Multiscale Analysis of Urban Vitality Based on Social Media Big Data: A Case Study of Shanghai

Zhiyan Wang

College of Marine Science and Ecological Environment, Shanghai Ocean University, Shanghai, 201306, China

Email: W_zhiyan2002@163.com

Abstract: As urban regionalization continues to deepen, urban vitality has become a crucial metric for assessing the quality of urban spatial morphology. This study focuses on a comprehensive multiscale analysis of urban vitality in Shanghai. Using techniques such as hotspot analysis (Getis-Ord Gi*) and kernel density estimation, we aim to evaluate the distribution and variations of urban vitality in Shanghai from both temporal and spatial perspectives. We compare urban vitality hotspots during weekdays and weekends to gain a deeper understanding of their patterns. The results indicate that from July to September 2022, Shanghai's overall urban vitality was concentrated in the central and eastern regions, exhibiting a spatial distribution characterized by high aggregation and low dispersion. High urban vitality values were predominantly observed in the core areas, particularly in Huangpu, Jing'an, Xuhui, Hongkou, and Changning districts. Interestingly, vitality hotspots demonstrated a diffused distribution on weekdays and weekends. During the study period, Shanghai exhibited higher urban vitality intensity on weekdays compared to weekends. The trend was marked by an initial decrease followed by a gradual increase across the 24 time periods. Notably, night-time vitality intensity was significantly higher than daytime, peaking between 21:00 and 23:00. Spatial vitality in Shanghai's core area exhibited a distinct clustered distribution during the specified study period, closely aligned with renowned landmarks and vibrant pedestrian streets. Additionally, during prescribed time periods, the evening witnessed a higher concentration of vitality compared to daytime. In conclusion, this study provides valuable insights into the distribution and variations of urban vitality in Shanghai from both temporal and spatial perspectives. The findings highlight the importance of considering urban vitality in urban planning and development efforts to enhance the quality of urban life.

Keywords: Urban vitality, Social media check-in data, Multi-scale analysis, Shanghai City

1. Introduction

Research on urban vitality originated from the urban vitality theory proposed by Jane Jacobs in 1961, which defined "urban vitality" as "street life over a 24-hour period" and pointed out that urban vitality generally refers to the ability to induce commercial activities and human activities [1]. In the subsequent studies, scholars expanded the concept of "urban vitality" from the macro population of the street to the space utilization of the street residents and the social activities carried out by the residents [2-5], which shows that the research object of urban vitality is a complex and huge economic, social, technological and cultural system with human beings as

ICBBEM 2024, March 29-31, Wuhan, People's Republic of China

DOI 10.4108/eai.29-3-2024.2347409

the main body. The interweaving of people's behaviors and activities with the places that carry them will generate and stimulate the vitality of urban space [6]. Meanwhile, the vitality of urban space is also affected by economic, social, cultural and other factors, which is reflected in the interaction between residents and various elements in the built environment [10]. Traditional research methods on "urban vitality" usually analyze the spatial characteristics of small and medium-sized urban areas, such as urban streets, communities, etc. [7], but the traditional methods have the limitations of small research space and small data volume, and it is difficult to fully reflect the overall spatial and temporal vitality characteristics of the city. Therefore, with the advent of the era of big data and the proposal of smart city construction, spatio-temporal big data, as a new type of geographic information data with the three data characteristics of time, space and attribute, can perceive the multi-dimension of the city from multiple perspectives, realize the large-scale and high-frequency collection of spatial dynamics during the distribution of urban population, broaden the spatial scope of traditional urban vitality research and increase the time scale. Xinyi Niu et al. summarized the origin, technological evolution, and prospect of current research topics of spatial-temporal big data in urban planning research by using urban activities and urban spatial clues. Starting from the relationship between spatio-temporal big data and urban planning, it is believed that spatio-temporal big data records large-scale urban activities under spatial constraints and provides a new data basis and new technology to support the research on the relationship between urban space and urban activities [8]. As a kind of spatio-temporal big data, social media data has the characteristics of large amount of data, user-centered, real-time data analysis function, etc., which can more intuitively reflect the vitality of the city. Visualization of spatio-temporal big data based on social media through GIS can more accurately and macrographically demonstrate the relationship between civic activities and urban vitality.

At present, domestic and foreign researches on urban vitality mainly focus on spatial and temporal big data such as mobile phone signaling data [9, 11], shared bicycle data [14] and social media data [15], and also combine spatial and temporal big data such as catering and transportation to analyze urban vitality from the perspective of multi-source data [16-19]. When studying the image factors of "urban vitality", geographical weighted regression model [11], data-driven method [12], Jacobs urban vitality theory [13] and other methods are mostly used.

It can be seen from the above studies that most studies on urban vitality take population density as an important index to measure the intensity of urban vitality. Therefore, taking Shanghai as an example, this study selects Weibo check-in data from July to September 2022, takes check-in density and the number of check-ins in each district of Shanghai as measurement indicators, divides the data into working days and rest days, and divides the data into time periods. Based on spatial analysis models such as hot spot analysis and nuclear density analysis, this study visualizes the spatial-temporal change characteristics of urban vitality, and evaluates its urban vitality. The paper further analyzes the factors that affect the vitality of Shanghai city, and compares the difference between the hot spots of urban vitality and the factors that affect the vitality of Shanghai city on weekdays and weekends.

2. Materials and Methods

2.1. Datasets and Pre-processing

The data in this study comprises two parts. The first part consists of social media check-in data, which is obtained through crawler technology. Specifically, the data selected for this paper pertains to microblog sign-ins in Shanghai. The second part of the data includes street vector data of Shanghai, which was downloaded from the website of China Satellite Resources Center and processed using ArcGIS software.

Regarding the first part, the social media check-in data is collected by accessing Weibo's location service dynamic application code interface through a crawler. This dataset encompasses check-in information within Shanghai's urban area from July to September 2022, including coordinates, timestamps, and user details. To ensure accuracy and reliability, we conducted data cleaning based on existing literature methods [20]. This involved excluding sign-in records outside Shanghai city limits and retaining only the initial record if a user checked in multiple times at the same location within 0.5 hours. Ultimately, we obtained POI (Point of Interest) data consisting of 166,925 sign-in records (Fig.1).

In this study, we focused on Shanghai's urban area as our research scope and utilized check-in density as an indicator to measure urban vitality [15]. When calculating check-in density, we established spatial units measuring 1km×1km with a total count of 6411 units; time units were set at one hour per unit with a total count of 24 units per day. By comparing working days and weekends' check-in densities in Shanghai, we aimed to analyze spatial and temporal characteristics related to overall vitality distribution during these periods.

Considering the complexity of Shanghai urban area, Shanghai urban area is divided into core area and non-core area according to the check-in density, which is convenient for further analysis of the spatial and temporal distribution characteristics of Shanghai urban vitality. In this paper, the top five administrative districts with the highest check-in density were selected as the core areas of Shanghai, while the rest were non-core areas.

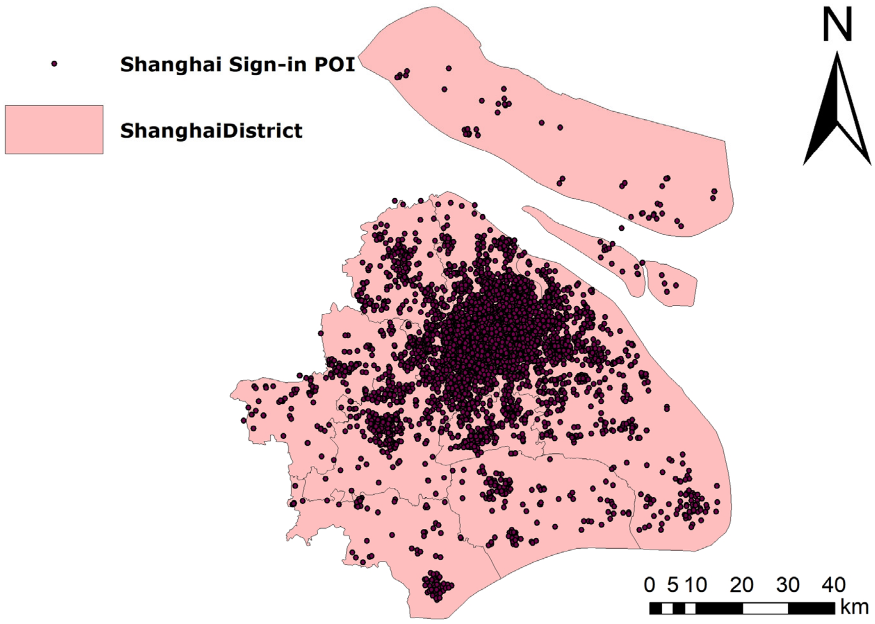


Figure 1: Weibo sign-in POI data of Shanghai urban area from July to September 2022.

2.2. Methods

2.2.1 Urban vitality evaluation index

In this study, crowd check-in density is selected as the metric for measuring urban vitality characteristics [15], with calculation methods outlined in Table 1. Specifically, district-level check-in density is utilized for core area selection, while street-level crowd check-in density is employed to assess both core area vitality and overall vitality of Shanghai.

Table 1: Description of characteristic values of urban vitality.

characteristic value	Description	formula mode
Sign-in Density of Area	Weekday Sign-in density of Area (per/h·km2)	$\text{Sign}-\text{in density of Area} = \frac{\text{District Sign}-\text{in Volume}}{\mathit{District\ Area}}$
	Weekend Sign-in density of Area (per/h·km2)	
Sign-in Density of Street	Weekday Sign-in density of Street (per/h·km2)	$\text{Sign}-\text{in density of Street} = \frac{\text{Street Sign}-\text{in Volume}}{\mathit{Community\ area}}$
	Weekend Sign-in density of Street (per/h·km2)	

2.2.2. Getis-Ord Gi* Index

Getis-Ord Gi* index method is a statistical method for spatial data analysis proposed by Getis and Ord, two American geographic scientists, mainly used to detect local clusters or heterogeneity in spatial data sets. This method analyzes the degree of clustering of target attribute values in local space by comparing the sum of the elements and their adjacent elements in a given range with the sum of all elements [21,22]. When the local sum is too different from the expected local sum to be a randomly generated result, a statistically significant Z-score is generated. The Z score is the Gi* statistic, calculated by:

$$G_I^* = \frac{\sum_{j=1}^{n} w_{i,j} x_j - \bar{X} \sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{\left[n\sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}} \quad (1)$$

$$\bar{X} = \frac{\sum_{j=1}^{n} x_i}{n} \quad (2)$$

$$S = \sqrt{-\bar{X}^2 + \frac{\sum_{j=1}^{n} x_j^2}{n}} \quad (3)$$

Where: x_i is the attribute value of element j, $w_{i,j}$ is the spatial weight between element i and j, and n is the total number of elements.

2.2.3. Kernel Density Estimation

Kernel Density Estimation (KDE) is a nonparametric statistical method used to estimate probability density functions, proposed by Rosenblatt (1955) and Emanuel Parzen(1962). Also known as the Parzen window [23]. The basic idea of kernel density estimation is to select the appropriate kernel function, generate the density contribution curve of each data point in the range of observed values, and then superposition and sum the contributions of all kernel functions to get the overall density estimation. The size of the observation range is determined by the bandwidth. This allows a smooth probability density function to be formed over the entire data range, reflecting the overall distribution of the data.

Mathematically, the kernel density is calculated by:

$$f(x) = \frac{1}{nh} \sum_{j=1}^{n} K\left(\frac{x - x^j}{h}\right) \quad (4)$$

Where, n is the number of data points, h is the bandwidth (that is, the search radius in ArcGIS kernel density analysis), and K is the selected kernel function.

The formula for calculating nuclear density chosen in this paper is as follows:

$$f(x, y) = \frac{1}{nh^2} \sum_{j=1}^{n} K\left(\frac{d(x,y)}{h}\right) \quad (5)$$

$$K(u) = \frac{1}{2u} e^{\frac{-u^2}{2}} \quad (6)$$

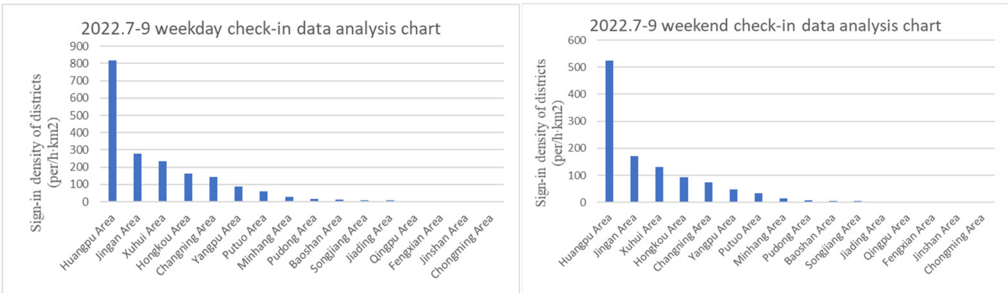
In the formula, d(x,y) represents the distance between the point (x_i, y_i) and the coordinate (x,y), n is the number of points in the range, h is the bandwidth, this paper stipulates that h=0.01, K (u) is the kernel function selected in this paper.

3. Results and Disscutions

3.1. Spatial and Temporal Characteristics of Urban Vitality in Shanghai

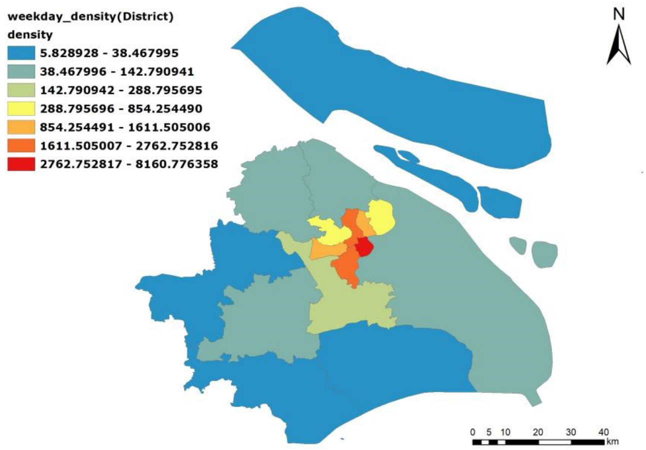
Before conducting a spatial-temporal analysis of urban vitality in Shanghai, it is essential to categorize the city into core and non-core areas based on administrative district check-in density.

As depicted in the statistical charts of administrative district check-in density on weekdays and weekends in Shanghai (Fig.2, Fig.3), the overall urban vitality during weekdays surpasses that of weekends. Spatially, the city's overall urban vitality exhibits a pattern of "high aggregation-low dispersion," with 65% of the urban vitality concentrated in the eastern part of the city, while areas with lower vitality are dispersed. Notably, Huangpu District, Jing'an District, Xuhui District, Hongkou District, and Changning District, ranking as the top five in regional vitality, collectively form the core area of urban vitality in Shanghai, while the remaining districts constitute the non-core areas.

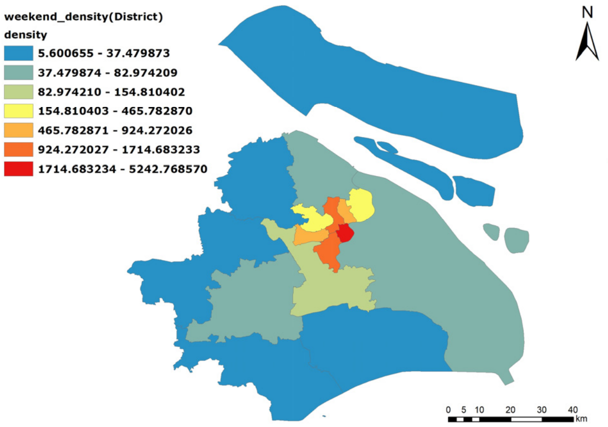


(a) (b)

Figure 2: Statistics of weekdays (a) and weekends (b) of Shanghai



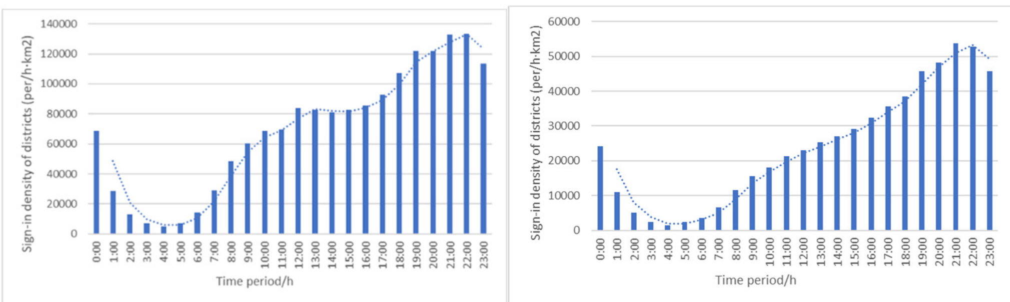
(a)



(b)

Figure 3: Shanghai district-level sign-in density classification map.

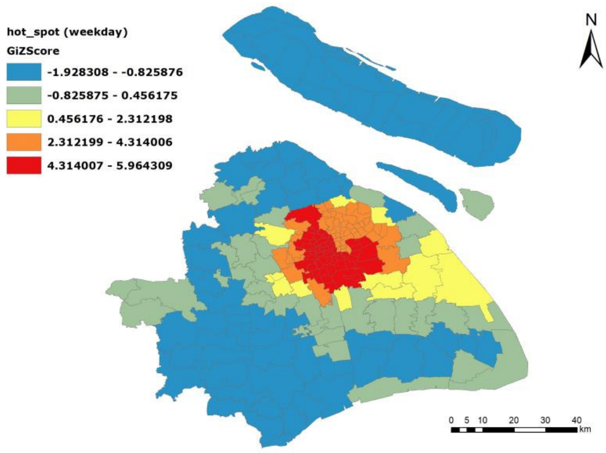
At the temporal scale, this study divides a day into 24 one-hour intervals, examining the changes in check-in density in Shanghai's urban areas during each interval. The statistical charts (Fig.4) reveal that the overall trend of vitality variation in Shanghai's urban areas during weekdays (Fig.4a) and weekends (Fig.4b) from July to September 2022 is roughly similar, showing a pattern of "initial decrease followed by gradual increase." High vitality values are concentrated during the nighttime from 18:00 to 22:00, while low values are concentrated between 1:00 and 6:00 in the early morning. On weekdays, a small peak in vitality occurs from 10:00 AM to 1:00 PM. This trend aligns with real-world scenarios, where the weekday vitality peak during late morning might indicate residents having meals near their workplaces. The nighttime vitality peak corresponds to residents' post-work check-ins, likely during commute or leisure activities. Considering the selected data falls within the summer vacation period, the higher daytime temperatures may influence travelers to plan activities during the cooler nighttime hours. The periods of low vitality align with most residents' rest time, resulting in lower check-in density and, consequently, lower urban vitality values.



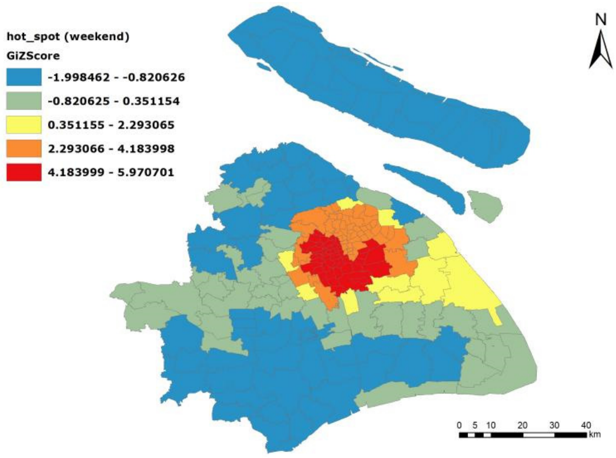
(a) (b)

Figure 4: Statistics of Sign-in density at 24 periods of time in Weekdays (a) and Weekends (b) of Shanghai

Analyzing the overall urban vitality of Shanghai at the spatial scale, the street-level vitality heat maps for weekdays (Fig.5a) and weekends (Fig.5b) reveal distinct patterns. High-confidence vitality heat values are concentrated in the eastern part of the city, aligning with the designation of the core area. Concurrently, the vitality values of streets on both weekdays and weekends exhibit a diffusion distribution. This implies a gradual decrease in vitality heat values from the core area towards the periphery. The farther a street is from the vital core, the lower the confidence level in vitality heat values, indicating lower vitality. Conversely, streets closer to the core exhibit higher confidence in vitality heat values, signifying increased vitality. This spatial analysis provides insights into the nuanced dynamics of urban vitality, with the identified core area playing a crucial role in shaping the overall urban vibrancy in Shanghai.



(a)



(b)

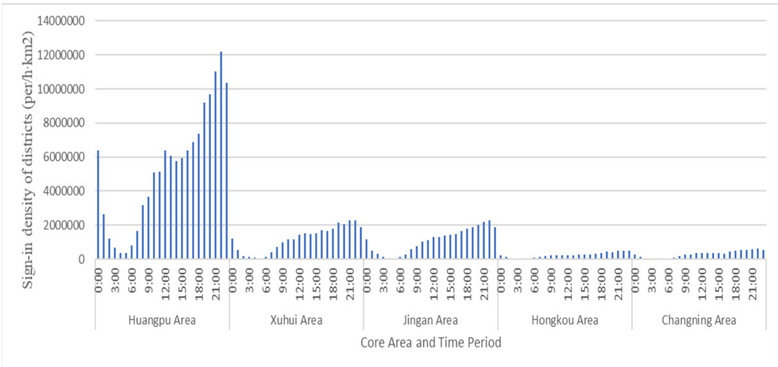
Figure 5: Hot-spots map of Weekdays (a) and Weekends (b) in Shanghai (Z-value)

3.2. Temporal and Spatial Characteristics of Urban Vitality in the Core Area of Shanghai

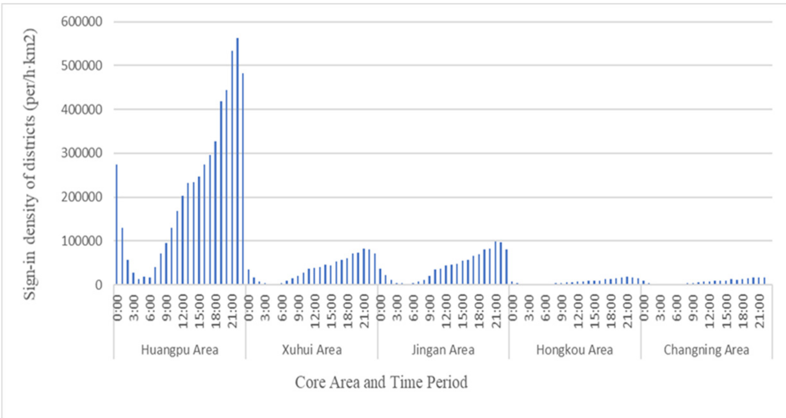
On a temporal scale, an in-depth analysis of the vitality metrics in the Shanghai core area was conducted. The statistical chart (Fig. 6a) illustrates the vitality values at various time periods, while Fig. 6b focuses on rest days. The observed distribution trend aligns with Shanghai's overall vitality pattern, peaking in the evening and reaching its nadir in the early morning.

Notably, on working days, the core area exhibits significantly higher vitality compared to rest days. Huangpu District stands out with the highest vitality, while Changning District records the lowest. This discrepancy underscores the influence of working days on the region's liveliness.

Examining the specific core areas, Huangpu District emerges as a focal point with elevated vitality. This surge may be attributed to the concentration of renowned landmarks, such as the Bund and Nanjing Road Pedestrian Street, resulting in substantial foot traffic, heightened check-in density, and increased vitality during the summer. These findings contribute valuable insights into the dynamic patterns shaping the vitality of Shanghai's core areas, aligning with international standards for high-quality journal publications.



(a)



(b)

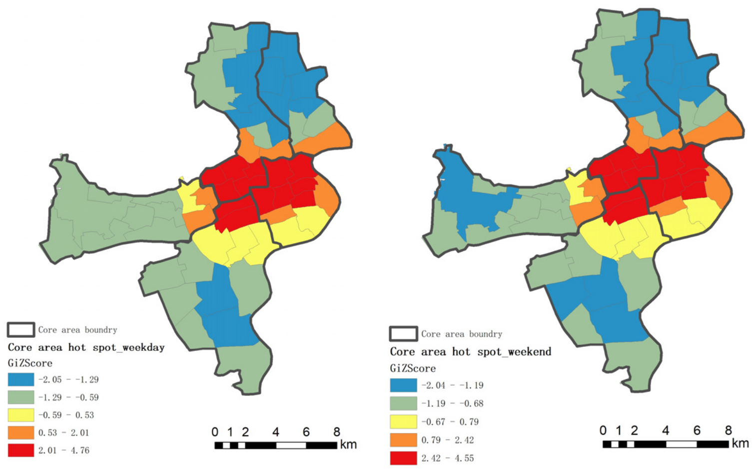
Figure 6: Statistics of Sign-in density at 24 periods of time in Weekdays (a) and Weekends (b) of Shanghai Core Area

Leveraging the check-in density within Shanghai's urban blocks as a metric for core area vitality, the spatial vitality heatmap (Fig. 7) unveils a consistent pattern of high-vitality zones on both weekdays (Fig. 7a) and weekends (Fig. 7b). These zones exhibit a clustered distribution, indicating their ability to stimulate vitality in surrounding areas, albeit within limited extents. Notably, the Bund in Huangpu District emerges as the highest vitality zone, boasting a check-in density of 2092.21 per hour per square kilometer on weekdays and 1384.66 on weekends. In contrast, Liangcheng Xincun Street in Hongkou District records the lowest vitality.

Cold spots during rest days outnumber those on working days, with Kangjian Xincun Street in Xuhui District and Xinjing Town in Changning District being notable examples. These areas, housing numerous commercial districts, contribute to higher urban vitality on weekdays compared to weekends.

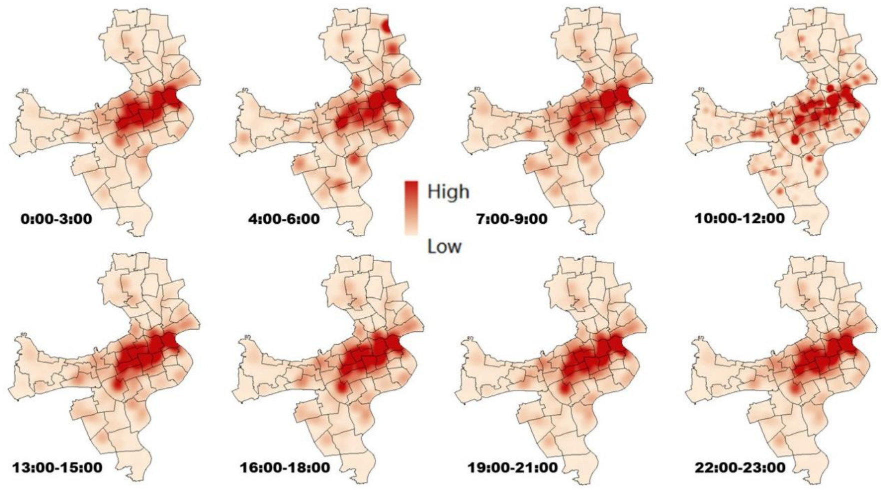
In alignment with the 24-hour urban vitality trend in Shanghai (Fig. 3) and residents' activity states, the 24 time periods are aggregated into eight segments: 0:00-3:00, 4:00-6:00, 7:00-9:00, 10:00-12:00, 13:00-15:00, 16:00-18:00, 19:00-21:00, and 22:00-23:00. Figure 8 illustrates that the core area's spatial vitality exhibits a highly clustered nature, predominantly around renowned landmarks and pedestrian streets like the Bund, Yuyuan, and Huaihai Middle Road. The spatial vitality in the core area is more dispersed during non-commuting hours and lunchtime, followed by an increasing trend in clustering. Notably, from 13:00-15:00 onwards, there is minimal variation in vitality concentration areas.

Combining the analyses of Figures 7 and 8, it is evident that tourists and commuters are the primary contributors to core area vitality, with the tourism sector notably influencing vitality in July 2022. Weekday vitality values consistently surpass those on weekends, displaying greater persistence and a stronger capacity to stimulate vitality in surrounding cold spot areas.

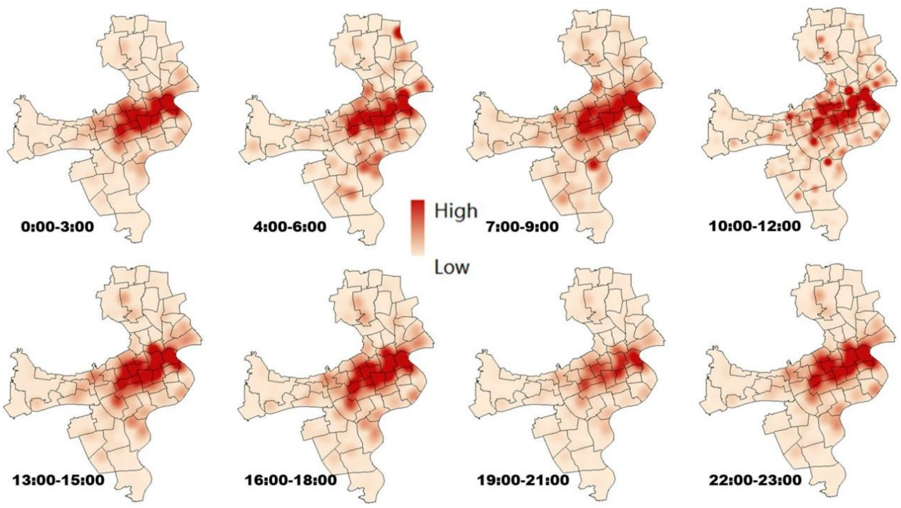


(a) (b)

Figure 7: Hot-spots map of Weekdays (a) and Weekends (b) in Shanghai Core Area from7-9 2022 (Z-value)



(a)



(b)

Figure 8: Results of Kernel density analysis in Shanghai core area during Weekdays (a) and Weekends(b)

4. Conclusions

This study focuses on Shanghai and utilizes Weibo check-in data from July to September 2022 to conduct a multiscale analysis of urban vitality in Shanghai from both temporal and spatial perspectives. The research reveals the following key findings:

(1) From July to September 2022, the overall urban vitality in Shanghai is concentrated in the central and eastern parts, exhibiting a "high agglomeration-low dispersion" spatial distribution. High values of urban vitality are concentrated in the core areas, particularly in Huangpu, Jing'an,

Xuhui, Hongkou, and Changning districts. On both weekdays and weekends, the vitality values on streets show a diffusion distribution, decreasing gradually from the core areas to surrounding neighborhoods. High-vitality streets exhibit strong driving capabilities on neighboring streets, but their influence is limited.

(2) During the study period, the vitality intensity on weekdays in Shanghai is higher than on weekends. It shows a trend of "initial decrease followed by gradual increase" across 24 time periods, with nighttime vitality significantly higher than daytime. The peak vitality occurs between 21:00 and 23:00. The analysis suggests that during this period, which coincides with the summer tourist season, high temperatures during the day prompt most visitors to choose nighttime activities. Additionally, evenings are the leisure time for most residents, resulting in higher check-in density than during the daytime.

(3) Within the study period, the spatial vitality in Shanghai's core areas exhibits a distinct agglomerative distribution. Aggregated areas align closely with famous attractions and high-temperature pedestrian streets. During specified time periods, vitality aggregation is higher in the evening than during the daytime, with lower aggregation during rest and meal times.

Recommendations for enhancing urban vitality in Shanghai based on this study include:

(1) Promoting regional collaboration to boost urban vitality. The analysis indicates that high-vitality areas in the core region have a certain driving effect on surrounding areas but with limited scope. Therefore, the government could consider regional collaborative approaches to expand the influence of vitality in the core areas, fostering overall urban vitality development.

(2) Developing tourism and catering industries in areas with lower vitality. Analysis reveals that high-vitality in core streets is due to cultural attractions, representative pedestrian streets, and diverse dining options, attracting visitors and contributing to a developed tourism industry. For areas with lower vitality, focusing on the development and promotion of pedestrian streets, cultural attractions, and improving tourism facilities could enhance vitality.

This study has certain limitations, such as the selected research data exhibiting characteristics of the peak summer tourist season. Its generalizability requires further research verification. Additionally, this study adopts a macroscopic perspective to analyze the spatiotemporal characteristics of Shanghai's urban vitality, and it does not provide a detailed discussion on the specific impacts of transportation, catering, and other industries on urban vitality. Researchers are encouraged to supplement and further explore the micro-level characteristics of Shanghai's urban vitality.

References

[1] Jacobs J. The death and life of great American cities[M]. New York: Random House, 1961

[2] Glaeser E. Cities, productivity, and quality of life[J]. Science, 2011, 333(6042): 592-594.

[3] Lynch K. The image of the city[M]. Boston: The MIT Press, 1960.

[4] Gehl J. Life between buildings: Using public space[M]. Washington: Island Press, 2011.

[5] Montgomery J. Making a city: Urbanity, vitality and urban design[J]. Journal of Urban Design, 1998, 3(1): 93-11

[6] Jiang, Y., Zhang, Y, et al. A review of urban vitality research in the Chinese world [J]. Transactions in Urban Data, Science, and Technology. Volume 2. Issue 2-3.2023. PP 81-99.
[7] Chai Y, Ta N, Ma J. The socio-spatial dimension of behavior analysis: Frontiers and progress in Chinese behavior geography[J]. Journal of Geographical Sciences, 2016, 26(8): 1243-1260.
[8] NIU Xinyi, LIN Shijia. Spatiotemporal Big Data in Urban Planning Researches: Technological Evolution, Research Topics, and Frontier Trends [J] China City Planning Review.,2022(06):50-57.
[9] SHI Yi, LI Tingting, YANG Junyan*. Research on Urban Waterfront Space Vitality Based on Mobile Phone Signaling Data — A Case Study of Jinji Lake in Suzhou [J]. Landscape Architecture,2021,28(01):31-38.
[10] Wang Bo, Lei Yaqin, Wang Chenggang et al. The spatio-temporal impacts of the built environment on urban vitality: A study based on big data. Scientia Geographica Sinica,2022,42(2):274-283.
[11] Yang Chen, Bingjie Yu, Bo Shu, et al. Exploring the spatiotemporal patterns and correlates of urban vitality: Temporal and spatial heterogeneity[J]. Sustainable Cities and Society, Volume 91,2023,104440, ISSN 2210-6707.
[12] Young-Long Kim.Data-driven approach to characterize urban vitality:how spatiotemporal context dynamically defins Seoul's nighttime[J].International Journal of Geographical Information Science, 2020,34(6):1235-1256.
[13] ZHAO LIU, FANG WANG, ANRONG DANG. Creating Engines of Prosperity: Spatiotemporal Patterns and Factors Driving Urban Vitality in 36 Key Chinese Cities[J]. Big Data,2022,10(6):528-546.
[14] Zeng, P.; Wei, M.; Liu, X. Investigating the Spatiotemporal Dynamics of Urban Vitality Using Bicycle-Sharing Data. Sustainability 2020, 12, 1714.
[15] Wang Bo, Lei Yaqin, Wang Chenggang et al. The spatio-temporal impacts of the built environment on urban vitality: A study based on big data. Scientia Geographica Sinica,2022,42(2):274-283.
[16] Chen, S., Lang, W., & Li, X. (2022). Evaluating Urban Vitality Based on Geospatial Big Data in Xiamen Island, China. SAGE Open, 12(4).
[17] Y. Zhang, L. Yang and X. Wang, "Analysis and Calculating of Comprehensive Urban Vitality Index by Multi-Source Temporal-Spatial Big Data and EW-TOPSIS," 2021 IEEE International Conference on Data Science and Computer Application (ICDSCA), Dalian, China, 2021, pp. 196-201.
[18] Li, Q.; Cui, C.; Liu, F.; Wu, Q.; Run, Y.; Han, Z. Multidimensional Urban Vitality on Streets: Spatial Patterns and Influence Factor Identification Using Multisource Urban Data. ISPRS Int. J. Geo-Inf. 2022, 11, 2. https://doi.org/10.3390/ijgi11010002
[19] Qingsong He, Weishan He, Yan Song, Jiayu Wu, Chaohui Yin, Yanchuan Mou,
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,Volume 78,2018,Pages 726-738,ISSN 0264-8377,https://doi.org/10.1016/j.landusepol.2018.07.020.
[20] Zhen F, Cao Y, Qin X et al. Delineation of an urban agglomeration boundary based on Sina Weibo microblog 'check-in' data: A case study of the Yangtze River Delta[J]. Cities, 2017, 60: 180-191.
[21] Getis A, Ord J K. The Analysis of Spatial Association by Use of Distance Statistics[J]. Geographical Analysis,1992, 24:189-206.
[22] Fotheringham, A. S., Brunsdon, C., & Charlton, M. (1996). Geographically Weighted Regression: A Natural Evolution of the Expansion Method for Spatial Data Analysis. Environment and Planning A, 28(11), 1905–1927.
[23] Rosenblatt, M. (1956). Remarks on Some Nonparametric Estimates of a Density Function. Annals of Mathematical Statistics, 27(3), 832–837.