PLoS One, 10*.

Lenormand, M., Louail, T., Cantú-Ros, O. G., Picornell, M., Herranz, R., Arias, J. M., ... Ramasco, J. J. (2015). Influence of sociodemographics on human mobility. *Scientific Reports, 5*, 10075.

Litman, T., & Steele, R. (2012). *Land use impacts on transport: How land use factors affect travel behavior.* BC., Canada: Victoria Transport Policy Institute Victoria.

Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., ... Shi, L. (2015). Social sensing: A new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers, 105*, 512–530.

Liu, Y., Wang, F., Xiao, Y., & Gao, S. (2012). Urban land uses and traffic 'source-sink areas': Evidence from GPS-enabled taxi data in Shanghai. *Landscape and Urban Planning, 106*, 73–87.

Louail, T., Lenormand, M., Picornell, M., García Cantú, O., Herranz, R., Frias-Martinez, E., ... Barthelemy, M. (2015). Uncovering the spatial structure of mobility networks. *Nature Communications, 6*, 6007.

Luo, F., Cao, G., Mulligan, K., & Li, X. (2016). Explore spatiotemporal and demographic characteristics of human mobility via Twitter: A case study of Chicago. *Applied Geography, 70*, 11–25.

Næss, P., & Jensen, O. B. (2002). Urban land use, mobility and theory of science: Exploring the potential for critical realism in empirical research. *Journal of Environmental Policy & Planning, 4*.

Newman, P. W., & Kenworthy, J. R. (1996). The land use—Transport connection: An overview. *Land Use Policy, 13*, 1–22.

Pei, T., Sobolevsky, S., Ratti, C., Shaw, S.-L., Li, T., & Zhou, C. (2014). A new insight into land use classification based on aggregated mobile phone data. *International Journal of Geographical Information Science, 28*, 1988–2007.

Ríos, S. A., & Muñoz, R. (2017). Land use detection with cell phone data using topic models: Case Santiago, Chile. *Computers, Environment and Urban Systems, 61*, 39–48.

Sagl, G., Delmelle, E., & Delmelle, E. (2014). Mapping collective human activity in an urban environment based on mobile phone data. *Cartography and Geographic Information Science, 41*, 272–285.

Shaw, S.-L., Tsou, M.-H., & Ye, X. (2016). Editorial: Human dynamics in the mobile and big data era. *International Journal of Geographical Information Science, 30*, 1687–1693.

Song, C., Qu, Z., Blumm, N., & Barabási, A.-L. (2010). Limits of predictability in human mobility. *Science, 327*, 1018–1021.

Sui, D., & Shaw, S.-L. (2018). Human dynamics in smart and connected communities. *Computers, Environment and Urban Systems, 72*, 1–3.

Ta, N., Kwan, M. P., Chai, Y., & Liu, Z. (2015). Gendered space-time constraints, activity participation and household structure: A case study using a GPS-based activity survey in suburban Beijing, China. *Tijdschrift voor Economische en Sociale Geografie, 107*, 505–521.

Toch, E., Lerner, B., Ben-Zion, E., & Ben-Gal, I. (2018). Analyzing large-scale human mobility data: A survey of machine learning methods and applications. *Knowledge and information systems*.

Toole, J. L., Ulm, M., Gonz, M. C., & Bauer, D. (2012). Inferring land use from mobile phone activity. *Proceedings of the ACM SIGKDD international workshop on urban computing* (pp. 1–8). Beijing, China: ACM.

Tu, W., Cao, J., Yue, Y., Shaw, S.-L., Zhou, M., Wang, Z., ... Li, Q. (2017). Coupling mobile phone and social media data: A new approach to understanding urban functions and diurnal patterns. *International Journal of Geographical Information Science, 31*, 2331–2358.

Tu, W., Cao, R., Yue, Y., Zhou, B., Li, Q., & Li, Q. (2018). Spatial variations in urban public ridership derived from GPS trajectories and smart card data. *Journal of Transport Geography, 69*, 45–57.

Wang, D., & Zhou, M. (2016). *The built environment and travel behavior in urban China: A literature review.* (*Transportation Research Part D*).

Xu, Y., Belyi, A., Bojic, I., & Ratti, C. (2018). Human mobility and socioeconomic status: Analysis of Singapore and Boston. *Computers, Environment and Urban Systems, 72*, 51–67.

Xu, Y., Shaw, S.-L., Zhao, Z., Yin, L., Fang, Z., & Li, Q. (2015). Understanding aggregate human mobility patterns using passive mobile phone location data: A home-based approach. *Transportation, 42*, 625–646.

Xu, Y., Shaw, S.-L., Zhao, Z., Yin, L., Lu, F., Chen, J., ... Li, Q. (2016). Another tale of two cities: Understanding human activity space using actively tracked cellphone location data. *Annals of the American Association of Geographers, 106*, 489–502.

Yang, X., Fang, Z., Xu, Y., Shaw, S.-L., Zhao, Z., Yin, L., ... Lin, Y. (2016). Understanding spatiotemporal patterns of human convergence and divergence using mobile phone location data. *ISPRS International Journal of Geo-Information, 5*, 177.

Yang, Z., Franz, M. L., Zhu, S., Mahmoudi, J., Nasri, A., & Zhang, L. (2018). Analysis of Washington, DC taxi demand using GPS and land-use data. *Journal of Transport Geography, 66*, 35–44.

Yin, L., Wang, Q., Shaw, S.-L., Fang, Z., Hu, J., Tao, Y., & Wang, W. (2015). Re-identification risk versus data utility for aggregated mobility research using mobile phone location data. *PLoS One, 10*(10), e0140589.

Yuan, Y., & Raubal, M. (2012). Extracting dynamic urban mobility patterns from mobile phone data. *International conference on geographic information science* (pp. 354–367). Springer.

Yuan, Y., & Raubal, M. (2016). Analyzing the distribution of human activity space from mobile phone usage: An individual and urban-oriented study. *International Journal of Geographical Information Science, 30*, 1594–1621.

Yue, Y., Lan, T., Yeh, A. G., & Li, Q.-Q. (2014). Zooming into individuals to understand the collective: A review of trajectory-based travel behaviour studies. *Travel Behaviour and Society, 1*, 69–78.

Zhang, S., Liu, X., Tang, J., Cheng, S., & Wang, Y. (2019). Urban spatial structure and travel patterns: Analysis of workday and holiday travel using inhomogeneous Poisson point process models. *Computers, Environment and Urban Systems, 73*, 68–84.