Tourism Management 107 (2025) 105064

Contents lists available at ScienceDirect

Tourism Management

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

Spatial concentration of intra-urban tourist activities and inter-group differences between Asian, European and North American travelers in Korean cities

Jingyi Cheng [a], Yang Xu [a,b,*], Izzy Yi Jian [c], Mengyao Ren [a], Sangwon Park [d]

[a] Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China
[b] Research Centre for Digital Transformation of Tourism (RCDTT), The Hong Kong Polytechnic University, Hong Kong, China
[c] Department of Social Sciences and Policy Studies, The Education University of Hong Kong, Hong Kong, China
[d] College of Hotel and Tourism Management, Kyung Hee University, Republic of Korea

ARTICLE INFO

Keywords:
Spatial concentration
Tourist activity
Travel behavior
Mobile phone data
Tourism big data
Urban tourism
Overtourism

ABSTRACT

With increasing concerns about overtourism and its associated impacts, there remains an imperative need to understand the spatial concentration of tourist activities within destinations. This study demonstrates that mobile phone data can be leveraged to accurately quantify the spatial concentration effect and potential heterogeneity among travelers with different social backgrounds. We introduce an analytical framework and demonstrate its efficacy through a case study of international travelers in three Korean cities (Seoul, Busan and Jeju). Results show that Asian travelers exhibit the highest spatial concentration, followed by European and North Americans, a finding that is consistent across all three cities and different times of the day. The Theil's entropy index effectively portrays spatial similarity of activities among the three groups, revealing that the spatial preferences of European travelers are intermediate between those of Asians and North Americans. The findings underscore the importance of considering travelers' social backgrounds in tourist flow management.

1. Introduction

The spatial concentration of tourist activities refers to the uneven distribution and spatial inequity of tourist visitation (Yang & Wong, 2013). It describes a phenomenon that a small number of areas attract a large number of tourist activities while many others receive minimal or no tourists (Shao et al., 2017). Due to the high concentration of tourism services and attractions, some places are more likely to be areas where tourist activities tend to concentrate, such as islands, coastal areas, and metropolitan areas, as well as historic and central districts within cities (Kotus et al., 2015; Mou et al., 2020; Su et al., 2016; Wu & Carson, 2008). This tendency of concentration may develop dynamically, such as the Matthew effect, which emphasizes the principle of cumulative advantage in tourism systems (Merton, 1968). Over time, well-known attractions may attract more tourists, while less-known attractions will gradually lose appeal. This self-reinforcing cycle may bring heavy tourist pressure on hotspots, which may have adverse impacts on balanced and sustainable tourism development (Gao et al., 2022). Excessive concentration of tourist activities can also trigger the perception of overcrowding, thus diminishing the attractiveness and livability for both residents and tourists (Namberger et al., 2021; Popp, 2012). In addition, previous studies have considered the degree of tourist concentration as an important symptom and indicator of overtourism (Buitrago & Yñiguez, 2021) and have emphasized the need to promote tourist dispersion in both space and time to alleviate this issue World Tourism Organization and NHL Stenden University of Applied Sciences, 2018). However, there is currently a lack of a systematic analytical framework to quantify the spatial concentration of tourist activities within cities.

The spatial concentration of tourist activities can be influenced and shaped by the aggregation of the tourism industry (Majewska, 2015). According to agglomeration theory, when businesses and industries are located in close proximity to one another in geographical space, they can benefit from various external economies of scale, leading to higher productivity and efficiency (Kim et al., 2021; Majewska, 2015). This agglomeration has a positive effect on tourism development, as confirmed in many regions and countries (Kim et al., 2021; Li & Liu,

* Corresponding author. Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China.
E-mail addresses: jingyi.cheng@connect.polyu.hk (J. Cheng), yang.ls.xu@polyu.edu.hk (Y. Xu), izzyjian@eduhk.hk (I.Y. Jian), mengyao.ren@connect.polyu.hk (M. Ren), sangwon.park@khu.ac.kr (S. Park).

https://doi.org/10.1016/j.tourman.2024.105064
Received 15 September 2024; Accepted 22 September 2024
0261-5177/© 2024 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

2022; Zhou et al., 2016), and results in an increasing intensity of tourism activities in specific geographical areas (Yang, 2012). The cumulative attraction theory also explains this phenomenon, suggesting that multiple attractions in a given area will attract more tourists than if those attractions are widely dispersed (Matthews et al., 2018). However, most of the research on tourism agglomeration focuses on the agglomeration of tourism industries and resources rather than that of tourist activities (Majewska, 2015; Yang, 2012; Yang & Wong, 2013). This may be due to the fact that the spatial distribution pattern of tourism industries is relatively static and stable, and the data reflecting the distribution of tourism resources are more readily available. Meanwhile, capturing the spatial concentration effect of tourist activities requires continuous and dynamic observation of tourist activities, which poses a challenge for data acquisition (Edwards et al., 2010). In fact, industrial agglomeration is not equivalent to the concentration of tourism activities, and there may be certain asymmetries or mismatches between supply and demand. Insufficient exploration of the spatial concentration effect from the perspective of tourism activities may hinder our comprehensive understanding of tourism agglomeration.

On the other hand, the interplay between nationality and tourist behavior is an essential research topic in the tourism literature (Lew et al., 2008; Pizam & Sussmann, 1995). For international tourists, individuals from diverse cultural backgrounds may exhibit varying travel interests, preferences, and expectations (You et al., 2000). This phenomenon aligns with Hofstede's cultural dimensions theory, which elucidates how a society's culture influences the values held by its individuals and how these values influence their behavior (Hofstede, 1983). Hofstede's cultural dimensions such as individualism-collectivism and uncertainty avoidance have been proven to be closely related to the heterogeneity of tourist behavior. This heterogeneity has been widely observed in tourist groups between nationalities, especially as past research has highlighted behavioral differences between Eastern and Western tourists (Maoz, 2007; Pizam & Fleischer, 2005; Pizam & Sussmann, 1995; Reisinger & Turner, 2012). For example, Asian tourists prefer to see themselves as group members and choose to travel in a group rather than independently. This can be interpreted as a coping mechanism to mitigate the impact of unfamiliar cultural environments and to maintain a sense of identity and security (Ritter, 1987). This characteristic can also be attributed to the collectivist cultural orientation prevalent in these societies (Pizam & Jeong, 1996). In contrast, Westerners, such as Americans and Italians, are often characterized as more sociable and individualistic, and adaptable to novel surroundings (Pizam & Sussmann, 1995). The American cultural ethos is frequently associated with a love of new things, a desire to be close to nature, intense individualism, and a low degree of uncertainty avoidance (Ritter, 1987). Some studies have also revealed distinctions between Europeans and Americans. For instance, a comparative analysis of 20 behavioral characteristics of tourists revealed that French-American is the most dissimilar pair after pairwise comparison (Pizam & Sussmann, 1995). Previous studies have provided valuable knowledge about behavioral differences among nationality groups in various aspects, including social interactions, activity preferences, and business transactions (Chen & Hsu, 2021; Özdemir & Yolal, 2017). However, the manifestation of this difference in spatial concentration remains unclear.

Emerging big data technologies have recently enabled access to unprecedented spatial and temporal data on tourist mobility patterns (Raun et al., 2016). Mobile phone data, in particular, offers a novel perspective to overcome the shortcoming by capturing tourists' digital footprints at fine-grained spatio-temporal resolutions and extensive spatial coverage (Blondel et al., 2015; Xu et al., 2021). From a data acquisition standpoint, traditional methods (e.g., surveys, questionnaires, or online diaries) face challenges in accurately capturing granular spatial and temporal distributions and flow characteristics of tourist populations at a fine-grained level, rendering much research generalized and lacking precise spatial insights. The application of mobile phone data analysis enables a more comprehensive capture of tourist flows and reveals intricate patterns and behaviors previously difficult to discern. The transformative potential of mobile phone data in delivering granular, precise, and spatially explicit insights has solidified its position as a recognized and prevalent approach in tourism research (Shoval & Ahas, 2016).

To address these gaps, this study aims to investigate the spatial concentration of intra-urban tourist activities using a mobile phone dataset in South Korea. The study covers three popular tourist destinations in South Korea: Seoul, Busan and Jeju Island. The three cities differ in their geographical location, city size, and tourism resources, and have a strong appeal to international travelers, making them suitable as study cases. Three major nationality groups of international travelers are identified: Asian, European, and North American travelers. The key research questions of this study include: How does the spatial distribution of intra-urban tourist activities vary among Asian, European and North American travelers? How does the degree of spatial concentration of tourist activities vary among these three nationality groups? What are the differences in spatial similarity of their tourist activities?

The major contributions of this research are as follows. First, by analyzing the movement trajectory of international travelers, we quantify the degree of spatial concentration based on the distribution of visitation frequency. Three types of tourist activities are identified, namely accommodation, daytime, and nighttime activities. These three activities cover important aspects during the travel. This analysis allows the spatial concentration degree of different nationality groups to be comparable across different context of tourist activities and cities. Second, we introduce the segregation index (Theil's Entropy) for evaluating the spatial similarity of tourist activities. It could reveal the unique spatial preferences of Asian, European, and North American travelers, and the similarity of their preferences across different activities. Third, we conduct a cross-city analysis of spatial distribution, concentration, and similarity of tourist activities, as well as summarize the general patterns and heterogeneities of different nationality groups in tourist activities.

We hope this study can enrich the comprehension of tourist intensity and how tourists use cities in space and time. The findings of this study are anticipated to illuminate areas that may be either overlooked or favored by tourists. This can help pinpoint regions experiencing excessive tourist pressure and identify locations where such flows could be redirected, thus creating new opportunities for the local tourism sector and economy. Furthermore, this research framework can be applied to various subgroups, such as travelers of different ages, genders, and socio-economic backgrounds, to better understand their spatial preferences and activity concentrations in tourism destinations.

2. Methodology

2.1. Study area and dataset

Seoul, Busan and Jeju Island are popular tourist destinations in South Korea with rich tourism resources (Fig. 1). Seoul (the Seoul Special City) is the capital and largest metropolis of South Korea. It has a population of 9.9 million and covers an area of 605.21 km^2, drawing tourists from around the globe with its rich historical heritage, unique Korean cuisine, and diverse shopping venues. Busan is the second-most populous city in South Korea, with a population of over 3.4 million residents and an area of 770.04 km^2. Renowned for the impressive collection of museums, historic landmarks and seaside scenery, Busan is the economic, cultural, and educational center of southeastern South Korea. Jeju Island is the largest island and largest tourist destination in the country. It has an area of 1847.2 km^2 and a population of more than 0.6 million. These three cities differ in geographic locations, city sizes, and tourism resources, making them ideal for studying the spatial concentration of tourist activities and inter-group differences. As the aim of this study is to explore potential similarities or disparities in tourist activity patterns across

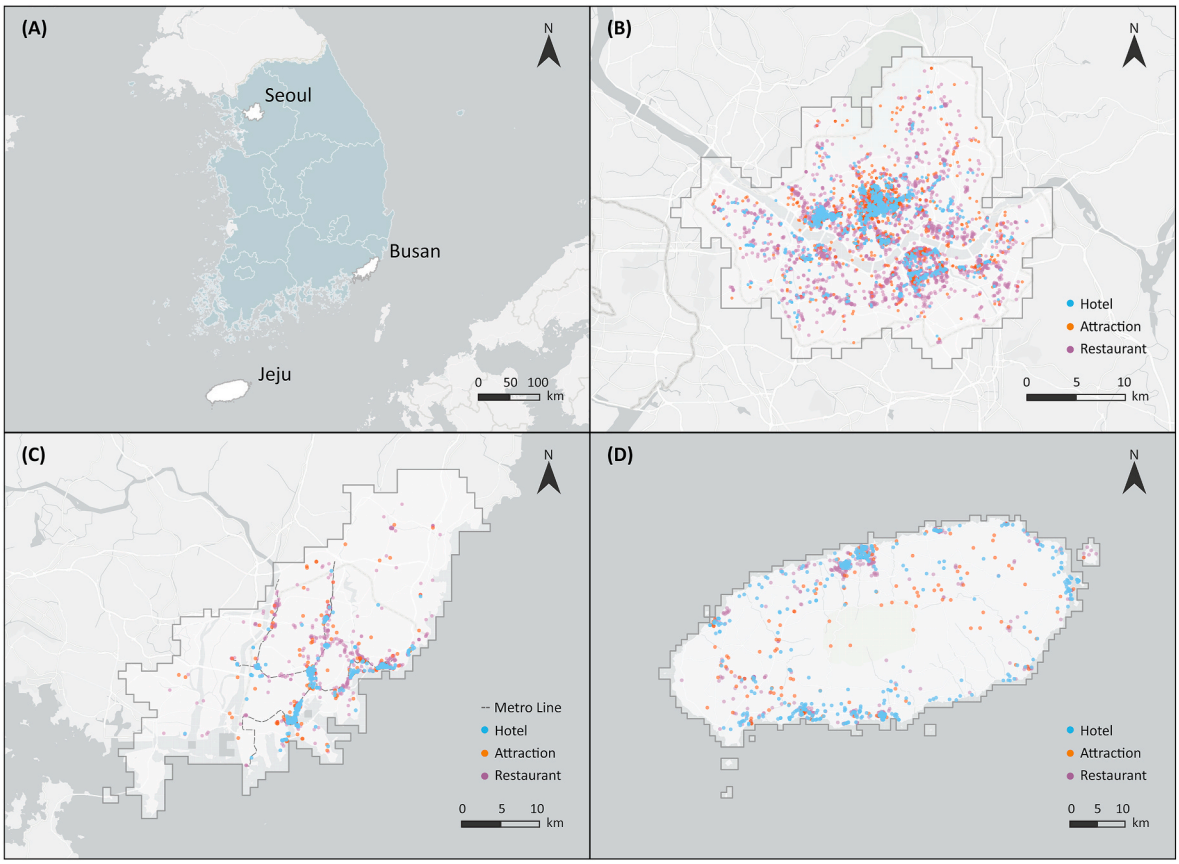


Fig. 1. (A) Locations of Seoul, Busan and Jeju in South Korea. (B–D) The spatial distribution of tourism resources in (B) Seoul, (C) Busan and (D) Jeju.

different city types, this study is conducted in these three cities. Fig. 1 shows the spatial distribution of tourism resources in Seoul, Busan and Jeju Island, respectively. The tourism resources include three broad categories of hotel, attraction and restaurant. The dataset was downloaded in 2021 from TripAdvisor, which is one of the major global online travel agencies.

This study utilizes an anonymized mobile phone dataset collected by a major cellular operator in South Korea, which holds a 40–50% market share. The dataset contains the footprints of international travelers who visited Seoul, Busan and Jeju between August 1, 2018, and August 15, 2018. In the dataset, a total of 86,561 travelers visited Seoul during the data collection period, while Busan and Jeju received 18,294 and 23,978 travelers, respectively. Note that no special events occurred in these three cities during the data collection period, indicating that the number of travelers recorded can be regarded as a typical volume on regular days.

The dataset tracked the sequence of locations at the individual level and was preprocessed by the data provider to generate estimates of the duration of stay. Each record in the dataset includes the unique ID of the user, nationality information, base station ID that was interacted with during the stay, longitude and latitude of the base station, date, starting time, and ending time that defined the corresponding stay period. To understand the composition of the traveler population, we calculate the number and proportion of travelers of different nationality groups in the dataset (Table 1). Due to South Korea's geographical location and cultural background, the majority of travelers are Asians, followed by Europeans and North Americans. Therefore, we divide international travelers into Asian (AS), European (EU), and North American (NA) groups to measure the behavioral heterogeneity among them.

2.2. Data processing

This section presents the approach for extracting tourist activity information from the mobile phone dataset. We generally follow the method proposed by Xu et al. (2021) for mobile phone data processing. The entire workflow of data processing and analysis is shown in Fig. 2. The initial step involves reconstructing each traveler's movement trajectory and mitigating the impacts of mobile phone signal shifting. An individual's trajectory is reconstructed by reading each of the record of the traveler in a chronological order. Signal drift refers to the condition where the signal of user's cell phone does not always communicate with a fixed base station, but instead drifts between several nearby base stations. To overcome this issue, a set of activity anchor points (an area containing several base stations) is established for each individual. Base

Table 1
Number and proportion of international travelers from each continent in three cities.

City	Statistics	Asian	European	North American	South American	Oceanian	African	Total
Seoul	Number of travelers	65,053	10,527	9974	46	859	102	86,561
	Proportion	75.2%	12.2%	11.5%	0.1%	1.0%	0.1%	100%
Busan	Number of travelers	13,781	2642	1623	7	212	29	18,294
	Proportion	75.3%	14.4%	8.9%	0%	1.2%	0.2%	100%
Jeju	Number of travelers	20,145	3041	700	5	80	7	23,978
	Proportion	84.0%	12.7%	2.9%	0%	0.3%	0%	100%

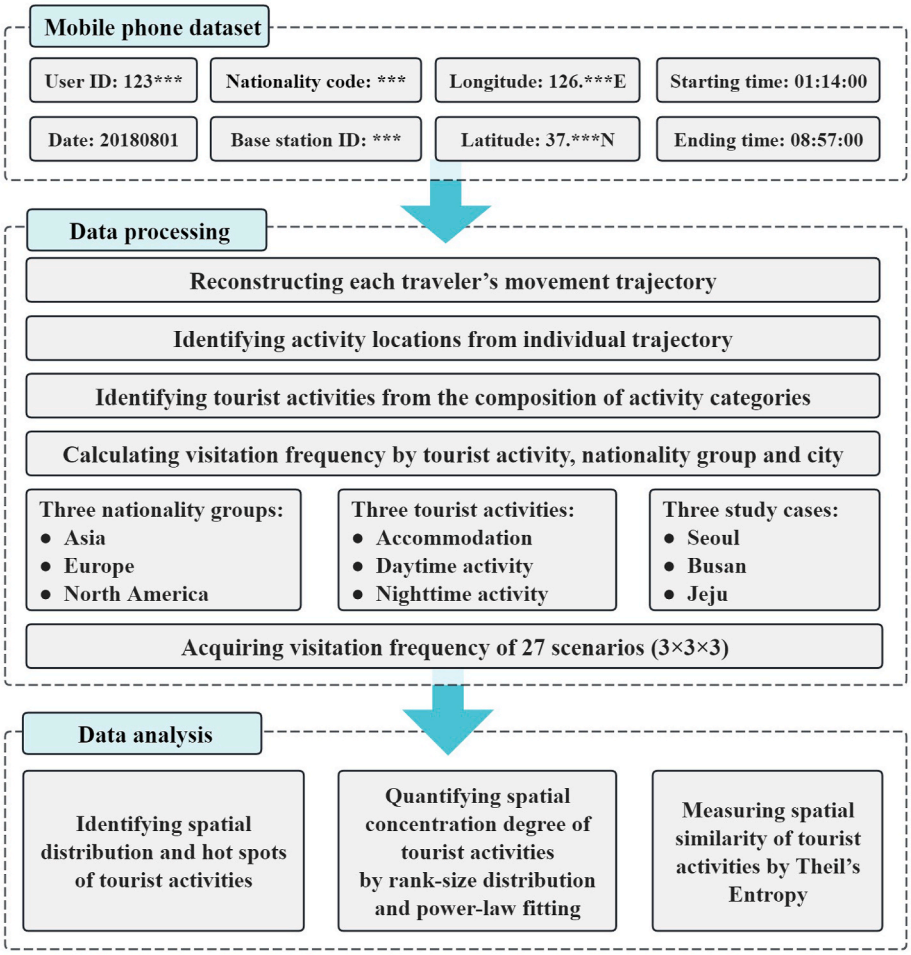


Fig. 2. Research framework of this study.

stations where signal drift has occurred are merged into the same anchor point.

Following the reconstruction of movement trajectories, a crucial task is to identify activity locations from the trajectory. For the anchor point contained in each individual's trajectory, we determine the activity semantics and corresponding locations based on when the stay occurs. We focus on the two categories of anchor points in an individual's trajectory: Nighttime Anchor Point (NAP) and Other Activity Locations (OTHERS). NAP refers to an anchor point where an individual stays most of the time from midnight to 7 a.m. It represents the potential place of accommodation of an individual. OTHERS refers to an anchor point which is not labeled as NAP but is still within the destination. In this case, travelers may be engaged in other activities, such as sightseeing and dining outside where they live.

After identifying activity locations and semantics, we divide a 24-h day into three periods: (a) 00:00–09:00, (b) 09:00–19:00, and (c) 19:00–24:00. This division takes into account the composition of activity categories in three cities, which can be found in the Appendix (Figure A.1). During the first period, we focus on travelers whose activity category is NAP to investigate their accommodation locations, and we refer to their activities as accommodation. In the second and third periods, we focus on the travelers who leave their accommodation places and visit other locations, namely those whose activity category is OTHERS. We refer to these two types of tourist activities as daytime activities and nighttime activities. By separating periods (b) and (c), our aim is to discern potential disparities between daytime and nighttime activities.

Subsequently, we identify three nationality groups—Asian, European, and North American travelers—in three time periods: accommodation, daytime, and nighttime activities—within Seoul, Busan, and Jeju, yielding a total of 27 analytical scenarios. For each scenario, we calculate visitation frequencies based on tourist activity, nationality group, and city. Each traveler who visits a spatial unit with a specific activity category contributes one visitation frequency. To mitigate the impact of uneven base station distribution within cities, we convert the visitation frequency calculations from Thiessen polygons to 1*1 km grids, as outlined in the appendix (Figure B.2). In this way, the visitation frequency of each grid in 27 scenarios can be obtained for further analysis.

2.3. Data analysis

In this section, we conduct the data analysis in the following three aspects. First, for each scenario, we map the spatial distribution of tourist visitation frequency of different nationality groups and conduct the hot spot analysis (Getis & Ord, 1992). This analysis helps us find the locations and ranges that are most visited by tourists and identify areas where tourists concentrate within the city. Second, for each scenario, we analyze the rank-size distribution using the proportion of visitation frequency and perform power law fitting. This step enables us to quantify the spatial concentration degree of tourist activities under different activity categories. The inter-group differences among Asian, European, and North American travelers will become comparable after quantification. Third, we further introduce Theil's Entropy (Iceland, 2004; Massey & Denton, 1988) to measure spatial similarity of tourist activities for the three nationality groups. Theil's Entropy is known to be

an indicator of quantifying residential segregation by considering the unevenness of different subgroups in urban space. It is defined by the following formulas:

$$E = \sum_{r=1}^{m} (\Pi_r) \cdot log\,(1/\Pi_r) \tag{1}$$

$$E_i = \sum_{r=1}^{m} (\Pi_{ri}) \cdot log\,(1/\Pi_{ri}) \tag{2}$$

$$H = \sum_{i=1}^{n} \frac{t_i(E - E_i)}{ET} \tag{3}$$

where Π_r refers to the proportion of a nationality group r in the city's population, Π_{ri} refers to the proportion of a nationality group r in the city's population in unit i, m refers to the total number of nationality groups, n refers to the number of units within a city, t_i refers to the total population of unit i, T refers to the total population of the city, E_i refers to the entropy of the unit i, and E refers to the entropy of the whole city. The entropy index quantifies the weighted average deviation of entropy for each spatial unit from the overall entropy of the city. The range of the entropy index varies from 0 to 1, with 0 representing that all the spatial units have the same composition, and 1 representing all the units containing one group only.

Given that Theil's Entropy allows both pairwise comparisons and multi-group measurements regardless of the size of each group, this index is suitable to measure the spatial similarity of tourist activities between each of the two groups, as well as among three nationality groups. For any pair of nationality groups, suppose that their activities exhibit completely different spatial distributions; for instance, group A concentrates in place A, while group B concentrates in place B, with no overlap between their activity locations. Consequently, we consider the spatial similarity between group A and group B to be very low. In such instances, the entropy index would be significantly high (close to 1), indicating that tourist activities of these two groups occur in entirely different locations. Using the spatial distribution of each nationality group in different time periods as inputs, we compute the entropy index between every pair of groups ($m = 2$), as well as among all three groups ($m = 3$). This approach allows us to identify the two groups with the most similar spatial distribution patterns and quantify the inter-group differences between them.

3. Analysis results

3.1. Spatial distribution of accommodation, daytime and nighttime activities

Fig. 3 shows the spatial distribution patterns of accommodation activities for three nationality groups in Seoul, Busan and Jeju. Grids of darker colors denote areas with higher visitation frequency. Asian travelers' accommodation activities in Seoul (Fig. 3A) exhibit a concentrated pattern, with the vast majority of travelers choosing Myeongdong, Jung-gu, and Mapo-gu as their preferred night residences.

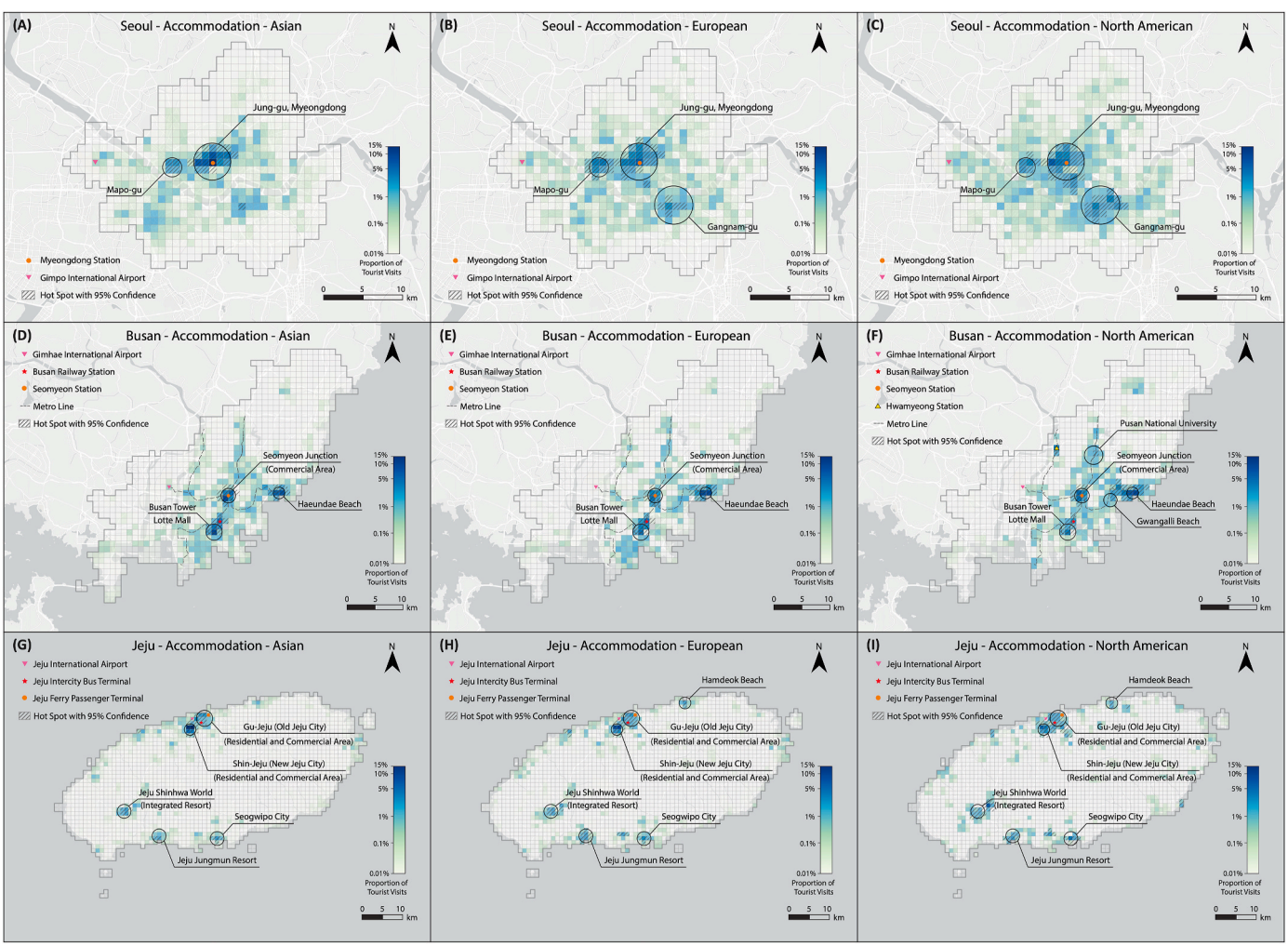


Fig. 3. Spatial distribution of accommodation activities of three nationality groups in: (A–C) Seoul, (D–F) Busan and (G–I) Jeju. Areas with hatched fills denote hot spots based on the Getis-Ord Gi*.

Notably, Jung-gu emerges as the most popular district for Asian travelers, accounting for only 3% of Seoul's total area but hosting an impressive 52% of their overnight stays. In contrast, European travelers (Fig. 3B) and North American travelers (Fig. 3C) have more diverse accommodation options compared to their Asian counterparts. In addition to the two popular districts of Jung-gu and Mapo-gu, European and North American travelers, especially North American travelers, also show a willingness to explore accommodation services in Gangnam-gu.

For Busan and Jeju, we also identify the preferred accommodation locations of each nationality group. As a seaside city with efficient public transportation, the tourist hot spots in Busan are linearly distributed along the metro line and southeastern coastline. Similar to Seoul, Asian travelers also exhibit concentrated accommodation activities in Busan, with three hot spots centered on Busan Tower, Seomyeon Junction, and Haeundae Beach (Fig. 3D). European travelers have similar choices as Asian travelers in terms of accommodation location (Fig. 3E). There are more small hot spots found among North American travelers (Fig. 3F). In Jeju Island, all three nationality groups show a clustered distribution in terms of accommodation activities. This may be due to the unique geography of Jeju Island, where accommodation resources are limited and concentrated in only two downtown areas: Jeju City in the north and Seogwipo City in the south. Asian travelers, in particular, favor New Jeju City for accommodation (Fig. 3G), which accommodates 42% of Asian travelers in a small area of 2 km * 2 km. New Jeju City is a new downtown area south of Jeju Airport, with a high concentration of hotels and duty-free shops, providing convenient accommodation and shopping options. The spatial distribution of European travelers' accommodation locations is similar to Asian travelers', but some small hot spots can also be observed (Fig. 3H). North American travelers have more dispersed accommodation options than Europeans (Fig. 3I), and they are not as interested in New Jeju City as Asian travelers.

Fig. 4 shows the spatial distribution and hot spots of daytime activities of three nationality groups in three cities. Asian travelers remain concentrated and show strong interest in Jung-gu and Mapo-gu during daytime (Fig. 4A), which is consistent with their preferences in accommodation locations. This may be due to the mixed land use and rich tourism resources in these areas, which can meet travelers' various needs such as accommodation, dining, and sightseeing. The spatial distribution patterns of European (Fig. 4B) and North American (Fig. 4C) travelers are concentrated in Jung-gu, Mapo-gu, and Gangnam-gu, spreading from the district center to the surrounding areas. These two groups are more evenly distributed across the city and have a broader range of activities and more diverse options within the group. In Busan, the spatial distribution of Asian tourists' daytime activities is similar to their distribution of accommodation. European and North American travelers are still relatively dispersed and can be observed in various places of the city along the coastline and metro line. In Jeju, during the daytime, the spatial distribution of tourist activities becomes more dispersed as travelers venture out to various attractions on the island. Among the three groups, North American travelers have the most diverse range of options, with hot spots detected in multiple locations along the coastline, suggesting their interests in multiple locations in Jeju during daytime activities. When analyzing the spatial distribution of tourist activities of the three groups, we notice that locations with rich

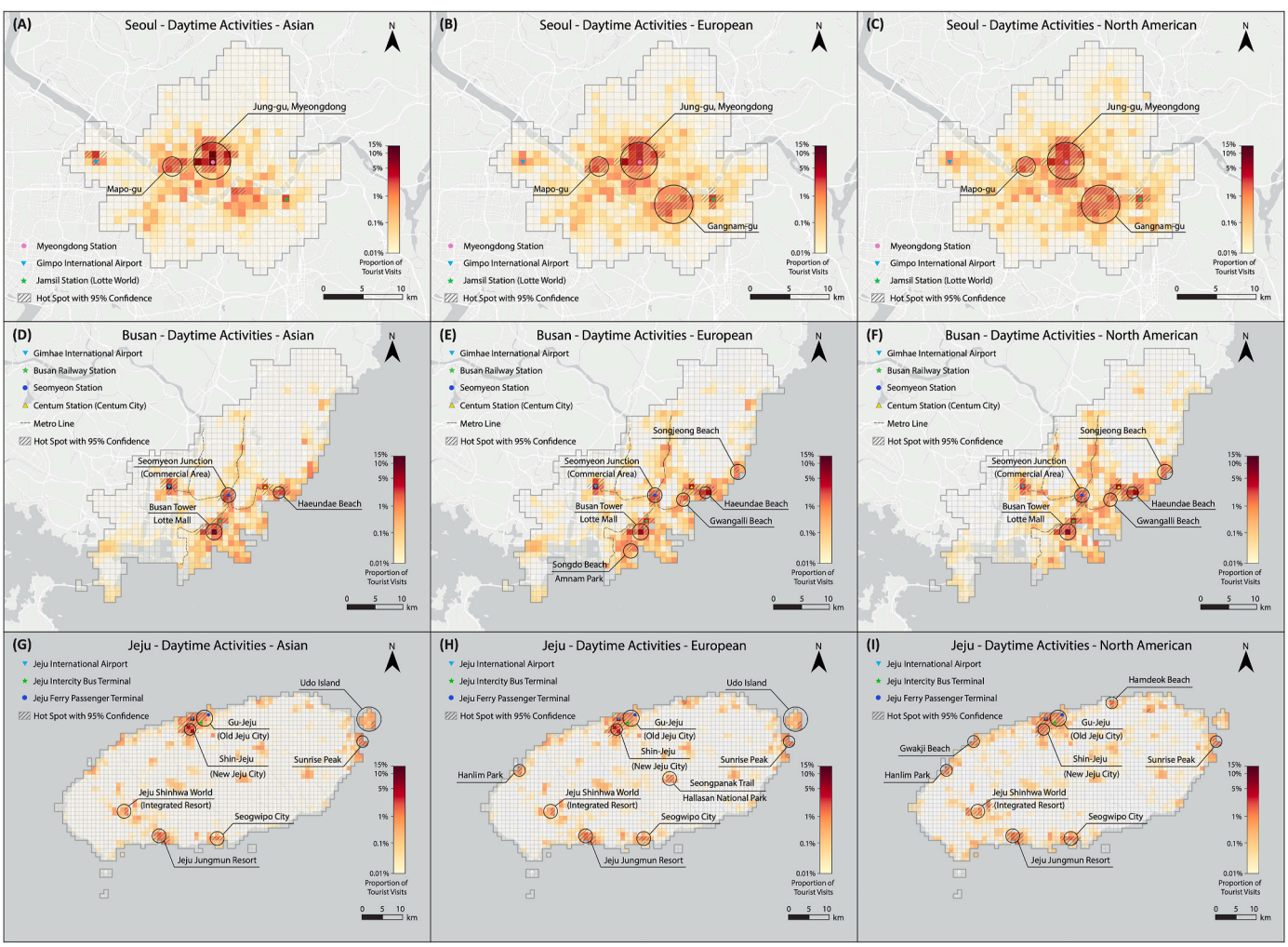


Fig. 4. Spatial distribution of daytime activities of three nationality groups in: (A–C) Seoul, (D–F) Busan and (G–I) Jeju.

tourism resources tend to have higher visitation frequencies, such as Jung-gu in Seoul, areas along the subway lines and coastlines in Busan, as well as New Jeju City and Old Jeju City in Jeju Island. This suggests that the spatial distribution of tourism resources may have a potential impact on the spatial distribution of tourist activities.

In terms of nighttime activities, the spatial distribution of the three nationality groups during nighttime activities (Fig. 5) shares similarities to both daytime activities and accommodation activities. This suggests that nighttime activities serve as a bridge between daytime activities and accommodation. Among all the scenarios, an interesting observation occurred on Jeju Island. As night falls, a notable shift occurs in the spatial distribution of Asian and European travelers, with their activities becoming more concentrated, potentially due to the early closure of natural attractions in Jeju and a limited nightlife scene. North American travelers remain relatively scattered along the west coast of the island, indicating a sustained interest in exploring different areas even during the night hours.

3.2. Spatial concentration degree of accommodation, daytime and nighttime activities

In order to further quantify the degree of spatial concentration effect in accommodation activities, we calculate the rank-size distribution of visitation frequency proportions and perform power law fitting for three cities (Fig. 6). A higher k–value indicates a greater rate of decay, which means that the lower the spatial unit is ranked, the greater the reduction in tourist visitations. The results show that Asian travelers are the most concentrated group with the largest exponent k, followed by European travelers and finally North American travelers. In other words, Asian travelers tend to concentrate their activities in the top-ranked hot spots and are less involved in other places. This finding is consistent across all cities and activities, suggesting that concentration may be an inherent attribute of Asian tourists. In contrast, European and North American travelers, especially North American travelers, can be observed outside the hot spots. When comparing the activities in three time periods, daytime activities usually have the smallest k-values, indicating that travelers are most dispersed during daytime activities. Nighttime activities and accommodation activities are more concentrated with a larger k-value. In terms of three cities, it is evident that travelers are most concentrated on Jeju Island, followed by Busan and finally Seoul. This observation implies that the spatial concentration of tourist activities may be linked to the type of destination and the available tourism resources.

3.3. Spatial similarity of tourist activities among three nationality groups

The Theil's Entropy between each pair of groups and among three groups provides the degree of spatial similarity of tourist activities (Fig. 7). In terms of activities, daytime activities exhibit the most significant similarity among the three nationality groups, while accommodation shows the most significant dissimilarity. Hence, the choice of accommodation location may best reflect the distinct preferences of travelers from diverse cultural backgrounds. Interestingly, daytime activities appear to bring the three groups together to create maximum

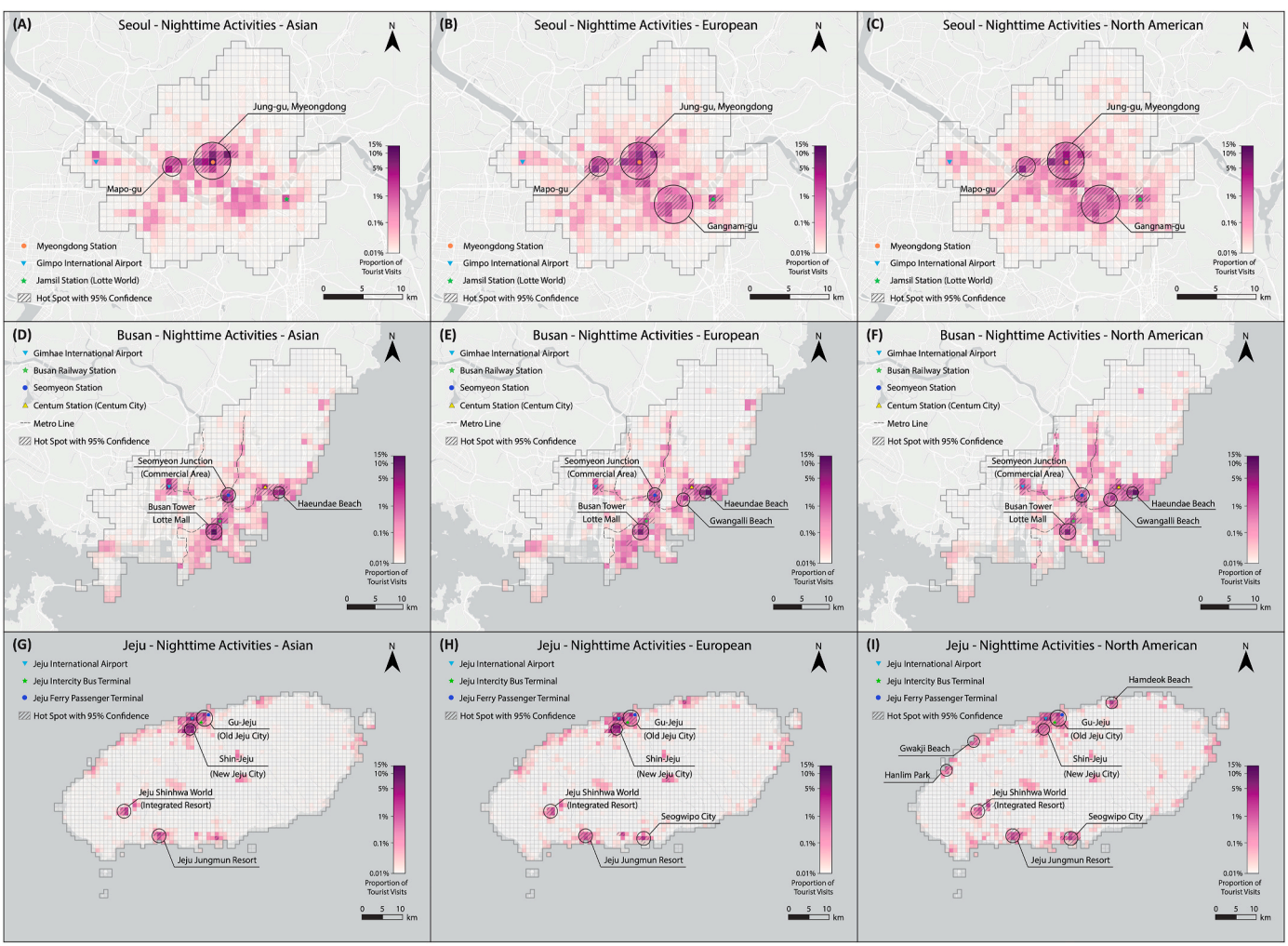


Fig. 5. Spatial distribution of nighttime activities of three nationality groups in: (A–C) Seoul, (D–F) Busan and (G–I) Jeju.

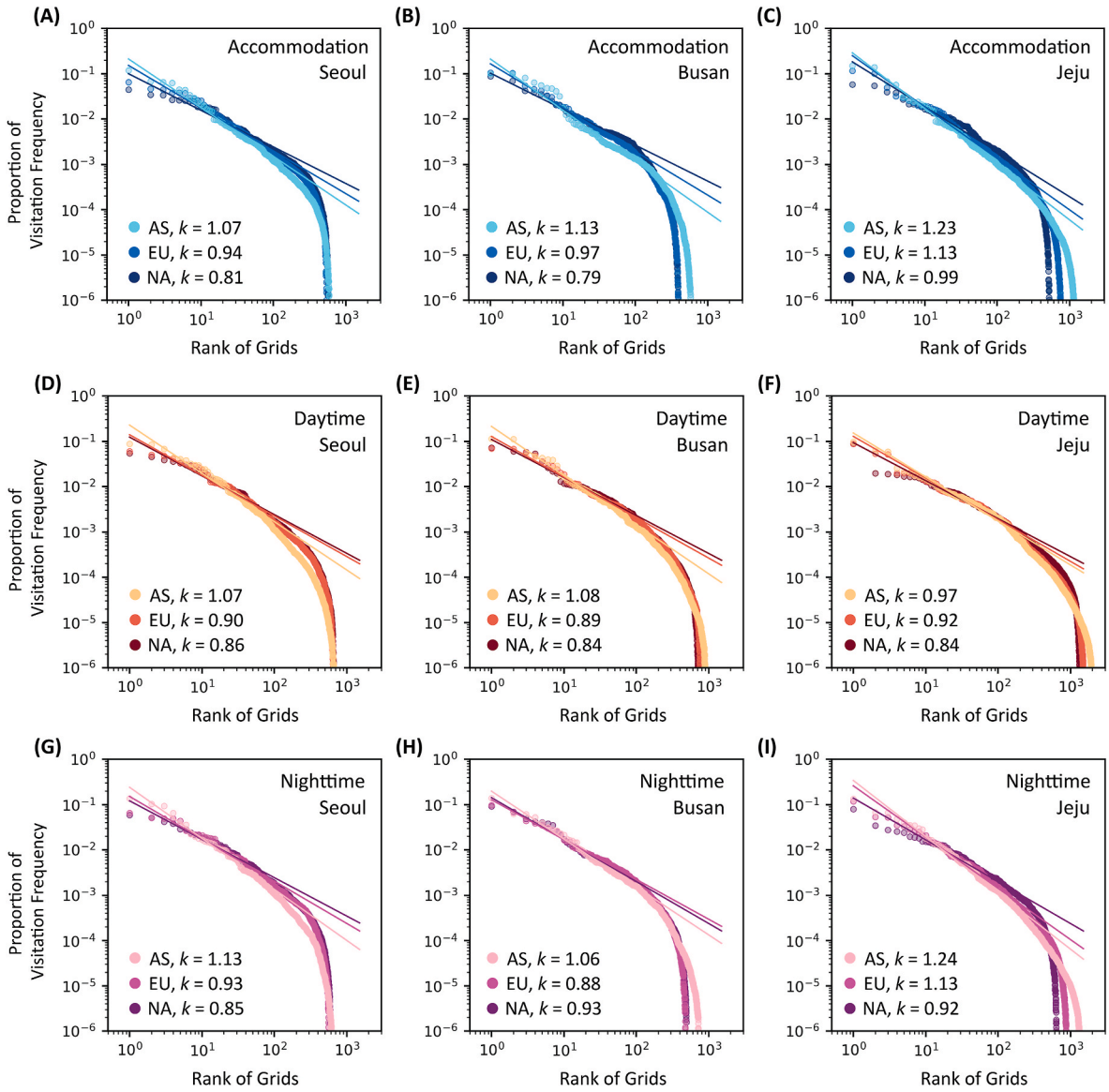


Fig. 6. Spatial concentration degree of: (A–C) accommodation, (D–F) daytime and (G–I) nighttime activities in three cities using rank-size distribution. Each dot in a subfigure denotes the frequency of accommodation activities in a given grid cell by the corresponding nationality group. The lines are fitted using power law distribution.

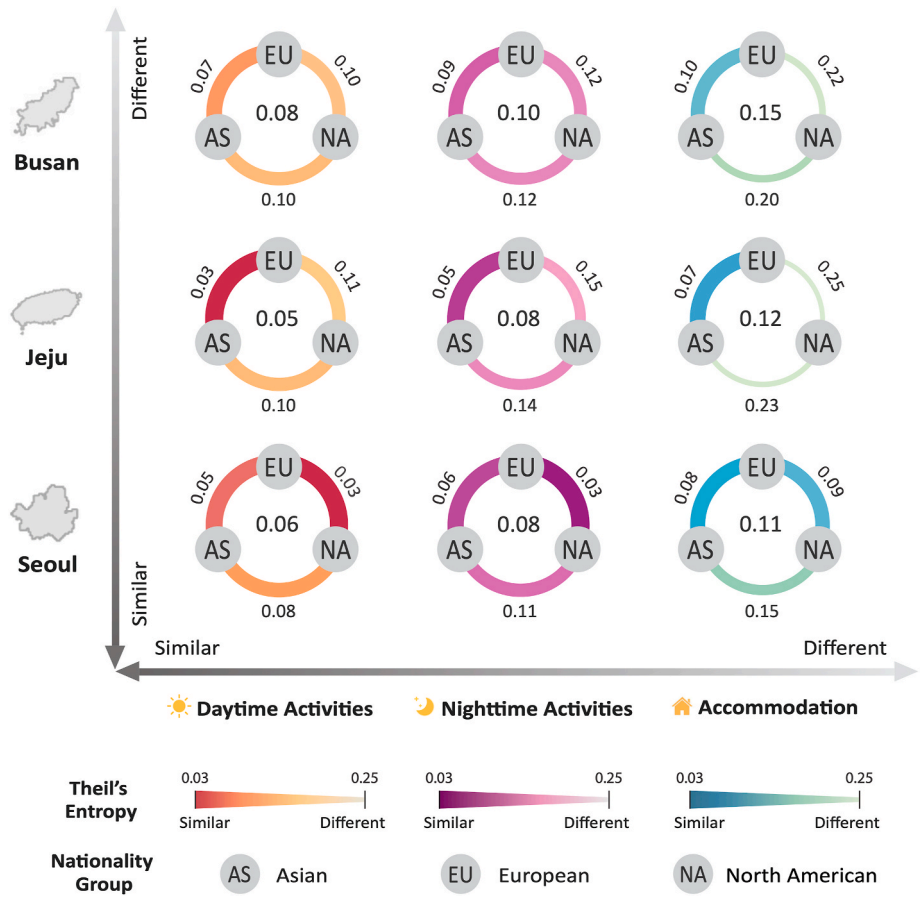


Fig. 7. Spatial similarity of tourist activities using Theil's entropy index. Each circle in the visualization represents the spatial similarity of tourist activities among Asian, European and North American travelers within a specific city and time period. Each linkage in the circle represents the spatial similarity between a pair of nationality groups. The entropy index among three groups is placed in the center of the circle. A wider linkage with darker color between the two groups represents more similar spatial distributions between two groups, with a value near 0. On the contrary, a narrower linkage with lighter color between two groups indicates a weak connection, representing less spatial similarity between the two groups, with a value close to 1. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

integration, possibly owing to their universal appeal to tourists of various cultural backgrounds. Nighttime activities, on the other hand, exhibit more variability among nationality groups than in daytime activities but less than in accommodation activities. This finding further implies the bridging role of nighttime activities. In terms of nationality groups, Europeans share some similarities with both North Americans and Asians, while Asian travelers are the least similar to North American travelers. In three cities, the spatial distribution patterns of tourist activities are more similar in Seoul, and tend to be different in Busan and Jeju. This implies that the attributes and characteristics of the target cities should be considered when analyzing the spatial similarity of tourist activities.

4. Conclusion and discussion

4.1. Conclusion

This study investigates the spatial concentration of intra-urban tourist activities and compares the inter-group differences among Asian, European, and North American travelers. A large-scale dataset of mobile phone records from Seoul, Busan and Jeju in South Korea is used to analyze the spatial distribution patterns, concentration degree and spatial similarity of tourist activities. By identifying activity locations from movement trajectories recorded in the dataset, we analyze how tourist activities of different nationality groups are distributed in different cities in terms of accommodation, daytime activities, and nighttime activities. We also employ the power law to fit the rank-size distribution of tourist visitation frequency, quantifying the concentration degree for each nationality group and activities in each time period. Furthermore, we adopt Theil's Entropy index to measure spatial similarity of tourist activities among three nationality groups.

The research findings reveal a notable trend: Asian travelers exhibit the highest level of concentration, followed by European travelers and lastly, North American travelers. This pattern consistently appears across various time periods and cities, indicating that Asian travelers tend to have more homogeneous preferences within the group. In contrast, Western travelers are more dispersed within cities, exhibiting a broader travel footprint and more diverse choices within groups. The spatial distribution of activities among Asian travelers differs the most from that of North Americans, while Europeans share some similarities with both North Americans and Asians. In terms of tourist activities, accommodation activities demonstrate the most significant concentration and spatial dissimilarity, while daytime activities witness a broader dispersion and integration.

4.2. Theoretical contributions

The findings of this study complement and extend previous research on spatial concentration effect of tourist activities. Our findings, as evidenced through the analysis of mobile phone data, reveal the strongest concentration of Asian travelers. This aligns with the previous observation that Asian tourists are more concentrated in a few

destination cities than Western tourists (Xu et al., 2021). While earlier studies discussed the spatial concentration effect of tourist activities at regional or national scales (Su et al., 2016), this study contributes to the literature by revealing the spatial concentration effect at the intra-city scale with a finer granularity. This tendency for spatial concentration among Asian tourist activities persists both regionally and within cities, suggesting that concentration may be an inherent attribute of Asian travelers. In many Asian countries, collectivism is a traditional value orientation, such as China, Japan, and South Korea (Cai et al., 2020; Pizam & Sussmann, 1995), which may lead Asian travelers to prefer traveling in groups and making decisions collectively. Given that individualism-collectivism is a key dimension in Hofstede's cultural dimensions theory, the findings of this study support the theory's explanation of the heterogeneity of tourist behavior of Asian, European and North American travelers. In addition, considering that existing tourism agglomeration studies generally focus on industries and resources, our study provides a novel perspective by examining spatial concentration effects through the lens of tourist activities. Furthermore, while prior research typically concentrates on individual cities, our study encompasses a diverse range of destination types, unveiling overarching patterns in tourist spatial dynamics.

4.3. Policy implications

Our findings have important implications for planning and policy recommendations. Given that differences in spatial concentration degree are found between different nationality groups, it is essential to consider the heterogeneity of tourists from different social backgrounds in destination and tourist flow management. In addition, the findings in section 3.3 indicate the highest spatial similarity in Seoul, particularly during daytime activities. This suggests that in metropolitan areas with comprehensive tourism resources, marketing strategies should be better adapted to the needs of tourists from diverse backgrounds to maintain competitiveness. However, the spatial preferences of three nationality groups exhibit more heterogeneity in Busan and Jeju, and also in the choice of accommodation locations. This emphasizes the necessity of customizing marketing strategies and accommodation suggestions in these cities to align with the preferences of the target tourist groups. Moreover, our findings reveal the degree of concentration and dispersion of tourist activities within cities, which can assist decision-making according to local conditions and offer insights into balanced spatial development of urban tourism. For example, in coastal cities, the concentration of tourists along the coast could limit tourism growth in inland destinations (Mou et al., 2020). Given the pronounced coastal concentration observed in both Busan and Jeju, local tourism authorities could consider implementing measures to monitor tourist flows in these areas.

4.4. Limitations and future direction

This study has some limitations. Firstly, although the mobile phone dataset utilized in this research holds a significant market share, it does not encompass all international travelers. For example, individuals who do not use cellular data services during their travels, such as children, seniors, and Wi-Fi-only travelers, are not represented in the dataset. This limitation may result in an underrepresentation of certain demographic groups and their travel patterns. Secondly, the timeframe of the data collection is restricted to a 15-day period, which poses challenges in isolating the influence of seasonality on travel behavior. Extending the observation period to a longer duration would enable a more comprehensive examination of seasonal effects. Thirdly, the characteristics of destination, such as urban renewal, spatial form, and transportation accessibility, are also related to the spatial concentration of tourist activities. Future studies could investigate how these factors influence the spatial concentration of tourist activities to unveil the drivers of this phenomenon. Future research could also consider analyzing the impact of nationality as well as various dimensions of culture on the spatial behavior of international tourists. In addition, given that past literature has observed the unique behavioral patterns of certain nationalities at the national level, such as behavioral differences between Japanese tourists and tourists from other countries, future research could provide an in-depth investigation of the major sources of international tourists at a national scale. Despite these limitations, our study offers valuable insights that enrich the comprehension of the spatial concentration of intra-urban tourist activities. The methodology used to measure the degree of spatial concentration can be applied in future studies to analyze the degree of agglomeration of economic activities and tourism resources. The research framework of this study can also be applied to other tourist destinations to understand spatial distribution patterns, concentration and similarity of different nationality groups.

Impact statement

This study develops a systematic analytical framework that utilizes mobile phone data to quantify the spatial concentration of tourist activities and potential heterogeneity among Asian, European and North American travelers. The framework can be used to analyze other location-based data and to investigate behavioral disparities among different social groups. While previous studies discussed the spatial concentration effect at regional or national scales, this study enriches the literature by revealing the spatial concentration effect at intra-city scale with a finer granularity. By introducing two indicators (rank-size distribution; Theil's Entropy) to quantify the spatial concentration and inter-group differences of tourist activities, the quantitative analysis results can provide a reference for the overtourism issues within cities. On the practical level, as variations in spatial concentration are observed among different nationality groups, this study's findings emphasize the necessity of considering travelers' social backgrounds in destination and tourist flow management.

CRediT authorship contribution statement

Jingyi Cheng: Conceptualization, Methodology, Visualization, Writing – original draft, Writing – review & editing, Formal analysis. **Yang Xu:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Project administration. **Izzy Yi Jian:** Writing – original draft, Writing – review & editing. **Mengyao Ren:** Writing – original draft, Writing – review & editing. **Sangwon Park:** Writing – review & editing, Data curation.

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

During the preparation of this work, the authors used ChatGPT in order to proofread the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Acknowledgement

The authors would like to thank the editors and anonymous reviewers for their valuable comments on earlier versions of the manuscript. This research was partly supported by a grant from the Research Center for Digital Transformation of Tourism, The Hong Kong Polytechnic University (Project No. 1-BBFG).

Appendix

Appendix A. Composition of tourist activities in an averaged 24-h day

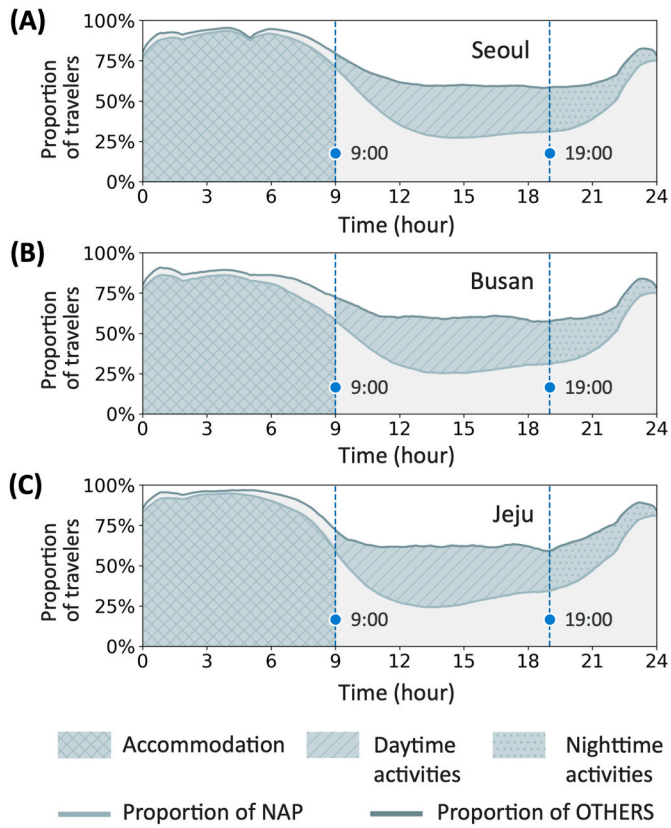


Fig. A.1. The composition of tourist activity categories in (A) Seoul, (B) Busan and (C) Jeju. The results are aggregated to the city level after extracting activity locations from travel trajectories at the individual level on an averaged 24-h day. NAP refers to the anchor point where tourists stay most of the time during the night, while OTHERS refers to the stay at non-NAP locations, indicating activities that occur outside of the residence location. For each city, we calculate the proportion of travelers whose activity semantics are NAP and OTHERS in each hour. The three cities display the similar pattern: the percentage of NAP decreases significantly during the day, while the percentage of OTHERS increases from 9:00 and decreases during the night. Accordingly, we identify the division of the three time periods and define three categories of tourist activities. For those travelers whose activity semantics is NAP from 0:00–9:00, we define their activities as accommodation. Daytime activities are defined as activities with semantics of OTHERS during 9:00–19:00. Nighttime activities are defined similar to daytime activities, but occur between 19:00–24:00.

Appendix B. Calculation of visitation frequency at grid level

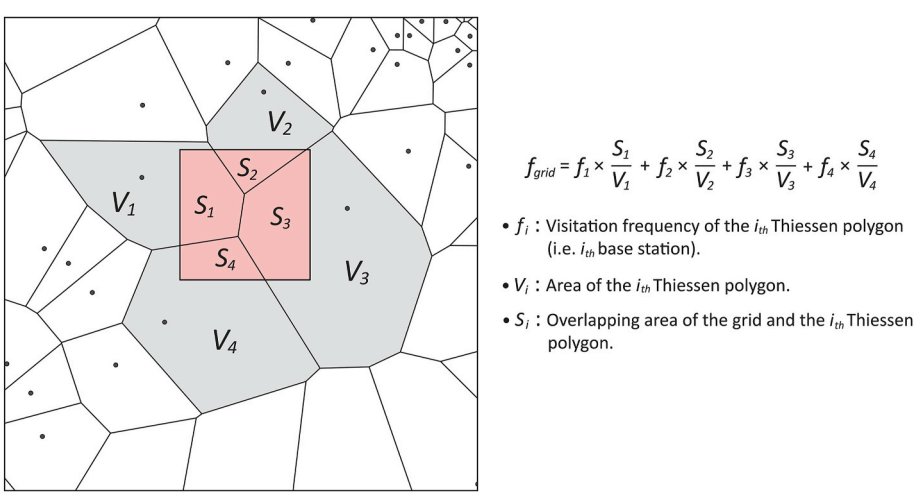


Fig. B.2. Conversion of visitation frequency from base station level to grid level. The service range of mobile phone base stations is approximated by Thiessen polygon.

References

Blondel, V. D., Decuyper, A., & Krings, G. (2015). A survey of results on mobile phone datasets analysis. *EPJ data science, 4*, 1–55.

Buitrago, E. M., & Yñiguez, R. (2021). Measuring overtourism: A necessary tool for landscape planning. *Land, 10*(9), 889.

Cai, Y., Ma, J., & Lee, Y. S. (2020). How do Chinese travelers experience the arctic? Insights from a hedonic and eudaimonic perspective. *Scandinavian Journal of Hospitality and Tourism, 20*(2), 144–165.

Chen, N., & Hsu, C. H. (2021). Tourist stereotype content: Dimensions and accessibility. *Annals of Tourism Research, 89*, Article 103212.

Edwards, D., Dickson, T., Griffin, T., & Hayllar, B. (2010). Tracking the urban visitor: Methods for examining tourists' spatial behaviour and visual representations. In *Cultural tourism research methods* (pp. 104–114). Wallingford UK: Cabi.

Gao, J., Peng, P., Lu, F., & Claramunt, C. (2022). A multi-scale comparison of tourism attraction networks across China. *Tourism Management, 90*, Article 104489.

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis, 24*(3), 189–206.

Hofstede, G. (1983). National cultures in four dimensions: A research-based theory of cultural differences among nations. *International Studies of Management & Organization, 13*(1–2), 46–74.

Iceland, J. (2004). The multigroup entropy index (also known as Theil's H or the information theory index). *US Census Bureau, 31*(2006), 60. *Retrieved July*.

Kim, Y. R., Williams, A. M., Park, S., & Chen, J. L. (2021). Spatial spillovers of agglomeration economies and productivity in the tourism industry: The case of the UK. *Tourism Management, 82*, Article 104201.

Kotus, J., Rzeszewski, M., & Ewertowski, W. (2015). Tourists in the spatial structures of a big polish city: Development of an uncontrolled patchwork or concentric spheres? *Tourism Management, 50*, 98–110.

Lew, A. A., Hall, C. M., & Williams, A. M. (Eds.). (2008). *A companion to tourism*. John Wiley & Sons.

Li, Z., & Liu, H. (2022). How tourism industry agglomeration improves tourism economic efficiency? *Tourism Economics, 28*(7), 1724–1748.

Majewska, J. (2015). Inter-regional agglomeration effects in tourism in Poland. *Tourism Geographies, 17*(3), 408–436.

Maoz, D. (2007). Backpackers' motivations the role of culture and nationality. *Annals of Tourism Research, 34*(1), 122–140.

Massey, D. S., & Denton, N. A. (1988). The dimensions of residential segregation. *Social Forces, 67*(2), 281–315.

Matthews, Y., Scarpa, R., & Marsh, D. (2018). Cumulative attraction and spatial dependence in a destination choice model for beach recreation. *Tourism Management, 66*, 318–328.

Merton, R. K. (1968). The Matthew effect in science: The reward and communication systems of science are considered. *Science, 159*(3810), 56–63.

Mou, N., Zheng, Y., Makkonen, T., Yang, T., Tang, J. J., & Song, Y. (2020). Tourists' digital footprint: The spatial patterns of tourist flows in Qingdao, China. *Tourism Management, 81*, Article 104151.

Namberger, P., Jackisch, S., Schmude, J., & Karl, M. (2021). Overcrowding, overtourism and local level disturbance: How much can Munich handle?. In *Travel and tourism in the age of overtourism* (pp. 100–120). Routledge.

Özdemir, C., & Yolal, M. (2017). Cross-cultural tourist behavior: An examination of tourists' behavior in guided tours. *Tourism and Hospitality Research, 17*(3), 314–324.

Pizam, A., & Fleischer, A. (2005). The relationship between cultural characteristics and preference for active vs. passive tourist activities. *Journal of Hospitality & Leisure Marketing, 12*(4), 5–25.

Pizam, A., & Jeong, G. H. (1996). Cross-cultural tourist behavior: Perceptions of Korean tour-guides. *Tourism Management, 17*(4), 277–286.

Pizam, A., & Sussmann, S. (1995). Does nationality affect tourist behavior? *Annals of Tourism Research, 22*(4), 901–917.

Popp, M. (2012). Positive and negative urban tourist crowding: Florence, Italy. *Tourism Geographies, 14*(1), 50–72.

Raun, J., Ahas, R., & Tiru, M. (2016). Measuring tourism destinations using mobile tracking data. *Tourism Management, 57*, 202–212.

Reisinger, Y., & Turner, L. (2012). *Cross-cultural behaviour in tourism*. Routledge.

Ritter, W. (1987). Styles of tourism in the modern world. *Tourism Recreation Research, 12*(1), 3–8.

Shao, H., Zhang, Y., & Li, W. (2017). Extraction and analysis of city's tourism districts based on social media data. *Computers, Environment and Urban Systems, 65*, 66–78.

Shoval, N., & Ahas, R. (2016). The use of tracking technologies in tourism research: The first decade. *Tourism Geographies, 18*(5), 587–606.

Su, S., Wan, C., Hu, Y., & Cai, Z. (2016). Characterizing geographical preferences of international tourists and the local influential factors in China using geo-tagged photos on social media. *Applied Geography, 73*, 26–37.

World Tourism Organization. (2018). *'Overtourism'? – Understanding and managing urban tourism growth beyond perceptions, executive summary*. Madrid: UNWTO. https://doi.org/10.18111/9789284420070 (UNWTO); Centre of Expertise Leisure, Tourism & Hospitality; NHTV Breda University of Applied Sciences; and NHL Stenden University of Applied Sciences.

Wu, C. L., & Carson, D. (2008). Spatial and temporal tourist dispersal analysis in multiple destination travel. *Journal of Travel Research, 46*(3), 311–317.

Xu, Y., Li, J., Belyi, A., & Park, S. (2021a). Characterizing destination networks through mobility traces of international tourists—a case study using a nationwide mobile positioning dataset. *Tourism Management, 82*, Article 104195.

Xu, Y., Li, J., Xue, J., Park, S., & Li, Q. (2021b). Tourism geography through the lens of time use: A computational framework using fine-grained mobile phone data. *Annals of the Association of American Geographers, 111*(5), 1420–1444.

Xu, Y., Xue, J., Park, S., & Yue, Y. (2021c). Towards a multidimensional view of tourist mobility patterns in cities: A mobile phone data perspective. *Computers, Environment and Urban Systems, 86*, Article 101593.

Yang, Y. (2012). Agglomeration density and tourism development in China: An empirical research based on dynamic panel data model. *Tourism Management, 33*(6), 1347–1359.

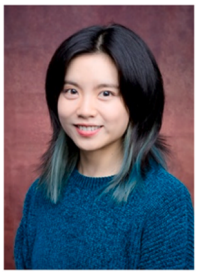
Yang, Y., & Wong, K. K. (2013). Spatial distribution of tourist flows to China's cities. *Tourism Geographies, 15*(2), 338–363.

You, X., O'leary, J., Morrison, A., & Hong, G. S. (2000). A cross-cultural comparison of travel push and pull factors: United Kingdom vs. Japan. *International Journal of Hospitality & Tourism Administration, 1*(2), 1–26.

Zhou, B., Qu, H., & Li, N. (2016). How tourist attraction agglomeration affects a regional tourism economy. *Tourism Analysis, 21*(6), 603–616.

Jingyi Cheng is a Ph.D. student in the Department of Land Surveying and Geo-Informatics at the Hong Kong Polytechnic University. Her research interests include Geographic Information Science (GIScience), urban informatics, urban tourism and spatio-temporal big data analysis.

Yang Xu is an Associate Professor in the Department of Land Surveying and Geo-Informatics at the Hong Kong Polytechnic University. His research interests include Geographic Information Science (GIScience), urban informatics and tourism big data.



Izzy Yi JIAN is an Assistant Professor in the Department of Social Sciences and Policy Studies at the Education University of Hong Kong. Her research interests include spatial justice, privatisation/liminality of public open space, inclusive/age-friendly design, community engagement and geomedia.

Mengyao Ren is a Ph.D. student in the Department of Land Surveying and Geo-Informatics at the Hong Kong Polytechnic University. Her research interests include human travel behaviour, urban informatics, urban resilience, public transport, and urban planning.

Sangwon Park is a Professor in the College of Hotel and Tourism Management at the Kyung Hee University. His research interests include tourism big data, information technology, and labour productivity in tourism and hospitality.