Cities 77 (2018) 104–116

Contents lists available at ScienceDirect

Cities

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

Check-in behaviour and spatio-temporal vibrancy: An exploratory analysis in Shenzhen, China

Chao Wu[a], Xinyue Ye[e], Fu Ren[a,b,c], Qingyun Du[a,b,c,d,*]

[a] School of Resources and Environmental Sciences, Wuhan University, 129 Luoyu Road, Wuhan 430079, China
[b] Key Laboratory of GIS, Ministry of Education, Wuhan University, 129 Luoyu Road, Wuhan 430079, China
[c] Key Laboratory of Digital Mapping and Land Information Application Engineering, National Administration of Surveying, Mapping and Geoinformation, Wuhan University, 129 Luoyu Road, Wuhan 430079, China
[d] Collaborative Innovation Center Of Geospatial Technology, Wuhan University, 129 Luoyu Road, Wuhan 430079, China
[e] Department of Geography, Kent State University, Kent, OH 44242, United States

ARTICLE INFO

Keywords:
Vibrancy
Spatio-temporal variation
Heterogeneity
Check-in data
POI
Shenzhen

ABSTRACT

Urban vibrancy describes the attraction, diversity and accessibility of a place and exhibits spatio-temporal variability. The relationships between urban vibrancy and land-use configurations are significant for governments, planners and residents. To date, it is challenging for traditional census datasets to support real-time analysis with detailed spatial and temporal granularity. This article takes advantage of emerging crowdsourcing data and adopts social media check-ins over a 24-h period as a proxy for urban vibrancy. A framework that incorporates kernel density estimation (KDE), geographically and temporally weighted regression (GTWR) and the Herfindahl-Hirschman index (HHI) is proposed to explore the spatio-temporal distribution characteristics of vibrancy and the spatio-temporal relationships with the influential factors. The results show that the evolution of vibrancy is influenced by various factors that are heterogeneous over space and time. With a new perspective and deeper understanding of the varying spatio-temporal relationships between vibrancy and point of interest (POI)-based configurations, this study can offers meaningful implications for policy makers and planners regarding the improvement of resource utilization and the rational design of neighbourhoods.

1. Introduction

Urban vibrancy describes the attraction, diversity and accessibility of a place. It can also reflect human activities and their interactions with spatial entities, which exhibits spatio-temporal variability. Vibrancy is an essential element for achieving urban quality of life, originating from good urban form, well-developed urban functions and sufficient urban activities (Jin et al., 2017). Therefore, scholars are keenly interested in urban vibrancy, a concept first described by Jacobs as follows: "liveliness and variety attract more liveliness; deadness and monotony repel life" (Jacobs, 1961, 1969). Lynch (1984) added that vibrancy has three main components: urban morphology, urban function and urban society. Then, March et al. (2012) noted that measuring vibrancy should consider the range of experiences required for a healthy life, including privacy, rest and contemplation.

Although the definitions and measurements of vibrancy differ slightly, most scholars and theories have stated that urban vitality is closely intertwined with land-use configurations and is defined as the recognized human use of land in a city (Coupland, 1997; Li et al., 2016). Many existing studies have attempted to prove that reasonable planning and mixed land-use configurations can increase urban functionality, prolong activity intensities and improve vibrancy at a city scale. Therefore, the increasing interest in improving vibrancy requires a deeper understanding of land-use configurations. While land-use configurations are important to vibrancy, there are also difficulties in measuring vibrancy due to a lack of appropriate data and effective means. Moreover, the question of how to allocate the land use and facilities to promote urban vibrancy has still not been answered articulately. Briefly, there are two main issues in studying urban vibrancy. On the one hand, it is important to find a suitable proxy to precisely measure vibrancy; on the other hand, how to explore the quantitative relationships between vibrancy and land-use configurations is vital.

Fortunately, the rapid development of information and communication technologies (ICTs) has transformed the focus of GIScience towards the spatial, temporal, and dynamic relationships of human behaviours and the environment while also filling many of the gaps of traditional statistical datasets (Shaw, Tsou, & Ye, 2016). As one of the most popular types of geo-tagged data, the spatio-temporal patterns of

* Corresponding author at: School of Resources and Environmental Sciences, Wuhan University, 129 Luoyu Road, Wuhan 430079, China.
E-mail addresses: chaowu@whu.edu.cn (C. Wu), xye5@kent.edu (X. Ye), renfu@whu.edu.cn (F. Ren), qydu@whu.edu.cn (Q. Du).

https://doi.org/10.1016/j.cities.2018.01.017
Received 30 March 2017; Received in revised form 12 January 2018; Accepted 16 January 2018
Available online 15 February 2018
0264-2751/ © 2018 Elsevier Ltd. All rights reserved.

social media check-in data from different locations in a city indicate population distribution, mobility and preference. Check-in data have great potential to provide significant information regarding people's daily activities at different locations and times, and these data have been widely used to analyse patterns of human mobility (Wu, Zhi et al., 2014; Jurdak et al., 2015) and urban structure (Liu et al., 2016; Wu et al., 2016; Zhi et al., 2016). Therefore, this study uses social media check-in data during all hours of a day as a proxy for vibrancy according to the definition of vibrancy (Jacobs, 1961, 1969). Moreover, this study focuses on exploring the quantitative relationships between vibrancy and land-use configurations.

Methodologically, a multitude of models devoted to revealing the relationships between variables have been developed, including multiple linear regression, geographically weighted regression (Fotheringham, Brunsdon, & Charlton, 2003) and spatial filtering model (Griffith, 2004). Notably, human mobility and aggregation show high degrees of temporal and spatial regularity because people often travel to places with certain intentions (Schafer, 2000; Wu, Zhi et al., 2014; Li et al., 2016). For example, many people work from sunrise to sundown. Thus, people are often at their workplaces during daytime hours and resting at their residences during the night. Check-in data that record users' locations over time reflect one dimension of urban vibrancy and dynamics. Thus, location and time are important determinants of urban vibrancy. A popular local regression technique called geographically and temporally weighted regression (GTWR) is adopted in this article to assess heterogeneous relationships over space and time (Huang, Wu, & Barry, 2010). GTWR examines local parameters, rather than global, thereby providing a method for studying the effects of local geography on the relationships between urban vibrancy and its influential factors considering spatial and temporal non-stationarity.

Therefore, this article aims to (1) verify the applicability of the potential data source as a proxy of urban vibrancy, (2) explore the relationship between vibrancy and land-use configurations, and (3) effectively visualize the estimated coefficients representing quantitative relationships between vibrancy and land use. Specifically, this research proposes a framework that integrates kernel density estimation (KDE), geographically and temporally weighted regression (GTWR) and the Herfindahl-Hirschman index (HHI) to characterize the dynamic vibrancy that is related to land-use configurations and other location-specific variables. KDE is applied to measure the spatio-temporal density variation of human behaviour and mobility and verify the validity of check-in data to represent urban dynamics. The results of KDE capture the structure of vibrancy and reflect the less evenly distributed vibrancy over space and time. The influential factors of city vibrancy have been a continuing topic in related research fields, and a thorough understanding of these significant factors is crucial for promoting vibrancy. Considering the fine spatial and temporal granularity of urban vibrancy, GTWR provides excellent advantages in addressing spatial-temporal heterogeneity and is chosen to model the relationships between vibrancy and influencing variables in this article. Although GTWR can address spatial-temporal varying relationships well, the results are difficult to visualize, and the positive and negative effects of variables may offset each other ineluctably if only averages are used to explore spatio-temporal variations. Therefore, HHI is applied to decompose the spatio-temporal varying relationships between vibrancy and factors.

This article introduces social media data to study urban vibrancy. The proposed framework is applied to Shenzhen, China. The results shed light on urban vibrancy in relation to land-use configurations. In principle, this framework can be popularized and transferred to any other study area. Our study can provide useful information to city governments, local organizations, planners, residents, visitors, and anyone seeking to learn more about a city. The remainder of this paper is organized as follows. Section 2 reviews related research on urban vibrancy, land-use/POI-based configurations and GTWR. Section 3 describes the study area and data, including check-in data and POI data. Section 4 describes the framework and corresponding methods in detail. Section 5 discusses the spatial and temporal variations in vibrancy and its spatio-temporal relationships. Section 6 concludes the paper and describes our future work.

2. Literature review

2.1. The measurements of vibrancy

The concept of vibrancy is closely associated with activity intensity and is an important characteristic of public spaces, particularly at the scale of streets and neighbourhoods. Jacobs (1961, 1969) claimed that urban vitality is street life over a 24-h period. Montgomery (1995, 1998) described urban vibrancy as the number of people in and around streets or neighbourhoods (i.e., pedestrian flow) at different times of day and night. With increasing interest in measuring vibrancy, different factors, such as housing/land prices (Nicodemus, 2013; Wu et al., 2016), cultural clusters (Stern & Seifert, 2010), night-time light data (Mellander et al., 2015), built-environment attributes (Winters et al., 2010), population census data, employment rates (Harvey, 2001), and accessibility and connectivity (Braun & Malizia, 2015), are being used to evaluate urban vibrancy. However, these previous studies have all neglected an important issue: One of the most striking features of urban vibrancy is the varying number of people in a location over time (Jacobs, 1961; Montgomery, 1998). Although traditional data collection methods, such as surveying and interviewing, provide detailed and authentic user profiles that include gender, age, and work, those indicators are costly and static and hardly represent population dynamics.

Geo-tagged data, such as social media data, global positioning system (GPS) data from taxis, bus smart-card data (SCD) and cell phone signal data, provide significant advantages that facilitate people to capture the diverse profiles of the urban structure from the perspective of social sensing (Liu et al., 2015). Compared with other location-based data (LBS) and traditional datasets, social media check-in data offer two prominent advantages: (1) Check-in data are spatial footprints of people's activities that can reflect individual travel demands and connect to POI types, and (2) check-in data are relatively easy to obtain through corresponding application program interfaces (APIs) without encountering privacy issues or data qualification. A number of novel studies have used check-in data to analyse a population's mobility (Lin & Cromley, 2015), reflect urban functional structures (Zhen et al., 2017; Zhi et al., 2016) or explore the effects on other urban geographical and economic elements (Shen & Karimi, 2016; Wu et al., 2016).In addition, a number of studies have attempted to correlate check-in data with population density information and DMSP/OLS night-time light image data (census data and remote sensing data, respectively) to assess urban vibrancy (Duggan & Brenner, 2013; Li, Goodchild, & Xu, 2013; Lin & Cromley, 2015; Jendryke et al., 2017). The aforementioned studies have confirmed that check-in data contain precise spatial and temporal information and can reveal activity patterns beyond the night-time residential geographies of conventional and static statistical data sources (Longley et al., 2015). However, these studies have mainly concentrated on using check-in data alone to reflect certain structures and lack analyses and discussions of internal mechanisms. Few studies have taken full advantage of social media check-in data to study urban vibrancy and explore its significant influencing factors.

2.2. Vibrancy and land use

The effect of land use on vibrancy is not a new area of inquiry. The earliest exploration of this topic can be traced back to Howard (1898) and Geddes (1915), who studied living-condition improvements through the optimization of land-use distribution. Before these studies, no clear definition of vitality was available. Jacobs (1961, 1969) was the first to describe vibrancy and implied a correlation between mixed land uses and city liveliness. Based on Jacobs's work, Montgomery (1995, 1998) improved the definition of vibrancy, suggesting that urban vibrancy could be achieved with a complex diversity of primary land uses over the long term. Previously, most research relied on theoretical and few quantitative methods. Sharkova and Sanchez (1999) applied ordinary least squares to measure the influence of neighbourhood types, land uses, socioeconomic characteristics, and urban accessibility on vibrancy. These authors concluded that density and land-use variables are important indicators of neighbourhood and residential quality. Chhetri, Stimson, and Western (2006) summarized three factors of neighbourhood attractiveness from the perspectives of aesthetic, amenity, and social interaction. Mehta (2007) used multivariate and factor analyses to explore the effects of eleven street characteristics on the liveliness index. In addition, many real-estate studies have revealed that land use significantly affects housing prices, potentially suggesting that vibrancy can attract people (Dai, Bai, & Xu, 2016; Jang & Kang, 2015; Wen & Tao, 2015). Thus, land use and facility layout may effectively affect urban vibrancy.

With the popularity of big data, some empirical studies have applied new data to study urban vibrancy and concluded that land-use configurations significantly affect urban vibrancy (Jacobs-Crisioni et al., 2014; Li et al., 2016; Ye, Li, & Liu, 2017; Yue et al., 2016). However, Jacobs-Crisioni et al. (2014) and Yue et al. (2016) measured mixed land use and its effects on vibrancy based on mobile phone data but did not address the existence of spatio-temporal heterogeneity. Li et al. (2016) used only the number of houses and consumption-related POIs without considering other types of POIs and mixed functions to study the mechanisms of spatio-temporal variation. Ye, Li and Liu (2017) used small catering businesses to study economic vibrancy and focused on the effects of urban morphology. Although the previous related study demonstrated that a thorough understanding of these significant factors is crucial for promoting vibrancy, the results of empirical studies have not reached a consensus either, and many of them have only focused on one or other specific factors. This article considers spatio-temporal heterogeneity and a comprehensive system of influential factors.

2.3. Shenzhen's vibrancy

At the forefront of China's reform and opening up policy, Shenzhen is one of the fastest growing cities in China and has high urban vibrancy. Recently, many scholars have studied the vibrancy in Shenzhen. Wu et al. (2014) and Wu et al. (2016) showed that households are willing to pay for properties and urban amenities, which have profound effects on vibrancy. Cheng et al. (2016) and Wu et al. (2017) explored the attraction of facilities by considering accessibility for residents in Shenzhen. Yue et al. (2016) explored the links between the proposed indices of mixed land use and vibrancy based on mobile phone data. Gong, Lin, and Duan (2017) captured the spatio-temporal structure of dynamic urban spaces by using metro smart card records and revealed the relationships between human activity and land-use types. Ye, Li, and Liu (2017) explored the relationship between urban morphology and urban vitality represented by small catering businesses in Shenzhen. Previous studies have offered significant insight for analysing vibrancy in Shenzhen. To date, no study has applied social media data to reflect Shenzhen's vibrancy, and the spatio-temporal heterogeneity has not addressed.

3. Study area and data

3.1. Case study: Shenzhen, China

This case study is based on Sina microblog check-in data in Shenzhen, China (22°27′ to 22°52′N, 113°46′ to 114°37′E). Upon implementing the Open Door Policy in 1979, Shenzhen became China's first Special Economic Zone (SEZ), and since then, this city has been one of the most successful examples of such zones. Currently, Shenzhen is one of the most important and developed cities in South China. Indeed, this city is listed as one of the top 20 cities in the City Momentum Index in the Jones Lang LaSalle (JLL) report of 2016.[1] Shenzhen has achieved this good performance based on its geographical advantages, urban radiation, growing population, open-door policy, and other characteristics. In this paper, our study area does not include Shekou, Neilingding Island, or other islands to ensure that the study area is spatially continuous and comparable.

The analysis unit for studying vibrancy is important (Yue et al., 2016). The boundaries of traditional municipal organization units, such as neighbourhoods and traffic analysis zones (TAZs), are too large to reflect the characteristics of vibrancy and ensure sufficient precision. In addition, the original check-in data consist of large numbers of discrete elements that are not advantageous for exploring spatial distribution patterns. Therefore, a 1 km × 1 km grid is used as the analysis unit to reflect vibrancy on a fine scale. Shenzhen can be divided into 2202 grids. Subsequently, we use a baseline grid unit to calculate the number of check-ins by the hour and correlate the results with POI data.

3.2. POI data and check-in data

In China, the POI categories are in accord with land-use classifications. Compared with land-use data, POI data have the following advantages: (1) POI data have greater flexibility for studying scale issues because point data can be transformed to arbitrary scales. (2) People's preferences and social functions can be represented by their interactions with POIs rather than by land-use type. (3) The statistical granularity of POI data is much finer; therefore, these data provide relatively useful information. POI data are thus used to reflect land use. POI data are obtained from AutoNavi, a web-mapping platform and location-based service provider in China, using API. A total of 510,635 POIs are obtained, and this article classifies these points into 12 types according to the classification of the AutoNavi POI and Code for Classification of Urban Land Use and Planning Standards of Development Land (GB50137-2011). The details of the POIs in each category are shown in Fig. 1. In this article, the effects of the mixing degree of POIs, housing-related POIs (HPOI), consumption-related POIs (CPOI) and traffic-related POIs (TPOI) on urban vibrancy are emphasized and explored. The functional mixing degree is calculated based on the original classification. The number of HPOIs, CPOIs, TPOIs and other POIs (OPOI) are measured in each grid. Here, HPOI includes residential districts and residential services, such as hotels. In contrast, CPOI includes catering services, shopping services, life services and sports leisure services. TPOI consists of bus stations, subway stations, railway stations, and airports, among others. Notably, workplaces are not classified as a separate category because the locations of workplaces cannot be chosen by most residents and therefore cannot reflect the residents' preferences. Fig. 2 describes the distribution of POIs and the road network in Shenzhen.

Check-in data are a type of geo-tagged data that have been widely studied in recent years. A check-in record can be used to log an individual activity at a specific time and location. As a popular variety of crowdsourcing data, check-in data can provide a new perspective for studying people's spatial and temporal preferences in urban locations (Batty, 2013; Shen & Karimi, 2016). Check-in records from January 2015 to December 2015 were collected from the social media Sina Visitor System. Check-in data have some limitations, such as sampling bias, location inconsistencies and the existence of zombie fans. Based on these limitations, we impose the following criteria to clean the check-in data (Wu, Zhi et al., 2014): (1) POIs outside of Shenzhen were deleted; (2) POIs with less than two check-ins were deleted and labelled invalid POIs; (3) users who checked in less than once were deleted; and (4) users who checked in at the same location more than once in a half-

[1] URL for reference: http://www.jll.com/research/165/city-momentum-index-2016

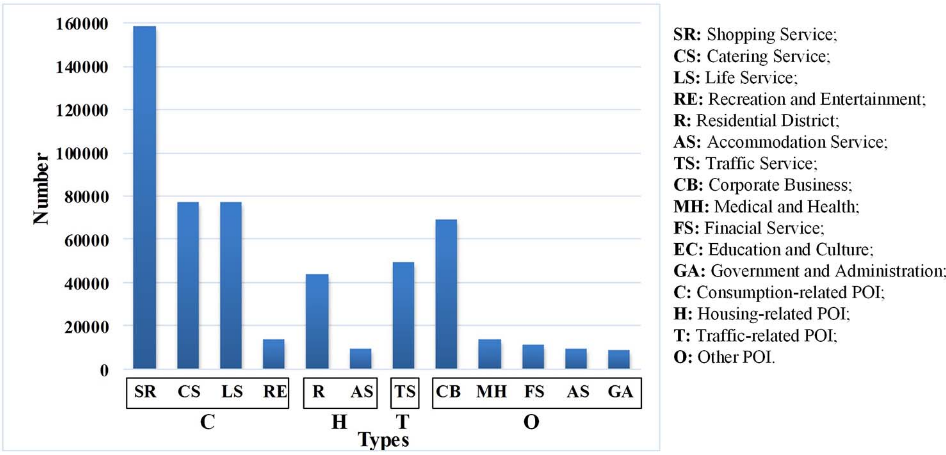

Fig. 1. Total number of POIs in each category.

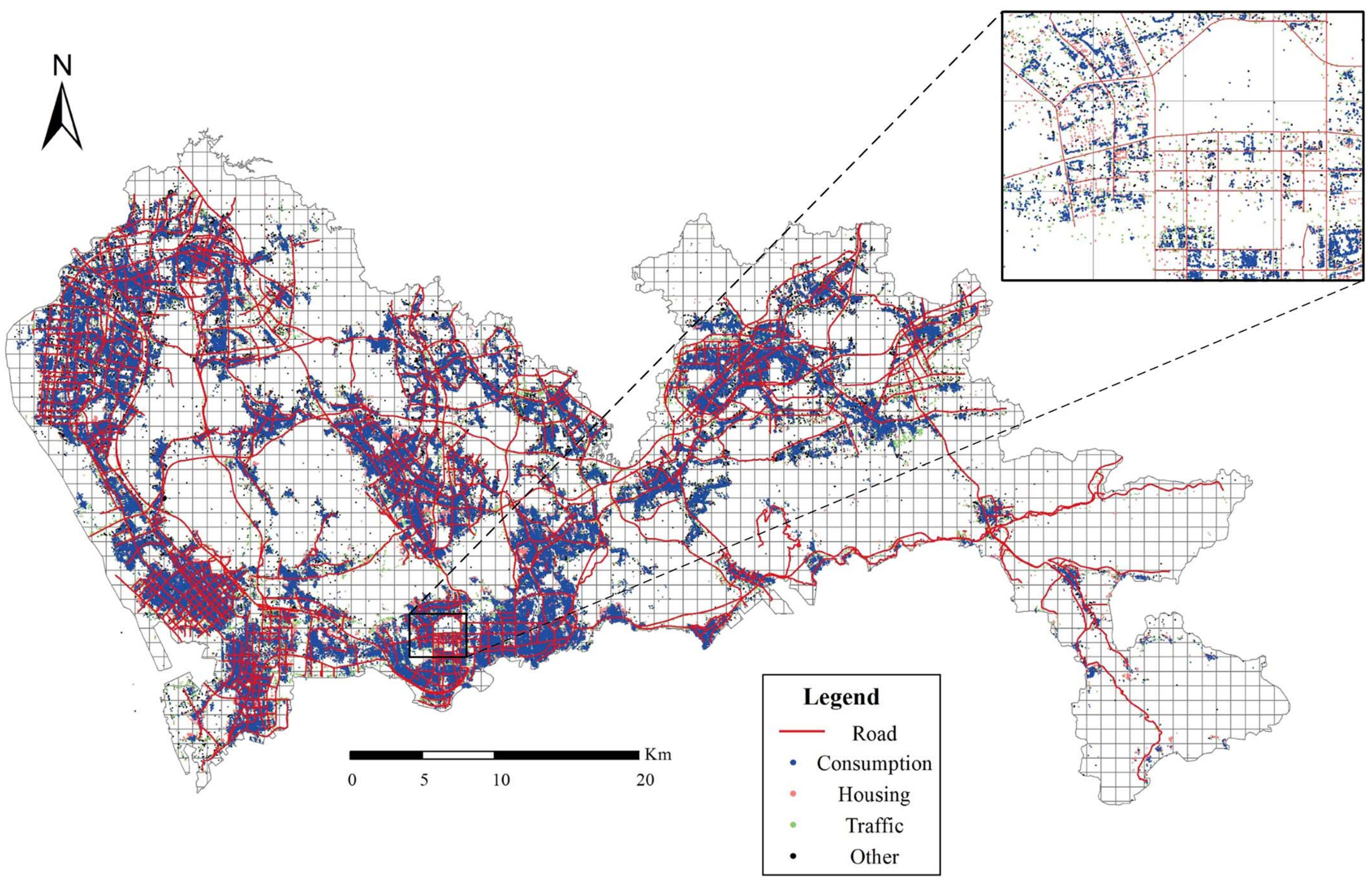

Fig. 2. Spatial grids and POIs in the study area.

hour period were treated as one check-in record. Finally, 1,159,748 check-in records are kept, with 133,825 users and 35,670 spots. Candia et al. (2008) showed that urban activities are well differentiated at particular times of day, days of the week, and between weekdays and weekend. Therefore, urban vibrancy is evaluated on weekdays and weekends. Specifically, 357,989 check-ins occurred on weekends, while 801,759 check-ins occurred on weekdays.

3.3. Descriptive statistics of the key variables

Based on previous studies of urban vibrancy and the context of Shenzhen, the independent variables in this article were divided into two types: POI-related variables and other variables. POI-related variables consist of HPOI, CPOI, TPOI, OPOI and the Shannon entropy (SE). SE is used to measure the mixing degree of the POIs. Other spatial variables are the distance to the city business district (DCBD) of Shenzhen (in this paper, the Civic Centre is defined as the CBD of Shenzhen), the distance to the district centre (DDC) and the road density (RD). Table 1 lists the detailed definitions and summary statistics of the variables.

4. Analytical framework

Focusing on the measurement of urban vibrancy and its relation to land-use configurations, this article proposes a framework to study urban vibrancy from the perspective of social sensing. Fig. 3 illustrates

Table 1
Definitions and summary statistics of the variables.

Variables	Max	Mean	SD
Average number of check-in records on a weekday	199.60	2.96	10.33
Average number of check-in records in a weekend	251.00	3.45	12.10
Number of CPOIs	2395	149.62	291.79
Number of HPOIs	440	24.21	50.78
Number of TPOIs	216	22.49	36.66
Number of OPOIs	842	50.03	89.69
Mixed land use	2.27	1.09	0.87
Distance to the CBD	58.51	26.03	13.04
Distance to the district centre	17.54	6.24	3.38
Road density	2.31	.18	.21

the general framework of urban vibrancy, which focuses on exploring the spatiotemporal relationships between vibrancy and land-use configurations, measured at grid units. First, the spatial and temporal distribution patterns of the check-in behaviour are investigated based on KDE, which illustrates the reasonability and validity of applying check-in data to represent vibrancy. Second, the spatio-temporal relationships are explored in-depth using a GTWR model considering spatial-temporal nonstationarity. The number of check-in records representing urban vibrancy is the dependent variable. The POI-related variables and other special variables are independent variables. Finally, based on the varying estimated coefficients of GTWR, the HHI is applied to portray spatio-temporal heterogeneity. A geographical information system (GIS) is applied to support data collection and processing. The details of the methodologies are described in the following sections.

4.1. Kernel density estimation

KDE is conducted to investigate the local distribution patterns of check-in behaviour from a spatial and temporal perspective, which is a method for analysing hot spots, estimating intensity and visualizing the distribution of points by creating a smooth surface based on a bivariate probability density function (Bailey & Gatrell, 1995; Silverman, 1986). KDE has been used to detect check-in hot spots (Hasan, Zhan, & Ukkusuri, 2013; Gerber, 2014; Wu, Zhi et al., 2014; Sun et al., 2016; Wu et al., 2016; Zhen et al., 2017). In this study, KDE is used to transform discrete check-in points with check-in numbers (i.e., POIs) into continuous surfaces that reflect their spatial density. According to Schabenberger and Gotway (2004), the KDE can be described by Eq. (1):

$$f(s) = \sum_{i=1}^{n} \frac{1}{h^2} k\left(\frac{d_{is}}{h}\right) \tag{1}$$

where $f(s)$ is the KDE function at location s, h denotes the bandwidth, d_{is} represents the distance from point i to s, and k is a space weight function. Previous studies have indicated that the choice of k barely affects the results and that the threshold of the distance decay is important. A grid size of 1 km is used to ensure consistency with the size of the study unit.

4.2. Geographically and temporally weighted regression

In this paper, GTWR is used to evaluate the spatio-temporal relationships between urban vibrancy and POI-based and other spatial

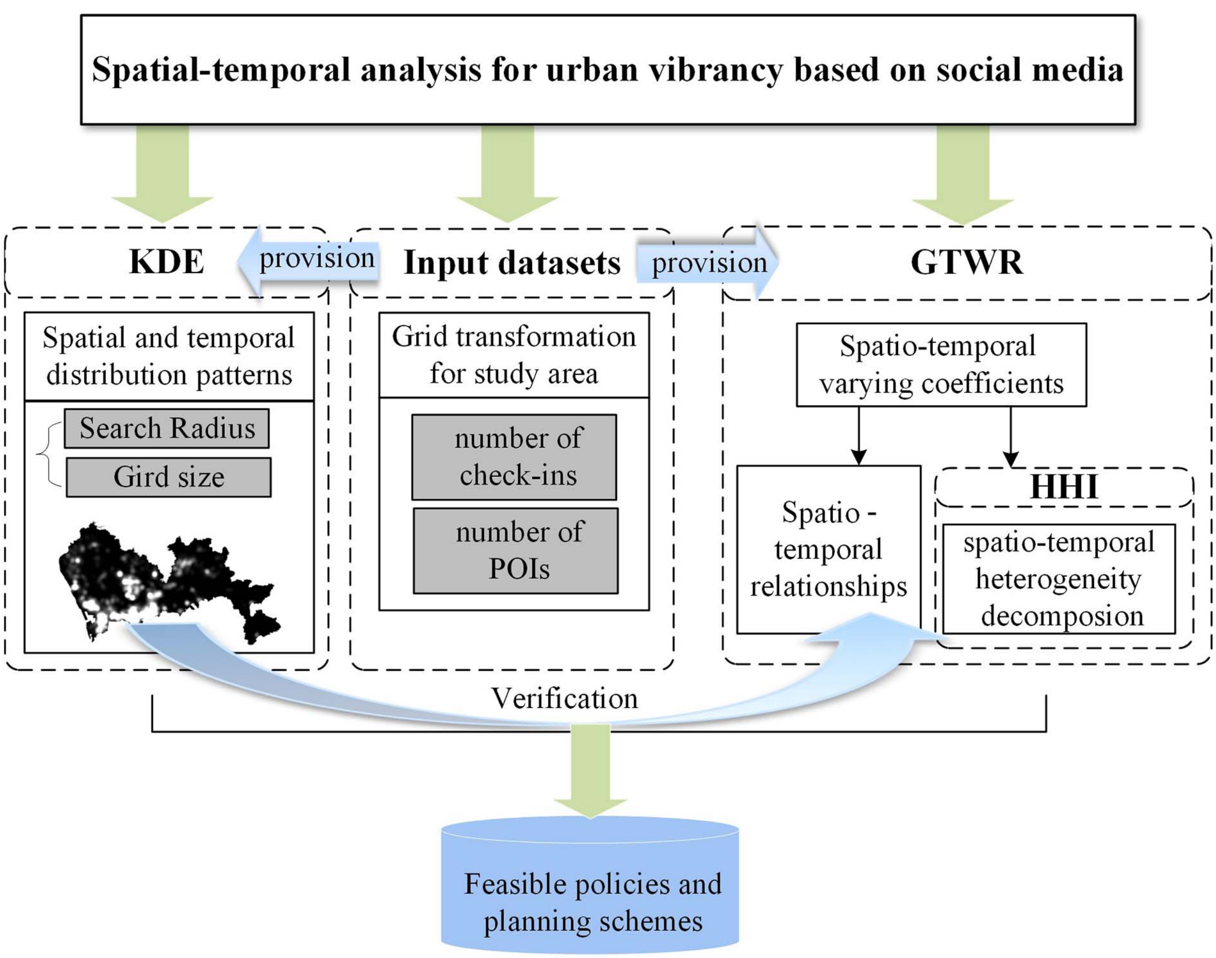

Fig. 3. Research framework of the spatio-temporal analysis for urban vibrancy.

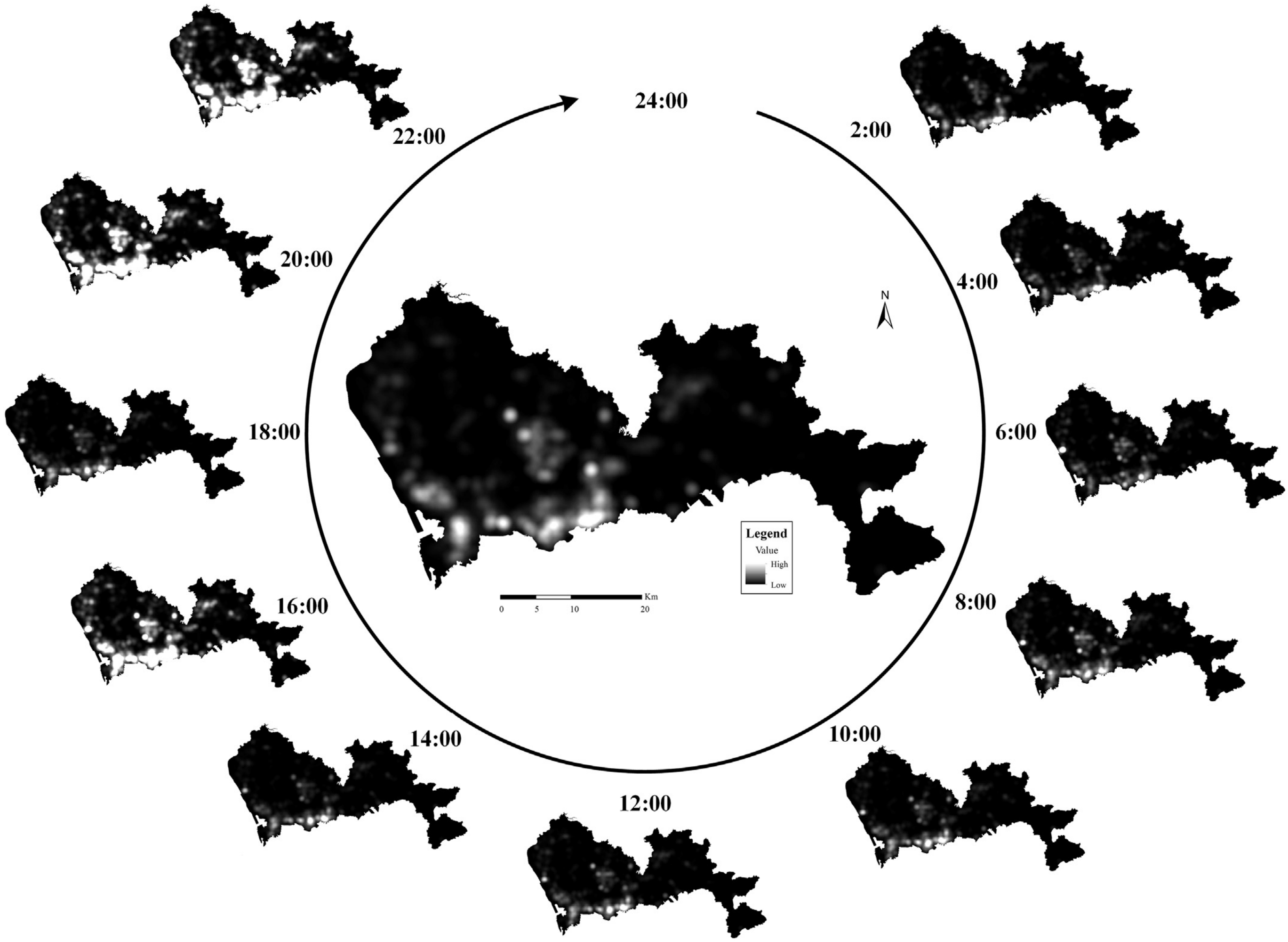

Fig. 4. Kernel estimation results of the check-in densities at different times.

variables (Huang, Wu, & Barry, 2010). As a temporal extension of geographically weighted regression (Brunsdon, Fotheringham, & Charlton, 1996), GTWR embeds time data into regression parameters to assess the local relationships between independent and dependent variables. The GTWR model can be defined as follows:

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_k \beta_k(u_i, v_i, t_i) X_{ik} + \varepsilon_i \quad i = 1, \ldots\ldots, n \tag{2}$$

where the dependent variable y_i refers to the vibrancy, using the number of check-in records as a proxy. X_i are the independent variables, including the POI-related variables and other spatial variables in Table 1. (u_i, v_i, t_i) are the space-time coordinates of grid i in the space-time dimensions (ST); u_i, v_i and t_i are the longitude, latitude and time, respectively; X_{ik} is the kth variable for grid i; $\beta_0(u_i, v_i, t_i)$ denotes the intercept value; and $\beta_k(u_i, v_i, t_i)$ represents a set of parameter values at grid i. The greatest strength of GTWR compared with other regression models is that $\beta(u_i, v_i, t_i)$ varies in the ST and GTWR can simultaneously capture spatial-temporal heterogeneity. Similarly to geographically weighted regression, the regression coefficients of GTWR are estimated based on local weighted least squares. The estimated parameters can be expressed by Eq. (3):

$$\widehat{\beta}(u_i, v_i, t_i) = [X^{\mathrm{T}} W(u_i, v_i, t_i) X]^{-1} X^{\mathrm{T}} W(u_i, v_i, t_i) Y \tag{3}$$

where the weighting matrix $W(u_i, v_i, t_i)$ is an $n \times n$ diagonal matrix and $W(u_i, v_i, t_i) = diag(W_{i1}, W_{i2} \ldots W_{ij} \ldots W_{in})$. $W_{ij}(1 \leq j \leq n)$ is the spatio-temporal distance decay function, which is determined by the spatio-temporal distance (d_{st}) and bandwidth h. In this study, a Gauss kernel function is employed to calculate the spatio-temporal weighting matrix with the greatest efficiency:

$$W_{ij} = exp[-(d_{ij}^{ST})^2/h^2] \tag{4}$$

According to Huang, Wu, and Barry (2010), the spatial-temporal distance is calculated as follows:

$$d^{ST} = \sqrt{\lambda[(u_i - u_j)^2 - (v_i - v_j)^2] + \mu(t_i - t_j)^2} \tag{5}$$

Here, h is a nonnegative parameter that produces a decay of influence with spatio-temporal distance $d_{ij}{}^{ST}$ between locations i and j. $W(u_i, v_i, t_i)$ depends on the bandwidth h, and the optimal bandwidth is chosen based on the minimum cross-validation (CV) value (Hurvich, Simonoff, & Tsai, 1998). The CV value is the sum of the squared error between the actual value y_i and predicted value $\widehat{y}_i(h)$:

$$\mathrm{CV}(h) = \sum_i (y_i - \widehat{y}_i(h))^2 \tag{6}$$

In this article, GTWR is applied to the data to assess the spatio-temporal relationships between vibrancy and POI-based configurations. The dependent variables are the average numbers of check-in records within 24 h in the 1 km × 1 km grid during weekdays and weekends. The independent variables are listed in Table 1. We apply logarithmic transformation of the dependent variables, CPOI, HPOI, TPOI and SE, to weaken the collinearity, eliminate heteroscedasticity and reduce the absolute values of the data. The estimated coefficients can be explained as "price elasticity", which is a concept from economics that is widely used in real estate studies (Liu & Wang, 2016; Wen & Tao, 2015; Yilmaz et al., 2008). A positive regression coefficient means that an increase of 1% in the variables induces a corresponding increase in the coefficient and vice versa.

4.3. Herfindahl-Hirschman index

It is difficult to visualize the estimated coefficients of GTWR in time and space. The local parameter estimates that denote local relationships are mapped using the average values. However, the positive and negative values may offset each other if only averages are used to explore spatio-temporal movement, this approach can create biases and fail to elucidate important potential mechanisms and disciplines. In industrial economics, HHI is an indicator that is widely applied to represent the sum of the squares of the market shares of different companies. Additionally, HHI can be used to measure diversity (Palan, 2010). From this perspective, 1-HHI can portray and decompose heterogeneity along the spatial and temporal dimensions. Therefore, this article applies the 1-HHI to decompose the spatio-temporal heterogeneity (Herfindahl, 1950). H_k is used to represent 1-HHI and can be calculated as follows:

$$H_k = 1 - p^2 - q^2 \tag{7}$$

where p and q are the proportions of positive and negative values, respectively. Palan (2010) noted that a large HHI value indicates greater heterogeneity and vice versa.

5. Results

5.1. Density variation of the vibrancy

KDE with a Gaussian kernel is used to estimate the check-in density distributions to determine the density of check-ins for each cell in 2-h intervals. The KDE results (Fig. 4) reveal that the patterns of people's activities are as follows: (1) In terms of the spatial distribution, main agglomeration areas with high frequencies and densities are evident in the central city at Futian, Luohu and Nanshan; sub-centres of Shenzhen at Baoan and Longgang; the Bao'an International Airport; and Shenzhen North Railway Station (a high-speed rail station). (2) From a temporal perspective, people's activity frequencies are relatively low during typical sleeping times and working times and relatively high during leisure times and at dinner time. Overall, distinct patterns in urban activities and people aggregation are associated with the natures of different types of typical activities. The information from KDE can facilitate studying the dynamic evolution of activities across both space and time. Additionally, the KDE results verify that vibrancy varies at fine temporal (i.e., a day) and spatial (i.e., a city) scales. The results also show that the check-in data can reflect more subtle phenomena and results than traditional data with fine time and spatial granularity.

5.2. Model specifications

Stepwise multiple linear regression is conducted to calculate the variance inflation factors (VIFs) to avoid multicollinearity and verify spatial correlations. We retain factors with VIF values below 10 and those that exert significant effects on the independent variables (Hair et al., 1995; Mason, Gunst, & Hess, 2003). Consequently, OPOI is removed from our model because it has a VIF value of 11.7 and significant correlations with other variables. After removing the OPOI, the remaining variables show no strong multicollinearity. In addition, the spatial autocorrelation in the dependent variable is a pre-condition for GTWR. This article utilizes Moran's I to evaluate the spatial autocorrelation (Cliff & Ord, 1981; Moran, 1950). The Moran's I values of vibrancy for weekdays and weekends are 0.778 and 0.776, respectively. The results indicate that people's activity intensity and check-in variation have positive spatial autocorrelation and distinct features of spatial clustering.

The GTWR results for weekdays and weekends are presented in Table 2 and Table 3, respectively. All selected variables are significant at the 1% level. The R^2 values are 0.869 and 0.847, denoting that the selected variables can explain 86.9% and 84.7% of the variation in the vibrancy on weekdays and weekends, with optimal spatio-temporal bandwidths of 0.100 and 0.110 for weekdays and weekends, respectively. The variation trends for all the coefficients for weekdays and weekends are roughly the same, although local differences can be observed. The details and corresponding mechanisms and analysis are introduced in the following section.

Table 2
Estimated GTWR parameters for weekdays.

	Min	Lower quartile	Median	Upper quartile	Max	SD	*p*-Value
Constant	−32.733	−0.569	−0.256	0.128	11.934	1.965	.000***
CPOI	−2.206	0.047	0.287	0.640	3.241	0.537	.000***
HPOI	−1.617	−0.021	0.127	0.370	4.049	0.437	.000***
TPOI	−2.939	−0.096	0.052	0.252	3.381	0.426	.000***
SE	−3.623	−0.270	−0.025	0.190	4.479	0.502	.000***
DCBD	−30.911	−0.596	−0.040	0.331	13.502	1.830	.000***
DDC	−4.482	−0.316	−0.038	0.068	7.145	0.592	.000***
RD	−1.033	−0.002	0.105	0.264	3.694	0.319	.000***

Diagnostic information	
Moran's I	0.7786
R^2	0.869
Residual sum of squares	6886.600
AIC	−107,681.900
Bandwidth	0.100

*** represents significance level of 1%.

Table 3
Estimated GTWR parameters for weekends.

	Min	Lower quartile	Median	Upper quartile	Max	SD	*p*-Value
Constant	−34.236	−0.590	−0.250	0.162	12.961	2.035	.000***
CPOI	−3.281	0.009	0.256	0.663	5.423	0.621	.000***
HPOI	−1.813	−0.068	0.097	0.343	3.628	0.457	.000***
TPOI	−5.888	−0.187	−0.008	0.153	2.322	0.428	.000***
SE	−2.309	−0.389	−0.176	−0.031	1.368	0.332	.000***
DCBD	−33.078	−0.629	−0.046	0.363	12.858	1.915	.000***
DDC	−3.966	−0.304	−0.036	0.095	8.317	0.624	.000***
RD	−1.618	0.004	0.132	0.310	5.803	0.439	.000***

Diagnostic information	
Moran's I	0.776
R^2	0.847
Residual sum of squares	8077.27
AIC	−99,253.884
Bandwidth	0.110

*** represents significance level of 1%.

5.3. Analysis of the spatio-temporal variation

The coefficient $\beta_k(u_i, v_i, t_i)$ varies in ST, which results in difficulties in visualization. In this paper, we choose a compromise to represent the spatial and temporal variation. Fig. 5 shows the temporal nonstationary nature of the coefficients. The solid line represents weekdays, and the dashed line represents weekends. The variation trends and degrees of all the variables are similar, except for the SE and TPOI. The tendencies of the SE on the weekdays and weekends are similar. The peak values appear at 11:00, 13:00 and night-time when people are engaged in eating dinner and entertainment activities. Greater potential exists to improve the capacity of SE for vibrancy on the weekends than on weekdays because the dashed line is located over the solid line. This result is consistent with people's lifestyles. On weekdays, most people,

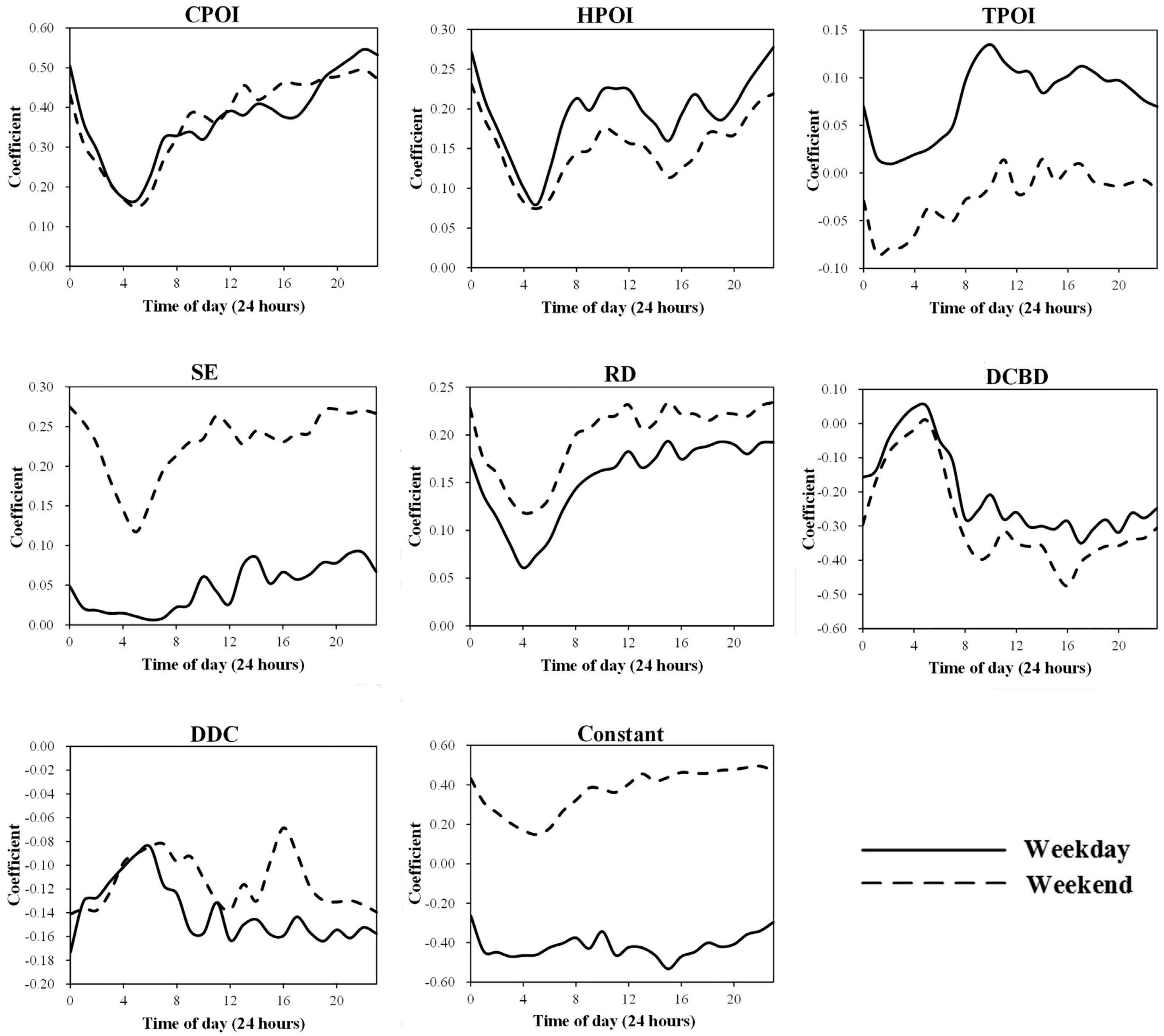

Fig. 5. Temporal trends of the variables in Shenzhen.

particularly office workers, are busy at work, and these people typically commute between their workplaces and homes and do little other than work. In contrast, these people want to relieve the fatigue of working on the weekends, and most of their activities are related to leisure, entertainment and dinner parties; thus, these people tend to frequent places where they can engage in these activities. Therefore, the average coefficients of SE on the weekends are all higher than those on weekdays.

For TPOI, the effects on vibrancy for weekdays and weekends are completely reversed. TPOI has positive effects on vibrancy on weekdays and negative effects on the weekends. Since the housing reform, the spatial patterns and distributions of living and employment locations in Shenzhen have been changing profoundly. Furthermore, the problem of home-work separation has become increasingly prominent with the formation of multi-centre city spatial structure (Bin, 2009; Liu, Zhang, & Chai, 2009). Therefore, people's travel activities depend on traffic facilities and transportation accessibility during the weekdays, and the varying trends of the coefficients in a day are consistent with the time characteristics of commuting, as reported in a previous study (Chai, 2014). On weekends, people's activities focus on resting, particularly at night, or engaging in entertainment and leisure activities near their home; thus, these people tend to reject the areas that are noisy and have high vehicular traffic.

The variation trends and degrees of the remaining variables are similar. CPOI has great elasticity on vibrancy, which is low from 0:00 to 08:00 (i.e., when people are sleeping). The secondary maximum of elasticity occurs during the afternoon and achieves its maximum value after work (18:00) on weekdays. The maximum value occurs at night because most people do not work at this time and tend to be engaged in after-work recreational activities. Notably, Shenzhen is called the "Sleepless City" because of its rich nightlife. The temporal variation in the elasticity of HPOI in this study differs from that described by Li et al. (2016). These authors concluded that the effects of the elasticity of the HPOI peak at approximately 4:00, 11:00 and 16:00 and can be attributed to the sleeping, working and living patterns of Beijing's residents. Li et al. (2016) noted that HPOI has positive effects on vibrancy. However, they neglected the positive and negative effects of HPOI because elasticity depends on the absolute value rather than the

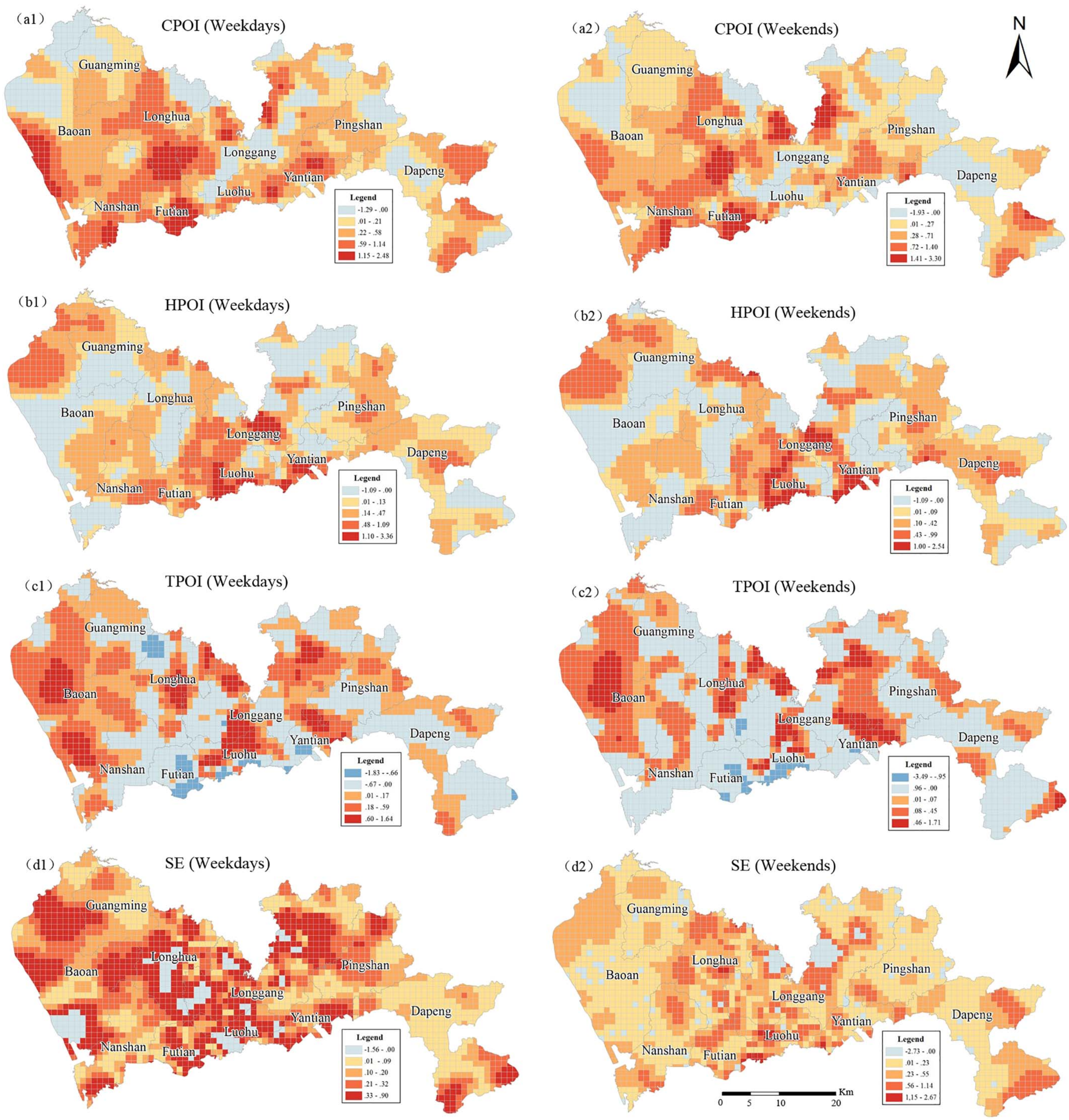

Fig. 6. Spatial distributions of the variables in Shenzhen.

sign. Given this city's proximity to Hong Kong, online stores and purchasing agents in Shenzhen are developing well; thus, working from home is increasingly popular. Hence, the effects of HPOI are relatively uniform throughout the day, except during sleeping times. For the selected variables, CPOI can produce the greatest premium for vibrancy. The temporal variation in consumption elasticity reaches its lowest value during sleeping times and is relatively high at other times. Li et al. (2016) explained that these results may be attributable to the business hours of shopping facilities and other consumption services.

Fig. 6 shows the average spatial change tendencies of the parameter estimates for POI-based variables, which are pivotal and focus factors in this article. Here, the natural breakpoint method is used to classify the estimated parameters, which are the maximum variance between groups and the minimum intra-class variance (De Smith, Goodchild, & Longley, 2007). This article also manually sets zero as a threshold to distinguish between the positive and negative effects. The general tendencies and nuances of four variables on weekdays and weekends are consistent with the overall trends of the spatial variation. The effects of CPOI on vibrancy in Shenzhen are mostly positive. Urban vibrancy commands a high premium if the density of CPOI is high. The positive effects mean that CPOI can be attractive to people. However, the effects of CPOI are the opposite in relatively remote and undeveloped areas. Most of these areas are ecological control areas in Shenzhen, which have low accessibility and low CPOI density. For

HPOI, urban areas, including Nanshan, Futian, Luohu and Yantian, all have positive effects on urban vibrancy. The other districts have mostly negative effects. Additionally, the HPOI areas with negative effects are consistent with ecological control areas, which have few settlements. The positive effects of HPOI on weekdays are greater than those on weekends, which match the differences in work-and-rest systems between weekdays and weekends. Unlike that of HPOI, the effect of TPOI is negative in central Shenzhen and positive in the sub-areas. Real estate in the urban areas of Shenzhen has well-equipped facilities, good accessibility and high housing prices that attract people. These areas are compact and have central locations with short commutes to nearby office buildings and commercial areas; therefore, people pay less for transportation. Inversely, excessive vehicular traffic may result in environmental pollution and congestion, which produce bad experiences that negatively affect vibrancy. Therefore, HPOI has positive effect and TPOI has negative effect in urban areas. In contrast, people must pay more to commute in the sub-areas of Shenzhen; TPOI thus has a positive influence in attracting people.

The influence of SE mainly reflects a difference in quantity. SE on the weekends has a greater elasticity (−2.73–2.67) than SE on the weekdays (−1.56–0.9), which is consistent with the temporal variation results based on the coefficients. The reasons are not repeated here. Moreover, the high elasticity of SE on weekdays is more disperse than on weekends. On weekdays, most people should go to workplaces, and they concern SE in nearby areas. On weekends, they go to popular places such as Dongmen Pedestrian Street. Generally, the elasticity of SE is mostly positive, particularly in the CBD and district centres. Clearly, high SE can effectively improve vibrancy in most locations in Shenzhen; thus, SE distinctly improves urban vibrancy, which is the same conclusion reached by previous studies. High SE refers to a combination of residential, commercial, cultural, institutional, or industrial uses, which can provide convenience and offer more attractions to people. Furthermore, the effects of SE in centres of city and districts are greater than in the suburbs of Shenzhen. Because the centres of Shenzhen or districts with a high degree of mixed land use may be more attractive to residents, in turn, the preference of residents can be improved by SE in relatively prosperous places. The most important findings may provide implications for developing compact urban structure and reasonable planning to improve mixed land use degree.

5.4. Evaluation of spatio-temporal heterogeneity

The aforementioned findings confirm that POI-based indices have significant spatio-temporally heterogeneous effects on urban vibrancy, as indicated by intensified human activities. Fig. 5 and Fig. 6 present visual representations of this heterogeneity in space and time. This article applies the HHI to decompose the degree of heterogeneity, and the results are described in Fig. 7 and Fig. 8. The periods and locations with high heterogeneity are the focus of this study. From the perspective of decomposed spatial heterogeneity over time, SE and TPOI on weekdays exhibit significant spatial heterogeneity over time. TPOI and HPOI on the weekends have the similar results. Their degrees remain fairly stable in each 24-h period, with relatively high HHI values. HPOI, CPOI on weekdays, and CPOs and SE on weekends have relatively low spatial heterogeneity and exhibit slight fluctuations within 24-h periods.

From the perspective of temporal heterogeneity over space, the HHIs of all the variables exhibit similar tendencies, except for SE on weekdays and weekends. CPOI and SE show significant temporal heterogeneity in the district centres on weekdays and weekends, while HPOI and TPOI show significant temporal heterogeneity across Shenzhen. The temporal heterogeneity of the different POI-based configuration variables is consistent with the temporal regularity and volatility of people's travel, work, leisure and entertainment. According to the relatively singular functions of HPOI and TPOI, commuting and rest are more attractive than during other periods of the day. Shenzhen is famous for its active nightlife and exhibits substantial variation in its local development levels. Therefore, the temporal heterogeneity of CPOI and SE is significant in most district centres. Indeed, the temporal heterogeneity of CPOI and SE is relatively steady in two types of locations: those with high effects during the day and night (relatively rapid development) and those with low effects during the day and night (relatively slow development).

6. Conclusions

In an era rich with crowdsourcing data, this study applies social media check-in data to reflect people's mobility and aggregation. These data can describe urban vibrancy, dynamics and evolution over space and time. Therefore, this article proposed a framework that integrates KDE, GTWR and the HHI to analyse urban vibrancy based on check-in data. Our work introduced social media check-in data as a proxy for characterizing urban vibrancy from a spatio-temporal continuum perspective. Then, we investigated its correlated structural variables and analysed the mechanisms that underlie its variations over space and time. The proposed framework to evaluate urban vibrancy aims for high spatio-temporal resolution in response to the increasing interest in introducing new quantitative thinking into urban structure studies (Ye, Li, & Liu, 2017).

The profiling results illustrate the trends for urban vibrancy and its influential factors. The recognition of spatial-temporal variations has important implications in public policies that target specific temporal and geographical units and can more effectively improve vibrancy and the planning of urban morphology, function and structure. The findings

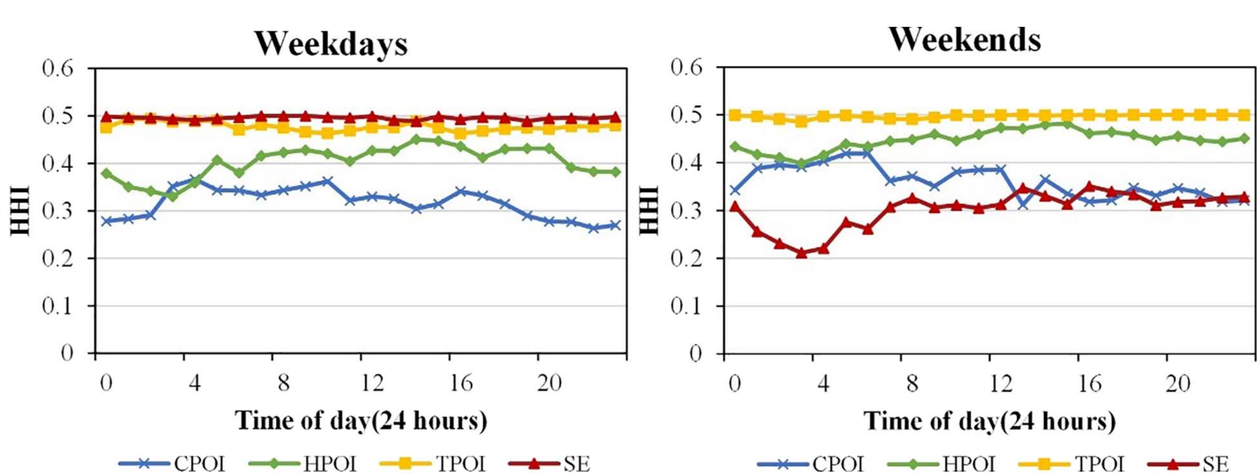

Fig. 7. Spatial heterogeneity in the temporal dimension.

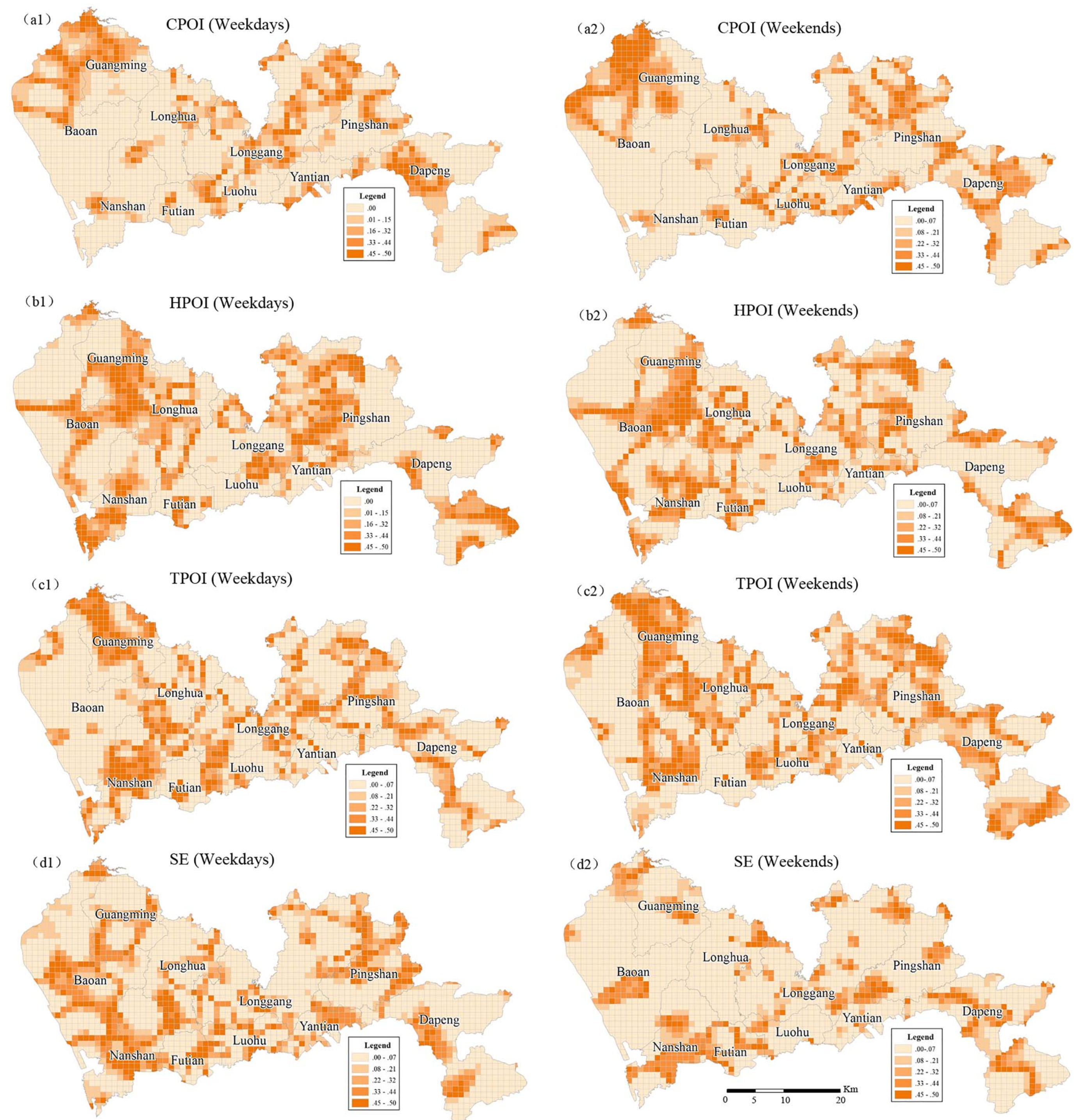

Fig. 8. Temporal heterogeneity in the spatial dimension.

of this empirical study can be utilized by urban managers and planners to explore the vibrancy and aggregation of people with respect to the ambient environment, to guide infrastructure and transportation planning, and to inform policy directions. For other researchers, this theoretical framework can explain the factors that contribute to vitality at the local level. For residents, the results can be used to schedule a trip and select the location of their home.

To the best of our knowledge, no other studies have addressed the varying spatio-temporal relationships between vibrancy using social media data. The results from this paper shed critical light on POI-based variables, which are significantly associated with variations in vibrancy over a day. This could improve our understanding of urban structure and land-use analysis. In addition, the HHI can be used to decompose spatial heterogeneity in time and temporal heterogeneity in space. Locations and times with greater HHI values are more distinctly heterogeneous and thus are key research areas for scholars and planners. The mechanisms that are exhibited by areas with greater HHI values will be examined in our future studies.

In this study, planning research can be conducted on a fine scale from the perspective of social sensing. The results show that this issue is also important from a policy perspective. We expect that these results will be adopted to guide land-use and facility planning, select appropriate sites for infrastructure services by considering local demand, and optimize the traffic environment. These results can be adopted to guide reasonable urban planning to improve the mixing degree and compactness of land use in Shenzhen. In addition, these findings can be applied to analyse land/housing values and to assess the living environment. Based on the relationships between vibrancy and related variables, this study provides relevant information for planners to optimize the urban function layout and neighbourhood design and pursue efficient approaches to promote city vitality. Finally, utilizing the predictability of a GTWR model, the relationships between vibrancy and land use can be used to predict future patterns of urban dynamics and estimate the spatio-temporal distribution of the population.

Studies on urban functions, structure and morphology should incorporate check-in data to represent and access urban vibrancy under the proposed framework. Social media data is a promising source for determining individual activities and location preferences because of their low cost, large sample size and instantaneity compared with those of static datasets. Although our work is one of the first comprehensive studies to understand the dynamics of human activities, interactions, aggregation and urban vibrancy, this framework has several limitations, which also highlight potential directions for future research. On the one hand, check-in data are not representative because young people are more likely to contribute check-in data. Therefore, check-in data do not reflect vibrancy based on all age groups. On the other hand, the GTWR model considers only the number of POIs without considering the supply and demand of POIs. In future research, we will apply at least two types of big data to verify our results and to reduce deviations based on the detailed classification of POIs by considering service levels (supply) and service populations (demand).

Acknowledgments

This study was supported by the National Natural Science Foundation of China (Project No. 41571438), the Open Fund of Key Laboratory of Urban Land Resources Monitoring and Simulation Ministry of Land and Resources (No. KF-2016-02-028) and the Hubei Key Laboratory of Regional Development and Environmental Response (Hubei University, 2015-B-002).

References

Bailey, T. C., & Gatrell, A. C. (1995). *Interactive spatial data analysis. Vol. 413*. Essex: Longman Scientific & Technical.

Batty, M. (2013). Big data, smart cities and city planning. *Dialogues in Human Geography, 3*(3), 274–279. http://dx.doi.org/10.1177/2043820613513390.

Bin, M. (2009). The spatial organization of the separation between jobs and residential locations in Beijing. *Acta Geographica Sinica, 64*(12), 1457–1466. http://dx.doi.org/10.11821/xb200912006.

Braun, L. M., & Malizia, E. (2015). Downtown vibrancy influences public health and safety outcomes in urban counties. *Journal of Transport & Health, 2*(4), 540–548. http://dx.doi.org/10.1016/j.jth.2015.08.005.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis, 28*(4), 281–298. http://dx.doi.org/10.1111/j.1538-4632.1996.tb00936.x.

Candia, J., González, M. C., Wang, P., Schoenharl, T., Madey, G., & Barabási, A. L. (2008). Uncovering individual and collective human dynamics from mobile phone records. *Journal of Physics A: Mathematical and Theoretical, 41*(22), 224015. http://dx.doi.org/10.1088/1751-8113/41/22/224015.

Chai, Y. (2014). *Spatial behavior and behavior space*.

Cheng, G., Zeng, X., Duan, L., Lu, X., Sun, H., Jiang, T., & Li, Y. (2016). Spatial difference analysis for accessibility to high level hospitals based on travel time in Shenzhen, China. *Habitat International, 53*, 485–494.

Chhetri, P., Stimson, R. J., & Western, J. (2006). Modelling the factors of neighbourhood attractiveness reflected in residential location decision choices. *Regional Studies, 36*(2), 393–417. http://dx.doi.org/10.2457/srs.36.393.

Cliff, A. D., & Ord, J. K. (1981). *Spatial processes: Models & applications. Vol. 44*. London: Pion.

Coupland, A. (1997). *An introduction to mixed use development. Reclaiming the city; mixed use development.* London: Spon.

Dai, X., Bai, X., & Xu, M. (2016). The influence of Beijing rail transfer stations on surrounding housing prices. *Habitat International, 55*, 79–88. http://dx.doi.org/10.1016/j.habitatint.2016.02.008.

De Smith, M. J., Goodchild, M. F., & Longley, P. (2007). *Geospatial analysis: A comprehensive guide to principles, techniques and software tools.* Troubador Publishing Ltd.

Duggan, M., & Brenner, J. (2013). *The demographics of social media users, 2012. Vol. 14.* Washington, DC: Pew Research Center's Internet & American Life Project.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2003). *Geographically weighted regression.* John Wiley & Sons, Limited.

Geddes, P. (1915). *City in evolution.* London: Williams and Norgate Ltd.

Gerber, M. S. (2014). Predicting crime using twitter and kernel density estimation. *Decision Support Systems, 61*, 115–125. http://dx.doi.org/10.1016/j.dss.2014.02.003.

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.

Griffith, D. A. (2004). A spatial filtering specification for the autologistic model. *Environment and Planning A, 36*(10), 1791–1811.

Hair, J. F., Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis* (3rd ed.). New York: Macmillan.

Harvey, L. (2001). Defining and measuring employability. *Quality in Higher Education, 7*(2), 97–109. http://dx.doi.org/10.1080/13538320120059990.

Hasan, S., Zhan, X., & Ukkusuri, S. V. (2013). Understanding urban human activity and mobility patterns using large-scale location-based data from online social media. *Proceedings of the 2nd ACM SIGKDD international workshop on urban computing* (pp. 6). ACM.

Herfindahl, O. C. (1950). *Concentration in the steel industry.* Columbia University (Doctoral dissertation).

Howard, E. (1898). *Tomorrow: A peaceful path to reform.* London: Swan Sonnenschein.

Huang, B., Wu, B., & Barry, M. (2010). Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International Journal of Geographical Information Science, 24*(3), 383–401. http://dx.doi.org/10.1080/13658810802672469.

Hurvich, C. M., Simonoff, J. S., & Tsai, C. L. (1998). Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society, Series B: Statistical Methodology, 60*(2), 271–293.

Jacobs, J. (1961). *The death and life of great American cities.* (Vintage).

Jacobs, J. (1969). *The life of cities.* Random House.

Jacobs-Crisioni, C., Rietveld, P., Koomen, E., & Tranos, E. (2014). Evaluating the impact of land-use density and mix on spatiotemporal urban activity patterns: An exploratory study using mobile phone data. *Environment and Planning A, 46*(11), 2769–2785. http://dx.doi.org/10.1068/a130309p.

Jang, M., & Kang, C. D. (2015). Retail accessibility and proximity effects on housing prices in Seoul, Korea: A retail type and housing submarket approach. *Habitat International, 49*, 516–528. http://dx.doi.org/10.1016/j.habitatint.2015.07.004.

Jendryke, M., Balz, T., McClure, S. C., & Liao, M. (2017). Putting people in the picture: Combining big location-based social media data and remote sensing imagery for enhanced contextual urban information in Shanghai. *Computers, Environment and Urban Systems, 62*, 99–112. http://dx.doi.org/10.1016/j.compenvurbsys.2016.10.004.

Jin, X., Long, Y., Sun, W., Lu, Y., Yang, X., & Tang, J. (2017). Evaluating cities' vitality and identifying ghost cities in China with emerging geographical data. *Cities, 63*, 98–109. http://dx.doi.org/10.1016/j.cities.2017.01.002.

Jurdak, R., Zhao, K., Liu, J., AbouJaoude, M., Cameron, M., & Newth, D. (2015). Understanding human mobility from twitter. *PLoS One, 10*(7), e0131469. http://dx.doi.org/10.1371/journal.pone.0131469.

Li, C., Wang, M., Wang, J., & Wu, W. (2016). *The geography of city liveliness and land use configurations: Evidence from location-based big data in Beijing (No. 0201).* Spatial Economics Research Centre, LSE.

Li, L., Goodchild, M. F., & Xu, B. (2013). Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and Geographic Information Science, 40*(2), 61–77. http://dx.doi.org/10.1080/15230406.2013.777139.

Lin, J., & Cromley, R. G. (2015). Evaluating geo-located Twitter data as a control layer for areal interpolation of population. *Applied Geography, 58*, 41–47. http://dx.doi.org/10.1016/j.apgeog.2015.01.006.

Liu, X., Kang, C., Gong, L., & Liu, Y. (2016). Incorporating spatial interaction patterns in classifying and understanding urban land use. *International Journal of Geographical Information Science, 30*(2), 334–350. http://dx.doi.org/10.1080/13658816.2015.1086923.

Liu, X., & Wang, M. (2016). How polycentric is urban China and why? A case study of 318 cities. *Landscape and Urban Planning, 151*, 10–20. http://dx.doi.org/10.1016/j.landurbplan.2016.03.007.

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*(3), 512–530. http://dx.doi.org/10.1080/00045608.2015.1018773.

Liu, Z., Zhang, Y., & Chai, Y. (2009). Home-work separation in the context of institutional and spatial transformation in urban China: Evidence from Beijing household survey data. *Urban Studies, 16*(9), 110–117.

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2015). *Geographic information science and systems.* John Wiley & Sons.

Lynch, K. (1984). *Good city form.* MIT press.

March, A., Rijal, Y., Wilkinson, S., & Firidin Özgür, E. (2012). Measuring building adaptability and street vitality. *Planning Practice and Research, 27*(5), 531–552. http://dx.doi.org/10.1080/02697459.2012.715813.

Mason, R. L., Gunst, R. F., & Hess, J. L. (2003). *Statistical design and analysis of experiments: With applications to engineering and science. Vol. 474.* John Wiley & Sons.

Mehta, V. (2007). Lively streets determining environmental characteristics to support social behavior. *Journal of Planning Education and Research, 27*(2), 165–187. http://dx.doi.org/10.1177/0739456X07307947.

Mellander, C., Lobo, J., Stolarick, K., & Matheson, Z. (2015). Night-time light data: A good proxy measure for economic activity? *PLoS One, 10*(10), e0139779. http://dx.doi.org/10.1371/journal.pone.0139779.

Montgomery, J. (1995). *Editorial urban vitality and the culture of cities*.

Montgomery, J. (1998). Making a city: Urbanity, vitality and urban design. *Journal of Urban Design, 3*(1), 93–116.

Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika, 37*(1/2), 17–23. http://dx.doi.org/10.2307/2332142.

Nicodemus, A. G. (2013). Fuzzy vibrancy: Creative placemaking as ascendant US cultural policy. *Cultural Trends, 22*(3–4), 213–222. http://dx.doi.org/10.1080/09548963.2013.817653.

Palan, N. (2010). *Measurement of specialization–the choice of indices (No. 062)*. FIW.

Schabenberger, O., & Gotway, C. A. (2004). *Statistical methods for spatial data analysis.* CRC press.

Schafer, A. (2000). Regularities in travel demand: An international perspective. *Journal of Transportation and Statistics, 3*(3), 1–31.

Sharkova, I. V., & Sanchez, T. W. (1999). *An analysis of neighborhood vitality: The role of local civic organizations*.

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*(9), 1687–1693.

Shen, Y., & Karimi, K. (2016). Urban function connectivity: Characterisation of functional urban streets with social media check-in data. *Cities, 55*, 9–21. http://dx.doi.org/10.1016/j.cities.2016.03.013.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis. Vol. 26*. CRC press.

Standardization Administration of the People's Republic of China (2011). *Code for classification of urban land use and planning standards of development land (GB 50137–2011)*.

Stern, M. J., & Seifert, S. C. (2010). Cultural clusters: The implications of cultural assets agglomeration for neighborhood revitalization. *Journal of Planning Education and Research, 29*(3), 262–279.

Sun, Y., Fan, H., Li, M., & Zipf, A. (2016). Identifying the city center using human travel flows generated from location-based social networking data. *Environment and Planning. B, Planning & Design, 43*(3), 480–498. http://dx.doi.org/10.1177/0265813515617642.

Wen, H., & Tao, Y. (2015). Polycentric urban structure and housing price in the transitional China: Evidence from Hangzhou. *Habitat International, 46*, 138–146. http://dx.doi.org/10.1016/j.habitatint.2014.11.006.

Winters, M., Brauer, M., Setton, E. M., & Teschke, K. (2010). Built environment influences on healthy transportation choices: Bicycling versus driving. *Journal of Urban Health, 87*(6), 969–993.

Wu, C., Ye, X., Du, Q., & Luo, P. (2017). Spatial effects of accessibility to parks on housing prices in Shenzhen, China. *Habitat International, 63*, 45–54.

Wu, C., Ye, X., Ren, F., Wan, Y., Ning, P., & Du, Q. (2016). Spatial and social media data analytics of housing prices in Shenzhen, China. *PLoS One, 11*(10), e0164553. http://dx.doi.org/10.1371/journal.pone.0164553.

Wu, J., Wang, M., Li, W., Peng, J., & Huang, L. (2014). Impact of urban green space on residential housing prices: Case study in Shenzhen. *Journal of Urban Planning and Development, 141*(4), 05014023.

Wu, L., Zhi, Y., Sui, Z., & Liu, Y. (2014). Intra-urban human mobility and activity transition: Evidence from social media check-in data. *PLoS One, 9*(5), e97010. http://dx.doi.org/10.1371/journal.pone.0097010.

Ye, Y., Li, D., & Liu, X. (2017). How block density and typology affect urban vitality: An exploratory analysis in Shenzhen, China. *Urban Geography,* 1–22.

Yilmaz, H., Bulut, Z., Temurlenk, M. S., & Yesil, P. (2008). Determination of the impact of parks and playgrounds on house prices in the city of Erzurum, Turkey. *International Journal of Natural and Engineering Sciences, 2*(1), 47–51.

Yue, Y., Zhuang, Y., Yeh, A. G., Xie, J. Y., Ma, C. L., & Li, Q. Q. (2016). Measurements of POI-based mixed use and their relationships with neighbourhood vibrancy. *International Journal of Geographical Information Science,* 1–18. http://dx.doi.org/10.1080/13658816.2016.1220561.

Zhen, F., Cao, Y., Qin, X., & Wang, B. (2017). Delineation of an urban agglomeration boundary based on Sina Weibo microblog 'check-in'data: A case study of the Yangtze River Delta. *Cities, 60*, 180–191. http://dx.doi.org/10.1016/j.cities.2016.08.014.

Zhi, Y., Li, H., Wang, D., Deng, M., Wang, S., Gao, J., ... Liu, Y. (2016). Latent spatio-temporal activity structures: A new approach to inferring intra-urban functional regions via social media check-in data. *Geo-spatial Information Science,* 1–12. http://dx.doi.org/10.1080/10095020.2016.1176723.