Article

Urban Analytics and City Science

EPB: Urban Analytics and City Science
2020, Vol. 47(9) 1543–1559

DOI: 10.1177/2399808319828730
journals.sagepub.com/home/epb
SAGE

Evaluating and characterizing urban vibrancy using spatial big data: Shanghai as a case study

Bo Huang and Yulun Zhou
The Chinese University of Hong Kong, China

Zhigang Li
Wuhan University, China

Yimeng Song and Jixuan Cai
The Chinese University of Hong Kong, China

Wei Tu
Shenzhen University, China

Abstract

Although people may recognize urban vibrancy when they see or sense it, developing direct and comprehensive measures of urban vibrancy remains a challenge. In the context of intense global competition, there is an increased realization that urban vibrancy is vital to the social and economic sustainability of cities. Such vibrancy may be significantly shaped by the urban built environment, yet we know little about the close connections between vibrancy and urban built environments. Empowered by newly available sources of spatial big data, which provide enormous amounts of information on both human dynamics and the built environment, this paper proposes a framework for evaluating and characterizing urban vibrancy. Thus far, vibrancy measures have mostly used single-source data that hardly reflect the multifaceted manifestations of urban vibrancy. Therefore, we propose a more comprehensive measure of urban vibrancy, extracted as the common latent factor from multiple surface attributes. Using the proposed framework, we evaluated and mapped the spatial dynamics of vibrancy in Shanghai, a typical large city in post-reform China, and investigated the associations between vibrancy and various urban built environment indicators. The evidence shows that the horizontal built-up density, rather than vertical height, is the leading generator of vibrancy in Shanghai, followed by the

Corresponding author:
Bo Huang, Department of Geography and Resources Management, The Chinese University of Hong Kong, Shatin, NT, Hong Kong, China.
Email: bohuang@cuhk.edu.hk

density and mixture of urban functions, accessibility, and walkability. In this vein, we contribute to current debates and future planning practices regarding vibrant spaces in large cities. This proposed evaluation framework, equipped with spatial big data, can benefit future urban studies.

Keywords
Urban vibrancy, urban built environment, spatial big data, factor analysis, Shanghai

Introduction

Creating and maintaining urban vibrancy is vital for sustainable urban development (Hall and Pfeiffer, 2013). In this neoliberal, entrepreneurial, and competitive urban world, vibrant cities attract more human and economic capital and thus become more productive and economically sustainable (Brenner, 2014; Brenner et al., 2010). Vibrancy also matters to the social sustainability of cities. Vibrant cities better satisfy the needs of urban residents and provide more comfortable living conditions (Couture, 2013). In particular, vibrancy promotes human activity and interaction (Jacobs, 1961), and it improves people's subjective feelings of urban spaces, which is crucial not only for the well-being of urbanites (Pinquart and Sörensen, 2000) but also for their innovation capabilities (Montgomery, 1998). Vibrant cities have more resilience and thrive in the midst of social, economic, and environmental change (Dale et al., 2010).

Nevertheless, the quantitative evaluation of the spatial dynamics of urban vibrancy remains a challenging issue, with even the definition of vibrancy *per se* being disputed (Montgomery, 1998; Still and Simmonds, 2000). For instance, Jacobs (1961) defined urban vibrancy as the active street life created by the presence of pedestrians at all times of the day. Gehl (1971) described urban vibrancy not as the number of people or residents, but as the *feeling* that a place is populated and being used. Dougal et al. (2015) defined urban vibrancy as a measure of the spillover effects that arise from the interactions of urban residents.

Since the 1960s, debates on urban vibrancy have been based mainly on anecdotal observations, and numerous qualitative theories have been proposed regarding the associations between vibrancy and various urban factors, including the urban built environment (UBE) (Gehl, 1971; Jacobs, 1961), population density (Simmel, 2002; Wirth, 1938), and safety (Jacobs, 1961). Other factors such as culture and social capital have also been reported as having the potential to generate economic and social benefits (Stern and Seifert, 2010). Due to a lack of data sources, quantitative measures of urban vibrancy were seldom discussed until the late 1990s. Since then, studies began to utilize information on population density, jobs, production, and urban services from then available census and survey data as proxies of urban vibrancy (Holian and Kahn, 2012). The census and survey data, while useful, are static information and inadequate for capturing the fast-changing urban dynamics in the modern world. Moreover, census and survey data usually have a small sample size and a slow update rate.

Since around 2010, newly available spatial big data have provided an opportunity to investigate the dynamics of urban vibrancy at an unprecedentedly fine spatial-temporal scale. So far, various spatial big data sources have been used as proxies of urban vibrancy. The call detailed records (CDR) data, mobile phone tracking data, and GPS trajectories are the most commonly used proxies of urban vibrancy because they directly reflect human presence on the streets (De Nadai et al., 2016; Yue et al., 2017; Kim, 2018; Delclòs-Alió et al., 2019; Jin et al., 2017; Wu et al., 2018). Social media check-in data have also proved useful for capturing the intensity of social and economic activity (Hasan et al., 2013; Xu et al., 2017), and such data have been used as proxies of urban vibrancy (Long and Huang, 2019; Ye et al., 2018; He et al., 2018). Other spatial big data sources, capturing various facets of urban dynamics, have also been proposed to approximate urban vibrancy, including public transport smart card data (Sulis et al., 2018), Wi-Fi access point data (Kim, 2018), and the presence of small catering services (Ye et al., 2018). These spatial big data sources, with their massive sample sizes and great spatial-temporal granularity, have enabled investigations of the detailed spatial-temporal heterogeneity in urban vibrancy.

However, recent studies have mostly used single-source data sets as proxies of vibrancy, even though urban vibrancy is multifaceted (Ednie-Brown, 2012), and single-source data may depict only one of those facets. Only a few new studies, such as those of Long and Huang (2019) and Jin et al. (2017), have used multi-source data. However, while the former used social media data and house price data as independent proxies of economic vitality, which led to contradictory results in modeling that are hard to explain, the latter simply multiplied the density of road junctions, points of interest, and records of human activity to generate a proxy of urban vibrancy, without necessary justification. Therefore, there is a demand for a comprehensive measure of multifaceted vibrancy that fuses multiple data sources in a logical, expandable, and adaptable manner.

The UBE plays a vital role in creating and sustaining vibrancy in urban spaces (Jacobs, 1961; Gehl, 1971). However, there are intensive debates over the associations between urban vibrancy and UBE elements. In her pioneering work, Jacobs (1961) stated that diversity is relevant to neighborhoods and that high-density environments are dialectically linked to the maintenance of urban vibrancy. Contemporary economists such as Porter (2009), Glaeser (2012), and Storper (2013) also articulated the relations between spatial characteristics and the "triumph of cities." Nevertheless, overcrowding, high density, and diversity may cause psychological strain and lead to negative health effects. Simmel (2002) and Wirth (1938) focused on industrial cities in the West and argued that the clustering of high-density populations could result in negative attitudes, hostility, and ignorance among residents. These ideas are crucial to modern planning, having legitimatized ongoing urban renewal, beautification, and gentrification processes. However, supporters of urban density, such as Jacobs, have criticized modern planning theories and strived to maintain a high level of diversity in urban neighborhoods. In fact, the current debates present two limitations. First, most of the evidence to date has been based on the cities of the Global North, and most of the ideas were generated in the last century. Today in the early 21st century, some cities in the Global South are burgeoning centers of urbanization, and new studies of these cities are needed to test the related arguments further (Robinson, 2016). Moreover, most of Jacobs's ideas on urbanism were based on her anecdotal observations. Currently, such observations can be quantitatively investigated by using the newly available "spatial big data."

Therefore, this study has three objectives. First, we propose a comprehensive framework for the evaluation and characterization of urban vibrancy that uses newly available multi-source spatial big data to reflect its multifaceted nature better. Second, we further address the hypothesis regarding the spatial dynamics of urban vibrancy and consider it in the context of Shanghai, which is set against a specific background of high-speed urbanization and market-oriented transition (Wu, 2007). Third, we empirically examine the associations between urban vibrancy and UBEs by using the evaluated vibrancy in Shanghai. The modality and urban form of Chinese cities are different from those of Western cities (Wu, 2016), and so it is essential to explore whether current theories apply to Chinese cities (Long and Huang, 2019; Yue et al., 2017).

The remainder of this paper is organized as follows. The following section introduces the methods, including the conceptual framework proposed for the evaluation and characterization of urban vibrancy, the design of the case study, and the data used. The subsequent section introduces the results, followed by the discussion and conclusions.

Materials and methods

Study area

Shanghai, China (Figure 1) makes a compelling case study. Since the market-oriented reforms of 1978, China has been experiencing tremendous economic reform and urbanization. As the "dragon head" of China's recent growth, Shanghai is a typical case; with more than 24 million residents, Shanghai is the largest city in post-reform China (Wu, 1999). Previous quantitative studies of urban vibrancy were conducted at different scales, from the municipal level (Montgomery, 1998) to the district level (De Nadai et al., 2016; Yue et al., 2017). We adopted a grid with a 1 km^2 spatial resolution (Zhou and Long, 2016) to provide neighborhood scale analysis of the spatial dynamics of urban vibrancy at the neighborhood level. Please refer to the online supplementary material for more details on the zoning system.

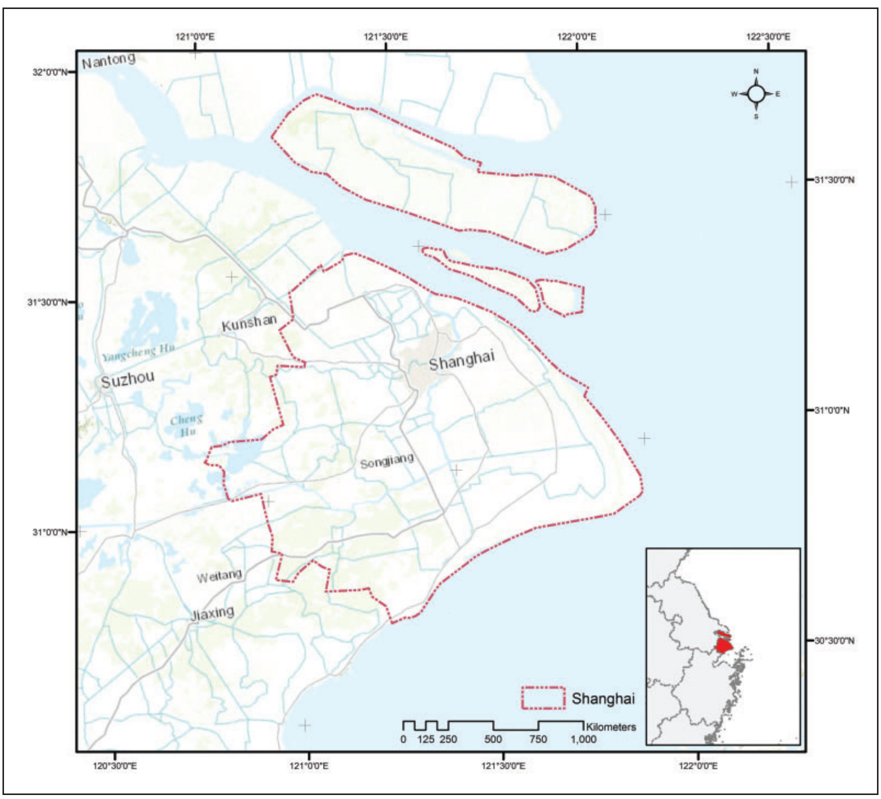

Figure 1. Shanghai, the study area.

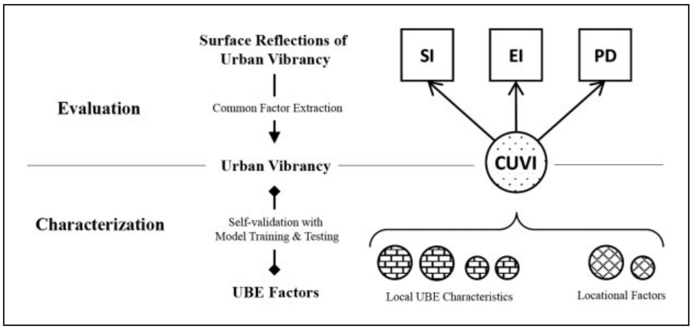

Figure 2. The conceptual framework proposed for evaluating and characterizing urban vibrancy based on factor analytical theory. CUVI, comprehensive urban vibrancy index, stands for the extracted proxy for urban vibrancy. UBE: urban built environment; SI: social activity intensity; EI: economic activity intensity; PD: pedestrian density.

Conceptual framework

The proposed framework (Figure 2) is composed of two parts that address the evaluation and characterization of urban vibrancy, respectively.

Evaluation of urban vibrancy. Urban vibrancy, given its multifaceted nature, is difficult to directly and comprehensively measure. However, the effects of urban vibrancy can be reflected and observed in a variety of so-called surface attributes. Being fully or at least partially attributable to vibrancy, these surface attributes collectively provide an opportunity to infer latent vibrancy. Therefore, we propose fusing multiple data sources to measure the surface attributes of multifaceted urban vibrancy.

The proposed measure of urban vibrancy is extracted as the common factor not only underlying a series of its surface attributes but also causing those attributes to co-vary. The surface attributes included in this manuscript were carefully selected based on the definition that vibrancy is defined not only by the number of people or residents, but also by the feeling that a place is populated and being used (Gehl, 1971), which increases population density (Jacobs, 1961), the intensity of social activity and interaction (Couture, 2013; Jacobs, 1961), and the intensity of economic activity (Brenner, 2014; Brenner et al., 2010). Based on this definition, the intensity of social and economic activity and pedestrian flow were selected as the surface attributes used to extract the measure of urban vibrancy.

The factor analysis (FA) method was used to extract urban vibrancy out of its chaotically correlated surface attributes. The FA method is a state-of-the-art method for studying the common latent factors that account for observed variations and covariations across a wide range of surface attributes (Fabrigar et al., 1999), and this approach has been widely applied in classical urban studies on factorial ecology (Park et al., 1925). In this paper, factor analysis was conducted by using the standardized surface attributes as inputs and principal axis factoring as the factor extraction method. The proxy of urban vibrancy was extracted as the common factor with the largest eigenvalue (Table 1), termed as the comprehensive urban vibrancy index (CUVI).

The derived vibrancy, based on multi-source data, renders a more comprehensive proxy of vibrancy than do the proxies that depict only one facet of vibrancy. The derived proxy is

Table 1. The solution of the factor analysis and descriptive statistics of the surface attributes.

	Descriptive statistics			Factor weights	Communalities	
Surface attributes	N	Mean	Std. deviation	1	Initial	Extracted
1. ln(EI)	741	5.16	0.90	.653	.42	.43
2. ln(SI)	741	6.31	2.13	.991	.67	.98
3. ln(PD)	741	6.16	0.79	.767	.56	.57

Table 2. Multi-source spatial big data used for evaluating urban vibrancy and quantifying UBE indicators.

Surface attributes		Data source	Year
Social activity intensity (SI)		Sina Weibo data	2016
Economic activity intensity (EI)		Dianping life services reviews	2016
Pedestrian density (PD)		Mobile phone GPS positioning requests	2016
UBE dimensions	Indicators	Data source	Year
Density	Building density	tianditu.com	2016
	Density of urban functions	map.baidu.com	2016
Intensity	Mean building height	lianjia.com	2016
Diversity	Diversity of urban functions	map.baidu.com	2016
	Diversity of building height	lianjia.com	2016
	Diversity of building age	lianjia.com	2016
	Diversity of house prices	lianjia.com	2016
Quality	Mean house price	lianjia.com	2016
Aging buildings	Mean building age	lianjia.com	2016
Walkability	Density of road junctions	SinoGrids (Zhou and Long, 2016)	2011

positively associated with all of the surface attributes, just as vibrancy should be. In doing so, the proposed measure is self-validated across multiple data sources.

Characterization of urban vibrancy. Various UBE elements may affect urban vibrancy, among which the most commonly discussed ones include density, diversity, and design, i.e., "the three Ds." Shorter street segments, aging buildings, and highly connected roads have also been identified as primary ingredients for urban vibrancy (Gehl, 1971; Jacobs, 1961; Montgomery, 1998). We selected density, intensity, diversity, quality, walkability, and aging buildings as the six major dimensions of the UBE to test their respective associations with urban vibrancy. The urban form in Chinese cities is characterized by a substantially greater use of vertical space compared with Western cities, so we differentiated density and intensity to examine their respective contributions. Density measures the use of horizontal space, while intensity measures the use of vertical space. Locational factors were also included to reflect the agglomeration effect of urban centers and the impact of public transportation infrastructures.

Measurements of surface attributes

We first measured the surface attributes to capture the effects of the underlying urban vibrancy from multiple perspectives utilizing multi-source spatial big data (Table 2). All of the measures were aggregated onto the uniform grid of Shanghai.

Social activity intensity. The social activity intensity (SI) was evaluated using a Weibo data set containing more than 50 million records for 2016, obtained from Weibo Open API (http://open.weibo.com/). Sina Weibo, the largest microblogging service in China, had 222 million monthly active users (MAU) by 2015. The SI was calculated as the density of Weibo posts for each grid cell, as follows

$$SI(x,y) = N_{post}(x,y) \tag{1}$$

where $SI(x,y)$ is the intensity of social activities at location (x, y) and $N_{post}(x, y)$ is the total number of geotagged microblogs within the cell at location (x, y).

Economic activity intensity. As suggested by Long and Huang (2019), the economic activity intensity (EI) was evaluated using a Dianping life services review data set including information from over 300,000 merchants in Shanghai in 2016. Dazhong Dianping (http://dianping.com) is one of the first and most widely used life services review websites in China. The EI in each grid cell was calculated as the density of merchants weighted by the popularity of each merchant, as follows

$$EI(x,y) = \sum\nolimits_{i}^{N_e(x,y)} P_i(x,y), \quad i \in [1, N_e(x,y)] \tag{2}$$

where $EI(x,y)$ is the intensity of economic activity with the cell (x, y), $N_e(x,y)$ is the total number of life service providers within the cell, and $P_i(x,y)$ is the popularity index of the ith life services provider within the cell.

Pedestrian density. Population density (PD) was measured using a mobile phone GPS positioning data set for a full week in April 2016, obtained from Tencent Map API (http://lbs.qq.com). The Tencent data record user locations when they request GPS positioning services. Tencent collects more than 36 billion daily positioning requests from more than 450 million users around the world.

Quantifications of UBE indicators

A list of indicators was selected to reflect the six dimensions of the UBE (Table 2). Multi-source spatial data sources that describe the UBE, including online open urban data from government-initiated, corporate-initiated and crowdsourcing platforms, and online house-trading posts, were utilized to quantify these UBE indicators. Please refer to the online supplementary material for the descriptive statistics of the UBE factors.

The building density was evaluated using the detailed building footprint data based on recent government surveying, which was acquired from Tianditu (http://www.tianditu.com/), China's first government-initiated free web map service. For UBE indicators related to urban functions and land use, the point of interest (POI) data were collected from the open API provided by Baidu Map (http://map.baidu.com), with seven categories of urban functions: government agencies, transportation, business, education, corporate enterprises, residential, and others. The density of urban functions was calculated as the density of POI points in each grid cell. The Shannon Index (Shannon, 1948) was utilized as the diversity metrics for the calculation of the diversity of urban functions and other diversity measures in this paper. Accessibility was also calculated using the POI data as the density of public transportation POIs in each grid. The road junction data set was retrieved from SinoGrids

(Zhou and Long, 2016), a crowdsourcing open urban data platform, for the evaluation of the walkability of UBE. We also obtained detailed and up-to-date information of the height, age, and price of urban buildings from 85,000 house-trading posts retrieved from Lianjia (http://lianjia.com) in 2016, one of the largest online real estate trading platforms in China.

The distances to the city center were calculated as the distance from each cell to its nearest city center. Shanghai is a polycentric city with one main city center and several sub-centers. The city center locations recognized by Cai et al. (2017) were used in this study.

Ordinary least squares regression with k-fold cross-validation

Ordinary least squared (OLS) regression was used to estimate the relationship between urban vibrancy and UBE indicators. To avoid the multicollinearity problem, the bivariate Pearson correlation coefficients (PCCs) among the explanatory variables were checked before the regression, and the variance inflation factor (VIF) was also calculated for each explanatory factor after the regression. The empirical bootstrap confidence interval was used to determine the statistical significance of each explanatory variable. Unlike the classical t-test, bootstrap confidence intervals are free from the assumptions of empirical distributions.

K-fold cross-validation was applied throughout the modeling process to self-validate the estimated relationship and avoid overfitting. Using the design of AB-testing, cross-validation trains the model on a randomly selected proportion of the samples and tests it on the remaining samples, providing insight into how the model generalizes to an unknown data set in practice. In this study, we applied 10-fold cross-validation, where all data samples were split into 10 shares; in each of the 10 rounds, one of the 10 shares was held for model testing while the other nine shares were used for model training.

Results

Evaluated urban vibrancy in Shanghai

We observed substantial spatial variances in the surface attributes of Shanghai, especially in the SI (Table 1). Like many other measures of urban dynamics, the three surface attributes follow heavy-tailed distributions, so logarithmic transformations were applied to the surface attributes before further analysis. Bivariate PCCs indicated the existence of moderate correlations among the surface attributes.

We evaluated urban vibrancy as the common latent factor that leads to covariance among SI, EI, and PD. As a linear combination of the SI, EI, and PD, the derived vibrancy also accounted for most (75.56%) of the communalities among the surface attributes and most (66.03%) of the total variances contained in the surface attributes.

We mapped the urban vibrancy in Shanghai to visualize its spatial dynamics (Figure 3) using the head/tail breaks technique (Jiang, 2013) as the classification scheme. Compared with other classification techniques, such as Jenk's natural break, the head/tail breaks method is especially useful for reclassifying heavy-tailed data, such as the derived vibrancy. The technique explicitly illustrated the polycentric structure of urban vibrancy in Shanghai with automatically defined parameters, including the number of classes and classification intervals.

It is challenging to validate the measure of vibrancy directly. As we observed significant agglomerations in the evaluated vibrancy, we conducted a partial validation of the proposed proxy by comparing the vibrancy centers identified in this study with those specified in local

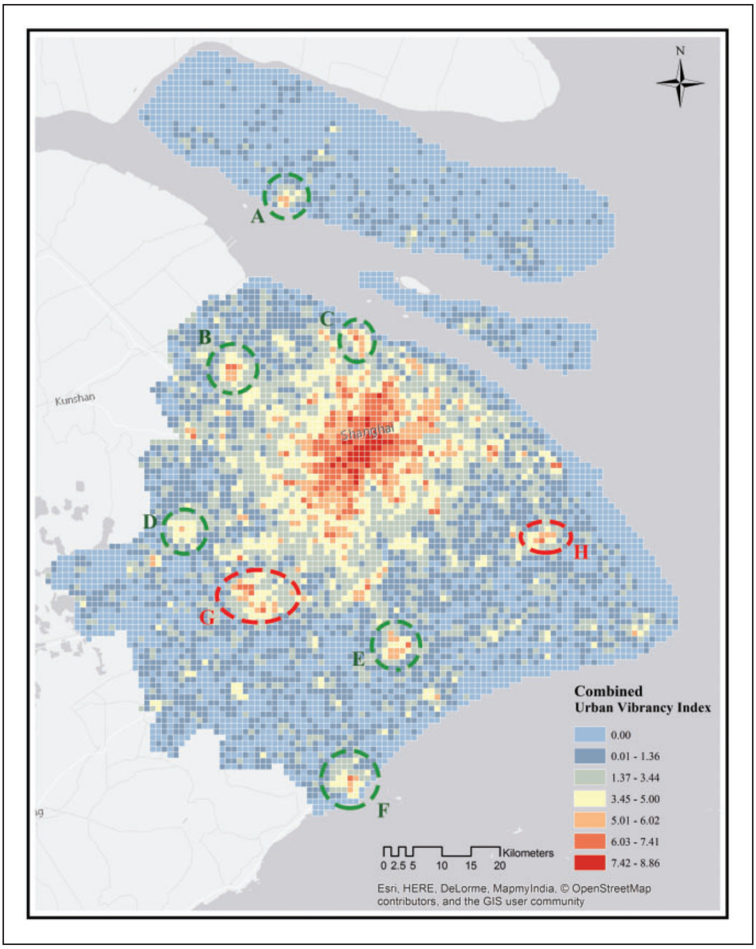

Figure 3. The evaluated urban vibrancy in Shanghai, classified using the head/tail breaks technique. Vibrancy centers in Shanghai identified using visual interpretation are marked with dashed circles. All centers identified in Hu et al. (2014) (in green) match well with our results, while two extra vibrancy centers are also discovered in this paper (in red).

urban development plans and a closely related study. In addition to the main city center, we identified eight vibrancy centers in Shanghai via visual interpretation (Figure 3). The identified vibrancy centers are mainly satellite town centers in Shanghai, including Qingpu, Songjiang, and Jiading. First, these identified centers matched well with the urban subcenters listed in the Shanghai Master Plan 2017–2035 (2018). Second, we further compared the centers with the mobility hotspots detected by Hu et al. (2014) (A to F in Figure 3) using taxi trajectory data. All the six mobility hotspots in Hu et al. (2014) matched the vibrancy centers identified in this paper, while two additional urban vibrancy centers (G and H in Figure 3) were identified in this paper. The two additional centers were vibrant areas with good accessibility to the public transportation system and relatively low taxi usage, such as the town of Songjiang (G in Figure 3) and the town of Huinan (H in Figure 3). In summary, the good matches between our results, the local urban development plan, and the related study served as an indirect validation of the evaluated urban vibrancy and the proposed framework in this paper.

Table 3. Results of the associations between urban vibrancy and UBEs ranked by standardized coefficients.

Dimension of UBE	UBE indicators	Abbr.	Beta	VIF
Density	Building density	BD	.31*	2.852
Density	Density of urban functions	POIDens	.19*	5.213
Diversity	Diversity of urban functions	POIDiv	.16*	1.511
Locational factors	Accessibility by public transportation	ACCE	.14*	2.491
Quality	Average house price	AVGP	.13*	2.700
Walkability	Road junction density	RJDens	.12*	2.200
Intensity	Average building height	AVGH	.11*	1.385
Locational factors	Distance to city center	DISC	.07*	1.193
Aging buildings	Average house age	AVGA	.05*	2.587
Diversity	Diversity of house prices	STDP	−.09*	2.668
Diversity		STDA	–	–
Diversity		STDH	–	–

N = 741.
VIF: variance inflation factor; UBE: urban built environment.
*Significant at the level of 0.005.

Urban vibrancy in association with UBE indicators

Online supplementary Table S1 shows the summary statistics and bivariate PCCs among the UBE indicators. No strong ($\geq$0.7) multi-collinearity among the UBE indicators was detected. In the feature selection process, two of the UBE indicators, diversity of house age and diversity of building height, were not found to be statistically significant ($p < 0.005$) so they were excluded from the final model.

We achieved similarly good training (R^2=0.764 $\pm$ 0.005) and testing accuracy (R^2=0.760 $\pm$ 0.047) in the 10-fold cross-validation, indicating no significant over fitting. The final model was then trained using all data samples and the significant UBE indicators (R^2=0.733, Table 3). The VIF values (Table 3) indicated no considerable redundancy among the explanatory variables.

The estimated coefficients (Table 3) in the final model provide evidence for the statistical associations between urban vibrancy and the UBE indicators.

Discussion

Evaluating urban vibrancy with multi-source spatial big data

Quantitative studies of urban vibrancy have been limited by the amount of available data across various contexts. In recent years, the increasing availability of spatial big data has empowered quantitative studies on urban vibrancy and UBEs. Web 2.0 applications, widespread mobile phone usage, and technological innovations in GPS have led to the development of various data sources, such as mobile phone CDR data, GPS positioning data, and social media data. These data sets have enabled the evaluation and analysis of human activity and urban functions in cities (Tu et al., 2017, 2018). Moreover, with the rapid development of GIS platforms, volunteer GIS sites and the open/big data movement, detailed spatial big data on urban form and function are also becoming increasingly accessible.

Compared with simple and straightforward urban concepts such as PD, urban vibrancy is a comprehensive concept that is more difficult to measure directly. Measures of urban vibrancy using single-source data can be regarded as partial and subjective. Urban vibrancy is described as the spillover effects that arise in urban contexts from the endogenous interactions of the people living in the city; these effects influence many aspects of urbanism. This timely study takes a step forward and proposes a comprehensive framework for evaluating urban vibrancy by using multi-source spatial big data. Although urban vibrancy cannot be directly measured, its effects can be reflected by an array of surface attributes that are fully or partially affected by it. FA is an effective method for measuring the latent quantities that cannot be measured directly, such as urban vibrancy, specifically by analyzing the communalities and the uniqueness of its multiple attributes, which are seemingly unrelated but positively correlated.

Urban vibrancy is not the only urban concept that is challenging to measure. The proposed framework has significant potential for evaluating other comprehensive urban concepts, such as urban resilience and innovation capability, for which single and direct measures are difficult to identify. These concepts influence a broad range of urban life, and their effects are reflected in an array of surface attributes.

The use of spatial big data sources offers enormous possibilities for understanding urban dynamics with large sample sizes and rich spatial and temporal details, and for detecting complex and non-linear variables (Lazar, 2014). However, big data researchers should always be aware of the biases of these data sources, such as sampling bias and heterogeneous sampling, and then try to minimize the effect of these biases. The proposed method of fusing multiple data sources also serves as a self-validation mechanism by generating measures from multiple spatial big data sources.

Characterizing urban vibrancy with UBEs

Existing studies showed conflicting results on the associations between vibrancy and some UBE elements (online supplementary Table S2). We focus on the context of the Global South and contribute to the understanding of these associations by providing quantitative evidence in the city of Shanghai, China. In Shanghai, seven UBE indicators, along with the two locational factors, contribute positively to urban vibrancy. The two density measures, building density (0.31) and density of urban functions (0.19), have the strongest positive associations with vibrancy, followed by the diversity of urban functions (0.16), accessibility by public transportation (0.14), average house price (0.13), and average building height (0.11). The diversity of house prices has a slightly negative association (−0.09) with urban vibrancy.

Density or intensity. High-density environments are closely associated with the maintenance of urban vibrancy (Jacobs, 1961). The spatial concentration is also linked to what economists have described as the "triumph of cities" (Glaeser, 2012).

While quantitative studies have disagreed over the significance of the density of urban functions (online supplementary Table S2), our results demonstrate a strong positive connection between vibrancy and the spatial concentrations of both the built environment and urban functions. Cities are the primary places for the production, distribution, and consumption of material goods, featuring a dense concentration of urban functions that satisfy the numerous and various needs of urban residents, thereby encouraging social and economic activity and increasing local vibrancy. The concentration of urban population further promotes urban vibrancy by making the urban environment safer and encouraging face-to-face interaction among people.

Compared with the Western urban form, large cities in the Global South, especially in post-reform China, show extensive development along their vertical dimension of urban space. Out of the world's 50 cities with the most skyscrapers, 23 (46%) are in China, and 37 (74%) are in Asia (Emporis, 2018). Evidence in Shanghai echoes De Nadai et al. (2016) that increased building height contributes positively (0.11) to vibrancy. On the one hand, building height directly determines the capacity of urban spaces to host human activities besides building density. On the other hand, the vertical structure of buildings has a psychological effect on people's desire to linger (Cattell et al., 2008).

However, the fact that the increased horizontal building density contributes to 2.8 times more positive changes in vibrancy compared to the increased building height suggests that for cities in China and other southern cities, the planning and design of vibrant urban spaces should focus primarily on filling horizontal space rather than vertical space.

Diversity. In theory, diversity is regarded as the primary generator of urban vibrancy (Jacobs, 1961). However, studies have reported conflicting results (online supplementary Table S2). For example, De Nadai et al. (2016) showed no significant correlation between vibrancy and the diversity of urban functions in Italian cities.

Our study contributes to the debate by providing evidence for Chinese cities. Our results show that among the diversity metrics, the mixture of urban functions is the most positively associated UBE indicator, which echoes Jacobs' (1961) theory that a diverse range of urban functions encourages people to linger in urban spaces and conduct various activities. We identified a slightly negative association (−0.09) between urban vibrancy and the diversity of housing prices, which contradicts the theory stating that a vibrant neighborhood should provide affordable shelter for people of different income levels. This discrepancy may be partly attributed to the disproportional accumulation of good transportation, hospitals, schools, and commercial services in high-priced locations within city centers. Also, vibrant urban areas tend to have higher and more consistent house prices. The dramatic increase in residential mobility could also contribute to the discrepancy. Due to well-developed public transportation systems, the service area of vibrant urban areas expands, and the local vibrancy level is no longer limited to residents from the local community. Meanwhile, we observed no significant associations between vibrancy and the diversity of building ages and building heights. In fact, most of today's Shanghai was built up over the last two or three decades, and most of the buildings in the central city have been renovated. Furthermore, due to most districts in the city center being full of newly built high-rise buildings or commodity housing, the disparities in building height among districts have also decreased substantially.

Walkability and accessibility. Good walkability promotes urban vibrancy. High road junction density is a direct reflection of road network density; the size of street blocks slows down the pace of urbanites' lives, giving them enough time for social and consumption activities. Small blocks increase the number of random human contacts, and Jacobs (1961) regarded these blocks as necessary for creating a vibrant urban space because they allow people to avoid lengthy and tedious walks. Also, a high density of road junctions provides multiple path options and attracts small and diverse neighborhood shops. Our conclusion for Shanghai is consistent with that of Jacobs (1961) and Long and Huang (2019); thus, compared with large rectangular blocks, small blocks not only create more opportunities for human contact but also improve the effective mixture of land uses, thereby leading to vibrant urban areas. Although Shanghai is evolving as a global city with an unprecedented rise in car ownership, street life and walkability still matter in the everyday lives of the Shanghainese.

With good accessibility, urban areas can attract people of mixed social-economic status from everywhere in the city, which explains accessibility by public transportation (0.14) having the fourth strongest positive association with urban vibrancy. The distance to the city center is also found to be significant (0.09).

Community quality and aging buildings. As a comprehensive reflection of the quality of the UBE according to life convenience, safety, and the accessibility and quality of nearby urban infrastructures, the average housing price (0.13) ranks fifth among the positive factors.

We observed a slightly positive correlation (0.05) between urban vibrancy and the average building age. Jacobs (1961) mentioned the role of aging buildings as one of the four key generators of urban diversity. Her theory is that "*a good lot of plain, ordinary, low-value old buildings*" should be affordable not only for high-profit, well-subsidized enterprises but also for smaller but diverse enterprises such as neighborhood bars. This theory partially explains the positive association between urban vibrancy and average building age in Shanghai. In addition, most post-reform Chinese cities, such as Shanghai, have recently undergone a process of massive growth, and their newly built neighborhoods need time to mature and become vibrant.

Conclusions

Being a magnet for both human and social capital, vibrancy may be the key to improving and sustaining the competitiveness and liveliness of cities. The value of this work is highlighted against the background of the increasingly intense global intercity competition and the newly available spatial big data. We compared qualitative analyses of the associations between vibrancy and UBEs and found conflicting results, making it necessary to add more quantitative evidence to the ongoing debate, especially in the context of the Global South. As China's economy is shifting gears after its unprecedented urbanization in recent decades, many cities are starting to show signs of shrinkage, with vacant urban spaces and underused urban infrastructure (Long and Wu, 2016). As a result, this study of sustainable urban development in China is especially timely.

Building on the advantages of previous studies in evaluating urban vibrancy, we proposed a framework to derive a more comprehensive measure of vibrancy that better reflects its multifaceted nature. Rather than using single-source data that may depict only one facet of vibrancy, we measured vibrancy as the latent common factor that leads to the co-variance among its multiple surface attributes, which are seemingly unrelated but positively correlated. Compared with analyses using multiple independent proxies of vibrancy, our method avoids potential conflict in the subsequent modeling process because the effects of vibrancy on multiple surface attributes are aggregated into one comprehensive proxy of vibrancy. The framework was applied in Shanghai, a typical post-reform city in China. The hypothesis of the spatial dynamics of urban vibrancy, based on the derived vibrancy, is addressed within the context of an urban area of the Global South.

Following the theories and analyses of the intimate connections between vibrancy and UBEs, we systematically investigated and cross-validated their associations based on the evaluated vibrancy in Shanghai. Prior qualitative studies of urban vibrancy were extended through this study's analysis of the association between urban vibrancy and UBEs. While previous studies have mainly been conducted in the cities of the Global North and while the relevant ideas were mainly generated in the last century, our case study tests current assumptions in the context of large cities in the Global South. Moreover, the urban form in Chinese cities is different from that of their Western counterparts, especially regarding the use of vertical spaces. We found that horizontal built-up density is the leading generator of vibrancy in Shanghai, while vertical height also has limited positive effects. The density and mixture of urban functions, accessibility, and walkability are also positive generators of vibrancy in Shanghai. Our results contribute to the debate and future planning practices regarding vibrant spaces in large cities.

The proposed framework for evaluating urban vibrancy, considering its multifaceted nature, can benefit future studies of urban vibrancy. This framework has considerable expandability and adaptability, thus allowing for the inclusion of a diverse range of surface attributes, such as jobs, safety, and urban services, based on specific definitions of vibrancy. For future work, it may be useful to apply the proposed framework in different cities or use different data sources, and then discuss the difference in results. Meanwhile, the relationship between vibrancy and factors other than the UBE factors, such as social capital and culture, can be further discussed under the proposed framework with its spatial big data. Various spatial techniques, such as geographically weighted regression and clustering algorithms, can be applied to understand the spatial-temporal patterns within the evaluated vibrancy. Finally, urban vibrancy is not the only urban concept that has a multifaceted nature and is challenging to measure. The framework can also help evaluate other comprehensive urban concepts, such as innovation and resilience, for which single and direct measures are hard to identify.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Yimeng Song http://orcid.org/0000-0001-9558-1220

Supplemental material

Supplemental material for this article is available online.

References

Brenner N (2014) *Implosions/Explosions: Towards a Study of Planetary Urbanization*. Berlin: Jovis.

Brenner N, Peck J and Theodore N (2010) Variegated neoliberalization: Geographies, modalities, pathways. *Global Networks: A Journal of Transnational Affairs* 10: 182–222.

Cai J, Huang B and Song Y (2017) Using multi-source geospatial big data to identify the structure of polycentric cities. *Remote Sensing of Environment* 202: 210–221.

Cattell V, Dines N, Gesler W, et al. (2008) Mingling, observing, and lingering: Everyday public spaces and their implications for well-being and social relations. *Health & Place* 14(3): 544–561.

Couture V (2013) *Valuing the consumption benefits of urban density*. University of California, Berkeley. Processed.

Dale A, Ling C and Newman L (2010) Community vitality: The role of community-level resilience adaptation and innovation in sustainable development. *Sustainability* 2(1): 215–231.

Delclós-Alió X, Gutiérrez A and Miralles-Guasch C (2019) The urban vitality conditions of Jane Jacobs in Barcelona: Residential and smartphone-based tracking measurements of the built environment in a Mediterranean metropolis. *Cities* 86: 220–228.

De Nadai M, Staiano J, Larcher R, et al. (2016) The death and life of great Italian cities: A mobile phone data perspective. In: *Proceedings of the 25th international conference on world wide web*. International World Wide Web Conferences Steering Committee, pp. 413–423. Available at: https://doi.org/10.1145/2872427.2883084 (accessed 22 January 2019).

Dougal C, Parsons CA and Titman S (2015) Urban vibrancy and corporate growth. *The Journal of Finance* 70(1): 163–210.

Ednie-Brown P (2012) Architecture on wire: Resilience through vitality. *Log* 25: 15–22.

Porter ME (2009) Location, competition, and economic development: Local clusters in a global economy. *Economic Development Quarterly* 14: 15–34.

Emporis (2018) Cities with the most skyscrapers. Available at: www.emporis.com/statistics/most-skyscraper-cities-worldwide (accessed 8 January 2018).

Fabrigar LR, Wegener DT, MacCallum RC, et al. (1999) Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods* 4(3): 272.

Gehl J (1971) *Life Between Buildings: Using Public Space*. Washington: Island Press.

Glaeser EL (2012) *Triumph of the City: How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier*. New York: Penguin Books.

Hall P and Pfeiffer U (2013) *Urban Future 21: A Global Agenda for Twenty-First Century Cities*. New York, NY: Routledge.

Hasan S, Zhan X and Ukkusuri SV (2013) Understanding urban human activity and mobility patterns using large-scale location-based data from online social media. In: *Proceedings of the 2Nd ACM SIGKDD International Workshop on Urban Computing*. New York: ACM, pp. 6:1–6:8.

He Q, He W, Song Y, et al. (2018) The impact of urban growth patterns on urban vitality in newly built-up areas based on an association rules analysis using geographical 'big data.' *Land Use Policy* 78: 726–738.

Holian MJ and Kahn ME (2012) The impact of center city economic and cultural vibrancy on greenhouse gas emissions from transportation (No. CA-MTI-12-1002). Mineta Transportation Institute, College of Business, San José State University.

Hu Y, Miller HJ and Li X (2014) Detecting and analyzing mobility hotspots using surface networks. *Transactions in GIS* 18(6): 911–935.

Jacobs J (1961) *The Death and Life of Great American Cities*. New York, NY: Vintage.

Jiang B (2013) Head/tail breaks: A new classification scheme for data with a heavy-tailed distribution. *The Professional Geographer* 65(3): 482–494.

Jin X, Long Y, Sun W, et al. (2017) 'Evaluating cities' vitality and identifying ghost cities in China with emerging geographical data. *Cities* 63: 98–109.

Kim YL (2018) Seoul's Wi-Fi hotspots: Wi-Fi access points as an indicator of urban vitality. *Computers, Environment and Urban Systems* 72: 13–24.

Lazar D (2014) The parable of Google Flu:Traps in big data analysis. *Science 343*.

Long Y and Huang C (2019) Does block size matter? The impact of urban design on economic vitality for Chinese cities. *Environment and Planning B: Urban Analytics and City Science* 46(3): 406–422.

Long Y and Wu K (2016) Shrinking cities in a rapidly urbanizing China. *Environment and Planning A* 48(2): 220–222.

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

Park RE, Burgess EW and McKenzie RD (1925) *The City*. Chicago: Chicago University Press.

Pinquart M and Sörensen S (2000) Influences of socioeconomic status, social network, and competence on subjective well-being in later life: A meta-analysis. *Psychology and Aging* 15(2): 187–224.

Robinson J (2016) Comparative urbanism: New geographies and cultures of theorizing the urban. *International Journal of Urban and Regional Research* 40: 187–199.

Shanghai Master Plan 2017-2035 (2018, January) Available at: www.shanghai.gov.cn/newshanghai/xxgkfj/2035001.pdf (accessed 22 January 2019).
Shannon CE (1948) A mathematical theory of communication. *Bell System Technical Journal* 27: 379–423, 623–656.
Simmel G (2002) The metropolis and mental life. In: Bridge G and Watson S (eds) *The Blackwell City Reader*. Oxford: Wiley-Blackwell, pp. 11–19.
Stern MJ and Seifert SC (2010) Cultural clusters: The implications of cultural assets agglomeration for neighborhood revitalization. *Journal of Planning Education and Research, 29*(3), 262–279.
Still B and Simmonds D (2000) Parking restraint policy and urban vitality. *Transport Reviews* 20(3): 291–316.
Storper M (2013) *Keys to the City: How Economics, Institutions, Social Interaction, and Politics Shape Development*. Princeton: Princeton University Press.
Sulis P, Manley E, Zhong C and Batty M (2018) Using mobility data as proxy for measuring urban vitality. *Journal of Spatial Information Science* 16(6): 137–162. https://doi.org/10.5311/JOSIS.2018.16.384
Tu W, Cao J, Yue Y, et al. (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(12): 2331–2358.
Tu W, Cao R, Yue Y, et al. (2018) Spatial variations in urban public ridership derived from GPS trajectories and smart card data. *Journal of Transport Geography* 69(3688): 45–57.
Wirth L (1938) Urbanism as a way of life. *American Journal of Sociology* 40: 1–24.
Wu C, Ye X, Ren F and Du Q (2018) Check-in behaviour and spatio-temporal vibrancy: An exploratory analysis in Shenzhen, China. *Cities* 77: 104–116.
Wu F (2007) *China's Emerging Cities: The Making of New Urbanism* London: Routledge.
Wu F (2016) Emerging Chinese cities: Implications for global urban studies. *Professional Geographer* 68: 1–11.
Wu W (1999) Shanghai. *Cities* 16: 207–216.
Xu Z, Liu Y, Yen N, et al. (2017) Crowdsourcing based description of urban emergency events using social media big data. *IEEE Transactions on Cloud Computing*. Available at: https://doi.org/10.1109/TCC.2016.2517638 (accessed 22 January 2019).
Ye Y, Li D and Liu X (2018) How block density and typology affect urban vitality: an exploratory analysis in Shenzhen, China. *Urban Geography* 39(4): 631–652.
Yue Y, Zhuang Y, Yeh AGO, et al. (2017). Measurements of POI-based mixed use and their relationships with neighbourhood vibrancy. *International Journal of Geographical Information Science* 31(4): 658–675.
Zhou Y and Long Y (2016) SinoGrids: A practice for open urban data in China. *Cartography and Geographic Information Science* 43(5): 379–392.

Bo Huang is a Professor in the Department of Geography and Resource Management, The Chinese University of Hong Kong, where he is also the associate director of Institute of Space and Earth Information Science (ISEIS). His research interests cover most aspects of GIScience, specifically the design and development of models and algorithms for spatial/spatio-temporal statistics, unified satellite image fusion and multi-objective spatial optimization, and their applications in environmental monitoring and sustainable land use and transportation planning. He currently serves as the Asia-Pacific Editor of International Journal of Geographical Information Science (Taylor & Francis).

Yulun Zhou is a PhD student in the Department of Geography and Resources Management at the Chinese University of Hong Kong, China. He works on quantified urban studies empowered by the geospatial big data analysis, especially those related to the urban built environment and its impacts on spatial behavior and urban health. Other areas of his academic interest include: spatial optimization methods and urban land use planning. He

received his MSc in Geo-Information Science from the Chinese University of Hong Kong and a B.Eng. from Fudan University, Shanghai, China.

Zhigang Li is a Professor at the School of Urban Design, Wuhan University. He also serves as the school dean. As both a geographer and planner, Professor Li works on the socio-spatial transformation of urban China, with a focus on such issues as residential segregation, community, migration, planning, etc. He services as the vice chairperson of China's Urban Geography Committee and a board member of Urban Studies (SAGE).

Yimeng Song is now a PhD candidate in the Department of Geography and Resource Management, The Chinese University of Hong Kong, Hong Kong, China. His research focuses on quantitative urban studies, geospatial big data analysis, dynamic urban environmental exposure assessment and environmental health impact assessment. He received his MSc in Geo-Information Science at The Chinese University of Hong Kong and received his BA in Economics and B.Eng. in Land Recourse Management at Peking University and China University of Geosciences (Beijing).

Jixuan Cai is currently a PhD candidate at The Chinese University of Hong Kong. He is now working towards the dissertation on analyzing the bias of big data, demographic representativeness and mapping fine resolution population distribution. Broadly, his methodological research focus is on human behavior simulation with urban environment modeling using spatial big data, which includes geo-tagged Internet data and earth observation imagery. Besides, he is also experienced in high performance computing for efficient geospatial data managing, processing and analyzing.

Wei Tu is a senior associate research fellow at the Lab of Spatial Smart Sensing and Service, Shenzhen University, China. He obtained the PhD from State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, China. His research interests focus on the urban informatics, including data-driven human activity-mobility (i.e., taxi GPS, bus GPS, smart card data, social media, data, mobile phone data, etc.), multi-source geospatial data fusion, and smart applications in the cities worldwide.