International Journal of
Geo-Information

MDPI

Article

Spatial Distribution of Urban Parks' Effect on Air Pollution-Related Health and the Associated Factors in Beijing City

Huimin Ji [1,2], Juan Wang [1,2,*], Yanrong Zhu [1,2], Changsheng Shi [1,2], Shaohua Wang [3,4,5], Guoqing Zhi [6] and Bin Meng [1,2]

1 College of Applied Arts and Sciences, Beijing Union University, Beijing 100191, China
2 Laboratory of Urban Cultural Sensing & Computing, Beijing Union University, Beijing 100191, China
3 International Research Center of Big Data for Sustainable Development Goals, Beijing 100094, China
4 Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China
5 State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China
6 China Electronic Technology Group Corporation Academy of Electronic and Information Technology, Beijing 100041, China
* Correspondence: wangjuan@buu.edu.cn

Abstract: Urban parks play an essential role in mitigating the effects of air pollution on human health in a healthy city construction process. However, due to the data limitations, little is known about the spatial distribution of real-time expressed air pollution-related health (APRH) across different urban parks and the contribution of the associated factors. To fill this research gap, this research was conducted based on social media Weibo data (Chinese Twitter) and other geographical data using semantic analyses and the Geo-Detector method by taking 169 urban parks in Beijing as the study area. The results showed that there were more Weibo items relating to APRH clustered within the third ring road and decreasing outward along the ring road. A total of 16 factors in three categories were introduced to analyze the driving forces of this spatial distribution. Accessibility was outstanding with a q-value of the number of subway stations (X14) as high as 0.79, followed by built environment and finally park attributes. Distinguished from those reports based on the traditional statistical data, this research demonstrated that although the urban parks improved the APRH, the exposure to air pollution also increased the health risks when visiting the urban park. It also provides a geographical understanding of the urban parks' effect on APRH and theoretical guidance for urban park planning and construction.

Keywords: urban park; air pollution-related health; geographical big data; Geo-Detector; Beijing

Citation: Ji, H.; Wang, J.; Zhu, Y.; Shi, C.; Wang, S.; Zhi, G.; Meng, B. Spatial Distribution of Urban Parks' Effect on Air Pollution-Related Health and the Associated Factors in Beijing City. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 616. https://doi.org/10.3390/ijgi11120616

Academic Editors: Wei Huang and Wolfgang Kainz

Received: 13 October 2022
Accepted: 5 December 2022
Published: 8 December 2022

1. Introduction

With the acceleration of global urbanization, more and more people are living in urban areas than rural areas [1]. On the one hand, urbanization has given rise to a series of environmental problems, especially air pollution, which severely threatens the human health [2,3]. For example, significant associations between short-term exposure to PM2.5 and allergic rhinitis outpatients were detected in Beijing [4]. On the other hand, cities also provide the possibility for obtaining public services for good health, among which urban parks were remarkable. As an important role in the development of sustainable cities and society [5], urban parks were labeled as the lungs of the city with the capacities to positively reduce air pollutants, restore capacities, and build capacities [6–8]. How urban parks impact air pollution-related health (APRH) and what factors contribute to the health effects of urban parks are issues that deserve the attention of researchers.

A large amount of previous research has demonstrated the positive impact of urban parks on APRH. For example, a study conducted in Kaunas, the second largest city in Lithuania, has shown that park users have a significantly lower prevalence of cardiovascular risk than non-users [9]. Another study in Guangzhou, China, showed similar results [10]. In Tehran, green space was found to have a significant mitigating effect in reducing air pollution and mortality from respiratory diseases [11]. Green space in cities have also been demonstrated to have a positive effect on reducing the mortality rate of residents [12]. Researchers have adequately examined the relationship between parks and APRH. Nevertheless, the effects of different parks on APRH could not be generalized. It was documented that the effect of urban parks on human health may be influenced by such factors as size [11], facilities [13], characteristics of the surrounding environment [14], and accessibility [9]. These factors were also closely associated with the visit [13], residents' well-being [15], and vitality [16] of the park. However, the role of these factors was usually assessed singularly in previous research. As the urban environment is a comprehensive system [17], the quantitative relative effect of various factors should not be ignored. Unfortunately, it was not clear how these factors impact the urban parks' effects on the APRH, which need to be further addressed.

From the perspective of the data to indicate health, hospital admission records [18], and mortality rates from the Medical Health Department [11,19], questionnaires for self-assessment of health status [9,20,21] have been largely employed. The data of hospital admission records and mortality rates reflect human health from the perspective of disease, which is a very serious part of the health effects caused by air pollution, and could be understood as being at the top of the health effects pyramid [22]. As for the questionnaire data, it was time-consuming and labor-intensive, which saw most of the current research only focusing on one or several specific big parks [21,23]. Moreover, this type of data cannot be used to reflect the spatial distribution of the health effects of the larger number of parks because of the limitation of data volume. With the advantage of rich contextual content and geographical location information [24], social media data has attracted much attention in indicating human health by an increasing number of research, and the validity has been strongly demonstrated [25]. Sina Weibo is a popular social media platform in China, which could be understood as Chinese Twitter. The monthly active users of Weibo reached 511 million in September 2020 [26]. It has been widely used to identify the health status and explore the sensitivity to air pollution [27,28].

From the above, this research was constructed to assess the spatial distribution of urban parks' effect on APRH based on the Weibo data by taking 169 urban parks in Beijing, China, as the research area. The impact of the associated factors on the spatial distribution were also quantified based on the geographical detectors. Specifically, this research has the following two objectives: (1) to map the spatial distribution of effects on APRH across different urban parks; (2) to quantify the associated factors that contributed to the spatial distribution of urban parks' effect on APRH. The remainder of the manuscript is organized as follows. In Section 2, a brief discussion about the importance on the chosen area is clarified, and the materials and methods are fully described. In Section 3, the results of this research were presented. Then, these findings were deeply discussed in Section 4. Finally, conclusions were brought up in Section 5.

2. Materials and Methods

2.1. Study Area

As the capital of China, Beijing covers an area of up to 16,410 km^2 and 78% of the population lives within the Sixth Ring Road [29]. The average concentration of PM2.5 in 2017 in Beijing was 116 μg/m^3 [30], which was ten times higher than the World Health Organization's air quality guidelines (WHO AQG: 10 μg/m^3). The poor air quality has a serious impact on the residents' health [31]. Beijing has been staying committed to building a livable city [32]. Therefore, this research was focused on the year 2017 in Beijing in order to explore the effect of urban parks on APRH. According to the list of urban parks published by the Beijing Municipal Bureau of Landscape Architecture and the area of interest (AOI) data in Beijing sourced from the Gaode Map platform, 169 urban parks within the sixth ring road of Beijing were selected as the study area in this research, which is shown in Figure 1.

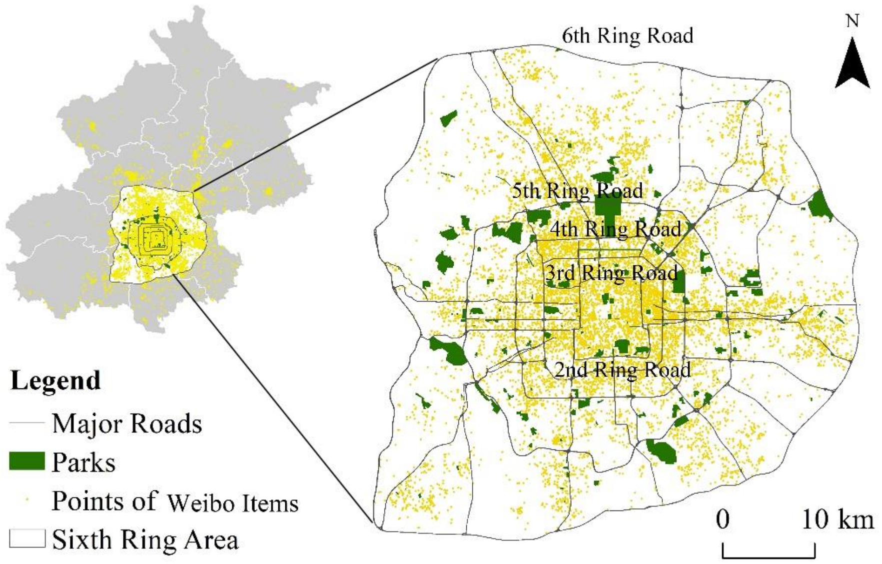

Figure 1. The study area in this research.

2.2. Deriving the Real-Time Expressed APRH Based on Weibo Data

In this research, 11.79 million Weibo items in Beijing in 2017 were obtained from the Sina Weibo Open Platform (SWOP) through web crawler technology. The period of Weibo data used in this research covers the period from 1 January to 31 December 2017. The Weibo data contains rich information such as text contents, time, and latitude and longitude. These data were firstly selected using keywords related to the health status associated with air pollution such as cough and sore throat. Then, the Bidirectional Encoder Representations from Transformers (BERT) model [33] was employed to obtain those Weibo items relating to APRH. For each item, if it was related to health status with air pollution, it was manually labeled as 1 (for example, "the smog makes me cough", "I can't stand my sore throat"), otherwise it was labeled as 0. By adjusting the learning rate and the number of iterations, the overall accuracy of the classifier by BERT was over 86%. More detailed information can be found in the research [27]. Finally, all the Weibo data in the study area were input into BERT and 21,372 Weibo items in total relating to APRH were obtained in the study area.

2.3. Assessing Urban Parks Effect on APRH

As it was indicated that urban parks were positively associated with the health benefits of residents within the 1 km buffer zone [34,35], the number of Weibo items relating to APRH within the 1 km buffer for each park was selected to indicate the urban parks' effect on APRH in this research. The more Weibo items about APRH indicate that more APRH of residents was expressed. The 1 km buffer of each park was generated through the buffer analysis of ArcGIS software. The boundaries of the park and its buffers were overlaid with the geocoded Weibo data to calculate the number of Weibo items relating to APRH within the 1 km buffer for each park, which were the dependent variable Y in this research.

2.4. Calculating the Associated Factors Based on Multisource Geographical Big Data

The selection of the factors was a very important issue in this research. According to previous research, the associated factors to the spatial distribution characteristics of parks' effect was varied, such as size [11], the facilities [13], the characteristics of the surrounding environment [14], and accessibility [9]. As this research was focused on the spatial distribution of urban parks' effect on air pollution-related health and the associated factors, three groups factors that were associated with the spatial distribution characteristics of parks' effect were introduced in this research (that was, independent variables X), namely the attributes of the parks, built environmental characteristics, and accessibility [36–38]. All independent variables X used in this research were summarized in Table 1.

Table 1. All independent variables (X) in this research.

Types of Factors	Factors	Type of Data	Source of Data	Mean	Standard Deviation
Attributes of the parks	X1 Area (km^2)	AOI	Gaode Map platform	0.53	1.15
	X2 Perimeter (km)			2.98	2.76
	X3 Ratio of perimeter to area (km/km^2)			17.39	15.15
	X4 Types of parks				
	X5 Annual average value of NDVI	Remote sensing	GEE	0.30	0.06
	X6 Visits	Weibo items	SWOP	2175.04	7320.57
Built environment	X7 Diversity of POIs	POI	Gaode Map platform	0.80	0.06
	X8 Density of catering services POI ($/km^2$)			0.16	0.04
	X9 Density of business residence POI ($/km^2$)			27.82	23.79
	X10 Density of tourism attraction POI ($/km^2$)			0.01	0.02
	X11 Density of science and education POI ($/km^2$)			40.80	34.02
	X12 Density of corporate business POI ($/km^2$)			0.15	0.07
Accessibility	X13 Number of bus stops	Basic geographic information data	Ministry of Transport of the People's Republic of China	22.03	13.21
	X14 Number of subway stations			2.29	2.65
	X15 Density of sidewalks (km/km^2)	Road Network	Open street map	1.08	1.67
	X16 Density of road network (km/km^2)			13.25	4.68

(1) The attributes of the parks. The area (X1), perimeter (X2), and perimeter-to-area ratio (X3) of the parks were calculated in ArcGIS based on the AOI of each park.

Parks were also grouped into four categories (X4) according to the size of their area: small neighborhood parks with an area less than 0.02 km^2; small regional parks between 0.02 and 0.2 km^2; large metropolitan parks between 0.2 and 1 km^2; and riverside belt parks based on the shape and with the area between 0.05 and 0.3 km^2 [39].

The normalized differential vegetation index (NDVI) (X5) was used to assess the vegetation coverage of the parks [15,40]. The NDVI data used in this research were provided by the National Aeronautics and Space Administration (NASA) MODIS Terra (MOD13Q1) 16-day vegetation index product with a spatial resolution of 250 m [41]. They were obtained through the Google Earth Engine (GEE), a platform powered by Google Cloud Computing [42]. The annual averaged NDVI of the parks was calculated by the Zonal Statistics tool in ArcGIS.

As the park visits were associated with residents' health [13], all the Weibo items within the park boundaries were used to indicate visits to the parks (X6) in this research [43].

(2) Built environmental factors. The diversity of POIs (X7) was an important variable that indicates the characteristics of the built environment. The POIs data in 2017 in the study area were sourced from the Gaode Map platform, and included 15 main categories (restaurants, shopping, company, etc.). This research draws on the methods of land use mixture and activity type diversity [44] to calculate the diversity of POIs. The formula was as follows:

$$M_j = \frac{-\sum_{i=1}^{k} q_i \ln q_i}{\ln(k)}$$

where M_j denotes the POI diversity index of Park j, q_i denotes the ratio of the number of type i POI in park j to the total number of POIs in the park j, and k denotes the number of POI types in Park j. The POI diversity index M_j has a value range from 0 to 1, and a larger value indicates a more diverse POIs in the park buffer.

In addition, the density of the four types of POIs (catering services (X8), business residence (X9), tourism attraction (X10) and science and education (X11), and corporate businesses (X12) within each park buffer was also calculated separately.

(3) Accessibility. Accessibility of a park was one important category of variables in this research and it was important to choose the suitable method in the research. There are many methods used in accessibility analysis of urban parks in the research, such as the statistical index method, simple buffer method, cost-weighted distance method, gravity potential energy analysis method, minimum distance method, network analysis method, etc [45]. These models usually calculate the accessibility based on such variables as population, distance, and time. On the one hand, each accessibility calculation method has its own shortcomings. For example, the gravitational model method produces gravitational values that only indicate the relative ease of using park services, reflecting differences within the study area, and cannot be used for comparisons between study areas [46]. On the other hand, in similar studies using multi-source geographic big data to research park use and associated factors, factors such as sidewalk density, road density, the number of bus, and subway stops were usually directly used to indicate park accessibility [36,38]. It was indicated that these factors were representative in characterizing park accessibility and were suitable for application in attribution analyses. Using these factors to reflect the accessibility was a more direct and clearer approach than using the results of the accessibility calculation model, which avoids the bias in the model calculation process because of the subjective choice of variables and parameters. Therefore, following this research, the density of bus stops (X13), subway stops (X14), sidewalks (X15), and roads (X16) within the 1 km buffer of the park were used to indicate the accessibility of the parks [36]. The data of bus and subway stations can be obtained through https://www.mot.gov.cn/shuju/ (accessed on 20 January 2022). The calculation of road density includes highways, national roads, provincial roads, primary urban roads, secondary roads, etc., can be obtained through https://www.openstreetmap.org/ (accessed on 15 January 2022).

2.5. Quantifying the Contribution Based on Geographical Detectors

Geographical detectors (http://geodetector.org/ (accessed on 15 April 2022)) are a group of statistical methods for detecting the spatial heterogeneity of research objects and revealing the driving forces [47]. A factor detector was employed in this research to quantify the impact of the associated factors. A factor detector was used to detect the spatial heterogeneity of the dependent variable Y and to judge the level of explanation of the independent variable X for Y. The level was quantified by the q-value. The formula for calculating q was as follows:

$$q = 1 - \frac{1}{N\sigma^2} \sum_{h=1}^{L} N_h \sigma_h^2 \quad (1)$$

where L was the stratification of Y or X; N_h and N were the number of units in stratum h and the whole region respectively; and σ_h^2 and σ^2 were the variances of Y values in stratum h and the whole region, respectively.

Usually, the value of q was within the range [0, 1]. Larger values of q indicated a stronger influence of factor X on the spatial distribution of the parks' effect on APRH. The q-value indicates that X explains $100 \times q\%$ of Y. In this research, three levels of P (0.01, 0.05, and 0.1) were used to test the significance of q [48]. When geographical detectors were applied in the research, all independent variables are required to be type variables. There were many ways to generate type variables. Different classification methods may have different q-values. The proposer of geographical detectors indicated that the effect of classification could be evaluated by the q-value, and the larger the q-value, the better the effect of classification [48]. Based on our experiments, it was optimal to classify independent variables into 5 categories according to the quantile classification method in this research. Therefore, all variables in this research except X4 were converted into 5 categories by quantile in ArcGIS.

The research framework for this research is shown in Figure 2.

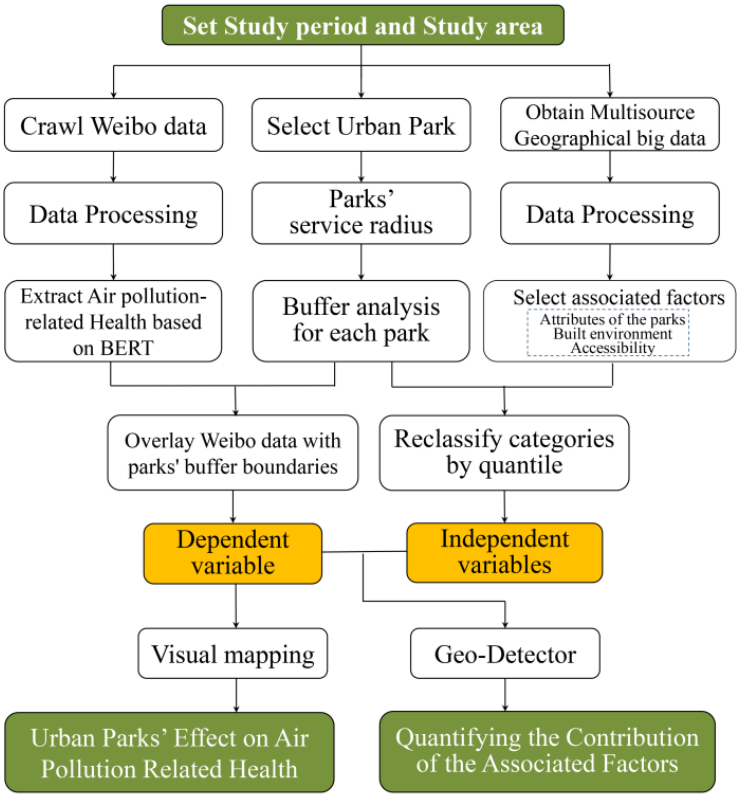

Figure 2. The framework in this research.

3. Results

3.1. Evaluation of the Weibo Items about APRH

The Weibo items relating to APRH in the year 2017 were extracted by semantic analysis in this research. To evaluate the validity of the extracted Weibo data, the health status of the residents in the social media space was compared with the air quality in the geographical space. Air Quality Index (AQI) was classified into six levels. When the AQI was greater than 100, it was defined as a polluted day. When the AQI was greater than 300, it was severely polluted. The correlation coefficient and significance of daily numbers of Weibo items relating to APRH and AQI were calculated by SPSS software. The correlation coefficient (CC) was 0.22, which was significant at the level of 0.01. The quantity and time trend correlated relationship are shown in Figure 3. It could be found that when AQI was high in the geographical space, there was also peak of public opinion in the social media space. For instance, an air pollution event in a certain geographic space occurred on the 1st of January and ended on the 7th, lasting 7 days. At the same time, the number of Weibo items relating to APRH in social media spaces increased sharply, reaching the highest value of the year (199) on the 4th of January. There was a two-day air pollution event in a certain geographic space on the 4th and 5th of May, but because of its short duration, there was a small increase in the number of Weibo items relating to APRH in social media spaces, which was not as strong of a response as was seen in January.

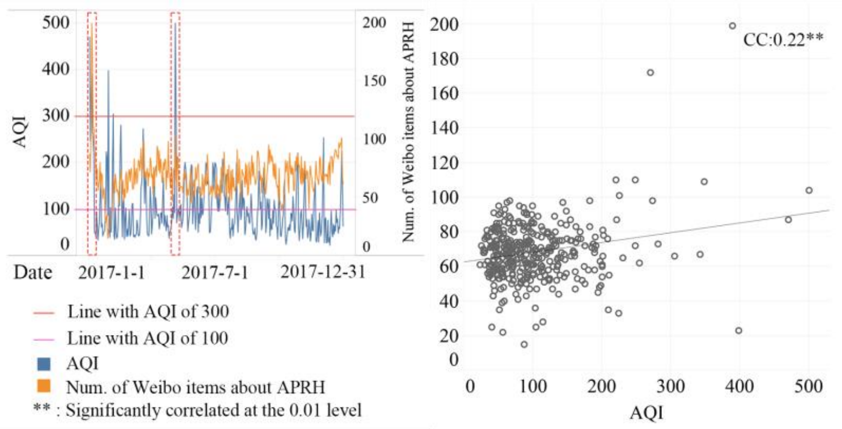

Figure 3. The quantity and time trend correlated relationship between AQI and the number of Weibo items relating to APRH.

3.2. Urban Parks' Effect on APRH

To assess the parks' effect on APRH, the number of Weibo items relating to APRH within and outside the 1 km buffer zone of the parks was compared. The results showed a significant difference at the 0.01 level, which indicated that the park has significant impact on the real-time expression of APRH. Based on this, the spatial distribution of different urban parks' effects on APRH were further assessed.

The spatial patterns of urban parks' effect on APRH is shown in Figure 4, which is indicated by the number and the kernel density of Weibo items relating to APRH within a 1 km buffer for each park. The parks with a larger number of Weibo items relating to APRH were clustered within the third ring and the northern part of the fifth ring of the study area.

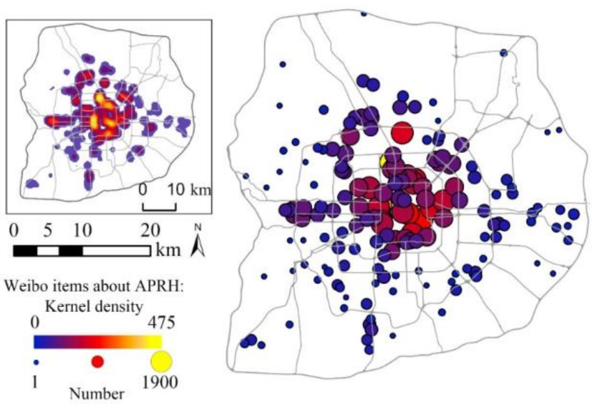

Figure 4. Visualization of the number of Weibo items relating to APRH in urban parks' buffer zone. Graph in the upper left corner is the density graph. Every circle in the figure represents an urban park.

When Weibo items relating to APRH were counted separately for the parks located between the ring roads (Figure 5), it roughly showed a gradual decrease around the ring road. Specifically, the average number of Weibo items relating to APRH was 373, 429, 244, 111, and 73 from the second ring to the sixth ring, in that order. In general, the spatial pattern of the parks with a larger number of Weibo items relating to APRH shows a "single center" pattern with decreasing distance from the main urban center.

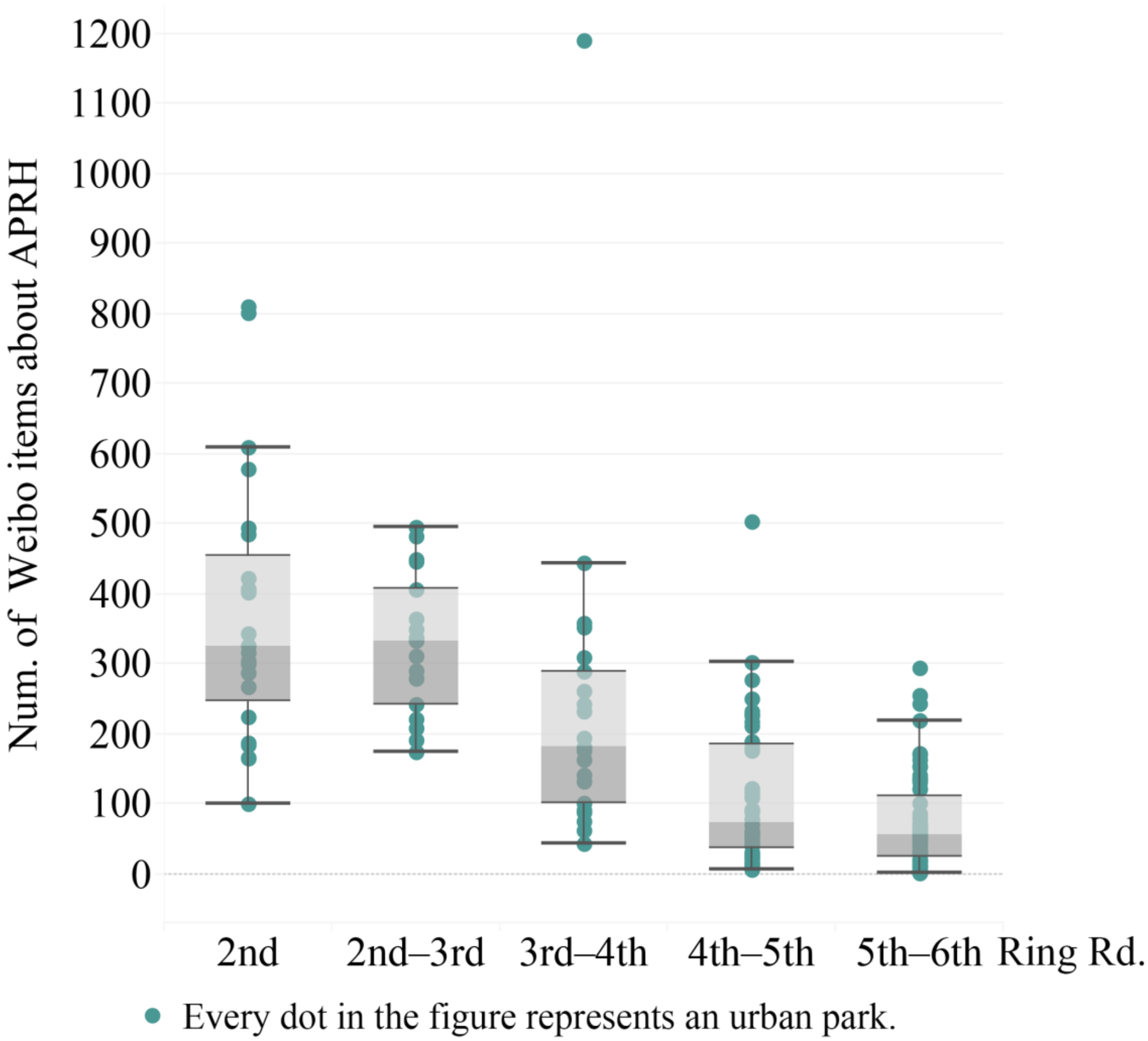

Figure 5. The number of Weibo items relating to APRH of parks between the Ring Roads.

3.3. Quantifying the Contribution of the Associated Factors

Sixteen independent variables selected in this research is shown in Figure 6. There were different spatial patterns across the variables. For example, parks with high values of X9 (Density of business residence) were distributed mainly within the fifth ring, whereas parks with high values of X10 (Density of tourism attraction) were mostly concentrated in the northwestern part between the fifth and sixth rings.

The contributions of the associated factors were indicated by the q value shown in Figure 7. The q-values of X6, X7, X9, X11, X13, X14, X15 and X16 passed the significance test of $p < 0.01$. The q-value of X14 (density of subway stations) was the largest among above factors, with q value of 0.79. Followed by X9 (density of business residence) and X11 (number of science and education), with q values reaching 0.55. The q-values of X13 (number of bus stops) and X16 (density of road network) over 0.3. X12 (Density of corporate business) passed the significance test of $p < 0.05$ with the q-value of 0.08. X5 (Annual average value of NDVI) passed the significance test of $p < 0.1$ with the q-value of 0.06. For the three types of factors, accessibility (X13–X16) had the largest q-value, followed by built environment (X7–X12) and finally attributes of the parks (X1–X6).

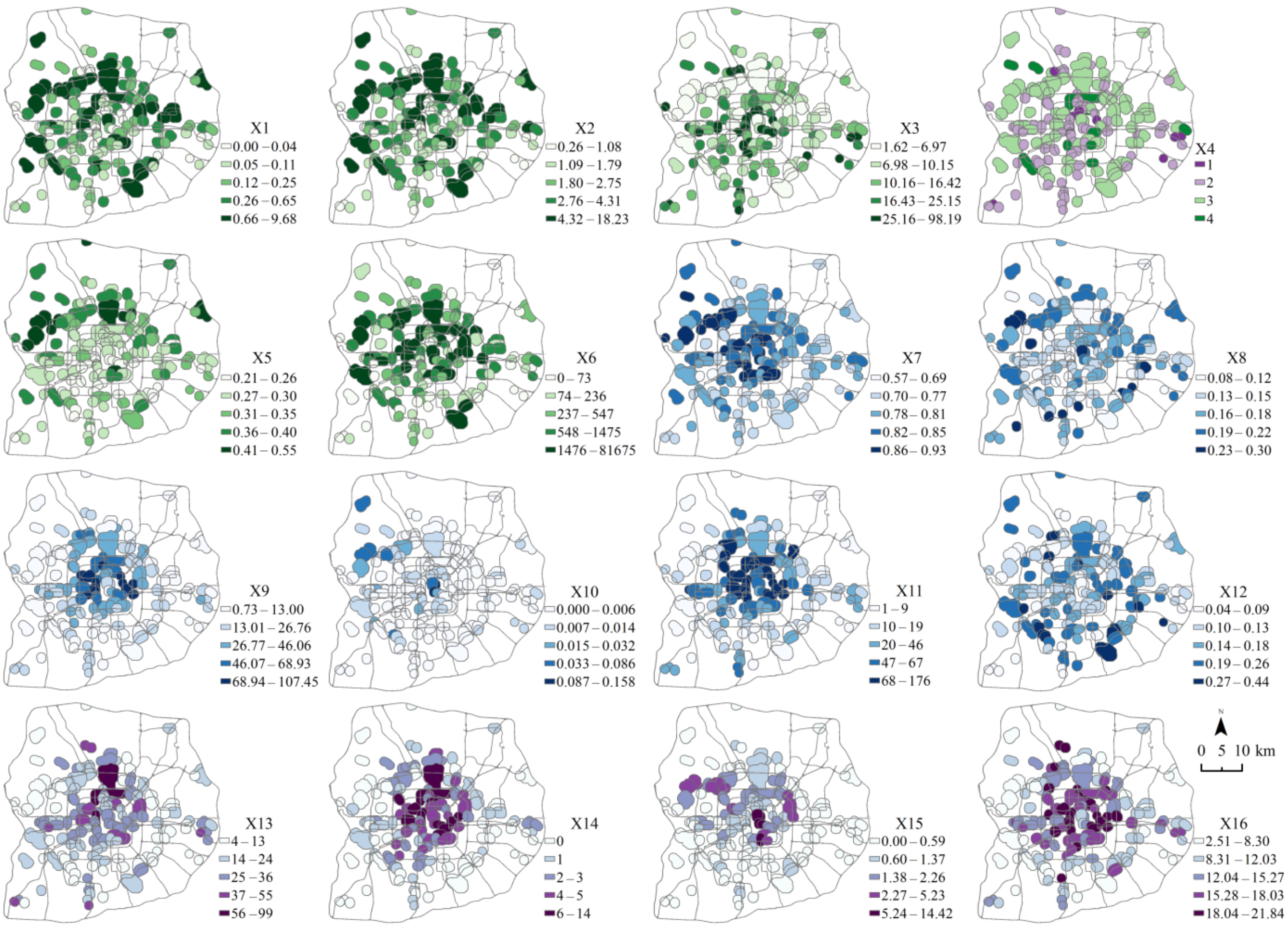

Figure 6. The spatial distribution of the selected independent variables in this research. The detailed information was listed in Table 1.

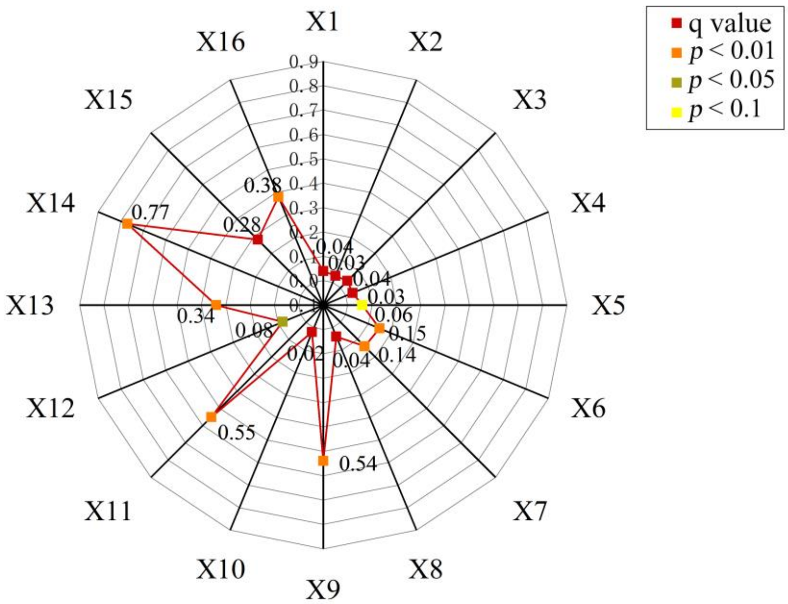

Figure 7. The results of factor detector.

4. Discussion

4.1. The Feasibility of Weibo Data

The rapid development of social media has provided a new data source for public health research [49,50]. Each user could be regarded as a sensor [51]. The social media data generated by them could be used to sense the health status of the residents. It was indeed the case that there were inevitable sample biases when using Weibo data [52]. Nevertheless, Weibo data are still valuable in park-related research. On the one hand, social media data were used in much park-related research, such as in research on recreational visits to urban parks and associated factors [43] and in understanding the use of urban green spaces [53]. The suitability of Weibo data in park-related research was demonstrated. On the other hand, although we think of parks as being used primarily by older people, there was a study in Cleveland that found that older adults were less likely to visit local parks than their younger counterparts [54]. Therefore, young people were also an important part of the park's visitors. Young people as the main users of social media were more educated, more active in public affairs, and are regarded as the mainstay in China [55], and so their voices are important for incentivizing governments [56]. Therefore, the results of this research were very meaningful despite the bias of the Weibo data.

It is also worth mentioning that compared with hospital admission data, social media data has many advantages. On the one hand, this type of data was generated directly by the users and could give a more real and comprehensive indication of the health of residents. It avoids the pyramid effect [22] of using data from the Centers for Disease Control. On the other hand, social media data is a typical type of big data [57]. Two typical characteristics of big data are its large volume and wide coverage [58]. Because of this, it is possible to break the limitation that only one or a few parks could be researched by using traditional data.

Weibo data as a typical social media data were employed in this research to reflect the APRH status of residents, and thereby explore the parks' effect on APRH. When we use Weibo data as health outcomes, there were discrepancies with actual disease data. For example, there was research found that higher park accessibility reduced older adults' risk of cardiovascular disease, joint disease, and endocrine disease [59]. However, higher accessibility affecting the health of residents by increasing their exposure to air pollution was found in this research. This will be discussed in detail below. Exploring the reasons behind this difference will be an interesting topic for our subsequent research. Moreover, the spatial distribution of 169 urban parks' effects on APRH was mapped, which shows the spatial pattern of clustering in the city center decreasing outward along the ring road. This spatial pattern was similar to the population distribution characteristics of Beijing [29]. This may be caused by the high number of Weibo items in places where people were densely distributed.

4.2. The Impact of Associated Factors

The results of factor detection (Figure 7) indicated that X5, X6, X7, X9, X11, X12, X13, X14, X15, and X16 significantly contributed to the spatial distribution of urban parks' effect on the APRH. To further analyze their relationships, we plotted the scatter plot and fitted the trend line, which shown in Figure 8.

For attributes of the parks, it could be found that X5 (Annual average value of NDVI) showed a negative correlation with the number of Weibo items relating to APRH. This was because vegetation could reduce air pollutants [60]. The capacity of trees to remove air pollution was documented by research [61]. X6 (visits) showed a positive correlation with the number of Weibo items relating to APRH. This was contrary to the results of previous studies that higher levels of park visits were associated with better health [13]. This may be because more park visits indicate that more humans are exposed to air pollution. A study has pointed out that no matter how poor the air quality was, a proportion of people (41–64%) put themselves at risk of exposure in order to enjoy the benefits of parks [23]. However, exposure to air pollution was associated with many diseases in people [62–64].

Therefore, the exposure of residents to air pollution while visiting the park should also be addressed.

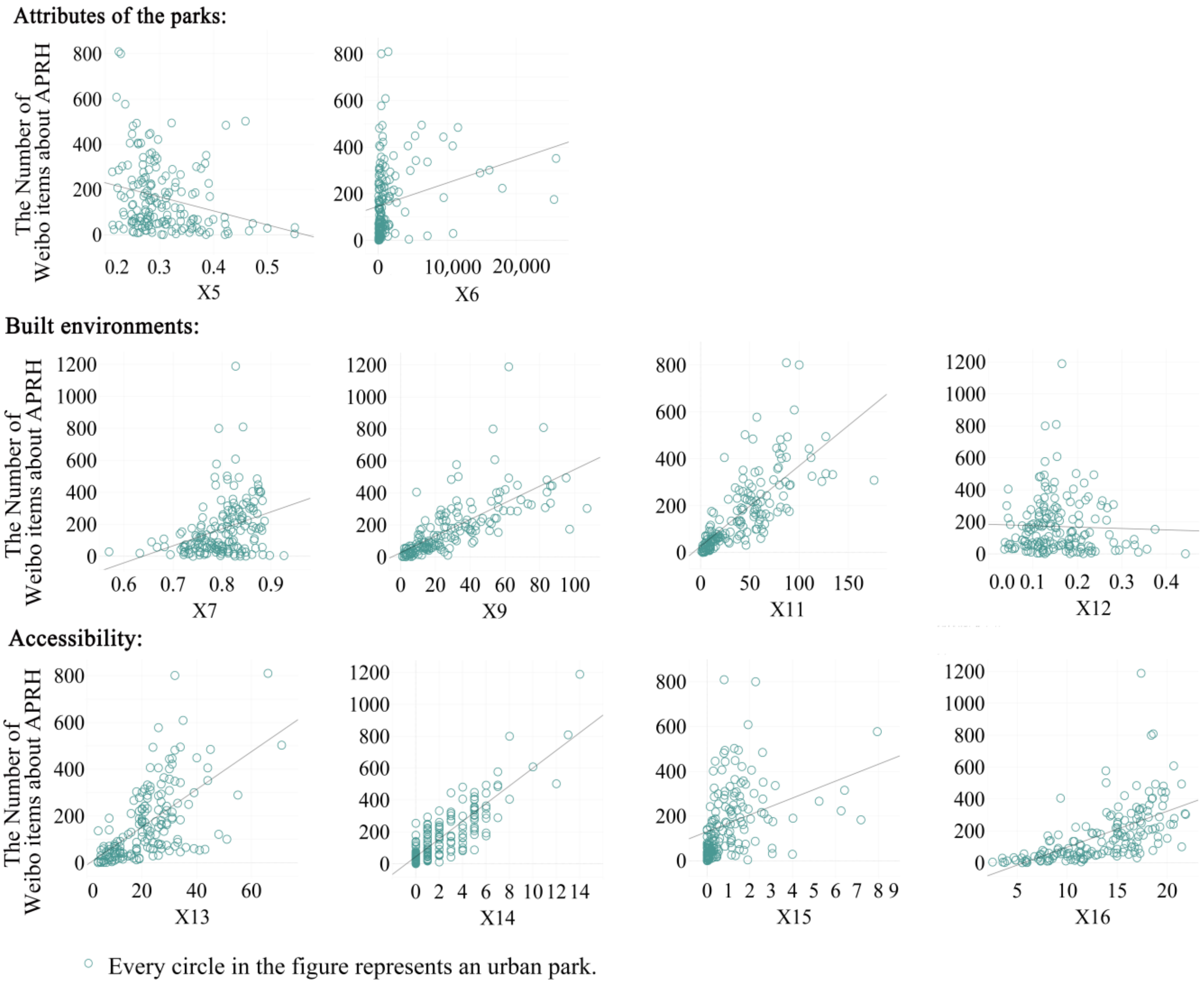

Figure 8. The correlation between the number of Weibo items relating to APRH with X5, X6, X7, X9, X11, X12, X13, X14, X15, and X16.

When the Geo-Detector was conducted to associate park visits (X6) with others independent variables (Figure 9), the results showed that many factors have an effect on park visitation. This indicates that these factors may indirectly affect the spatial distribution of parks' effect on APRH by influencing park visits. For example, the previous study has shown that the attributes of green space such as size, shape, and patch density were associated with human health [11]. Yet the area (X1), perimeter (X2), and perimeter-to-area ratio (X3) of the parks did not have significant impacts on the spatial distribution of parks' effect on APRH in this research. A significant impact on park visits was found in Figure 9, X1, X2, and X3. Therefore, we speculate that the attributes of the park area, perimeter, and perimeter-to-area ratio could indirectly affect the health effects of parks by influencing park visits.

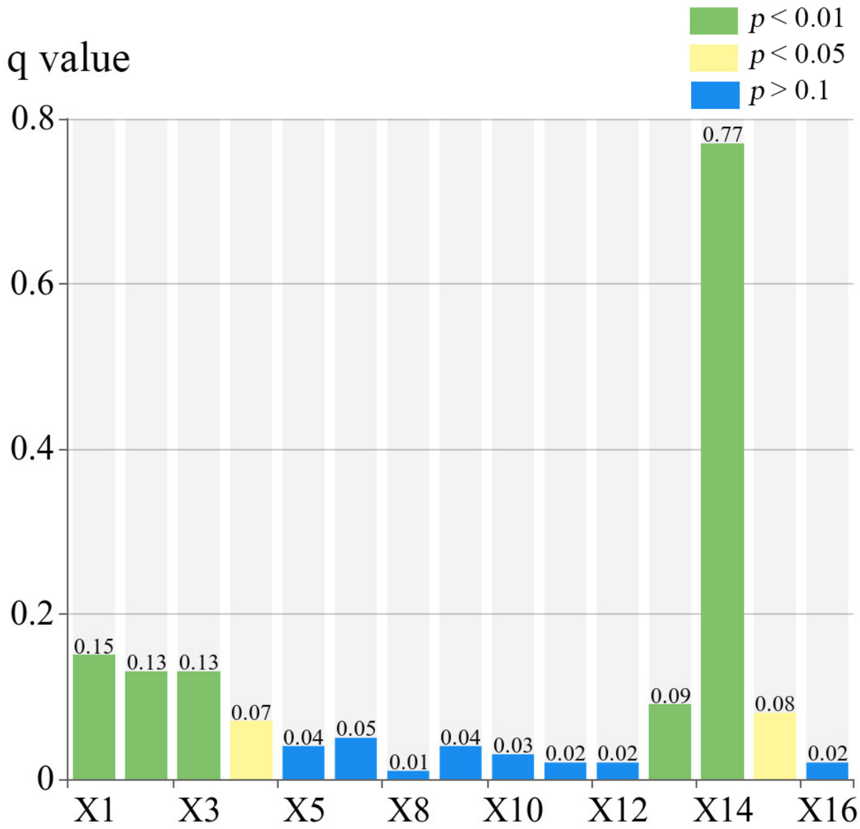

Figure 9. The results of factor detection with the park visits (X6) as the dependent variable and the other factors as independent variables.

For the built environment, X7 (Diversity of POIs), X9 (Density of business residence POI), and X11 (Density of science and education POI) had significant positive correlations with the number of Weibo items relating to APRH. POI was widely used to identify the functional areas of the city [65,66]. The high values of X7, X9, and X11 indicated that there were diversity and convenient urban facilities in the 1 km buffer zone of the park, which could attract human activity and increase potential park visits. This may expose more people to air pollution when visiting the park. For accessibility, X13 (Number of bus stops), X14 (Number of subway stations), X15 (Density of sidewalks), and X16 (Density of road network) showed a strong positive correlation with the number of Weibo items relating to APRH. This may be caused by the high accessibility to the park, means a greater convenience for getting to the park. This would increase the willingness of residents to go to parks and increase the visits of parks [67]. This may also increase residents' exposure to air pollution, resulting in more Weibo items relating to APRH. This echoes the discussion above.

5. Conclusions

The Sina Weibo data could monitor the urban residents' real-time expressed health status from a more comprehensive level. It provided powerful support to explore urban parks' effect on APRH and the associated factors across various parks. By combining the semantic analysis and Geo-Detector method based on multi-source geographical big data, this research assessed the spatial distribution of urban parks' effect on APRH and the associated factors in Beijing city. It was found that the parks with a larger number of Weibo items relating to APRH were clustered within the third rd Ring and north of the fifth Ring in this study area, and it roughly showed a gradual decreasing around the ring road. Ten associated factors (X5, X6, X7, X9, X11, X12, X13, X14, X15, and X16) with significant contributions also clearly showed in this research. It was implied that although the positive effects of parks on APRH have been demonstrated, exposure to air pollution is also a subject that cannot be ignored when assessing the urban parks' effect on APRH. The difference between the health outcomes from Weibo data and from the actual disease data also was an interesting topic for our subsequent research. This research provides a new perspective for understanding the urban parks' effect on APRH and stresses the importance of paying attention to air pollution exposure. This was very meaningful for scientific prevention and was also valuable for developing healthy and sustainable cities.

This research focuses on the spatial distribution of parks' effect on APRH in Beijing from a geographical perspective, enriching the theoretical study of urban parks. Furthermore, the relatively quantitative research of the associated factors was a highlight of this research, which helps to further elucidate the internal logic of spatial distribution. The results of the research provide insights into the sustainable optimization of urban parks and the spatially rational layout of urban green spaces in the process of building a livable city in Beijing.

Author Contributions: Conceptualization, Huimin Ji and Juan Wang; Data curation, Huimin Ji and Juan Wang; Funding acquisition, Shaohua Wang, Juan Wang and Bin Meng; Investigation, Huimin Ji, Yanrong Zhu, Changsheng Shi and Guoqing Zhi; Methodology, Huimin Ji and Juan Wang; Project administration, Shaohua Wang, Juan Wang and Bin Meng; Software, Huimin Ji, Yanrong Zhu, Changsheng Shi and Guoqing Zhi; Visualization, Huimin Ji, Yanrong Zhu, Changsheng Shi and Guoqing Zhi; Writing—original draft, Huimin Ji; Writing—review and editing, Huimin Ji, Shaohua Wang, Juan Wang and Bin Meng. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Academic Human Resources Development in Institutions of Higher Learning under the Jurisdiction of Beijing Municipality (Grant Nos.CIT&TCD201904070), the R&D program of the Beijing Municipal Education Commission (KM202211417015), Academic Research Projects of Beijing Union University (No. ZK10202209), and the Hundred Talents Program Youth Project (Category B) of the Chinese Academy of Sciences (Grant Nos. E2Z10501).

Data Availability Statement: The data is available from the authors upon reasonable request.

Acknowledgments: The authors would like to thank the editor and anonymous reviewers for their comments on the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. 68% of the World Population Projected to Live in Urban Areas by 2050, Says UN. Available online: https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html (accessed on 10 January 2022).
2. Geng, G.; Zheng, Y.; Zhang, Q.; Xue, T.; Zhao, H.; Tong, D.; Zheng, B.; Li, M.; Liu, F.; Hong, C. Drivers of $PM_{2.5}$ air pollution deaths in China 2002–2017. *Nat. Geosci.* **2021**, *14*, 645–650. [CrossRef]
3. Song, C.; Wu, L.; Xie, Y.; He, J.; Chen, X.; Wang, T.; Lin, Y.; Jin, T.; Wang, A.; Liu, Y. Air pollution in China: Status and spatiotemporal variations. *Environ. Pollut.* **2017**, *227*, 334–347. [CrossRef] [PubMed]
4. Wang, M.; Wang, S.; Wang, X.; Tian, Y.; Wu, Y.; Cao, Y.; Song, J.; Wu, T.; Hu, Y. The association between $PM_{2.5}$ exposure and daily outpatient visits for allergic rhinitis: Evidence from a seriously air-polluted environment. *Int. J. Biometeorol.* **2020**, *64*, 139–144. [CrossRef] [PubMed]
5. Chiesura, A. The role of urban parks for the sustainable city. *Landsc. Urban Plan.* **2004**, *68*, 129–138. [CrossRef]
6. Markevych, I.; Schoierer, J.; Hartig, T.; Chudnovsky, A.; Hystad, P.; Dzhambov, A.M.; De Vries, S.; Triguero-Mas, M.; Brauer, M.; Nieuwenhuijsen, M.J. Exploring pathways linking greenspace to health: Theoretical and methodological guidance. *Environ. Res.* **2017**, *158*, 301–317. [CrossRef]
7. Urban Green Spaces and Health (No. WHO/EURO: 2016-3352-43111-60341). Available online: https://apps.who.int/iris/handle/10665/345751 (accessed on 30 January 2022).
8. Xing, Y.; Brimblecombe, P. Role of vegetation in deposition and dispersion of air pollution in urban parks. *Atmos. Environ.* **2019**, *201*, 73–83. [CrossRef]

9. Tamosiunas, A.; Grazuleviciene, R.; Luksiene, D.; Dedele, A.; Reklaitiene, R.; Baceviciene, M.; Vencloviene, J.; Bernotiene, G.; Radisauskas, R.; Malinauskiene, V. Accessibility and use of urban green spaces, and cardiovascular health: Findings from a Kaunas cohort study. *Environ. Health* **2014**, *13*, 20. [CrossRef]
10. Cao, Z.; Wu, Z.; Guo, G.; Ma, W.; Wang, H. Quantifying spatial associations between effective green spaces and cardiovascular and cerebrovascular diseases by applying volunteered geo-referenced data. *Environ. Res. Lett.* **2022**, *17*, 014055. [CrossRef]
11. Jaafari, S.; Shabani, A.A.; Moeinaddini, M.; Danehkar, A.; Sakieh, Y. Applying landscape metrics and structural equation modeling to predict the effect of urban green space on air pollution and respiratory mortality in Tehran. *Environ. Monit. Assess.* **2020**, *192*, 412. [CrossRef]
12. Villeneuve, P.J.; Jerrett, M.; Su, J.G.; Burnett, R.T.; Chen, H.; Wheeler, A.J.; Goldberg, M.S. A cohort study relating urban green space with mortality in Ontario, Canada. *Environ. Res.* **2012**, *115*, 51–58. [CrossRef]
13. Grilli, G.; Mohan, G.; Curtis, J. Public park attributes, park visits, and associated health status. *Landsc. Urban Plan.* **2020**, *199*, 103814. [CrossRef]
14. Huang, H.; Yang, H.; Chen, Y.; Chen, T.; Bai, L.; Peng, Z.-R. Urban green space optimization based on a climate health risk appraisal–A case study of Beijing city, China. *Urban For. Urban Green.* **2021**, *62*, 127154. [CrossRef]
15. Cheng, Y.; Zhang, J.; Wei, W.; Zhao, B. Effects of urban parks on residents' expressed happiness before and during the COVID-19 pandemic. *Landsc. Urban Plan.* **2021**, *212*, 104118. [CrossRef]
16. Zhu, J.; Lu, H.; Zheng, T.; Rong, Y.; Wang, C.; Zhang, W.; Yan, Y.; Tang, L. Vitality of urban parks and its influencing factors from the perspective of recreational service supply, demand, and spatial links. *Int. J. Environ. Res. Public Health* **2020**, *17*, 1615. [CrossRef] [PubMed]
17. Lidin, K. Urbanship: In search for a comprehensive definition of urban environment. Проект Байкал **2015**, 84–89. [CrossRef]
18. Seo, S.; Choi, S.; Kim, K.; Kim, S.M.; Park, S.M. Association between urban green space and the risk of cardiovascular disease: A longitudinal study in seven Korean metropolitan areas. *Environ. Int.* **2019**, *125*, 51–57. [CrossRef]
19. Shen, Y.-S.; Lung, S.-C.C. Can green structure reduce the mortality of cardiovascular diseases? *Sci. Total Environ.* **2016**, *566*, 1159–1167. [CrossRef]
20. Wu, J.; Yang, M.; Xiong, L.; Wang, C.; Ta, N. Health-oriented vegetation community design: Innovation in urban green space to support respiratory health. *Landsc. Urban Plan.* **2021**, *205*, 103973. [CrossRef]
21. Li, X.; Chen, C.; Wang, W.; Yang, J.; Innes, J.L.; Ferretti-Gallon, K.; Wang, G. The contribution of national Parks to human health and well-being: Visitors' perceived benefits of Wuyishan National Park. *Int. J. Geoheritage Park.* **2021**, *9*, 1–12. [CrossRef]
22. Royé, D.; Tobías, A.; Figueiras, A.; Gestal, S.; Taracido, M.; Santurtun, A.; Iniguez, C. Temperature-related effects on respiratory medical prescriptions in Spain. *Environ. Res.* **2021**, *202*, 111695. [CrossRef]
23. Jiang, Y.; Huang, G.; Fisher, B. Air quality, human behavior and urban park visit: A case study in Beijing. *J. Clean. Prod.* **2019**, *240*, 118000. [CrossRef]
24. Ning, P.; Wan, Y.; Ren, F. Sina Weibo for the data of time and space hot event detection method. *Geomat. Spat. Inf. Technol.* **2017**, *40*, 33–43.
25. Achrekar, H.; Gandhe, A.; Lazarus, R.; Yu, S.-H.; Liu, B. Predicting flu trends using twitter data. In Proceedings of the 2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Shanghai, China, 10–15 April 2011; pp. 702–707.
26. Weibo User Development Report for 2020. Available online: https://data.weibo.com/report/reportDetail?id=456 (accessed on 13 January 2022).
27. Ji, H.; Wang, J.; Meng, B.; Cao, Z.; Yang, T.; Zhi, G.; Chen, S.; Wang, S.; Zhang, J. Research on adaption to air pollution in Chinese cities: Evidence from social media-based health sensing. *Environ. Res.* **2022**, *210*, 112762. [CrossRef] [PubMed]
28. Wang, S.; Paul, M.J.; Dredze, M. Exploring health topics in Chinese social media: An analysis of Sina Weibo. In Proceedings of the Workshops at the Twenty-Eighth Aaai Conference on Artificial Intelligence, Québec City, QC, Canada, 27–28 July 2014.
29. 2020 Beijing Transport Development annual REPORT. Available online: http://www.bjtrc.org.cn/List/index/cid/7.html (accessed on 13 April 2022).
30. Wen, W.; Ma, X.; Tang, Y.; Wei, P.; Wang, J.; Guo, C. The impacts of meteorology on source contributions of air pollution in winter in Beijing, 2015–2017 changes. *Atmos. Pollut. Res.* **2020**, *11*, 1953–1962. [CrossRef]
31. Xie, Y.-B.; Chen, J.; Li, W. An assessment of $PM_{2.5}$ related health risks and impaired values of Beijing residents in a consecutive high-level exposure during heavy haze days. *Huan Jing Ke Xue = Huanjing Kexue* **2014**, *35*, 1–8.
32. Zhang, S.; Liu, J.; Song, C.; Chan, C.-S.; Pei, T.; Wenting, Y.; Xin, Z. Spatial-temporal distribution characteristics and evolution mechanism of urban parks in Beijing, China. *Urban For. Urban Green.* **2021**, *64*, 127265. [CrossRef]
33. Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
34. Dong, Y.; Liu, H. Health-oriented evaluation of the spatial distribution of urban green space in the Wuhan inner city area of China. In Proceedings of the WIT Transactions on the Built Environment, Seventh International Conference on Harmonization between Architecture and Nature, Brockenhurst, UK, 2–4 October 2018; pp. 37–48.
35. Maas, J.; Verheij, R.A.; Groenewegen, P.P.; De Vries, S.; Spreeuwenberg, P. Green space, urbanity, and health: How strong is the relation? *J. Epidemiol. Community Health* **2006**, *60*, 587–592. [CrossRef]

36. Chen, Y.; Liu, X.; Gao, W.; Wang, R.Y.; Li, Y.; Tu, W. Emerging social media data on measuring urban park use. *Urban For. Urban Green.* **2018**, *31*, 130–141. [CrossRef]
37. Li, F.; Li, F.; Li, S.; Long, Y. Deciphering the recreational use of urban parks: Experiments using multi-source big data for all Chinese cities. *Sci. Total Environ.* **2020**, *701*, 134896. [CrossRef]
38. Lyu, F.; Zhang, L. Using multi-source big data to understand the factors affecting urban park use in Wuhan. *Urban For. Urban Green.* **2019**, *43*, 126367. [CrossRef]
39. Espey, M.; Owusu-Edusei, K. Neighborhood parks and residential property values in Greenville, South Carolina. *J. Agric. Appl. Econ.* **2001**, *33*, 487–492. [CrossRef]
40. Wang, J.; Zhou, W.; Wang, J.; Qian, Y. From quantity to quality: Enhanced understanding of the changes in urban greenspace. *Landsc. Ecol.* **2019**, *34*, 1145–1160. [CrossRef]
41. Jin, S.; Sader, S.A. MODIS time-series imagery for forest disturbance detection and quantification of patch size effects. *Remote Sens. Environ.* **2005**, *99*, 462–470. [CrossRef]
42. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]
43. Zhang, S.; Zhou, W. Recreational visits to urban parks and factors affecting park visits: Evidence from geotagged social media data. *Landsc. Urban Plan.* **2018**, *180*, 27–35. [CrossRef]
44. Liu, J.; Meng, B.; Wang, J.; Chen, S.; Tian, B.; Zhi, G. Exploring the Spatiotemporal Patterns of Residents' Daily Activities Using Text-Based Social Media Data: A Case Study of Beijing, China. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 389. [CrossRef]
45. Wang, K.; Shi, H.; Xu, Y.; Xian, M.; Zhang, Q. Accessibility analysis of urban parks based on GIS. In Proceedings of the 2012 Fifth International Conference on Information and Computing Science, Liverpool, UK, 24–25 July 2012; pp. 56–59.
46. Liu, C.; Li, X.; Han, D. Accessiblity analysis of urban parks: Methods and key issues. *Acta Ecol. Sin.* **2010**, *30*, 5381–5390. (In Chinese)
47. Wang, J.-F.; Li, X.-H.; Christakos, G.; Liao, Y.-L.; Zhang, T.; Gu, X.; Zheng, X.-Y. Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun Region, China. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 107–127. [CrossRef]
48. Wang, J.F.; Xu, C.D. Geodetector: Principle and prospective. *Acta Geogr. Sin.* **2017**, *01*, 116–134. (In Chinese)
49. Kass-Hout, T.A.; Alhinnawi, H. Social media in public health. *Br. Med. Bull.* **2013**, *108*, 5–24. [CrossRef] [PubMed]
50. Dredze, M. How social media will change public health. *IEEE Intell. Syst.* **2012**, *27*, 81–84. [CrossRef]
51. Sakaki, T.; Okazaki, M.; Matsuo, Y. Earthquake shakes twitter users: Real-time event detection by social sensors. In Proceedings of the 19th International Conference on World Wide Web, Rome, Italy, 5–8 June 2017; pp. 851–860.
52. Cai, J.; Huang, B.; Song, Y. Using multi-source geospatial big data to identify the structure of polycentric cities. *Remote Sens. Environ.* **2017**, *202*, 210–221. [CrossRef]
53. Heikinheimo, V.; Tenkanen, H.; Bergroth, C.; Järv, O.; Hiippala, T.; Toivonen, T. Understanding the use of urban green spaces from user-generated geographic information. *Landsc. Urban Plan.* **2020**, *201*, 103845. [CrossRef]
54. Payne, L.L.; Mowen, A.J.; Orsega-Smith, E. An examination of park preferences and behaviors among urban residents: The role of residential location, race, and age. *Leis. Sci.* **2002**, *24*, 181–198. [CrossRef]
55. Kay, S.; Zhao, B.; Sui, D. Can social media clear the air? A case study of the air pollution problem in Chinese cities. *Prof. Geogr.* **2015**, *67*, 351–363. [CrossRef]
56. Zheng, S.; Wang, J.; Sun, C.; Zhang, X.; Kahn, M.E. Air pollution lowers Chinese urbanites' expressed happiness on social media. *Nat. Hum. Behav.* **2019**, *3*, 237–243. [CrossRef]
57. Davenport, T.H.; Barth, P.; Bean, R. How "big data" is different. *Manag. Rev.* **2012**, *54*, 21–25.
58. Anuradha, J. A brief introduction on Big Data 5Vs characteristics and Hadoop technology. *Procedia Comput. Sci.* **2015**, *48*, 319–324.
59. Xie, B.; An, Z.; Zheng, Y.; Li, Z. Healthy aging with parks: Association between park accessibility and the health status of older adults in urban China. *Sustain. Cities Soc.* **2018**, *43*, 476–486. [CrossRef]
60. Yin, S.; Shen, Z.; Zhou, P.; Zou, X.; Che, S.; Wang, W. Quantifying air pollution attenuation within urban parks: An experimental approach in Shanghai, China. *Environ. Pollut.* **2011**, *159*, 2155–2163. [CrossRef]
61. Pace, R.; De Fino, F.; Rahman, M.A.; Pauleit, S.; Nowak, D.J.; Grote, R. A single tree model to consistently simulate cooling, shading, and pollution uptake of urban trees. *Int. J. Biometeorol.* **2021**, *65*, 277–289. [CrossRef] [PubMed]
62. Johnson, D.; Parker, J.D. Air pollution exposure and self-reported cardiovascular disease. *Environ. Res.* **2009**, *109*, 582–589. [CrossRef] [PubMed]
63. Zhao, X.; Lu, M.; An, Z.; Li, J.; Li, H.; Zhao, Q.; Wu, Z.; Wu, W.; Liu, Y.; Song, J. Acute effects of ambient air pollution on hospital outpatients with chronic pharyngitis in Xinxiang, China. *Int. J. Biometeorol.* **2020**, *64*, 1923–1931. [CrossRef] [PubMed]
64. Wu, S.; Ni, Y.; Li, H.; Pan, L.; Yang, D.; Baccarelli, A.A.; Deng, F.; Chen, Y.; Shima, M.; Guo, X. Short-term exposure to high ambient air pollution increases airway inflammation and respiratory symptoms in chronic obstructive pulmonary disease patients in Beijing, China. *Environ. Int.* **2016**, *94*, 76–82. [CrossRef] [PubMed]
65. Zhai, W.; Bai, X.; Shi, Y.; Han, Y.; Peng, Z.-R.; Gu, C. Beyond Word2vec: An approach for urban functional region extraction and identification by combining Place2vec and POIs. *Comput. Environ. Urban Syst.* **2019**, *74*, 1–12. [CrossRef]

66. Wang, Z.; Ma, D.; Sun, D.; Zhang, J. Identification and analysis of urban functional area in Hangzhou based on OSM and POI data. *PLoS ONE* **2021**, *16*, e0251988. [CrossRef]
67. Tu, X.; Huang, G.; Wu, J.; Guo, X. How do travel distance and park size influence urban park visits? *Urban For. Urban Green.* **2020**, *52*, 126689. [CrossRef]