



ARTICLE



<https://doi.org/10.1057/s41599-024-04262-0>

OPEN

Spatial and temporal evolution of tourism flows among 296 Chinese cities in the context of COVID-19: a study based on Baidu Index

Yibo Tang¹, Gangmin Weng^{1✉}, Simeng Qin² & Yue Pan¹

COVID-19 repeated outbreaks in China have significantly impacted residents' lives, particularly in the tourism sector. Accurately assessing the spatio-temporal evolution of tourism flow under the disturbance of COVID-19 is essential for scientific tourism planning and management. This study analyzes the tourism information search index of 296 cities in China to represent tourist flow tendencies and explores the distribution characteristics and formation reasons of tourism flow from both tourist source and destination perspectives. The results are as follows. (1) The COVID-19 has severely affected Chinese tourists' willingness to travel, with a decline of over 50% from 2020 to 2022 compared to 2019. Although a rebound is observed in 2023, it has not yet returned to pre-epidemic levels. (2) COVID-19 has impacted the spatial distribution pattern without altering the essential characteristic of "high in the East and low in the West." The distribution of core tourist sources is "bow-shaped", while that of popular destinations is "U-shaped". (3) In both tourism information inflow and outflow, the number of "head" and "tail" cities is relatively small, with most cities situated in the "middle" echelon. Additionally, the "Matthew effect" has emerged among tourist destination cities, which may aggravate the unbalanced nature of tourism distribution in China. Therefore, future tourism development should be strategically planned, respecting the principles of evolutionary patterns.

¹School of Economics and Management, Yanshan University, Qinhuangdao 066004 Hebei, China. ²School of Management, Northeastern University at Qinhuangdao, Qinhuangdao 066004 Hebei, China. ✉email: wenggangmin@ysu.edu.cn

Introduction

Tourism is often referred to as a “happiness industry” and plays an immeasurable role in the UN Sustainable Development Goals (World Tourism Organization 2022). Previous studies indicate that the tourism industry has experienced rapid, high-quality growth and has emerged as a critical driving force for global economic and social development (Danish and Wang 2018; Dogru and Bulut 2018). Recent reports from the World Travel & Tourism Council show that, prior to COVID-19, tourism supported approximately 10.3% of all jobs (334 million) and 10.4% of global GDP (\$10 trillion), while directly or indirectly generating nearly 1 in 5 new jobs worldwide. However, due to its inherent fragility and sensitivity, tourism is particularly vulnerable to external factors, especially the impacts of various emergencies (Bulchand-Gidumal 2022; Li et al. 2022).

Globally, tourism flows are constantly disrupted by the coronavirus pandemic (Gossling et al. 2021). According to the latest data from the United Nations World Tourism Organization, compared to 2019, the number of international tourists declined by 72% in 2020 and 70% in 2021, and began recovering in 2022; however, by 2023, it had only recovered to 88% of its original level (United Nations World Tourism Organization 2024). In various countries, accurately assessing temporal and spatial variations in tourism flows facilitates the implementation of differentiated planning and management to achieve sustainable tourism development (Matsuura and Saito 2022).

In particular, China, which experienced the earliest and most prolonged outbreak, has seen its tourism industry significantly impacted (Qiang MM, 2022; Hu and Zhu 2023). According to the latest data from the National Bureau of Statistics, compared to 2019, the number of domestic tourists in 2020, 2021 and 2022 fell by 52, 46 and 57%, respectively; however, it is encouraging to note that in 2023, Chinese tourism began to pick up, recovering to 81% of its original level (National Bureau of Statistics 2024). The year 2023 marks a full recovery for tourism. The removal of COVID-19 control measures, such as health codes and travel codes, is expected to significantly enhance people’s mobility, providing a positive boost to the tourism industry (Central People’s Government of the People’s Republic of China 2022). However, underlying challenges remain, as the popularity of cities depends on various factors. For instance, while Beijing and Shanghai consistently remain popular, Zibo and Xi’an have recently risen to prominence, whereas Lijiang, Nyingchi, and others have surprisingly declined in popularity. Thus, assessing the evolution of temporal and spatial characteristics from a macro perspective is highly beneficial for facilitating the orderly recovery of tourism.

Accordingly, this study presents the main contents and contributions as follows: Grounded in a national city-level perspective, it focuses on the unquied period of the COVID-19. From a broad-scope, small-scale and long-term perspective, this study constructs the tourism flow network among 296 cities in China from 2019 to 2023 based on Baidu Index. Utilizing GIS spatial analysis and other methods, this research explores the spatial distribution and agglomeration characteristics of tourism flow across China. Additionally, it examines both the inflow and outflow of tourism, clarifying the dual role of cities as sources and destinations for tourists to accurately evaluate their ability to generate and attract visitors, thereby providing a basis for scientifically formulating tourism plans. Furthermore, this study integrates theory with practice by incorporating policy analysis into the push-pull theory, thoroughly exploring the mechanisms of tourism flow formation in the context of the COVID-19 pandemic. In doing so, it reveals flow patterns previously overlooked in traditional research.

Literature review

The impact of COVID-19 on tourism flow. Tourism flow broadly encompasses passenger, logistics, technology, and information flows, which is an essential issue in both tourism economics and geography (Shoval and Isaacson 2007). Understanding and predicting the temporal and spatial characteristics of tourism flows can assist managers in identifying coordination or mismatches between supply and demand, thereby facilitating the formulation of effective tourism planning and management strategies (Mou et al. 2020; Chen et al. 2022; Zheng et al. 2022). Consequently, there is a substantial body of research on tourism flows, which is increasingly moving beyond static models to emphasize dynamic comparative analysis (Dou et al. 2024). Regarding research scope, the broadest perspective is international, while the narrowest focuses on specific scenic spots (Huang et al. 2017; Liu et al. 2022; Xu et al. 2022). From the perspective of the research content, most studies focus on the spatial and temporal distribution (Mou et al. 2020; Lau et al. 2016; Wang et al. 2021), network structure (Peng et al. 2016; Wang et al. 2021; Park et al. 2022; Tian et al. 2023), influence mechanism (Kim et al. 2016; Chen et al. 2022; Khalid et al. 2022; Li et al. 2022), and trend prediction of tourism flow (Schmöcker 2021; Zhang et al. 2021), and also focuses on the relationship between tourism flow and other systems, such as the regional economy (Ruan and Zhang 2021), population (Okafor et al. 2022), transportation (Alderighi and Gaggero 2019; Lin et al. 2022), and environment (Danish and Wang 2018; Paramati et al. 2017). From a methodological perspective, most studies primarily use social network analysis, econometric statistics, and geographical GIS analysis to examine tourism flows. These studies significantly contribute to clarifying the concept of tourism flow and establishing the research direction and paradigm of tourism flow.

In addition, informatization and digitization have had a revolutionary impact on the tourism industry, with information, as a new production factor, further participating in the optimization of the tourism market (Navio-Marco et al. 2018; Zhou and Chen 2021). With the rapid development of the Internet and big data, an increasing number of tourists are posting their itineraries and experiences online through texts, pictures, videos, and other formats (Kim et al. 2017). As a result, travelers are developing a habit of using social platforms to obtain travel information (Ho et al. 2012). To some extent, tourism information flows reflect trends in the flow of potential tourists as well as differences in attractiveness between destinations, serving as a barometer for tourist flows (Zillinger 2020).

Both tourism flow and tourism information flow are highly susceptible to external policies and crisis events (Jin et al. 2019). Especially since COVID-19, the mobility of tourists has been limited, resulting in the tourism information flow attracting increasing attention from scholars (Liu et al. 2019; Lu et al. 2022). According to the research paradigm, these studies can be classified into two types: theoretical research and empirical research (Yu et al. 2023). Theoretical studies are primarily explanatory analyses that explain the causality and mechanisms of changes in tourism flow under the influence of an epidemic (Rather 2021; Sharma et al. 2021; Sigala 2020). Empirical studies, on the other hand, are predominantly data-driven, and the quality of these studies depends on their data and models (Arbulú et al. 2021). The data sources for most of these studies are official statistical agencies, which are authoritative but often lag behind. In some studies, data sources are actively collected by researchers, including extracting routes from content posted by tourists (such as travel tips, travelogues), asking tourists to wear GPS sensors to record their routes, or directly interviewing tourists (Lewis et al. 2021; Xu et al. 2022). This type of data collection and processing

requires significant time and effort, limiting most of these studies to small samples (such as a single scenic spot or city) or short durations (like special holidays) (Mou et al. 2020; Weng and Pan 2021; Li et al. 2022; Park et al. 2022).

Research gaps analysis. Overall, most scholars have conducted a series of studies on the spatiotemporal changes in tourism flows; however, certain limitations remain:

First, previous scholars have devoted considerable attention to tourism flows, but their research has often focused on small to medium-sized regions, individual provinces or cities, or specific scenic areas. Research on tourism flows at the national level (broad scope), particularly at the city scale (small scale), has been relatively scarce. This scarcity is primarily due to challenges associated with obtaining large datasets, particularly regarding geographic spatial information. Thus, to address the shortcomings of statistical data (time-lag) and user-generated data (hard-earned), online big data (Baidu Index) is utilized to represent tourist mobility trends. This approach enables the study to observe the impact of COVID-19 on inter-city tourism flows in China over an extended period and from a broad perspective.

Second, tourism flow refers to the flow of tourists or tourism information between sources and destinations. Its distribution characteristics and network structure reflect the interaction among urban tourism industries. However, it is regrettable that current scholars focus more on the study of tourism flow volume, while further exploration of its direction requires enhancement. Ignoring the dual identity of tourist sources and destinations for each city will lead to incomplete tourism planning and unscientific tourism management. Therefore, this study uses the outflow of tourist information to represent the attractiveness of cities as destinations, and the inflow of tourist information to represent the potential of cities as sources. It analyzes the tourism development pattern of Chinese cities from this fresh perspective.

Third, due to the influence of COVID-19, existing studies focus, on the one hand, on tourists and their psychological and behavioral changes, while, on the other hand, they focus on the tourism industry, emphasizing the economic impact and countermeasures. However, research on the changes in tourism spatial structure under the normalization of COVID-19 is insufficient. Investigating the new dynamics in tourism flow resulting from the pandemic will help identify emerging hotspots in tourism demand and shape a new tourism supply pattern. Therefore, this study focuses on inter-city tourist mobility tendencies and thoroughly discusses the spatial pattern of China's tourism to guide differentiated tourism planning across various cities.

Methodology

Data sources and processing. This study focuses on the spatial-temporal evolution of tourism flows among all cities in China within the context of COVID-19. In the narrow sense, tourism flow mainly includes both tourist flow and tourist information flow (Shoval and Isaacson 2007). Unfortunately, during the COVID-19, many cities suspended their statistical operations, making it difficult to obtain actual tourist flow data at the city scale (Shao et al. 2020). However, previous studies have fully demonstrated a high correlation between tourist information flow and tourist flow (Sun et al. 2019; Lin et al. 2021). Therefore, this study uses the Baidu Index as the data source and the flow of inter-city tourism information as the entry point to analyze the impact of COVID-19 on China's tourism flow.

On the one hand, recent data show that China has 1.067 billion Internet users, with Baidu maintaining a commanding market share of 85.48% (China Internet Network Information Center 2023). Baidu Index is based on the search behavior of Internet users in

Baidu, which scientifically calculates the weighted sum of the frequency of each search keyword (Liu and Liao 2021). As a result, the Baidu Index, calculated from these extensive search records, can provide a timely and objective mapping of social hotspots (Cao et al. 2023; Yu and Jin 2022). On the other hand, Baidu is able to recognize user addresses and filter them to specific areas. Accordingly, by combining different keywords, we can obtain the search frequency of Internet users in a given city for information related to other cities; thus, it can not only characterize the flow but also indicate the direction of the flow (Ruan and Zhang 2021).

Although the Baidu Index cannot cover all the tourist information flow, it remains the most timely and comprehensive database available. It is feasible and representative to use the Baidu Index to measure tourist information flow between cities (Huang et al. 2017; Kang et al. 2022). Therefore, we selected the Baidu index as our primary database, to obtain a long-term (five-year), high-resolution (daily), and large-scale (296 cities) travel search index in bulk, enabling us to examine the impact of COVID-19 on China's tourism geographical pattern from a broader perspective. The following outlines the data collection and processing methods.

Firstly, the permanent population of Sansha City is just over 2000 people, which is too small compared to other cities, leading to the exclusion of the Baidu Index. Therefore, we selected 296 cities at the prefecture level and above as the research focus. Secondly, we utilized "city + tourism" as the keyword to obtain daily data on inter-city tourism information searches from January 1, 2019, to December 31, 2023, in bulk. Thirdly, the data were retrieved and calculated using PyCharm software to obtain annual and monthly averages. Finally, according to the direction of flow, the outflow index (O) and inflow index (I) of tourism information of a certain city were summarized. The data matrix is presented below.

$$T = \begin{matrix} & \begin{matrix} I_1 & I_2 & \dots & I_{j-1} & I_j \end{matrix} \\ \begin{matrix} O_1 \\ O_2 \\ \vdots \\ O_{i-1} \\ O_i \end{matrix} & \begin{bmatrix} T_{11} & T_{12} & \dots & T_{1(j-1)} & T_{1j} \\ T_{21} & T_{22} & \dots & T_{2(j-1)} & T_{2j} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ T_{(i-1)1} & T_{(i-1)2} & \dots & T_{(i-1)(j-1)} & T_{(i-1)j} \\ T_{i1} & T_{i2} & \dots & T_{i(j-1)} & T_{ij} \end{bmatrix} \end{matrix} \quad (1)$$

In the above formula, T_{ij} represents the volume of tourism information searches from city j to city i . O_i denotes the total outflow of tourism information from city i ; that is, the sum of tourism searches from other cities directed at city i , representing the tourism attraction of city i as a destination. I_j denotes the total inflow of tourism information to j city; that is, the sum of tourism searches conducted by city j for other cities, which is used to represent the travel potential of city j as a tourist source.

Research method

Jenks classification and Kriging interpolation. ArcGIS is the most widely utilized geographic information system software, known for its powerful capabilities in spatial data storage, processing, and analysis, primarily for visualization and analysis (Zhang et al. 2019). In this study, Jenks natural breakpoint method and the Kriging interpolation method are employed to identify the spatiotemporal pattern evolution features of tourism information flows. Jenks identifies classification intervals based on the inherent characteristics of the data, maximizing the differences between categories; thus, this study applies this method to classify tourism information data (Bai et al. 2022). On this basis, ordinary kriging interpolation is employed to fit the spatial relationships of

tourism information based on variance functions, combined with unbiased estimation and optimal estimation (Paparrizos et al. 2018).

Rank-size law. The rank-size law is a concept within fractal theory that accurately reveals the relationship between the scale and order of tourism flows (Lau et al. 2016). The research approach is as follows: rank the outflow (inflow) of tourism information from cities in descending order, and derive the relationship between scale and ranking. If a straight-line segment appears on the scale-free space, it indicates that the rank-size law is satisfied, and the slope of the line is the Zipf index (q), which describes the structural differences within the city system (Guo et al. 2016; Xu 2022). The mathematical model for Zipf is:

$$\ln P_K = \ln P_1 - q \ln K \quad (2)$$

Here, P_k represents the amount of tourism information in a city, K denotes the city's order among all cities, P_1 indicates the tourist information of the first city under ideal conditions, and q is the Zipf index, also known as the concentration index.

The scale distribution of tourism information can be divided into three categories based on the value of q : (1) when $q \geq 1.2$, this is the "primacy type", indicating that high-order cities significantly influence tourist mobility trends, while middle-low cities remain undeveloped, potentially leading to polarization phenomena. (2) When $0.85 < q < 1.2$, this is the "centralized type", where the ideal state occurs when $q = 1$, indicating that the distribution structure of tourism information is optimal. (3) When $q \leq 0.85$, this is the "dispersed type", depicting that the structure of cities in each level is considerably fragmented, and cities of the "primacy type" lack the radiation-driving ability.

Kernel density estimation. Kernel density estimation (KDE) is a non-parametric estimation method that, unlike traditional estimation techniques, does not require prior assumptions about the distribution type or specific parameters (Zhao and Zhu 2024). Consequently, it has lower dependence on the data and model, demonstrating strong robustness, and is widely used in studies of

dynamic evolution (Wang et al. 2019). It utilizes probability density functions to generate continuous curves, characterizing the evolutionary trends of random variables through variations in the curve's position, shape and ductility. The calculation formula is presented as follows:

$$f(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x_i - \bar{x}}{h}\right) \quad (3)$$

Where n represents the number of cities, with $n = 296$. x_i denotes the logarithmic value of tourism information flow, while \bar{x} represents the mean value across all cities. $K(\cdot)$ represents the kernel density function, with common examples including the Epanechnikov kernel and Gaussian kernel, etc. Additionally, h refers to the bandwidth, the smaller the bandwidth, the more accurate the estimation. However, the number of samples involved in the calculation will also be correspondingly reduced, leading to a less smooth curve. Therefore, to enhance the accuracy of the statistical tests, the kernel functions and bandwidths were automatically generated using Stata software.

Results and discussion

Analysis of tourism flow driving mechanisms under COVID-19 perturbation. The push-pull theory has a natural explanatory adaptability to the formation of tourism flow, dividing the driving forces of tourism inflow and outflow into two aspects: the intrinsic push from tourism demand at the source and the extrinsic pull from tourism supply at the destination (Lewis et al. 2021; Ayoub and Mohamed 2024). Meanwhile, considering the frequent disturbance of tourism flow caused by the COVID-19, the current policy analysis is added to the traditional analysis of push and pull theory, and the driving mechanism is constructed from the three dimensions of "push-pull-resistance" (Fig. 1).

The push from the tourist source is the driving force behind tourism flows. The population size and economic development determine the quantity and quality of potential tourists, while the willingness and purchasing power of tourists form the basis of tourist flow (Liu and Liao 2021). The convenience of transportation is a necessary condition for determining whether travel can

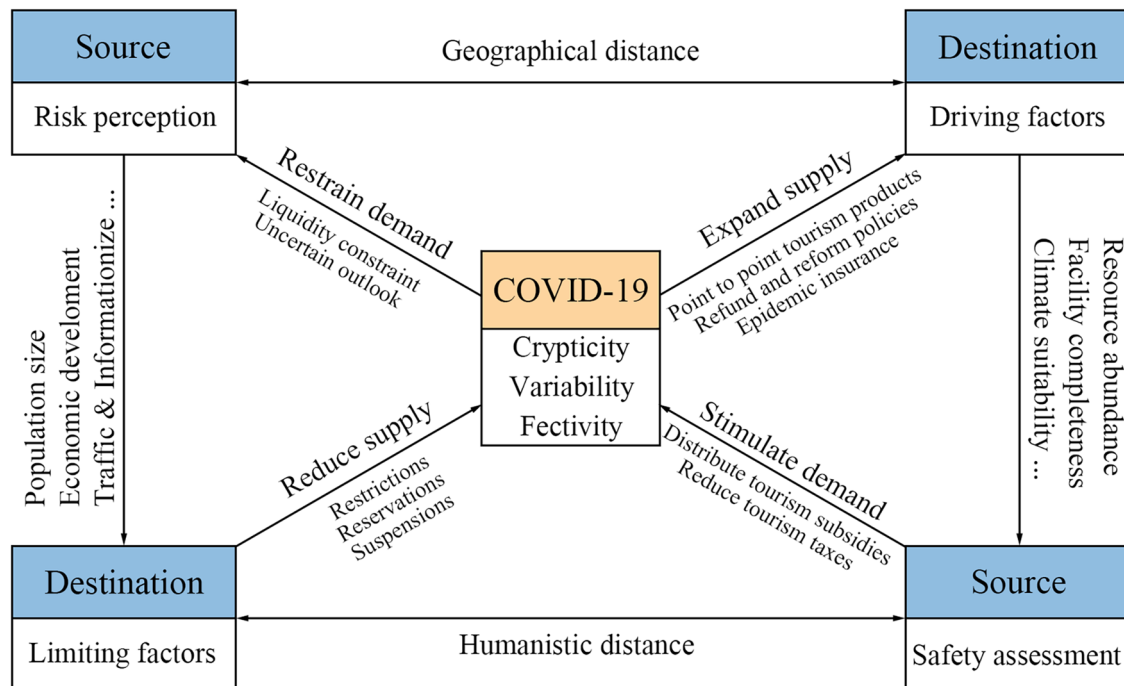


Fig. 1 The driving mechanism of tourism flow. The figure shows the formation mechanism of tourism flows in the context of COVID-19.

be completed, and informatization and digitization serve as technical guarantees for the dissemination of tourism information (Ruan and Zhang 2021; Shi et al. 2022).

The pull from the tourist destination is another driving force behind tourism flows. The abundance of resources directly determines the attractiveness of a destination and market awareness, while the availability of facilities reflects the destination's capacity to accommodate tourists and its level of hospitality (Mou et al. 2020). In addition, climate suitability is an essential factor influencing tourists' decisions to travel (Medina et al. 2022).

Resistance factors between the source and destination are the controlling forces of tourism flows. According to distance decay theory, both spatial and cultural distances between the source and destination are critical factors determining the scope and direction of tourism flow (Tan et al. 2022). Among these factors, spatial distance primarily affects tourism decision-making under objective conditions. This is because, as spatial distance increases, so do transportation, accommodation, and other expenses. However, tourists have limited affordability and leisure time. Conversely, cultural distance primarily affects tourism decision-making under subjective conditions. Although cultural distance can satisfy tourists' needs for novelty and difference, it is important to note that as cultural distance increases, travelers' sense of unfamiliarity with the destination intensifies. Out of a desire to avoid risks, their motivation to travel is subsequently weakened (Ma and Tong 2019).

Moreover, risk perception and safety assessments of non-habitual environments significantly constrain tourism flows (Xie et al. 2023; Park et al. 2022). These perceptions become particularly critical during periods of frequent policy adjustments, which can exacerbate uncertainties and influence travelers' decision-making (Wu et al. 2023). Therefore, this study closely integrates theory and practice by incorporating relevant policy analysis into push-pull theory to further refine the mechanisms of tourism flow formation. This integration not only provides a more comprehensive understanding of the factors driving tourism flows but also enhances the applicability of push-pull theory in dynamic and uncertain environments.

Specifically, in December 2019, the COVID-19 pandemic emerged in Wuhan, China. The National Health Commission of China responded swiftly by initiating an epidemiological investigation and announcing a "lockdown" in January 2020. Other cities soon followed suit similar lockdown measures. In 2020, during the initial phase of the pandemic, community lockdowns, venue closures, and traffic disruptions were implemented, greatly restricting mobility and gatherings, directly choking the "throat" of the tourism industry (Hao et al. 2021). By 2021, with the nationwide mutual recognition of health codes and the adoption of targeted, zone-based control measures, production and daily life gradually resumed, and the tourism industry began to make efforts toward recovery. However, due to recurrent outbreaks and policy uncertainties, the willingness to travel among Chinese tourists remained low during this period. Moving into 2022, with the widespread deployment of vaccines, China adopted a "dynamic zero-COVID" policy. Although this policy was more flexible, rapid control measures were implemented in any city experiencing an outbreak, significantly affecting the continuity of tourism and hindering the full recovery of the industry (Yu et al. 2023; Fan et al. 2023). By 2023, China officially optimized its pandemic prevention policies, downgrading COVID-19 from "Class A, Category I" to "Class B, Category II" management. This shift marked the beginning of a reopening phase, characterized by a significant relaxation of various prevention measures and a comprehensive restart of the economy and society (Dou et al. 2024). As a result, the tourism industry capitalized on this opportunity, establishing the groundwork for a robust recovery after years of intermittent challenges.

Meanwhile, in response to the impact of COVID-19, the Chinese tourism industry actively implemented a series of remedial measures. In 2020, facing widespread lockdowns and travel restrictions, the industry sought to reduce operational costs through various measures, such as limiting visitor numbers, requiring reservations, and even suspending operations (Li et al. 2021). To preserve customer goodwill and prepare for future recovery, businesses negotiated with clients, offered refunds, and managed cancellations (Liu et al. 2021). As the pandemic gradually came under control in 2021 and 2022, though with occasional regional resurgences, the industry innovatively introduced "Safe Travel," focusing on flexible product combinations (e.g., point-to-point intercity travel), simplified transportation methods (e.g., self-driving), and robust cancellation policies and pandemic insurance to alleviate travelers' concerns (Qiang MM 2022; Zhang et al. 2023). Concurrently, the government increased support by issuing tourism subsidies, reducing tourism taxes, and other measures to stimulate tourism demand and ensure supply (Wong and Lai 2022). Despite these efforts, the recurring nature of the pandemic led to intermittent disruptions in the tourism sector. By 2023, with a significant relaxation of pandemic prevention policies, both tourism enterprises and local governments intensified their market promotion efforts. This coordinated push led to a rapid rebound in interprovincial and intercity travel, setting the stage for a robust recovery in the industry.

To sum up, these policy shifts, therefore, serve as crucial intermediaries in the push-pull dynamics, shaping the balance between tourism demand in the tourist source and tourism supply in the tourist destination, while simultaneously moderating resistance factors that influence travel decisions. As the epidemic draws to a close, cities are now faced with the challenge of leveraging these dynamics to ensure tourists can access reliable travel information, rebuild confidence in travel, and stimulate tourism motivation. In this context, understanding the temporal evolution and spatial distribution of tourism flows between cities in China becomes very essential.

Temporal changes of travel information search. Since 2019, the repeated outbreaks of COVID-19 have significantly impacted the tourism industry, especially in China, where the pandemic began the earliest and lasted the longest (Li et al. 2022). Figure 2 illustrates the daily average of tourism information searches and the

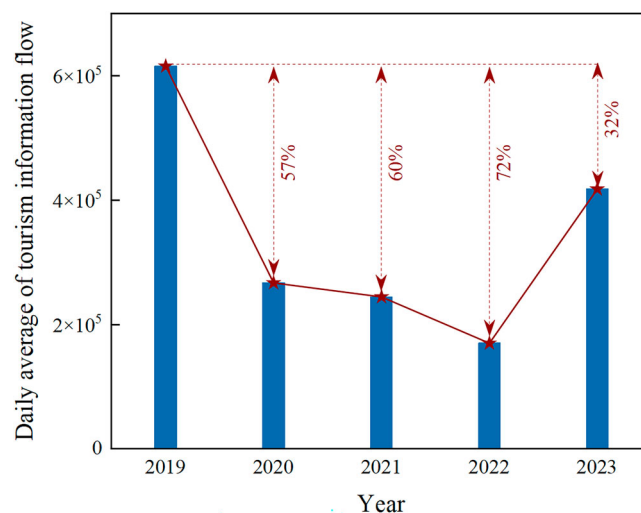


Fig. 2 The daily average of tourism information searches and the damage rates in each year from 2019 to 2023. The figure shows that tourism information searches declined by more than 50% in 2020–2022 and rebounded in 2023.

corresponding damage rates for each year from 2019 to 2023. It is important to note that using the tourism information search volume in 2019 as the initial value, and the search volume from 2020 to 2023 as the current value, the damage rate can be calculated as: $\text{damage rate} = (\text{initial value} - \text{current value}) / \text{initial value} * 100$. This allows for an analysis of the impact of COVID-19 on tourism flow. The data shows that before the COVID-19 pandemic, the annual average of tourism information searches reached as high as over 60,000. However, the initial phase of the pandemic in 2020 led to a sharp decrease in tourism-related searches, resulting in a decline of 57%. This demonstrates the strong constraints imposed on tourism flows by strict control policies. As policies were continuously optimized, the pandemic was brought under control in 2021 and 2022, resulting in a smaller decline in tourism information searches; however, they remained 60 and 72% lower than the initial levels in 2019. This suggests that the “partial silence and dynamic clearance” policy had a positive effect, although it was insufficient to reverse the downward trend in tourism. By 2023, with the significant relaxation of COVID-19 prevention measures and the shift towards reopening, there was a noticeable recovery in tourism information searches. However, while there was a positive uptick compared to previous years, search volumes remain at 68% of pre-pandemic levels, reflecting the lingering effects of the pandemic on consumer confidence and travel behavior. Additionally, from 2019 to 2023, the trend in China’s annual travel information search volume is largely consistent with the trend in the number of domestic tourists. This further indicates that, to some extent, it is scientifically valid to use tourism information flow as a representation of tourism flow.

In Fig. 3, the tourism information search index is calculated as the monthly average, with the dotted line representing the average value of tourism information flow across three time periods: before COVID-19 (2019), during COVID-19 (2020–2022), and after COVID-19 (2023). As demonstrated in Fig. 3, there were three peaks in tourism information searches in China prior to COVID-19, appearing in February, April, and July, respectively.

Due to the precursor effect of tourism information, it can be reasonably inferred that there is clear seasonality in China’s tourism, with peak seasons concentrated around the Spring Festival, May Day, and Summer Vacation (Li et al. 2023). Additionally, the months whose daily average in this month fell below the daily average in the year were mostly concentrated from September to December, which is the winter with cold weather. During this period, the number of national tourist information searches decreased significantly, indicating a close correlation between tourism activities and climatic conditions (Zhong and Chen 2019).

During the outbreak of COVID-19 in Wuhan from January to March 2020, the volume of tourism information searches dropped by half due to panic. Consequently, tourism information in February plummeted, marking a transition from peak to trough and entering a prolonged downturn. It resumed by September 2020, becoming the only peak of the year, signaling the National Day Golden Week. In 2021, two anomalous months were observed: February and August, attributable to the COVID-19 outbreaks in North China during January and February and in Hunan-Jiangsu during July and August. In 2022, there was only one peak in July, attributed to the COVID-19 outbreak in Shaanxi during January and February and in Shanghai from March to May. All of these outbreaks were prolonged (lasting more than one month) and widespread (spanning over ten provinces), once again underscoring the significant negative impact of COVID-19 on tourism and tourism information searches (Sun and Xing 2023). Additionally, the volatility trend of the tourism information index in 2023 mirrors patterns observed prior to the epidemic, albeit with a lower peak value. This suggests a cautious resurgence in the willingness of Chinese tourists to travel. While the temporal effects of COVID-19 on tourism flow have been tentatively identified, its impact on the spatial distribution of tourism in China remains ambiguous. Further empirical studies are necessary to address these questions and provide a comprehensive understanding of spatial implications.

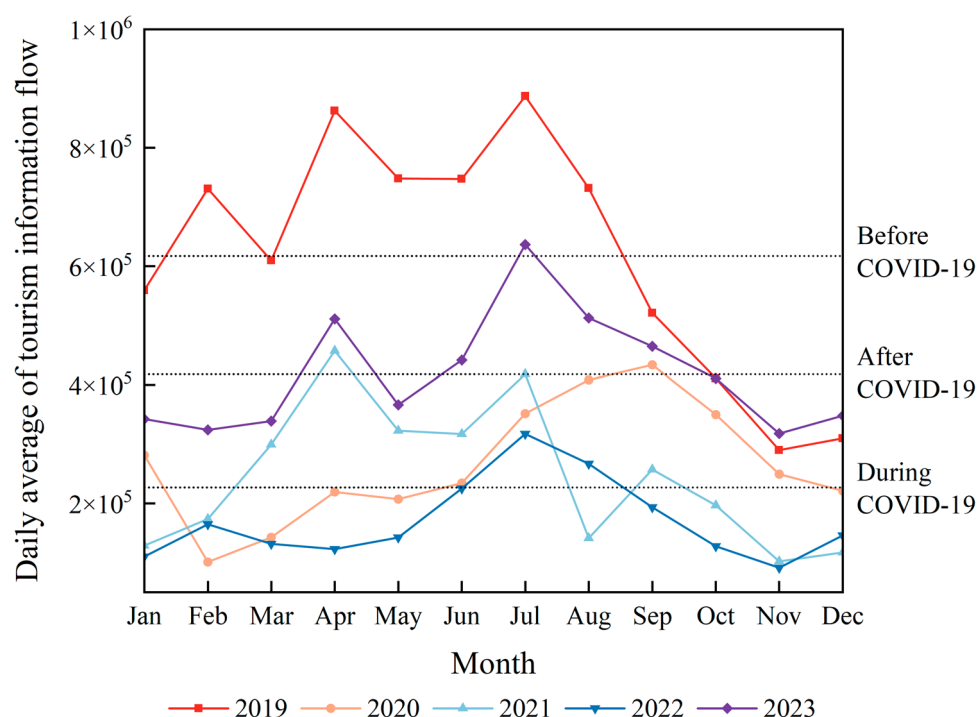


Fig. 3 The daily average of tourism information searches in each month from 2019 to 2023. The figure shows that tourism information searches in China exhibited seasonality, with a clear disruption in recent years due to the impact of the pandemic.

Spatial distribution of tourist flow tendency. Based on the push-pull theory, we classify city attributes into tourist source and destination cities, analyzing the spatial patterns of tourism flows in China amid the pandemic's perturbations from a dual perspective. Additionally, three specific periods—before COVID-19, during COVID-19, and after COVID-19—are selected to illustrate the temporal evolution of tourism flows in China (Figs. 4–6).

Regarding the travel potential of tourist sources (Fig. 4), China's tourism source market has fully developed a “bow-shaped” spatial distribution structure prior to COVID-19. Specifically, it is centered on three high-value agglomerations: Beijing-Tianjin-Hebei, the Yangtze River Delta, and the Pearl River Delta, with the coastal areas forming the arc, the Beijing-Guangzhou link as the string, and the Yangtze River as the arrow. Although the coverage of core tourist source areas contracted during COVID-19, there was a recovery and expansion after the pandemic.

The three core regions are not only the key development zones for China's regional economy but also the primary promotion zones for China's new urbanization, which have already become the most prosperous cities in China. Consequently, these regions can respond quickly and effectively to public health emergencies, such as COVID-19. Most cities in these regions have relatively high-quality economies with high levels of technologization, informatization, and internationalization, as well as a strong supportive capacity for the Internet economy and modern service industry. This enables them to maintain significant tourism potential, even during the COVID-19 pandemic, reflecting the positive impact of a highly active economy on the tourism source market.

Regarding the attractiveness of destinations (Fig. 5), prior to COVID-19, China's tourism hotspots formed a “U-shaped” spatial distribution structure, with Beijing-Tianjin-Hebei, the Yangtze River Delta and Chengdu-Chongqing as the core. The attractiveness of tourism destinations in the western and north-eastern regions is relatively insufficient, while the central region exhibits a characteristic of “local collapse.” During and after COVID-19, tourism destinations exhibited a Matthew effect, where the strong became stronger while the weak became weaker. In addition to the original popular areas, hotspots expanded northward into Inner Mongolia and extended southward into Hainan. Similarly, the coverage of the cold spot region in the west and northeast has expanded, demonstrating the poor resilience and the low risk tolerance of these less developed regions in the face of the epidemic, which requires improvement.

The Beijing-Tianjin-Hebei region, the Jiangsu-Zhejiang-Shanghai region and the Sichuan-Chongqing region in China possess abundant and high-quality tourism resources, making them popular tourism destinations (Fang et al. 2022). However, for the western provinces and the three northeastern provinces, tourism attention is constrained by factors such as location, transportation, and infrastructure. These regions should concentrate on overcoming the challenges posed by inadequate infrastructure and the seasonality affecting their tourism development.

In the six central provinces, the collapse area of tourism attraction is mainly concentrated in the north of Anhui, southeast of Henan, and southwest of Shandong, which overlaps with the “Yellow River Flooding Area” (Wang et al. 2020). The reason for the overlaps is the superimposed of ecological environment constraints, economic development dilemmas, and other unfavorable factors, which also result in the weakness of the regional tourism economy. Furthermore, coupled with the homogenization, it further reflects the “central collapse” in tourism attention.

Given the complexity of the national tourism flow spatial network, this study selects the core routes with the top 500 traffic

volume for visual analysis to further demonstrate the flow characteristics of Chinese tourists (Fig. 6). The core routes of China's tourism are distributed in a diamond shape, with Beijing, Shanghai, Sanya, and Chongqing at the four corners, and the cross intersects with Xi'an. Within this diamond structure, a secondary network has formed to connect Chengdu, Hangzhou, Tianjin, Nanjing, Shenzhen, Zhangjiajie and other regional core cities. This pattern was not disrupted during the pandemic; instead, it intensified. Additionally, in 2023, Lhasa emerged as a new destination outside the existing diamond-shaped structure of China's core tourism flows. This shift is attributed to the fact that, following the high-pressure period of the pandemic, people are increasingly seeking to alleviate stress through travel (Fan et al. 2023). With its expansive natural landscapes and rich cultural heritage, Lhasa has become an ideal choice.

In general, regardless of the inflow and outflow of tourism information, the overall spatial patterns closely resemble the economic and population distribution (Zhang et al. 2023). Specifically, a “southeast-northwest” high and low-value spatial distribution exists, with the “Hu Huanyong Line” serving as the boundary. This indicates that city strength remains the primary factor driving the spatial differentiation of tourism. Additionally, a spatial dislocation exists between tourist sources and destinations, suggesting that factors such as resource abundance and convenient transportation are also critical in affecting the spatial differentiation of tourism. In summary, COVID-19 has impacted the spatial patterns of city tourism in China but has not altered the essential characteristics.

Characteristics of the structural distribution of tourist cities.

To further analyze the concentration of tourism flows, this paper ranks the volumes of inflow and outflow of tourism information from each city and explores the evolutionary characteristics of the distribution structure of tourist cities during COVID-19.

Figure 7 illustrates that before COVID-19, the Zipf index for population size (number of permanent residents) and economic development (per capita GDP) in China's cities was 0.6670 and 0.5023, respectively. This indicates that the economic development and population size of cities in China remain dispersed and that development is unbalanced and insufficient. Additionally, the Zipf index for tourism information inflow and outflow was 0.9503 and 1.1882, respectively. This indicates that tourism information inflow is relatively reasonable, approaching an ideal equilibrium state. And the Zipf index for tourism information outflow is higher than the others, indicating a primacy distribution.

Tourism motivation theory posits that the transition from tourism willingness to behavior is influenced by various factors, including consumption capacity, motivational preferences, resource endowments, transportation, and climate. To maximize the utility of tourists, tourist destinations will undoubtedly concentrate on cities with unique resources, convenient transportation, pleasant climates, and comprehensive facilities. The highly integrative nature of tourism also facilitates spatial agglomeration of tourism flows, which surpasses that of the economy and population, making it possible to gap the huge discrepancy of tourism development in different regions.

As shown in Fig. 8, the rank-size curve shifts left annually from 2019 to 2022 and shifts right for the first time in 2023. Among these, the distance between the curves for 2020, 2021, and 2022 is relatively narrow, reflecting a continued slower decline in tourism popularity during the normalization period of COVID-19. In addition, the Zipf index for tourism information inflows rose from 0.9503 to 1.1194 before subsequently declining to 0.8736, indicating a fluctuating trend. This suggests that the balanced distribution of potential tourism sources was disrupted during the

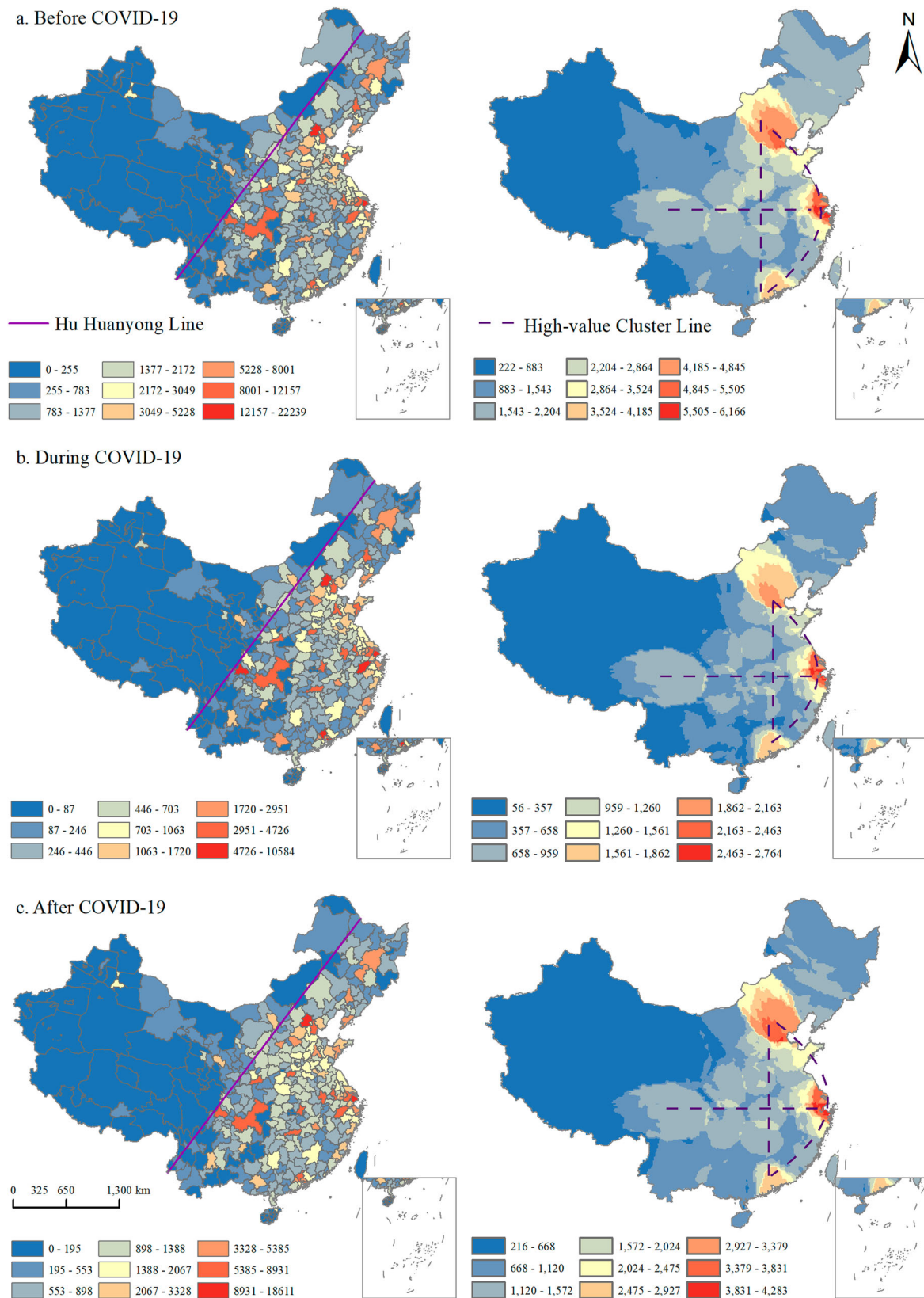


Fig. 4 Evolution of the spatial pattern of tourism information inflow. The figure shows the spatial distribution of tourism information inflow before, during and after the COVID-19 respectively from top to bottom, reflecting the changes in China's core tourist sources.

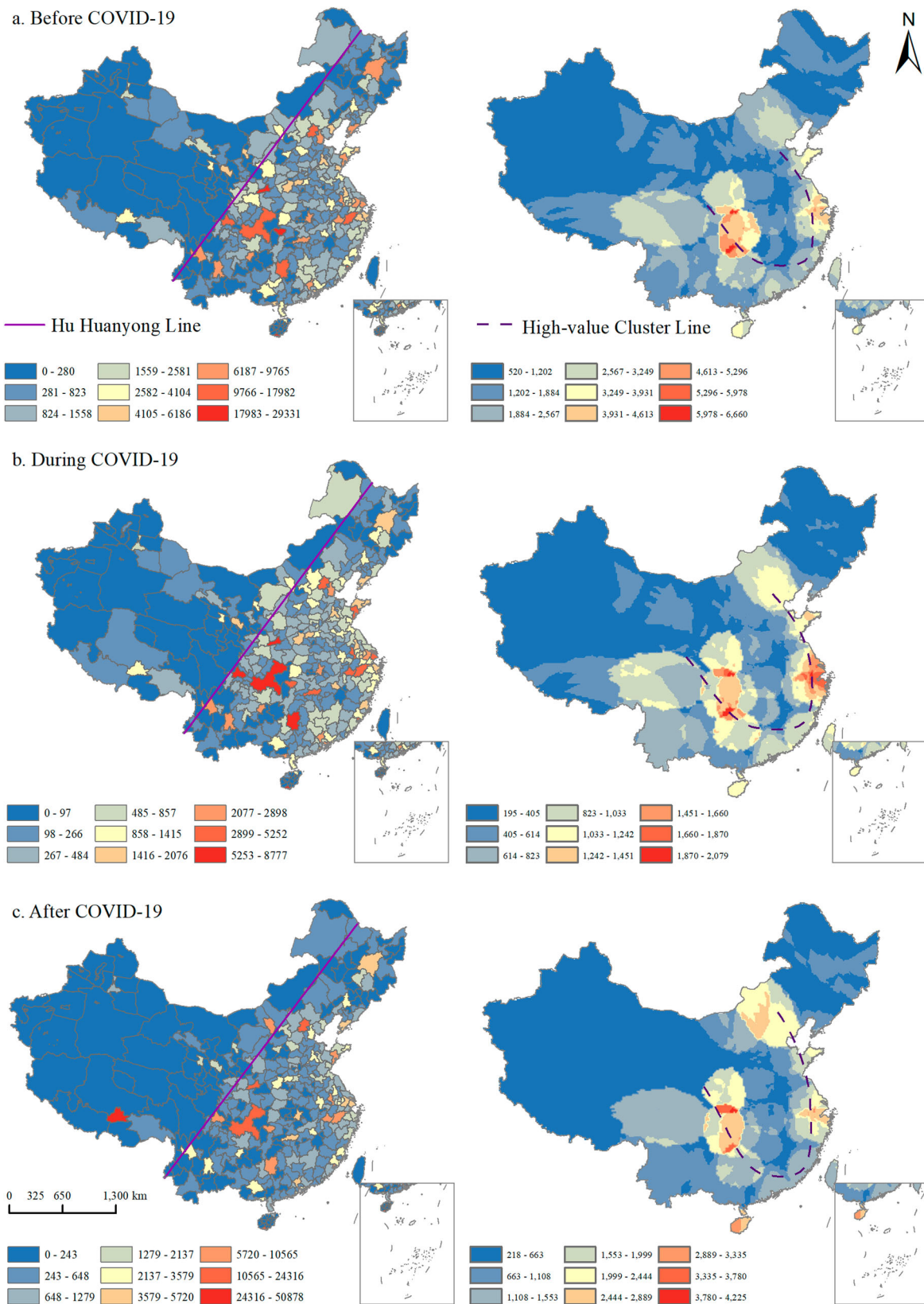


Fig. 5 Evolution of the spatial pattern of tourism information outflow. The figure shows the spatial distribution of tourism information outflow before, during and after the COVID-19 respectively from top to bottom, reflecting the changes of popular destinations in China.

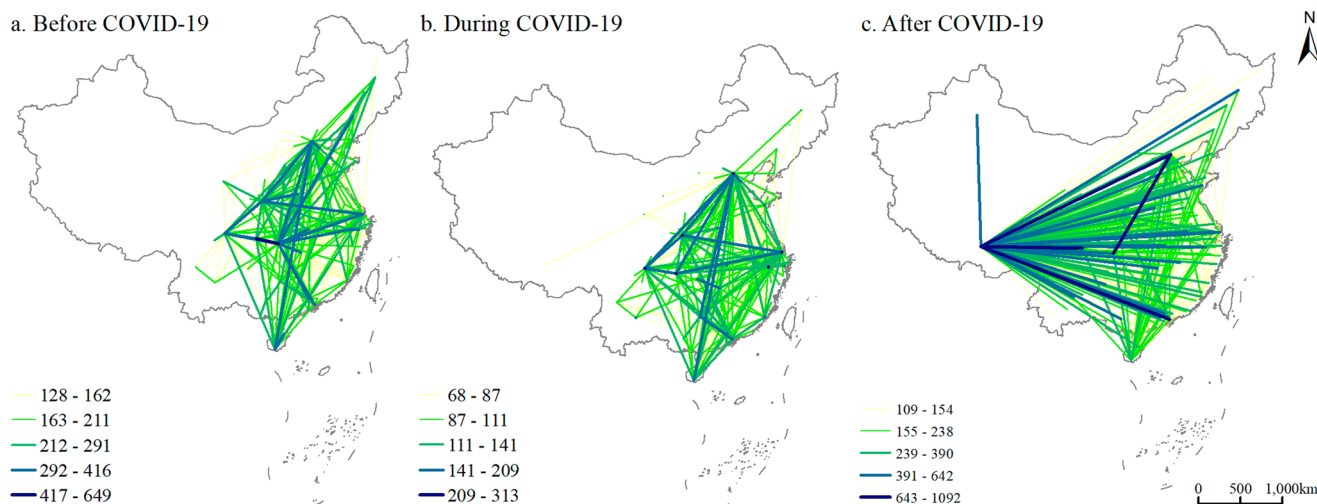


Fig. 6 Tourism flow spatial network structure. The figure shows from left to right the core travel routes of China's top 500 before, during and after the COVID-19.

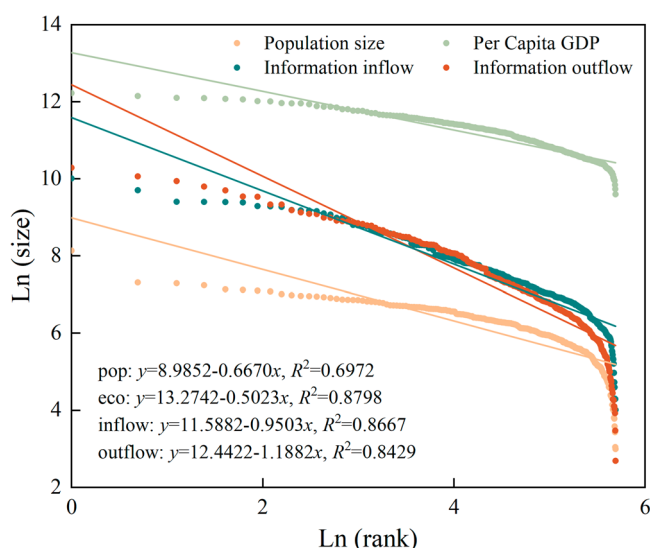


Fig. 7 2019 National population, economy, tourism information flow rank-size. The figure shows the rank-size chart of China's population, economy, tourism information inflow and outflow in 2019, with Zipf index of 0.6670, 0.5023, 0.9503, and 1.1882, respectively.

pandemic, shifting towards high-level cities before returning to normal. Meanwhile, the Zipf index of tourism information outflows consistently fluctuated around 1.2, indicating that the primacy of major tourism destinations remained stable.

Based on the intersection of scatter points and straight lines, cities can be roughly classified into three types: “head,” “middle,” and “tail.” As seen in Fig. 8, both information outflow and inflow exhibit similar patterns. In both inflow and outflow, the “head” cities exhibit a discrete distribution, with both quantity and quality requiring further improvement. A large number of cities are clustered in the middle of the curve with minimal spacing, reflecting the weak hierarchy of cities in the “middle.” Cities at the bottom of the ranking, such as Haidong and Shannan, are primarily underdeveloped areas with lagging local economic and social development. They curve downwards almost vertically, indicating that cities in the “tail” receive low attention in tourism.

To directly reflect the distribution and differences in tourism flows before and after COVID-19, a kernel density estimation plot (Fig. 9) is presented. As shown in Fig. 9, first, regarding the

position of the curve, whether for the inflow or outflow of tourism information, the main body of the kernel density curve is shifting to the left. This further confirms that China's urban tourism has not yet recovered to its pre-COVID-19 state, indicating that further efforts are needed. Second, regarding the shape of the curve, both the inflow and outflow of tourism information exhibit a unimodal distribution, with the peak becoming higher and narrower. This indicates a decrease in the standard deviation (representing the difference between data points and the mean) of tourism information inflow and outflow intensity, suggesting that after COVID-19, the cities in the “middle” of the chart above are more crowded. It is essential to consider how these cities can differentiate themselves in future competition. Finally, regarding the ductility of curve, the starting point of the kernel density curve of tourism information inflow moves to the right, while the end point moves to the left, indicating a clustering trend. This reflects a decrease in the extreme deviation (representing the difference between the maximum and minimum values) of tourism information inflow following COVID-19. Conversely, the starting point of the kernel density curve of tourism information outflow shifts to the left while the end point shifts to the right, indicating an expansion trend. This reflects an increase in the extreme deviation of tourism information outflow following COVID-19. This further confirms the aforementioned Matthew effect observed in tourist destinations, where the strong become stronger and the weak become weaker.

Based on the above classification and analysis, a Sankey diagram (Fig. 10) is presented to further illustrate the distribution and flow of tourism information among the “head,” “middle,” and “tail” cities. As illustrated in Fig. 10, after COVID-19, the inflow ratio of tourism information increased in both the “head” echelon and the “tail” echelon, although the change was relatively modest, rising from 25.6% to 26.2% and from 4.7% to 5.7%, respectively. In contrast, regarding the outflow ratio of tourism information, only the “head” echelon experienced an increase, rising significantly from 40.2 to 55%. This indicates that cities with strong tourism foundations are more resilient to external shocks.

Specific to individual cities, Fig. 11 clearly shows that, among the 296 cities in China, few cities can combine the attributes of both tourist source and destination. Typical examples include Chengdu, Hangzhou, Chongqing, and Qingdao, all of which are “new first-tier cities” with solid economic foundations and rich cultural heritage. Therefore, these cities should maintain their

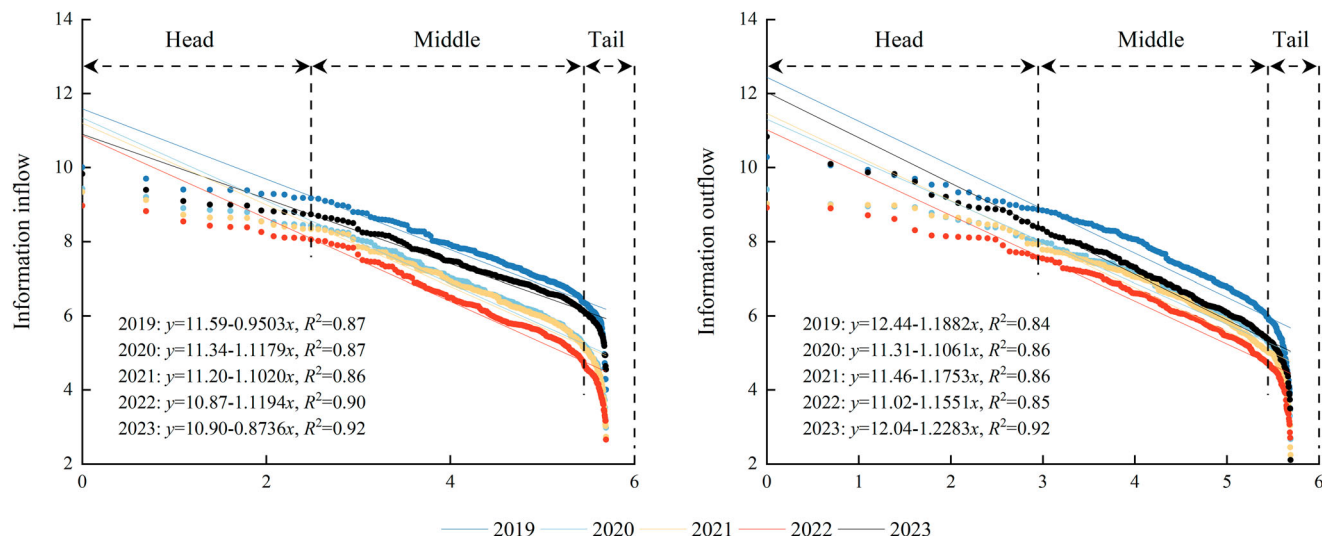


Fig. 8 Tourist information flow rank-size. The left side of the figure shows the rank-size chart of tourism information inflow from 2019 to 2023, while the right side shows the rank-size chart of tourism information outflow from 2019 to 2023.

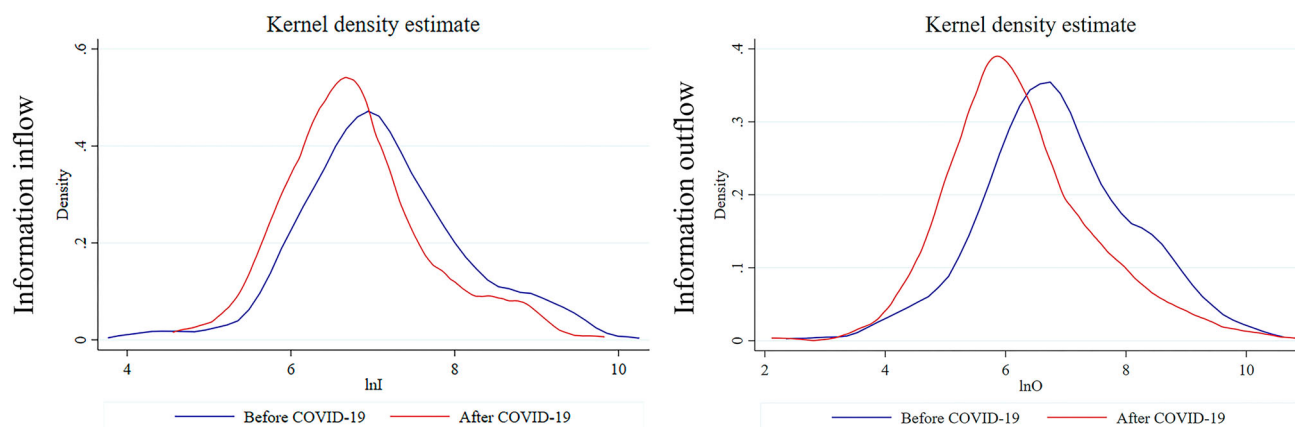


Fig. 9 Kernel density estimation of tourism information flow. The left side of the figure shows the estimated kernel density of tourism information inflow before and after the COVID-19, while the right side shows the estimated kernel density of tourism information outflow before and after the COVID-19.

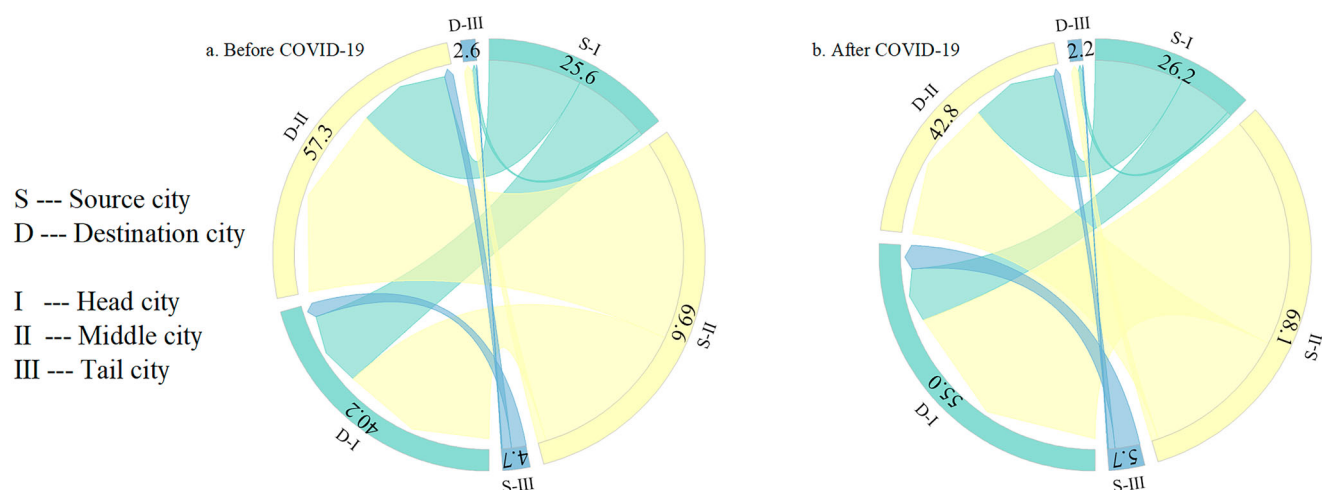


Fig. 10 Sankey diagram of tourism information flow. The left side of the figure shows the proportion of tourism information flow from different types of tourist sources to destinations before the COVID-19, while the right side shows the proportion after the COVID-19.

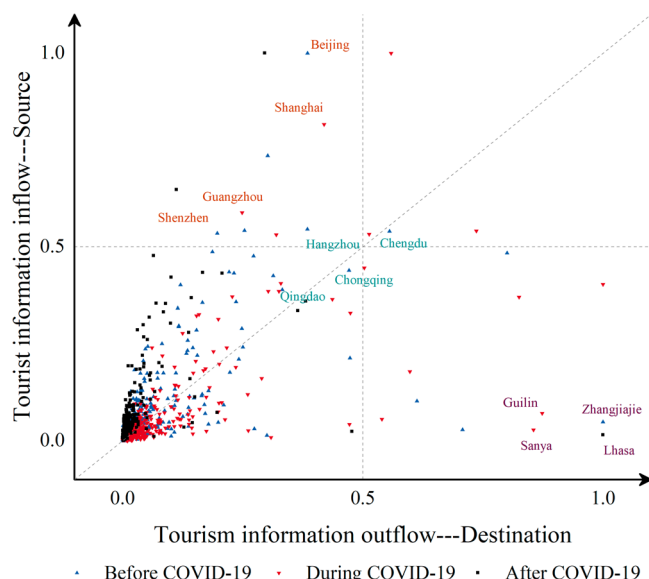


Fig. 11 The evolution of the source and destination property of 296 cities in China. The figure shows the tourism attributes of 296 cities in China.

current status while striving for higher quality. Additionally, some prosperous “first-tier cities” with strong single attribute, including Beijing, Shanghai, Guangzhou and Shenzhen, exhibit a higher tourist source attribute than destination attribute. In the future, the advantages conferred by a prosperous economy will significantly drive tourism development, thereby enhancing its attractiveness to tourists. Guilin, Sanya, Zhangjiajie, and Lhasa are cities where the tourist destination attribute is stronger than source attribute; they are classified as “third-tier,” “fourth-tier,” or even “fifth-tier” cities, with relatively underdeveloped economies. Thus, the tourism industry in these cities can serve as a critical support mechanism for economic development.

Conclusion, recommendation, limitations, and future research

In line with the fact that China conducted prevention measures for COVID-19 considering “city” as a unit, this paper examines 296 cities in China as the research area, focusing on tourist mobility tendencies as the subject of study, and spans the period from 2019 to 2023 to explore the temporal and spatial evolution characteristics and distributions of tourism information inflows and outflows using GIS spatial analysis, rank-size law and kernel density estimation.

Firstly, in terms of temporal characteristics, COVID-19 has significantly decreased tourism information flows, particularly from 2020 to 2022, during which the damage rate exceeded 50%. Although there has been a rebound in 2023, it has not yet returned to pre-pandemic levels. Regarding spatial characteristics, the distribution of tourism flow exhibits a clear pattern of “high in the east and low in the west”. The distribution of popular tourism sources is “bow-shaped”, while that of popular tourism destinations is “U-shaped.” Additionally, there is a significant spatial dislocation between tourism sources and destinations.

Secondly, the results of tourism flow network structure indicate that the distribution of tourism sources is relatively well-ordered, while the distribution of tourism destinations tends to follow a “primacy-type” distribution. Beijing and Shanghai attract a strong inflow of tourism information, and Zhangjiajie and Lhasa experience a substantial outflow of tourism information. In contrast, Chengdu and Chongqing exhibit a relatively balanced inflow and outflow of tourism information. Furthermore, the

“head” cities present a sporadic distribution, the “middle” cities are crowded, and the “tail” cities have a reasonable representation. However, following the impact of COVID-19, the Matthew effect emerged in tourist destinations in Chinese cities, indicating that cities with strong tourism foundations are more resilient to external shocks.

Finally, the findings of this study both align with and differ from previous conclusions. Some scholars similarly focused on tourism flows within Chinese cities; however, their original data was derived from travelogues on travel websites, such as Ctrip. A comparison revealed that actual tourism routes also follow the distribution pattern of higher numbers in the east and lower numbers in the west, and, to some extent, conform to the rank-size rule. Interestingly, in their research, the Yangtze River Delta and Chengdu-Chongqing regions are popular areas, while Beijing-Tianjin-Hebei and the Pearl River Delta are less prominent (Wang et al. 2020; Fang et al. 2022). However, in this study, these two regions are recognized as significant tourist sources. This illustrates that different data sources can yield different results; however, they can mutually validate and complement each other. Once again, this demonstrates the necessity and importance of exploring development patterns and tourism planning based on urban attributes.

Therefore, to promote the scientific planning and sustainable development of tourism in the post-epidemic era, it is necessary to facilitate the rational mobility of tourism elements and enhance inter-city cooperation. The recommendations presented in this paper are as follows:

The first recommendation is to adhere to the law of spatial and temporal evolution of tourism flow and continue leveraging the radiating and driving role of core regions such as the Beijing-Tianjin-Hebei, Yangtze River Delta, Pearl River Delta, and Chengdu-Chongqing regions. Strengthening cooperation between cities at different levels and promoting the equitable distribution of tourism production factors is also crucial. Additionally, given the remarkable rise of Lhasa as a tourist destination in the new era, we recommend promoting the integration of culture and tourism as a strategic approach to enhance the appeal of the destination. This integration not only strengthens the cultural identity of regions but also enriches the tourism experience, playing a crucial role in attracting visitors from all directions in the post-pandemic era.

The second recommendation is to emphasize the multiple push and pull effects of the economy, resources and policies, and to establish a virtuous cycle of mutual benefit between tourist sources and destinations. For instance, Guizhou has signed strategic agreements with Shanghai, Jiangsu, and several other provinces and cities to exchange resources and has implemented preferential measures, such as mutual exemptions for tickets. These measures have significantly enhanced the connection between tourist sources and destinations. Other cities, particularly those classified as “head” and “tail” cities, should learn from these measures to find strategic partners according to their specific attributes. In addition, for numerous “middle” cities, it is crucial to leverage their unique advantages and pursue differentiated competition, striving to embody the principle of “what others lack, we have; what others have, we excel”.

The third recommendation is to recognize the tourism characteristics of the city, building on strengths and mitigating weaknesses. Tourist sources, particularly megacities, should focus on tourism product development and accurately identify tourism demands, while adhering to the principles of innovative, green, and shared development. Tourist destinations, particularly those in less developed regions, should emphasize marketing and establishing strong tourism brands. For example, cities in the

Northeast and Northwest, which have relatively weak economic and transportation foundations, should vigorously develop unique tourism routes, such as desert, and ice and snow tourism, to leverage tourism for economic development.

Overall, the study of the spatiotemporal distribution and network structure of tourism flow is a crucial issue in tourism geography and a practical necessity for the high-quality development of tourism. This study constructs a tourism flow network among 296 cities in China based on the Baidu Index. From the perspective of both source and destination attributes, it reveals the evolutionary patterns of tourism flows in China, characterized by “broad scope, small scale and long term”. This study provides a scientific basis for tourism planning and management, and promotes sustainable development in urban tourism.

However, this study has several limitations. (1) Due to limited data collection capacity, the data sources in this study may be somewhat simplistic, and future research should aim to expand these sources to include more comprehensive data from questionnaire surveys and user-generated content. (2) Due to the lack of statistical data on economic, demographic, and other indicators in most cities, along with the difficulty in measuring cultural distance, this study lacks empirical tests on the formation mechanisms of tourism flows. Future research should explore alternative indicators for deeper analysis. (3) While this study focuses on inter-city tourism flows from 2019 to 2023, future studies should analyze China’s domestic tourism flows over a longer timeframe and in smaller units to provide more comprehensive insights for more effective tourism planning and management.

Data availability

The datasets generated during and/or analyzed during the current study are available in <https://github.com/Data-loading/123>.

Received: 19 October 2023; Accepted: 13 December 2024;

Published online: 04 January 2025

References

- Alderighi M, Gaggero AA (2019) Flight availability and international tourism flows. *Ann Tour Res* 79:102642
- Arbulú I, Razumova M, Rey-Maqueira J, Sastre F (2021) Can domestic tourism relieve the COVID-19 tourist industry crisis? The case of Spain. *J Destin Mark Manag* 20:100568
- Ayoub D, Mohamed DNHS (2024) The impact of push-pull motives on internal tourists’ visit and revisit intentions to Egyptian domestic destinations: the mediating role of country image. *Humanit Soc Sci Commun* 11:358
- Bai ZF, Han L, Jiang XH et al. (2022) Spatiotemporal evolution of desertification based on integrated remote sensing indices in Duolun County, Inner Mongolia. *Ecol Inform* 70:101750
- Bulchand-Gidumal J (2022) Post-COVID-19 recovery of island tourism using a smart tourism destination framework. *J Destin Mark Manag* 23:100689
- Cao XT, Zhang HL, Wang MQ et al. (2023) Stay or travel? Spatial heterogeneity impact of air pollution on travel intention. *Curr Issues Tour* 27(7):1064–1077
- Central People’s Government of the People’s Republic of China (2022) Nucleic acid, health code and travel code of inter-regional mobile persons will no longer be checked. https://www.gov.cn/xinwen/2022-12/07/content_5730576.html
- Chen Q, Chen WX, Wu D et al. (2022) Spatiotemporal evolution and influencing factors of tourism development efficiency in the Yangtze River Economic Belt, China. *J Clean Prod* 379:134722
- China Internet Network Information Center (2023) Statistical Report on Internet Development in China. <https://www.cnnic.net.cn/n4/2023/0303/c88-10757.html>
- Danish, Wang ZH (2018) Dynamic relationship between tourism, economic growth, and environmental quality. *J Sustain Tour* 26(11):1928–1943
- Dogru T, Bulut U (2018) Is tourism an engine for economic recovery? Theory and empirical evidence. *Tour Manag* 67:425–434
- Dou WK, Zhang HL, Xu CX et al. (2024) Spatial evolution pattern of tourism flow in China: case study of the May Day Holiday based on Baidu migration data. *Curr Issues Tour*. <https://doi.org/10.1080/13683500.2024.2345179>
- Fan XC, Lu JY, Qiu MX et al. (2023) Changes in travel behaviors and intentions during the COVID-19 pandemic and recovery period: a case study of China. *J Outdoor Recreat Tour* 41:100522
- Fang YL, Huang ZF, Li JL et al. (2022) Research on the spatial differentiation and effects of network structure in tourism flow in Chinese cities: big data mining based on Ctrip. *J Nat Res* 37(01):70–82
- Gossling S, Scott D, Hall CM (2021) Pandemics, tourism and global change: a rapid assessment of COVID-19. *J Sustain Tour* 29(1):1–20
- Guo YR, Zhang J, Zhang HL (2016) Rank-size distribution and spatio-temporal dynamics of tourist flows to China’s cities. *Tour Econ* 22(3):451–465
- Hao Y, Bai HY, Sun SW (2021) How does COVID-19 affect tourism in terms of people’s willingness to travel? Empirical evidence from China. *Tour Rev* 76(4):892–909
- Ho CI, Lin MH, Chen HM (2012) Web users’ behavioural patterns of tourism information search: from online to offline. *Tour Manag* 33(6):1468–1482
- Hu ZL, Zhu ST (2023) Impact of the COVID-19 outbreak on China’s tourism economy and green finance efficiency. *Environ Sci Pollut R* 30(17):49963–49979
- Huang XK, Zhang LF, Ding YS (2017) The Baidu Index: uses in predicting tourism flows—a case study of the Forbidden City. *Tour Manag* 58:301–306
- Jin X, Qu MY, Bao JG (2019) Impact of crisis events on Chinese outbound tourist flow: a framework for post-events growth. *Tour Manag* 74:334–344
- Kang JF, Guo XY, Fang L et al. (2022) Integration of Internet search data to predict tourism trends using spatial-temporal XGBoost composite model. *Int J Geogr Inf Sci* 36(2):236–252
- Khalid U, Okafor LE, Burzynska K (2022) Do regional trade agreements enhance international tourism flows? Evidence from a cross-country analysis. *J Travel Res* 61(6):1391–1408
- Kim SE, Lee KY, Shin SI et al. (2017) Effects of tourism information quality in social media on destination image formation: the case of Sina Weibo. *Inform Manage* 54(6):687–702
- Kim S, Prideaux B, Timothy D (2016) Factors affecting bilateral Chinese and Japanese travel. *Ann Tour Res* 61:80–95
- Lau PL, Koo TTR, Wu CL (2016) Network analysis of tourist flows: a cross-provincial boundary perspective. *Tour Geogr* 18(5):561–586
- Lewis C, Prayag G, Pour S (2021) Linking travel motives to identity and travel behavior of the Australian LGBT market during COVID-19. *J Travel Tour Mark* 38(7):725–741
- Li L, Tao ZM, Lu L et al. (2022) The impact of COVID-19 on the regional tourism flow network: an empirical study in Hubei Province. *Curr Issues Tour* 25(2):287–302
- Li S, Liu JJ, An K (2023) Research on network structure characteristics and spatial effects of provincial tourism flow in China based on online data. *J Stat Inform* 38(1):116–128
- Li SJ, Ding JQ, Zheng X et al. (2021) Beach tourists behavior and beach management strategy under the ongoing prevention and control of the COVID-19 pandemic: a case study of Qingdao, China. *Ocean Coast Manag* 215:105974
- Li Y, Li YP, Li LL et al. (2022) Tourism demand forecasting from the perspective of mobility: a brand-new predictive variable generated from intercity population mobility big data. *Asia Pac J Tour Res* 27(5):526–546
- Lin HH, Hsu IC, Lin TY et al. (2022) After the epidemic, is the smart traffic management system a key factor in creating a green leisure and tourism environment in the move towards sustainable urban development? *Sustainability* 14(7):3762
- Lin H, Liu Y, Li G et al. (2021) Tourism demand nowcasting using a LASSO-MIDAS model. *Int J Contemp Hosp Manag* 33(6):1922–1949
- Liu PX, Zhang HL, Zhang J et al. (2019) Spatial-temporal response patterns of tourist flow under impulse pre-trip information search: from online to arrival. *Tour Manag* 73:105–114
- Liu XY, Fu X, Hua C et al. (2021) Crisis information, communication strategies and customer complaint behaviours: the case of COVID-19. *Tour Rev* 76(4):962–983
- Liu YW, Liao W (2021) Spatial characteristics of the tourism flows in China: a study based on the Baidu Index. *ISPRS Int J Geo Inf* 10(6):378
- Liu YP, Yu J, Wang FQ (2022) Inbound tourism flows and foreign exchange revenue in the ASEAN from China and the world. *Curr Issues Tour* 25(4):524–540
- Lu JY, Xiao X, Xu ZX et al. (2022) The potential of virtual tourism in the recovery of tourism industry during the COVID-19 pandemic. *Curr Issues Tour* 25(3):441–457
- Matsuura T, Saito H (2022) The COVID-19 pandemic and domestic travel subsidies. *Ann Tour Res* 92:103326
- Ma Y, Tong Y (2019) From region to field: reconceptualising the relationship between culture and tourism. *Tour Tribune* 34(04):7–9

- Medina RMP, Medina RMP, Martínez JMG et al. (2022) Analysis of the role of innovation and efficiency in coastal destinations affected by tourism seasonality. *J Innov Knowl* 7(1):100163
- Mou NX, Yuan RZ, Yang RZ et al. (2020) Exploring spatio-temporal changes of city inbound tourism flow: the case of Shanghai, China. *Tour Manag* 76:103955
- Mou NX, Zhang YH, Makkonen T et al. (2020) Tourists' digital footprint: the spatial patterns of tourist flows in Qingdao, China. *Tour Manag* 81:104151
- National Bureau of Statistics (2024) National data. <https://data.stats.gov.cn/easyquery.htm?cn=C01>
- Navio-Marco J, Ruiz-Gómez LM, Sevilla-Sevilla C (2018) Progress in information technology and tourism management: 30 years on and 20 years after the internet—revisiting Buhalis & Law's landmark study about eTourism. *Tour Manag* 69:460–470
- Okafor LE, Khalid U, Khalid U (2022) The effect of migration on international tourism flows: the role of linguistic networks and common languages. *J Travel Res* 61(4):818–836
- Paparrizos S, Schindler D, Potouridis S et al. (2018) Spatio-temporal analysis of present and future precipitation responses over South Germany. *J Water Clim Chang* 9(3):490–499
- Paramati SR, Alam MS, Chen CF (2017) The effects of tourism on economic growth and CO2 emissions: a comparison between developed and developing economies. *J Travel Res* 56(6):712–724
- Park S, Kim YR, Ho CST et al. (2022) Analysis of travel mobility under Covid-19: application of network science. *J Travel Tour Mark* 39(3):335–352
- Peng HS, Zhang JH, Liu ZH et al. (2016) Network analysis of tourist flows: a cross-provincial boundary perspective. *Tour Geogr* 18(5):561–586
- Qiang MM (2022) Quantifying the loss of China's tourism revenue induced by COVID-19. *Curr Issues Tour* 25(24):3919–3924
- Rather RA (2021) Demystifying the effects of perceived risk and fear on customer engagement, co-creation and revisit intention during COVID-19: a protection motivation theory approach. *J Destin Mark Manag* 20:100564
- Ruan WQ, Zhang SN (2021) Can tourism information flow enhance regional tourism economic linkages? *J Hosp Tour Manag* 49:614–623
- Schmöcker JD (2021) Estimation of city tourism flows: challenges, new data and COVID. *Transp Rev* 41(2):137–140
- Shao YH, Huang SS, Wang YY et al. (2020) Evolution of international tourist flows from 1995 to 2018: a network analysis perspective. *Tour Manag Perspect* 36:100752
- Sharma GD, Thomas A, Paul J (2021) Reviving tourism industry post-COVID-19: a resilience-based framework. *Tour Manag Perspect* 37:100786
- Shi XT, Wu JF, Wu BQ (2022) Distribution characteristics and spatial conjugate relationship of heterogeneous tourist flows from the perspective of multiple transportation. *Sci Geogr Sci* 42(9):1546–1554
- Shoval N, Isaacson M (2007) Tracking tourists in the digital age. *Ann Tour Res* 34(1):141–159
- Sigala M (2020) Tourism and COVID-19: impacts and implications for advancing and resetting industry and research. *J Bus Res* 117:312–321
- Sun GN, Xing D (2023) The impact of COVID-19 on China's household consumption and consumption structure in 2020–2022. *Acta Geogr Sin* 78(04):894–912
- Sun SL, Wei YJ, Tsui KL et al. (2019) Forecasting tourist arrivals with machine learning and internet search index. *Tour Manag* 70:1–10
- Tan T, Zha JP, Tang JY et al. (2022) Spatial-temporal disparities in the impact of the coronavirus disease 2019 on domestic tourism demand: a study of Emeishan National Park in Mainland China. *J Vacat Mark* 28(3):261–284
- Tian SZ, Jiang JL, Li H et al. (2023) Flow space reveals the urban network structure and development mode of cities in Liaoning, China. *Human Soc Sci Commun* 10(1):257
- United Nations World Tourism Organization (2024) International Tourism and Covid-19. <https://www.unwto.org/tourism-data/international-tourism-and-covid-19>
- Wang HJ, Su JX, Li XX et al. (2020) The spatial agglomeration of tourist routes and the scheme of China's tourism geographical division. *Econ Geogr* 40(09):213–221
- Wang SJ, Huang YY, Zhou YQ (2019) Spatial spillover effect and driving forces of carbon emission intensity at the city level in China. *J Geogr Sci* 29(2):231–252
- Wang YW, Chen H, Wu XY (2021) Spatial structure characteristics of tourist attraction cooperation networks in the Yangtze River delta based on tourism flow. *Sustainability* 13(21):12036
- Weng GM, Pan Y (2021) Tourism gravitation and its response relationship with tourist inflows propensity of Beijing-Tianjin-Hebei region in May Day Holiday. *Appl Spat Anal Policy* 14(1):1–29
- Wong JWC, Lai IKW (2022) The mechanism influencing the residents' support of the government policy for accelerating tourism recovery under COVID-19. *J Hosp Tour Manag* 52:219–227
- World Tourism Organization (2022) Tourism in the 2030 agenda. <https://www.unwto.org/tourism-in-2030-agenda>
- Wu LL, Liu Y, Liu K et al. (2023) Analysis of tourist market structure and its driving factors in small cities before and after COVID-19. *ISPRS Int J Geo Inf* 12(6):243
- Xie CW, Li BB, Zhang JC et al. (2023) Effects of risk message on loneliness and travel intention: roles of conflictive family atmosphere and risk propensity in the context of COVID-19 pandemic. *Curr Issues Tour* 27(6):954–971
- Xu J, Yang Y, Jin C (2022) Tracking discrepancies between expected and actual flows of tourists in an urban destination: An application of user-generated data. *J Hosp Tour Manag* 52:29–38
- Xu XH (2022) The impacts of COVID-19 on the rank-size distribution of regional tourism central places: a case of Guangdong-Hong Kong-Macao Greater Bay Area. *Sustainability* 14(19):12184
- Yu L, Zhao PJ, Tang JQ et al. (2023) Changes in tourist mobility after COVID-19 outbreaks. *Ann Tour Res* 98:103522
- Yu WH, Jin X (2022) Does environmental information disclosure promote the awakening of public environmental awareness? Insights from Baidu keyword analysis. *J Clean Prod* 375:134072
- Zhang HY, Song HY, Wen L et al. (2021) Forecasting tourism recovery amid COVID-19. *Ann Tour Res* 87:103149
- Zhang K, Chen Y, Li CL (2019) Discovering the tourists' behaviors and perceptions in a tourism destination by analyzing photos' visual content with a computer deep learning model: the case of Beijing. *Tour Manag* 75:595–608
- Zhang SQ, Sun TW, Lu Y (2023) The COVID-19 pandemic and tourists' risk perceptions: tourism policies' mediating role in sustainable and resilient recovery in the new normal. *Sustainability* 15(2):1323
- Zhang YM, Guo XY, Su YY et al. (2023) Changes in spatiotemporal pattern and network characteristics in population migration of China's cities before and after COVID-19. *Humanit Soc Sci Commun* 10:673
- Zhao QS, Zhu XJ (2024) Provincial trade high-quality development level measurement, regional differences and time-space evolution. *Stat Decis* 40(9):140–144
- Zheng WM, Li ML, Lin ZB et al. (2022) Leveraging tourist trajectory data for effective destination planning and management: a new heuristic approach. *Tour Manag* 89:104437
- Zhong LS, Chen DJ (2019) Progress and prospects of tourism climate research in China. *Atmosphere* 10(11):701
- Zhou XH, Chen WW (2021) The impact of informatization on the relationship between the tourism industry and regional economic development. *Sustainability* 13(16):9399
- Zillinger M (2020) The curious case of online information search. *Curr Issues Tour* 23(3):276–279

Acknowledgements

This work was supported by the National Social Science Foundation of China (23BJY143).

Author contributions

YT: conceptualization, methodology, data curation, visualization, formal analysis, writing—original draft and writing—review & editing. GW: supervision, validation and funding acquisition. SQ: data curation, writing—review & editing. YP: writing—review & editing.

Competing interests

The authors declare no competing interests.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Informed consent

This article does not contain any studies with human participants performed by any of the authors.

Additional information

Correspondence and requests for materials should be addressed to Gangmin Weng.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025