Eastern Journal of European Studies
DOI: 10.47743/ejes-2023-SI08 ● November 2023 ● VOLUME 14, (SI) SPECIAL ISSUE

The geography of social media platform attention for tourist attractions - spatial digital data analytics of scenic hot spots in China

Huiqin Li[a], Jingyan Pan[a], Yujie Hui[a], Jingjing Liu[a], Peter Nijkamp[bc*]

[a]China University of Geosciences, China; [b]Alexandru Ioan Cuza University, Iasi, Romania; [c]Open University, Heerlen, The Netherlands

Abstract
Based on the geo-spatial distribution and rich social media data of many important scenic tourist places (high-level scenic spots in China), this study presents a quantitative analysis using GIS technology and several spatial statistical tools to examine the geographical distribution and network attention of these spots. It is found that there is a clear geographical imbalance in the spatial distribution of these scenic spots in China, primarily concentrated in the lower-lying and densely populated eastern regions. Using spatial autocorrelation methods to assess the degree of match between these two spatial patterns, it is observed that the spatial network attention and geographical distribution of hotspots are mutually correlated only in major coastal cities. The results enhance our understanding of effective tourism network marketing instruments and provide further insight into the geographical layout of scenic spots in the country.

Keywords: tourist scenic attractions, social media platform, geographical distribution, network attention, destination marketing

Introduction

With the development of the Internet, utilizing it to gather pertinent information about tourist destinations has become a crucial tool in tourism decision-making (Kourtit, 2019). When tourists browse the web, search engines usually record and tally the extensive network attention data related to tourist spots, defining what is known as "network attention" (Pan et al., 2012; Choi et al., 2012). Based on research into the network effects of high-level scenic spots in China, it has been observed that the network attention these spots receive has a significant "advance effect" (Li et al., 2008). Online attention can also predict travel flows (Artola et al., 2015) or analyse tourist behaviour trajectories to enhance destination system personalization (Dietz et al., 2020).

* ✉ Peter Nijkamp, Professor, Alexandru Ioan Cuza University, Iasi, Romania and Open University, Heerlen, The Netherlands; e-mail: pnijkamp@hotmail.com.

The concept of network attention has been frequently employed in predicting tourist flows to scenic spots, aiding scenic spot managers in devising appropriate response measures to enhance spot capacity. Consequently, the connection between online information flow and real-world tourism flows has garnered significant attention from tourism researchers over the years (Lu et al., 2007). While numerous studies have explored the geographical spatial distribution of scenic spots (Wu, 2003), recent studies have achieved noteworthy results in tourist flow analysis and the geographical spatial distribution structure of tourist scenic spots, considering network attention. Evidently, China's tourism industry exhibits a distinct regional development pattern, with high concentrations in the East and South, and lower ones in the West and North (Zhang et al., 2020). China's national 5A tourist attractions also display a significant pattern with one high-attention area, three main centres, and five sub-centers (Li et al., 2019).

Understanding the relationship between geographical spatial distribution and network attention for scenic spots is undeniably vital for tourism managers aiming to pursue effective destination marketing and management, an area that remains underexplored. Therefore, our study seeks to explore this relationship to assist scenic spots in utilizing online marketing strategically, enhancing brand influence, and fostering the healthy and sustainable development of Chinese tourism.

5A-grade scenic spots represent the highest standard for tourist destinations in China and significantly influence regional tourism development (Zhang et al., 2019). Investigating the spatial distribution of 5A scenic spots can guide future planning and marketing efforts for these spots in China. The network attention of tourist destinations serves as a reliable indicator to gauge their allure and development trends. It can also support online marketing in these destinations, improve the local environment, and enable rational resource allocation (Cai et al., 2016; Nadotti et al., 2019). Therefore, jointly studying the geographical spatial distribution and network attention of 5A scenic spots in China holds great significance for tourism management.

The present study aims to investigate whether the actual geographical distribution of tourist amenities in a given country or region, such as scenic spots, aligns with the spatial distribution of digital attention on social media platforms. This research aims to enrich our understanding of resource distribution and the influence of tourist destinations, ultimately aiding in the management and marketing of these scenic spots.

To achieve this, the study utilizes Baidu Index, a tool capable of accurately analysing online interest in Chinese attractions, to gauge network attention. This approach allows us to delve into the perceived appeal of Chinese tourist attractions in both a geographical and functional context. Consequently, this paper quantitatively analyses major tourist attractions across China, specifically the 5A-level scenic spots.

The analysis employs various quantitative tools, including a GIS-based quantitative spatial dispersion index, the nearest neighbour index (Cong et al., 2020), along with the use of a Gini coefficient (Cong et al., 2020), density analysis (Wei et al., 2023), and a standard deviation ellipse (Wei et al., 2023). Subsequently, spatial autocorrelation analysis (Ma et al., 2023) is employed to explore the spatial distribution pattern of network attention in China.

These analytical methods comprehensively depict the spatial distribution pattern of network attention for high-level scenic spots in China, offering valuable insights for further research and tourism planning in the country.

1. Literature review

A few tourism studies have concentrated on the geographical distribution of scenic spots, as well as on the temporal distribution and prediction of network attention for tourist scenic spots (Pan & Fesenmaier, 2006). Our study specifically examines the geographical distribution characteristics of tourist scenic spots, the distribution characteristics of network attention, and the spatial coupling relationship between geographical spatial distribution and network attention for tourist scenic spots in China. This paper emphasizes the significance of spatial pattern research and illustrates the correlation between online information flows and real-world tourism flows. To provide context for our research, we will begin with a brief literature review.

1.1. The geographical distribution structure of scenic spots

Tourist attractions are recognized as a crucial factor in shaping the local tourism economy, significantly impacting its ability to attract tourists (Pascariu et al., 2021). The analytical study of spatial relations and patterns is gradually replacing earlier descriptive approaches, a trend that extends to tourism research, including scenic spots.

The geographical distribution of scenic spots refers to the degree and state of spatial concentration resulting from the interaction of attractive tourism-economic elements in a given area. It captures the spatial attributes and interrelationships of tourism activities and essentially reflects tourism activities in geographical space (Bian, 2003). Tourist scenic spots play a pivotal role in the tourism industry, guiding, supporting, and ensuring regional economic development. Analysing the spatial distribution of scenic spots to inform planning and layout can better facilitate the strategic marketing of scenic spots and promote sustainable tourism development. Thus, a geo-spatial analysis of scenic spots holds substantial practical significance and supports tourism management and marketing efforts.

The spatial structure of scenic spots is a focal point in tourism research. It predominantly showcases the core-edge distribution structure, an unbalanced centralized distribution profile, and hierarchical characteristics (Guedes & Jiménez, 2015; Kang et al., 2018; Yuan et al., 2010). The uneven geographical spatial distribution of scenic spots is influenced by various factors, including regional tourism resources, topography, population distribution, economic development level, and government strategies (Huang et al., 2010). Additionally, the spatial distribution structure of scenic spots varies with their level. For instance, A-level tourist scenic spots in Beijing exhibit a "dumbbell structure," featuring dense urban areas, outer suburbs, and sparser near-suburbs, driven primarily by resource and market factors (Mao et al., 2011). These 4A-level scenic spots display a cohesive distribution in space, indicating a strong spatial connection with China's regional economy and urban development level (Ma & Yang, 2003).

It's essential to note that a 5A grade scenic spot, as mentioned, represents the highest quality tourist destination in China, signifying the pinnacle of tourist destination excellence. Therefore, a study on the spatial distribution structure of 5A scenic spots in China carries critical significance in guiding tourism development.

1.2. The spatial distribution of network attention to scenic spots

With the rapid proliferation of IT in China, the Internet has emerged as a crucial marketing and e-commerce tool for the tourism industry. Online entertainment marketing has become a pivotal method for stimulating travel interest (Nishijima, 2020). It has become customary to acquire tourism information, select travel destinations, formulate travel plans, and organize itineraries with the assistance of the Internet (Cen & Liang, 2007). In recent years, the connection between information flows represented by cyberspace attention and tourism flows represented by actual tourist numbers has garnered the attention of numerous researchers. Some scholars have specifically examined the access to tourist destination information to expand the visibility of scenic spots (Éber et al., 2018).

The online nature of network data can effectively compensate for the delays associated with traditional prediction methods. Typically, there exists a cointegration relationship between network attention and tourist traffic. Several researchers have endeavoured to explore this relationship using the Baidu Index (Yang et al., 2015). Huang et al. (2013) employed econometric methods to scrutinize the relationship between the Baidu Index and visitor numbers at the Imperial Palace Scenic Spot in Beijing. They also applied cointegration theory and Granger causality tests to examine the connection between network attention and passenger traffic. Furthermore, they utilized ARMA models and VAR models to forecast tourist volumes (Sun et al., 2017). These experimental findings demonstrated that, compared to benchmark models, the proposed kernel extreme learning machine (KELM) models, which integrate tourist volume series with the Baidu Index and the Google Index, significantly enhance forecasting performance in terms of both accuracy and robustness analysis (Sun et al., 2019).

While numerous studies have concentrated on the analysis of network attention and the predicted tourist numbers at scenic spots, only a few have delved into the spatial distribution characteristics of network attention and the factors influencing such distribution. This information holds valuable guidance for scenic spot management and aids websites in supporting tourism destination marketing activities.

1.3. Coupling of geographical distribution and spatial pattern of network attention to scenic spots

Research into tourism network attention, geographical distribution characteristics, and their relationship with tourist scenic spots holds significant importance in guiding destination network marketing, optimizing resource spatial layout, and fostering sustainable tourist destination development (Ju et al., 2017). In studies focusing on spatial distribution, researchers often refer to the Baidu Index search platform to access network attention data (Feng & Li, 2014; Wang et al., 2014).

Various statistical inference methods, including data comparison methods (Lu et al., 2010; Ma et al., 2011), correlation analysis methods (Long et al., 2013), least square methods (Long et al., 2011), autoregressive distribution lag models, MIDAS models, and vector autoregressive models, are employed to examine the relationship between tourist volumes and network attention (Prosper & Ryan, 2015). These methods aim to determine whether a positive correlation exists between them.

However, the aforementioned studies are predominantly data-centric and overlook the real-world geographical space upon which scenic spots and network attention are contingent. Clearly, the actual geo-spatial distribution is intricate, and the underlying relationship has been insufficiently explored. Consequently, delving into the connection between the spatial distribution structure of scenic spots and the spatial distribution of network attention among potential clients poses a challenging research task. The following section will provide a description of our database and the set of research methods employed to address the aforementioned research objectives.

2. Methodology

In this section, we present the data collected and the research methodology of this paper. The database primarily consists of information related to scenic spots and network attention. The methodology predominantly involves analysing the spatial distribution of scenic spots, the spatial distribution of network attention, and the correlation between them.

2.1. Data collection

5A-level scenic spots represent the highest standard for tourist destinations in China, embodying their brand value and core competitiveness. Additionally, 5A scenic spots tend to have elevated brand influence and reputation, making their network attention indicative of tourism development trends. Hence, this paper selects 5A-level scenic spots and their network attention as the focal points of our research.

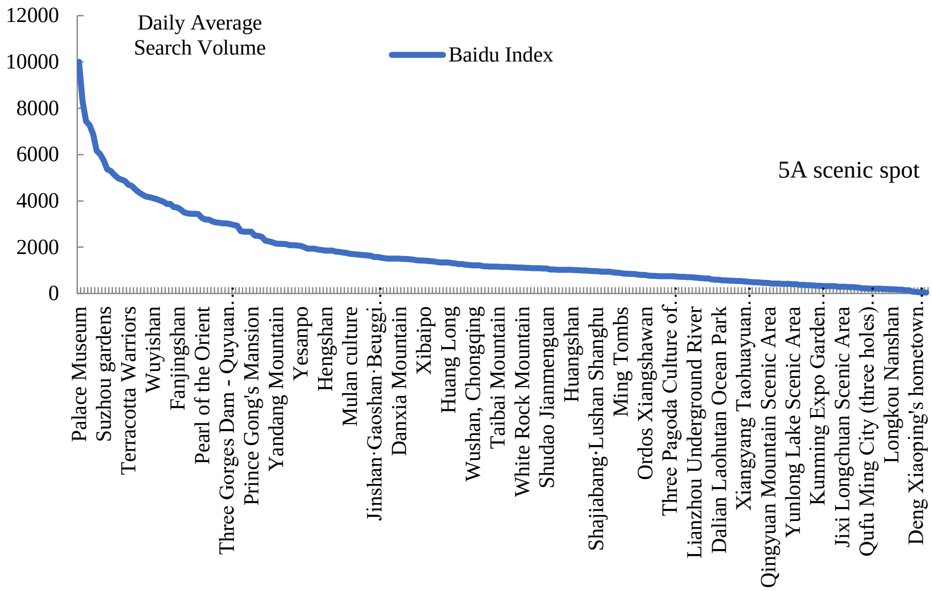
Our data collection in China is divided into two main components: national 5A-level scenic spots (in 2018) and digital network attention. These data encompass all 31 provinces in China, excluding the Hong Kong, Macao, and Taiwan regions. We employ the Baidu Index, a search engine tool, to identify and analyse the level of online attention among internet users towards various user concepts during a specified period. It accurately characterizes the online attention dynamics of Chinese attractions. Consequently, we utilize the Baidu Index to gauge the network attention received by the 5A-level scenic spots. The Baidu Index is acquired based on the keywords associated with each scenic spot within the selected timeframe.

Geographical data for the 31 Chinese provinces are primarily sourced from the National Basic Geographic Information Centre's 14 million entries of basic geographic information. This data allows us to access national boundaries, provincial boundaries, and other essential geographical information. The statistical tables containing longitude and latitude coordinates of 5A-level scenic spots are generated using Baidu Maps' coordinate selection function. Other network attribute data mainly originate from the National Tourism Administration (http://www.cnta.gov.cn), the National Statistical Bureau (http://www.stats.gov.cn/), and the relevant statistical yearbooks of the 31 provinces. After compilation and processing, we are able to construct a comprehensive database encompassing all network attributes of 5A-level tourist attractions in China.

To obtain the Baidu Index for 5A scenic spots, the author inputs the names and relevant keywords of the main scenic spots into the platform, along with the specified time period (from January 1, 2018, to December 31, 2018). This process allows us to retrieve all the data for the specified period. Figure 1 illustrates the daily average Baidu Index among the top 35 national 5A scenic spots in 2018.

We should note that the database does not include 16 scenic spots, such as the Baishahu Scenic Area of the 10th Division of Xinjiang Production and Construction Corps, the Qingzhou Ancient City Tourist Areas of Weifang, Jiangxi, the Ashatu Stonehenge Tourist Area of Chifeng City in Inner Mongolia Autonomous Region, and others. Additionally, the Shanhaiguan scenic spot in Hebei Province only regained its 5A-level status in November 2018. Consequently, data from these 17 scenic spots have been excluded. In the final count, out of the 2,595 A-level attractions in China, only 242 received attention on the Internet. These will be further examined in the remaining sections of the paper.

Figure 1. Distribution of Baidu Index of national 5A tourist attractions



Source: authors' representation

2.2. Spatial distribution characteristics of 5A class tourist scenic spots

Overall distribution characteristics and spatial distribution types

Using ArcGIS software, we can present and visualize the national 5A tourism scenic spots on a map. This allows us to obtain and examine the overall distribution characteristics of these 5A tourist attractions in China. The nearest-neighbour index is a spatial distribution measure used to describe the arrangement of point elements within a specific area (Cong et al., 2020). Thus, in this paper, we employ it to analyse the spatial distribution patterns of 5A scenic spots in China. The formula for calculating the nearest neighbour index is as follows:

$$R=r/r_E \tag{1}$$

where, r represents the average observation distance, and r_E the average expected distance; its formula is as follows:

$$r_E=1/(2\sqrt{n/a}) \tag{2}$$

where n represents the number of point elements in the region, and a represents the surface area of the region.

Spatial agglomeration characteristics

5A-level scenic spots are often geographically unequally distributed. The Gini coefficient (Cong et al., 2020) is a commonly used indicator for measuring regional disparities and revealing spatial patterns in the distribution of relevant geographical phenomena. The formula for calculating the Gini coefficient is as follows:

$$G=(-\sum_{i=1}^{n} P_i \times \ln P_i)/\ln N \tag{3}$$

$$C=1-G \tag{4}$$

Where G represents the Gini coefficient, C signifies the degree of distribution equilibrium, and P_i represents the proportion of 5A-level scenic spots in the i-th province relative to the total number of 5A-level scenic spots in the country. N denotes the number of zones, which in this case is 31 provinces, cities, and districts. To further elucidate the spatial distribution characteristics of the 5A tourist attractions, we can also employ the complementary Lorenz curve, which is based on the number of 5A tourist attractions and their proportion in the national total.

Spatial distribution density

To investigate the specific distribution of 5A scenic spots in China, it’s essential to identify high-density and low-density areas, assess the spatial dispersion (or agglomeration) of these spots, and introduce the nuclear density analysis method. The nuclear density analysis method is employed to examine the clustering of entities in space (Wei et al., 2023). Assuming that there are certain data points within an area, the probability density of a data point “x” is calculated as follows:

$$f(x)=\left(1/_{nh^d}\right)\sum_{i=1}^{n} k\left({}^{x-x_i}/_{h}\right) \tag{5}$$

where d represents the dimension of the point data, $k\left({}^{x-x_i}/_{h}\right)$ denotes the calculation equation of the kernel function ; h denotes the specified search radius, which is essentially the core bandwidth ; $(x\text{-}x_i)$ signifies the distance from the hypothetical point to x_i. By utilizing the ArcGIS kernel density analysis tool, a spatial distribution density map of China’s 5A tourist attractions can be generated.

2.3. Overall characteristics and spatial pattern of network attention

We will now conduct a more comprehensive analysis of the geographical pattern of the network attention value for China's 5A-level tourist attractions, aiming to gain a holistic understanding of their network attention. In this context, we can utilize the ArcGIS kernel density analysis tool to create both the network attention density map and the spatial distribution map of China's 5A scenic spots.

2.4. Matching analysis of spatial distribution patterns and network attention patterns of scenic spots

In this section, we will conduct the standard deviation ellipse analysis and the spatial autocorrelation analysis to confirm the relationship between the spatial distribution pattern of national 5A scenic spots and the network attention pattern.

Standard deviation ellipse analysis

The standard deviation ellipse (Wei et al., 2023) is employed to summarize the spatial characteristics of regional tourism spots, encompassing aspects such as the central tendency, dispersion, and directional trends. The formula for calculating the standard deviation ellipse is as follows:

$$SDE_X=\sqrt{\sum_{i=1}^{n}(x_i-\bar{X})^2/n} \tag{6}$$

$$SDE_y=\sqrt{\sum_{i=1}^{n}(y_i-\bar{Y})^2/n} \tag{7}$$

where (x_i,y_i) represents the coordinates of the spatial position of each geographical element; $\bar{X}$ and $\bar{Y}$ denote the arithmetic average centre; and SDE_x and SDE_y represent the centre of an ellipse To determine the orientation of the ellipse, we use the x-axis as the reference, with the North (12-point direction) considered as 0 degrees with a clockwise rotation. The calculation formula is as follows:

$$tan\,\theta = \left(\sum_{i=1}^{n}\bar{x}_i^2-\sum_{i=1}^{n}\bar{y}_i^2\right)+\sqrt{\left(\sum_{i=1}^{n}\bar{x}_i^2-\sum_{i=1}^{n}\bar{y}_i^2\right)^2+4\left(\sum_{i=1}^{n}\bar{x}_i\bar{y}_i\right)^2}\Bigg/2\sum_{i=1}^{n}\bar{x}_i\bar{y}_i \tag{8}$$

where $\bar{x}_i$ and $\bar{y}_i$ represent the difference between the mean centre and the $\bar{X}$ and $\bar{Y}$ coordinate. Finally, the length of the semi-major axis of the standard deviation ellipse is calculated, and then the standard deviation ellipse is drawn. Based on ArcGIS software, the standard deviation ellipse can be generated directly using the directional distribution tool.

Spatial autocorrelation analysis

Spatial autocorrelation (Ma et al., 2023) essentially pertains to the potential interdependence of certain key variables within the same geographical distribution area. This interdependence can be quantified using the standard Moran's Index, which is calculated using the following formula:

$$I = N\sum_i \Sigma_j W_{ij}(x_i-\bar{x})(x_j-\bar{x}) \Big/ (\Sigma_i\Sigma_j W_{ij})\Sigma_i(x_i-\bar{x})^2 \tag{9}$$

where x_i and x_j represent the specific attribute values of the location space unit; $\bar{x}$ denotes the mean value; N represents the number of points within the sample; W_{ij} denotes the spatial weight matrix; i and j denote the spatial relationship representing the sum. Additionally, our study utilizes GEODA software to establish the spatial weight matrix based on the .shp file of 5A tourist attractions distributed across the 31 administrative regions of the country. We then conduct a global autocorrelation analysis. To further investigate the correlation between the spatial distribution of the 5A scenic spots and their network attention, our analysis also includes a bivariate spatial autocorrelation analysis (see Section 4).

3. Results

3.1. Spatial distribution characteristics of 5A class tourist scenic spots

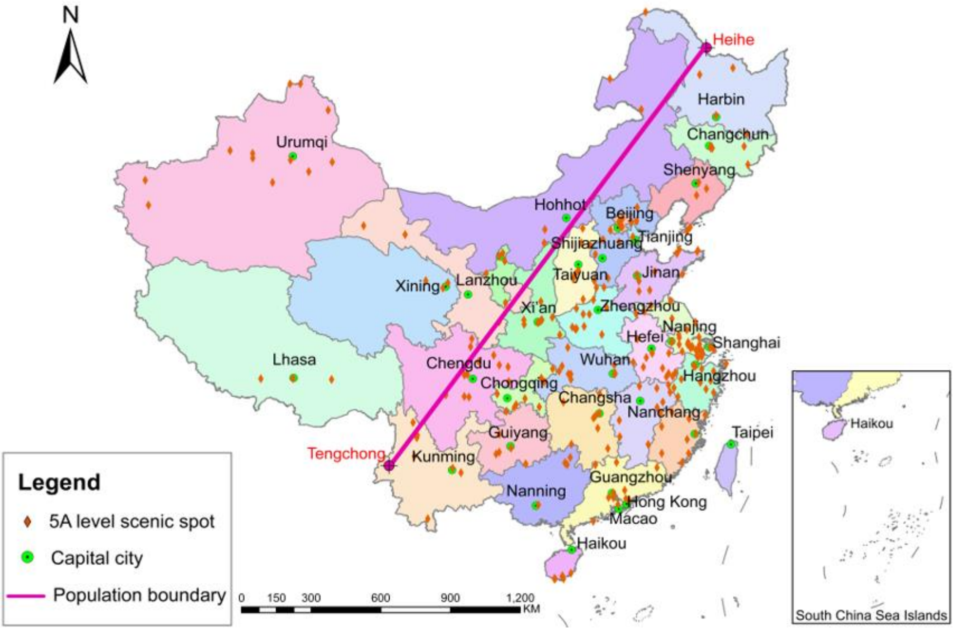
Overall distribution characteristics and types

Using ArcGIS software, the national 5A tourism scenic spots can be visualized on a map (see Figure 2). From this graph, we observe that, overall, the majority of China's 5A tourism scenic spots are concentrated in the East, with fewer in the West. These spots radiate from East to West, displaying a gradient from dense to sparse. Specifically, scenic spots are primarily concentrated in the Pearl River Delta, the plain of the middle and lower reaches of the Yangtze River, the North China Plain, the Sichuan Basin, and the Guanzhong Plain. In contrast, the number of scenic spots in Northwest and Southwest China is less than five, encompassing regions such as Tibet, Southern Xinjiang, Qinghai, Inner Mongolia, and Guangxi.

It's worth noting that the high terrain in the east and the low terrain in the west may influence this spatial distribution. The land terrain can be categorized into three layers: the first layer, which includes the Qinghai-Tibet Plateau and the Qaidam Basin, is the least prevalent. The third layer, primarily consisting of the Northeast Plain, the North China Plain, and the middle and lower reaches of the Yangtze River, is the most densely populated. There is a notable similarity between the spatial distribution of the 5A tourist attractions and the population density distribution across the entire country.

Along the Hu Line, which serves as a boundary, the distribution of 5A tourist attractions is sparse in the West and dense in the East (Figure 2).

Figure 2. Spatial distribution of national 5A scenic spots



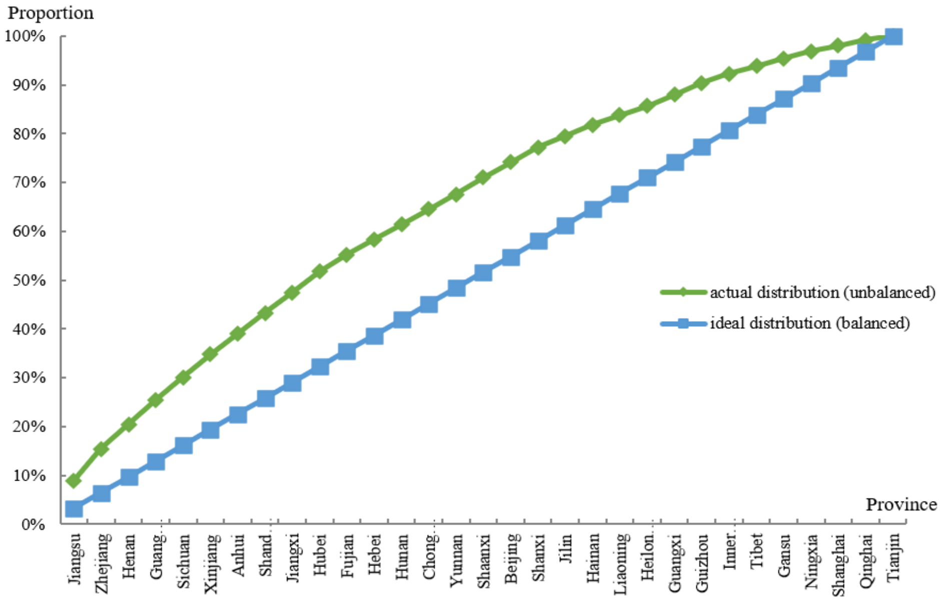
Source: based on the Ministry of Natural Resources standard map service website GS (2019) No. 1825 standard map production, the basemap boundary has not been modified.

Using ArcGIS software, we have determined that the average observation distance r for 5A tourist attractions is estimated to be 73.451 km. According to formula (2), we can now calculate the theoretical average expected distance. In this formula, "n" refers to the 242 5A tourist attractions, and a represents the total land area of China, which is 9.6 million square kilometres. Therefore, the theoretical average expected distance can be computed as 128.162 km. With a proximity index R of 0.5731, it's evident that R< 1, indicating a clear spatial concentration pattern in the distribution of 5A scenic spots across the entire country.

Spatial agglomeration characteristics

The calculation of the Gini coefficient yields the following values: G=0.5948; C=0.4052. Therefore, 259 5A scenic spots are highly concentrated across China's 31 provinces. Referring to the 5A spatial distribution map of tourist attractions in Figure 3, we can observe the varying presence of scenic spots in different regions.

Figure 3. The Lorenz curve of spatial distribution of national 5A tourist attractions



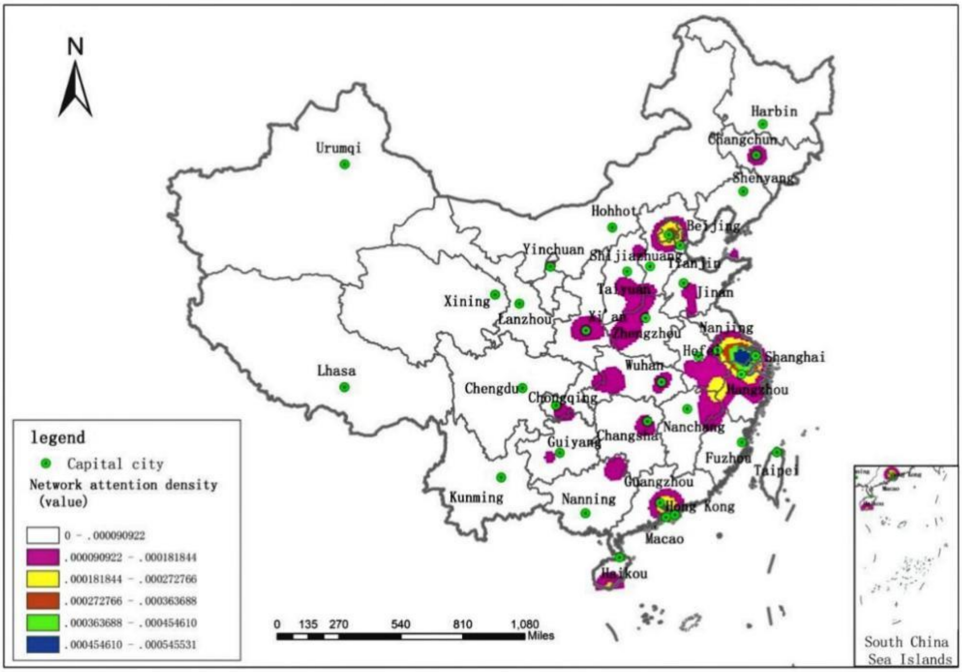
Source: authors' representation

For instance, in Jiangsu Province, there are 23 scenic spots, constituting 8.88% of the total. Following Jiangsu are Zhejiang, Henan, Guangdong, Sichuan, Xinjiang, Anhui, Shandong, Jiangxi, and Hubei. The number of 5A-level scenic spots in these ten provinces accounts for 51.74% of the total number of scenic spots in China. Regions with robust economic conditions, such as Shanghai and Tianjin, have more 5A scenic spots, while areas with weaker economic foundations, such as Tibet and Qinghai, have fewer scenic spots. This pattern is further affirmed by the Lorenz curve results, where the pronounced arc of the Lorenz curve reflects the uneven spatial distribution of 5A scenic spots in China.

Spatial distribution density

With the assistance of the ArcGIS nuclear density analysis tool, we generated the density map depicting the spatial distribution of 5A-level tourist attractions in China, displayed in Figure 4. The high-density accumulation areas are primarily concentrated in the middle and lower reaches of the Yangtze River, Beijing, Tianjin, Hebei, the middle and lower reaches of the Yellow River, Central China, and Southeast China. This distribution appears to resemble a well-known '328' pattern, characterized by 'three enrichment zones, two main enrichment centers, and eight sub-enrichment centers.' Notably, the Yangtze River Delta, the Beijing-Tianjin Rim, and the Pearl River Delta exhibit higher distribution density, as outlined in Table 1.

Figure 4. Distribution density of network space in national 5A tourist attractions



Source: based on the Ministry of Natural Resources standard map service website GS (2019) No. 1825 standard map production, the basemap boundary has not been modified.

Table 1. Spatial distribution pattern of 5A scenic spots nationwide

Distribution pattern	Name of 5A scenic spots nationwide
Three enrichment zones	Taihang Mountains-Funiu Mountains Enrichment Zone Central-South Shandong Enrichment Zone Anhui-Zhejiang-Jiangxi enrichment zone
Two main enrichment centres	Centre of Yangtze River Delta Beijing-Tianjin-Hebei Centre
Eight enrichment Sub-centres of Concern enrichment	Xi'an Wuhan Chongqing Changsha Guangzhou North Guangxi Southern Hainan Changchun

Source: authors' representation

3.2. Quantitative analysis of social media network attention

Overall characteristics of network attention

Based on the processed and sorted network attention values of 5A-level tourist attractions in China, the top five scenic spots in 2018 were the Forbidden City (10003), Wuzhen (8293), Huashan (7437), Qingchengshan-Dujiangyan (7274), and Mount Tai (6876). Conversely, the five least popular scenic spots were the Shenzhen Overseas Chinese City (77), Deng Xiaoping's hometown tourist area (71), Happy Land Resort World (47), and the Mutual Aid Turkish Hometown (37).

In 2018, the Palace Museum launched a substantial cultural program called 'Shangxin, the Palace Museum' to break away from the former stereotype and allow tourists to experience a 'zero distance' connection with the history and culture of the Forbidden City. This program's immense success led to a significant increase in the Palace Museum's attention.

The Fifth World Internet Congress held in Wuzhen also attracted more attention to the area as a tourist attraction. Among the top 20 scenic spots in terms of network attention, the majority (13) are mountain and river destinations, including Huashan, Mount Tai, Putuo, and Emeishan. These scenic spots are highly popular, boast well-developed tourism infrastructure, effective marketing, and nationwide recognition.

As the tourism market continues to evolve, tourist demands are diversifying, providing a wider range of choices. However, the popularity of various theme parks among tourists has noticeably declined, trailing far behind the allure of scenic spots.

The spatial layout of network attention

The ArcGIS nuclear density analysis tool was employed to generate the density map illustrating the network attention and spatial distribution of China's 5A-level scenic spots (Figure 5). The network attention of these scenic spots exhibits distinct high-density and low-density areas, offering a clear depiction of concentrated agglomeration and dispersed distribution.

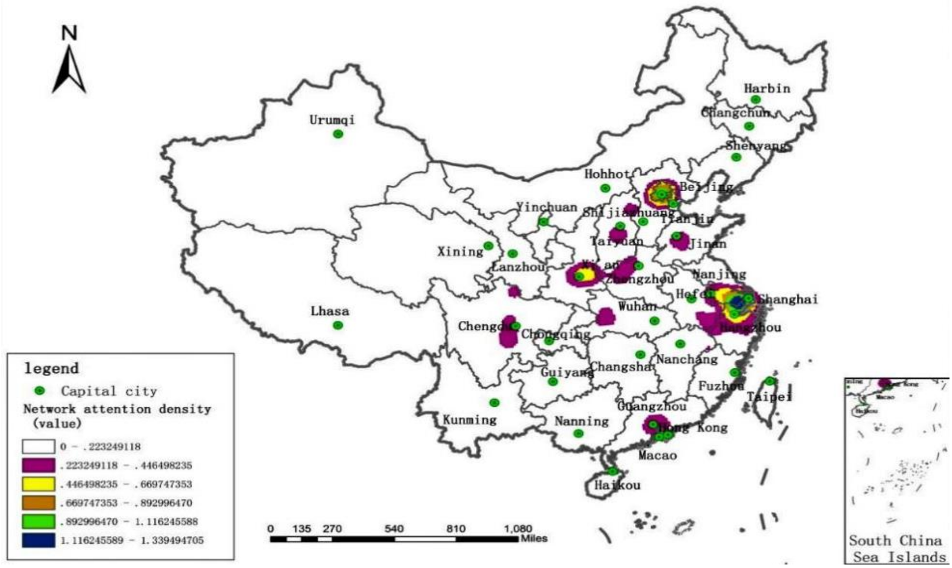
It's worth noting that the current spatial pattern, often referred to as the 'Three Five-Year Plan,' is being implemented nationwide. This pattern comprises 'three main centers of attention and five sub-centers of attention.' High-density areas of attention are primarily concentrated in the Yangtze River Delta, around Beijing and Tianjin, in the Guanzhong Plain, and a few other locations, as detailed in Table 2.

Table 2. National 5A scenic area network attention spatial pattern

Distribution pattern	Name of national 5A scenic area
Three main centres of concern	Centre of Yangtze River Delta Beijing-Tianjin-Hebei Centre Central Guanzhong Plain
Five sub-centres of concern	Guangzhou Sub-Centre Chengdu Sub-Centre Zhengzhou Sub-Centre Jinan Sub-Centre Taiyuan Sub-Centre

Source: authors' representation

Figure 5. Spatial distribution density of network attention in national 5A-level tourist attractions



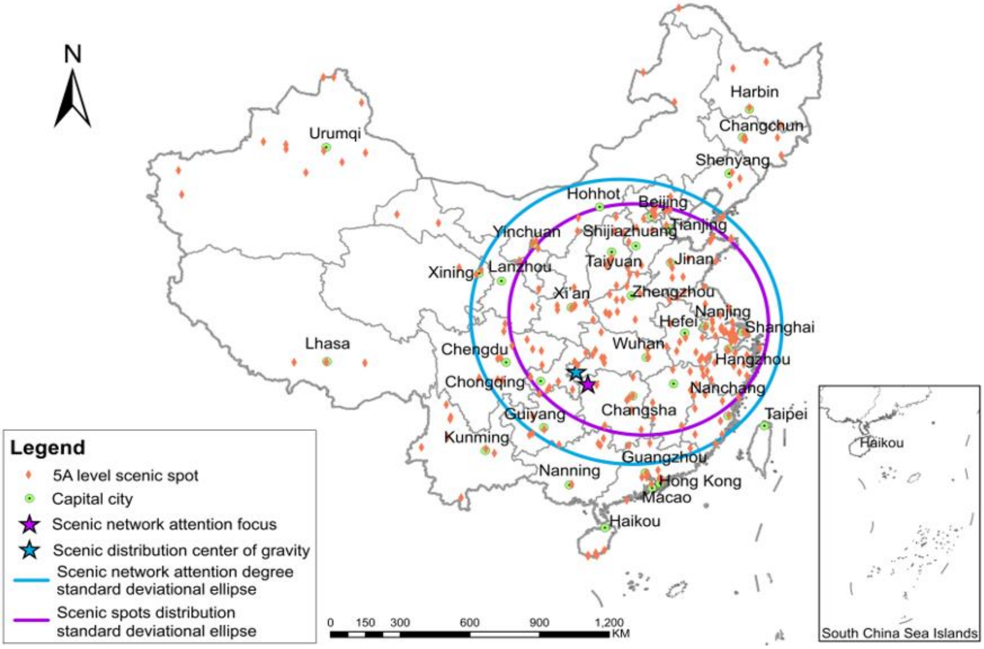
Source: based on the Ministry of Natural Resources standard map service website GS (2019) No. 1825 standard map production, the basemap boundary has not been modified.

3.3. Matching analysis of spatial distribution and network attention patterns of scenic spots

Standard deviation ellipse analysis

The standard deviational ellipse between the spatial distribution of the 5A tourist attractions and their network attention is shown (Figure 6).

Figure 6. National 5A-level tourist scenic and network attention distribution standard deviation ellipse



Source: based on the Ministry of Natural Resources standard map service website GS (2019) No. 1825 standard map production, the basemap boundary has not been modified.

The shape of the standard deviation ellipse for China's 5A scenic spots appears to be similar to that of the standard deviation ellipse for network attention. The results for each parameter also exhibit similarities, indicating a strong spatial coupling. However, when considering the distribution direction, the primary trend direction of the spatial distributional ellipse and the network attention ellipse is not immediately evident. In contrast, the network attention ellipse appears to be more circular, and there is no standard deviational ellipse for spatial distribution. This suggests that the network attention area for 5A tourist attractions is more concentrated. Examining the location of the distribution center, the center of the spatial distributional ellipse for 5A tourist attractions is situated near Laifeng County, Enshi Autonomous Prefecture, Hubei Province.

The geographical coordinates of the centre of spatial distribution are 109 degrees 13'17 E and 29 degrees 41'52 N. The distribution centre of the network attention ellipse of the 5A tourist attractions is located near Longshan County, Xiangxi Tujia and Miao Autonomous Prefecture, Hunan Province. The distribution centre of the network attention ellipse is 109 degrees 37'42 E and 29 degrees 19'31 N, located in the southeast direction of the spatial distribution ellipse distribution centre. The straight-line distance between the two places is about 59 km, which indicates that the

5A scenic spots are more concentrated in the central area. In terms of the ratio of the length to the short half-axis of the standard deviational ellipse, the length of the long half-axis and the short half-axis of the spatial distributional ellipse for 5A scenic spots is 1127.84 km and 1008.59 km, respectively; the length of the long half-axis and that of the short half-axis of the network attention ellipse for the 5A scenic spots are 970.49 km and 879.02 km, respectively. This shows that the difference between the length of the long half-axis and the short half-axis of the 5A scenic spots is relatively small. Overall, both the spatial distribution of scenic spots and network attention are relatively centralized. When considering the angle of the standard deviational ellipse, the spatial distributional ellipse of the 5A scenic spots has an angle of approximately 110.73 degrees, while the angle of the network attention ellipse for the 5A scenic spots is around 96.02 degrees. The difference is relatively small, approximately 14 degrees, indicating that the network attention pattern and the spatial distribution pattern of the 5A scenic spots in China align along an East-West direction in space.

Spatial autocorrelation analysis

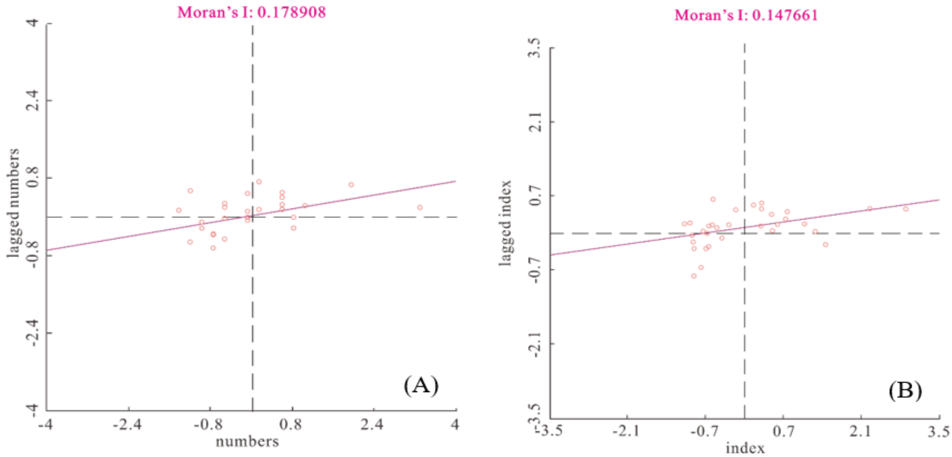
Finally, a global autocorrelation analysis was conducted using GEODA software (Figure 7A). The computed Moran's I value is 0.178908, which is greater than 0 for the 5A scenic spots. This suggests a positive spatial correlation among different provinces, indicating that the distribution of scenic spots is linked to provincial locations. However, the overall Moran's I for the 5A scenic spots is relatively low, implying that while there is a positive correlation between different provinces and regions, it is weak, and there is limited interaction among them.

Through 999 Monte Carlo tests on Moran's I, the calculated p-value for Moran's I was determined to be 0.005, indicating a high level of reliability in this method's calculation of Moran's I.

Using the same method, the spatial autocorrelation analysis of the network attention for national 5A scenic spots (Figure 7B) reveals a Moran's I value of 0.147661. This indicates a positive spatial correlation in the network attention for these scenic spots across different provinces, with some level of network attention coming from neighboring provinces and regions. However, the overall connection is relatively weak (Moran's I < 0.5). Likewise, the p-value obtained after 999 Monte Carlo tests for the network attention of 5A scenic spots is 0.014, signifying a clear spatial autocorrelation at a 98.6% confidence level. Nevertheless, this aggregation is not considered significant. To further investigate the correlation between the spatial distribution of 5A-level scenic spots and their network attention, a bivariate spatial autocorrelation analysis was conducted, as depicted in Figure 8. In this figure, we observe that Moran's I is 0.145699, signifying a spatial correlation between the distribution of 5A-level scenic spots across all provinces and the distribution of network attention across all provinces. This correlation trend is positively oriented.

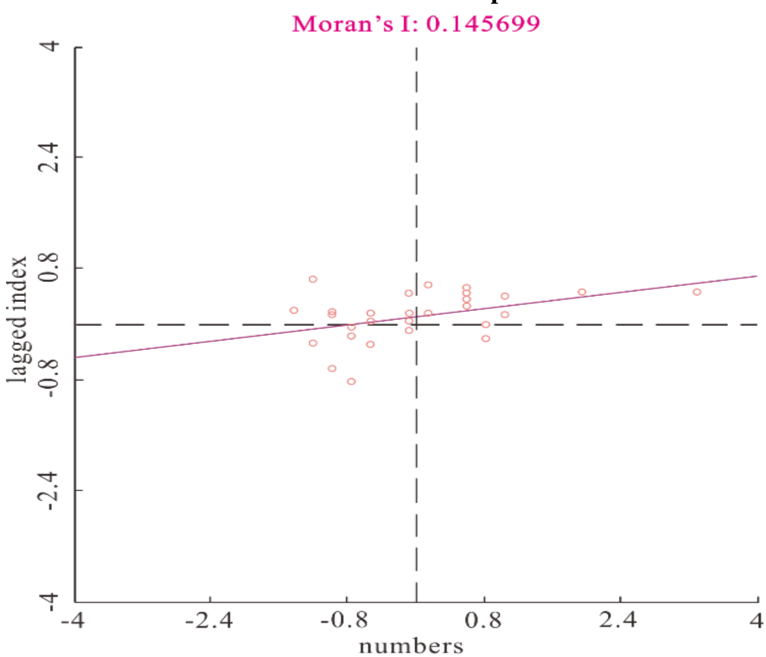
Subsequently, after testing, the p-value was determined to be 0.011, indicating its significance at a confidence level of 98.9%.

Figure 7. 5A-grade scenic spot distribution scatter plot (A) and its network attention scatter plot (B)



Source: authors' representation

Figure 8. Bivariate autocorrelation Moran scatter plot



Source: authors' representation

4. Discussion and lessons

In the era of big data, when tourists search for information about tourist destinations online, the digital footprints they leave behind are referred to as "network attention" (see, for example, Choi & Varian, 2012; Jordan et al., 2013; Vuylsteke et al., 2010). Existing research demonstrates a strong connection between the network attention of tourism and actual tourist traffic (Lin et al., 2012; Long et al., 2011). Therefore, it holds significant importance to analyse the network attention garnered by tourism scenic spots to develop intelligent marketing and attraction strategies.

Our research focuses on utilizing "big data" information related to 242 prominent scenic spots (5A attractions) across 31 Chinese provinces (excluding Hong Kong, Macao, and Taiwan). Employing GIS technology and various statistical tools, we conducted a quantitative analysis with the aim of uncovering the correlation between the spatial distribution pattern of these key attractions in China and their network attention. We discovered that the spatial distribution of 5A scenic spots in China exhibits a noticeable geographical imbalance. Social media networks tend to concentrate on China's major urban hubs, such as the Yangtze River Delta, the Beijing-Tianjin-Hebei region, and the Guanzhong Plain. However, the spatial pattern of China's 5A scenic spots is primarily aligned with the spatial distribution of social media network attention in central areas, particularly in the Beijing-Tianjin area, the Yangtze River Delta, and the Pearl River Delta

In this study, we assessed the spatial structure of China's 5A-level tourist attractions using the nearest neighbour index and the Gini coefficient. Our findings reveal that the spatial distribution structure of these attractions in China is notably concentrated, primarily situated in the eastern region within cities with robust economic prowess. The clustering is predominantly located within the low-lying "third step" region of China, mirroring the population distribution pattern across the country. This conclusion aligns with prior research (Wu & Tang, 2003). Several scholars have also noted that the unequal spatial distribution of China's 5A scenic spots can be attributed to variations in socio-economic and tourism development levels among different regions (Zhang et al., 2019). Economically prosperous and densely populated areas tend to offer superior financial and material support for the development and construction of 5A scenic spots, coupled with a robust tourist market boasting substantial consumption capacity (Wang et al., 2013).

It's crucial to highlight that the western part of China possesses relatively abundant tourism resources, accounting for approximately 40% of the nation's total. Consequently, there is merit in establishing 5A scenic spots to actively promote the sustainable development of these western tourism resources and, in turn, help narrow the economic disparities between the eastern and western urban hubs.

We identified the top five scenic spots in China's 5A category with the highest network attention in 2018 as follows: the Forbidden City, Wuzhen, Huashan, Qingchengshan-Dujiangyan, and Mount Tai. Interestingly, in 2014, the top five 5A-level scenic spots in China were led by natural mountain destinations, with Jiuzhaigou at the forefront. The reason behind the Forbidden City's top ranking in 2018 can be attributed to a series of innovative tourism marketing strategies implemented since 2013, such as cross-border marketing involving cosmetics, cultural variety program marketing, and innovation in tourism cultural and creative products. These multifaceted measures significantly increased the scenic spot's exposure to the public and enhanced its overall popularity. However, Jiuzhaigou did not make the top five in 2018, primarily due to the closure of the scenic area following a magnitude 7.0 earthquake in 2017, which lasted until September 2019.

Furthermore, our analysis revealed that the spatial distribution of 5A scenic spots and their network attention patterns exhibit a certain degree of correlation, notably in regions like Beijing-Tianjin, the Yangtze River Delta, and the Pearl River Delta. In contrast, this correlation is weaker in other provinces across China. These results emphasize the importance of regional hubs leveraging their magnetic appeal to stimulate tourism. Scenic spots within these regions should capitalize on the internet as an effective marketing tool, engaging in collaborative marketing efforts to enhance the overall visibility of regional attractions.

Our study delved into the spatial distribution patterns of China's 5A scenic spots and their network attention, shedding light on the coupling relationship between them. While previous research primarily focused on individual provinces, specific scenic spots, or analysed the potential tourist influx based on network attention, few studies have examined the national and spatial perspective of this relationship. Thus, our study fills this gap in the existing literature.

5. Limitation and future research

This study exclusively relies on data from Chinese social media platforms and does not account for the internet attention of international tourists. In the future, it would be beneficial to validate our findings using different internet search tools, such as Google Trends. Additionally, our study focuses solely on the annual data of the Baidu Index for 2018, which represents a relatively limited time frame. Further investigation is required to delve into the coupling relationship between the spatial distribution pattern of 5A scenic spots and their network attention in China. It's worth noting that the accessibility and level of transportation development between scenic spots or regions are crucial factors influencing tourist flow. Future research can incorporate these practical factors into a comprehensive analysis.

Conclusion and recommendations

This paper has primarily analysed the spatial distribution characteristics and patterns of network attention for China's 5A scenic spots using the "Baidu index" across 242 5A scenic spots in 31 provinces, excluding Hong Kong, Macao, and Taiwan. Our quantitative investigation, employing spatial statistical analysis tools, has yielded the following key findings:

- Significant geographic disparities exist in the distribution of major scenic attractions across China.
- High-density areas of network attention partly align with densely distributed scenic spots, with a strong focus on major urban agglomerations.
- The spatial distribution pattern of Chinese scenic spots shows similarities to the pattern of network attention, with a higher degree of alignment in specific regions, primarily the Beijing-Tianjin region, the Yangtze River Delta, and the Pearl River Delta.

In light of these findings, we offer the following strategic recommendations for future research and tourism planning: (i) Optimize the distribution pattern of 5A scenic spots. (ii) Balance regional tourism-economic differences. (iii) Utilize the Internet as an effective tool in the marketing of scenic spots to enhance their visibility. (iv) Strengthen regional tourism cooperation to achieve win-win situations. (v) Innovate tourism marketing tools. (vi) Enhance the national and international network attention of scenic spots. These guidelines provide valuable insights for shaping the future of China's tourism industry and its scenic attractions.

Acknowledgment: This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS - UEFISCDI, project number PN-III-P4-PCE-2021-1878, within PNCDI III "Institutions, digitalization and regional development in the European Union".

References:

Artola, C., Pinto, F., & Pedraza, P. D. (2015). Can internet searches forecast tourism inflows? *International Journal of Manpower, 36*(1), 103-116. https://doi.org/10.1108/ijm-12-2014-0259

Bian, X. (2003). Research on Urban Tourism Spatial Structure. *Geography and Geo-Information Science*, *19*(1), 106-108. http://dx.chinadoi.cn/10.3969/j.issn.1672-0504.2003.01.026

Cai, W., Peng, J., & Qin, J. (2016). A Study on National Network Attention Heat Matrix andPromotion Strategy in Shaoshan. *Tourism Science, 30*(4), 61-72. http://dx.chinadoi.cn/10.16323/j.cnki.lykx.2016.04.005

Cen, C., & Liang, T. (2007). Research on Internet Information Search Behaviors of Young Travelers in China: a Case Study of the University Students in Guangzhou. *Tourism Science*, *21*(1), 56-62. http://dx.chinadoi.cn/10.16323/j.cnki.lykx.2007.01.011

Choi, H., & Varian, H. (2012). Predicting the Present with Google Trends. *Economic Record, 88*(1), 2-9. https://doi.org/10.1111/j.1475-4932.2012.00809.x

Cong, L., Yu, J., Wang, L. (2020). Spatiotemporal evolution and its influencing factors of semi-consumptive wildlife tourist attractions in China. *Journal of Natural Resources*, *35*(12), 2831-2847. https://doi.org/10.31497/zrzyxb.20201202

Dietz, L. W., Sen, A., Roy, R., & Wörndl, W. (2020). Mining trips from location-based social networks for clustering travelers and destinations. *Information Technology & Tourism*, *22*(1), 131-166. https://doi.org/10.1007/s40558-020-00170-6

Éber, F. Z., Baggio, R., & Fuchs, M. (2018). Hyperlink network analysis of a multi destination region: the case of Halland, South Sweden. *Information Technology & Tourism*, *20*(1), 181-188. https://doi.org/10.1007/s40558-018-0108-9

Feng, N., & Li, J. (2014). A couple analysis of the extraversion online tourism information and inbound tourist flow. A case of the American and Canadian inbound tourist flow. *Tourism Tribune*, *29*(4). http://dx.chinadoi.cn/10.3969/j.issn.1002-5006.2014.04.009

Huang, X., Zhang, L., & Ding, Y. (2013). Study on the Predictive and Relationship between Tourist Attractions and theBaidu Index: A Case Study of the Forbidden City. *Tourism Tribune*, *28*(11), 93-100. http://dx.chinadoi.cn/CNKI:SUN:LYXK.0.2013-11-016

Huang, Y., Chen, G., & Wu, X. (2010). Spatial structure of tourist attractions in Fujian province-statistic analysis based on the national 3A level-above tourist attractions. *Economic Geography*, *30*(7), 1195-1199. http://dx.chinadoi.cn/10.15957/j.cnki.jjdl.2010.07.031

Jordan, E. J., Norman, W. C., & Vogt, C. A. (2013). A cross-cultural comparison of online travel information search behaviors. *Tourism Management Perspectives*, 6, 15-22. https://doi.org/10.1016/j.tmp.2012.11.002

Ju, S., Tao, Z., & Han, Y. (2017). Coupling coordination degree between rural scenic tourist network attentionand gravity in Nanjing city. *Economic Geography*, *37*(11), 220-228. http://dx.chinadoi.cn/10.15957/j.cnki.jjdl.2017.11.027

Kang, S., Lee, G., Kim, J., & Park, D. (2018). Identifying the spatial structure of the tourist attraction system in South Korea using GIS and network analysis: An application of anchor-point theory. *Journal of Destination Marketing & Management*, 9, 358-370. https://doi.org/10.1016/j.jdmm.2018.04.001

Kourtit, K. (2019). Cultural heritage, smart cities and digital data analytics. *Eastern Journal of European Studies*, *10*(1), 151-159.

Li, H., Li, D., Dong, X., & Xu, N. (2019). Spatial patterns of 5A-level tourist attractions and their nework attention de-grees in China. *Journal of Arid Land Resources and Environment*, *33*(10), 179-184. http://dx.chinadoi.cn/10.13448/j.cnki.jalre.2019.305

Li, S., Xuqiu, R., & Chen, L. (2008). Cyberspace Attention of Tourist Attractions Based on Baidu Index: Temporal Distribution and Precursor Effect. *Geography and Geo-Information Science*, *24*(6), 102-107. http://dx.chinadoi.cn/CNKI:SUN:DLGT.0.2008-06-027

Long, M., Sun, G., & Long, Z. (2013). Tourist flow′s response to degree of consumer network attention to Zunyi tourism. *Geography and Geo- Information Science, 29*(5), 102-105. http://dx.chinadoi.cn/10.7702/dlydlxxkx20130521

Long, M., Sun, G., Ma, L., & Wang, J. (2011). An analysis on the variation between the degree of consumer attention of travel network and tourist flow in regional tourism: A Case of Sichuan province. Areal Research and Development, 30(3). http://dx.chinadoi.cn/10.3969/j.issn.1003-2363.2011.03.019

Lu, Z., Li, X., Yang, X., Zhang, Q., & Wang, S. (2010). Multiple time dimensions of visitors′ behavior based on the interactive function of tourism websites. *Economic Geography*, *30*(12), 2100-2103. http://dx.chinadoi.cn/10.15957/j.cnki.jjdl.2010.12.001

Lu, Z., Zhao, Y., Wu, S., & Han, B. (2007). The Time Distribution and Guide Analysis of Visiting Behavior ofTourism Website Users. *Acta Geographica Sinica*, *62*(6), 621-630. http://dx.chinadoi.cn/10.3321/j.issn:0375-5444.2007.06.007

Ma, L., Sun, G., Huang, Y., & Zhou, R. (2011). A correlative analysis on the relationship between domestic tourists and network attraction. *Economic Geography*, *31*(4), 680-685. http://dx.chinadoi.cn/10.15957/j.cnki.jjdl.2011.04.026

Ma, X., & Yang, X. (2003). A study on the 4A tourism area（spots）in China：spatial characters and industrial distribution. *Economic Geography*, *23*(5), 713-716. http://dx.chinadoi.cn/10.3969/j.issn.1000-8462.2003.05.029

Ma, X., Zhang, Z., Gong, W., Zhao, X., Guo, Y. (2023). Spatial pattern evolution and driving factors identification of manufacturing industry in LanZhou. *Scientia Geographica Sinica*, *43*(03), 519-529. https://doi.org/10.13249/j.cnki.sgs.2023.03.014

Mao, X., Song, J., & Yu, W. (2011). Space structure and its evolution of A-grade tourist attractions in Beijing. *Economic Geography*, *31*(8), 1381-1386. http://dx.chinadoi.cn/10.15957/j.cnki.jjdl.2011.08.025

Mitchell, L. S., & Smith, R. V. (1985). Recreational geography: inventory and prospect. *Professional Geographer*, *37*(1), 6-14. https://doi.org/10.1111/j.0033-0124.1985.00006.x

Nadotti, L., & Vannoni, V. (2019). Cultural and event tourism: an interpretative key for impact assessment[J]. *Eastern journal of European studies*, *10*(1), 115.

Nishijima, R. (2020). A Taiwanese pilgrim’s daytrip into the scenes of Your Name. *Journal of Tourism and Cultural Change*, *18*(1), 27-41. https://doi.org/10.1080/14766825.2020.1707462

Pan, B., Wu, D., & Song. H. (2012). Forecasting hotel room demand using search engine data. *Journal of Hospitality & Tourism Technology*, *3*(3), 196-210. https://doi.org/10.1108/17579881211264486

Pascariu, G. C., Ibanescu, B. C., Nijkamp, P., & Kourtit, K. (2021). Tourism and economic resilience: implications for regional policies, tourism and regional science. In S. Suzuki, K. Kourtit & P. Nijkamp (Eds.), *Tourism and Regional Science* (129-148), Berlin, Tokyo. http://dx.doi.org/10.1007/978-981-16-3623-3_8

Prosper, F. B., & Ryan, W. S. (2015). Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. *Tourism Management*, 46, 454-464. https://doi.org/10.1016/j.tourman.2014.07.014

Sun, S., Wang, S., Wei, Y., Yang, X., & Tsui, K. L. (2019). Forecasting tourist arrivals with machine learning and internet search index. *Tourism Management*, 70, 1-10. https://doi.org/10.1016/j.tourman.2018.07.010

Sun, Y., Zhang, H., Liu, P., & Zhang, J. (2017). Forecast of tourism flow volume of tourism attraction based on degree of tourist attraction of travel network: A case study of baidu index of different clients. *Human Geography*, *32*(3), 152-160. http://dx.chinadoi.cn/10.13959/j.issn.1003-2398.2017.03.020

Vuylsteke, A., Wen, Z., Baesens, B., & Poelmans, J. (2010). Consumers' Search for Information on the Internet: How and Why China Differs from Western Europe. *Journal of Interactive Marketing*, *24*(4), 309-331. https://doi.org/10.1016/j.intmar.2010.02.010

Wang, M., Chen, N., & Huang, H. (2013). Spatial distribution structure of 5A tourist attractions in China. *Geo spatial Information*, *11*(2). http://dx.chinadoi.cn/10.11709/j.issn.1672-4623.2013.02.032

Wang, K., Guo, F., Li, R., & Fu, X. (2014). Tourism attention degree about China from overseas and its spatial patterns based on Tripadvisor. *Progress in Geography*, *33*(11), 1462-1473. http://dx.chinadoi.cn/10.11820/dlkxjz.2014.11.004

Wei, Z., Chen, X., Liu, Y., Huang, S. (2023). Spatial distribution pattern of urban community group buying pickup points and its influencing factors: the casa of Guangzhou. *Economic Geography*, *43*(07), 109-118.

Wu, B., & Tang, Z. (2003). A study on spatial structure of national 4A grade tourism attractions in China. *Human Geography*, *18*(1), 1-5. http://dx.chinadoi.cn/10.3969/j.issn.1003-2398.2003.01.001

Yang, X., Pan, B., Evans, J. A., & Lv, B. (2015). Forecasting Chinese tourist volume with search engine data. *Tourism Management*, 46, 386-397. https://doi.org/10.1016/j.tourman.2014.07.019

Yuan, J., Yu, R., Liu, C., & Jiang, Y. (2010). The comparative study on status of energy saving of star hotel among different regions in Guangdong. *Economic Geography*, *30*(2), 325-328. http://dx.chinadoi.cn/10.15957/j.cnki.jjdl.2010.02.031

Zhang, H., Shi, T., & Bao, H. (2019). The Spatial Structure Characteristics of China's 5A-Level Tourist Attractions. *Journal of Huaqiao University* (Philosophy and Social Sciences), 4, 80-90. http://dx.chinadoi.cn/10.16067/j.cnki.35-1049/c.2019.04.010

Zhang, C., Weng, S., & Bao, J. (2020). The changes in the geographical patterns of China's tourism in 1978-2018: Characteristics and underlying factors. *Journal of Geographical Sciences*, *30*(3), 487-507. https://doi.org/10.1007/s11442-020-1739-2