


## Article

# Factors Aggregating Ability and the Regional Differences among China's Urban Agglomerations

Chengliang Liu <sup>1,2,3,†</sup> , Tao Wang <sup>1</sup> and Qingbin Guo <sup>4,\*,†</sup>

<sup>1</sup> School of Urban and Regional Sciences, East China Normal University, Shanghai 200062, China; clliu@re.ecnu.edu.cn (C.L.); wtao1990@163.com (T.W.)

<sup>2</sup> Institute for Global Innovation and Development, East China Normal University, Shanghai 200062, China

<sup>3</sup> Institute of Eco-Chongming, East China Normal University, Shanghai 200062, China

<sup>4</sup> School of Economics and Management, Hainan University, Haikou 570228, China

\* Correspondence: gqbhust@aliyun.com; Tel.: +86-139-7109-8589

† These authors contributed equally to this work.

Received: 24 September 2018; Accepted: 9 November 2018; Published: 13 November 2018



**Abstract:** Continuous aggregation of socioeconomic factors is the key issue of sustainable development in urban agglomerations. To date, more attention has been paid to single urban agglomeration than to multiple agglomerations. In this paper, China's 19 urban agglomerations were selected as the case study and their spatial differences in factors aggregating ability were portrayed comparatively. Firstly, the spatial pattern of urban factors aggregating ability is relatively well distributed in all China's cases, most noticeably in the Yangtze River Delta urban agglomeration, closely followed by the Beijing-Tianjin-Hebei and the Pearl River Delta urban agglomerations. However, more significant differences on factors aggregating ability are noticeably seen between cities than among urban agglomerations. Meanwhile, the rank-size structure distribution of factors aggregating ability in China's 19 cases is in line with the Zipf's law of their urban systems, and divided into three types: Optimized, balanced, and discrete. Furthermore, the urban factors aggregation ability in one urban agglomeration is roughly negatively correlated with its primacy ratio of factors aggregating ability distribution. Lastly, urban agglomerations with higher average values of factors aggregating ability are concentrated on the three major urban agglomerations: The Yangtze River Delta, the Beijing-Tianjin-Hebei and the Pearl River Delta. Otherwise, high-high clusters in the three urban agglomerations are distinctly observed as well.

**Keywords:** factors aggregating ability; POIs; rank-size structure; spatial difference; urban agglomerations

## 1. Introduction

Currently, a magnitude of socioeconomic factors, such as capital, labor, technology and information etc., are accelerating centralization in densely-populous and developed urbanized regions, which results in a higher ability to aggregate factors in these regions. Furthermore, the factors aggregating process positively accelerates the emergence and growth of various urban agglomerations [1], noticeably in high-speed developing countries like China. Therefore, factors aggregating ability is widely regarded as a way of identifying their sustainable development.

An urban agglomeration is a highly developed urbanized area consisted of integrated cities [2], similar to various terms including megalopolis [3], metropolitan belt [4], urban clusters [5], metropolitan regions [6], metropolitan areas [7], and others. Although it is still difficult to obtain a unified definition, it is generally viewed as an urbanized area filled by an integration of cities with huge population density and close interaction flows. In China, urban agglomerations are

gradually becoming vital growth poles of new-style urbanization development and the main body of China's participation in global competition, which depends on the support of continuous factor flows. Driven by "rent-seeking behavior" [8], variety of factors such as talents, technology, knowledge, information, and capital obey profit-seeking law, which leads to their continuous concentration on advantageous places [9]. As a result, the more competitive urban agglomerations tend to occupy a higher position in national urban systems and thus gain more developing opportunities. Moreover, their competitiveness and potential are accumulatively enlarged under the cyclical cumulative causal effect or Matthew effect. Therefore, the factors aggregating ability has become a key criterion to measure the sustainable competitiveness of urban agglomerations in national and inter-country competition contexts.

Although factors aggregating ability is specially measured and documented by few scholars at present, much more attention is paid to theoretically refer to factors and their aggregation, such as factor endowment theory and factor flow theory. Generally, factors were classified to primary production factors (i.e., general natural resources such as land, water, forests, minerals) and created production factors (including knowledge and technology, innovation, capital and facility, etc.) by Porter [10]. Oskenbayev et al. further pointed out that the role of the first factors such as land, oil and coal in promoting regional development has gradually weakened, while the later including capital, science and technology, talent, information, and institution factors has become increasingly decisive for regional development [11]. In the 1930s, Ohlin and Heckscher [12] proposed the factor endowment theory (H-O model), which indicates that one region with abundant low-cost production factors has an obvious cost advantage and competitive advantage. Russo & Musolino [13], and Trippi et al. [14] argued that urban economic growth depends on whether the city has factor endowments that are different from other cities. Then, the significant role of factors flows in economic development was extensively revealed. Chang & Oxley found that the inflow of young, energetic and skilled labor outside the region would form an agglomeration advantage and increase the growth rate of the entire national economy [15]. Hochman's research on inter-city factor flows evidenced a significant difference in factor aggregating abilities between different levels and different regional cities [16]. Fu & Gabriel [17] certified that that in a stage with rapid development, the ability of a city to gather various factors was gradually enhanced, particularly in the clustering of capital and human resources. In recent decades, the agglomeration of factors has been well documented. More concerns are focused on created production factors than on primary factors, such as financial factors [18], talent factors [19], and technological innovation factors [20], among others. Some quantitative methods are introduced into the field. Alfaro & Chen [21] used the EDA method to construct a factors aggregation benefit evaluation model. Ellison [22] constructed the agglomerating index to comprehensively calculate and analyze the inter-urban differences of factors aggregating ability. Alternatively, a network analysis was introduced to visualize the factors clustering pattern in geospatial by Catini et al. [23].

In summary, the factors aggregating ability, referring to the ability of a city or urban region to aggregate socioeconomic factors flows, is extensively viewed as an indicator evaluating the competitiveness and sustainability of urban agglomerations. It has been well documented in the nexus with urban growth and regional development, along with national competition. However, to date, its measurement and mapping in urban agglomerations has been little done in detail. On the one hand, because of the limits of data mining, the calculation of factors aggregating ability needs still be improved. On the other hand, as pointed out by Hochman [16], there widely exists differences of ability between cities or regions, which isn't portrayed well in spatial forms. Less attention has been paid to the geographical differences among urban agglomerations with different levels. So, several questions need to be answered urgently: How to comprehensively measure the factors aggregating ability based on the context of China's urban agglomerations? What are their differences and on what grounds are these differences viewed? How can this be explained? Answers to these questions will bring new insights to sustainable development in urban agglomerations. Therefore, in this paper, first of all, we constructed a set of comprehensive index system for evaluating factors aggregating

ability in urban agglomerations. Then, 19 major urban agglomerations in China were introduced to comparatively investigate their differences of factors aggregating ability from three perspectives, including hierarchical distribution, rank-size distribution, and spatial distribution.

## 2. Study Area and Data

### 2.1. Study Area

China's "13th Five-Year Plan" issued by the National Development and Reform Commission of China (<http://www.ndrc.gov.cn/zcfb/zcfbghwb/>), has been proposed to promote the sustainable development of 19 key urban agglomerations, such as the Beijing-Tianjin-Hebei, the Yangtze River Delta, the Pearl River Delta, the Shandong Peninsula, West Coast of the Straits, Harbin-Changchun, Central-southern of Liaoning, Central Plains, the middle reaches of the Yangtze River, Chengdu-Chongqing, Central Shaanxi Plain, Beibu Gulf, the Jinzhong regions, Hohhot-Baotou-Ordos-Yulin, Central Guizhou, Central Yunnan, Lanzhou-Xining, and the areas along the Huanghe River in Ningxia, and the northern slopes of the Tianshan Mountains urban agglomerations (Figure 1), which include a total of 198 cities. In 2016, the 19 urban agglomerations had a population of 907 million, accounting for 65.58% of total GDP in China; their total GDP was 63.57 trillion CNY and 85.44% of total China. These urban agglomerations are the most fundamental areas supporting China's land development and also play a vital role in China's participation in global competition.

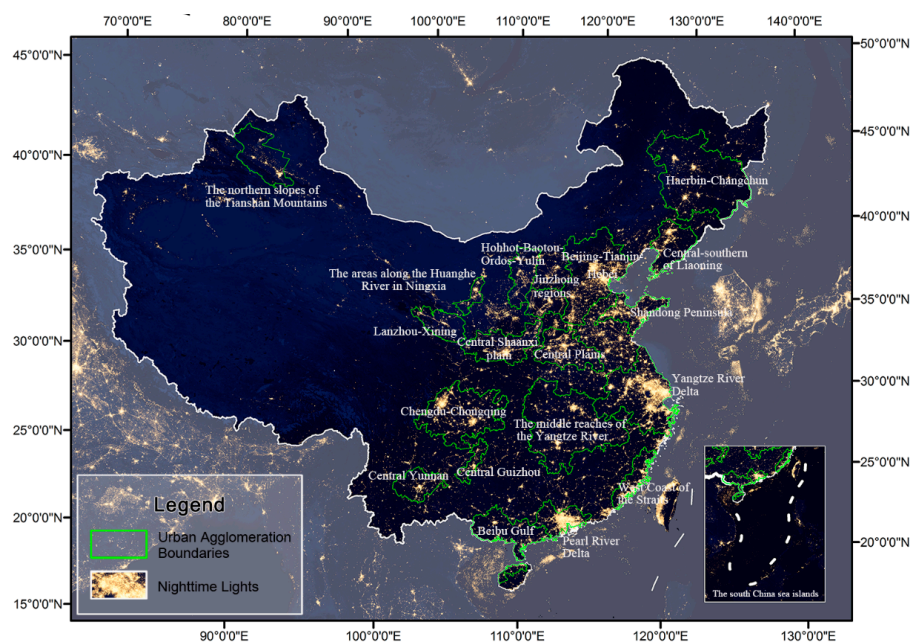


Figure 1. Location of 19 urban agglomerations in China.

### 2.2. Data Sources

In the paper, the data of index system originated from two sources. Some statistical data were derived from the authoritative statistical yearbooks, including the 2017 China Urban Statistical Yearbook and the 2017 Provincial and Municipal Statistical Yearbook. The remainder of the data came from the Point of Interest (POI) data from AMAP (open electronic map) of the Alibaba Group (<https://lbs.amap.com/console/show/picker>), similar to Google. A POI data refers to a spatial landmark point on an electronic map based on Location Based Services (LBS) [24], and includes attribute features such as name, category, latitude and longitude, etc. As some of the most essential data for urban spatial analysis, POIs data could directly and effectively reflect the aggregation of various urban factors [25]. Since a POI data has a position attribute, it can be regarded as a point of

a certain kind of factors, and the number of POIs (factors) in different areas can be obtained by the use of spatial statistics. For that reason, this paper used ArcGIS software to count the number of POIs of a certain kind of factors in China and an urban agglomeration, and calculated the proportion of the number of factors of the urban agglomeration to the total number of factors in China. As such, we obtained the index value of this kind of factors in the urban agglomeration. In the same way, the calculation of urban factors aggregating ability also adopted this method, that is, the values of evaluation indices were obtained by calculating the proportion of the number of factors of the city to the total number of factors in the urban agglomeration. By using data mining technology of web crawler, a total of 1.479 million POIs data was captured and retrieved in December 2016. To data processing, the basic data of each city were collected from the statistical yearbook and the electronic map POIs, and then aggregated into the data of the urban agglomerations. While, the national level data was directly obtained from the national statistical yearbook and the POIs in the country.

The vector data used in the mapping of this paper is derived from the 1:1 million basic geographic database (2015) of the National Basic Geographic Information Center in China (<http://www.webmap.cn/commres.do?method=dataDownload>), and the map projection type is Gauss-Kruger projection.

### 3. Methodologies

The purpose of this paper is to analyze the differences of factors aggregating ability among 19 urban agglomerations in China. In order to achieve this goal, as shown in Figure 2, we mainly designed the research framework from three methods: The evaluation model of factors aggregating ability, the Rank-size model and the Kernel Density Estimation model. Firstly, we constructed the evaluation model of factors aggregating ability by designing a comprehensive evaluation index system, in order to obtain the basic calculation result of the factors aggregating ability in urban agglomerations. Secondly, we used the calculation result data as the analysis source and introduced the Rank-size model and the Kernel Density Estimation model, to achieve the main research objective of this paper—interpreting the rank-size and the spatial differences of the factors aggregating ability among China’s urban agglomerations. Finally, we draw our conclusions. In the research process, the construction of the evaluation model of factors aggregating ability is the precondition and analysis basis for the use of the Rank-size model and the Kernel Density Estimation model.

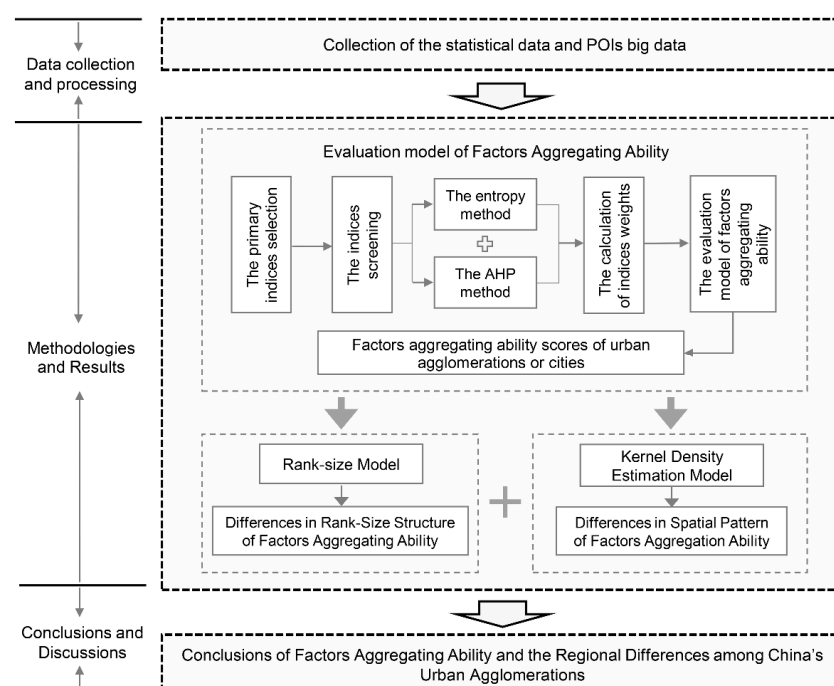


Figure 2. Research framework.

### 3.1. Evaluation Model of Factors Aggregating Ability

In this paper, we constructed a measurement model of factors aggregating ability by designing a comprehensive evaluation indicator system and determining indices weights as follows: the primary indices selection, the indices screening, and the calculation of indices weights. First of all, in order to avoid subjectivity, the frequency statistics and expert consultation methods were implemented to design the index system. Plenty of indicators with high-frequency adoption or occurrence in literature were firstly selected and constructed a preliminary index system by using the frequency statistical method. Then, the preliminary indices were further modified based on expert comments in the expert consultation method. As a result, we initially identified an evaluation index system with 40 indices, including population, land, economy, facility, ecology, science and technology, and finance aspects, and so on.

We also re-screened and reduced the primary indices by calculating correlation coefficients, coefficient of variation and factor analysis, to avoid interdependence and duplication between indices. First, the indices value was standardized. Secondly, the correlation coefficient and coefficient of variation of each index were calculated separately. Then, those indices with correlation coefficient greater than 0.9 and lower coefficient of variation were excluded. Thirdly, factor analysis was used to determine the principal components of the index system that cannot be deleted. Finally, on the basis of the above, we deleted 14 indices, and constructed the comprehensive index system containing 26 indices (Table 1).

Last but not least, we obtained the weights of indices by integrating subjective AHP weighting method and objective entropy weighting method, for the sake of avoiding their own disadvantages. Amongst them, the AHP method exists subjective judgments and random scoring problems despite of certain rationality in judging information importance. The entropy method may avoid the subjective problem, but easily ignores the difference in the actual importance of the indices.

First, the entropy method was used to calculate the index weights. The proportion  $y_{ij}$  of sample  $i$  under the  $j$  indicator was calculated by:

$$y_{ij} = x''_{ij} / \sum_{i=1}^n x''_{ij} \quad (i = 1, 2, \dots, n; j = 1, 2, \dots, m) \quad (1)$$

where  $n$  denotes the number of samples,  $m$  denotes the number of indicators, and  $x''_{ij}$  represents the value of sample  $i$  of index  $j$ .

The information entropy  $e_j$  of the  $j$  indicator was defined as:

$$e_j = -k \sum_{i=1}^n y_{ij} \ln(y_{ij}) \quad (2)$$

where  $k = 1/\ln(n)$ , with  $0 \leq e_j \leq 1$ .

The difference coefficient  $g_i$  of the  $j$  indicator  $g_i$  was denoted as follows:

$$g_i = 1 - e_i$$

After normalizing the difference coefficient of each indicator, the entropy weight of the  $j$  indicator was calculated by [26]:

$$\omega_{ij} = g_j / \sum_{j=1}^m g_j \quad (3)$$

Besides, on the basis of the entropy weights, the AHP method is used to calculate the relative importance of each index to the criterion layer index. The weight coefficient  $\omega_{bj}$  of each index is determined.



Furthermore, the index weights  $\omega_{aj}$  and  $\omega_{bj}$  obtained from the two methods further determined the final weight of the index, defined as:

$$F = \sum_{j=1}^m \omega_j (\ln \omega_j - \ln \omega_{aj}) + \sum_{j=1}^m \omega_j (\ln \omega_j - \ln \omega_{bj}) \quad (4)$$

where  $\sum_{j=1}^m \omega_j = 1$ ,  $\omega_j > 0$ . By using the Lagrangian multiplier method, the final weight  $\omega_j$  was obtained as:

$$\omega_j = \sqrt{\omega_{aj}\omega_{bj}} / \sum_{j=1}^m \sqrt{\omega_{aj}\omega_{bj}} \quad (5)$$

where  $\omega_{aj}$  is the index weight determined by the entropy method, and  $\omega_{bj}$  is the index weight determined by the AHP method.

On this basis, the evaluation model of factors aggregating ability was constructed as follows:

$$A = \sum_{j=1}^m \omega_j x''_{ij} \quad (j = 1, 2, \dots, m) \quad (6)$$

Factors aggregating ability of urban agglomerations represented the factors aggregating ability of one urban agglomeration in China compared to that of other urban agglomerations. Its main significance was reflected in the relative advantages or disadvantages of factors aggregation ability between urban groups [27]. The factors aggregating ability of urban agglomerations ranged from 0 to 1. The larger the  $A$  value is, the greater is the factors aggregating ability of urban agglomerations, and the stronger is the concentration of various factors.

### 3.2. Rank-Size Model

An urban agglomeration is a typical multi-level urban system composed of large, medium, and small cities. Consequently, in this study, the Lotka rank-size model was introduced into revealing the hierarchical distribution and the scale structure of urban factors aggregating ability [28–30], written by:

$$P_i = P_1 \times R_i^{-q} \quad (i = 1, 2, \dots, n) \quad (7)$$

where  $n$  represents the number of cities and  $R_i$  represents the rank of factors aggregating ability of the city  $i$  in the urban agglomeration.  $P_i$  denotes the value of urban factors aggregating ability with the rank of  $R_i$  after sorting from the large to the small.  $P_1$  denotes the factors aggregating ability value of the first city. The parameter  $q$  denotes the Zipf's coefficient. Through the logarithm processing, equator (7) was further transformed to:

$$\ln P_i = \ln P_1 - q \ln R_i \quad (8)$$

where if  $q = 1$ , the aggregating characteristics of the urban factor aggregating ability is in an optimal distribution satisfying the Zipf's criterion at this time. If  $q < 1$ , the scale distribution of urban factors aggregating ability is relatively concentrated with many middle-rank cities. If  $q > 1$ , the urban factors aggregating ability tends to be dispersed, and the differences among urban agglomerations are quite significant. The first city has a strong monopoly of factors aggregating ability as well.

### 3.3. Kernel Density Estimation

Kernel Density Estimation (KDE) analysis is a process of interpolating by the discrete point data or line data using the kernel function [31]. In this study, KDE was used to estimate the density distribution of urban factors aggregating ability. Let  $x_1, \dots, x_n$  be independent distribution samples

extracted from the population with a distribution density function of  $f$ . Estimate the value  $f(x)$  of  $f$  at a point  $x$ , usually with the Rosenblatt-Parzen kernel estimation equation [32]:

$$f_n(x) = \frac{1}{nh^d} \sum_{i=1}^n k\left(\frac{x - x_i}{h}\right) \quad (9)$$

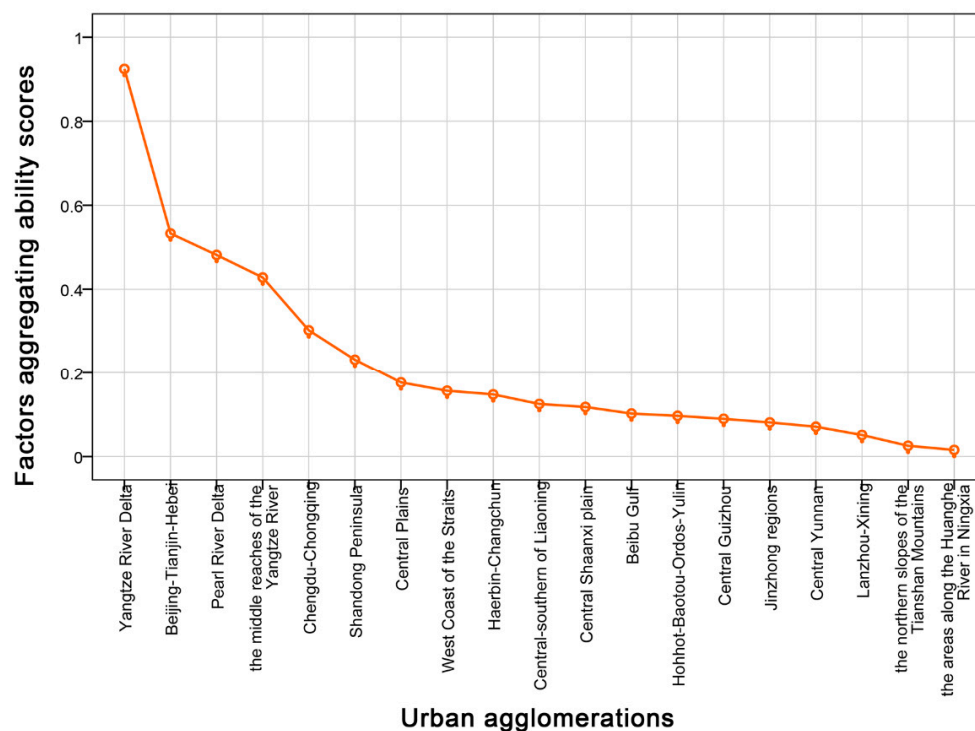
where  $k$  denotes the kernel function;  $h > 0$  denotes the bandwidth;  $(x - x_i)$  denotes the distance from the estimated point to the sample  $x_i$ ;  $n$  is the number of known points in the bandwidth; and  $d$  is the dimension of the data.

In the KDE analysis, the bandwidth  $h$  has a greater impact on final calculation results. Generally, the smaller bandwidth is suitable to reflect local variation of the density distribution, and the larger bandwidth can effectively reflect spatial variation at global scale. In this study, the distributional difference of China's 19 urban agglomerations needs be considered. The smallest length among the 19 urban agglomerations is about 400–500 km, therefore, it is necessary to avoid the spatial differentiation being over-enhanced or weakened. According to the study of bandwidth [33], we repeated test with several bandwidth of 80, 100, 120, 150 and 180 km, respectively and found that the KDEs remain stable between 100–180 km. Accordingly, a bandwidth of 120 km is more reasonable and appropriate to be adopted in this study.

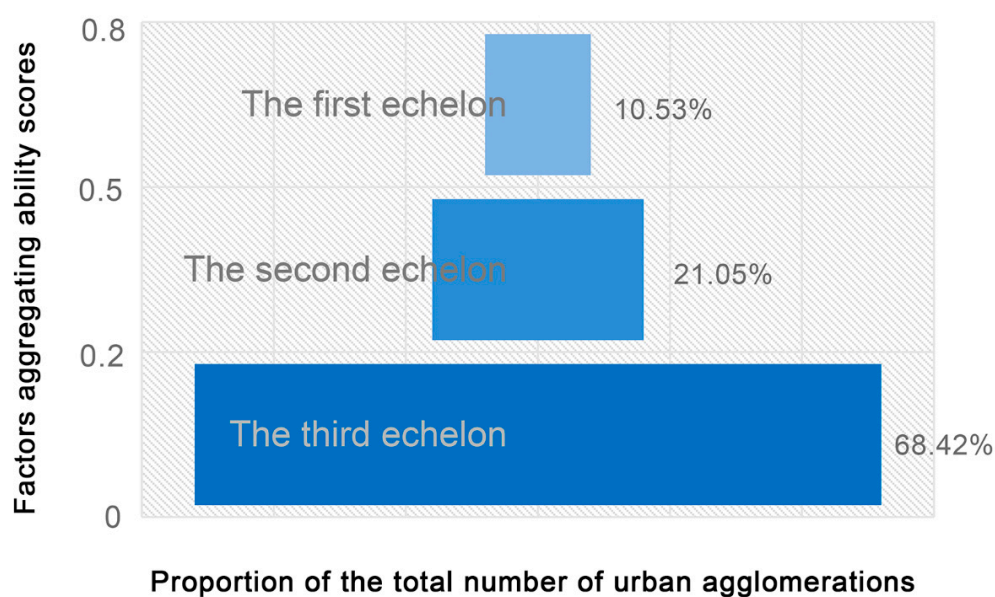
## 4. Results

### 4.1. Differences in Hierarchical Distribution of Factors Aggregating Ability

Figure 3 shows the difference and hierarchical distribution of the factors aggregating ability among 19 urban agglomerations in China based on Equation (6). Statistically, as shown in Figure 3a, the difference of factors aggregating ability among the 19 cases isn't relatively significant, except for the first-rank Yangtze River Delta urban agglomeration. The top three, consistent with current development status, are the three national-level urban agglomerations of the Yangtze River Delta, Beijing-Tianjin-Hebei and the Pearl River Delta. They are economically developed, densely populated, and open to the outside world. The cases in the middle rank include the middle reaches of the Yangtze River, Shandong Peninsula, Chengdu-Chongqing, West Coast of the Straits, Central Plains, Central-southern of Liaoning, and Harbin-Changchun, which are the key development areas in national land functional planning. The group with lower orders consisted of Central Shaanxi Plain, the northern slopes of Tianshan Mountains, Hohhot-Baotou-Ordos-Yulin, Lanzhou-Xining, Central Guizhou province, the Jinzhong regions, Central Yunnan, and the areas along the Huanghe River in Ningxia province. The vast majority of these areas are located in the western regions of China, and are viewed as the growth poles of their local regions.



(a) Factors aggregating ability scores of 19 urban agglomerations



(b) The hierarchical distribution of factors aggregating ability scores

**Figure 3.** The hierarchical distribution of factors aggregating ability among the 19 urban agglomerations.

On the basis of the comprehensive factors aggregating ability scores, China's 19 major urban agglomerations can be roughly divided into three tiers (Figure 3b): In the first echelon, there are two cases accounting for 10.53%, that's, the Yangtze River Delta and the Beijing-Tianjin-Hebei urban agglomerations, and their scores of the factors aggregating ability are above 0.5; the second echelon is composed of four urban agglomerations whose average scores are between 0.2 and 0.5, accounting for 21.05% of total number; the third echelon includes the rest of 19 cases with average score below



0.2, accounting for 68.42%. According to their proportion, the three echelons are visualized as the “pyramid” distribution as shown in Figure 3b, which is in accordance with the current development level of China’s urban agglomerations. Most of urban agglomerations are still in the early and middle industrialization stages, with relatively lower factors aggregating ability. It will take a long period of improvement to continuously gather the various factors and further to promote the integration of urban agglomeration.

Table 2 illustrates the aggregating abilities between comprehensive factors and various factors. In the same way as the comprehensive factors, the aggregating ability of all kinds of factors among the 19 cases follows a decreasing gradient trend from the east to the west. That is, the factors aggregating ability of urban agglomerations in the eastern coastal areas is mostly at the highest level, which is followed by the cases in central China and in the western regions in turn. Meanwhile, the pyramid structure consisted of the three echelons is also observed in the aggregating ability distribution of all kinds of factors. Amongst them, the first echelon group of all factors is almost entirely composed of the three-big national urban agglomerations of the Yangtze River Delta, the Pearl River Delta, and the Beijing-Tianjin-Hebei regions. However, the number and distribution of urban agglomerations in the second echelon and the third echelon exhibit some differences among various factors. Specifically, there are larger differences in the second and third echelons of these factors including population, land, economy, public facilities, and ecological environment. Similar findings are seen in the third echelon of such factors as finance, scientific and technological innovation, economic openness, and culture.

**Table 1.** Comprehensive evaluation index system and index weight of factors aggregating ability.

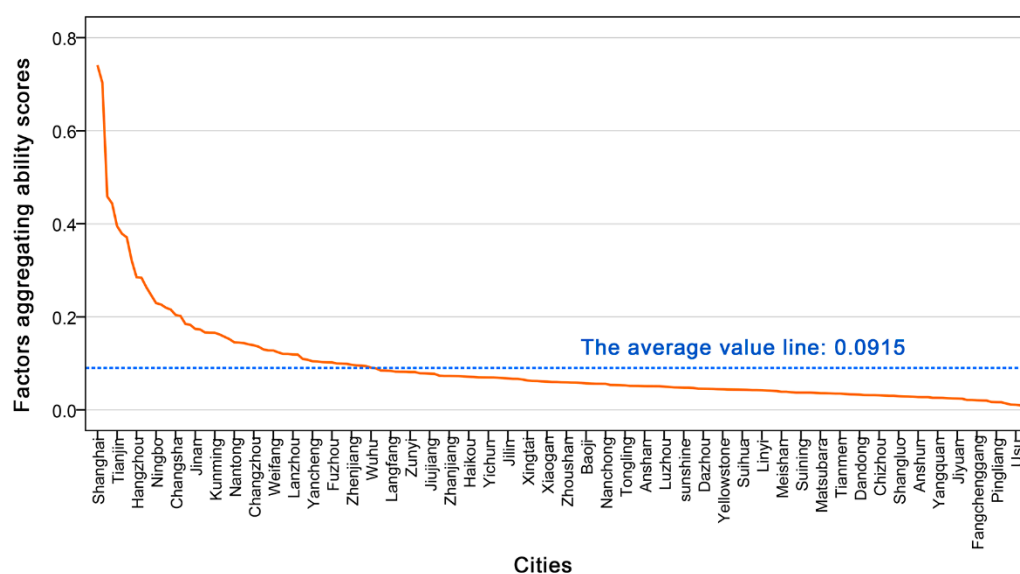
Target Layer	Criteria Layer	Indicators Layer	Indicator Weight
Factors Aggregating Ability	Population Factor Aggregating Ability (0.1084)	The proportion of the population in the urban agglomeration to that in the country	0.0672
		The ratio of college students per 10,000 students in the urban agglomeration to the national indicator level	0.0412
	Land Factor Aggregating Ability (0.0734)	The proportion of the built-up area in the urban agglomeration to that in the country	0.0570
		The ratio of unit land area output rate in the urban agglomeration to the national indicator level	0.0164
	Economy Factor Aggregating Ability (0.1163)	The proportion of GDP in the urban agglomeration to that in the country	0.0786
		The ratio of fixed asset investment intensity in the urban agglomeration to the national indicator level	0.0377
	Financial Factor Aggregating Ability (0.099)	The proportion of the deposits of financial institutions at the end of the year in the urban agglomeration to those in the country	0.0494
		The proportion of bank outlets in the urban agglomeration to those in the country	0.0496
	Technological innovation Factor Aggregating Ability (0.1568)	The ratio of professional and technical personnel per 10,000 people in the urban agglomeration to the national indicator level	0.0373
		The proportion of high-tech enterprises in the urban agglomeration to those in the country *	0.0572
		The proportion of invention patent applications in the urban agglomeration to those in the country	0.0623
	Public Facility Factor Aggregating Ability (0.1500)	The ratio of road network density in the urban agglomeration to the national indicator level *	0.0256
		The proportion of bus stations in the urban agglomeration to those in the country *	0.0181
		The proportion of commercial supermarket service agencies in the urban agglomeration to those in the country *	0.0196
		The proportion of catering service agencies in the urban agglomeration to those in the country *	0.0191
		The proportion of hotel service agencies in the urban agglomeration to those in the country *	0.0198
		The proportion of primary and secondary schools in the urban agglomeration to those in the country *	0.0235
		The proportion of medical institutions in the urban agglomeration to those in the country *	0.0243
	Cultural Factor Aggregating Ability (0.0569)	The ratio of cultural industry revenue as a share of GDP in the urban agglomeration to the national indicator level	0.0273
		The proportion of art institutions in the urban agglomeration to those in the country *	0.0296
	Ecological Environment Factor Aggregating Ability (0.0614)	The ratio of environmental governance investment as a share of fiscal expenditure in the urban agglomeration to the national indicator level	0.0285
		The proportion of parks in the urban agglomeration to those in the country *	0.0329
	Public Facility Factor Aggregating Ability (0.0614)	The proportion of government public finance income and expenditure ratio in the urban agglomeration to the national indicator level	0.0151
		The ratio of private economy added value as a share of GDP in the urban agglomeration to the national indicator level	0.0463
	Opening-Up Factor Aggregating Ability (0.1164)	The proportion of actual inflow of foreign investment in the urban agglomeration to that in the country	0.0556
		The proportion of inbound tourists received in the urban agglomeration to those in the country	0.0608

Note: The data of these indices marked with a star symbol (\*) after the text are derived from POIs big data.

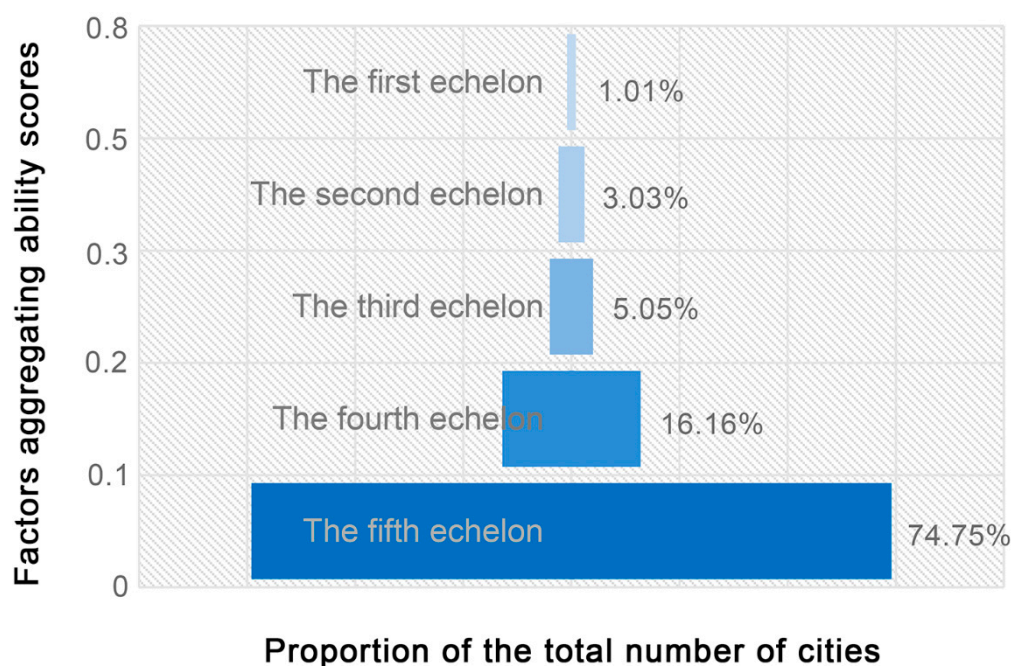
**Table 2.** Factors aggregating ability scores of 19 urban agglomerations in China.

Urban Agglomerations	Comprehensive Factors	Population Factor	Land Factor	Economic Factor	Financial Factor	Technological Innovation Factor	Public Facility Factor	Cultural Factor	Ecological Environment Factor	Policy Factor	Opening-Up Factor
Yangtze River Delta	0.9245	0.0832	0.0677	0.1163	0.0990	0.1449	0.1400	0.0509	0.0472	0.0588	0.1164
Beijing-Tianjin-Hebei	0.5330	0.0566	0.0415	0.0617	0.0761	0.0815	0.0708	0.0453	0.0283	0.0146	0.0564
Pearl River Delta	0.4817	0.0448	0.0441	0.0440	0.0384	0.0699	0.0658	0.0186	0.0501	0.0405	0.0655
The middle reaches of the Yangtze River	0.4280	0.0779	0.0349	0.0663	0.0339	0.0380	0.0826	0.0340	0.0162	0.0119	0.0324
Chengdu-Chongqing	0.3023	0.0408	0.0294	0.0420	0.0308	0.0255	0.0769	0.0200	0.0111	0.0074	0.0183
Shandong Peninsula	0.2318	0.0526	0.0238	0.0341	0.0177	0.0149	0.0257	0.0111	0.0098	0.0226	0.0195
Central Plains	0.1763	0.0352	0.0193	0.0271	0.0154	0.0107	0.0350	0.0089	0.0043	0.0087	0.0118
West Coast of the Straits	0.1565	0.0287	0.0222	0.0176	0.0092	0.0077	0.0191	0.0060	0.0092	0.0188	0.0182
Harbin-Changchun	0.1478	0.0267	0.0201	0.0199	0.0142	0.0087	0.0233	0.0118	0.0067	0.0035	0.0129
Central-southern of Liaoning	0.1249	0.0188	0.0133	0.0104	0.0142	0.0095	0.0251	0.0080	0.0095	0.0105	0.0057
Central Shaanxi Plain	0.1178	0.0220	0.0146	0.0159	0.0120	0.0103	0.0241	0.0100	0.0016	0.0007	0.0065
Beibu Gulf	0.1021	0.0226	0.0137	0.0124	0.0081	0.0048	0.0191	0.0056	0.0074	0.0063	0.0020
Hohhot-Baotou-Ordos-Yulin	0.0967	0.0369	0.0165	0.0091	0.0027	0.0026	0.0035	0.0030	0.0087	0.0103	0.0033
Central Guizhou	0.0895	0.0461	0.0078	0.0035	0.0020	0.0022	0.0162	0.0017	0.0008	0.0077	0.0014
Jinzhong regions	0.0809	0.0440	0.0099	0.0036	0.0046	0.0021	0.0042	0.0038	0.0022	0.0052	0.0012
Central Yunnan	0.0707	0.0214	0.0135	0.0049	0.0027	0.0039	0.0094	0.0035	0.0026	0.0079	0.0009
Lanzhou-Xining	0.0511	0.0223	0.0079	0.0027	0.0025	0.0024	0.0064	0.0031	0.0010	0.0003	0.0026
The northern slopes of the Tianshan Mountains	0.0255	0.0036	0.0030	0.0008	0.0010	0.0018	0.0031	0.0007	0.0012	0.0102	0.0003
The areas along the Huanghe River in Ningxia	0.0156	0.0015	0.0002	0.0004	0.0001	0.0001	0.0006	0.0007	0.0092	0.0026	0.0001

Similarly, an obvious heterogeneity (polarization and hierarchy) of the factors aggregating ability among 198 cities is observed as well, as shown in Figure 4. The urban factors aggregating ability tends to exhibit a scale-free trait with better-fitting power-law distributions (Figure 4a), that's, a few cities have higher scores of factors aggregating ability, which is similar to the rule of Pareto's distribution. Only two megacities, Shanghai and Beijing, score above 0.5. Almost all of the cities have a lower factors aggregating ability. Obviously, the scores of 141 cities are much lower than the average of 0.0915 and there exists a huge gap with a few megacities such as Shanghai, Beijing, Guangzhou, and Shenzhen. As exhibited as Figure 3b, according to urban scores, these cities are divided into five echelons (greater than 0.5, 0.3–0.5, 0.2–0.3, 0.1–0.2, and less than 0.1), accounting for 1.01%, 3.03%, 5.05%, 16.16%, and 74.75%, respectively (Figure 4b).



(a) Urban factors aggregating ability scores



(b) The hierarchical distribution of urban factors aggregating ability scores

**Figure 4.** The hierarchical distribution of factors aggregating ability among the 198 cities.

To more intuitively compare the agglomeration characteristics of different factors, we calculated the variation coefficient of the 10 factors aggregating ability scores among 19 urban agglomerations and among 198 cities. Whether at the inter-city or inter-group level, the variation coefficient of the ability of technological innovation, opening-up, and finance factors are higher, indicating the pattern to be unevenly distributed, especially in larger cities and urban groups. Because of convenient interactions, well-developed public facilities, developed economies, and high input-output efficiency, they are appropriate to obtain more advantageous competitiveness when these factors are concentrated. At the same time, other kinds of factors with lower aggregating ability tend toward decentralization, such as population, economy, and public facilities, due to low requirements for location conditions and environmental factors.

#### 4.2. Differences in Rank-Size Structure of Factors Aggregating Ability

Table 3 indicates the rank-size structure of factors aggregating ability in the 19 cases. In Table 3, most of the  $R^2$  values are mostly above 0.9, and the minimum  $R^2$  value is 0.764, with a good fitting degree, which indicates that the hierarchical distribution is basically in line with the Zipf's law (Table 3). In accordance with their  $q$  values, the 19 urban agglomerations are divided into the five types with different characteristics of statistical distribution as follows. The first is optimized type structure, in line with  $0.9 < q \leq 1$ . There are 5 urban agglomerations, including the Beijing-Tianjin-Hebei, Chengdu-Chongqing, Central-southern of Liaoning, Central Shaanxi Plain, and Jinzhong urban agglomerations. It represents a more optimal rank-size structure that exhibits a well-fitting power-law distribution. The number of cities with high, medium, and low factors aggregating ability in urban agglomerations are relatively appropriate. The second is balanced type structure with  $q < 0.9$ . This type includes 5 urban agglomerations located in the western inland such as the northern slopes of the Tianshan Mountains, Lanzhou-Xining, Central Guizhou, Central Yunnan, and the areas along the Huanghe River in Ningxia. For this type, there are a majority of medium-and-small with weaker degree factors aggregating ability in these urban agglomerations, which is related to their development levels. The third is discrete type structure ( $q > 1$ ) and includes 9 urban agglomerations such as the Yangtze River Delta, Pearl River Delta, the middle reaches of the Yangtze River, Shandong Peninsula, West Coast of the Straits, Central Plains, Harbin-Changchun, Beibu Gulf, Hohhot-Baotou-Ordos-Yulin. The distribution of urban factors aggregating ability in these urban agglomerations tends to be discrete. Few hub cities have stronger factors aggregating ability in the region, and greater magnetic forces to attract socioeconomic factors or flows, resulting in the cumulative Matthew effect.

From the perspective of the primacy ratio, the factors aggregation ability of one urban agglomeration is negatively related to its primacy ratio (Table 3). Specifically, those urban agglomerations with larger factors aggregation ability, such as the Yangtze River Delta, the Pearl River Delta, Beijing-Tianjin-Hebei, the middle reaches of the Yangtze River, Shandong Peninsula, Chengdu-Chongqing, West Coast of the Straits, Central Plains, Central-southern of Liaoning, and Harbin-Changchun, their primacy ratios are much lower. Generally, one urban agglomeration with higher primacy ratio is often underdeveloped, because magnitude of factors is over aggregative in its primacy or central city, defined as the Siphon effect between its primacy city and its peripheral regions.



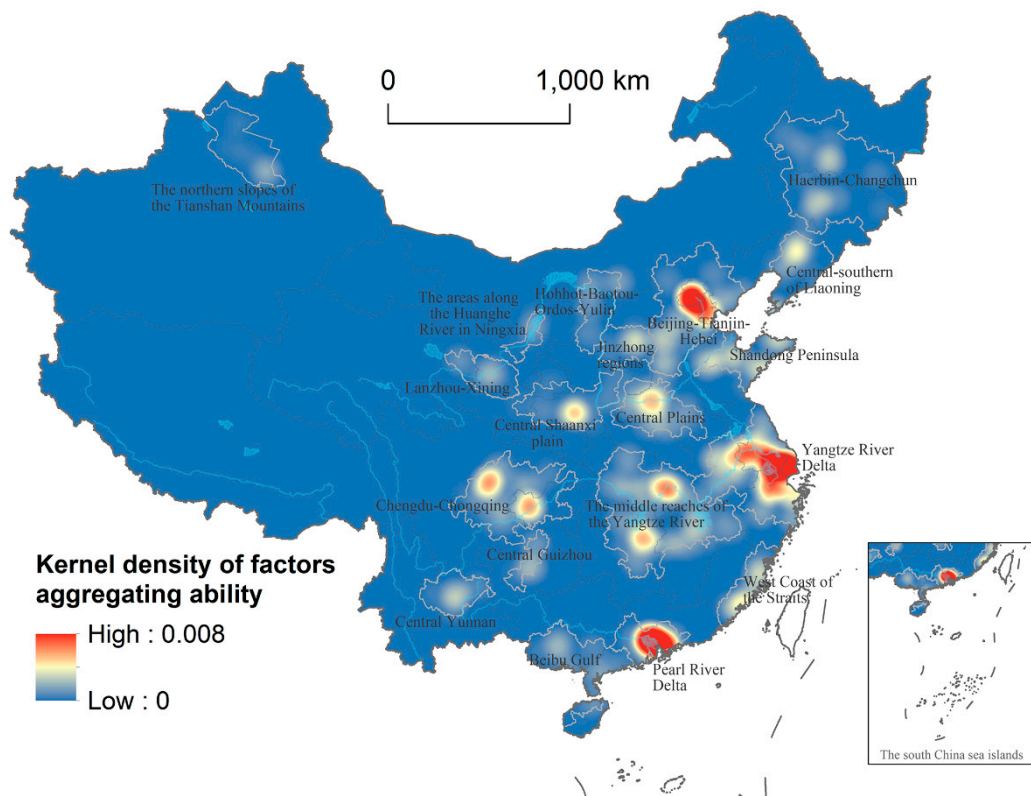
**Table 3.** Rank-size structures of factors aggregating ability of 19 urban agglomerations in China.

Urban Agglomerations	<i>q</i> Value	<i>R</i> <sup>2</sup>	Equation	Primacy Ratio
Yangtze River Delta	1.144	0.947	$y = -1.144x + 5.151$	1.19
Pearl River Delta	1.001	0.925	$y = -1.001x + 4.199$	1.01
Beijing-Tianjin-Hebei	0.941	0.966	$y = -0.941x + 3.969$	1.16
the middle reaches of the Yangtze River	1.631	0.942	$y = -1.631x + 5.495$	1.11
Shandong Peninsula	1.312	0.885	$y = -1.312x + 4.434$	1.08
Chengdu-Chongqing	0.997	0.858	$y = -0.997x + 3.673$	1.05
West Coast of the Straits	1.057	0.764	$y = -1.057x + 3.459$	1.04
Central Plains	1.328	0.922	$y = -1.328x + 4.064$	1.33
Central-southern of Liaoning	0.981	0.892	$y = -0.981x + 3.028$	1.01
Harbin-Changchun	1.075	0.929	$y = -1.075x + 3.362$	1.01
Beibu Gulf	1.300	0.933	$y = -1.300x + 3.548$	1.29
Central Shaanxi Plain	0.944	0.925	$y = -0.944x + 0.925$	1.59
The northern slopes of the Tianshan Mountains	0.766	0.924	$y = -0.766x + 1.74$	2.45
Hohhot-Baotou-Ordos-Yulin	2.267	0.849	$y = -2.267x + 5.366$	1.06
Lanzhou-Xining	0.858	0.958	$y = -0.858x + 2.09$	1.51
Central Guizhou	0.641	0.880	$y = -0.641x + 1.818$	1.84
Jinzhong regions	0.992	0.995	$y = -0.992x + 2.515$	1.42
Central Yunnan	0.803	0.941	$y = -0.803x + 2.216$	1.73
The areas along the Huanghe River in Ningxia	0.749	0.998	$y = -0.749x + 1.481$	1.98

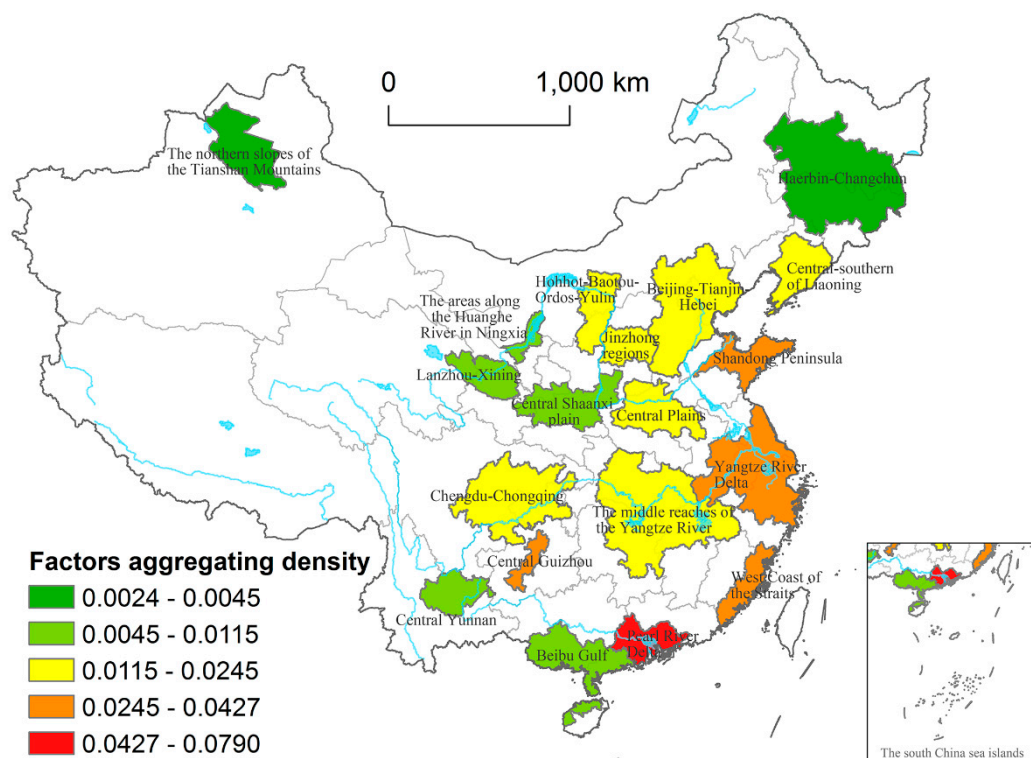
#### 4.3. Differences in Spatial Pattern of Factors Aggregation Ability

Figure 5 portrays the spatial distribution of kernel density and factors aggregating density of factors aggregating ability in the 19 groups. From Figure 5, a declining gradient from the East to the West is observed clearly. In detail, the urban agglomerations with the greatest values of factors comprehensive aggregating ability are mainly centralized in the eastern coastal regions of the Yangtze River Delta, the Pearl River Delta, and Beijing-Tianjin-Hebei urban agglomerations. Following the three-big urban groups, the urban agglomerations, such as the middle reaches of the Yangtze River and Central Plains in Central China, along with Chengdu-Chongqing in Western China, hold a higher average ability of factors aggregation. At the same time, almost all of the western urban agglomerations have the lowest comprehensive ability, including Central Guizhou, Central Yunnan, the areas along the Huanghe River in Ningxia, the northern slopes of the Tianshan Mountains, Lanzhou-Xining, and Hohhot-Baotou-Ordos-Yulin urban agglomerations. Overall, the declining gradient pattern from east to west is consistent with the regional differences of China's socioeconomic geography, such as GDP and population distributions.

Figure 6 maps the spatial pattern of urban factors aggregating ability at the city level. It can be seen that the cities with higher factors aggregating ability are densely distributed in the urban agglomerations such as Yangtze River Delta, Pearl River Delta, Beijing-Tianjin-Hebei, the middle reaches of the Yangtze River, Shandong Peninsula, Chengdu-Chongqing, Central Plains, and Central-southern of Liaoning. This aggregative pattern is basically consistent with the spatial distribution of China's population density. Most of cities with higher-degree factors aggregating ability are concentrating on the southeastern orientation of the "Hu Huanyong population dividing line" (In a paper published in 1935 entitled "Population Distribution in China," Hu Huanyong found that spatial distribution of population in China is divided into two basic areas: Southeast China and Northwest China, with the dividing line named as the Heihe-Tengchong line (or Hu line for short)), where about 95% of inhabitants reside. Recently in China, various socioeconomic factors are accelerative flowing into densely-populated cities, to seek greater economic efficiency under a consumer-based orientation.

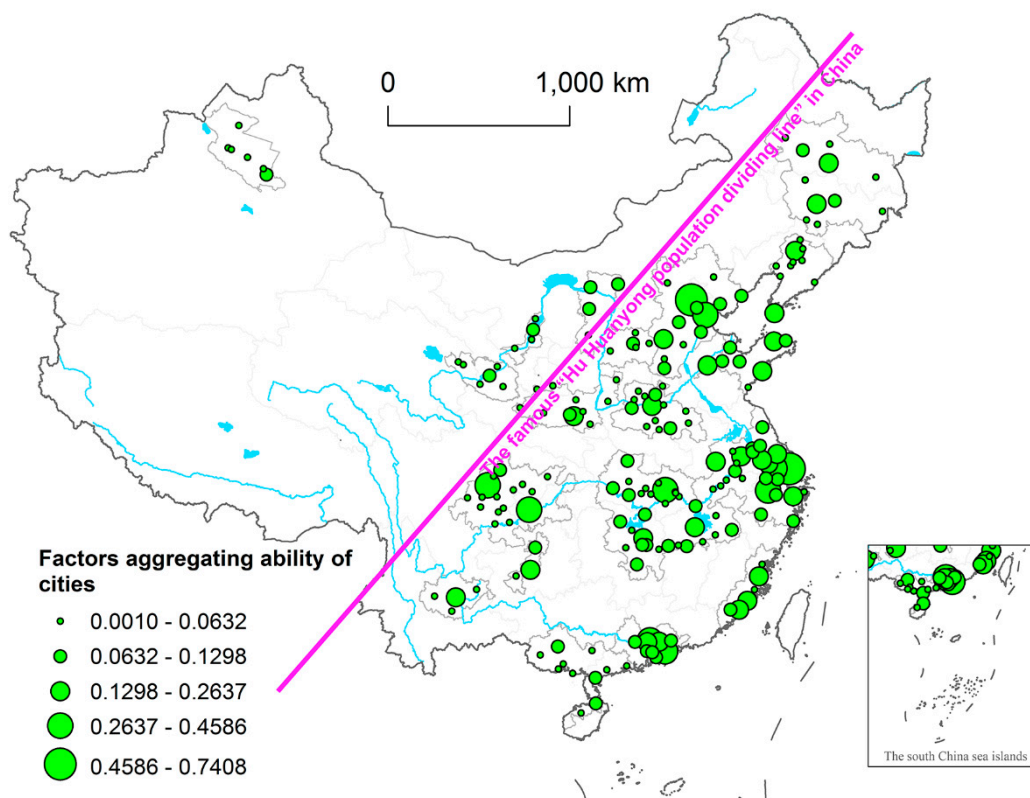


(a) The KDEs of factors aggregating ability



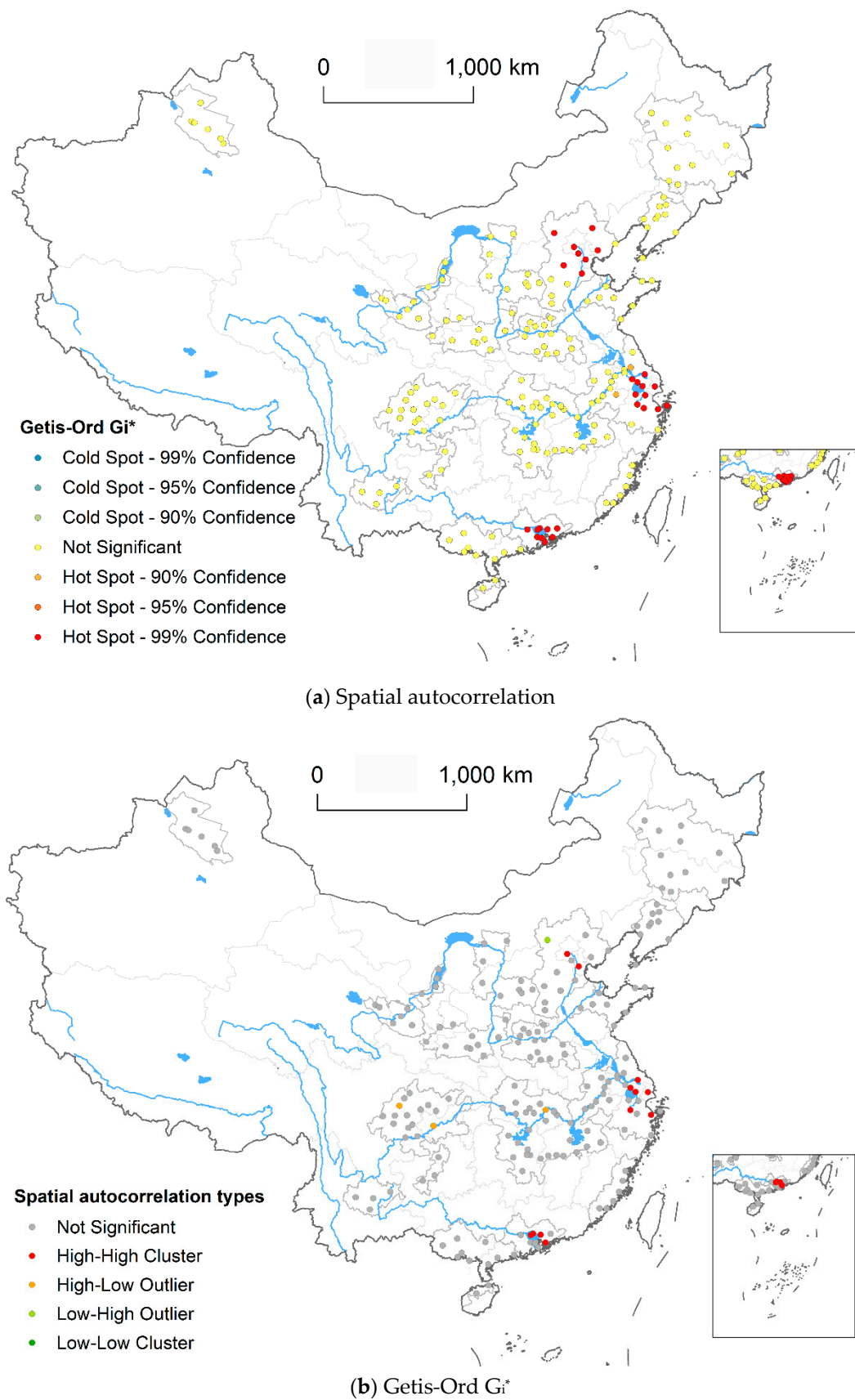
(b) Spatial distribution of factors aggregating ability

**Figure 5.** Spatial patterns of factors aggregating ability of 19 urban agglomerations in China.



**Figure 6.** Spatial patterns of factors aggregating ability of 198 cities in China.

Figure 7 depicts the local clusters of factors aggregating ability of 198 cities by the local spatial autocorrelation analysis. At a glance, there is not significant spatial interdependence of factors aggregating ability in many cities. A few cities have a stronger spatial autocorrelation, and exhibit mainly high-high agglomeration. They are mainly centralized in the Yangtze River Delta, Beijing-Tianjin-Hebei, and the Pearl River Delta urban agglomerations (Figure 7a). In addition, from the cold spots and hotspots distribution as shown in Figure 7b, it is clear that hotspots (high-high clusters) under 99% confidence are polarizing in the Yangtze River Delta, the Pearl River Delta, and the Beijing-Tianjin-Hebei urban agglomerations, with higher factors aggregating ability. However, cold spots (low-low clusters) and other outliers are almost absent in China, which indicates further that the socioeconomic factors are continuously clustering in the three-big urban agglomerations, but that other urban agglomerations are not significant at the factors aggregation.



**Figure 7.** Spatial clusters of factors aggregation ability of 198 cities in China.

## 5. Conclusions and Discussion

### 5.1. Conclusions

Currently, the relationship between factors aggregation and sustainable development in urban agglomerations has been an important issue in urban sustainability research. Much less attention is paid to comprehensive measurements and geographical differences of factors aggregating ability among urban agglomerations. Then, the two questions need to be answered: How to evaluate comprehensively factors aggregating ability? What are the differences among urban agglomerations with different geographical contexts?

To fill this gap, in this paper, China's 19 major urban agglomerations were selected as the case study. On the basis of POIs data mining, we constructed a comprehensive index system to measure factors aggregating ability. A comparative analysis of factors aggregating ability among the 19 cases were further conducted from the perspectives of structural and spatial differences.

Firstly, the statistical differences of factors aggregating ability among 19 urban agglomerations in China aren't relatively significant. The most powerful cases are the three state-level urban agglomerations of the Yangtze River Delta, Beijing-Tianjin-Hebei, and the Pearl River Delta. Otherwise, the average factors aggregating ability of urban agglomerations can be roughly divided into three echelons, exhibiting the stable "pyramid" structure. The aggregating ability of all kinds of factors among urban agglomerations exhibit differences as well, noticeably in the technological innovation, economic openness, and financial factors. At city level, the hierarchical distribution of urban factors aggregating ability is more obviously polarized. Higher values of factors aggregating ability are centralized in a few megacities such as Shanghai, Beijing, Guangzhou and Shenzhen. At the same time, most of cities have far lower factors aggregating ability.

Secondly, the rank-size distribution of factors aggregating ability is in line with the rank-size rule of the urban system in China. The rank-size structure of factors aggregating ability in urban agglomerations can be divided into three types: Optimized, balanced, and discrete. Beijing-Tianjin-Hebei, Chengdu-Chongqing, Central-southern of Liaoning, Central Shaanxi Plain and Jinzhong urban agglomerations are of the optimized structure type. The distribution of factors aggregating ability in five urban agglomerations in the western China is balanced. Nine urban agglomerations in the Yangtze River Delta, the Pearl River Delta, the middle reaches of the Yangtze River, Shandong Peninsula, West Coast of the Straits, Central Plains, Haerbin-Changchun, Beibu Gulf, Hohhot-Baotou-Ordos-Yulin have a discrete structure. Furthermore, urban factors aggregating ability in one urban agglomeration is roughly negatively correlated with its primacy ratio of factors aggregating ability distribution.

Thirdly, the spatial distribution of factors aggregating ability presents a gradient declining from the East to the West in China, forming the three-big centers with higher values of factors aggregating ability including the Yangtze River Delta, the Pearl River Delta, and Beijing-Tianjin-Hebei urban agglomerations. Moreover, the distribution of factors aggregating ability at city level coincides with the geographical distribution law of "Hu's Line", which is essentially consistent with the spatial distribution of China's population density. Furthermore, there is slightly spatial interdependence of factors aggregating ability, particularly in the three-big urban agglomerations. The high-high clusters and hotspots are concentrated in the Yangtze River Delta, the Pearl River Delta, and the Beijing-Tianjin-Hebei urban agglomerations.

### 5.2. Discussion

To date, factors flows and aggregation have played the key role in sustainable development in urban agglomerations. Due to different geographical context of urban agglomerations, there are obvious differences of factors aggregating ability between urban agglomerations. So, a comparative analysis of the differences on factor aggregating ability among China's 19 major urban agglomerations brings new insights to sustainable development in urban agglomerations. However, further research



is needed as follows. The first is the construction of the comprehensive evaluation index system. Although the POIs big data mining is introduced to supplement the lack of indicators data, it still needs further improvement at city and county scales. The second is the dynamic analysis of factors aggregating ability. Due to the huge difficulty in data collection, we have not conducted the spatiotemporal evolution of factors aggregating ability on long-term dynamics. Additionally, the changing differences of factors aggregating ability and their reasons or drivers need to be revealed in detail, which will facilitate the understanding of driving mechanism of factors aggregating ability under different urban agglomerations. The third is the spatial effect of factors aggregating ability. In future research, it would be interesting and important to examine the spatial spillover effects of factors aggregation in urban agglomerations, and to explore its impact on economic development, spatial expansion, environmental sustainability, and urbanization, etc.

**Author Contributions:** Conceptualization, C.L.; Data curation, T.W.; Formal analysis, T.W.; Funding acquisition, Q.G.; Investigation, C.L.; Methodology, C.L.; Project administration, Q.G.; Software, T.W.; Supervision, Q.G.

**Funding:** This research was jointly funded by the [National Nature Science Foundation of China] grant number [41571123], [National Social Science Fund of China] grant number [18BJL056], and [Shanghai Pujiang Program] grant number [17PJC030].

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Fang, C.L.; Yu, D.L. Urban agglomeration: An evolving concept of an emerging phenomenon. *Landsc. Urban Plan.* **2017**, *162*, 126–136. [\[CrossRef\]](#)
2. Polyan, P.M. Large urban agglomerations of the Soviet Union. *Sov. Geogr.* **1982**, *23*, 707–718. [\[PubMed\]](#)
3. Gottmann, J. Megalopolis, or the urbanization of the North-eastern seaboard. *Econ. Geogr.* **1957**, *33*, 189–200. [\[CrossRef\]](#)
4. Kunzmann, K.R.; Wegener, M. The pattern of urbanization in Western Europe. *Ekistics* **1991**, *50*, 156–178.
5. Portnov, B.A.; Schwartz, M. Urban clusters as growth foci. *J. Region. Sci.* **2009**, *49*, 287–310. [\[CrossRef\]](#)
6. Chapple, K. The fiscal trade-off: Sprawl, the conversion of land, and wage decline in California's metropolitan regions. *Landsc. Urban Plan.* **2018**, *177*, 294–302. [\[CrossRef\]](#)
7. Bagheri, B.; Tousi, S.N. An explanation of urban sprawl phenomenon in Shiraz Metropolitan Area (SMA). *Cities* **2018**, *73*, 71–90. [\[CrossRef\]](#)
8. Katherine, H. Urban infrastructure and natural resource flows: Evidence from Cape Town. *Sci. Total Environ.* **2013**, *461*, 839–845.
9. Guo, Q.B.; Liu, Q. Research progress review on urban factors aggregating ability. *J. Hubei Univ.* **2016**, *43*, 122–129. (In Chinese)
10. Porter, M.E. Location, competition and economic development: Local clusters in a global economy. *Econ. Dev. Quart.* **2000**, *14*, 15–34. [\[CrossRef\]](#)
11. Oskenbayev, Y.; Yilmaz, M.; Abdulla, K. Resource concentration, institutional quality and the natural resource curse. *Econ. Syst.* **2013**, *37*, 254–270. [\[CrossRef\]](#)
12. Ohlin, B. *Interregional and International Trade*; Harvard University Press: Cambridge, MA, USA, 1933.
13. Russo, F.; Musolino, G. A unifying modelling framework to simulate the spatial economic transport interaction process at urban and national scales. *J. Transp. Geogr.* **2012**, *24*, 189–197. [\[CrossRef\]](#)
14. Tripp, M.; Grillitsch, M.; Isaksen, A.; Sinozic, T. Perspectives on cluster evolution: Critical review and future research issues. *Eur. Plan. Stud.* **2015**, *23*, 2028–2044. [\[CrossRef\]](#)
15. Chang, C.; Oxley, L. Industrial agglomeration, geographic innovation and total factor productivity: The case of Taiwan. *Math. Comput. Simul.* **2009**, *79*, 2787–2796. [\[CrossRef\]](#)
16. Hochman, O. Efficient agglomeration of spatial clubs. *J. Urban Econ.* **2011**, *69*, 118–135. [\[CrossRef\]](#)
17. Fu, Y.M.; Gabriel, S.A. Labor migration, human capital agglomeration and regional development in China. *Region. Sci. Urban Econ.* **2012**, *42*, 473–484. [\[CrossRef\]](#)
18. Papi, L.; Sarno, E.; Zazzaro, A. The Geographical Network of Bank Organizations: Issues and Evidence for Italy. In *Handbook of Geographies of Money and Finance*; Martin, R., Pollard, J., Eds.; CSEF Working Paper, 403; Edward Elgar: Cheltenham, UK, 2015.

19. Bhaumik, S.K.; Dimova, R. Does Human Capital Endowment of Foreign Direct Investment Recipient Countries Really Matter Evidence from Cross-country Firm Level. *Data Rev. Dev. Econ.* **2013**, *17*, 559–570. [[CrossRef](#)]
20. Yang, C.; Lin, H.; Li, H. Influences of production and R&D agglomeration on productivity: Evidence from Chinese electronics firms. *China Econ. Rev.* **2013**, *27*, 162–178.
21. Alfaro, L.; Chen, M.X. The global agglomeration of multinational firms. *J. Int. Econ.* **2014**, *94*, 263–276. [[CrossRef](#)]
22. Ellison, G. Geographic concentration in U. S. manufacturing industrial: A dartboard approach. *J. Urban Econ.* **2000**, *9*, 48–67.
23. Catini, R.; Karamshuk, D.; Penner, O.; Riccaboni, M. Identifying geographic clusters: A network analytic approach. *Res. Policy* **2015**, *44*, 1749–1762. [[CrossRef](#)]
24. Scioscia, F.; Binetti, M.; Ruta, M.; Ieva, S.; Di Sciascio, E. A framework and a tool for semantic annotation of POIs in Open Street Map. *Procedia* **2014**, *111*, 1092–1101.
25. Zhang, X.Y.; Du, S.H.; Wang, Q. Hierarchical semantic cognition for urban functional zones with VHR satellite images and POI data. *ISPRS J. Photogramm. Remote Sens.* **2017**, *132*, 170–184. [[CrossRef](#)]
26. He, Y.X.; Yang, J.J. Comprehensive evaluation of global clean energy development index based on the improved entropy method. *Ecol. Indic.* **2018**, *88*, 305–321. [[CrossRef](#)]
27. Guo, Q.B.; Zhang, Z.H. Spatial-temporal evolution of factors aggregating ability in urban agglomeration in the middle reaches of the Yangtze River. *Acta Geogr. Sin.* **2017**, *72*, 1746–1761. (In Chinese)
28. Arshad, S.; Hu, S.; Ashraf, B.N. Zipf's law and city size distribution: A survey of the literature and future research agenda. *Physica A* **2018**, *492*, 75–92. [[CrossRef](#)]
29. Chen, Y.G. The evolution of Zipf's law indicative of city development. *Physica A* **2016**, *443*, 555–567. [[CrossRef](#)]
30. Giesen, K.; Südekum, J. Zipf's law for cities in the regions and the country. *J. Econ. Geogr.* **2011**, *11*, 667–686. [[CrossRef](#)]
31. Spencer, C.J.; Yakymchuk, C.; Ghaznavi, M. Visualizing data distributions with kernel density estimation and reduced chi-squared statistic. *Geosci. Front.* **2017**, *8*, 1247–1252. [[CrossRef](#)]
32. Hu, Y.J.; Wang, F.H.; Guin, C.; Zhu, H.J. A spatial-temporal kernel density estimation framework for predictive crime hotspot mapping and evaluation. *Appl. Geogr.* **2018**, *99*, 89–97. [[CrossRef](#)]
33. Heidenreich, N.B.; Schindler, A.; Sperlich, S. Bandwidth selection for kernel density estimation: A review of fully automatic selectors. *Adv. Stat. Anal.* **2013**, *97*, 403–433. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).