Analysis on spatial distribution pattern of football fields: A case study in Wuhan, China

Yuwei Wang^{1,*}, Kaicheng Wu¹, and Wang Gao¹

¹School of Artificial Intelligence, Jiangnan University, Wuhan 430056, China

Abstract. In order to promote the sustainable development of sports facilities, the statistical data of football fields in 13 districts of Wuhan city were selected. Based on geographic information system (GIS) and exploratory spatial data analysis (ESDA) technology, this paper analyzed the global and local autocorrelation of the number of football fields per capita in Wuhan city. The results indicate that: (1) Comparing to the new districts in Wuhan, the number of football fields per capita in the central districts was relatively few. (2) In the global scale, the number of football fields per capita in Wuhan presented a spatial aggregation pattern. (3) In the local scale, the regional discrepancy and spatial clusters were confirmed in the local spatial autocorrelation analysis for football fields in Wuhan. At the same time, the research results could provide reasonable references for the planning of football field facilities.

1 Introduction

Football fields are the material basis for the development of football and are important parts of the national activity venues [1,2]. In recent years, people's enthusiasm for participating in football has gradually increased. Consequently, improving the degree of participation in football and the number of people involved are important approaches to build China into a world sports power. But at present, our country is facing the problems of insufficient supply and unreasonable spatial distribution for football field facilities, which restricts the sustainable development of football [3,4].

With the help of GIS and ESDA technology, the spatial distribution of different sports venues can be mapped, and the spatial dependence and heterogeneity distribution characteristics of sports venues can be defined, quantified and explained [5–7]. Nevertheless, few studies have discussed the distribution patterns of football fields, especially at county level.

Wuhan is one of the first batches of China's football reform pilot cities. Therefore, it is of great value to explore the spatial distribution pattern of football fields in Wuhan. This study applied ESDA method and Moran's I index to measure the spatial association patterns of football field distribution in 13 districts of Wuhan city. Firstly, from a global perspective, the Moran's I index was used to detect the spatial distribution pattern of the number of football fields in Wuhan. Secondly, from the local point of view, the local Moran's I analysis were exerted to identify different local clustering patterns in the study area.

2 Study area, data and methodology

2.1 Study Area

Wuhan is located in central China, between 29°58'N-31°22'N and 113°41'E -115°05'E. The total area is about 8569.15 km², including 13 administrative districts. Among them, Jiang'an District (Marked as 1 in Figure 1, the same below), Jiangnan District(2), Qiaokou District(3), Hanyang District(4), Hongshan District(5), Wuchang District(6) and Qingshan District(7) are the central districts, while Jiangxia District(8), Hannan District(9), Caidian District(10), Dongxihu District(11), Huangpi District(12) and Xinzhou District(13) are the new districts [8]. With the rise of national fitness as national strategy, the number of football fields has reached a certain extent in Wuhan, the spatial distribution pattern of football fields needs to be further examined.

2.2 Data

The data used in this study includes three sections, namely, Wuhan football field survey data, Wuhan administrative data and demographic data. (1) Wuhan football field survey data were originated from field census (10 years as a census cycle). The data were based on the 5th national general survey of sports fields (2004), the 6th national general survey of sports fields (2014), and general survey of football fields in Wuhan (2018). The data contain the detailed name, location, and year of completion for each football field. (2) The data of Wuhan administrative division was from the vector map of Wuhan City (scale 1:10000) drawn by Wuhan Planning Institute, and the data

* Corresponding author: weberwang@jhu.edu.cn

format is Shapefile. It includes 13 regions under the jurisdiction of Wuhan. (3) The demographic data was obtained from Hubei statistical yearbook 2019, including the number of permanent residents in 13 districts of Wuhan in 2018.

In this study, the number of football fields per capita was applied as the index to explore the spatial distribution patterns of football fields in Wuhan. The number of football fields per capita is calculated by dividing the number of football fields per district by the permanent residents (Unit: Number/10,000 person, the same below).

Figure 1. The administrative division of the study area.

2.3 Methodology

2.3.1 The analysis on global spatial autocorrelation

The first law of geography is the basis of spatial data analysis, which points out that everything is related to other things, but the similar things are more closely related. Spatial autocorrelation is an important index to study the degree of spatial data association, which is used to measure the similarity between the same spatial attributes. Spatial autocorrelation can reveal three distribution associations of spatial data, namely positive association (clustered distribution), negative association (distributed distribution), non-association (random distribution) [9,10].

Two metrics, i.e., global Moran's I index and local Moran's I index, were adopted to estimate the global and local spatial autocorrelation for football fields respectively. Global Moran's I index is used to measure the overall distribution for a certain spatial attribute in the whole research area. The calculation formula is as follows:

$$I = \frac{n}{s_0} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (1)$$

Where n represents the number of spatial units in the study area; x_i (x_j) represents the observation value of the i (j)th space unit in the study area; w_{ij} is the spatial weight matrix, which represents the adjacency of spatial units; $\bar{x}$ is the average value; s_0 is the sum of the spatial units of the spatial weight matrix [11].

The statistical test of global Moran's I usually adopts two-sided test, and the calculation formula of Z value is as follows:

$$Z = \frac{I - E(I)}{\sqrt{V(I)}} \quad (2)$$

Where $E(I)$ and $V(I)$ denote the expected values and variances of global Moran's I index.

Global Moran's I value generally ranges from -1 to 1, and greater than 0 indicates that there is a positive spatial association in the distribution of football fields in the study area. On the contrary, less than 0 indicates that there is a negative spatial association in the distribution of football fields. The closer the value is to 1, the higher the spatial positive association degree is, and revealing the distribution of football fields tends to be concentrated. Particularly, when the value reaches 0, the football fields are randomly distributed in the study area.

2.3.2 The analysis on local spatial autocorrelation

Global spatial autocorrelation is essentially a global average measure in space, which can't reflect the spatial association among local regions. Local spatial autocorrelation is often employed to access the spatial association of every spatial unit in the study area. The measurement of local spatial autocorrelation usually depends on local indicators of spatial association (LISA). LISA transforms global spatial association into local form. The idea of LISA is introduced into local Moran's I index, which can define the local Moran's I index for any spatial observation i in the study area [12]. For spatial observation i , the LISA definition for Moran's I is as follows:

$$I_i = z_i \sum_j w_{ij} z_j \quad (3)$$

z_i and z_j record the deviation between the observed values and the mean values of all observed values. The addition operation of j targets at all the observed values in the neighborhood set j_i of i , and the definition of spatial neighborhood is expressed by the spatial weight matrix w_{ij} .

In literature, Moran scatter plot and LISA cluster map are often used in local spatial autocorrelation analysis. The combination of the two methods can reveal the spatial association characteristics of the spatial objects from different perspectives.

Moran scatter plot depends on Cartesian coordinate system, where the X-axis is the normalized attribute value

of each spatial unit, Y-axis is the spatial lag value. Moran scatter plot contains four quadrants, representing four local spatial clustering patterns between observation i and its neighbourhood observations, namely, first quadrant: local high-high clustering pattern, high value observation is encompassed by high value neighbors; second quadrant: local high-low clustering pattern, high value observation is encompassed by low value neighbors; third quadrant: local low-low clustering pattern, low value observation is encompassed by low value neighbors; fourth quadrant: local low-high clustering pattern, low value observation is encompassed by high value neighbors.

In addition, through the LISA cluster map, the statistical significance of four local spatial clustering patterns can be inspected, and the statistically significant autocorrelation patterns can be highlighted and visualized on the LISA cluster map. Generally, the statistical significance level is set at $\alpha = 0.05$.

3 Result

3.1 The spatial distribution of football fields in each district in Wuhan

By cleaning and integrating the census data of sports fields in three years (2004, 2014 and 2018), a total of 1022 football fields in Wuhan were obtained in this study. Table 1 summarizes the number of football fields and population of 13 districts in Wuhan.

Table1. The number of football fields and population of 13 districts in Wuhan.

District	The number of football fields	Population (Unit:10,000)
Jiang'an	77	96.27
Jianghan	64	72.97
Qiaokou	53	86.87
Hanyang	43	66.42
Wuchang	77	128.28
Hongshan	128	167.73
Qingshan	13	52.89
Dongxihu	66	58.48
Caidian	72	76.16
Jiangxia	201	96.20
Huangpi	70	101.19
Xinzhou	88	91.06
Hannan	70	13.58

Figure 2 is a hierarchical map to display the number of football fields per capita in 13 districts in Wuhan. The number of football fields was classified into 5 categories. As is displayed in the map, the spatial distribution of football fields in different districts shows regional

variations. The areas with less football fields were mainly concentrated in central districts, including Jiang'an District, Qiaokou District, Hanyang District, Wuchang District, Hongshan District and Qingshan District. Correspondingly, the new districts occupied more football fields, including Dongxihu District, Caidian District, Jiangxia District, Xinzhou District and Hannan District.

Figure 2. The hierarchical map of the number of football fields per capita in Wuhan.

3.2 Global spatial autocorrelation analysis for number of football fields per capita

The global Moran's I index and its z-score for the number of football fields per capita were calculated by using the GeoDa software (Version 1.18.0), and the pseudo p-value was obtained by using 99, 199, 499, 999 and 9999 random permutations respectively. In the calculation, the queen contiguity matrix was adopted as the spatial matrix in this study. The Global Moran's I index, z-value and p-value for the number of football fields per capita are listed in Table 2.

Table2. Estimated value of Morans' I Index of football fields in Wuhan.

The number of football fields per capita	Global Moran's I index	z-value	p-value
	0.206	3.32	0.006**

**significant at 1% level

The results indicate that the global Moran's I index of the number of football fields per capita in Wuhan is significant at the level of 0.01. The global Moran's I index

is 0.206, which implies that the number of football fields per capita is positively aggregated in space. With the proximity of geographical positions, the spatial association of the football fields is more significant. That is, with the aggregation of spatial position, the spatial association becomes more significant. High values tend to be surrounded with high values and low values tend to be surrounded with low values.

3.3 Local spatial autocorrelation analysis for number of football fields per capita

Figure 3. The Moran scatter plot of football fields in Wuhan.

The Moran scatter plot (Figure 3) displays the local spatial distribution of the number of football fields per capita in Wuhan. Two districts were distributed in the first quadrant, including Jiangxia District and Hannan District, indicating a high-high clustering pattern for the number of football fields per capita. At the same time, most of the districts in Wuhan were distributed in the third quadrant, i.e., Jiang'an District, Jianghan District, Qiaokou District, Hanyang District, Wuchang District, Hongshan District, Qingshan District, Dongxihu District, Huangpi District and Xinzhou District, indicating a low-low clustering pattern for the number of football fields per capita. In comparison, the Caidian District was the only region located in the fourth quadrant. It could be derived that the number of football fields per capita in Caidian District was lower than the surrounding areas.

Figure 4 shows the univariate LISA cluster map of football fields in Wuhan. The map depicts the clustering characteristics of the number of football fields per capita in Wuhan in terms of local scale. The red part (Wuchang District) in Figure 4 represents that the spatial association mode of football fields for the region and its surrounding areas is a local low-low value clustering pattern, that is, the number of football fields per capita in this region and its neighborhood are both low, and the low value are encompassed by the low values. The green part is a low-high value clustering pattern, namely, the number of football fields in the region is low, while the number of football fields in the neighborhood area is high, the low

value are encompassed by the high values. The area showing low-high pattern is Caidian District.

Figure 4. The univariate LISA cluster map of football fields in Wuhan.

4 Discussion

This paper tried to analyze spatial distribution pattern of football fields in Wuhan through ESDA and GIS technology. ESDA technology has been proven to be an effective tool to analyze spatial data, with spatial association measurement as the core, which can reveal the spatial dependence and spatial heterogeneity of data. GIS provides technical support for ESDA analysis. Based on GIS technology, spatial data can be visualized, stored, simulated and predicted. By adding spatial location information into football field data, the field data can be regarded as spatial data.

First of all, the hierarchical map for the number of football fields per capita in Wuhan provides evidence for the uneven distribution of football fields. We have discovered that central districts in Wuhan tend to own less football fields. On the contrary, the number of football fields per capita in new districts increases compared to the central districts. With the acceleration process of urbanization and population migration, the construction areas are expanding obviously [13,14]. In the central areas of Wuhan, most of the lands are occupied by urban areas, the areas planned for sports facilities are relatively rare. Correspondingly, in the new districts of Wuhan, the level of urbanization is comparatively low, and therefore more land has been planned for the construction of football fields.

Secondly, the results of global spatial autocorrelation, i.e., calculation of global Moran's I index, have confirmed

the positive spatial association of football fields in whole area of Wuhan city. It means that the football field allocation level of Wuhan city is in a state of clustering distribution among neighboring districts. In the vast majority of cases, when the per capita number of football fields in a region is high, the number of football fields in the surrounding region is also high [15,16]. In line with the high-high clustering pattern, when the number of football fields per capita in a certain area is low, the number of neighboring areas is also low. These two clustering patterns are further confirmed in local spatial autocorrelation subsequently.

Last but not least, the Moran scatter plot and univariate LISA cluster map of football fields in Wuhan provide contribution to the study of local spatial autocorrelation. The local Moran's I index in most districts mainly falls into the third quadrant (low-low clustering pattern). This provides some insights that there exists a clustering trend for the districts with lower number of football fields [17]. When the number of football fields per capita in a region is small, the number of fields in the surrounding regions tends to be small with high probabilities.

Under the statistical test level α of 0.05, only two districts were displayed on the LISA cluster map. The number of football fields per capita in Caidian District and its surrounding districts showed a low-high spatial clustering pattern. This reminds the policymaker to pay much attention to the unbalanced development of football fields in Wuhan. To cope with this, for the areas with less football fields, the number of football fields per capita should be expanded. During the implementation of sports venues planning, these areas need to be given priority during the development process. Meanwhile, the significant low-low spatial distribution pattern for Wuchang District and its surrounding districts was found. This could be explained by the spatial diffusion effect between Wuchang District and its surrounding areas, the spatial discrepancies between these regions were decreasing. It reflects the lack of implementation of the planning policies, indicating that the government's management and planning of sports facilities need to be strengthened.

This study only targets to analyse the global and local spatial distribution patterns of football fields in Wuhan city. We call for future research to study the driving factors of these distribution patterns, including socio-economic factors, natural factors and policy factors.

5 Conclusion

Sports facilities play an important role on physical fitness and health. Our research found that comparing to the new districts in Wuhan, the number of football fields per capita in the central districts is relatively less. To address the uneven distribution of football fields in Wuhan city, the spatial distribution pattern of football fields should be captured in global and local scales.

The results of global spatial autocorrelation for football fields indicate that the number of football fields per capita tended to aggregate at district level. The spatial

positions of football fields have potential mutual influence, the football fields tend to gather in adjacent areas. In addition, the results of local spatial analysis indicate that there are regional disparities for the number of football fields per capita in Wuhan.

Therefore, the land use planning policy for football fields should be initiated. For instance, more football fields should be planned and built in the central areas of Wuhan.

References

1. C. Vater, S. Luginbühl, and L. Magnaguagno, *J. Sports Sci.* **37**, 2789 (2019)
2. Y. Hosokawa, A. J. Grundstein, and D. J. Casa, *J. Athl. Train.* **53**, 860 (2018)
3. J. Connell, *Geogr. Res.* **56**, 5 (2018)
4. R. Izzo and E. Belpassi, *Risk* **178**, 2016 (2018)
5. E. C. A. Nordbø, H. Nordh, R. K. Raanaas, and G. Aamodt, *Landsc. Urban Plan.* **177**, 19 (2018)
6. F. He, *Arab. J. Geosci.* **14**, 1 (2021)
7. G. Rybarczyk, S. Banerjee, M. D. Starking-Szymanski, and R. R. Shaker, *J. Locat. Based Serv.* **12**, 40 (2018)
8. Y. Liu, T. Luo, Z. Liu, X. Kong, J. Li, and R. Tan, *Habitat Int.* **47**, 113 (2015)
9. G. Geniaux and D. Martinetti, *Reg. Sci. Urban Econ.* **72**, 74 (2018)
10. M. Ceci, R. Corizzo, D. Malerba, and A. Rashkovska, *Data Min. Knowl. Discov.* **33**, 698 (2019)
11. H. Pu, K. Luo, P. Wang, S. Wang, and S. Kang, *Environ. Sci. Pollut. Res.* **24**, 4457 (2017)
12. R. S. Bivand and D. W. S. Wong, *TEST* (2018)
13. H. Li, Y. Liu, Q. He, X. Peng, and C. Yin, *J. Urban Plan. Dev.* **144**, 5018002 (2018)
14. T. Luo, R. Tan, X. Kong, and J. Zhou, *Sustainability* **11**, 2207 (2019)
15. L. Anselin, *Geogr. Anal.* **51**, 133 (2019)
16. L. Anselin, *Geogr. Anal.* **27**, 93 (1995)
17. T. N. TİEN, *Maden Tetk. ve Aram. Derg.* **156**, 179 (2018)