COMPREHENSIVE ANALYSIS OF THE VITALITY OF URBAN CENTRAL ACTIVITIES ZONE BASED ON MULTI-SOURCE DATA

Case studies of Lujiazui and other sub-districts in Shanghai CAZ

CHENYU HUANG[1], PIXIN GONG[2], RUI DING[3], SHUYU QU[4] and XIN YANG[5]
[1,2,3,4,5]*School of Architecture and Art, North China University of Technology*
[1,3,4]*{huangchenyu303|dingrui418|qushuyuu}@163.com*
[2]*253501297@qq.com* [5]*bjyangxin@126.com*

Abstract. With the use of the concept Central Activities Zone in the Shanghai City Master Plan (2017-2035) to replace the traditional concept of Central Business District, core areas such as Shanghai Lujiazui will be given more connotations in the future construction and development. In the context of today's continuous urbanization and high-speed capital flow, how to identify the development status and vitality characteristics is a prerequisite for creating a high-quality Central Activities Zone. Taking Shanghai Lujiazui sub-district etc. as an example, the vitality value of weekday and weekend as well as 19 indexes including density of functional facilities and building morphology is quantified by obtaining multi-source big data. Meanwhile, the correlation between various indexes and the vitality characteristics of the Central Activities Zone are tried to summarize in this paper. Finally, a neural network regression model is built to bridge the design scheme and vitality values to realize the prediction of the vitality of the Central Activities Zone. The data analysis method proposed in this paper is versatile and efficient, and can be well integrated into the urban big data platform and the City Information Modeling, and provides reliable reference suggestions for the real-time evaluation of future urban construction.

Keywords. Multi-source big data; Central Activities Zone; Vitality; Lujiazui.

1. Introduction

How to enhance the vitality of cities through design is one of the key issues that have long plagued urban scholars. One of the solutions is that introducing the concept of Central Activities Zone (CAZ)in the *Shanghai City Master Plan (2017-2035)* (hereinafter referred to as *Master Plan*)to replace the previous concept of Central Business District (CBD).

PROJECTIONS, Proceedings of the 26th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2021, Volume 2, 549-558. © 2021 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

The concept of the CAZ first appeared in *the Greater London Plan* in 2004, whose original intention was to solve the problem of lack of vitality in the central area caused by the focus on functional zoning in urban planning. The main functions of the CAZ in the *Master Plan* are defined as "core bearer of global urban functions including highly integrated finance, commerce, business, culture, leisure, tourism and other functional areas." In addition, the *Master Plan* also emphasizes the establishment of a dynamic monitoring and a timely maintenance mechanism for the implementation of the plan to realize the dynamic update. In the context of today's continuous urbanization and high-speed capital flow, how to identify the development status and vitality characteristics of the CAZ is the prerequisite for realizing the dynamic update of the plan. While urban big data and artificial intelligence technology will play an important role in it.

In fact, in the past few years, urban big data has underlain a great number of city researches including CAZ development. LIU Liu et al. (2018)use Point of Interest (POI) data to identify potential CAZ areas in the city, and proposes a method of defining CAZ boundary based on functional blending degree. Based on urban big data and multiple modeling methods, LONG et al. (2014) propose a method to build a refined urban model to support policy formulation. NIU and SONG (2014) use mobile phone data to identify the spatial structure of Shanghai as well as the city's different functional areas and mixed functions. In the research field of urban vitality, in addition to the research on architectural form derived from Jane Jacobs's theory of diversity and vitality, data research has gradually become a more convincing way. Bosselmann (2012) believes that urban vitality can be characterized by crowd flow to a certain extent. On this basis, a method for urban vitality research using real-time crowd flow distribution data from Baidu heat map has been developed in China. Wu and YE (2016) use Baidu heat map to identify the periodic characteristics of Shanghai's population gravity center.

Based on the above background, 8 typical areas are taken in Shanghai such as Lujiazui CAZ as examples, the hourly Baidu heat map data is applied as an explicit representation of the CAZ vitality. Meanwhile, commercial quality data, functional POI, Second-hand house data, geographic data etc. are respectively obtained, on which establishes a total of 19 indexes, the potential correlation between each index and vitality is explored. Finally, a shallow neural network is built to establish a regression model for the 8 most relevant indexes, in order to reveal the relationship between the indexes and the vitality of CAZ, so as to guide future urban design and planning.

2. Methods

Shanghai is a representative megacity in China, and its spatial structure and development mode have been extensively studied worldwide. According to the *Master Plan*, 7 sub-districts in the CAZ are selected including Lujiazui, The Bund, East Nanjing Road, West Nanjing Road, Middle Huaihai Road, Jing'an Temple and Zikawei. Besides, as a traditional CBD, Wujiaochang is a representative area to be selected. Sub-district selection is shown in Figure 1.

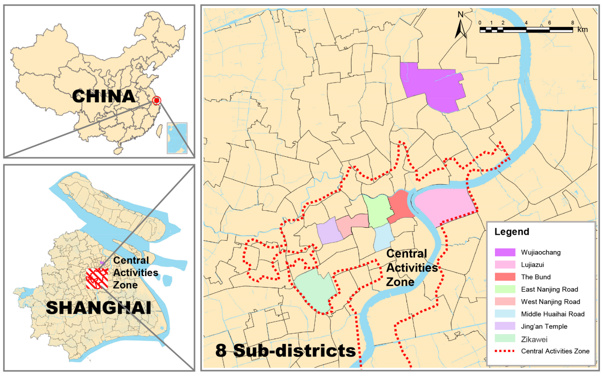

Figure 1. Geographical location of the 8 sub-districts studied.

Multi-source data has been obtained in this study and reliable index calculation and data analysis is conducted on the above 8 sub-districts. Figure 2 shows the data used and the indexes set in this paper, including four types of data: heat map data, functional data, geographic data and street view data respectively sourced from 5 different data platforms. Based on these data, 2 vitality values and 19 multi-dimensional indexes were established, including Commercial Quality Indexes, Housing Conditions Indexes, Functional Facilities Indexes, Morphological Indexes, Perceptible Indexes. Finally, an all-round understanding of the site is achieved, and potential indexes that affecting vitality are explored. The data processing and calculation methods of a series of indexes will be described in detail in this section.

2.1. CALCULATION OF CAZ VITALITY VALUE

Baidu heat map data of Shanghai on weekday (Dec.10, 2020) and weekend (Dec.12, 2020) is downloaded, and the data sampling accuracy is 1h from 7:30 in the morning to 23:30 in the evening, 17 data per day. The cleaned data was imported into ArcGIS software, and the map with coordinates (WGS84 Universal Transverse Mercator Zone51) was used for geographic calibration and processed into vector image. The relative value of the density can only be applied because of the open data source not providing accurate legends. In this study, the vector heat map is classified according to the grid value, and a heat value ranging from 0-1 is defined as the representation of CAZ vitality. It is shown in the Figure 3 that the heat value of Shanghai has been graded by 10 levels of color, and the accuracy of the data can be seen from the figure on the right.

Given that the indexes created in this study are mostly static such as functional facilities and building forms, 17 hourly heat maps of weekend and weekday are superimposed on a 100*100m fishing net and averaged.

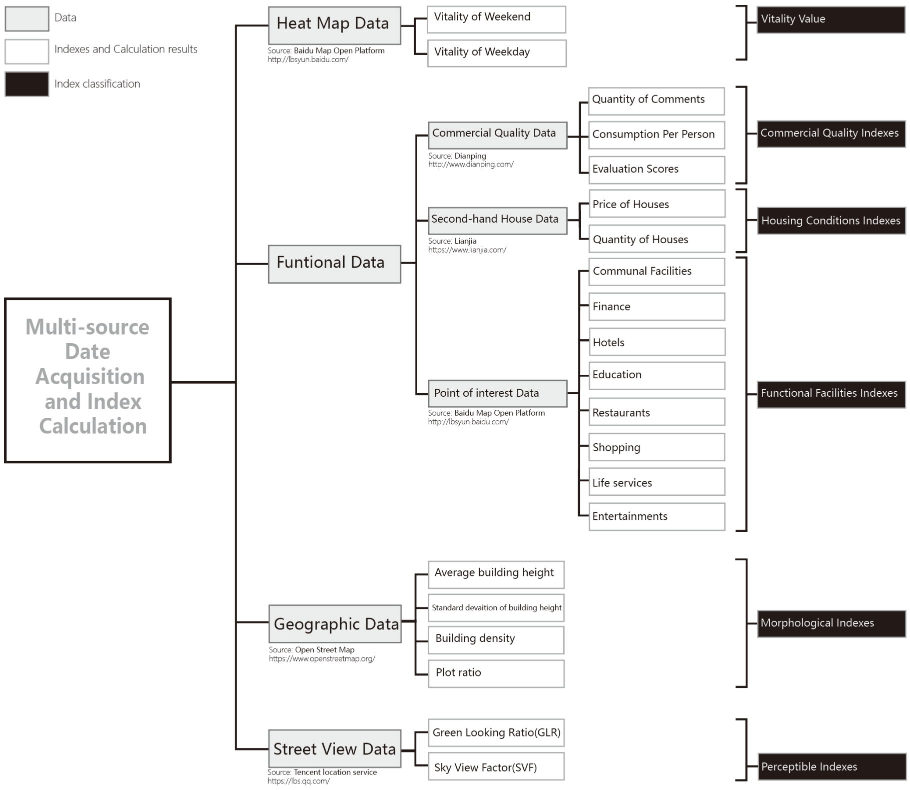

Figure 2. Multi-source data acquisition and index calculation framework.

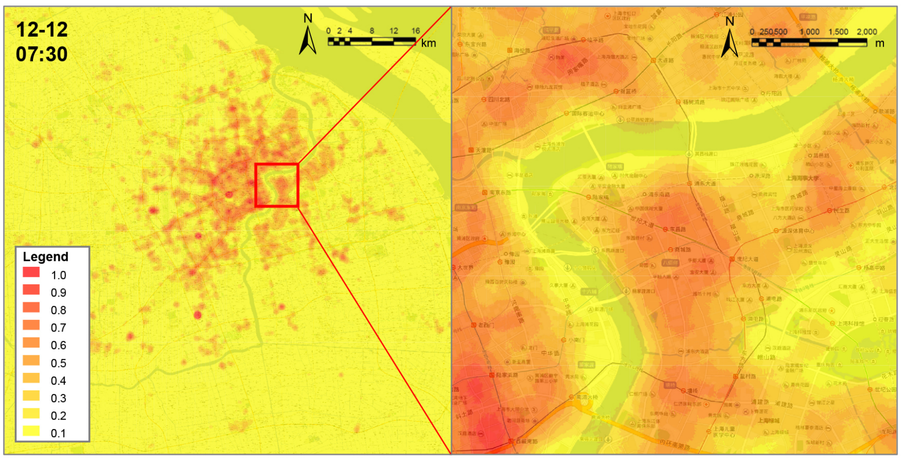

Figure 3. The vector heat map displayed hierarchically by heat value.

2.2. CALCULATION OF FUNCTIONAL INDEXES

The *Master Plan* puts forward comprehensive requirements for the functions of the CAZ. In order to achieve the comprehensive goal from financial business to cultural education to tourism and leisure, commercial quality data (quantity of comments, consumption per person and evaluation scores), second hand house data (price of houses and quantity of houses) and POI data of functions and facilities (communal facilities, finance, hotels, education, restaurants, shopping, life services, entertainments) are obtained by Python respectively from the Dianping, Lianjia and Baidu Map Open Platform.

Similarly, a 100*100m fishing net is used to cut 8 sub-district, counting the number of functional facilities (such as the number of restaurants or the number of entertainment) or the average value of indexes (such as the average value of house prices) on each 100*100m land . The calculated results will be given in Section 3.

2.3. CALCULATION OF MORPHOLOGICAL INDEXES

According to Jane Jacobs's theory, urban form factors such as small-scale compact urban texture, short and winding streets can trigger activities and interactions, which are essential to the formation of a city's diversity and vitality. In fact, by performing spatial and geographic operations on geographic data, more indexes that can describe the urban morphology can be obtained, which allows us to continuously quantify the theoretical results of Jacobs. Based on the geographic data (buildings, roads, etc.) obtained from the Open Street Map combined with ArcGIS software, this study created four morphological parameters to describe cities that have potential effects on urban vitality form. Several basic parameters are determined by the definitions in Figure 5, and the calculation formulas of the morphological indexes are given in Table 1.

Table 1. Equations of morphological indexes.

Building morphology index	**symbol**	**Unit**	**Equation**
Average building height	H_{ave}	m	$H_{ave} = \frac{1}{n}\sum_{i=1}^{n} H_i$
Standard deviation of building height	σ_H	m	$\sigma_H = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(H_i - H_{ave})^2}$
Building density	λ_P	%	$\lambda_P = (\sum_{i=1}^{n} A_{Pi})/A_T$
Plot ratio	PR		$RP = \sum_{i=1}^{n} A_{Pi} \times F_i / A_T$

In the table above, H_i is the height of the building, n is the number of the building, A_P is the footprint area of the building, A_T is the lot area of the urban lot, F_i is the number of the floor of a building. In this study, the average values of the four morphological indexes for each 100*100m of land for 8 sub-districts are calculated. The results will be given in the Section 3.

2.4. CALCULATION OF PERCEPTIBLE INDEXES

It is assumed that Green Looking Ratio(GLR) and Sky View Factor(SVF) have potential effects on attracting people and enhancing urban vitality, which are defined as perceptible indexes. The reason is that the green of the vegetation in the field of vision plays an important role in regulating people's psychology, which has been proved by many studies. On the other hand, The reason why SVF can be used as an important morphological index is that it has a strong description of urban spatial enclosure. In other words, a space with a higher GLR may cause people to gather; a space with a lower SVF indicates a high degree of spatial enclosure, which conforms to Jacobs's theory that small-scale space may be contribute to the promotion of urban vitality.

A deep learning technique called image semantic segmentation is used in the calculation of these two indexes. A semantic segmentation network SegNet based on VGG-16 net, a convolutional neural network (CNN), is established to classify each pixel in the image to generate a segmented image by category. The street view image is acquired by calculating the center point coordinate parameters of the 100*100m fishing nets of 8 sub-districts from Tencent webservice API of Tencent Location Service.

360-degree street view images are used for stitching and polar coordinate transformation to generate fisheye images. The above network is also used to segment and calculate the proportion of the sky. The calculation results of the two indexes of GLR and SVF will be given in Section 3.

3. Calculation results and analysis

3.1. CALCULATION RESULTS

Figures 4-9 show the calculation results in Section 2 taking Lujiazui as an example, by mining multi-sources of big data, an accurate quantitative understanding of the site can be achieved.

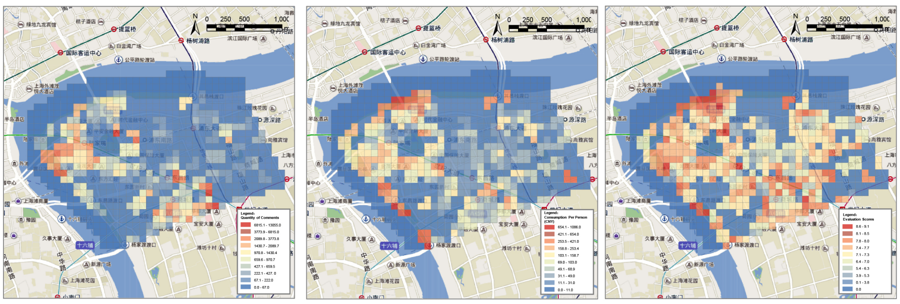

Figure 4. Calculation results of commercial quality indexes of Lujiazui (from left to right are quantity of comments, consumption per person and evaluation scores).

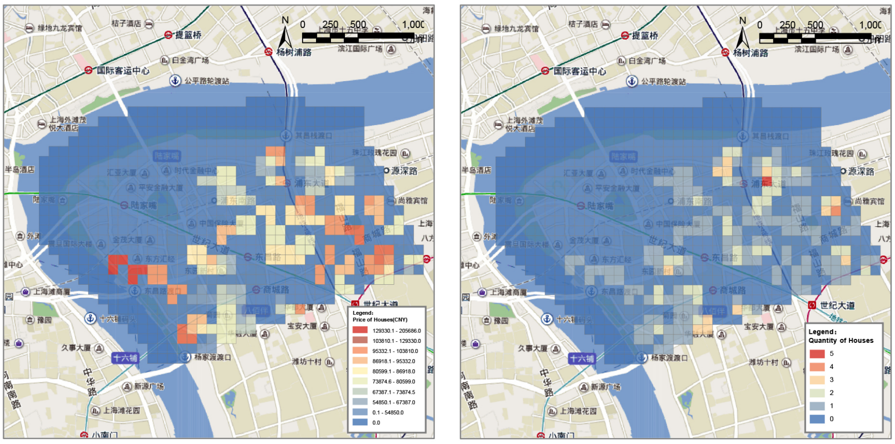

Figure 5. Calculation results of housing conditions indexes of Lujiazui (from left to right are price of houses and quantity of houses).

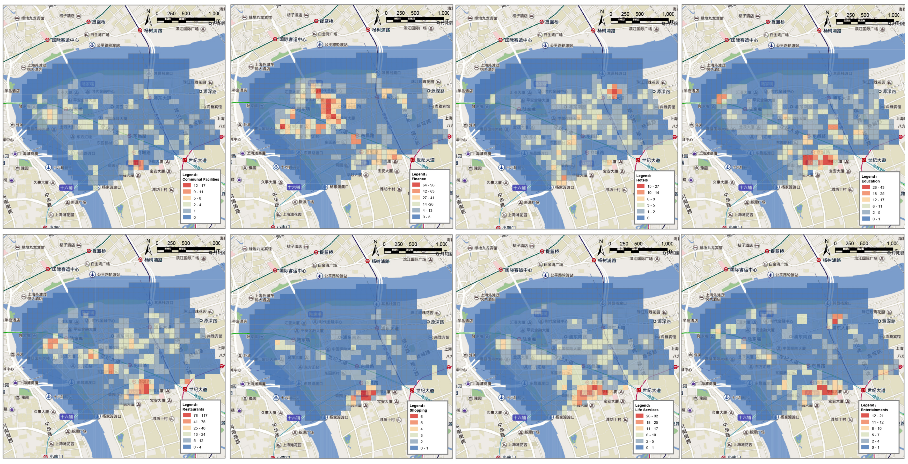

Figure 6. Calculation results of functional facilities indexes of Lujiazui (from top left to bottom right are communal facilities, finance, hotels, education, restaurants, shopping, life services and entertainments).

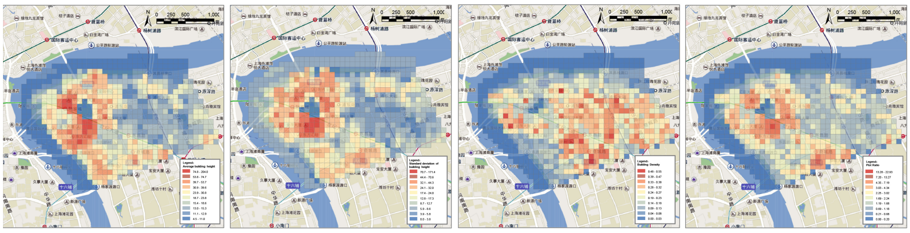

Figure 7. Calculation results of morphological indexes of Lujiazui (from left to right are average building height, standard deviation of building height, building density and plot ratio).

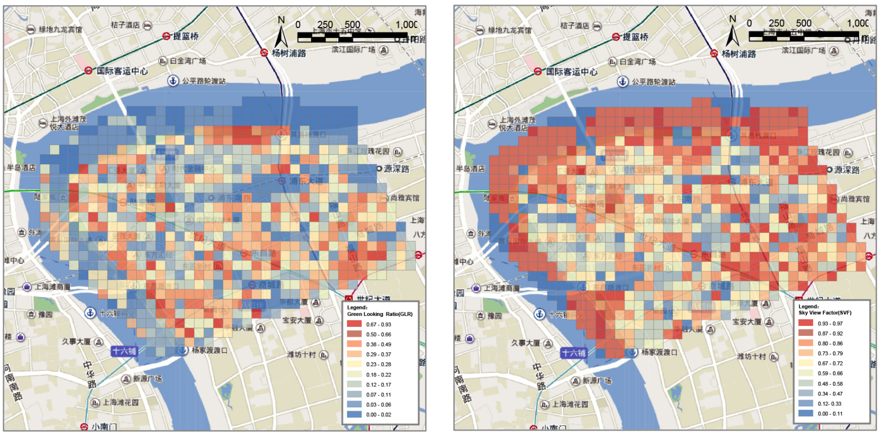

Figure 8. Calculation results of perceptible indexes of Lujiazui (from left to right are GLR and SVF).

East Nanjing Road and other 7 sub-districts are calculated based on the previous method. By counting the average and standard deviation of the indexes of the 8 sites, we can compare CAZ in a quantitative manner. And taking the highly dynamic sub-districts as an positive example, improve the overall vitality of CAZ by means of supplementing functional facilities, optimizing building morphology or improving commercial quality.

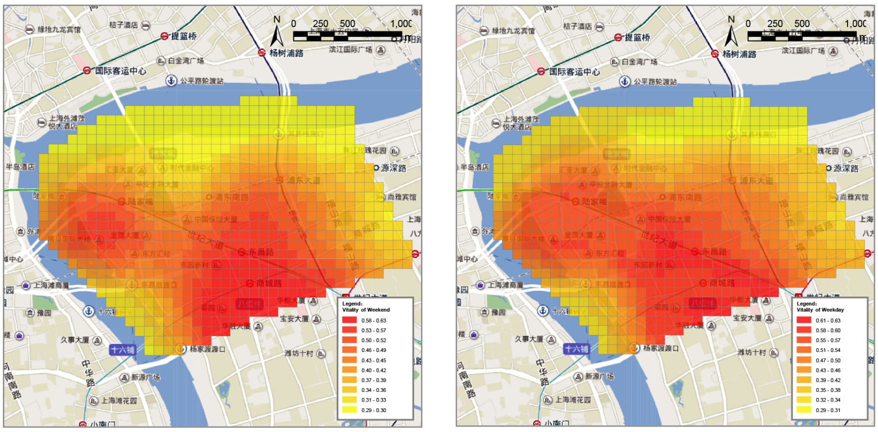

Figure 9. Calculation results of vitality value (from left to right are vitality of weekend and vitality of weekday).

3.2. CORRELATION ANALYSIS

In order to reveal the correlation between various indexes and vitality, correlation analysis was adopted. After preliminary calculation, there is not a linear correlation between the index and the vitality value, so the Spearman method should be used to analyze the correlation of each variable. Spearman correlation coefficient reflects the correlation between index and vitality value. A positive value indicates a positive correlation, and a negative value indicates a negative correlation. The greater the absolute value, the stronger the correlation. Through correlation analysis we can draw some interesting conclusions. The 19 indexes are different from the correlation between the vitality of the weekday and weekend For example, the positive correlation of the quantity of comments on the vitality of the weekday is stronger than that of the vitality of weekend, while life services, shopping and entertainments has a stronger correlation with vitality of weekend. It is also shown that the GLR has a weak influence on the vitality of CAZ and can be ignored. There is a negative correlation between SVF and vitality, which may be because the smaller the SVF, the better the sense of enclosure, which is conducive to the gathering of people and is consistent with empirical judgment.

In order to be able to effectively use the correlation between various indexes and CAZ vitality in future urban design and planning, we selected the eight indexes with the greatest correlation with vitality, and used shallow neural networks to build a regression model to predict the city how the design scheme and planning scheme influence the vitality of CAZ and provide effective support for dynamic planning. They are: plot ratio (ρ=0.528/0.532), building density (ρ=0.439/0.405), entertainments (ρ=0.439/0.426), life services (ρ=0.489/0.470), restaurants (ρ=0.457/0.456), hotels (ρ=0.394/0.404), finance (ρ=0.389/0.427) and quantity of comments(ρ=0.488/0.500).

Using MATLAB Statistics and Machine Learning Toolbox, a shallow neural network is built with 10 hidden layers, 70% of the sample data divided into the training set, 15% for verification, and 15% for independent generalization testing. Levenberg-Marquardt algorithm is used for training. In the entire sample data, the regression effect of the model is acceptable, which is $R^2 = 0.925$. The CAZ vitality prediction model created by using 8 indexes has great application potential. We hope to create an open model for CAZ vitality, which can continuously load new sample data and indexes for dynamic training. Through the way of vitality prediction, the real-time update and dynamic judgment of urban design and planning are forced.

4. Conclusions

With the use of the concept of Central Activities Zone in the *Shanghai City Master Plan (2017-2035)*, core areas such as Shanghai Lujiazui will be given more connotations in the future construction and development. In this study, Baidu heat data is used to characterize the city's vitality, function data, geographic data, and street view data are relatively obtained from several sources, and 19 indexes are established for 8 representative CAZ (or potential) sub-districts at a spatial resolution of 100*100m. Correlation analysis of the calculation results of the indexes yields 8 indexes that are most relevant to CAZ vitality. These indexes are used to build a vitality prediction model. However, the research in this paper is still inadequate, such as the lack of data acquisition and analysis of traffic facilities and traffic usage that may affect the vitality of CAZ. We will do more research and improvement in the future. Using multi-source big data can create a multi-dimensional accurate city portrait of the CAZ area, which is of great significance for comprehensive site identification. This paper not only provides an application method for multi-source big data, but also studies the potential related factors that affect the vitality of CAZ and draws interesting conclusions. The regression model created in this paper will become a tool to improve the disconnect between planning and design solutions and actual results in the foreseeable future and will provide a powerful tool for refined urban governance and dynamic planning.

Acknowledgments: This research was financially supported by Beijing Natural Science Foundation (8202017) and 2018 Beijing municipal university Academic Human Resources Development- youth talent support program (PXM2018_014212_000043)

References

LIU, L. and LIU, Z.Q.: 2018, *The Recognition of CAZ in Shanghai Based on Evaluated POI*, Springer.

LONG, Y., MAO, M.G. and MAO, Q.: 2014, FINE-SCALE URBAN MODELING AND ITS OPPORTUNITIES IN THE "BIG DATA" ERA: METHODS, DATAAND EMPIRICAL STUDIES, *HUMAN GEOGRAPHY*, **29**, 7-13.

NIU, X.Y., DING, L. and SONG, X.D.: 2014, Understanding Urban Spatial Structure of Shanghai Central City Based on Mobile Phone Data, *Urban Planning Forum*, **6**, 61-67.

WU, Z.Q. and YE, Z.N.: 2016, RESEARCH ON URBAN SPATIAL STRUCTURE BASED ON BAIDU HEAT MAP: A CASE STUDY ON THE CENTRAL CITY OF SHANGHAI, *City Planning Review*, **4**, 33-40.