

Using Mobile Phone Data to Understand the Demographic Characteristics and Behavioral Patterns of Park Visitors in a Megacity, Beijing, China

Yihang Wu

Chinese Research Academy of Environmental Sciences

Wenhao Zhao

Chinese Research Academy of Environmental Sciences

JIN MA (✉ majin@craes.org.cn)

Chinese Research Academy of Environmental Sciences <https://orcid.org/0000-0002-0394-7238>

Liu Yaxi

Institute of Geographic Sciences and Natural Resources Research Chinese Academy of Sciences

Tao Pei

Institute of Geographic Sciences and Natural Resources Research Chinese Academy of Sciences

Qiyuan Liu

Chinese Research Academy of Environmental Sciences

Yajing Qu

Chinese Research Academy of Environmental Sciences

Research

Keywords: behavioral patterns, mobile phone data, demographic characteristics, urban park, Beijing

Posted Date: October 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-942007/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Urban parks are important places that allow urban residents to experience nature but are also associated with the risk of exposure to contaminated soil. Local researches on demographic characteristics and population behavior patterns are the basis of soil exposure assessment. The objectives were to determine park visitors' demographic characteristics and behavioral patterns.

Methods

A total of 86 urban parks in Beijing were selected, and mobile phone data were used to quantify the park visits, identify the demographic characteristics of the visitors, and determine the visitors' stay times. A Kruskal-Wallis test and kernel density estimation were used for difference analysis and to describe the spatial aggregation of visitors, respectively. A random forest model was used to analyze the factors influencing the visitors' stay times.

Results

Visitor numbers at comprehensive parks and theme parks decreased significantly in winter, whereas seasonal variations in visitor numbers at community parks and country parks were small. Almost half of the visitors to Beijing's urban parks were foreign visitors, who, unlike local visitors, preferred to visit parks located in the city's center. Parks were mostly used by males and visitors aged 31–45. Most visitors stayed in the park for 1–2 h, and the distance from a given visitor's home to the park was the most important factor affecting stay time, contributing 80.65% of the difference in stay time.

Conclusions

This study provides a quantitative analysis of the patterns relating to urban park use by different groups of people. Changes in behavior patterns of different groups of people should be considered when assessing soil exposure.

Highlights

1. Mobile phone data is an effective way to study population behavior patterns
2. Urban parks in Beijing serves not only local visitors but also foreign visitors
3. The behavior patterns by different groups of park visitors were quite different
4. Findings were translated into pragmatic suggestions for refined risk assessment

1. Introduction

With the rapid development of urbanization, 55% of the world's population lived in urban areas in 2018 (United Nations, 2018). Urban space is typically characterized by artificial environments, architecture, traffic, and air pollution, so that urban residents have few opportunities to experience nature (Orioli et al., 2019). As important public resources, urban parks are loved by urban residents because of the good natural environment, benefit to human health, quality of life and social interaction (Fischer et al., 2018; Kazmierczak, 2013). Nevertheless, urban park soils are frequently enriched in contaminants owing to the development of urbanization and industrialization (Frimpong and Koranteng, 2019; Khan et al., 2016), and people may be exposed to contaminated soil through ingestion, dermal absorption, and inhalation while engaging in various activities in urban parks (Fig. S1). Consequently, the health risks associated with soil pollution in urban parks have attracted widespread attention (Huang et al., 2021; Penteadó et al., 2021; Qu et al., 2020), and the Chinese government published the first risk control standard for soil contamination of park land in 2018.

Human activity patterns are important for exposure modeling (Brandon and Price, 2020; Wu et al., 2011). The selection of sensitive receptors and the study of population behavior patterns are the premise of fine soil exposure assessment and health risk assessment. Different groups of people have different recreational activities and levels of contact with the soil in the park, so that it is vital to select appropriate sensitive receptors when conducting soil exposure assessment. Identifying the demographic characteristics of park users is the first step in selecting sensitive receptors. In addition, the population behavior pattern significantly affects the soil exposure assessment results. For example, the average soil ingestion rate, key parameter for soil exposure assessment, is correlated to amount of time spent outdoors (CL: AIRE, 2014). Therefore, local park visiting patterns, the demographic characteristics and behavioral patterns of urban park visitors are of great importance to the study of the soil exposure levels of park visitors. Barrio-Parra et al. (2019) also emphasizes the importance of characterizing urban gardeners' local activity patterns when assessing human health risks in urban garden scenario. What's more, Wu et al. (2011) suggested that changes in human activity patterns should be taken into account in exposure modeling.

Until now, traditional research methods, such as questionnaire surveys, field observation, and interviews, have been widely used to study the behavioral patterns and demographic characteristics of park visitors (Bertram et al., 2017; Fischer et al., 2018; Kabisch et al., 2021; Subramanian and Jana, 2018). However, it is well known that traditional methods are typically subjective, costly and laborious (Li et al., 2019). In recent years, the use of big data, such as social media and mobile phone data, has gradually gained traction as a new method of studying park visitation patterns in view of the limitations of traditional methods and the rapid development of computer science- and internet-related techniques (Chen et al., 2018; Li et al., 2018; Lin et al., 2021; Song et al., 2020a). Many studies using big data focus on park accessibility (Guo et al., 2019a; Hamstead et al., 2018), the spatial and temporal distribution of urban park users (Chen et al., 2018; Li et al., 2017; Liang and Zhang, 2021; Ullah et al., 2020), and the factors influencing park usage (Fan et al., 2021; Lyu and Zhang, 2019), with these studies aiming to provide references for the planning and construction of urban parks. Although some previous studies based on questionnaires and interviews have revealed that park visit patterns are influenced by the visitors' age and

gender, unfortunately, such visitor data cannot be easily obtained (Song et al., 2020b; Ullah et al., 2019). Consequently, studies that use big data methods to analyze the demographic characteristics of urban park visitors and the behavioral patterns of different groups of people in parks are rare, which limits the ability to obtain more objective results on population behavior patterns and highlights the importance of this study.

In this study, based on the mobile phone location data which can identify the demographic characteristics of users, we aim to determine the demographic characteristics and behavioral patterns of urban park visitors to provide a reference for the refined assessment of park visitors' soil exposure levels. To achieve our research purpose, Beijing, a megacity, was chosen as the study site. The study's aims are as follows: to determine the visitation patterns of park visitors in Beijing based on mobile phone data; to identify the demographic characteristics of urban park visitors, such as gender, age, and local or foreign status; and to determine the visitors' length of stay in parks and its influencing factors. Our findings will be helpful for the assessments of urban park visitors' soil exposure levels and the refined risk assessment in urban park scenario.

2. Methods

2.1 Brief description of the study area

Beijing is China's capital and the city with the largest built-up area. The latitudinal range of the territory of Beijing is 115°25'–117°30' E, and its longitudinal range is 39°26'–41°03' N (Feng et al., 2019). According to the Beijing Municipal Bureau of Statistics, at the end of 2019, the city's permanent population amounted to 21.54 million, and the total permanent population in the central urban area is 52.17% of that in Beijing overall. In recent years, the main residential areas have expanded to between the Fifth and Sixth Ring Roads. To ensure an environmentally friendly and livable city, Beijing is vigorously developing and constructing urban parks. By the end of 2019, there were 344 urban parks in Beijing covering a total area of 35,157 hectares. Moreover, Beijing has the greatest area of green coverage among China's built-up environments (National Bureau of Statistics of China, 2020).

In this study, 86 parks within Beijing's Sixth Ring Road, an area with high population density (Guo et al., 2019a; Liu et al., 2018), were selected. In view of the diversity of urban parks in Beijing, we divided the 86 urban parks into four categories based on the Standard for Classification of Urban Green Space, the parks' functions, and previous studies (Guo et al., 2019b; Li et al., 2017). The four categories are community parks, comprehensive parks, country parks and theme parks, with 17, 22, 21, and 26 urban parks classified into these four categories, respectively. Figure 1 illustrates the geographic area of the study and the distribution of parks within it.

2.2 Data sources

This study involved several data sources. First, the list of parks was obtained from the Beijing Gardening and Greening Bureau's website. Second, the parks' boundaries were extracted from Google Earth and then

processed using ArcGIS 10.7 to obtain the parks' shapefiles. Third, we determined the park visit volumes and visitors' stay time in parks using mobile phone data obtained from one of the three communications operators in China. In anticipation of seasonal variations in the visitors' behaviors, 16 days' worth of data were collected, and these 16 days were evenly distributed across all four seasons in 2019. To avoid the impact of bad weather on park visits, study dates with good weather conditions were selected (Table S1). The method of identifying park visitors was as follows:

- a) By overlaying park boundaries and communication operator base stations for analysis, stations located within the boundary of a given park were extracted.
- b) For each study date, users whose trajectories between 6 a.m. and 9 p.m. include these stations located within the park boundary were extracted as candidate visitors, and their movement trajectories were extracted.
- c) According to the candidate visitors' trajectories and excluding the park staff, we designated candidate visitors who stayed in the park for 0.5–5 h as park visitors. To ensure that the data were representative, we removed parks that attracted fewer than 100 visitors on each study date (Ullah et al., 2020), and visitor volume data were obtained for 86 parks.

We further determined each park visitor's gender and age. The locations of the visitors' homes could also be determined by extracting the most frequent locations of the visitors at night (00:00–06:00 and 21:00–24:00). Local and foreign visitors could be distinguished because gender and age data could not be obtained for foreign visitors.

Finally, to study the factors influencing visitors' stay time in parks and based on existing research regarding park visitation (Aikoh et al., 2012; Bertram et al., 2017; Chen et al., 2018; Jiang et al., 2019; Subramanian and Jana, 2018), we selected three categories of variables that may influence visitors' stay times. The first category was the parks' attributes and surrounding environmental characteristics and included four independent variables: park area, number of dwellings and number of transportation facilities within 2 km of the park, and the distance from a given visitor's home to the park. We chose 2 km as the threshold distance as people are particularly likely to visit parks within 2 km from their homes (Lin et al., 2021; Tu et al., 2020). Park area data were obtained from the website of the Beijing Gardening and Greening Bureau. The numbers of dwellings and transportation facilities were expressed by the number of points of interest, which was obtained by crawling the Baidu map of the area. The distances from visitors' homes to the parks were calculated based on latitude and longitude. The second category included the characteristics of park visitors, such as the gender and age of the visitors. The third category was temperature, which represented the environmental variable. Temperature was acquired from an open platform (<https://www.tianqi.com/>).

2.3 Statistical analysis

We first used Excel 2016 and SPSS 25.0 for basic data processing and statistical analysis. Because the data did not follow a normal distribution, the Kruskal-Wallis test was used to compare differences in visitors' gender and age across different seasons and different park types. It was also used to analyze the differences in the proportions of local and foreign visitors in different seasons. A Dunn-Bonferroni test

was used for pairwise multiple comparisons. Kernel density estimation (KDE) was used in ArcGIS 10.7 to study the spatial differences in park visits between foreign and local visitors. KDE is an effective measurement tool used to describe the spatial structure of the density of visitors within each selected urban park (Li et al., 2017). Random forest (RF) was used to analyze the factors influencing visitors' stay time in parks and their relative importance. RF is an integrated machine-learning method that assigns importance measures to variables to facilitate the evaluation of each variable's importance. The RF model was implemented in Python 3.7 software. The optimized parameters of the RF model were as follows: $n_estimators = 80$, $max_depth = 5$, $min_samples_split = 2$.

3. Results

3.1 Seasonal variation in park visitor volumes

Figure S2 illustrates the seasonal variation in the average visitor numbers at different parks. In general, the average visitor numbers were highest in autumn and lowest in winter, although few differences in the average visitor numbers were observed between spring, summer, and autumn. The average visitor numbers of comprehensive parks and theme parks were similar and significantly higher than those of community parks and country parks for any season. However, for different park types, the seasonal variation in average visitor volume was not exactly the same. The seasonal variation in average visitor volume for theme parks was the largest, followed by comprehensive parks. Regarding the above two park types, the visitor numbers in winter were significantly lower than those in other seasons. For community parks and country parks, the average visitor volumes exhibited little difference across the seasons.

3.2 Park visit patterns of local and foreign visitors

As Figure 2 illustrates, local and foreign visitors were almost equally divided in Beijing's urban parks. The respective proportions of local and foreign visitors in different seasons were as follows: spring, 50.36% and 49.64%; summer, 48.38% and 51.62%; autumn, 48.28% and 51.72%; and winter, 53.20% and 46.80%. The number of foreign visitors in winter was significantly lower than those in other seasons. Results of a Kruskal-Wallis test showed statistically significant differences in the proportions of local visitors among different seasons ($p < 0.05$) as well as for foreign visitors. Bonferroni tests revealed that the proportions of local visitors and foreign visitors differed significantly between autumn and winter ($p < 0.05$), whereas the proportions in spring, summer, and autumn were similar ($p > 0.05$).

The KDE method was used to analyze the spatial distribution of local and foreign visitors, and core agglomeration was used to reflect the gathering places where high-density spots of visitors in urban parks were located. As Figure 3 illustrates, the spatial agglomeration characteristics of local and foreign visitors differed significantly. For local visitors, two single-core agglomerations were observed with a high kernel density of visitor numbers. One was located near north Fifth Ring Road and mainly included the Olympic Forest Park and its surrounding parks (such as Yangshan Park). The other was located between west Fourth Ring Road and west Fifth Ring Road and included Beijing International Sculpture Park and Wukesong Olympic Park. Additionally, some differences in the spatial aggregations of local visitors

across the seasons were observed. The differences in seasonal variation were reflected in the disappearance of two single-core agglomerations with low kernel density of visitor numbers in winter. The two single-core agglomerations were located within the Second Ring Road and mainly included Taoranting Park, Shichahai Park, Jingshan Park, Huangchenggen Ruins Park, Temple of Heaven, and Zhongshan Park. However, for foreign visitors, only one single-core agglomeration was located in the city center, and it included Zhongshan Park, Jingshan Park, and Shichahai Park. Furthermore, the spatial gathering of foreign visitors was almost identical across seasons.

3.3 Differences in park visits by gender and age

Figure 4a shows the ratio of male to female (M/F) visitors to urban parks. Because the gender and age of foreign visitors could not be identified, we only analyzed age, gender, and length of stay in the park for local visitors. Overall, it was clear that the gender ratio was greater than 1 for different types of parks, with an average ratio of 1.585. Moreover, according to the results of the Kruskal-Wallis test ($p < 0.001$), there were statistically significant differences in the M/F ratios for visitors to different types of parks. The ratio was highest in country parks (1.847) and lowest in theme parks (1.401). No statistically significant difference was observed in the M/F ratio between community parks and comprehensive parks, which were associated with average ratios of 1.559 and 1.57, respectively. Regarding the influence of season on the M/F ratio, significant differences were observed across different seasons based on the Kruskal-Wallis test ($p < 0.01$). The ratios in spring, summer, autumn, and winter were 1.589, 1.601, 1.543, and 1.605, respectively. The ratio was lowest in autumn, with this value significantly different from the ratios in summer and winter.

The age distribution of park visitors was also considered (Figure 4b). Given that the proportion of visitors under the age of 18 was low and the data may not be representative, we only considered visitors older than 18. Figure 4b illustrates the differences in the proportion of visitors across different age groups. The proportion of visitors aged 31–45 was the highest, followed by visitors aged 46–60. In general, the proportions of park visitors aged 19–30, 31–45, 46–60, and > 60 were 11.23%, 41.40%, 31.52%, and 15.85%, respectively. As Figure 4b shows, no significant seasonal change emerged in the proportions of visitors of different ages, but significant differences in the proportions of visitors of different ages were found for different park types. Specifically, the proportions of visitors of different ages varied significantly between country parks and theme parks. Compared with theme parks, the proportion of visitors aged 31–45 was higher at country parks, whereas the proportion of visitors aged over 60 was lower. The average proportion of visitors aged over 60 at country parks was 13.5%, whereas the proportion of visitors aged over 60 at theme parks was 20%.

3.4 Visitors' length of stay in a park and its influencing factors

First, we observed that the visitors' length of stay in the parks ranged widely, with the shortest time being close to 0.5 h and the longest being close to 5 h (Figure 5). The median length of stay in a park was approximately 1.54 h, and the highest proportion of visitors (42%) stayed in a park for 1–2 h. A small

number of people (around 5%) remained for 4 h or longer. In addition, we observed significant differences in the time spent by visitors of different age groups in the parks ($p < 0.001$). Post hoc test comparisons indicated that visitors aged 31–45 and 46–60 did not differ in terms of their stay times. The median visitor stay times for the 19–30, 31–45, 46–60, and > 60 age groups were 1.594, 1.525, 1.539, and 1.627 h, respectively. The stay time of visitors over 60 was slightly higher than those of visitors in other age groups. Visitor stay times were also compared across different park types, and the visit duration varied among different types of parks ($p < 0.001$). The results showed that visitors stayed slightly longer in community parks (median 1.591 h) and theme parks (median 1.59 h) than in comprehensive parks (median 1.513 h) and country parks (median 1.520 h). Although the differences are not obvious (Figure 5), the visitors' length of stay did differ significantly across different seasons ($p < 0.001$).

To explore the reasons for changes in visitor stay times, the RF method was used to quantify the influences of seven factors on stay time. As can be seen from Figure 6, the distance from a visitor's home to the park was the most important factor affecting their stay time, contributing 80.65% to the change in stay time. Compared with the distance from home to the park, other factors had less influence on the stay time. Park area, the number of dwellings with 2 km of the park, temperature, the number of transportation facilities within 2 km of the park, the age and gender of visitors contributed 7.33%, 6.91%, 3.12%, 1.81%, 0.18%, and 0%, respectively. Figure 7 illustrates how some factors contribute to the variation in stay time. Given that gender had almost no effect on the length of stay, it was not included in the figure. Figure 7a demonstrates that the stay time and the distance from home to the park were significantly negatively correlated when the distance was short, whereas a significantly positive correlation was found when the distance was far. Park area, the number of dwellings, and the number of transportation facilities were all positively correlated with stay time (Figure 7). Moreover, the correlation between park area and stay time increased significantly when park area is larger, which indicates that visitors typically stay longer in larger parks. Lower temperatures were negatively correlated with stay time, whereas no correlation was observed at higher values. By contrast, no correlation between age and stay time was observed for younger groups, but a weak positive correlation was noted for older groups.

4. Discussion

4.1 Park users' demographic characteristics and visit patterns

In proportional terms, almost half of the visitors to Beijing's urban parks were from elsewhere. This may first be attributed to the fact that Beijing is a metropolis with a large number of immigrant population. Statistics show that at the end of 2019, Beijing's permanent migrant population amounted to 7.456 million, accounting for 34.6% of the permanent resident population (Beijing Municipal Bureau of Statistics, 2020). In addition, Beijing is a famous historical city with many places of interest. According to statistics recorded by the Beijing Municipal Bureau of Culture and Tourism, there were 227 scenic spots with star ratings in 2020 that attracted large numbers of tourists. According to the World Tourism Cities

Development Report (2019), Beijing ranks first in the list of tourist destination cities in China, and so foreign tourists may constitute a significant proportion of visitors to Beijing's urban parks.

Local and foreign visitors may have different park visiting preferences, while few studies have paid attention to this. We noted that most of the visitors who went to the parks located in the city center were foreign visitors. Similar results have been reported by previous studies. For example, a study of Beijing tourism found that the central city is typically the most densely populated area with respect to foreign tourists because the downtown district is home to several traditional scenic spots and modern shopping areas (Zhang et al., 2019). The same explanation applies to the distribution of foreign visitors in urban parks (Li et al., 2017; Song et al., 2020b). Regarding the influence of season on visitor distribution, only the geographic distribution of local visitors was weakly associated with seasonal changes, with the results showing that parks located in the center of Beijing are less attractive to local visitors during winter. This is because the parks located in Beijing's center are mainly theme parks, and their attraction for local visitors declines during the winter owing to the low temperatures and changes in scenery. This is consistent with the significant decline in the numbers of visitors to theme parks in winter, as reflected in Section 3.1. However, seasonal changes did not affect the geographical distribution of foreign visitors. This indicates that parks located downtown are attractive to foreign visitors across all seasons because of their popularity and proximity to the city center.

It also emerged that the number of visitors varied with gender and age, which is consistent with previous studies (Azagew and Worku, 2020; Lin et al., 2014; Liu et al., 2017). In this study, the overall M/F ratio was 1.585, which is higher than that of the Beijing population (1.032) (Beijing Municipal Bureau Statistics, 2020). This indicates that men are more likely than women to visit parks in Beijing, as some researchers have reported in other contexts. For example, Cohen et al. (2007) found that the M/F ratio for visitors to Los Angeles' urban parks was around 1.63, based on field observation and face-to-face interviews. Lin et al. (2014) found that the number of male visitors was slightly higher than that of female visitors. Women are more likely to be discouraged from visiting parks owing to fear and a lack of companions or interest (Derose et al., 2018; Liu et al., 2017). The largest proportion of urban park visitors was aged 31–60, whereas the proportions of visitors aged 19–30 and over 60 were low. From examining the age distribution of park visitors who live in Beijing, we found that younger visitors aged 19–30 were not inclined to visit parks. However, with advancing age, people tend to visit parks more frequently to promote their health (Grilli et al., 2020). Liu et al. (2017a) found that the parents of children aged under 7 years were more likely to visit parks, so accompanying children may be another reason for middle-aged people over 30 to go to a park. Other studies have also reported on park visitors' age distributions. Veitch et al. (2015) found that the proportion of visitors aged 21–59 was 53%. Interestingly, Azagew and Worku (2020) found that the 16–25 age group used the park most, whereas those aged over 55 exhibited the least park use in Ethiopia, which is significantly different from our results. This suggests that the age distribution of park visitors may vary from region to region and may be related to the specific age distribution of the population in each region.

Additionally, significant differences in the gender and age of visitors were observed among different types of parks, which are likely related to the services provided by the park, the proximity of the park (Cohen et al., 2020), and the activities of visitors in the park. Women and elderly people are generally regarded as marginalized populations, and they are more likely to be reluctant to visit parks for multiple reasons. In this study, the median distances from a park visitor's home to different types of parks are as follows: community parks, 2.14 km; comprehensive parks, 3.09 km; country parks, 2.84 km; and theme parks, 3.23 km. The median distance from home to a country park is not much shorter than that to a comprehensive park or a theme park. However, the services and landscapes provided by country parks to visitors are not as good as those provided by comprehensive and theme parks. Furthermore, women and older people are less likely to participate in moderate-to-vigorous physical activities (Cohen et al., 2020; Subramanian and Jana, 2018). Sang et al. (2016) found that older people prefer to engage in activities related to nature, in contrast to younger people, and so, they are more likely to visit parks with better scenery for entertainment and sightseeing purposes. Gu et al. (2020) also found that age has a negative impact on the likelihood of visiting country parks. In summary, country parks are less attractive to marginalized populations. Theme parks encompass many park types, such as historical parks, memorial parks and parks with famous scenery, which have beautiful scenery and/or diverse functions, so they can attract different groups of people.

4.2 Factors influencing stay time and their relative importance

In this study, we found that most visitors spent 1–2 h in a park. To determine whether visitors from different regions spent different amounts of time in urban parks, we summarized the stay times of visitors from different regions (Table S2). The majority of visitors from most areas spent 1–2 h in the park visited. However, visitors from several areas such as Mianyang, Singapore, and Cape Town were observed to stay at a park for less than 15 min. Therefore, it is necessary to analyze the factors that influence visitor stay time.

Based on the results of RF modeling, the distance from the visitor's home to the park is the factor that exerts the greatest influence on stay time. This result is similar to that reported by Cohen et al. (2020), but it is not exactly the same. Cohen et al. (2020) only found that the average duration of stay in parks reached by active commuting tended to be shorter than that in parks reached by motor transport, and visitors who lived closer to parks were more likely to use active commuting, which means that the greater the distance from home to the park, the longer the stay time. Actually, visitors are also likely to stay longer in a park when the distance from their home to the park is short. Longer stays in parks closer to home may be due to visitor behavior and activities in the park. Parks close to residential areas are typically designed to serve the surrounding residents, and it may be that visitors engage in activities in parks that take some time, such as accompanying children, dancing, playing chess, and exercising (Figure S1; Wang et al., 2021). Pham et al. (2019) showed that practicing sports and exercising were predominant in explaining longer stay times in parks. In Kansas City, 89.9% of park visitors participated in physical activities in parks, and they spent a mean of 77.1 min being physically active (Bai et al., 2013).

Excessively long distances from an individual's homes to the park typically reduces the frequency of visits (Bertram et al., 2017; Cohen et al., 2020). Therefore, visitors usually choose to stay longer in the park to compensate for the lack of visits to the park and obtain sufficient exposure to nature.

Park area was the second most important factor impacting stay time, although only 7.33% of the variation in stay time could be attributed to it. Clearly, one of the reasons that visitors stay longer in larger parks is that it usually takes longer to visit the entire park. Kazmierczak (2013) confirmed that visit length was associated with park size and that visitors tended to spend more time in large parks. Dwelling numbers can represent the population density near a park. Koohsari et al. (2020) found that population density was positively associated with sedentary behaviors, which may increase visitor stay time in a park. Temperatures was negatively associated with stay time at low values, which means that people usually spend less time outdoors because of the cold. Transportation facilities may affect visitor stay time by influencing their travel behaviors. For example, visitors can stay late in parks with convenient transportation, because they don't have to worry about the lack of transportation facilities to go home. In addition, a weak positive correlation was observed between age and stay time in parks with older people, consistent with findings from a study in Hong Kong, China (Wong, 2009). Given that 60 is the legal retirement age in China, this may be related to the fact that older people have more leisure time.

4.3 Implications for soil exposure assessment

To explore the behavioral patterns of urban park visitors, park visits, the visitors' demographic characteristics, and visitor length of stay were analyzed based on mobile phone data. The variations in the number of park visitors across different seasons and different park types suggest that season and park type affect visitor behaviors and may thus impact visitor soil exposure. In particular, no significant seasonal variation was evident in the number of visitors to community and country parks, suggesting that visitors are likely to visit these parks at a similar frequency and the frequency of visiting to these parks may be close to 365 days per year. However, the number of visitors to comprehensive and theme parks dropped significantly in winter, indicating that many visitors may reduce the frequency of visits to these parks in winter. It can be inferred that the frequency at which visitors go to comprehensive parks and theme parks is likely to be lower than that for community and country parks, with a corresponding lower frequency of soil exposure. In China, the frequency of soil exposure experienced by people who visit community parks and other types of parks are distinguished in technical guidelines for risk assessment of soil contamination of land for construction. Unfortunately, many researchers have failed to consider differences in visitors' frequency of exposure at different types of parks. To ensure a more accurate risk assessment, it is necessary to use different exposure frequencies for different types of parks in future research.

The visitors' demographic characteristics indicate that Beijing's urban parks not only provide services for local people but also for large numbers of migrants, although their park visit behavioral patterns differ considerably. For the sake of conservatism, the parameters used in risk assessment are usually based on local populations. However, for a city like Beijing with many immigrants and the exposure duration for local people may be longer than that for foreign people, it may be more accurate to consider the soil

exposure of visitors to parks dominated by foreign visitors. In addition, the number of male visitors is higher than that of female visitors in Beijing. In other words, it's too conservative to select only women as sensitive receptors in risk assessment. A better approach would be to take both men and women into account and some exposure parameters, such as body weight and height, be weighted according to the proportion of male and female visitors.

Although the durations of visitors' stay in the park varies widely and are significantly affected by the distance from home to the park, the occupancy period of 1.5 h may be a reference exposure parameter for risk assessment in a park scenario. In addition, most visitors spent 1–2 h in parks, significantly less time than in their residences (usually a minimum of 8 h) and visitors are often exposed at a lower frequency to parkland than to residential land. As a matter of fact, in China, no strict distinction is made between residential land and parkland when assessing the risk of contaminated sites, which may lead to an overestimation of the health risks faced by people in parkland. Therefore, to ensure more realistic results, different exposure parameters should be used for risk assessments of residential land and parkland.

4.4 Limitations

There may be a few limitations to this study. First, considering the availability of data and the complexity of human behavioral patterns, we only used data collected over 16 days to represent seasonal variation in park visits and visitor stay time in a park as an indicator to represent their behavioral patterns. In future research, more data that reflect seasonal changes and characterize visitors' behavioral patterns should be collected, if possible. Second, the mobile phone data used in this study cover children insufficiently and cannot accurately reflect patterns in children's visits to parks. Considering children's sensitivity to contaminated-soil exposure, additional research on children's park visits is warranted. Finally, although mobile phone data have the advantage of wide coverage, yield large amounts of information, and can more objectively reflect park visit trends, park visitation preferences are subjective and vary among human individuals. Thus, it is usually necessary to investigate individual subjective willingness to visit a park once human behavioral patterns are involved. For example, visitors' attitudes toward parks influence their use of the park, but such data are usually only available through visitors' self-reports (Zhang and Tan, 2019). Therefore, it may be more fruitful to use mobile phone data and data obtained from questionnaire surveys or field observations in combination.

5. Conclusions

To our knowledge, this study is the first to use mobile phone data to analyze the park visiting behavior with different demographic characteristics in detail. In this study, we found that males and visitors aged 31–45 are active users of urban parks and foreign visitors are also a part of Beijing urban park visitors that cannot be ignored. Differences were evident in the park visitation behavior of visitors with different demographics. Most visitors stay in a park for 1–2 h, and the distance from a visitor's accommodation to the park was the most important factor impacting the stay time of tourists. Big data offer a new means of studying the behavioral patterns of park visitors but should also be combined with traditional survey

methods to be more accurate and comprehensive. A more refined risk assessment may be achieved, for example, by distinguishing between residential land and parkland, different types of parks, and different groups of people.

Declarations

Ethics approval and consent to participate

Not Applicable.

Consent for publication

Not Applicable.

Availability of data and materials

The mobile phone data used in this study are available from one of China's three major telecommunications operators. They were used under license for the current study, and so are not publicly available.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by the National Key Research and Development Program of China (2019YFC1804601) and the National Natural Science Foundation of China (42177221).

Authors' contributions

Yihang Wu: Software, Formal analysis and Writing-Original draft. Wenhao Zhao: Software and Formal analysis. Jin Ma: Conceptualization and Supervision. Yaxi Liu: Data Curation. Tao Pei: Data Curation. Qiyuan Liu: Validation. Yajing Qu: Validation.

References

Aikoh, T., Abe, R., Kohsaka, R., et al., 2012. Factors influencing visitors to suburban open space areas near a northern Japanese city. *Forests*, 3(2), 155-165.

Azagew, S., Worku, H., 2020. Socio-demographic and physical factors influencing access to urban parks in rapidly urbanizing cities of Ethiopia: The case of Addis Ababa. *Journal of Outdoor Recreation and Tourism*, 31, 100322.

- Bai, H., Stanis, S.A.W., Kaczynski, A.T., et al., 2013. Perceptions of neighborhood park quality: associations with physical activity and body mass index. *Annals of Behavioral Medicine*, 45, S39-S48.
- Barrio-Parra, F., Izquierdo-Diaz, M., Dominguez-Castillo, A., et al., 2019. Human-health probabilistic risk assessment: the role of exposure factors in an urban garden scenario. *Landscape and Urban Planning*, 185, 191-199.
- Beijing Municipal Bureau of Statistics, 2020. Beijing Statistical Yearbook 2020. <http://nj.tjj.beijing.gov.cn/nj/main/2020-tjnj/zk/indexch.htm>.
- Bertram, C., Meyerhoff, J., Rehdanz, K., et al., 2017. Differences in the recreational value of urban parks between weekdays and weekends: A discrete choice analysis. *Landscape and Urban Planning*, 159, 5-14.
- Brandon, N., Price, P.S., 2020. Calibrating an agent-based model of longitudinal human activity patterns using the Consolidated Human Activity Database. *Journal of Exposure Science and Environmental Epidemiology*, 30(1), 194-204.
- Chen, Y.Y., Liu, X.P., Gao, W.X., et al., 2018. Emerging social media data on measuring urban park use. *Urban Forestry and Urban Greening*, 31, 130-141.
- CI: AIRE, 2014. Development of category 4 screening levels for assessment of land affected by contamination.
- Cohen, D.A, Mckenzie, T.L., Sehgal, A., et al., 2007. Contribution of public parks to physical activity. *American Journal of Public Health*, 97(3), 509-514.
- Cohen, D.A., Williamson, S., Han, B., 2020. Gender differences in physical activity associated with urban neighborhood parks: Findings from the National Study of Neighborhood Parks. *Women's Health Issues*. 31(3), 236-244.
- Derose, K.P., Han, B., Williamson, S., et al., 2018. Gender disparities in park use and physical activity among residents of high-poverty neighborhoods in Los Angeles. *Women's Health Issues*, 28(1), 6-13.
- Fan, Z.X., Duan, J., Lu, Y., et al., 2021. A geographical detector study on factors influencing urban park use in Nanjing, China. *Urban Forestry and Urban Greening*, 59, 126996.
- Feng, S., Chen, L.D., Sun, R.H., 2019. The distribution and accessibility of urban parks in Beijing, China: Implications of social equity. *International Journal of Environmental Research and Public Health*, 16(24), 4894.
- Fischer, L.K., Honold, J., Botzat, A., et al., 2018. Recreational ecosystem services in European cities: Sociocultural and geographical contexts matter for park use. *Ecosystem Services*, 31, 455-467.

- Frimpong, S.K., Koranteng, S.S., 2019. Levels and human health risk assessment of heavy metals in surface soil of public parks in Southern Ghana. *Environmental Monitoring and Assessment*, 2019, 191(9), 588.
- Grilli, G., Mohan, G., Curtis, J., et al., 2020. Public park attributes, park visits, and associated health status. *Landscape and Urban Planning*, 199, 103814.
- Gu, X.K., Li, Q., Chand, S., 2020. Factors influencing residents' access to and use of country parks in Shanghai, China. *Cities*, 2020, 97, 102501.
- Guo, S.H., Song, C., Pei, T., et al., 2019a. Accessibility to urban parks for elderly residents: Perspectives from mobile phone data. *Landscape and Urban Planning*, 191, 103642.
- Guo, S.H., Yang, G.G., Pei, T., 2019b. Analysis of factors affecting urban park service area in Beijing: Perspectives from multi-source geographic data. *Landscape and Urban Planning*, 181, 103-117.
- Hamstead, Z.A., Fisher, D., Illieva, R.T., et al., 2018. Geolocated social media as a rapid indicator of park visitation and equitable park access. *Computers Environment and Urban Systems*, 72, 38-50.
- Huang, J., Wu, Y., Sun, J., et al., 2021. Health risk assessment of heavy metal(loid)s in park soils of the largest megacity in China by using Monte Carlo simulation coupled with Positive matrix factorization model. *Journal of Hazardous Materials*, 415, 125629.
- Jiang, Y.Q., Huang, G.L., Fisher, B., 2019. Air quality, human behavior and urban park visit: A case study in Beijing. *Journal of Cleaner Production*, 240, 118000.
- Kabisch, N., Kraemer, R., Masztalerz, O., et al., 2021. Impact of summer heat on urban park visitation, perceived health and ecosystem service appreciation. *Urban Forestry and Urban Greening*, 60, 127058.
- Kazmierczak, A., 2013. The contribution of local parks to neighbourhood social ties. *Landscape and Urban Planning*, 109(1), 31-44.
- Khan, S., Munir, S., Sajjad, M., 2016. Urban park soil contamination by potentially harmful elements and human health risk in Peshawar City, Khyber Pakhtunkhwa, Pakistan. *Journal of Geochemical Exploration*, 165, 102-110.
- Koohsari, M.J., Shibata, A., Ishii, K., et al., 2020. Built environment correlates of objectively-measured sedentary behaviours in densely-populated areas. *Health and Place*, 66, 102447.
- Li, F.Z., Zhang, F., Li, X., et al., 2017. Spatiotemporal patterns of the use of urban green spaces and external factors contributing to their use in central Beijing. *International Journal of Environmental Research and Public Health*, 14(3), 237.

- Li, J.J., Xu, L.Z., Tang, L., 2018. Big data in tourism research: A literature review. *Tourism Management*, 68, 301-323.
- Liang, H.L., Zhang, Q.P., 2021. Temporal and spatial assessment of urban park visits from multiple social media data sets: A case study of Shanghai, China. *Journal of Cleaner Production*, 297, 126682.
- Lin, B.B., Fuller, R.A., Bush, R., et al., 2014. Opportunity or orientation? Who uses urban parks and why? *PLoS One*, 9(1), e87422.
- Lin, Y.Y., Zhou, Y.H., Lin, M.S., et al., 2021. Exploring the disparities in park accessibility through mobile phone data: Evidence from Fuzhou of China. *Journal of Environmental Management*, 281, 111849.
- Liu, H.X., Li, F., Xu, L.F., et al., 2017. The impact of socio-demographic, environmental, and individual factors on urban park visitation in Beijing, China. *Journal of Cleaner Production*, 163, S181-S188.
- Liu, Z., Ma, T., Du, Y.Y., et al., 2018. Mapping hourly dynamics of urban population using trajectories reconstructed from mobile phone records. *Transaction in GIS*, 22(2), 494-513.
- Lyu, F.N., Zhang, L., 2019. Using multi-source big data to understand the factors affecting urban park use in Wuhan. *Urban Forestry and Urban Greening*, 43, 126367.
- National Bureau of Statistics of China, 2020. *China Statistical Yearbook 2020*.
<http://www.stats.gov.cn/tjsj/ndsj/2020/indexch.htm>.
- Orioli, R., Antonucci, C., Scortichini, M., et al., 2019. Exposure to residential greenness as a predictor of cause-specific mortality and stroke incidence in the Rome longitudinal study. *Environmental Health Perspectives*, 127(2), 027002.
- Penteado, J.O., Brum, R.D., Ramires, P.F., et al., 2021. Health risk assessment in urban parks soils contaminated by metals, Rio Grande city (Brazil) case study. *Ecotoxicology and Environmental Safety*, 208, 111737.
- Pham, T.T.H., Labbe, D., Lachapelle, U., et al., 2019. Perception of park access and park use amongst youth in Hanoi: How cultural and local context matters. *Landscape and Urban Planning*, 189, 156-165.
- Qu, Y.J., Gong, Y.W., Ma, J., et al., 2020. Potential sources, influencing factors, and health risks of polycyclic aromatic hydrocarbons (PAHs) in the surface soil of urban parks in Beijing, China. *Environmental Pollution*, 260, 114016.
- Sang, A.O., Knez, I., Gunnarsson, B., et al., 2016. The effects of naturalness, gender, and age on how urban green space is perceived and used. *Urban Forestry and Urban Greening*, 18, 268-276.
- Song, X.P., Richards, D.R., He, P.J., et al., 2020a. Does geo-located social media reflect the visit frequency of urban parks? A city-wide analysis using the count and content of photographs. *Landscape and Urban*

Planning, 203, 103908.

Song, X.P., Richards, D.R., Tan P.Y., 2020b. Using social media user attributes to understand human–environment interactions at urban parks. *Scientific Reports*, 10(1), 808.

Subramanian, D., Jana, A., et al., 2018. Assessing urban recreational open spaces for the elderly: A case of three Indian cities. *Urban Forestry and Urban Greening*, 35, 115-128.

Tu, X.Y., Huang, G.L., Wu, J.G., et al., 2020. How do travel distance and park size influence urban park visits? *Urban Forestry and Urban Greening*, 52, 126689.

Ullah, H., Wan, W.G., Haidery, S.A., et al., 2019. Analyzing the spatiotemporal patterns in green spaces for urban studies using location-based social media data. *Isprs International Journal of Geo-information*, 8(11), 506.

Ullah, H., Wan, W.G., Haidery, S.A., et al., 2020. Spatiotemporal patterns of visitors in urban green parks by mining social media big data. *IEEE Access*, 8, 39197-39211.

United Nations, 2018. *World Urbanization Prospects Report (2018)*.
<https://population.un.org/wup/Publications/>.

Veitch, J., Carver, A., Abbott, G., et al., 2015. How active are people in metropolitan parks? An observational study of park visitation in Australia. *BMC Public Health*, 15, 610.

Wang, P.W., Zhou, B., Han, L.R., et al., 2021. The motivation and factors influencing visits to small urban parks in Shanghai, China. *Urban Forestry and Urban Greening*, 60, 127086.

Wong, K.K., 2009. Urban park visiting habits and leisure activities of residents in Hong Kong, China. *Managing Leisure*, 14(2), 125-140.

Wu, X.M., Bennett, D.H., Lee, K., et al., 2011. Longitudinal variability of time-location/activity patterns of population at different ages: a longitudinal study in California. *Environmental Health*, 10:80.

Zhang, K., Chen, Y., Li, C.L., 2019. 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. *Tourism Management*, 75, 595-608.

Zhang, J.Y., Tan, P.Y., et al., 2019. Demand for parks and perceived accessibility as key determinants of urban park use behavior. *Urban Forestry and Urban Greening*, 44, 126420.

Figures

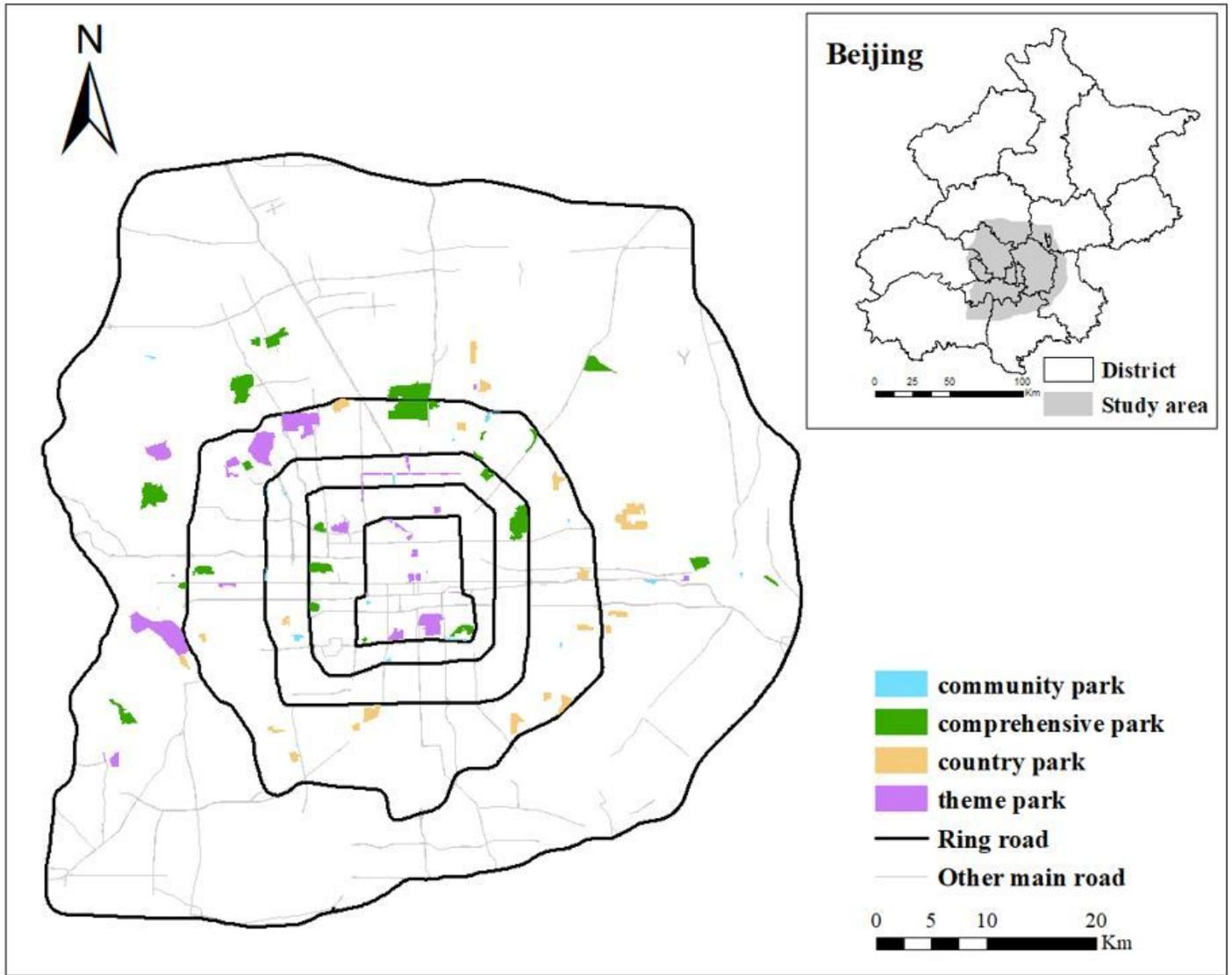


Figure 1

Study area and distribution of parks in the present study

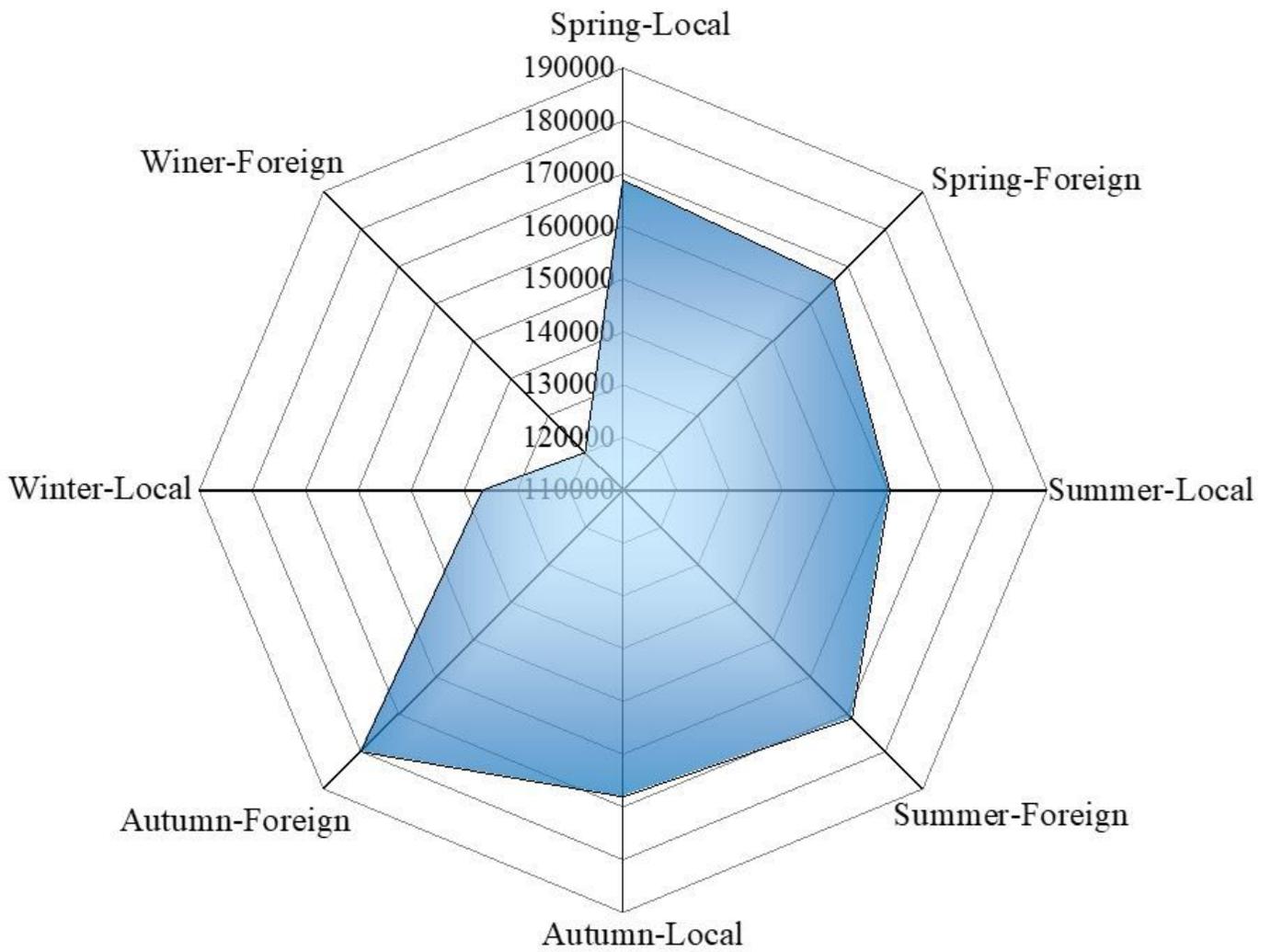
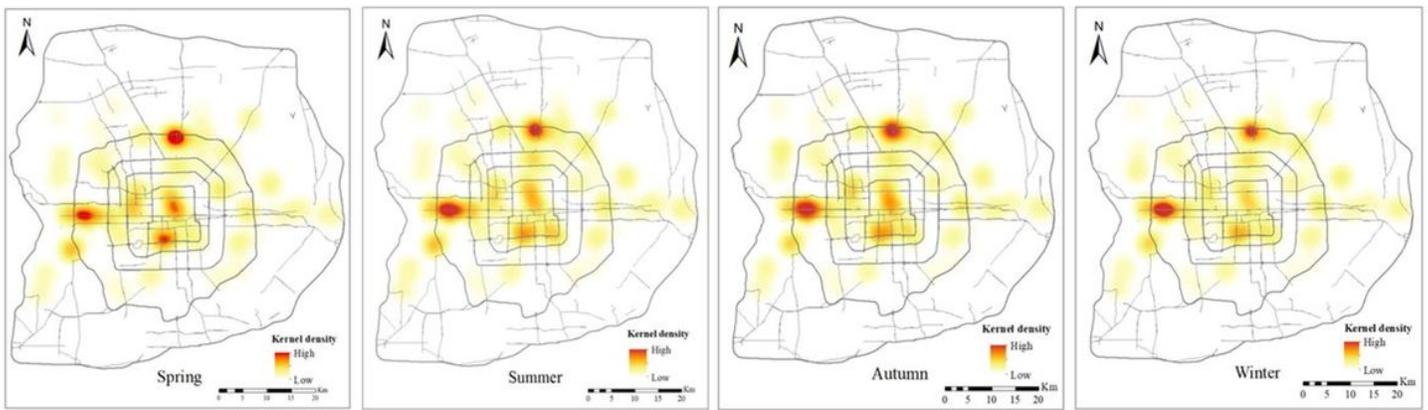


Figure 2

The number of local and foreign visitors in different seasons note: Spring-Local means the number of local visitors in spring

(a) Local visitors



(b) Foreign visitors

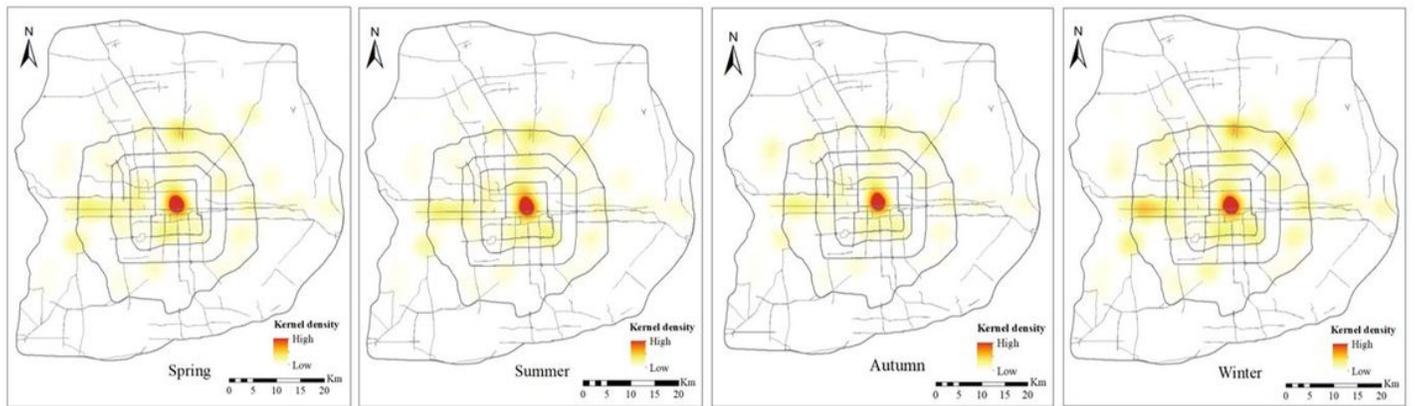


Figure 3

Kernel density relating to visitor numbers: (a) local tourists; (b) foreign tourists

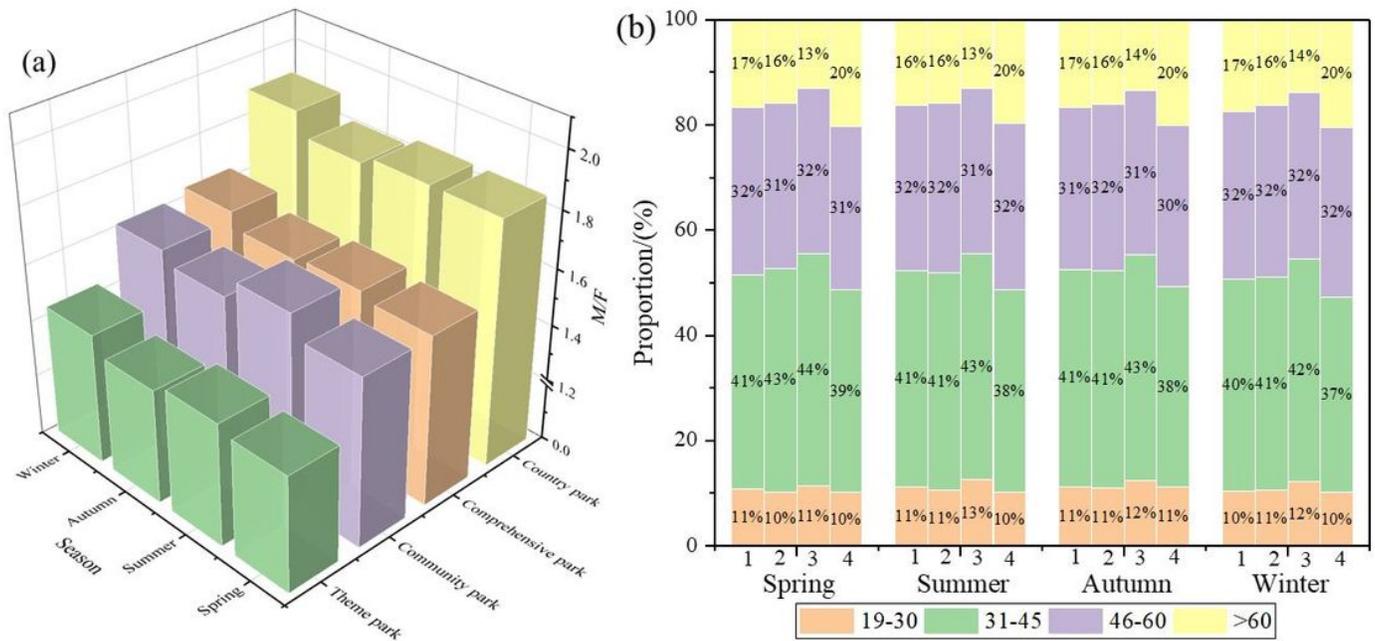


Figure 4

Gender and age distribution of park visitors: (a) gender; (b) age note:1-community park, 2-comprehensive park, 3-country park, 4-theme park

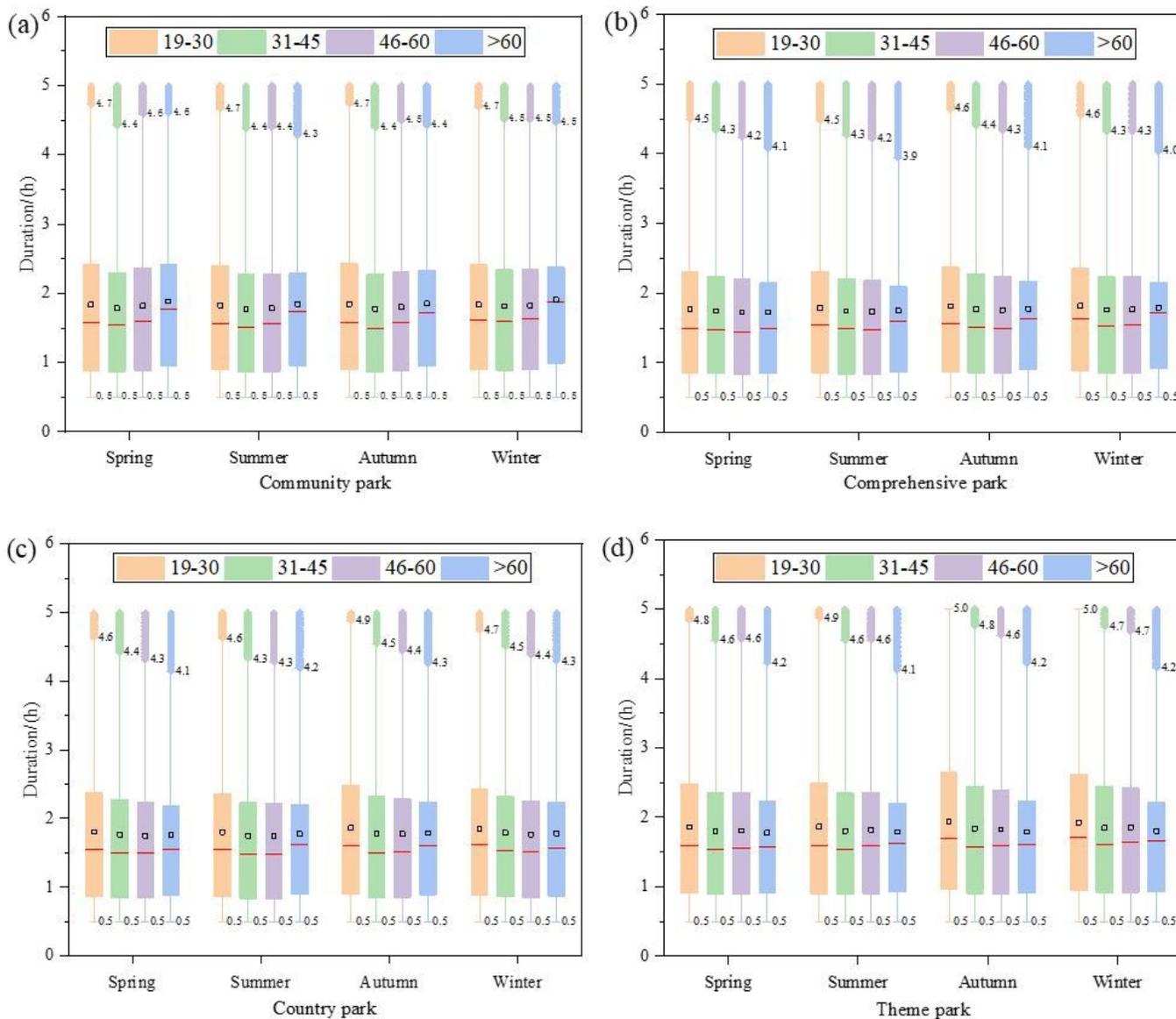


Figure 5

Stay time of visitors in different types of parks

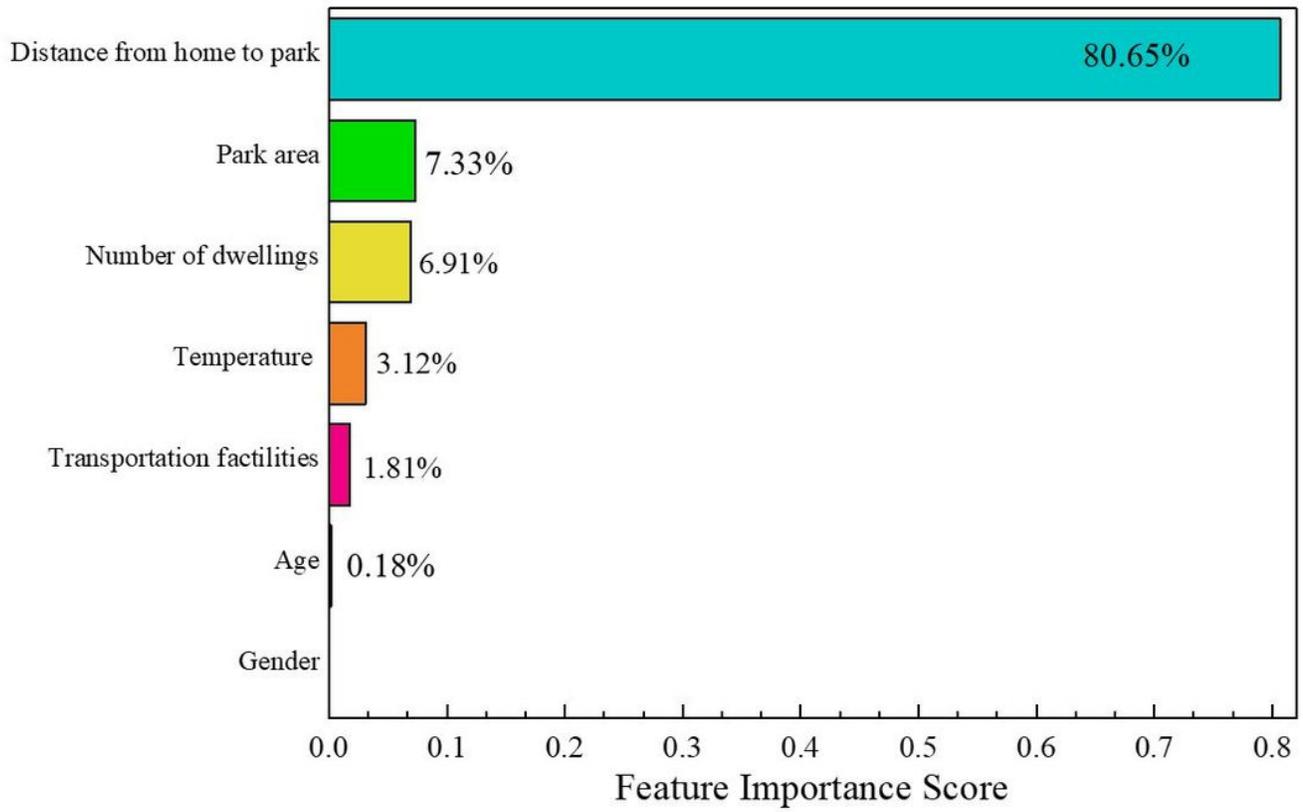


Figure 6

Importance of factors affecting visitors' stay time

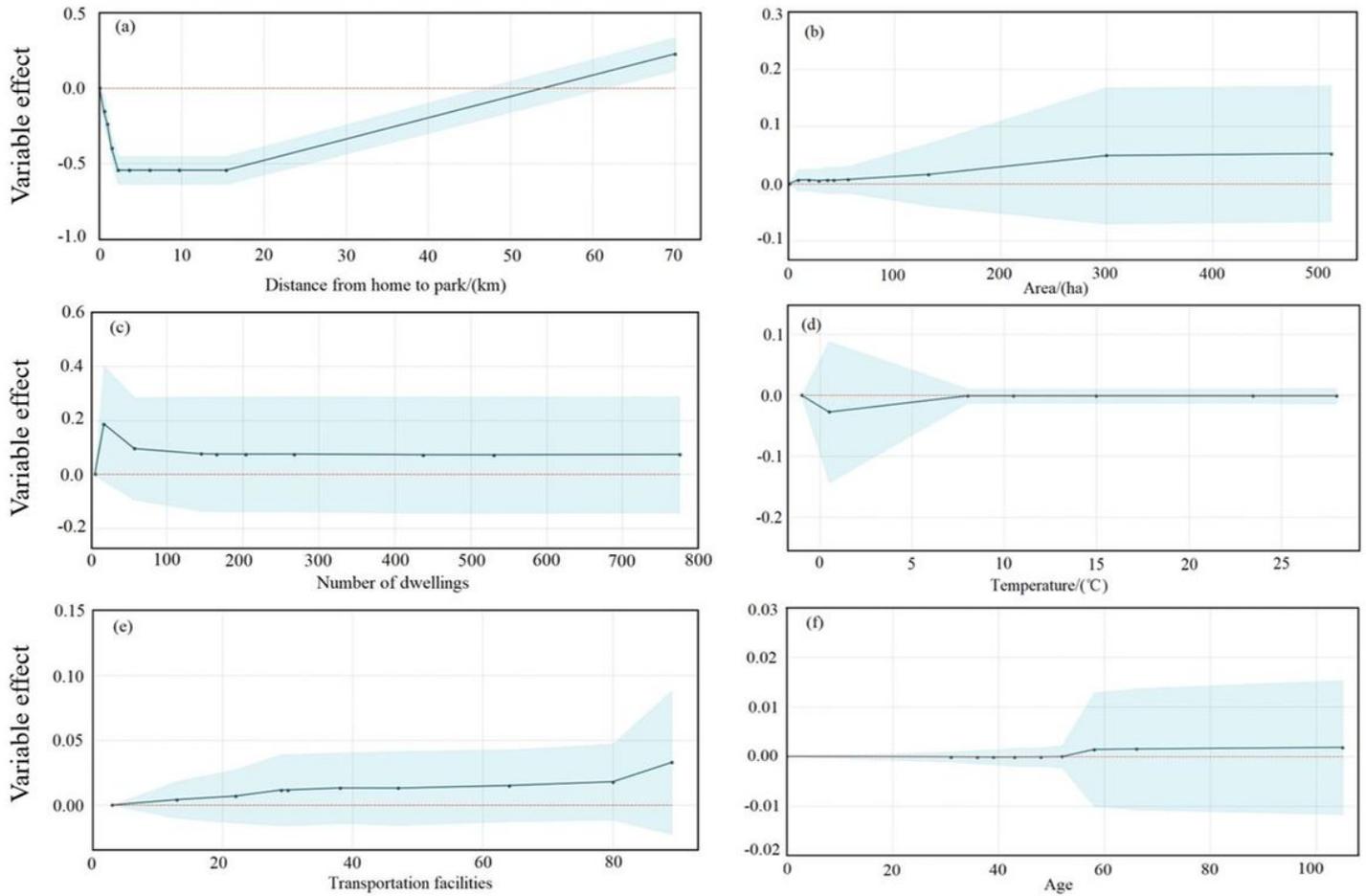


Figure 7

Single factor partial dependence plot for the random forest model of stay time in park

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterial.docx](#)