

Article

Understanding the Tourists' Spatio-Temporal Behavior Using Open GPS Trajectory Data: A Case Study of Yuanmingyuan Park (Beijing, China)

Qian Yao ¹ , Yong Shi ^{2,3,*}, Hai Li ², Jiahong Wen ⁴, Jianchao Xi ³ and Qingwei Wang ¹

¹ Department of Tourism Management and Planning, School of Tourism Management, Zhengzhou University, Zhengzhou 450001, China; yaoqian@gs.zzu.edu.cn (Q.Y.); wangqingwei@zzu.edu.cn (Q.W.)

² Department of Economics and Management, Shanghai University of Sport, Shanghai 200438, China; lh@sus.edu.cn

³ Institute of Geographic Sciences and Natural Research, Chinese Academy of Sciences, Beijing 100101, China; xijc@igsnr.ac.cn

⁴ Department of Geography, School of Environmental and Geographical Sciences, Shanghai Normal University, Shanghai 200234, China; jhwen@shnu.edu.cn

* Correspondence: shiyong@sus.edu.cn

Abstract: The visit paths, dwell time, and taking pictures are all variables of great significance to our understanding of tourists' spatio-temporal behavior. Does having a large number of visitors mean that tourists are interested in a tourist location? What is the relationship between the dwell time and taking pictures? Are there differences in tourist behavior in different seasons? These issues are of great significance to tourism research but they have not been rigorously analyzed yet. This paper aims to understand the relationship between tourists' visit path, dwell time, and taking pictures, and test whether there are differences in tourist behavior in different seasons. We used open global positioning systems (GPS) trajectory data at Yuanmingyuan Park from January 2014 to August 2020. Using Python and ArcGIS tools, we found hot spots of tourist passing, hot spots of tourist gathering, high average dwell time areas, and tourist interest areas. It is further found that: (1) passenger flow strongly explains dwell time, while the correlation between passenger flow and average dwell time is weak. (2) There was a close relationship between tourists' stay and photo-taking behavior, which provided a theoretical basis for defining tourist photo behavior as tourists' stay behavior. (3) Seasons did not significantly affect tourist behavior in Yuanmingyuan Park. This study presents a grid-based open GPS trajectory data processing framework that clarified the potential of an open GPS trajectory in tourist behavior research. Furthermore, our study explored the relationship between essential indicators and found that there is a strong consistency in tourist behavior across seasons.

Keywords: tourist behavior; GPS; big data; spatio-temporal behavior; theme park



Citation: Yao, Q.; Shi, Y.; Li, H.; Wen, J.; Xi, J.; Wang, Q. Understanding the Tourists' Spatio-Temporal Behavior Using Open GPS Trajectory Data: A Case Study of Yuanmingyuan Park (Beijing, China). *Sustainability* **2021**, *13*, 94. <https://dx.doi.org/10.3390/su13010094>

Received: 25 November 2020

Accepted: 18 December 2020

Published: 24 December 2020

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



Copyright: © 2020 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

In the context of human-centered development, studies of tourists' spatio-temporal behavior have been increasingly scrutinized by tourism, urban planning, geography, and other scholars. Generally, tourist behavior research can be divided into two categories: inter-destination and intra-destination tourist behavior [1]. Inter-destination tourist behavioral research mainly focuses on the movement behavior between destinations, such as how tourists travel from one city to another. Intra-destination tourist behavioral research focuses on the movement trajectory of tourists within a destination. For example, this research may explore the travel routes of tourists between various attractions in a city.

Recently, however, researchers have focused on the micro-scale of tourist behavior, namely intra-attraction tourist behavior [2–5]. Intra-attraction tourist behavioral research mainly focuses on smaller places such as national parks, theme parks, historical sites, etc. An understanding of tourist movement and behavior can assist targeted marketing [6,7],

help to manage impacts associated with overuse or crowding [8,9], guide adjustments to transport systems [1], and improve the visitor experience [7,10].

Global positioning systems (GPS) data have been widely used in tourist behavior research [2,11–15]. Scholars often collect trajectories through handheld GPS. Because the cost of traditional GPS data acquisition is still high, the in-depth study of micro-scale tourist behavior is greatly restricted. However, the emergence of big data provides an opportunity to overcome this problem.

With the rapid development of computer science and Internet techniques, many open and accessible forms of User Generation Content (UGC) have been generated, including location photos, GPS tracks, travel logs, web reviews, etc. [16–20]. In the study of tourist behavior, big data has become the mainstream data source. Many scholars have used check-in records, geo-tagged photographs, and travel logs to explore tourist flows between and within destinations [21–24]. Yang argued that the massive scale of open data could compensate for the limitation of sample size issues that survey data users face, thereby providing a new way to understand tourist behavior [25]. However, due to the particularity of open GPS trajectory data, its value in intra-attraction tourist behavior research is still not fully understood.

On the one hand, tourist behavior always takes place in time and space. Changes in visitor travel patterns are a function of the locations visited and of time. Therefore, it is crucial to understand how space and time affect visitor travel patterns [26,27]. On the other hand, with the popularity of mobile devices, tourists can record the places they are interested in by taking photos, which also provides a new way for researchers to identify tourists' behavior patterns [28,29]. The length of time a visitor spends at a single location (dwell time) is a key factor in their enjoyment of exhibits, as it takes time to absorb information and relate it to what is being observed [30]. Most tourists take photographs, creating their visit records as final proof of their experiences [31]. Nevertheless, does a large number of visitors mean that tourists are interested in a tourist location? What is the relationship between the length of time a tourist spends at a single location (the dwell time) and the behavior of taking pictures? Are there differences in tourist behavior in different seasons? This research will try to answer the above questions.

In an attempt to resolve these issues and verify the feasibility of open GPS trajectory data in intra-attraction research, this paper proposes a grid-based open GPS trajectory data processing framework. On this basis, this paper used Yuanmingyuan Park, China, as a case study to identify intra-attraction tourist behavior characteristics. The study aimed to determine the relationship between tourists' visit path, dwell time, and behavior of taking pictures, as well as to test the differences in tourist behavior in different seasons in order to broaden the horizons of existing intra-attraction tourist behavior research.

Given this aim, this paper opens with a review of extant literature related to intra-destination tourism flow and intra-attraction tourist behavior research. This is followed by an introduction of our research framework, study area, collected data, methodology, and a presentation of the results of the analysis. The paper concludes with a discussion and conclusion.

2. Literature Review

2.1. Intra-Destination Tourism Flow

There are many commercial or non-commercial nodes in the destination, which play a large or small role in the destination system and are connected by tourism flows [32,33]. Understanding the number, direction, and duration of tourism flows within a destination is of positive significance for improving destination management.

When the city is a tourist destination, researchers have mainly attended to tourist movements and flows [34] and the network patterns and characteristics created by natural tourism flows in a destination [32,35]. For example, Lew and McKercher divided the tourism flow patterns of Hong Kong and identified tourism flows were influenced by six interrelated factors of territoriality [34]. Liu, Huang, and Fu applied network analysis to

research on the tourism flow of destinations in Xinjiang and found that tourist attractions of the same level in the tourist destination system mainly compete with each other [36]. Mou and Yuan et al. identified the spatio-temporal changes of city inbound tourism flow in Shanghai, extracted the Area of Interest (AOI), and found that the inbound tourism flow network of Shanghai has small-world characteristics, and also indicated that the distribution of its AOI (nodes) and tourist routes (edges) has general power-law features, which have been influenced by the World Expo [37].

When the tourist attraction is a tourist destination, the system is more obvious. In the planning of tourist attractions, functional zoning is often considered. Different areas serve as nodes for serving, transit, recreation, or attracting tourists. Roads connect different functional areas within the attraction and affect the tourism flow. Relevant research is devoted to revealing the causes, patterns, and laws of tourism flow within attractions. For example, East and Osborne et al. found that time affects the tourism flow within theme parks [3]. At different times of the day, the tourist gathering area in a theme park is different. Li and Xie et al. used the social network analysis method to calculate the social network centrality index of 54 areas in Gulangyu and found areas with development potential in Gulangyu [2]. Peterson and Perry et al. identified which areas are the most likely to gather in attractions and explored the differences in tourism flows in different seasons [26]. Li and Yang et al. reveals the underlying behavioral mechanism of choice within tourist destinations, confirming that proximity, history, and attractiveness significantly influence tourism flow [38].

Although many scholars have studied intra-destination tourism flow, the literature regarding tourist destinations as cities or attractions often ignores the relationship between the direction and duration of tourism flows. The relationship between the two is significant. Revealing the relationship between tourism flow direction and duration is one of the interests of this research.

2.2. Intra-Attraction Tourist Behavior

Most studies in tourist behavior have only focused on inter-destination behavior, whereas discussion involving intra-destination tourist behavior remains relatively scarce [4]. Intra-destination space refers to an enclosed space with defined boundaries, in which tourist behavior is much more controllable [5]. To analyze tourist behavior within a destination, some scholars have used traditional methods such as travel diaries and observations.

Recently, various devices such as mobile phones and handheld GPS devices have been used to collect high-resolution data, laying the foundation for micro-scale research. A great deal of previous research into intra-attraction tourist behavior has focused on feasibility analysis, tourist behavior characteristics, tourist behavior classification, and tourist behavior prediction, among other topics. For example, Xiao-Ting developed a combination approach to clarify such behavioral patterns quantitatively and qualitatively using the concept of the space-time path of time geography to explain the patterns in terms of temporal behavior factors, spatial behavior factors, activity choice factors, and path characteristics [5]. East combined tracking data and survey data and found that most tourists visiting the zoo follow a similar route, revealing a strong dependence on the main road [5]. Li taking Gulangyu as an example, combined GPS data and survey data and developed a multinomial logit (MNL) model for identifying factors that affect tourists' destination choices [2]. Zheng combined dynamic time warping and the earth mover's distance to accurately measure the similarity in tourist trajectories [14]. Huang given the marked differences in demographic and emotional characteristics, identified three spatial-temporal behavior clusters via density center clustering, consisting of four factors: path length, tour time, coverage area, and oval circumference [4].

In general, existing research still has some shortcomings. First, in terms of data sources, most of the GPS data used in research are collected by the researchers. The high upfront cost of GPS data limits the popularization of their use in research. Second, with regards to the research content, most studies have only focused on the classification of tourist

behavior. To date, few studies have investigated the association between some indicators in the course of tourist behavior, such as dwell time, number of tourists, and a number of geo-tagged photographs. As a result, the relationship between these indicators is still unclear. Although some studies have proved the uniformity of tourist behavior, it is still unclear whether there are differences in tourist behavior across seasons due to the short duration of existing research. Therefore, this study will propose a processing framework for open GPS trajectory data, explore the relationship between essential indicators, and identify seasonal differences in tourist behavior to resolve the shortcomings mentioned above.

3. Data and Methods

3.1. Study Area

Yuanmingyuan Park is located in Haidian District, the western suburb of Beijing, and is adjacent to the Summer Palace. Built in the 46th year of Emperor Kangxi's reign (1707), Yuanmingyuan includes three gardens: Yuanmingyuan, Changchunyuan, and Qichunyuan. With a floor area of 350 hectares and a building area of almost 200,000 square meters, Yuanmingyuan was a vast royal palatial garden established and operated for over 150 years by emperors during the Qing Dynasty. Yuanmingyuan is also a waterscape garden, and the water surface occupies more than half of the entire area of the park. The Qing emperor came to Yuanmingyuan, which was also called the "Summer Palace," every summer to escape the heat, listen to politics, and handle military affairs. In November 1976, the Administrative Office of Yuanmingyuan was established. In January 1988, Yuanmingyuan Park was announced as a key cultural relic protection site at the national level. On 29 June 1988, Yuanmingyuan Park was officially opened to the public. The ticket offices of Yuanmingyuan Park are mainly located at the gate of Qichunyuan Palace, the East Gate of Changchunyuan, Zaoyuan Gate, and Zhengjue Temple Gate. The entrance ticket is 10 yuan per person. It is an important place for both Beijing residents and for tourism, leisure, and science education. Yuanmingyuan Park is a suitable year-round destination, and the scenic management office recommends various tourist routes for the different seasons (spring, summer, autumn, and winter). Please see Figure 1.



Figure 1. Map of Yuanmingyuan Park. (source: <https://map.baidu.com/>).

3.2. Data

Due to the popularization of mobile internet and the use of GPS, tourists produce many digital footprints during their travel. These publicly available footprints include location photos, GPS tracks, travel logs, online reviews, etc. These text data and location photo information have been abundantly mined but results based on open GPS trajectory data are rare. Although many scholars have devoted themselves to tourist GPS data mining, they often collect trajectory data through handheld GPS, but collect tourist population and emotional data through questionnaires.

There are many advantages to using open GPS trajectory data. First, traditional research is limited by cost and the process of collecting data is often concentrated within two weeks [4], but the timeframe is much wider when collecting open GPS data. Second, open GPS trajectory data access does not require researchers to purchase equipment and does not require traveling to the case site for data collection. The cost and convenience of research are thus greatly improved. Third, tourists often update their trajectories as they upload location photos [3,39]. These photos reflect the interests of tourists and can broaden the scope of research. However, open GPS trajectory data still has many disadvantages when compared to traditional GPS data. First, handheld GPS data is highly accurate, and the time distribution between points is relatively even [40]. However, open GPS trajectory data is subject to interference from mobile phone signals and other factors. Therefore, its distribution is more scattered. Second, the starting and ending points of open GPS trajectory data records are not controlled by the researcher. General tourist spatio-temporal behavioral research indicators such as average speed, travel length, and total travel time are difficult to obtain directly. Third, as in the case of general open data, open GPS trajectory data cannot obtain the demographic characteristics of the uploader, which limits further analysis. Even with these shortcomings, the potential of open GPS trajectory data in tourist behavior research is still huge.

We selected the GPS tracking data of tourists who used the outdoor route websites 2bulu.com and foouooot.com, which are widely used in China. We used Python to write scripts to collect all the trajectory and photo data uploaded by tourists in the study area on the two platforms and imported the trajectory data into ArcGIS after coordinate conversion. The original GPS trajectory was obtained by displaying the XY coordinates and Points to Line. The first step was to remove low-quality trajectory, including strokes that were too short or repeated as well as excessively abnormal points. The second step was to use Pandas (Pandas is a powerful toolset for analyzing structured data based on Numpy, which is used for data mining and data analysis; Pandas also provides data cleaning functions.) to calculate the time spent between two points on the same track. The calculation method used the absolute value of the difference between two adjacent points, and the distance between two points was calculated as follows:

$$d = \sum_{i=2}^n \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}. \quad (1)$$

The third step was to find points that were more than 30 s apart from another point. If the distance between the two points exceeded 30 m, it was considered an abnormality and deleted. In step four, we located the outliers and eliminated the influence of extreme values. For example, we deleted the points that were far apart but took too little time to reach. In step five, we imported the processed data into ArcGIS and removed the points outside the study area.

A total of 1219 trajectories data were available from January 2014 to July 2020, of which 906 trajectories were used, including 595,344 valid GPS points, representing a validity percentage of 74.3%. Example data are shown in Table 1. Each point's attribute information includes the route number, the sequence number of the point in the route, longitude, latitude, time spent, the distance between two points, and timestamp information. Additionally, the geographical location data of check-in points for visitors from the same

website was adapted to reflect the spatial distribution of tourist preferences in Yuanmingyuan Park.

Table 1. Data attribute information.

Number	n	Longitude	Latitude	Dwell time	Distance/m	Time	Timestamp
1	12	116.3064	39.99998	9	11	31 January 2018 13:51:18	1517377878
1	13	116.3065	40.00004	8	11	31 January 2018 13:51:26	1517377886
1	14	116.3064	39.99998	34	11	31 January 2018 13:52:00	1517377920
1	15	116.3063	39.99996	6	11	31 January 2018 13:52:06	1517377926
1	16	116.3062	39.99999	9	11	31 January 2018 13:52:15	1517377935
1	17	116.3061	40.00006	8	10	31 January 2018 13:52:23	1517377943
1	18	116.306	40.0001	8	11	31 January 2018 13:52:31	1517377951
1	19	116.3058	40.0001	7	12	31 January 2018 13:52:38	1517377958
1	20	116.3057	40.00011	8	11	31 January 2018 13:52:46	1517377966
1	21	116.3056	40.0001	8	12	31 January 2018 13:52:54	1517377974

Source: <http://www.foooot.com/> and <https://www.2bulu.com/>.

3.3. Data Analysis

Data analysis included four steps. The first step was to grid the spatial data. This method has been proven by many scholars to be an effective means of processing spatial data [2,10,40]. By dividing the research area into grids of a specific size, visualization and statistics can be conveniently performed. As mentioned earlier, the accuracy of open GPS data is affected by the user's mobile device, and therefore, there are often huge inaccuracies. By converting spatial data to grids, the spatially uneven data was processed according to a specific geometric grid. Statistics should be combined to more easily analyze the research data's coupling characteristics and geographic space [41]. Because grid statistics rely on GPS trajectory points instead of trajectory lines, they can quickly summarize information such as the number of tourists, dwell time, and the number of geo-tagged photographs. GPS trajectory technology can also remove the influence that a single tourist's extreme behaviors may have on the overall characteristics, and therefore it is highly reliable. More importantly, we can obtain many statistical units to facilitate correlation testing and modeling features of different attributes by gridding the spatial data. An example is provided in Figures 2 and 3. Considering the study area's environmental characteristics and the accuracy of GPS trajectories, we assumed that it was appropriate to set the grid at 50×50 m. Tourist trajectories and location photos were collected on grid cells to analyze tourists' spatial distribution characteristics in Yuanmingyuan Park, the differences in tourists' spatial distribution in different seasons, and the relationship between tourist interest districts and dwell time. (Figure 2)

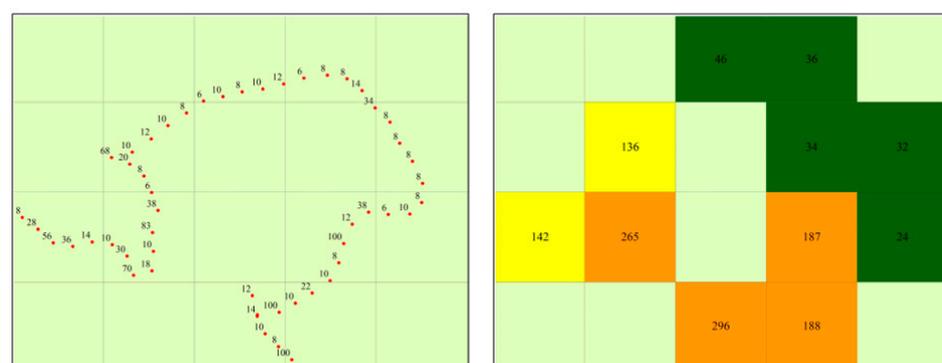


Figure 2. An example of dwell time calculation (source: author self-painted).



Figure 3. An example of global positioning systems (GPS) points and grid cells. (source: author self-painted).

The second step was to use the spatial join function of ArcGIS to join the location and attribute information of the trajectory data to the grid and add up the time of all trajectory points in the grids to obtain the total dwell time in grids. We calculated the number of tourists passing by and visitors taking pictures in the grid, performed the merge rule-join operation on the track number field, and removed duplicate values to produce the number of tourists and geo-tagged photographs. Since the researcher does not control the starting and ending point of the open GPS trajectory data, it is unreasonable to use the total tour time and path length without cropping to reflect tourists' actual behavior characteristics. Using the method proposed, the calculated tourist behavior indicators include path length, tour time, and average speed, which can better reflect tourist behavior characteristics. The open GPS trajectory data process framework proposed in this paper overcomes the difficulties associated with using the trajectory as line data and the accuracy of kernel density estimation directly on the trajectory point data is not high. This framework will significantly improve the researchers' control over the experiment (Figure 3).

The third step was to visualize the data based on the indicators obtained above. To improve the visualization, we used the Feature to Point tool in ArcGIS software to extract each grid's center point to make it more convenient to use Kernel density. Kernel density analysis was used to highlight the spatial distribution characteristics in time spent, visit path and region of interest in particular areas of the park for all visitors. The kernel density maps were produced using a default output raster cell size and a search radius of 100 m. To reduce the impact of small samples on the results, we selected a grid with a total number of people greater than four as the statistical cell. By completing this step, we found hot spots for tourists passing, hot spots for tourist gathering, high average dwell time areas, and tourist interest areas.

We can intuitively discover the geographical features of tourists' behavior in Yuanmingyuan Park through the visualization method. To further understand and explain the tourists' spatio-temporal behavior characteristics, the fourth step involved correlation analysis to determine the relationship between dwell time, number of tourists, and number of geo-tagged photographs. After completing this step, we used one-way analyses of variance (ANOVAs) and spatial data visualization methods to test the differences in tourist behavior in different seasons.

4. Analysis and Findings

4.1. Spatio-Temporal Behavior Characteristics and Interrelationships

4.1.1. Tourist Visit Path

The kernel density analysis, which is based on the number of tourists passing by, reveals the areas which most tourists pass through (Figure 4a). Most tourists enter the park through the gate of Qichunyuan Palace. As Fuhai lies at the center of the Yuanmingyuan Park, the northern and southern areas are the main roads for tourists. Fewer tourists enter Changchun Garden, as most choose to pass through the north of Changchun Garden. Two agglomeration centers are also located at the intersection of major roads, including

Dequange and Xieqiqu at the northern and southern ends of the Yuanmingyuan and Changchunyuuan.

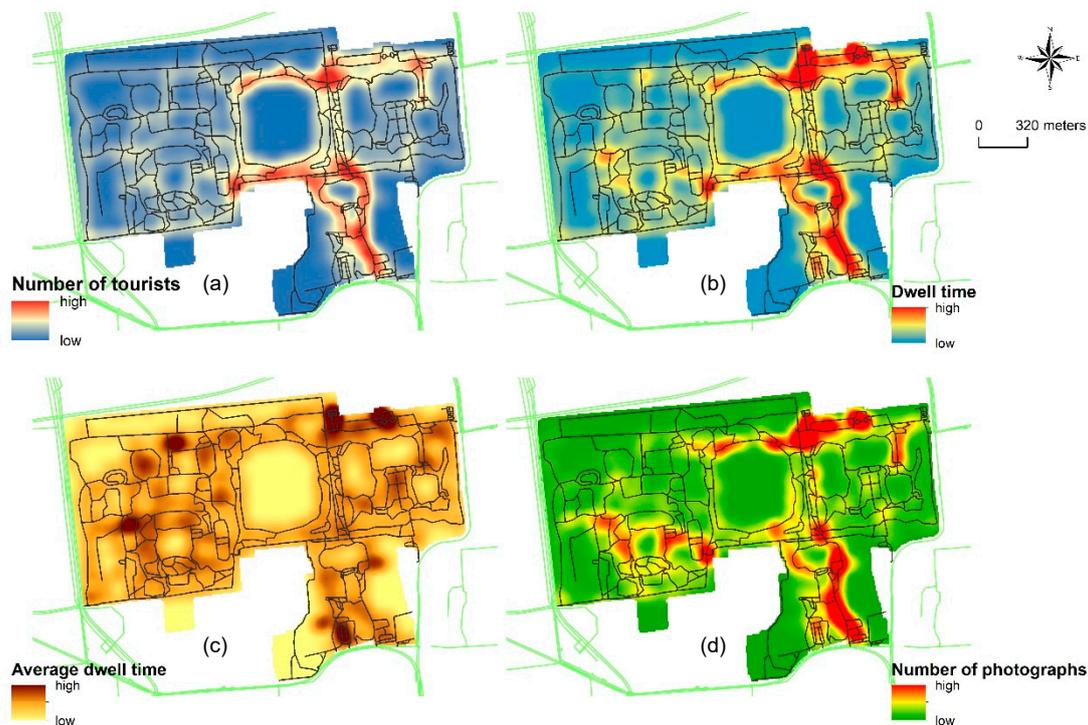


Figure 4. The kernel density map of tourist behavior characteristics. (a) tourist number, (b) dwell time, (c) average dwell time, and (d) number of geo-tagged photographs (source: author self-painted).

4.1.2. Tourist Dwell Time

The kernel density analysis, weighted by tourists' dwell time, reveals the areas where tourists stay the longest (Figure 4b). In terms of aggregation characteristics, the analysis shows a trend similar to that of tourists' visit paths. The agglomeration centers where tourists travel are all high-value areas of tourists' dwell time, but the latter's degree of agglomeration is weaker than that of the former. The analysis results show that there are several areas where tourists spend a long time, even where there are not many tourists. They mainly include the Zhengjue Temple in Qichun Garden, the Huanghuazhen, Dashuifa, and Yuanmingyuan panoramic sand table scenic area in Changchun Garden. The correlation coefficient between the number of tourists and the dwell time is calculated ($R = 0.87$, $p < 0.01$), which shows that the passenger flow has a strong ability to explain the hot spot where tourists spend more time.

4.1.3. Tourist Average Dwell Time

To eliminate the tourist numbers' influence on dwell time, the average dwell time was calculated to find the places where tourists spent longer times. The kernel density analysis, which uses the average dwell time of tourists as the weight, reveals the areas where tourists spend longer times, on average (Figure 4c). We found that the main scenic roads and road intersections had the highest number of tourists, but the average dwell time of tourists was short. The correlation coefficient between the number of tourists and the average dwell time ($R = 0.13$, $p < 0.01$) demonstrates that there was a weak correlation between the tourist number and the average dwell time.

4.1.4. Number of Geo-Tagged Photographs

The number of geo-tagged photographs as the weight for kernel density analysis reveals the places tourists were most interested in (Figure 4d). From the gathering center's spatial location, the analysis reveals that the gathering area highly coincided with the tourist dwell time. Both the correlation coefficient between the number of tourists and the number of geo-tagged photographs ($R = 0.73, p < 0.01$) and the correlation coefficient between the dwell time and the number of geo-tagged photographs ($R = 0.88, p < 0.01$) were calculated. The coefficients demonstrate that compared with the number of tourists, there was a stronger correlation between dwell time and the number of geo-tagged photographs in the spatial distribution.

4.2. Tourists Behave Differently in Different Seasons

According to the travel time, tourists' trajectory data was divided into four categories: spring, summer, autumn, and winter. One-way analysis of variance (ANOVAs) was used to investigate whether there were significant differences in tourist behavior in different seasons. The selected variables are path length, tour time, and average speed. Post-hoc tests were conducted when ANOVA results were significant at the $p < 0.05$ level. If the assumption of homogeneity of variance was valid, a Tukey test was performed; otherwise, Tamhane's T2 was used [4].

There was no significant difference in the tour time and average speed in different seasons (Table 2). There was a significant difference in the path length in spring and autumn ($p < 0.05$), but there was no significant difference in the path length in other seasons. This demonstrates that the path length, tour time, and average speed in the Yuanmingyuan Park show substantial homogeneity in different seasons. To further understand the seasonal differences of tourist behaviors within the attraction, we used tourists' dwell time in different seasons as weights to draw a kernel density map (Figure 5).

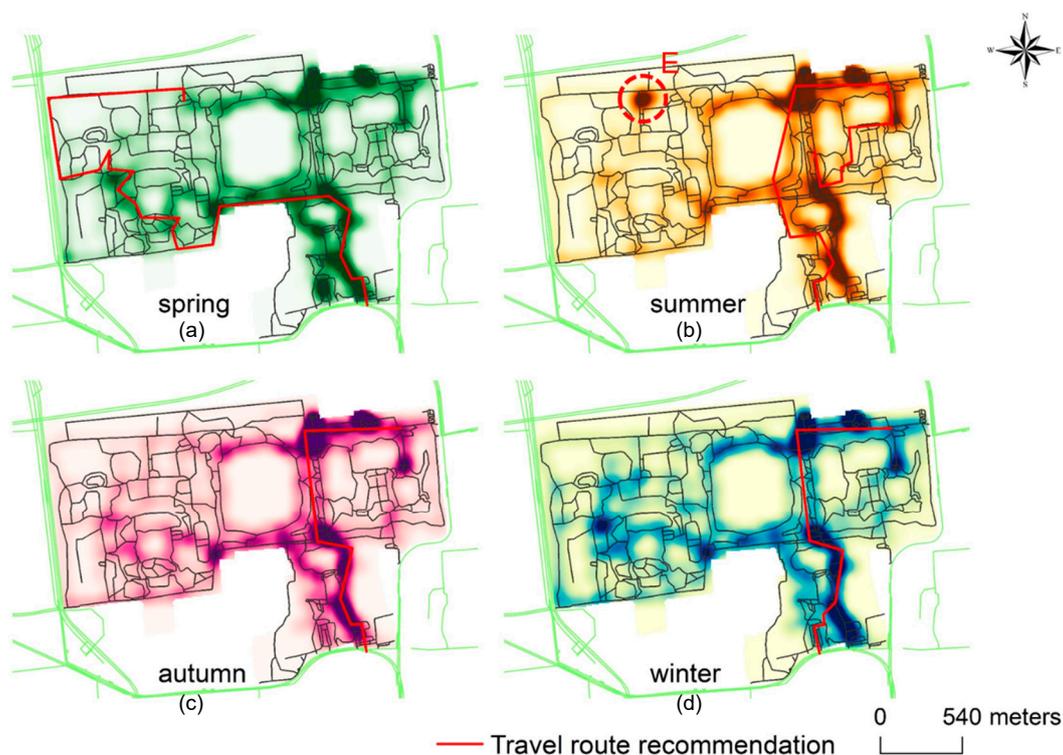


Figure 5. The kernel density map of visitor's dwell time in (a) spring, (b) summer, (c) autumn, and (d) winter (source: author self-painted).

Table 2. Analyses of variance (ANOVA) results for seasonal spatial-temporal patterns. (source: author self-painted).

	Spring <i>n</i> = 288	Summer <i>n</i> = 250	Autumn <i>n</i> = 251	Winter <i>n</i> = 117	Test Score
length	7056	6412	6286	6552	F = 3.399 <i>p</i> = 0.017
Time	7822.233	7653.928	7512.072	7507.154	F = 0.326 <i>p</i> = 0.807
Speed	59.08	57.00	56.05	56.75	F = 0.708 <i>p</i> = 0.548

Figure 5 shows that, although the Administrative Office of Yuanmingyuan provides different recommended sighting lines for different seasons, tourists always choose the same paths, which again proves the homogeneity of tourists' behavior. The number of tourists in winter declined compared to the other seasons. However, further analysis shows that the season does not significantly affect tourists' behavior, regardless of whether it is the average travel time, average speed, travel length, or tourist movement behavior. Although tourist behavior appears to be homogeneous, this does not necessarily mean that there is no difference in tourists' behavior choices in different seasons. As shown in Figure 5b, there is an apparent high-value area for tourists to stay in the lotus area (zone E) during the summer. The figure shows that although zone E is not highly accessible, the behavior of tourists is affected by attractions in summer.

5. Conclusions and Future Research

5.1. Conclusions

This study set out to determine if open GPS trajectory data could be used to accurately understand intra-attraction tourist behavior. This study visually demonstrated the tourists' spatio-temporal behavior, analyzed the relationship between several essential indicators in tourists' behavior and investigated whether different seasons significantly affected tourists' behavior. The data was obtained from two websites, fooooooot.com and 2bulu.com, with 906 trajectory data obtained in the Yuanmingyuan Park. The paper demonstrated that open GPS trajectory data are reliable in tourist behavior research. The research results show that open GPS trajectory data are significantly better than traditional self-collected GPS data in both quantity and efficiency. Nevertheless, the accuracy of the data is relatively low. We have identified tourist interest areas and found that the indicator of dwell time could explain the relationship between tourist stay behavior and tourist photo behavior. We used ANOVA to compare the differences in tourist behavior (path length, tour time, and average speed) in different seasons. We also drew a kernel density map of tourists' dwell time in different seasons and found that tourist behavior is consistent regardless of the season.

Our results significantly contribute to current theory and provide a set of guidelines for managers. First, the study demonstrated that although UGC is widely used in tourism research and destination marketing and management, the role of open GPS trajectory data in tourist behavior research is underestimated. We then clarified the potential of open GPS trajectory data in tourist behavior research and compared their accuracy and processing to traditional self-collected trajectory data. The study proposed a framework for processing analysis that will provide a new perspective for tourist behavior research. Second, this paper found that the *R* between the average dwell time and the number of tourists was only 0.13, suggesting that the area of real interest to tourists may not be the areas with the highest number of trips made by tourists. This is different from people's intuitive experience, but it provides important guidance for the tourist attractions' functional planning. Tourist attractions often build shopping stores or other public service facilities at the tourist passages that tourists must pass. Because of the large flow of people in these areas, the potential benefits of layout facilities may be higher. However, this is not necessarily the case. Our research proves that the correlation between the number

of tourists and the average stay time of tourists is weak, indicating that the layout of service facilities in areas with a high number of tourists does not significantly increase the average stay time of tourists, and the revenue target of the tourist attraction may not be well achieved. Furthermore, the COVID 19 pandemic has changed the way that people travel. More and more people hope to reduce their contact with others during travel and to feel nature more. However, suppose facilities that require tourists to stay are built at the tourist passages that tourists must pass. In that case, it may slow down the flow of people and cause many tourists to gather in a small space, which intensifies the discomfort of tourists and increases the risk of virus transmission.

On the other hand, the study found that the R of dwell time and number of geo-tagged photographs was as high as 0.88. This indicated that there was a close relationship between tourists' stay and photo-taking behavior. It also implied that if more tourists take photos in an area, it is often an easy place for tourists to gather. In previous research on tourist behavior, due to the difficulty of obtaining data on tourists' staying time, scholars often use the number of photos to reverse the concentration of tourists. They believe that the more photos are taken of a particular area, the stronger the concentration of tourists. Although this concept has obtained support in the past, there was very little evidence to prove the quantitative connection between the number of photos taken and tourists' length of stay. This article proves the strong correlation between the two. It also provides a theoretical basis for the research of tourist behavior. Third, the study found that seasons do not significantly affect tourist behavior in Yuanmingyuan Park. Although the reasons are still worth exploring, they raise a question about tourism research and destination management. Does the seasonality of tourism correspond with tourist behavior at the level of intra-attraction? Is it necessary for tourism attractions to set up different recommended routes in different seasons?

In the context of big data, this research has the following implications for destination management: (1) The conclusions can enhance destinations' management capabilities, including tourism marketing, landscape design, potential crowding prevention, etc. (2) The proposed method provides decision-makers with powerful tools to optimize resource allocation and service facility layout. (3) The conclusions can enrich the understanding of tourist behaviors during different seasons to improve attraction and enhance the tourist experience.

5.2. Future Research

This paper's findings provide the following insights for future research: On the one hand, the high cost of collecting traditional trajectory data makes it difficult to use when conducting long-term investigations. It is still unclear whether there are differences in tourist behavior in different seasons. Although this article finds that tourist behavior is very consistent across seasons, which is consistent with the results of related studies [42,43], there are inconsistent examples [44]. Urban parks were the object of study of this article. It is still unclear whether the same conclusion will be obtained for a natural scenic location. Future research should pay more attention to the differences in the tourist behavior in multi-type and multi-scale destinations over the course of either a year or a month. Open trajectory data are often accompanied by tourists' picture-taking behavior, making the refined research of destination images possible. In the past, the research of destination images was often based on cities or attractions. Future research can try to study the attraction's internal image and use deep learning and image recognition technology to analyze and identify the destination image's characteristics in different regions. Nevertheless, it is difficult for open GPS trajectory data to collect tourists' demographic and emotional characteristics, limiting the identification of causal relationships and further analysis of tourist classification. In future studies, the combination of multi-source data should be reasonably selected or configured for a variety of research purposes. For example, the approach devised for this study can be extended to understand if the utilization of short-term facilities set up for festivals has an impact on tourist behavior. Future research can combine the long-term nature of open GPS trajectory data with the short-term concen-

tration of self-acquired GPS data to identify tourist behavior characteristics during such specific activities.

Author Contributions: Q.Y. analyzed the data and wrote the paper; Y.S. conceived and designed the research and edited the paper; H.L., J.X., J.W., and Q.W. reviewed and edited the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Grant number: 41,601,566, 41,601,568, 41,671,151, 41,771,540, and 71,603,168), National Social Science Foundation of China (Grant number: 18BJY164), National Key Research and Development Plan (Grant number: YS2017YFC1503001, 2018YFC0406505), The Education Department of Henan Province (Grant number: 2018-YYZD-16), The Humanities and Social Sciences Plan Project of Henan Provincial Education Department (Grant number: 2020-ZZJH-457), Key Laboratory of Regional Sustainable Development Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences (Grant number: KF2018-06) and the Humanities and Social Science Project of Education Ministry (Grant number: 14YJCZH128).

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

References

1. Edwards, D.; Griffin, T. Understanding tourists' spatial behaviour: GPS tracking as an aid to sustainable destination management. *J. Sustain. Tour.* **2013**, *21*, 580–595. [\[CrossRef\]](#)
2. Li, Y.; Xie, J.; Gao, X.; Law, A. A Method of selecting potential development regions based on GPS and social network models—from the perspective of tourist behavior. *Asia Pac. J. Tour. Res.* **2018**, 1–17. [\[CrossRef\]](#)
3. East, D.; Osborne, P.; Kemp, S.; Woodfine, T. Combining GPS & survey data improves understanding of visitor behaviour. *Tour. Manag.* **2017**, *61*, 307–320. [\[CrossRef\]](#)
4. Huang, X.; Li, M.; Zhang, J.; Zhang, L.; Zhang, H.; Yan, S. Tourists' spatial-temporal behavior patterns in theme parks: A case study of Ocean Park Hong Kong. *J. Destin. Mark. Manag.* **2020**, *15*, 100411. [\[CrossRef\]](#)
5. Xiao-Ting, H.; Bi-Hu, W. Intra-attraction tourist spatial-temporal behaviour patterns. *Tour. Geogr.* **2012**, *14*, 625–645. [\[CrossRef\]](#)
6. Hungenberg, E.; Gray, D.; Gould, J.; Stotlar, D. An examination of motives underlying active sport tourist behavior: A market segmentation approach. *J. Sport Tour.* **2016**, *20*, 81–101. [\[CrossRef\]](#)
7. Johns, N.; Gyimóthy, S. Market segmentation and the prediction of tourist behavior: The case of Bornholm, Denmark. *J. Travel Res.* **2002**, *40*, 316–327. [\[CrossRef\]](#)
8. Shi, B.; Zhao, J.; Chen, P. Exploring urban tourism crowding in Shanghai via crowdsourcing geospatial data. *Curr. Issues Tour.* **2017**, *20*, 1186–1209. [\[CrossRef\]](#)
9. Shoal, N.; Isaacson, M. Tracking tourists in the digital age. *Ann. Tour. Res.* **2007**, *34*, 141–159. [\[CrossRef\]](#)
10. De Cantis, S.; Ferrante, M.; Kahani, A.; Shoal, N. Cruise passengers' behavior at the destination: Investigation using GPS technology. *Tour. Manag.* **2016**, *52*, 133–150. [\[CrossRef\]](#)
11. Galí, N.; Aulet, S. Tourists' space-time behavior in heritage places: Comparing guided and nonguided visitors. *Int. J. Tour. Res.* **2019**, *21*, 388–399. [\[CrossRef\]](#)
12. Grinberger, A.Y.; Shoal, N.; McKercher, B. Typologies of tourists' time-space consumption: A new approach using GPS data and GIS tools. *Tour. Geogr.* **2014**, *16*, 105–123. [\[CrossRef\]](#)
13. McKercher, B.; Shoal, N.; Ng, E.; Birenboim, A. First and Repeat Visitor Behaviour: GPS Tracking and GIS Analysis in Hong Kong. *Tour. Geogr.* **2012**, *14*, 147–161. [\[CrossRef\]](#)
14. Zheng, W.; Zhou, R.; Zhang, Z.; Zhong, Y.; Wang, S.; Wei, Z.; Ji, H. Understanding the tourist mobility using GPS: How similar are the tourists? *Tour. Manag.* **2019**, *71*, 54–66. [\[CrossRef\]](#)
15. Sugimoto, K.; Ota, K.; Suzuki, S. Visitor Mobility and Spatial Structure in a Local Urban Tourism Destination: GPS Tracking and Network analysis. *Sustainability* **2019**, *11*, 919. [\[CrossRef\]](#)
16. Li, J.; Xu, L.; Tang, L.; Wang, S.; Li, L. Big data in tourism research: A literature review. *Tour. Manag.* **2018**, *68*, 301–323. [\[CrossRef\]](#)
17. Pantano, E.; Priporas, C.; Stylos, N. 'You will like it!' using open data to predict tourists' response to a tourist attraction. *Tour. Manag.* **2017**, *60*, 430–438. [\[CrossRef\]](#)
18. Lyu, F.; Zhang, L. Using multi-source big data to understand the factors affecting urban park use in Wuhan. *Urban Forest. Urban Green.* **2019**, *43*, 126367. [\[CrossRef\]](#)
19. Liu, Y.; Teichert, T.; Rossi, M.; Li, H.; Hu, F. Big data for big insights: Investigating language-specific drivers of hotel satisfaction with 412,784 user-generated reviews. *Tour. Manag.* **2017**, *59*, 554–563. [\[CrossRef\]](#)
20. Khan, M.S.; Woo, M.; Nam, K.; Chathoth, P.K. Smart city and smart tourism: A case of Dubai. *Sustainability* **2017**, *9*, 2279. [\[CrossRef\]](#)
21. Zhang, K.; Chen, Y.; Li, C. Discovering the tourists' behaviors and perceptions in a tourism destination by analyzing photos' visual content with a computer deep learning model: The case of Beijing. *Tour. Manag.* **2019**, *75*, 595–608. [\[CrossRef\]](#)

22. Luo, J.M.; Vu, H.Q.; Li, G.; Law, R. Tourist behavior analysis in gaming destinations based on venue check-in data. *J. Travel Tour. Mark.* **2018**, *36*, 107–118. [[CrossRef](#)]
23. Mou, N.; Zheng, Y.; Makkonen, T.; Yang, T.; Tang, J.J.; Song, Y. Tourists' digital footprint: The spatial patterns of tourist flows in Qingdao, China. *Tour. Manag.* **2020**, *81*, 104151. [[CrossRef](#)]
24. 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](#)]
25. Yang, X.; Pan, B.; Evans, J.A.; Lv, B. Forecasting Chinese tourist volume with search engine data. *Tour. Manag.* **2015**, *46*, 386–397. [[CrossRef](#)]
26. Peterson, B.A.; Perry, E.E.; Brownlee, M.T.J.; Sharp, R. The transient nature of concentrated use at a national park: A spatiotemporal investigation into visitor behavior. *J. Outdoor Recreat. Tour.* **2020**, *31*, 100310. [[CrossRef](#)]
27. Beeco, J.A.; Hallo, J.C. GPS tracking of visitor use: Factors influencing visitor spatial behavior on a complex trail system. *J. Park Recreat. Adm.* **2014**, *32*, 43–61.
28. Miah, S.J.; Vu, H.Q.; Gammack, J.; McGrath, M. A big data analytics method for tourist behaviour analysis. *Inf. Manag. Amster.* **2017**, *54*, 771–785. [[CrossRef](#)]
29. Prebensen, N.K.; Foss, L. Coping and co-creating in tourist experiences. *Int. J. Tour. Res.* **2011**, *13*, 54–67. [[CrossRef](#)]
30. Moss, A.; Francis, D.; Esson, M. The relationship between viewing area size and visitor behavior in an immersive Asian elephant exhibit. *Visit. Stud.* **2008**, *11*, 26–40. [[CrossRef](#)]
31. Urry, J. The tourist gaze “revisited”. *Am. Behav. Sci.* **1992**, *36*, 172–186. [[CrossRef](#)]
32. Shih, H. Network characteristics of drive tourism destinations: An application of network analysis in tourism. *Tour. Manag.* **2006**, *27*, 1029–1039. [[CrossRef](#)]
33. Aranburu, I.; Plaza, B.; Esteban, M. Sustainable cultural tourism in urban destinations: Does space matter? *Sustainability* **2016**, *8*, 699. [[CrossRef](#)]
34. Lew, A.; McKercher, B. Modeling tourist movements: A local destination analysis. *Ann. Tour. Res.* **2006**, *33*, 403–423. [[CrossRef](#)]
35. Leung, X.Y.; Wang, F.; Wu, B.; Bai, B.; Stahura, K.A.; Xie, Z. A social network analysis of overseas tourist movement patterns in Beijing: The impact of the Olympic Games. *Int. J. Tour. Res.* **2012**, *14*, 469–484. [[CrossRef](#)]
36. Liu, B.; Huang, S.S.; Fu, H. An application of network analysis on tourist attractions: The case of Xinjiang, China. *Tour. Manag.* **2017**, *58*, 132–141. [[CrossRef](#)]
37. Mou, N.; Yuan, R.; Yang, T.; Zhang, H.; Tang, J.J.; Makkonen, T. Exploring spatio-temporal changes of city inbound tourism flow: The case of Shanghai, China. *Tour. Manag.* **2020**, *76*, 103955. [[CrossRef](#)]
38. Li, Y.; Yang, L.; Shen, H.; Wu, Z. Modeling intra-destination travel behavior of tourists through spatio-temporal analysis. *J. Destin. Mark. Manag.* **2019**, *11*, 260–269. [[CrossRef](#)]
39. Chrysanthi, A.; Earl, G.P.; Pagi, H. Visitor movement and tracking techniques. A visitor-sourced methodology for the interpretation of archaeological sites. *Int. J. Herit. Digit. Era* **2012**, *1*, 33–37. [[CrossRef](#)]
40. Ferrante, M.; De Cantis, S.; Shoval, N. A general framework for collecting and analysing the tracking data of cruise passengers at the destination. *Curr. Issues Tour.* **2018**, *21*, 1426–1451. [[CrossRef](#)]
41. McKercher, B.; Shoval, N.; Park, E.; Kahani, A. The [Limited] Impact of Weather on Tourist Behavior in an Urban Destination. *J. Travel Res.* **2014**, *54*, 442–455. [[CrossRef](#)]
42. Birenboim, A.; Anton-Clavé, S.; Russo, A.P.; Shoval, N. Temporal Activity Patterns of Theme Park Visitors. *Tour. Geogr.* **2013**, *15*, 601–619. [[CrossRef](#)]
43. Ingrid, Z.; Malin, Z. Emotions in motion: Tourist experiences in time and space. *Curr. Issues Tour.* **2012**, *15*. [[CrossRef](#)]
44. XU, X.; Hu, J. Study on Spatiotemporal Behavior of Urban Park Tourists Based on GPS Data: A Case Study of Wuhan East Lake Scenic Area. *Econ. Geogr.* **2020**, *40*, 224–232. (In Chinese)