Exploring the relationship between the built environment and block vitality based on multi-source big data: an analysis in Shenzhen, China

Geyu Lv, Sheng Zheng and Wei Hu

Department of Land Management, Zhejiang University, Hangzhou, China; Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control (AEMPC), Nanjing University of Information Science & Technology, Nanjing, China; School of Resource and Environment Science, Wuhan University, Wuhan, China

ABSTRACT
Improving the vitality of cities has long been considered an important goal of planning. However, people’s understanding of how complex and diverse built environment factors affect urban vitality is still limited. In recent years, the emergence of new data provides a new perspective for the study of urban vitality. In this article, the spatio-temporal variation of urban vitality was quantitatively measured by using Baidu heat map, and the influence of built environment factors on urban vitality is further analyzed by geographically weighted regression model. The analysis was conducted at the block level, taking into account differences between weekdays and weekends. The results show that Shenzhen presents a vitality pattern of three centers and two sub-centers, and the average vitality level of weekdays is higher than that of weekends. Distance to subway station, road density, residential density, land mixed use, and compactness have significant influence on block vitality, but the influence varies from block to block, showing strong spatial heterogeneity. Commercial facility density and floor space ratio show significance only on weekends and weekdays, respectively. The findings reveal that we need to take regional differences into consideration and develop more targeted urban planning policies to facilitate block vitality.

ARTICLE HISTORY
Received 11 March 2022
Accepted 14 June 2022

KEYWORDS
Urban vitality; urban built environment; multi-source big data; geographically weighted regression; Baidu heat map

1. Introduction
Since Jane Jacobs put forward the concept of urban vitality, people’s understanding of this concept has been deepening, and creating a vibrant city has become the basic goal of urban planning and development, especially in the center of the city (Jacobs 1961; Montgomery 1998). However, since this concept is mainly prevalent in western
countries before, and early urbanization in China was mainly rapid and low-quality expansion, improving the vitality of cities have received little attention in China’s urban planning before the twenty-first century (Ye et al. 2018). Large-scale demolition, reconstruction, and expansion not only destroy the diversity of some urban fabric but also cause problems such as residential vacancy and social segregation, which threaten the sustainability of the current urban system (Batty 2016). In recent years, with the increase of income level and housing demand, the public began to pay more attention to the quality of urban space, and the government also started a new urbanization plan aimed at improving the quality of urbanization (Wang et al. 2015). This transformation requires planners and designers to play a more active role in stimulating urban vitality and improving the quality of the built environment (Lang et al. 2016). Therefore, further study on the distribution of urban vitality and understanding of the complex relationship between urban vitality and the built environment is of great significance for urban development.

There are two pathways used by researchers to measure urban vitality. One way is more traditional which uses field investigation and observation vitality (Azmi and Karim 2012). But it is almost impossible to make large-scale measurements using this method because of the amount of manpower and time required. With the development of science and technology, multi-source big data provides a new possibility to measure urban vitality. Geo-tagged data, such as points of interest (POIs) (Long and Huang 2017; Zhang et al. 2021), social media data (Wu et al. 2018a; Chen et al. 2019), global positioning system data from taxis, buses, or metro smart-card data (Sulis et al. 2018), Wi-Fi access points (Kim 2018) and mobile signaling data (Tu et al. 2017) allow for a more nuanced portrait of human activity. Mobile signaling data is an ideal data source, but it is difficult to obtain due to privacy problems and the spatial accuracy is not high because the data are obscured by a series of confidentiality requirements, which means small-scale analysis cannot be carried out (Tang et al. 2018). Other data either do not reflect time changes or only reflect the movements of some groups. Therefore, it is still a problem to be solved which data can better represent urban vitality.

There are many factors that affect urban vitality, among which socioeconomic factors such as population and employment are often studied (Meng and Xing 2019). However, the physical built environment of the city cannot be ignored either, and some studies have proved that there is a close connection between urban vitality and the built environment. Although the choices of influencing factors vary in different researches, the measurement of the built environment is usually based on density, diversity, and design, namely ‘3Ds’ (Cervero and Kockelman 1997). Later, some scholars also proposed ‘5Ds’, adding destination accessibility and distance to transit (Ewing and Cervero 2010). Specific influencing factors include land function, mixed land use, number of intersections, building density, proportion of sky and greenery, etc. (Lu et al. 2019; Li et al. 2022; Yue et al. 2021). With the advent of the big data era, many mapping platforms provide free and open-source urban spatial information data, including POIs, road networks, architectural vectors, and street view images, researchers can more accurately quantify the built environment factors and perceive the physical and social spaces of cities from a new dimension. Therefore, we use multi-source big data to better delineate the urban built environment in this study.
How to accurately portray the relationship between urban vitality and the built environment has always been an important issue. Urban vitality has strong heterogeneity in both spatial and temporal dimensions, which means the impact of each factor on urban vitality will be different in different regions at different times. Therefore, global regression such as ordinary least squares (OLS) used in many studies is difficult to accurately reveal the spatial effect of the built environment on urban vitality (Sung and Lee 2015). As a local regression model that can better reflect spatial heterogeneity, geographically weighted regression (GWR) is a better approach that has been used more frequently in recent research (Liu et al. 2020; Fan et al. 2021; Yang et al. 2021b). Also, since the mobility patterns of the population vary greatly between weekdays and weekends, we considered both scenarios to make our analysis more targeted.

In this study, Shenzhen, a highly urbanized city, is selected as the research area in this article. Based on Baidu heat map, POIs, Open Street Map (OSM), and other open source big data, the spatio-temporal changes of Shenzhen’s urban vitality on weekdays and weekends are analyzed. GWR is used to further reveal the influence mechanism of land use, road network, buildings, and other urban built environment factors on urban vitality.

2. Study area and data

2.1. Study area

Shenzhen is located in the south of China, which is close to Hong Kong. Shenzhen was designed as China’s first special economic zone in 1980, and the city has changed dramatically over the past 40 years. In 2020, the city’s GDP reached 2.77 trillion yuan, ranking third in China, and has a permanent population of 17.56 million. Shenzhen has 9 administrative districts and 1 new district (shown in Figure 1), with a total area of 1997 km².

In this study, blocks are used as analysis units to ensure that the spatial heterogeneity of urban vitality can be accurately expressed to the maximum extent. The whole of Shenzhen was divided into 2644 blocks. The specific method for generating blocks...
is available in Supplementary Material (Section 2) and the result of the formation is shown in Figure S1.

2.2. Data sources and preprocessing

The main data used in this study include Baidu heat map, POIs, OSM road network, and gross floor area, which are mainly from open source websites and the government. The specific data types and sources are shown in Table 1.

2.2.1. Baidu heat map

Baidu heat map data comes from the location information carried by smartphone users when they visit Baidu products. By counting the location service requests initiated in a given area and making further calculations, the real-time population activity intensity of the region can be obtained. The population intensity in different areas is color-coded on the map and the heat map will refresh every 15 min. Due to concerns about user privacy, Baidu platform does not provide original user data. The data we can obtain through the open interface is the heat value of sampling points after calculation. The higher the heat value, the higher the intensity of population activity, indicating that the region is more dynamic.

As the most famous Internet company in China, Baidu has hundreds of millions of users, so the crowd coverage of Baidu map is very high (more than 130 billion location service requests every day). Compared with other data, Baidu heat map data can be obtained through open interfaces with lower acquisition difficulty and higher temporal and spatial resolution, which can more precisely reflect the characteristics of spatio-temporal dynamic distribution of population and is a reasonable data reflecting urban vitality (Yang et al. 2021a). Some studies have used the Baidu heat map to describe changes in urban vitality (Feng et al. 2019; Li et al. 2019). Therefore, we try to use Baidu heat map to measure urban vitality in this study.

Data in this study include Baidu heat map data of Shenzhen city from 2 July to 8 July 2020. The data acquisition interval was 1 h, and the heat value data sets of 168 time points were obtained. Each time point data set, including more than 20,000 sampling points. Each point of the record includes four properties, respectively for the longitude, latitude, time, and heat value.

2.2.2. POIs dataset

POIs data are very useful in accurately estimating land use classification. Compared with traditional land use data, POIs have finer statistical granularity, higher

| Data                        | Data source                  | Year |
|-----------------------------|------------------------------|------|
| Road network                | Open Street Map              | 2020 |
| Administrative boundaries   | National Geomatics Center of China | 2020 |
| and water systems           |                              |      |
| POIs data                   | Amap                         | 2020 |
| Bus and subway stations     | Amap                         | 2020 |
| Gross floor area            | Shenzhen Municipal Bureau of Planning and Natural Resources | 2020 |
| Baidu Heat Map              | Baidu map                    | 2020 |
resolution, and low cost (Jiang et al. 2015). Therefore, this study uses POIs data to reflect the land use status within the block.

The POIs data used in this study came from Amap (https://www.amap.com), one of the largest map search engines and suppliers in China (Lu et al. 2019). A total of 1,064,476 POIs of Shenzhen in 2020 were obtained from Amap. 862,306 POIs were obtained after screening, and each POIs contained latitude and longitude, name, address, category, and other information. Amap’s POIs adopt a three-level classification system, in which there are 23 first-level classifications, 267 second-level classifications, and 904 third-level classifications. According to research needs and referring to relevant studies (Zhang et al. 2019b; Xia et al. 2020), this study finally reclasses POIs into 6 categories. The number and proportion of POIs by category are provided in the Supplementary Material (Table S1).

2.2.3. Other complementary data
As one of the most popular volunteered geographic information projects, OSM has reliable quality, and the data quality of OSM in China is constantly improving, especially in large cities, which is significantly better than that in small and medium-sized cities (Jiang and Liu 2012; Brovelli et al. 2017). We obtained national road network vector data of China in May 2020 from the OSM, and extracted the data in Shenzhen city. The original road classification system of OSM was relatively complex. According to relevant studies (Liu and Long 2016) and the actual situation of Shenzhen, roads were reclassified into three levels: express road, main road, and secondary road. Urban branch roads, internal roads, and pedestrian roads were excluded. We also obtained the 1:250,000 vector data of Shenzhen city from National Geomatics Center of China and extracted the water system layer from it.

3. Methods
3.1. Block vitality measurement
3.1.1. Kernel density analysis
Kernel density analysis can estimate the intensity and visualize the distribution of points by creating a smooth surface based on a bivariate probability density function (Wu et al. 2018b; Fu et al. 2021). In this study, this method is used to transform discrete heat points into continuous surfaces that reflect the vitality distribution of the whole Shenzhen city. Furthermore, we calculate the average heat value of each block.

3.1.2. Hot spot analysis
Getis-ord Gi* hot spot analysis is a method used to identify high-value or low-value spatial clusters with statistical significance. It takes neighboring factors into consideration and measures whether spatial clusters are high-value clusters or low-value clusters by calculating Getis-ord Gi* statistics and combining z-score and p-value (Getis and Ord 2010). This method has been applied to the study of population distribution and urban vitality identification (Fang et al. 2020; Zeng et al. 2020). We use this method to identify the urban vitality pole of Shenzhen.
3.2. Vitality impact factors calculation

Considering the representativeness of factors and the accessibility of data, this study mainly includes 13 indicators. The specific factors and descriptions are shown in Table 2.

3.2.1. Mixed use

Since Jacobs, the mixed use of land has been considered important for improving the vitality of cities. Many researchers have conducted research on this issue and found that higher land-use density and land-use mix were closely related to higher vitality and increased the attractiveness of the region (Jacobs-Crisioni et al. 2014).

There are many methods to measure land mixed use, including Richness, Shannon Entropy, Simpson diversity index, etc. (Yue et al. 2017). Some researchers analyzed the effects of 14 different mixed measurement methods and recommended that Entropy is appropriate when the land use type is greater than 2 (Song et al. 2013). And this index has been widely used in current research (Huang et al. 2019). Therefore, our measure of functional mixture applies the Shannon entropy function (Shannon 1948). Generally speaking, the higher the entropy is, the higher the degree of land mixed use is. And is calculated as follows:

\[ M = -\sum_{i=1}^{n} (p_i \times \ln p_i) \]  

(1)

where \( p_i \) is the proportion of a certain type of POIs in the number of all POIs in the block.

3.2.2. Floor space ratio

Building data can be used to measure the density and strength of a building. Commonly used indicators are mainly measured from a two-dimensional or three-
dimensional perspective, including floor space ratio (FSR), ground space index, height, Open space, and other indicators (Ye et al. 2018; Li et al. 2021). Since FSR takes into account both two-dimensional and three-dimensional characteristics of buildings, we choose it to evaluate the building intensity of blocks.

3.2.3. Compactness

The physical compactness of urban space plays an important role. With compact urban space, urban transportation energy consumption will be greatly reduced and the efficiency of land use will greatly improve.

The physical characteristic indexes commonly used to measure the outer contour of urban built-up areas include Gibbs compactness, Cole compactness, and Richardson compactness (Lan et al. 2021). According to relevant theories of urban economics, Richardson et al. proved that the most efficient external space form of urban construction is circular. When the city is round, the compactness is 1. The larger the Richardson index, the more compact the spatial shape. This article refers to relevant studies (Lu et al. 2019; Meng and Xing 2019) and uses Richardson index to express the compactness of the block. And is calculated as follows:

$$C = \frac{2\sqrt{\pi*A}}{P}$$  \hspace{1cm} (2)

where $A$ is the area of the block, $P$ is the perimeter of the block.

3.3. Geographically weighted regression

Based on the idea of local smoothness and the assumption of spatial heterogeneity, Fotheringham proposed geographically weighted regression, introduced the spatial location of data into the model parameters, and used the local weighted least square method to estimate the parameters (Fotheringham et al. 2003). Thus, the variable changes with the position of space, which solves the problem of space non-stationarity and makes the result more consistent with the actual situation. The GWR model is formulated as follows:

$$y_i = \beta_0(u_i, v_i) + \sum_{j=1}^{m} \beta_j(u_i, v_i)x_{ji} + \epsilon_i$$  \hspace{1cm} (3)

where $i$ represents the spatial unit $i$, $y_i$ denotes the value of urban vitality of the spatial unit $i$, $x_{ji}$ is the $j$th built environment indicators of the spatial unit $i$, $m$ stands for the total number of spatial units, $\epsilon_i$ denotes the random error term of the spatial unit $i$, $(u_i, v_i)$ signifies the geographic coordinate of spatial spatial unit $i$ where $u_i$ is the longitude of unit $i$ and $v_i$ is the latitude of unit $i$, $\beta_0(u_i, v_i)$ stands for the intercept at the location $i$, and $\beta_j(u_i, v_i)$ represents the local estimated coefficient of the built environment variable $x_{ji}$.

In this study, the adaptive Gaussian kernel function is used to model geographic weights. The best kernel bandwidth is set according to the corrected Akaike information criterion (AICs).
4. Results and discussion

4.1. Spatial pattern of urban vitality

4.1.1. Spatial distribution and variation

Figure 2 shows the average vitality distribution on weekdays and weekends. The natural breaks (Jenks) method is used to classify the block vitality into five categories based on weekdays’ vitality value: highest (759–1388), higher (510–758), medium (326–509), lower (161–325), and lowest (0–160).

On the whole, the spatial distribution of urban vitality shows a pattern of three centers and two sub-centers. Weekday highest vitality blocks are mainly distributed in the south (Futian district and Luohu District), southwest (Nanshan District), and middle (Longhua District) part of Shenzhen, forming three vitality centers (black circles in Figure 2(a)). At the same time, two sub-centers with higher vitality were formed in the northwest (Baoan District) and northeast (Longgang District), respectively (red circles in Figure 2(a)). And the sub-center in the northwest is has larger higher vitality area than that in the northeast.

The vitality distribution is restricted by the natural terrain conditions of Shenzhen city, and different vitality centers are separated by low vitality areas, which are in blue. Lowest vitality blocks are mainly distributed in the eastern part of Shenzhen, which is mainly due to the large distribution of natural mountains, so these areas are sparsely populated and undeveloped. The other lowest vitality blocks formed in the western part are also due to the distribution of natural mountains and water bodies such as Tiegang Reservoir, Xili Reservoir, Yangtai Mountain Forest Park, and Tanglang Mountain Park. These natural areas hinder the further integration of the three vitality centers.

The overall vitality distribution pattern of weekends is similar to that of weekdays (Figure 2(b)). But the number of highest vitality and higher vitality blocks decreases, especially in the two vitality centers of the south and southwest (black circles in Figure 2(b)), which are the CBD of the city with a large number of businesses located. However, there were slight increase of the higher and medium vitality blocks in two sub-centers.

Table 3 shows the number and proportion of blocks with different vitality types. In terms of quantity, the number of lowest vitality blocks is the largest, accounting...
for nearly 30%. The number of highest vitality blocks was the lowest, less than 5%. There are some differences in the amount of block vitality between weekdays and weekends. On weekdays, compared with weekends, the number of highest and higher vitality blocks is more, while the number of medium, lower and lowest vitality blocks is less. In particular, the difference between the number of highest vitality blocks is significant. From the area point of view, the area of highest and higher vitality blocks accounts for a small proportion of about 5%, while the area of lowest vitality blocks accounts for nearly 70%. This is mainly because the high vitality area in the downtown area is usually covered with road network, so the block area is small, while the low vitality area is usually the mountain, water, and other natural areas, and the area is large. Therefore, the area difference of different vitality blocks is greater than the quantity difference.

### 4.1.2. Hot and cold spots

Hot spot analysis can be used to identify the spatial distribution of high vitality and low activity clusters, which represent a highly concentrated or sparsely distributed population. Figure 3(a) shows that there are four high vitality clusters on weekdays, which are located in the northwest, southwest, south-central, and central regions, among which the southern and central regions form contiguous hot spots. This distribution pattern is basically consistent with the distribution of vitality centers and sub-centers mentioned above, but it can be noted that compared with the northwest sub-center (red circle in Figure 2(a)), the northeast sub-center (yellow circle in Figure 2(a)) does not form hot spots. A large area of low vitality agglomeration is formed in the east of Shenzhen, and a long and narrow low vitality agglomeration is formed on the west coast.

Figure 3(b) shows that the distribution of weekends is roughly the same as that of weekdays, which indicates the relative stability of Shenzhen’s vitality pole. However, on weekends, the hot spot agglomeration area in the northwest sub-center (red circle in Figure 2(b)) increased significantly, and the hot spot agglomeration area in central Shenzhen also spread to the surrounding area. Meanwhile, some cold spot agglomeration areas in the northeast sub-center (yellow circle in Figure 2(b)) decreased significantly, which was transformed into an insignificant area. This may be because Baoan district in the northwest and Longgang district in the northeast are home to a large number of people who work in other areas of the city center, and these people, by not having to go to work on weekends, are active in the vicinity of their residence, thus enhancing the vitality of these areas. On the whole, the concentration of vitality

### Table 3. The number and proportion of blocks with different vitality types.

| Vitality type | Weekdays | | | Weekends | | |
|---------------|----------|----------|----------|----------|----------|
|               | Counts   | Percentage of count | Percentage of area | Counts   | Percentage of count | Percentage of area |
| Highest       | 125      | 4.73     | 1.06     | 73       | 2.76     | 0.59     |
| Higher        | 437      | 16.53    | 4.65     | 401      | 15.17    | 4.42     |
| Medium        | 593      | 22.43    | 9.11     | 624      | 23.60    | 9.40     |
| Lower         | 710      | 26.85    | 18.36    | 756      | 28.59    | 18.06    |
| Lowest        | 779      | 29.46    | 66.82    | 790      | 29.88    | 67.53    |
on weekdays is smaller, mainly because the distribution of population on weekdays is more concentrated in working places, while the distribution of population on weekends is more dispersed because people have more diverse activities thus not travel regularly.

4.2. Temporal changes of urban vitality

Calculate the average vitality of 24 h for each block. Figure 4 shows the vitality distribution of 8 time points in 24 h, at an interval of 3 h. It can clearly show the spatial changes of vitality during a day.

By counting the number of different types of blocks every hour with a cycle of 24 h, Figure 5 shows the changing trend of the number of different types of blocks. The change of block vitality within a day is closely related to human activity patterns. On weekdays, the number of highest vitality blocks forms two peaks at 13:00 and 19:00. The number of lowest vitality blocks increases rapidly after 23:00 and peaks at 6:00. After 6:00, the number of lowest vitality blocks decreases rapidly, and highest vitality blocks appear at around 9:00, which is consistent with people’s working time.

The change in the number of blocks of different types on weekends is moving backward in the time dimension. The time of the emergence of highest vitality and higher vitality block is obviously late, higher vitality block appeared at 9:00, highest vitality block appeared at 11:00. At the same time, there were not two peaks but only one peak at 19:00 in the highest vitality blocks. Activity declines more slowly on weekends than on weeknights, and there are fewer lowest vitality blocks after 23:00.

4.3. Results of regression models

In order to eliminate the influence of different data ranges and units of different variables in the analysis and improve the comparability of data, the zero-mean normalization method was adopted to standardize the variables (Gong, Lin, & Duan, 2017).

Since collinearity may exist between different factors, we constructed 2 stepwise multiple regression models to avoid collinearity and select variables of significance before building the GWR models (Wu et al., 2018a; Fu et al., 2021). Finally, 7 independent variables were retained on weekdays and weekends respectively. The analysis
Figure 4. (a) Block vitality at 0:00; (b) Block vitality at 3:00; (c) Block vitality at 6:00; (d) Block vitality at 9:00; (e) Block vitality at 12:00; (f) Block vitality at 15:00; (g) Block vitality at 18:00; (h) Block vitality at 21:00.
results are shown in Table 4. The analysis results are shown in Table 4. Then we put these variables into GWR models and Table 5 represents the coefficients of each independent variable in the GWR model, respectively. The GWR model parameter test results obtained show that the adjusted $R^2$ of the weekday model is 0.626, and that of
the weekends is 0.589, both of which are significantly improved compared with the stepwise regression models.

By further visualizing the coefficients of each independent variable in the GWR model, Figure 6 shows the coefficient changes of different influencing factors in different regions.

4.3.1. Distance to subway stations

Theoretically, the closer the distance to the subway station is, the higher the accessibility is and the higher the urban vitality is. Figure 6(a) shows that this relationship exists in most areas, especially in the east part of Shenzhen (blue area in Figure 6(a)). This shows that the planning and construction of subway lines in these areas can effectively improve the vitality of blocks, especially for the northeast sub-center, where only one line passes through at present. The new line will help new development areas in the east better integrate with the urban core in the west.

However, in some of the areas in the south, southwest, northwest, and middle part of Shenzhen, subway stations did not play a positive role (red area in Figure 6(a)). As the center of the city, the southern area of Shenzhen, mainly the Futian district (blue circle in Figure 6(a)), has the densest subway stations, but it has no positive effect on the vitality of the city. This may mean that denser subway stations are not absolute better when a city develops to a certain stage. In addition, in the southwest region, subway stations have significant positive and negative effects on urban vitality at the

### Table 4. Stepwise multiple regression results of urban vitality on weekdays and weekends.

| Variable | Weekdays |  |  |  | Weekends |  |  |  |
|----------|----------|---|---|---|----------|---|---|---|
|          | Estimate | Standard error | t |  | Estimate | Standard error | t |  |
| Constant | 323.17*** | 3.87 | 83.44 |  | 305.58*** | 3.63 | 84.17 |  |
| DSS      | −33.87*** | 4.22 | −8.03 |  | −33.40*** | 3.94 | −8.47 |  |
| ROD      | 22.44*** | 4.63 | 4.84 |  | 14.33*** | 4.37 | 3.28 |  |
| RI       | 17.10*** | 4.23 | 4.04 |  | 16.57*** | 3.89 | 4.26 |  |
| MU       | 19.09*** | 4.42 | 4.32 |  | 17.47*** | 4.33 | 4.03 |  |
| RD       | 25.29*** | 4.26 | 5.93 |  | 28.13*** | 4.03 | 6.98 |  |
| TD       | 56.06*** | 5.26 | 10.65 |  | 39.37*** | 5.51 | 7.15 |  |
| FSR      | 26.25*** | 4.94 | 5.31 |  | – | – | – |  |
| CD       | – | – | – |  | 24.78*** | 5.01 | 4.95 |  |
| $R^2$    | 0.274 |  |  |  | 0.238 |  |  |  |
| Adjusted $R^2$ | 0.272 |  |  |  | 0.236 |  |  |  |

***Represents significance level of 1%.

---

### Table 5. Regression results of urban vitality by GWR model on weekdays and weekends.

| Variable | Weekdays |  |  |  | Weekends |  |  |  |
|----------|----------|---|---|---|----------|---|---|---|
|          | Mean    | STD | Min | Max | Range | Mean | STD | Min | Max | Range |
| DSS      | 9.51    | 145.04 | −257.26 | 663.20 | 920.46 | 14.00 | 143.28 | −196.97 | 592.52 | 789.50 |
| ROD      | 0.26    | 27.37 | −65.39 | 77.28 | 142.67 | −0.60 | 27.05 | −65.09 | 89.98 | 155.07 |
| RI       | 10.22   | 27.37 | −61.26 | 102.30 | 163.56 | 8.81 | 23.72 | −67.13 | 81.70 | 148.83 |
| MU       | 3.60    | 22.33 | −66.15 | 64.54 | 130.68 | 7.49 | 19.90 | −50.00 | 76.11 | 126.11 |
| RD       | 24.56   | 36.82 | −42.59 | 136.94 | 179.52 | 21.21 | 35.94 | −50.60 | 145.63 | 196.23 |
| TD       | 35.08   | 36.28 | −67.46 | 119.24 | 186.70 | 24.23 | 30.73 | −70.15 | 120.86 | 191.01 |
| FSR      | 17.64   | 35.70 | −65.08 | 135.98 | 201.06 | – | – | – | – | – |
| CD       | – | – | – |  | 28.45 | 32.32 | −40.33 | 139.59 | 179.92 |  |
| $R^2$    | 0.662 |  |  |  | 0.627 |  |  |  |
| Adjusted $R^2$ | 0.626 |  |  |  | 0.589 |  |  |  |
Figure 6. Spatial characteristics of estimated coefficients of built environment variables on weekdays: (a) The coefficient of distance to subway stations; (b) The coefficient of road density; (c) The coefficient of compactness; (d) The coefficient of mixed use; (e) The coefficient of residential density; (f) The coefficient of traffic density; (g) The coefficient of floor space ratio.
same time, indicating that the role of subway stations may need to be in conjunction with other factors.

Central and northwest areas in red which are mainly the Baoan district and Longhua district are closely connected with the city center in the south (black circles in Figure 6(a)). A large number of people in these areas work in the two city centers in the south and tend to live near the subway stations. Therefore, subway lines may transport people to the city center, resulting in the decline of the vitality of the region near subway stations. On the other hand, if compared with the coefficient distribution of bus stations, the distribution of bus stations in these regions has a significant positive effect. This may be related to the travel mode of the regional population. For residents in these areas who mainly live and work locally, they tend to take buses rather than take the subway. Therefore, urban vitality is positively correlated with the distribution of bus stations in these areas.

From the analysis above we can conclude that the setting of subway stations needs to consider the local economic development stage and the local population’s commuting pattern and travel preference.

4.3.2. Road density

Figure 6(b) shows that the increase of road network density has a significant positive effect on the block vitality in the northwest, northeast, and southwest corner of the western region (blue circles in Figure 6(b)). These areas are just the periphery of the central area of each district, which are under development, indicating that higher road network density has a positive effect in the stage of urban expansion.

In the three centers of Shenzhen, the earliest developed areas, the increase of road network density has no positive effect (red circles in Figure 6(b)). This is significantly different from previous studies (Tu et al. 2020). Given that previous studies used data from 5 or 6 years ago, it is possible that as the road network density in core urban areas continues to increase, more roads mean more noise and pollution, which in turn reduces the attractiveness of blocks to residents. This means that when the city develops to a certain stage, the density of the road network needs to be controlled within a reasonable range. As for the northeastern area of Shenzhen, which has developed rapidly in recent years, road network density also has no positive effect (black circle in Figure 6(b)). On the one hand, it may be because this area is far away from the center of the city, and the population mainly relies on the subway rather than cars for commuting. More dense roads do not attract more people from other areas. On the other hand, it may be because this area has a large natural area. A denser road network does not mean greater development intensity or human activities.

4.3.3. Compactness

Compactness often means higher land use efficiency and is positively correlated with urban vitality. Compactness plays a positive role in most areas of Shenzhen, especially in the urban core area in the south (blue circle in Figure 6(c)). As for the area showing a negative relationship, the western area along the coast is mainly composed of airports and undeveloped areas, and the northeast sub-center area has a large number of development parks (black circles in Figure 6(c)). The development mode of these
areas is exactly the opposite of compact development, so compactness cannot be the explanatory factor for the vitality of these areas.

4.3.4. Mixed use
Land use diversity is considered to have a positive effect on urban vitality, but the distribution of this effect has no obvious rule according to the analysis results (Figure 6(d)). For example, the mixed use of land in the south and middle part of Shenzhen has completely opposite effects, though these areas are urban centers. This indicates that land mixed use needs to be combined with the actual situation of different regions and consider a synergistic effect with other influencing factors, rather than taking it as a rule that applies to all regions.

4.3.5. Residential density
According to Figure 6(e), residential density has a positive effect in most areas of Shenzhen, especially in the northeast and northwest regions. This means that more new housing in these areas will have a positive effect on the vitality of the city. Some areas with no positive effect are mainly natural areas, such as a large number of mountains and reservoirs in the north and the east of Shenzhen. However, there is no positive correlation in some areas in the southwest, which may be because this region, as the central area of Shenzhen, has a large number of people working in this area from other areas every day. Therefore, the vitality of this area is not mainly determined by local residents.

4.3.6. Traffic density
The density of traffic facilities has an obvious positive effect on the whole of Shenzhen, except for Shenzhen Airport, Phoenix Mountain, Nanshan Park, and Pinghu Park, which are in blue in Figure 6(f). In other areas, the higher the density of traffic facilities, the higher the vitality of the city. Although bus and subway stations are included in the traffic facilities in this study, 80% of the points are parking lots. Therefore, parking lots can be considered as a reliable predictor of urban vitality. Parking lots reflect where people end up or where they will stay for a relatively long time. Therefore, we can speculate that in an era when people increasingly rely on cars to travel, people may be more inclined to choose places with convenient parking services as travel destinations. We can also explain it in another way that dynamic places attract a large number of people, thus leading to more parking lots. So parking lots may be a cause or a reason for the vitality, but there is definitely a relationship between the two, which means it can still be considered as a reliable predictor of urban vitality. What we want to emphasize is that we need to pay more attention to the comprehensive consideration of data analysis results and realistic logic. For example, it is important to note that we cannot simply interpret the result as building more parking lots in an area will definitely improve the vitality of one area.

4.3.7. Floor space ratio
On weekdays, higher building intensity has a positive effect on urban vitality as a whole (shown in Figure 6(g)), but it should be noted that it doesn’t mean the higher
building intensity is, the better is. In some areas which are in blue in Figure 6(g), there have been negative effects, which means that building intensity needs to be controlled within a reasonable range and coordinated with local development conditions.

### 4.4. Comparison between weekdays and weekends

The coefficient distribution of most factors on weekends is similar to that on weekdays, but there are differences in floor space ratio, commercial density, and mixed use.

On weekends, floor space ratio did not become a significant factor, it is probably because weekday office is often high vitality region, and office areas generally have a higher intensity of construction, so on weekdays, construction is closely related to the vitality intensity. But on weekends, office areas without people working will cause a decrease in vitality, so the construction strength is hard to be the interpretation of the vitality factors.

The density of commercial services shows significance only on weekends (shown in Figure 7(a)). Part of the northeast, western, and northwest areas has significant positive effects (black circles in Figure 7(a)). A large percentage of the population in these areas needs to go to work in the city center on weekdays, so the demand for commercial services is not large. However, these people stay in the local area on weekends and have more demands for leisure, entertainment, and life services. Therefore, commercial service facilities in these areas are strongly correlated with the vitality of the block. However, it can be expected that with the increase of local jobs in these regions and the balanced development between regions, this model may disappear in the future.

The coefficient difference of mixed use is mainly reflected in the sub-center in the northeast (blue circles in Figure 7(b)). The vitality of this area is almost completely positively correlated with diversity on weekends, while diversity has almost no positive effect on weekdays. This phenomenon may be the same as that of commercial services, where more people stay in the area on weekends rather than going to work out of the area, so diverse areas tend to attract more people. The area with a positive

![Figure 7](image_url). Spatial characteristics of estimated coefficients of built environment variables on weekends: (a) The coefficient of commercial density; (b) The coefficient of mixed use.
effect of diversity also increased in the northwest regions on weekends, which also had a large number of people working in city center areas. On the other hand, the positive area of mixed use decreased in the southwest during the weekends. This may mean that diversity alone is not enough to keep a block alive, but needs to be supported by enough people moving around the blocks.

5. Conclusion

With all kinds of new data, we have more ways to accurately quantify the vitality of cities. This study uses Baidu heat map data to measure the urban vitality of Shenzhen city and explores the spatial distribution pattern of Shenzhen city’s urban vitality and the rule of vitality changing with time within a day. On the whole, the high vitality area is consistent with the urban core area planned by Shenzhen city, showing a vitality pattern of three centers and two sub-centers. But the concentration degree of the two sub-centers in the northeast and northwest should be strengthened. In particular, it is necessary to promote the integration of the new development areas in the east and the original urban core areas in the west. The vitality changes of the blocks in 24 h are closely related to the patterns of human activity. In addition, we further compare the vitality difference between weekdays and weekends. Compared with the stepwise multiple model, the GWR model can better reflect the spatial heterogeneity, so we build the GWR model to analyze the influence of built environment factors on the urban vitality and found that the distance from the subway station, road density, compactness, mixed use, residential density, and traffic density have significant effect, but the effects of different factors vary from block to block. Meanwhile, commercial facility density and floor space ratio showed significance only on weekends and weekdays, respectively. Some of the factors thought to be influential did not show significance in the model such as block area and public service density. Overall, urban vitality remains a complex issue. Therefore, how to improve the vitality of the city may be better suited for small scales, such as block scale, to make targeted planning recommendations, rather than trying to find a rule that applies to all areas.

There are still limitations in our research. First of all, although urban big data provides us with a new perspective to quantify vitality, different data sources may produce different results (Xia et al. 2020), so data reliability is very important. A single data source is likely to produce deviations, so it is necessary to consider the comprehensive vitality measurement of a variety of data to make up for the defects of different data to the greatest extent. In addition, the criteria of block identification still need some further study. For example, adopting a different road grade division system, selecting different road buffer distances, or taking three-dimensional characteristics of blocks into consideration will affect the number or size of the blocks and ultimately affect the analysis results (Zhang et al. 2019a, 2022). At the same time, the number of blocks generated is large because we want to reflect the vitality differences of the blocks in a more detailed way, which makes our analysis can’t focus on single blocks. If more detailed case analysis can be conducted on some representative blocks in the future, such as some blocks in the Longgang district or Baoan district mentioned in our study, researchers may obtain richer conclusions. What’s more, there may be interactions between built environment
factors, so the synergy between different factors should be considered in future studies, which may help to develop more comprehensive and effective policies.

**Funding**

This research was funded by the National Natural Science Foundation of China (Grant No. 42007194); and the open fund by Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control (KHK1907).

**Data availability statement**

The data that support the findings of this study are available from the corresponding author, Sheng Zheng, upon reasonable request.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**References**

Azmi DI, Karim HA. 2012. Implications of walkability towards promoting sustainable urban neighbourhood. Procedia Soc Behav Sci. 50:204–213.

Batty M. 2016. Empty buildings, shrinking cities and ghost towns. Environ Plann B Plann Des. 43(1):3–6.

Brovelli MA, Minghini M, Molinari M, Mooney P. 2017. Towards an automated comparison of OpenStreetMap with authoritative road datasets. Trans in GIS. 21(2):191–206.

Cervero R, Kockelman K. 1997. Travel demand and the 3Ds: density, diversity, and design. Transp Res D Transp Environ. 2(3):199–219.

Chen T, Hui EC, Wu J, Lang W, Li X. 2019. Identifying urban spatial structure and urban vibrancy in highly dense cities using georeferenced social media data. Habitat Int. 89:102005.

Ewing R, Cervero R. 2010. Travel and the built environment. J Am Plan Assoc. 76(3):265–294.

Fan Z, Duan J, Luo M, Zhan H, Liu M, Peng W. 2021. How did built environment affect urban vitality in urban waterfronts? A case study in Nanjing reach of Yangtze river. Int J Geo-Inf. 10(9):611.

Fang L, Huang J, Zhang Z, Nitivattananon V. 2020. Data-driven framework for delineating urban population dynamic patterns: case study on Xiamen island, China. Sustain Cities Soc. 62:102365.

Feng D, Tu L, Sun Z. 2019. Research on population spatiotemporal aggregation characteristics of a small city: a case study on Shehong county based on Baidu heat maps. Sustainability. 11(22):6276.

Fotheringham AS, Brunsdon C, Charlton M. 2003. Geographically weighted regression: the analysis of spatially varying relationships. Chichester, England:John Wiley & Sons.

Fu R, Zhang X, Yang D, Cai T, Zhang Y. 2021. The relationship between urban vibrancy and built environment: an empirical study from an emerging city in an arid region. Int. J. Environ. Res. Public Health. 18(2):525.

Getis A, Ord JK. 2010. The analysis of spatial association by use of distance statistics. Perspectives on spatial data analysis. New York, NY:Springer; p. 127–145.

Gong Y, Lin Y, Duan Z. 2017. Exploring the spatiotemporal structure of dynamic urban space using metro smart card records. Comput Environ Urban Syst. 64:169–183.
Huang B, Zhou Y, Li Z, Song Y, Cai J, Tu W. 2019. Evaluating and characterizing urban vibrancy using spatial big data: Shanghai as a case study. Environ Plann B Urban Anal City Sci. 47(9):1543–1559.

Jacobs J. 1961. The death and life of great American cities. New York, NY: Vintage Book Company.

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. Environ Plan A. 46(11):2769–2785.

Jiang S, Alves A, Rodrigues F, Ferreira J, Pereira FC. 2015. Mining point-of-interest data from social networks for urban land use classification and disaggregation. Comput Environ Urban Syst. 53:36–46.

Jiang B, Liu X. 2012. Scaling of geographic space from the perspective of city and field blocks and using volunteered geographic information. Int J Geog Inf Sci. 26(2):215–229.

Kim Y-L. 2018. Seoul’s Wi-Fi hotspots: Wi-Fi access points as an indicator of urban vitality. Comput Environ Urban Syst. 72:13–24.

Lan T, Shao G, Xu Z, Tang L, Sun L. 2021. Measuring urban compactness based on functional characterization and human activity intensity by integrating multiple geospatial data sources. Ecol Indic. 121:107177.

Lang W, Chen T, Li X. 2016. A new style of urbanization in China: transformation of urban rural communities. Habitat Int. 55:1–9.

Li X, Li Y, Jia T, Zhou L, Hijazi IH. 2022. The six dimensions of built environment on urban vitality: Fusion evidence from multi-source data. Cities. 121:103482.

Li M, Liu J, Lin Y, Xiao L, Zhou J. 2021. Revitalizing historic districts: identifying built environment predictors for street vibrancy based on urban sensor data. Cities. 117:103305.

Li J, Li J, Yuan Y, Li G. 2019. Spatiotemporal distribution characteristics and mechanism analysis of urban population density: a case of Xi’an, Shaanxi, China. Cities. 86:62–70.

Liu X, Long Y. 2016. Automated identification and characterization of parcels with OpenStreetMap and points of interest. Environ Plann B Plann Des. 43(2):341–360.

Liu S, Zhang L, Long Y, Long Y, Xu M. 2020. A new urban vitality analysis and evaluation framework based on human activity modeling using multi-source big data. Int J Geo-Inf. 9(11):617.

Long Y, Huang CC. 2017. Does block size matter? The impact of urban design on economic vitality for Chinese cities. Environ Plann B Urban Anal City Sci. 46(3):406–422.

Lu S, Shi C, Yang X. 2019. Impacts of built environment on urban vitality: regression analyses of Beijing and Chengdu, China. Int J Environ Res Public Health. 16(23):4592.

Meng Y, Xing H. 2019. Exploring the relationship between landscape characteristics and urban vibrancy: a case study using morphology and review data. Cities. 95:102389.

Montgomery J. 1998. Making a city: urbanity, vitality and urban design. J Urban Des. 3(1):93–116.

Shannon CE. 1948. A mathematical theory of communication. Bell Syst Tech J. 27(3):379–423.

Song Y, Merlin L, Rodriguez D. 2013. Comparing measures of urban land use mix. Comput Environ Urban Syst. 42:1–13.

Sulis P, Manley E, Zhong C, Batty M. 2018. Using mobility data as proxy for measuring urban vitality. J Spat Inf Sci. 16(16):137–162.

Sung H, Lee S. 2015. Residential built environment and walking activity: empirical evidence of Jane Jacobs’ urban vitality. Transp Res D Transp Environ. 41:318–329.

Tang L, Lin Y, Li S, Li S, Li J, Ren F, Wu C. 2018. Exploring the influence of urban form on urban vibrancy in shenzhen based on mobile phone data. Sustainability. 10(12):4565.

Tu W, Cao J, Yue Y, Shaw SL, Zhou M, Wang Z, Chang X, Xu Y, Li Q. 2017. Coupling mobile phone and social media data: a new approach to understanding urban functions and diurnal patterns. Int J Geogr Inf Sci. 31(12):2331–2358.

Tu W, Zhu T, Xia J, Zhou Y, Lai Y, Jiang J, Li Q. 2020. Portraying the spatial dynamics of urban vibrancy using multisource urban big data. Comput Environ Urban Syst. 80(101428):101428.
Wang XR, Hui ECM, Choguill C, Jia SH. 2015. The new urbanization policy in China: which way forward? Habitat Int. 47:279–284.
Wu J, Ta N, Song Y, Lin J, Chai Y. 2018b. Urban form breeds neighborhood vibrancy: a case study using a GPS-based activity survey in suburban Beijing. Cities. 74:100–108.
Wu C, Ye X, Ren F, Du Q. 2018a. Check-in behaviour and spatio-temporal vibrancy: an exploratory analysis in Shenzhen, China. Cities. 77:104–116.
Xia C, Yeh AGO, Zhang A. 2020. Analyzing spatial relationships between urban land use intensity and urban vitality at street block level: a case study of five Chinese megacities. Landsc Urban Plann. 193:103669.
Yang J, Cao J, Zhou Y. 2021a. Elaborating non-linear associations and synergies of subway access and land uses with urban vitality in Shenzhen. Transp Res A Policy Pract. 144:74–88.
Yang Y, Ma Y, Jiao H. 2021b. Exploring the correlation between block vitality and block environment based on multisource big data: taking Wuhan city as an example. Land. 10(9):984.
Ye Y, Li D, Liu X. 2018. How block density and typology affect urban vitality: an exploratory analysis in Shenzhen, China. Urban Geogr. 39(4):631–652.
Yue W, Chen Y, Thy PTM, Fan P, Liu Y, Zhang W. 2021. Identifying urban vitality in metropolitan areas of developing countries from a comparative perspective: Ho Chi Minh City versus Shanghai. Sustain Cities Soc. 65:102609.
Yue Y, Zhuang Y, Yeh AGO, Xie JY, Ma CL, Li QQ. 2017. Measurements of POI-based mixed use and their relationships with neighbourhood vibrancy. Int J Geogr Inf Sci. 31(4):658–675.
Zeng P, Wei M, Liu X. 2020. Investigating the spatiotemporal dynamics of urban vitality using bicycle-sharing data. Sustainability. 12(5):1714.
Zhang Y, Li Q, Tu W, Mai K, Yao Y, Chen Y. 2019b. Functional urban land use recognition integrating multi-source geospatial data and cross-correlations. Comput Environ Urban Syst. 78:101374.
Zhang A, Li W, Wu J, Lin J, Chu J, Xia C. 2021. How can the urban landscape affect urban vitality at the street block level? A case study of 15 metropolises in China. Environ Plann B Urban Anal City Sci. 48(5):1245–1262.
Zhang A, Xia C, Chu J, Lin J, Li W, Wu J. 2019a. Portraying urban landscape: a quantitative analysis system applied in fifteen metropolises in china. Sustain Cities Soc. 46:101396.
Zhang A, Xia C, Li W. 2022. Relationships between 3D urban form and ground-level fine particle matter at street block level: Evidence from fifteen metropolises in china. Build Environ. 211:108745.