Article

Predict the Suitable Places to Run in the Urban Area of Beijing by Using the Maximum Entropy Model

Liuyi Song ^{1,2} and An Zhang ^{1,2,*}

¹ State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; songly.20s@igsrr.ac.cn

² College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100190, China

* Correspondence: zhangnan@igsrr.ac.cn

Abstract: Many people in the world do not have enough physical activities to maintain good health, which has recently become a threat to public health. In addition to individual genetic and social factors, we considered the geographical environment of the city as a factor that affects these healthy physical activities. We used the location-based data in social media combined with the open geographic data to explore the impact mechanism of urban environmental factors on human running behaviors. This study collected nine urban environmental variables and preference tracks in Beijing's main urban area. We used the Maximum Entropy Model (MaxEnt) to analyze the relationship between running behaviors and environmental variables and identify suitable areas for running in Beijing. The results showed that: firstly, the variables of attractions, sports and sidewalk density contributed the most to running suitability. Secondly, 47.5% of the main urban areas in Beijing are suitable for running, mainly in the main urban areas with better economic development. Thirdly, the distribution of suitable places for running is unfair in that some places with large populations do not have a matching running environment.

Keywords: running routes; physical activities; urban environment; MaxEnt

Citation: Song, L.; Zhang, A. Predict the Suitable Places to Run in the Urban Area of Beijing by Using the Maximum Entropy Model. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 534.

<https://doi.org/10.3390/ijgi10080534>

Academic Editor: Wolfgang Kainz

Received: 21 May 2021

Accepted: 4 August 2021

Published: 9 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

There are pieces of evidence to show that physical activities such as walking, running and cycling can bring many health benefits to humans [1–3], especially in the etiology of many chronic diseases, including cancer, cardiovascular disease, hypertension and obesity [4,5]. Globally, many adults and children do not have enough physical activity to maintain good health [6]. In traditional socio-economic research, human behavior is usually assumed as a steady-state random Poisson process. However, related research on human behavior dynamics has found that the spatial distribution of human motion may have complex dynamic mechanisms. There are generally non-Poisson characteristics [7–9], driven by various factors. Environmental ethology combines the relationship and interaction between human behavior and its surrounding material environment, studies the factors of humans and the environment, explores the influences of environmental factors on human behavior and considers the demand of human activities in the urban environment. In addition to the genetic factors that determine the tendency of physical activity [10,11] and personal social factors [12–14], the city's geographical environment is also considered a factor affecting these healthy physical activities in a broad sense. Many environmental health studies have explored whether and how social and physical environment affect health and found that the degree of its impact is controlled by the type of physical environment [15–18]. In geographic and health research, which seeks to understand how environmental factors affect health behaviors and outcomes, ecological models can explore the causal relationship between healthy behaviors and urban physical environment [19]. At present, there are many discoveries about the interaction between people's physical activities and urban

environmental factors: for example, residents' walking activities (including walking routes and walking experience) are significantly affected by the number of stores, the number of bus stops, road shading rate and sidewalk density [20,21]; residents' activity experience in the urban environment is affected by sidewalk environment, street facilities, retail building floor area ratio and residential building density [22,23]; the level of physical activity of urban residents is significantly related to the characteristics of the built-up environment, and the availability of natural and semi-natural spaces (such as parks) [24–26], the environmental attributes of public open space will affect when and where residents engage in physical activity. Many previous types of research have used field research methods to record athletes' preferences for urban street routes through interviews, questionnaires or network scoring systems, to dig out the influencing environmental factor [27]. For example, Walkability App launched by Walkonomic company allows users to score the walkability of each street from eight aspects according to the open data, and mark it on Google map with color-coding, which can be used as a reference for residents to choose commuting routes and places of residence. In addition, compared with the traditional questionnaire survey method, the current development of new technology provides more options for physical activity researches: combining the physical activity sensor (such as pedometer, accelerometer) or wearable camera [28,29] with GPS technology [30,31], we can evaluate the physical activity in a specific environment and specific behavior [32] through the activity log; for example, Patrick Norman analyzed the walking, running and cycling data in a sports APP MapMyfitness, and did a simple qualitative analysis and comparison with traditional data [33].

In summary, the current research mainly focuses on the walking environment, discussing the impacts of the urban material environment on residents' walking activities. Most of the samples are small-range questionnaires or individual experiments. However, most of the research on running trajectory is still in the visualization stage. The data in the sports app has not been further analyzed, and the investigations on its spatial behavior are still more minor.

In China, with the rapid development of urbanization, automobile-led urban streets have a significant impact on the outdoor activities of residents. The residents' physical activities have essentially transformed into indoor, while the public space for physical activities decreases. However, outdoor natural environments significantly positively affect participants [34,35]. Therefore, understanding the causes of outdoor physical activities is essential for developing and improving public health interventions [36]. Encouraging more physical activities through urban planning measures can reduce the residents' chronic disease risk and also help to improve the vitality of urban streets. Due to the rapid development of location-based services and Volunteered Geographic Information (VGI), many sports track data with geographic tags (including walking, running and riding) are uploaded and shared on the Internet by users. At the same time, points of interest (POIs) and OpenStreetMap (OSM), which are based on map services, have become the data sources of urban human activity research. These big data have the characteristics of a large amount of data, fast updates and easy access, providing detailed spatiotemporal information for human activity research on a large scale. This study takes the main urban area of Beijing as the research area. In order to explore the urban street environmental factors that drive the running activities, we use the maximum entropy model to predict the running suitability in the main urban area of Beijing based on the running track data and urban geographical environment data. We try to explain how the running behavior is affected by the urban environmental factors and propose suggestions for urban running road planning in the future. This study provides a reference basis for designing a healthy city and reflects the "human-centered" urban concept.

2. Materials and Methods

2.1. Materials

2.1.1. Running Routes

This paper selects the central urban area of Beijing as the research area, including six districts: Xicheng, Dongcheng, Haidian, Chaoyang, Shijinshan and Fengtai. Beijing is the capital of China and the economic, cultural and political center. The total area comprises 1381 square kilometers with a population of 11 million in 2019. The main urban area of Beijing is basically in the plain area, the elevation of 95% of the area is between 20–60 m and the average elevation is 43.5 m. Beijing has put forward the goal of a healthy city for residents through urban space. Adidas Runtastic is a running fitness app that monitors and records the user's activity information and provides the query function for other user's trajectory information. In this study, 153 routes were selected from the Adidas Runtastic' website in the main urban area of Beijing, with a total of 527 users participating. The number of some popular running route users reached 25, as shown in Table 1. It includes route description, tags, average speed, number of users and the average age of users, indicating that the running trajectory data can reflect the residents' running preference areas to a certain extent.

Table 1. The sample data of Adidas Runtastic.

ID	Route Descriptions	Tags	Avg Speed (km/h)	Users Count	Avg Age
1	Regular exercise	[road, exercise, flat, high, marked]	7.85	25	33
2	Behai park around once	[asphalt, exercise, flat, high, accessible]	9.14	16	31
3	10 km including north and south park	[trail, exercise, flat, normal, accessible]	10.18	11	34
4	From the company to Forest Park	[road, exercise, undulating, less, accessible]	12.78	7	24
5	Two half marathon in North South Park of Orson	[road, exercise, undulating, normal, marked]	6.7	7	30

2.1.2. Natural Environment Variables

Outdoor running activities rely on natural space, and green spaces have apparent support for these linear physical activities [37]. The multispectral satellite (raster information) can monitor the green space in urban areas, and the Normal Difference Vegetation Index (NDVI) is a good indicator of remote sensing measurement results [38]. Therefore, NDVI is selected as one of the environmental variables that may drive running behavior. Due to the Urban Heat Island effect (UHI), the Land Surface Temperature (LST) of the city center is often higher than that of the surrounding area [39]. The impact of heat island can lead to serious health problems and even increase the mortality rate in some cities. For this reason, we choose LST as the second environmental factor that may drive running behavior.

2.1.3. Street Variables

In addition to taking on the role of transportation, streets are also the living spaces for urban residents; the physical environment of the street provides the place for the residents to live in daily life and meets people's needs for space, environment and activities. For example, residents in blocks with exclusive sidewalks will likely engage in more recreational walking and cycling activities. Streets that support public transportation will reduce traffic congestion and help people shorten commuting time; the streets with high accessibility of convenience stores and restaurants can provide people with convenient food sources. The spatial types with vital health service functions include green space, natural landscape, community sports facilities. Research shows that human health services are most clearly reflected in these urban spaces [37]. In addition, the number of bus stops and subway stations can reflect the traffic convenience and traffic volume of the street to a

certain extent, which will also affect the route choice of runners [27]. The buildings also provide the athletes with a view of the landscape and a sense of security [40]. In this study, the seven types of POI facilities in the street with attribute information were selected as environmental variables that may drive running behavior. Including bus stops, subway stations, life service POI, scenic spots POI, sports and leisure POI, the length of sidewalks and the area of buildings (Table 2). Among them, life service facilities include personal care, convenience stores, express delivery points and other categories, scenic spots include scenic spots, squares, parks and schools, and sports and leisure include gyms, gymnasiums and residents' activity centers.

Table 2. The information of running routes data and variables.

Variables	Data Descriptions	Data Size	Data Sources
Running routes	Preferred running routes, data acquisition year is 2017	153 routes	Adidas Runtastic: https://www.Runtastic.com/ , accessed on 1 April 2020
LST	Land Surface Temperature, 2017	1537 grids (1 km × 1 km)	https://developers.google.com/earth-engine/datasets/catalog/MODIS_006_MOD11A1 , accessed on 1 April 2020
NDVI	Normal Difference Vegetation index, 2017	1537 grids (1 km × 1 km)	https://developers.google.com/earth-engine/datasets/catalog/MODIS_MOD09GA_006_NDVI , accessed on 1 April 2020
Buildings	The density of the area of buildings, 2018	52,221 polygons	Open Street Map: https://download.geofabrik.de/asia.html , accessed on 1 April 2020
Length of Sidewalks	The density of sidewalks used for exercise, 2018	26,889 lines	Gaode Map: https://lbs.amap.com/demo/jsapi-v2/example/layers/roadnet/ , accessed on 1 June 2019
Attractions	The density of scenic spots, squares, parks, 2018	3786 points	Points of interest (POI) from Gaode Map: https://lbs.amap.com/api/android-sdk/guide/map-data/poi/ , accessed on 1 June 2019
Bus stops	The density of bus stops	3424 points	
Service	The density of personal cares, convenience business supers, express points, 2018	28,390 points	
Sports	The density of gyms, gymnasiums, activity centers, 2018	15,893 points	
Subway stations	The density of subway stations, the data date is 2018	329 points	

2.2. Methods

The overview of the methodology is shown in Figure 1. First of all, it is necessary to preprocess the obtained data and unify the calculation to the same scale (1 km × 1 km). Then, the correlation analysis of environmental variables is carried out to select maximum entropy model calculation variables. After that, the running route is divided into the training set and validation set. Finally, the suitability of running is calculated by the maximum entropy model.

Figure 1. The overview of the methodology.

2.2.1. MaxEnt Modeling

Mathematician and physicist E. T. Jaynes propose the principle of maximum entropy (MaxEnt), and this method has been successfully applied in many fields [41], such as the distribution prediction of species. The principle of maximum entropy is a Bayesian inference method [42], which is one of the criteria for learning probabilistic models; when the entropy of the model is the maximum, it is considered the best model.

Assuming the probability distribution $P(X)$ of the discrete random variable X , its entropy is:

$$H(X) = - \sum_{i=1}^N P(x_i) \log P(x_i) \quad (1)$$

Entropy satisfies the following inequalities:

$$0 \leq H(P) \leq \log|X| \quad (2)$$

When X is uniformly distributed, the equal sign on the right holds and entropy is maximum. Firstly, the prior distribution [43] of the distribution with the smallest amount of information in the problem is determined. Then the relative entropy is maximized as far as possible in combination with the known constraint conditions. In this way, the true probability distribution can be calculated. The time complexity is $O(n \log n)$. $N \log n$ times need to be compared to eliminate the uncertainty of information and obtain the optimal result. The finally selected probability model satisfies the known constraints.

Both the distribution of human activities and the distribution of species are affected by the surrounding environment. Therefore, the maximum entropy model was used to predict the running behaviors of residents in the main urban area of Beijing. We used the MaxEnt software (Version 3.4.4) for MaxEnt Modeling, the MaxEnt Modeling can be download in the MaxEnt home page (http://biodiversityinformatics.amnh.org/open_source/ accessed on 1 April 2020), and the software was produced and released by the Center for Biodiversity and Conservation(CBC) at the American Museum of Natural History. In this study, the running hotspots were used to input the maximum entropy model simulation. The input format is CSV. If there is a preferred running track, the grid value is 1. Otherwise, it is 0, including longitude and dimension information; nine variables are input as environment variables in ASCII format, including the same coordinate information. The range of different variables needs to be consistent; Bootstrap sampling was used to form pseudo samples to estimate the population. The maximum iteration term was set to 5000. The regularization multiplier was 1.5; due to the randomness of sampling, the model results fluctuated; a total of 30 model operations were carried out, and the average value was taken as the final result of the model simulation. The computing environment of the model is Windows10 (64-bit), 16 GB memory, RTX2060 graphics card and Java1.8.0. It takes 10 min to calculate the model once, and 300 min to calculate 30 times.

2.2.2. Ripley's K Function

Ripley's K function, also known as multi distance spatial clustering analysis, is an analysis method of the point data pattern. Its statistics can reveal the specific pattern of point feature distribution under different spatial observation scales. In this study, we use a standard transformation of Ripley's K function, which is usually called $L(d)$ [44]:

$$L(d) = \sqrt{\frac{A \sum_{i=1}^n \sum_{j=1, j \neq i}^n k_{ij}}{\pi n(n-1)}} \quad (3)$$

d is the distance, n is the total number of elements, A is the total area of elements and k_{ij} is the weight. If there is no edge correction when the distance between i and j is less than d , the weight will be equal to 1; otherwise, the weight will be 0.

Ripley's K-function was used to analyze the running track data to explore its spatial distribution pattern. Choose 50 km as the maximum search radius of Ripley's K function,

the number of distance segments is 100, and each time increase 500 m; The confidence interval is 99%; Use the bounding rectangle range as the bounding area.

2.2.3. Variables Calculation

According to formula (4), the indexes of POI points of bus stops, subway stations, sports facilities, service facilities and attractions facilities were calculated, and assigned the attributes of the total number of POIs for each grid; if the POI point is in a specific grid, w_i is 1, otherwise w_i is 0; S_{grid} is the area of a single grid. In this paper, the area is 1 km².

$$\text{Count} = \frac{\sum_{i=1}^n w_i}{S_{grid}} \quad (4)$$

For the data of the plane-shaped buildings, firstly, we identified the unique ID of each grid to each building. Next, we summarized the total area of the buildings on each grid according to formula (5); S_i is a single building in the grid, S_{grid} is the area of a single grid. In this paper, the area is 1 km².

$$\text{Area} = \frac{\sum_{i=1}^n S_i}{S_{grid}} \quad (5)$$

For the data of linear sidewalks, the unique ID of each grid was also marked on each sidewalk. Moreover, the total length of each grid sidewalk was summarized according to formula (6). L_i is the length of a single sidewalk in the grid, S_{grid} is the area of a single grid. In this paper, the area is 1 km².

$$\text{Length} = \frac{\sum_{i=1}^n L_i}{S_{grid}} \quad (6)$$

2.2.4. Variables' Correlation

We selected natural environmental variables, and seven socio-economic variables were analyzed to explore a linear correlation. Choose the Pearson Correlation analysis method. The Pearson correlation coefficient is used to judge whether the two are linear.

2.2.5. Model Validation

As shown in Figure 2, these preferred running routes are mainly concentrated in the center, including some parks, such as Orson Park, the Summer Palace, Yuyuantan and the famous Chang'an Street. The central area of Beijing has been divided into 1537 grids of 1 km × 1 km, which is more than 1381 km² to include all the areas around the urban boundary line. We defined the running hotspots as the grids which contain the running route. Furthermore, there are 388 running hotspots in the central area of Beijing, and the rest of the blank grid has no preference for running routes. In order to measure the validity of the model results, 80% of the running hotspots were randomly selected as the training set of the model, the number of input models was $388 \times 80\% = 310$, the remaining 20% (78) hot spots were used as the validation set of the model results (Figure 2). According to formula (7), the higher the hit rate successfully predicted by the model, the higher the accuracy.

$$P = \frac{\sum_{i=1}^n 1}{388} \quad (7)$$

Among them, P is the hit rate of the model, and $\sum_{i=1}^n 1$ are the verification grids that successfully fall into the prediction range.

Figure 2. The distribution map of running routes.

3. Results

3.1. Results

3.1.1. Spatial Aggregation Analysis of Running Track Points

Ripley's K has been widely used to measure the aggregation of spatial units [45,46]. We acquire Expected K and Observed K through Ripley's K function analysis on the data of running track points. The Observed K value refers to the actual density value calculated by us, while the Expected K value refers to the expected distribution in random distribution. If the Observed K at a certain distance is greater than the Expected K, the distribution has a higher degree of clustering than the random distribution of that distance (analysis scale); if the Observed K is less than the Expected K, the distribution is more discrete than the random distribution of the distance. The results show that when the Expected K is 10 km, the running track in the main urban area of Beijing reaches the highest degree of aggregation. It indicates that the running track points have been significantly aggregated in space within 10 km. Therefore, in the subsequent calculation of the Maximum Entropy Model, all variables are processed on a grid of $1 \text{ km} \times 1 \text{ km}$.

3.1.2. Variables Analysis

The spatial distributions of nine environmental variables were shown as grids of $1 \text{ km} \times 1 \text{ km}$ in Figure 3. Due to the western mountainous area, the values of NDVI and LST in the central and eastern urban areas were lower than those of the southwest. The environmental variables of each street showed a trend of gradually spreading from the urban center to the outside: the high-density areas were concentrated in Dongcheng District, Xicheng District, the southeast of Haidian District, the west of Chaoyang District

and the northeast of Fengtai District, which are consistent with the level of regional economic development; sports variable, services variable and attractions variable are mainly concentrated at the junction of the central part of Dongcheng District and Chaoyang District, Beijing railway station, Sanlitun and other commercial intensive areas, as well as the temple of Heaven Park, the Forbidden City and other scenic attractions are located here; the density distribution of bus stops is the same as that of Beijing's ring roads, while the density of subway stops extends from the center to the periphery.

Figure 3. Spatial distribution of 9 environmental variables, all variables are in 1 km grid after calculation.

Calculated the correlation matrix of 9 kinds of environmental variables and perform correlation analysis. As shown in Figure 4, there was a strong positive correlation between the sports variable and services variable (0.80), indicating that where the number of sports facilities was high, the number of public services facilities was also high.

Figure 4. Correlation matrix of 9 environmental variables.

3.1.3. Running Suitability Map

The performance of the MaxEnt model was evaluated according to the area under the curve (AUC) of receiver operating characteristic (ROC) [47], the higher the AUC value is, the more reliable the prediction is, and it is far from the random distribution of 0~1. Generally, the AUC division standards are as follows: excellent (0.9–1.0), good (0.8–0.9), acceptable (0.7–0.8), poor (0.6–0.7) and insufficient (0.5–0.6) [48,49]. The AUC of the prediction model of running suitability in the central urban area of Beijing is 0.79 (Figure 5), which indicates that the prediction accuracy of the model obtained is good. Residents' running behavior is not randomly distributed in space but related to nearby environmental factors. Therefore, this model is adopted, and the result can distinguish the probability of running behavior in different areas well.

Figure 5. The AUC values derived from the Maxent models.

Figure 6 shows the contribution of each environmental variable to the model, which is the average result of overlapping tests of different variable importance. The green bar refers to the impact of a single environmental variable not included in the model, and

the dark blue bar refers to the individual contribution of the variable to the model. The density of attractions, sports and sidewalks are the three most significant gains when we use each environmental variable alone. The density of building areas and bus stations can also promote running activities to a certain extent. NDVI and LST contribute little to the model. The contribution of the density of subway stations is in the least.

Figure 6. Jackknife of regularized training gain for running.

The probability P of running behavior in the central urban area of Beijing obtained by the maximum entropy model is between 0 and 1. The higher the p -value is, the higher the probability of residents' running behavior is. Moreover, the higher the suitability of running is, the more suitable the environmental conditions in this area are for running activities. From the comprehensive influence of nine environmental variables, the occurrence probability of running behavior is the reflection of the surrounding environment of the region: the most suitable area for running in Beijing is in the Summer Palace, with the suitability of 0.80, and the lowest area is at the junction of the South Fifth Ring and Jingkai Expressway, with the suitability of 0.05. The results of running fitness in Beijing were divided into five levels according to the natural breakpoint method (Level 1: 0.05–0.20, Level 2: 0.20–0.32; Level 3: 0.32–0.46, Level 4: 0.46–0.59, Level 5: 0.59–0.80), as shown in Figure 7.

In Beijing, the area of level 1 is 503 km², accounting for 32.73%; the area of level 2 is 304 km², accounting for 19.78%; the area at level 3 is 280 km², accounting for 18.22%; the area of level 4 is 249 km², accounting for 16.20%; and the area of level 5 is 201 km², accounting for 13.08%; the area of level 3 or greater is 730 km², accounting for 47.50% of the total area. The running suitability levels in the old central districts of Beijing, including Dongcheng District and Xicheng District, are almost at levels 4 and 5, which are very high. Sports facilities, attractions and sidewalks are also high in these areas. The areas close to the center of the central city, including the southeast of Haidian District and the west of Chaoyang District, are also high suitability areas for running. These areas have a high level of comprehensive development and gather many universities and scientific research institutes. However, the running suitability of the marginal areas of the central urban area, such as the north of Haidian District and Shijingshan District, the south of Fengtai District and the east of Chaoyang District, is at level 2 and below. From the perspective of environmental factors, the infrastructure of these areas is inferior. There are few places suitable for running. It is not suitable for running, from the perspective of environmental conditions in these areas, the infrastructures in these areas are lacking. The area of suitable places for running are less and are not suitable for running. Therefore, these areas are not suitable for running. There are mountains in the west and southwest of Haidian District and the northwest of Fangshan District. The natural conditions here are good; thus, some areas have high suitability for running.

Figure 7. Running suitability map based on maximum entropy model prediction.

3.1.4. The Relationship between Environmental Variables and Running Suitability Values

Even though the street material environmental variables positively affect the model, the relationship between them and the suitability values for running is not a simple linear increase. As shown in Figure 8, the top six environmental variables contributing to the model and the running suitability values are selected to draw the response curve, respectively: the suitability value for running can reach the highest value when the density of the three types of environmental variables, attractions, buildings and service is moderate, these three types of environmental variables can reflect the built-up area environment and the conditions of public facilities; when the density of these three variables is too high, the living environment may be more crowded, which is no longer suitable for running activities; the higher the density of bus stops, sidewalks and sports, the greater the suitability values for running, the density of bus stops and sidewalks reflects the accessibility level of the region, and high accessibility tends to lead to high urban activities [50], and sports facilities are the best places for sports activities.

Figure 8. Response curves of running suitability values to predominant factors derived from the model.

3.1.5. Spatial Mismatch between High Population Density and Low Running Suitability Values

Previous research has found that physical activity distribution is not uniform, and running activities are more active in economically developed areas [51]. In this study, we also found that the running activities in Beijing depend on the developed urban environment. However, this phenomenon reflects the unfairness of the spatial distribution of regional running suitability: although the population density of some areas is very high and there is a tremendous demand for running places, the value of predicted running suitability is not high, the environmental conditions are not suitable for residents to run; finding out these unfair areas can provide a theoretical basis for urban renewal, improve the running service function of these areas and promote the use of public space by residents. In the main urban area of Beijing, the area with a population density higher than 25,000 people/km² but the value of predicted running suitability lower than 0.5 is 69 km², accounting for 4.49% of the main urban area. As shown in Figure 9, such unbalanced development areas are mainly concentrated in Haidian District, Fengtai District and Chaoyang District: for example, Xierqi, in the northeast of Haidian District, is the gathering place of many large Internet companies in China, and is the employment-oriented functional block in the city, the population density here is high, but the environmental conditions are not suitable for running activities; the primary function of the Summer Palace in the south of Haidian District and the area around Yuyuantan is tourism, lacking sports and service facilities, and the density of sidewalks is low; Zhangyicun and Beidadi in Fengtai District are in the stage of development, what's more, the Xinfadi area with comprehensive function and Songjiazhuang area with residential function orientation also have the problem that the

surrounding environment conditions cannot meet the demand of running activities; in addition, this problem also exists around Wangjing and Jinsong in Chaoyang District.

Figure 9. Areas with unfair running conditions. The source of the basemap: Esri, World Imagery, <https://www.arcgis.com/home/item.html?id=10df2279f9684e4a9f6a7f08febac2a9>, accessed on 1 April 2021.

4. Discussions

4.1. Validation

This study aims to predict the suitability of running in the main urban area of Beijing. Therefore, in addition to evaluating the accuracy of the result prediction through the ROC curve output by the model, the simulated running suitability results are also verified through the verification points randomly selected in advance. Among the 78 verification points, there are 10 verification points with a running suitability value at level 1, accounting for $10/78 = 12.82\%$. Eight verification points with a running suitability value at level 2, accounting for $8/78 = 10.26\%$. There are 13 verification points with a running suitability value of level 3, accounting for $13/78 = 16.67\%$; 19 verification points with a running suitability value of level 4, accounting for $19/78 = 24.36\%$; and 28 verification points with a running suitability value of level 5, accounting for $28/78 = 35.90\%$; finally, 60 verification points with a running suitability value higher than level 3, accounting for $60/78 = 76.92\%$. The results further verified the reliability of the model. They reflected the effects of the nine environmental variables, which indicated that the model results could predict the running suitability of the main urban area of Beijing to a certain extent and reflect the residents' running preference.

4.2. Advice on Running Routes Planning

To create a more livable urban environment and advance residents to run in urban public spaces, urban space planning should pay more attention to improving space quality. Therefore, urban construction should not be ordinary and one-sided expansion, but rather design different urban space scientifically and accurately according to local conditions. In 2007, WHO determined 11 elements of outdoor space design of cities in the global era, including green space, outdoor seats, sidewalks, transportation, circulation path, architecture [52]; these design elements of space planning can be used to build a more suitable public space for running, Beijing is a highly compact and dense city, the renovation of urban facilities in old areas must be flexible: for residential areas with high population

density, we should pay more attention to the coordination between the physical activities such as running and the overall urban spatial planning, for example, Tiantongyuan is a massive community in Beijing with a high population density, we can increase the number and length of sidewalks and the safety of intersections, reduce the speed limit of motor vehicles and improve the walking accessibility and improve the convenience of running activities. In addition, some community centers can be set up to provide running places for residents. In employment-oriented office areas, such as Zhongguancun, the suitability of running activities should be improved based on meeting the regional functions, more outdoor stadiums should be added, set up some rest areas to maintain the structure and quality of the plants and maintain the beauty of the scenery, increase the possibility of running after work. The diversity of spatial functions should be promoted. For mixed areas, such as Qinghe and other areas, the accessibility of attractions, sports facilities and others should be enhanced to attract more residents to participate in healthy running activities, transform or build new urban green spaces and promote the connectivity between streets and parks. In summary, to realize the inevitable demand of the refinement and humanization of urban planning, we need to plan urban space reasonably, realize the compound and intensive utilization of space, provide sufficient and suitable public space for residents' running activities, and promote a healthy way of life.

4.3. Limitations

VGI data provides the possibility for this study. However, at the same time, VGI data may face some problems: firstly, the data is incomplete and regional heterogeneity. For example, the amount of data in a densely populated city is significantly better than a sparsely populated small city, which cannot cover all the population. Secondly, VGI data may be messy, and sometimes it takes much time to clean the data to make it analyzable. In addition, the quality and reliability of the data need to be verified. Finally, there are ethical issues to consider with VGI data, such as the ownership of the data and whether the data contributors are informed [41,53,54]. In addition, there are some software and websites that provide similar running track information, such as Strava heat map (<https://www.strava.com/> accessed on 1 April 2021), which provides an overview of all the runs carried out by athletes and occasional runners in the world, differently by Runtastic users that decide to voluntarily upload single traces on the platform to be made available. Even though the data density in Strava is higher, it cannot obtain accurate trajectory and users' information. In contrast, the users in Adidas Runstatic are mostly the general public, which can represent the general level of residents better and provide more information about users. For example, through the number of users, we can know that this running route is trendy. In addition, Adidas Runtastic is a smartphone app. The user group may be young and modern. There is user bias. Due to privacy settings, we cannot acquire the user's personal information. The data obtained from the fitness app is different from the ordinary counter, field investigation and other methods. Therefore, this running preference data may not be extended to the general population. Hence it cannot be very accurate to reflect the running activities of all residents. In addition, the acquisition time of running track data is 2017. Even though we selected the environmental variable data of the same period as far as possible, the data of environmental variables did not match it very well in time. Even though the update speed of each selected environmental variable is not quick, it may have an inevitable impact on the accuracy of the results.

5. Conclusions

In this study, based on the VGI data and the maximum entropy model, we predicted the regional distribution of running suitability in the main urban area of Beijing and the influence mechanism of the built-up environment variables of each city on the running behavior was determined. The shortcomings of the existing planning in some areas were found, which shows that this method has great potential for studying spatial behavior.

The results show that the running activities of residents did not show a simply random Poisson distribution but were driven by various urban environmental variables. Firstly, the areas with high suitability for running in the main urban area of Beijing are mainly concentrated in the developed areas with relatively complete public infrastructure. Secondly, the suitability for running is higher in the western mountainous areas with good natural conditions, while running suitability is lower in the fringe areas of the main urban area with inadequate public facilities. Thirdly, the density of attractions, buildings and services with moderate density can promote running activities to the greatest extent. The density of bus stops, sidewalks and sports has a monotonically increasing trend in the effect of running suitability values. The higher the traffic accessibility, the more running activities can be triggered; some areas have insufficient planning of existing facilities, and the environmental conditions cannot provide a good running service function for many people. Some countries and regions have launched urban renewal projects, hoping to promote the healthy life of residents from the perspective of the urban environment, and some promising results have been achieved. For example, Knox County in the United States spent 2.1 million dollars in 2005 to build an 8-foot-wide and 2.9-mile-long asphalt greenway in the intervention community; through the comparative experiments, it was found that two years later (2007), residents in the intervention area spent significantly more time on physical activities [55]. The Roald region in Denmark also spent about 35 million euros in large-scale urban renewal in 2009: four new urban green spaces and playgrounds were created in the region, and large public parks were renovated, which also significantly increased the activity time of teenagers in the area through comparative experiments [56]. In 2021, the Beijing Municipal Government issued the Guidance Opinions of the Beijing Municipal People's Government on the Implementation of Urban Renewal Action, marking the official start of the urban renewal work in Beijing. In addition to the renovation methods mentioned above, we can learn from the existing experience of developed countries. Combined with the results of this study, consider adding sidewalks, stadiums and other measures according to local conditions to create a suitable living environment for runners and contribute to public health.

Author Contributions: Conceptualization, Liuyi Song and An Zhang; methodology, An Zhang; software, Liuyi Song; validation, Liuyi Song; formal analysis, An Zhang; investigation, An Zhang; resources, An Zhang; data curation, Liuyi Song; writing—original draft preparation, Liuyi Song; writing—review and editing, An Zhang; visualization, Liuyi Song; supervision, An Zhang; project administration, An Zhang; funding acquisition, An Zhang. Both authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China (Grant No. 2017YFB0503500).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. US Department of Health Human Services. *US Department of Health and Human Services 2008 Physical Activity Guidelines for Americans*; J. Elite Esthetics: Hyattsville, MD, USA; Washington, DC, USA, 2008; pp. 1–40. [CrossRef]
2. WHO. *Global Recommendations on Physical Activity for Health*; World Health Organization: Geneva, Switzerland, 2010. Available online: <https://www.who.int/publications/i/item/9789241599979> (accessed on 1 April 2021).
3. Warburton, D.E.; Charlesworth, S.; Ivey, A.; Nettlefold, L.; Bredin, S.S. A systematic review of the evidence for Canada's Physical Activity Guidelines for Adults. *Int. J. Behav. Nutr. Phys. Act.* **2010**, *7*, 39. [CrossRef] [PubMed]
4. Friedenreich, C.M.; Neilson, H.K.; Lynch, B. State of the epidemiological evidence on physical activity and cancer prevention. *Eur. J. Cancer* **2010**, *46*, 2593–2604. [CrossRef]

5. Nocon, M.; Hiemann, T.; Muller-Riemenschneider, F.; Thalau, F.; Roll, S.; Willich, S.N. Association of physical activity with all-cause and cardiovascular mortality: A systematic review and meta-analysis. *Eur. J. Cardiovasc. Prev. Rehabil.* **2008**, *15*, 239–246. [[CrossRef](#)] [[PubMed](#)]
6. Hallal, P.C.; Andersen, L.B.; Bull, F.C.; Guthold, R.; Haskell, W.; Ekelund, U. Global physical activity levels: Surveillance progress, pitfalls, and prospects. *Lancet* **2012**, *380*, 247–257. [[CrossRef](#)]
7. Brockmann, D.; Hufnagel, A.L.; Geisel, T. The scaling laws of human travel. *Nat. Cell Biol.* **2006**, *439*, 462–465. [[CrossRef](#)] [[PubMed](#)]
8. González, M.C.; Hidalgo, C.A.; Barabasi, A. Understanding individual human mobility patterns. *Nat. Cell Biol.* **2008**, *453*, 779–782. [[CrossRef](#)] [[PubMed](#)]
9. Rhee, I.; Shin, M.; Hong, S.; Lee, K.; Kim, S.J.; Chong, S. On the Levy-Walk Nature of Human Mobility. *IEEE/ACM Trans. Netw.* **2011**, *19*, 630–643. [[CrossRef](#)]
10. Ornstein, K.; Galea, S. Teaching Epidemiology: A Guide for Teachers in Epidemiology, Public Health and Clinical Medicine. Third Edition: Edited by Jorn Olsen, Rodolfo Saracci, and Dimitrios Trichopoulos. *Am. J. Epidemiol.* **2010**, *172*, 1207–1208. [[CrossRef](#)]
11. Rhodes, E.R.; Smith, N.E.I. Personality correlates of physical activity: A review and meta-analysis. *Br. J. Sports Med.* **2006**, *40*, 958–965. [[CrossRef](#)]
12. Craggs, C.; Corder, K.; van Sluijs, E.; Griffin, S.J. Determinants of Change in Physical Activity in Children and Adolescents: A Systematic Review. *Am. J. Prev. Med.* **2011**, *40*, 645–658. [[CrossRef](#)]
13. Rhodes, R.E.; Martin, A.D.; Taunton, J.E.; Rhodes, E.C.; Donnelly, M.; Elliot, J. Factors Associated with Exercise Adherence Among Older Adults. *Sports Med.* **1999**, *28*, 397–411. [[CrossRef](#)]
14. Sallis, J.F.; Prochaska, J.J.; Taylor, W.C. A review of correlates of physical activity of children and adolescents. *Med. Sci. Sports Exerc.* **2000**, *32*, 963–975. [[CrossRef](#)] [[PubMed](#)]
15. Cummins, S.; Curtis, S.; Diez-Roux, A.V.; Macintyre, S. Understanding and representing ‘place’ in health research: A relational approach. *Soc. Sci. Med.* **2007**, *65*, 1825–1838. [[CrossRef](#)]
16. Matthews, S.A. The salience of neighborhood: Some lessons from sociology. *Am. J. Prev. Med.* **2008**, *34*, 257–259. [[CrossRef](#)] [[PubMed](#)]
17. Myers, C.A.; Denstel, K.D.; Broyles, S.T. The context of context: Examining the associations between healthy and unhealthy measures of neighborhood food, physical activity, and social environments. *Prev. Med.* **2016**, *93*, 21–26. [[CrossRef](#)] [[PubMed](#)]
18. Widener, M.J.; Farber, S.; Neutens, T.; Horner, M.W. Using urban commuting data to calculate a spatiotemporal accessibility measure for food environment studies. *Heal Place* **2013**, *21*, 1–9. [[CrossRef](#)] [[PubMed](#)]
19. Sallis, J.F.; Cervero, R.B.; Ascher, W.; Henderson, K.A.; Kraft, M.K.; Kerr, J. An Ecological Approach to Creating Active Living Communities. *Annu. Rev. Public Heal.* **2006**, *27*, 297–322. [[CrossRef](#)] [[PubMed](#)]
20. Krizek, K.J. Residential Relocation and Changes in Urban Travel: Does Neighborhood-Scale Urban Form Matter? *J. Am. Plan. Assoc.* **2003**, *69*, 265–281. [[CrossRef](#)]
21. Seneviratne, P.N.; Morrall, J.F. Analysis of factors affecting the choice of route of pedestrians. *Transp. Plan. Technol.* **1985**, *10*, 147–159. [[CrossRef](#)]
22. Badland, H.M.; Keam, R.; Witten, K.; Kearns, R. Examining Public Open Spaces by Neighborhood-Level Walkability and Deprivation. *J. Phys. Act. Heal.* **2010**, *7*, 818–824. [[CrossRef](#)]
23. Frank, L.D.; Sallis, J.F.; Saelens, B.; Leary, L.; Cain, K.; Conway, T.L.; Hess, P.M. The development of a walkability index: Application to the Neighborhood Quality of Life Study. *Br. J. Sports Med.* **2009**, *44*, 924–933. [[CrossRef](#)] [[PubMed](#)]
24. Dallat, M.A.T.; Soerjomataram, I.; Hunter, R.F.; Tully, M.A.; Cairns, K.J.; Kee, F. Urban greenways have the potential to increase physical activity levels cost-effectively. *Eur. J. Public Health* **2013**, *24*, 190–195. [[CrossRef](#)] [[PubMed](#)]
25. Thompson, C.W. Activity, exercise and the planning and design of outdoor spaces. *J. Environ. Psychol.* **2013**, *34*, 79–96. [[CrossRef](#)]
26. White, M.P.; Pahl, S.; Ashbullby, K.; Herbert, S.; Depledge, M.H. Feelings of restoration from recent nature visits. *J. Environ. Psychol.* **2013**, *35*, 40–51. [[CrossRef](#)]
27. Kelly, C.; Tight, M.; Hodgson, F.; Page, M. A comparison of three methods for assessing the walkability of the pedestrian environment. *J. Transp. Geogr.* **2011**, *19*, 1500–1508. [[CrossRef](#)]
28. Doherty, A.R.; Kelly, P.; Kerr, J.; Marshall, S.J.; Oliver, M.; Badland, H.; Foster, C. Use of wearable cameras to assess population physical activity behaviours: An observational study. *Lancet* **2012**, *380*, S35. [[CrossRef](#)]
29. Kelly, P.; Doherty, A.; Berry, E.; Hodges, S.; Batterham, A.M.; Foster, C. Can we use digital life-log images to investigate active and sedentary travel behaviour? Results from a pilot study. *Int. J. Behav. Nutr. Phys. Act.* **2011**, *8*, 1–9. [[CrossRef](#)]
30. Dunton, G.F.; Almanza, E.; Jerrett, M.; Wolch, J.; Pentz, M.A. Neighborhood Park Use by Children. *Am. J. Prev. Med.* **2014**, *46*, 136–142. [[CrossRef](#)]
31. Evenson, K.R.; Wen, F.; Hillier, A.; Cohen, D.A. Assessing the Contribution of Parks to Physical Activity Using Global Positioning System and Accelerometry. *Med. Sci. Sports Exerc.* **2013**, *45*, 1981–1987. [[CrossRef](#)]
32. Kaczynski, A.T.; Potwarka, L.R.; Smale, B.J.A.; Havitz, M.E. Association of Parkland Proximity with Neighborhood and Park-based Physical Activity: Variations by Gender and Age. *Leis. Sci.* **2009**, *31*, 174–191. [[CrossRef](#)]
33. Norman, P.; Pickering, C.M.; Castley, G. What can volunteered geographic information tell us about the different ways mountain bikers, runners and walkers use urban reserves? *Landsc. Urban Plan.* **2019**, *185*, 180–190. [[CrossRef](#)]
34. Laforteza, R.; Carrus, G.; Sanesi, G.; Davies, C. Benefits and well-being perceived by people visiting green spaces in periods of heat stress. *Urban For. Urban Green.* **2009**, *8*, 97–108. [[CrossRef](#)]

35. Pretty, J.; Peacock, J.; Sellens, M.; Griffin, M. The mental and physical health outcomes of green exercise. *Int. J. Environ. Health Res.* **2005**, *15*, 319–337. [[CrossRef](#)] [[PubMed](#)]
36. Sallis, J.F.; Owen, N.; Fotheringham, M.J. Behavioral epidemiology: A systematic framework to classify phases of research on health promotion and disease prevention. *Ann. Behav. Med.* **2000**, *22*, 294–298. [[CrossRef](#)] [[PubMed](#)]
37. Yang, Y.; TANG, X.; Liu, J.; Lu, S. Urban Street Health Service Function Based on Mobile Fitness Data. *Landsc. Archit.* **2018**, *25*, 18–23. [[CrossRef](#)]
38. Moreno, R.; Ojeda, N.; Azócar, J.; Venegas, C.; Inostroza, L. Application of Ndvi for Identify Potentiality of the Urban Forest for the Design of a Green Corridors System in Intermediary Cities of Latin America: Case Study, Temuco, Chile. *Urban For. Urban Green.* **2020**, *55*, 126821. [[CrossRef](#)]
39. Kappes, H.; Katzschner, L.; Nowak, C. Urban summer heat load: Meteorological data as a proxy for metropolitan biodiversity. *Meteorol. Z.* **2012**, *21*, 525–528. [[CrossRef](#)]
40. Brown, B.B.; Werner, C.M.; Amburgey, J.W.; Szalay, C. Walkable Route Perceptions and Physical Features. *Environ. Behav.* **2007**, *39*, 34–61. [[CrossRef](#)]
41. Ricker, B. *International Encyclopedia of Human Geography*, 2nd ed.; Elsevier: Amsterdam, The Netherlands, 2020; pp. 187–195. [[CrossRef](#)]
42. Jaynes, E.T. Information Theory and Statistical Mechanics. II. *Phys. Rev.* **1957**, *108*, 171–190. [[CrossRef](#)]
43. Shipley, B. Limitations of entropy maximization in ecology: A reply to Haegeman and Loreau. *Oikos* **2009**, *118*, 152–159. [[CrossRef](#)]
44. Dixon, P.M. Ripley's K function. *Encycl. Environ.* **2001**, *3*, 1796.
45. Garrocho-Rangel, C.; Álvarez-Lobato, J.A.; Chávez, T. Calculating Intraurban Agglomeration of Economic Units with Planar and Network K-Functions: A Comparative Analysis. *Urban Geogr.* **2013**, *34*, 261–286. [[CrossRef](#)]
46. Lu, Y.; Chen, X. On the false alarm of planar K-function when analyzing urban crime distributed along streets. *Soc. Sci. Res.* **2007**, *36*, 611–632. [[CrossRef](#)]
47. Ardestani, E.G.; Mokhtari, A. Modeling the lumpy skin disease risk probability in central Zagros Mountains of Iran. *Prev. Veter Med.* **2020**, *176*, 104887. [[CrossRef](#)] [[PubMed](#)]
48. Greiner, M.; Pfeiffer, D.; Smith, R. Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Prev. Veter Med.* **2000**, *45*, 23–41. [[CrossRef](#)]
49. Li, Q.; Ren, H.; Zheng, L.; Cao, W.; Zhang, A.; Zhuang, D.; Lu, L.; Jiang, H. Ecological Niche Modeling Identifies Fine-Scale Areas at High Risk of Dengue Fever in the Pearl River Delta, China. *Int. J. Environ. Res. Public Heal* **2017**, *14*, 619. [[CrossRef](#)]
50. Karimi, K. A configurational approach to analytical urban design: 'Space syntax' methodology. *Urban Des. Int.* **2012**, *17*, 297–318. [[CrossRef](#)]
51. Liu, K.; Siu, K.W.M.; Gong, X.Y.; Gao, Y.; Lu, D. Where do networks really work? The effects of the Shenzhen greenway network on supporting physical activities. *Landsc. Urban Plan.* **2016**, *152*, 49–58. [[CrossRef](#)]
52. WHO. *Global Age-Friendly Cities: A Guide*; World Health Organization: Geneva, Switzerland, 2007. [[CrossRef](#)]
53. Elwood, S.; Goodchild, M.F.; Sui, D.Z. Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice. *Ann. Assoc. Am. Geogr.* **2012**, *102*, 571–590. [[CrossRef](#)]
54. Goodchild, M.F.; Li, L. Assuring the quality of volunteered geographic information. *Spat. Stat.* **2012**, *1*, 110–120. [[CrossRef](#)]
55. Fitzhugh, E.C.; Bassett, D.R.; Evans, M.F. Urban Trails and Physical Activity. *Am. J. Prev. Med.* **2010**, *39*, 259–262. [[CrossRef](#)] [[PubMed](#)]
56. Andersen, H.B.; Christiansen, L.B.; Klinker, C.D.; Ersbøll, A.K.; Troelsen, J.; Kerr, J.; Schipperijn, J. Increases in Use and Activity Due to Urban Renewal: Effect of a Natural Experiment. *Am. J. Prev. Med.* **2017**, *53*, e81–e87. [[CrossRef](#)] [[PubMed](#)]