Applied Geography 167 (2024) 103300

Contents lists available at ScienceDirect

Applied Geography

journal homepage: www.elsevier.com/locate/apgeog

Visitation-based classification of urban parks through mobile phone big data in Tokyo

Yichun Zhou [a,b], ChengHe Guan [a,c,**], Longfeng Wu [d,*], Ying Li [a,c], Xuanyi Nie [e], Jihoon Song [f], Seung Kyum Kim [g], Yuki Akiyama [h]

[a] Shanghai Key Laboratory of Urban Design and Urban Science, NYU Shanghai, Shanghai, China
[b] Department of Civil and Urban Engineering, Tandon School of Engineering, New York University, Brooklyn, NY, USA
[c] Division of Arts and Sciences, NYU Shanghai, Shanghai, China
[d] College of Urban and Environmental Sciences, Peking University, Beijing, China
[e] The University at Buffalo's School of Architecture and Planning, NY, USA
[f] Ojeong Resilience Institute, Korea University, South Korea
[g] Korea Advanced Institute of Science and Technology, Graduate School of Future Strategy, Daejeon, South Korea
[h] Faculty of Architecture and Urban Design, Tokyo City University, Japan

ARTICLE INFO

Keywords:
Urban parks
Mobile phone data
Visitation patterns
Clustering
Trip distance

ABSTRACT

Urban parks, pivotal in fostering physical activity, mental well-being, and environmental stewardship, are integral to green infrastructure planning. Despite advances in georeferenced data applications, existing park classifications often overlook actual visitation patterns. This study reclassifies urban parks using over 5.9 million records from approximately 330 thousand visitors across 300 Tokyo parks, comparing with size-based park categorizations. We employed a range of analytical tools, including principal component analysis, Isolation Forest algorithm, various clustering algorithms, and the Gini index. Our findings first revealed four key visitation indicators, activity intensity, utilization efficiency, temporal occupancy, and revisit volume. These indicators uncovered parks with atypical visitation patterns within each size category, leading to three novel park classifications, everyday leisure parks, social destination parks, and seasonal activity parks. Moreover, we discovered notable disparities in distances traveled to parks, particularly during nights, weekends, and holidays, with pronounced inequalities in seasonal activity parks and smaller parks. The findings advocate for a nuanced park management strategy, prioritizing maintenance and amenity development aligned with observed visitation patterns to enhance recreational potential, thereby contributing insights to urban park research that support the advancement of green infrastructure planning and policy aimed at improving park utility and enjoyment.

1. Introduction

Urban parks represent more than just scenic vistas within the concrete landscape of cities, they are integral to individuals' daily movement by fostering physical health and psychological well-being (Csomos et al., 2023; Doll et al., 2022; Ma et al., 2023). The intricate behaviors exhibited by park visitors, including movement patterns, frequency and timing of visits, and duration of stays, underscore the complex role that urban parks fulfill in the lives of city residents (Guan et al., 2020; Song et al., 2020; Tu et al., 2020; Zhang & Zhou, 2018). While size-based categorizations provide a structural framework (Csomos et al., 2023), they may overlook these detailed engagement patterns within park visitation. (Rigolon, 2016). A refined approach to classifying urban parks that incorporates visitation behaviors is essential to inform effective urban green infrastructure planning and management (Doll et al., 2022; Ibes, 2015; Talal & Santelmann, 2020).

This study seeks to delve into visitation patterns by leveraging mobile phone data to glean insights into park utilization and visitor behavior. We aim to explore the urban park dynamics, examine following questions: (1) What visitation indicators can be derived from mobile data to accurately reflect park usage? (2) To what extent does the size of an urban park influence its visitation dynamics, and can we identify parks that defy the norm within each size category? (3) How can these indicators inform a visitation-based classification of urban parks

* Corresponding author.
** Corresponding author. Shanghai Key Laboratory of Urban Design and Urban Science, NYU Shanghai, Shanghai, China.
E-mail addresses: chenghe.guan@nyu.edu (C. Guan), longfengwu@pku.edu.cn (L. Wu).

https://doi.org/10.1016/j.apgeog.2024.103300
Received 7 February 2024; Received in revised form 9 May 2024; Accepted 17 May 2024
Available online 20 May 2024

and highlight disparities in park utilization? Our inquiry is motivated by the need to transcend size-based categorization, which may not always align with the realities of visitor engagement and park activities.

2. Literature review

2.1. Urban park classification systems

Urban parks are not mere embellishments in the cityscape but serve as vital components for enhancing urban livability, promoting ecological sustainability, and offering recreational and restorative environments for city residents (Talal & Santelmann, 2020; Zhang & Zhou, 2018). As the understanding of their multifaceted roles deepens, so does the complexity of classifying these spaces. Historically, classifications have been contingent upon quantifiable characteristics such as size, utility, and location (Ibes, 2015). However, emerging critiques highlight that these paradigms are insufficient for capturing resident's actual usage of urban parks, which function as nodes for social interaction, biodiversity conservation, and as barometers for urban health (Nielsen et al., 2013; Palliwoda et al., 2017).

A call for a more inclusive classification approach is emergent, one that integrates the physical dimensions with social values. This approach recognizes that urban parks are not uniform but rather diverse in their offerings and significance to different community groups (Peters et al., 2010). The burgeoning field of geospatial technologies has introduced a dynamic element to the study of urban parks. High-resolution imagery and georeferenced data from mobile devices provide nuanced insights into how these spaces are utilized, marking a shift from static classification to a dynamic understanding based on real-world usage (Ren & Guan, 2022; Shahtahmassebi et al., 2021). However, this shift brings to the fore a research gap: the need for classification systems that are grounded in the lived experiences of urban residents. Current literature often limits the scope to broad city or regional levels, bypassing the granular details of interaction within individual urban parks.

2.2. Urban park visitation behavior

The study of urban park visitation behavior is paramount in the realm of urban planning and public space management, providing essential insights into the multifaceted use of these crucial city landscapes (Donahue et al., 2018). Traditional methodologies such as surveys and direct observations have been instrumental in highlighting general park usage patterns; however, they are often constrained by their sporadic nature and limited capacity to grasp the continuous and dynamic flux of visitor interactions (Whiting et al., 2012). In contrast, the advent of mobile technology and geospatial analytics has revolutionized this field, delivering an unprecedented depth of detailed, real-time data that encapsulates a wide spectrum of user behaviors and park engagements (Chen et al., 2018; Lyu & Zhang, 2019).

This granular mobile data is invaluable for park studies, as it is in aligning park management strategies with the genuine rhythms and preferences of visitors. Such alignment ensures that parks effectively serve the community's needs throughout varying times of day and changing seasons (Chen et al., 2015; Guan et al., 2020; Ullah et al., 2020). Despite these technological strides, a research gap persists, an in-depth examination into the temporal nuances of park visitation behaviors is lacking. Existing studies often fail to dissect how engagement with parks varies during daily cycle, holidays, and across seasons, which are critical for understanding the full scope of park utilization (Ngesan et al., 2012; Tinsley et al., 2010). This oversight underscores the need for detailed spatiotemporal analyses of park visitation patterns.

2.3. Disparities in trip distance to parks

The evaluation of urban park utilization is shifting from static accessibility measures to dynamic analyses that account for temporal variations in park use. Initial assessments primarily concentrated on the parks' physical presence, geographic distribution, and the surrounding population density, using methods such as the gravity model (Guo et al., 2019; Wang et al., 2015). They have illuminated critical aspects such as the availability and proximity to residential areas, incorporating considerations like modes of transport (Xiao et al., 2017; Zhang & Tan, 2019). Yet, such measures often remain static, lacking the capacity to reflect the dynamic, temporal nature of park visitation. Additionally, they do not fully capture how parks are used, the diversity of activities within them, or the motivations behind visits.

Research has shown that the distances traveled to parks often surpass recommended accessibility guidelines (Schindler et al., 2022), underscoring the importance of including temporal considerations in park usage evaluations (Li et al., 2021; Ullah et al., 2020). Recent research has been advocating for innovative approaches that capture various spectrum of actual distances traveled to park. The advent of mobile data analytics represents a significant step forward, offering detailed insights into individual and collective behaviors within urban parks (Guan et al., 2020; Guo et al., 2019, 2022; Ren & Guan, 2022). Prior research often lacked the spatiotemporal aspects of travel to parks, relying on static assessments that do not capture fluctuating usage trends. Through real-time tracking of movements, analytics can illuminate the intricate patterns of park visitation, revealing disparities in trip distances to urban parks across varying times.

3. Data and methods

3.1. Study area

The focus of this study is Tokyo's 23 special wards, center of political, economic, and cultural activities in Japan and one of the most densely populated metropolitan areas globally, with a population of 9.71 million as of 2021 (Tokyo Metropolitan Government, 2021). Covering an area of 628 km^2, these wards provide a unique setting for analyzing urban park visitation patterns within densely populated areas. Tokyo's well-established urban park system, bolstered by the 2020 Action Plan (2016) for parkland development, grew to approximately 2,028 ha by 2017, an increase of 6.4 ha from the previous year (Bureau of Construction). To gather park geographic data, we utilized the QuickOSM plugin with the QGIS software. Additionally, we collected data on the opening year of each park. For parks with missing opening years, we conducted manual searches to obtain this information. Details on the area size and opening year of each park are provided in Table S1 in the Supplementary Materials. Fig. 1 shows the geographic distribution of the 300 selected parks, organized into five size-based categories, 30 city block parks (0.25–1 ha), 137 neighborhood parks (1–3 ha), 78 district parks (3–10 ha), 45 comprehensive parks (10–50 ha), and 10 national parks (50+ ha), aligning with previous researches and administrative park classifications (Csomos et al., 2023; Ministry of Land, 2006).

3.2. Mobile phone data

Mobile phone data were sourced from ZENRIN DataCom Co., LTD, utilizing the Konzatsu-Tokei® Data. This dataset compiles location signals from mobile phone users who have consented to share their data, with individual privacy protected through statistical anonymization. Originating from NTT DOCOMO, INC.'s "docomo map navi" service, the dataset encompasses the entire study area for 2012, providing up to 288 anonymized geolocation entries per user per day. These entries include user ID, datetime stamps, and georeferenced locations, all devoid of personal information, ensuring anonymity.

To ascertain park visitation, we first confirmed the existence of the 300 selected parks in 2012. We then aligned geolocation data with park boundaries, treating any data point within these boundaries as a potential visit. A visit was classified as such if it included a minimum stay

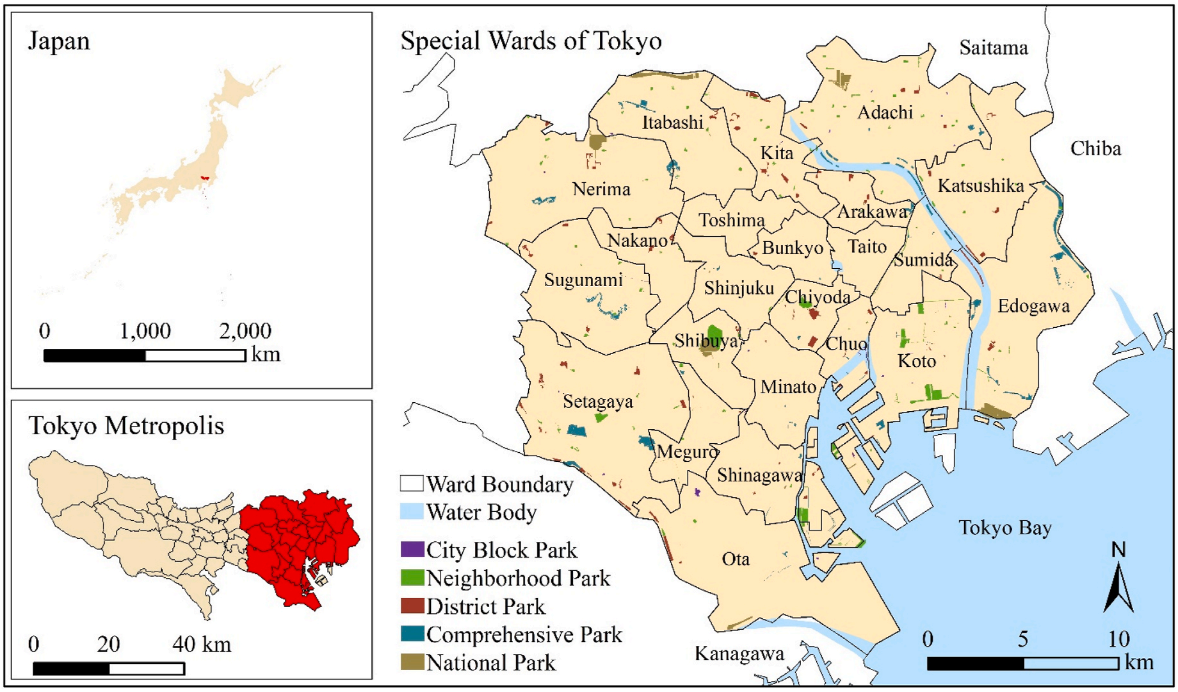

Fig. 1. Study area and location of parks.

of 5 min within a park, determined using Global Navigation Satellite System (GNSS) signal data (Akiyama et al., 2013). In total, the study collected 5,967,384 mobile records from 330,404 unique users throughout 2012, with positional accuracy within 5 to10 meters.

For the estimation of visitors' home locations, we analyzed regular overnight GNSS signals, a consistent overnight presence between 8 p.m. and 7 a.m. over a minimum of four nights a week for more than half the year was indicative of a residential location. This identification facilitated the estimation of travel distance from home to park for each visitor. Table 1 displays a sample of the mobile phone data, and Table 2 provides a summary of the observations across different park sizes and time categories. National parks recorded the highest number of observations in most categories, whereas city block and neighborhood parks had fewer observations, especially during weekends and holidays.

3.3. Analytic methods

Fig. 2 delineates the analytical framework of the study, encompassing a four-part workflow: (1) Identifying visitation indicator (2) Atypical parks detection, (3) Parks classification, and (4) Analysis of disparities in trip distances.

Table 1
Sample of mobile phone dataset.

Use ID	Date	Coordinates set	Time set	Travel distance
847	1/5	139.722XXX:35.631XXX | 139.723XXX:35.631XXX	32170 | 32473	17.536
1356	1/5	139.723XXX:35.632XXX | 139.724XXX:35.632XXX	63769 | 64371	1.202
592	4/5	139.724XXX:35.632XXX | 139.723XXX:35.632XXX	50623 | 50925	1.202
8690	7/5	139.723XXX:35.632XXX | 139.724XXX:35.632XXX	44416 | 45919	3.286
1312	7/5	139.721XXX:35.632XXX | 139.724XXX:35.632XXX	63654 | 64257	26.525
1293	8/5	139.723XXX:35.632XXX | 139.722XXX:35.632XXX	47473 | 51681	4.609

3.3.1. Identifying visitation indicators

For each park, feature extraction was conducted on the mobile phone data to ascertain visitor behaviors. Four primary activity metrics were utilized: visit frequency, calculated by averaging the number of visits per unique user ID; stay duration, representing the average time spent in the park measured in hours; average number of movements, quantifying the average movements within the park per visit; and average movement distance, calculating the average distance covered per visit. (Ren & Guan, 2022). These variables were derived by aggregating the data points associated with unique visitor IDs to quantify park usage. The median of these aggregated values was then used to characterize the typical visitor behavior for each park. Temporal aspects were also incorporated into the analysis of these variables to capture the dynamics of park visitation under different temporal conditions. This distinguished between daytime (6 a.m.–6 p.m.) and nighttime (6 p.m.–6 a.m.) periods, as well as between weekdays and weekends. Seasonal variations were systematically categorized into spring (March–May), summer (June–August), fall (September–November), and winter (December–February), in addition to considering holiday periods. A comprehensive set of 40 spatiotemporal visitation variables was extracted.

Principal component analysis (PCA) is then utilized to refine the feature set from the visitation variables previously identified. PCA serves as an effective technique for dimensionality reduction (Maćkiewicz & Ratajczak, 1993), crucial for the outlier detection and clustering stages that follow. This statistical method transforms a dataset defined by possibly correlated variables into a set of linearly uncorrelated variables, termed principal components. The adoption of PCA is predicated on its capacity to distill data into fewer dimensions, thereby exposing the intrinsic patterns of visitation behaviors. These principal components capture the majority of the information contained within the original dataset, greatly simplifying data analysis without substantial loss of information. By concentrating on principal components that account for the bulk of the variation, the analysis can focus on the most impactful factors influencing park visitation.

3.3.2. Atypical parks detection

The Isolation Forest method was utilized to identify parks with atypical visitation patterns, followed by T-tests to determine which

Table 2
Distribution summary of mobile records across park size categories and time periods.

	Size-based Categories	City block park	Neighborhood park	District park	Comprehensive park	National park
	Total number of mobile records	309,249	393,314	1,015,329	3,013,925	1,235,567
	Overall Average	10,308	2,850	13,017	65,520	123,557
	Overall Std. Dev	31,121	5,975	33,620	115,521	144,996
	Overall Median	495	1,323	5,044	27,708	76,007
Day and Night						
	Daytime Average	5,723	1,925	9,231	42,987	93,264
	Daytime Std. Dev	17,446	4,181	25,456	83,269	108,774
	Daytime Median	309	926	3,679	16,456	56,599
	Nighttime Average	4,586	925	3,786	22,534	30,293
	Nighttime Std. Dev	13,733	1,931	8,370	37,247	37,108
	Nighttime Median	178	392	1,429	10,247	20,976
Weekday and Weekend						
	Weekday Average	9,041	2,401	11,254	54,960	100,633
	Weekday Std. Dev	27,391	4,932	30,650	94,183	117,934
	Weekday Median	397	1,105	4,445	23,461	64,496
	Weekend Average	1,268	449	1,763	10,560	22,924
	Weekend Std. Dev	3,736	1,171	3,313	21,718	27,262
	Weekend Median	80	217	795	4,229	11,693
Seasonal						
	Spring Average	2,892	823	3,845	18,580	38,497
	Spring Std. Dev	8,729	1,708	9,497	32,056	45,046
	Spring Median	140	400	1,500	7,582	24,571
	Summer Average	2,551	700	3,207	17,616	29,503
	Summer Std. Dev	7,636	1,486	8,561	35,962	34,226
	Summer Median	101	335	1,231	7,237	18,907
	Fall Average	2,279	634	2,873	14,215	26,558
	Fall Std. Dev	6,868	1,311	7,488	23,187	32,824
	Fall Median	116	317	1,196	6,899	15,991
	Winter Average	2,586	693	3,092	15,109	28,999
	Winter Std. Dev	7,894	1,506	8,109	24,908	33,263
	Winter Median	112	303	1,189	6,369	16,813
Holiday						
	Holiday Average	425	161	586	3,807	8,684
	Holiday Std. Dev	1,268	475	1,173	8,287	10,693
	Holiday Median	20	68	274	1,492	4,328

visitation indicators differ significantly between typical and atypical groups. The Isolation Forest is an ensemble-based, unsupervised learning algorithm that operates under the premise that anomalies are few and different in the data space, thus they are easier to 'isolate' compared to normal points (Liu et al., 2008). It constructs numerous binary trees, or clusters, from the dataset, and anomalies are expected to have shorter path lengths on these trees. The algorithm evaluates each data point, assigning an anomaly score that reflects its probability of being an outlier. This score is then used to discern the degree of atypicality of park visitation patterns, separating out the outliers for focused analysis.

To validate the findings from the Isolation Forest, T-tests were employed as a statistical method to compare the average of visitation indicators between the identified atypical and typical park groups. This test assesses whether the observed differences are statistically significant, thereby confirming the distinctiveness of the atypical visitation patterns. The combination of Isolation Forest and T-test methodologies in this context provides a robust approach to distinguishing atypical parks. Isolation Forest effectively isolates and flags potential anomalies, while T-tests offer a rigorous statistical basis to confirm the relevance and significance of the differences in visitation patterns, thereby challenging the categorization based on park size.

3.3.3. Parks classification with visitation indicator

We employed an optimization model that integrates multiple clustering algorithms drawn from the scikit-learn package. This model integrates a suite of algorithms, K-means, Mini Batch K-means, Gaussian Mixture, Bisecting K-means, Agglomerative Clustering, and Spectral Clustering, chosen for their robust performance in unsupervised learning scenarios (Rodriguez et al., 2019). These algorithms excel in the categorization of unlabeled datasets into distinct and meaningful clusters, making them especially suitable for discerning complex visitation patterns within urban park data. This study tests different combinations of visitation indicators, ranging from individual pairs to the entire set, to determine the most significant contributors to park visitation profiles. The application of multiple clustering methods is a strategic choice designed to harness their combined strengths in identifying patterns within complex datasets. This multifaceted approach ensures a comprehensive and nuanced classification of parks, with the Silhouette Score serving as the key metric for assessing the quality of clustering (Shahapure & Nicholas, 2020). This score measures the similarity of an object to its own cluster compared to others, helping to determine the optimal cluster count and the most appropriate clustering model for the dataset at hand. High silhouette scores indicate a clustering outcome that truly reflects the underlying visitation behaviors, thus affirming the validity of our research approach.

3.3.4. Analyzing disparities in trip distances to parks

To gauge disparities in trip distances to parks, we calculate the Euclidean distance from home to park for each visit (Li et al., 2021; Schindler et al., 2022). The median travel distance served as a key indicator for each park, offering a robust measure of central tendency that reduces the impact of extreme values. To enhance the analysis, we factored in temporal variations, distinguishing between daytime and nighttime, weekdays and weekends, and across different seasons and holidays. The comprehensive assessment facilitates a thorough examination of the dynamic patterns of trip distances, elucidating variations influenced by time, day, and season.

To evaluate the distribution equality of trip distances across different park categories, including both size-based groups and the proposed visitation-driven classification, we utilized the Gini coefficient. The Gini coefficient is a well-established statistical measure of distribution that has been widely applied in economic studies to evaluate income disparity (Gastwirth, 1972). In the context of trip distances disparities, it

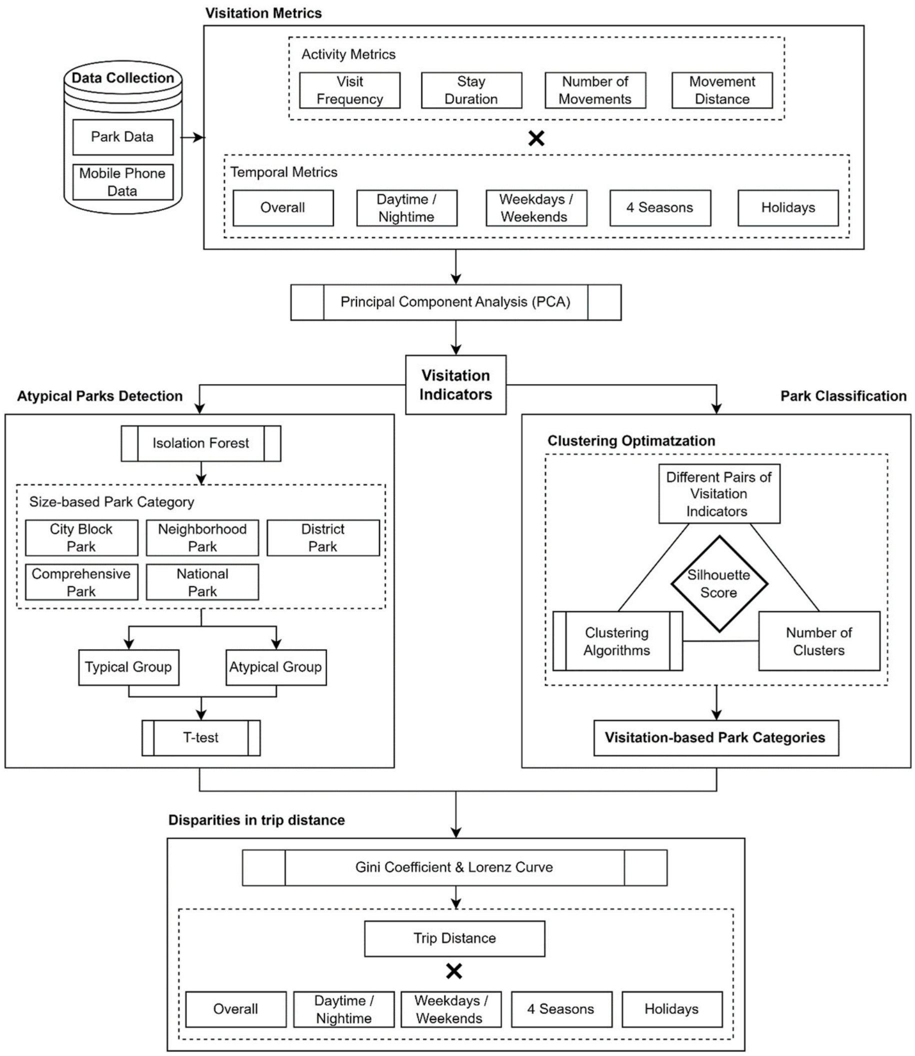

Fig. 2. Research framework.

serves as a quantifier of the inequality present within the distribution of travel distances. The Lorenz curve, employed alongside the Gini coefficient, visually represents this distribution by plotting the cumulative percentage of parks against the cumulative percentage of travel distances. This combination of measures allows for a detailed analysis of trip distance disparities, highlighting potential inequities within and between park groups.

4. Results

4.1. Summary statistics and PCA

Table 3 provides a comparison of park visitation metrics across five park size categories. The overall visit frequency exhibits an increasing trend with park size, with city block parks recording an average frequency of 1.183 and national parks at 2.000. However, the overall stay duration does not follow this trend. Comprehensive parks report the longest average stay duration of 0.700, which is higher than that of national parks at 0.583. In the metrics of overall movement count and distance, there is an increase as park size grows. National parks lead in these metrics with an average count of 1.404 and a distance of 201.279.

Temporal factors show that nighttime and weekdays generally see higher visit frequencies and stay durations across all park size categories except national park. In terms of seasonal variations, the fall season sees a noticeable increase in movement distance. This is particularly true for city block parks (214.459) and national parks (205.912). The holiday metrics show similar visit frequencies to regular weekdays, but with longer stay durations and higher movement counts. This trend is most evident in larger parks, such as national parks, suggesting these may be popular holiday destinations. However, the movement distance does not consistently increase with park size. For example, the overall movement distance is highest for national parks, but comprehensive parks report a lower average movement distance (186.563) than both the city block parks (197.653) and neighborhood parks (196.433). Similarly, the nighttime stay duration is highest for city block parks (2.233), contradicting the notion that larger parks have longer stay durations. Moreover, the stay duration on weekends shows an unexpected trend. The city block parks have the highest average stay duration (1.024) while national parks, despite having the highest overall visit frequency, report a lower average stay duration of 0.608. Supplementary Material Table S2 provides a comprehensive statistical analysis of all parks' visitation patterns.

PCA significantly reduced the data dimensionality from 40 to four principal components (PC), accounting for approximately 71.9% of the total variance. Supplementary Material Table S3 presents the detailed PC loadings for each park visitation metrics. The first PC (PC1) explains 28.6% of the variance. It predominantly measures a visitor's activity level in the park, with significant positive loadings on overall, daytime, and weekday visit frequency and movement count. This component suggests that higher activity levels are associated with frequent visits

Table 3
Summary statistics of visitation metrics by size-based park category (average and standard deviation).

	Visitation metrics	City block park	Neighborhood park	District park	Comprehensive park	National park
	Overall VF	1.183 (0.382)	1.147 (0.350)	1.308 (0.465)	1.609 (0.537)	2.000 (0.471)
	Overall SD	0.609 (0.211)	0.593 (0.315)	0.628 (0.301)	0.700 (0.311)	0.583 (0.208)
	Overall MC	1.100 (0.201)	1.158 (0.255)	1.243 (0.258)	1.258 (0.264)	1.404 (0.238)
	Overall MD	197.653 (41.839)	196.433 (35.977)	196.452 (34.571)	186.563 (29.145)	201.279 (42.171)
Day and Night						
	Daytime VF	1.100 (0.275)	1.118 (0.323)	1.250 (0.432)	1.543 (0.546)	2.000 (0.471)
	Daytime SD	0.581 (0.219)	0.584 (0.317)	0.643 (0.318)	0.698 (0.319)	0.577 (0.199)
	Daytime MC	1.124 (0.276)	1.196 (0.294)	1.294 (0.295)	1.31 (0.297)	1.455 (0.244)
	Daytime MD	202.684 (42.688)	198.918 (34.48)	204.959 (30.477)	195.971 (28.760)	206.064 (40.136)
	Nighttime VF	1.317 (0.533)	1.368 (0.535)	1.365 (0.52)	1.413 (0.498)	1.200 (0.422)
	Nighttime SD	2.233 (2.287)	1.431 (1.972)	0.894 (1.164)	0.807 (0.482)	0.560 (0.320)
	Nighttime MC	1.050 (0.103)	1.053 (0.126)	1.044 (0.138)	1.070 (0.156)	1.096 (0.129)
	Nighttime MD	197.771 (48.341)	195.850 (49.605)	177.309 (45.382)	164.803 (30.507)	176.422 (49.228)
Weekday and Weekend						
	Weekday VF	1.133 (0.346)	1.140 (0.348)	1.237 (0.424)	1.533 (0.542)	1.700 (0.483)
	Weekday SD	0.636 (0.230)	0.597 (0.314)	0.607 (0.282)	0.706 (0.309)	0.578 (0.190)
	Weekday MC	1.109 (0.254)	1.128 (0.225)	1.178 (0.226)	1.211 (0.257)	1.364 (0.221)
	Weekday MD	198.107 (43.243)	196.954 (36.825)	195.608 (35.32)	184.868 (29.867)	197.215 (45.759)
	Weekend VF	1.033 (0.127)	1.103 (0.305)	1.205 (0.398)	1.478 (0.505)	1.900 (0.316)
	Weekend SD	1.024 (0.648)	0.837 (0.700)	0.780 (0.454)	0.830 (0.363)	0.608 (0.287)
	Weekend MC	1.140 (0.211)	1.305 (0.357)	1.344 (0.331)	1.371 (0.313)	1.493 (0.284)
	Weekend MD	198.529 (34.909)	195.177 (37.701)	197.814 (35.736)	194.483 (28.858)	217.884 (32.227)
Seasonal						
	Spring VF	1.167 (0.379)	1.169 (0.371)	1.333 (0.474)	1.543 (0.546)	1.900 (0.316)
	Spring SD	0.742 (0.323)	0.655 (0.385)	0.662 (0.307)	0.716 (0.305)	0.557 (0.216)
	Spring MC	1.147 (0.271)	1.166 (0.26)	1.274 (0.294)	1.258 (0.264)	1.429 (0.215)
	Spring MD	196.079 (51.126)	194.99 (37.823)	196.224 (33.931)	184.107 (31.237)	201.015 (41.97)
	Summer VF	1.200 (0.385)	1.136 (0.341)	1.167 (0.375)	1.435 (0.544)	1.700 (0.483)
	Summer SD	0.813 (0.421)	0.685 (0.421)	0.709 (0.359)	0.816 (0.347)	0.669 (0.261)
	Summer MC	1.088 (0.190)	1.135 (0.226)	1.176 (0.249)	1.213 (0.297)	1.384 (0.264)
	Summer MD	192.970 (45.276)	200.509 (39.574)	196.019 (36.955)	189.392 (29.912)	202.458 (39.357)
	Fall VF	1.183 (0.425)	1.173 (0.377)	1.231 (0.416)	1.457 (0.504)	1.800 (0.422)
	Fall SD	0.708 (0.316)	0.707 (0.408)	0.676 (0.335)	0.722 (0.293)	0.646 (0.253)
	Fall MC	1.128 (0.224)	1.164 (0.272)	1.179 (0.248)	1.243 (0.287)	1.373 (0.233)
	Fall MD	214.459 (55.851)	203.783 (38.364)	203.189 (37.538)	194.725 (26.164)	205.912 (38.949)
	Winter VF	1.117 (0.313)	1.143 (0.344)	1.199 (0.398)	1.391 (0.493)	1.600 (0.516)
	Winter SD	0.895 (0.646)	0.708 (0.481)	0.621 (0.322)	0.692 (0.281)	0.531 (0.203)
	Winter MC	1.065 (0.174)	1.123 (0.219)	1.136 (0.206)	1.144 (0.196)	1.229 (0.208)
	Winter MD	200.353 (47.999)	192.136 (39.844)	189.734 (38.451)	177.474 (32.490)	194.558 (50.266)
Holiday						
	Holiday VF	1.100 (0.305)	1.099 (0.332)	1.122 (0.324)	1.391 (0.493)	1.600 (0.516)
	Holiday SD	1.864 (1.902)	1.144 (1.324)	0.884 (0.596)	0.961 (0.462)	0.605 (0.290)
	Holiday MC	1.183 (0.358)	1.248 (0.362)	1.302 (0.332)	1.347 (0.328)	1.425 (0.262)
	Holiday MD	203.131 (71.368)	195.993 (48.437)	198.39 (41.486)	194.38 (32.315)	219.960 (30.065)

Note: VF, Visit Frequency; SD, Stay Duration; MC, Movement Count; MD, Movement Distance.

and movement within the park during daytime and weekdays, we propose to name "activity intensity". The second PC (PC2), accounts for 21.4% of the variance. It's primarily associated with negative loadings on the movement distance and stay duration variables during all time periods except nighttime and holidays. PC2 is proposed to serve as the "utilization efficiency" indicator. The third PC (PC3) encapsulates 12.6% of the variance. It's predominately tied to the stay duration variable across all time periods, except nighttime, reflecting the duration of a visitor's park visit, we proposed to term "temporal occupancy". The fourth PC (PC4) captures 9.4% of the variance. It's characterized by strong positive loadings for visit frequency across all time periods, counterbalanced by relatively high negative loadings for movement count across the same periods. This component implies that more frequent park visits are associated with less movement within the park, possibly indicating that repeat visitors may have more focused or purposeful visits, termed "revisit volume".

4.2. Identification of parks with atypical visitation patterns

Isolation Forest algorithm has identified parks with atypical visitation patterns across varying park size categories. Fig. 3 delineates these patterns, depicting atypical parks in red and typical parks in green, providing an insightful visualization of their differences across the four principal visitation indicators. Notably, 3 atypical parks were identified in the city block park category, 14 in neighborhood parks, 8 in district parks, 5 in comprehensive parks, and 1 in national parks. This disparity is further illustrated through pairwise comparisons, emphasizing the stark contrasts in visitation indicators between atypical and typical parks. Furthermore, the three-dimensional plot accentuates the peripheral positioning of atypical parks, underscoring their distinct visitation characteristics.

Despite the lack of discernible spatial patterns for atypical parks, T-test analyses (Table S4) have substantiated significant variances in visitation indicators among the park categories. For city block parks, atypical parks registered a higher activity intensity (3.758), in stark contrast to the typical group's (−1.418). They also exhibited higher utilization efficiency (3.277), compared to the typical group (−0.782). Conversely, the typical group demonstrated a stronger tendency for revisits, with a revisit volume (0.775), much higher than the atypical group (−4.003). In neighborhood parks, atypical parks showed not only increased activity intensity (5.553), but also enhanced temporal occupancy (1.909), suggesting longer stays. This is compared to the typical group's utilization efficiency (0.195) against the atypical group (−1.907). District parks' atypical group outperformed the typical group with a higher activity intensity (3.736). Comprehensive parks' atypical group had a substantial activity intensity (6.793) but less utilization

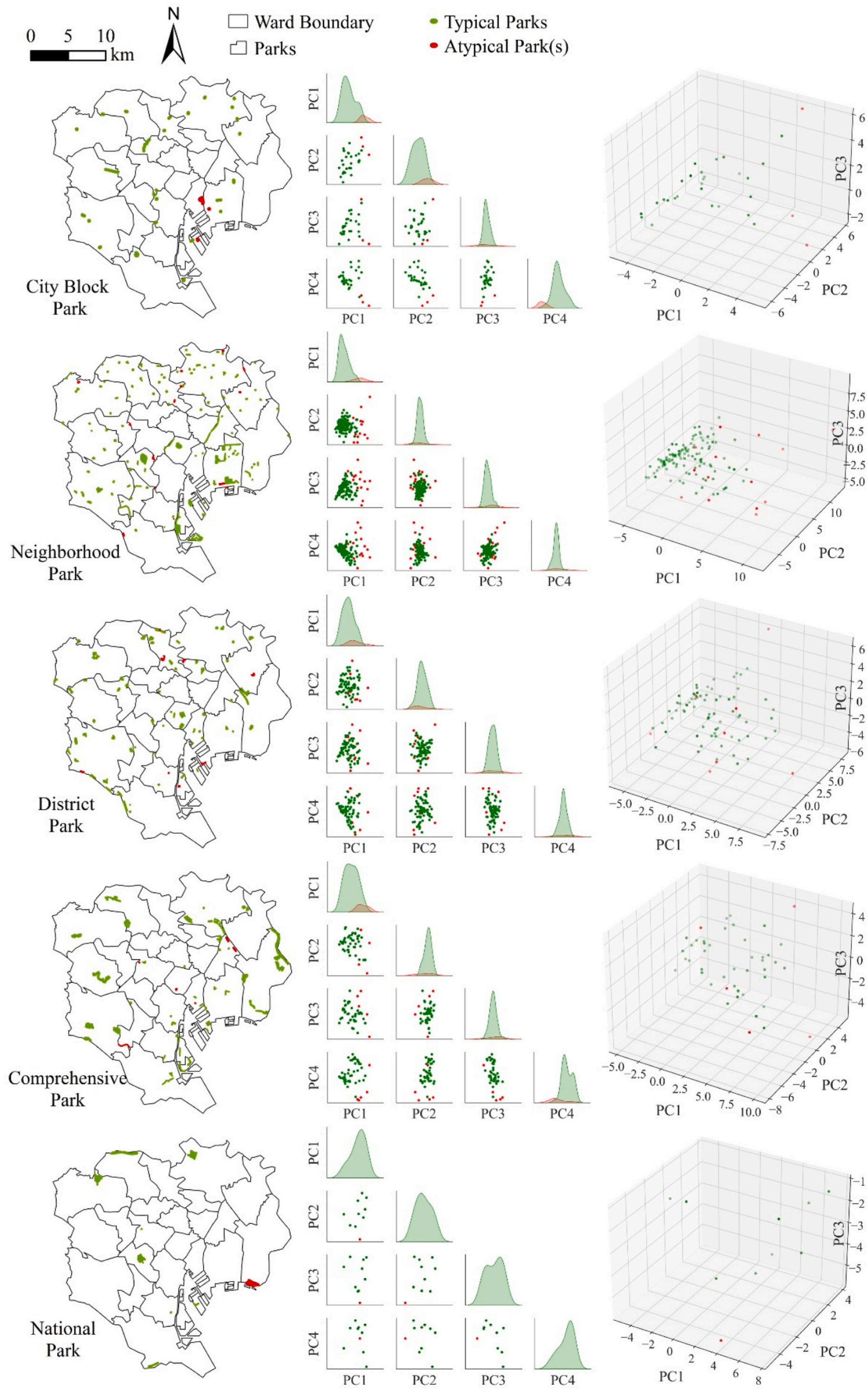

Fig. 3. Parks with atypical visitation patterns in each size-based park category.

efficiency (−1.319). Nonetheless, the typical group in comprehensive parks had a higher revisit volume (1.163). Finally, in national parks, the typical group's revisit volume was 1.933, indicating a more pronounced pattern of repeat visitation, as opposed to the atypical group's −1.586.

4.3. Classification of parks based on visitation patterns

Applying the Mini Batch K-means algorithm to the park visitation indicators, movement intensity (PC1) and visit frequency (PC4), yielded an optimal cluster solution with a silhouette score of 0.520, designating three clusters as the most coherent grouping. The statistical analysis of visitation variables for each park classification is meticulously detailed in Table S5. From the observed visitation patterns across various times, we have delineated three distinct visitation-based park classifications, each named to reflect their unique characteristics.

Everyday leisure parks (Cluster 1), includes 165 parks, exhibit consistent visitation patterns across various seasons. Characterized by brief stay duration (0.508 h), few movement count (1.032), and modest visit frequency (1.070), these parks stand out for considerable high movement distance (205.508 m). Notably, the duration of stays during nighttime and weekends, at 1.063 and 0.714 h respectively, is considerably longer compared to daytime and weekday visits, which average 0.504 and 0.514 h. This pattern underscores the parks' role in providing spaces for a range of activities, reflecting their importance in urban leisure throughout the year.

Seasonal activity parks (Cluster 2) consist of 71 parks characterized by a moderate visit frequency of 1.176, movement distance (183.603 m), and an extended average stay duration of 0.768 h. Activity within these parks intensifies during spring and fall, with movement counts escalating to 1.430 in spring and 1.361 in fall, accompanied by longer stay durations of 0.807 and 0.841 h, respectively. These patterns highlight the parks' appeal as prime locations for diverse seasonal activities, attracting visitors for more prolonged and engaging visits during these peak periods.

Social destination parks (Cluster 3), includes 64 parks distinguished by their high visit frequency (1.992). Coupled with a moderate stay duration and movement count of 0.744 and 1.380, and the shortest movement distance of 181.053 m. Notably, there is a marked decrease in visitation frequency, movement count, and movement distance during nighttime compared to daytime. Specifically, daytime metrics show a visitation frequency of 1.922, stay duration of 0.743 h, movement count of 1.457, and a movement distance of 191.673 m, whereas nighttime visits exhibit a frequency of 1.789, longer stay durations of 1.647 h, a movement count of 1.094, and a reduced movement distance of 162.562 m. These patterns indicate that these parks are preferred spot for frequent social gatherings and stationery activities.

The spatial distribution of the classified parks does not show a significant spatial pattern (Fig. 4a); parks under each category are scattered throughout the study area. Fig. 4b offers a visual interpretation of clusters against the backdrop of activity intensity and revisit volume. Additional analytical dimensions are presented in Fig. 4c and d through a radar chart and a boxplot, which represent the multivariate data of the parks' visitation profiles. From these analyses, it is evident that everyday leisure parks are marked by very low activity intensity, while exhibiting moderate utilization efficiency, temporal occupancy, and revisit volume. Social destination parks have very low revisit volume, but moderate scores for activity intensity, utilization efficiency, and temporal occupancy. Seasonal activity parks, on the other hand, show significantly higher activity intensity and temporal occupancy, marginally higher utilization efficiency, and lower temporal occupancy.

4.4. Proportional analysis and disparities in trip distance to parks

Fig. 5 presents a proportional analysis that highlights the intricate relationship between park visitation types and park size categories within Tokyo's urban landscape. The analysis reveals that neighborhood

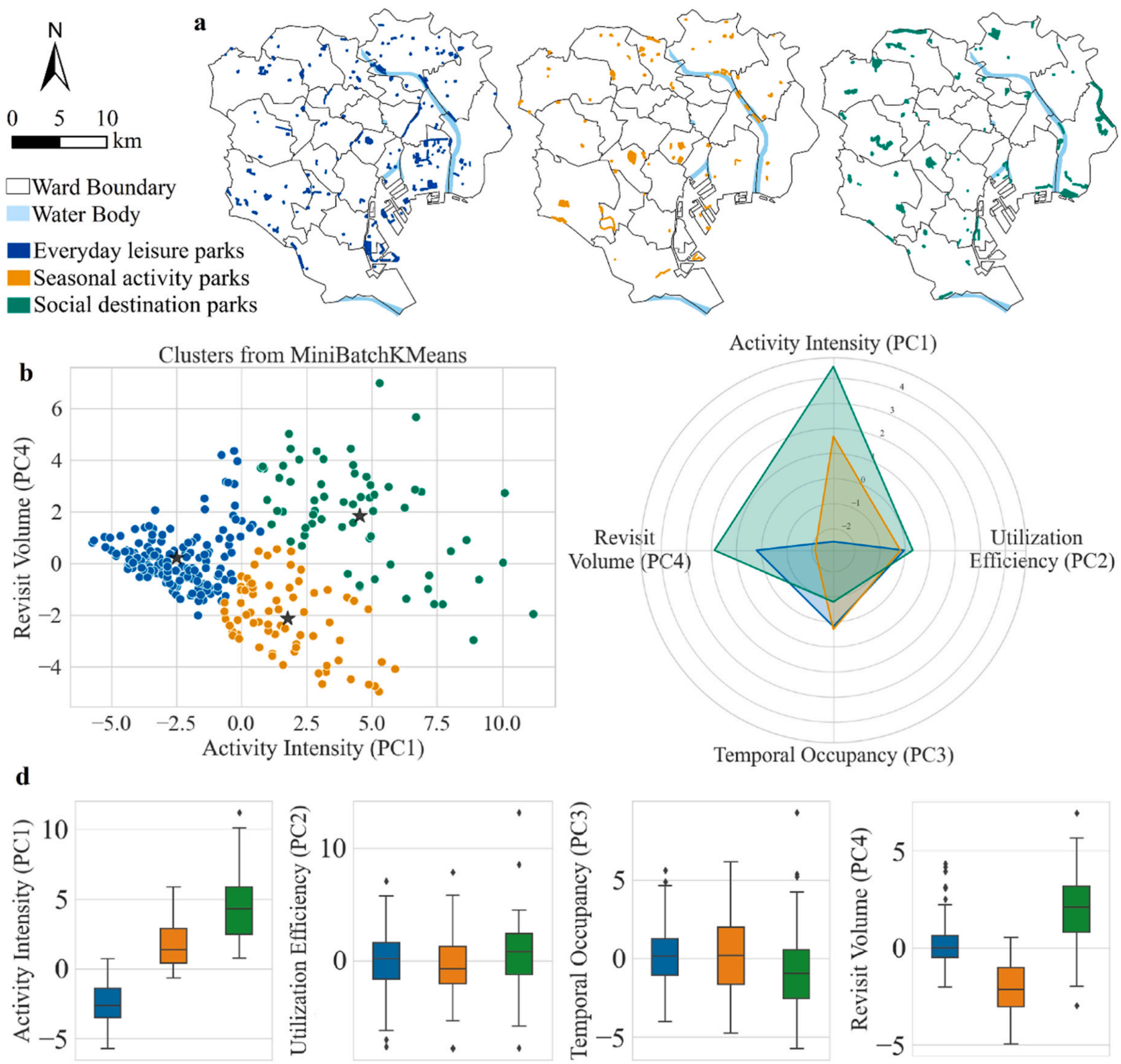

Fig. 4. Clustering result of visitation-based park groups and spatial distribution.

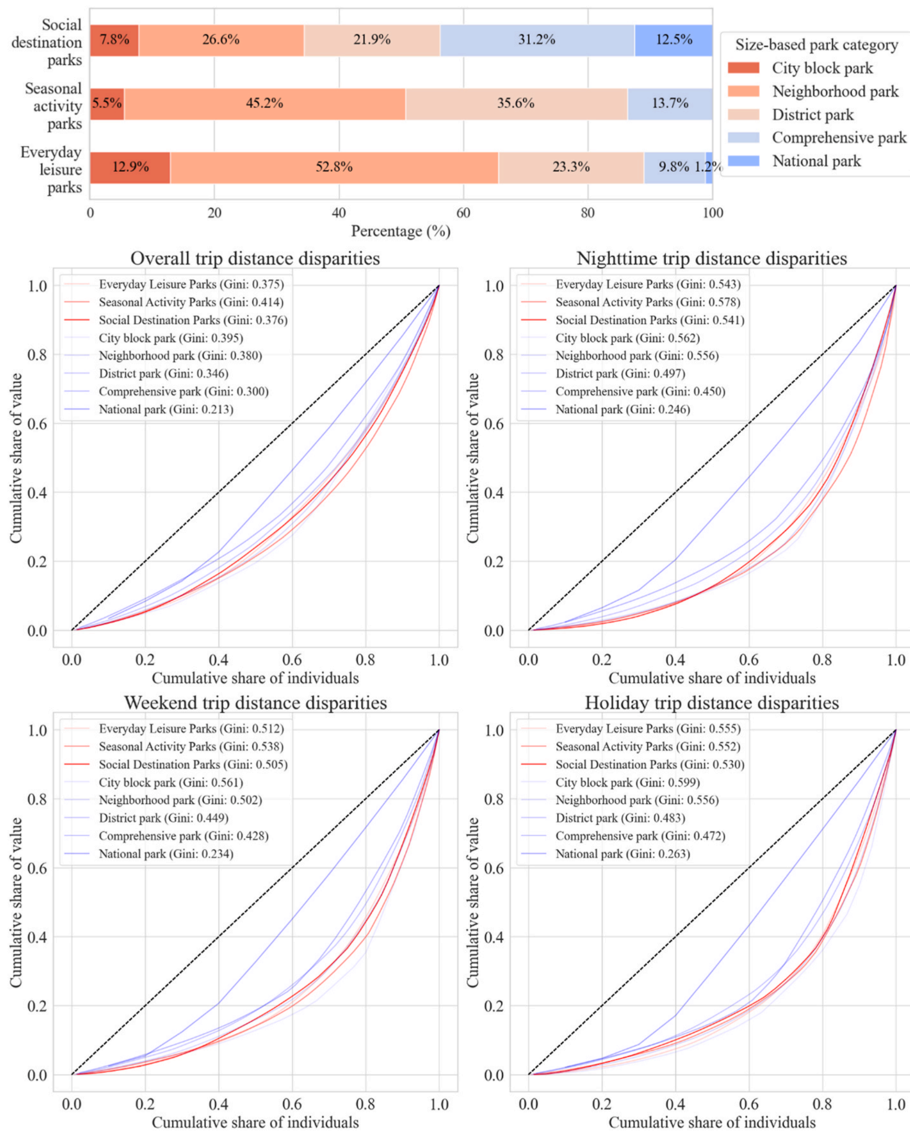

Fig. 5. Proportional analysis and disparities in trip distance to parks.

parks, constituting 53.3% of everyday leisure parks, play a pivotal role in the daily lives of the community. District parks account for 23.0%, city block parks for 12.7%, and comprehensive parks for 9.7%, while national parks, at a minimal 1.2%, are the least represented. Social destination parks are more evenly spread out, with 21.9% in district parks, 31.2% in comprehensive parks, and 26.6% in neighborhood parks, followed by 12.5% in national and 7.8% in city block parks. Seasonal activity parks which make up 43.7% neighborhood parks, and district parks at 36.6%. Comprehensive parks (14.1%) and city block parks (5.6%) also contribute to this category, however there is a notable absence of national parks from this category.

Table S6 presents a summary statistics of trip distances by visitation-based and size-based park categories. The Lorenz curves in Fig. 5 offer insight into trip distance variances across different park categories, indicating marked disparities in trip distances between visitation-based (indicated in red) and size-based park categories (in blue) over various time periods. The Gini index values presented in supplementary materials Table S7 further reflect disparities in trip distance to parks within different categories. Everyday leisure parks show a moderate level of inequality (Gini index: 0.375) but exhibit significant increases during nighttime (0.543), weekends (0.512), and holidays (0.555). Social destination parks mirror this trend with an overall Gini index of 0.376, experiencing a rise in inequality at nighttime (0.541), weekend (0.505), and holidays (0.530). Seasonal activity parks, with a higher Gini index of 0.414, indicate greater inequality, which, as with other categories, peaks during nighttime (0.578), weekends (0.538), and holidays (0.552). Size-based categories reveal that smaller parks such as city block parks and neighborhood parks tend to have higher inequality during holidays (Gini index: 0.599 and 0.556, respectively), while larger parks such as national parks (0.263) demonstrate less inequality.

5. Discussion

5.1. *Unraveling the dynamics of park visitation*

In this study, PCA has identified four key indicators of park visitation behavior: activity intensity, utilization efficiency, temporal occupancy, and revisit volume. These indicators collectively unravel the complex web of factors representing park visitation patterns, emphasizing the necessity for a spatiotemporal approach to understanding park visitation patterns (Guo et al., 2022; Hughes et al., 2009; Lyu & Zhang, 2019).

Activity intensity is characterized by frequent and active movement within the park, particular during daylight and weekday periods, suggests that enhancing the quality of experiences in these specific times may catalyze elevated activity levels (Jeon & Hong, 2015; Ullah et al., 2020). Utilization efficiency underscores the depth of park exploration, highlighting both the extent of visitor movement and the length of visits, which together offer a nuanced view of how parks are utilized for various activities. Temporal occupancy relates to the length of visits, especially during daylight hours, highlight the potential of park features or programs that incentivize prolonged stay (Guo et al., 2022). Revisit volume underscores the behavior of regular visitors, who often engage in specific activities with minimal park traversal (Ren & Guan, 2022; Xiao et al., 2017).

Informed by these indicators, strategic interventions could have the potential to refine park experiences, fostering increased visitation and enhanced visitor satisfaction (Ngesan et al., 2012; Ullah et al., 2020). Employing strategies to enhance spatial utilization and visitor experiences during non-peak periods, such as specialized fitness programs and creative workshops in photography and painting, can effectively redistribute the visitor load and mitigate congestion issues during peak hours (Rigolon, 2016; Wang et al., 2015; Zhang et al., 2021). Customizing experiences and amenities is pivotal in maximizing park occupancy and visitor enjoyment (Guo et al., 2022). A deeper understanding of the motivations and preferences of frequent visitors is vital, not only propels revisit frequencies but also underscores the importance of a bespoke strategy in the operational and promotional planning of park activities (Xiao et al., 2017).

5.2. Refining urban park through visitation-based park classification

This study introduced three novel urban park categories - everyday leisure parks, social destination parks, and seasonal activity parks - based on distinct visitation patterns. This approach transcends size-based classifications, uncover subtle distinctions in park utilization. The detection of parks with atypical visitation patterns across various size categories underscore the limitations of size-based categories in capturing the dynamic usage of parks (Guan et al., 2020; Guo et al., 2022; Ullah et al., 2020). This is also evident in the presence of all size categories within everyday leisure and social destination parks, but the absence of national parks in seasonal activity category, which emphasizes the significant role that smaller, more localized parks in accommodating seasonal activities (Guo et al., 2022; Ibes, 2015; Zhang et al., 2021). These findings provide a foundation for developing strategic recommendations tailored to each distinct visitation-based park category.

Everyday leisure parks, marked by significant movement distances and uniform visitation throughout various seasons, demonstrate a pattern of brief engagements. The distinction in visit durations between nighttime and weekends versus daytime and weekdays underlines the adaptive use of these parks beyond typical working hours, highlighting their critical function within the urban leisure fabric (Tinsley et al., 2010; Tu et al., 2020). In light of these trends, adopting management strategies that adjust to these temporal variations is essential. Practical steps include the adjustment of park management to better suit the fluctuations in usage, with a focus on enhancing visitor experiences during peak times, such as after hours and weekends (Guo et al., 2022; Palliwoda et al., 2017; Ullah et al., 2020; Zhang et al., 2021). Recommendations involve a detailed review and adaptation of park programming and operational hours to match the observed usage patterns more closely. This could include introducing or highlighting activities and amenities that cater to evening and weekend visitors, such as guided night walks or weekend family events, to leverage the parks' full potential as leisure destinations.

Seasonal activity parks, experiencing heightened visitation during spring and fall, suggest a preference for engaging in park activities when temperatures are moderate (Guo et al., 2019; Zhang et al., 2021). Further analysis of the 11 selected parks identified as atypical in each size-based group and within this category (Table S8), reveals some characteristics in common (Guan et al., 2021). They are equipped with youth sports facilities, with some located by rivers or adorned with flora. To cater to these preferences, the development of responsive park programming is essential. Initiatives like organizing sports leagues and hosting seasonal events (Guo et al., 2019; Zhang et al., 2021), particularly cherry blossom viewings, can significantly enhance park usage during peak seasons. Moreover, adopting adaptive landscaping strategies that embrace Tokyo's seasonal diversity, such as planting various seasonal flora and the strategic placement of temporary seating and lighting around sports fields can make these spaces more inviting throughout the year (Talal & Santelmann, 2020). Additionally, maintaining sports facilities ensures they remain functional and appealing for users. Tailoring park amenities and events to the seasonal patterns of park usage not only aligns with community preferences but also maximizes the utility and enjoyment of these urban parks.

Social destination parks, marked by high visitation during the day and a notable reduction at night. An analysis of the selected 13, which identified as atypical within the size-based categories (Table S9), reveals they commonly feature children's playgrounds, expansive lawns, and water bodies, making them attractive as community gathering spots. These elements underscore the parks' potential as key sites for community events and social interaction. (Donahue et al., 2018). To enhance their function as community engagement hubs, prioritizing regular maintenance and volunteer involvement is essential for ensuring the safety and upkeep of key areas, such as playgrounds and communal spaces like barbecue zones. Introducing features like amphitheaters, picnic areas, and versatile open spaces for cultural events, family activities, and community markets can significantly boost their appeal as central social hubs. (Talal & Santelmann, 2020; Tinsley et al., 2010). Maintaining these spaces through consistent cleaning and care is crucial for preserving their welcoming and safe atmosphere. Furthermore, integrating modern amenities such as Wi-Fi hotspots and mobile charging stations addresses the needs of contemporary urban residents. Such improvements can make social destination parks more dynamic and inclusive, strengthening community bonds and enhancing social cohesion.

5.3. Enhancing urban park utilization through data-informed strategies

The investigation into park utilization underscores the need to adopt a classifications model that reflects actual visitation patterns, incorporating with the size-based park classifications (Ren & Guan, 2022). Significant disparities in trip distance to parks have been uncovered in smaller parks and those with high seasonal usage, especially during nights, weekends, and holidays. Allocating resources effectively and developing amenities for peak usage are crucial (Ren & Guan, 2022; Xiao et al., 2017; Zhang et al., 2021). Smaller parks with high-activity level, sometimes overlooked, require focused maintenance and development resources. This could include frequent upkeep of pathways and the installation of energy-efficient, robust lighting for improved safety and usability during evening hours (Ngesan et al., 2012). Additionally, the introduction of flexible, modular amenities to cater to increased visitor numbers during weekends and holidays can significantly enhance the park experience (Donahue et al., 2018; Guan et al., 2020; Talal & Santelmann, 2020; Zhang & Zhou, 2018). Deploying temporary facilities like foldable play areas, portable seating, and seasonal pop-up stalls offering refreshments or activities can meet the fluctuating demand.

Engaging the local community in park management decisions can lead to more user-centered outcomes. This can be achieved through community-led forums, digital platforms for feedback, volunteer-led maintenance initiatives, and community-driven events (Guo et al., 2022; Ullah et al., 2020). Such participatory approaches not only foster a sense of ownership but also ensure that the park's development aligns with the community's needs. Furthermore, an adaptive, data-driven management approach is essential (Guo et al., 2022; Song et al., 2020; Ullah et al., 2020). Regular analysis of park visitation data should inform the adjustment of maintenance schedules, amenity provisions, and event planning. This dynamic strategy, supported by a feedback loop from park visitors through surveys and mobile app interactions, ensures that the parks continually evolve to meet the changing needs of their users.

6. Conclusion and future work

This research has demonstrated the importance of mobile phone derived visitation patterns to enhance our understanding of urban park classification and utilization. Through the analysis of over 5.9 million records across 300 parks in Tokyo, we have identified key visitation indicators that challenge traditional size-based classifications. The refined categorization into everyday leisure parks, social destination parks, and seasonal activity parks provides insights into park usage patterns and visitor preferences. Our findings also reveal significant disparities in trip distance to parks, particularly during nights, weekends, and holidays, with seasonal activity parks and smaller parks experiencing the most pronounced inequalities. This emphasizes the importance of aligning park maintenance and amenity development with empirical usage data to enhance the recreational potential of urban parks. The adoption of such a data-driven approach to park management can lead to more responsive policies that cater to the needs of urban residents.

Despite the contributions of this study, there are certain limitations. The mobile phone data, although comprehensive, may not encompass the entire population, especially considering individuals without mobile devices or those who choose not to share their data. Additionally, the dataset of 330,404 unique visitors, representing only 3.4% of the population in Tokyo's special wards, might not accurately reflect broader visitation patterns. Moreover, concentrating on the 23 special wards of Tokyo, areas characterized by high urban density, may not accurately reflect park visitation patterns in less urbanized and rich natural featured areas, such as Tokyo's Tama Area. Furthermore, the spatial precision of the positioning data, ranging from 5 to 10 meters, may compromise the reliability of park visitation metrics, particularly in determining whether devices near park boundaries are within park limits. Furthermore, the dataset, collected in 2012, does not capture societal change in recent years, such as COVID-19 pandemic, or park redesigns, which have likely influenced park visitation behaviors.

Future research should extend this study to various urban contexts, validating our findings and refining the clustering optimization model. This will help determine if visitation indicators like utilization efficiency and temporal occupancy consistently influence cluster formation across different study sites. In addition, integrating the physical feature of parks with the visitation patterns could lead to a more comprehensive park classification system. Long-term tracking of park visitation changes would also be invaluable for informing urban planning and sustainability initiatives. Future research could further broaden its scope by integrating mobile data with other urban sensing data, such as environmental monitoring and user-generated content, along with census information. This multifaceted approach would yield a richer understanding of park usage and its broader impact on community well-being.

CRediT authorship contribution statement

Yichun Zhou: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Formal analysis, Conceptualization. **ChengHe Guan:** Writing – review & editing, Supervision, Investigation, Funding acquisition, Formal analysis, Data curation. **Longfeng Wu:** Writing – review & editing, Validation, Supervision. **Ying Li:** Supervision, Project administration. **Xuanyi Nie:** Writing – review & editing, Validation. **Jihoon Song:** Writing – review & editing, Validation. **Seung Kyum Kim:** Writing – review & editing, Validation. **Yuki Akiyama:** Writing – review & editing, Validation, Investigation, Data curation.

Declaration of generative AI and AI-assisted technologies in the writing process

Statement: During the preparation of this work, the authors used ChatGPT to improve readability and language. After using the tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgments

We gratefully acknowledge the Shanghai Nature and Health Foundation (Grant No. 20230701 SNHF CH_Guan) for their financial sponsorship. We also thank the Preparation Fund of Shanghai Key Laboratory of Urban Design and Urban Science (Grant No. 10407_Key Lab Preparation Fund) for their support. We are also indebted to the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (Grant No. 20230111 SMEC CH_Guan) and the National Foreign Young Talents Program from the State Administration of Foreign Experts Affairs (Grant No. 10109_ ChengHe Guan _Special Grant) for their backing. Finally, we appreciate the support from the Spring 2022 Climate Change Initiative Seed Grants (Grant No. 22-2126). This collective funding has been instrumental in facilitating this research and we extend our sincere thanks to all contributing bodies.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apgeog.2024.103300.

References

Akiyama, Y., Horanont, T., & Shibasaki, R. (2013). *Time series analysis of the number of visitors in commercial areas using large-scale human flow: 大規模人流データを用いた商業地域における来訪者数の時系列分析*.

Chen, Y., Liu, X., Gao, W., Wang, R. Y., Li, Y., & Tu, W. (2018). Emerging social media data on measuring urban park use. *Urban Forestry and Urban Greening, 31*, 130–141. https://doi.org/10.1016/j.ufug.2018.02.005

Chen, L., Wen, Y., Zhang, L., & Xiang, W.-N. (2015). Studies of thermal comfort and space use in an urban park square in cool and cold seasons in Shanghai. *Building and Environment, 94*, 644–653. https://doi.org/10.1016/j.buildenv.2015.10.020

Construction, B. o. Parks in Tokyo. https://www.kensetsu.metro.tokyo.lg.jp/english/jigyo/park/01.html.

Csomos, G., Borza, E. M., & Farkas, J. Z. (2023). Exploring park visitation trends during the Covid-19 pandemic in Hungary by using mobile device location data. *Scientific Reports, 13*(1), Article 11078. https://doi.org/10.1038/s41598-023-38287-3

Doll, C., Polyakov, M., Pannell, D. J., & Burton, M. P. (2022). Rethinking urban park irrigation under climate change. *Journal of Environmental Management, 314*, Article 115012. https://doi.org/10.1016/j.jenvman.2022.115012

Donahue, M. L., Keeler, B. L., Wood, S. A., Fisher, D. M., Hamstead, Z. A., & McPhearson, T. (2018). Using social media to understand drivers of urban park visitation in the Twin Cities, MN. *Landscape and Urban Planning, 175*, 1–10. https://doi.org/10.1016/j.landurbplan.2018.02.006

Gastwirth, J. L. (1972). The estimation of the Lorenz curve and Gini index. *The Review of Economics and Statistics*, 306–316.

Government, T. M. (2021). TOKYO'S history, geography, and population. https://www.metro.tokyo.lg.jp/english/about/history/history02.html.

Guan, C., Song, J., Keith, M., Akiyama, Y., Shibasaki, R., & Sato, T. (2020). Delineating urban park catchment areas using mobile phone data: A case study of Tokyo. *Computers, Environment and Urban Systems, 81*. https://doi.org/10.1016/j.compenvurbsys.2020.101474

Guan, C., Song, J., Keith, M., Zhang, B., Akiyama, Y., Da, L., Shibasaki, R., & Sato, T. (2021). Seasonal variations of park visitor volume and park service area in Tokyo: A mixed-method approach combining big data and field observations. *Urban Forestry and Urban Greening, 58*. https://doi.org/10.1016/j.ufug.2020.126973

Guo, H., Luo, Z., Li, M., Kong, S., & Jiang, H. (2022). A literature review of big data-based urban park research in visitor dimension. *Land, 11*(6). https://doi.org/10.3390/land11060864

Guo, S., Song, C., Pei, T., Liu, Y., Ma, T., Du, Y., Chen, J., Fan, Z., Tang, X., Peng, Y., & Wang, Y. (2019). Accessibility to urban parks for elderly residents: Perspectives from

mobile phone data. *Landscape and Urban Planning, 191*. https://doi.org/10.1016/j.landurbplan.2019.103642
Hughes, M., Ham, S., & Brown, T. (2009). Influencing park visitor behavior, a belief based approach. *Journal of Park and Recreation Administration, 27*(4), 38–53.
Ibes, D. C. (2015). A multi-dimensional classification and equity analysis of an urban park system: A novel methodology and case study application. *Landscape and Urban Planning, 137*, 122–137. https://doi.org/10.1016/j.landurbplan.2014.12.014
Jeon, J. Y., & Hong, J. Y. (2015). Classification of urban park soundscapes through perceptions of the acoustical environments. *Landscape and Urban Planning, 141*, 100–111. https://doi.org/10.1016/j.landurbplan.2015.05.005
Li, X., Huang, Y., & Ma, X. (2021). Evaluation of the accessible urban public green space at the community-scale with the consideration of temporal accessibility and quality. *Ecological Indicators, 131*. https://doi.org/10.1016/j.ecolind.2021.108231
Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). *Isolation forest. 2008 eighth ieee international conference on data mining*.
Lyu, F., & Zhang, L. (2019). Using multi-source big data to understand the factors affecting urban park use in Wuhan. *Urban Forestry and Urban Greening, 43*. https://doi.org/10.1016/j.ufug.2019.126367
Ma, L., Shi, W., & Wu, L. (2023). Mediating roles of perceptions and visiting behavior in the relationship between urban greenspace accessibility and personal health: Evidence from Lanzhou, China. *Applied Geography, 159*. https://doi.org/10.1016/j.apgeog.2023.103085
Maćkiewicz, A., & Ratajczak, W. (1993). Principal components analysis (PCA). *Computers & Geosciences, 19*(3), 303–342.
Ministry of Land, I. (2006). Transport and tourism. In *Current status of provision of city parks in Japan*. https://www.mlit.go.jp/english/2006/d_c_and_r_develop_bureau/03_parks-and-green/.
Ngesan, M. R., Karim, H. A., & Zubir, S. S. (2012). Human behaviour and activities in relation to shah alam urban park during nighttime. *Procedia - Social and Behavioral Sciences, 68*, 427–438. https://doi.org/10.1016/j.sbspro.2012.12.239
Nielsen, A. B., van den Bosch, M., Maruthaveeran, S., & van den Bosch, C. K. (2013). Species richness in urban parks and its drivers: A review of empirical evidence. *Urban Ecosystems, 17*(1), 305–327. https://doi.org/10.1007/s11252-013-0316-1
Palliwoda, J., Kowarik, I., & von der Lippe, M. (2017). Human-biodiversity interactions in urban parks: The species level matters. *Landscape and Urban Planning, 157*, 394–406. https://doi.org/10.1016/j.landurbplan.2016.09.003
Peters, K., Elands, B., & Buijs, A. (2010). Social interactions in urban parks: Stimulating social cohesion? *Urban Forestry and Urban Greening, 9*(2), 93–100. https://doi.org/10.1016/j.ufug.2009.11.003
Ren, X., & Guan, C. (2022). Evaluating geographic and social inequity of urban parks in Shanghai through mobile phone-derived human activities. *Urban Forestry and Urban Greening, 76*. https://doi.org/10.1016/j.ufug.2022.127709
Rigolon, A. (2016). A complex landscape of inequity in access to urban parks: A literature review. *Landscape and Urban Planning, 153*, 160–169. https://doi.org/10.1016/j.landurbplan.2016.05.017
Rodriguez, M. Z., Comin, C. H., Casanova, D., Bruno, O. M., Amancio, D. R., Costa, L. D. F., & Rodrigues, F. A. (2019). Clustering algorithms: A comparative approach. *PLoS One, 14*(1), Article e0210236. https://doi.org/10.1371/journal.pone.0210236
Schindler, M., Le Texier, M., & Caruso, G. (2022). How far do people travel to use urban green space? A comparison of three European cities. *Applied Geography, 141*. https://doi.org/10.1016/j.apgeog.2022.102673
Shahapure, K. R., & Nicholas, C. (2020). Cluster quality analysis using silhouette score. In *2020 IEEE 7th international conference on data science and advanced analytics (DSAA)*.
Shahtahmassebi, A. R., Li, C., Fan, Y., Wu, Y., lin, Y., Gan, M., Wang, K., Malik, A., & Blackburn, G. A. (2021). Remote sensing of urban green spaces: A review. *Urban Forestry and Urban Greening, 57*. https://doi.org/10.1016/j.ufug.2020.126946
Song, X. P., Richards, D. R., He, P., & Tan, P. Y. (2020). 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*. https://doi.org/10.1016/j.landurbplan.2020.103908
Talal, M. L., & Santelmann, M. V. (2020). Vegetation management for urban park visitors: A mixed methods approach in portland, Oregon. *Ecological Applications, 30*(4), Article e02079. https://doi.org/10.1002/eap.2079
Tinsley, H. E. A., Tinsley, D. J., & Croskeys, C. E. (2010). Park usage, social milieu, and psychosocial benefits of park use reported by older urban park users from four ethnic groups. *Leisure Sciences, 24*(2), 199–218. https://doi.org/10.1080/01490400252900158
Tu, X., Huang, G., Wu, J., & Guo, X. (2020). How do travel distance and park size influence urban park visits? *Urban Forestry and Urban Greening, 52*. https://doi.org/10.1016/j.ufug.2020.126689
Ullah, H., Wan, W., Haidery, S. A., Khan, N. U., Ebrahimpour, Z., & Muzahid, A. A. M. (2020). Spatiotemporal patterns of visitors in urban green parks by mining social media big data based upon WHO reports. *IEEE Access, 8*, 39197–39211. https://doi.org/10.1109/access.2020.2973177
Wang, D., Brown, G., Liu, Y., & Mateo-Babiano, I. (2015). A comparison of perceived and geographic access to predict urban park use. *Cities, 42*, 85–96. https://doi.org/10.1016/j.cities.2014.10.003
Whiting, J. W., Larson, L. R., & Green, G. T. (2012). Monitoring visitation in Georgia state parks using the system for observing play and recreation in communities (SOPARC). *Journal of Park and Recreation Administration, 30*(4), 21–37.
Xiao, Y., Wang, Z., Li, Z., & Tang, Z. (2017). An assessment of urban park access in Shanghai – Implications for the social equity in urban China. *Landscape and Urban Planning, 157*, 383–393. https://doi.org/10.1016/j.landurbplan.2016.08.007
Zhang, R., Sun, F., Shen, Y., Peng, S., & Che, Y. (2021). Accessibility of urban park benefits with different spatial coverage: Spatial and social inequity. *Applied Geography, 135*. https://doi.org/10.1016/j.apgeog.2021.102555
Zhang, J., & Tan, P. Y. (2019). Demand for parks and perceived accessibility as key determinants of urban park use behavior. *Urban Forestry and Urban Greening, 44*. https://doi.org/10.1016/j.ufug.2019.126420
Zhang, S., & Zhou, W. (2018). Recreational visits to urban parks and factors affecting park visits: Evidence from geotagged social media data. *Landscape and Urban Planning, 180*, 27–35. https://doi.org/10.1016/j.landurbplan.2018.08.004